From the Department of Medical Biochemistry and Biophysics
Karolinska Institutet, Stockholm, Sweden

TRANSCRIPTOMIC CELLULAR DIVERSITY OF THE EARLY HUMAN DEVELOPING BRAIN

Emelie Paula Braun

Karolinska Institutet

Stockholm 2023

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2023

ISBN 978-91-8016-929-5
Cover illustration: *"A growing brain in an embryo, drawn in Chagall style"* by Midjourney AI-based art generator

TRANSCRIPTOMIC CELLULAR DIVERSITY OF THE EARLY HUMAN DEVELOPING BRAIN

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Emelie Braun

The thesis will be defended in public on **Friday February 17th, 2023, at 09:30**
At **Eva & Georg Klein Lecture Hall**, Solnavägen 9, Karolinska Institutet

Principal Supervisor:
Prof. Sten Linnarsson
Karolinska Institutet
Department of Medical Biochemistry and Biophysics
Division of Molecular Neurobiology

Co-supervisor(s):
Prof. Ernest Arenas
Karolinska Institutet
Department of Medical Biochemistry and Biophysics
Division of Molecular Neurobiology

Kimberly Siletti, PhD
Karolinska Institutet
Department of Medical Biochemistry and Biophysics
Division of Molecular Neurobiology

Opponent:
Assistant Prof. Aparna Bhaduri
University of California Los Angeles
Department of Biological Chemistry

Examination Board:
Prof. Thomas Perlmann
Karolinska Institutet
Department of Cell and Molecular Biology

Prof. Taija Mäkinen
Uppsala University
Department of Immunology, Genetics and Pathology
Division of Vascular Biology

Prof. András Simon
Karolinska Institutet
Department of Cell and Molecular Biology

To my parents,

and my grandparents

Every object that biology studies is a system of systems.

Francois Jacob (1974)

POPULAR SCIENCE SUMMARY OF THE THESIS

The mammalian brain is one of the most complex organs consisting of around 200 billion brain cells. These cells are diverse in terms of their physical appearance and function in the body which is manifested in the characters and behaviors of an animal. Extensive work in the neuroscience field has been done to categorize this cell diversity into distinct cell types based on the features described above. Broadly, all brain cells in mammals can be grouped into two classes: neurons and glial cells. Neurons are responsible for the electrical impulses that drives the communication between neurons and different brain regions, while glial cells are traditionally described to serve as support and protection for neurons across the brain.

To understand when and how brain cell type diversity emerges we have to study the brain in its developing phase, at the embryonic stage. Most of our knowledge about brain development originates from studies in mouse, chicken and zebrafish (vertebrates), but also fruit flies and worms (invertebrates). Despite partial species-conservation, human-specific differences are to be addressed. Some of these differences include higher cognitive functions and it is clear that human brain development differs markedly from that of common model organisms, considering their long gestational period. The brain derives from a specific cell-layer of the embryo, *ectoderm*, and appears during the third week after fertilization (i.e. postconception) in humans. This cellular structure undergoes morphological changes during development upon exposure to molecular stimuli, until finally reaching its mature state in the form of a brain.

In this thesis, we studied the emergence of cell type diversity in the human developing brain during the 5 to 14 postconceptional weeks i.e. during the first and beginning of second trimester. We used single-cell RNA-sequencing to address this question, a technique commonly used for assessing cell type composition and diversity in tissues. This method relies on the fact that specific cell types have distinct signatures of gene activity – whether a gene is turned "on or off" –, which relates to their function in the tissue they derive from. Using this technique, we isolated individual single cells from different brain regions and measured the gene activities in each of these cells. Thereafter, we computationally grouped cells together based on their similarities in gene activity, and thus displayed their cell diversity. We collected nearly two million cells from all major brain regions. Overall, we mostly found immature cell types, in which some of them had partially acquired their characteristic mature gene signatures. We identified major cell types known to exist in the brain, with neurons appearing at week 5 postconception. Surprisingly, we found a large number of immature glial cells that were region-specific and emerged around week 6 postconception. Some of these glial cell patterns were observed in the adult mouse brain, which is highlighted in the first study of this thesis. Moreover, gene activity patterns change over time as the brain develops, and we show that this is measurable by resolving the cell division dynamics of specific cell types in the forebrain that are undergoing maturation. The second study in this thesis describes how to quantify how fast these changes occur in single cells. To summarize, we show that single-cell RNA sequencing is a powerful tool to study cell dynamics and that our findings may have future implications for understanding the human developing brain in both a healthy and diseased context.

POPULÄRVETENSKAPLIG SAMMANFATTNING

Däggdjurets hjärna är ett av de mest komplexa organen och består av omkring 200 miljarder hjärnceller. Hjärncellerna har olika egenskaper baserat på sina fysiska utseenden och vilka funktioner de har i kroppen. Dessa egenskaper ligger till grund för olika beteendemönstren hos djur. Omfattande studier inom neurovetenskapen har gjorts för att klassificera celler som distinkta celltyper baserat på de ovannämnda egenskaperna. Översiktligt, kan alla mammaliska hjärnceller delas in i två grupper: nervceller och gliaceller. Nervcellerna kommunicerar med andra nerverceller genom att skicka elektriska impulser mellan olika hjärnregioner, medan gliacellerna traditionellt fungerar som stöd och skydd för nervcellerna i hjärnan.

För att förstå när och hur hjärnans celltyper uppstår, måste vi studera hjärnan i dess utvecklingsfas på embryonalstadiet. Majoriteten av vår kunskap om hjärnans utveckling härrör från studier på mus, kyckling och zebrafisk (ryggradsdjur), men även från bananflugor och rundmaskar (ryggradslösa djur). Trots likheter mellan olika arter, är de människospecifika skillnaderna inte fullt kartlagda. Några av dessa skillnader utgörs av våra högre kognitiva förmågor och med tanke på människans långa dräktighetsperiod är det påtagligt att mänsklig hjärnutveckling skiljer sig markant från den hos de ovannämnda modellorganismerna. Hjärnan utvecklas från ett specifikt cellager på embryot, *ektodermen*, som uppstår under den tredje veckan efter befruktningen. Denna cellstruktur genomgår åtskilliga morfologiska förändringar under utvecklingen och exponeras för olika molekylära stimuli tills den slutligen når sitt mogna tillstånd i form av en hjärna.

Denna avhandling avser att studera uppkomsten av celltyper i människans hjärna under utvecklingen från vecka 5 till 14 (efter befruktningen), dvs. under den första- och början på andra trimestern. Vi använde oss av enkelcell-RNA-sekvensering för att studera denna process, en välkänd teknik som användas för att studera kompositionen av celltyper i olika vävnader. Metoden bygger på att specifika celltyper har olika gener som är aktiva – huruvida en gen är ”på eller av” – vilket återspeglas i deras funktion och vävnaden de kommer ifrån. Med denna teknik isolerade vi enskilda celler från alla huvudregioner i hjärnan och mätte genaktiviteten i varje enskild cell. Därefter grupperade vi celler med lika genaktivitet digitalt, dessa grupper kom sedan att definiera alla celltyper. Totalt mätte vi nära två miljoner celler. Överlag, såg vi mestadels omogna celler, varav en del hade antagit mognare karaktäristiska egenskaper baserat på sina genetiska drag. Vi identifierade majoriteten av celltyper som förekommer i hjärnan och såg att nervceller kan observeras vid vecka 5. Oväntat nog, såg vi ett större antal omogna regionspecifika gliaceller som framträdde vid vecka 6. Dessa gliacell-mönster påminde om de vi såg i den vuxna mushjärnan som berörs i den första artikeln i denna avhandling. Vidare, kan genernas aktivitet ändras över tid som hjärnan utvecklas. Vi mätte denna förändring under celldelningsfaserna hos specifika celltyper i framhjärnan som genomgick en mognadsprocess. Den andra studien i denna avhandling beskriver en metod för att mäta hastigheten på denna förändring i genaktivitet hos enskilda celler. Sammanfattningsvis, utgör enkelcell-RNA-sekvensering en kraftfull teknik för att kunna studera cellers dynamik. Våra upptäckter hoppas kunna öka förståelsen för mänsklig hjärnutveckling, både hos den friska och sjuka hjärnan.

ABSTRACT

The complexity of the mammalian brain is partly reflected in its cell type diversity which influences the function of neurons that encode the behavior of animals. Brain cell type diversity emerges during embryonic stages, a critical period when neurons start to become functionally active and establish their connectivity across the brain. Since the pioneering of single-cell RNA-sequencing (scRNA-seq), we can question *when* and *how* cellular diversity arises in the brain in a large-scale manner.

This thesis aims to study the human brain during the first trimester by using scRNA-seq to obtain a global view of the basic principles of the developing brain. First, I introduce human embryology from a historical perspective and summarize key concepts in central nervous system (CNS) development. I review few gaps in the field related to our findings, followed by current approaches and nomenclatures used in the field of single-cell genomics that applies to development. To put our work into perspective, I present an overview of the latest efforts to study human brain development at the single-cell level, both in the healthy and diseased brain.

Then I present the following two papers and a manuscript:

In **Paper I** we used scRNA-seq to construct a cell taxonomy of the adult mouse nervous system. We describe two major groups: neuronal- and non-neuronal cells that were subdivided into distinct cell types. Overall, the neurons were transcriptionally similar across brain regions, whereas non-neuronal cells such as astrocytes, formed subgroups and were regionally distinct. The whole dataset revealed an organization that reflects the developmental origin of all cell types.

Paper II describes a method, *RNA velocity*, that infers temporal changes from static scRNA-seq gene expression measurements. By realigning sequencing reads, this method detects and makes use of the unspliced and spliced mRNA, whose relative abundance is used to measure the change of rate in gene expression (the time derivative) in different tissues. This method is particularly suitable for developmental lineages, which was shown and validated both *in vitro* and *in situ* in this study.

Paper III presents a single-cell atlas of the human developing CNS across all major brain regions during postconceptional weeks (p.c.w.) 5 to 14. We observe that major cell classes emerge during this period, most of them being regionally diverse and to a surprisingly high degree among glial cells. We display the high resolution of this data by resolving several lineages in the forebrain and validated the spatial location of transcriptional cell types at 5 p.c.w. by using single-molecule FISH. As a whole, this study represents a reference of human brain development during the first critical period in life.

To tie these studies together, our findings on glial diversity were partially shared between the adult mouse and developing human CNS. We further showed that an RNA velocity-based method can be used to model the cell cycle dynamics in cortical tissue. To conclude, I discuss advantages and limitations of single-cell transcriptomics, its future challenges and how using this technology sheds light on the early human developing brain as is described in this thesis.

LIST OF SCIENTIFIC PAPERS

I. *Amit Zeisel, Hannah Hochgerner, Peter Lönnerberg, Anna Johnsson, Fatima Memic, Job van der Zwan, Martin Häring,* ***Emelie Braun****, Lars E. Borm, Gioele La Manno, Simone Codeluppi, Alessandro Furlan, Ka Wai Lee, Nathan Skene, Kenneth D. Harris, Jens Hjerling-Leffler, Ernest Arenas, Patrik Ernfors, Ulrika Marklund, Sten Linnarsson*

Molecular Architecture of the Mouse Nervous System
Cell 2018 August, 174 (4), 999-1014

II. *Gioele La Manno, Ruslan Soldatov, Amit Zeisel,* ***Emelie Braun****, Hannah Hochgerner, Viktor Petukhov, Katja Lidschreiber, Maria E. Kastriti, Peter Lönnerberg, Alessandro Furlan, Jean Fan, Lars E. Borm, Zehua Liu, David van Bruggen, Jimin Guo, Xiaoling He, Roger Barker, Erik Sundström, Goncalo Castelo-Branco, Patrick Cramer, Igor Adameyko, Sten Linnarsson, Peter Kharchenko*

RNA Velocity of Single Cells
Nature 2018 August, 560 (7719), 494–98

III. ***Emelie Braun****, *Miri Danan-Gotthold**, Lars E. Borm, Elin Vinsland, Ka Wai Lee, Peter Lönnerberg, Lijuan Hu, Xiaofei Li, Xiaoling He, Žaneta Andrusivová, Joakim Lundeberg, Ernest Arenas, Roger A. Barker, Erik Sundström, Sten Linnarsson*

Comprehensive Cell Atlas of the First-Trimester Developing Human Brain
bioRxiv*,* 2022 October (In revision)

*equal contribution

PUBLICATIONS NOT INCLUDED IN THIS THESIS

Muus C, Luecken M. D., Eraslan G, Sikkema L, Waghray A, Heimberg G, Kobayashi Y, Dhaval Vaishnav E, Subramanian A, Smillie C, Jagadeesh K. A., Thu Duong E, Fiskin E, Torlai Triglia E, Ansari M, Cai P, Lin B, Buchanan J, Chen S, Shu J, Haber A. L., Chung H, Montoro D. T., Adams T, Aliee H, Allon S. J., Andrusivova Z, Angelidis I, Ashenberg O, Bassler K, Bécavin C, Benhar I, Bergenstråhle J, Bergenstråhle L, Bolt L, ***Braun E****, Bui L. T., Callori S, Chaffin M, Chichelnitskiy E, Chiou J, Conlon T. M., Cuoco M. S., Cuomo A. S. E., Deprez M, Duclos G, Fine D, Fischer D. S., Ghazanfar S et al.… Ziegler C. G. K., The NHLBI LungMap Consortium and The Human Cell Atlas Lung Biological Network*

Single-Cell Meta-Analysis of SARS-CoV-2 Entry Genes across Tissues and Demographics
Nature Medicine, 2021 March, 27 (3), 546–59

Sountoulidis A, Marco Salas S, ***Braun E****, Avenel C, Bergenstråhle J, Theelke J, Vicari M, Czarnewski P, Liontos A, Abalo X, Andrusivová Ž, Mirzazadeh R, Asp M, Li X, Hu L, Sariyar S, Martinez Casal A, Ayoglu B, Firsova A, Michaëlsson J, Lundberg E, Wählby C, Sundström E, Linnarsson S, Lundeberg J, Nilsson M, Samakovlis C*

A Topographic Atlas Defines Developmental Origins of Cell Heterogeneity in the Human Embryonic Lung
Nature Cell Biology, 2022 January

Li X, Andrusivová Ž, Czarnewski P, Mattsson Langseth C, Andersson A, Liu Y, Gyllborg D, ***Braun E****, Larsson L, Hu L, Alekseenko Z, Lee H, Avenel C, Kopp Kallner H, Åkesson E, Adameyko I, Nilsson M, Linnarsson S, Lundeberg J, Sundström E*

Decoding Spatiotemporal Gene Expression of the Developing Human Spinal Cord and Implications for Ependymoma Origin
bioRxiv, 2022 September (Accepted in Nature Neuroscience)

CONTENTS

LIST OF ABBREVIATIONS

BMP	Bone Morphogenetic Protein
CNS	Central Nervous System
EEL-FISH	Enhanced ELectric in situ Hybridization
EGFR	Epidermal Growth Factor Receptor
ISVZ	Inner Subventricular Zone
IZ	Intermediate Zone
KNN	*k*-Nearest Neighbour
nIPC	neuronal Intermediate Progenitor Cell
OPC	Oligodendrocyte Precursor Cell
OSVZ	Outer Subventricular Zone
PC	Principal Component
PCA	Principal Component Analysis
p.c.w.	postconceptional week
PNS	Peripheral Nervous System
scATAC-seq	Single-cell Assay for Transponsase-Accessible Chromatin sequencing
scRNA-seq	Single-cell RNA sequencing
SHH	Sonig Hedgehog
SVZ	Subventricular Zone
TF	Transcription Factor
t-SNE	t-distributed Stochastic Neighbor Embedding
UMAP	Uniform Manifold Approximation and Projection
UMI	Unique Molecular Identifier
VZ	Ventricular Zone

1 MODERN EMBRYOLOGY AND DEVELOPMENTAL NEUROSCIENCE

1.1 INTRODUCTION

Almost all animals on Earth, excluding porifera (sponges) and placozoans (Trichoplaxes) have a nervous system (Moroz 2009; Kristan 2016). The most primitive nervous systems are found in the *Cnidarians*, such as the Hydra (Ji and Flavell 2017). They have so called "diffuse nerve nets", a group of nerve cells that assemble into ganglia, but lack a cephalic structure with a CNS. The evolutionary benefit of a nervous system is highlighted in the behaviours of animals, such as response to various stimuli, communication skills and physical interactions. The brain is arguably the most complex organ of the human body. This is supported by the fact that the nervous system is composed of a vast number of heterogeneous cell populations that signal through electrochemical impulses and together orchestrates the body to function. A major effort has been applied to understand how the collective activity of neurons gives rise to the behaviour of an animal and modern neuroscience has allowed us to couple anatomical, physiological and molecular features of neurons and glia (Südhof 2017). Developmental neuroscience is important for the understanding of how the mature brain gains its properties and functions, but can also teach us about the conservation of the evolutionary traits of the nervous system (Kristan 2016). Due to the nature of the embryology field, the difficulty in obtaining biological specimens has resulted in knowledge mostly about development of specific vertebrates and invertebrates used as model organisms, such as *Gallus gallus*, *Danio rerio*, *D. melanogaster* or *C. elegans*.

1.1.1 Historical perspectives

The interest in embryology can be dated back to the ancient Greeks, like Aristotle and Hippocrates who laid the groundwork for what we refer to as *epigenesis* today (Horder 2010). The development of an organism has continued to captivate many scientists throughout history, taking Leonardo Da Vinci's anatomical drawings of a foetus as an example and later scientists during the Enlightenment period (Horder 2010). However, it was not until the time of Robert Remak, Christian Pander and Karl Ernest von Baer that the germ layer theory was confirmed and modern embryology began to develop (Horder 2010; Dupont 2017; Hopwood 2007). Studies of embryos had been mostly descriptive until the development of new staining techniques and more advanced microscopes. Embryology was not unified as a scientific discipline during most of the 19th century, but was for a long time divided into comparative and experimental embryology (even referred to as 'developmental mechanics') (Horder 2010; Hopwood 2007; 2019a). With improved preservation techniques and the foundation of cell theory, development could be studied in terms of cells and not only germ layers (Dupont 2017; Schleiden et al. 1847). Discoveries made in genetics during the early 1900's, also contributed to the understanding of development, as many genes were found vital for the development of an organism (Gluecksohn-Schoenheimer 1940; Waddington 1939; Gilbert 2000a). During the

same time period, the Carnegie Institution made an immense effort to systematically make a descriptive collection of human development that are still used today (Hopwood 2019b; de Bakker et al. 2016). The Kyoto collection founded in 1961 along with newly founded tissue resources have contributed greatly to our understanding of human development (de Bakker et al. 2016; Hill 2018; Yasuda 2018; Lindsay et al. 2016; Kerwin et al. 2010).

Not surprisingly, embryology and developmental neurobiology – often termed neuroembryology – were tightly linked since the nervous system develops very early. The discovery of 'growth cones' in human and chicken revealed that the axonal ends of neurons were guided towards its target in a chemotactic manner as a result of secreting substances by surrounding cells along the way (Hamburger 1981; Louis and Stapf 2001; Garcia-Marin, Garcia-Lopez, and Freire 2009; Morange 2009). This is in line with today's knowledge about axon guidance and the growth of neurons, where secreted proteins like Netrins, Ephrins and Semaphorins either attract or repel growth cones by interacting with their respective receptors (Raper 2009). Modern developmental neurobiology has become more niched since the discovery of the *organizer region,* that induces the formation of the neural plate and gives rise to the nervous system (Spemann and Mangold 1924). Although we know more about the patterning mechanisms, genetics and anatomy of brain development today, there is still a lack of knowledge of the fundamental processes that govern the morphogenesis of the CNS.

1.1.2 Experimental approaches throughout the years

Traditionally, neurodevelopmental studies have been conducted by transplanting tissues within the same or between different species to identify a particular organizer region (Solini, Dong, and Saha 2017; Martinez Arias and Steventon 2018). Fate mapping experiments such as the generation of chick-quail chimeras have also been of importance to study the ontogeny of cells (Balaban, Teillet, and Douarin 1988; Matsushita 1996). Today, chromogenic staining techniques, immunocyto- and immunohistochemistry to detect neural-specific antigens *in situ*, the study of cellular interactions and the potency of neural stem cells (NSC) *in vitro* or knocking out or overexpressing genes one by one, are all common practices in developmental biology (Mullen, Buck, and Smith 1992; Vernes et al. 2011; Ivanov et al. 2004; Heisenberg, Brennan, and Wilson 1999; Honig, Herrmann, and Shatz 1996; Breunig, Haydar, and Rakic 2011; Tsien et al. 1996; Nüsslein-volhard and Wieschaus 1980). These methodologies have enabled the creation of genetically engineered lines in various animal models, which has given us more insight into important genetic markers in the CNS and made it possible to study cellular lineages and trace different cell populations over a given period of time (Kretzschmar and Watt 2012). As development is a highly dynamic process, dependent on both spatial and temporal organization, reductionistic models such as *C. elegans* have served useful to study lineages and cell fate. It has an advantage because of its transparency, which allows for a continuous direct observation of cell fates in an intact whole organism embryo which is also non-invasive. In contrast to other species, *C. elegans* make autonomous cell-fate decisions which facilitates lineage tracing (Kretzschmar and Watt 2012; Byrum, Wijesena, and Wikramanayake 2012; Sulston et al. 1983). However, direct imaging of embryogenesis in an intact mammalian

embryo is not feasible in that way. Therefore, fate mapping and lineage tracing of the CNS in 'higher level' organisms like mice or rats, have classically often been performed through the isolation of NSCs from early stage embryos that were cultured and grown *in vitro* (Temple 1989; J. Liu et al. 2013). As a natural consequence, fewer similar studies have been done in human, but recent advances in growing brain organoids and organotypic slice cultures of human embryonic brain tissue may provide valuable insights into brain development and eventual pathologies (Andrews and Nowakowski 2019; Bershteyn and Kriegstein 2013).

Today, developmental neurobiology is facing a new field of high-throughput technologies like single-cell genomics and spatial transcriptomics that may reveal new answers to biological questions that were not possible to answer before (Tanay and Regev 2017; Haniffa et al. 2021). In contrast to reductionistic models, these techniques go along with systems biology that describes biological information at the level of a whole organism, tissue or cell, often without *a priori* knowledge or hypotheses. Thus, developmental neurobiology has become more hypothesis-generating allowing for a top-down approach in order to answer relevant questions in a more unbiased manner.

2 DEVELOPMENT OF THE CENTRAL NERVOUS SYSTEM

2.1 PATTERNING AND MORPHOGENESIS OF THE NEURAL TUBE

The vertebrate CNS originates from the *ectoderm*, the outermost germ layer of the embryo, that under different gene expression programs either develops into epidermal cells or progresses into the formation of the central- or peripheral nervous system (PNS). Early in development, part of the ectoderm becomes specialized into the *neuroectoderm,* consisting of columnar neuroepithelial cells that undergo developmental programs to form diverse groups of neurons and glia (Kandel et al. 2012). Ectodermal cells commit to the neuronal lineage after being induced by signalling molecules – also termed *morphogens* – by surrounding cells. Thereafter, the neuroectoderm forms the *neural plate,* the most dorsal part of the ectoderm, that is made up by the neural precursors. Through the process of *neurulation*, the neural plate makes an invagination towards the *mesoderm* and closes dorsally at its midline, resulting in the *neural tube* (Figure 1A). Neurulation is initiated by the secretion of *bone morphogenetic proteins* (BMPs) that are released from ectodermal cells which promotes the differentiation of epidermal cells. Neuronal differentiation will occur as a result of inhibition of BMP signalling (Kandel et al. 2012). Cells from the organizer region suppress BMP signalling through the release of BMP antagonists (chordin, noggin and follistatin) and will instead drive ectodermal cells into neuronal differentiation, which has been shown to be the "default" program (Grove and Monuki 2013). There are several local so-called *organizers,* these are constituted by cells that secrete morphogens that form concentration gradients along the neuraxis and through diffusion induce morphogenesis of the neural tube throughout development. Hence, the fate of a cell in

the progression of development is determined by its position along the neuraxis as well as its competence, how it responds to secreted signals through receptors, that in turn activates transcription factors (TFs) and activate gene expression programs for the differentiation of diverse cell types.

As the neural tube closes, waves of neural proliferation and differentiation take place (Figure 1A, B). Neural tube closure occurs around p.c.w. 3 in human, with the anterior to posterior part developing into forebrain and hindbrain respectively (Kandel et al. 2012). By the action of morphogenetic signalling and cellular interactions, the neural tube becomes patterned and shaped into the distinct brain regions with various cell types. Vesicle formation of the tube-like structure arises during p.c.w. 4, giving rise to three brain vesicles: forebrain (prosencephalon), midbrain (mesencephalon) and hindbrain (rhombencephalon). At a later stage, two more vesicles bud out, where the forebrain divides into telencephalon and diencephalon whereas hindbrain will split into metencephalon (pons and cerebellum) and myelencephalon (medulla). These main structures together with the spinal cord, make up the mature adult CNS and obtain their specialized functions through the continuous exposure to secreted signals in the patterning events during development.

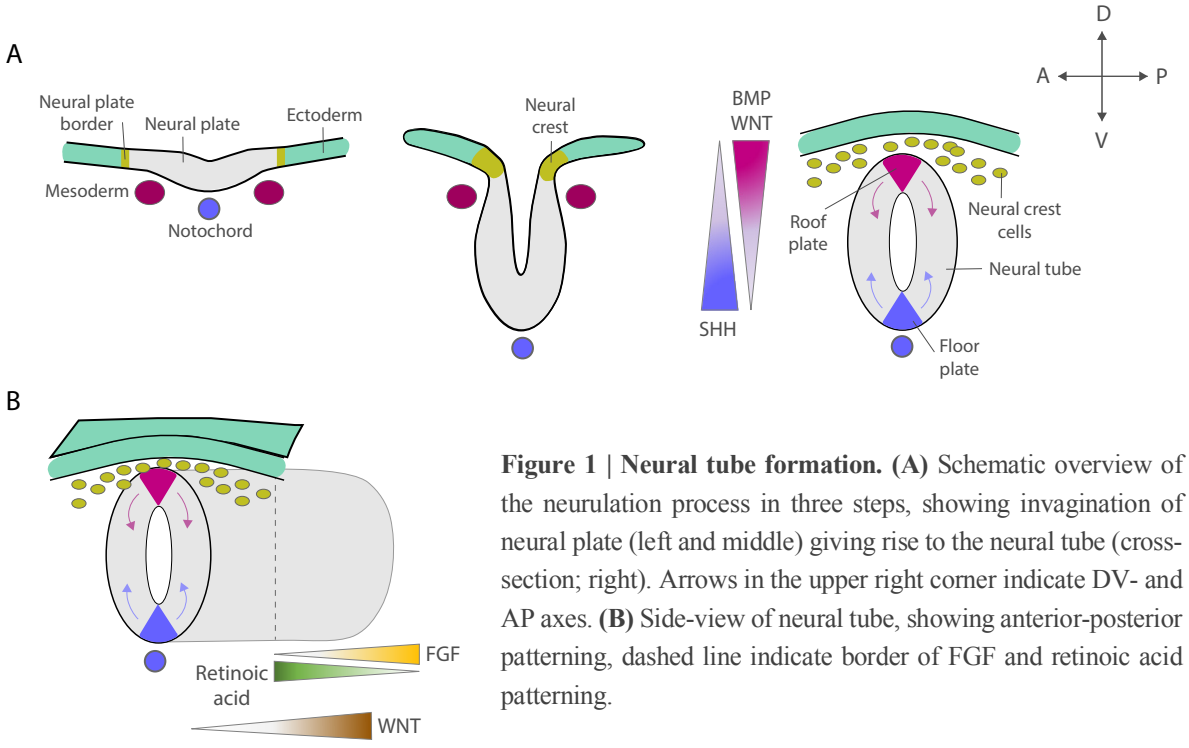

Figure 1 | Neural tube formation. (A) Schematic overview of the neurulation process in three steps, showing invagination of neural plate (left and middle) giving rise to the neural tube (cross-section; right). Arrows in the upper right corner indicate DV- and AP axes. **(B)** Side-view of neural tube, showing anterior-posterior patterning, dashed line indicate border of FGF and retinoic acid patterning.

2.2 ANTERIOR–POSTERIOR AND DORSAL–VENTRAL PATTERNING

Regionalization of the neural tube is specified along its anteroposterior (AP) and dorsoventral (DV) axes early in development (Figure 1). Before closure, the neural plate is already beginning to take shape along its AP axis. This process is directed by mesodermal and endodermal cells. AP patterning is partially established by a gradient of WNT signalling proteins with a low concentration anteriorly that is gradually increasing towards the posterior part of the neural

tube. The AP axis is further refined by *secondary organizers* that are involved in patterning the subregions of the brain. The *anterior neural ridge*, located in the most anterior part of the tube, is one of the organizers secreting FGF8 and is involved in the patterning of telencephalon whereas retinoic acid has been described in posteriorizing the neural tube (Figure 1B; Tole and Hébert 2013). As a result of patterning signals, the neural tube is divided into regionally defined TF domains, where cells in the forebrain and midbrain express *Otx2* while the hindbrain domain is defined by *Gbx2* expression. A midbrain/hindbrain border is established that forms the *isthmic organizer* which is defined by its expression of *Engrailed 1* (*En1*) (Nakamura 2013). Another organizer, *zona limitans intrathalamica* secretes *sonic hedgehog* (SHH) and is involved in the formation of thalamus (Grove and Monuki 2013).

While the AP axis is forming the anatomical subdivisions of the future brain, the DV patterning defines the neural progenitors that will mature into distinct classes of neurons. This mechanism has been extensively described in the vertebrate spinal cord (Wilson and Maden 2005). The DV patterning is initiated by the formation of the *notochord*, a ventral mesoderm originating structure, extending along the medial axis of the neural tube. The notochord secretes SHH and induces floor plate cells to secrete and form a dorsal-to-ventral gradient of this morphogen, with the highest concentration ventrally (Figure 1A). The dorsal signalling centre in the midline of the neural tube is made up by the roof plate which controls the specification of dorsal neuronal subtypes. The roof plate cells which derive from epidermal ectoderm, secrete BMP4, BMP5 and BMP7 as well as WNT (Figure 1A; Wilson and Maden 2005). BMP proteins also induce the formation of neural crest cells, positioned between the neural plate and the non-neuronal ectoderm, that upon neural tube closure migrate from the dorsal side towards the periphery destined to become cells of the PNS, but also melanocytes, craniofacial cartilage, bone and smooth muscle (Figure 1A, B; Gilbert 2000b). DV patterning results in the formation of different domains of TFs that will give rise to distinct cell types. The dorsal spinal cord will differentiate into dorsal interneurons, important for relaying somatosensory information, while the ventral side produces motor- and ventral interneurons. This is largely controlled by the SHH gradient forming along the DV axis which induces activation of early TFs such as *Nkx2.2*, *Olig2*, *Irx3* and *Dbx2* ventrally. The dorsal side expresses another set of TFs, some of them being *Pax2*, *Lbx1* and *Lhx1* (Wilson and Maden 2005; Hernandez-Miranda, Müller, and Birchmeier 2017). Many of these TFs function in a cell-autonomous manner, meaning that they repress each other's function. In addition, the development of motor neuron subtypes is also dependent on the *Hox* gene family that is expressed along the AP axis of the neural tube (Tümpel, Wiedemann, and Krumlauf 2009). The DV patterning mechanism of the spinal cord is thought to remain similar in other brain regions.

The action of morphogens on CNS patterning is a very complex biological process. Many studies have provided more insight into morphogenetic processes, but there is still a lack of a mechanistic understanding of how a tissue can establish and maintain concentration gradients over a longer distance as well as conduct sharp changes in gene expression programs (Rogers and Schier 2011; Grove and Monuki 2013). A few studies have measured the effect of

morphogenetic gradients on human embryonic stem cells, using different strategies to administer controlled doses of WNT or gradually induce SHH through optic stimulation (Rifes et al. 2020; De Santis et al. 2021). Furthermore, earlier controversial models explaining this process, such as the Turing model, have regained attention and become more widely accepted among developmental biologists with the accumulated experimental evidence supporting this model (Kondo 2022). The single-cell genomic field has opened up a new approach to study these dynamical processes. Single-cell RNA sequencing (scRNA-seq) is starting to become widely used in order to comprehend and build models of cellular states and transitions. This could lead to an increased understanding of the effect of morphogens by measuring the gene regulatory response in tissues.

2.3 THE PROSOMERIC MODEL

Different models that describe the neuroanatomical development of the neural tube have evolved from the past (Hidalgo-Sánchez et al. 2021). Today, the prosomeric model constitutes a fundamental concept to explain the anatomy of the developing nervous system. The prosomeric model, that originally describes the structural division of the embryonic forebrain (Rubenstein et al. 1994; Bulfone et al. 1993; Luis Puelles and Rubenstein 1993) based on distinct morphological segments and expression patterns of several homeobox genes in vertebrates, have been extended to the mammalian brain (Thompson et al. 2014; Luis Puelles et al. 2013). In this model, the neural tube is recognized as a segmental structure that is subdivided into a transverse- and longitudinal axis. The transversal segments, also referred to as neuromeres, are distinct units of different regions along the neural tube: prosomeres (prosencephalon), mesomeres (midbrain), rhombomeres (hindbrain) and myelomeres (spinal cord) (Figure 2). These developmental units have traditionally been described anatomically, but have later been found to be molecularly distinct. The model also describes a shared DV pattern along the longitudinal axis of the neural tube, including zones like the floor, roof, alar and basal plates that extends throughout the neural tube. Extensive work has been done to develop and update the prosomeric model with new findings such as gene expression patterns specific to substructures of the neural tube. As explained in Paper III, we have partially relied on The Allen Developing Mouse Brain Atlas – who also base their anatomical annotation on the prosomeric model – to define anatomical units in our human brain developmental data as well (Figure 2). This kind of anatomical model has served to be particularly important and useful to understand the species-specific and evolutionary aspect of brain development. Furthermore, this model has been further developed to serve as an ontology of mammalian brain development (Luis Puelles et al. 2013). Moreover, the gene expression specificity seen in certain brain regions is helpful in scenarios where precise dissections are difficult to obtain. This is illustrated in Paper III, where some dissected regions were overlapping and we use the distinct molecular patterns to computationally dissect pallium (*FOXG1, EMX1*) and subpallium (*FOXG1, DLX2*) for region-specific analyses of several cell lineages.

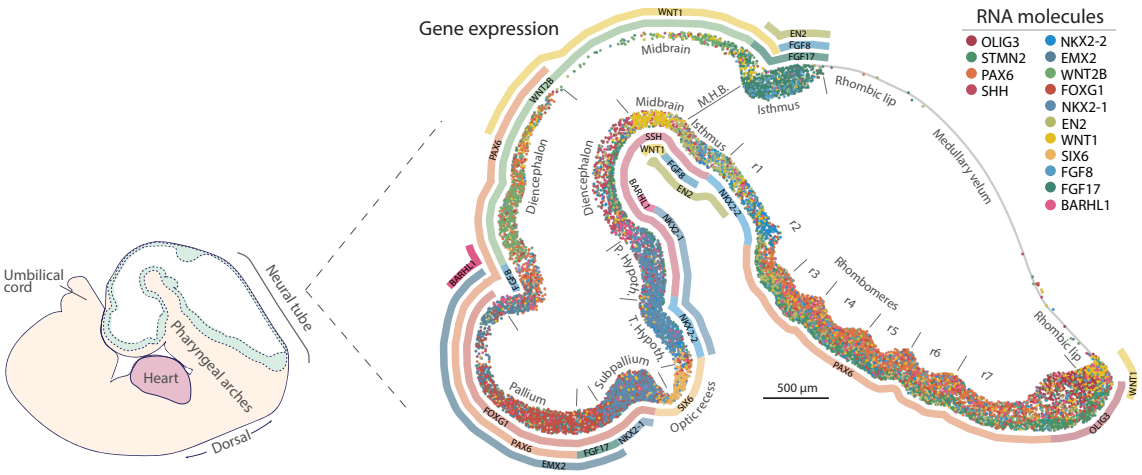

Figure 2 | Neural tube at 5 p.c.w. in human. Right: illustration of a whole embryo at 5 p.c.w. (sagittal view), the neural tube is highlighted in lightgreen (empty/ventricular space in white). Left: zoom-in of the neural tube highlighting 16 genes measured by Enhanced Electric in situ Hybridization (EEL-FISH; dots representing mRNA molecules colored by gene) that were used for annotating regions according to the prosomeric model Ribbons indicate gene expression domains. M.H.B.; midbrain-hindbrain border, P. Hypoth.; peduncular hypothalamus T. Hypoth.; terminal hypothalamus. Illustrations adapted and reused form Braun, Danan-Gotthold et al. 2022 (Paper III).

2.4 RADIAL GLIA AND NEUROGENESIS

Radial glia descends from the ectodermal-derived neuroepithelial cells and are the neuronal progenitors, i.e. neural stem cells in the brain. They were already morphologically described in the 1800s and are named after their characteristic bipolar-shaped appearance with extended long radial processes, their radial way of migrating and dividing (as opposed tangential cell migration in the brain) and their resemblance to glial cells during development (Kriegstein and Alvarez-Buylla 2009). They have long attracted the attention of developmental neuroscientists because they give rise to both neurons and macroglial cells such as astrocytes, oligodendrocytes and ependymal cells that lines the ventricles and produces cerebrospinal fluid in the brain (Kriegstein and Alvarez-Buylla 2009). During early brain development, the neural tube is mostly comprised of a hollow cavity, with the ventricular zones lining this cavity that later will form the brain ventricles. Radial glia are located in the ventricular zones of the neural tube and are thought to undergo neurogenesis in all brain regions during development (Anthony et al. 2004; Malatesta et al. 2003). However, most studies have focused on the radial glial cells in cortical development. The cerebral cortex has long sparked interest in neuroscientists as this is the region thought to underlie our higher cognitive capabilities and is considered to be the origin of many neurodevelopmental disorders (Subramanian, Calcagnotto, and Paredes 2020; Juric-Sekhar and Hevner 2019; Rubenstein 2011). Moreover, cortical development follows a well-organized plan that is favourable in experimental settings. The cerebral cortex is the last to reach maturation in relation to other brain regions which naturally means studying later

timepoints, usually including the end of the first- up to the third trimester in human (Cadwell et al. 2019).

The cortical cytoarchitecture stands out with its many cell-specific layers that have distinct functions. In the adult brain, these layers span from the most superficial layer beneath the *pia mater* of the meninges, to the deepest layer bordering the subcortical white matter (Palomero-Gallagher and Zilles 2019). In the adult, these layers are molecularly distinct (Codeluppi et al. 2018), but in development some of them overlap as they have not fully emerged and are not completely transcriptionally distinct. Depending on the developmental stage, only a few layers are present in the cortex (Kriegstein and Alvarez-Buylla 2009). These layers are referred to as the *germinal zones* during development and varies in their structure depending on the region of the neural tube (Gilbert 2000c; Nieuwenhuys and Puelles 2015). Some of these zones are further described in Paper III. As the cortex expands, these layers become more pronounced (Fernández, Llinares-Benadero, and Borrell 2016). The main layers referred to in the cortex are the ventricular-, subventricular- and intermediate- zones (VZ, SVZ, IZ), including the cortical plate (CP) and the marginal zone (MZ; see Figure 3). The marginal zone resides on top of the cortical plate and remains from the early cortical preplate (Tkachenko et al. 2016). The SVZ can further be histologically divided into the inner- and outer subventricular zone (ISVZ; OSVZ), with the latter being specific to gyrencephalic animals (Fernández, Llinares-Benadero, and Borrell 2016). Radial glia reside in the VZ and SVZ and have apical processes that extends all the way to the pial surface. The cortical tissue architecture is expanded through the asymmetric and symmetric cell divisions of radial glia. It has been shown that radial glia undergo asymmetric cell division in the VZ, giving rise to a "self-renewed" radial glia and a neuronal intermediate progenitor cell (nIPC) or a neuron directly, whereas nIPCs usually migrate to the SVZ generating pairs of neurons by dividing symmetrically (Kriegstein and Alvarez-Buylla 2009). This process of indirect- and direct neurogenesis is thought to occur in order to preserve the stem cell pool in the cortex during development. The nIPCs have historically not been so well-defined and often overlap with neuroblasts transcriptionally. In this thesis, in accordance with Paper III, we refer to the nIPCs as a cycling population and neuroblasts (*NHLH1+)* as post-mitotic (Martínez-Cerdeño et al. 2018). Lately, a couple of studies have focused on dissecting these layer-derived radial glia, both anatomically and transcriptionally (Beattie and Hippenmeyer 2017; De Juan Romero and Borrell 2015). In particular, one study suggest that that two kinds of radial glia exist in the VZ and the OSVZ: ventricular- and outer radial glia (Nowakowski et al. 2016), which are transcriptionally distinct with the outer one being more involved in maintaining the stemness (Pollen et al. 2015). Although, cortical radial glia have been and are still extensively studied as new technologies are emerging, they have attracted so much interest among many labs that radial glia cell-research have formed its own niche in the neurodevelopmental field (Malatesta and Götz 2013).

Other tissues such as the meninges that make contact with the radial glia end feet, are also thought to regulate survival of radial glia during cortical development (Radakovits et al. 2009). With that said, it is also worth mentioning that radial glia are thought to give rise to a pool of

astrocytes that are involved in maintaining adult neural stem cells in the brain that are present in the SVZ of the lateral ventricle and the subgranular zone in the hippocampal dentate gyrus (Kriegstein and Alvarez-Buylla 2009). The SVZ in the adult mouse brain also produce neuroblasts that migrate into the olfactory bulb, a process referred to as the rostral migratory stream (Lennington, Yang, and Conover 2003). While one study claims a similar migratory stream in humans (Curtis et al. 2007), this finding was contradicted (Bergmann et al. 2012) and later supported by the finding that neuroblasts and newly produced neurons locate to the striatum likely originating from the SVZ (Ernst et al. 2014).

In summary, these radial glia transitions have mostly been described in cortex, but it is likely that other brain regions undergo similar processes (Anthony et al. 2004; Malatesta et al. 2003). Our findings in Paper III suggest that there are strong regional differences between radial glia that are driven by classical regional patterning genes. Not many studies have looked into the radial glia lineage in other brain regions, but this may be facilitated with the many single-cell datasets that are publicly available and contains different regions of the developing brain in mammals (Manno et al. 2020; Eze et al. 2020; Braun et al. 2022).

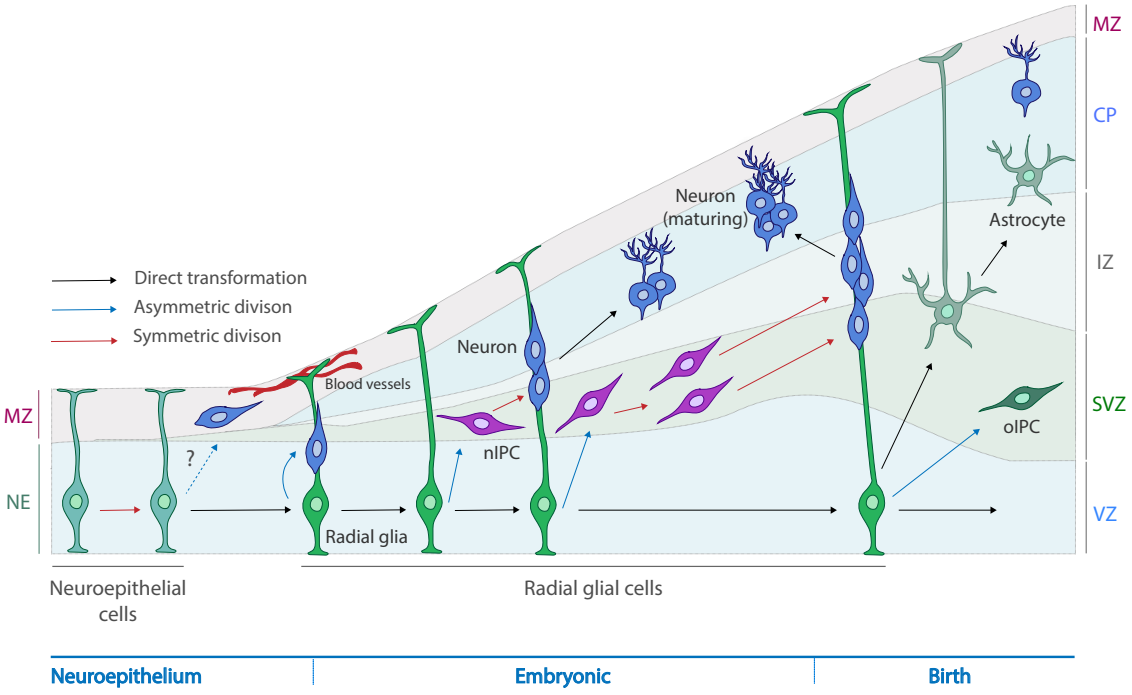

Figure 3 | Mammalian cortical development. Illustration of the radial glia lineage, arrows indicating their cell division modalities resulting in different cell fates. NE; neuroepithelium, MZ; marginal zone, VZ; ventricular zone, SVZ; subventricular zone, IZ; intermediate zone, CP; cortical plate, nIPC; neuronal intermediate progenitor cell, oIPC; oligodendrocyte intermediate progenitor cell. Redrawn from Kriegstein & Alvarez-Buylla, 2009.

2.5 GLIOGENESIS

Radial glia not only give rise to neuronal cells but also glial cells as previously mentioned. For the sake of this thesis, emphasis will be on glial cells derived from the CNS in contrast to neural crest-derived cells from the PNS. Glial cells have its name from *glia* ("glue" in Greek), because they are tightly surrounding neurons and were thought to originally glue neurons together as support (Purves et al. 2001). Gliogenesis is the process in which glial cells are produced by radial glia and go through an intermediate *glioblast* stage to produce astrocytes and oligodendrocytes (Kessaris, Pringle, and Richardson 2008; Delaunay et al. 2008). Microglia belongs to a separate lineage as the innate immune cells of the brain and derive from the yolk sac before migrating to the brain which is thought to occur before the start of gliogenesis during human fetal development (Lanjewar and Sloan 2021). Similar to neuronal progenitors, glioblasts emerge in the SVZ and migrate out and generate astrocytic precursors or oligodendrocyte precursor cells (OPCs) before finally reaching their mature forms (Marshall, Suzuki, and Goldman 2003). However, the transition from radial glia to glial cells are not well understood and many times these populations overlap in the gene expression of certain genes (*GLAST (SLC1A3)*, *BLBP (FABP7*), *TNC, GFAP)*, which complicates the analytical interpretation even with current single-cell technologies. In addition, some studies that have used certain markers like, *GFAP* for lineage tracing of astrocytes, have been shown to not be exclusive for this cell population, but also target radial glia at earlier stages and have been shown to drive the induction of neurons and oligodendrocytes under a human GFAP promoter in mice (Malatesta et al. 2003; Anthony and Heintz 2008; Anthony et al. 2004; Casper and McCarthy 2006).

During development, radial glia switch from neurogenesis into producing glial cells predominantly (Miller and Gauthier 2007; Rowitch and Kriegstein 2010). The mechanism behind the maintenance of neurogenesis during this time is not completely known, although some studies have implicated involvement of changes in the chromatin accessibility through chromatin regulators, microRNAs as well as the influence of BMP-signaling (Nakagawa et al. 2020; Katada et al. 2021; Bronstein et al. 2017). In mice, the neurogenic-to-gliogenic switch occurs around E12 to E16 with gliogenesis peaking mostly at postnatal timepoints (E17.6 to P10) (Bronstein et al. 2017; Manno et al. 2020). In humans on the other hand, neurogenesis can be observed starting roughly at 5 p.c.w. where radial glia switch to a predominantly gliogenic fate around 7.5, 9.5 and 10.5 p.c.w. in the hindbrain, midbrain and forebrain respectively, as we show in Paper III. Recent studies have shown that increased expression of *epidermal growth factor receptor (EGFR)* coincides with the start of gliogenesis in human development during the second trimester (Huang et al. 2020; van Bruggen et al. 2022; Manno et al. 2020; Fu et al. 2021; Burrows et al. 1997; Sun, Goderie, and Temple 2005). These *EGFR*-expressing cells are thought to comprise a heterogenous group of glial progenitors in human. This was recently seen in a study enriching for EGFR-positive cells in the human cerebral cortex during gestational week 21 to 26, that found two groups of progenitors with similar gene expression patterns to astrocytes and were concentrated in the OSVZ suggesting they may derive from outer radial glia (Fu et al. 2021). Pre-astrocytic cells as well as pre-OPCs have

been observed to derive from these *EGFR*-positive cells and the co-expression of *NOTCH*-ligands suggest that they capture transient stage from neurogenesis to gliogenesis (Fu et al. 2021; Huang et al. 2020; van Bruggen et al. 2022). This is in accordance with our data where *EGFR*-expressing cells are identified as pre-OPCs, but bridge the "commited" pre-astrocytes with the OPCs (see Paper III, Fig. 5A), suggesting that this cluster of cells may still undergo different gliogenic fate-decisions.

The gliogenic potential of progenitors is also relevant for understanding how glial heterogeneity arises in astrocytes. Since the discovery of *protoplastmic* and *fibrous* astrocytes, astrocytes have been described in different regions of the CNS, such as the Müller glia in the retina and Bergmann glia in cerebellum (Westergard and Rothstein 2020; Andriezen 1893). A recent study described the spatial location of cortical layer-specific astrocytes in both mouse a human (Bayraktar et al. 2020). Further, astrocyte heterogeneity was observed as seven distinct groups in the adult mouse brain described in our transcriptomic survey of the whole brain in Paper I. At a high level, these groups were defined as telencephalic and non-telencephalic astrocytes, with a clear border observed based on their transcriptional profiles *in situ* and the single-cell data (see Paper I, Fig. 3). Interestingly, this heterogeneity was also found in glioblasts in mouse and recently in our human developmental data, suggesting that the observed adult astrocytic heterogeneity originates from a developmental program (see Paper III, Fig. 5A). Very recently, a single-cell data from our group confirmed that telencephalic and non-telencephalic astrocytes can be found in the adult human brain (Siletti et al. 2022), which we think correspond to the glioblasts heterogeneity seen during development that later will give rise to maturing astrocytes.

Another field of glial development include OPCs, the precursor of myelin-producing oligodendrocytes that insulate neuronal axons. This cell type has historically been of great interest as it is involved in myelin-degenerating disease such as Multiple sclerosis and evidence of OPC turnover in the adult brain are becoming more prominent (Siletti et al. 2022; Fernandez-Castaneda and Gaultier 2016). However, studies have shown that OPCs also are involved in maintaining homeostasis in the brain's neuronal circuitries, not least during development (Akay, Effenberger, and Tsai 2021). More recently, OPCs were shown to guide migrating cortical interneurons through contact repulsion and even have a role in the fine-tuning of synaptic circuits through phagocytic clearance in the developing pre- and postnatal brain respectively (Lepiemme et al. 2022; Buchanan et al. 2022; Auguste et al. 2022). It is known that OPCs develop from multiple anatomical regions during development and in both the murine brain and spinal cord OPCs are generated ventrally from distinct germinal zones followed by dorsal source that arises just before birth (Kessaris et al. 2006; Boda et al. 2022; Hashimoto et al. 2016; Tripathi et al. 2011; Cai et al. 2005; Bergles and Richardson 2016; Winkler et al. 2018). Oligodendrocytes are morphologically, regionally and transcriptionally heterogeneous, which in contrast has been less observed among OPCs. One study showed that OPCs become functionally heterogeneous with age, a phenotype that was intra- and interregional (Spitzer et al. 2019). Despite OPCs being produced in a region-specific manner,

they have not been found to be heterogeneous at the transcriptional level. As discussed in Paper I, we saw patterns indicating that mature oligodendrocytes or OPCs in the mouse may retain a developmental imprint through the expression of anteroposterior patterning genes despite them not diverging into distinct region-specific cell types. However, these patterns could not be fully confirmed due to potential contamination by other cells (see Paper I Fig. 3D, E). Thus, it is unknown whether the heterogeneity seen in mature oligodendrocytes is encoded in developing OPCs, or is due to regional environmental factors in the adult brain. This further leads to a long-standing question: is the heterogeneity seen in mature oligodendrocytes derived from development? This question is highlighted in Paper III, where we found that human OPCs are regionally distinct and emerge earlier than thought during the first trimester, demonstrating that OPCs are transcriptionally heterogeneous between brain regions. Unpublished data from our group shows that this finding holds true for the adult human brain as well (Siletti et al. 2022).

Altogether, the process of gliogenesis continues to spark interest, especially within research related to brain cancer. The transcriptional heterogeneity seen in glial progenitors is of particular interest for understanding different kind gliomas of glial-originated tumors. One example is a study by Filbin's group who showed that different cellular states emerge in glioblastoma that mimic the developmental program of certain glial and neural progenitors (Neftel et al. 2019). It also believed that some of these tumors arise in a region-specific manner as previously seen in histone-mutated gliomas, a reason to why glial heterogeneity may be of interest for the future studies.

3 SINGLE-CELL RNA TRANSCRIPTOMICS

With the discovery of the cell and foundation of the *cell theory* (Mazzarello 1999), we learnt that there are distinct cell types that are tissues-specific and have different features. Until the DNA was first described followed by the concept of gene expression there was a significant gap in biology in explaining the process of gene regulation (McClintock 1950; Jacob and Monod 1961). Today, modern biology has rapidly expanded through the influence of interdisciplinary areas, such an example being the use of gene regulatory network models (Elowitz et al. 2002; Kauffman 1969; Karlebach and Shamir 2008). We can partially explain cellular diversity at the molecular level knowing that gene expression is a major factor that determines cell type identity. All cells in the body theoretically share the same set of DNA, yet, a muscle cell differs in its intrinsic and extrinsic properties from a nerve cell, which is simplistically due to its transcriptional regulations of a specific subset of genes whose expression levels are typically maintained within the cell in order to keep its identity. Although, a clear consensus is still lacking, a cell type is usually defined by the expression of these specific set of genes that are related to its function (Zeng 2022). Other definitions of what a cell type is, raise the question of whether cell type ontology is a consequence of an evolutionary process that is conserved between species (Arendt et al. 2016). It is reasonable to believe that both gene expression-specificity and evolutionary aspects can be used to classify cell types, as many

genes and cellular functions have evolved into having generic functions among species. scRNA-seq has proven to be a powerful method that outputs gene expression profiles of single cells within different tissues and enables experimental scenarios where we aim to understand molecular heterogeneity in targeted and global settings in both healthy as well as pathological conditions. Furthermore, scRNA-seq has not only contributed to an increased knowledge of the biological heterogeneity between cells, but also to the dynamics within them. No two single cells are transcriptionally identical, and – due to intrinsic transcriptional noise – this would be true even if measurement noise could be completely avoided (A. Raj and van Oudenaarden 2008).

3.1 EARLY ATTEMPTS AND THE BEGINNING OF SINGLE-CELL BIOLOGY

3.1.1 Historical introduction

The desire to measure gene expression patterns in single cells has its roots in neuroscience. Often, neuronal cell types overlap in their physiological and morphological properties, although they may differ in their molecular composition. With that reasoning, experimentalists identified the need to further elucidate the molecular basis of cellular heterogeneity in the brain beyond the standard classification schemes, such as probing cellular morphology and neurotransmitter/receptor specificity (coupled to their electrophysiological characteristics). In the 1990s a few groups published protocols on how to extract, preserve and amplify the intracellular contents of cells using PCR and *in vitro transcription* (Eberwine et al. 1992; Van Gelder et al. 1990; Brady, Barbara, and Iscove 1990). James Eberwine's group in particular developed a method used on dissociated rat hippocampal neurons to inject reaction components for cDNA synthesis directly into the cell through intracellular dialysis following patch-clamp recordings (Eberwine et al. 1992), which describes the first attempts to analyse transcriptional species in single cells. However, the readout using this approach was limited and detected by using Southern blot. Since then, other labs have been inspired by these protocols and further development have lead to the performance of RT-PCR and sequencing of single neurons (Matsunami and Buck 1997; Malnic et al. 1999), which is nowadays dominated by the systemized version of all these techniques, namely the Patch-sequencing (Patch-seq) method. Patch-seq allows simultaneous measurements of morphology, gene expression (from total RNA) and electrophysiological features in single neurons which has been applied on human brain as well (Cadwell et al. 2017; Fuzik et al. 2016; Lipovsek et al. 2021).

3.1.2 The technical evolution of single-cell biology

These above-mentioned techniques introduced a ground-breaking concept of detecting individual transcripts in single cells and ever since, the field of single-cell biology has continued to develop at a breakneck speed. Since the first scRNA-seq study using a modern protocol on individual mouse blastomeres (Tang et al. 2009; Guo et al. 2010), the field has moved from manually collecting dozens of individual single cells with micropipettes (or by

mouth pipetting!), to hundreds of thousands of cells in an automated manner. scRNA-seq had its real breakthrough when the multiplexing strategy was developed in our group (Islam et al. 2011), allowing for the quantification of absolute numbers of mRNA molecules, first by including known amounts of reference RNA in the experimental setup and later by applying a barcoding logic to trace back individual transcripts to their cell of origin (Islam et al. 2014; 2011).

In theory, scRNA-seq can be applied to any cell of interest, but the tissue dissociation procedure differs a lot between tissues and is a crucial step in order to obtain a clean and viable single-cell suspension. In addition, some cells are simply more fragile than others which can lead to an unavoidable loss of certain cell types. Therefore, knowing the tissue composition is important. As mRNA is very unstable and only exists at the scale of picograms in a cell, all single-cell protocols require the conversion of mRNA to cDNA. A few methods use the direct detection of mRNA, but it has not caught on with the large-scale interest in the field (Depledge et al. 2019; Ozsolak et al. 2009). The key steps in a scRNA-seq protocol involve the capture of mRNA molecules in single cells through oligo(dT) priming that anneal to the poly(A)-tails of mRNA and via reverse transcription produce cDNA that is amplified. Next, the cDNA is further amplified and processed in several steps into a complete library with correct fragment length ready for sequencing. Even if this main step is similar among protocols, they vary in which parts of the transcripts are targeted. For example, some protocols enrich for either the 5'or 3' end of transcripts during cDNA library preparation, resulting in biased sequencing read coverage. This may influence the downstream analysis and enable detections of upstream elements such as transcription start sites (TSS) when using a 5'-based protocol (Islam et al. 2011; Kouno et al. 2019). Other methods, such as Smart-seq, which capture the full length of transcripts have a more even distribution of the 5' to 3' read coverage (Hagemann-Jensen et al. 2020; Picelli et al. 2014). Many technical strategies in capturing single cells have evolved, ranging from plate-based techniques – that sorts single cells into individual wells – to microfluidic devices using droplet-encapsulation of single cells (Svensson, Vento-Tormo, and Teichmann 2018a; Mereu et al. 2020). Some of these platforms were developed in-house, but are also commercially available today (Svensson, Vento-Tormo, and Teichmann 2018a; Ziegenhain et al. 2017). Largely, most scRNA-seq protocols have greatly advanced to facilitate the capture of an increased number of cells more efficiently, leading to the generation of larger datasets and thus improving the statistical power. However, there is still a compromise between sensitivity and throughput as is seen in protocols like Smart-seq, with high sensitivity and low throughput compared to high throughput platforms with lower sensitivity as in Chromium by 10X Genomics.

3.1.3 The Chromium platform

The leading single-cell technology on the market today is the Chromium platform offered by 10X Genomics, which is similar to the in-house version inDrop (Klein et al. 2015) and enables large-scale experiments with high throughput. A third method, Drop-seq, works with the similar principle as inDrop, but differs slightly in the mRNA capturing and processing steps

(Macosko et al. 2015). Briefly, this technology is based on the generation of oil-in-water emulsions in a capillary microfluidic device that tightly regulates the flow of water and oil, forming jet streams. Firstly, two water-based streams are merged into one channel: I) one that carries the cells and reaction components (master mix for reverse transcription) necessary for cDNA synthesis, II) the other carrying hydrogel beads with attached oligonucleotide (oligo) sequences. At third stream with an oil-based solution intersects with the stream of the cells and beads resulting in the formation of oil-droplets. In a successful capture, single cells and hydrogel beads are partitioned and concurrently encapsulated into these oil-droplets. The oligo sequences on the hydrogel beads contain poly-dT primers with attached cell- and transcript-specific sequences referred to as the 'cell barcode' and the unique molecular identifier (UMI) and is key to the multiplexing of unbiased cell capture of a tissue. These sequences allow for the identification of cell-specific molecules and the quantification of the number of molecules per cell. This thesis is focusing on the commercial Chromium Single Cell 3' solution that were used for the major part of the projects, with the exemption for Paper II, where already existing data was reanalysed from inDrop and Smart-seq2 experiments.

3.1.4 How many cells should we sequence?

With the commercialisation of scRNA-seq platforms and facilities, any lab in principle, is now able to generate this kind of data. Although the price per cell has reduced drastically over the years, being around $2 per cell in the beginning (excluding sequencing; Islam et al. 2012) and today about $0.5 per cells, the limiting factor is still the high cost of the reagent kits including the expensive sequencing costs that follows. Another obstacle is the number of cells needed to answer a particular question. The single-cell community has always strived to collecting more cells in the most efficient way as possible, as this increases the chance of detecting 'all' cell types globally represented in a tissue, but it also entails a better distinction of cell types and the transcriptional phenomena they are involved in (although the sequencing depth and method used also has an impact). For example: sampling more cells from a specific lineage that are still undergoing cell division, will probably result in a better capture and visual representation of the different cell cycle phases, as seen with the increased sensitivity of current methods (Riba et al. 2022) and shown in the cycling cells of the forebrain in Paper III (Fig. 3D). However, the bottleneck of maximizing the number of cells in an experiment is the increased sequencing cost and compromised read depth. A large number of cells is also computationally costly on the analysis side. Importantly, there is a trade-off between the chosen number of cells sequenced and which read depth to consider, depending on the biological question asked and the budget at disposal (Svensson, da Veiga Beltrame, and Pachter 2019; Haque et al. 2017; M. J. Zhang, Ntranos, and Tse 2020). In our experience, increasing the number of cells may also increase the biological resolution of important trajectories and lineages when studying developmental tissues. However, we also do not always fully understand whether certain cell types are specific targets for under-sampling.

3.2 AN OVERVIEW OF SINGLE-CELL RNA-SEQUENCING ANALYSIS

The output of a scRNA-seq experiment after sequencing which is the input for analysis, is a *M* by *N* sparse count matrix of RNA molecules UMIs, corresponding to *M* genes and *N* cells. Single-cell data is therefore viewed as "high-dimensional" because of the high number of variables or features (genes) measured for each observation (cell), typically thousands of genes per cell. Once obtaining the scRNA-seq data, the real difficulty lies in the data analysis. In parallel with the technical evolution of the single-cell field, a number of analysis pipelines have been developed along the way (Luecken and Theis 2019). Many of them are based on a common framework: read-alignment to reference genome, quality control, feature selection (selection of informative genes), normalization of count data, dimensionality reduction, clustering algorithms and graphical visualization of the high-dimensional data (Luecken and Theis 2019). Despite building on a similar concept of analysis, they may differ a bit in each step depending on the choice of program, however the steps that have most effect on the downstream analysis is generally the strategy for read-alignment, feature selection, normalization and the clustering parameters including the visualization tools used. In this thesis, I will refer to our in-house pipeline developed in our group, *Cytograph*, unless otherwise stated.

3.2.1 Quality control

3.2.1.1 General quality metrics

The very first step is deciding on a relevant read-alignment approach (i.e. mapping transcript reads to matching genomic loci) which are done on the FASTQ files containing the raw sequencing data that also undergoes quality assessment. Different methods such as pseudoalignment of reads have developed to speed up this process as this step is also computationally exhausting, that skips the alignment of sequencing reads on the nucleotide-level completely and instead uses a set of predefined rules to map the query to a matching transcript reference based on *k*-mer algorithms (Bray et al. 2016; Brüning et al. 2022). Nowadays, mapping of intronic reads are commonly included during read-alignment to distinguish unspliced and spliced RNA molecules as *RNA velocity*-based programs have increased in popularity. This step is also included as an optional part of the available STAR/STARsolo alignment pipeline (La Manno et al. 2018; Kaminow, Yunusov, and Dobin 2021; Dobin et al. 2013).

The viability of cells is assessed early on during tissue dissociation procedures, usually by the application of dyes like Trypan blue, that permeates the cell membrane of dead cells. However, there are always sensitive cells that react to stress or start undergoing apoptosis during sample preparation, and will be included until the final step in the experimental procedure. Therefore, the first steps once the count matrix is generated, is to filter out bad quality cells. Typical metrics include filtering out cells with high expression of mitochondrial or immediate early genes, indicating a stress response in cells. It is also conventional to remove cells that have very

low detection of gene and molecule counts, which is commonly assessed by having a fixed upper or lower threshold of the number of genes and molecules per cell. This should of course be carefully designed to avoid filtering out specific cell types that on average are naturally richer or poorer in the number of genes or molecules, a feature often linked to the cell size.

3.2.1.2 Unspliced UMIs as an indicative measure of cell quality

The approach to filter out bad-quality cells in *Cytograph* has varied from time to time depending on the cell populations analysed and so forth. Typically, cells that have less than 1000 total UMIs and above 1 % mitochondrial UMIs would be discarded. As we obtain unspliced UMIs for each cell – which make up part of the total molecules in the count matrix – we started monitoring the fraction of unspliced UMIs of the total UMIs. The idea is that a high unspliced fraction reflects the viability and functionality of a good-quality cell, suggesting that the cell is still metabolically healthy and produces a lot of new RNA. When examining the distribution of the unspliced UMIs ratio with data generated by Chromium, a minor "peak" is frequently observed in the beginning of the histogram (Figure 4A). We later identified this group of cells as being of bad quality because they have a low unspliced UMI ratio (normally falling below 10 %) and thus also a decreased total number of genes and molecule counts (Figure 4B).

3.2.1.3 Doublet detection

A phenomenon occurring in droplet-based scRNA-seq platforms is the occurrence of 'doublets' in the data, i.e. droplets where two cells were captured and thus ended up with the *same* 'cell barbode'. There are also scenarios leading to the retrieval of empty droplets, or droplets containing ambient RNA that is derived from leaky or lysed nearby cells in the suspension. A cell may be captured together with 'free-floating' ambient RNA, thereby causing contamination by other cells or surrounding RNA. Since they share the same barcode, the cell origin of the transcripts of the two cells cannot be distinguished. Nevertheless, there are some features that characterize doublets which makes it possible to computationally describe and remove them before the downstream analysis. A characteristic of doublets is their resemblance of large cells as they contain twice the cellular material. However, unfortunately the variation in RNA content among normal cells is too large to use the double RNA content alone to identify doublets. Most doublet-detection software identifies doublets based on their high molecule number, gene co-expression and other features or by the generation of artificial doublets that are compared to the cells in the dataset (Luecken and Theis 2019). In a study comparing doublet-detection software, *DoubletFinder* outperformed other existing methods in terms of accuracy in detecting doublets, but was not among the most computationally efficient ones (Xi and Li 2021). DoubletFinder is also the method that was used in our latest human developmental study, as described in Paper III. Note that this approach aims to the detect *heterotypic* doublets, that remain from two transcriptionally different cells, in contrast to *homotypic* doublets that are formed by similar cell types and may be discovered via SNP

analysis if they derive from different individuals (Kang et al. 2018; Xi and Li 2021). DoubletFinder relies on the *k*-nearest neighbour (KNN) algorithm and works by randomly sampling pairs of cells from the single-cell data by averaging their gene expression profiles to generate artificial doublets, as to best having them represent the underlying data structure (McGinnis, Murrow, and Gartner 2019). The generated doublets (of a decided amount) are included in the analysis together with the real cells and standard parameters for dimensionality reduction is performed using principal component analysis (PCA). Next, a KNN graph of all cells is constructed in principal component space and the proportion of artificial nearest neighbours can be estimated for every real cell. Assuming that real and artificial cells co-localize in the KNN graph, all cells get a score depending on the proportions of artificial nearest neighbours and can be flagged for removal if they exceed a scoring-threshold. Cells below this threshold will be kept for the downstream analysis. The doublet estimations can also be compared with the theoretical expected number of doublets present in a sequenced sample, which are based on the loading densities of cells on the Chromium device (see 10X Genomics Chromium Single Cell 3'-based solution). These probabilities can be estimated as droplet-encapsulation follows a Poisson distribution. However, it should be noted that the true number of doublets may exceed the Poisson expectation, for example if cells are physically sticky or difficult to dissociate. Removing doublets improves cell type discovery as they are confounding factors in the analysis. Note, that even if major groups of doublets are removed using these software packages, it is common that some doublets remain. These can sometimes be identified by inconsistent gene markers, or may be visible by eye on a t-distributed stochastic neighbour embedding (t-SNE) embedding and can be manually removed from the dataset.

3.2.1.4 Classification of 'droplet classes'

As previously mentioned, the unspliced UMI ratio may indicate if a cell is of good quality. In the beginning of our pipeline, each sample (i.e. sequencing library) undergoes a quality control to assess the above-mentioned metrics. If the sample does not pass our requirements, it is excluded from any analysis later on. In this step, we noticed that the cells that have very low total UMIs and gene counts, also have a very high unspliced UMI ratio (Figure 4C). This violates our usual assumptions: that high unspliced UMI ratio indicates good quality and low total UMIs indicate worse quality. We hypothesized that these may be nuclei from single cells, that should be rich in unprocessed, unspliced mRNA. Therefore, the unspliced UMI ratio either entails cells of lower quality or that they are bare nuclei. This group of bare nuclei is unnoticed when observing the regular distribution of the unspliced UMI ratio. When visualizing the total UMIs as a function of the unspliced UMI ratio of the cells in a sample, the group of nuclei can be distinguished. In fact, this graph reveals three distinct cell populations or "clouds" (Figure 5A, left). The first group is the small peak of 'bad cells' observed in the histogram, with low total UMIs and low unspliced UMI ratio, denoted as cytoplasmic debris. The second group contains the 'whole cells', that are higher in these UMI numbers. The potential nuclei are in the final, third group with generally lower molecule counts than the whole cells and a much higher unspliced UMI ratio on average. According to our own observations, this pattern seems

to be conserved for all Chromium-based embryonic brain tissue used for our study, but we did observe a similar pattern in other datasets of mouse and human. In contrast, in a dataset of adult human brain where nuclei where exclusively used, unspliced ratios were generally higher and putatively whole cells were not observed (Figure 5A, right). These observations strengthen our explanation that a substantial part of the cell barcodes captured in droplets are indeed bare nuclei.

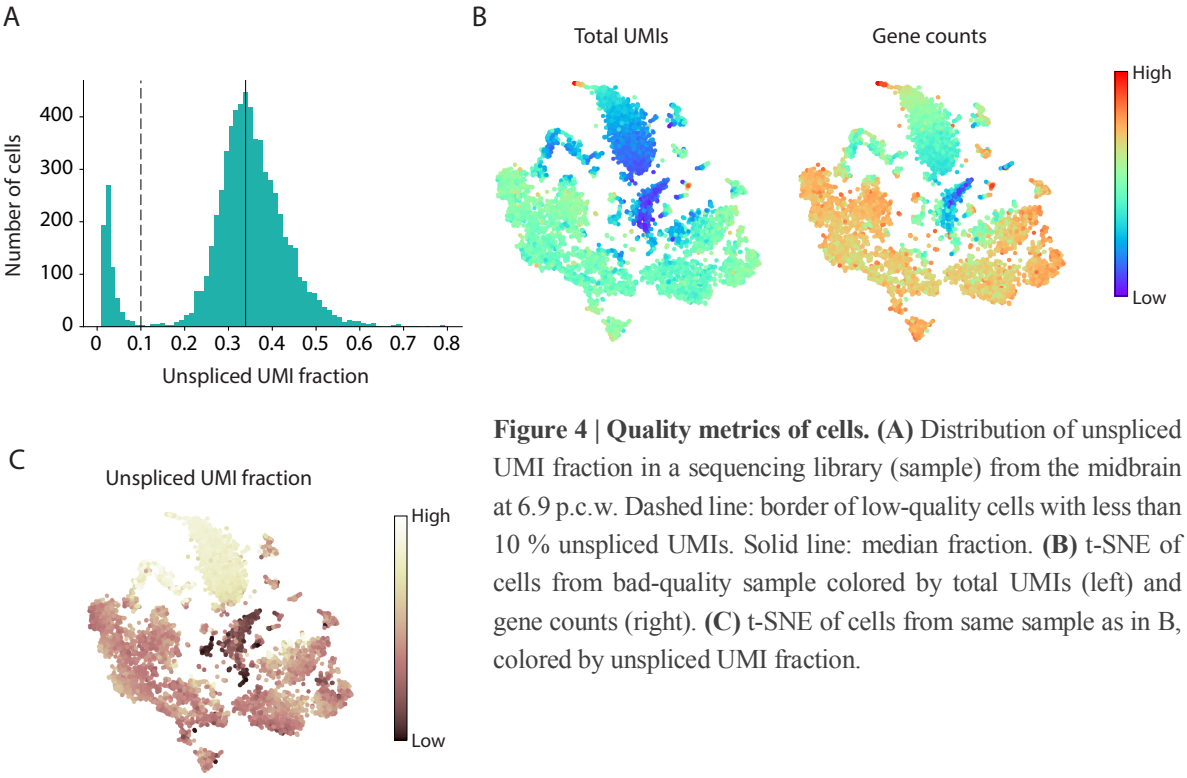

Figure 4 | Quality metrics of cells. (A) Distribution of unspliced UMI fraction in a sequencing library (sample) from the midbrain at 6.9 p.c.w. Dashed line: border of low-quality cells with less than 10 % unspliced UMIs. Solid line: median fraction. **(B)** t-SNE of cells from bad-quality sample colored by total UMIs (left) and gene counts (right). **(C)** t-SNE of cells from same sample as in B, colored by unspliced UMI fraction.

Based on these observations of droplet groups, it is possible to estimate *droplet classes* based on the total UMI counts and the unspliced UMI fraction of a sample (Figure 5B). The whole approach is based on modelling the parameters that best fit the underlying multivariate Gaussian distributions (here defined by the total UMI counts and unspliced UMI fraction). First, all droplets that are likely single cells, are selected if they meet fixed thresholding values for the unspliced UMI fraction, total UMIs and the logarithm of total UMIs (for details, see methods in Paper III). Using these thresholding criteria of the unspliced UMI fraction and the logarithm of total UMIs, a two-dimensional Gaussian maximum likelihood estimate (Anderson and Olkin 1985) is fit to find the μ (mean) and σ^2 (variance) that best describes this droplet-data and is later used to calculate the probability density function of the droplets. All droplets with a probability greater than 0.1 are retained, except for the doublets that are still flagged by DoubletFinder. Finally, this results in the classification of the following seven droplet classes: 1. Cells (kept for analysis) 2. Doublets (by DoubletFinder) 3. Large cells (might be multiplets) 4. Cytoplasmic debris (low total UMIs and unspliced UMI fraction) 5. Cellular debris (low total UMI, normal unspliced UMI fraction) 6. Nuclear debris (bare nuclei; low total UMIs, high

unspliced UMI fraction) and 7. Mitochondrial debris (high fraction of mitochondrial UMIs) (Figure 5B). One study designed a model to detect empty droplets including their ambient RNA in single-cell datasets (Lun et al. 2019), but only one method: *DropletQC*, uses a similar approach to ours in order to detect droplets based on nuclear fraction (Muskovic and Powell 2021). This method was developed independently and although it uses the same principles, it is based on a kernel density estimate. Overall, this droplet classification based on unspliced fractions results in a more stringent filtering of cells than our earlier strategies, which impacts the later analysis steps.

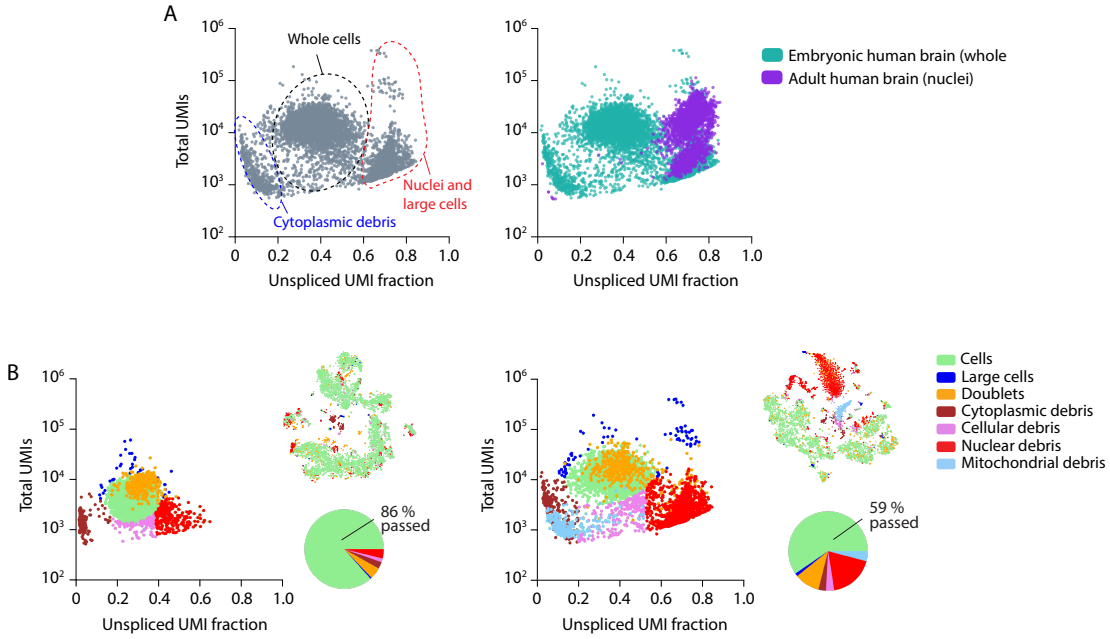

Figure 5 | Droplet classes. (A) Left: scatterplot of a bad sample of human embryonic brain, showing three groups, "clouds", of cells: cytoplasmic debris, whole cells and potential nuclei. Right: scatterplot displaying cells of two samples from embryonic and adult human brain. **(B)** Droplet classes for a good sample (left) and a bad sample (right), showing total UMIs as function of unspliced UMI fraction. Fraction of good-quality cells that passed are indicated in green in the corresponding t-SNE plots and pie charts.

3.2.2 Normalization, feature selection and dimensionality reduction

The subsequent step after retrieving all the good-quality cells, is normalization of the data. This is done to account for the sequencing depth of all the cells, to avoid the influence of variability seen between identical populations because of sampling effects. A common approach is to adjust all cells to the same count depth which can be achieved by scaling each cell to the median total UMIs, followed by a $\log(x + 1)$-transformation to stabilize the variance, that is required for the downstream PCA analysis. Another commonly used strategy today is using Pearson residuals (Hafemeister and Satija 2019; Lause, Berens, and Kobak 2021) that models the technical noise in the data using a regularized negative binomial regression and accounts for overfitting scenarios. This considers individual gene-specific differences and cell-to-cell variation arising from technical effects that might confound the biological heterogeneity

observed. This method does not require prior normalization as mention above for computing the PCA, as it models the "true" nature of single-cell data considering that gene counts are assumed to follow a negative binomial distribution (Hafemeister and Satija 2019; Kharchenko, Silberstein, and Scadden 2014).

After reducing the technical effects on gene expression by a suitable normalization method, feature selection is performed. This is one of the most crucial steps in scRNA-seq analysis. Despite measuring thousands of genes per cell, not all of these genes are informative in the downstream analysis. Many genes are expressed either at very low level or not at all, resulting in very sparse gene count matrices, mostly containing zeros due to gene dropouts (zero-inflated). Whether these gene dropouts are true biological signals or a technical artefacts is still an ongoing discussion (Svensson 2020; Choi et al. 2020; Jiang et al. 2022). Regardless, genes that are very low or not expressed at all are excluded in the following steps. The goal is to find the genes that are the biological meaningful, in other words the most variable genes reflecting the cell heterogeneities. The selection of genes is traditionally performed by fitting a noise-model using the coefficient-of-variation (CV) versus the mean expression. Genes that disperse the most around the mean are the most variable ones and represents the biological heterogeneity seen between cells. Usually the top 1000 to 2000 most highly variable genes are selected for downstream analysis, but it depends on the study design as well. Alternatively, feature selection by Pearson residuals can be used to choose meaningful genes (Hafemeister and Satija 2019). By modelling the observed versus expected expression of genes, the most deviant residuals, i.e. genes are chosen. Another example that has been implemented in our pipeline, is the approach of selecting features by performing a preliminary clustering of the data followed by a gene enrichment analysis of those clusters, as an alternative way of finding genes informative for cell type-specificity. Note, however that this occurs in an iterative fashion, as it relies on the initial feature selection of genes for the preliminary clustering to find genes which in turn can be used for downstream analysis.

Finally, dimensionality reduction is performed on the selected genes. As noted, not all genes are informative in the analysis for cell type discovery. At the same time, the number of dimensions is reduced which diminishes the complexity of the data and partially circumvents the problem of having a high number of dimensions ("curse of dimensionality"), that otherwise complicates the analysis and interpretation of the data and its visualization. PCA, a matrix factorization method, is commonly used to reduce the number of dimensions. Simplified, PCA orthogonally transforms, using the selected variable genes, the count matrix into a smaller set of n principal component (PC)s. These PCs condenses the gene information that best explains the variability between the cells in the data. The PCs are ranked from the highest to lowest variability they explain, where typically the top 50 components are kept. A minimum threshold can be set to filter out non-informative PCs, for example by excluding PCs that explain less than 30 % of the variance in the data. Out of the 50, the top number of components explain the most variability in the data. Thus, we have reduced the number of dimensions by keeping the most biologically informative variables in the analysis.

3.2.3 Batch correction, clustering and visualization

There are different data integration tools available today. These are often used to either correct for batch effects present in the dataset or to integrate multiple datasets from different origins. Different approaches exist for integration of single-cell data and should be assessed individually and decided upon depending on the method that best suits the relevant data. A typical example of batch effects is the one observed in our human developmental dataset in Paper III. We sampled cells using two different chemistry versions of the Chromium protocol (v2 and v3), because 10X Genomics stopped producing the older one as we were in the middle of data collection. Hence, we see batch effects between the two chemistries, simply because the recent chemistry resulted in more efficient cell lysis and overall higher sensitivity. In theory, these effects should not be there, as these are the result of technical effects and do not reflect true biological differences. This means that we observe higher gene and molecule counts per cell with the newer version. The differences between the chemistries do not scale linearly for the majority of genes and standard normalization applied early in the analysis do not remove this effect. In these cases, it can be useful to perform batch-aware feature selection in the earlier steps. Furthermore, one strategy is to integrate cells of two chemistry versions in order to have a complete and integrated dataset for the next analysis steps. Our strategy included using an iterative batch-correction algorithm: Harmony, which is a linear embedding model and uses the previously calculated PCs (Korsunsky et al. 2019). A soft clustering is performed in this reduced PC space to assign cells to multiple clusters using a modified *k*-means clustering that maximises the diversity of cells from different datasets in each cluster. For each cluster, Harmony calculates a global centroid and dataset-specific centroid which are used to make a linear correction factor that is used for the integration. Similar methods have been developed that identifies and integrates matching mutual nearest neighbours (Tran et al. 2020; Luecken et al. 2022; Haghverdi et al. 2018). Other batch-removal methods include graph-based- and deep learning methods (Luecken and Theis 2019). In Paper III, we show that different integration methods can be used for the integration of multiple timepoints during development (Lopez et al. 2018; Hie, Bryson, and Berger 2019). Currently, there is no standardized approach of integrating single-cell data and each dataset requires individual assessments of the chosen model. Preferably, the integration should be done in the most suitable way that removes batch effects, but preserves the biological information. It is disputed how much biological information is actually lost during batch integrations, especially for developmental datasets where essential, yet very subtle dynamical changes may be sensitive to the model used. Many of the current integration tools have been evaluated by Luecken et al. 2022 and Tran et al. 2022.

In one of the final steps, clustering is applied on the integrated data. Through clustering algorithms, it is possible to, in an unsupervised fashion, organize groups of cells by similarity in terms of gene expression in the reduced gene expression space. Among the different approaches, one is to perform clustering on a constructed KNN graph. KNN is a manifold learning algorithm that embeds cells using the reduced PC space into a graphical structure, then based on an Euclidean distance measure it connects nodes (cells) to its *k* nearest neighbours, where *k* is defined by the user depending on the size of the dataset. This yields a graph where

the nodes that are densely connected are cells that share transcriptional similarity and the edges indicate the distances between them. Different variants of this algorithm exist, such as the radius-nearest neighbours (RNN) that find neighbours within a fixed radius or weight-adjusted approach, balanced KNN that balances the distribution of the neighbourhoods. A subsequent clustering step is performed, like the Leiden clustering algorithm that uses the KNN graph to refine the structure by locally moving and partitioning the nodes into new communities (Traag, Waltman, and van Eck 2019). Leiden clustering has been shown to outperform the Louvain algorithm in terms of speed and community detection (Traag, Waltman, and van Eck 2019). Finally, t-SNE or its equivalent: uniform approximation and projection (UMAP) method is often used to visualize this high-dimensional single-cell data

into a 2- or 3-dimensional graphical representation and are constructed on the chosen PCs (Becht et al. 2019; Van Der Maaten and Hinton 2008).

Feature selection on a large and diverse dataset like the human brain will necessarily emphasize common cell types, and may miss genes informative for subtle cellular subtypes in specific brain regions. To refine the clustering, we have implemented a semi-automatic iterative splitting of clusters into sub-clusters. In other words, we first obtain a number of clusters from a given dataset, construct a dendrogram of all the clusters, split the set of clusters based on the dendrogram, and then apply the same algorithm to each cluster subset. Then, based on a certain criterion, a number of clusters is split into another subgroup. Meaning that in each iteration our dataset will undergo a new feature selection, thus revealing a new set of informative genes that – together with other clustering parameters – refines the resolution and identification of clusters and thereby cell type detection.

To conclude this chapter, it should be highlighted that even though most single-cell pipelines share a similar computational backbone, they differ in terms of the input parameters used and the individual algorithms used for the cell type discovery, for example the clustering steps.

3.2.4 Lineage tracing and trajectory inference

As scRNA-seq methods have evolved over the past years, focus has shifted from improving the efficiency of capturing single cells to developing custom-made pipelines for data analysis and maximizing the computational efficiency required for these analyses. Because scRNA-seq is capturing a snapshot of a cell's gene expression profile at the time of the experiment, you mainly retrieve static measurements of the biological information from tissues. Lineage tracings of a cell population and its progeny has usually been done partially through live imaging using reporters (Kretzschmar and Watt 2012). However, there are limitations to the scalability of these experiments. Recently, techniques and algorithms have been developed in order to measure the fate of a cell and further reconstruct its developmental lineage at the larger scale. One approach is to introduce actual barcodes or 'scars' into the genome, for example by using CRISPR-Cas9 to make small insertions or deletions that can be traced and identified in the single-cell data later on to assess the clonality of cell types (Alemany et al. 2018; B. Raj et

al. 2018; Spanjaard et al. 2018). This technology allows the reconstruction of the shared the clonal ancestry of cells that cannot be studied through conventional scRNA-seq. Endogenous mutations can also be used; for example it has been shown that somatic mutations in mitochondrial DNA can be detected in scRNA-seq or single-cell assay for transposase accessible chromatin sequencing (scATAC-seq) and used to infer clonal relationships of cells (Ludwig et al. 2019).

Another approach has been to infer the dynamics of gene expressing data using regular scRNA-seq and computationally reconstruct a trajectory of cell populations by temporally ordering them based on similarities in their gene expression profile (Lederer and La Manno 2020). By using trajectory inference methods every cell can be assigned a *pseudotime*, an arbitrary measure of how far a cell is within a trajectory in a dynamic process (Saelens et al. 2019). This can be done on datasets containing single cells from a mixed population or by systematically collecting cells from different time points (e.g. during a differentiation protocol). A number of different trajectory inference methods has been systematically compared and some of the differences lies in the performance of dimensionality reduction of the gene expression matrix (Saelens et al. 2019; Lederer and La Manno 2020). It is important to know that pseudotime estimation is a statistical measure and does not reveal the true path of an individual cell's trajectory. Other models like RNA velocity, estimates the abundance of unspliced and spliced RNA molecules from scRNA-seq data and model the time derivative of the gene expression state to predict a future state of cells in a developmental phenomenon like embryogenesis (La Manno et al. 2018). RNA velocity was recently expanded by another group by refining assumptions about transcriptional dynamics and taking into account that not all genes share a common splicing rate on a cellular level (Bergen et al. 2020). Other techniques allows the discrimination of newly synthesized and pre-existing RNA transcripts in single cells, which can serve as a good estimation of RNA turnover to further improve existing dynamic models (Hendriks et al. 2019; Erhard et al. 2019). Models that utilize the unspliced and spliced mRNA counts are constantly developing. In Paper III, we explored DeepCycle, a deep-learning approach using a variational autoencoder, that based on the unspliced and spliced mRNA models the cell cycle dynamics by targeting the cell cycle genes (Riba et al. 2022). Using this model, we could dissect the cell cycle trajectories of radial glia and nIPCs in the pallium (see Paper III, Fig 3).

3.3 HUMAN BRAIN DEVELOPMENT AT THE SINGLE-CELL LEVEL

3.3.1 Creating atlases of brain development

As touched upon earlier, the increasing use and development of scRNA-seq technologies has enabled large-scale studies of millions of cells representative of a tissue in an unbiased approach (Tang et al. 2009; Svensson, Vento-Tormo, and Teichmann 2018b; Macosko et al. 2015; Mereu et al. 2020; Klein et al. 2015; Islam et al. 2014; Kivioja et al. 2012). A number of studies have been conducted on the adult mouse brain, where the heterogeneity of brain regions and cells have been profiled transcriptionally, morphologically and also with respect to their physiological function (Zeisel et al. 2018; Fuzik et al. 2016; Cadwell et al. 2017; Tasic et al. 2018). Although focusing on the adult mouse brain, our study in Paper I describes the conservation of development-related genes, particularly in neurons, astrocytes and ependymal cells that were found to be heterogeneous and have their own regional signature (Zeisel et al. 2018). This suggests that there is a developmental structure reflecting the molecular architecture of the adult brain.

For a long time, tissue resources have from human and even non-human primates have been limited. With the increased use of tissue banks and ethical regulations today, both human adult and prenatal brain is becoming more available. High-throughput technologies have also made it possible to extract as much biological information as possible from these rare tissues. Applying scRNA-seq on human embryonic brain tissue allows us to sample a representation of cells from the entire brain in an unbiased manner and profile the whole transcriptome of these cell populations. By combining this information with the previous knowledge on important genes, TFs and morphogens in brain development, the effect and response of these factors can be measured on a larger scale. To give an example, scRNA-seq makes it possible to study the ventralizing effect of SHH on surrounding cells on a global transcriptional level, as opposed to identifying cell fates based on single gene markers (Manno et al. 2020). Note however, that that this kind of gene expression data is not measuring the actual morphogen concentration or ligand-receptor interactions, but rather the expression level of these genes and genes encoding their respective receptors. However, new single-cell technologies have emerged that detect physically interacting cells like PIC-seq, even though it has only been restricted to the use of immune cells so far (Giladi et al. 2020). Computational methods have also been developed that infer the cell-cell interactions on already existing single-cell data (Efremova et al. 2020; Z. Liu, Sun, and Wang 2022; Ghaddar and De 2022).

Of the few scRNA-seq datasets that exist of the human embryonic brain, the majority of them has focused on human cortical development. These studies have examined timepoints ranging from the first to third trimester (GW8 to GW26) and described major cell populations arising during neurogenesis in the developing cortex (Fan et al., 2018; Nowakowski et al., 2017; Polioudakis et al., 2019; Pollen et al., 2015; Zhong et al., 2018). All studies have provided early insights into cortical development on the single-cell level, where cortical layers could be identified by genes specific to radial glia, nIPCs and maturing post-mitotic neurons originating

from the VZ and SVZ. As described earlier, one study found two kinds of radial glia: ventricular radial glia and outer radial glia arising from the VZ and OSVZ in the cortex which are transcriptionally distinct (Pollen et al. 2015). Here, the authors suggest that outer radial glia are more involved in maintaining the stemness according to their gene expression profile, in contrast to ventricular radial glia. There have also been efforts to measure the physiological response of human embryonic cortical neurons through patch-clamp recordings or by performing calcium imaging of the same set of cells subjected to RNA sequencing, showing that there is neuronal activity during mid-gestation (Mayer et al., 2019; Zhong et al., 2018). Although, these studies have reported evidence about transcriptional heterogeneity of the human developing cortex, there is a considerable variability in the methodology that were used, which is partially explained by the rapid development of the different platforms within the scRNA-seq field. This fact is highlighted by a single-cell study where a dataset of 40 000 cells – an order of magnitude more cells compared to previous datasets – was generated of the human neocortical development. This study shows that the depth and number of cells sequenced is profoundly increasing the resolution of the cell populations and their subgroups that emerge in the cortical layers during development (Polioudakis et al. 2019). Here, the authors also compared bulk tissue with single-cell data and performed gene regulatory network analysis of TFs essential for specific cell types, which can serve useful for the understanding of neurological disorders arising during development. This study also showed that cell fate decisions in cortical radial glia and nIPCs seem occur before the S-phase, an observation that is also confirmed by us in pallial nIPCs in Paper III.

Currently, not many human studies exist that focus on earlier timepoints, due to the difficulty in receiving this type of tissue and varying ethical regulations in different countries. As mentioned, cortical tissue has been of great interest for the scientific community, and in some countries, later timepoints are more frequently acquired in clinical settings. In total, about 30 single-cell studies have been published (including preprints) on human CNS development with one of them including a few cells from the PNS (Vinsland and Linnarsson 2022; Bocchi et al. 2021; Herb and Glover 2022; Fan et al. 2018b; Rayon et al. 2021; Wälchli et al. 2021). The majority of these studies involves cortical development or telencephalon-derived tissues (i.e. pallium, subpallium/ganglionic eminences, hippocampus), while the remaining, nearly dozen studies, focus on tissues like spinal cord (Rayon et al. 2021), ventral midbrain (La Manno et al. 2016), hypothalamus (Herb and Glover 2022), cerebellum (Aldinger et al. 2021; Cao et al. 2020; Eze et al. 2021), pituitary gland (S. Zhang et al. 2020), inner ear (Yu et al. 2019) and the retina (Lu et al. 2020; Sridhar et al. 2020; Y. Hu et al. 2019; Eze et al. 2021). Others have targeted cell types like microglia or vascular cells through FACS (Kracht et al. 2020; Wälchli et al. 2021). A few additional studies have been conducted on very early timepoints from gastrulation up to 4 p.c.w., but these datasets are either too early to observe the first signs of brain-tissue specification or focus on other questions like whole embryogenesis or malignancies, yet they are extremely insightful for the earliest stages of development (Dong et al. 2020; Tyser et al. 2021; Xu et al. 2021).

To summarize some of the already mentioned studies, the very first single-cell datasets published on neocortical development included mainly a couple hundreds of cells that were captured with "full-length" mRNA protocols and most of them using the plate-based single-cell platforms that were popular at the time (Fluidigm C1 Single-Cell AutoPrep System). In some cases, the specific cortical germinal zones were dissected out for the single-cell experiments and revealed early insights into radial glia and neuronal maturation in the cortical layers including radial glia-enriched non-long cording RNAs (Pollen et al. 2015; S. J. Liu et al. 2016; Pollen et al. 2014). Studies from a few years later, show an increased resolution in the datasets capturing most cell types in the cortical tissues: progenitors, maturing neurons, glial and vascular cells. Overall, the current single-cell protocols in combination with the computational analyses, have become better in detecting transitioning cell states, such as nIPCs and phases of the cell cycle, which aid trajectory analysis. In addition, the growing use of single-cell "multiome"-sequencing which is now available by 10X Genomics as well and enables simultaneous measurements of gene expression and TF-binding motifs in nuclei, will increase the understanding for the gene regulatory mechanisms behind cell fate commitments and lineage specifications, as well as in neurogenic and gliogenic programs. The first dataset to use this methodology on human fetal cortex during mid-gestation was published in 2021 by Greenleaf's lab, where they identified two progenitor types of astrocyte-like glial precursors and trained a neural network model to infer base-pair-resolved cell-type specific chromatin accessibility profiles to detect mutations in *de novo* mutations in TF binding sites of ASD individuals (Trevino et al. 2021). Another atlas on human cortical development, applying scATAC-seq emphasizes the importance of chromatin state in cell type specificity to increase the robustness in cortical organoids (Ziffra et al. 2021).

Altogether, and as reviewed earlier (Vinsland and Linnarsson 2022) other brain regions beyond cortex – with the few earlier mentioned exceptions – are understudied in human development. Most of these studies also focused on later timepoints with the earliest starting from the middle or end of the first trimester. The largest published dataset that covers a wider age-span in the first trimester, looked at multiple brain regions in human development including: telencephalon, diencephalon, midbrain, hindbrain and cerebellum. It includes a molecular and spatial single-cell atlas of 289 000 whole cells (using the Chromium protocol) on early stages of the neuroepithelium with timepoints from Carnegie Stage (CS) 12 to 22, which corresponds to approximately p.c.w 4 – 8 (Eze et al. 2020). This study highlights the early transition from neuroepithelial cells to early radial glia showing that even samples at very early timepoints of cortical tissue contains a small number of neurons. Moreover, nine distinct progenitor populations were identified anatomically close to the telencephalon, including a mesenchymal-like population (*ALX* and *LUM*-expressing), which disappeared at the start of neurogenesis. Although, early cortical development was the centre of this study, the other brain regions in this dataset remain to be further explored in relation to the few other studies existing on similar tissues.

A second study, used single-cell combinatorial indexing (sci-RNAseq) to sample nuclei from the cerebrum and cerebellar tissue, with 2 005 512 and 1 372 010 nuclei respectively, but from older samples ranging from 72 days to 129 days post-conception (~10 – 18 p.c.w.) (Cao et al. 2020). This was part of a bigger dataset of 4 million single cells and nuclei from 15 different organs that was sequenced at a depth of ~ 14 000 reads per cell and a recovery of 863 molecules and 524 genes on median, in contrast to other single-cell protocols that enables the detection of a higher molecule and gene counts (Mereu et al. 2020). Here, the main focus was not on the brain, but will serve a useful reference for future single-cell datasets on the developing brain.

To conclude, of all single-cell studies that exist on human brain development, most of them have found major cell types and states like neuronal progenitors, intermediate progenitors, maturing neurons (excitatory and inhibitory), various glial populations as well as brain vascular cells, with emphasis on cortical tissues. Yet, many questions remain to be answered with respect to each cell type and between brain regions. Among these cell types, interneurons have been a topic of discussion since their origin is not revealed. They are known to undergo tangential migration from the ganglionic eminences into the cortex during development and occurs around E11.5 to E16 in mouse, but when this process starts in human is not entirely clear (J. S. Hu et al. 2017; Lim et al. 2018; Reinchisi et al. 2012). The question is whether interneurons acquire their identity in their birth place (ganglionic eminences), during or after migration. An interesting finding in mice shows that xenografted human cortical progenitors can give rise to both excitatory and inhibitory cortical neurons (Delgado et al. 2022). The growing use of human brain organoids are now being expanded into other tissue-specific models than cortex and will serve as a complementary model system to research that mainly rely on the direct use of human primary tissues such as scRNA-seq (Uzquiano and Arlotta 2022). Lately, optimization of cortical organoid protocols has shown the impact of cellular stress and how that influences cell-type specification *in vitro* (Bhaduri et al. 2020). In addition to rodents, studies on non-human primate brain development also serve as an important role in understanding species-conservations in the brain (Y. Zhu et al. 2018; Pollen et al. 2019).

3.3.2 Transcriptomic diversity in the context of neurodevelopmental disorders and diseases

Studying human brain development is also of interest in order to gain understanding behind the cause of certain disorders, such as neuropsychiatric disorders or brain malformations that are thought to arise during development. A few related examples are schizophrenia, autism spectrum disorders or epilepsy (Subramanian, Calcagnotto, and Paredes 2020). However, human study designs of neurodevelopmental disorders are even more difficult as some of them manifest in the adult brain even if the foundation of the cause is thought to arise during development (Owen et al. 2011). Recently, a few publications using scRNA-seq, assessed the role of specific developmental-related diseases and disorders in human brain such as the risk for developing neuropsychiatric disorders. One of them attempted to link gene signatures from a human schizophrenia GWAS database to specific cell types in the brain (Cameron et al. 2022). Here, the authors analysed single nuclei of fetal brain from the second trimester to assess

whether common risk alleles for schizophrenia were enriched in genes specific to certain cell types and found schizophrenia genetic risk were enriched in genes highly expressed primarily in a group of excitatory neurons in the forebrain and hippocampus as well as selected group of inhibitory neurons in the ganglionic eminence. This is one example among few other studies that previously used the same approach (Trevino et al. 2021; Polioudakis et al. 2019). It should be added that the underlying cause for schizophrenia is still debatable within the field.

Other difficulties in studying developmental-related disorders lie in their polygenic nature or idiopathic incidents where genetic factors cannot be established at all, which is the case in certain cerebellar malformations. This was previously shown in an analysis of the physiological and genetic components in a quite extensive cohort with cerebellar malformations (282 individuals) from 100 families. By applying SPLiT-seq on prenatal second-trimester cerebellar tissue, the authors related some genetic components observed in the patient cohort to fetal cerebellar neuronal and vascular cell types (Aldinger et al. 2019). This led them to hypothesize that abnormal vasculogenesis in cerebellum may cause cerebellar malformations, which underlines the importance of vasculogenesis during development (Ross et al. 2020; Tata, Ruhrberg, and Fantin 2015; Potente and Mäkinen 2017). In addition to these studies on the aberrant brain, innovative research has paved the way for novel findings in adult and pediatric brain tumors. Similar to other cancers, brain malignancies are heterogeneous in their nature and origin of the brain (Lulla, Saratsis, and Hashizume 2016). Many brain-derived tumors have been proven to mimic developmental programs of neurons and glia, which partially reflects their severity. This developmental pattern which predominantly includes a major part of cycling cells and together with other tumor environmental features enables brain tumor progression (C. Liu and Zong 2012; Azzarelli, Simons, and Philpott 2018). Thus, understanding how the healthy human brain develops has many implications besides the aspiration to learn the basics of neurodevelopmental mechanisms.

4 RESEARCH AIMS

As a continuation of the previous the cell-atlas publications in our group; the overall aim of this thesis was to investigate the emergence of cell types and their transcriptional heterogeneity in the developing human brain during the critical first trimester through the use of scRNA-seq.

The specific aims for the presented papers were the following (contributions highlighted in bold):

Paper I

- Perform large-scale scRNA-seq on major CNS and PNS regions of juvenile mice.
- Develop and perform analysis for cell type discovery of the whole nervous system.
- Describe the cellular axonomy in relation to known an unknown biological phenomena.
- **Validate some of the findings spatially, *in situ* (which was done on the discovered astrocyte heterogeneity)**.

Paper II

- Develop a method and tool to identify and realign sequencing reads to intronic regions.
- Develop a theoretical model, differential equation to calculate first time-derivative, i.e. RNA velocities on the retrieved gene expression matrix.
- Develop the computational pipeline that practically uses this realignment strategy and model to calculate velocities on the input gene expression matrices.
- Validate the model on different cell- types, lineages and states including bulk-RNA-seq data.
- **Spatially validate some inferred velocity trajectories (which was done in the human embryonic cortex at 10 p.c.w)**.

Paper III

- Perform large-scale scRNA-seq experiments (on the million-scale), on major brain regions from human developing brain in the first trimester.
- Develop a strategy and framework for analysing all regions together for cell- type and state discovery.
- Spatially validate the single-cell data by multiplex smFISH (done with the EEL-FISH).

5 ETHICAL CONSIDERATIONS

The experimental use of human embryonic tissue is a highly sensitive topic. This section intends to emphasize the ethical assessments and precautions that have been undertaken for this thesis. Paper III includes human embryonic samples that have been obtained because of elective terminated gestations. According to EU regulations, the destruction of embryos is not allowed. This is not violated by our project, as the tissue is already destined to be discarded since the donors are not aware of the option of donation until the decision of abortion has been made. In addition, the tissue is not saved after the experimental procedures and cannot be further used for other research purposes. This project has been approved by the Swedish Ethical Review Authority as well as by the National Research Ethics Service in England (see Paper III). Before the elective abortion, information is provided about the donation and a written consent from the donor is required for participating in the research study. An important issue is that the donor's genetic profile can be mapped from the gene expression data obtained from RNA sequencing of the embryo. There is a limitation of how much information can be traced back to the donor from the embryo/fetus. However, since it is theoretically possible to obtain a DNA profile of the donor, this has to be accounted for in our ethical considerations.

An issue that has been brought up and discussed regards the possibility of reidentification of the male parent of an aborted fetus and whether he needs to consent to the tissue donation. Consent from the father is hard to implement as he may not always be aware of the pregnancy. Additionally, it may also violate the mother's autonomy to decide about the abortion, as in Sweden, the fetus is not a subject, but is legally and ethically considered part of the woman's body. Another hypothetical scenario may occur where both parents have independently participated in genomic profiling via commercially available services (e.g. 23andMe), which may reveal highly sensitive and private information about the parents' relation and that a child was conceived who was later subjected to abortion. These ethical scenarios have been weighed and the Ethics Review Board agreed with our measures taken which includes: **i)** consent of donation from the mother is enough, **ii)** the raw sequence data should be available only via controlled access repositories that have legal and technical safeguards to prevent reidentification and **iii)** the personal data should be handled through the GDPR legislation.

The anonymity of every donor is ensured by the pseudonymization of the patient's social security number. It is mainly the hospital and the people directly in contact with the delivery of the tissue, who can trace the donor ID to the patient's social security number. The research groups in this study do not have access to this information, but mainly receives the pseudonymized donor ID: s to keep track of the samples which are stored among other metadata in a database. The donor can withdraw from the study at any time, whereby the data must be removed from all sources that had access to it. All the subprojects related to this study goes under the same ethical permission. Ethical regulations for this project are of great importance as the aim is to make the final dataset openly accessed. Only the anonymized data (such as gene expression matrices) will be openly accessed in consent with the donors. Our research

project focuses mainly on gene expression, but our ethical permit does cover the profiling of DNA for looking at chromatin modifications.

Finally, research on this kind of human tissue entails a great responsibility, not only because of the ethical regulations and legislations, but also because it is a loaded topic of conversation in society and is heavily politically discussed. Therefore, it is our obligation to communicate the reason behind our research and the impact it has on society, both within and outside science.

6 RESULTS

6.1 PAPER I: MOLECULAR ARCHITECTURE OF THE MOUSE NERVOUS SYSTEM

The brain is one of the most heterogeneous tissues in terms of their cell composition. As 10X Genomics came out with a commercially available high-throughput method for collecting single cells using scRNA-seq, our group set out on a large-scale project and sampled above 500 000 cells of the adolescent mouse nervous system that resulted in 492 949 single cells after quality assessments. This was done through careful dissociation of cells from the major regions in the CNS and PNS. More precisely, the nervous system was dissected into broad anatomical regions including the brain, spinal cord and the peripheral sensory, enteric and sympathetic nervous systems. In total, 19 regions were collected covering most of the mouse nervous system except for the retina, olfactory epithelium, vomeronasal organ, inner ear and the parasympathetic ganglia that were excluded from the study. All tissues were subjected to cell sampling without sorting, except for in the intestine, where neural-crest-derived cells were selectively isolated by fluorescent-activated cell sorting (FACS) using a *Wnt1-Cre* transgenic mouse model. Cortical and hippocampal inhibitory neurons were also enriched using FACS from a transgenic *vGat-Cre* mouse (vGat being the vesicular GABA transporter, encoded by the Slc32a1 gene).

Next, an analysis pipeline, *Cytograph* was developed for cell type discovery in a mostly automated fashion. After reducing the number of oligodendrocytes that are highly abundant in the hindbrain and spinal cord, the remaining analysis of around 160 000 cells were subjected to Louvain clustering on the constructed KNN graph. The whole analysis was done in 5 major steps, by initially pooling cells from the same tissue and clustering them followed by a selective splitting procedure in the next step. In each step, some clusters where manually curated, to remove remaining doublets or clusters that were over-split. In the last step, all the cells were pooled and subjected to gene enrichment analysis and gene set discovery. Then a dendrogram of all clusters was built to construct a cellular taxonomy of the whole nervous system. For cluster-specific marker gene a “trinarization” score was developed to assess the likelihood of a gene being expressed or not. This score was used together with the calculated gene enrichment per cluster in order to find marker genes that would be unique for each cluster. On the whole, this yielded 265 high-quality clusters. In addition, a classifier was trained on the cluster labels to assess the cluster robustness of the analysis pipeline.

Broadly, this dataset revealed a hierarchical organization of almost the whole nervous system that could be traced back due to its developmental origin. Neuronal and non-neuronal populations comprised two major groups, that further subdivided into specific cell types from the CNS and PNS. All clusters could be defined by their neurotransmitter properties. Major cell classes such as: neurons, astrocytes, ependymal, oligodendrocytes, vascular and immune cells were described. The diversity among neurons was mostly driven by genes encoding cell type-specific TFs, synaptic proteins, neurotransmitters and membrane conductance proteins

(ion channels, solute carriers etc.). We found that this neuronal diversity was similar across all brain regions. Another major discovery included seven distinct clusters of region-specific astrocytes that could be labeled as telencephalic and non-telencephalic groups. Some of these region-specific genes included *Mfge8* (forebrain) and *Agt* (from diencephalon to hindbrain). Furthermore, neurons from the CNS showed distinct spatial distributions when mapped to the Allen Mouse Brain Atlas. Altogether, this dataset comprises a detailed overview of the transcriptional diversity in the mouse CNS that can be used as a reference to understand the normal brain or aid in the design of transgenic mouse models to target specific cell populations.

6.2 PAPER II: RNA VELOCITY OF SINGLE CELLS

With scRNA-seq data we typically measure a snapshot of a cell's transcriptional profile at a specific point in time. Unless a such an experiment is designed in a way that a given group of cells are sampled with a shift in time, this kind of data lacks a time component. Other studies using bulk RNA data, have previously shown that the splicing and degradation rate of genes can be determined by the relative abundance of nascent (unspliced) and mature (spliced) mRNA (Zeisel et al. 2011; Gray et al. 2014; Gaidatzis et al. 2015). Building on this concept, RNA velocity was developed in order to estimate the temporal dynamics in single cells on already existing scRNA-seq data. Briefly, the main framework includes counting unspliced RNA in single-cell data by distinguishing intronic reads during the genome alignment. In theory, only mature polyadenylated mRNA should be captured with the use of oligo(dT)-primers in an experiment. In reality, 15-25 % of unspliced intronic reads were present in a number of single-cell datasets processed with different protocols (SMART-seq2, STRT/C1, inDrop, Chromium). This is likely due to the priming of internal poly-A sequences in the mRNA transcript.

With these observations, a computational pipeline was developed to realign scRNA-seq data, now including intronic reads based on a stringent definition of the genome mapping. Next, a theoretical model was developed that estimates the first time derivative of the spliced mRNA abundance, in order to predict the future state of a cell. This derivative is determined by the relative abundance of spliced and unspliced mRNA, under the assumption of constant rates of mRNA splicing and degradation. With these assumptions, RNA velocity can be used to infer the transcriptional dynamics that are specific to each gene.

RNA metabolism usually occurs on the time-scale of hours, making RNA velocity informative on a timescale comparable to the gene expression changes taking place during the differentiation of a cell. To demonstrate the applicability of RNA velocity, we examined bulk RNA-seq data from mouse liver cells measured during a time-course experiment of 24 hours. This showed that, for circadian genes, the unspliced mRNA abundance was predictive of the spliced mRNA at the next timepoint. Furthermore, this model was tested on different published single-cell datasets of cells that undergo cell differentiation or maturation. Using a dataset of the developing mouse hippocampus, RNA velocity was able to describe the cellular fates of

hippocampal neuronal lineages, shown by projected velocity fields on a t-SNE originating from radial glia cells that transitioned into either glial and neuronal fates. In human cortex at 10 p.c.w., the trajectory dynamics of differentiating cortical excitatory neurons could be inferred, predicting the future states of radial glia into neuroblasts, immature neurons and neurons. Using human embryonic tissue from the corresponding timepoints, we validated these findings *in situ* using RNA *in situ* hybridization with known gene markers that label these cortical cellular states.

Taken together, RNA velocity yields vector fields with the size and direction indicating the speed and fate of differentiating cells. It estimates the future gene expression state of cells, particularly in developmental settings and was shown to work in datasets generated with different scRNA-seq protocols. Furthermore, this method was demonstrated to work in multiple cell lineages in both mouse and human. Thus, RNA velocity provides a tool that illuminates the transition from one cell state to another in cellular trajectories that are otherwise not recognizable in single-cell data.

6.3 PAPER III: COMPREHENSIVE CELL ATLAS OF THE FIRST-TRIMESTER DEVELOPING HUMAN BRAIN

As our lab has been interested in understanding the cellular diversity of the brain, this study was a follow up on the previous cell-atlases generated in the mouse. Here, we aimed to understand how the human brain develops and differ from the mouse. In this study, we performed scRNA-seq, using the droplet-based Chromium platform (10X Genomics) on all major brain regions of the human developing brain during the first trimester and collected over 2 million cells. We targeted timepoints from 5 to 14 p.c.w., which spans the first to the beginning of the second trimester. The whole dataset resulted in a collection of 26 donors (embryos, fetuses), of which 15 were female. After stringent quality control, we obtained 1 665 937 cells that were included in the downstream analysis. Our analysis strategy consisted of pooling all the cells of the dataset and followed by integration at the chemistry level (Chromium v2 and v3) using Harmony. In short, we used the Leiden clustering algorithm on the KNN graph, which primarily yielded 40 main clusters. By re-clustering each of the main clusters into new subsets we increased the cluster resolution further. All final clusters were pooled into a complete dataset that resulted in 616 robust clusters of all cellular states and types. We identified 12 major cell classes including: radial glia, nIPCs, neuroblasts, neurons, glioblasts, OPCs, immune, vascular, placode- and neural crest-derived cells, erythrocytes and fibroblasts. At a superior level, the whole dataset was grouped into radial glia and neurons, i.e. half of the cells were clearly present in cell cycle whereas the others were post-mitotic.

As expected in development, we found that most of the cell classes were regionally diverse, with the strongest regionalization seen in the neuronal and glial lineages. More surprisingly, was the finding of highly regionalized glioblasts that were organized as telencephalic and non-telencephalic progenitors, reminiscent of the major astrocyte subtypes previously seen in the

adult mouse in **Paper I**. We further defined a maturing group of glioblasts that expressed *GJA1* and *AQP4* as pre-astrocytes. We also confirmed the presence of regionalized OPCs that seem to originate from *EGFR*-positive cells, bridging OPCs and the pre-astrocytic populations. This contrasts with the adult mouse brain (**Paper I**), where only one OPC population was identified, but agrees with the findings in the adult human brain (Siletti et al 2022, paper in review).

This dataset revealed an unprecedented high resolution in some different brain regions, allowing us to resolve lineages particularly in the forebrain. In the telencephalon, where we had fairly equal sampling of all timepoints, we identified the lineage progression of cortical excitatory neurons in the pallium by extracting *EMX1*-positive cells and were able to dissect the cell cycle phases of radial glia and nIPCs transitioning into maturing neurons.

Finally, we applied EEL-FISH (Borm et al. 2022) on a human whole embryo at 5 p.c.w. targeting over 400 genes that revealed the spatial distribution of the early patterning of the neural tube. This allowed us to map the single-cell clusters of the corresponding timepoint to their spatial location in the neural tube, by using a modified version of the Tangram algorithm (Biancalani et al. 2021). Based on a selective collection of gene expression patterns from the EEL-FISH genes and the prosomeric model (L Puelles 2009), we anatomically annotated all major segments including the germinal zones in the hindbrain of the neural tube. Altogether, this serves as a comprehensive study of the early human developing brain during the critical first trimester that could be partially confirmed *in situ*. We believe that it will become a useful resource for other developmental-related research areas and aid the understanding of brain development in human and other species.

7 CONCLUSIONS AND PERSPECTIVES

The existing and almost endlessly emerging biomedical high-throughput technologies have unsealed new concepts and opportunities for answering fundamental biological questions. It is therefore worth emphasizing the impact that these technologies have on our approach to design and form new experiments and hypotheses. scRNA-seq has especially contributed to the large-scale, unbiased discoveries of cell types, their dynamics and tissue heterogeneity, all of which have an apparent role of systems biology today. Despite, systems biology seeming like a young field, the concept has been around for quite some time with the modern terminology introduced in the late 1960s (Trewavas 2006; Kesić 2016). This concept is fundamental to how we explore science as it goes against the reductionistic way of thinking. However, biology at that time, that was dependent on other disciplines such as chemistry and physics (e.g. the discovery of the DNA structure), had to practically be approached in a reductionist view. Furthermore, many pioneering ideas during the mid-20th century could not be put into practice because of the technical evolution lagging behind. As time has gone by, many theories such as constructing gene regulatory networks of a cell's molecular machinery, have advanced in parallel with the technical development of experimental biology. Today, we are able to compose large-scale

experiments and measure thousands of genes in thousands of single cells, all in one experiment. As a result, we are facing an anti-reductionistic view of approaching biology with all the emerging technologies related to single-cell genomics. Data-driven and hypothesis-generating research can be overwhelming, many times because the research question is not always well-defined beforehand, as is the case with many cell-atlasing projects. On the other hand, observing biological phenomenon representative in a whole biological system (a cell, whole organism etc.) in an unbiased way, may lead to new and exciting findings that sometimes even contradict traditional views or explanations. In that sense, this thesis exemplifies the contribution of scRNA-seq to the overall discoveries in tissues with limited access, such as the developing human brain, to give an overview of species-conserved and unique developmental mechanisms as well as illuminating the potential of these findings in the context of disease.

To summarize and put our findings into perspective, the adult mouse brain atlas in **Paper I** demonstrates that it is possible to generate a cellular and molecular taxonomy of the nervous system that can further be used to describe anatomical landmarks in the brain. This also shows that single-cell gene expression profiles at the cluster level can be mapped to a spatial reference atlas of the same species, which was done by using gene expression voxels from *in situ* hybridization data from the Allen Mouse Brain Atlas. Furthermore, this kind of data indicates the possibility of relating a molecular organization of cell types to its developmental origin. Altogether, this study shows that scRNA-seq provides a tool for understanding cell type heterogeneity in a complex tissue and is able to capture fine gene expression signatures that are possible to locate *in situ* based on a few combinations of genes. This was partially demonstrated by revealing the spatial preference of astrocytes through RNA *in situ* hybridization by using a small combinatorial gene set from the single-cell data, in order to identify their telencephalic and non-telencephalic populations. As a whole, this cell taxonomy atlas of the mouse nervous system constitutes a resource for basic and mechanistic biology, which has been widely used by others.

The development of RNA velocity in **Paper II** marks a good example of how theories usually evolve in conjunction with technological advancement. By exploring existing sequencing data, RNA velocity shows how "side-effects" in sequencing protocols can lead to a new conceptual framework for analyzing scRNA-seq data. Inferring transcriptional dynamics in single cells has been enabled with the realization that internal priming of poly-A stretches of RNA molecules is a consequence of several single-cell protocols and results in the detection of intronic sequences in about a quarter of the sequencing reads. Furthermore, this computational method has not only opened up a new way of analyzing transcriptional cellular states, but also highlighted previously unnoticed phenomena occurring in widely used sequencing protocols which is now being further investigated and mapped (Svoboda, Frost, and Bosco 2022). Another strength with this method is highlighted by the possibility to estimate the future state of a cell on already existing single-cell data. Validations on bulk RNA-seq data of the mouse liver targeting circadian genes, various tissues undergoing cellular differentiation including human tissue, demonstrate the robustness of the method. We showed that applying RNA

velocity on single-cell data of embryonic human cortical tissue, captures major cells states of the excitatory neuronal lineage. By RNA *in situ* hybridization, we confirmed that these cell states are in fact present in corresponding human tissue of matching age. RNA velocity has thereby contributed to a new way of extrapolating transcriptional profiles of single-cells sampled at one specific timepoint, into their future states. The caveats of this model have been rapidly addressed by others in the community (Bergen et al. 2020; 2021), that have refined it. For example, RNA velocity postulates that all genes have a similar splicing rate while scVelo redefines previous assumptions in the model in order to estimate gene-specific splicing rates (Bergen et al. 2020). We are continuously learning more about the applications and pitfalls of velocity-based algorithms and will gain further knowledge by complementary datasets measuring other modalities, such as observing cell fate decisions at the chromatin level which was recently shown by Tedesco et al. 2022 through Chromatin Velocity.

Our last effort in **Paper III**, shows the power of sampling a great number of cells from access-limited tissues like the human developing brain, which have provided profound insights into specific developmental events. We observed major patterning events across the neural tube and maturing brain. Moreover, the astrocytic heterogeneity found in the adult and developing mouse brain (Zeisel et al. 2018; Manno et al. 2020) was confirmed and extended in human with the presence of telencephalic and non-telencephalic glioblasts. The observation of region-specific OPCs confirms an early heterogeneity emerging during development that hitherto was difficult to detect, demonstrating the importance of consistent sampling of multiple timepoints and regions during a developmental time window. We delineated the cortical excitatory neuronal trajectory in the pallium of the telencephalon, including their gradual maturation and cell cycle phases in both radial glia and nIPCs, proving the detailed resolution of this dataset. The cell cycle modeling in pallium also shows the application of RNA velocity through the use of DeepCycle. In the subpallium, we identified the lateral-, medial- and caudal- ganglionic eminences (LGE; MGE; CGE) and describe the migration of telencephalic interneurons to cortical tissues. In addition, we were able to see patterns of a minor group of CGE- and MGE-like neurons in the thalamus and hypothalamus that were not expressing *FOXG1,* thus these cells are thought to originate elsewhere than telencephalon. Finally, we provide a unique spatial dataset of the neural tube at 5 p.c.w., revealing the distribution of early patterning genes and germinal zones in the hindbrain. We were able to spatially map individual single-cell clusters containing neural progenitors and maturing neurons of the matching age. All in all, this dataset contributes to the further understanding of human brain development during the critical first trimester and underlines the emerging complexity of the brain at early timepoints.

Altogether, our studies provide significant insights into a global view of the brain's cellular complexity. As human brain development and the emergence of cell type identity is still relatively unexplored, this kind of dataset serve as a good reference for other developmental-related research. However, the opposite is also true. One among many difficulties when constructing cell-atlases of developing tissues, especially in human, lie in the annotation process of cell types and cell states. In contrast to other animal models such as mice,

reproducibility is often compromised in study designs encompassing rare human tissues which may confound results that is beyond our control. Individual variations influence the accuracy of estimating the gestational age in human and the precision in distinguishing anatomical landmarks in the brain is especially difficult during early development. In addition, human studies of this kind lack internal controls in the sense that there is no standardized way of estimating how well our sampling represents the whole brain for each timepoint. We can overcome some of these limitations by repeatedly measuring a large number of cells with scRNA-seq, assuming that we get a representative depiction of the developing brain. Despite that, we still heavily rely on previous knowledge not only of human, but also of other species whose developmental time course have been densely studied. Early studies on human brain development have typically included histological examinations using chromogenic or immunohistological stainings, in one example targeting as early as 28 days of gestation (Müller and O'Rahilly 1988) and we are continuously finding new anatomical patterns that has not been described before (Nascimento et al. 2022). These histological studies are essential for the validation of our findings based on the single-cell data. On the contrary, we are still exploring the specificity of these histological data that usually measure a targeted set of proteins that may overlap between multiple cell types. Our findings together with other studies of similar kind, complement our understanding of cell type-specificity and the transient cell states during development. As important are the spatial distributions of emerging cell types during development, which is why we also depend on spatial transcriptomics among other spatial techniques to assess the tissue composition.

With the vast number of datasets that have been generated with single-cell technologies, the future of single-cell biology is facing the need and challenge to integrate this great collection of cells across different tissues. The Human Developmental Cell Atlas is one among several existing consortia striving to do this with the aim to generate a reference map of all cells during human development (Haniffa et al. 2021). Our data on human brain development is part of this consortium and will hopefully provide new insights into the developing brain together with other present datasets. Moreover, our dataset may serve as a useful reference for validating *in vitro* experiments including many cell differentiation protocols of various kinds such as human induced pluripotent stem cells used for generating midbrain dopaminergic neurons as a treatment strategy for Parkinson's disease (Studer 2012; Ásgrímsdóttir and Arenas 2020). However, while an assembly of millions of cells obtained from various sources increases the chances to observe biological phenomena in different tissues, it is as important to be able to interpret the integration of all datasets in a profound way. Beyond being able to successfully integrate data, the question remains as to which extent resulting integrations masks any biological features unique to a dataset. Particularly problematic is the integration of multiple datasets containing cells that undergo dynamical transformations as is seen during development. Comparing clusters (i.e. cell types) between datasets is not always informative in the sense that these often represent a gradual change of a cell population, as is noted in our developmental data. Similarly, aligning datasets – containing cells from different experimental conditions or a disease state – to a reference dataset, is also a challenge. To deal with this,

efforts have been made to align cell trajectories between datasets (Sugihara et al. 2022; Alpert et al. 2022). While there are multiple tools for aligning query datasets to reference atlases, many are still being developed and some of them specifically addressing strategies to integrate datasets between species.

Developmental biology has entered a new era where technological advances have made it possible to design large-scale studies without predefined hypotheses. As sequencing costs has drastically decreased and improvements in single-cell genomics methods have dramatically improved, we have seen a great number of studies in human tissue and will continue to see an increase in such studies of different disease models including patient-derived specimens. Current available single-cell datasets in mouse development will serve as a template for understanding and translating findings in human development. The rise of multi-modal single-cell protocols that captures not only RNA, but also chromatin modifications and protein levels, will contribute to a better understanding of gene regulations that can be correlated to phenotypes observed during developmental processes of an organism (C. Zhu, Preissl, and Ren 2020; Kashima et al. 2020).

Science cannot solve the ultimate mystery of nature. And that is because, in the last analysis, we ourselves are a part of the mystery that we are trying to solve.

Max Planck

8 ACKNOWLEDGEMENTS

This thesis is really a product by the mentorship and support I have received during this journey, so I wish to express my gratitude here.

First of all, **Sten**, thank you so much for giving me this opportunity and for the chance to learn completely new things and showing that nothing is impossible even for the inexperienced like me. You really opened up a world to me that I didn't know existed and for that I am immensely grateful. I especially admire your philosophical spirit and creativity in the scientific process and I'm always fascinated by how fast you turn ideas into practice. I also enjoyed the spontaneous moments along our projects and your encouragement to leave room for unexpected findings and projects. Thank you for your patience, your mentorship and support in the good, but also in the tough and stressful moments and for being fair and responsive during those times. It's emotional to leave the lab after so long, but I'm thankful for what you have taught me and the knowledge I have gained so far, not the least how to conduct science in an honest, open and solid way. Also, thank you for all the nice dinners and for being a great cook!

Ernest, thank you for your kindness and for being my co-supervisor. I have learnt a lot from you, not only about developmental biology, but also about the scientific process. Thanks for always having an open door, for sparing time and giving me critical feedback in different moments, and for doing so even when I'm running late with deadlines! I'm always impressed by your dedication to science and I'm really thankful for your support and contribution on our final project and happy how well it turned out!

Thank you, **Aparna Bhaduri**, for taking your time to be my opponent and for travelling all the way to Sweden in the middle of February. It's really an honor and I'm looking forward to our discussion. To **Thomas Perlmann**, **Taija Mäkinen** and **András Simon,** thank you very much for bringing a great combination of expertise and serving as the examination board during my dissertation.

To all the other PI's of Mol Neuro: **Patrik**, **Ulrika**, **Gonçalo**, **Jens**, **Per** and **Onur**, thank you for making the unit such a nice workplace, promoting collaborations between the groups and enabling us to do great science while keeping up the good atmosphere in the unit in many ways.

Erik Sundström, this project would not have been possible without you. Thanks for collaborating with us, being on call, organizing deliveries and your invaluable knowledge in embryology. **Xiaofei**, it was really fun working with you. Thanks for all the support on the project, your hard work on optimizing protocols and for reminding me that the CNS extends to the spinal cord…

To the Linnarsson lab, thanks you so much, you're all incredibly smart and creative! **Anna,** thanks for always supporting and being there for us when in need. I can't imagine the lab without you. Thanks for your spontaneous friendly visits to our office, and more lately, for

your input on how to best succeed with growing kale! **Miri,** I can't express enough how happy I am that we got to work together and I wouldn't have made it through without you. I have learnt so much from you and your positivity is contagious. You are really an inspiration and I look up to you both as a scientist and person. I'm always impressed of your eye for details and your way of materializing complicated tasks. Thank you for supporting me, for teaching and complementing me along the project. Also, thanks again for introducing me to my neighbours! Since then, I have started to question "coincidences" in life… **Peter** thanks for bringing science and music to the lab and for your invaluable knowledge in bioinformatics. I'm really grateful for your contributions and for bearing with my constant computer-related questions. You play the harpsichord as well as you play the (computer) keyboard. Also, thanks for letting me experience Händel's Messiah once and for all. **Lijuan**, how many libraries did we process again? You're my partner in crime for collecting cells in this project. Thank you for teaching me, providing mental support and helping me during all the experiments and for putting up with the spontaneous tissue deliveries, and sorry for always having you cancelling your laundry slots because of that… **Kimberly** thanks for being a great desk friend and more lately my co-supervisor. I have learnt so much from you, thanks for your scientific input, support and for letting me constantly disturbing you with numerous Python- and non-Python-related questions during these years. I'm excited to see what your future lab brings! **Alejandro** you've already paid a telepathic visit as I'm writing. But I'll say it anyway and keep it short (for the long version, check your desk). You're really an inspiration to science. Thanks for bringing me to laughter without saying a word, for managing my crises, for brightening the room when visiting our office and more importantly; for your prophecies as the 'Gene Oracle'. **Camiel**, thanks for taking brain development to a new modality and for all the interesting discussions, I have learnt a lot from you including unexpected things that I didn't know about! I appreciate you always listening to my scientific obsession about cell size. Thanks for your thoughtfulness and for the really fun times in Japan! **Razieh**, thanks for your utterly sweetness, for always spreading positivity and being a great listener. I admire your dedication to science and philosophical view on things. Also, thanks for calming me down whenever I decide to install something in R. **Jesper**, it's great to have an MD in the lab. I'm still not sure how you manage time, I believe you exist in a parallel universe where you do all experiments while being in the clinics. I suspect Pepsi Max has something do with it, but I guess I'll never know. Thanks for all the fun! **Jokubas**, I think you were born in the lab. Your dry sense of humor is really something extra and your scientific critical view on things. I'm looking forward to meet the future you. **Egle**, we didn't overlap for long, but you're just one of those people who spread good vibes. Thanks for your inspiration, for emphasizing the importance of macrophages and bringing back my interest for immunology. I'm so thankful for you chairing my defense! **Ivana** thank you for your kindness, fun stories and for making the new generation in the lab. Also, thanks a lot for being the stand-in for Alejandro as the crisis manager. **Elin V**, thanks for brining your anatomy skills and keeping up the single-cell spirit in the lab. And for sharing interest in big earrings! **David FG** for your passion about gene regulatory networks. After the late evenings I've passed by you and Jokubas, I'm convinced that you two could start a philosophical talk show in Biomedicum. **Marina** for organizing and bringing new knowledge to the lab.

To the former lab members: **Lars**, I'm so happy that we shared this journey! You're an amazing scientist and I've learnt so much from you during our work together. Thanks for always being there, your great taste in music, ironic humor and for bringing a chilled view on things whenever I'm overthinking. Also, thanks for having any kind of equipment prepared for any unexpected moment. **Dieke**, you were indirectly part of the lab, so thanks for all the fun and for your great taste and expertise in design. **Simone**, thanks for your limitless knowledge and for listening to my constant nonsense. You've have been like a dictionary in the lab, always having a solution to things and viewing things with a critical eye for us to improve. Thanks for teaching me in the early and late days, for sharing great stories and not laughing at my fake-Italian. And thanks for great pizza! **Gioele**, I'll never forget your exciting pitch about the Linnarsson lab. Thanks for teaching and supporting me, for giving advice on everything from scientific problems to career advice and for the opportunity to contribute with small things in the lab. I enjoyed the endless philosophical discussions in the office, and I often find myself thinking what your opinion would be whenever I'm reading a paper or learning something new. **Job**, it was great to be your desk neighbor in the very beginning and get tips and help with programming. Also, I'll never forget the impact a Fairphone can have. **Hannah**, thanks for making me feel welcome in the lab in the early days, for bearing with my experimental questions and for taking time and teaching me single-cell biology in the lab. I'm so grateful for what you have taught me. Also, thanks for timing your defense with my master's graduation and for leaving behind some treasures on your shelf in the office, I found a lot of nice papers there! **Amit**, I have so much to thank you for and I owe you a great debt of gratitude. You're a role model and I wouldn't be here without you. Thanks for being a great mentor, having patience with me and all my broken micropipettes and flat Bioanalyzer readouts, for not making me feel stupid when I made mistakes and for constantly pushing me into the right direction. I value your input and I often hear you in the back of my head whenever I'm doing things in the lab (even though I usually have to fail in my own way before I listen to you…). Also, thanks for introducing me to Matti Caspi…!

To my other friends in Mol Neuro: **Ka Wai**, I have so much fun with you! I don't know how many dinners I owe you by now, but they're coming… I'm constantly learning from you both scientifically and in the kitchen. I'm so happy that we got to work a bit together in the end, despite the short time, so thanks for digging out the dopamine in our data, for making the best cakes and pastries, like ever, for feeding me during stressful moments and for your endless support in the good and bad times. I really admire your creativity and humor for things. **Emilia**, thanks for all our great times, wine evenings and for always bearing with me in any kind of situation. I love how our weird conversations go anywhere from morphogenetic gradients (usually WNTs…) to "what kind of dress should I wear to this party? " I'm grateful for everything, for encouraging my interest in Western blots and for teaching me the importance of having a baking dish at home. I'm always impressed by your cell differentiations, and I'm excited to see what you'll be differentiating next! **Ibrahim** thanks for your kindness, for sharing the Seinfeld humor and for having a relaxed view on things when I don't. Thanks for the great times in the US, I had so much fun! And for being a technical hero and support during

panicky moments. I owe you, and maybe you will get to try that lobster one day… I also like how you always have a life hack to things! **David vB**, thanks for advertising single-cells already in the early days with the RNA Club. It was fun to share the office with you and I enjoyed our conversations a lot. Thanks for the great times in Japan and more lately fun times in the US. You're an amazing host, even without major furniture and a coffee table made out of cardboard. **Alessandro**, thanks for all the nice dinners and culinary recommendations! I appreciate your critical view on things and detailed knowledge in the lab. Thanks for thinking outside the box and never trivialize even the smallest things in science or my naïve questions. **Kasra**, it's funny how we ended up in Mol Neuro together after being in high school. Thanks for always bringing a good atmosphere around and for sharing papers on cell size. **Petra**, thanks for providing me with all kinds of snacks and chocolate when in need. And for sharing the passion for music and recommendations on violin concertos. One day we'll get through that Czardas… **Chao**, for letting Petra share your snacks with me and for your humor. **Bastien**, for bringing the French cuisine to Mol Neuro. Now I know what a chinois is. **Eneritz**, thanks for being a great office mate and for sharing your incredible knowledge OPCs and bioinformatics. **Lisa**, for your kindness, nice discussions and for bringing the mathematical intellect to Mol Neuro. **Tony**, thanks for managing and bringing class to the unit. Your opinions on fashion and art means a lot to me. **Wei**, for your sweetness and always having a positive view on things. **Natalie**, for being a sweet office neighbor and keeping the Sunday spirit in the lab. Soon we'll go to Faschin… **Chiara**, I only know you for a short time, but it's great that you joined the unit. Maybe one day you'll understand my Italian. **Mukund**, for your inspiring work in the lab and your kindness. **Dmitry**, for your kindness, honesty and wisdom. **Carmen S**, for always being so caring and always having a positive view on things. It was great to be neighbors for a short time. **Roland**, for your Friday essays and emphasizing the importance of Swedish fika. Thanks for caring for us in Mol Neuro. **Leslie**, thanks for the fun times and great input on conferences. **Neemat**, for your interesting stories and contagious joy. **Hayley**, for embracing the Swedish culinary kitchen. **Komgrit**, for sharing interest in the RNA club before. **Müge**, for bringing the new generation of PhD students. **Elin G**, for the nice lunch conversations and sharing the Täby hoods. **Alessandra R**, for always being so sweet. **Yizhou**, for your single-cell enthusiasm. To **Milda**, **Lisbeth**, **Carmen A**, **Leonardo**, **Anqi**, **Jie**, **Mingdong**, **Songbai**, **Michael**, **Milda**, **Zoe**, **Jussi**, **Karl**, **Dongoh**, **Yonglong** thank you for making Mol Neuro a nice place. **Connla** for your funny view on things and sharing interesting stories. **Göran** for always being so "hurtig" and sharing exciting stories as well. Thanks to the old Mol Neuro members: **Shanzheng**, **Anneke**, **Jose**, **Martin**, **Ana M**, **Ana F**, **Rakesh**, **Puneet**, **Daouha**. **Lily**, for the nice chats in the office and for your kindness. **Daniel**, for sharing your knowledge in the lab and your funny Halloween costumes. **Mandy**, for hosting great cheese evenings and board game nights. **Dagmara**, for welcoming me in the early days and insisting on me joining the ski trips. And for embracing the cold Swedish Valborg tradition that one time. **Carolina**, for your sweetness and for letting me bombarding you with Patch-seq-related questions when I was a master student. I'm happy we overlapped sometime later as well. **Viktoria**, thanks for also insisting on me going on the ski conference with Dagmara. I had a lot of fun in Trysil. **Hermany**, for spreading joy and great dance moves. **Sam**, for sharing the experience of being

an exam guard. That was really something… **Elisa**, for showing what spontaneity actually means. How can one miss a *DNA tour* with Backstreet Boys? **Lauri**, for the great discussions and philosophical view on things. **Ivar**, for welcoming me to the office, sharing inside information and random funny facts.

To the **Silberberg group**, thank you for teaching me e-phys, the location of striatum and the importance of a good hummus. **Gilad**, I've learnt so much from being in your group. Thank you for giving me the chance, your mentorship and for letting me play around and design my own project in your lab. I didn't produce any useful data, but I learnt a ton in the process. Thanks for listening to my interests back then, which also led me to the Linnarsson group. And big thanks for sending me to your friend Amit, to learn molecular biology. **Maya**, you're an *in vivo* rockstar, thanks for always supporting me, for making great hummus and picking delicate wines. **Yvonne**, I had so much fun and thanks for always answering my questions when in need. **Matthijs**, thanks for teaching me how to patch-clamp and for emphasizing suit-up Fridays. **Ania**, thanks for letting me in your project, teaching me e-phys and having patience with me. **Roberto**, for always being so friendly and having a chilled view on things. **Zach**, for sharing late evenings in the lab, giving me advice and listening to my scientific blabbing. **Elin**, for supporting us with all necessary things in the lab back then!

To the first lab in Zürich: **Martin**, thanks for welcoming me to your lab as a newbie and including me in all the activities and conferences. You've had a great influence on me even later on. You were a great mentor and I learnt so much being in your group. **Corinna**, it was so nice to work with you. Thanks for teaching me, your patience and for sparking my interest in running! Big thanks, **Jenny**, **Vanessa, Oleg** and **Walter** for making the best out of my exchange. I had so much fun!

University friends: **Gio** your positive energy is contagious. I admire your enormous drive in life, and your way of looking at things. You have a mind of steel and always see the bright side of things. I'm so grateful for everything and for all your support. I'm sorry that I didn't visit you as much as I wanted to, but I hope to track you down on your new adventure in Australia! **Linda** communicating with you is a joy, especially when it comes to food. I always enjoy our discussions and your critical thinking. Thanks for your chilled view on certain things and you bringing me to sense whenever I'm feeling overwhelmed. I failed to visit you after you moved, but thank you for coming back to Sweden. I guess we'll plan a trip to a new continent soon! **Laura**, thanks for your sweetness and for always spreading joy. You're such an ambitious person and thanks for reminding me that the gym exists. **Dominyka**, it was fun to share part of the PhD journey. I'm happy to see you so ambitious and succeeding with so many things! Thanks for all the fun and for embracing the Swedish culture nowadays. **Ilke**, thanks for always being such a great host and the fun parties! It's funny that we became neighbors in Biomedicum and I'm looking forward to see result of the hard work and what you've been up to (as I always see you through the lab window). **Emma**, I had so much fun with you. Thanks for correcting my English grammar and for sharing the love for Nutella! **Helena**, it's funny that we both ended up in single-cell genomics. Thanks for your sweetness, always being rational and

enjoying life to the fullest. **Anton**, what have you not studied? I enjoyed your interdisciplinary approach in the undergrads already. I don't often come out of my nest, but I appreciate you and Ilke bringing me to after works occasionally. **Bobbi**, I never get tired of your hilarious stories. You're a great storyteller and thanks for introducing me to the Bulgarian culture. **Benjamin**, I don't think we've ever had a chat that lasted less than 30 min. It's always great talking to you and I've never seen anyone so caught up in politics. **Sara**, thanks for always being so sweet, a power woman and for coming back to KI. **Lucas**, thanks for the hilarious moments and for just being you. **Nathaniel**, for being weirdly fun in a really nerdy way and old times. **Karolina**, for putting Gotland on the map to a Stockholm snob like me. **Emmie**, even though we didn't overlap for long, thanks for sharing the burden (and notes) in organic chemistry. I'll never forget that. To the **CHaSE team 2014**: **Tiago**, **Francesca**, **Thibaud**, **Moritz**, **Cécile**, **Christian, Michael**, **Su**, **Liza**, **Florian** and **Daniel.** Thanks for being my role models at the time and still. And for the great parties and events. I have no idea what I was doing back then, but now I have a hunch of why you were all so stressed doing your PhDs… To some people at **6B**: **Leo**, for the fun times in Japan. You're surprisingly good at Karaoke. Thanks for always spreading good vibes even when you're annoyed. **Christoph**, because of you I have to keep up with Smart-seq even if I don't use it. I still think you and Michael should develop a mathematical constant. **Ionut**, thanks for always being so sweet and for the nice party.

Josefin, thanks for keeping in touch during all year, I always enjoy discussions about art. **Freja**, thanks for also staying in touch, for keeping the music alive and introducing me to vegan restaurants.

To my family: **Lila**, thanks for being the best aunt ever, for your amazing food and for always looking after me during all these years. **Jerzyk**, for your kindness and passion for medicine! **Vicky**, your sweetness and being the best-dressing cousin. **Marc**, for always listening to my meaningless babbling. **Oscar**, for matching Vicky's outfits. **Nike** and **Ruben**, you're both just supercute. **Emanuel**, thanks for the fun times in Israel.

My brothers, **David** and **Daniel**, you're always there for me in any moment. Thanks for bearing with my indecisiveness and silly issues throughout the years. I have so much to thank you for. **Camilla**, thanks for always welcoming me to your home and caring for me. **Denice** and **Alice**, I can't believe you grew so fast, thanks for being the best nieces!

Mamma and **Pappa**, thanks for introducing me to the best of two worlds: *science* and *music*. Behind every thesis is a mother who cooks and provides lovely food, and sometimes a father who constantly challenges your scientific critical thinking. Thanks for giving me absolutely everything in life, for supporting me in all moments and for believing in me. You're both amazing with so many ambitions. Thanks for teaching me the important things in life and for always showing interest in what I do. Mamma, thanks for being a wise 'mama bear', so artistic and philosophical at the same time. Pappa, thanks for your great sense of humor, infinite science-obsession and the sourdough bread of course.

9 REFERENCES

Akay, Leyla Anne, Audrey H. Effenberger, and Li Huei Tsai. 2021. "Cell of All Trades: Oligodendrocyte Precursor Cells in Synaptic, Vascular, and Immune Function." *Genes and Development*. Cold Spring Harbor Laboratory Press. https://doi.org/10.1101/GAD.344218.120.

Aldinger, Kimberly A., Zachary Thomson, Ian G. Phelps, Parthiv Haldipur, Mei Deng, Andrew E. Timms, Matthew Hirano, et al. 2021. "Spatial and Cell Type Transcriptional Landscape of Human Cerebellar Development." *Nature Neuroscience* 24 (8): 1163–75. https://doi.org/10.1038/s41593-021-00872-y.

Aldinger, Kimberly A., Andrew E. Timms, Zachary Thomson, Ghayda M. Mirzaa, James T. Bennett, Alexander B. Rosenberg, Charles M. Roco, et al. 2019. "Redefining the Etiologic Landscape of Cerebellar Malformations." *American Journal of Human Genetics* 105 (3): 606–15. https://doi.org/10.1016/j.ajhg.2019.07.019.

Alemany, Anna, Maria Florescu, Chloé S. Baron, Josi Peterson-Maduro, and Alexander Van Oudenaarden. 2018. "Whole-Organism Clone Tracing Using Single-Cell Sequencing." *Nature* 556 (7699): 108–12. https://doi.org/10.1038/nature25969.

Alpert, Ayelet, Ornit Nahman, Elina Starosvetsky, Michal Hayun, Tyler J. Curiel, Yishai Ofran, and Shai S. Shen-Orr. 2022. "Alignment of Single-Cell Trajectories by TuMap Enables High-Resolution Quantitative Comparison of Cancer Samples." *Cell Systems* 13 (1): 71-82.e8. https://doi.org/10.1016/j.cels.2021.09.003.

Anderson, T. W., and I. Olkin. 1985. "Maximum-Likelihood Estimation of the Parameters of a Multivariate Normal Distribution." *Linear Algebra and Its Applications* 70 (C): 147–71. https://doi.org/10.1016/0024-3795(85)90049-7.

Andrews, Madeline G., and Tomasz J. Nowakowski. 2019. "Human Brain Development through the Lens of Cerebral Organoid Models." *Brain Research*. Elsevier B.V. https://doi.org/10.1016/j.brainres.2019.146470.

Andriezen, W. Lloyd. 1893. "The Neuroglia Elements in the Human Brain." *British Medical Journal* 2 (1700): 227–30. https://doi.org/10.1136/bmj.2.1700.227.

Anthony, Todd E., and Nathaniel Heintz. 2008. "Genetic Lineage Tracing Defines Distinct Neurogenic and Gliogenic Stages of Ventral Telencephalic Radial Glial Development." *Neural Development* 3 (1): 30. https://doi.org/10.1186/1749-8104-3-30.

Anthony, Todd E., Corinna Klein, Gord Fishell, and Nathaniel Heintz. 2004. "Radial Glia Serve as Neuronal Progenitors in All Regions of the Central Nervous System." *Neuron* 41 (6): 881–90. https://doi.org/10.1016/S0896-6273(04)00140-0.

Arendt, Detlev, Jacob M. Musser, Clare V.H. Baker, Aviv Bergman, Connie Cepko, Douglas H. Erwin, Mihaela Pavlicev, et al. 2016. "The Origin and Evolution of Cell Types." *Nature Reviews Genetics*. Nature Publishing Group. https://doi.org/10.1038/nrg.2016.127.

Ásgrímsdóttir, Emilía Sif, and Ernest Arenas. 2020. "Midbrain Dopaminergic Neuron Development at the Single Cell Level: In Vivo and in Stem Cells." *Frontiers in Cell and Developmental Biology*. Frontiers Media S.A. https://doi.org/10.3389/fcell.2020.00463.

Auguste, Yohan S.S., Austin Ferro, Jessica A. Kahng, Andre M. Xavier, Jessica R. Dixon, Uma Vrudhula, Anne Sarah Nichitiu, et al. 2022. "Oligodendrocyte Precursor Cells Engulf Synapses during Circuit Remodeling in Mice." *Nature Neuroscience* 25 (10): 1273–78. https://doi.org/10.1038/s41593-022-01170-x.

Azzarelli, Roberta, Benjamin D. Simons, and Anna Philpott. 2018. "The Developmental Origin of Brain Tumours: A Cellular and Molecular Framework." *Development (Cambridge)*. Company of Biologists Ltd. https://doi.org/10.1242/dev.162693.

Bakker, Bernadette S. de, Kees H. de Jong, Jaco Hagoort, Karel de Bree, Clara T. Besselink, Froukje E. C. de Kanter, Tyas Veldhuis, et al. 2016. "An Interactive Three-Dimensional Digital Atlas and Quantitative Database of Human Development." *Science* 354 (6315): aag0053. https://doi.org/10.1126/science.aag0053.

Balaban, Evan, Marie Aimée Teillet, and Nicole L.E. Douarin. 1988. "Application of the Quail-Chick Chimera System to the Study of Brain Development and Behavior." *Science* 241 (4871): 1339–42. https://doi.org/10.1126/science.3413496.

Bayraktar, Omer Ali, Theresa Bartels, Staffan Holmqvist, Vitalii Kleshchevnikov, Araks Martirosyan, Damon Polioudakis, Lucile Ben Haim, et al. 2020. "Astrocyte Layers in the Mammalian Cerebral Cortex Revealed by a Single-Cell in Situ Transcriptomic Map." *Nature Neuroscience* 23 (4): 500–509. https://doi.org/10.1038/s41593-020-0602-1.

Beattie, Robert, and Simon Hippenmeyer. 2017. "Mechanisms of Radial Glia Progenitor Cell Lineage Progression." *FEBS Letters*. Wiley Blackwell. https://doi.org/10.1002/1873-3468.12906.

Becht, Etienne, Leland McInnes, John Healy, Charles Antoine Dutertre, Immanuel W.H. Kwok, Lai Guan Ng, Florent Ginhoux, and Evan W. Newell. 2019. "Dimensionality Reduction for Visualizing Single-Cell Data Using UMAP." *Nature Biotechnology* 37 (1): 38–47. https://doi.org/10.1038/nbt.4314.

Bergen, Volker, Marius Lange, Stefan Peidli, F. Alexander Wolf, and Fabian J. Theis. 2020. "Generalizing RNA Velocity to Transient Cell States through Dynamical Modeling." *Nature Biotechnology* 38 (12): 1408–14. https://doi.org/10.1038/s41587-020-0591-3.

Bergen, Volker, Ruslan A Soldatov, Peter V Kharchenko, and Fabian J Theis. 2021. "RNA Velocity—Current Challenges and Future Perspectives." *Molecular Systems Biology* 17 (8): e10282. https://doi.org/10.15252/msb.202110282.

Bergles, Dwight E., and William D. Richardson. 2016. "Oligodendrocyte Development and Plasticity." *Cold Spring Harbor Perspectives in Biology* 8 (2): a020453. https://doi.org/10.1101/cshperspect.a020453.

Bergmann, Olaf, Jakob Liebl, Samuel Bernard, Kanar Alkass, Maggie S.Y. Yeung, Peter Steier, Walter Kutschera, et al. 2012. "The Age of Olfactory Bulb Neurons in Humans." *Neuron* 74 (4): 634–39. https://doi.org/10.1016/j.neuron.2012.03.030.

Bershteyn, Marina, and Arnold R. Kriegstein. 2013. "Cerebral Organoids in a Dish: Progress and Prospects." *Cell*. Cell Press. https://doi.org/10.1016/j.cell.2013.09.010.

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. https://doi.org/10.1038/s41586-020-1962-0.

Biancalani, Tommaso, Gabriele Scalia, Lorenzo Buffoni, Raghav Avasthi, Ziqing Lu, Aman Sanger, Neriman Tokcan, et al. 2021. "Deep Learning and Alignment of Spatially Resolved Single-Cell Transcriptomes with Tangram." *Nature Methods* 18 (11): 1352–62. https://doi.org/10.1038/s41592-021-01264-7.

Bocchi, Vittoria Dickinson, Paola Conforti, Elena Vezzoli, Dario Besusso, Claudio Cappadona, Tiziana Lischetti, Maura Galimberti, et al. 2021. "The Coding and Long Noncoding Single-Cell Atlas of the Developing Human Fetal Striatum." *Science* 372 (6542). https://doi.org/10.1126/science.abf5759.

Boda, Enrica, Martina Lorenzati, Roberta Parolisi, Brian Harding, Gianmarco Pallavicini, Luca Bonfanti, Amanda Moccia, Stephanie Bielas, Ferdinando Di Cunto, and Annalisa Buffo. 2022. "Molecular and Functional Heterogeneity in Dorsal and Ventral Oligodendrocyte Progenitor Cells of the Mouse Forebrain in Response to DNA Damage." *Nature Communications* 13 (1): 1–17. https://doi.org/10.1038/s41467-022-30010-6.

Borm, Lars E., Alejandro Mossi Albiach, Camiel C.A. Mannens, Jokubas Janusauskas, Ceren Özgün, David Fernández-García, Rebecca Hodge, et al. 2022. "Scalable in Situ Single-Cell Profiling by Electrophoretic Capture of MRNA Using EEL FISH." *Nature Biotechnology*, September, 1–10. https://doi.org/10.1038/s41587-022-01455-3.

Brady, Gerard, Mary Barbara, and Norman N Iscove. 1990. "Representative in Vitro CDNA Amplification From Individual Hemopoietic Cells and Colonies." *METHODS IN MOLECULAR AND CELLULAR BIOLOGY*. Vol. 2.

Braun, Emelie, Miri Danan-Gotthold, Lars E Borm, Elin Vinsland, Ka Wai Lee, Peter Lönnerberg, Lijuan Hu, et al. 2022. "Comprehensive Cell Atlas of the First-Trimester Developing Human Brain." *BioRxiv*, January, 2022.10.24.513487. https://doi.org/10.1101/2022.10.24.513487.

Bray, Nicolas L., Harold Pimentel, Páll Melsted, and Lior Pachter. 2016. "Near-Optimal Probabilistic RNA-Seq Quantification." *Nature Biotechnology* 34 (5): 525–27. https://doi.org/10.1038/nbt.3519.

Breunig, Joshua J., Tarik F. Haydar, and Pasko Rakic. 2011. "Neural Stem Cells: Historical Perspective and Future Prospects." *Neuron*. NIH Public Access. https://doi.org/10.1016/j.neuron.2011.05.005.

Bronstein, Robert, Jackson Kyle, Ariel B. Abraham, and Stella E. Tsirka. 2017. "Neurogenic to Gliogenic Fate Transition Perturbed by Loss of HMGB2." *Frontiers in Molecular Neuroscience* 10 (May): 153. https://doi.org/10.3389/fnmol.2017.00153.

Bruggen, David van, Fabio Pohl, Christoffer Mattsson Langseth, Petra Kukanja, Hower Lee, Alejandro Mossi Albiach, Mukund Kabbe, et al. 2022. "Developmental Landscape of Human Forebrain at a Single-Cell Level Identifies Early Waves of Oligodendrogenesis." *Developmental Cell* 57 (11): 1421-1436.e5. https://doi.org/10.1016/j.devcel.2022.04.016.

Brüning, Ralf Schulze, Lukas Tombor, Marcel H. Schulz, Stefanie Dimmeler, and David John. 2022. "Comparative Analysis of Common Alignment Tools for Single-Cell RNA Sequencing." *GigaScience* 11 (December): 1–12. https://doi.org/10.1093/gigascience/giac001.

Buchanan, JoAnn, Leila Elabbady, Forrest Collman, Nikolas L. Jorstad, Trygve E. Bakken, Carolyn Ott, Jenna Glatzer, et al. 2022. "Oligodendrocyte Precursor Cells Ingest Axons in the Mouse Neocortex." *Proceedings of the National Academy of Sciences* 119 (48): e2202580119. https://doi.org/10.1073/pnas.2202580119.

Bulfone, A., L. Puelles, M. H. Porteus, M. A. Frohman, G. R. Martin, and J. L.R. Rubenstein. 1993. "Spatially Restricted Expression of Dlx-1, Dlx-2 (Tes-1), Gbx-2, and Wnt-3 in the Embryonic Day 12.5 Mouse Forebrain Defines Potential Transverse and Longitudinal Segmental Boundaries." *Journal of Neuroscience* 13 (7): 3155–72. https://doi.org/10.1523/jneurosci.13-07-03155.1993.

Burrows, Robert C., Deborah Wancio, Pat Levitt, and Laura Lillien. 1997. "Response Diversity and the Timing of Progenitor Cell Maturation Are Regulated by Developmental Changes in EGFR Expression in the Cortex." *Neuron* 19 (2): 251–67. https://doi.org/10.1016/S0896-6273(00)80937-X.

Byrum, Christine A, Naveen M Wijesena, and Athula H Wikramanayake. 2012. "Autonomous Cell Fate Specification: Overview." In *ELS*. Chichester: John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470015902.a0001148.pub3.

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*. Cell Press. https://doi.org/10.1016/j.neuron.2019.07.009.

Cadwell, Cathryn R., Federico Scala, Shuang Li, Giulia Livrizzi, Shan Shen, Rickard Sandberg, Xiaolong Jiang, and Andreas S. Tolias. 2017. "Multimodal Profiling of Single-Cell Morphology, Electrophysiology, and Gene Expression Using Patch-Seq." *Nature Protocols* 12 (12): 2531–53. https://doi.org/10.1038/nprot.2017.120.

Cai, Jun, Yingchuan Qi, Xuemei Hu, Min Tan, Zijing Liu, Jianshe Zhang, Qun Li, Maike Sander, and Mengsheng Qiu. 2005. "Generation of Oligodendrocyte Precursor Cells from Mouse Dorsal Spinal Cord Independent of Nkx6 Regulation and Shh Signaling." *Neuron* 45 (1): 41–53. https://doi.org/10.1016/j.neuron.2004.12.028.

Cameron, Darren, Da Mi, Ngoc Nga Vinh, Caleb Webber, Meng Li, Oscar Marín, Michael C. O'Donovan, and Nicholas J. Bray. 2022. "Single-Nuclei RNA Sequencing of 5 Regions of the Human Prenatal Brain Implicates Developing Neuron Populations in Genetic Risk for Schizophrenia." *Biological Psychiatry* 93 (2): 157–66. https://doi.org/10.1016/j.biopsych.2022.06.033.

Cao, Junyue, Diana R. O'Day, Hannah A. Pliner, Paul D. Kingsley, Mei Deng, Riza M. Daza, Michael A. Zager, et al. 2020. "A Human Cell Atlas of Fetal Gene Expression." *Science* 370 (6518). https://doi.org/10.1126/science.aba7721.

Casper, Kristen B., and Ken D. McCarthy. 2006. "GFAP-Positive Progenitor Cells Produce Neurons and Oligodendrocytes throughout the CNS." *Molecular and Cellular Neuroscience* 31 (4): 676–84. https://doi.org/10.1016/j.mcn.2005.12.006.

Choi, Kwangbom, Yang Chen, Daniel A. Skelly, and Gary A. Churchill. 2020. "Bayesian Model Selection Reveals Biological Origins of Zero Inflation in Single-Cell Transcriptomics." *Genome Biology* 21 (1): 183. https://doi.org/10.1186/s13059-020-02103-2.

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. https://doi.org/10.1038/s41592-018-0175-z.

Curtis, Maurice A., Monica Kam, Ulf Nannmark, Michelle F. Anderson, Mathilda Zetterstrom Axell, Carsten Wikkelso, Stig Holtås, et al. 2007. "Human Neuroblasts Migrate to the Olfactory Bulb via a Lateral Ventricular Extension." *Science* 315 (5816): 1243–49. https://doi.org/10.1126/science.1136281.

Delaunay, Delphine, Katharina Heydon, Ana Cumano, Markus H. Schwab, Jean Léon Thomas, Ueli Suter, Klaus Armin Nave, Bernard Zalc, and Nathalie Spassky. 2008. "Early Neuronal and Glial Fate Restriction of Embryonic Neural Stem Cells." *Journal of Neuroscience* 28 (10): 2551–62. https://doi.org/10.1523/JNEUROSCI.5497-07.2008.

Delgado, Ryan N., Denise E. Allen, Matthew G. Keefe, Walter R. Mancia Leon, Ryan S. Ziffra, Elizabeth E. Crouch, Arturo Alvarez-Buylla, and Tomasz J. Nowakowski. 2022. "Individual Human Cortical Progenitors Can Produce Excitatory and Inhibitory Neurons." *Nature* 601 (7893): 397–403. https://doi.org/10.1038/s41586-021-04230-7.

Depledge, Daniel P., Kalanghad Puthankalam Srinivas, Tomohiko Sadaoka, Devin Bready, Yasuko Mori, Dimitris G. Placantonakis, Ian Mohr, and Angus C. Wilson. 2019. "Direct RNA Sequencing on Nanopore Arrays Redefines the Transcriptional Complexity of a Viral Pathogen." *Nature Communications* 10 (1): 1–13. https://doi.org/10.1038/s41467-019-08734-9.

Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. "STAR: Ultrafast Universal RNA-Seq Aligner." *Bioinformatics* 29 (1): 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Dong, Rui, Ran Yang, Yong Zhan, Hua Dong Lai, Chun Jing Ye, Xiao Ying Yao, Wen Qin Luo, et al. 2020. "Single-Cell Characterization of Malignant Phenotypes and Developmental Trajectories of Adrenal Neuroblastoma." *Cancer Cell* 38 (5): 716-733.e6. https://doi.org/10.1016/j.ccell.2020.08.014.

Dupont, Jean Claude. 2017. "Wilhelm His and Mechanistic Approaches to Development at the Time of Entwicklungsmechanik." *History and Philosophy of the Life Sciences* 39 (3): 21. https://doi.org/10.1007/s40656-017-0148-z.

Eberwine, James, Hermes Yeh, Kevin Miyashiro, Yanxiang Cao, Suresh Nair, Richard Finnell, Martha Zettel, and Paul Coleman. 1992. "Analysis of Gene Expression in Single Live Neurons." *Neurobiology*. Vol. 89.

Efremova, Mirjana, Miquel Vento-Tormo, Sarah A. Teichmann, and Roser Vento-Tormo. 2020. "CellPhoneDB: Inferring Cell–Cell Communication from Combined Expression of Multi-Subunit Ligand–Receptor Complexes." *Nature Protocols* 15 (4): 1484–1506. https://doi.org/10.1038/s41596-020-0292-x.

Elowitz, Michael B., Arnold J. Levine, Eric D. Siggia, and Peter S. Swain. 2002. "Stochastic Gene Expression in a Single Cell." *Science* 297 (5584): 1183–86. https://doi.org/10.1126/science.1070919.

Erhard, Florian, Marisa A.P. Baptista, Tobias Krammer, Thomas Hennig, Marius Lange, Panagiota Arampatzi, Christopher S. Jürges, Fabian J. Theis, Antoine Emmanuel Saliba, and Lars Dölken. 2019. "ScSLAM-Seq Reveals Core Features of Transcription Dynamics in Single Cells." *Nature* 571 (7765): 419–23. https://doi.org/10.1038/s41586-019-1369-y.

Ernst, Aurélie, Kanar Alkass, Samuel Bernard, Mehran Salehpour, Shira Perl, John Tisdale, Göran Possnert, Henrik Druid, and Jonas Frisén. 2014. "Neurogenesis in the Striatum of the Adult Human Brain." *Cell* 156 (5): 1072–83. https://doi.org/10.1016/j.cell.2014.01.044.

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 (4): 584–94. https://doi.org/10.1038/s41593-020-00794-1.

Eze, Ugomma C, Aparna Bhaduri, Maximilian Haeussler, Tomasz J Nowakowski, and Arnold R Kriegstein. 2020. "Single-Cell Atlas of Early Human Brain Development Highlights Heterogeneity of Human Neuroepithelial Cells and Early Radial Glia." *BioRxiv*. https://doi.org/10.1101/2020.03.06.981423.

Fan, Xiaoying, Ji Dong, Suijuan Zhong, Yuan Wei, Qian Wu, Liying Yan, Jun Yong, et al. 2018a. "Spatial Transcriptomic Survey of Human Embryonic Cerebral Cortex by Single-Cell RNA-Seq Analysis." *Cell Research* 28 (7): 730–45. https://doi.org/10.1038/s41422-018-0053-3.

Fernandez-Castaneda, Anthony, and Alban Gaultier. 2016. "Adult Oligodendrocyte Progenitor Cells – Multifaceted Regulators of the CNS in Health and Disease." *Brain, Behavior, and Immunity*. Academic Press Inc. https://doi.org/10.1016/j.bbi.2016.01.005.

Fernández, Virginia, Cristina Llinares-Benadero, and Víctor Borrell. 2016. "Cerebral Cortex Expansion and Folding: What Have We Learned?" *The EMBO Journal* 35 (10): 1021–44. https://doi.org/10.15252/embj.201593701.

Fu, Yuanyuan, Ming Yang, Hongmin Yu, Yicheng Wang, Xinglong Wu, Jun Yong, Yunuo Mao, et al. 2021. "Heterogeneity of Glial Progenitor Cells during the Neurogenesis-to-Gliogenesis Switch in the Developing Human Cerebral Cortex." *Cell Reports* 34 (9). https://doi.org/10.1016/j.celrep.2021.108788.

Fuzik, Janos, Amit Zeisel, Zoltan Mate, Daniela Calvigioni, Yuchio Yanagawa, Gabor Szabo, Sten Linnarsson, and Tibor Harkany. 2016. "Integration of Electrophysiological Recordings with Single-Cell RNA-Seq Data Identifies Neuronal Subtypes." *Nature Biotechnology* 34 (2): 175–83. https://doi.org/10.1038/nbt.3443.

Gaidatzis, Dimos, Lukas Burger, Maria Florescu, and Michael B. Stadler. 2015. "Analysis of Intronic and Exonic Reads in RNA-Seq Data Characterizes Transcriptional and Post-Transcriptional Regulation." *Nature Biotechnology* 33 (7): 722–29. https://doi.org/10.1038/nbt.3269.

Garcia-Marin, Virginia, Pablo Garcia-Lopez, and Miguel Freire. 2009. "The Growth Cone as Seen through Cajal's Original Histological Preparations and Publications." *Journal of the History of the Neurosciences*. J Hist Neurosci. https://doi.org/10.1080/09647040801961430.

Gelder, R. N. Van, M. E. Von Zastrow, A. Yool, W. C. Dement, J. D. Barchas, and J. H. Eberwine. 1990. "Amplified RNA Synthesized from Limited Quantities of Heterogeneous CDNA." *Proceedings of the National Academy of Sciences of the United States of America* 87 (5): 1663–67. https://doi.org/10.1073/pnas.87.5.1663.

Ghaddar, Bassel, and Subhajyoti De. 2022. "Reconstructing Physical Cell Interaction Networks from Single-Cell Data Using Neighbor-Seq." *Nucleic Acids Research* 50 (14): e82–e82. https://doi.org/10.1093/nar/gkac333.

Giladi, Amir, Merav Cohen, Chiara Medaglia, Yael Baran, Baoguo Li, Mor Zada, Pierre Bost, et al. 2020. "Dissecting Cellular Crosstalk by Sequencing Physically Interacting Cells." *Nature Biotechnology* 38 (5): 629–37. https://doi.org/10.1038/s41587-020-0442-2.

Gilbert, Scott F. 2000a. "The Embryological Origins of the Gene Theory." In *Developmental Biology*, 6th ed. Sunderland (MA): SInauer Associates.

Gilbert Scott F. 2000b. "The Neural Crest." In , 6th ed. Sinauer Associates. https://www.ncbi.nlm.nih.gov/books/NBK10065/.

Gilbert, Scott F. 2000c. "Tissue Architecture of the Central Nervous System." https://www.ncbi.nlm.nih.gov/books/NBK10047/.

Gluecksohn-Schoenheimer, S. 1940. "THE EFFECT OF AN EARLY LETHAL (T0) IN THE HOUSE MOUSE." *Genetics* 25 (4).

Gray, Jesse M., David A. Harmin, Sarah A. Boswell, Nicole Cloonan, Thomas E. Mullen, Joseph J. Ling, Nimrod Miller, et al. 2014. "SnapShot-Seq: A Method for Extracting Genome-Wide, In Vivo MRNA Dynamics from a Single Total RNA Sample." Edited by Zhuang Zuo. *PLoS ONE* 9 (2): e89673. https://doi.org/10.1371/journal.pone.0089673.

Grove, E A, and E S Monuki. 2013. *Morphogens, Patterning Centers, and Their Mechanisms of Action: Patterning and Cell Type Specification in the Developing CNS and PNS. Pattering and Cell Type Specification in the Developing CNS and PNS*. Vol. #volume#. Elsevier Inc. https://doi.org/10.1016/B978-0-12-397265-1.00041-1.

Guo, Guoji, Mikael Huss, Guo Qing Tong, Chaoyang Wang, Li Li Sun, Neil D. Clarke, and Paul Robson. 2010. "Resolution of Cell Fate Decisions Revealed by Single-Cell Gene Expression Analysis from Zygote to Blastocyst." *Developmental Cell* 18 (4): 675–85. https://doi.org/10.1016/j.devcel.2010.02.012.

Hafemeister, Christoph, and Rahul Satija. 2019. "Normalization and Variance Stabilization of Single-Cell RNA-Seq Data Using Regularized Negative Binomial Regression." *Genome Biology* 20 (1): 296. https://doi.org/10.1186/s13059-019-1874-1.

Hagemann-Jensen, Michael, Christoph Ziegenhain, Ping Chen, Daniel Ramsköld, Gert Jan Hendriks, Anton J.M. Larsson, Omid R. Faridani, and Rickard Sandberg. 2020. "Single-Cell RNA Counting at Allele and Isoform Resolution Using Smart-Seq3." *Nature Biotechnology* 38 (6): 708–14. https://doi.org/10.1038/s41587-020-0497-0.

Haghverdi, Laleh, Aaron T.L. Lun, Michael D. Morgan, and John C. Marioni. 2018. "Batch Effects in Single-Cell RNA-Sequencing Data Are Corrected by Matching Mutual Nearest Neighbors." *Nature Biotechnology* 36 (5): 421–27. https://doi.org/10.1038/nbt.4091.

Hamburger, Viktor. 1981. "Historical Landmarks in Neurogenesis." *Trends in Neurosciences* 4 (C): 151–55. https://doi.org/10.1016/0166-2236(81)90051-5.

Haniffa, Muzlifah, Deanne Taylor, Sten Linnarsson, Bruce J. Aronow, Gary D. Bader, Roger A. Barker, Pablo G. Camara, et al. 2021. "A Roadmap for the Human Developmental Cell Atlas." *Nature*. Nature Research. https://doi.org/10.1038/s41586-021-03620-1.

Haque, Ashraful, Jessica Engel, Sarah A Teichmann, and Tapio Lönnberg. 2017. "A Practical Guide to Single-Cell RNA-Sequencing for Biomedical Research and Clinical Applications." *Genome Medicine*. https://doi.org/10.1186/s13073-017-0467-4.

Hashimoto, Ryoya, Kei Hori, Tomoo Owa, Satoshi Miyashita, Kenichi Dewa, Norihisa Masuyama, Kazuhisa Sakai, et al. 2016. "Origins of Oligodendrocytes in the Cerebellum, Whose Development Is Controlled by the Transcription Factor, Sox9." *Mechanisms of Development* 140 (May): 25–40. https://doi.org/10.1016/j.mod.2016.02.004.

Heisenberg, C.P., C. Brennan, and S.W. Wilson. 1999. "Zebrafish Aussicht Mutant Embryos Exhibit Widespread Overexpression of Ace (Fgf8) and Coincident Defects in CNS Development." *Development* 126 (10).

Hendriks, Gert Jan, Lisa A. Jung, Anton J.M. Larsson, Michael Lidschreiber, Oscar Andersson Forsman, Katja Lidschreiber, Patrick Cramer, and Rickard Sandberg. 2019. "NASC-Seq Monitors RNA Synthesis in Single Cells." *Nature Communications* 10 (1): 1–9. https://doi.org/10.1038/s41467-019-11028-9.

Herb, Brian R, and Hannah J Glover. 2022. "Title: Single-Cell Genomics Reveals Region-Specific Developmental Trajectories Underlying Neuronal Diversity in the Human Hypothalamus." *Arnold R. Kriegstein*, no. 2 (October): 2021.07.20.453090. https://doi.org/10.1101/2021.07.20.453090.

Hernandez-Miranda, Luis R., Thomas Müller, and Carmen Birchmeier. 2017. "The Dorsal Spinal Cord and Hindbrain: From Developmental Mechanisms to Functional Circuits." *Developmental Biology*. Elsevier Inc. https://doi.org/10.1016/j.ydbio.2016.10.008.

Hidalgo-Sánchez, Matías, Marco Onorati, Andre Goffinet, and Luis Puelles. 2021. "Recollections on the Origins and Development of the Prosomeric Model." https://doi.org/10.3389/fnana.2021.787913.

Hie, Brian, Bryan Bryson, and Bonnie Berger. 2019. "Efficient Integration of Heterogeneous Single-Cell Transcriptomes Using Scanorama." *Nature Biotechnology* 37 (6): 685–91. https://doi.org/10.1038/s41587-019-0113-3.

Hill, Mark Anthony. 2018. "Developing the Digital Kyoto Collection in Education and Research." *Anatomical Record* 301 (6): 998–1003. https://doi.org/10.1002/ar.23783.

Honig, Lawrence S., Kathrin Herrmann, and Carla J. Shatz. 1996. "Developmental Changes Revealed by Immunohistochemical Markers in Human Cerebral Cortex." *Cerebral Cortex* 6 (6): 794–806. https://doi.org/10.1093/cercor/6.6.794.

Hopwood, Nick. 2007. “A History of Normal Plates, Tables and Stages in Vertebrate Embryology.” *International Journal of Developmental Biology* 51 (1): 1–26. https://doi.org/10.1387/ijdb.062189nh.

Hopwood, Nick. 2019a. “Inclusion and Exclusion in the History of Developmental Biology.” *Development (Cambridge)* 146 (7). https://doi.org/10.1242/dev.175448.

Horder, Tim. 2010. “History of Developmental Biology.” In *Encyclopedia of Life Sciences*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470015902.a0003080.pub2.

Hu, Jia Sheng, Daniel Vogt, Magnus Sandberg, and John L. Rubenstein. 2017. “Cortical Interneuron Development: A Tale of Time and Space.” *Development (Cambridge)*. Company of Biologists Ltd. https://doi.org/10.1242/dev.132852.

Hu, Yuqiong, Xiaoye Wang, Boqiang Hu, Yunuo Mao, Yidong Chen, Liying Yan, Jun Yong, et al. 2019. “Dissecting the Transcriptome Landscape of the Human Fetal Neural Retina and Retinal Pigment Epithelium by Single-Cell RNA-Seq Analysis.” *PLoS Biology* 17 (7). https://doi.org/10.1371/journal.pbio.3000365.

Huang, Wei, Aparna Bhaduri, Dmitry Velmeshev, Arturo Alvarez-Buylla, David H Rowitch, Arnold R Kriegstein, Shaohui Wang, Li Wang, and Catherine A Rottkamp. 2020. “Origins and Proliferative States of Human Oligodendrocyte Precursor Cells In Brief Ll Origins and Proliferative States of Human Oligodendrocyte Precursor Cells.” *Cell* 182: 594-608.e11. https://doi.org/10.1016/j.cell.2020.06.027.

Islam, Saiful, Una Kjällquist, Annalena Moliner, Pawel Zajac, Jian Bing Fan, Peter Lönnerberg, and Sten Linnarsson. 2011. “Characterization of the Single-Cell Transcriptional Landscape by Highly Multiplex RNA-Seq.” *Genome Research* 21 (7): 1160–67. https://doi.org/10.1101/gr.110882.110.

Islam, Saiful, Kjällquist Una, Moliner Annalena, Zajac,Pawel, Fan Jian Bing, Lönnerberg Peter, Linnarsson Sten. 2012. “Highly Multiplexed and Strand-Specific Single-Cell RNA 5′ End Sequencing.” *Nature Protocols* 7 (5): 813–28. https://doi.org/10.1038/nprot.2012.022.

Islam, Saiful, Amit Zeisel, Simon Joost, Gioele La Manno, Pawel Zajac, Maria Kasper, Peter Lönnerberg, and Sten Linnarsson. 2014. “Quantitative Single-Cell RNA-Seq with Unique Molecular Identifiers.” *Nature Methods* 11 (2): 163–66. https://doi.org/10.1038/nmeth.2772.

Ivanov, Andrej I., Alessandra C. Rovescalli, Paola Pozzi, Siuk Yoo, Brian Mozer, Hsi Ping Li, Shu Hua Yu, et al. 2004. “Genes Required for Drosophila Nervous System Development Identified by RNA Interference.” *Proceedings of the National Academy of Sciences of the United States of America* 101 (46): 16216–21. https://doi.org/10.1073/pnas.0407188101.

Jacob, François, and Jacques Monod. 1961. “Genetic Regulatory Mechanisms in the Synthesis of Proteins.” *Journal of Molecular Biology*. Academic Press. https://doi.org/10.1016/S0022-2836(61)80072-7.

Ji, Ni, and Steven W. Flavell. 2017. “Hydra: Imaging Nerve Nets in Action.” *Current Biology*. Cell Press. https://doi.org/10.1016/j.cub.2017.03.040.

Jiang, Ruochen, Tianyi Sun, Dongyuan Song, and Jingyi Jessica Li. 2022. “Statistics or Biology: The Zero-Inflation Controversy about ScRNA-Seq Data.” *Genome Biology*. BioMed Central Ltd. https://doi.org/10.1186/s13059-022-02601-5.

Juan Romero, Camino De, and Víctor Borrell. 2015. “Coevolution of Radial Glial Cells and the Cerebral Cortex.” *Glia* 63 (8): 1303–19. https://doi.org/10.1002/glia.22827.

Juric-Sekhar, Gordana, and Robert F. Hevner. 2019. “Malformations of Cerebral Cortex Development: Molecules and Mechanisms.” *Annual Review of Pathology: Mechanisms of Disease*. Annual Reviews Inc. https://doi.org/10.1146/annurev-pathmechdis-012418-012927.

Kaminow, Benjamin, Dinar Yunusov, and Alexander Dobin. 2021. “STARsolo: Accurate, Fast and Versatile Mapping/Quantification of Single-Cell and Single-Nucleus RNA-Seq Data.” *BioRxiv*, May, 2021.05.05.442755. https://doi.org/10.1101/2021.05.05.442755.

Kandel, R. Eric, H. James Schwartz, M. Thomas Jessell, A. Steven Siegelbaum, and J. A. Hudspeth. 2012. *Principles of Neural Science*. 5th ed. McGraw-Hill.

Kang, Hyun Min, Meena Subramaniam, Sasha Targ, Michelle Nguyen, Lenka Maliskova, Elizabeth McCarthy, Eunice Wan, et al. 2018. “Multiplexed Droplet Single-Cell RNA-Sequencing Using Natural Genetic Variation.” *Nature Biotechnology* 36 (1): 89–94. https://doi.org/10.1038/nbt.4042.

Karlebach, Guy, and Ron Shamir. 2008. “Modelling and Analysis of Gene Regulatory Networks.” *Nature Reviews Molecular Cell Biology*. Nature Publishing Group. https://doi.org/10.1038/nrm2503.

Kashima, Yukie, Yoshitaka Sakamoto, Keiya Kaneko, Masahide Seki, Yutaka Suzuki, and Ayako Suzuki. 2020. “Single-Cell Sequencing Techniques from Individual to Multiomics Analyses.” *Experimental and Molecular Medicine*. Springer Nature. https://doi.org/10.1038/s12276-020-00499-2.

Katada, Sayako, Jun Takouda, Takumi Nakagawa, Mizuki Honda, Katsuhide Igarashi, Takuya Imamura, Yasuyuki Ohkawa, Shoko Sato, Hitoshi Kurumizaka, and Kinichi Nakashima. 2021. “Neural Stem/Precursor Cells Dynamically Change Their Epigenetic Landscape to Differentially Respond to BMP Signaling for Fate Switching during Brain Development.” *Genes and Development* 35 (21–22): 1431–44. https://doi.org/10.1101/GAD.348797.121.

Kauffman, Stuart. 1969. “Homeostasis and Differentiation in Random Genetic Control Networks.” *Nature*. Nature Publishing Group. https://doi.org/10.1038/224177a0.

Kerwin, Janet, Yiya Yang, Paloma Merchan, Subrot Sarma, Jessica Thompson, Xunxian Wang, Juan Sandoval, Luis Puelles, Richard Baldock, and Susan Lindsay. 2010. “The HUDSEN Atlas: A Three-Dimensional (3D) Spatial Framework for Studying Gene Expression in the Developing Human Brain.” *Journal of Anatomy* 217 (4): 289–99. https://doi.org/10.1111/j.1469-7580.2010.01290.x.

Kesić, Srdjan. 2016. “Systems Biology, Emergence and Antireductionism.” *Saudi Journal of Biological Sciences* 23 (5): 584–91. https://doi.org/10.1016/j.sjbs.2015.06.015.

Kessaris, Nicoletta, Matthew Fogarty, Palma Iannarelli, Matthew Grist, Michael Wegner, and William D. Richardson. 2006. “Competing Waves of Oligodendrocytes in the Forebrain and Postnatal Elimination of an Embryonic Lineage.” *Nature Neuroscience* 9 (2): 173–79. https://doi.org/10.1038/nn1620.

Kessaris, Nicoletta, Nigel Pringle, and William D. Richardson. 2008. “Specification of CNS Glia from Neural Stem Cells in the Embryonic Neuroepithelium.” *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society. https://doi.org/10.1098/rstb.2006.2013.

Kharchenko, Peter V., Lev Silberstein, and David T. Scadden. 2014. “Bayesian Approach to Single-Cell Differential Expression Analysis.” *Nature Methods* 11 (7): 740–42. https://doi.org/10.1038/nmeth.2967.

Kivioja, Teemu, Anna Vähärautio, Kasper Karlsson, Martin Bonke, Martin Enge, Sten Linnarsson, and Jussi Taipale. 2012. “Counting Absolute Numbers of Molecules Using Unique Molecular Identifiers.” *Nature Methods* 9 (1): 72–74. https://doi.org/10.1038/nmeth.1778.

Klein, Allon M., Linas Mazutis, Ilke Akartuna, Naren Tallapragada, Adrian Veres, Victor Li, Leonid Peshkin, David A. Weitz, and Marc W. Kirschner. 2015. “Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells.” *Cell* 161 (5): 1187–1201. https://doi.org/10.1016/j.cell.2015.04.044.

Kondo, Shigeru. 2022. “The Present and Future of Turing Models in Developmental Biology.” https://doi.org/10.1242/dev.200974.

Korsunsky, Ilya, Nghia Millard, Jean Fan, Kamil Slowikowski, Fan Zhang, Kevin Wei, Yuriy Baglaenko, Michael Brenner, Po ru Loh, and Soumya Raychaudhuri. 2019. “Fast, Sensitive and Accurate Integration of Single-Cell Data with Harmony.” *Nature Methods* 16 (12): 1289–96. https://doi.org/10.1038/s41592-019-0619-0.

Kouno, Tsukasa, Jonathan Moody, Andrew Tae Jun Kwon, Youtaro Shibayama, Sachi Kato, Yi Huang, Michael Böttcher, et al. 2019. “C1 CAGE Detects Transcription Start Sites and Enhancer Activity at Single-Cell Resolution.” *Nature Communications* 10 (1): 3. https://doi.org/10.1038/s41467-018-08126-5.

Kracht, L., M. Borggrewe, S. Eskandar, N. Brouwer, S. M. Chuva de Sousa Lopes, J. D. Laman, S. A. Scherjon, J. R. Prins, S. M. Kooistra, and B. J.L. Eggen. 2020. “Human Fetal Microglia Acquire Homeostatic Immune-Sensing Properties Early in Development.” *Science* 369 (6503): 530–37. https://doi.org/10.1126/science.aba5906.

Kretzschmar, Kai, and Fiona M. Watt. 2012. “Lineage Tracing.” *Cell*. Elsevier. https://doi.org/10.1016/j.cell.2012.01.002.

Kriegstein, Arnold, and Arturo Alvarez-Buylla. 2009. “The Glial Nature of Embryonic and Adult Neural Stem Cells.” *Annual Review of Neuroscience*. https://doi.org/10.1146/annurev.neuro.051508.135600.

Kristan, William B. 2016. “Early Evolution of Neurons.” *Current Biology*. Cell Press. https://doi.org/10.1016/j.cub.2016.05.030.

Lanjewar, Samantha N., and Steven A. Sloan. 2021. “Growing Glia: Cultivating Human Stem Cell Models of Gliogenesis in Health and Disease.” *Frontiers in Cell and Developmental Biology*. Frontiers Media S.A. https://doi.org/10.3389/fcell.2021.649538.

Lause, Jan, Philipp Berens, and Dmitry Kobak. 2021. “Analytic Pearson Residuals for Normalization of Single-Cell RNA-Seq UMI Data.” *Genome Biology* 22 (1): 258. https://doi.org/10.1186/s13059-021-02451-7.

Lederer, Alex R., and Gioele La Manno. 2020. “The Emergence and Promise of Single-Cell Temporal-Omics Approaches.” *Current Opinion in Biotechnology*. Elsevier Ltd. https://doi.org/10.1016/j.copbio.2019.12.005.

Lennington, Jessica B., Zhengang Yang, and Joanne C. Conover. 2003. “Neural Stem Cells and the Regulation of Adult Neurogenesis.” *Reproductive Biology and Endocrinology*. BioMed Central. https://doi.org/10.1186/1477-7827-1-99.

Lepiemme, Fanny, Julie Stoufflet, Míriam Javier-Torrent, Gabriel Mazzucchelli, Carla G. Silva, and Laurent Nguyen. 2022. “Oligodendrocyte Precursors Guide Interneuron Migration by Unidirectional Contact Repulsion.” *Science* 376 (6595). https://doi.org/10.1126/science.abn6204.

Lim, Lynette, Da Mi, Alfredo Llorca, and Oscar Marín. 2018. “Development and Functional Diversification of Cortical Interneurons.” *Neuron*. Cell Press. https://doi.org/10.1016/j.neuron.2018.10.009.

Lindsay, Susan J., Yaobo Xu, Steven N. Lisgo, Lauren F. Harkin, Andrew J. Copp, Dianne Gerrelli, Gavin J. Clowry, et al. 2016. “HDBR Expression: A Unique Resource for Global and Individual Gene Expression Studies during Early Human Brain Development.” *Frontiers in Neuroanatomy* 10 (October): 86. https://doi.org/10.3389/fnana.2016.00086.

Lipovsek, Marcela, Cedric Bardy, Cathryn R. Cadwell, Kristen Hadley, Dmitry Kobak, and Shreejoy J. Tripathy. 2021. “Patch-Seq: Past, Present, and Future.” In *Journal of Neuroscience*, 41:937–46. Society for Neuroscience. https://doi.org/10.1523/JNEUROSCI.1653-20.2020.

Liu, Chong, and Hui Zong. 2012. “Developmental Origins of Brain Tumors.” *Current Opinion in Neurobiology*. NIH Public Access. https://doi.org/10.1016/j.conb.2012.04.012.

Liu, Jia, Cecilia Götherström, Magda Forsberg, Eva Britt Samuelsson, Jiang Wu, Cinzia Calzarossa, Outi Hovatta, Erik Sundström, and Elisabet Åkesson. 2013. “Human Neural Stem/Progenitor Cells Derived from Embryonic Stem Cells and Fetal Nervous System Present Differences in Immunogenicity and Immunomodulatory Potentials in Vitro.” *Stem Cell Research* 10 (3): 325–37. https://doi.org/10.1016/j.scr.2013.01.001.

Liu, Siyuan John, Tomasz J. Nowakowski, Alex A. Pollen, Jan H. Lui, Max A. Horlbeck, Frank J. Attenello, Daniel He, et al. 2016. “Single-Cell Analysis of Long Non-Coding RNAs in the Developing Human Neocortex.” *Genome Biology* 17 (1): 67. https://doi.org/10.1186/s13059-016-0932-1.

Liu, Zhaoyang, Dongqing Sun, and Chenfei Wang. 2022. “Evaluation of Cell-Cell Interaction Methods by Integrating Single-Cell RNA Sequencing Data with Spatial Information.” *Genome Biology* 23 (1): 218. https://doi.org/10.1186/s13059-022-02783-y.

Lopez, Romain, Jeffrey Regier, Michael B. Cole, Michael I. Jordan, and Nir Yosef. 2018. “Deep Generative Modeling for Single-Cell Transcriptomics.” *Nature Methods* 15 (12): 1053–58. https://doi.org/10.1038/s41592-018-0229-2.

Louis, Elan D., and Christian Stapf. 2001. “Unraveling the Neuron Jungle: The 1879-1886 Publications by Wilhelm His on the Embryological Development of the Human Brain.” *Archives of Neurology*. American Medical Association. https://doi.org/10.1001/archneur.58.11.1932.

Lu, Yufeng, Fion Shiau, Wenyang Yi, Suying Lu, Qian Wu, Joel D. Pearson, Alyssa Kallman, et al. 2020. "Single-Cell Analysis of Human Retina Identifies Evolutionarily Conserved and Species-Specific Mechanisms Controlling Development." *Developmental Cell* 53 (4): 473-491.e9. https://doi.org/10.1016/j.devcel.2020.04.009.

Ludwig, Leif S., Caleb A. Lareau, Jacob C. Ulirsch, Elena Christian, Christoph Muus, Lauren H. Li, Karin Pelka, et al. 2019. "Lineage Tracing in Humans Enabled by Mitochondrial Mutations and Single-Cell Genomics." *Cell* 176 (6): 1325-1339.e22. https://doi.org/10.1016/j.cell.2019.01.022.

Luecken, Malte D., M. Büttner, K. Chaichoompu, A. Danese, M. Interlandi, M. F. Mueller, D. C. Strobl, et al. 2022. "Benchmarking Atlas-Level Data Integration in Single-Cell Genomics." *Nature Methods* 19 (1): 41–50. https://doi.org/10.1038/s41592-021-01336-8.

Luecken, Malte D, and Fabian J Theis. 2019. "Current Best Practices in Single-cell RNA-seq Analysis: A Tutorial." *Molecular Systems Biology* 15 (6): e8746. https://doi.org/10.15252/msb.20188746.

Lulla, Rishi R., Amanda Muhs Saratsis, and Rintaro Hashizume. 2016. "Mutations in Chromatin Machinery and Pediatric High-Grade Glioma." *Science Advances*. American Association for the Advancement of Science. https://doi.org/10.1126/sciadv.1501354.

Lun, Aaron T.L., Samantha Riesenfeld, Tallulah Andrews, The Phuong Dao, Tomas Gomes, and John C. Marioni. 2019. "EmptyDrops: Distinguishing Cells from Empty Droplets in Droplet-Based Single-Cell RNA Sequencing Data." *Genome Biology* 20 (1): 63. https://doi.org/10.1186/s13059-019-1662-y.

Maaten, Laurens Van Der, and Geoffrey Hinton. 2008. "Visualizing Data Using T-SNE." *Journal of Machine Learning Research*. Vol. 9.

Macosko, Evan Z., Anindita Basu, Rahul Satija, James Nemesh, Karthik Shekhar, Melissa Goldman, Itay Tirosh, et al. 2015. "Highly Parallel Genome-Wide Expression Profiling of Individual Cells Using Nanoliter Droplets." *Cell* 161 (5): 1202–14. https://doi.org/10.1016/j.cell.2015.05.002.

Malatesta, Paolo, and Magdalena Götz. 2013. "Radial Glia - From Boring Cables to Stem Cell Stars." *Development (Cambridge)* 140 (3): 483–86. https://doi.org/10.1242/dev.085852.

Malatesta, Paolo, Michael A. Hack, Eva Hartfuss, Helmut Kettenmann, Wolfgang Klinkert, Frank Kirchhoff, and Magdalena Götz. 2003. "Neuronal or Glial Progeny: Regional Differences in Radial Glia Fate." *Neuron* 37 (5): 751–64. https://doi.org/10.1016/S0896-6273(03)00116-8.

Malnic, Bettina, Junzo Hirono, Takaaki Sato, and Linda B. Buck. 1999. "Combinatorial Receptor Codes for Odors." *Cell* 96 (5): 713–23. https://doi.org/10.1016/S0092-8674(00)80581-4.

Manno, Gioele La, Daniel Gyllborg, Simone Codeluppi, Kaneyasu Nishimura, Carmen Salto, Amit Zeisel, Lars E. Borm, et al. 2016. "Molecular Diversity of Midbrain Development in Mouse, Human, and Stem Cells." *Cell* 167 (2): 566-580.e19. https://doi.org/10.1016/j.cell.2016.09.027.

Manno, Gioele La, Kimberly Siletti, Alessandro Furlan, Daniel Gyllborg, Elin Vinsland, Christoffer Mattsson Langseth, Irina Khven, et al. 2020. "Molecular Architecture of the Developing Mouse Brain." *BioRxiv*. bioRxiv. https://doi.org/10.1101/2020.07.02.184051.

Manno, Gioele La, 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. https://doi.org/10.1038/s41586-018-0414-6.

Marshall, Christine A.G., Satoshi O. Suzuki, and James E. Goldman. 2003. "Gliogenic and Neurogenic Progenitors of the Subventricular Zone: Who Are They, Where Did They Come from, and Where Are They Going?" *Glia* 43 (1): 52–61. https://doi.org/10.1002/glia.10213.

Martínez-Cerdeño, Verónica, Stephen C Noctor, Hans J Ten Donkelaar, Christoph Viebahn, and Richard S Nowakowski. 2018. "Neural Progenitor Cell Terminology." https://doi.org/10.3389/fnana.2018.00104.

Martinez Arias, Alfonso, and Ben Steventon. 2018. "On the Nature and Function of Organizers." *Development (Cambridge)*. Company of Biologists Ltd. https://doi.org/10.1242/dev.159525.

Matsunami, Hiroaki, and Linda B. Buck. 1997. "A Multigene Family Encoding a Diverse Array of Putative Pheromone Receptors in Mammals." *Cell* 90 (4): 775–84. https://doi.org/10.1016/S0092-8674(00)80537-1.

Matsushita, Susumu. 1996. "Fate Mapping Study of the Endoderm of the 1.5-Day-Old Chick Embryo." *Arch Dev Biol*. Vol. 205.

Mayer, Simone, Jiadong Chen, Tomasz J Nowakowski, Alex A Pollen, and Arnold R Kriegstein. 2019. "Multimodal Single-Cell Analysis Reveals Physiological Maturation in the Developing Human Neocortex." https://doi.org/10.1016/j.neuron.2019.01.027.

Mazzarello, Paolo. 1999. "A Unifying Concept: The History of Cell Theory." *Nature Cell Biology* 1 (1): E13–15. https://doi.org/10.1038/8964.

McClintock, B. 1950. "The Origin and Behavior of Mutable Loci in Maize." *Proceedings of the National Academy of Sciences of the United States of America* 36 (6): 344–55. https://doi.org/10.1073/pnas.36.6.344.

McGinnis, Christopher S., Lyndsay M. Murrow, and Zev J. Gartner. 2019. "DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors." *Cell Systems* 8 (4): 329-337.e4. https://doi.org/10.1016/j.cels.2019.03.003.

Mereu, Elisabetta, Atefeh Lafzi, Catia Moutinho, Christoph Ziegenhain, Davis J. McCarthy, Adrián Álvarez-Varela, Eduard Batlle, et al. 2020. "Benchmarking Single-Cell RNA-Sequencing Protocols for Cell Atlas Projects." *Nature Biotechnology* 38 (6): 747–55. https://doi.org/10.1038/s41587-020-0469-4.

Miller, Freda D., and Andrée S. Gauthier. 2007. "Timing Is Everything: Making Neurons versus Glia in the Developing Cortex." *Neuron*. Cell Press. https://doi.org/10.1016/j.neuron.2007.04.019.

Morange, Michel. 2009. “What History Tells Us XVI. A Third Pillar for Molecular Biology: Molecular Embryology.” *Journal of Biosciences* 34 (1): 17–20. https://doi.org/10.1007/s12038-009-0004-8.

Moroz, Leonid L. 2009. “On the Independent Origins of Complex Brains and Neurons.” *Brain, Behavior and Evolution*. Karger Publishers. https://doi.org/10.1159/000258665.

Mullen, R.J., C.R. Buck, and A.M. Smith. 1992. “NeuN, a Neuronal Specific Nuclear Protein in Vertebrates.” *Development* 116 (1).

Müller, F., and R. O’Rahilly. 1988. “The Development of the Human Brain from a Closed Neural Tube at Stage 13.” *Anatomy and Embryology* 177 (3): 203–24. https://doi.org/10.1007/BF00321132.

Muskovic, Walter, and Joseph E. Powell. 2021. “DropletQC: Improved Identification of Empty Droplets and Damaged Cells in Single-Cell RNA-Seq Data.” *Genome Biology* 22 (1): 329. https://doi.org/10.1186/s13059-021-02547-0.

Nakagawa, Takumi, Yoshikuni Wada, Sayako Katada, and Yusuke Kishi. 2020. “Epigenetic Regulation for Acquiring Glial Identity by Neural Stem Cells during Cortical Development.” *GLIA* 68 (8): 1554–67. https://doi.org/10.1002/glia.23818.

Nakamura, H. 2013. “Midbrain Patterning: Isthmus Organizer, Tectum Regionalization, and Polarity Formation.” In *Patterning and Cell Type Specification in the Developing CNS and PNS*, 45–60. Elsevier Inc. https://doi.org/10.1016/B978-0-12-397265-1.00020-4.

Nascimento, Marcos Assis, Sean Biagiotti, Vicente Herranz-Pérez, Raymund Bueno, Chun J Ye, Taylor Abel, Juan S Rubio-Moll, et al. 2022. “Persistent Postnatal Migration of Interneurons into the Human Entorhinal Cortex.” *BioRxiv*, March, 2022.03.19.484996. https://doi.org/10.1101/2022.03.19.484996.

Neftel, C., J. Laffy, Mariella G. Filbin, Toshiro Hara, Marni E. Shore, Gilbert J. Rahme, Alyssa R. Richman, et al. 2019. “An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma.” *Cell* 178 (4): 835-849.e21. https://doi.org/10.1016/j.cell.2019.06.024.

Nieuwenhuys, Rudolf, and Luis Puelles. 2015. *Towards a New Neuromorphology*. *Towards a New Neuromorphology*. https://doi.org/10.1007/978-3-319-25693-1.

Noctor, Stephen C., Verónica Martinez-Cerdeño, Lidija Ivic, and Arnold R. Kriegstein. 2004. “Cortical Neurons Arise in Symmetric and Asymmetric Division Zones and Migrate through Specific Phases.” *Nature Neuroscience* 7 (2): 136–44. https://doi.org/10.1038/nn1172.

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. https://doi.org/10.1126/science.aap8809.

Nowakowski, Tomasz J., Alex A. Pollen, Carmen Sandoval-Espinosa, and Arnold R. Kriegstein. 2016. “Transformation of the Radial Glia Scaffold Demarcates Two Stages of Human Cerebral Cortex Development.” *Neuron* 91 (6): 1219–27. https://doi.org/10.1016/j.neuron.2016.09.005.

Nüsslein-volhard, Christiane, and Eric Wieschaus. 1980. “Mutations Affecting Segment Number and Polarity in Drosophila.” *Nature* 287 (5785): 795–801. https://doi.org/10.1038/287795a0.

Owen, Michael J., Michael C. O’Donovan, Anita Thapar, and Nicholas Craddock. 2011. “Neurodevelopmental Hypothesis of Schizophrenia.” *British Journal of Psychiatry*. Royal College of Psychiatrists. https://doi.org/10.1192/bjp.bp.110.084384.

Ozsolak, Fatih, Adam R. Platt, Dan R. Jones, Jeffrey G. Reifenberger, Lauryn E. Sass, Peter McInerney, John F. Thompson, Jayson Bowers, Mirna Jarosz, and Patrice M. Milos. 2009. “Direct RNA Sequencing.” *Nature* 461 (7265): 814–18. https://doi.org/10.1038/nature08390.

Palomero-Gallagher, Nicola, and Karl Zilles. 2019. “Cortical Layers: Cyto-, Myelo-, Receptor- and Synaptic Architecture in Human Cortical Areas.” *NeuroImage*. Academic Press Inc. https://doi.org/10.1016/j.neuroimage.2017.08.035.

Picelli, Simone, Omid R. Faridani, Åsa K. Björklund, Gösta Winberg, Sven Sagasser, and Rickard Sandberg. 2014. “Full-Length RNA-Seq from Single Cells Using Smart-Seq2.” *Nature Protocols* 9 (1): 171–81. https://doi.org/10.1038/nprot.2014.006.

Polioudakis, Damon, Luis De La Torre-Ubieta, Justin Langerman, Mark B Gerstein, Kathrin Plath, and Daniel H Geschwind Correspondence. 2019. “A Single-Cell Transcriptomic Atlas of Human Neocortical Development during Mid-Gestation In Brief.” *Neuron* 103: 785-801.e8. https://doi.org/10.1016/j.neuron.2019.06.011.

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 (4): 743-756.e17. https://doi.org/10.1016/j.cell.2019.01.017.

Pollen, Alex A., Tomasz J. Nowakowski, Jiadong Chen, Hanna Retallack, Carmen Sandoval-Espinosa, Cory R. Nicholas, Joe Shuga, et al. 2015. “Molecular Identity of Human Outer Radial Glia during Cortical Development.” *Cell* 163 (1): 55–67. https://doi.org/10.1016/j.cell.2015.09.004.

Pollen, Alex A., Tomasz J. Nowakowski, Joe Shuga, Xiaohui Wang, Anne A. Leyrat, Jan H. Lui, Nianzhen Li, et al. 2014. “Low-Coverage Single-Cell MRNA Sequencing Reveals Cellular Heterogeneity and Activated Signaling Pathways in Developing Cerebral Cortex.” *Nature Biotechnology* 32 (10): 1053–58. https://doi.org/10.1038/nbt.2967.

Potente, Michael, and Taija Mäkinen. 2017. “Vascular Heterogeneity and Specialization in Development and Disease.” *Nature Reviews Molecular Cell Biology*. Nature Publishing Group. https://doi.org/10.1038/nrm.2017.36.

Puelles, L. 2009. “Forebrain Development: Prosomere Model.” In *Encyclopedia of Neuroscience*, 315–19. https://doi.org/10.1016/B978-008045046-9.01076-7.

Puelles, Luis, Megan Harrison, George Paxinos, and Charles Watson. 2013. “A Developmental Ontology for the Mammalian Brain Based on the Prosomeric Model.” *Trends in Neurosciences*. Elsevier Current Trends. https://doi.org/10.1016/j.tins.2013.06.004.

Puelles, Luis, and John L.R. Rubenstein. 1993. “Expression Patterns of Homeobox and Other Putative Regulatory Genes in the Embryonic Mouse Forebrain Suggest a Neuromeric Organization.” *Trends in Neurosciences*. Elsevier Current Trends. https://doi.org/10.1016/0166-2236(93)90080-6.

Purves, Dale, George J Augustine, David Fitzpatrick, Lawrence C Katz, Anthony-Samuel LaMantia, James O McNamara, and S Mark Williams. 2001. “Neuroglial Cells.” https://www.ncbi.nlm.nih.gov/books/NBK10869/.

Radakovits, Randor, Claudia S. Barros, Richard Belvindrah, Bruce Patton, and Ulrich Müller. 2009. “Regulation of Radial Glial Survival by Signals from the Meninges.” *Journal of Neuroscience* 29 (24): 7694–7705. https://doi.org/10.1523/JNEUROSCI.5537-08.2009.

Raj, Arjun, and Alexander van Oudenaarden. 2008. “Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences.” *Cell*. Elsevier. https://doi.org/10.1016/j.cell.2008.09.050.

Raj, Bushra, Daniel E. Wagner, Aaron McKenna, Shristi Pandey, Allon M. Klein, Jay Shendure, James A. Gagnon, and Alexander F. Schier. 2018. “Simultaneous Single-Cell Profiling of Lineages and Cell Types in the Vertebrate Brain.” *Nature Biotechnology* 36 (5): 442–50. https://doi.org/10.1038/nbt.4103.

Raper, J. A. 2009. “Growth Cones.” In *Encyclopedia of Neuroscience*, 981–85. Elsevier Ltd. https://doi.org/10.1016/B978-008045046-9.00340-5.

Rayon, Teresa, Rory J. Maizels, Christopher Barrington, and James Briscoe. 2021. “Single-Cell Transcriptome Profiling of the Human Developing Spinal Cord Reveals a Conserved Genetic Programme with Human-Specific Features.” *Development (Cambridge)* 148 (15). https://doi.org/10.1242/DEV.199711.

Reinchisi, Gisela, Kumiko Ijichi, Nicole Glidden, Igor Jakovcevski, and Nada Zecevic. 2012. “COUP-TFII Expressing Interneurons in Human Fetal Forebrain.” *Cerebral Cortex* 22 (12): 2820–30. https://doi.org/10.1093/cercor/bhr359.

Riba, Andrea, Attila Oravecz, Matej Durik, Sara Jiménez, Violaine Alunni, Marie Cerciat, Matthieu Jung, Céline Keime, William M. Keyes, and Nacho Molina. 2022. “Cell Cycle Gene Regulation Dynamics Revealed by RNA Velocity and Deep-Learning.” *Nature Communications* 13 (1): 1–13. https://doi.org/10.1038/s41467-022-30545-8.

Rifes, Pedro, Marc Isaksson, Gaurav Singh Rathore, Patrick Aldrin-Kirk, Oliver Knights Møller, Guido Barzaghi, Julie Lee, et al. 2020. “Modeling Neural Tube Development by Differentiation of Human Embryonic Stem Cells in a Microfluidic WNT Gradient.” *Nature Biotechnology* 38 (11): 1265–73. https://doi.org/10.1038/s41587-020-0525-0.

Rogers, Katherine W., and Alexander F. Schier. 2011. “Morphogen Gradients: From Generation to Interpretation.” *Annual Review of Cell and Developmental Biology* 27 (October): 377–407. https://doi.org/10.1146/annurev-cellbio-092910-154148.

Ross, Jayden M., Chang Kim, Denise Allen, Elizabeth E. Crouch, Kazim Narsinh, Daniel L. Cooke, Adib A. Abla, Tomasz J. Nowakowski, and Ethan A. Winkler. 2020. “The Expanding Cell Diversity of the Brain Vasculature.” *Frontiers in Physiology*. Frontiers Media S.A. https://doi.org/10.3389/fphys.2020.600767.

Rowitch, David H., and Arnold R. Kriegstein. 2010. “Developmental Genetics of Vertebrate Glial-Cell Specification.” *Nature*. Nature Publishing Group. https://doi.org/10.1038/nature09611.

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*. NIH Public Access. https://doi.org/10.1111/j.1469-7610.2010.02307.x.

Rubenstein, John L.R., Salvador Martinez, Kenji Shimamura, and Luis Puelles Nina. 1994. “The Embryonic Vertebrate Forebrain: The Prosomeric Model.” *Science*. https://doi.org/10.1126/science.7939711.

Saelens, Wouter, Robrecht Cannoodt, Helena Todorov, and Yvan Saeys. 2019. “A Comparison of Single-Cell Trajectory Inference Methods.” *Nature Biotechnology* 37 (5): 547–54. https://doi.org/10.1038/s41587-019-0071-9.

Santis, Riccardo De, Fred Etoc, Edwin A. Rosado-Olivieri, and Ali H. Brivanlou. 2021. “Self-Organization of Human Dorsal-Ventral Forebrain Structures by Light Induced SHH.” *Nature Communications* 12 (1): 1–11. https://doi.org/10.1038/s41467-021-26881-w.

Schleiden, M. J., Theodor Schwann, Henry Smith, and M. J. Schleiden. 1847. *Microscopical Researches into the Accordance in the Structure and Growth of Animals and Plants. Tr. from the German of Dr. Th. Schwann ... by Henry Smith ...* London,: The Sydenham Society,. https://doi.org/10.5962/bhl.title.11431.

Siletti, Kimberly, Rebecca Hodge, Alejandro Mossi Albiach, Lijuan Hu, Ka Wai Lee, Peter Lönnerberg, Trygve Bakken, et al. 2022. “Transcriptomic Diversity of Cell Types across the Adult Human Brain.” *BioRxiv*, October, 2022.10.12.511898. https://doi.org/10.1101/2022.10.12.511898.

Solini, Grace E, Chen Dong, and Margaret Saha. 2017. “Embryonic Transplantation Experiments: Past, Present, and Future.” *Trends in Developmental Biology* 10: 13–30. http://www.ncbi.nlm.nih.gov/pubmed/31631952.

Spanjaard, Bastiaan, Bo Hu, Nina Mitic, Pedro Olivares-Chauvet, Sharan Janjuha, Nikolay Ninov, and Jan Philipp Junker. 2018. “Simultaneous Lineage Tracing and Cell-Type Identification Using CrIsPr-Cas9-Induced Genetic Scars.” *Nature Biotechnology* 36 (5): 469–73. https://doi.org/10.1038/nbt.4124.

Spemann, Hans, and Hilde Mangold. 1924. “Facsimile Reproduction of the Cover of an Original Reprint of the 1924 Article by Hans Spemann and Hilde Mangold, with a Handwritten Dedication by H. Spemann Which Reads ‘With Best Regards, H.S.’ (Courtesy of K. Sander, Freiburg).” *Archiv*.

Spitzer, Sonia Olivia, Sergey Sitnikov, Yasmine Kamen, Kimberley Anne Evans, Deborah Kronenberg-Versteeg, Sabine Dietmann, Omar de Faria, Sylvia Agathou, and Ragnhildur Thóra Káradóttir. 2019. “Oligodendrocyte Progenitor Cells Become Regionally Diverse and Heterogeneous with Age.” *Neuron* 101 (3): 459-471.e5. https://doi.org/10.1016/j.neuron.2018.12.020.

Sridhar, Akshayalakshmi, Akina Hoshino, Connor R. Finkbeiner, Alex Chitsazan, Li Dai, Alexandra K. Haugan, Kayla M. Eschenbacher, et al. 2020. "Single-Cell Transcriptomic Comparison of Human Fetal Retina, HPSC-Derived Retinal Organoids, and Long-Term Retinal Cultures." *Cell Reports* 30 (5): 1644-1659.e4. https://doi.org/10.1016/j.celrep.2020.01.007.

Studer, Lorenz. 2012. "Derivation of Dopaminergic Neurons from Pluripotent Stem Cells." In *Progress in Brain Research*, 200:243–63. Elsevier B.V. https://doi.org/10.1016/B978-0-444-59575-1.00011-9.

Subramanian, Lakshmi, Maria Elisa Calcagnotto, and Mercedes F. Paredes. 2020. "Cortical Malformations: Lessons in Human Brain Development." *Frontiers in Cellular Neuroscience*. Frontiers Media S.A. https://doi.org/10.3389/fncel.2019.00576.

Südhof, Thomas C. 2017. "Molecular Neuroscience in the 21st Century: A Personal Perspective." *Neuron*. Cell Press. https://doi.org/10.1016/j.neuron.2017.10.005.

Sugihara, Reiichi, Yuki Kato, Tomoya Mori, and Yukio Kawahara. 2022. "Alignment of Single-Cell Trajectory Trees with CAPITAL." *Nature Communications* 13 (1): 1–11. https://doi.org/10.1038/s41467-022-33681-3.

Sulston, J. E., E. Schierenberg, J. G. White, and J. N. Thomson. 1983. "The Embryonic Cell Lineage of the Nematode Caenorhabditis Elegans." *Developmental Biology*. Academic Press. https://doi.org/10.1016/0012-1606(83)90201-4.

Sun, Yu, Susan K Goderie, and Sally Temple. 2005. "Asymmetric Distribution of EGFR Receptor during Mitosis Generates Diverse CNS Progenitor Cells." *Neuron* 45: 873–86. https://doi.org/10.1016/j.neuron.2005.01.045.

Svensson, Valentine. 2020. "Droplet ScRNA-Seq Is Not Zero-Inflated." *Nature Biotechnology*. Nature Research. https://doi.org/10.1038/s41587-019-0379-5.

Svensson, Valentine, Eduardo da Veiga Beltrame, and Lior Pachter. 2019. "Quantifying the Tradeoff between Sequencing Depth and Cell Number in Single-Cell RNA-Seq." *BioRxiv*. bioRxiv. https://doi.org/10.1101/762773.

Svensson, Valentine, Roser Vento-Tormo, and Sarah A. Teichmann. 2018a. "Exponential Scaling of Single-Cell RNA-Seq in the Past Decade." *Nature Protocols*. Nature Publishing Group. https://doi.org/10.1038/nprot.2017.149.

Svoboda, Marek, H Robert Frost, and Giovanni Bosco. 2022. "Internal Oligo(DT) Priming Introduces Systematic Bias in Bulk and Single-Cell RNA Sequencing Count Data." *NAR Genomics and Bioinformatics* 4 (2). https://doi.org/10.1093/nargab/lqac035.

Tanay, Amos, and Aviv Regev. 2017. "Scaling Single-Cell Genomics from Phenomenology to Mechanism." *Nature*. Nature Publishing Group. https://doi.org/10.1038/nature21350.

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. https://doi.org/10.1038/nmeth.1315.

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. https://doi.org/10.1038/s41586-018-0654-5.

Tata, Mathew, Christiana Ruhrberg, and Alessandro Fantin. 2015. "Vascularisation of the Central Nervous System." *Mechanisms of Development*. Elsevier Ireland Ltd. https://doi.org/10.1016/j.mod.2015.07.001.

Tedesco, Martina, Francesca Giannese, Dejan Lazarević, Valentina Giansanti, Dalia Rosano, Silvia Monzani, Irene Catalano, et al. 2022. "Chromatin Velocity Reveals Epigenetic Dynamics by Single-Cell Profiling of Heterochromatin and Euchromatin." *Nature Biotechnology* 40 (2): 235–44. https://doi.org/10.1038/s41587-021-01031-1.

Temple, Sally. 1989. "Division and Differentiation of Isolated CNS Blast Cells in Microculture." *Nature* 340 (6233): 471–73. https://doi.org/10.1038/340471a0.

Thompson, Carol L., Lydia Ng, Vilas Menon, Salvador Martinez, Chang Kyu Lee, Katie Glattfelder, Susan M. Sunkin, et al. 2014. "A High-Resolution Spatiotemporal Atlas of Gene Expression of the Developing Mouse Brain." *Neuron* 83 (2): 309–23. https://doi.org/10.1016/j.neuron.2014.05.033.

Tkachenko, Lyubov A., Pavel A. Zykin, Ruslan A. Nasyrov, and Elena I. Krasnoshchekova. 2016. "Distinctive Features of the Human Marginal Zone and Cajal–Retzius Cells: Comparison of Morphological and Immunocytochemical Features at Midgestation." *Frontiers in Neuroanatomy* 10 (MAR). https://doi.org/10.3389/fnana.2016.00026.

Tole, S., and J. Hébert. 2013. "Telencephalon Patterning." *Patterning and Cell Type Specification in the Developing CNS and PNS* 1: 3–24. https://doi.org/10.1016/B978-0-12-397265-1.00018-6.

Traag, V. A., L. Waltman, and N. J. van Eck. 2019. "From Louvain to Leiden: Guaranteeing Well-Connected Communities." *Scientific Reports* 9 (1): 1–12. https://doi.org/10.1038/s41598-019-41695-z.

Tran, Hoa Thi Nhu, Kok Siong Ang, Marion Chevrier, Xiaomeng Zhang, Nicole Yee Shin Lee, Michelle Goh, and Jinmiao Chen. 2020. "A Benchmark of Batch-Effect Correction Methods for Single-Cell RNA Sequencing Data." *Genome Biology* 21 (1): 12. https://doi.org/10.1186/s13059-019-1850-9.

Trevino, Alexandro E., Fabian Müller, Jimena Andersen, Laksshman Sundaram, Arwa Kathiria, Anna Shcherbina, Kyle Farh, et al. 2021. "Chromatin and Gene-Regulatory Dynamics of the Developing Human Cerebral Cortex at Single-Cell Resolution." *Cell* 184 (19): 5053-5069.e23. https://doi.org/10.1016/j.cell.2021.07.039.

Trewavas, Anthony. 2006. "A Brief History of Systems Biology." *Plant Cell*. American Society of Plant Biologists. https://doi.org/10.1105/tpc.106.042267.

Tripathi, Richa B., Laura E. Clarke, Valeria Burzomato, Nicoletta Kessaris, Patrick N. Anderson, David Attwell, and William D. Richardson. 2011. "Dorsally and Ventrally Derived Oligodendrocytes Have Similar Electrical Properties but Myelinate Preferred Tracts." *Journal of Neuroscience* 31 (18): 6809–19. https://doi.org/10.1523/JNEUROSCI.6474-10.2011.

Tsien, Joe Z., Dong Feng Chen, David Gerber, Cindy Tom, Eric H. Mercer, David J. Anderson, Mark Mayford, Eric R. Kandel, and Susumu Tonegawa. 1996. "Subregion- and Cell Type-Restricted Gene Knockout in Mouse Brain." *Cell* 87 (7): 1317–26. https://doi.org/10.1016/S0092-8674(00)81826-7.

Tümpel, Stefan, Leanne M. Wiedemann, and Robb Krumlauf. 2009. "Chapter 8 Hox Genes and Segmentation of the Vertebrate Hindbrain." *Current Topics in Developmental Biology*. Academic Press. https://doi.org/10.1016/S0070-2153(09)88004-6.

Tyser, Richard C.V., Elmir Mahammadov, Shota Nakanoh, Ludovic Vallier, Antonio Scialdone, and Shankar Srinivas. 2021. "Single-Cell Transcriptomic Characterization of a Gastrulating Human Embryo." *Nature* 600 (7888): 285–89. https://doi.org/10.1038/s41586-021-04158-y.

Uzquiano, Ana, and Paola Arlotta. 2022. "Brain Organoids: The Quest to Decipher Human-Specific Features of Brain Development." *Current Opinion in Genetics and Development*. Elsevier Ltd. https://doi.org/10.1016/j.gde.2022.101955.

Vernes, Sonja C., Peter L. Oliver, Elizabeth Spiteri, Helen E. Lockstone, Rathi Puliyadi, Jennifer M. Taylor, Joses Ho, et al. 2011. "Foxp2 Regulates Gene Networks Implicated in Neurite Outgrowth in the Developing Brain." Edited by Joshua M. Akey. *PLoS Genetics* 7 (7): e1002145. https://doi.org/10.1371/journal.pgen.1002145.

Vinsland, Elin, and Sten Linnarsson. 2022. "Single-Cell RNA-Sequencing of Mammalian Brain Development: Insights and Future Directions." *Development (Cambridge, England)* 149 (10). https://doi.org/10.1242/dev.200180.

Waddington, C. H. 1939. "Preliminary Notes on the Development of the Wings in Normal and Mutant Strains of Drosophila." *Proceedings of the National Academy of Sciences* 25 (7): 299–307. https://doi.org/10.1073/pnas.25.7.299.

Wälchli, Thomas, Moheb Ghobrial, Marc Schwab, Shigeki Takada, Hang Zhong, Samuel Suntharalingham, Sandra Vetiska, et al. 2021. "Gelareh Zadeh 2,24 , Viviane Tabar 25." *Ralph Schlapbach* 16: 32. https://doi.org/10.1101/2021.10.18.464715.

Westergard, Thomas, and Jeffrey D. Rothstein. 2020. "Astrocyte Diversity: Current Insights and Future Directions." *Neurochemical Research* 45 (6): 1298–1305. https://doi.org/10.1007/s11064-020-02959-7.

Wilson, Leigh, and Malcolm Maden. 2005. "The Mechanisms of Dorsoventral Patterning in the Vertebrate Neural Tube." *Developmental Biology*. Academic Press Inc. https://doi.org/10.1016/j.ydbio.2005.02.027.

Winkler, Caitlin C., Odessa R. Yabut, Santiago P. Fregoso, Hector G. Gomez, Brett E. Dwyer, Samuel J. Pleasure, and Santos J. Franco. 2018. "The Dorsal Wave of Neocortical Oligodendrogenesis Begins Embryonically and Requires Multiple Sources of Sonic Hedgehog." *Journal of Neuroscience* 38 (23): 5237–50. https://doi.org/10.1523/JNEUROSCI.3392-17.2018.

Xi, Nan Miles, and Jingyi Jessica Li. 2021. "Benchmarking Computational Doublet-Detection Methods for Single-Cell RNA Sequencing Data." *Cell Systems* 12 (2): 176-194.e6. https://doi.org/10.1016/j.cels.2020.11.008.

Xu, Yichi, Tengjiao Zhang, Qin Zhou, Mengzhu Hu, Yao Qi, Yifang Xue, Lihui Wang, Zhirong Bao, and Weiyang Shi. 2021. "A Single-Cell Transcriptome Atlas of Human Early Embryogenesis China." *Regenerative Medicine*. https://doi.org/10.1101/2021.11.30.470583.

Yasuda, Mineo. 2018. “The Beginning of Kyoto Collection: Focus on Memories.” *The Anatomical Record* 301 (6): 951–54. https://doi.org/10.1002/ar.23788.

Yu, Kevin Shengyang, Stacey M. Frumm, Jason S. Park, Katharine Lee, Daniel M. Wong, Lauren Byrnes, Sarah M. Knox, Julie B. Sneddon, and Aaron D. Tward. 2019. “Development of the Mouse and Human Cochlea at Single Cell Resolution.” *BioRxiv*.

Zeisel, Amit, Hannah Hochgerner, Peter Lönnerberg, Anna Johnsson, Fatima Memic, Job van der Zwan, Martin Häring, et al. 2018. “Molecular Architecture of the Mouse Nervous System.” *Cell* 174 (4): 999-1014.e22. https://doi.org/10.1016/j.cell.2018.06.021.

Zeisel, Amit, Wolfgang J. Köstler, Natali Molotski, Jonathan M. Tsai, Rita Krauthgamer, Jasmine Jacob-Hirsch, Gideon Rechavi, et al. 2011. “Coupled Pre-MRNA and MRNA Dynamics Unveil Operational Strategies Underlying Transcriptional Responses to Stimuli.” *Molecular Systems Biology*. https://doi.org/10.1038/msb.2011.62.

Zeng, Hongkui. 2022. “What Is a Cell Type and How to Define It?” *Cell*. Elsevier B.V. https://doi.org/10.1016/j.cell.2022.06.031.

Zhang, Martin Jinye, Vasilis Ntranos, and David Tse. 2020. “Determining Sequencing Depth in a Single-Cell RNA-Seq Experiment.” *Nature Communications* 11 (1): 1–11. https://doi.org/10.1038/s41467-020-14482-y.

Zhang, Shu, Yueli Cui, Xinyi Ma, Jun Yong, Liying Yan, Ming Yang, Jie Ren, Fuchou Tang, Lu Wen, and Jie Qiao. 2020. “Single-Cell Transcriptomics Identifies Divergent Developmental Lineage Trajectories during Human Pituitary Development.” *Nature Communications* 11 (1): 1–16. https://doi.org/10.1038/s41467-020-19012-4.

Zhong, Suijuan, Shu Zhang, Xiaoying Fan, Qian Wu, Liying Yan, Ji Dong, Haofeng Zhang, et al. 2018. “A Single-Cell RNA-Seq Survey of the Developmental Landscape of the Human Prefrontal Cortex.” *Nature* 555 (7697): 524–28. https://doi.org/10.1038/nature25980.

Zhu, Chenxu, Sebastian Preissl, and Bing Ren. 2020. “Single-Cell Multimodal Omics: The Power of Many.” *Nature Methods*. Nature Research. https://doi.org/10.1038/s41592-019-0691-5.

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.

Ziegenhain, Christoph, Beate Vieth, Swati Parekh, Björn Reinius, Amy Guillaumet-Adkins, Martha Smets, Heinrich Leonhardt, Holger Heyn, Ines Hellmann, and Wolfgang Enard. 2017. “Comparative Analysis of Single-Cell RNA Sequencing Methods.” *Molecular Cell* 65 (4): 631-643.e4. https://doi.org/10.1016/j.molcel.2017.01.023.

Ziffra, Ryan S., Chang N. Kim, Jayden M. Ross, Amy Wilfert, Tychele N. Turner, Maximilian Haeussler, Alex M. Casella, et al. 2021. “Single-Cell Epigenomics Reveals Mechanisms of Human Cortical Development.” *Nature* 598 (7879): 205–13. https://doi.org/10.1038/s41586-021-03209-8.