Title
Identifying Intrinsic Regulators of Areal Patterning in the Neocortex by Single-Cell RNA-seq and Organoid Arealization

Author
Sandoval Espinosa, Carmen del Rocio

Publication Date
2021

Peer reviewed|Thesis/dissertation

Identifying Intrinsic Regulators of Areal Patterning in the Neocortex by Single-Cell RNA-seq and Organoid Arealization

by

Carmen Sandoval Espinosa

DISSERTATION

Submitted in partial satisfaction of the requirements for degree of

DOCTOR OF PHILOSOPHY

in

Neuroscience

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:

Mercedes Paredes

404BC180B179406...

Mercedes Paredes

Chair

DocuSigned by:

Xianhua Piao

DocuSigned by:

Arnold Kriegstein

DocuSigned by:

Georgia Panagiotakos

613938B56C88470...

Xianhua Piao

Arnold Kriegstein

Georgia Panagiotakos

Committee Members

To my mom and dad, for allowing me to always be a wild spirit in search for answers.

Acknowledgements

I owe my gratitude, first and foremost, to my mentor, Arnold Kriegstein. Thank you for trusting me work in your lab without knowing much about me except the fact that I was obsessed with radial glia, as I made it clear in my email to you as a recent college graduate in 2016. I have grown not only as a scientist, but as a person, throughout the innumerable experiences that spun out of that email and your trust in me. The knowledge and experiences I have gained, I will cherish for a lifetime.

Thank you to my committee members, Georgia Panagiotakos, Mercedes Paredes, and Xian Piao— it has been a true pleasure to learn from you. Your expertise and advice have instrumentally guided my work, and I have seen all the further because of it. To the Neuroscience program coordinators, Pat Veitch and Lucita Nacionales, thank you for your superb administrative support throughout my far-from-straightforward PhD journey.

I would like to extend a heartfelt thank you to all the members of the Kriegstein Lab, past and present. I would especially like to thank William Walantus, Maureen Galvez, Shaohui Wang, Lizzie Di Lullo, Ugomma Eze, Madeline Andrews, Aparna Bhaduri, Tanzila Mukhtar, Laskshmi Subramanian, Tom Nowakowski, Alex Pollen and Jiadong Chen. This work would not have been possible without your knowledge and endless support, and I am grateful for your mentorship and friendship.

Ugomma Eze, your wit, laughter and endless compassion made the lab a wonderful place to work. I will miss our time together as graduate students tremendously.

Lizzie Di Lullo, thank you for showing me how to be an integral researcher, to work thoughtfully and to care deeply about the process and not just the results. It has truly been an honor to learn from you how to do research.

Madeline Andrews, your dedication, hard work, and equanimity are what I aspire to. Thank you for setting the bar so high for future women scientists.

Aparna Bhaduri, thank you for showing me how to work hard and stay focused, for encouraging me to keep going and for your extensive patience when I thought I couldn't.

I also would like to thank my best friend, Oscar, for accompanying me every step of the way. From helping me figure out San Francisco, to being by my side many times patiently waiting for the storm to pass, to celebrating the good moments–I know you are always there, and I cannot thank you enough for it.

To my partner, Brian, thank you for seeing in me what at times I forget about myself. I am lucky beyond belief to have crossed paths with you–adventurous, kind, thoughtful and unpredictably fun *you*. I am not exaggerating when I say the final stretch was all possible because you were by my side. Thank you for being so patient, understanding, and for always encouraging me to look forward and think bigger.

I am eternally grateful for my parents, Carmen and Rodrigo, for showing me, from the day I was born, how to believe in myself. For never doubting my capacity to pull off whatever crazy endeavor I decide to get myself into (this PhD included). Thank you for seeding in me the unshakeable belief that I will eventually figure things out– it goes everywhere with me, just as you do. Thank you for the endless love you have for me, and for the support and encouragement you have ceaselessly provided my whole life. And to my brother, Rodrigo: thank you for being a constant source of inspiration and laughter in

my life. I secretly think we're twins and they tricked us into thinking we were born two years apart. Thank you for always making me feel that I'm not alone in the world.

To my "second parents", Gaby and Felipe, for always being there, tirelessly supporting me even in the worst of times. Thank you for always coming to the rescue. I could not have done this without you. To Sebastián, for somehow being wise beyond your years and always reminding me that everything is going to be OK.

It really does take a village. I am forever indebted to every single person who has supported me on this long road. It's hard to believe I made it all the way. But when I stop to think of all the people who surround me, it's even harder to understand how I ever thought I wouldn't. Thank you.

Contributions to Published Work

All of the work described in this dissertation was done under the supervision and guidance of Arnold Kriegstein.

The content of Chapter 2 was modified from the publication:

Bhaduri A*, **Sandoval-Espinosa C***, Otero-Garcia M, Oh I, Yin R, Eze UC, et al. An Atlas of Cortical Arealization Identifies Dynamic Molecular Signatures. *Nature,* 2021.

Identifying Intrinsic Regulators of Areal Patterning in the Human Neocortex by Single-Cell RNA-seq and Organoid Arealization

by Carmen Sandoval Espinosa

ABSTRACT

The human neocortex is the largest, most evolutionarily recent structure of the human brain, composed of billions of neurons and glial cells of still incompletely characterized diversity. Arguably the most complex structure of the human body, the cortex is often referred to as "the crowning jewel of evolution", and it is the structure that most distinguishes us from other species. Evolved from the dorsal cortex of reptiles, this exquisitely organized six-layered structure is unique to mammals and is responsible for our higher-order brain functions. It enables our cognitive abilities and is a key biological substrate for consciousness. Accordingly, the neocortex is also a vulnerable target of many neurological and neuropsychiatric disorders.

One of the most prominent characteristics of the neocortex is its organization into distinct cytoarchitectonic areas, cortical regions with distinct cellular organization, connectivity and function. For over a century, developmental neurobiologists have sought to understand how the neocortex is patterned into these distinct, functionally specialized areas throughout development. To date, large scale sequencing efforts have enabled unprecedented insights into the emergence of cellular diversity in the developing cortex. However, there remains a paucity of studies interrogating how areal identity, a key determinant of cortical circuit development, emerges.

In this thesis, I describe our efforts to better understand the intrinsic factors that establish molecular differences across areas of the developing neocortex. Following an

introductory chapter, I first describe our efforts using single-cell transcriptome profiling and single molecule fluorescence RNA *in situ* hybridization to identify molecular subtypes of progenitor cells and excitatory neurons specific to prospective cortical areas during mid-fetal developmental stages. We determined unique genetic markers and expression signatures of these populations, with particular emphasis on area-specific transcription factors, and we found unexpectedly dynamic brain region- and area-specific gene expression signatures across developmental time and lineage progression. These findings offer new insights into the dynamics and specificity of areal identity across distinct cell types of the excitatory lineage and across developmental stages of mid-gestation, shedding light into the intrinsic factors that shape cellular diversity across areas of the human neocortex. Our results suggest an integrative view of two prominent and opposing hypotheses for cortical patterning, the protomap and protocortex hypothesis: We find strong evidence for a partial early cortical protomap between cell populations, including progenitors, at the frontal and occipital poles of the neocortex, while cell populations located in between these two poles are less specified towards a particular areal identity at early stages, but become more specified over time.

In chapter 3 of this thesis, I describe my efforts to harness some of what we learned from our transcriptomic profiling of developing cortical areas to direct the differentiation of human embryonic stem (ES) and induced pluripotent stem (iPS) cell-derived cortical neurons towards a frontal or caudal neocortical identity. I discuss the need for improved cortical organoid models aimed at generating neurons of a specific areal identity and the implications of this endeavor for unraveling the etiology of neurodevelopmental psychiatric disorders that selectively affect particular cortical areas, including autism spectrum disorder and schizophrenia.

The overarching goal of the work described in this thesis is to improve our understanding of how distinct areas arise in the neocortex and to use that knowledge to generate cells with distinct areal identities in organoid models of the human brain. Understanding how area-specific cell types are determined during development is important for the accurate and reproducible modelling of human neocortex development *in vitro*. It is especially important for the development of therapies for neuropsychiatric disorders that disproportionately affect specific neocortical areas.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

LIST OF TABLES

Chapter 1

CHAPTER 1

Areal Patterning of the Human Neocortex

Building the human neocortex

The brain begins to form during the third week of development in a human embryo. In the ectoderm, rapidly dividing neuroepithelial (NE) cells give rise to the neural plate, which then invaginates to become the neural tube. Beginning as a monolayer of pseudostratified NE cells, the neural tube will eventually give rise to the entire central nervous system (CNS), which encompasses the brain, retina, and spinal cord. Within the brain, three distinct brain regions will form: the forebrain, midbrain, and hindbrain. The forebrain, also known as the prosencephalon, will then differentiate into the diencephalon (comprised of the thalamus and hypothalamus) and the telencephalon (Figure 1). Lastly, the telencephalic anlage is subdivided into two major regions: the Pax6-expressing dorsal region (the pallium), which forms the cortex, cortical hem and choroid plexus, and the Pax6-negative ventral region (the subpallium), which forms the basal ganglia (Guillemot 2005).

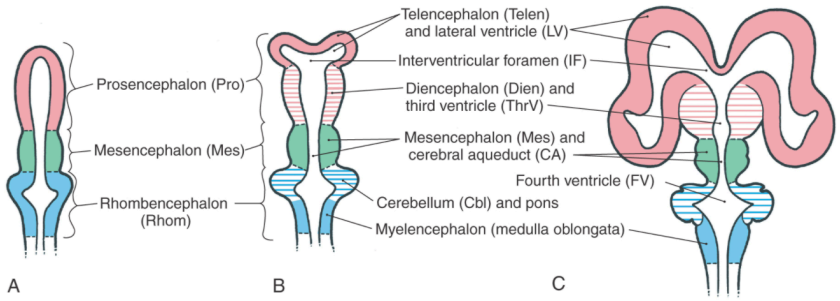


Figure 1.1. Schematic shows the early development of the brain and ventricular system, arising from the neural tube. Reproduced with permission from Fundamental Neuroscience for Basic and Clinical Applications. (Haines 2012)

There are between 14 and 16 billion neurons in the fully developed human neocortex, which are arranged horizontally into layers and radially into columns. During development, the NE cells of the dorsal telencephalon function as stem cells, undergoing

symmetrical divisions to self-renew and asymmetrical divisions to generate postmitotic newborn neurons. Neuroepithelial cells subsequently give rise to ventricular radial glial (vRG) cells, which lack some of the epithelial features of NE cells. vRGs are bipolar cells that span the entire cerebral wall via a long process that contacts the pial surface and another that contacts the ventricular surface. Like NE cells, vRGs also have the capacity to self-renew, as well as to give rise to outer radial glia (oRG) and intermediate progenitor cells (IPCs). In turn, IPCs can self-renew, albeit to a more limited extent, and eventually differentiate into excitatory glutamatergic neurons. Newly generated neurons born in the germinal zones migrate along radial glia processes and into their respective laminar positions along the radial span of the neocortex. Early-born neurons occupy the deeper layers first, and later-born neurons form the more superficial layers in a precisely organized inside-out arrangement. This exquisitely orchestrated, sequential neurogenic process takes place throughout the developing neocortex, and all neocortical excitatory neurons stem from a common progenitor pool.

One of the most prominent characteristics of the neocortex is its organization into distinct cytoarchitectonic areas, cortical regions with distinct cellular and physiological characteristics, as well as unique connectivity patterns. Perhaps most importantly, each of these areas has highly specialized functions, implying that the cellular and molecular differences between cortical areas are a critical substrate for the proper formation of area-specific circuits. For example, the caudal-most portion of the neocortex, the primary visual cortex (V1), specializes in processing visual information. The rostral-most portion, the prefrontal cortex (PFC), on the other hand, is crucial for higher-level cognitive functions that distinguish humans, including reasoning, decision-making, planning and language. In between these two areas with vastly different functions are dozens of other

cortical areas devoted to different cognitive and behavioral functions (e.g., motor, sensory, auditory, associative, etc.). The functional specialization of distinct cortical areas is evident in the effects of localized brain damage, in their distinct activity patterns revealed with modern functional imaging techniques, and in their histological characteristics, as was first described more than 100 years ago.

Brodmann's areas

In 1909, the German neurologist Korbinian Brodmann published "Vergleichende Lokalisationslehre der Großhirnrinde (Localisation in the Cerebral Cortex)" (Brodmann 1909, 2006). This book was the result of years of meticulous observation and comparison of the cortices of several species, including humans[1]. Brodmann noted significant histological differences across regions of the cortical sheet, presenting the first map of the human neocortex, in which he delineated and numbered 52 areas grouped into 11 histological areas[2]. The delineation of these areas, which would later come to be known as "Brodmann areas" (BAs), was based on gross anatomical features, as well as on finer structural and histological characteristics. These characteristics included the cytoarchitectural organization of neurons, which he observed and documented using Nissl staining (Figure 2). Although not proven at the time, Brodmann proposed that structurally-distinct cortical areas would likely be involved in different brain functions within a larger network, a hypothesis that indeed turned out to be the case. Many of the areas he defined based only on their cytoarchitecture have since been associated to

[1] For the English translation of Brodmann's original work, see Garey's translation (Brodmann 2006).

[2] Several years later, in 1925, Constantin von Economo and Georg N. Kosinkas published a more detailed cortical map.

specific cortical functions. For example, Brodmann areas 1, 2 and 3 constitute the primary somatosensory cortex, while area 4 corresponds to the primary motor cortex; area 17 comprises the primary visual cortex, and areas 41 and 42 correspond closely to primary auditory cortex. Higher order functions of associational cortices also robustly and consistently localize to the same Brodmann areas by neurophysiological and functional imaging methods. Brodmann areas have been discussed, refined, and renamed exhaustively for over a century, but more than 100 years after their initial publication, Brodmann's maps of the human neocortex continue to be universally used to locate neuropsychological functions in the human cortex.

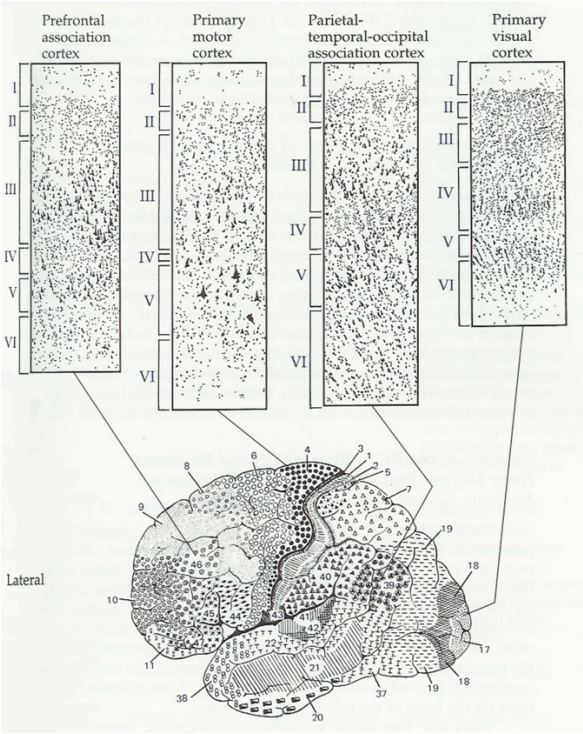


Figure 1.2. The cortical areas of the lateral and medial surfaces of human cerebral cortex printed in the original edition of Brodmann's monograph. Several relevant functional areas have been highlighted.

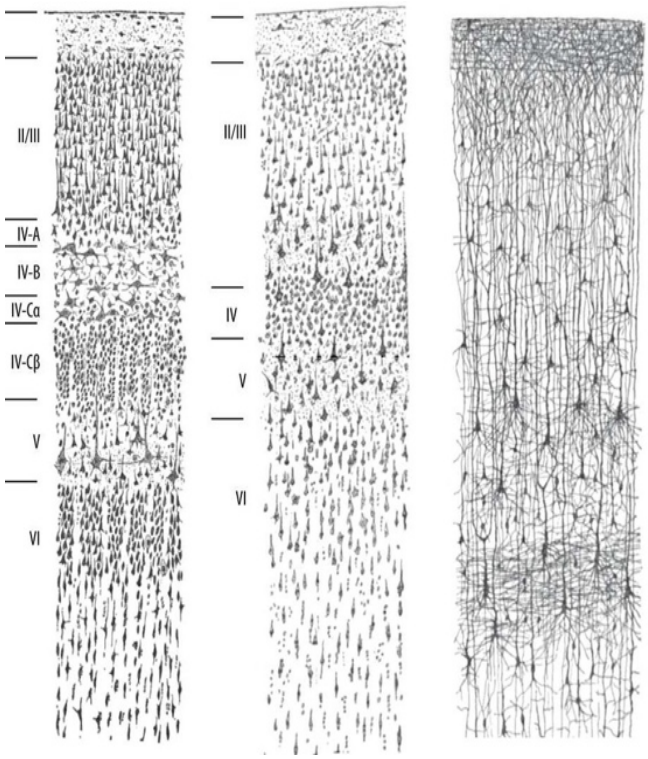


Figure 1.3. Drawings of neocortical cytoarchitecture in the human brain by Santiago Ramon y Cajal, 1911. Left: Nissl staining of human adult visual cortex. Middle: Nissl staining of human adult motor cortex. Right: Golgi staining of the infant (1 month and ½) human neocortex.

It is currently thought that the human neocortex consists of about 120 areas, but this number is far from final. Even within a single cortical region or domain, there are

many related areas. For example, in primates, much of the neocortex is devoted to visual processing, with over twenty distinct visual areas subserving different functions. Some of these areas are quite well understood, while others remain a mystery.

Distinct cortical areas differ strikingly in their cellular architecture. Though the neocortex is generally considered a six-layered structure, there are also striking differences in the laminar organization of different areas. For example, primary motor cortex (M1, BA4) lacks a prominent layer IV, also known as the internal granular layer, which typically contains stellate and small pyramidal cells that receive incoming thalamic inputs. M1 is therefore classically considered an *agranular* area. In contrast, the primary visual cortex (V1) of primates has six readily distinguishable major layers, including a highly developed layer IV, which in turn contains distinct sublayers that are only present in this area.

Additionally, neurons that reside across different areas of the neocortex differ in their connectivity patterns to other cortical areas and brain regions. For example, areas dedicated to processing visual information are located in the caudal, or occipital, lobe of the neocortex and receive input from the lateral geniculate nucleus (LGN) of the thalamus, which is in turn targeted by retinal afferents. In contrast, cortical areas that process auditory stimuli are located in the temporal cortex and receive input from the medial geniculate nucleus (MGN) of the thalamus.

Lastly, there exists a large heterogeneity of cell types, including types of excitatory neurons, across cortical areas. This cellular heterogeneity was already recognized by early neuroanatomists, most notably Santiago Ramón y Cajal, who used the Golgi staining technique (the "*reazione nera*") to generate intricately detailed morphological depictions

of individual cells in neocortex (Figure 1.3). Ramón y Cajal's extensive body of work uncovered a high degree of cellular diversity across the neocortex (Ramón y Cajal 1909).

These neuronal subtypes are still only partly classified, and their distinct functions are far from being well understood, but tremendous progress has been made in uncovering the extent of this diversity over the course of the past century using histological and transcriptomic approaches. Cellular diversity encompasses many features, morphology being only one of them. Glutamatergic neurons in the neocortex show several other divergent cellular properties in relation to their spatial location in the cortex, including their electrophysiological, hodological and gene expression properties.

An intriguing question arises when attempting to explain the origin of the extensive heterogeneity evident across distinct regions of the fully formed neocortex: If indeed a common progenitor pool gives rise to all neocortical excitatory neurons, how then is the neocortex patterned into distinct functional areas during development, each one with distinct morphological and functional characteristics? This remains one of the fundamental questions in developmental neurobiology, and is a key motivator for the work described in this thesis.

The Protomap and Protocortex Views of Cortical Arealization

Over the past four decades, two main hypotheses attempting to explain how neocortical areas are specified have been proposed and intensely debated. In 1988, Rakic proposed that progenitor cells are predetermined, based on their tangential location across the cortical sheet, to acquire area-specific identities and, thus, the cortical neuroepithelium encodes a *protomap* of the future neocortex (P. Rakic 1988). In this

view, progenitors are intrinsically primed to generate neurons with area-specific identities.

Shortly thereafter, in 1989, O'Leary proposed the existence of an initially uniform protocortex–a *"tabula rasa"* that only acquires area-specific identities upon the influence of extrinsic factors, most prominently thalamic afferents (D. D. O'Leary 1989). According to this idea, progenitors across regions of the cortical sheet are all the same. It is now widely accepted that the truth lies somewhere in the middle. A series of elegant studies have shown that the early cortical primordium undergoes intrinsic molecular regionalization, which precedes the subsequent development of more mature areal properties, while later refinement of areal differences results largely from the functional integration of extrinsic thalamocortical afferents.

Classical Studies of Early Patterning Factors in the Mouse Neocortex

A series of seminal studies focused on elucidating the role of transcription factors in early cortical patterning. Using genetic loss-of-function mutant mice, these studies demonstrated that cortical arealization is regulated by a small number of transcription factors expressed in overlapping gradients across the cortical sheet, leading to the establishment of an early cortical map. These transcription factors include COUP-TFI, EMX2, SP8 and PAX6 (Armentano et al. 2007; Polleux 2004; Hamasaki et al. 2004; Muzio 2003). Subsequent studies uncovered FGF8, a diffusible morphogen, as a key player in establishing transcription factor gradients responsible for shaping areal profiles in the neocortex (Storm et al. 2006; Assimacopoulos et al. 2012; Fukuchi-Shimogori and Grove 2001). The vast majority of these early studies on cortical arealization focused on rodent cortex development. Although many mechanisms of cortical development

identified in the rodent are conserved in humans, cortical organization and connectivity is very distinct between these two species. The human brain is 3,800 times larger in terms of mass than that of the mouse, and the cortical sheet of humans is calculated to be 1000 times larger by surface area. As mentioned earlier, the human neocortex comprises about 120 areas, with anthropoid primates having many unique prefrontal areas compared to rodents, and the prefrontal cortex as a whole comprises at least 30% of the entire cortical sheet in humans. In addition to being gyrated, the human neocortex comprises a much greater number of local area identities with differences in structure and function. In stark contrast, distinct parcellations of the mouse cortex range from 40 to 50 areas, dominated by large somatosensory, motor areas, followed by visual, auditory, gustatory, and olfactory areas (Van Essen et al. 2019). Of particular relevance are the mechanisms that underlie the development and expansion of the prefrontal cortex in primates, which remain largely unknown. A recent study reported that the developing neocortex of humans and macaques is characterized by a PFC-enriched gradient of retinoic acid (RA), as well as several potential sources of RA, which are not seen in mice. This distinct PFC-localized RA signaling has been proposed to play a critical role in the evolutionary expansion of the PFC in humans. These pronounced differences between human and rodent cortical development underscore a need for the investigation of cortical areal patterning in human tissue and human model systems.

Transcriptomic Atlases of the Developing Human Neocortex

Beginning in the late 2000's, a number of landmark large-scale studies leveraged gene expression arrays and bulk RNA sequencing to shed light on transcriptomic differences across the developing human brain and neocortex. (Table 1). These studies consistently

found extensive differences in gene expression profiles across areas of the prenatal human neocortex. Each area displayed significant area-specific gene expression and alternative splicing patterns, and rostro-caudal, mediolateral, or frontotemporal gradients of gene expression are common throughout the mid-fetal neocortex. Genes upregulated posteriorally in the human developing neocortex include *EMX2*, *COUPTFI* and *FGFR3*, and frontally-enriched genes include *CNTNAP2*, *PCDH17*, *ROBO1*, and *BCL11B*. In addition, these transcriptomic studies revealed that, while some of the areal markers and graded expression patterns that exist in the mouse developing cortex are also present in human cortices, there are some important differences. For example, Ip et al. found that in developing human cortex, *PAX6* showed a high anterior/lateral, low posterior/medial gradient, as in rodents, but only up until gestational week (GW) 8; no gradient was found afterwards. Of note, these studies also found that, in comparison to developmental timepoints, transcriptional profiles in the adult neocortex are relatively homogeneous throughout cortical areas, although there is still some degree of heterogeneity. Kang et al. produced a comprehensive time course study exploring gene expression in 11 cortical areas, and showed that the largest spatiotemporal variability occurs before birth, with transcriptomes in brain regions converging as we age. The inter-areal heterogeneity seen in adulthood was driven predominantly by the medial prefrontal cortex and V1, without any obvious gradient-like patterns. The genes that are differentially expressed between cortical areas in the adult cortex were mostly related to processes associated with neuronal and synaptic function, whereas fetal differentially expressed genes were mostly related to developmental processes including, neuron differentiation, cell cycle, cell morphogenesis, mitosis, cell adhesion, and cellular component morphogenesis (Kang et al. 2011; Pletikos et al. 2014). Interestingly, comparisons of adult human and non-human

primate brain showed that primates share common, closely matched gene expression patterns across most of the neocortex, but the transcriptional signatures of the human frontal cortex are much more complex (Pletikos et al., 2014). The key findings of these transcriptomic studies of prenatal human neocortex are summarized in Table 1.1.

These seminal transcriptomic studies provided robust evidence that cortical areas are molecularly distinct during human development. However, much remains to be uncovered about the mechanisms that give rise to neuronal diversity across cortical areas during development. The brain is a complex structure made up of a myriad of cell types, and differential expression signatures derived from bulk RNA preparations marked the starting point for explaining the distinct composition of these area-specific profiles. These early studies paved the way for single cell transcriptomic analysis of cellular diversity in the developing brain.

Table 1.1. Transcriptomic Studies of Areal Differences Across the Developing Human Neocortex

Year	Study	Timepoints Sampled	Areas Sampled	Key Findings
2009	Functional and evolutionary insights into human brain development through global transcriptome analysis. Johnson, Kawasawa, Geschwind, State, Sestan, et al. Technology: Affymetrix GeneChip Human Exon 1.0 ST Arrays with Genome-wide whole transcript coverage. Associated Dataset: https://hbatlas.org/	Mid-fetal human: 18, 19, 21, 23 GW.	9 neocortical areas 4 distinct areas of PFC: orbital (OPFC) dorsolateral (DLPFC) medial (MPFC) ventrolateral (VLPFC)	• Found a large number of specific gene expression and alternative splicing patterns, as well as co-expression networks, associated with distinct regions and neuro-developmental processes. • Across 13 cortical regions, 6% of genes are expressed, and 44% of these are differentially regulated. • Differentially expressed genes are more frequently associated with human-specific evolution of putative cis-regulatory elements.
2010	Investigating gradients of gene expression involved in early human cortical development. Ip, B. K., Wappler, I., Peters, H., Lindsay, S., Clowry, G. J., & Bayatti, N. Technology: Affymetrix Human Genome U133 Plus 2.0 Array, qPCR Associated Dataset: N/A, but see Supplementary Tables 1 and 2.	Embryonic and early fetal 8 to 12.5 PCW 8 PCW 10 10.5PCW 12 PCW 12.5 PCW	Cortices were sectioned into 5mm coronal slices along the anterior–posterior axis. RNA was extracted from 2 slices per brain: the anterior-most and posterior-most slices, for comparison of RNA expression levels.	• Identified several genes that exhibit gradient along the anterior–posterior axis of the human neocortex. • Genes upregulated posteriorally vs anteriorally include EMX2, COUPTFI and FGF receptor 3. • Genes upregulated anteriorally include cell adhesion molecules such as cadherins and protocadherins. • Identified potential motor cortex markers and frontal markers in human including CNTNAP2, PCDH17, ROBO1, and CTIP2. • PAX6 showed a high anterior/lateral, low posterior/medial gradient, as in rodent, but only up until GW8; no gradient was found afterwards.
2011	Spatio-temporal transcriptome of the human brain. Kang, H. J., Kawasawa, Y. I., Sousa, A. M. M., Sousa, A., Pletikos, M., Johnson, M. B, Sestan, N. et al. Technology: Affymetrix GeneChip Human Exon 1.0 ST Arrays, genome-wide whole transcript coverage. Associated Dataset: https://hbatlas.org/	Embryonic; early-, mid-, late- fetal. 4PCW 8PCW 10PCW 13PCW 16PCW 19PCW 24 PCW 38PCW	11 neocortical areas: • *For samples 4-8pcw:* FC, frontal cerebral wall PC, parietal cerebral wall TC, temporal cerebral wall • *For samples 10-38 pcw*: OFC, orbital prefrontal DFC, dorsolateral PFC VFC, ventrolateral PFC MFC, medial PFC M1C, primary motor (M1) S1C, primary somatosensory IPC, posterior inferior parietal A1C, primary auditory (A1) STC, superior temporal ITC, inferior temporal	• The majority of spatiotemporal differences were detected before birth, with subsequent increases in similarity among regional transcriptomes: 83% of genes showed spatiotemporal differences across fetal development; 0.9% and 1.4% were temporally regulated across postnatal development and adulthood, respectively. • V1 cortex had the most distinctive transcriptional profile of neocortical areas throughout development and adulthood. • ANKRD32 was transiently expressed in a gradient along the anterior–posterior axis of the mid-fetal frontal cortex, with the highest expression in OFC and the lowest in M1C. The longer isoform of ANKRD32 (ANKRD32a) was equally expressed across fetal cortical areas, while the shorter isoform (ANKRD32b) had dynamic areal patterns. • These spatiotemporal patterns disappeared after birth, when only ANKRD32 was expressed, and were not found in the mouse NCX of equivalent ages.
2011	Temporal dynamics and genetic control of transcription in the human prefrontal cortex. Colantuoni, C., Weinberger, D. R., Kleinman, J. E., et al. Technology: Custom-spotted microarrays, Illumina Oligoset (HEEBO7). Associated Dataset: GSE30272	Mid-fetal: GW 14 GW 15 GW 16, GW 17 GW 18 GW 20	DFC (Dorsolateral PFC)	• Discovered a wave of gene expression changes occurring during fetal development which are reversed in early postnatal life. • This pattern of reversals is mirrored in ageing and in neurodegeneration.

2014	Transcriptional landscape of the prenatal human brain. Miller, J. A., Butler, S., Lein, E. S., et al. **Technology:** Custom Agilent Human 8360K arrays **Associated Dataset:** Brainspan	Mid-fetal: 15pcw 16pcw 21 pcw	~25 neocortical areas spanning all lobes; subdivisons of: Prefrontal cortex Orbital frontal cortex Posterior frontal cortex (motor) Primary motor-sensory cortex Primary somatosensory cortex Subcentral cortex Posterior parietal cortex Temporal cortex Occipital cortex Insular cortex Cingulate neocortex Periallocortex	• Both germinal and post-mitotic cortical layers exhibit fronto-temporal gradients, with particular enrichment in the frontal lobe. • Several neurodevelopmental disorder- and human-evolution-related genes show patterned expression, potentially underlying unique features of human cortical formation. • Spatiotemporal differences exhibit a temporal hourglass pattern, dividing the human neocortical development into three major phases. The first phase, corresponding to prenatal development, shows the highest number of differentially expressed genes among areas and gradient-like expression patterns, including those that are different between human and macaque.
2014	Temporal Specification and Bilaterality of Human Neocortical Topographic Gene Expression. Pletikos, M., Sousa, A., Sedmak, G., Sedmak, G., Meyer, K. A., Zhu, Y., Cheng, F., Li, M., Kawasawa, Y., & Sestan, N **Technology:** Custom Exon Arrays and RNAseq **Associated Dataset:** https://hbatlas.org/	10 – 38 pcw	Medial prefrontal cortex Orbital prefeontal cortex Dorsolateral prefrontal cortex Ventrolateral prefrontal cortex Primary motor cortex Primary somatosensory cortex Parietal cortex Primary auditory cortex Superior temporal cortex Inferior temporal cortex Primary visual cortex	• Interareal differences have a temporal hourglass pattern, dividing human neocortical development into three major phases. Phase 1, prenatal development, shows the highest number of differentially expressed genes among areas and gradient-like expression patterns, including some that are different between human and macaque. • Certain prenatal expression patterns differ between human and macaque. • Population-level areal transcriptomes are globally symmetric across time.
2018	Integrative functional genomic analysis of human brain development and neuropsychiatric risks. BrainSpan & PsychENCODE Consortium Jeremy Willsey, A., Oldre, A., Szafer, A., Camarena, A., Cherskov, A., Charney, A. W., Abyzov, A., Kozlenkov, A., Safi, A., Jones, A. R., Ashley-Koch, A. E., Ebbert, A., Price, A. J., Sekijima, A., Kefi, A., Bernard, A., Amiri, A., Sboner, A., Clark, A., ... Li, Z. (2018). **Technology:** Bulk and single-cell RNAseq; ATAC-seq **Associated Dataset:** Psychencode	PCW 5 PCW 9 PCW 12 PCW15 PCW18 PCW22 PCW27	11 neocortical areas and 5 additional brain regions Prefrontal cortex Primary motor cortex Somatosensory cortex Inferior parietal cortex Auditory cortex Superior temporal cortex Primary visual cortex	• Found a widespread transcriptomic transition beginning during late fetal development, consisting of sharply decreased regional differences. This reduction coincided with increases in mature neuron transcriptional signatures and increased expression of genes associated with dendrite development, synapse development, and neuronal activity. These three functional categories appeared temporally synchronous across neocortical areas, while myelination and oligodendrocyte maturation appeared to be asynchronous across areas.

Studies of Cortical Development at Single-cell Resolution

Although gene expression differences across developing cortical areas were indeed revealed in the studies described above, the extent to which those differences establish distinct area-specific neuronal cell types remained unclear. The possibility of exploring the transcriptome of individual cells from distinct areas of the developing cortex with single-cell RNA profiling has opened up a new era of scientific inquiry into the question of cortical patterning. Historically, disentangling spatiotemporal gene expression changes has been particularly challenging, as many distinct cell types coexist in developing tissues, and cells are at different stages of differentiation and maturation. This heterogeneity can obscure differences in transcriptional programs that exist between cells of each type, particularly when the cell type in question is rare or not a predominant population in the tissue being studied. Hence, it was difficult to determine the extent to which distinct area-specific cell types exist based on the extensive transcriptomic differences seen across the developing neocortex from bulk RNA studies of highly heterogeneous tissue. Single-cell sequencing makes possible a more granular interrogation of the molecular composition of the developing cortex, enabling the study of how distinct cells might be from one area to another. Are there radial glia or intermediate progenitor subpopulations specific to a certain area or domain of the developing cortex? If so, what genes distinguish them from other subpopulations, and what do these genes do? We can now focus on a specific cell population, and explore, for example, the extent to which newborn neurons of a given area are transcriptionally distinct from each other. It is then possible to investigate when areal identities are cemented during the progression from progenitors to neurons and to explore distinct areal-specific transcriptomic profiles at the progenitor stage. These

questions have been the focus of several recent studies, including the central work of this thesis, described in the following chapter.

Since its initial introduction (Tang et al. 2009), single-cell RNA-sequencing (scRNA-seq) technology has developed extensively, along with methods and best practices for its analysis, and has been applied broadly to investigate the cellular composition and heterogeneity of the mammalian brain. Several pioneering studies have demonstrated the power of this technology to unravel the complexity of cell types in the adult and developing brain. In one of the first landmark scRNA-seq papers, over 3000 single cells were sequenced from the adult mouse somatosensory cortex and CA1 region of the hippocampus (Zeisel et al. 2015). The authors identified nine major brain cell types that could be further partitioned into 49 subpopulations. This study fundamentally expanded the classical view of brain cell taxonomy. In a later study, Darmanis et al performed scRNA-seq on 466 cells from adult and fetal human brain to identify genes that are differentially expressed between fetal and adult neurons, as well as genes with expression gradients reflective of the transition between replicating and quiescent fetal populations (Darmanis et al. 2015).

A later study (Tasic et al. 2018) sequenced 23,822 single cells from the mouse primary visual cortex and anterior lateral motor cortex, defining 133 transcriptomic cell types. They found that nearly all types of inhibitory (GABA+) neurons are shared across both areas, whereas most types of excitatory neurons were found in only one of the two areas. Additionally, the authors combined scRNA-seq with retrograde labeling to match excitatory neuron cell types to their long-range projection-specific targets. In a 2017 study (Nowakowski et al. 2017), we analyzed the transcriptomes of single cells from developing human cortex, comparing homologous cell types from the prefrontal cortex (PFC) and

primary visual cortex (V1) from 13 paired specimens. We found that between these two developing areas, modest transcriptional differences among radial glia cascade into robust distinctions between maturing neurons, suggesting that distinct spatially restricted excitatory sublineages might emerge in the rostral and caudal neocortex. To study this question in greater detail and with much more power, we profiled cells from six distinct areas of the human neocortex throughout the entire second trimester. The findings of this study are discussed in the following chapter.

CHAPTER 2

Arealization of the Mid-Fetal Human Neocortex at Single-Cell Resolution

The content of this chapter was modified from the following publication:

Bhaduri A*, **Sandoval-Espinosa C***, Otero-Garcia M, Oh I, Yin R, Eze UC, et al. An Atlas of Cortical Arealization Identifies Dynamic Molecular Signatures. Nature, 2021.

SUMMARY

The human brain is subdivided into distinct anatomical structures. The neocortex, one of these structures, enables higher-order sensory, associative, and cognitive functions, and in turn encompasses dozens of distinct specialized cortical areas. Early morphogenetic gradients are known to establish an early blueprint for the specification of brain regions and cortical areas. Furthermore, recent studies have uncovered distinct transcriptomic signatures between opposing poles of the developing neocortex. However, how early, broad developmental patterns result in finer and more discrete spatial differences across the adult human brain remains poorly understood.

Here, we use single-cell RNA-sequencing to profile ten major brain structures and six neocortical areas during peak neurogenesis and early gliogenesis. Our data reveal that distinct cell subtypes are predominantly brain-structure specific. Within the neocortex, we find that even early in the second trimester, a large number of genes are differentially expressed across distinct cortical areas in all cell types, including radial glia, the neural progenitors of the cortex. However, the abundance of areal transcriptomic signatures increases as radial glia differentiate into intermediate progenitor cells and ultimately give rise to excitatory neurons. Using an automated, multiplexed single-molecule fluorescent in situ hybridization (smFISH) approach, we validated the expression pattern of area-specific neuronal genes and also discover that laminar gene expression patterns are highly dynamic across cortical regions.

Together, our data suggest that early cortical areal patterning is defined by strong, mutually exclusive frontal and occipital gene expression signatures, with resulting gradients giving rise to the specification of areas between these two poles throughout successive developmental timepoints.

Introduction

The specification of the brain into distinct regions has long been of interest to neuroscientists. Explaining how functionally distinct neural structures and cortical areas emerge bridges brain development with adult brain function. Understanding when brain regions acquire their unique features and how this specification occurs has broad implications for the study of human brain evolution, including species-specific developmental differences that may be responsible for the expansion of cortical areas such as the prefrontal cortex (PFC). It is also crucial for characterizing the pathology of neurodevelopmental and neuropsychiatric disorders, that often preferentially impact specific brain regions and/or cortical areas (Rubenstein 2011). Moreover, understanding brain patterning is essential for the accurate recapitulation of human brain development in *in vitro* models, including pluripotent stem cell-derived organoids. Early patterning of the developing telencephalon is orchestrated by morphogenetic gradients of growth factors including bone morphogenetic proteins (Bmps), Wnts, Sonic hedgehog (Shh), and, most prominent in the cortex, fibroblast growth factor (FGF) (Cadwell et al. 2019). However, the molecular patterns that arise as a result later in development are less understood. Many of the seminal studies of cortical arealization took place prior to the widespread availability of next-generation sequencing and were performed in rodent models. Thus, numerous questions remain concerning human brain patterning and cortical arealization, such as the areal specificity of distinct cell populations, changes in gene expression throughout differentiation and maturation, and the processes by which these differences arise in the developing human brain.

Results

To characterize the emergence of cellular diversity across major regions of the developing human brain and across cortical areas, we sequenced the transcriptome of single cells from distinct microdissected regions of developing human brain tissue samples along a ten-week window during the second gestational trimester, which encompasses peak stages of neurogenesis (Manuel et al. 2011). In this study, we refer to the subdivisions of the cerebrum and cerebellum as "regions", and to subdivisions of the cerebral cortex as "areas". We sampled cells from 10 distinct major forebrain, midbrain, and hindbrain regions from 13 individuals (see Methods). Where available, we sampled: neocortex, proneocortex (cingulate), allocortex (hippocampus), claustrum, ganglionic eminences (GE), hypothalamus, midbrain, striatum, thalamus, and cerebellum (Fig 2.1A). In addition, we sampled six neocortical areas from the same individuals: prefrontal (PFC), motor, somatosensory, parietal, temporal, and primary visual (V1) cortex. We used stringent quality control (QC) parameters (see Methods), resulting in 698,820 high-quality cells for downstream analysis.

Regional Identity is Highly Pervasive Across Distinct Cell Types

We began by exploring the cell types and gene expression signatures in a whole brain dataset. Using an iterative clustering approach to mitigate batch effects (see Methods), we characterized the broad cell types and states across the entire dataset. We found expected cell populations of the developing brain, including excitatory neurons, intermediate progenitor cells (IPCs), radial glia, mitotic cells, astrocytes, oligodendrocytes, inhibitory neurons, microglia, and vascular cells (including endothelial cells and pericytes) (Fig 2.1B). Unbiased clustering of cells, as well as their visual

representation in UMAP space, were driven primarily by cell type rather than biological age, except for more mature glial populations, including astrocytes and oligodendrocytes, which were enriched in older samples (Fig 2.1B, SFig 2.1A).

We next sought to identify marker genes specific to or significantly enriched in distinct cell populations in each brain region. We found genes that were region-specific across all cell types, as well as genes that were region-specific for individual cell types. We detected previously described markers of brain regions, including *FOXG1* (cortex) (Manuel et al. 2011), *ZIC2* (cerebellum, also observed in the neocortex) (Aruga et al. 2002), and *NRP1* (allocortex) (Bakken et al. 2016) (SFig 2.1B), as well as other region-specific genes, many of which encode transcription factors. We also identified numerous genes encoding cell type- and brain region-specific transcription factors, including *OTX2, GATA3, LHX9* and *PAX3*. The distribution of cell types across regions was as expected, with progenitor and differentiated cell types identified in each region (SFig 2.1C), leading us to ask whether brain region or cell type is a stronger component of regional identity during the second trimester.

At earlier developmental timepoints, we and others have noted that regional signatures are not broadly pervasive and do not yet reflect area-specific identities of unique brain substructures (Bakken et al. 2016; Eze et al. 2021). To perform this analysis, we first used hierarchical clustering to group individual cell type subclusters. As expected, cluster branches were primarily organized by cell type, validating our annotation approach and highlighting the robustness of cell type in driving cluster similarity. However, by quantifying the proportion of cells from each region contributing to each cluster, it became apparent that the majority (115 of 192) of clusters were strongly enriched for a single brain region (i.e. cortex, thalamus, etc.) or for several related regions

(e.g. forebrain) (SFig 2.1D). A small number of inhibitory and excitatory neuron clusters bridged across regions, while a larger proportion of glial and vascular cells could be identified across brain regions, indicating a strong regional signature for each individual cell type.

To further interrogate the interplay between cell type and brain region identity, we generated constellation plots using cells' combined brain region and cell type annotations. Constellation plots are a powerful tool to visualize the relationships between groups of cells, by quantifying the proportion of cells in a group or cluster with an above-threshold number of nearest neighbors in other groups, while preserving the UMAP topography of the dataset (see Methods). We find that across the whole brain, cell type is the primary source of segregation (Fig 2.1C). However, in certain cases, such as the GE, cells of distinct types from a common region are drawn together in UMAP space, suggesting that the regional identity conferred upon groups of cells of different types can also be a strong source of variation (Fig 2.1C). A heatmap of area-specific gene score enrichments (see Methods) shows that some region-specific genes are present across multiple cell types within a given region. This suggests that some regional gene expression signatures are highly penetrant across cell types. Regionalization is stronger in glial populations at the developmental timepoints we studied (SFig 2.1E). Thus, we identify strong, cell-type independent regional signatures, as well as signatures that are restricted to specific cell types.

The neocortex, allocortex, and proneocortex are evolutionarily closely related, and physically proximal (Supèr et al. 1998). However, we sought to identify distinct regional gene expression programs among these three closely related regions by co-clustering these samples in isolation from the rest of the brain. We observed extensive integration

of cells from the neocortex and allocortex, but segregation of the proneocortex, which was sampled at a later timepoint. We set out to identify the similarities and differences between cell-type and regional expression signatures between these three cortical structures. Using differential gene expression in each excitatory lineage cell type, we again used a gene score annotation paired with hierarchical clustering. Surprisingly, even within these closely related cortical structures, region was still the primary driving force, and again, regional signatures bridged multiple cell types. These analyses indicate that during the second trimester of development, regional signatures are sufficiently established to distinguish cells across brain structures, with some of these signatures extending beyond an individual cell type. We provide analysis of these region-specific gene signatures, including a cell-type specific analysis.

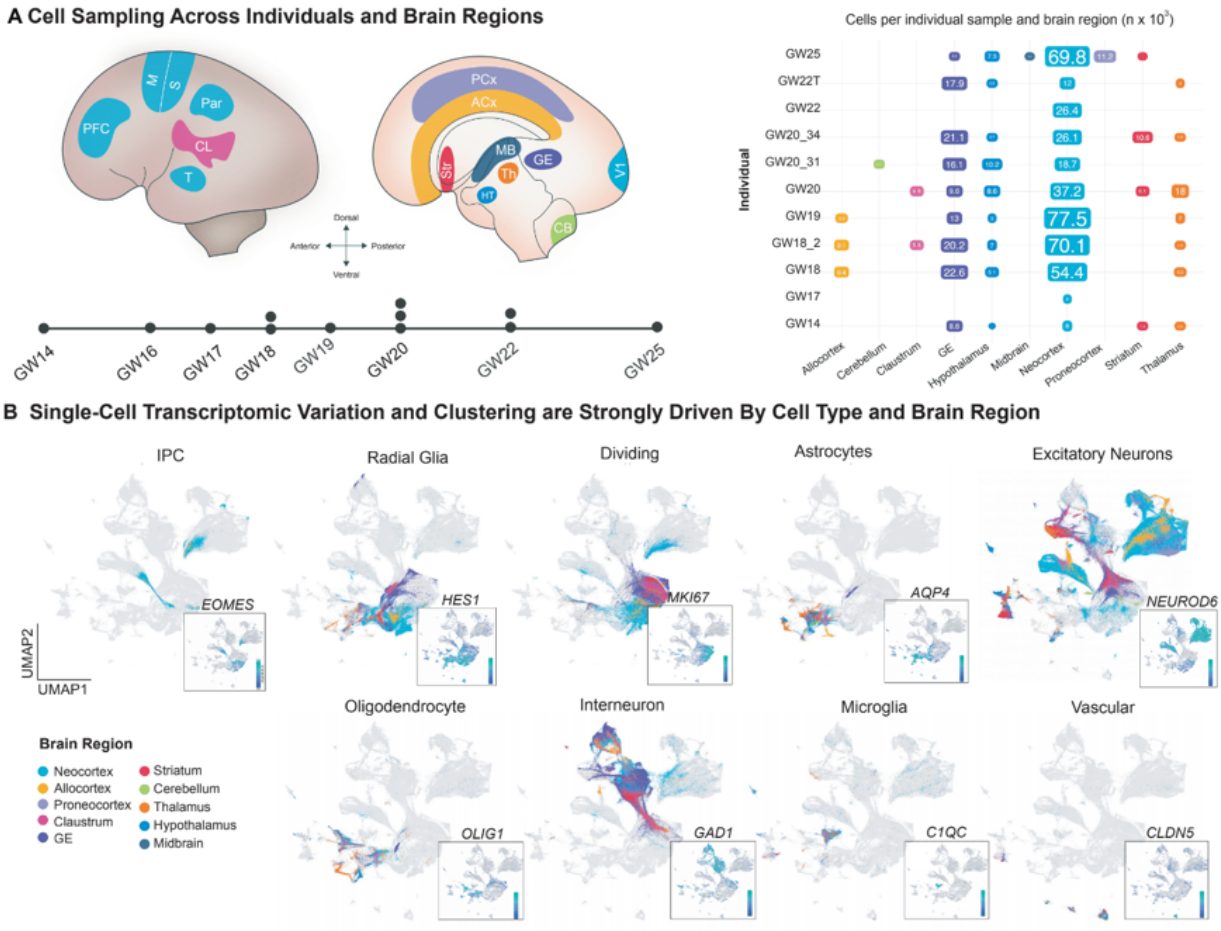

A Cell Sampling Across Individuals and Brain Regions
Cells per individual sample and brain region (n x 10³)
Individual
GW25
GW22T
GW22
GW20_34
GW20_31
GW20
GW19
GW18_2
GW18
GW17
GW14
69.8
26.4
37.2
77.5
70.1
54.4
Allocortex
Cerebellum
Claustrum
GE
Hypothalamus
Midbrain
Neocortex
Proneocortex
Striatum
Thalamus
PFC
M
S
Par
CL
T
PCx
ACx
MB
GE
Th
Str
HT
V1
CB
Dorsal
Anterior
Posterior
Ventral
GW14
GW16
GW17
GW18
GW19
GW20
GW22
GW25
B Single-Cell Transcriptomic Variation and Clustering are Strongly Driven By Cell Type and Brain Region
IPC
Radial Glia
Dividing
Astrocytes
Excitatory Neurons
EOMES
HES1
MKI67
AQP4
NEUROD6
UMAP2
UMAP1
Oligodendrocyte
Interneuron
Microglia
Vascular
OLIG1
GAD1
C1QC
CLDN5
Brain Region
Neocortex
Allocortex
Proneocortex
Claustrum
GE
Striatum
Cerebellum
Thalamus
Hypothalamus
Midbrain
C Constellation Plots Reveal Close Relationships Between Cells of the Same Type Across Distinct Regions
Number of Cells
1,000
5,000
10,000
25,000
63,000
235,000
Shared Connectivity
(fraction of cluster)
Cell Type/State
Excitatory neurons
IPC
Radial Glia
Dividing
Astrocytes
Interneurons
Microglia
Outliers
Vascular
Oligodendrocytes/
OPCs
Brain Region
ACx Allocortex
CB Cerebellum
CL Claustrum
GE Ganglionic
Eminences
HT Hypothalamus
MB Midbrain
NCx Neocortex
PCx Proneocortex
Str Striatum
Th Thalamus

Figure 2.1

Gene Expression Signatures are Highly Cell Type-Specific Across Cortical Areas

The neocortex is made up of dozens of functional areas that specialize in an astounding range of cognitive processes, from sensory perception to reasoning, decision-making and language (Rakic 2009). Longstanding, juxtaposed hypotheses propose the existence of either a cortical protomap (Rakic 1988), where the areal identity of cortical progenitors, and consequently that of their progeny, is cell-intrinsic and genetically predetermined, or a protocortex (D. D. M. O'Leary 1989), where newborn neurons are not areally specified until extrinsic signals, such as those from thalamocortical afferents, reach the developing cortex. Single-cell RNA-sequencing has the power to deconvolute areal gene expression signatures and therefore may provide a means to test these models. Recent work has shown that while neurons are distinct between V1 and PFC soon after their birth (Nowakowski et al. 2017), other cell types do not show such large area-specific differences. Studies in the adult mouse have additionally shown that neuronal cell types of the anterior lateral motor cortex (ALM) and V1, are transcriptionally distinct from each other but that denser sampling of areas between the ALM and V1 reveals a gradient-like transition between cell type profiles (Tasic et al. 2018).

We sought to expand upon these findings by profiling single cells from distinct cortical areas in order to clarify how and when distinctions between areas begin to emerge. To do so, we performed single-cell RNA-sequencing of six cortical regions subdissected from the samples described above, yielding 387,141 high-quality cells, after filtering (see Methods) (SFig 2.2A-B), for subsequent analysis. We applied the same iterative clustering and annotation methods used for our analysis of the whole brain. We found expected cell types, including Cajal-Retzius neurons, dividing cells, excitatory

neurons, inhibitory neurons, IPCs, microglia, oligodendrocyte precursor cells, radial glia/astrocytes, and vascular cells (Fig 2.2A). Hierarchical clustering of 138 neocortical clusters grouped cells by cell type (Fig 2.2B). Additionally, it revealed that most clusters (104/138) are composed of cells from multiple cortical areas. To explore how distinct each cluster was from others of the same cell type or lineage, we again used a constellation plot approach (Fig 2.2C). We found strong connectivity between clusters of the same cell type, suggesting that borders between clusters are fluid. We also found some connectivity, albeit to a lesser degree, between cells from the excitatory lineage. To quantify intra-cell type and inter-cell type connectivity, we calculated the magnitude of transcriptional proximity between nodes (see Methods), and found, not surprisingly, that clusters of each cell type connected most strongly to each other (Fig 2.2D). However, we also found that when comparing inter-cell type connectivity, IPC subclusters connected much more strongly with excitatory neurons than with radial glia subclusters (Fig 2.2D).

We then defined gene signatures characteristic of radial glia, IPCs, and excitatory neurons in the neocortex using a differential gene expression approach (see Methods). These signatures were scored using a module eigengene calculation (see Methods) across the major excitatory lineage cell types (radial glia, IPC, excitatory neurons). We found that radial glia had the highest up-regulation of the progenitor signature, but downregulation of the IPC and neuronal signatures. In contrast, excitatory neurons had the lowest up-regulation of the neuronal signature but strongest downregulation of other programs, perhaps reflecting lack of neuronal maturation. Together with the differences in connectivity strength between cell types seen in the constellation plots, this suggests that there is a strong break between radial glia and excitatory neuron identities, but interestingly, that IPCs track more closely with neurons than radial glia, despite their

progenitor nature. These differences in cell type gene signature strength and cell type connectivity reveal a cascading differentiation program along the excitatory lineage in the neocortex.

In addition to systematically exploring gene expression signatures between excitatory lineage cells along the axis of differentiation, we sought to understand how cortical area influences the identity of cells in the developing cortex, as well as the relationships between cell clusters. We validated our cortical sub-dissections by visualizing the expression of *NR2F1*, a posterior-high to anterior-low expression gradient marker (Harrison-Uy et al. 2013), in our dataset (Fig 2.3A) and previously described cortical area-specific genes (SFig 2.2C). Next, we built constellation plots using a different cell grouping approach, with each group corresponding to a specific area and cell type (Fig 2.3B). Several striking patterns emerged. First, we noted that radial glia nodes were connected primarily to other radial glia, while IPCs and excitatory neurons were frequently mutually connected to one another (Fig 2.3B - C). This break between radial glia versus IPCs and glutamatergic neurons suggests large differences in cell type signature and possibly also areal signatures between the groups. Radial glia from distinct cortical areas were highly interrelated and formed a tight, insular network, with sparse edges to their descendant cell types, IPCs and pyramidal cells (Fig 2.3B). In contrast, excitatory neurons from distinct cortical areas were less interrelated and formed a looser network when compared to radial glia. Of note, neuronal nodes of different areas show robust area-specific transcriptional proximity to their IPC counterparts, suggesting that some degree of the areal specification seen in neurons is already present at the IPC stage (Fig 2.3B). We did not find edges between PFC and V1 cell type nodes (Fig 2.3C), suggesting a model of strong mutual exclusivity between these two gene expression

programs. These patterns are persistent when including cell subtype annotations, as well as when analyzing individual developmental stages separately (SFig 2.3A – H). This was consistent with previous observations of early specification of PFC and V1 identity, so we additionally quantified the number of differentially expressed genes across each cell type by area (Fig 2.3D, Supp Fig 2.2D). Consistent with prior observations (Tasic et al. 2018), the specificity of neuronal areal markers was significantly higher compared to radial glia (Fig 2.3D) but surprisingly, more genes were differentially expressed in radial glia than in neurons (Fig 2.3D). These two observations indicate that markers of areal identity are already detectable in radial glia but become more pronounced as differentiation proceeds, and importantly, there is a significant overlap between the area-specific genes we describe here and those in previous studies (SFig 2.2E).

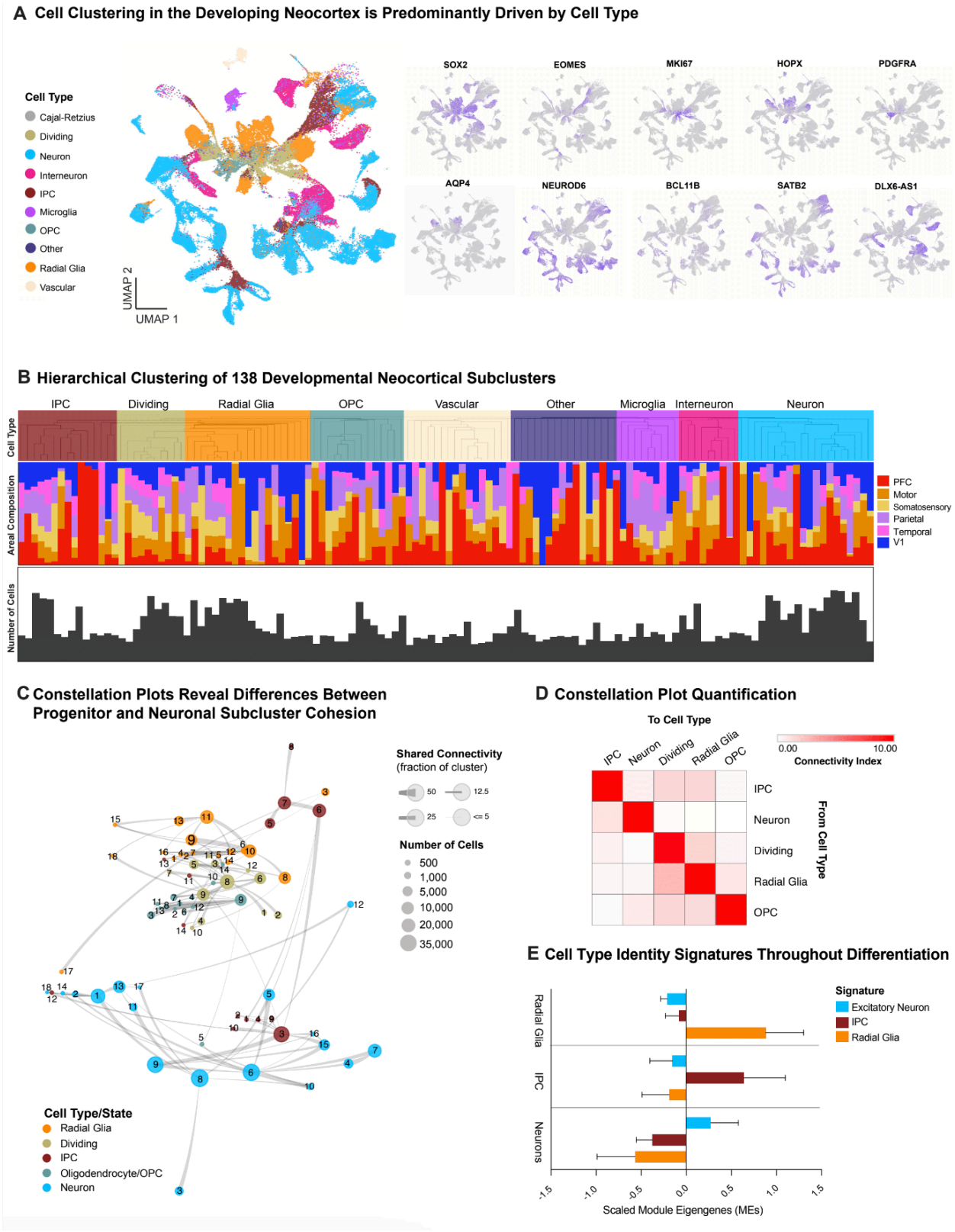

A Cell Clustering in the Developing Neocortex is Predominantly Driven by Cell Type
Cell Type
Cajal-Retzius
Dividing
Neuron
Interneuron
IPC
Microglia
OPC
Other
Radial Glia
Vascular
UMAP 2
UMAP 1
SOX2
EOMES
MKI67
HOPX
PDGFRA
AQP4
NEUROD6
BCL11B
SATB2
DLX6-AS1
B Hierarchical Clustering of 138 Developmental Neocortical Subclusters
IPC
Dividing
Radial Glia
OPC
Vascular
Other
Microglia
Interneuron
Neuron
Cell Type
Areal Composition
Number of Cells
PFC
Motor
Somatosensory
Parietal
Temporal
V1
C Constellation Plots Reveal Differences Between Progenitor and Neuronal Subcluster Cohesion
Shared Connectivity
(fraction of cluster)
50
12.5
25
<= 5
Number of Cells
500
1,000
5,000
10,000
20,000
35,000
Cell Type/State
Radial Glia
Dividing
IPC
Oligodendrocyte/OPC
Neuron
D Constellation Plot Quantification
To Cell Type
IPC
Neuron
Dividing
Radial Glia
OPC
0.00
10.00
Connectivity Index
From Cell Type
E Cell Type Identity Signatures Throughout Differentiation
Radial Glia
IPC
Neurons
Signature
Excitatory Neuron
IPC
Radial Glia
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
Scaled Module Eigengenes (MEs)

Figure 2.2

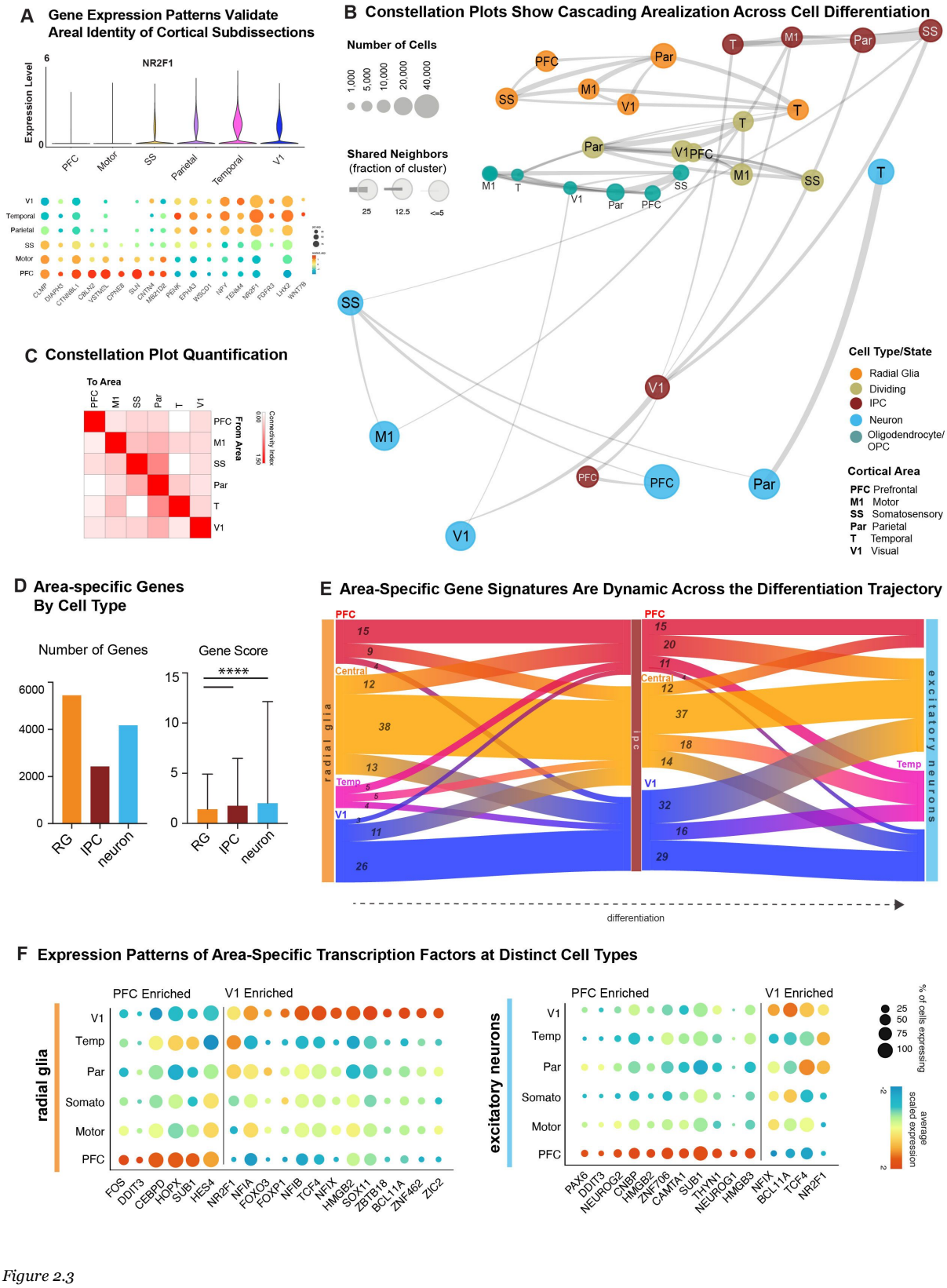

A Gene Expression Patterns Validate Areal Identity of Cortical Subdissections
NR2F1
Expression Level
PFC
Motor
SS
Parietal
Temporal
V1
B Constellation Plots Show Cascading Arealization Across Cell Differentiation
Number of Cells
1,000
5,000
10,000
20,000
40,000
Shared Neighbors
(fraction of cluster)
25
12.5
<=5
Cell Type/State
Radial Glia
Dividing
IPC
Neuron
Oligodendrocyte/ OPC
Cortical Area
PFC Prefrontal
M1 Motor
SS Somatosensory
Par Parietal
T Temporal
V1 Visual
C Constellation Plot Quantification
To Area
From Area
Connectivity Index
D Area-specific Genes By Cell Type
Number of Genes
Gene Score
****
RG
IPC
neuron
E Area-Specific Gene Signatures Are Dynamic Across the Differentiation Trajectory
radial glia
IPC
excitatory neurons
Central
Temp
V1
differentiation
F Expression Patterns of Area-Specific Transcription Factors at Distinct Cell Types
PFC Enriched
V1 Enriched
radial glia
excitatory neurons
% of cells expressing
average scaled expression
FOS
DDIT3
CEBPD
HOPX
SUB1
HES4
NR2F1
NFIA
FOXO3
FOXP1
NFIB
TCF4
NFIX
HMGB2
SOX11
ZBTB18
BCL11A
ZNF462
ZIC2
PAX6
NEUROG2
CNBP
ZNF706
CAMTA1
THYN1
HMGB3
TCF4
NR2F1

Figure 2.3

Dynamics of Area-Specific Gene Expression Signatures

To further interrogate the relationship between differentiation and the transformation of regional signatures across development, we inferred lineage trajectories across cells using RNA velocity (La Manno et al. 2018). This uncovered strong trajectories along known differentiation gradients, primarily between radial glia, IPCs, and excitatory neurons (SFig 2.4A). For each individual cortical area, we identified the most dynamic genes across the differentiation cascade as the top loading velocity genes (SFig 2.4B). Hierarchical clustering of the underlying counts values of these genes across excitatory lineage cells depicted an expected segregation by cell type, and within the excitatory neuron populations, showed small clusters of unique enriched genes across all areas. The gene enrichment in excitatory neuron groups, but not in radial glia or IPC populations led us to ask how areal signatures might change during differentiation. We defined areal signatures of excitatory neurons as gene networks and evaluated their representation in radial glia across cortical areas by calculating their module eigengene scores (SFig 2.4C). We found an early strong binary V1 expression, while PFC gene signature emerged only at later developmental timepoints (SFig 2.4C).

We sought to explore how areal gene signatures change throughout neuronal differentiation by constructing Sankey diagrams to display the overlap between area-specific markers of each cell type. We found that area-specific gene expression signatures change substantially across cell types, with small numbers of areal markers preserved throughout the differentiation trajectory (Fig 2.3E). This is in stark contrast to the pervasive quality of regional signatures which were present across distinct cell types across the whole brain. This suggests that cell type identity is a much stronger contributor to a cell's transcriptional profile than cortical areal identity at these developmental stages.

Area-and-Cell -Type-Specific Transcription Factors

Within each set of areal marker genes, we identified several transcription factors that were robustly enriched in cells of a specific area relative to all other areas, as well as TFs with a broader frontal or caudal enrichment. These transcription factors are shown in a dot plot format to clearly depict both expression level and fraction of cells across cortical areas (Fig 2.3F). These observations suggest that each cortical area is anchored with core transcriptional programs, some of which might persist across cell type boundaries (Fig 2.3F, SFig 2.5A). A subset of marker TFs show consistent area specificity throughout the entire developmental window we studied, i.e. at early, middle and late second trimester (SFig 2.5A).

We detected TFs with established roles in arealization, such as *NR2F1*, which confers positional identity across the rostro-caudal axis19, and *BCL11A*, which interacts with NR2F1 and was shown to repress motor identity in the cortex (Chan et al. 2013). Both genes are also implicated in neurodevelopmental disease (Bertacchi et al. 2019; Simon, Wiegreffe, and Britsch 2020). We also detected TFs that, to our knowledge, have not been previously described in the context of cortical arealization. In V1, these markers include members of the Nuclear Factor I family, *NFIA*, *NFIB* and *NFIX*. These genes are important regulators of brain development and have been implicated in developmental disorders including macrocephaly and severe cognitive impairment29. They also include *ZBTB18/RP58*, which encodes a transcription factor involved in neuron differentiation and cortical migration and a putative driver of brain expansion (Xiang et al. 2012). In the PFC, area-specific TFs we identified include members of the HMGB family, *HMGB2* and *HMGB3*, which are differentially expressed by neural stem cells at distinct stages of development (Abraham et al. 2013) and are thought to be key regulators of differentiation.

Of note, *HMGB3* mutations can result in severe microcephaly. We also found *NEUROG1* and *NEUROG2*, encoding the TFs Neurogenin 1 and 2, to be upregulated in PFC neurons. As in V1, while these genes have been found to be important regulators of neuronal differentiation, they have not been previously implicated or studied in the process of cortical arealization.

In addition to exploring differences in areal gene signatures at the cell type level, we investigated these signatures change throughout development. Consistent with proposed models of extensive transcriptional remodeling during the second and third trimesters (Jeremy Willsey et al. 2018), we observed that while area-specific gene signatures are composed of significant and specific marker genes, they also changed substantially throughout the early, middle and late second trimester (SFig 2.5 C). Concordantly, we only found a small overlap of area-specific gene signatures, and low cluster correspondence, between this dataset and that of the adult brain (SFig 2.6 A-C).

Our analysis indicates that after the second trimester, neurons in the neocortex remain largely immature, and that while the extreme rostral and caudal identities are already determined in early progenitor populations, further areal specification is refined at later time-points and may depend on sensory inputs to determine terminal identity. We thus find strong evidence for a partial early cortical protomap, which is then further refined throughout development as proposed by the protocortex model. Our single-cell data uncovers a large diversity of cell types and transcriptional profiles across six areas of the developing human cortex.

High throughput single-molecule RNA *in situ* hybridization reveals spatial remodeling across cortical areas

We selected candidate markers of excitatory neuron clusters that were enriched in one or more sampled areas for validation by multiplexed single-molecule fluorescent *in situ* hybridization (smFISH) (Fig 2.4A). We used the Rebus Esper spatial omics platform (Rebus Biosystems, Inc), a novel, automated system based on synthetic aperture optics (SAO) imaging. We quantified the expression level of 31 RNA transcripts per tissue section at four cortical regions from a GW20 sample (Fig 2.4B). This resulted in data from an additional 608,960 cells, enabling us to visualize marker gene expression levels, their spatial distributions, and co-expression patterns. We used DAPI staining along with kernel density expression (KDE) plots of canonical cell type marker genes *SOX2*, *SATB2*, and *BCL11B* to identify the ventricular zone and cortical plate (Fig 2.4C). Additionally, we confirmed the previously described areal pattern dynamics between the neuronal genes *SATB2* and *BCL11B*, which are co-expressed in frontal regions but mutually exclusive in occipital regions1 (Fig 2.4C). These spatial datasets are available for exploration and analysis at kriegsteinlab.ucsf.edu/datasets/arealization.

Across all areas, we explored novel candidate markers of subpopulations, including genes that are also subplate markers (*NEFL*, *SERPINI1*). For these KDE plots, we also plotted the canonical subplate marker *NR4A2*. All three markers could be found at roughly equal intensities across cells in the PFC, somatosensory, temporal and V1 cortex, but their spatial distribution relative to one another changed substantially (Fig 2.4D). These genes were co-expressed in the PFC but were mutually exclusive across all other regions. However, in the somatosensory cortex, we found these markers to be expressed in upper cortical layers rather than in the subplate. We similarly identified three frontally-

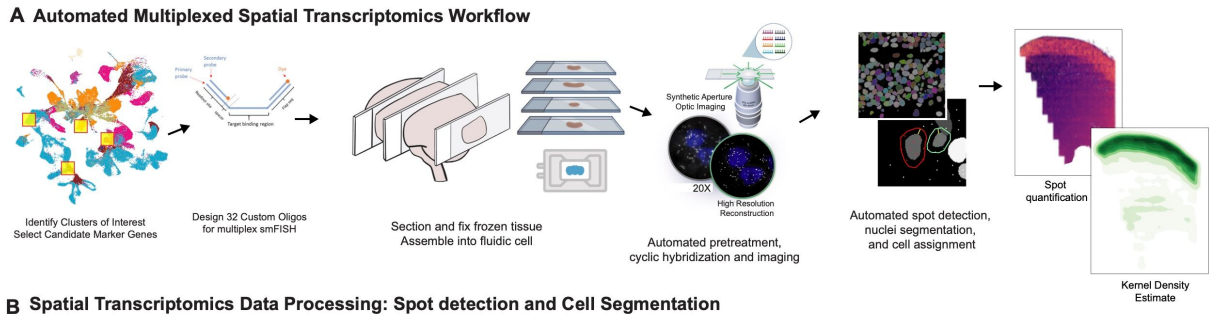
enriched marker genes, *PPP1R1B*, *CBLN2*, and *CPLX3*, whose quantification also revealed higher signal in the PFC and somatosensory tissue sections after normalization (Fig 2.4D). Caudally, we observed higher intensities of *LOH12CR12*, *ZFPL1*, and *PALMD* (Fig 2.4D).

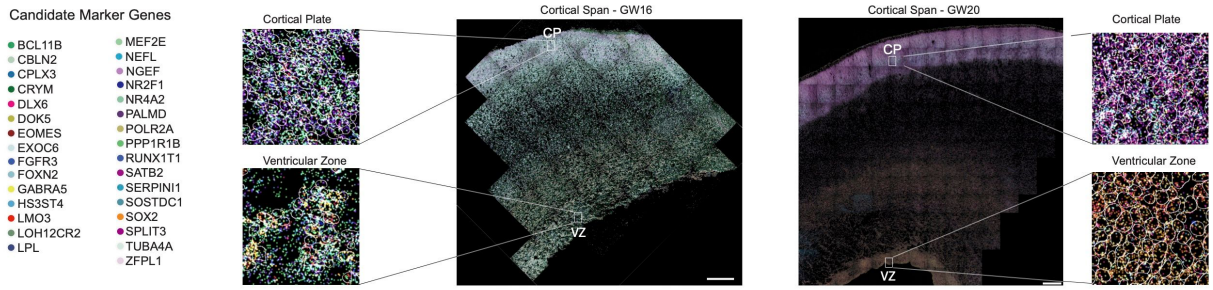
For both sets of markers, we found striking differences in the laminar distribution of gene expression, suggesting that not only are gene expression levels different across the cortical rostro-caudal axis, but that laminar cell type distribution might also change. While this observation may be reflective of differences in maturation states across the developing cortex, cell types may express genes in a different manner across distinct cortical areas. We leveraged cell segmentation aspects of our spatial transcriptomics analysis to investigate how the co-expression patterns of the genes queried in this experiment changed across areas. Transcripts detected with smFISH were automatically assigned to nuclei by proximity with the Rebus Esper imaging processing software.

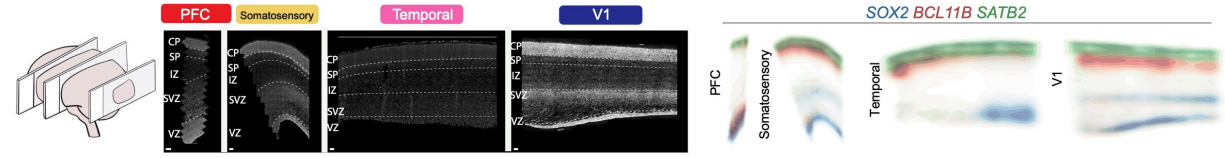
We calculated co-expression relationships between single cells to generate networks show the frequency of two genes expressed the same cell. The resulting networks highlight that the most stringent markers of areal identity are binary, i.e., they are either included or excluded from the gene network. In most cases, however, we found remodeled co-expression patterns across cortical areas rather than elimination or inclusion of a gene from the network. Even when using all 31 genes to construct the networks, we see substantial co-expression remodeling across cortical areas. We repeated this spatial analysis in a second individual (GW16) and validated some of the area specific aspects of gene expression, while also continuing to observe dynamic laminar redistribution across cortical areas. In sum, these data provide *in situ* evidence that

reinforces our observation from single-cell RNA seq analysis that there exist area-specific and regional-specific cell populations.

Some areal identity markers are mutually exclusive in their expression patterns across neocortical areas and even present a distinctive laminar distribution, as observed analyzing the set of 31 marker genes using smFISH. For example, subplate markers *NR4A2*, *NEFL* and *SERPINI1* show co-expression in prefrontal cortex, but differential laminar distribution in the three other regions examined (somatosensory, temporal and visual cortex) (see Figure 2.4 below). Overall, the networks of co-expressed genes substantially change throughout neocortical areas, and those strongly associated to specific area identities are mutually exclusive (not co-expressed in the same cell).


A Automated Multiplexed Spatial Transcriptomics Workflow
Identify Clusters of Interest
Select Candidate Marker Genes
Design 32 Custom Oligos
for multiplex smFISH
Section and fix frozen tissue
Assemble into fluidic cell
Synthetic Aperture
Optic Imaging
20X
High Resolution
Reconstruction
Automated pretreatment,
cyclic hybridization and imaging
Automated spot detection,
nuclei segmentation,
and cell assignment
Spot
quantification
Kernel Density
Estimate


B Spatial Transcriptomics Data Processing: Spot detection and Cell Segmentation
Candidate Marker Genes
BCL11B
CBLN2
CPLX3
CRYM
DLX6
DOK5
EOMES
EXOC6
FGFR3
FOXN2
GABRA5
HS3ST4
LMO3
LOH12CR2
LPL
MEF2E
NEFL
NGEF
NR2F1
NR4A2
PALMD
POLR2A
PPP1R1B
RUNX1T1
SATB2
SERPINI1
SOSTDC1
SOX2
SPLIT3
TUBA4A
ZFPL1
Cortical Plate
Ventricular Zone
Cortical Span - GW16
CP
VZ
Cortical Span - GW20
CP
VZ
Cortical Plate
Ventricular Zone


C Spatial Transcriptomics Validates Known Patterns of Arealization
PFC
Somatosensory
Temporal
V1
CP
SP
IZ
SVZ
VZ
SOX2 BCL11B SATB2
PFC
Somatosensory
Temporal
V1

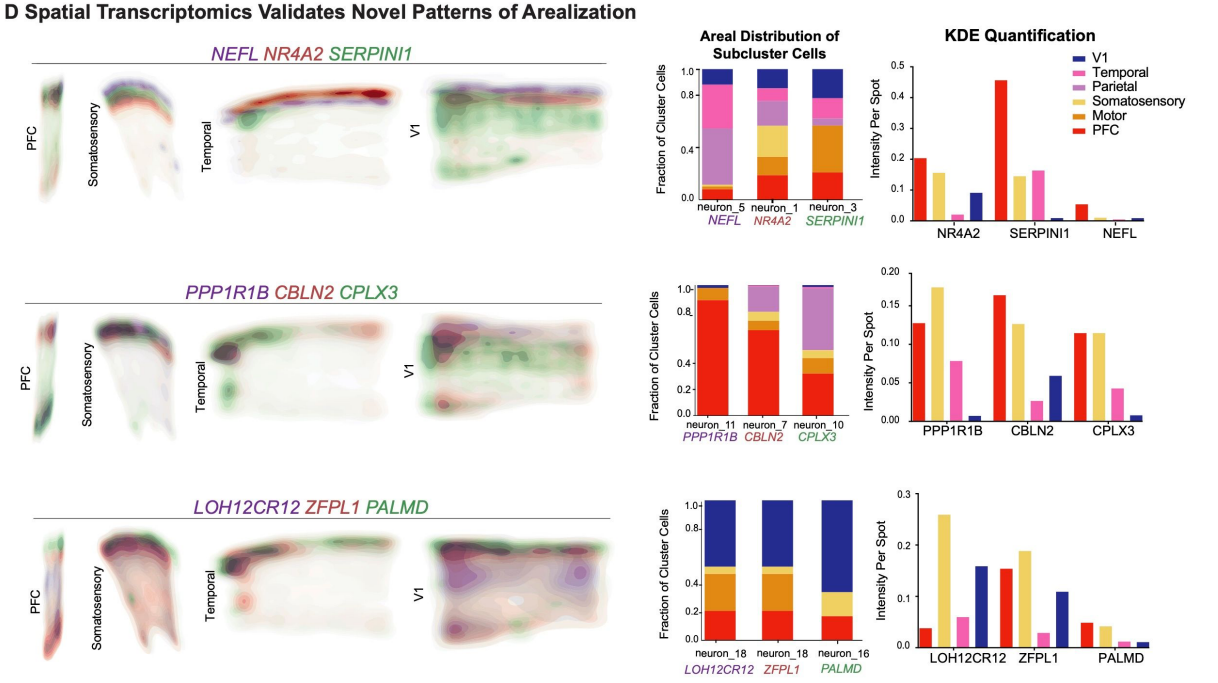

D Spatial Transcriptomics Validates Novel Patterns of Arealization
NEFL NR4A2 SERPINI1
PFC
Somatosensory
Temporal
V1
Areal Distribution of
Subcluster Cells
Fraction of Cluster Cells
neuron_5
NEFL
neuron_1
NR4A2
neuron_3
SERPINI1
KDE Quantification
Intensity Per Spot
V1
Temporal
Parietal
Somatosensory
Motor
PFC
NR4A2
SERPINI1
NEFL
PPP1R1B CBLN2 CPLX3
neuron_11
PPP1R1B
neuron_7
CBLN2
neuron_10
CPLX3
PPP1R1B
CBLN2
CPLX3
LOH12CR12 ZFPL1 PALMD
neuron_18
LOH12CR12
neuron_18
ZFPL1
neuron_16
PALMD
LOH12CR12
ZFPL1
PALMD

Figure 2.4

Discussion

In this study, we performed an in-depth characterization of the transcriptomic patterns and gene expression dynamics of the developing human brain during mid-gestation. We first analyzed molecular signatures of brain regionalization across ten major brain regions, and subsequently focused on the neocortex and its specification into distinct areas. Our results provide a granular understanding of the gene expression signatures of distinct cell types across neocortical areas, and at eight different timepoints throughout the second trimester of development.

We find that across major brain structures, regional identity is highly pervasive among distinct cell types. In contrast, areal identity in the neocortex is highly specific and restricted to individual cell types. Furthermore, we find that that in addition to cell type identity, the developmental stage of cells (i.e. gestational week) is a strong determinant of gene expression signature composition. Together, these observations suggest that the dynamics of area-specific gene expression signatures are surprisingly fast moving and cell type-specific (Fig 2.5 C, D). This is in contrast with previous models of areal patterning, where gene expression programs have been generally assumed to be persistent once established at a given region. We find strong evidence for the presence of a partial early cortical protomap between cell populations, including progenitors, at the frontal and occipital poles of the neocortex (Fig 2.5 A, B). We see evidence of transcriptional regulation programs that may prime more differentiated and mature cells to acquire either a rostral or caudal identity. For example, even though progenitor clusters in the neocortex show little molecular diversity in relation to the multiple cortical regions that will eventually emerge, we observe strong specification of PFC and V1 molecular identity among cells of this cell type.

In a previous study, we noted that radial glia were characterized by a small number of transcriptional differences that cascade into strong area-specific excitatory neurons (Nowakowski et al. 2017). Here, we present an analysis of nearly 400,000 cells compared to the 4,000 cells in the previous study. The analysis of a much greater number of cells and more cortical areas enabled us to uncover the strong difference between PFC and V1 radial glia, while confirming that glutamatergic neurons are even more distinct between cortical areas. Our data suggests that cells located in between the prefrontal and occipital poles are less specified towards a particular areal identity, an observation that is more consistent with the protocortex hypothesis. Together, these observations suggest that the expression gradients that establish early neocortical areal patterning may be propagated throughout development through transcriptional or epigenetic memory. Importantly, we find strikingly distinct spatial distributions of excitatory neuronal marker genes across areas *in situ,* as well as gene co-expression patterns unique to specific cortical areas. These differences may help explain the distinct morphological features of cortical areas, as well as the distinct connectivity patterns seen between areas and from different areas to other parts of the brain.

Characterizing the dynamic diversity of cell populations throughout the development of a structure as complex as the brain involves disentangling multiple axes of variation. The data we present here provides a spatially and temporally detailed molecular atlas of human brain and neocortex specification upon which future experimental characterizations can expand. These findings can additionally improve our understanding and development of in vitro models of cortical formation, including those used in the study of neurodevelopmental disorders.

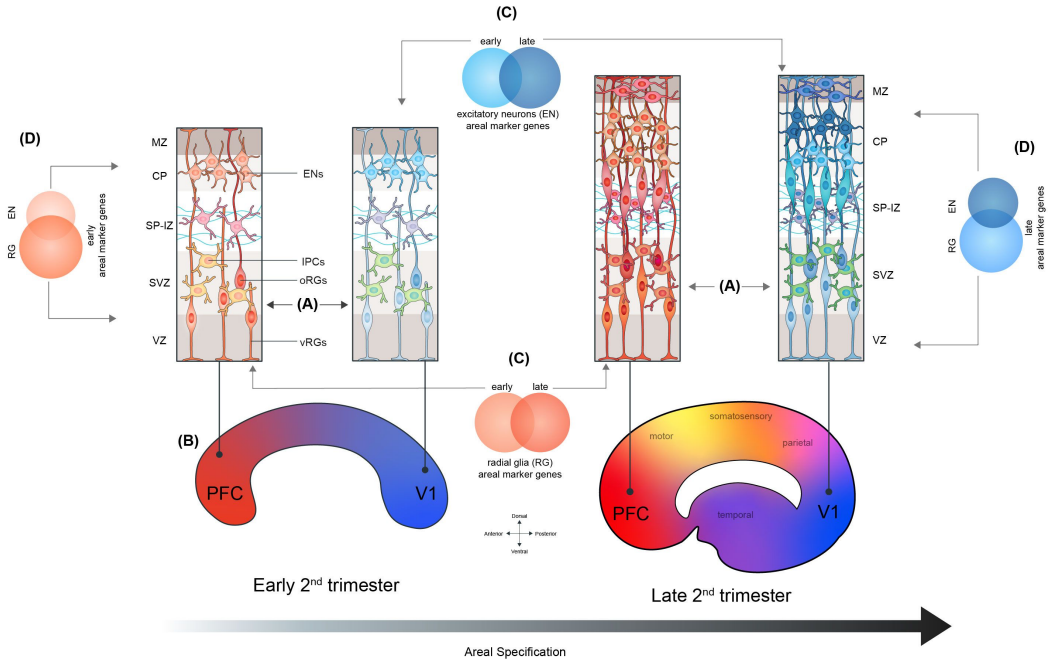
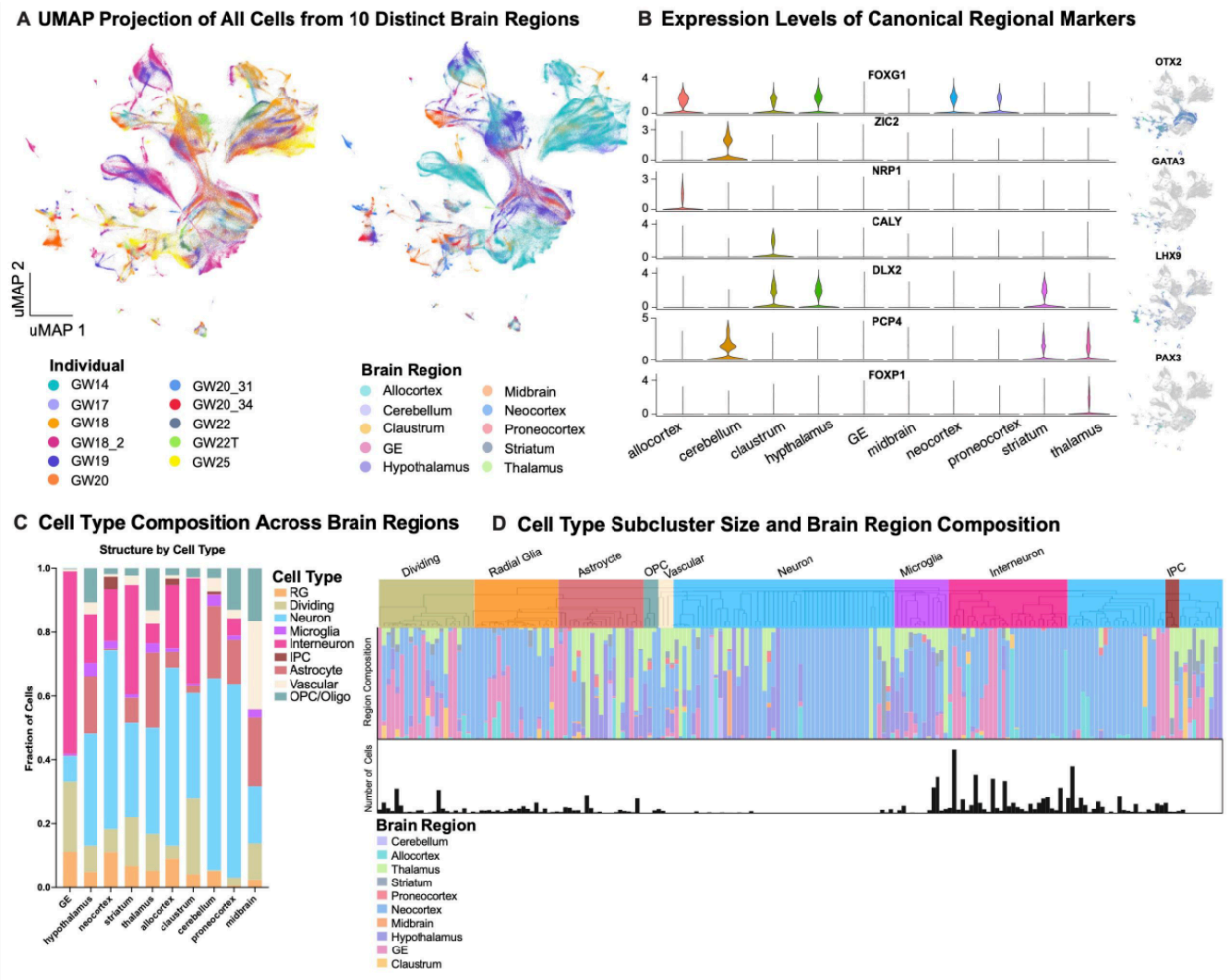
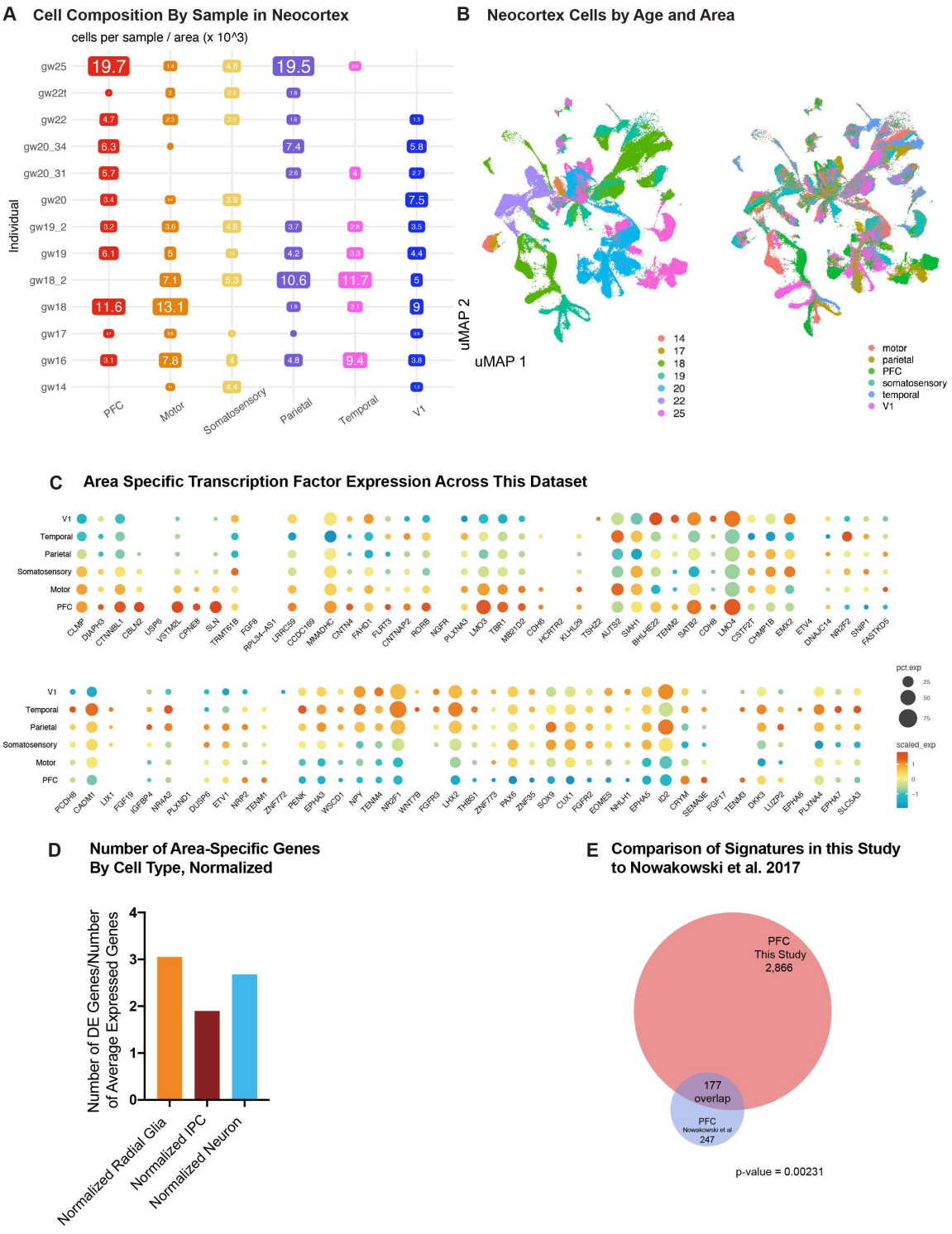

(C)
early
late
excitatory neurons (EN)
areal marker genes
(D)
EN
RG
early
areal marker genes
MZ
CP
SP-IZ
SVZ
VZ
ENs
IPCs
oRGs
vRGs
(A)
(B)
PFC
V1
(C)
early
late
radial glia (RG)
areal marker genes
Dorsal
Anterior
Posterior
Ventral
somatosensory
motor
parietal
temporal
PFC
V1
(A)
MZ
CP
SP-IZ
SVZ
VZ
(D)
EN
RG
late
areal marker genes
Early 2nd trimester
Late 2nd trimester
Areal Specification

Figure 2.5

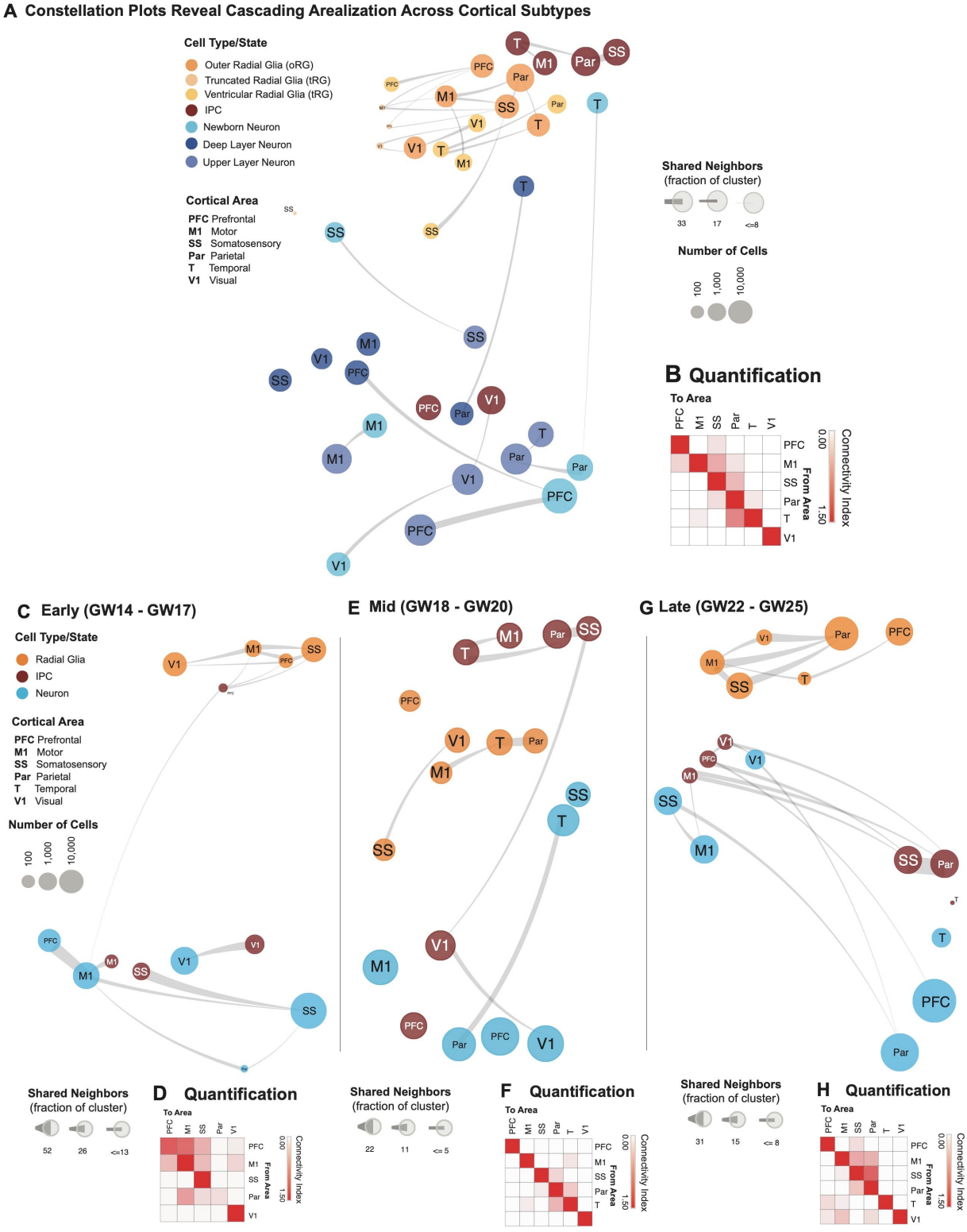
Supplementary Figures



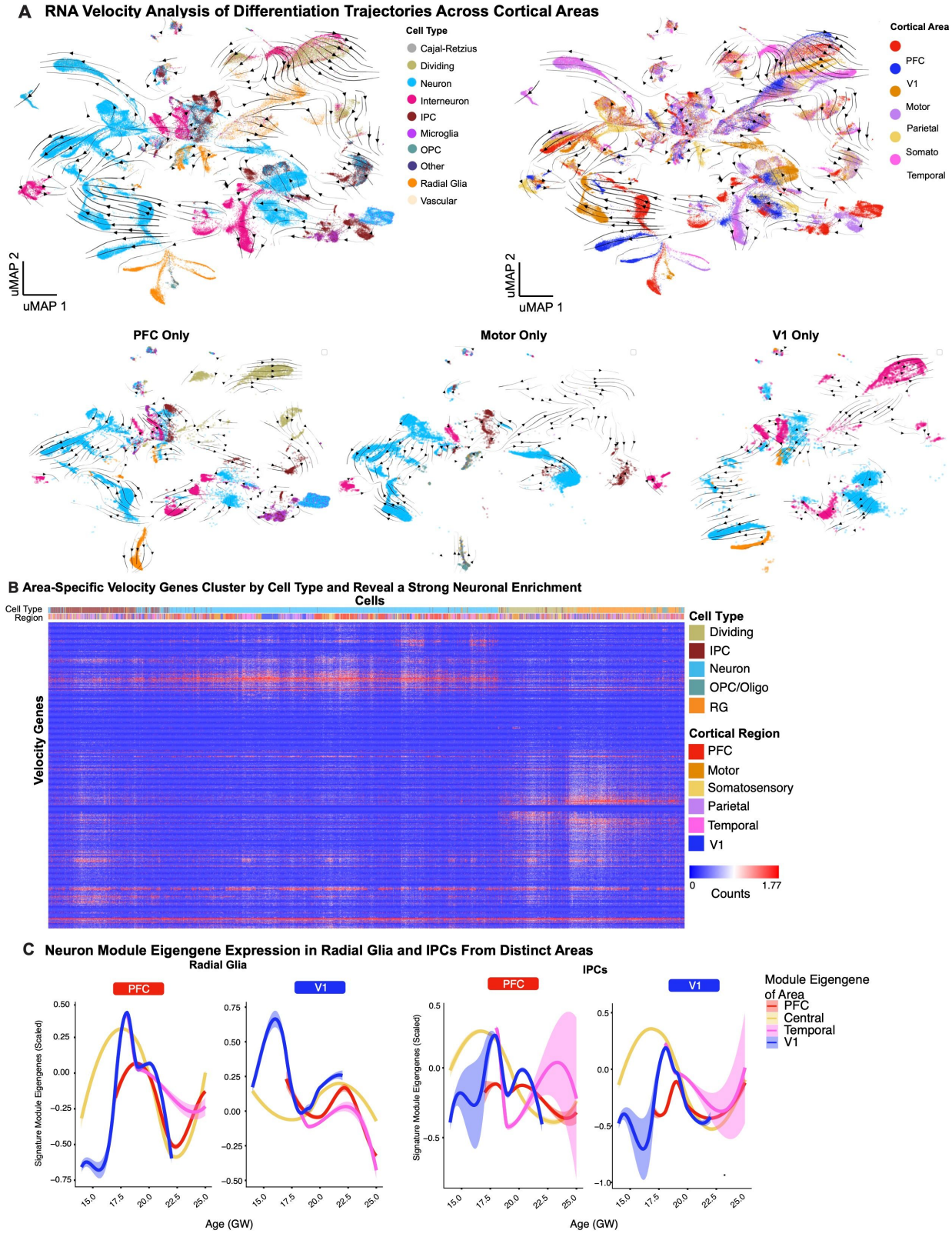
Supplementary Figure 2.1



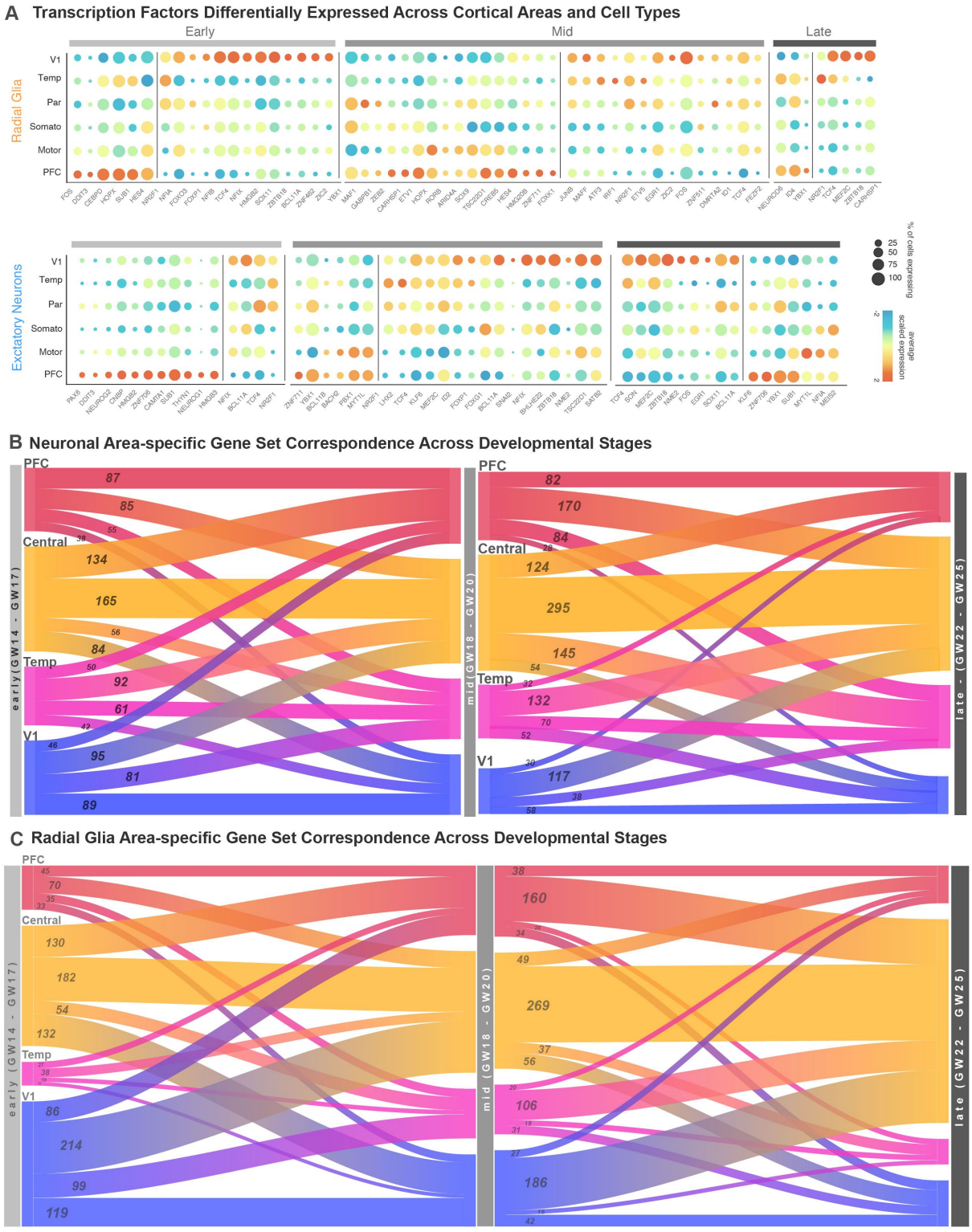
Supplementary Figure 2.2



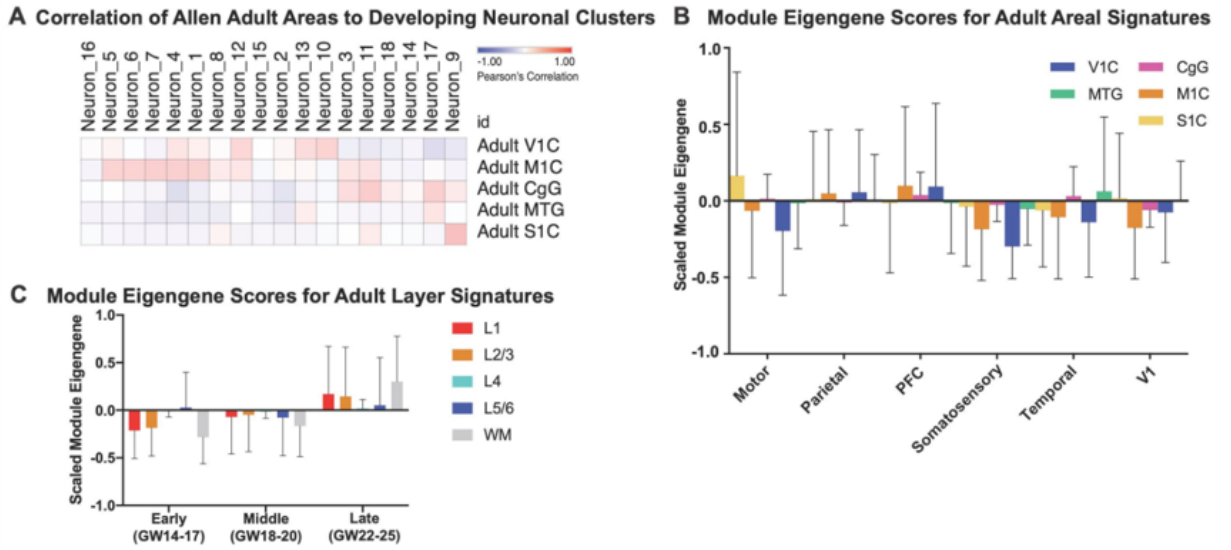
Supplementary Figure 2.3



Supplementary Figure 2.4



Supplementary Figure 2.5



Supplementary Figure 2.6

Methods

Sample Acquisition

De-identified 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, Embryo, and Stem Cell Research Committee (institutional review board) at the University of California, San Francisco. Two sets of samples included twins: GW20_31 and GW20_34; GW22 and GW22T.

Single-cell RNA Sequencing Capture and Processing

Brain dissections were performed under a stereoscope with regards to major sulci to identify cortical regions. Importantly, all dissections were performed by the same individual (T.J.N) to enable reproducibility and comparison between samples. Tissue was incubated in 4 ml of papain/DNAse solution (Worthington) for 20 min at 37C, after which

it was carefully triturated with a glass pipette, filtered through a 40um cell strainer and washed with HBSS. The GW22 and GW25 samples were additionally passed through an ovomucoid gradient (Worthington) in order to minimize myelin debris in the captures. The final single-cell suspension was loaded onto a droplet-based library prep platform Chromium (10X Genomics) according to the manufacturer's instructions. Version 2 was used for all samples except for GW19_2, GW16, and GW18_2 for which version 3 chemistry was utilized. cDNA libraries were quantified using an Agilent 2100 Bioanalyzer and sequenced with an Illumina NovaSeq S4.

Quality Control and Filtering

We filtered cells using highly stringent QC metrics. Briefly, we discarded potential doublets using the R package scrublet (Wolock, et al; PMID: 30954476) for each individual capture lane, then required at least 750 genes per cell and removed cells with high levels (>10%) of mitochondrial gene content. These strict metrics for quality control preserved no more than 40% of cells for downstream analysis, and re-analysis of the data for specific brain structures or cell types may benefit from less stringent QC for additional discovery. Our goal was to obtain clean populations with a high validation rate for a better understanding of arealization signatures. The resulting ~700,000 cells passing all thresholds were used in downstream analyses.

Clustering Strategy

We used a recursive clustering workflow to understand the cell types present in our dataset. In order to minimize potential batch effects and to increase detection sensitivity of potential rare cell populations, we performed Louvain-Jaccard clustering on each

individual sample first. After initial cell type classification, we sub-clustered all the cells belonging to a cell type to generate the most granular cell subtypes possible. We then correlated subtypes between individuals based upon the gene scores in all marker genes to bridge any batch effects, and iteratively combined clusters across all individuals and cell types. For this study, we combined the clusters within a single cell type across all individuals once, and again with all clusters from all individuals and cell types, resulting in two iterative combinations. The annotations at each step are preserved in the supplementary tables to enable reconstruction at any point in the pipeline. Hierarchical Clustering of Clusters Cluster hierarchies are generated from matrices correlating all clusters to one another using Pearson's correlation in the space of gene scores for all marker genes across all groups. Hierarchical clustering is performed within Morpheus (https://software.broadinstitute.org/morpheus) across all rows and columns using one minus the Pearson correlation for the distance metric.

Constellation plots

To visualize and quantify the global relationships and connectedness between cell types, cell type subclusters, or cell type-area groups, we implemented the constellation plots described in Tasic, 2018, by adapting the code made available at https://github.com/AllenInstitute/scrattch.hicat/. Briefly, we represented each group of cells as a node, whose size is proportional to the number of cells contained within it. Each node is positioned at the centroid of the UMAP coordinates of its cells. Edges represent relationships between nodes, and were calculated by obtaining the 15 nearest neighbors for each cell in PCA space (PCs 1:50), then determining, at each cluster, the fraction of neighbors belonging to a different cluster. An edge is drawn between 2 nodes if >5% of

nearest neighbors belong to the opposite cluster in at least one of them. An edge's width at the node is proportional to the fraction of nearest neighbors belonging to the opposite node, with the maximum fraction of out-of-node neighbors across all clusters represented as an edge width of 100% and equal to node width. The full code adaptation and implementation of this analysis is described in the function buildConstellationPlot in this paper's Github repository.

Quantification of Constellation Plots

Constellation plots were quantified by using a summary of the input values described above. For each cell type or area connection, the number of edges between two groups was multiplied by the average fraction of cells meeting the threshold for a connection within the group. This resulting number was called the connectivity index. Module Eigengene Calculations Module eigengenes were calculated for numerous gene sets using the the R package WGCNA (Langfelder and Horvath 2008). Scores were generated for each set of up to 10,000 randomly subsetted cells from the group using the function moduleEigengene function, Scores were calculated based on the intersection of the gene set of interest and genes expressed in the subset of cells. For the area-specific signatures, differential expression was performed as described above, and the gene signatures from late stage neurons across all areas were used to calculate module eigengenes for the radial glia and intermediate progenitor populations.

Area-specific markers / Gene Score Calculations

The expression profiles of cells from each subcluster or cortical area were compared to those of all other cells using the two-sided Wilcoxon rank-sum test for differential gene

expression implemented by the function FindAllMarkers in the R package Seurat and selected based on an adjusted p-value cutoff of 0.05. Adjusted pvalues were based on Bonferroni correction using all features in the dataset. We performed this step separately for each cell type and each individual, since we observed that gene specificity was highly dynamic throughout the developmental process. We then combined the individual gene lists of each cell type and area, and annotated the stage(s) at which each gene appeared to be specific. We binned individuals into three stages: early (GW14, 16 and 17), middle (GW18, 19, 20), and late (GW22, GW25). We ranked upregulated genes by specificity by calculating their gene score, which we defined as the result of a gene's average log fold-change x its enrichment ratio, in turn defined as the percentage of cells expressing the gene in the cluster of interest / percentage of cells expressing in the complement of all cells. Dot Plots used to visualize the expression of distinct marker genes across cell types and/or cortical areas were generated the custom function makeDotPlot available in our code repository, which makes use of the Seurat function DotPlot.

Briefly, for each gene, the average expression value of all non-zero cells from each group (cortical area) is scaled using the base R function scale(), yielding obtaining z-scores. Scaling is done in order to enable the visualization of genes across vastly different expression ranges on the same color scale. ich Transcription Factor Annotation Areally enriched marker genes obtained as described above were annotated against a comprehensive list of 1,632 human transcription factors described in36 and downloaded from the transcription factor database AnimalTFDB, available at http://bioinfo.life.hust.edu.cn/static/AnimalTFDB3/download/Homo_sapiens_TF.

Gene signature overlap and Sankey Diagrams

In order to quantify the degree of (dis)similarity of molecular signatures across distinct cell types, cortical areas, and/or developmental stages, we calculated the overlap between all sets of cell type and area-specific gene markers at each stage, and visualized these comparisons using Sankey diagrams using the function ggSankey from the ggvis R package. We then calculated the proportion of genes for each node shared with every other node, and clustered nodes hierarchically by calculating their euclidean distances based on their proportions of shared genes. The code used to construct the overlap matrices, create the plots and quantify the results is described in the functions buildSankey and buildHeatmap in our Github repository.

RNA Velocity

Velocity estimates were calculated using the Python 3 packages Velocyto v0.1720 and scVelo v0.2.221. Reads that passed quality control after clustering were used as input for the Velocyto command line implementation. The human expressed repeat annotation file was retrieved from the UCSC genome browser. The genome annotation file used was provided by CellRanger. The output loom files were merged and used in scVelo to estimate velocity. For the combined cortical analysis, cells underwent randomized subsampling (fraction=0.5), and were filtered based on the following parameters: minimum total counts = 200, minimum spliced counts = 20 and minimum unspliced counts = 10. The final processed object generated a new UMI count matrix of 18,970 genes across 195,775 cells, for which the velocity embedding was estimated using the stochastic model. The embedding was visualized using Uniform Manifold and Approximation and Projection of dimension reduction. The velocity genes were matched by cortical area and were

estimated using the rank velocity genes function in scVelo. Computational analysis of the transcriptomic data described in detail above were performed using R 4.038 and Python 3, the R packages Seurat (version 2 and version 3), googleVis, dplyr and ggplot2, the Python packages Velocyto v0.1720 and scVelo v0.2.221 as well as the custom-built R functions described. Our reproducible code is available in the Github repository associated with this manuscript.

Validation Marker Gene Selection

Marker genes for validation with the spatial omics platform were chosen first by identify useful cell type markers within the dataset. *SOX2* was chosen to mark radial glia, *EOMES* was chosen to mark IPCs, and *BCL11B* and *SATB2* were chosen to marker excitatory neuronal populations with previously validated changing co-expression patterns. *POLR2A* was used as a positive control for the technology. The remaining genes were selected based upon their status as a specific marker gene for excitatory neuron clusters of interest. Rebus Esper Spatial Omics Platform Samples for spatial transcriptomics were dissected from primary tissue as described above. Samples were flash frozen in OCT following the protocol described in the osmFISH protocol (Codeluppi et al. 2018). Samples were then mounted to APS coated coverslips, and fixed for 10 minutes in 4% PFA. Samples were then washed with PBS, and processed for spatial analysis. The spatially resolved, multiplexed in situ RNA detection and analysis was performed using the automated Rebus Esper spatial omics platform (Rebus Biosystems, Inc., Santa Clara, CA). The system integrates synthetic aperture optics (SAO) microscopy (Ryu et al. 2006), fluidics and image processing software and was used in conjunction with single-molecule fluorescence in situ hybridization (smFISH) chemistry. Individual transcripts from target

genes were automatically detected, counted, and assigned to individual cells, generating a Cell x Feature matrix that contains gene expression and spatial location data for each individual cell, as well as registered imaging data, as follows: Rebus Biosystems proprietary software was used to design primary target probes (22-96 oligos) and corresponding unique readout probes (assigned and labeled with Atto dyes) for each gene. The oligos were purchased from Integrated DNA Technologies and resuspended at 100μM in TE buffer. Coverslips (24 x 60 mm, No. 1.5, Cat. # 1152460, Azer Scientific) were functionalized as previously published43. Fresh frozen brain tissue sections (10 μm) were cut on a cryostat, mounted on the treated coverslips and fixed for 10 min with 4% paraformaldehyde (Alfa Aesar, Cat#43368) in PBS at room temperature, rinsed twice with PBS at room temperature and stored in 70% ethanol at 4°C before use. The sample section on the coverslip was assembled into a flow cell, which was then loaded onto the instrument. The hybridization cycles and imaging were done automatically under the instrumental control software. Briefly, primary probes for all target genes were initially hybridized for 6 hours and probes not specifically bound were washed away. Readout probes labeled with Atto532, Atto594 and Atto647N dyes for the first 3 genes were then hybridized, washed, counterstained with DAPI and then imaged with an Andor sCMOS camera (Zyla 4.2 Plus, Oxford Instruments) through 20xC, 0.45NA dry lens (CFI S Plan Fluor ELWD, Nikon) with 365nm LED for DAPI and 532nm, 595nm and 647nm lasers configured for SAO imaging. Multiple fields of view (FOVs) were imaged for each channel within the region of interest (ROI). Single Z-planes with 2.8μm depth of field were acquired for each field of view. After imaging, the first 3 readout probes were stripped and the readout probes for the next 3 genes were then hybridized, imaged, and stripped. This process was repeated until readout was completed for all genes. Using the Rebus Esper

image processing software, the raw images were reconstructed to generate high resolution images (equivalent or better than images obtained with a 100x oil immersion lens). RNA spots were automatically detected to generate high fidelity RNA spot tables containing xy positions and signal intensities. Nuclei segmentation software based on StarDist45 identified individual cells by finding nuclear boundaries from DAPI images. The detected RNA spots were then assigned to each cell using maximum distance thresholds. The resulting CellxFeature matrix contains gene counts per cell along with annotations for cell location and nuclear size. Kernel Density Estimation Plots Kernel density estimation plots were created from individual gene spot location maps retrieved from the spatial transcriptomics pipeline. They were created using the seaborn kdeplot function in Python with shading and cmap coloring. They were merged together for Figure 4 with Adobe Illustrator's overlay and darken features, using 50% opacity.

Spatial Co-expression Analysis

Using the cell by feature matrices, we eliminated all spots with less than 10 counts for signal. Pearson's correlations were then performed across the genes within each dataset and filtered for self-correlation. Positive control (*POL2RA*) and non-excitatory neuron cell type markers (*SOX2*, *EOMES*, *DLX6*) were removed from the analysis. Interactions of 0.05 or more were preserved and visualized with Cytoscape v3.8.2 using a force-directed ayout. Individual nodes were colored by their color in the merged image file in Figure 2.4B.

CHAPTER 3

Specification of Areal Identity in Human Stem Cell-Derived Cortical Organoids by Modulation of Fgf8 Signaling

SUMMARY

Despite the fact that the neocortex is organized into cytoarchitecturally and functionally distinct areas, the control of cortical area identity has not been extensively studied in pluripotent stem cell-derived models, including cortical organoids.

In this chapter, I discuss the current state of cortical organoid models, and focus on one major limitation of current cortical organoid protocols: the lack of anteroposterior patterning and the accompanying paucity of clear and reproducible areal identities. As many neurodevelopmental psychiatric disorders, including autism spectrum disorder (ASD) and schizophrenia (SCZ), appear to preferentially affect certain cortical areas, I place special emphasis on the importance of controlling areal identity in the context of neurodevelopmental psychiatric disease modelling. Finally, I describe recent efforts to further direct sub-lineage and spatial identities in organoids and a study I conducted in an effort to control the specification of human ES and iPSC-derived cortical organoids towards a homogeneous frontal or caudal neocortical identity.

Introduction

For decades, scientists have used model organisms, including fruit flies, frogs, and mice, as well as *in vitro* cell culture systems, to study developmental processes. These systems have indisputably provided fundamental insights into the biology of brain development. However, animal models and traditional stem cell culture assays do not fully recapitulate human brain development and studying these processes in humans is challenging for a variety of reasons: carrying out studies using human embryos poses ethical concerns, and studies on human primary tissue are not possible for most researchers due to a lack of tissue availability. As a result, we do not fully understand how cell specification and organization occurs in humans.

All of these factors underscore the critical need for better *in vitro* models for the study of human brain development–a need that organoids promise to address. Cerebral organoids are three-dimensional cell culture models of the developing brain that can be derived from embryonic stem cells (ES cells) or from induced pluripotent stem cells (iPSCs) reprogrammed from somatic cells. Organoids promise to close the gap between reductionist stem-cell-derived cortical monolayer cultures and non-human *in vivo* studies by providing a relevant species-specific background, as well as the complexity of cell organization in a 3D tissue context. Cortical organoids, specifically, represent a promising tool to study human-specific aspects of early cortical development in a reproducible and quantifiable manner. However, cortical organoid protocols are not without fault and have serious limitations that must be addressed. Many of these limitations are being extensively discussed, explored, and tackled by the scientific community, while others have largely been overlooked.

Limitations of Traditional Model Systems

Non-human animal models and traditional 2D stem cell culture, while instrumental to our current understanding of mammalian brain development, do not fully recapitulate the processes that give rise to the human brain. Studying these processes in humans is challenging for ethical and practical reasons. In this section, I aim to underscore the reasons why traditional model systems might not be sufficient to fill the gaps in our understanding of human cortical development, and how organoid models promise to address some of these shortcomings.

Mouse Models

Mice have been the model of choice for the study of mammalian brain development for a variety of historical reasons, most notably their genetic tractability (Goffinet & Rakic, 2000; Molnar & Price, 2016). Throughout development, both the mouse and human brain follow similar cellular processes, including cell proliferation and migration, and both species share a generally comparable timeline of developmental events. Additionally, mice and humans share comparable mechanisms of neuroplasticity. All of this has made the mouse an exceptional model for understanding fundamental principles of brain development and connectivity. Mouse models are invaluable for the study of brain development and have provided fundamental insights into the principles of this complex biological process. We have learned that in all mammals, brain development involves an intricately organized choreography of progenitor proliferation, neurogenesis, and migration of newborn neurons. There are however, as discussed below, significant differences between the brains of non-human animal models and humans, especially in

the neocortex, and not all observations can be transferred to our understanding of human brain development.

The human neocortex possesses a number of distinct genetic, molecular, cellular, and anatomical features, which simply cannot be studied using rodents or other animal models. Cognitive functioning in the human brain has undergone an accelerated evolution in comparison to the brains of other mammalian species, and there are underlying fundamental differences in the neural circuitry of the human cerebral cortex and its associated structures that distinguish us from the mouse, our currently favored model organism, and even from other primates. The mouse cortex is the result of its own evolutionary process, which favored a small body and a smooth, or lissencephalic, brain. Other large-scale differences with the human include, in the thalamus, an unlaminated lateral geniculate nucleus, and limited binocular vision, along with lateral-set eyes.

Humans have a much higher brain to body size ratio, a greater number of neurons, and a much higher degree of brain lateralization (Silver et al. 2019). In relation to the mouse, the human neocortex displays greatly enlarged supragranular layers and a significantly enlarged transitory subplate (Kostović, Išasegi, and Krsnik 2019). Outside of the cortex, other brain structures, including the cerebellum and higher-order thalamic nuclei, which mediate transthalamic cortico-cortical interactions, are massively enlarged in primates (Molnár et al. 2019). The primary mechanisms that have been proposed to contribute to primate cortical evolution are an increased proliferation of stem and progenitor cells, as well as the generation of outer radial glia cells (oRGs), which are only rarely found in rodents (Hansen et al. 2010). Additionally, human radial glia have greater mTOR activation compared to non-human primates (Pollen et al. 2019).

The most striking distinction between the neocortex of different mammals is the overall surface area of the cortical sheet, which can vary by a factor of 100 or more between species (Herculano-Houzel 2016). Compared to the mouse neocortex, the human cortex is over 1,000 times larger by area and by neuron number (Herculano-Houzel 2009), Even within the primate lineage. the human cerebral cortex underwent a massive, nearly three-fold expansion. This is one of the most salient features distinguishing humans from other great apes (Herculano-Houzel, 2012). This size difference is already apparent at mid-gestation, before neurogenesis is complete (Sakai et al., 2012).

The brains of most large mammals, including humans and some other primates, have a folded cortical surface, In fact, in the human, most of the cerebral cortex is not visible from the outside, but is buried in the sulci, and the insular cortex is completely hidden. In contrast, the brain of many other species, such as mice, rats[3] and new-world monkeys, are smooth, The surface convolutions of the cortex, sulci and gyri, appear during development and continue maturing after birth through the process of gyrification, providing a greater surface area in the confined volume of the skull. In addition to minimizing brain and cranial volume, folding is crucial for the extensive wiring of neurons, and ultimately for the functional organization of the cortex. For example, pathological alterations in the folding of the human cortex lead to severe intellectual disability (Walsh 1999) and intractable epilepsy (Barkovich et al. 2012).

Another striking difference between human cortex development and that of other mammals has to do with its timing. Human brain development is characterized by being much slower and protracted relative to other species. It is thought that this slower

[3] There are examples of rodents with gyrification, such as the capybara *(Hydrochoerus hydrochaeris)*.

development allows for the formation of relatively more cortical tissue, in particular of frontal cortex. This slower rate of development is not restricted to the longer gestational period of humans. Primates have a greatly extended period of postnatal development, including a more extended adolescence. Even between humans and other primate species, there is a large difference in the duration of developmental time frames, with human postnatal neocortex development extending longer than other primates, by close to a factor of four (Huttenlocher, 1994). It is thought that this protracted postnatal development in humans is results in more evident regional differences across the neocortex. Indeed, a recent report identified numerous genes and gene coexpression modules with temporally and/or spatially divergent expression patterns between humans and macaques (Zhu et al. 2018).

Although the basic architecture of the neocortex appears to be generally conserved in mammalian species, there are important differences in the cellular composition of the human cortex, including human-specific or human-enriched cell types and cellular features. One such example are the von Economo neurons–large, bipolar neurons located in layers 3 and 5 of the frontoinsular (FI) and anterior cingulate cortex (ACC) in great apes and humans, but not in other primates[4] (Nimchinsky et al. 1999). Interestingly, it has been proposed that specific aspects of VEN biology confer selective vulnerability to the ACC and FI in fronto-temporal dementia (FTD). It has also been reported that human cortical astrocytes are larger and more structurally complex and diverse than those of rodents (Colombo 1996; Oberheim et al. 2009). Compared to the mouse, the human brain

[4] Von Economo neurons are also present in other large brained social mammals including elephants (Hakeem et al. 2009) and several cetaceans, including the beluga whale and some dolphin species (Butti et al. 2009).

has a much greater number of neurons–recent reports estimate the human brain to be made up of about 86 billion neurons, and the mouse brain of about 70 million (Herculano-Houzel 2009; Landhuis 2020). The human brain also appears to have much greater neuronal diversity. Specifically, recent studies suggest that there has been a large increase in interneuron diversity in humans, and some interneuron populations are much more numerous in human prefrontal cortex compared with rodent model organisms (Lim et al. 2018). Using scRNA-seq to classify the cell types in the middle temporal gyrus of human neocortex, a recent study revealed 75 distinct cell types, including 6 non-neuronal, 24 excitatory and 45 inhibitory cells (Hodge et al. 2019). When comparing these human cell types to existing mouse single-cell RNA-sequencing datasets, the authors found that most, but not all, of the cell types identified in humans had corresponding cell types in mice. However, there were substantial differences in the transcriptional signatures between corresponding cell types. Highlighting the challenges posed by the use of mouse models for the study of neuropsychiatric disease, serotonin receptors, which have been implicated in a variety of neuropsychiatric disorders, were the second most divergent gene family between the two species.

Along with a massive expansion of the cortical sheet, the diversification of cortical architecture, connectivity, and functional specification across regions of the neocortex has been an important target of evolutionary adaptation in primates. This differential expansion of cortical areas and the emergence of new functional modules throughout evolution may have resulted from changes in cortical progenitors. It has been proposed that neuronal progenitors diversified in human to give rise to a larger variety of cortical neurons (Llorca and Marín 2020). This functional diversification is especially evident in humans, which possess numerous higher-order cortical functions. (Burkhalter, 2008;

Forbes & Grafman, 2010). The human cortex shows a wide divergence from the mouse in this respect, although the two species indeed share many similarities. The size and complexity of higher order association areas of the neocortex have increased substantially in primates. The human neocortex has a significantly larger association cortex and increased connectivity between cortical regions, as well as an enlarged thalamus compared to mouse (Molnar 2020). The associative frontal and parietal areas of the neocortex are thought to be unique to, or substantially more developed in primates. The frontal associative area, also known as the prefrontal cortex (PFC), is the largest of the two, extending throughout the anterior frontal lobe, and covering almost 80% of the entire frontal lobe. The PFC comprises one third of the total cortical surface (Teffer et al. 2013; Hladnik et al. 2014) and is considered to be the foremost substrate for "higher-order" cognitive functions in humans, including language, decision-making, working memory and social behavior (O'Rahilly and Müller 2008; Perani et al. 2011; Roth and Dicke 2012).

Several additional important differences between mouse and human in the mechanisms of cortical patterning and areal specifications have been elegantly demonstrated in recent studies (Alzu'bi et al. 2019; Hodge et al. 2019; Shibata et al. 2019; Zhu et al. 2018). In humans, the transcription factors SP8 and COUP-TFI, which are regarded as frontal and caudal-most TFs during development, respectively, in fact overlap extensively in the VZ of visual, auditory and somatosensory cortex, while in mouse, COUP-TFI and SP8 show very little overlap (Alzu'bi et al. 2017a). Adding to the differences in the developmental timeline of the human brain, it was recently shown that thalamocortical afferents (TCAs) invade the cortical subplate much earlier in primates than in rodent models (Alzu'bi et al. 2019). This has important implications for the role

of TCAs in cortical areal specification in humans. Specifically, in rodents, thalamocortical afferents (TCAs) innervate the subplate at the end of the neurogenic window. This suggests that intrinsic instructions, rather than thalamic input, have the largest role in establishing the identity of cortical neurons. However, Al'zubi et al found that TCA afferents occupy most of the human cortex as early as 12 PCW, suggesting that pioneer thalamic afferents may well have a role in determining cortical neuron characteristics by contributing to early cortical circuitry in humans.

Many aspects of mammalian brain development and function are yet unknown. Mouse models have been an essential tool in getting us to where we are now in our understanding of brain development, and they will undoubtedly remain a powerful model for brain development research in the future. However, as discussed, there are multiple features of brain development that are unique to humans or that have diverged significantly from other mammals along the human evolutionary lineage. Humans have developed unique ways of coordinating organ development, many of which underpin the etiology of psychiatric disorders, that simply cannot be fully explored in other species. This is where mouse models fall short in their ability to aid our understanding of brain development, and where human organoid models promise to fill the gap.

Primary Human Tissue

The study of human brain development traditionally relied on descriptive histological studies of postmortem specimens, rendering the potential to explore mechanistic processes very limited (Clowry et al. 2010; Molnar, 2011). An important system used to study the development of the human brain is primary fetal tissue. However, human tissue is scarce, and *ex vivo* slice cultures generally not viable for long-

term experimentation, while organoids have been shown to survive for more than a year *in vitro*. Organoid cultures allow the observation and manipulation of many aspects of human brain development with nearly unlimited temporal resolution and fewer ethical and practical concerns.

Developing *in vitro* systems to study human brain development

There is tremendous interest in modeling the impact of human-specific genetic changes on brain development (Pollard et al. 2006, Florio et al. 2016) and of gaining a better understanding of the human-specific mechanisms of brain development. Moreover, there is increasing interest in understanding how genetic mutations lead to cellular changes that underlie neurodevelopmental disorders in the human. Brain organoids offer enormous potential in this regard, particularly with respect to modeling neurodevelopmental disorders that have been difficult to model in rodents. Indeed, brain organoids were the first organoids derived from patient iPSCs and were used to model microcephaly, a developmental disorder of the brain (Lancaster et al. 2013). Brain organoids have also been used to study distinctive mechanisms of division of human neural stem cells (Iefremova et al. 2017; Bershteyn et al. 2017), as well as to identify potential cellular mechanisms underlying genetically-defined neurodevelopmental and neuropsychiatric disorders (Birey, Andersen, and Pașca 2017; Ilieva et al. 2018). For example, prior studies using mouse models had confirmed a role for the LIS1/NDEL1/14.3.3ε protein complex, which is deleted in patients with Miller-Dieker syndrome (MDS), in several cellular processes. However, they had failed to recapitulate the severity of the microcephalic phenotype seen in humans with the mutation (Yingling, Toyo-Oka, and Wynshaw-Boris 2003; Yingling et al. 2008). This phenotype was

successfully recapitulated in patient-specific organoid models of MDS, which in turn revealed the involvement of Wnt signaling in MDS and better defined the role of *LIS1* and *14.3.3ε* in maintaining the cortical niche (Iefremova et al. 2017). These findings illustrate the need to study the human brain more directly, through *in vitro* models that closely recapitulate the context of human brain development, in addition to using model organisms.

Lastly, by improving organoid models, we may be able to reduce our reliance on animals for research, reserving them for targeted and more complex studies that require intact circuit and whole-organism readouts.

Traditional 2D Cell Culture Models

Cell culture systems are an indispensable tool for a wide range of biomedical research fields. The traditional method of cell culture, two-dimensional (2D) cell culture, has been used since the early 20th century (Ferreira et al., 2018). Studies performed by Harrison on the development of frog nerve fibers *in vitro* in the early 1900s, followed by the establishment of aseptic technique and subculture methods by Carrell and Ebeling in the 1920s, and the successful isolation and maintenance of the first immortalized human cell line (HeLa cells) in the 1950s, all laid the groundwork for growing cells in an artificial environment and led to the emergence of the fields of cell and molecular biology [Taylor and Taylor, 2014]. While 2D cultures are simple and convenient, and indisputably play a vital role in research and drug development, they have a number of crucial limitations that ultimately result in a poor representation of tissue architecture *in vitro*.

2D cell culture systems are simplified representations of the *in vivo* environment, in which cells develop in a complex and dynamic three-dimensional (3D) context. In 2D cell

culture systems, cells grow as a monolayer on flat dishes usually made of plastic or glass. Cells are plated on coated surfaces, which enables them to adhere and spread. This is not representative of real cell microenvironments – *in vivo*, cells are surrounded by other cells and extracellular matrix in 3D space. 2D cell cultures exist in two dimensions, and this is an inaccurate representation of how cells grow and are affected by disease and injury, and how tissues develop. Without this tissue environment and architecture, physical signals for cell-cell and cell-matrix communication, which are required for a multitude of cellular activities, are absent. These include critical developmental processes like mitosis, self-renewal, and differentiation. Additionally, these physical limitations prevent cells from spontaneously organizing and spreading vertically, compelling them to flatten out and develop as monolayers (Fitzgerald et al., 2015). As a result, gene expression, protein synthesis, and cytoskeletal structure are changed in the 2D culture setting, leading to the loss of critical cell phenotypes, cell behaviors, and ultimately function (Fontoura et al. 2020; Birgersdotter, Sandberg, and Ernberg 2005). Additionally, the absence of oxygen and nutrient gradients in 2D cultures impairs cell response to physiological stimuli, further limiting basic cellular processes and cell-cell communication. While two-dimensional cultures can be highly heterogeneous, even when not intended to be (Buettner et al. 2015), they generally give rise to a less broad range of cell types than do 3D models (Duval et al. 2017), especially when the latter are cultured for protracted periods of time. This reduced cellular diversity limits the ability of 2D cultures to model more complex tissue structures. It is worth mentioning, however, cell diversity in organoids remains to be further enriched.

Ultimately, these limitations of 2D cell culture systems result in failure to truly understand cell activity in developing tissue, as well as healthy and diseased tissue.

The importance of tissue microenvironment and 3D culturing techniques was first proposed in the early 1980s for cancer research. Throughout the last few decades, the research community has sought alternative technologies that would enable the creation of *in vitro* models that more closely resemble the complexity of whole tissue (Kapałczyńska et al., 2018). Though 2D cell cultures continue to be used for most studies, 3D cell culture, and particularly organoids, have gained widespread adoption, particularly in cancer and stem cell research. Several factors contribute to the growing adoption of 3D cell culture. Compared to adherent cell culture, 3D cultures exhibit a greater degree of structural complexity, are able to survive for much longer periods of time, and better recreate the interaction between different cell types. 3D culture additionally improves the recapitulation of epithelial tissue, which serves as a barrier between organ compartments, interacts with the environment, and protects the organism from it. Correct functioning of epithelial tissue is critical for survival, and barrier dysfunction is implicated in a wide variety of illnesses. 3D cell cultures can be used to create an *in vivo*-like microenvironment that is a well-controlled and can be highly customized for each application (Chen, 2016, Koledova, 2017). In comparison to monolayer or neurosphere cultures, brain organoids more closely mimic the makeup, variety, and structure of cell types in the developing cortex. Cortical organoids show greater cell diversity (Nascimento et al. 2019), replicate more complex and dynamic cell-cell interactions (Jimenez-Palomares et al. 2021), grow to later stages than 2D cultures (Giandomenico et al. 2021), and can simulate brain disorders using patient-derived iPSCs (Bershteyn et al. 2017). In a broad sense, 3D cell culture enables cells to grow more realistically than do monolayer cell cultures.

The Current State of Human Cortical Organoids

Early studies in the field of stem cell neurobiology suggested that when cultured at low densities, ES cells give rise to sphere-forming neural stem cells, albeit at a relatively low efficiency. This autonomous differentiation in the absence of external inhibitory factors is consistent with the 'neural default' model. Pioneering work from the lab of Yoshiki Sasai developed organoid protocols with increased control and greater efficiency of cortical differentiation from ES cells. Watanabe, 2005 developed a suspension culture technique of 3D stem-cell aggregates termed "serum-free embryoid bodies (SFEB)". SFEB cultures derived from mouse ES cells cultured under these conditions already showed selective neural differentiation, but when treated with the Wnt antagonist Dkk1 and the Nodal antagonist LeftyA, neural differentiation rate was further enhanced. The authors noted that treatment with Wnt3a during the late culture period increased the proportion of pallial telencephalic cells (Pax6+), while treatment with Shh increased the proportion of basal telencephalic cells. This study showed that ES cells could be differentiated into "naive" telencephalic precursors, which could further acquire subregional identities in response to additional patterning signals. Shortly thereafter, in 2007, Sasai's group published an improved version of this protocol, showing that addition of a selective Rho-associated kinase (ROCK) inhibitor ROCK inhibitor, Y-27632, to human ES cells markedly diminished dissociation-induced apoptosis, even in serum-free suspension (SFEB) culture. Human ES cell cultured using this protocol formed floating aggregates and differentiated into BF1+ cortical and basal telencephalic progenitors, as SFEB-cultured mouse ES cells did in their earlier study. Together, these studies set the stage for further exploration of human cortical development in an in-depth study of the composition of human cortical organoids derived using a variation of the SFEB method.

This culture method, SFEBq, used low adhesion 96-well plates to induce reaggregation of dissociated cells more quickly and to more efficiently induce neural differentiation of uniformly sized stem cell aggregates. Under these conditions, a larger proportion of telencephalic (FOXG1+) progenitors was reproducibly generated than with the original SFEB method (~75% of total cells, compared to ~50%). Moreover, the majority of FOXG1+ cells generated coexpressed the dorsal telencephalon (cortical) marker EMX1. This modified quick-aggregation procedure was a big milestone in improving the reproducibility of cortical organoids. It was much more effective for inducing the differentiation of cortical cells than the original SFEB culture, which relied on slow reaggregation. Sasai's group reported that these organoids were apico-basally polarized, recapitulating an essential feature of *bona fide* cortical tissue during development. Specifically, they found that cortical organoids derived from mouse and human ESCs formed a self-organized structure with four distinct zones along the apico-basal direction: a ventricular zone, early and late cortical plates, and a Cajal-Retzius cell zone. Notably, the human ESC-derived organoids, while also showing this laminar organization, formed at a considerably slower rate than their mouse counterparts. Moreover, they displayed rudimentary separation of the cortical plate (CP) into distinct layers.

More than a decade later, human stem cell-derived cortical organoids have become widely adopted by the field of neurodevelopment. Organoids are a powerful tool, but they are not without limitations. There is a high degree of variability between cell lines, with some lines predominantly giving rise to off-target lineages, including hindbrain, choroid plexus, retina, and mesenchymal cells, while others show strong cell fate bias towards the cortical lineage (Nazareth et al. 2013). It is worth noting that this has historically been a widespread problem with pluripotent stem cell lines. For example, the human ES cell line

H9 has been known to be exceptionally efficient for neural induction (Lu et al. 2015). There is also a high degree of variability between organoids of the same cell line within a culture (intra-culture organoid-to-organoid variability), most notably with respect to organoid size and to relative proportions of cell types (Quadrato et al. 2017; Velasco et al. 2019), making reproducibility a challenge. Importantly, current protocols exploit Matrigel, an ill-defined and poorly-controlled substrate matrix, to support organoid growth that requires more advanced 3D spatial control and stimulation.

Perhaps one of the most important limitations of cortical organoids is an incomplete recapitulation of cell type diversity. Cortical development involves the emergence of diverse communities of cells, including excitatory neuron lineage cells (radial glia, intermediate progenitor cells, excitatory neurons), inhibitory neurons that migrate into the cortex from the ventral telencephalon, glial cells (astrocytes, oligodendrocytes, and microglia), and vascular cells. Organoid models to date capture a subset of these cell types, while also introducing additional sources of variation in cell types, cell states, and gene expression (Quadrato and Arlotta 2017). Interestingly, however, the signature of differentiation from radial glia to excitatory neurons seems to be highly correlated between primary cells and organoid models, with very few genes deviating from the *in vivo* trajectory (Camp et al. 2015). It has been reported that organoids ectopically activate cellular stress pathways, which in turn is thought to impair full cell-type specification (Bhaduri et al. 2020). It is worth noting, however, that stress pathway genes are also expressed *in vivo*, and although they do seem to be expressed at higher levels *in vitro* than *in vivo*, a recent study found that they show a flat trajectory over time, suggestive of a homeostatic state, as opposed to progressive stress or dysfunction (Gordon et al. 2021).

Cortical organoids do not show surface folding, although it has been reported that *PTEN* mutant human, but not mouse, cerebral organoids display some surface folding and contain expanded ventricular and outer radial glia? progenitors. Organoids also seem to incompletely recapitulate the epigenomic landscape of the developing cortex. The chromatin landscape of cerebral organoids was reported to have broad cell type-specific enhancer accessibility patterns similar to the developing cortex, but lacked many putative regulatory elements identified in homologous primary cell types (Ziffra et al. 2020). This would suggest a potential crucial limitation for cortical organoids to test epigenetic regulation hypotheses, as many non-coding regulatory elements, in particular distal enhancers, may not be recapitulated in these models. Lack of vascularization is generally also an issue with organoids, resulting in limited oxygen and nutrient supply, especially to their innermost regions. The lack of vasculature also limits our ability to investigate processes that rely on neurovascular interactions or barrier functions. Because of limitations in nutrient supply, organoids have a limited growth potential, which can also affect their maturation (Cakir et al. 2019). Recently, efforts have been made to implement vascular structure in human brain organoids. For example, ETV2-induced reprogramming of endothelial in organoids was shown to generate vascularized cortical organoids with a functional vasculature-like network (Cakir et al. 2019). These vascularized organoids acquired several blood-brain barrier characteristics, and showed enhanced functional maturation.

Finally, current protocols for cortical organoid differentiation do not generate spatially organized tissue. While it is true that organoids are largely self-organizing structures that exhibit remarkably similar tissue architecture to their *in vivo* counterparts, they are highly heterogeneous: each organoid is unique and exhibits relative

positioning of tissue regions that are often random, possibly because of a lack of embryonic axis formation. In fact, in cerebral organoids, which give rise to various brain regions, each region individually develops quite similarly to those *in vivo*, but regions are not reliably organized relative to one another (Lancaster et al. 2013) due to the lack of anterior-posterior and dorsal-ventral axes. Although some molecular signatures of cortical areas do emerge in organoid neurons, they are not spatially segregated (Bhaduri et al. 2020), adding to the random organization of organoids. This heterogeneity also makes it difficult to generate pure populations of single cell types.

Cortical organoids derived from current differentiation protocols are promising systems to study human-specific cortical development, but further optimizations of culture conditions are required to improve the diversity and fidelity of subtype-identities, reduce metabolic stress and recapitulate *in vivo* patterns of neuronal migration and maturation. There is a great need to create improved 3D culturing protocols that create spatially organized brain organoids and include morphogen gradients that establish a dorso-ventral, medial-lateral, and anterior-posterior axis (Armstrong & Stevens, 2019). Ideally, organoid models should also reflect the asymmetries of the developing brain, such as the topography of those gradients. Future improvements to organoid protocols should show increased reproducibility of cell composition and better control of subregional identities in space.

Sub-Lineage Specification in Current Cortical Organoid Protocols

A handful of studies (Kadoshima et al. 2013; Eiraku et al. 2008) have shown that the regional identity of pallial cortical organoids could be further selectively controlled into olfactory bulb (the rostral-most part of the cortex) and rostral and caudal cortex by introducing additional patterning factors, including Fgf, Wnt or BMP, and by treating FOXG1+ aggregates with Fgf8, an inducer of rostral cortical regions (Shimogori and Grove, 2005), beginning on day 7. The addition of Fgf8 strongly suppressed COUP-TF1 expression. Conversely, when aggregates are treated with the Fgf inhibitor FGFR3-Fc, this resulted in a large increase in COUP-TF1+ cells (REF). These findings provided evidence that Fgf signaling could induce rostralization of early cortical tissue *in vitro*, as it does *in vivo*. Together, the results from these studies showed that spatial and temporal aspects of early corticogenesis were indeed recapitulated and could be further controlled in cortical organoids derived from mouse and human ES cells.

In a later study, the Sasai group took perhaps the most detailed look at the capacity of cortical organoids to acquire more specific regional identities, specifically rostral and caudal neocortical identities. This time, using only human ES-derived cortical organoids, the authors found that organoids frequently acquired an axial polarity along the dorsocaudal and ventrodorsal axes and underwent what they described as “region-specific rolling morphogenesis”. Specifically, expression of COUP-TF1, which is expressed in a dorsocaudal-to-rostroventral gradient in the fetal brain, was stronger on one side of the cortical neuroepithelium, whereas the ventrorostral marker Sp8 was expressed in the opposite pattern. The authors observed cortical hem-like tissue (OTX2+) flanking the region of strong dorsocaudal marker expression (COUPTFI, LHX2) and a higher level of pERK opposite to the region with high COUPTFI expression. These findings suggested,

unexpectedly, that a wide range of self-organizing events are driven by intrinsic programs during early human neocortex development. Moreover, they seemed to indicate that hESC-derived cortical tissue spontaneously acquires an "intracortical dorsocaudal-ventrorostral polarity". In this study, they again examined the effects of adding exogenous Fgf8 to the culture media. Treatment with Fgf8 suppressed COUP-TFI expression and markedly expanded the expression domain of Sp8. Of note, the authors detected the presence of outer radial glia, which are abundantly present in the human developing neocortex, but not in the mouse.

Further studies from the Sasai lab focused on the subregional specification of other cortical domains (ventral and dorsomedial telencephalon) through the addition of extrinsic signals. In a 2011 study, Danjo & Eiraku systematically recapitulated patterning of the ventral telencephalon, or subpallium, using mouse ES cell-derived organoids (Danjo et al. 2011). Here, they showed that the dorsoventral specification of ESC-derived telencephalon can be directed by Shh in a dose dependent manner, with high concentrations of Shh giving rise to subpallial tissue. This region in the developing brain becomes subdivided into three: the lateral, medial and ganglionic eminences (LGE, MGE and CGE, respectively). Django & Eiraku showed that early Shh treatment, even before the onset of FOXG1 expression, is critical to specify LGE progenitors, whereas high doses of the hedgehog agonist SAG directed cells towards an MGE and CGE identity. Furthermore, they found that in addition to Hedgehog signaling, Fgf8 and Fgf15/19 play distinct and opposing roles in the specification of the MGE and CGE. In a 2015 study, Sasai's group investigated the sub-regional specification of dorsomedial telencephalon, which includes the medial pallium (Sakaguchi et al. 2015). Using Wnt and BMP signaling, they induced choroid plexus, the most dorsomedial portion of the telencephalon, from

human ES cells. Then, by titrating BMP and Wnt, they reported self-organization of medial pallium tissues, including the hippocampus, and following long-term dissociation culture, they derived functional hippocampal granule and pyramidal neurons (ZBTB20+/PROX1+ and ZBTB20þ/KA1+, respectively).

Finally, (Motono et al. 2016) revisited the idea of controlling subregional identity in neocortical (dorsal telencephalon) organoids. Using a small molecule Wnt inhibitor, C59, which was found to inhibit both canonical and non-canonical Wnt signaling, the authors reported the derivation of cortical motor neurons.

Novel Methods to Control Spatial Identity in Organoids

The development of organoid differentiation protocols that include further spatial specification, including polarized or spatially organized organoid models, is crucial for the study of human neurodevelopment (Arlotta & Pasca, 2019). New methods for systematically controlling the generation and organization of distinct cell types would dramatically improve tissue modeling and developmental research. Recent efforts to improve cortical organoid models have proposed the adoption of new technologies to reproduce specific patterns of specification *in vitro*.

In this section, I review several recent approaches to control subregional specificity in cerebral organoids. While most of these approaches have not been used directly to control areal identity in cortical tissue, they could in principle be applied to this specific problem.

Scaffold Bioengineering

Biomaterials are increasingly being adopted for stem cell research and have been shown to enhance 3D stem cell culture, including neural organoids. Several studies have shown the application of biomaterials towards high-resolution patterning and organization of cellular architecture in organoids. Biomaterial scaffolds are composed of natural or synthetic matrix molecules that mimic characteristics of the extracellular matrix, and several types of biomaterials have been shown to influence the survival and function of developing stem cells. This scaffolding can serve a variety of functions, including guiding the structural organization and patterning of cells by providing both molecular and mechanical cues, creating attachment points for cells to promote anchorage-dependent survival, and preventing cells from being washed away. Biomaterial scaffolds can also

improve the recapitulation of cell-cell interactions, matrix properties, and signaling gradients that occur during development (McMurtrey 2016; Tan et al. 2015).

Novel designs have been described for functional biomaterial constructs that guide tissue development towards targeted regional identities. These designs typically comprise compartmentalized regions of patterning factors that form concentration gradients throughout the entire construct established according to classical diffusion models. Ultimately, the gradients created by these compartmentalized designs can establish axis patterning, including rostral/caudal, ventral/dorsal, or medial/lateral identities of the developing brain. Signaling factors, which can come in the form of differentiation cues, patterning factors, growth factors, and axon guidance molecules) are loaded into specific compartments of the biomaterial scaffold at desired initial concentrations, and these then diffuse to give rise to specific concentration gradients throughout the construct.

Organizer Cells

Another strategy to generate spatial organization within organoids was recently developed by the Studer group. This approach entails embedding morphogen-secreting cells at one of the poles of an aggregate to simulate *in vivo* organizers (Cederquist et al. 2019).

Morphogen-Soaked Beads

Another recent study used morphogen-soaked beads to influence spatial identities within hESC-derived brain organoids (Ben-Reuven and Reiner 2020). Using a combination of morphogens and synthetic molecules, beads were embedded near positionally-fixed organoids. These beads were effectively interpreted as local organizers, and key

transcription factor expression levels changed as a function of the distance of cells from the bead, akin to distance from a morphogen source during normal *in vivo* development. Using this strategy, the authors showed that titration of the WNT agonist CHIR99201 and of BMP4 directed the expression of telencephalon and medial pallium genes; dorsal and ventral midbrain markers; and isthmus-related genes. The highest concentration tested resulted in the expression of markers involved in patterning of the mesencephalon: *FGF8*, *WNT1*, *EN1/2,* and *PAX2* were detected, all of which are required for midbrain and cerebellar development.

Much work remains to be done on the appropriate combinations of biomaterials, scaffolding designs, signaling molecules, and genome engineering required to effectively recreate cortical area specification *in vitro*. The ability to recapitulate these innate neurodevelopmental processes and three-dimensional organization in a spatially controlled manner with finely tuned gradients and localized concentrations of signaling factors is crucial for directed specification and accurate study of tissue development. Moreover, this has critical implications for the study of neurodevelopmental disorders, as well as the targeted regeneration of functional neural tissue.

Specification of Areal Identity in Human Stem Cell-Derived Cortical Organoids by Modulation of Fgf8 Signaling

Embryonic tissues *in vivo* are exposed to gradients of developmental signals and spatial cues, i.e. morphogens, which typically need to be introduced *in vitro*. Gradients of several morphogens, including FGFs, Wnts, and BMPs, determine areal identity during neural development, with Fgf8 the most well studied as a central regulator of cortical area patterning. Fgf8 is secreted from the anterior neural ridge (ANR), located in the rostral-most part of the neural tube, establishing a gradient along the rostro-caudal axis and modulating the expression of transcription factors that further establish rostro-caudal expression gradients.

With these factors in mind, I designed and performed a study aimed at determining whether areal patterning can be better controlled in organoids. In particular, my goal was to develop a system to control areal identity along the rostro-caudal axis of pluripotent stem cell-derived cortical organoids by modulating fibroblast growth factor 8 (Fgf8) signaling (Figure 3.1). As described in Chapter 3, it was previously reported that Fgf8 treatment changes the expression of several markers with rostro-caudal expression profiles in both mouse and human pluripotent stem cell (PSC)-derived cortical organoids (Eiraku et al., 2008; Kadoshima et al., 2013). However, these studies only examined the expression of three markers as a proxy for cortical area identity.

Inspired by prior work performed in iPSC-derived neurospheres (Imaizumi et al. 2018), the main objective of this study was to establish a proof of principle system for selective areal identity in cortical organoids. I hypothesized that the rostro-caudal identity of progenitors and newborn neurons in PSC-derived cortical organoids could be controlled by treating cells with exogenous Fgf8, or with exogenous FgfR3, which

sequesters existing Fgf8. Specifically, I hypothesized that addition of Fgf8 would direct cortical organoids toward a more frontal-like identity, whereas FgfR3 would direct them towards a more caudal-like one. What follows is a description of the study design and methods, as well as preliminary results from bulk RNA sequencing conducted on treated and untreated cortical organoids. Further work should be build on these observations and devise an informed, improved follow-up study based on these results.

Materials and Methods

Pluripotent Stem Cell Culture and Expansion

Two pluripotent (PSC) stem cell lines were used: human embryonic stem cell line H1 (WiCell) and iPSC line 13234 (Conklin Laboratory, Gladstone Institutes), which have been shown by our lab and others to reliably differentiate into dorsal telencephalon at high rates (Kadoshima et al. 2013). Stem cell expansion was performed as described previously (Bhaduri et al. 2020). Briefly, cells were expanded on plastic six-well plates coated with growth factor reduced Matrigel (Corning, Cat. 354230). Stem cells were thawed in StemFlex feeder-free medium (Gibco, Cat. A3349401) containing 10 μM Rock inhibitor (Y-27632). Culture media was changed every other day and lines were passaged when colonies reached about 70% confluency. Stem cells were passaged using ReLesR (Stem Cell Technologies, Cat. 05872) following the vendor's protocol. All lines used for this study were between passage 22 and 40.

Cortical organoid differentiation protocol

Cortical organoids were generated using a minimally directed differentiation protocol described in (Kadoshima et al. 2013). In brief, PSC lines were expanded and dissociated to single cells using Accutase. After dissociation, cells were reconstituted in neural induction medium at 10,000 cells per well in a 96-well v-bottom low-adhesion plate.

The neural induction medium from days 0-18 was GMEM-based and included 20% knockout serum replacement (KSR), 1× non-essential amino acids (NEAA), 0.11 mg/ml sodium pyruvate, 1× penicillin–streptomycin, 0.1 mM β-mercaptoethanol, 5 μM SB431542 (TGF-β inhibitor) and 3 μM IWR1-endo (Wnt pathway inhibitor). Medium was supplemented with 20 μM Rock inhibitor (Y-27632) for the first 6 days to enhance cell

survival. After 18 days, the medium was changed to DMEM/F12 medium containing 1× Glutamax, 1× N2, 1× CD lipid concentrate and 1× penicillin–streptomycin, and organoids were transferred from 96-well to 6-well low-adhesion plates and placed on an orbital shaker rotating at 90 rpm. After 35 days, organoids were moved into DMEM/F12-based medium containing 10% FBS, 5 µg/ml heparin, 1× N2, 1× CD lipid concentrate and 0.5% matrigel (BD). Throughout the duration of the experiment, organoids were fed every other day. Organoids were collected for immunohistochemistry and RNA extraction after 6 or 8 weeks.

Fgf8 or FgfR3 Treatment

Fgf8 signaling was modulated by adding recombinant Fgf8 protein (50ng/ml) or soluble Ffg receptor 3 (FgfR3) (50ng/ml), which sequesters endogenous Fgf8, directly to the culture media every other day. With Day 0 as the day of seeding, Fgf8 treatment was initiated at week 3 or week 5 and continued for 1, 2 or 4 weeks. Organoids were collected at weeks 6 and week 8 for RNA extraction and immunofluorescence staining (Figure 3.1).

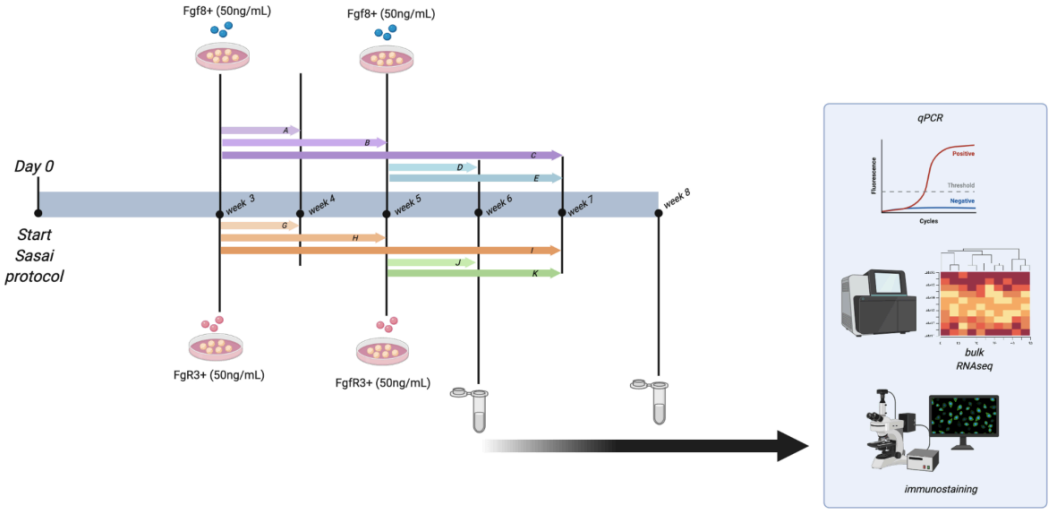


Figure 3.1 Schematic of the study design. Fgf8 or FgfR3 are added on week 3 or week 5 of organoid culture. Treatment is continued for 1, 2 or 4 weeks and samples are collected at week 5 and week 7.

Results

Using bulk RNA sequencing to profile the effects of exogenous Fgf8 addition on gene expression revealed that each treatment group showed a highly distinct differential gene expression signature (Figure 3.2 A, C), with gene expression changes exclusive to each treatment condition and some shared across treatment conditions. This suggests that gene expression changes are highly dependent on treatment duration and starting timepoint. A subset of genes were strongly upregulated in all groups, whether treatment was started at week 3 (duration of 1, 2 or 4 weeks) or week 5 (duration of 2 weeks) (Figure 3.2b). These included the interneuron marker genes *GAD1*, *GAD2*, *DLX1*, *DLX2* , *DLX5* and *DLX6* (Pla et al. 2018; Wang et al. 2010). They also included *SP8*, which encodes a transcription factor that displays a high rostral/low caudal expression gradient that induces and is in turn induced by Fgf8 (Sahara et al. 2007).

Conversely, the expression of many genes was only significantly affected when Fgf8 treatment was started either at week 3 or at week 5, respectively, including *MAPK6*, *NOTCH1*, *SOX9*, *ID4* (upregulated, week 3 start), *HES6* (upregulated, week 5 start) and *SLC7A11*, *SLC39A7*, and *SLC1A5* (downregulated, week 5 start). We and others have noted that newborn neurons start to become abundant in cortical organoids at around week 5. It is possible that delaying exposure to exogenous Fgf8 results in a distinct cell type population pool that sees the effects of the treatment. For example, if treatment with Fgf8 increased the pool of progenitors or decreased the rate of differentiation of organoid cells, organoids that have been treated for 2 weeks at week 5 would be composed of a larger proportion of progenitors than their naïve counterparts, which would comprise a larger proportion of neurons. The gene expression outcome of these two groups with distinct histories upon treatment with the same factor at week 5 would then be likely very

different. The majority of genes, however, appeared to be most prominently up or downregulated in a single treatment condition. (Fig 3.2 A, C).

RNA sequencing revealed that addition of exogenous Fgf8 has a wide range of effects on the expression of known cortical areal markers, which are at least partially dependent on the duration of treatment. Treatment with Fgf8 resulted in markedly decreased expression of the caudal marker *COUPTF1* (*NR2F1*), and the degree of expression reduction was correlated with extended duration of treatment (Figure 3.2 D). Fgf8-treated organoids exposed to the small molecule at 3 weeks for 1 week showed decreased *COUPTF1* expression relative to control. Organoids treated from week 3 to week 5 showed even lower *COUPTF1* expression, and organoids treated from week 3 – 7 showed the lowest *COUPTF1* expression of all 3 treatment groups. However, organoids treated starting at week 5 for 2 weeks (wk5-7) did not display *COUPTF1* expression levels as low as their week 3-7 counterparts, suggesting that the low expression level seen in the latter group was at least partly dependent on the longer exposure to Fgf8, or the earlier timepoint at which exposure started, or both. A strikingly similar expression pattern was seen for FgfR3, another caudal marker gene, but also for the rostral markers AUTS2.

Interestingly, the caudal marker *EMX2* showed a different downregulation pattern: its expression initially decreased with Fgf8 treatment, most pronounced after 2 weeks of exposure, but after 4 weeks of treatment, *EMX2* levels were higher than in untreated controls. This is suggestive of a non-linear expression dynamic, which could potentially be explained by a negative feedback loop in the regulatory network affected by Fgf8 treatment. It is known that Fgf8 represses EMX2 expression, and EMX2 in turn represses Fgf8 (Cholfin and Rubenstein 2007). It is thus possible that the initial repression of EMX2 caused by ectopic Fgf8 results in the de-repression of endogenous Fgf8, which would in

turn result in a stronger repression of EMX2. Thus, the observed rebounding of EMX2 expression after 4 weeks of treatment likely involves other regulatory mechanisms at play.

A handful of genes showed a similar expression dynamic, rebounding after treatment for 4 weeks after being downregulated with treatment for 1-2 weeks. These genes included LHX2, a marker of hippocampal identity; *SEZ6L*, an ASD-associated gene, the WNT-pathway gene *LEF1*, and the G2/M gene *HMGA2*, among others (Figure 3.2D) Expression levels of Fgf8 were undetectable in all samples by RNA sequencing and by qPCR. Other downstream targets of Fgf8 appeared to be upregulated with Fgf8 treatment in addition to *SP8*, including *ETV1* (wk3-7, wk5-7) and *ETV3* (wk5-7). However, *ETV4* and *ETV5* appeared to be downregulated in all treatment conditions and most prominently in the wk5-7 group.

We also observed a substantial increase in the expression of progenitor and cell division markers, including *SOX2* and *MKI67*, in all treatment conditions, suggesting that Fgf8 treatment is markedly increasing proliferation. This is a known effect of FGFs and has been demonstrated for *FGF8* specifically. (Pei et al. 2017; Cruz-Martinez et al. 2014). Of note, the morphology of Fgf8-treated organoids was substantially different than that of control organoids. Neural rosettes, which are very prominent under a light microscope at 6 weeks in control cortical organoids, were not as readily observable in the treated group, which showed a much denser composition (Figure 3.4). It would be interesting to understand if there is a relationship between this observed difference in composition and the apparent increase in cell proliferation.

The MAPK pathway is a major downstream target of Fgf8. *MAPK3* (*ERK1)* becomes activated upon Fgf8 binding to its receptors and in turn promotes proliferation. Although our data cannot reveal whether MAPK3 activation is increased, it is interesting

to note that there is a substantial increase in *MAPK3* expression levels at two of the week 3 treatment conditions (wk3-4, wk3-5) (Figure 3.4 B, right). Several members of the *FGF*, *BMP* and *MAPK* family were also differentially expressed with Fgf8 treatment (Figure 3.4 B).

Along with the upregulation of several interneuron marker genes as discussed above, I also observed a downregulation of *FOXG1*, a telencephalic marker, and of *EMX1*, a dorsal telencephalon marker. This is concerning, as a previous study using neurospheres to assess the effects of Fgf8 treatment reported no reduction in *FOXG1* expression, instead reporting a slight upregulation (Imaizumi et al. 2018). A similar observation was made in mouse embryos upon Fgf8 overexpression (Shimamura and Rubenstein 1997). This discrepancy suggests that the context in which cells exist in these organoids plays an important role in the effect of Fgf8 signaling and might not be conducive to their anteriorization as intended.

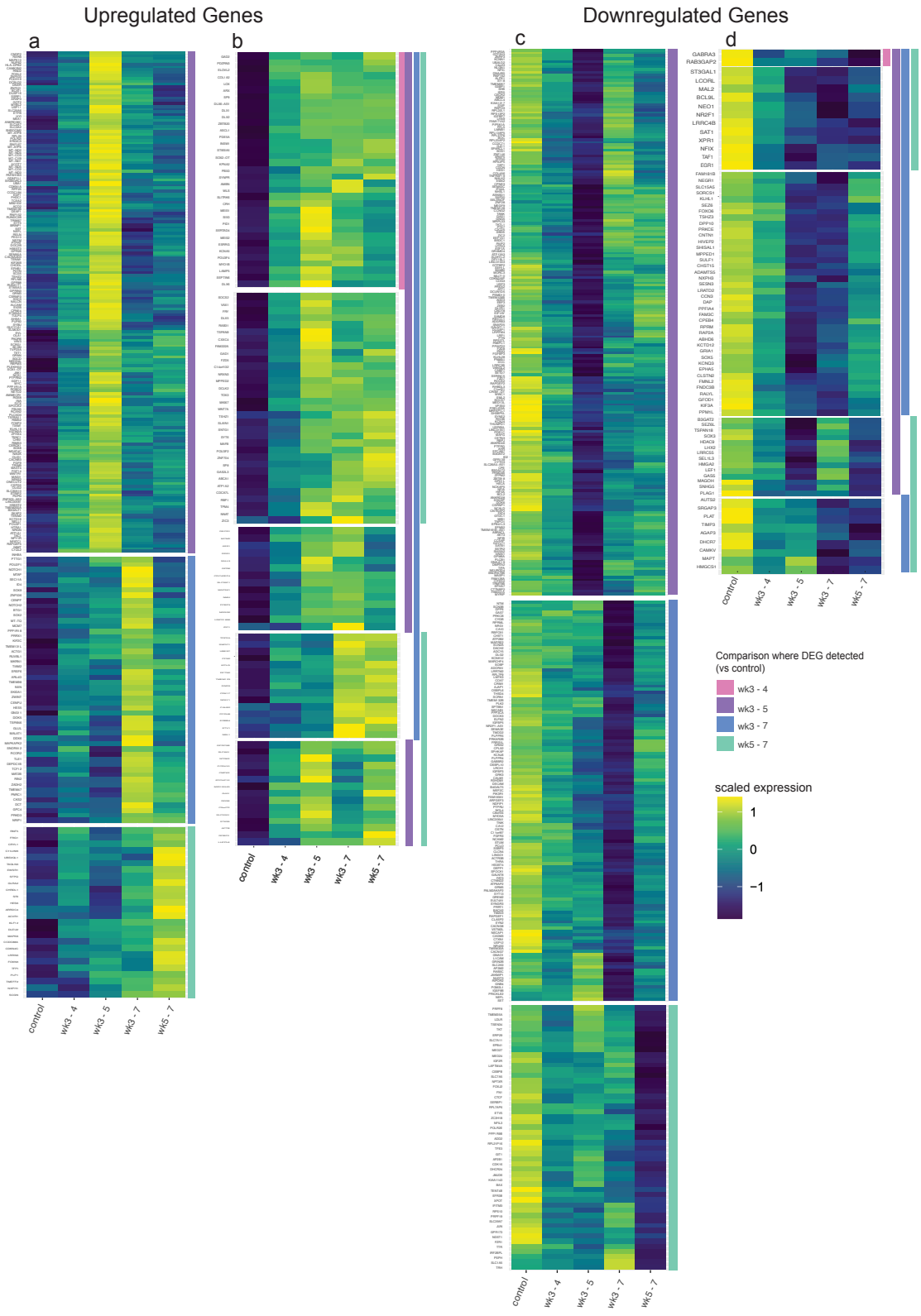


Figure 3.2. Differentially expressed genes detected in week 8 organoids treated with Fgf8 at distinct timepoints..

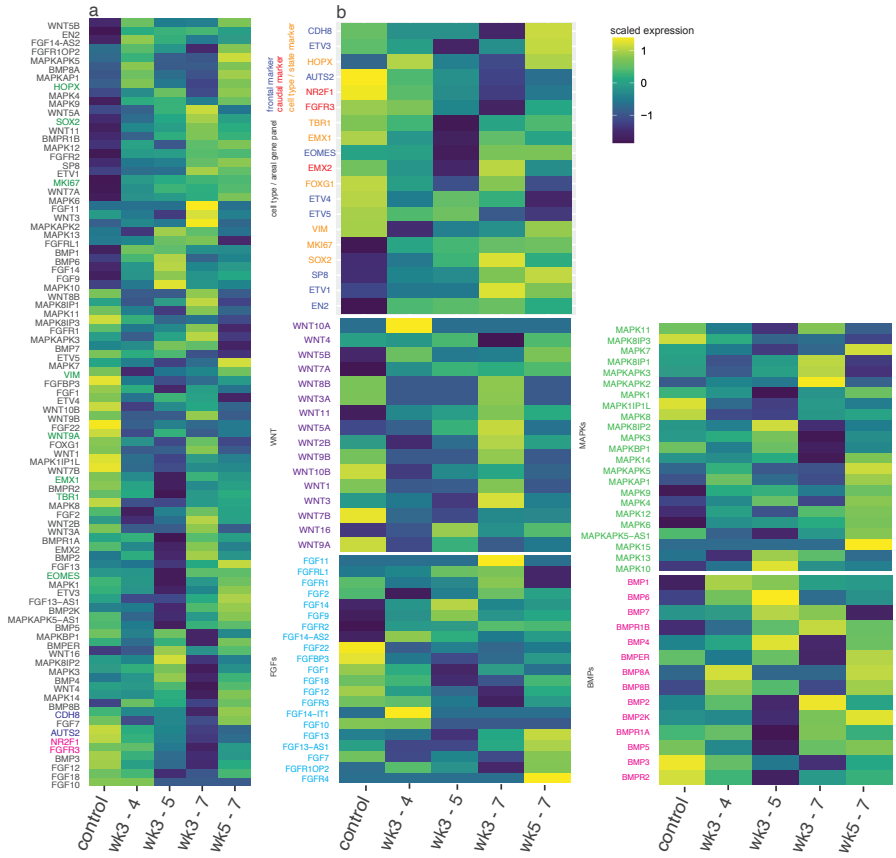


Figure 3.3. Expression of selected gene panels across treatment groups.

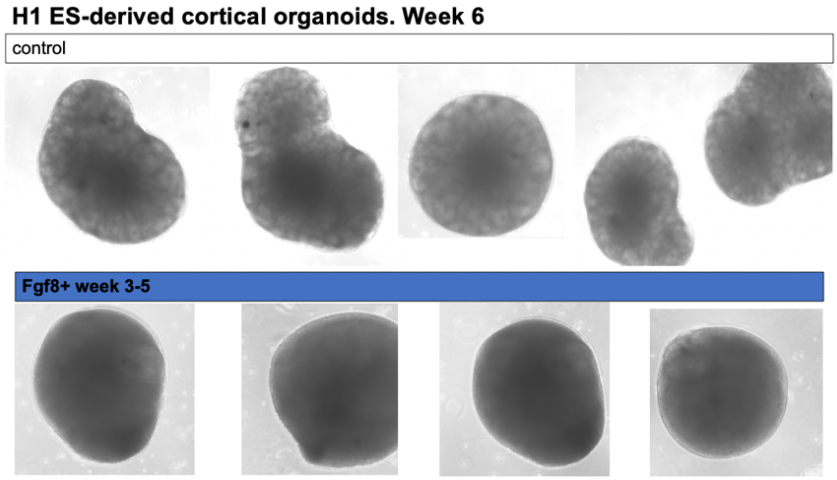


Figure 3.4 Control ES-derived cortical organoids (top) show a less dense composition, with neural rosettes readily detectable throughout at week 6. In contrast , Fgf8-treated organoids treated for 2 weeks (3-5) showed a much denser composition.

Discussion

Controlling areal identity is crucial for accurate modeling of neurodevelopmental psychiatric disorders. In a previous study, treatment with Fgf8 from days 12-18 conferred rostral identity (sensorimotor-like) to cortical progenitors, whereas progenitors displayed a caudal identity (temporal-like) in the absence of Fgf8 (Imaizumi et al. 2018).The results I report here were not consistent with what was reported in this previous study, which could be attributed to several important factors. First, the neuronal differentiation, aggregate induction and culture protocol (referred to as 'neurosphere culture system') used in the previous study was substantially different from the cortical organoid differentiation protocol that our lab uses, the SFEBq system, which is described in (Kadoshima et al. 2013). One critical difference is that the previous study does not include any Wnt inhibition during the neural induction phase (days 0-18), but instead effects Wnt activation with GSK-3. Additionally, that protocol uses dual SMAD inhibition with SB43412 and LDN193189, whereas our protocol only includes SB43412.

Neural induction in the neurosphere study consisted of pretreating PSCs for 6 days with 3 μM SB431542 (TGF-β inhibitor) and 150nM LDN193189 (BMP inhibitor) before seeding cells, effectively subjecting them to dual SMAD inhibition. Following day 6, cells were seeded for neurosphere culture, and additional growth factors and inhibitors were added from days 6-12: FGF-2, B27 supplement without vitamin A, SB431542, IWR-1e, and 10 μM Y-27632 (ROCK inhibitor). On day 12, cells were reseeded with minimal hormone media (MHM) including B27 and Y-27632. In contrast, our protocol does not include a dual SMAD inhibition pretreatment stage, and neural induction is achieved with a single SMAD inhibitor (SB431542, a TGF-B inhibitor) and a Wnt-inhibitor (IWR-1e), which are added to culture media on days 0-18 (day 0 is defined as the day when cells are

seeded in 96-well plates). Thus, our cortical differentiation protocol and organoid culture system are substantially different from the previous study, and areal specification has not been shown in this system. Additionally, Imaizumi et al only maintained neurospheres in culture for 18 days, and later timepoints were not assayed. Fgf8 and FgfR3 were added at only once at day 12, and cells were maintained for 6 days prior to analysis. In this study, treatment with Fgf8 was started at distinct timepoints along the organoid differentiation protocol and continued for varied durations, as described in Methods.

CHAPTER 4

Concluding Remarks and Future Directions

The adult cerebral cortex has striking area-specific differences in cellular composition, lamination pattern, and circuit structure. Proposed around 4 decades ago, the view that the cortical primordium is initially patterned in similar ways to the rest of the embryo, that is, controlled by a system of signaling centers that provide positional information and influence cell fate, was a conceptual breakthrough in the field of neurodevelopment. Emerging single-cell transcriptomic technologies have created an unprecedented opportunity for investigating gene expression programs at high-resolution and have enabled unbiased comparisons of the molecular signatures of homologous cell types across areas.

The work described in this thesis contributes to an updated view cortical arealization, in which molecular distinctions across broad "proto-regions" of the developing cortex are established early in development and refined into functional areas through activity-dependent mechanisms. Consistent with this, we find that a subset of areal identity profiles, particularly those seen at the anterior and posterior poles of the cortical primordium, are mutually exclusive in their expression patterns even at early timepoints in the second trimester of human gestation. One of the most striking findings of our work is that the expression patterns of areal marker genes are highly dynamic, with only a small fraction of markers present across cell types at all time points examined (between 10 to 25 gestational weeks). In addition, we observed that area-specific gene signatures observed in neurons at these developmental stages differed from those seen in neurons of the adult human brain, suggesting that gene expression programs are further refined across the lifespan. Our findings are consistent with previous transcriptomic studies of arealization performed using bulk RNA sequencing, but our study adds an important layer of complexity by examining these transcriptional signatures at a cell type

resolution. Aside from dynamic areal-specific gene regulation, the expression patterns of some genes even display distinct laminar distribution across developing cortical areas. For example, subplate markers *NR4A2, NEFL* and *SERPINI1* are co-expressed in prefrontal cortex, but show differential laminar distribution in somatosensory, temporal and visual cortices. Overall, networks of co-expressed genes change substantially across neocortical areas, and genes strongly associated to specific area identities are remarkably mutually exclusive – rarely or never co-expressed in the same cell.

Today, the field supports an integrative view of the protomap and protocortex hypotheses, two views originally proposed to explain how the cortex is patterned into functional areas during development. After the early, initial establishment of areal identity by a system of signaling centers, axons from the developing thalamus arrive at their correct cortical areal destination by axon guidance mechanisms and begin to form synapses. The activity of these thalamocortical axons are then thought to drive the finer maturation of each area. Prior to our study, the degree to which progenitors from distinct "proto-areas" of the developing neocortex are molecularly distinct from each other was not well understood. In the work described in this thesis, we used the power of single-cell transcriptomics in combination with single-molecule fluorescence *in situ* hybridization to uncover new insights into this key question of developmental neurobiology. While many studies over the past three decades have uncovered key transcription factors that are expressed in a gradient-like manner across the many axes of the developing dorsal telencephalon, our work shows that combinatorial transcriptional programs define distinct subpopulations of progenitors across developing cortical areas. We present new evidence for a model that combines and extends the two prevailing hypotheses and propose, based on this evidence, that during the early second trimester, neural progenitor

cells are primed toward either of two fates: frontal or occipital identities. This essentially reflects a partial protomap scenario, or what has been referred to as “molecular pre-patterning”. Later in development, these transcriptional differences become more pronounced and areal identities throughout the rostrocaudal axis gradually emerge.

The maturation rate of brain regions and cortical areas is known to vary across the rostrocaudal axis during development. While we did not focus on this point in this study, it may explain a subset of the patterns of differentially expressed genes that we observe. It should be noted, however, that differences in maturation rate are difficult to control for as they are not constant across development. Our lab has previously used gene expression network analysis to uncover genes that are correlated with maturation signatures (Nowakowski et al. 2017), and when examining the list of differentially expressed genes across areas generated by the studies included in this thesis, we do not find an enrichment for those maturation genes. Although certain genes may influence maturation, we do not have evidence for whether these may also play roles in establishing areal identities of the cortex. Additionally, we did not find any evidence for over-representation of any molecular functions or biological processes among the gene markers contributing to area-specific signatures. This may be due to the fact that the functions of many of the genes we identified in area-specific signatures have not been well described or annotated in a developmental context, which presents an exciting area for future work.

Differences in areal gene signatures are detected as early as in the radial glia from the poles of the rostrocaudal axis, the prefrontal and primary visual cortices, respectively. Those transcriptional differences then change, but also strengthen with lineage progression in area-specific neuronal populations. At a stage in between radial glia and

neurons, the intermediate progenitor cells, we also see evidence for area-specific transcriptional programs. We find that these areal signatures are dynamic across cell types along the differentiation trajectory, as well as across developmental stages (i.e. gestational weeks). Interestingly, the areal identities of IPCs were more similar to those of neurons than to their radial glia counterparts, despite their progenitor identity. This suggests that IPCs are more restricted in their fate potential, and that areal identity of neurons might already be determined, or "locked in", by the transcriptional programs expressed in IPCs, which are in turn primed, but to a looser extent, by the area-specific programs expressed in RG. Future studies comparing the epigenetic configuration of these cell types across cortical areas should help us understand why there seem to be different degrees of similarity between cell types across the differentiation trajectory. For example, it is possible that the gene expression differences we see between RG from distinct areas have fewer epigenetic consequences, and are therefore more fluid, whereas the expression differences that emerge from the RG stage have more stable and permanent, chromatin-related effects. This has been proposed as a general mechanism of cell differentiation and lineage commitment, in which the level of DNA methylation increases as stem cell differentiation progresses, along with an increase in inactive heterochromatin. This would also suggest that while area-specific expression signatures are seen as early as RG, heterotopic transplants of cells of this type could potentially still give rise to neuronal progeny with the target areal identity, but that transplanting cells further along the differentiation trajectory, namely IPCs, would not, given their less malleable identity.

The area-specific transcriptional programs uncovered in this study should enable future, data-driven studies of mechanisms of arealization. How do area-specific

transcriptional programs expressed in RG become refined in intermediate progenitor cells, and further refined in their neuronal progeny? What are the molecular switches that trigger changes in areal signatures from one cell type to another? It will also be interesting to investigate how the transcriptional differences uncovered here set the stage for the proper guidance of thalamocortical axons from corresponding thalamic nuclei to their respective cortical destination. Differential thalamocortical innervation across cortical areas is a well-studied neurodevelopmental process, but little is known about the intrinsic molecular factors within the cortex that establish the signaling conditions to guide thalamic axons to their appropriate destinations as they innervate the cortex. The transcriptional signatures uncovered in this study should provide a valuable starting point to begin exploring these mechanisms. It will also be especially interesting to investigate the transcriptional programs that are present in the progenitors of the developing prefrontal cortex, an area that has undergone a disproportionally massive expansion in the evolutionary lineage of humans, and which is responsible for the higher-order cognitive abilities that distinguish us as a species. These efforts are also of significant translational relevance, as the prefrontal is disproportionately affected in debilitating neurodevelopmental disorders like schizophrenia and ASD.

References

Abraham, Ariel B., Robert Bronstein, Avanish S. Reddy, Mirjana Maletic-Savatic, Adan Aguirre, and Stella E. Tsirka. 2013. "Aberrant Neural Stem Cell Proliferation and Increased Adult Neurogenesis in Mice Lacking Chromatin Protein HMGB2." *PloS One* 8 (12): e84838.

Alzu'bi, Ayman, Ayman Alzu'bi, Jihane Homman-Ludiye, James A. Bourne, and Gavin J. Clowry. 2019. "Thalamocortical Afferents Innervate the Cortical Subplate Much Earlier in Development in Primate than in Rodent." *Cerebral Cortex* 29. https://doi.org/10.1093/cercor/bhy327.

Armentano, Maria, Shen-Ju Chou, Giulio Srubek Tomassy, Axel Leingärtner, Dennis D. M. O'Leary, and Michèle Studer. 2007. "COUP-TFI Regulates the Balance of Cortical Patterning between Frontal/Motor and Sensory Areas." *Nature Neuroscience* 10 (10): 1277–86.

Aruga, Jun, Takashi Inoue, Jun Hoshino, and Katsuhiko Mikoshiba. 2002. "Zic2 Controls Cerebellar Development in Cooperation with Zic1." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 22 (1): 218–25.

Assimacopoulos, Stavroula, Tina Kao, Naoum P. Issa, and Elizabeth A. Grove. 2012. "Fibroblast Growth Factor 8 Organizes the Neocortical Area Map and Regulates Sensory Map Topography." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 32 (21): 7191–7201.

Bakken, Trygve E., Jeremy A. Miller, Song-Lin Ding, Susan M. Sunkin, Kimberly A. Smith, Lydia Ng, Aaron Szafer, et al. 2016. "A Comprehensive Transcriptional Map of Primate Brain Development." *Nature* 535 (7612): 367–75.

Barkovich, A. James, Renzo Guerrini, Ruben I. Kuzniecky, Graeme D. Jackson, and William B. Dobyns. 2012. “A Developmental and Genetic Classification for Malformations of Cortical Development: Update 2012.” *Brain: A Journal of Neurology* 135 (Pt 5): 1348–69.

Ben-Reuven, Lihi, and Orly Reiner. 2020. “Toward Spatial Identities in Human Brain Organoids-on-Chip Induced by Morphogen-Soaked Beads.” *Bioengineering (Basel, Switzerland)* 7 (4). https://doi.org/10.3390/bioengineering7040164.

Bershteyn, Marina, Tomasz J. Nowakowski, Alex A. Pollen, Elizabeth Di Lullo, Aishwarya Nene, Anthony Wynshaw-Boris, and Arnold R. Kriegstein. 2017. “Human IPSC-Derived Cerebral Organoids Model Cellular Features of Lissencephaly and Reveal Prolonged Mitosis of Outer Radial Glia.” *Cell Stem Cell* 20 (4): 435-449.e4.

Bertacchi, Michele, Agnès Gruart, Polynikis Kaimakis, Cécile Allet, Linda Serra, Paolo Giacobini, José M. Delgado-García, Paola Bovolenta, and Michèle Studer. 2019. “Mouse Nr2f1 Haploinsufficiency Unveils New Pathological Mechanisms of a Human Optic Atrophy Syndrome.” *EMBO Molecular Medicine* 11 (8): e10291.

Bhaduri, Aparna, Madeline G. Andrews, Walter Mancia Leon, Diane Jung, David Shin, Denise Allen, Dana Jung, et al. 2020. “Cell Stress in Cortical Organoids Impairs Molecular Subtype Specification.” *Nature* 578 (7793): 142–48.

Birey, Fikri, Jimena Andersen, and Sergiu P. Paşca. 2017. “Generation and Assembly of Forebrain Spheroids.” *Protocol Exchange*, April. https://doi.org/10.1038/protex.2017.044.

Birgersdotter, Anna, Rickard Sandberg, and Ingemar Ernberg. 2005. “Gene Expression Perturbation in Vitro--a Growing Case for Three-Dimensional (3D) Culture Systems.” *Seminars in Cancer Biology* 15 (5): 405–12.

Brodmann, K. 1909. “Vergleichende Lokalisationslehre Der Grosshirnrinde in Ihren Prinzipien Dargestellt Auf Grund Des Zellenbaues.”

———. 2006. *Brodmann’s: Localisation in the Cerebral Cortex*. Springer.

Buettner, Florian, Kedar N. Natarajan, F. Paolo Casale, Valentina Proserpio, Antonio Scialdone, Fabian J. Theis, Sarah A. Teichmann, John C. Marioni, and Oliver Stegle. 2015. “Computational Analysis of Cell-to-Cell Heterogeneity in Single-Cell RNA-Sequencing Data Reveals Hidden Subpopulations of Cells.” *Nature Biotechnology* 33 (2): 155–60.

Butti, Camilla, Chet C. Sherwood, Atiya Y. Hakeem, John M. Allman, and Patrick R. Hof. 2009. “Total Number and Volume of Von Economo Neurons in the Cerebral Cortex of Cetaceans.” *The Journal of Comparative Neurology* 515 (2): 243–59.

Cadwell, Cathryn R., Aparna Bhaduri, Mohammed A. Mostajo-Radji, Matthew G. Keefe, and Tomasz J. Nowakowski. 2019. “Development and Arealization of the Cerebral Cortex.” *Neuron* 103 (6): 980–1004.

Cakir, Bilal, Yangfei Xiang, Yoshiaki Tanaka, Mehmet H. Kural, Maxime Parent, Young-Jin Kang, Kayley Chapeton, et al. 2019. “Engineering of Human Brain Organoids with a Functional Vascular-like System.” *Nature Methods*, October. https://doi.org/10.1038/s41592-019-0586-5.

Camp, J. Gray, Farhath Badsha, Marta Florio, Sabina Kanton, Tobias Gerber, Michaela Wilsch-Bräuninger, Eric Lewitus, et al. 2015. “Human Cerebral Organoids Recapitulate Gene Expression Programs of Fetal Neocortex Development.” *Proceedings of the National Academy of Sciences of the United States of America* 112 (51): 15672–77.

Cederquist, Gustav Y., James J. Asciolla, Jason Tchieu, Ryan M. Walsh, Daniela Cornacchia, Marilyn D. Resh, and Lorenz Studer. 2019. “Specification of Positional Identity in Forebrain Organoids.” *Nature Biotechnology* 37 (4): 436–44.

Chan, Chun Ming, Joel Fulton, Cristina Montiel-Duarte, Hilary M. Collins, Neetu Bharti, Frances R. Wadelin, Paula M. Moran, Nigel P. Mongan, and David M. Heery. 2013. “A Signature Motif Mediating Selective Interactions of BCL11A with the NR2E/F Subfamily of Orphan Nuclear Receptors.” *Nucleic Acids Research* 41 (21): 9663–79.

Cholfin, Jeremy A., and John L. R. Rubenstein. 2007. “Genetic Regulation of Prefrontal Cortex Development and Function.” *Novartis Foundation Symposium* 288: 165–73; discussion 173-7, 276–81.

Codeluppi, Simone, Lars E. Borm, Amit Zeisel, Gioele La Manno, Josina A. van Lunteren, Camilla I. Svensson, and Sten Linnarsson. 2018. “Spatial Organization of the Somatosensory Cortex Revealed by OsmFISH.” *Nature Methods* 15 (11): 932–35.

Colombo, J. A. 1996. “Interlaminar Astroglial Processes in the Cerebral Cortex of Adult Monkeys but Not of Adult Rats.” *Acta Anatomica* 155 (1): 57–62.

Cruz-Martinez, Pablo, Almudena Martinez-Ferre, Jesus Jaramillo-Merchán, Alicia Estirado, Salvador Martinez, and Jonathan Jones. 2014. “FGF8 Activates Proliferation and Migration in Mouse Post-Natal Oligodendrocyte Progenitor Cells.” *PloS One* 9 (9): e108241.

Danjo, Teruko, Mototsugu Eiraku, Keiko Muguruma, Kiichi Watanabe, Masako Kawada, Yuchio Yanagawa, John L. R. Rubenstein, and Yoshiki Sasai. 2011. “Subregional Specification of Embryonic Stem Cell-Derived Ventral Telencephalic Tissues by Timed and Combinatory Treatment with Extrinsic Signals.” *The Journal of*

Neuroscience: The Official Journal of the Society for Neuroscience 31 (5): 1919–33.

Darmanis, Spyros, Steven A. Sloan, Ye Zhang, Martin Enge, Christine Caneda, Lawrence M. Shuer, Melanie G. Hayden Gephart, Ben A. Barres, and Stephen R. Quake. 2015. "A Survey of Human Brain Transcriptome Diversity at the Single Cell Level." *Proceedings of the National Academy of Sciences of the United States of America* 112 (23): 7285–90.

Duval, Kayla, Hannah Grover, Li-Hsin Han, Yongchao Mou, Adrian F. Pegoraro, Jeffery Fredberg, and Zi Chen. 2017. "Modeling Physiological Events in 2D vs. 3D Cell Culture." *Physiology (Bethesda, Md.)* 32 (4): 266–77.

Eiraku, Mototsugu, Kiichi Watanabe, Mami Matsuo-Takasaki, Masako Kawada, Shigenobu Yonemura, Michiru Matsumura, Takafumi Wataya, Ayaka Nishiyama, Keiko Muguruma, and Yoshiki Sasai. 2008. "Self-Organized Formation of Polarized Cortical Tissues from ESCs and Its Active Manipulation by Extrinsic Signals." *Cell Stem Cell* 3 (5): 519–32.

Eze, Ugomma C., Aparna Bhaduri, Maximilian Haeussler, Tomasz J. Nowakowski, and Arnold R. Kriegstein. 2021. "Single-Cell Atlas of Early Human Brain Development Highlights Heterogeneity of Human Neuroepithelial Cells and Early Radial Glia." *Nature Neuroscience* 24. https://doi.org/10.1038/s41593-020-00794-1.

Fontoura, Julia C., Christian Viezzer, Fabiana G. Dos Santos, Rosane A. Ligabue, Ricardo Weinlich, Renato D. Puga, Dyeison Antonow, Patricia Severino, and Cristina Bonorino. 2020. "Comparison of 2D and 3D Cell Culture Models for Cell Growth, Gene Expression and Drug Resistance." *Materials Science & Engineering. C, Materials for Biological Applications* 107 (110264): 110264.

Fukuchi-Shimogori, T., and E. A. Grove. 2001. “Neocortex Patterning by the Secreted Signaling Molecule FGF8.” *Science* 294 (5544): 1071–74.

Giandomenico, Stefano L., Magdalena Sutcliffe, and Madeline A. Lancaster. 2021. “Generation and Long-Term Culture of Advanced Cerebral Organoids for Studying Later Stages of Neural Development.” *Nature Protocols* 16 (2): 579–602.

Gordon, Aaron, Se-Jin Yoon, Stephen S. Tran, Christopher D. Makinson, Jin Young Park, Jimena Andersen, Alfredo M. Valencia, et al. 2021. “Long-Term Maturation of Human Cortical Organoids Matches Key Early Postnatal Transitions.” *Nature Neuroscience* 24 (3): 331–42.

Guillemot, François. 2005. “Cellular and Molecular Control of Neurogenesis in the Mammalian Telencephalon.” *Current Opinion in Cell Biology* 17 (6): 639–47.

Haines, Duane E. 2012. *Fundamental Neuroscience for Basic and Clinical Applications*. 4th ed. London, England: W B Saunders.

Hakeem, Atiya Y., Chet C. Sherwood, Christopher J. Bonar, Camilla Butti, Patrick R. Hof, and John M. Allman. 2009. “Von Economo Neurons in the Elephant Brain.” *Anatomical Record (Hoboken, N.J.: 2007)* 292 (2): 242–48.

Hamasaki, Tadashi, Axel Leingärtner, Thomas Ringstedt, and Dennis D. M. O’Leary. 2004. “EMX2 Regulates Sizes and Positioning of the Primary Sensory and Motor Areas in Neocortex by Direct Specification of Cortical Progenitors.” *Neuron* 43 (3): 359–72.

Hansen, David V., Jan H. Lui, Philip R. L. Parker, and Arnold R. Kriegstein. 2010. “Neurogenic Radial Glia in the Outer Subventricular Zone of Human Neocortex.” *Nature* 464 (7288): 554–61.

Harrison-Uy, Susan J., Julie A. Siegenthaler, Andrea Faedo, John L. R. Rubenstein, and Samuel J. Pleasure. 2013. “CoupTFI Interacts with Retinoic Acid Signaling during Cortical Development.” *PloS One* 8 (3): e58219.

Herculano-Houzel, Suzana. 2009. “The Human Brain in Numbers: A Linearly Scaled-up Primate Brain.” *Frontiers in Human Neuroscience* 3 (November): 31.

———. 2016. “The Evolution of Evolutionary Neuroscience.” In *The Wiley Handbook of Evolutionary Neuroscience*, 21–37. Chichester, UK: John Wiley & Sons, Ltd.

Hladnik, Ana, Domagoj Džaja, Sanja Darmopil, Nataša Jovanov-Milošević, and Zdravko Petanjek. 2014. “Spatio-Temporal Extension in Site of Origin for Cortical Calretinin Neurons in Primates.” *Frontiers in Neuroanatomy* 8 (June): 50.

Hodge, Rebecca D., Trygve E. Bakken, Jeremy A. Miller, Kimberly A. Smith, Eliza R. Barkan, Lucas T. Graybuck, Jennie L. Close, et al. 2019. “Conserved Cell Types with Divergent Features in Human versus Mouse Cortex.” *Nature* 573 (7772): 61–68.

Iefremova, Vira, George Manikakis, Olivia Krefft, Ammar Jabali, Kevin Weynans, Ruven Wilkens, Fabio Marsoner, et al. 2017. “An Organoid-Based Model of Cortical Development Identifies Non-Cell-Autonomous Defects in Wnt Signaling Contributing to Miller-Dieker Syndrome.” *Cell Reports* 19 (1): 50–59.

Ilieva, Mirolyba, Åsa Fex Svenningsen, Morten Thorsen, and Tanja Maria Michel. 2018. “Psychiatry in a Dish: Stem Cells and Brain Organoids Modeling Autism Spectrum Disorders.” *Biological Psychiatry* 83 (7): 558–68.

Imaizumi, Kent, Koki Fujimori, Seiji Ishii, Asako Otomo, Yasushi Hosoi, Hiroaki Miyajima, Hitoshi Warita, et al. 2018. “Rostrocaudal Areal Patterning of Human PSC-Derived Cortical Neurons by FGF8 Signaling.” *ENeuro* 5 (2). https://doi.org/10.1523/ENEURO.0368-17.2018.

Jeremy Willsey, A., Aaron Oldre, Aaron Szafer, Adrian Camarena, Adriana Cherskov, Alexander W. Charney, Alexej Abyzov, et al. 2018. “Integrative Functional Genomic Analysis of Human Brain Development and Neuropsychiatric Risks.” *Science*, December. https://science.sciencemag.org/content/362/6420/eaat7615.abstract.

Jimenez-Palomares, Margarita, Alba Cristobal, and Mª Carmen Duran Ruiz. 2021. “Organoids Models for the Study of Cell-Cell Interactions.” In *Cell Interaction - Molecular and Immunological Basis for Disease Management*. IntechOpen.

Kadoshima, Taisuke, Hideya Sakaguchi, Tokushige Nakano, Mika Soen, Satoshi Ando, Mototsugu Eiraku, and Yoshiki Sasai. 2013. “Self-Organization of Axial Polarity, inside-out Layer Pattern, and Species-Specific Progenitor Dynamics in Human ES Cell–Derived Neocortex.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (50): 20284–89.

Kang, Hyo Jung, Yuka Imamura Kawasawa, Feng Cheng, Ying Zhu, Xuming Xu, Mingfeng Li, André M. M. Sousa, et al. 2011. “Spatio-Temporal Transcriptome of the Human Brain.” *Nature* 478. https://doi.org/10.1038/nature10523.

Kostović, Ivica, Iris Žunić Išasegi, and Željka Krsnik. 2019. “Sublaminar Organization of the Human Subplate: Developmental Changes in the Distribution of Neurons, Glia, Growing Axons and Extracellular Matrix.” *Journal of Anatomy* 235 (3): 481–506.

La Manno, Gioele, Ruslan Soldatov, Amit Zeisel, Emelie Braun, Hannah Hochgerner, Viktor Petukhov, Katja Lidschreiber, et al. 2018. “RNA Velocity of Single Cells.” *Nature* 560 (7719): 494–98.

Lancaster, Madeline A., Magdalena Renner, Carol-Anne Martin, Daniel Wenzel, Louise S. Bicknell, Matthew E. Hurles, Tessa Homfray, Josef M. Penninger, Andrew P.

Jackson, and Juergen A. Knoblich. 2013. “Cerebral Organoids Model Human Brain Development and Microcephaly.” *Nature* 501 (7467): 373–79.

Landhuis, Esther. 2020. “Probing Fine-Scale Connections in the Brain.” *Nature* 586 (7830): 631–33.

Langfelder, Peter, and Steve Horvath. 2008. “WGCNA: An R Package for Weighted Correlation Network Analysis.” *BMC Bioinformatics* 9 (1): 559.

Llorca, Alfredo, and Oscar Marín. 2020. “Orchestrated Freedom: New Insights into Cortical Neurogenesis.” *Current Opinion in Neurobiology* 66 (October): 48–56.

Lu, David, Eric Y. T. Chen, Philip Lee, Yung-Chen Wang, Wendy Ching, Christopher Markey, Chase Gulstrom, Li-Ching Chen, Thien Nguyen, and Wei-Chun Chin. 2015. “Accelerated Neuronal Differentiation toward Motor Neuron Lineage from Human Embryonic Stem Cell Line (H9).” *Tissue Engineering. Part C, Methods* 21 (3): 242–52.

Manuel, Martine N., Ben Martynoga, Mike D. Molinek, Jane C. Quinn, Corinne Kroemmer, John O. Mason, and David J. Price. 2011. “The Transcription Factor Foxg1 Regulates Telencephalic Progenitor Proliferation Cell Autonomously, in Part by Controlling Pax6 Expression Levels.” *Neural Development* 6 (March): 9.

McMurtrey, Richard J. 2016. “Multi-Compartmental Biomaterial Scaffolds for Patterning Neural Tissue Organoids in Models of Neurodevelopment and Tissue Regeneration.” *Journal of Tissue Engineering* 7 (January): 2041731416671926.

Molnár, Zoltán, Gavin J. Clowry, Nenad Šestan, Ayman Alzu’bi, Trygve Bakken, Robert F. Hevner, Petra S. Hüppi, et al. 2019. “New Insights into the Development of the Human Cerebral Cortex.” *Journal of Anatomy*, August. https://doi.org/10.1111/joa.13055.

Motono, Makoto, Yoshihiko Ioroi, Takenori Ogura, and Jun Takahashi. 2016. "WNT-C59, a Small-Molecule WNT Inhibitor, Efficiently Induces Anterior Cortex That Includes Cortical Motor Neurons From Human Pluripotent Stem Cells." *Stem Cells Translational Medicine* 5 (4): 552–60.

Muzio, L. 2003. "Emx1, Emx2 and Pax6 in Specification, Regionalization and Arealization of the Cerebral Cortex." *Cerebral Cortex*. https://doi.org/10.1093/cercor/13.6.641.

Nascimento, Juliana Minardi, Verônica M. Saia-Cereda, Rafaela C. Sartore, Rodrigo Madeiro da Costa, Clarissa S. Schitine, Hercules Rezende Freitas, Michael Murgu, Ricardo A. de Melo Reis, Stevens K. Rehen, and Daniel Martins-de-Souza. 2019. "Human Cerebral Organoids and Fetal Brain Tissue Share Proteomic Similarities." *Frontiers in Cell and Developmental Biology* 7 (November): 303.

Nazareth, Emanuel J. P., Joel E. E. Ostblom, Petra B. Lücker, Shreya Shukla, Manuel M. Alvarez, Steve K. W. Oh, Ting Yin, and Peter W. Zandstra. 2013. "High-Throughput Fingerprinting of Human Pluripotent Stem Cell Fate Responses and Lineage Bias." *Nature Methods* 10 (12): 1225–31.

Nimchinsky, E. A., E. Gilissen, J. M. Allman, D. P. Perl, J. M. Erwin, and P. R. Hof. 1999. "A Neuronal Morphologic Type Unique to Humans and Great Apes." *Proceedings of the National Academy of Sciences of the United States of America* 96 (9): 5268–73.

Nowakowski, Tomasz J., Aparna Bhaduri, Alex A. Pollen, Beatriz Alvarado, Mohammed A. Mostajo-Radji, Elizabeth Di Lullo, Maximilian Haeussler, et al. 2017. "Spatiotemporal Gene Expression Trajectories Reveal Developmental Hierarchies of the Human Cortex." *Science* 358 (6368): 1318–23.

Oberheim, Nancy Ann, Takahiro Takano, Xiaoning Han, Wei He, Jane H. C. Lin, Fushun Wang, Qiwu Xu, et al. 2009. "Uniquely Hominid Features of Adult Human Astrocytes." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 29 (10): 3276–87.

O'Leary, D. D. 1989. "Do Cortical Areas Emerge from a Protocortex?" *Trends in Neurosciences* 12 (10): 400–406.

O'Leary, Dennis D. M. 1989. "Do Cortical Areas Emerge from a Protocottex?" *TINS* 12 (10).

O'Rahilly, Ronan, and Fabiola Müller. 2008. "Significant Features in the Early Prenatal Development of the Human Brain." *Anatomischer Anzeiger [Annals of Anatomy]* 190 (2): 105–18.

Pei, Yuanmin, Xueling Sun, Xiwei Guo, Huashan Yin, Le Wang, Fugu Tian, Hongxi Jing, Xiaobo Liang, Jun Xu, and Pengcheng Shi. 2017. "FGF8 Promotes Cell Proliferation and Resistance to EGFR Inhibitors via Upregulation of EGFR in Human Hepatocellular Carcinoma Cells." *Oncology Reports* 38 (4): 2205–10.

Perani, Daniela, Maria C. Saccuman, Paola Scifo, Alfred Anwander, Danilo Spada, Cristina Baldoli, Antonella Poloniato, Gabriele Lohmann, and Angela D. Friederici. 2011. "Neural Language Networks at Birth." *Proceedings of the National Academy of Sciences of the United States of America* 108 (38): 16056–61.

Pla, Ramon, Amelia Stanco, Mackenzie A. Howard, Anna N. Rubin, Daniel Vogt, Niall Mortimer, Inma Cobos, et al. 2018. "Dlx1 and Dlx2 Promote Interneuron GABA Synthesis, Synaptogenesis, and Dendritogenesis." *Cerebral Cortex (New York, N.Y.: 1991)* 28 (11): 3797–3815.

Pletikos, Mihovil, André M. M. Sousa, Goran Sedmak, Kyle A. Meyer, Ying Zhu, Feng Cheng, Mingfeng Li, Yuka Imamura Kawasawa, and Nenad Sestan. 2014. "Temporal Specification and Bilaterality of Human Neocortical Topographic Gene Expression." *Neuron* 81 (2): 321–32.

Pollen, Alex A., Aparna Bhaduri, Madeline G. Andrews, Tomasz J. Nowakowski, Olivia S. Meyerson, Mohammed A. Mostajo-Radji, Elizabeth Di Lullo, et al. 2019. "Establishing Cerebral Organoids as Models of Human-Specific Brain Evolution." *Cell* 176. https://doi.org/10.1016/j.cell.2019.01.017.

Polleux, Franck. 2004. "Generation of the Cortical Area Map; Emx2 Strikes Back." *Neuron* 43 (3): 295–97.

Quadrato, Giorgia, and Paola Arlotta. 2017. "Present and Future of Modeling Human Brain Development in 3D Organoids." *Current Opinion in Cell Biology* 49 (December): 47–52.

Quadrato, Giorgia, Tuan Nguyen, Evan Z. Macosko, John L. Sherwood, Sung Min Yang, Daniel R. Berger, Natalie Maria, et al. 2017. "Cell Diversity and Network Dynamics in Photosensitive Human Brain Organoids." *Nature* 545 (7652): 48–53.

Rakic, P. 1988. "Specification of Cerebral Cortical Areas." *Science* 241 (4862): 170–76.

Rakic, Pasko. 2009. "Evolution of the Neocortex: A Perspective from Developmental Biology." *Nature Reviews. Neuroscience* 10 (10): 724–35.

Ramón y Cajal, Santiago. 1909. *Histologie Du Système Nerveux de l'homme & Des Vertébrés*. Paris : Maloine,.

Roth, Gerhard, and Ursula Dicke. 2012. "Evolution of the Brain and Intelligence in Primates." *Progress in Brain Research* 195: 413–30.

Rubenstein, John L. R. 2011. “Annual Research Review: Development of the Cerebral Cortex: Implications for Neurodevelopmental Disorders.” *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 52 (4): 339–55.

Ryu, J., S. S. Hong, B. K. P. Horn, D. M. Freeman, and M. S. Mermelstein. 2006. “Multibeam Interferometric Illumination as the Primary Source of Resolution in Optical Microscopy.” *Applied Physics Letters* 88 (17): 171112.

Sahara, Setsuko, Yasuhiko Kawakami, Juan Carlos Izpisua Belmonte, and Dennis D. M. O’Leary. 2007. “Sp8 Exhibits Reciprocal Induction with Fgf8 but Has an Opposing Effect on Anterior-Posterior Cortical Area Patterning.” *Neural Development* 2 (May): 10.

Sakaguchi, Hideya, Taisuke Kadoshima, Mika Soen, Nobuhiro Narii, Yoshihito Ishida, Masatoshi Ohgushi, Jun Takahashi, Mototsugu Eiraku, and Yoshiki Sasai. 2015. “Generation of Functional Hippocampal Neurons from Self-Organizing Human Embryonic Stem Cell-Derived Dorsomedial Telencephalic Tissue.” *Nature Communications* 6 (November): 8896.

Shibata, Mikihito, Kartik Pattabiraman, Belen Lorente-Galdos, David Andrijevic, Xiaojun Xing, Andre M. M. Sousa, Gabriel Santpere, and Nenad Sestan. 2019. “Regulation of Prefrontal Patterning, Connectivity and Synaptogenesis by Retinoic Acid.” *BioRxiv*. https://doi.org/10.1101/2019.12.31.891036.

Shimamura, K., and J. L. Rubenstein. 1997. “Inductive Interactions Direct Early Regionalization of the Mouse Forebrain.” *Development (Cambridge, England)* 124 (14): 2709–18.

Simon, Ruth, Christoph Wiegreffe, and Stefan Britsch. 2020. “Bcl11 Transcription Factors Regulate Cortical Development and Function.” *Frontiers in Molecular Neuroscience* 13 (April): 51.

Storm, Elaine E., Sonia Garel, Ugo Borello, Jean M. Hebert, Salvador Martinez, Susan K. McConnell, Gail R. Martin, and John L. R. Rubenstein. 2006. “Dose-Dependent Functions of Fgf8 in Regulating Telencephalic Patterning Centers.” *Development* 133 (9): 1831–44.

Supèr, H., A. Martínez, J. A. Del Río, and E. Soriano. 1998. “Involvement of Distinct Pioneer Neurons in the Formation of Layer-Specific Connections in the Hippocampus.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 18 (12): 4616–26.

Tan, Kenneth K. B., Jason Y. Tann, Sharvari R. Sathe, Seok Hong Goh, Dongliang Ma, Eyleen L. K. Goh, and Evelyn K. F. Yim. 2015. “Enhanced Differentiation of Neural Progenitor Cells into Neurons of the Mesencephalic Dopaminergic Subtype on Topographical Patterns.” *Biomaterials* 43 (March): 32–43.

Tang, Fuchou, Catalin Barbacioru, Yangzhou Wang, Ellen Nordman, Clarence Lee, Nanlan Xu, Xiaohui Wang, et al. 2009. “MRNA-Seq Whole-Transcriptome Analysis of a Single Cell.” *Nature Methods* 6 (5): 377–82.

Tasic, Bosiljka, Zizhen Yao, Lucas T. Graybuck, Kimberly A. Smith, Thuc Nghi Nguyen, Darren Bertagnolli, Jeff Goldy, et al. 2018. “Shared and Distinct Transcriptomic Cell Types across Neocortical Areas.” *Nature* 563 (7729): 72–78.

Teffer, Kate, Daniel P. Buxhoeveden, Cheryl D. Stimpson, Archibald J. Fobbs, Steven J. Schapiro, Wallace B. Baze, Mark J. McArthur, et al. 2013. “Developmental Changes

in the Spatial Organization of Neurons in the Neocortex of Humans and Common Chimpanzees." *The Journal of Comparative Neurology* 521 (18): 4249–59.

Van Essen, David C., Chad J. Donahue, Timothy S. Coalson, Henry Kennedy, Takuya Hayashi, and Matthew F. Glasser. 2019. "Cerebral Cortical Folding, Parcellation, and Connectivity in Humans, Nonhuman Primates, and Mice." *Proceedings of the National Academy of Sciences of the United States of America* 116 (52): 26173–80.

Velasco, Silvia, Silvia Velasco, Amanda J. Kedaigle, Amanda J. Kedaigle, Sean Simmons, Allison Nash, Allison Nash, et al. 2019. "Individual Brain Organoids Reproducibly Form Cell Diversity of the Human Cerebral Cortex." *Nature* 570. https://doi.org/10.1038/s41586-019-1289-x.

Walsh, C. A. 1999. "Genetic Malformations of the Human Cerebral Cortex." *Neuron* 23 (1): 19–29.

Wang, Yanling, Catherine A. Dye, Vikaas Sohal, Jason E. Long, Rosanne C. Estrada, Tomas Roztocil, Thomas Lufkin, Karl Deisseroth, Scott C. Baraban, and John L. R. Rubenstein. 2010. "Dlx5 and Dlx6 Regulate the Development of Parvalbumin-Expressing Cortical Interneurons." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 30 (15): 5334–45.

Xiang, C., V. Baubet, S. Pal, L. Holderbaum, V. Tatard, P. Jiang, R. V. Davuluri, and N. Dahmane. 2012. "RP58/ZNF238 Directly Modulates Proneurogenic Gene Levels and Is Required for Neuronal Differentiation and Brain Expansion." *Cell Death and Differentiation* 19 (4): 692–702.

Yingling, Jessica, Kazuhito Toyo-Oka, and Anthony Wynshaw-Boris. 2003. “Miller-Dieker Syndrome: Analysis of a Human Contiguous Gene Syndrome in the Mouse.” *The American Journal of Human Genetics* 73 (3): 475–88.

Yingling, Jessica, Yong Ha Youn, Dawn Darling, Kazuhito Toyo-Oka, Tiziano Pramparo, Shinji Hirotsune, and Anthony Wynshaw-Boris. 2008. “Neuroepithelial Stem Cell Proliferation Requires LIS1 for Precise Spindle Orientation and Symmetric Division.” *Cell* 132 (3): 474–86.

Zeisel, Amit, Ana B. Muñoz-Manchado, Simone Codeluppi, Peter Lönnerberg, Gioele La Manno, Anna Juréus, Sueli Marques, et al. 2015. “Brain Structure. Cell Types in the Mouse Cortex and Hippocampus Revealed by Single-Cell RNA-Seq.” *Science (New York, N.Y.)* 347 (6226): 1138–42.

Zhu, Ying, André M. M. Sousa, Tianliuyun Gao, Mario Skarica, Mingfeng Li, Gabriel Santpere, Paula Esteller-Cucala, et al. 2018. “Spatiotemporal Transcriptomic Divergence across Human and Macaque Brain Development.” *Science* 362 (6420). https://doi.org/10.1126/science.aat8077.

Ziffra, Ryan S., Chang N. Kim, Amy Wilfert, Tychele N. Turner, Maximilian Haeussler, Alex M. Casella, Pawel F. Przytycki, et al. 2020. “Single Cell Epigenomic Atlas of the Developing Human Brain and Organoids.” *Cold Spring Harbor Laboratory*. https://doi.org/10.1101/2019.12.30.891549.

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:
Carmen Sandoval Espinosa
7661B1A1C562427...
Author Signature

12/15/2021
Date