Epigenomic Programming in Early Fetal Brain Development

by

Luolan Li

B.Sc., Nanjing University, 2012

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Genome Science and Technology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

November 2022

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Epigenomic programming in early fetal brain development

submitted by Luolan Li in partial fulfillment of the requirements for

the degree of Master of Science

in Genome Science and Technology

Examining Committee:

Martin Hirst, Professor, Microbiology and Immunology, UBC
Supervisor

Matt Lorincz, Professor, Medical Genetics, UBC
Supervisory Committee Member

Stephen Yip, Associate Professor, Pathology and Laboratory Medicine, UBC
Supervisory Committee Member

Additional Supervisory Committee Members:

Paul Pavlidis, Professor, Psychiatry, UBC
Supervisory Committee Member

Abstract

A comprehensive understanding of gene regulatory networks in the developing human brain provides a foundation for interpreting pathogenic deregulation. Here we analyzed the complete epigenomes and transcriptomes of dissected brain regions and primary neural progenitor cells (NPCs) derived from cortex and ganglionic eminence (GE) of four human fetuses including a pair of monozygotic twins. Epigenetic regulatory states were compared between NPCs derived from cortex and GE, across developmental stages, and between monozygotic twins. Comparisons across developmental stages reveal an increase in active epigenetic states, transcription factor activities and gene transcription with increasing developmental stage. NPCs derived from different brain regions retained brain region and gestational week specific regulatory states. We also found evidence of divergent epigenetic signatures between monozygotic twins before midgestation.

Lay Summary

During cell differentiation, the DNA sequence that is largely the same throughout all the cells in body stays the same but chemical modifications to your DNA occur that instruct the cell to turn genes on or off and thus determine cell fate. These chemical changes to DNA and the special proteins that package it collectively form the epigenome. The early fetal period is a critical time for human brain development. To study the epigenetic changes during this period, we profiled the epigenomes of neural progenitor cells of different human brain regions at different developmental stages. We found a more active epigenetic state in later developmental stages. We also showed that the epigenome of monozygotic twins were different as early as midgestation. Finally, we provide an epigenetic roadmap for specific brain regions in early development providing a comprehensive reference for future studies in brain development and disease.

Preface

I was responsible for major areas of concept, data analysis of the sequencing data, and manuscript composition. Martin Hirst initially designed the project with Joseph F. Costello and Keith L. Ligon, and oversaw the progression of the project at all stages. Cecile L. Maire, Chibo Hong and Raman P. Nagarajan dissected brain tissues from human samples, prepared the neurospheres in culture and extracted their nucleic acids. Angela Tam, Baljit Kamoh and Stephanie Cho prepared the sequencing libraries. Marco A Marra, Steven JM Jones, Andrew J Mungall, Richard Moore and Martin Hirst oversaw the generation of sequencing data. Misha Bilenky, Annaick Carles, Alireza Heravi-Moussavi, Dorothy Cheung, Irene Li, Tina Wong and Ting Wang were involved in the initial stage of data processing including aligning sequencing data, generating quality control reports and generating DNA methylation levels. Claudia L. Kleinman and Nada Jabado provided additional datasets for validation. Use of the brain tissues to conduct this study was approved by UBC committee under ethics certificate H12-01767.

A version of this manuscript is published at:

Li L, Maire CL, Bilenky M, Carles A, Heravi-Moussavi A, Hong C, Tam A, Kamoh B, Cho S, Cheung D, Li I, Wong T, Nagarajan RP, Mungall AJ, Moore R, Wang T, Kleinman CL, Jabado N, Jones SJ, Marra MA, Ligon KL, Costello JF, Hirst M. Epigenomic programming in early fetal brain development. *Epigenomics*. 2020 Jun;12(12):1053-1070. doi: 10.2217/epi-2019-0319. Epub 2020 Jul 17. PMID: 32677466; PMCID: PMC7857341.

A part of Chapter 1 is adapted from:

Li L, Lorzadeh A, Hirst M. Regulatory variation: an emerging vantage point for cancer biology. *Wiley Interdiscip Rev Syst Biol Med*. 2014 Jan-Feb;6(1):37-59. doi: 10.1002/wsbm.1250. Epub 2013 Nov 19. PMID: 24254976.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

2HG – 2-hydroxyglutarate

5hmC – 5-hydroxymethylcytosine

5mC – 5-methylcytosine

AR – androgen receptor

BCL6 – B-Cell CLL/Lymphoma 6

bp – base pair

CGI – CpG island

CHD1 – Chromodomain Helicase DNA Binding Protein 1

ChIP-seq – Chromatin immunoprecipitation sequencing

CP – cortical plate

DE – differential expression

DMR – differential methylated regions

DNA – deoxyribonucleic acid

DNMT– DNA methyltransferase

EPHA3 – EPH Receptor A3

FEZF1 – FEZ family Zinc Finger 1

FOX – Forkhead Box

GE – ganglionic eminences

GFAP – glial fibrillary acidic protein

GW – gestational week

H3K27ac – histone H3 lysine 27 acetylation

H3K27me3 – histone H3 lysine 27 tri-methylation

H3K36me3 – histone H3 lysine 36 tri-methylation

H3K4me1 – histone H3 lysine 4 mono-methylation

H3K4me3 – histone H3 lysine 4 tri-methylation

H3K9me2 – histone H3 lysine 9 di-methylation

H3K9me3 – histone H3 lysine 9 tri-methylation

HAT – histone acetyl transferases

HDAC – histone deacetylases

HIPAA – Health Insurance Portability and Accountability Act

hiPSC – human induced pluripotent stem cell

HP1 – Heterochromatin protein 1

IHEC – International Human Epigenome Consortium

JMJD2A – Jumonji domain 2

KMT – lysine methyltransferases

LHX – LIM Homeobox

LTR – long terminal repeats

MBD – methyl-CpG-binding domain

MeCP2 – methyl CpG binding protein 2

MeDIP– methylated DNA immunoprecipitation

miRNA – microRNA

MRE – methylation sensitive restriction enzyme

MZ – monozygotic

NFIX – nuclear factor I/X

NPC – neural progenitor cells

NTN – netrin

OLIG – oligodendrocyte lineage transcription factor

OTX – Orthodenticle Homeobox proteins

PCR – polymerase chain reaction

QC – quality control

RNA – ribonucleic acid

RPKM – reads per kilobase of transcript per million reads mapped

SAM – S-adenosylmethionine

SFRP – Secreted Frizzled-Related Protein

SLIT2 – Slit Homolog 2

SNP – single nucleotide polymorphism

SNV – single nucleotide variant

SOP – standard operating procedures

SOX3 – SRY-Box Transcription Factor 3

TET – ten-eleven-translocation

TF – transcription factors

TFBS – transcription factor binding site

TSS – transcription start site

UMR – unmethylated (hypomethylated) regions

VAX1 – Ventral Anterior Homeobox 1

VZ – ventricular zone

WGBS – whole-genome bisulphite sequencing

WNT – Wingless-type MMTV Integration Site

Acknowledgements

I offer my gratitude to the faculty, staff, and peers at the University of British Columbia who have facilitated intellectual discourse and inspired my work. I extend my thanks to Dr. Matt Lorincz, Dr. Stephen Yip and Dr. Paul Pavlidis for being on my thesis advisory committee and enriching my knowledge within the realms of epigenetics and bioinformatics.

I thank Dr. Martin Hirst for his continuous support and direction during difficult times. I thank Michelle Moksa for her support as our lab manager. I thank Dr. Misha Bilenky, Annaick Carles and Dr. Alireza Lorzadeh for their expertise in bioinformatic analysis, and for active engagement in discussions related to my work. I thank the other members of my lab for providing an encouraging space for scientific inquiry and intellectual pursuit. I would also like to thank the production and technical staff at Canada's Michael Smith Genome Sciences Centre for the generation of the reference epigenomes.

I gratefully acknowledge the funding provided by the following grants: US National Institutes of Health (NIH) Roadmap Epigenomics Program (NIH grant 5U01ES017154-02), Genome British Columbia and the Canadian Institutes of Health Research as part of the Canadian Epigenetics, Environment and Health Research Consortium Network (CIHR-262119) and Genome Science and Technology Graduate Program Fellowship, University of British Columbia.

Dedication

To my family

Chapter 1: Introduction

1.1 Epigenetics

In the 1940s Conrad Waddington first used the term "epigenetic landscape" to describe *Drosophila Melanogaster* embryo development as "a ball running down an uneven hill with saddle points representing differentiation decision points" (Waddington, 1942). In broad terms, epigenetics describes anything that influences mitotically heritable phenotypic changes without alterations in the DNA sequence itself (Cavalli and Heard, 2019). The genome of an organism carries chemical marks covalently bonded to DNA bases or to the chromatin proteins that alter local chromatin structure and regulate gene expression, thus leading to phenotypical changes (Boyes and Bird, 1991a; Local et al., 2018; Skourti and Dhillon, 2022; Tompkins et al., 2012). These chemical marks, such as modifications of DNA or amino acids on histone tails, collectively form the epigenome (**Figure 1**). With the exception of specialized immune cells, the genome of an organism remains the same throughout the normal process of development while its epigenome undergoes significant changes during differentiation (Adithya et al., 2022; Moltrasio et al., 2022; Tompkins et al., 2012; Wu et al., 2022).

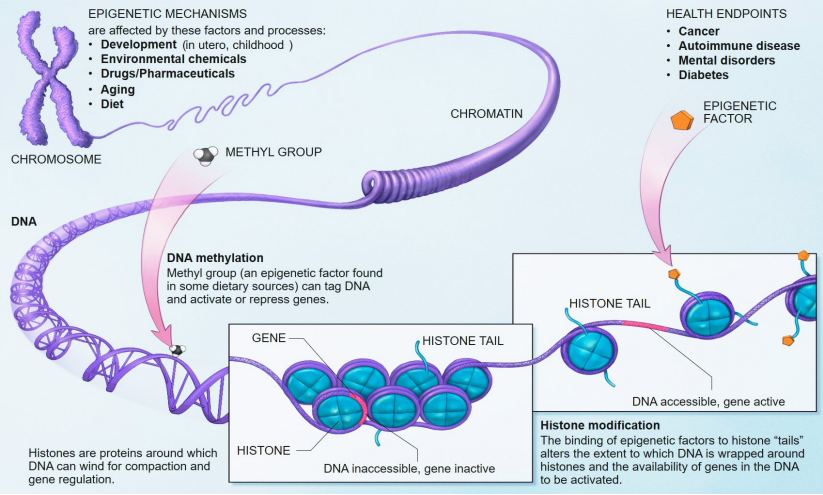

Figure 1. Human epigenome

DNA wraps around histone complex and package into the complicated structure of chromatin. Epigenetic factors, such as DNA methylation and histone modifications, are bound to DNA or the histone tails and regulate gene expression. From Laura, B. (2008) Epigenomics: The new tool in studying complex diseases. *Nature Education* 1(1):178. Used with permission from *Nature Education.*

1.2 DNA methylation

DNA methylation is one of the most abundant and extensively studied chromatin modification. In somatic mammalian cells, 70-80% of cytosines are methylated (5mC) when found in the context of CpG dinucleotides (Ehrlich et al., 1982), and to a lesser extent in non-CpG contexts in specific cell types (Ziller et al., 2011). CpGs are found in various densities along the genome, including CpG islands (CGIs), defined as CpG-rich (>50%) regions in 1kb or greater length, that are largely unmethylated when located in gene promoters (Bird et al., 1985; Ziller et al., 2013). DNA methylation is not essential for cell viability but is essential for cellular differentiation (Chattopadhyaya and Ghosal, 2022; Tsumura et al., 2006) and plays important roles in the regulation of gene transcription, maintenance of genome stability, repression of endogenous retroviral transcripts and X chromosome inactivation (Al Adhami et al., 2022; He et al., 2022; Li et al., 2022; Robertson, 2005).

DNA methylation is catalyzed by the DNA methyltransferase (*DNMT*) family (**Figure 2**). There are four members in this family: *DNMT1*, responsible for the maintenance of DNA methylation during replication, *DNMT3A* and *DNMT3B*, responsible for *de novo* DNA methylation, and *DNMT3L*, a cofactor of *DNMT3A/B* with non-catalytic functions (Klose and Bird, 2006). The ten-eleven-translocation family (*TET1-3*) catalyze the conversion of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) in an α-ketoglutarate-dependent manner (Tahiliani et al., 2009); 5hmC is thought to be an intermediate in the pathway that actively (Wu and Zhang,

2011) or passively (Bhutani et al., 2011) demethylates 5mC (**Figure 2**). Regulated demethylation through 5hmC offers an important mechanism for dynamic reprogramming of the human methylome and thus transcriptional regulation (Charlton et al., 2020; Ginno et al., 2020).

DNA methylation is regulated and maintained by an interactive network of effector proteins. Alterations to components of the DNA methylation machinery have been identified and shown to compromise network balance, global DNA methylation, and transcriptional output. One outcome of this imbalance is the activation of long terminal repeats (LTRs) by intergenic hypomethylation (Gimenez et al., 2010). Transcriptionally active transposable elements can serve as enhancers or promoters to drive the expression of nearby oncogenes (Howard et al., 2008; Lamprecht et al., 2010; Sin et al., 2006a; Sin et al., 2006b) and through this process a pathogenic transcriptional network can emerge. These examples highlight the importance of studying the DNA methylation modifiers in a network context.

DNA methylation machinery also interacts directly with other regulatory components such as transcription factors (TFs). This relationship can be either in the context of the methylation status affecting TF binding or TFs mediating the DNA methylation/demethylation process. Methylated cytosines can alter the structure and accessibility of chromatin and thus affect TF binding depending on the affinity of TFs to 5mC (Mattei et al., 2022; Wiench et al., 2011; Yin et al., 2017). CpG methylation within gene promoter regions is linked with repression of transcription initiation through the inhibition of the binding of transcriptional activators (Peng et al., 2022; Watt and Molloy, 1988) or through the recruitment of other repressive regulatory complexes (Boyes and Bird, 1991b; Nan et al., 1998). Emerging studies suggest a second scenario where in distal regulatory regions such as enhancers, the binding of certain tissue specific DNA-binding factors is both necessary and sufficient for the hypomethylation state of the regions, indicating that TF

binding can mediate DNA methylation states by depleting *DNMT*s from the region (Detilleux et al., 2022; Kremsky and Corces, 2020; Miyajima et al., 2022; Stadler et al., 2011; Thurman et al., 2012). In addition, TFs are also believed to be involved in the establishment of localized hypermethylated states during development and in *DNMT* mutant cancers by recruiting Dnmts to the promoter regions of specific genes (Gu et al., 2011; Hervouet et al., 2009).

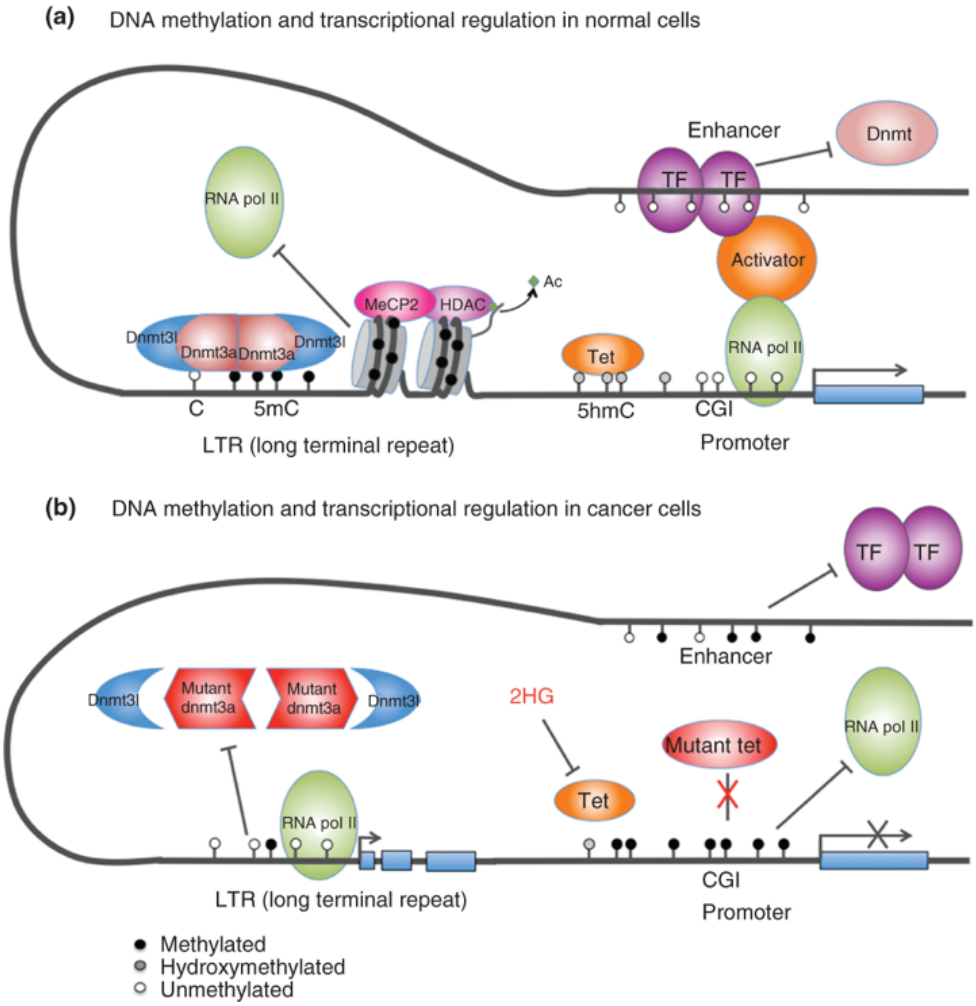

Figure 2. DNA methylation and transcriptional regulation

a) DNA methylation and transcriptional regulation in normal cells. Dnmt3a and Dnmt3l form a tetramer to catalyze DNA methylation. MeCP2, member of MBD family, binds to methylated cytosine and recruits histone modifiers and chromatin remodelers to the site resulting in a more compact chromatin structure. These two mechanisms together keep the long terminal repeats (LTRs) silenced in the genome. Tet proteins catalyze 5mC (black circles) into 5hmC

(gray circles) and lead to loss of methylation state in CGIs of promoters, thus activate transcription. TF binding at enhancer sites depletes Dnmts resulting in hypomethylation of the region. It may also recruit other activators that interact with the promoters and regulate transcription initiation. **b**) DNA methylation and transcriptional regulation in cancer cells. Mutations in DNMT3A interrupt the tetramerization of Dnmt3a-Dnmt3l and cause hypomethylation and aberrant activation of LTRs. Loss-of-function mutations or inhibition by oncometabolite 2HG of Tet will lead to loss of 5hmC and global hypermethylation, resulting in an altered methylation status and disrupted transcription initiation.

1.3 Histone modifications

Covalent chemical modification to histones is an additional form of epigenetic modification that influences transcriptional regulation and chromatin structure. Histone modifications and deregulation of the pathways which control them collude with other factors to drive human diseases including cancer (Li et al., 2013; Love et al., 2012; Maiques-Diaz et al., 2012; Schwartzentruber et al., 2012; Wang et al., 2013; Wei et al., 2013). These chemical modifications occur primarily on the N-terminal tail of histones and generate a code that reinforces active or repressive chromatin states (**Figure 3**). Examples of such modifications are histone acetylation (Jiang et al., 2012), methylation (Wang et al., 2001), phosphorylation (Lohse et al., 2013), ubiquitylation (Yuan et al.), sumoylation (Shiio and Eisenman, 2003), formylation (Jiang et al., 2007), and hydroxylation (He et al., 2018; Unoki et al., 2013). Among them, histone acetylation and methylation are the most extensively studied and commonly used marks associated with specific chromatin states. These modifications together regulate gene expression either by influencing the 3D structure and thus chromatin accessibility directly or by affecting the binding of other transcriptional regulators (Bannister and Kouzarides, 2011; Wang et al., 2008).

1.3.1 Histone acetylation

Acetylated histones are found associated with actively transcribed genomic regions and are thought to act, at least in part, through destabilization of chromatin-DNA interactions by

neutralizing the positive charge of lysine (K) residues (Jiang et al., 2012). Histone acetylation levels are determined through the concerted actions of both histone acetyl transferases (HAT) and histone deacetylases (HDAC). HATs transfer the acetyl group of acetyl-coenzyme A to a histone lysine residue weakening the interaction between DNA and histones leading to chromatin expansion and gene activation, while HDACs remove the acetyl group from the histone tail leading to chromatin contracting and compacting, and thus gene repression (Gansen et al., 2015; Kurdistani et al., 2004; Lombardi et al., 2011). Histone H3 acetylation is found associated with enhancer regions and active transcriptional start sites whereas acetylated H4 is found largely on promoter transcribed regions (Creyghton et al., 2010). In particular, H3K27ac marks active enhancers and can form broad domains called super enhancers that facilitate distinct gene expression patterns during differentiation (Hnisz et al., 2013; Khan and Zhang, 2016; Loven et al., 2013).

1.3.2 Histone methylation

Histone methylation occurs at both arginine and lysine residues of histone H3 and H4 (Cao et al., 2002; Wang et al., 2001). Analogous to histone acetylation, histone methylation levels are determined by the combined actions of histone methyltransferase and histone demethylases. Lysine methyltransferases (KMTs) are responsible for methylating histone lysine residues. KMTs contain a catalytic SET domain which catalyzes the transferring of a methyl group from S-adenosylmethionine (SAM) to the ε-amine on the side chain of lysine residue (Guo and Guo, 2007).

The characteristics of histone lysine methylation depends not only on the degree of the methylation, such as mono- (me), di- (me2), and tri-methylation (me3), but also on the location of the methylation. Histone lysine methylation can be recognized by distinct domains of many

chromatin factors thus recruits specific chromatin-modifying enzymes. For example, Histone H3 lysine 4 tri-methylation (H3K4me3) can be bound by CHD1, a nucleosome repositioning enzyme (Sims et al., 2005), and by JMJD2A, a histone demethylase (Huang et al., 2006), or prevent the binding of NuRD complex, a general transcriptional repressor (Zegerman et al., 2002); H3K9me3 binding to HP1, a chromatin binder in its dimer form, is essential for forming repressive heterochromatin structures (Bannister et al., 2001).

Specific histone marks are associated with distinct genomic regions and can be used to predict chromatin states. For instance, H3K4me1 is enriched at active transcriptional enhancers; H3K4me3 is a hallmark of promoters of active genes and poised genes associated with differentiation; and H3K36me3 is associated with actively transcribed gene body regions (Bannister et al., 2005; Heintzman et al., 2007a; Hon et al., 2009; Mikkelsen et al., 2007; Schneider et al., 2004; Venkatesh et al., 2016). In contrast, methylation of H3K27 and H3K9 are generally associated with repressive chromatin states. H3K9me2 and H3K9me3 are associated with gene repression and mark the heterochromatin regions (Black et al., 2012; Pinheiro et al., 2012; Towbin et al., 2012). KMTs for H3K9 can interact with chromatin remodelers, transcriptional repressors and histone deacetylases to facilitate heterochromatin formation (Salton et al., 2014; Shan et al., 2020; Yamada et al., 2005; Zhang et al., 2008). H3K27me3 forms broad domains at the promoters of silenced and poised genes and at poised enhancers with low H3K4me1, and is essential in repressing genes associated with differentiation and development (Banaszynski et al., 2013; Coward et al., 2018; Rada-Iglesias et al., 2011; Simon and Kingston, 2013). Some genomic regions are co-occupied by more than one histone modifications and have distinct biological functions, such as H3K4me3 and H3K27me3 bivalency found in key developmental gene promoters can

poise genes for rapid activation or repression during development (Bernstein et al., 2006; Kinkley et al., 2016; Sanulli et al., 2015).

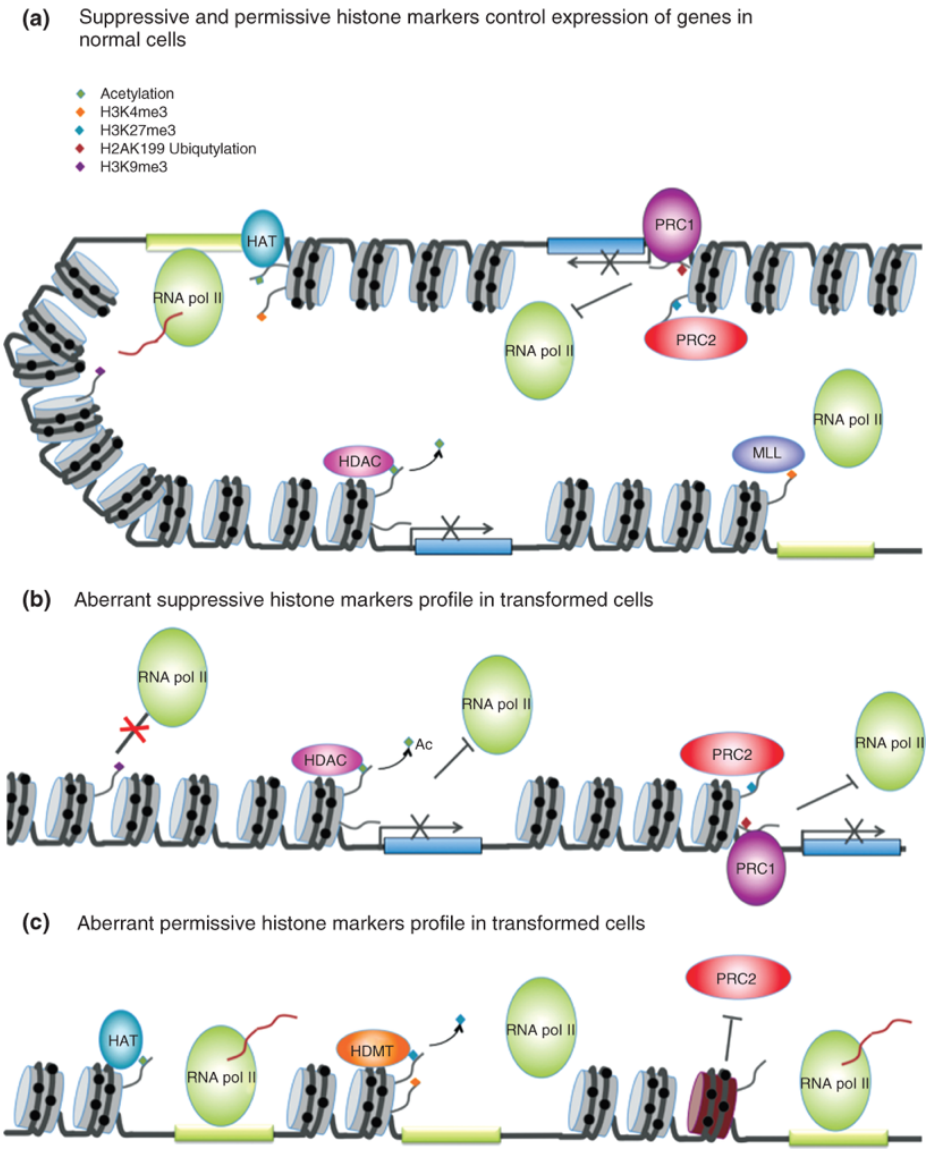

Figure 3. Histone modifications and transcriptional regulation

a) Regulation of gene expression by histone suppressive and permissive markers and associated regulatory factor in nontransformed cells. **b**) Upregulation of suppressive markers and their associated factors or downregulation of permissive markers and their associated protein leads to aberrant expression of genes, such as tumor suppressor genes in transformed cells. **c**) Upregulation of permissive histone marker and their associated regulatory factors or downregulation of suppressive histone markers and their associated regulatory factors leads to aberrant expression of

genes in transformed cells. Variation in histone associated genes (red histone) can lead to abnormal expression of genes through disruption of regulatory factors.

1.4 Epigenetic landscape of the brain during the early fetal period

Transcription is the driver of cell differentiation and specification involving multilateral interactions between epigenetic modifications and transcription factors. While significant progress has been made in defining transcriptional programs within distinct brain cell types (Darmanis et al., 2015; Eze et al., 2021; La Manno et al., 2021; Llorens-Bobadilla et al., 2015; Luo et al., 2015; Prajapati et al., 2019; Zhu et al., 2018), we still lack comprehensive knowledge of how these programs are shaped by mitotically inherited epigenetic states during brain development. Current models, based primarily on studies of the murine hematopoietic system, posit that specific histone epigenetic states are 'primed' in early development and subsequently engaged by the transcriptional machinery following cellular division and differentiation. *In vitro* models suggest that epigenetic priming and stage-specific transcription factor engagement may occur during neuronal differentiation (Ziller et al., 2014). Studies on CpG methylation during *in vitro* differentiation and in cortical tissue during fetal brain development revealed developmental specific methylation states in both human and mouse (Lister et al., 2013; Numata et al., 2012; Siegmund et al., 2007). An *in vitro* study on enhancer and transcriptome landscape during early fetal cortex development using human induced pluripotent stem cell (hiPSC) - derived cortical organoids revealed a major role of enhancer landscape in regulating radial glial cell growth and cortical neuron specification (Amiri et al., 2018). Several studies on histone modifications also showed a critical role of dynamic H3K27me3 reprogramming at neuronal-specific genes in stimulating neurogenesis and neural differentiation (Burgold et al., 2008; Cacci et al., 2017; Desai et al., 2020; Geng et al., 2021; Guan et al., 2021; Wu et al., 2022).

The early fetal period, gestational week 8 (GW8) to mid-gestation (GW18), is a critical period in neocortex development involving extensive neurogenesis and cellular migration (Stiles and Jernigan, 2010). This period is transcriptionally dynamic within and across different brain regions (Kang et al., 2011; Zhu et al., 2018), and is associated with a directional loss of DNA methylation (Spiers et al., 2015) and dynamic changes in enhancer gene regulation with gestational week (Amiri et al., 2018). However, our current understanding of the dynamic DNA methylation states associated with fetal brain development are based on reduced representation techniques that have focused on CpG-rich regions within gene promoters leaving intergenic regulatory regions largely unexplored. Furthermore, the spatiotemporal landscape of histone modifications during fetal brain development and their interplay with DNA methylation states remains to be defined.

Primary neural progenitor cells (NPCs) can be enriched from dissected fetal brain tissue and propagated as neurospheres in culture (Reynolds and Weiss, 1996). NPCs are heterogeneous populations of progenitor cells at different stages of pluripotency, including neuroepithelial cells, radial glial cells and other intermediate progenitors, that differentiate and migrate in a precise spatiotemporal manner(Koo et al., 2022). During the neurogenesis period, NPCs originating from the germinative area including ganglionic eminences (GE) migrate to the cortical layers and NPCs originating from outer cortical layers are characterized by a more differentiated neuronal phenotype compared to GE derived NPCs (Florio and Huttner, 2014; Lavdas et al., 1999; Ulfig, 2002). Indeed, NPCs derived from different brain regions have distinct cellular compositions based on their level of specification and differentiation (Betizeau et al., 2013; Götz and Huttner, 2005; Hansen et al., 2010). However, the regulatory and transcriptional states that drive NPC development in different brain regions during early fetal development is largely unknown.

To address these gaps, we analyzed the transcriptional and epigenetic state of brain tissue and matched NPCs from four human foetuses. From these data we construct regulatory network models for primary human NPCs and the brain regions from which they were derived at three developmental stages.

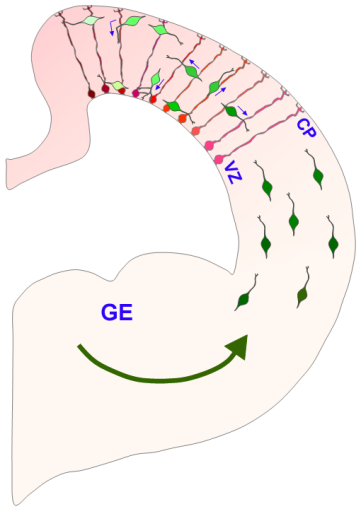

Figure 4. Fetal brain development during neurogenesis.

NPCs originating from ganglionic eminences (GE) migrate to the cortical layers. VZ: ventricular zone. CP: cortical plate. From Yokota Y, Ghashghaei HT, Han C, Watson H, Campbell KJ, Anton E (2007) Radial Glial Dependent and Independent Dynamics of Interneuronal Migration in the Developing Cerebral Cortex. *PLoS ONE* 2(8): e794. Used under Creative Commons Attribution License.

1.5 Objectives

- Construct regulatory network models for primary human NPCs at different developmental stages in early fetal period.
- Identify distinct epigenetic signatures in NPCs derived from different fetal brain regions and at different developmental stages.
- Compare regulatory states in fetal brain derived NPCs from monozygotic (MZ) twins as early as midgestation.

Chapter 2: Materials and Methods

2.1 Human tissue samples and cell culture

Human brain tissues were obtained from disease-free fetuses at GW13, GW15, and GW17 after informed consent. Samples were identified through unlinked codes in accordance with the federal Health Insurance Portability and Accountability Act (HIPAA) guidelines. Specific brain regions were carefully dissected based on landmark and immediately frozen in liquid nitrogen or processed for cell culture. Cortical and GE regions were clearly identified and confirmed by HE section on adjacent tissue. Cells were extracted by gentle tissue dissociation with the Neural Tissue Dissociation kit with papain (Milteny) and were plated at low density in ultra-low attachment cell culture flasks (Corning) in Neural stem cell media (Stemcell technologies) supplemented with EGF (20 ng/ml), bFGF (10 ng/ml) and heparin (2 ug/ml). Cells growing as neurospheres were fed weekly by adding fresh media and passaged every 2 weeks by gentle manual dissociation. After 2 to 3 passages, neurospheres were harvest and frozen in liquid nitrogen.

2.2 Whole-genome Bisulphite Sequencing (WGBS)

Qubit quantified genomic DNA (1–5 mg) was utilized for library construction. Unmethylated Lambda DNA (Promega, Cat no. D1521) was added to genomic DNA for a 0.1% final concentration. DNA was sonicated to a fragment size of B300 bp using a Bioruptor sonicator (Diagenode). End-repair, addition of 30 A bases and adapter ligation was performed as per the Illumina PE genomic DNA sample prep kit protocol except that methylated cytosine PE adapters were used. Bisulfite conversion of purified adapter-ligated DNA was performed using the EZ DNA Methylation Gold kit (ZymoResearch Cat.D5005) according to the manufacturer's instructions. The DNA was then purified with the Qiagen Qiaquick kit, followed by PCR enrichment using

Kapa HiFi Hot Start Uracil þ Ready (Kapa Biosystems, Cat no. KK2801) for five cycles with PCR PE primers 1.0 and 2.0. PCR products were purified with the Qiagen Minelute kit and size selected with PAGE gel purification. DNA libraries were checked for quantity by Qubit (Life Technologies) and quality by Agilent DNA Bioanalyzer (Agilent). Libraries were sequenced using paired-end 100 nt sequencing chemistry on an Illumina HiSeq 2000 following the manufacturer's protocols (Illumina).

Sequence reads were examined for quality, sample swap and reagent contamination using custom in house scripts, and directionally aligned to the human genome (GRCh37-lite) as described (Kundaje et al., 2015). Fractional methylation calls for each CpG were generated using bismark(Krueger and Andrews, 2011).

2.3 Methylated DNA immunoprecipitation sequencing (MeDIP)

DNA (2–5 mg) was sonicated to B100–500bp with a Bioruptor sonicator (Diagenode). Sonicated DNA was end-repaired, A-tailed and ligated to single-end adapters following the standard Illumina protocol. After agarose sized-selection to remove unligated adapters, adapter-ligated DNA was used for each immunoprecipitation using a mouse monoclonal anti-methylcytidine antibody (1 mg/ml, Eurogentec, catalogue no. BI-MECY-0100). DNA was heat-denatured at 95 °C for 10 min, rapidly cooled on ice, and immunoprecipitated with 1 ml primary antibody per microgram of DNA overnight at 4 °C with rocking agitation in 500 ml IP buffer (10 mM sodium phosphate buffer, pH 7.0, 140 mM NaCl and 0.05% Triton X-100). To recover the immunoabsorbed DNA fragments, 1 ml of rabbit anti-mouse IgG secondary antibody (2.5 mg/ml, Jackson Immunoresearch) and 100 ml Protein A/G beads (Pierce Biotechnology) were added and incubated for an additional 2h at 4 °C with agitation. After immunoprecipitation, a total of six IP washes were performed with ice cold IP buffer. A non-specific mouse IgG IP (Jackson

Immunoresearch) was performed in parallel to methyl DNA IP as a negative control. Washed beads were resuspended in Tris-EDTA buffer (TE) with 0.25% SDS and 0.25 mg/ml proteinase K for 2 h at 55 °C and then allowed to cool to room temperature. MeDIP and supernatant DNA were purified using Qiagen MinElute columns and eluted in 16 ml EB (Qiagen, USA). Fifteen cycles of PCR were performed on 5 ml of the immunoprecipitated DNA using the single-end Illumina PCR primers. The resulting reactions were purified over Qiagen MinElute columns, after which a final size selection (192–392 bp) was performed by electrophoresis in 2% agarose. Libraries were quality controlled by spectrophotometry and Agilent DNA Bioanalyzer analysis. An aliquot of each library was diluted in EB to 5 ng/ml and 1 ml used as template in four independent PCR reactions to confirm the enrichment of methylated and de-enrichment of unmethylated sequences, compared with 5 ng of the input (sonicated DNA). Two positive controls (SNRPN and MAGEA1 promoters) and two negative controls (a CpG-less sequence on Chr15 and GAPDH promoter) were amplified. Cycling was 95 °C for 30 s, 58 °C for 30 s, 72 °C for 30 s with 30 cycles. PCR products were visualized by 1.8% agarose gel electrophoresis.

Raw MeDIP-seq sequences were examined for quality, sample swap and reagent contamination using custom in house scripts. Sequence reads were aligned to NCBI GRCh37-lite reference using BWA 0.5.7(Li and Durbin, 2010) and default parameters. To transform aligned MeDIP-seq sequences to single CpG fractional calls for each library we calculated the MeDIP coverage signal for CpGs genome wide and the average coverage in all genomic regions, which were >= 500bp away from a CpG. The latter was used as a background. Next, we convert signal and background values into the MeDIP methylation score: a continuous value between 0 and 1 with a distribution very similar to the one of WGBS fractional methylation. During this process

we exclude locations with mappability < 0.5 and correct for mappability for the locations with mappability between 0.5 and 1.

2.4 Methylation sensitive restriction enzyme sequencing (MRE-seq)

Three parallel digests were performed (HpaII, AciI and Hin6I; Fermentas), each with 1 mg of DNA. Five units of enzyme per microgram DNA were added and incubated at 37 °C in Fermentas 'Tango' buffer for 3h. A second dose of enzyme was added (five units of enzyme per microgram DNA) and the DNA was incubated for an additional 3h. Digested DNA was precipitated with sodium acetate and ethanol, and 500ng of each digest were combined into one tube. Combined DNA was size selected by electrophoresis on a 1% agarose Tris–borate–EDTA gel. A 100–300bp gel slice was excised using a sterile scalpel and gel-purified using Qiagen Qiaquick columns, eluting in 30ml of Qiagen EB buffer. Library construction was performed using the Illumina Genomic DNA Sample Kit (Illumina Inc., USA) with single-end adapters, following the manufacturer's instructions with the following changes. For the end repair reaction, T4 DNA polymerase and T4 polynucleotide kinase were excluded and the Klenow DNA polymerase was diluted 1:5 in water, and 1 ml was used per reaction. For single-end oligo adapter ligation, adapters were diluted 1:10 in water, and 1ml was used per reaction. After the second size selection, DNA was eluted in 36ml EB buffer using Qiagen Qiaquick columns, and 13ml was used as a template for PCR, using Illumina reagents and cycling conditions with 18 cycles. After cleanup with Qiagen MinElute columns, each library was examined by spectrophotometry (Nanodrop, Thermo Scientific, USA) and Agilent DNA Bioanalyzer (Agilent, USA). Sequence reads were aligned to NCBI GRCh37-lite reference using BWA 0.5.7(Li and Durbin, 2010) and default parameters.

2.5 Chromatin immunoprecipitation sequencing (ChIP-seq)

Standard operating procedures for ChIP-seq library construction are available (http://www.roadmapepigenomics.org/protocols/type/experimental/) or by request. ChIP-seq library construction involves the following standard operating procedures (SOPs) in order: (1) Crosslinking of frozen cell pellet; (2) DNA sonication using Sonic Dismembrator 550; (3) SLX-PET protocol for Illumina sample preparation. Antibodies used in this study were subjected to rigorous quality assessment to meet the reference epigenome mapping quality standards (http://www.roadmapepigen-omics.org/protocols) including western blot of whole-cell extracts, 384 peptide dot blot (Active Motif MODified Histone Peptide Array) and ChIP-seq using control cell pellets (HL60). Antibody vendor, catalogue number and lot are provided along with ChIP-seq library construction details as part of the metadata associated with all ChIP-seq data sets and available through GEO and the NCBI epigenomics portals (for example, http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM613886). Final library distributions were calculated using an Agilent Bioanalyzer and quantified by fluorometric quantification (Qubit, Life Technologies). Libraries were sequenced using single-end 76nt sequencing chemistry on an Illumina GAiix or HiSeq 2000 following the manufacturer's protocols (Illumina) as either single or multiplex using custom index adapters added during library construction.

Raw sequences were examined for quality, sample swap and reagent contamination using custom in house scripts. Sequence reads were aligned to NCBI GRCh37-lite reference using BWA 0.5.7(Li and Durbin, 2010) and default parameters, and assessed for overall quality using Findpeaks (Fejes et al., 2008). Aligned reads were directionally extended by the average insert size of the DNA fragments for a given library estimated from Agilent Bioanalyzer (Agilent) profiles

measured during library construction and varied between B130 and 250bp. Custom java program (BAM2WIG) was used to generate wig files for downstream analysis and visualization. Reads with BWA mapping quality scores <5 were discarded and reads that aligned to the same genomic coordinate were counted only once in the profile generation.

2.6 RNA sequencing (RNA-seq)

Standard operating procedures for RNA-seq library construction are available (http://www.roadmapepigenomics.org/protocols/type/experimental/) or by request. RNA-seq library construction involves the following standard operating procedures (SOPs) in order: (1) Purification of polyA + mRNA and mRNA(-) flowthrough total RNA using MultiMACS 96 separation unit; (2) strand-specific 96-well complementary DNA (cDNA) synthesis; (3) strand-specific 96-well library construction for Illumina sequencing. In brief, polyA + RNA was purified using the MACS mRNA isolation kit (Miltenyi Biotec, Bergisch Gladbach, Germany), from 2–10 mg of total RNA with a RIN >= 7 (Agilent Bioanalyzer) as per the manufacturer's instructions. The process included on-column DNaseI treatment (Invitrogen, Carlsbad, CA, USA). Double-stranded cDNA was synthesized from the purified polyA + RNA using the Superscript II Double-Stranded cDNA Synthesis kit (Invitrogen) and 200 ng random hexamers (Invitrogen). After first strand synthesis, dNTPs were removed using 2 volumes of AMPure XP beads (Beckman Genomics, Danvers, MA, USA). GeneAmp (Invitrogen) 12.5 mM dNTPs blend (2.5 mM dCTP, 2.5 mM dGTP, 2.5 mM dATP, 5.0 mM dUTP) was used in the second-strand synthesis mixture in the presence of 2 mg of ActinomycinD. Double-stranded cDNA was purified using 2 volumes of Ampure XP beads, fragmented using Covaris E series shearing (20% duty cycle, Intensity 5, 55 s), and used for paired-end sequencing library preparation (Illumina). Before library amplification uridine digestion was performed at 37 °C for 30 min following with 10 min at 95 °C in Qiagen

Elution buffer (10 mM Tris–Cl, pH 8.5) with 5 units of Uracil-N- Glycosylase (UNG: AmpErase). The resulting single-stranded sequencing library was amplified by PCR (10–13 cycles) to add Illumina P5 and P7 sequences for cluster generation. PCR products were purified on Qiaquick MinElute columns (Qiagen, Mississauga, ON) and assessed and quantified using an Agilent DNA 1000 series II assay and Qubit fluorometer (Invitrogen, Carlsbad, CA), respectively. Libraries were sequenced using paired-end 76 nt sequencing chemistry on a cBot and Illumina GAiix or HiSeq 2000 following the manufacture's protocols (Illumina).

RNA-seq pair-end reads are aligned to a transcriptome reference consisting of the reference genome extended by the annotated exon–exon junctions(Morin et al., 2008). To generate transcriptome reference we used the JAGuaR v 1.7.6 pipeline(Butterfield et al., 2014). Reads aligned to a custom transcriptome reference (build from NCBI GRCh37-lite reference and Ensembl v65 (GenCode v10) annotations) were then 'repositioned' on to genomic coordinates, transforming reads that span exon–exon junctions into large-gapped alignments. Using repositioned reads we generated genome-wide coverage profiles (wiggled files) using a custom BAM2WIG java program (http:// www.epigenomes.ca/tools.html) for further analysis and visualization in genome browsers. To generate profiles we included pairs that are marked as duplicated as well as pairs that are mapped in multiple genomic locations. A custom RNA-seq QC and analysis pipeline was applied to the generated profiles and a number QC metrics were calculated to assess the quality of RNA-seq library such as intron–exon ratio, intergenic reads fraction, strand specificity, 30-50 bias, GC bias and reads per kilobase of transcript per million reads mapped (RPKM) discovery rate. To quantify the exon and gene expression we calculated modified RPKM metrics(Mortazavi et al., 2008). For the normalization factor in RPKM calculations we used the total number of reads aligned into coding exons and excluded reads from

mitochondrial genome as well as reads falling into genes coding for ribosomal proteins as well as reads falling into top 0.5% expressed exons. RPKM for a gene was calculated using the total number of reads aligned into its all merged exons normalized by total exonic length. The resulting files contain RPKM values for all annotated exons and coding and noncoding genes, as well as introns. We also report the coordinates of all significant intergenic RNA-seq clusters not overlapping the annotated genes.

2.7 Integrative data analysis

Single nucleotide polymorphisms (SNPs) were called using SAMtools mpileup with –C50 to reduce the effect of reads with excessive mismatches, and BCFtools with –D100 to exclude sites with read depth greater than 100 for cortex and GE NPCs from all four subjects (Li, 2011). SNPs not included in dbSNP build 138 were filtered out. SNPs were compared between each of the six pairs among the four subjects to obtain discordant variants in NPCs. The overlap of the discordant variants from the two cell types were taken as the discordant SNPs between the pair.

Differential methylated regions (DMRs) for WGBS were identified from differential methylation analysis using a C++ tool with $p < 0.005$(Raineri et al., 2014) and dynamic growth approach to concatenate adjacent differential methylated CpGs within 500bp. Regions with less than 3 CpGs were filtered. On average, WGBS DMRs are ~ 250bp across all samples. For MeDIP, differential methylated CpGs were identified with fractional methylation call differences greater than 0.6 and hypermethylation greater than 0.75. DMRs were concatenated from differential methylation CpGs with the same dynamic growth approach used for WGBS. Regions with less than 4 CpGs were filtered. On average, MeDIP DMRs are ~ 100bp across all samples.

Histone modifications enriched regions were identified with FindER v0.9.3b (http://www.epigenomes.ca/finder.html) using minER = 300bp for H3K4me3, and the overlap of minER = 250bp, 500bp, 1000bp for all other marks to capture narrow and broad peaks. Differentially marked promoters were identified by calculating ChIP-seq signals for promoter regions (TSS +/- 1.5 kb), ranking promoters for all 52418 protein-coding and non-coding genes based on their signal level. Promoters with >= 5000 differences in rank in pairwise comparisons were identified as differentially marked.

Differentially expressed (DE) genes were identified by an in house matlab tool DEfine as previously described (Gascard et al., 2015) using entropy-based thresholds with FDR = 0.01. Exon-level-normalized RPKM were used to identify cell type-specific isoforms. Pairwise comparisons were performed to identify DE exons using DEfine (FDR = 0.01). Expressed exons were defined as those with RPKM > 10% of gene RPKM; unexpressed exons were defined as those with RPKM < 1% of gene RPKM; and the exons in between (1-10% gene RPKM) were discarded to filter out most false positives. Isoforms for each pairwise comparison were identified as genes with DE exons expressed in only one of the two samples, excluding DE genes.

GREAT(McLean et al., 2010) and DAVID(Huang et al., 2009) were employed to study biological processes and pathways associated with regulatory regions and genes using default parameters. To study the transcription factor activities in regulatory regions, Homer was used to identify enriched transcription factors with q-value < 0.01 and Percent of regions with motif > 20%(Heinz et al., 2010). Transcription factor target genes were identified with Homer annotated transcription factor binding site within TSS +/- 1.5 kb. Transcriptional regulation network was analyzed using Cytoscape(Shannon et al., 2003).

Chapter 3: Results

3.1 Epigenome profiling of human fetal brain

To explore the gene regulatory landscape of the developing human brain, dissected brain tissue and primary neurospheres derived from cortex and GE were obtained from four phenotypically normal human fetuses at GW13, GW15 and GW17 (**Figure 5A**). Among these four samples, only Subject 3 from GW15 is male and all other three are females, so we excluded chromosome X and Y when comparing this subject to the others. Immunohistochemistry analysis of primary neurospheres cultures showed cell-type specificity with cortex derived neurospheres displaying an increased neuronal phenotype compared to GE derived neurospheres (**Figure 6** and **Figure 7A**). To confirm the progenitor status of the neurosphere cultures, they were differentiated into neuroblasts (DCX+, CALRETININ+, TUJ1+) and astrocytes (GFAP+) (**Figure 6** and **Figure 7A**). Massively parallel sequencing based assays were employed to generate chromatin immunoprecipitation sequencing (ChIP-seq), whole genome bisulfite sequencing (WGBS), methylated DNA immunoprecipitation sequencing (MeDIP-seq), methylation sensitive restriction enzyme sequencing (MRE-seq), mRNA-sequencing (RNA-seq) and microRNA-sequencing (miRNA-seq) datasets from the dissected primary brain tissue and the neurospheres derived from cortex and GE following the International Human Epigenome Consortium (IHEC) guidelines (**Figure 5B** and **Table 1**) (Stunnenberg and Hirst, 2016).

To establish genetic relationships between twin fetuses we called single nucleotide variants (SNVs) from RNA-seq reads (Morin et al., 2010) and performed pair-wise comparisons across subjects. On average, 11,227 discordant SNVs were identified across all pairs, except the GW17 twins for which only 1,784 were found (**Figure 7B**). Interestingly discordant SNVs identified

between the GW17 twins were homozygous in one twin and heterozygous in the other, hence can be either the result of genetic differences or the result of allele specific expression. To distinguish these possibilities we examined discordant SNVs using MeDIP-seq reads and found that all examined discordant sites were heterozygous across all three tissue and cell types, suggesting that the discordance observed in RNA-seq at these positions was due to allele specific expression. We conclude that the GW17 twins (referenced as Subject1 and Subject2) were genetically identical and thus monozygotic.

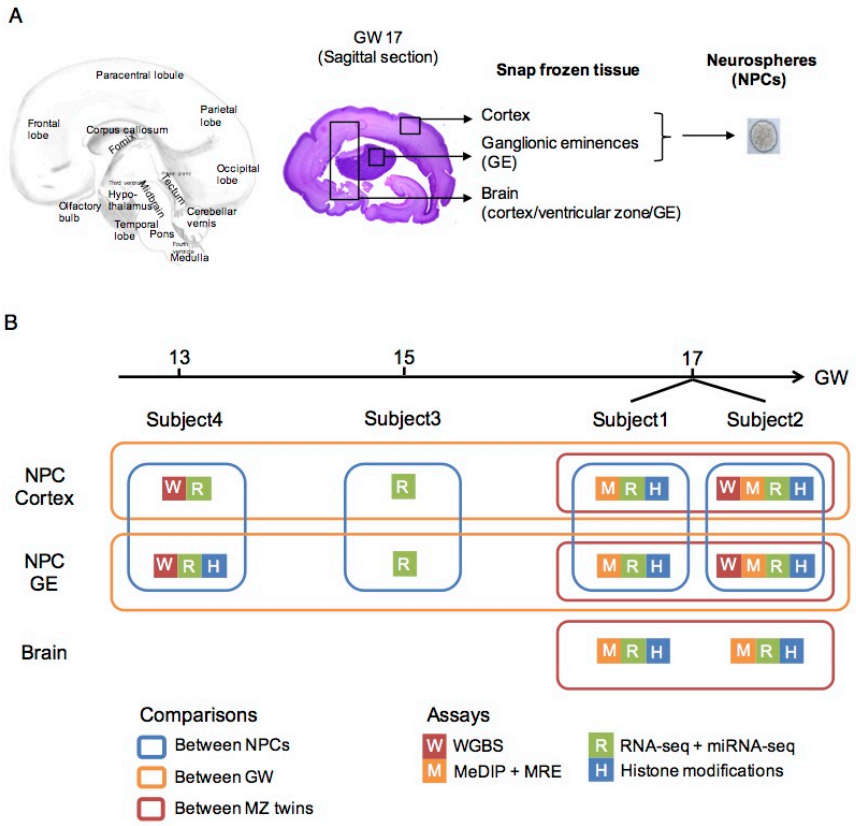

Figure 5. Experimental design and overview

(A). Human brain coronal section highlighting the section plan used during dissection to isolate specific regions as pictured by the boxed area on the hematoxylin eosin section. (B). Summary of sample information and bioinformatics analysis design. Boxes outline the pairwise comparisons, including comparisons between cortex and GE derived NPCs (blue), across three gestational weeks (orange), and between MZ twins (red).

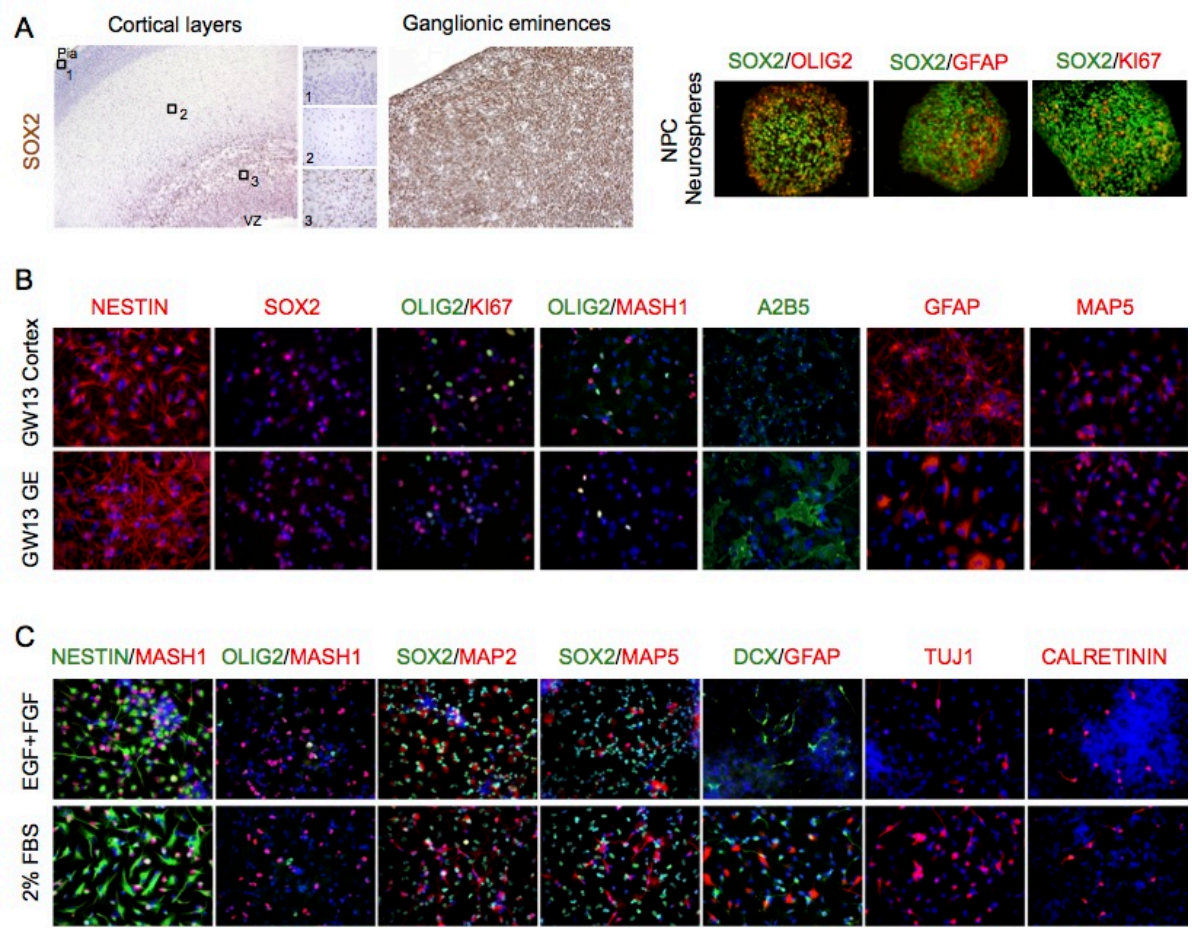

Figure 6. Characterization of fetal brain derived NPCs

(A). Cells isolated from the cortex or GE are able to proliferate when cultured as neurospheres and maintain a stem/progenitor phenotype, confirmed by the expression of SOX2 and OLIG2. (B). When grown in presence of growth factors, NPCs express a mixture of stem cell markers (NESTIN+, SOX2+, OLIG2+), glial progenitors (OLIG2+, A2B5+, MASH1-), astrocytes (GFAP+) and neuroblasts (MAP5+) with cortex derived NPCs presenting a more neuronal phenotype (MASH1 high, MAP5 high) while the GE derived NPCs have a more glial phenotype (A2B5 high, GFAP more differentiated). (C). In 2% serum, NPCs are differentiate into a neuronal lineage, from an immature neuroblast (MAP2+, MAP5+) to a migrating neurons (Doublecortin DCX+, TUJ1+) to a postmitotic neuron (CALRETININ+), as well as into GFAP+ astrocytes.

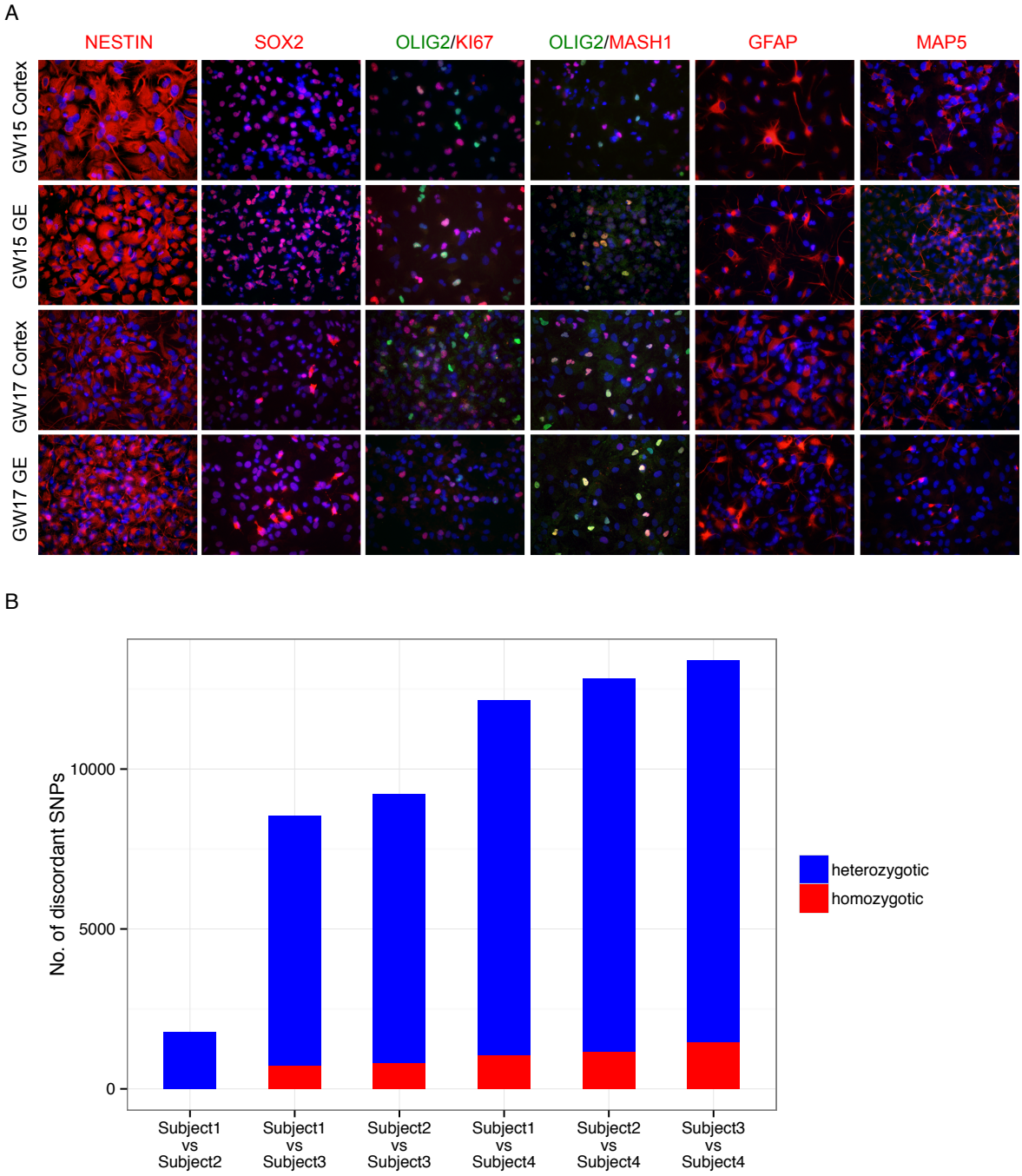

Figure 7. Cell markers and validate monozygotic twins

(A). NPCs characterization from GE and cortical regions isolated from 15 and 17 GW human fetuses. (B). Discordant SNPs were called from all pairs among the four subjects. Red represents number of homozygotic discordant SNPs, and blue shows heterozygotic SNPs.

Table 1. Human fetal brain samples

Original Source	Subject ID	Cell Type	Sex
HuFNSC01	Subject1	Brain	F
HuFNSC01	Subject1	Primary Cell Culture Neurospheres, Cortex Derived	F
HuFNSC01	Subject1	Primary Cell Culture Neurospheres, Ganglionic Eminence Derived	F
HuFNSC02	Subject2	Brain	F
HuFNSC02	Subject2	Primary Cell Culture Neurospheres, Cortex Derived	F
HuFNSC02	Subject2	Primary Cell Culture Neurospheres, Ganglionic Eminence Derived	F
HuFNSC03	Subject3	Primary Cell Culture Neurospheres, Cortex Derived	M
HuFNSC03	Subject3	Primary Cell Culture Neurospheres, Ganglionic Eminence Derived	M
HuFNSC04	Subject4	Primary Cell Culture Neurospheres, Cortex Derived	F
HuFNSC04	Subject4	Primary Cell Culture Neurospheres, Ganglionic Eminence Derived	F

3.2 NPCs derived from different fetal brain regions and at different developmental stages are epigenetically distinct

To identify common regulatory states of NPCs, we defined enhancer regions within the genomes of cortex and GE derived NPCs as regions marked by H3K4me1(Heintzman et al., 2007b). This analysis identified 39,033 core enhancers that are shared across all NPCs regardless of development stage or brain region from which they were derived. Intersecting NPC core enhancers with transcription factor binding sites (TFBSs) predicted by Homer(Heinz et al., 2010) revealed 13 transcription factors significantly enriched (q-value < 0.01 and > 20% enhancers with motif) (**Figure 8**), including the master regulators of pluripotency *NANOG* and *SOX3*, as well as key regulators of brain development such as LIM Homeobox 3 (*LHX3*), a transcription factor involved in the specification of motor neurons and interneurons(Thaler et al., 2002), and oligodendrocyte lineage transcription factor 2 (*OLIG2*), implicated in the early stages of oligodendrocyte specification and maturation(Bouvier et al., 2003; Cai et al., 2007; Jakovcevski and Zecevic, 2005; Lu et al., 2001; Rivera et al., 2006; Zhou et al., 2001). This core set of

enhancers is predicted to play a role in defining the regulatory state of NPCs within the developing fetal brain (**Figure 8**).

We next sought to establish whether cortical and GE tissue derived NPCs retained distinct regulatory states following the *in vitro* expansion associated with their generation. We first identified DNA methylation signatures of cortex and GE NPCs derived at each developmental stage independently. We identified differentially methylated regions genome wide (DMRs; see methods) between cortex and GE derived NPCs by pairwise comparisons of WGBS datasets. DMRs were validated using MeDIP-seq and MRE-seq datasets(Beck, 2010; Li et al., 2015) generated from matching NPCs (**Figure 9**). On average 1412 DMRs (average length 327bp) covering 20,008 CpGs were identified between cortex and GE derived NPCs, with 646 DMRs identified at GW13 and 2178 in GW17 suggesting increased divergence in DNA methylation with developmental age (**Figure 10A**). Directionality was also observed between NPCs with 3.5-fold more DMRs hypomethylated in cortex derived NPCs compared with GE derived NPCs. Consistent with differences in cell migration and fates between cortex and GE derived NPCs(Nadarajah and Parnavelas, 2002), genes associated with cortex-specific hypomethylated DMRs were found to be enriched in forebrain regionalization, while those associated with GE-specific hypomethylated DMRs were enriched in neuronal migration and differentiation (**Figure 10B**). These findings suggest that NPCs derived from distinct brain regions retain the epigenetic signatures specific to these regions.

To examine the transcriptional output of NPCs alone and in association with regulatory states, we generated RNA-seq libraries. Pairwise comparisons (see methods) identified 382 cortex NPC specific and 456 GE NPC specific genes differentially expressed in at least two subjects

(hypergeometric p-value < 10^{-26}). Consistent with an increased divergence in DNA methylation with gestational week, NPCs derived from GW17 exhibited nearly twice the number of differential expressed genes as GW13 (**Figure 8**). In addition, exon level differential expression analysis also identified twice as many putative gene isoforms in GW17 as in GW13, suggesting more divergent transcriptomes between cortex and GE NPCs in the later developmental age (**Figure 8**). Associating differential expressed protein coding genes with H3K4me3 promoter density (TSS +/- 1.5kB; see methods) identified 159 differentially expressed genes with gains/losses in H3K4me3 promoter density, a majority of which (74.8%) showed expected orientation, i.e. gain of H3K4me3 with up-regulation and vice versa. Gains and losses in H3K27me3 promoter density were associated with 508 differentially expressed genes with a majority (75.4%) in the expected orientation, i.e. gain of H3K27me3 with down-regulation and vice versa.

Functional analysis of genes differentially expressed in both GW13 and GW17 between cortex and GE derived NPCs revealed enrichment in genes implicated in neurogenesis and cell morphogenesis regulation (**Figure 8**), including glial fibrillary acidic protein (*GFAP*, **Figure 10C**), an astrocyte marker associated with brain disorders including gliomas(Rodriguez et al., 2001; Rutka and Smith, 1993; Zhao et al., 2004), nuclear factor I/X (CCAAT-binding transcription factor), (*NFIX,* **Figure 10C**), an active transcription factor essential for the development of a number of organ systems including brain(Campbell et al., 2008; Mason et al., 2009), FEZ family Zinc Finger 1 (*FEZF1*), a transcription repressor involved in the axonal projection and proper termination of olfactory sensory neurons (Shimizu and Hibi, 2009), Orthodenticle Homeobox proteins (*OTX1* and *OTX2*), essential in patterning the developing brain(Acampora et al., 1999), and Ventral Anterior Homeobox 1 (*VAX1*), involved in forebrain and visual system development(Hallonet et al., 1999).

Taken together, our analysis between cortex and GE derived NPCs supports a model of distinct epigenetic and transcriptional states in fetal brain regions as early as GW13 that become increasingly divergent with developmental age. We also provide evidence that NPCs derived from distinct brain regions retain region of origin signatures when cultured as neurospheres.

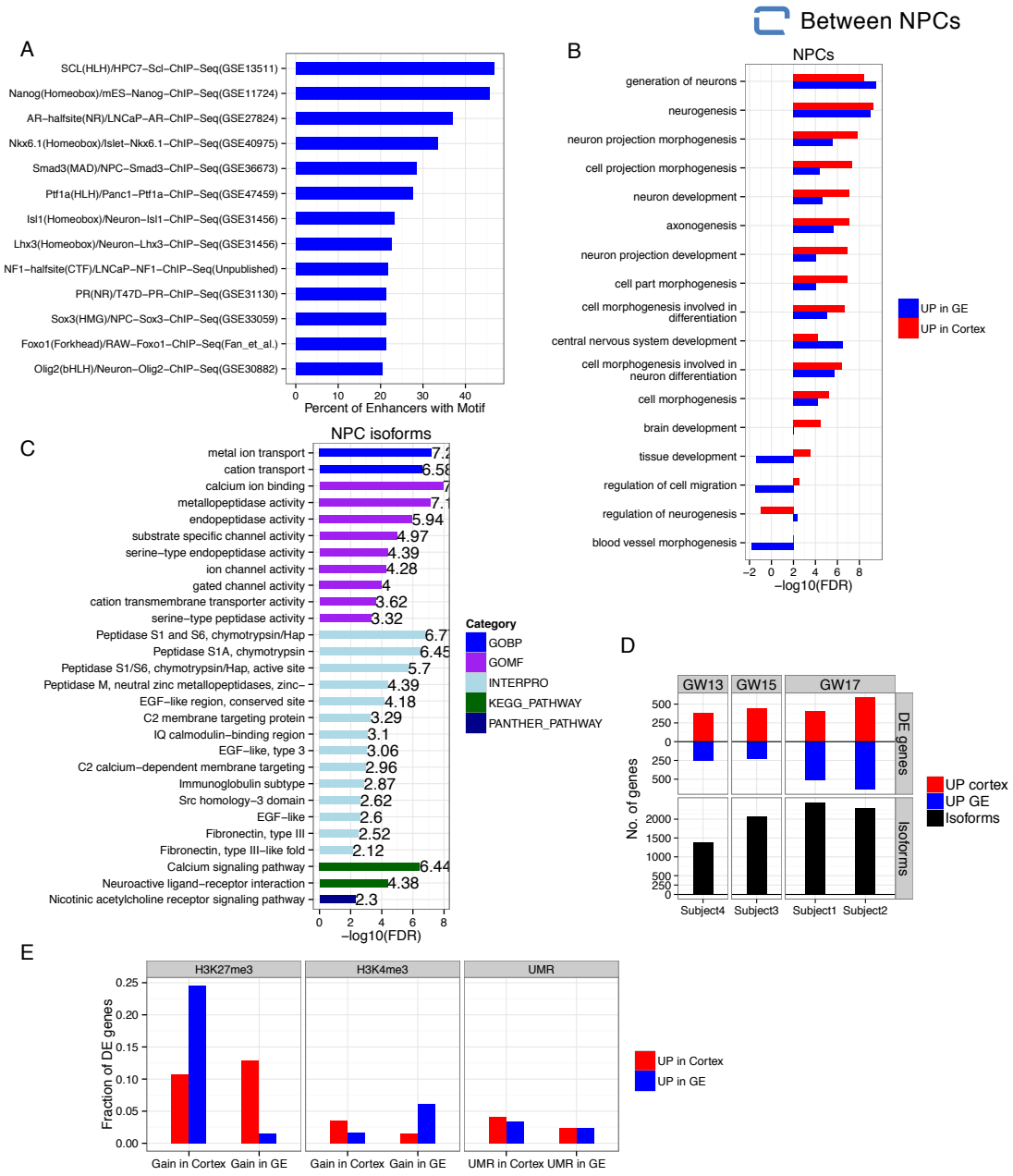

Figure 8. Epigenetic profiles of NPCs

(A). Transcription factors enriched in NPC core enhancers. (B). DAVID GOBP analysis of differential expressed genes between NPCs showed enrichment in neurogenesis. (C). DAVID functional enrichment of NPC isoform suggested enrichment for signalling related terms. (D). Number of differential expressed (top) / isoform (bottom) genes between cortex and GE NPCs in four subjects. (E). Fraction of differentially expressed genes associated with differential methylation of H3K4me3 and H3K27me3 in promoter regions, as well as proximal UMRs showed expected orientation in H3K4me3 and H3K27me3 (upregulation upon gain of H3K4me3 or loss of H3K27me3, vice versa), but no orientation for UMRs (hypomethylated regions).

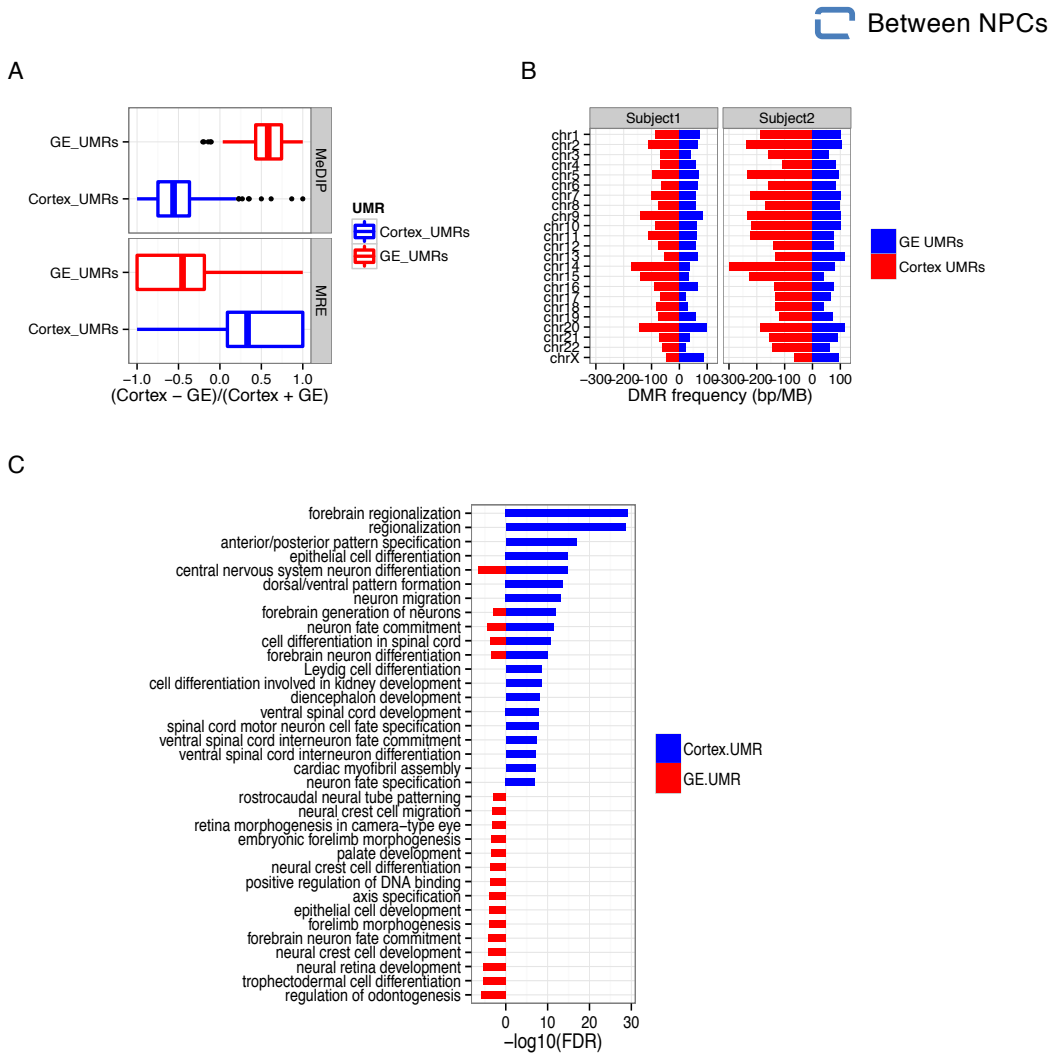

Figure 9. Epigenetic and transcriptional differences between NPCs

(A). For each hypomethylated DMR (UMR) called from WGBS, calculate corresponding MeDIP and MRE signal levels. The differences between two NPC cell types were then calculated as differences in signal levels divided by sum of signal levels in two cell types. (B). MeDIP UMR frequency (bp/MB) across all chromosomes supported the asymmetry observed from WGBS UMRs, with more cortex UMRs than GE UMRs. (C). GREAT GOBP enrichment for MeDIP UMRs supported results from WGBS UMRs.

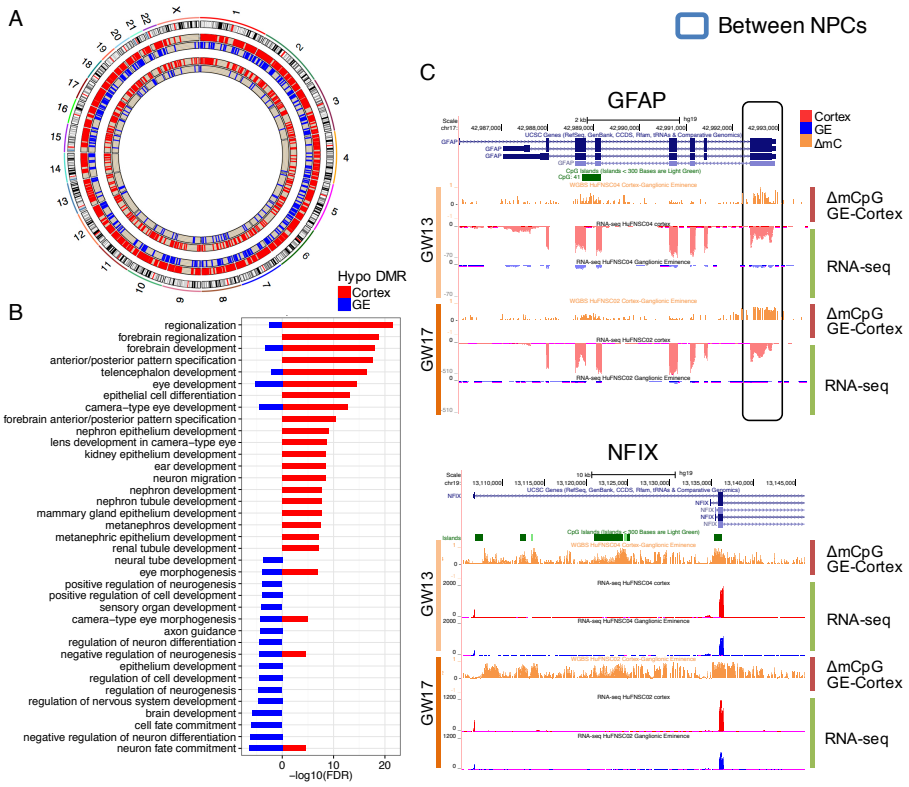

Figure 10. Cortex and GE derived NPCs were epigenetically distinct

(A). DMRs between cortex and GE derived NPCs. More hypomethylated DMRs were identified in cortex NPCs (red) compared to GE NPCs (blue). Moreover, there were fewer DMRs found in GW13 (Subject4, inner circles) than GW17 (Subject2, outer circles). (B). GREAT Gene Ontology biological processes showed enrichment (region-based binomial and hypergeometric FDR < 0.05) in forebrain development terms in cortex hypomethylated DMRs (red), and neuron differentiation and cell fate commitment in GE hypomethylated DMRs (blue). (C). UCSC genome browser tracks of examples of key regulators that were hypomethylated in the promoters and up-regulated in cortex derived NPCs, top panel: Glial Fibrillary Acidic Protein (GFAP), bottom panel: Nuclear Factor I/X (CCAAT-Binding Transcription Factor, NFIX).

3.3 NPCs undergo epigenetic and transcriptional activation during early fetal brain development

The classical view of epigenetic patterning during development posits a model whereby cellular differentiation is accompanied by epigenetic restriction(Waddington, 1942). To explore whether evidence of such epigenetic restriction could be found within the genomes of cells separated by 4 weeks of fetal brain development, hypomethylated DMRs between GW13 and GW17 were identified in cortex and GE derived NPCs respectively (**Figure 11**). In agreement with a previously reported developmentally associated genome hypomethylation in grossly dissected human fetal brain tissue(Spiers et al., 2015), we observed 3.4-fold increase in the number of CpGs within hypomethylated DMRs in GW17 NPCs compared to GW13 (**Figure 11A**). Functional enrichment analysis of genes associated with GW17 hypomethylated DMRs revealed enrichment in cell fate commitment and brain cell differentiation (**Figure 12**). GW17-specific hypomethylated CpGs within DMRs were significantly enriched in core NPC enhancers (hypergeometric p-value $< 10^{-26}$) while GW13-specific DMRs showed no such enrichment (**Figure 11B**), suggesting that a loss of CpG methylation at active enhancers follows the establishment of histone mediated enhancer states during brain development. To validate this observation we examined the DNA methylation states of fetal brain regions across developmental stages in an independent cohort generated by 450K array. Consistent with our observation a directional loss of methylation with development stages we found an increase in the number of probes that were hypomethylated in GW17 (difference in beta value > 0.15) compared to those that were hypomethylated in any of the GW13 individuals (**Figure 13**). Furthermore, 241 out of the 667 probes within hypomethylated regions identified by WGBS were also hypomethylated in 450K array, a 12-fold enrichment of all hypomethylated probes.

We next explored gene promoters (TSS +/- 1.5 kb) differentially marked by H3K4me3 and H3K27me3 between GW13 and GW17. H3K4me3 differentially marked the promoters of 526 protein-coding genes with 2.6-fold more genes gaining H3K4me3 in GW17 compared with GW13. We observed the opposite trend in H3K27me3 with a 6-fold decrease in H3K27me3 marked gene promoters in GW17 compared to GW13, consistent with a more active promoter histone modification landscape in GW17. A subset of the genes associated with the differentially marked promoters were also differentially expressed with a majority (73%) in the expected orientation and significantly enriched in genes related to neurogenesis and axonogenesis (**Figure 12**).

To explore the transcription factor regulatory network of GW13 and GW17 NPCs, we overlapped gestational week specific enhancers, differentially marked by H3K4me1 between GW13 and GW17, with TFBSs and identified five transcription factors exclusively enriched in GW17 specific enhancers, three of which were expressed in both cortex and GE derived NPCs (**Figure 11C**). The three expressed transcription factors play important roles in brain development: *OLIG2*, *FOXO1*, involved in cell specification(Zhou et al., 2015); and androgen receptor (*AR*), involved in sex differences in brain morphology and behaviour(Zuloaga et al., 2008). Enhancers containing these TFBSs were also enriched in hypomethylated CpGs(Feldmann et al., 2013) at GW17 (**Figure 11D**), and genes associated with these hypomethylated enhancers(Heinz et al., 2010) were transcriptional up-regulated in GW17 compared to GW13 (**Figure 11D**). The *OLIG2* promoter itself was also hypomethylated and the gene up-regulated in both cortex and GE NPCs at GW17 (**Figure 13** and **Figure 14**). Moreover, 60% of *OLIG2* target genes were both hypomethylated in their promoters and up-regulated (**Figure 14B**), and were enriched (FDR < 0.05) in genes implicated in neurogenesis in both cortex and GE derived NPCs (**Figure 14C**). Among the predicted 132 target genes, 25 were differentially expressed between GW13 and

GW17, of which 22 genes were up-regulated at GW17 (**Figure 14D**). These up-regulated genes were highly enriched (FDR < 0.05) in genes critical for axon guidance and cell migration, including EPH Receptor A3 (*EPHA3*)(Egea and Klein, 2007), netrin-1 (*NTN1*)(Hamasaki et al., 2001), Netrin G1 (*NTNG1*)(Kennedy, 2000), and Slit Homolog 2 (*SLIT2*)(Hu, 1999), reinforcing the critical role of *OLIG2* in regulating the transcriptional and epigenetic programming of neurogenesis in early human development.

Taken together, our analysis suggests that epigenetic and transcriptional activation, rather than restriction, is the dominant signature associated with the specification of transcriptional programs associated with early brain development.

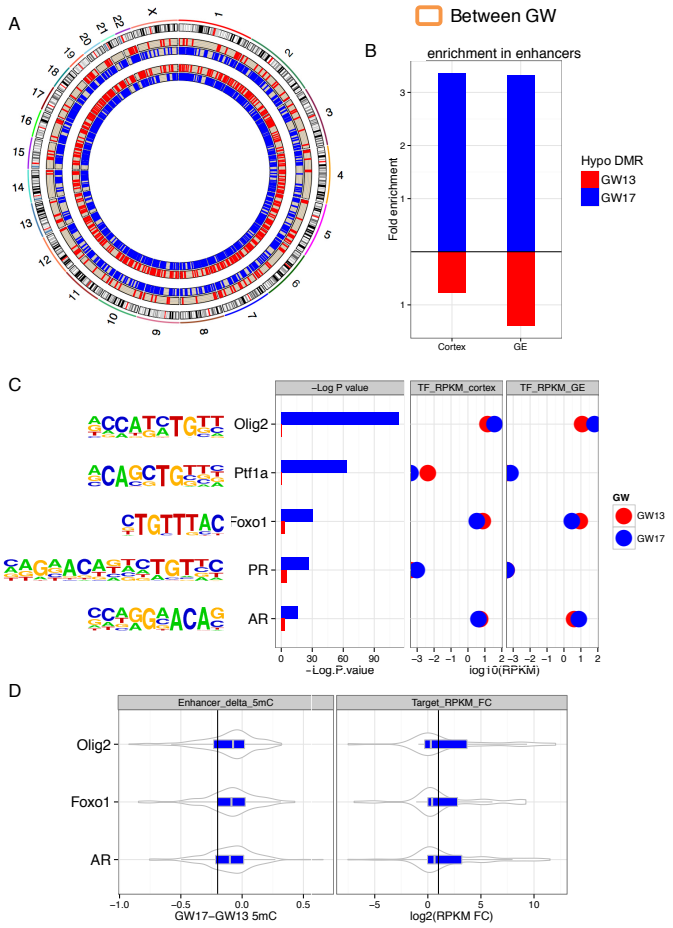

Figure 11. Temporal changes revealed a more active epigenetic state in the later developmental stages

(A). DMRs between GW13 and GW17. We identified more hypomethylated DMRs in GW17 (blue) than GW13 (red) in both GE derived NPCs (inner circles) and cortex derived NPCs (outer circles). (B). Fold enrichment of DMRs between GW13 and GW17 in core NPC enhancers. GW13 hypomethylated DMRs, shown in red, were not enriched in the enhancers, while GW17 hypomethylated DMRs, shown in blue, were enriched. (C). Transcription factors exclusively enriched (Benjamini corrected p-value < 0.01, left panel, and percent of enhancers with motif > 20%) in GW17 unique enhancers, and their transcription levels in cortex derived NPCs (middle panel) and GE derived NPCs (right panel). GW13 expression values are shown in red, and GW17 are shown in blue. Out of the five transcription factors, *OLIG2*, *FOXO1*, and *AR* were expressed. (D). Methylation differences between GW17 and GW13 in enhancers with the particular transcription factor binding sites showed hypomethylation in GW17 (left panel). The vertical line, indicating hypomethylation greater than 0.2, was used as a cut-off for hypomethylated enhancers. Fold change of transcriptional levels between GW17 and GW13 for the nearest genes of the hypomethylated enhancers identified in the left panel showed up-regulation in GW17 (right panel). The vertical line indicates 2-fold increase in expression.

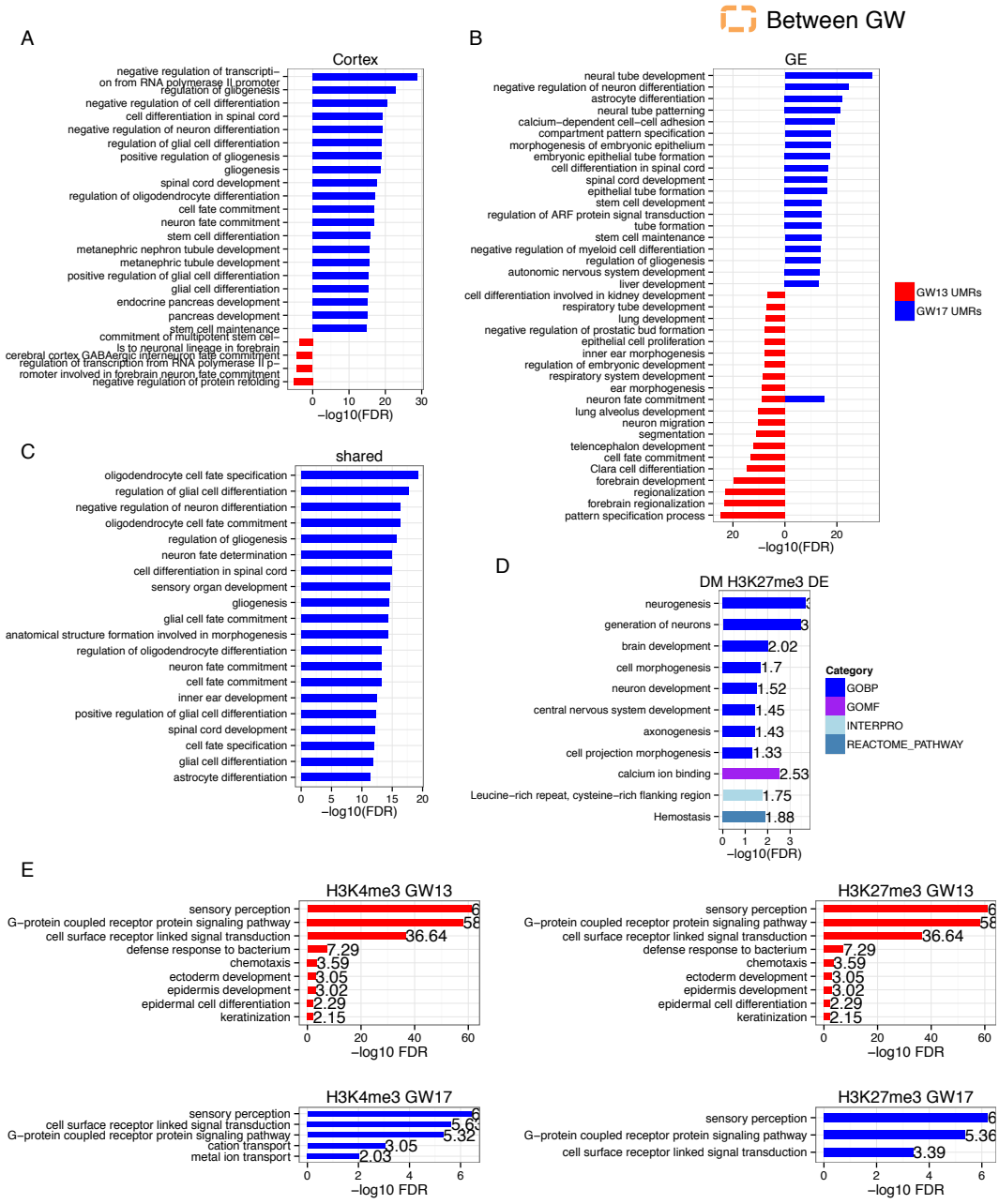

Figure 12. Epigenetic differences across developmental stages

GREAT GOBP functional enrichment analysis for genes associated with GW13 UMRs (red) and GW17 UMRs (blue) in cortex NPCs (A), GE NPCs (B), and UMRs shared by the two cell types (C). (D). DAVID functional enrichment analysis of genes differentially expressed and differentially methylated by H3K27me3 in promoters between GW13 and GW17. (E). DAVID GOBP analysis for genes differentially marked by H3K4me3 or H3K27me3 in promoter regions between GW13 and GW17 showed enrichment in sensory perception.

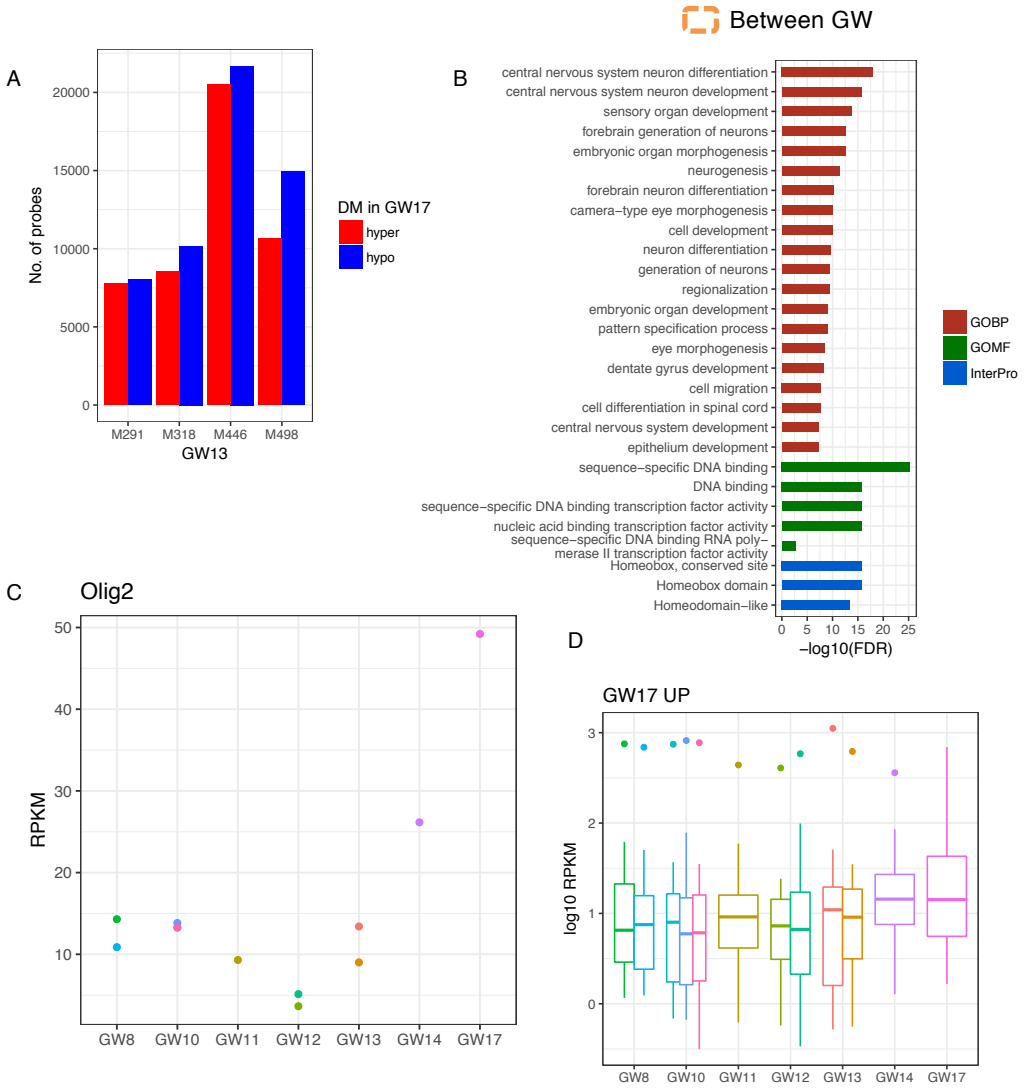

Figure 13. Validating changes between GW13 and GW17 with validation datasets

(A). Identify differentially methylated probes between GW13 and GW17 by pairwise comparison with difference in beta value > 0.15. There were more hypomethylated probes in GW17 compared to all GW13 individuals. (B). Functional enrichment for the 241 GW17 hypomethylated probes also identified in WGBS using GREAT. (C). Expression levels of *OLIG2* in the validation sets showed transcriptional upregulation in later stages. (D). Expression levels in validation sets of genes identified as upregulated in GW17 in NPCs showed upregulation in GW17 as well.

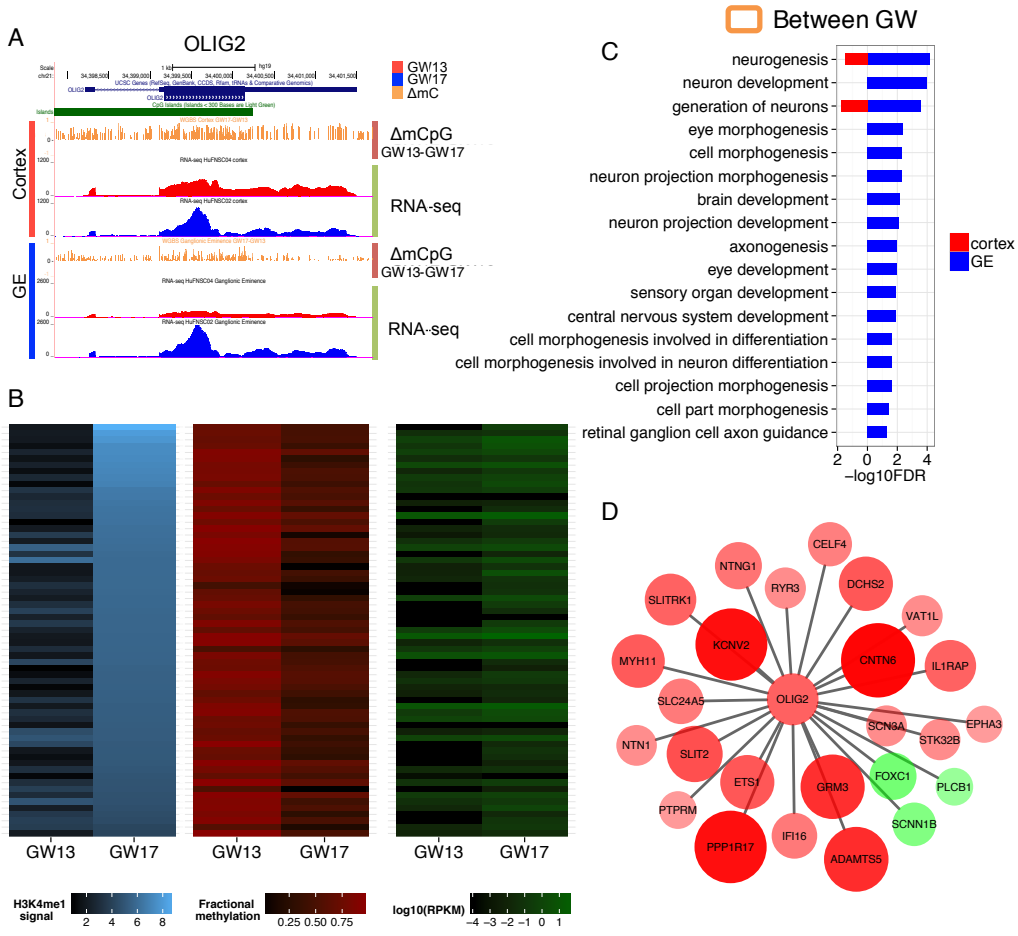

Figure 14. Regulatory network of OLIG2

(A). UCSC genome browser tracks showed Oligodendrocyte Lineage Transcription Factor 2 (*OLIG2*) hypomethylated in promoter region and up-regulated in GW17. (B). Heatmap of enhancer signal level (H3K4me1) of GW17 specific enhancers with *OLIG2* binding sites (left), their corresponding DNA fractional methylation (middle), and transcription levels of their nearest genes (right) showed hypomethylation of the enhancers and up-regulation of the predicted target genes in GW17. (C). DAVID (Huang et al., 2009) GO biological processes enrichment analysis of *OLIG2* target genes indicating significant enrichment for brain development related biological processes. (D). Cytoscape network of *OLIG2* target genes that are differential expressed. 22 of genes were up-regulated in GW17 (red) while only 3 were down-regulated (green). The size of the circles corresponds to the fold change of expression levels between GW17 and GW13. Among the up-regulated genes, *EPHA3, NTN1, NTNG1, and SLIT2* were associated with axon guidance and cell migration.

3.4 Temporal transcriptional programs are specific for NPCs derived from different brain regions

Unsupervised clustering of protein-coding gene expression showed that GW13 NPCs from different brain regions clustered together and with GW15 GE derived NPCs, while GW15 cortex derived NPCs clustered with GW17 cortex NPCs (**Figure 15A**). These relationships were also observed by independent clustering of the expression of exons, non-coding genes, and miRNAs (**Figure 16**). It suggests that NPCs derived from different regions were transcriptionally similar at GW13 but that by GW15 region-specific signatures emerge. Differentially expressed genes between GW13, GW15, and GW17 were identified, and categorized into eight profile groups based on their temporal expression pattern (up-regulated, down-regulated, and not differentially expressed) between GW13-GW15 and GW15-GW17. While overall, we observed a trend for increasingly gene expression with GW in NPCs (**Figure 15B**), differentially expressed genes shared by both cortex and GE derived NPCs showed no discernable pattern across gestational weeks (**Figure 15C**). In contrast an increase in expression was observed between GW13-GW15 in cortex derived NPCs (**Figure 15C**). In GE derived NPCs a similar up-regulation in gene expression was observed however it occurred between GW15-GW17 (**Figure 15C**). Several key factors in brain development followed this pattern (**Figures 15D**), including *OLIG2*, as well as LIM homeobox 4 (*LHX4*), a transcription factor critical for nervous system development(Hunter and Rhodes, 2005); *WNT5A*, involved in regulating axon growth and guidance(Blakely et al., 2011); and Forkhead Box H1 (*FOXH1*), playing key roles in patterning of the forebrain(Silvestri et al., 2008). These differentially expressed genes were also confirmed in our independent validation dataset (**Figure 13**).

Taken together, our analysis provides a high-resolution view of the key transcriptional and epigenetic network changes occurring during early fetal brain development and supports a model of increased transcriptional activity associated with different developmental stages in NPCs derived from different brain regions.

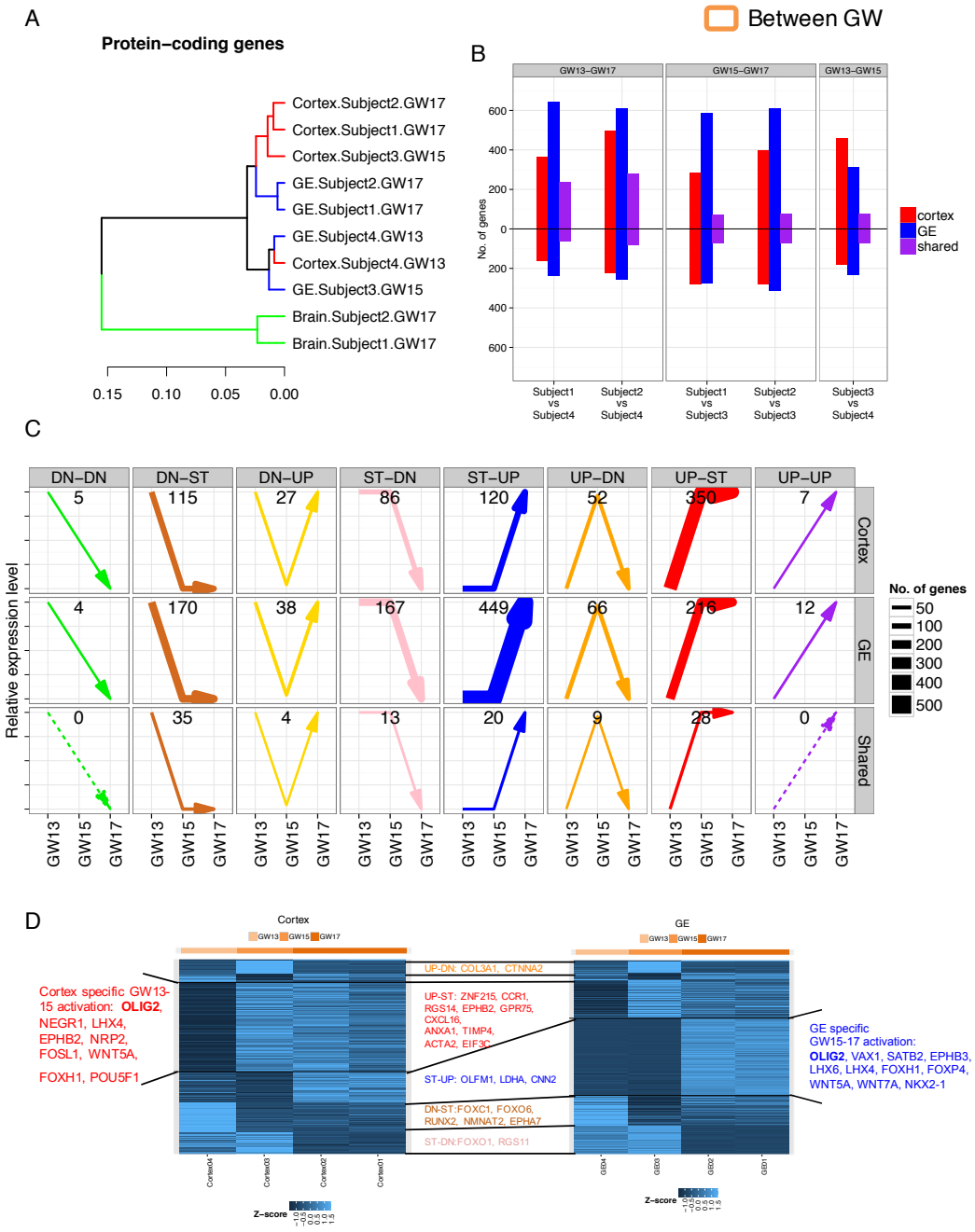

Figure 15. Stage-specific differential expression

(A). Unsupervised clustering of expression levels of protein coding genes. NPCs derived from different regions clustered together at GW13. However, by GW15, GE derived NPCs still clustered with GW13 NPCs, while cortex derived NPCs branched off and clustered with GW17 cortex NPCs. (B). Number of differentially expressed genes between different gestational weeks (left panel: GW13 vs. GW17, middle panel: GW15 vs. GW17, right panel: GW13 vs. GW15) in cortex derived NPCs (red), GE derived NPCs (blue), and shared by two cell populations (purple). Bars pointing up showed up-regulation in the later stages, and bars pointing down showed down-regulation, suggesting a general trend of up-regulation in later stages. (C). Patterns of expression for genes differential expressed between gestational weeks in cortex derived NPCs (top panel), in GE derived NPCs (middle panel), and in genes shared by both cortex and GE derived NPCs (bottom panel). Genes are divided into eight expression profile groups according to differential expression analysis of GW13 vs. GW15 and GW15 vs. GW17, represented by the arrows of eight different colours. The thickness of the line and the number in each panel corresponded to the number of genes within each category, and the dashed lines means no genes in that category. Cortex derived NPCs showed dominant transcription activation in GW13-GW15, while GE derived NPCs showed similar activation but in GW15-GW17. (D). Heatmap for expression values of stage-specific differential expressed genes in cortex NPCs and GE NPCs showed stage-specific expression patterns for key regulators of brain development. UP: up-regulated; DN: down-regulated; ST: stable, not differentially expressed.

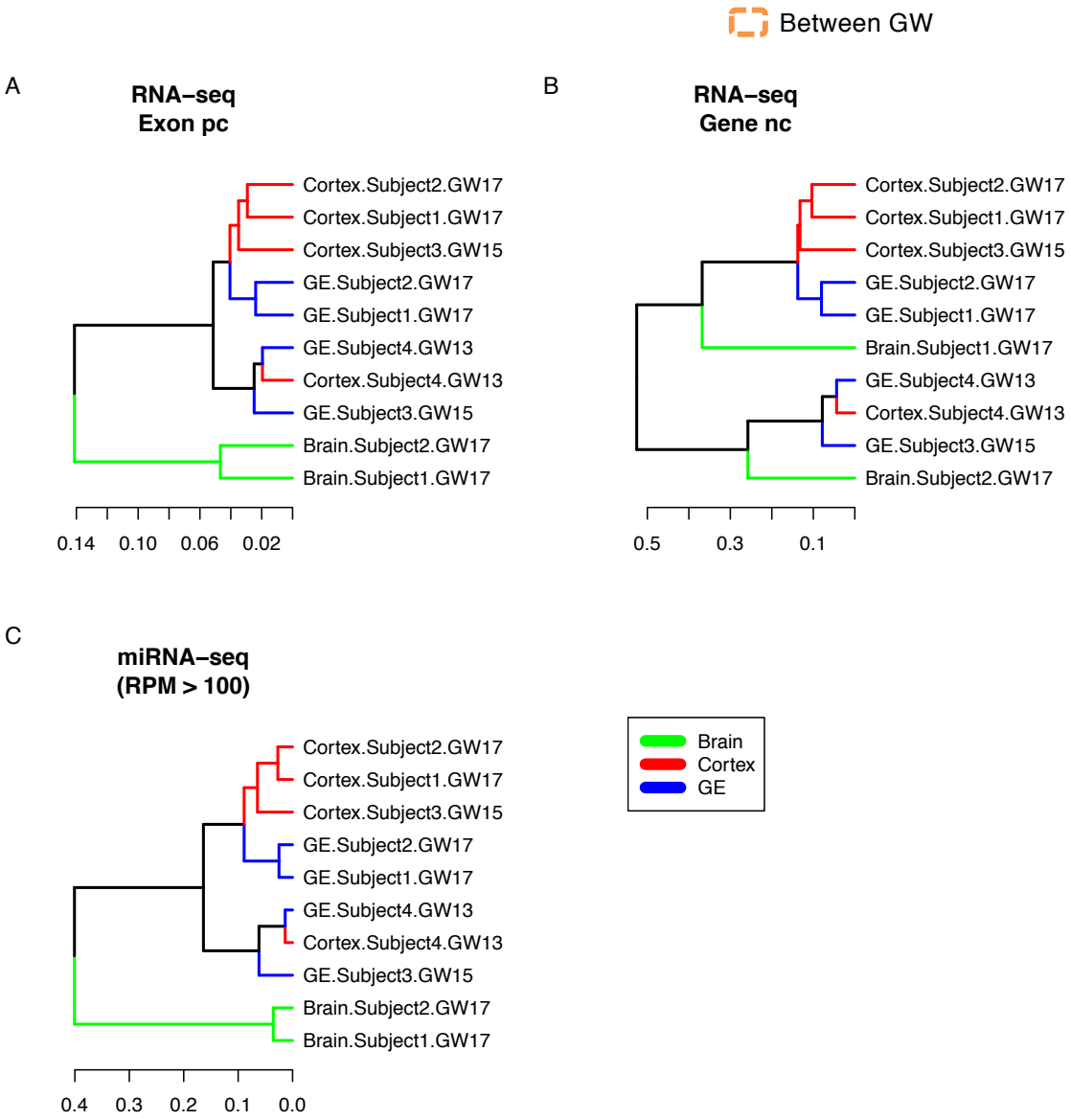

Figure 16. Hierarchical clustering of NPC transcriptomes

Non-supervised clustering of expression levels of protein-coding exons (A), non-coding genes (B), and highly expressed (RPM > 100) microRNAs (C) showed similar clustering for NPCs, with GW13 NPCs clustering together with GE GW15 NPC, while cortex GW15 NPCs clustering with GW17 NPCs. The clusters were generated with spearman correlation.

3.5 Epigenetic and transcriptional differences between monozygotic twins arise before midgestation

The genomes of MZ twins are genetically identical providing a unique opportunity to study developmental processes that give rise to epigenetic variations during normal development within the womb. To investigate epigenetic differences arising between MZ twins during fetal brain development, we analyzed DNA methylation and histone modification data from brain tissue, and cortex and GE derived NPCs of the GW17 MZ twins. MZ twin DMRs were identified through pairwise differential methylation in each tissue and cell type respectively (see methods). In dissected brain tissue and cortex derived NPCs there was significant asymmetry in the orientation of these DMRs with an average of 2-fold more hypomethylated DMRs present in subject 2 compared to subject 1 (**Figure 17A**). As all samples were processed in the same batch using standardized protocols, we considered it unlikely that the asymmetry is a result of sampling bias or technical differences (see methods). To explore the possible functional significance of these asymmetric MZ twin DMRs, we examined their relationship to protein coding genes. On average, we identified 405 DMRs (12% of total) within the promoters of protein-coding genes, a three-fold enrichment compared to that expected by chance (hypergeometric p-value ~0; **Figure 18**). We associated DMRs to genes, and gene enrichment analysis of the resulting gene sets(McLean et al., 2010) revealed an enrichment for Homeobox genes (FDR < 10^{-3}), many of which have been implicated in relevant biological processes such as forebrain regionalization (**Figure 17B** and **Figure 18**). Intersecting DMRs between the MZ twins from all tissue and cell populations, we found 6 regions consistently differentially methylated in brain tissue and both NPCs, including one region at the promoter of *CCDC169* (**Figure 17C**), suggesting individual specific epigenetic signatures between MZ twins that are common across cell types as early as midgestation.

Gene expression analysis revealed an average of 470 differentially expressed genes in pairwise comparisons across twins in the same cell types enriched in pathways involved in neurogenesis and brain development (**Figure 18**). 108 differentially expressed genes were also found to be differentially marked by H3K4me3 with 56% showed expected orientation; 191 by H3K27me3, 60% showed expected orientation. Genes differentially methylated in their promoters and differentially expressed were highly enriched in brain developmental processes (**Figure 17D**), including Oligodendrocyte Transcription Factor 1 (*OLIG1*) and Forkhead Box O1 (*FOXO1*), as well as members of *WNT* pathway such as Secreted Frizzled-Related Protein family (*SFRP1* and *SFRP2*), Wingless-type MMTV Integration Site family (*WNT3* and *WNT7A*). The *WNT* signalling pathway is critical in maintaining normal neurogenesis in developing brain(Malaterre et al., 2007) and its deregulation is important for promoting glioma cell proliferation (Sandberg et al., 2013). Intersecting differentially expressed gene between the MZ twins in different tissue and cell types found B-Cell CLL/Lymphoma 6 (*BCL6*) consistently higher expressed in Subject2 in brain tissue and both NPCs. *BCL6* has been shown to control neurogenesis through epigenetic repression of Notch target genes (Tiberi et al., 2012).

In summary, our findings suggest that epigenetic differences between MZ twins arise as early as GW17 in fetal brain tissue, and that there are individual specific signatures that are consistent across different cell types. Furthermore, these epigenetic differences are associated with differential expression of genes critical to brain development.

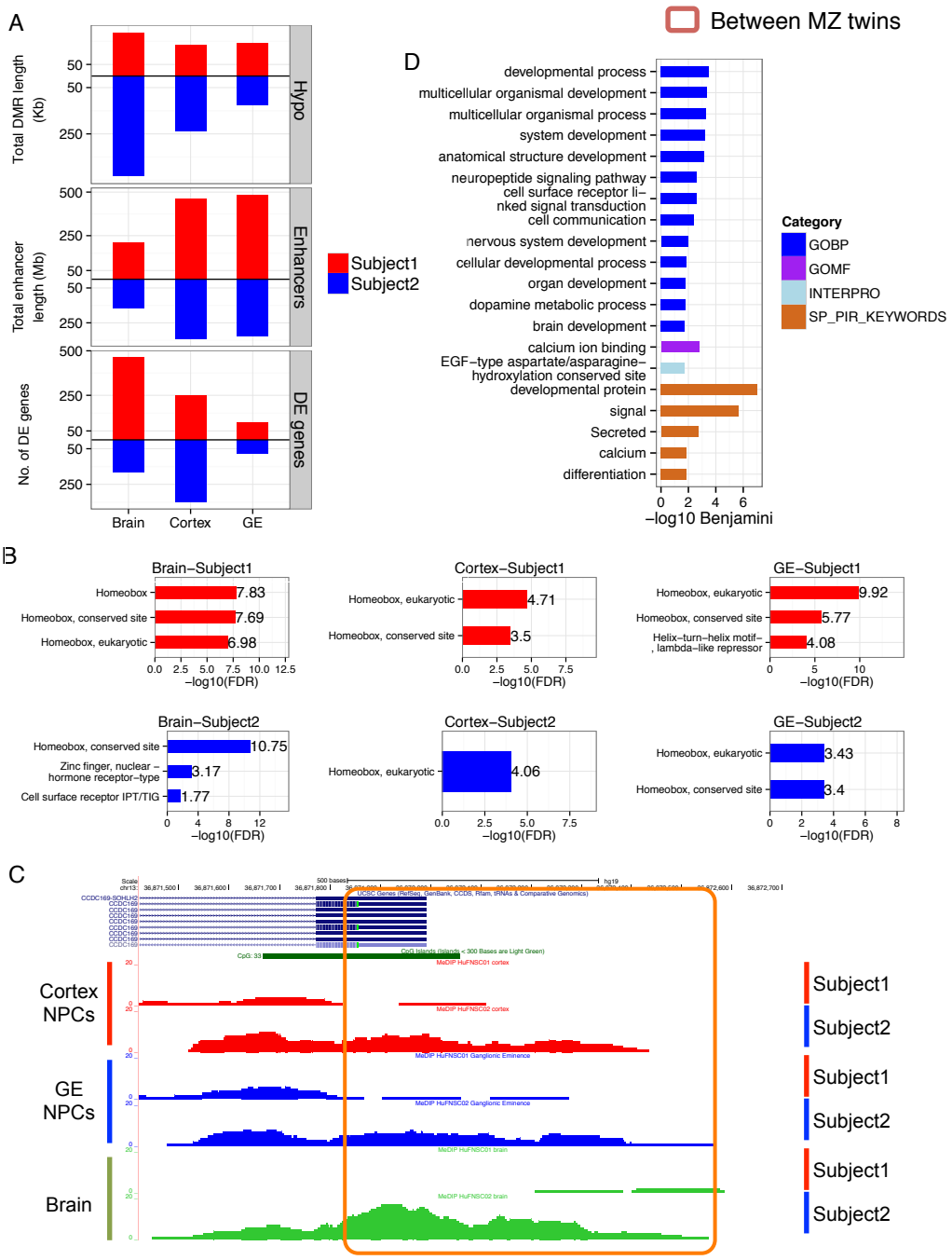

Figure 17. Epigenetic and transcriptional differences between MZ twins arise early in fetal brain development

(A). Total DMR length (top panel), number of enriched bases in enhancers (middle panel), and number of differential expressed genes (bottom panel) between MZ twins. Subject1-specific hypomethylated DMRs/enhancers and up-regulated genes were shown in red, and Subject2-specific ones were shown in blue. (B). DMRs between MZ twins in three tissue and cell populations from GREAT (McLean et al., 2010) functional enrichment analysis all showed

enrichment in Homeobox protein domains. (C). UCSC genome browser MeDIP-seq tracks showed hypermethylation in Subject2 compared to Subject1 at the promoter region of CCDC169 in dissected brain tissue as well as cortex and GE derived NPCs. (D). DAVID (Huang et al., 2009) functional enrichment analysis for genes differentially methylated in their promoters and differentially expressed between the MZ twins showed enrichment in brain developmental processes.

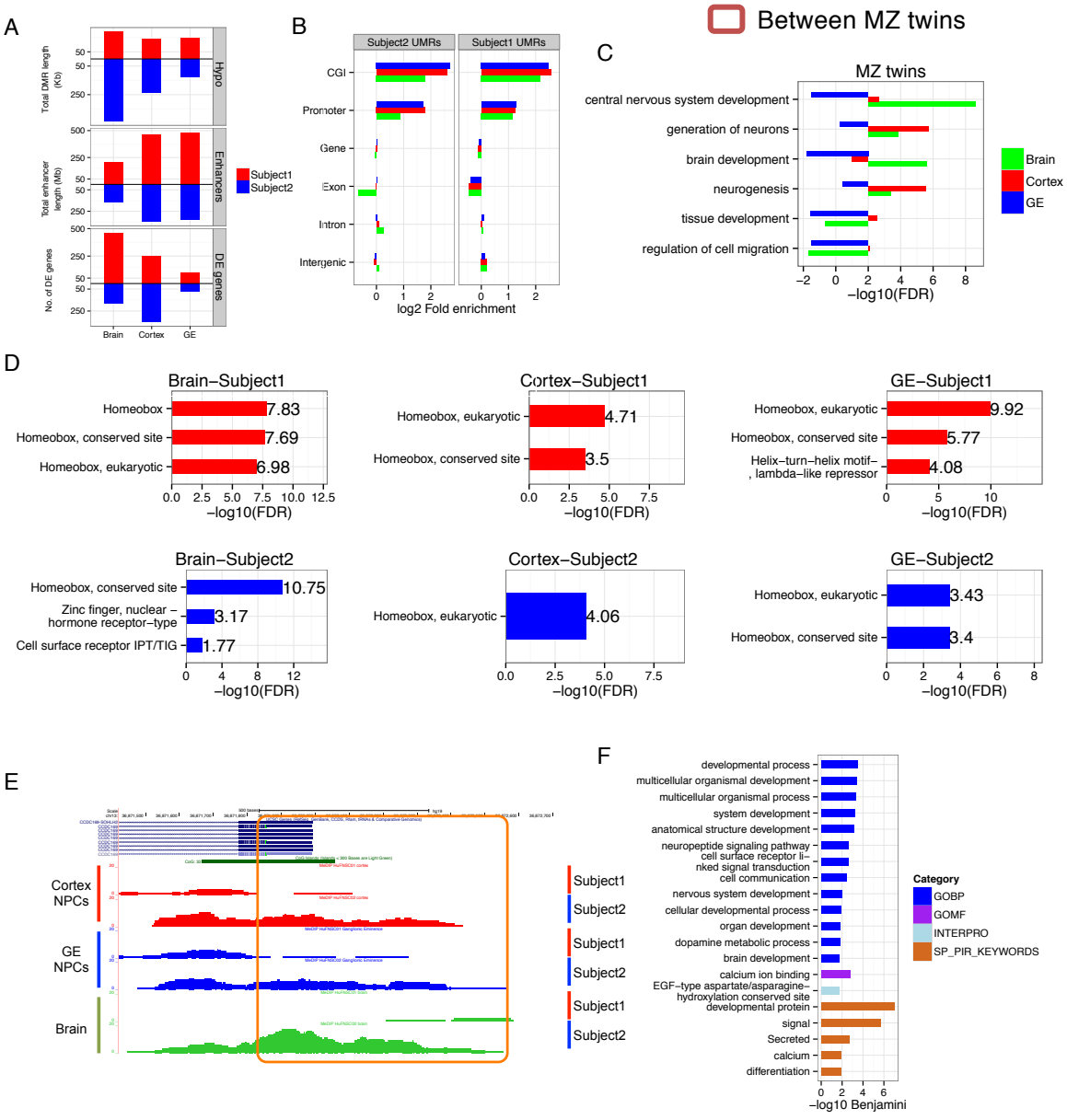

Figure 18. Differences in epigenomes and transcriptomes of MZ twins

(A). Total DMR length (top panel), number of enriched bases in enhancers (middle panel), and number of differential expressed genes (bottom panel) between MZ twins. Subject1-specific hypomethylated DMRs/enhancers and up-regulated genes were shown in red, and Subject2-specific ones were shown in blue. (B). UMR enrichment in genomic features. (C). DAVID functional enrichment of genes differentially expressed between MZ twins. (D). DMRs between MZ twins in three tissue and cell populations from GREAT (McLean et al., 2010) functional enrichment analysis all showed enrichment in Homeobox protein domains. (E). UCSC genome browser MeDIP-seq tracks showed hypermethylation in Subject2 compared to Subject1 at the promoter region of CCDC169 in dissected brain tissue as well as cortex and GE derived NPCs. (F). DAVID (Huang et al., 2009) functional enrichment analysis for genes differentially methylated in their promoters and differentially expressed between the MZ twins showed enrichment in brain developmental processes.

Chapter 4: Discussion

Reference epigenome mapping efforts (Kundaje et al., 2015; Stunnenberg and Hirst, 2016) have largely focused on hematopoietic and epithelial cell types, and there remains a significant gap in our understanding of difficult-to-obtain tissues and cells including those within the human brain. My analysis presented in this thesis provides a valuable resource of epigenomic and transcriptomic signatures and derived regulatory networks of the developing human fetal brain and comprehensive annotation of epigenetic and transcriptional differences between MZ twins, NPCs and different developmental stages in early fetal brain development.

The early fetal period (GW8-GW18) is characterized by a dynamic neurogenesis phase involving proliferation and migration of NPCs (Florio and Huttner, 2014) and significant effort has been directed at the characterization of the epigenetic and transcriptional dynamics of NPCs (Llorens-Bobadilla et al., 2015; Miller et al., 2014; Ziller et al., 2014). Neural migration patterns have not been completely deciphered in the developing human brain, but murine NPCs derived from different brain regions differentiate into neurons and glial cells that colonize distinct brain regions (Kriegstein and Noctor, 2004; Nadarajah and Parnavelas, 2002). This cellular migration is strictly regulated, including by factors involved in cell-cell adhesion, cell cycle control, and interaction with extracellular matrix (Gressens, 2000). We compared the epigenetic states of NPCs derived from two distinct brain regions and identified DNA methylation signatures that could be functionally associated with cell migration patterns specific to these brain regions. Furthermore, tissue specific comparisons suggest that cortex derived NPCs may be at a more epigenetically and transcriptionally active state compared to GE derived NPCs. Taking into account temporal analysis of NPCs which suggests an increase in active epigenetic and transcription states with

developmental stage, the more active epigenetic state of the cortex derived NPCs suggests that cortex derived NPCs are developmentally more advanced than GE derived NPCs, consistent with the more neuronal phenotype that we observed in cortex derived NPCs.

Targeted CpG methylation array studies have demonstrated that the genomes of human fetal brain tissue undergo DNA hypomethylation with age (Spiers et al., 2015), and that the bulk of temporal transcriptional changes occur at prenatal stages with the activation of critical regulators during early-mid fetal development (Kang et al., 2011). Our analysis confirmed a developmentally associated hypomethylation between GW13-17 and extended this to show that this hypomethylation, in particular in enhancer regions, was associated with increased transcription factor binding activities of critical transcription factors, such as *OLIG2*. The activation of these transcription factors may lead to activation of genes critical for axon guidance and cell migration, suggesting an essential role of epigenetic regulation in increased cell migration activities during this critical developmental period. Furthermore, by comparing the transcription profiles in two distinct NPC populations, we found that the transcriptional activation occurred in waves, and at different developmental stages in NPCs derived from different brain regions. This pattern supports a model where NPCs from both brain regions share a common epigenetic state, possibly inherited from a common progenitor cell, and that cortex derived NPCs diverged earlier and became more differentiated than GE derived NPCs, providing additional insight into the spatio-temporal developmental trajectory within the fetal brain.

Post-natal morphological differences in the gross anatomy of the brain tissue of human MZ twins are well documented (Biondi et al., 1998; Cavanna et al., 2010; Steinmetz et al., 1995). Previous studies on post-natal MZ twins suggested that the epigenetic differences between the

twins arise primarily through exposure to divergent environments(Lévesque et al., 2014), and may contribute to phenotypical discordance in development and disease onset(Galetzka et al., 2012; Singh et al., 2002; Townsend et al., 2005; Weksberg et al., 2002). However, these studies were unable to find distinct epigenetic signatures at the age of three due to technical limitations (Fraga et al., 2005). Whether or when these differences arise in pre-natal MZ twins and their consequences are still poorly understood. Our analysis of the epigenomes of pre-natal MZ twins provides evidence of epigenetic differences as early as GW17 that are associated with the activities of master regulators of brain development such as the WNT family of proteins.

In summary, my analysis provides an unprecedented annotation of epigenetic and transcriptional states of early fetal brain development. My analysis across developmental stages supports a model of epigenetic and transcriptional activation with developmental age, and suggested that cortex derived NPCs may be at a more advanced epigenetic state than GE derived NPCs. My results reveal epigenetic differences between MZ twins during early fetal periods. I provide a high-resolution regulatory network for NPCs from different brain regions, providing a comprehensive reference for future studies in brain development and disease.

Bibliography

Acampora, D., Barone, P., and Simeone, A. (1999). Otx Genes in Corticogenesis and Brain Development. Cerebral Cortex *9*, 533-542.

Adithya, S.P., Balagangadharan, K., and Selvamurugan, N. (2022). Epigenetic modifications of histones during osteoblast differentiation. Biochim Biophys Acta Gene Regul Mech *1865*, 194780.

Al Adhami, H., Bardet, A.F., Dumas, M., Cleroux, E., Guibert, S., Fauque, P., Acloque, H., and Weber, M. (2022). A comparative methylome analysis reveals conservation and divergence of DNA methylation patterns and functions in vertebrates. BMC Biol *20*, 70.

Amiri, A., Coppola, G., Scuderi, S., Wu, F., Roychowdhury, T., Liu, F., Pochareddy, S., Shin, Y., Safi, A., Song, L.*, et al.* (2018). Transcriptome and epigenome landscape of human cortical development modeled in organoids. Science *362*.

Banaszynski, L.A., Wen, D., Dewell, S., Whitcomb, S.J., Lin, M., Diaz, N., Elsasser, S.J., Chapgier, A., Goldberg, A.D., Canaani, E.*, et al.* (2013). Hira-dependent histone H3.3 deposition facilitates PRC2 recruitment at developmental loci in ES cells. Cell *155*, 107-120.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res *21*, 381-395.

Bannister, A.J., Schneider, R., Myers, F.A., Thorne, A.W., Crane-Robinson, C., and Kouzarides, T. (2005). Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. J Biol Chem *280*, 17732-17736.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-124.

Beck, S. (2010). Taking the measure of the methylome. Nature Biotechnology *28*, 1026-1028.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K.*, et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

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

Bhutani, N., Burns, D.M., and Blau, H.M. (2011). DNA demethylation dynamics. Cell *146*, 866-872.

Biondi, A., Nogueira, H., Dormont, D., Duyme, M., Hasboun, D., Zouaoui, A., Chantôme, M., and Marsault, C. (1998). Are the brains of monozygotic twins similar? A three-dimensional MR study. AJNR Am J Neuroradiol *19*, 1361-1367.

Bird, A., Taggart, M., Frommer, M., Miller, O.J., and Macleod, D. (1985). A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. Cell *40*, 91-99.

Black, J.C., Van Rechem, C., and Whetstine, J.R. (2012). Histone lysine methylation dynamics: establishment, regulation, and biological impact. Mol Cell *48*, 491-507.

Blakely, B.D., Bye, C.R., Fernando, C.V., Horne, M.K., Macheda, M.L., Stacker, S.A., Arenas, E., and Parish, C.L. (2011). Wnt5a regulates midbrain dopaminergic axon growth and guidance. PLoS One *6*, e18373.

Bouvier, C., Bartoli, C., Aguirre-Cruz, L., Virard, I., Colin, C., Fernandez, C., Gouvernet, J., and Figarella-Branger, D. (2003). Shared oligodendrocyte lineage gene expression in gliomas and oligodendrocyte progenitor cells. J Neurosurg *99*, 344-350.

Boyes, J., and Bird, A. (1991a). DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. Cell *64*, 1123-1134.
Boyes, J., and Bird, A. (1991b). DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. Cell *64*, 1123-1134.
Burgold, T., Spreafico, F., De Santa, F., Totaro, M.G., Prosperini, E., Natoli, G., and Testa, G. (2008). The histone H3 lysine 27-specific demethylase Jmjd3 is required for neural commitment. PLoS One *3*, e3034.
Butterfield, Y.S., Kreitzman, M., Thiessen, N., Corbett, R.D., Li, Y., Pang, J., Ma, Y.P., Jones, S.J., and Birol, I. (2014). JAGuaR: junction alignments to genome for RNA-seq reads. PLoS One *9*, e102398.
Cacci, E., Negri, R., Biagioni, S., and Lupo, G. (2017). Histone Methylation and microRNA-dependent Regulation of Epigenetic Activities in Neural Progenitor Self-Renewal and Differentiation. Curr Top Med Chem *17*, 794-807.
Cai, J., Chen, Y., Cai, W.-H., Hurlock, E.C., Wu, H., Kernie, S.G., Parada, L.F., and Lu, Q.R. (2007). A crucial role for Olig2 in white matter astrocyte development. Development *134*, 1887-1899.
Campbell, C.E., Piper, M., Plachez, C., Yeh, Y.-T., Baizer, J.S., Osinski, J.M., Litwack, E.D., Richards, L.J., and Gronostajski, R.M. (2008). The transcription factor Nfix is essential for normal brain development. BMC Dev Biol *8*, 52.
Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039-1043.
Cavalli, G., and Heard, E. (2019). Advances in epigenetics link genetics to the environment and disease. Nature *571*, 489-499.
Cavanna, A.E., Stecco, A., Rickards, H., Servo, S., Terazzi, E., Peterson, B., Robertson, M.M., Carriero, A., and Monaco, F. (2010). Corpus callosum abnormalities in Tourette syndrome: an MRI-DTI study of monozygotic twins. J Neurol Neurosurg Psychiatry *81*, 533-535.
Charlton, J., Jung, E.J., Mattei, A.L., Bailly, N., Liao, J., Martin, E.J., Giesselmann, P., Brandl, B., Stamenova, E.K., Muller, F.J.*, et al.* (2020). TETs compete with DNMT3 activity in pluripotent cells at thousands of methylated somatic enhancers. Nat Genet *52*, 819-827.
Chattopadhyaya, S., and Ghosal, S. (2022). DNA methylation: a saga of genome maintenance in hematological perspective. Hum Cell *35*, 448-461.
Coward, W.R., Brand, O.J., Pasini, A., Jenkins, G., Knox, A.J., and Pang, L. (2018). Interplay between EZH2 and G9a Regulates CXCL10 Gene Repression in Idiopathic Pulmonary Fibrosis. Am J Respir Cell Mol Biol *58*, 449-460.
Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., and Sharp, P.A. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences *107*, 21931-21936.
Darmanis, S., Sloan, S.A., Zhang, Y., Enge, M., Caneda, C., Shuer, L.M., Gephart, M.G.H., Barres, B.A., and Quake, S.R. (2015). A survey of human brain transcriptome diversity at the single cell level. Proceedings of the National Academy of Sciences *112*, 7285-7290.
Desai, D., Khanna, A., and Pethe, P. (2020). PRC1 catalytic unit RING1B regulates early neural differentiation of human pluripotent stem cells. Exp Cell Res *396*, 112294.

Detilleux, D., Spill, Y.G., Balaramane, D., Weber, M., and Bardet, A.F. (2022). Pan-cancer predictions of transcription factors mediating aberrant DNA methylation. Epigenet Chromatin *15*.
Egea, J., and Klein, R. (2007). Bidirectional Eph–ephrin signaling during axon guidance. Trends in Cell Biology *17*, 230-238.
EHRLICH, M., GAMASOSA, M.A., HUANG, L.H., MIDGETT, R.M., KUO, K.C., MCCUNE, R.A., and GEHRKE, C. (1982). AMOUNT AND DISTRIBUTION OF 5-METHYLCYTOSINE IN HUMAN DNA FROM DIFFERENT TYPES OF TISSUES OR CELLS. NUCLEIC ACIDS RESEARCH *10*, 2709-2721.
Eze, U.C., Bhaduri, A., Haeussler, M., Nowakowski, T.J., and Kriegstein, A.R. (2021). Single-cell atlas of early human brain development highlights heterogeneity of human neuroepithelial cells and early radial glia. Nat Neurosci *24*, 584-594.
Fejes, A.P., Robertson, G., Bilenky, M., Varhol, R., Bainbridge, M., and Jones, S.J. (2008). FindPeaks 3.1: a tool for identifying areas of enrichment from massively parallel short-read sequencing technology. Bioinformatics *24*, 1729-1730.
Feldmann, A., Ivanek, R., Murr, R., Gaidatzis, D., Burger, L., and Schübeler, D. (2013). Transcription factor occupancy can mediate active turnover of DNA methylation at regulatory regions. PLoS Genet *9*, e1003994.
Figueroa, M.E., Abdel-Wahab, O., Lu, C., Ward, P.S., Patel, J., Shih, A., Li, Y., Bhagwat, N., Vasanthakumar, A., Fernandez, H.F*., et al.* (2010). Leukemic {IDH1} and {IDH2} Mutations Result in a Hypermethylation Phenotype, Disrupt {TET2} Function, and Impair Hematopoietic Differentiation. Cancer Cell *18*, 553 - 567.
Florio, M., and Huttner, W.B. (2014). Neural progenitors, neurogenesis and the evolution of the neocortex. Development *141*, 2182-2194.
Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suñer, D., Cigudosa, J.C., Urioste, M., Benitez, J*., et al.* (2005). Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci U S A *102*, 10604-10609.
Galetzka, D., Hansmann, T., El Hajj, N., Weis, E., Irmscher, B., Ludwig, M., Schneider-Rätzke, B., Kohlschmidt, N., Beyer, V., Bartsch, O*., et al.* (2012). Monozygotic twins discordant for constitutive BRCA1 promoter methylation, childhood cancer and secondary cancer. Epigenetics *7*, 47-54.
Gansen, A., Toth, K., Schwarz, N., and Langowski, J. (2015). Opposing roles of H3- and H4-acetylation in the regulation of nucleosome structure--a FRET study. Nucleic Acids Res *43*, 1433-1443.
Gascard, P., Bilenky, M., Sigaroudinia, M., Zhao, J., Li, L., Carles, A., Delaney, A., Tam, A., Kamoh, B., Cho, S*., et al.* (2015). Epigenetic and transcriptional determinants of the human breast. Nat Commun *6*, 6351.
Geng, H., Chen, H., Wang, H., and Wang, L. (2021). The Histone Modifications of Neuronal Plasticity. Neural Plast *2021*, 6690523.
Gimenez, J., Montgiraud, C., Pichon, J.-P., Bonnaud, B., Arsac, M., Ruel, K., Bouton, O., and Mallet, F. (2010). Custom human endogenous retroviruses dedicated microarray identifies self-induced HERV-W family elements reactivated in testicular cancer upon methylation control. Nucleic Acids Res *38*, 2229-2246.
Ginno, P.A., Gaidatzis, D., Feldmann, A., Hoerner, L., Imanci, D., Burger, L., Zilbermann, F., Peters, A., Edenhofer, F., Smallwood, S.A*., et al.* (2020). A genome-scale map of DNA

methylation turnover identifies site-specific dependencies of DNMT and TET activity. Nat Commun *11*, 2680.
Götz, M., and Huttner, W.B. (2005). The cell biology of neurogenesis. Nat Rev Mol Cell Biol *6*, 777-788.
Gressens, P. (2000). Mechanisms and Disturbances of Neuronal Migration. Pediatr Res *48*, 725-730.
Gu, P., Xu, X., Le Menuet, D., Chung, A.C.-K., and Cooney, A.J. (2011). Differential Recruitment of Methyl CpG-Binding Domain Factors and DNA Methyltransferases by the Orphan Receptor Germ Cell Nuclear Factor Initiates the Repression and Silencing of Oct4. STEM CELLS *29*, 1041-1051.
Guan, X., Chen, X., Dai, L., Ma, J., Zhang, Q., Qu, S., Bai, Y., and Wang, Y. (2021). Low Maternal Dietary Intake of Choline Regulates Toll-Like Receptor 4 Expression Via Histone H3K27me3 in Fetal Mouse Neural Progenitor Cells. Mol Nutr Food Res *65*, e2000769.
Guo, H.-B., and Guo, H. (2007). Mechanism of histone methylation catalyzed by protein lysine methyltransferase SET7/9 and origin of product specificity. Proceedings of the National Academy of Sciences *104*, 8797-8802.
Hallonet, M., Hollemann, T., Pieler, T., and Gruss, P. (1999). Vax1, a novel homeobox-containing gene, directs development of the basal forebrain and visual system. Genes Dev *13*, 3106-3114.
Hamasaki, T., Goto, S., Nishikawa, S., and Ushio, Y. (2001). A role of netrin-1 in the formation of the subcortical structure striatum: repulsive action on the migration of late-born striatal neurons. J Neurosci *21*, 4272-4280.
Hansen, D.V., Lui, J.H., Parker, P.R.L., and Kriegstein, A.R. (2010). Neurogenic radial glia in the outer subventricular zone of human neocortex. Nature *464*, 554-561.
He, H., Hu, Z., Xiao, H., Zhou, F., and Yang, B. (2018). The tale of histone modifications and its role in multiple sclerosis. Hum Genomics *12*, 31.
He, L., Huang, H., Bradai, M., Zhao, C., You, Y., Ma, J., Zhao, L., Lozano-Duran, R., and Zhu, J.K. (2022). DNA methylation-free Arabidopsis reveals crucial roles of DNA methylation in regulating gene expression and development. Nat Commun *13*, 1335.
Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A*., et al.* (2007a). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet *39*, 311-318.
Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A*., et al.* (2007b). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet *39*, 311-318.
Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.
Hervouet, E., Vallette, F.M., and Cartron, P.-F. (2009). Dnmt3/transcription factor interactions as crucial players in targeted DNa methylation. Epigenetics *4*, 487-499.
Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-Andre, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. Cell *155*, 934-947.
Hon, G.C., Hawkins, R.D., and Ren, B. (2009). Predictive chromatin signatures in the mammalian genome. Hum Mol Genet *18*, R195-201.

Howard, G., Eiges, R., Gaudet, F., Jaenisch, R., and Eden, A. (2008). Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. Oncogene *27*, 404-408.
Hu, H. (1999). Chemorepulsion of Neuronal Migration by Slit2 in the Developing Mammalian Forebrain. Neuron *23*, 703-711.
Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.
Huang, Y., Fang, J., Bedford, M.T., Zhang, Y., and Xu, R.M. (2006). Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. Science *312*, 748-751.
Hunter, C.S., and Rhodes, S.J. (2005). LIM-homeodomain genes in mammalian development and human disease. Mol Biol Rep *32*, 67-77.
Jakovcevski, I., and Zecevic, N. (2005). Olig transcription factors are expressed in oligodendrocyte and neuronal cells in human fetal CNS. J Neurosci *25*, 10064-10073.
Jiang, J., Lu, J., Lu, D., Liang, Z., Li, L., Ouyang, S., Kong, X., Jiang, H., Shen, B., and Luo, C. (2012). Investigation of the acetylation mechanism by GCN5 histone acetyltransferase. PloS one *7*, e36660.
Jiang, T., Zhou, X., Taghizadeh, K., Dong, M., and Dedon, P.C. (2007). N-formylation of lysine in histone proteins as a secondary modification arising from oxidative DNA damage. Proceedings of the National Academy of Sciences *104*, 60-65.
Kang, H.J., Kawasawa, Y.I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A.M.M., Pletikos, M., Meyer, K.A., Sedmak, G.*, et al.* (2011). Spatio-temporal transcriptome of the human brain. Nature *478*, 483-489.
Kennedy, T.E. (2000). Cellular mechanisms of netrin function: long-range and short-range actions. Biochem Cell Biol *78*, 569-575.
Khan, A., and Zhang, X. (2016). dbSUPER: a database of super-enhancers in mouse and human genome. Nucleic Acids Res *44*, D164-171.
Kinkley, S., Helmuth, J., Polansky, J.K., Dunkel, I., Gasparoni, G., Frohler, S., Chen, W., Walter, J., Hamann, A., and Chung, H.R. (2016). reChIP-seq reveals widespread bivalency of H3K4me3 and H3K27me3 in CD4(+) memory T cells. Nat Commun *7*, 12514.
Klose, R.J., and Bird, A.P. (2006). Genomic {DNA} methylation: the mark and its mediators. Trends in Biochemical Sciences *31*, 89 - 97.
Koo, B., Lee, K.H., Ming, G.L., Yoon, K.J., and Song, H. (2022). Setting the clock of neural progenitor cells during mammalian corticogenesis. Semin Cell Dev Biol.
Kremsky, I., and Corces, V.G. (2020). Protection from DNA re-methylation by transcription factors in primordial germ cells and pre-implantation embryos can explain trans-generational epigenetic inheritance. Genome Biology *21*.
Kriegstein, A.R., and Noctor, S.C. (2004). Patterns of neuronal migration in the embryonic cortex. Trends Neurosci *27*, 392-399.
Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics *27*, 1571-1572.
Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M.J.*, et al.* (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317-330.
Kurdistani, S.K., Tavazoie, S., and Grunstein, M. (2004). Mapping global histone acetylation patterns to gene expression. Cell *117*, 721-733.

La Manno, G., Siletti, K., Furlan, A., Gyllborg, D., Vinsland, E., Mossi Albiach, A., Mattsson Langseth, C., Khven, I., Lederer, A.R., Dratva, L.M., *et al.* (2021). Molecular architecture of the developing mouse brain. Nature *596*, 92-96.
Lamprecht, B., Walter, K., Kreher, S., Kumar, R., Hummel, M., Lenze, D., Köchert, K., Bouhlel, M.A., Richter, J., Soler, E., *et al.* (2010). Derepression of an endogenous long terminal repeat activates the CSF1R proto-oncogene in human lymphoma. Nat Med *16*, 571-579, 571p following 579.
Lavdas, A.A., Grigoriou, M., Pachnis, V., and Parnavelas, J.G. (1999). The medial ganglionic eminence gives rise to a population of early neurons in the developing cerebral cortex. J Neurosci *19*, 7881-7888.
Lévesque, M.L., Casey, K.F., Szyf, M., Ismaylova, E., Ly, V., Verner, M.-P., Suderman, M., Brendgen, M., Vitaro, F., Dionne, G., *et al.* (2014). Genome-wide DNA methylation variability in adolescent monozygotic twins followed since birth. Epigenetics *9*, 1410-1421.
Li, D., Zhang, B., Xing, X., and Wang, T. (2015). Combining MeDIP-seq and MRE-seq to investigate genome-wide CpG methylation. Methods *72*, 29-40.
Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics *27*, 2987-2993.
Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics *26*, 589-595.
Li, K., Liu, C., Zhou, B., Bi, L., Huang, H., Lin, T., and Xu, K. (2013). Role of EZH2 in the Growth of Prostate Cancer Stem Cells Isolated from LNCaP Cells. Int J Mol Sci *14*, 11981-11993.
Li, S., Peng, Y., Landsman, D., and Panchenko, A.R. (2022). DNA methylation cues in nucleosome geometry, stability and unwrapping. Nucleic Acids Res *50*, 1864-1874.
Lister, R., Mukamel, E.A., Nery, J.R., Urich, M., Puddifoot, C.A., Johnson, N.D., Lucero, J., Huang, Y., Dwork, A.J., Schultz, M.D., *et al.* (2013). Global epigenomic reconfiguration during mammalian brain development. Science *341*, 1237905.
Llorens-Bobadilla, E., Zhao, S., Baser, A., Saiz-Castro, G., Zwadlo, K., and Martin-Villalba, A. (2015). Single-Cell Transcriptomics Reveals a Population of Dormant Neural Stem Cells that Become Activated upon Brain Injury. Cell Stem Cell *17*, 329-340.
Local, A., Huang, H., Albuquerque, C.P., Singh, N., Lee, A.Y., Wang, W., Wang, C., Hsia, J.E., Shiau, A.K., Ge, K., *et al.* (2018). Identification of H3K4me1-associated proteins at mammalian enhancers. Nat Genet *50*, 73-82.
Lohse, B., Helgstrand, C., Kristensen, J.B., Leurs, U., Cloos, P.A., Kristensen, J.L., and Clausen, R.P. (2013). Posttranslational Modifications of the Histone 3 Tail and Their Impact on the Activity of Histone Lysine Demethylases In Vitro. PloS one *8*, e67653.
Lombardi, P.M., Cole, K.E., Dowling, D.P., and Christianson, D.W. (2011). Structure, mechanism, and inhibition of histone deacetylases and related metalloenzymes. Current opinion in structural biology *21*, 735-743.
Love, I.M., Sekaric, P., Shi, D., Grossman, S.R., and Androphy, E.J. (2012). The histone acetyltransferase PCAF regulates p21 transcription through stress-induced acetylation of histone H3. Cell Cycle *11*, 2458-2466.
Loven, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell *153*, 320-334.

Lu, Q.R., Park, J.K., Noll, E., Chan, J.A., Alberta, J., Yuk, D., Alzamora, M.G., Louis, D.N., Stiles, C.D., Rowitch, D.H.*, et al.* (2001). Oligodendrocyte lineage genes (OLIG) as molecular markers for human glial brain tumors. Proc Natl Acad Sci U S A *98*, 10851-10856.
Luo, Y., Coskun, V., Liang, A., Yu, J., Cheng, L., Ge, W., Shi, Z., Zhang, K., Li, C., Cui, Y.*, et al.* (2015). Single-cell transcriptome analyses reveal signals to activate dormant neural stem cells. Cell *161*, 1175-1186.
Maiques-Diaz, A., Chou, F.S., Wunderlich, M., Gomez-Lopez, G., Jacinto, F.V., Rodriguez-Perales, S., Larrayoz, M.J., Calasanz, M.J., Mulloy, J.C., Cigudosa, J.C.*, et al.* (2012). Chromatin modifications induced by the AML1-ETO fusion protein reversibly silence its genomic targets through AML1 and Sp1 binding motifs. Leukemia *26*, 1329-1337.
Malaterre, J., Ramsay, R.G., and Mantamadiotis, T. (2007). Wnt-Frizzled signalling and the many paths to neural development and adult brain homeostasis. Front Biosci *12*, 492-506.
Mason, S., Piper, M., Gronostajski, R.M., and Richards, L.J. (2009). Nuclear factor one transcription factors in CNS development. Mol Neurobiol *39*, 10-23.
Mattei, A.L., Bailly, N., and Meissner, A. (2022). DNA methylation: a historical perspective. Trends Genet *38*, 676-707.
McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol *28*, 495-501.
Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P.*, et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-560.
Miller, J.A., Ding, S.-L., Sunkin, S.M., Smith, K.A., Ng, L., Szafer, A., Ebbert, A., Riley, Z.L., Royall, J.J., Aiona, K.*, et al.* (2014). Transcriptional landscape of the prenatal human brain. Nature *508*, 199.
Miyajima, Y., Noguchi, S., Tanaka, Y., Li, J.R., Nishimura, H., Kishima, M., Lim, J., Furuhata, E., Suzuki, T., Kasukawa, T.*, et al.* (2022). Prediction of transcription factors associated with DNA demethylation during human cellular development. Chromosome Res *30*, 109-121.
Moltrasio, C., Romagnuolo, M., and Marzano, A.V. (2022). Epigenetic Mechanisms of Epidermal Differentiation. Int J Mol Sci *23*.
Morin, R., Bainbridge, M., Fejes, A., Hirst, M., Krzywinski, M., Pugh, T., McDonald, H., Varhol, R., Jones, S., and Marra, M. (2008). Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. Biotechniques *45*, 81-94.
Morin, R.D., Johnson, N.A., Severson, T.M., Mungall, A.J., An, J., Goya, R., Paul, J.E., Boyle, M., Woolcock, B.W., Kuchenbauer, F.*, et al.* (2010). Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nat Genet *42*, 181-185.
Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods *5*, 621-628.
Nadarajah, B., and Parnavelas, J.G. (2002). Modes of neuronal migration in the developing cerebral cortex. Nat Rev Neurosci *3*, 423-432.
Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature *393*, 386-389.

Numata, S., Ye, T., Hyde, T.M., Guitart-Navarro, X., Tao, R., Wininger, M., Colantuoni, C., Weinberger, D.R., Kleinman, J.E., and Lipska, B.K. (2012). DNA methylation signatures in development and aging of the human prefrontal cortex. Am J Hum Genet *90*, 260-272.
Peng, V., Xing, X.Y., Bando, J.K., Trsan, T., Di Luccia, B., Collins, P.L., Li, D.F., Wang, W.L., Lee, H.J., Oltz, E.M.*, et al.* (2022). Whole-genome profiling of DNA methylation and hydroxymethylation identifies distinct regulatory programs among innate lymphocytes. Nat Immunol *23*, 619-+.
Pinheiro, I., Margueron, R., Shukeir, N., Eisold, M., Fritzsch, C., Richter, F.M., Mittler, G., Genoud, C., Goyama, S., Kurokawa, M.*, et al.* (2012). Prdm3 and Prdm16 are H3K9me1 methyltransferases required for mammalian heterochromatin integrity. Cell *150*, 948-960.
Prajapati, B., Fatma, M., Maddhesiya, P., Sodhi, M.K., Fatima, M., Dargar, T., Bhagat, R., Seth, P., and Sinha, S. (2019). Identification and epigenetic analysis of divergent long non-coding RNAs in multilineage differentiation of human Neural Progenitor Cells. RNA Biol *16*, 13-24.
Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279-283.
Raineri, E., Dabad, M., and Heath, S. (2014). A note on exact differences between beta distributions in genomic (Methylation) studies. PLoS One *9*, e97349.
Reynolds, B.A., and Weiss, S. (1996). Clonal and population analyses demonstrate that an EGF-responsive mammalian embryonic CNS precursor is a stem cell. Dev Biol *175*, 1-13.
Rivera, F.J., Couillard-Despres, S., Pedre, X., Ploetz, S., Caioni, M., Lois, C., Bogdahn, U., and Aigner, L. (2006). Mesenchymal stem cells instruct oligodendrogenic fate decision on adult neural stem cells. Stem Cells *24*, 2209-2219.
Robertson, K.D. (2005). DNA methylation and human disease. NATURE REVIEWS GENETICS *6*, 597-610.
Rodriguez, D., Gauthier, F., Bertini, E., Bugiani, M., Brenner, M., N'guyen, S., Goizet, C., Gelot, A., Surtees, R., Pedespan, J.M.*, et al.* (2001). Infantile Alexander disease: spectrum of GFAP mutations and genotype-phenotype correlation. Am J Hum Genet *69*, 1134-1140.
Rutka, J.T., and Smith, S.L. (1993). Transfection of human astrocytoma cells with glial fibrillary acidic protein complementary DNA: analysis of expression, proliferation, and tumorigenicity. Cancer Res *53*, 3624-3631.
Salton, M., Voss, T.C., and Misteli, T. (2014). Identification by high-throughput imaging of the histone methyltransferase EHMT2 as an epigenetic regulator of VEGFA alternative splicing. Nucleic Acids Res *42*, 13662-13673.
Sandberg, C.J., Altschuler, G., Jeong, J., Strømme, K.K., Stangeland, B., Murrell, W., Grasmo-Wendler, U.-H., Myklebost, O., Helseth, E., Vik-Mo, E.O.*, et al.* (2013). Comparison of glioma stem cells to neural stem cells from the adult human brain identifies dysregulated Wnt- signaling and a fingerprint associated with clinical outcome. Exp Cell Res *319*, 2230-2243.
Sanulli, S., Justin, N., Teissandier, A., Ancelin, K., Portoso, M., Caron, M., Michaud, A., Lombard, B., da Rocha, S.T., Offer, J.*, et al.* (2015). Jarid2 Methylation via the PRC2 Complex Regulates H3K27me3 Deposition during Cell Differentiation. Mol Cell *57*, 769-783.
Schneider, R., Bannister, A.J., Myers, F.A., Thorne, A.W., Crane-Robinson, C., and Kouzarides, T. (2004). Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. Nat Cell Biol *6*, 73-77.

Schwartzentruber, J., Korshunov, A., Liu, X.Y., Jones, D.T., Pfaff, E., Jacob, K., Sturm, D., Fontebasso, A.M., Quang, D.A., Tonjes, M.*, et al.* (2012). Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature *482*, 226-231.
Shan, C.M., Bao, K., Diedrich, J., Chen, X., Lu, C., Yates, J.R., 3rd, and Jia, S. (2020). The INO80 Complex Regulates Epigenetic Inheritance of Heterochromatin. Cell Rep *33*, 108561.
Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res *13*, 2498-2504.
Shiio, Y., and Eisenman, R.N. (2003). Histone sumoylation is associated with transcriptional repression. Proceedings of the National Academy of Sciences *100*, 13225-13230.
Shimizu, T., and Hibi, M. (2009). Formation and patterning of the forebrain and olfactory system by zinc-finger genes Fezf1 and Fezf2. Dev Growth Differ *51*, 221-231.
Siegmund, K.D., Connor, C.M., Campan, M., Long, T.I., Weisenberger, D.J., Biniszkiewicz, D., Jaenisch, R., Laird, P.W., and Akbarian, S. (2007). DNA methylation in the human cerebral cortex is dynamically regulated throughout the life span and involves differentiated neurons. PLoS One *2*, e895.
Silvestri, C., Narimatsu, M., von Both, I., Liu, Y., Tan, N.B.J., Izzi, L., McCaffery, P., Wrana, J.L., and Attisano, L. (2008). Genome-Wide Identification of Smad/Foxh1 Targets Reveals a Role for Foxh1 in Retinoic Acid Regulation and Forebrain Development. Developmental Cell *14*, 411-423.
Simon, J.A., and Kingston, R.E. (2013). Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. Mol Cell *49*, 808-824.
Sims, R.J., 3rd, Chen, C.F., Santos-Rosa, H., Kouzarides, T., Patel, S.S., and Reinberg, D. (2005). Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. J Biol Chem *280*, 41789-41792.
Sin, H.-S., Huh, J.-W., Kim, D.-S., Kang, D.W., Min, D.S., Kim, T.-H., Ha, H.-S., Kim, H.-H., Lee, S.-Y., and Kim, H.-S. (2006a). Transcriptional control of the HERV-H LTR element of the GSDML gene in human tissues and cancer cells. Arch Virol *151*, 1985-1994.
Sin, H.-S., Huh, J.-W., Kim, D.-S., Kim, T.-H., Ha, H.-S., Kim, W.-Y., Park, H.-K., Kim, C.-M., and Kim, H.-S. (2006b). Endogenous retrovirus-related sequences provide an alternative transcript of MCJ genes in human tissues and cancer cells. Genes Genet Syst *81*, 333-339.
Singh, S.M., Murphy, B., and O'Reilly, R. (2002). Epigenetic contributors to the discordance of monozygotic twins. Clin Genet *62*, 97-103.
Skourti, E., and Dhillon, P. (2022). Cancer epigenetics: promises and pitfalls for cancer therapy. FEBS J *289*, 1156-1159.
Spiers, H., Hannon, E., Schalkwyk, L.C., Smith, R., Wong, C.C.Y., O'Donovan, M.C., Bray, N.J., and Mill, J. (2015). Methylomic trajectories across human fetal brain development. Genome Res *25*, 338-352.
Stadler, M.B., Murr, R., Burger, L., Ivanek, R., Lienert, F., Schöler, A., Wirbelauer, C., Oakeley, E.J., Gaidatzis, D., Tiwari, V.K.*, et al.* (2011). DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature.
Steinmetz, H., Herzog, A., Schlaug, G., Huang, Y., and Jäncke, L. (1995). Brain (A) symmetry in monozygotic twins. Cereb Cortex *5*, 296-300.
Stiles, J., and Jernigan, T.L. (2010). The basics of brain development. Neuropsychol Rev *20*, 327-348.

Stunnenberg, H.G., and Hirst, M. (2016). Cell. In The International Human Epigenome Consortium: A Blueprint for Scientific Collaboration and Discovery (United States), pp. 1897.
Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., *et al.* (2009). Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. Science *324*, 930-935.
Thaler, J.P., Lee, S.-K., Jurata, L.W., Gill, G.N., and Pfaff, S.L. (2002). LIM factor Lhx3 contributes to the specification of motor neuron and interneuron identity through cell-type-specific protein-protein interactions. Cell *110*, 237-249.
Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., *et al.* (2012). The accessible chromatin landscape of the human genome. Nature *489*, 75-82.
Tiberi, L., van den Ameele, J., Dimidschstein, J., Piccirilli, J., Gall, D., Herpoel, A., Bilheu, A., Bonnefont, J., Iacovino, M., Kyba, M., *et al.* (2012). BCL6 controls neurogenesis through Sirt1-dependent epigenetic repression of selective Notch targets. Nature Neuroscience *15*, 1627-1635.
Tompkins, J.D., Hall, C., Chen, V.C., Li, A.X., Wu, X., Hsu, D., Couture, L.A., and Riggs, A.D. (2012). Epigenetic stability, adaptability, and reversibility in human embryonic stem cells. Proc Natl Acad Sci U S A *109*, 12544-12549.
Towbin, B.D., Gonzalez-Aguilera, C., Sack, R., Gaidatzis, D., Kalck, V., Meister, P., Askjaer, P., and Gasser, S.M. (2012). Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. Cell *150*, 934-947.
Townsend, G.C., Richards, L., Hughes, T., Pinkerton, S., and Schwerdt, W. (2005). Epigenetic influences may explain dental differences in monozygotic twin pairs. Aust Dent J *50*, 95-100.
Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S.-i., Sakaue, M., Matsuoka, C., Shimotohno, K., Ishikawa, F., Li, E., Ueda, H.R., *et al.* (2006). Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes to Cells *11*, 805-814.
Ulfig, N. (2002). Ganglionic eminence of the human fetal brain--new vistas. Anat Rec *267*, 191-195.
Unoki, M., Masuda, A., Dohmae, N., Arita, K., Yoshimatsu, M., Iwai, Y., Fukui, Y., Ueda, K., Hamamoto, R., and Shirakawa, M. (2013). Lysyl 5-Hydroxylation, a Novel Histone Modification, by Jumonji Domain Containing 6 (JMJD6). Journal of Biological Chemistry *288*, 6053-6062.
Venkatesh, S., Li, H., Gogol, M.M., and Workman, J.L. (2016). Selective suppression of antisense transcription by Set2-mediated H3K36 methylation. Nat Commun *7*, 13610.
Waddington, C.H. (1942). Canalization of Development and the Inheritance of Acquired Characters. Nature *150*, 563-565.
Wang, H., Huang, Z.Q., Xia, L., Feng, Q., Erdjument-Bromage, H., Strahl, B.D., Briggs, S.D., Allis, C.D., Wong, J., Tempst, P., *et al.* (2001). Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. Science *293*, 853-857.
Wang, Y., Zhang, R., Wu, D., Lu, Z., Sun, W., Cai, Y., Wang, C., and Jin, J. (2013). Epigenetic change in kidney tumor: downregulation of histone acetyltransferase MYST1 in human renal cell carcinoma. J Exp Clin Cancer Res *32*, 8.
Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Peng, W., Zhang, M.Q., *et al.* (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet *40*, 897-903.

Watt, F., and Molloy, P.L. (1988). Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. Genes Dev *2*, 1136-1143.
Wei, Y., Chen, R., Dimicoli, S., Bueso-Ramos, C., Neuberg, D., Pierce, S., Wang, H., Yang, H., Jia, Y., Zheng, H.*, et al.* (2013). Global H3K4me3 genome mapping reveals alterations of innate immunity signaling and overexpression of JMJD3 in human myelodysplastic syndrome CD34+ cells. Leukemia.
Weksberg, R., Shuman, C., Caluseriu, O., Smith, A.C., Fei, Y.-L., Nishikawa, J., Stockley, T.L., Best, L., Chitayat, D., Olney, A.*, et al.* (2002). Discordant KCNQ1OT1 imprinting in sets of monozygotic twins discordant for Beckwith-Wiedemann syndrome. Hum Mol Genet *11*, 1317-1325.
Wiench, M., John, S., Baek, S., Johnson, T.A., Sung, M.-H., Escobar, T., Simmons, C.A., Pearce, K.H., Biddie, S.C., Sabo, P.J.*, et al.* (2011). DNA methylation status predicts cell type-specific enhancer activity. The EMBO Journal *30*, 3028-3039.
Wu, H., and Zhang, Y. (2011). Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. Genes Dev *25*, 2436-2452.
Wu, Q., Shichino, Y., Abe, T., Suetsugu, T., Omori, A., Kiyonari, H., Iwasaki, S., and Matsuzaki, F. (2022). Selective translation of epigenetic modifiers affects the temporal pattern and differentiation of neural stem cells. Nat Commun *13*, 470.
Yamada, T., Fischle, W., Sugiyama, T., Allis, C.D., and Grewal, S.I. (2005). The nucleation and maintenance of heterochromatin by a histone deacetylase in fission yeast. Mol Cell *20*, 173-185.
Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., Das, P.K., Kivioja, T., Dave, K., Zhong, F.*, et al.* (2017). Impact of cytosine methylation on DNA binding specificities of human transcription factors. Science *356*.
Yuan, G., Ma, B., Yuan, W., Zhang, Z., Chen, P., Ding, X., Feng, L., Shen, X., Chen, S., Li, G.*, et al.* (2013). Histone H2A ubiquitination inhibits the enzymatic activity of H3 Lysine 36 methyltransferases. J Biol Chem.
Zegerman, P., Canas, B., Pappin, D., and Kouzarides, T. (2002). Histone H3 lysine 4 methylation disrupts binding of nucleosome remodeling and deacetylase (NuRD) repressor complex. J Biol Chem *277*, 11621-11624.
Zhang, K., Mosch, K., Fischle, W., and Grewal, S.I. (2008). Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. Nat Struct Mol Biol *15*, 381-388.
Zhao, W., Bian, X.-w., Shi, J.-q., and Jiang, X.-f. (2004). [Effects of ectopic glial fibrillary acidic protein/green fluorescent protein gene expression on cellular differentiation and proliferation of human glioma cell line]. Zhonghua Bing Li Xue Za Zhi *33*, 449-453.
Zhou, J., Li, H., Li, X., Zhang, G., Niu, Y., Yuan, Z., Herrup, K., Zhang, Y.-W., Bu, G., Xu, H.*, et al.* (2015). The roles of Cdk5-mediated subcellular localization of FOXO1 in neuronal death. J Neurosci *35*, 2624-2635.
Zhou, Q., Choi, G., and Anderson, D.J. (2001). The bHLH transcription factor Olig2 promotes oligodendrocyte differentiation in collaboration with Nkx2. 2. Neuron *31*, 791-807.
Zhu, Y., Sousa, A.M.M., Gao, T., Skarica, M., Li, M., Santpere, G., Esteller-Cucala, P., Juan, D., Ferrandez-Peral, L., Gulden, F.O.*, et al.* (2018). Spatiotemporal transcriptomic divergence across human and macaque brain development. Science *362*.

Ziller, M.J., Edri, R., Yaffe, Y., Donaghey, J., Pop, R., Mallard, W., Issner, R., Gifford, C.A., Goren, A., Xing, J.*, et al.* (2014). Dissecting neural differentiation regulatory networks through epigenetic footprinting. Nature.
Ziller, M.J., Gu, H., Muller, F., Donaghey, J., Tsai, L.T., Kohlbacher, O., De Jager, P.L., Rosen, E.D., Bennett, D.A., Bernstein, B.E.*, et al.* (2013). Charting a dynamic DNA methylation landscape of the human genome. Nature *500*, 477-481.
Ziller, M.J., Müller, F., Liao, J., Zhang, Y., Gu, H., Bock, C., Boyle, P., Epstein, C.B., Bernstein, B.E., Lengauer, T.*, et al.* (2011). Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. PLoS Genet *7*, e1002389.
Zuloaga, D.G., Puts, D.A., Jordan, C.L., and Breedlove, S.M. (2008). The role of androgen receptors in the masculinization of brain and behavior: what we've learned from the testicular feminization mutation. Horm Behav *53*, 613-626.