EPIGENETIC ANALYSES OF HUMAN BRAIN EVOLUTION AND AGING

A Dissertation
Presented to
The Academic Faculty

by

Hyeonsoo Jeong

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Bioinformatics

Georgia Institute of Technology
December 2021

EPIGENETIC ANALYSES OF HUMAN BRAIN EVOLUTION AND AGING

Approved by:

Dr. Soojin V. Yi, Advisor
School of Biological Sciences
Georgia Institute of Technology

Dr. Jeffrey Todd Streelman, Advisor
School of Biological Sciences
Georgia Institute of Technology

Dr. I. King Jordan
School of Biological Sciences
Georgia Institute of Technology

Dr. Genevieve Konopka
Department of Neuroscience
UT Southwestern Medical Center

Dr. John McDonald
School of Biological Sciences
Georgia Institute of Technology

Date Approved: December 5, 2021

ACKNOWLEDGEMENTS

First and foremost, I am deeply grateful to my research advisor, Prof. Soojin Yi for her unwavering support and invaluable advice during my academic research and daily life. My gratitude also extends to my thesis committee members (alphabetized by last names) – Dr. King Jordan, Dr. Genevieve Konopka, Dr. John McDonald, and Dr. Todd Streelman. Their plentiful experience and knowledge have motivated and encouraged me in all the time of my Ph.D. study.

I would like to express my sincere gratitude to my lab members (alphabetized by last names) – Dr. Nicole Baran, Paramita Chatterjee, Taylor Hoyt, Dr. Iksoo Huh, Thomas Layman, Dr. Isabel Mendizabal, Robert Morgan, Devika Singh, Dr. Dan Sun, Dr. Xin Wu for their insightful comments and suggestions at every stage of my research project. I am extremely grateful to my collaborators – Dr. Young Jang and Dr. Donna Maney for their research insights and career advice.

I would like to express my sincere gratitude to my family and friends. Without their encouragement and support during my graduate studies, it would not be possible for me to finish this long journey. Finally, my appreciation goes out to my wife (soon-to-be Dr.) Jeongmoon Choi for her tremendous understanding and encouragement through the toughest time of my graduate studies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

BA46	Brodmann area 46
BMI	Body mass index
CH	Non-CG contexts
DEGs	Differentially expressed genes
DML	Differentially methylated loci
DMRs	Differentially methylated regions
DNMTs	DNA methyltransferases
FANS	Fluorescence-activated nuclei sorting
FDR	False discovery rate
GWAS	Genome-wide association studies
H3K4me3	Histone H3-trimethyl-lysine 4
hmC	hydroxymethylcytosines
LD	Linkage disequilibrium
mC	methylcytosines

MTG	Medial temporal gyrus
OR	Odds ratio
PMI	Postmortem interval
RRBS	Reduced representation bisulfite sequencing
WGBS	Whole genome bisulfite sequencing
WGS	Whole-genome sequencing

SUMMARY

One of the central questions in modern biology is to understand the molecular basis of phenotypic variation. This question can be partially answered by examining (epi)genomic variation between species or between individuals within the same species exhibiting phenotypic variation. The aim of my thesis is to study the implication of epigenetic modifications on phenotypic traits and evolution, using DNA methylation data.

The human brain is a great example of evolutionary innovation at multiple levels, having undergone dramatic expansion accompanied by structural and molecular reorganization in a relatively short geological time. At the same time, it is becoming clear that some human brain specific traits such as disease vulnerability can be better understood in the context of their evolutionary origins. In chapter 2 and 3, we elucidated the evolutionary origins of cell-type-specific epigenetic modifications in the human brain, by performing a comparative whole genome methylome analysis of human, chimpanzee, and rhesus macaque brains. We used fluorescence-activated sorted nuclei of neurons and oligodendrocytes from the prefrontal cortex, a region involved in cognition and neuropsychiatric diseases, to compare cell-type specific methylomes of human and non-human primates.

In chapter 2, we show that human brains have overall reduced CG methylation compared to brains of non-human primates. The reduction of CG methylation in human brains contributed to human brain-specific active regulatory landscape. We discovered that human neuron-specific CG hypomethylation significantly contributes to susceptibility to schizophrenia. Interestingly, cytosine methylation in different contexts (CG vs. CH) has

played distinctive roles during human brain evolution. Specifically, in chapter 3, we show that CH methylation has increased along the evolution of human brains. The increase of CH methylation in human brains is associated with the epigenetic definition of neuronal subtypes. These novel findings link the epigenetic evolution of human brains to regulation and disease susceptibility.

It has been known for several decades that aging has a profound influence on DNA methylation. However, genomic patterns of brain DNA methylation with aging at cell-type resolution remain not well understood. In chapter 4, we examined aging-associated DNA methylation changes at the whole-genome scale and at cellular resolution. Our comprehensive analyses discovered that age explains a substantial proportion of DNA methylation variation observed in human brains. Moreover, we show that CpG methylation that varies with aging exhibits highly cell-type-specific patterns. The cell-type-specific age-associated methylated loci significantly contribute to genetic risk for brain disorders.

Since molecular mechanisms of aging may share common evolutionary characteristics across closely related species, aging studies from non-human primates can be useful for studying human aging. In chapter 5, using DNA methylation data from baboons, we developed a DNA methylation-based age predictor. Using a few hundred of CpG positions, we show that the baboon age predictor accurately predicts DNA methylation ages similar to the chronological ages. Also, we showed that early social rearing experiences differentially affect rates of epigenetic aging.

Together, these comprehensive epigenome studies will shed light on our understanding of the epigenetic evolution of the human brain and aging epigenetic programs in human brains.

CHAPTER 1. INTRODUCTION

One of the most distinguishing traits of the human lineage is its large and complex brain. The size of the human brain is more than three times larger than that of the chimpanzee when considering body size (Herculano-Houzel 2012). A number of comparative studies have shown that the human brain exhibits the species-specific shapes of neuroepithelial cells and neuron morphology (Rakic 1995, DeFelipe, Alonso-Nanclares et al. 2002, Rakic 2007). Also, the human brain shows a substantial increase in a total number of neurons compared to chimpanzees and gorillas (de Sousa, Sherwood et al. 2010, Benito-Kwiecinski, Giandomenico et al. 2021).

Even though the brains of great apes have been increasing during evolution, the human brain has experienced a recent acceleration of size expansion, considering divergence time between humans and chimpanzees (approximately 5-6 million years ago) (Herculano-Houzel 2012). Analyses of hominin skull fossils show that the average brain size of human ancestors has gradually and continuously increased over the past 3 million years, and the size of the brain became similar to that of modern humans from about 500,000 years ago (Kappelman 1996, Du, Zipkin et al. 2018). Thus, the expansion of the human brain has occurred in a relatively short geological time.

Genes expressed in the brain are known to evolve slowly compared to those in the other regions. Because the human brain had evolved rapidly, when the analysis of the sequence evolution of the human brain first began, an initial hypothesis was that brain-expressed genes in the human brain might show substantial changes in DNA sequence level. Some genes showed this pattern. For example, the SLIT-ROBO Rho GTPase-activating

protein 2 has experienced multiple rounds of human-specific duplications and has an important role in brain size and cognitive abilities in mammals (Guerrier, Coutinho-Budd et al. 2009, Sudmant, Kitzman et al. 2010, Dennis, Nuttle et al. 2012). Also, many human accelerated regions reside on the genes that produce proteins that contribute to the genetic risk of neuropsychiatric disorders and neurodegenerative diseases (Pollard, Salama et al. 2006, Capra, Erwin et al. 2013, Hubisz and Pollard 2014). However, the brain genes show the overall slow rate of evolution in humans than in chimpanzees (Wang, Chien et al. 2007). This result might appear to be contradictory to the finding of the human brain-expressed genes showing faster evolution in gene expression.

Previous studies have shown that genomic differences between humans and other apes are enriched in non-coding parts of the genome (Shulha, Crisci et al. 2012, Prescott, Srinivasan et al. 2015, Mendizabal, Shi et al. 2016). Thus, it has been proposed that human-specific traits are enriched within regulatory elements in which gene expression is regulated by various epigenetic modifications (Mendizabal, Shi et al. 2016, Vermunt, Tan et al. 2016, Castelijns, Baak et al. 2020). Epigenetic change, which is a chemical modification of the genome and of proteins that package the genome, is another important aspect of understanding the human brain evolution and genetic and gene expression changes (Razin and Riggs 1980, Vaillant and Paszkowski 2007, Lister, Mukamel et al. 2013). Since epigenetic modification is critical for many regulatory processes, epigenetic changes along the evolution of the human brain can reveal yet unknown information on important regulatory changes. In addition, epigenetic changes are often implicated in neuropsychiatric disorders, including some thought to be specific to humans, such as schizophrenia (Feng, Chang et al. 2005, Feng and Fan 2009). To understand the

contribution of epigenetics to the origins and molecular mechanisms of brain-specific gene regulation and neuropsychiatric disorders, we need to extend our knowledge of evolutionary changes in epigenome during human evolution.

In this thesis, we specifically focused on one type of epigenetic modification called DNA methylation. This refers to the addition of methyl group to cytosines in the genome, which is facilitated by DNA methyltransferases (DNMTs) (Moore, Le et al. 2013). DNA methylation is believed to affect gene expression, directly and indirectly (Holliday and Pugh 1975). DNA methylation is also known to suppress transposable elements and maintain tight packaging of genomic DNA (Hutnick, Huang et al. 2010). Dysregulation of DNA methylation has been reported in many diseases, including cancer and neuropsychiatric disorders (Aran, Sabato et al. 2013, Mendizabal, Berto et al. 2019). Many human cells are heavily modified by DNA methylation, and patterns of DNA methylation vary across different cell types (Mendizabal, Berto et al. 2019). A limitation of many previous studies is that they often used bulk tissues while DNA methylation between different cell types is known to show substantially divergent patterns. Consequently, comparing DNA methylation from bulk tissue can be biased toward specific cell types and generate false positives and false negatives due to the heterogeneity of brain tissue and the different relative cell compositions of other species. Thus, it is necessary to perform a cell-type-specific DNA methylation study to clearly understand the functional roles and implications of DNA methylation changes that occurred in the human lineage.

In chapter 2, we perform DNA methylation analyses of neurons and oligodendrocytes sorted from bulk brain samples. For single nucleotide resolution methylation maps, we utilized the whole genome bisulfite sequencing (WGBS) approach.

We perform comprehensive analyses of whole-genome methylomes from humans, chimpanzees, and rhesus macaques, thus elucidating evolutionary changes of DNA methylation during human brain evolution with unprecedented cell-type resolution. Evolutionary alterations in DNA methylation of closely related species can be an extremely powerful tool for deciphering the genotype-to-phenotype connections and providing the causal link between evolution and disease vulnerability. We hypothesize that DNA methylation changes in the human brain since the divergence of humans and chimpanzees may have affected human brain-specific regulation and contributed to genetic risk to neuropsychiatric disorders. To test our hypotheses, we investigate the association of species-specific DNA methylation changes with neuropsychiatric diseases by examining genetic risk loci resulting from genome-wide association studies.

Although DNA methylation at CG context is a dominant form of methyl-cytosines, DNA methylation at non-CG contexts is also abundant in a few cell types, especially enriched in neurons (Lister, Mukamel et al. 2013, Stroud, Su et al. 2017). In chapter 3, we investigate DNA methylation changes in a non-CG context, namely CH methylation, during recent human brain evolution. Using WGBS data, we compare the human brains to chimpanzee and rhesus macaque brains both in neurons and oligodendrocytes. Because CH methylation is often found in neurons, we aim to understand whether human-specific CH methylation changes are enriched in regions that play a role in neuronal functions. In addition, we present comprehensive multi-omics analyses of human brain evolution by integrating DNA methylome data with transcriptome data from the same individuals (Berto, Mendizabal et al. 2019) and recent data from studies of bulk and cell-type-specific

epigenetic and transcriptomic modifications of human brains (Shulha, Crisci et al. 2012, Luo, Keown et al. 2017, Zhu, Sousa et al. 2018, Price, Collado-Torres et al. 2019).

Although DNA methylation is a relatively stable epigenetic mark, DNA methylation is known to co-vary strongly with aging in diverse species (Horvath 2013, Lu, Fei et al. 2021). Interestingly, it has been observed for several decades that DNA methylation changes constantly with chronological age. A term 'epigenetic drift' is often used to refer to this change of DNA methylation during aging (Teschendorff, West et al. 2013). Epigenetic drift involves both the increase and decrease of DNA methylation, suggesting that epigenetic drift might be due to gradual dysregulation of epigenetic maintenance (Teschendorff, West et al. 2013, Sun and Yi 2015).

Despite the development of age-associated DNA methylation research, the extent of genomic patterns of DNA methylation changes with age is relatively little understood. For instance, almost all of the current epigenetic aging studies analyze DNA methylation from methylation arrays. A limitation of array-based methods is that these methods capture DNA methylation from only a subset of CpG positions in the genome. Even high-density methylation array methods can only cover less than 2% of the total CpG sites in the human genome. Also, the selected CpG positions captured by the methylation array are often biased towards promoter regions. Therefore, it is necessary to extend the study of aging-associated DNA methylation changes to the whole genome, using methods developed to examine all genomic CpGs. Also, another key missing piece of information is understanding DNA methylation at cellular resolution (Horvath and Raj 2018, Bell, Lowe et al. 2019). As epigenomic studies show heterogeneity of cellular epigenetic programs, we need to evaluate how aging-associated DNA methylation changes occur in different cell

types. To address these issues, in chapter 4, we present comprehensive analyses of age-associated DNA methylation changes using 127 WGBS data sets from postmortem brains of individuals from neonate to old adults. We take advantage of cell-type-specific epigenetic data, including neurons, non-neuronal cells, and oligodendrocytes, to investigate the genomic patterns of epigenetic drift for the cell types. In addition, by integrating transcriptome data from the same tissues as well as the disease risk genetic variants resulting from genome-wide association studies (GWAS), we aim to examine functional consequences and disease enrichment of age-associated DNA methylation changes.

Previous studies have demonstrated that methylated CpG positions that are subject to age-associated changes can be used to predict chronological and phenotypic ages (Hannum, Guinney et al. 2013, Horvath 2013). This age predictor is known to be remarkably robust and estimates the biological age using DNA methylation levels of only a few hundred CpG positions. Because of its high accuracy, the DNA methylation-based age predictor is often called the 'epigenetic clock'. Despite the fact that most epigenetic clock studies focus on human aging, epigenetic clocks have been also developed for non-human species (Anderson, Johnston et al. 2021, Bors, Baker et al. 2021, Horvath, Zoller et al. 2021). Although the epigenetic clocks are generated to predict species-specific age-related conditions, DNA methylation levels of the clock CpGs exhibit a high age correlation at least within mammalian species (Lu, Fei et al. 2021). These highly concordant aging DNA methylation patterns across species suggest that molecular mechanisms of aging may share common evolutionary characteristics across species, emphasizing the potential significance of epigenetic research from non-human primates in

aging studies. In chapter 5, by utilizing newly sequenced reduced representation bisulfite sequencing (RRBS) data of 140 anubis baboons from a broad age span to build epigenetic clocks, which can be used to study health and aging in this population. In this study, we used two groups of baboons that experience different early rearing experiences (mother-raised or nursery-raised). Thus, another goal of this study is to test if the social rearing experience during early life differentially affects rates of epigenetic aging.

In summary, these studies highlight that comprehensive epigenome analyses can be a compelling approach to uncovering genotype-to-phenotype connections and provide the causal link between evolution and disease vulnerability.

CHAPTER 2. EVOLUTION OF DNA METHYLATION IN THE HUMAN BRAIN

DNA methylation is a critical regulatory mechanism implicated in development, learning, memory, and disease in the human brain. Here we have elucidated DNA methylation changes during recent human brain evolution. We demonstrate dynamic evolutionary trajectories of DNA methylation in cell-type-specific manner. DNA methylation in CG context shows pronounced reduction (hypomethylation) in human brains, notably in cis-regulatory regions, leading to up-regulation of downstream genes. We show that the majority of differential CG methylation between neurons and oligodendrocytes originated before the divergence of hominoids and catarrhine monkeys, and harbors strong signal for genetic risk for schizophrenia. Remarkably, a substantial portion of differential CG methylation between neurons and oligodendrocytes emerged in the human lineage since the divergence from the chimpanzee lineage and carries significant genetic risk for schizophrenia. Therefore, recent epigenetic evolution of human cortex has shaped the cellular regulatory landscape and contributed to the increased vulnerability to neuropsychiatric diseases. Contents in this chapter have been published in *Nature Communications*, as Jeong et al. 2021 "Evolution of DNA methylation in the human brain".

2.1 Introduction

DNA methylation is a stable epigenetic modification of genomic DNA with critical roles in brain development (Lister, Mukamel et al. 2013, Spiers, Hannon et al. 2015, Price, Collado-Torres et al. 2019). To understand the contribution of DNA methylation to human

brain-specific gene regulation and disease susceptibility, it is necessary to extend our knowledge of evolutionary changes in DNA methylation during human brain evolution. It was previously suggested that human brain specific CG methylation may be associated with human brain specific regulation of gene expression (Zeng, Konopka et al. 2012, Mendizabal, Shi et al. 2016). However, these studies used bulk tissues, while DNA methylation is known to vary substantially between cell types. Cell-type specific epigenetic marks, including DNA methylation and histone modifications, are implicated in cell-type specific gene expression and disease susceptibility in humans (Girdhar, Hoffman et al. 2018, Mendizabal, Berto et al. 2019). Data from bulk tissues can be biased toward specific cell types and consequently, underpowered to detect cell-type specific evolutionary changes (Berto, Mendizabal et al. 2019, Khrameeva, Kurochkin et al. 2020). Therefore, to fully understand the role of DNA methylation in human brain evolution, it is necessary to study cell type specific changes of DNA methylation.

In this work, we present comparative analyses of neuron- and oligodendrocyte-specific whole-genome DNA methylomes of humans, chimpanzees, and rhesus macaques. We further integrated these data with transcriptome data from the same individuals (Berto, Mendizabal et al. 2019) and recent data from studies of bulk and cell-type specific epigenetic and transcriptomic modifications of human brains (Shulha, Crisci et al. 2012, Luo, Keown et al. 2017, Zhu, Sousa et al. 2018, Price, Collado-Torres et al. 2019). By doing so, we show that dramatic changes of DNA methylation have occurred in a cell-type-specific manner during human brain evolution.

2.2 Results

2.2.1 Distinctive methylomes of neurons and oligodendrocytes in human and non-human primate prefrontal cortex

We generated cell-type specific DNA methylomes of sorted nuclei from post-mortem brain samples of humans (Mendizabal, Berto et al. 2019), chimpanzees (Pan troglodytes), and rhesus macaques (Macaca mulatta). We selected Brodmann area 46 (BA46) from dorsolateral prefrontal cortex (also referred to as 'prefrontal cortex' or 'cortex' henceforth), which is involved in higher-order cognitive functions that have likely undergone marked changes in human evolution (Sherwood, Subiaul et al. 2008, Preuss 2011). Neuronal (NeuN+) and oligodendrocyte (OLIG2+) cell populations were isolated using fluorescence-activated nuclei sorting (FANS) as previously described (Berto, Mendizabal et al. 2019, Mendizabal, Berto et al. 2019). We used whole-genome bisulfite sequencing (WGBS) to generate DNA methylomes at nucleotide resolution for NeuN+ and OLIG2+ populations (Figure A. 1). Altogether, we compared 25, 11, 15 NeuN+ methylomes and 20, 11, 13 OLIG2+ methylomes from human, chimpanzee, and rhesus macaque, respectively (Table A. 1 and Table A. 2). We also performed whole-genome sequencing (WGS) of the same individuals (Table A. 3). Polymorphic sites at cytosines (i.e. C to T for forward strand and G to A for reverse strand) were excluded to avoid spurious methylation calls due to the technical limitation of distinguishing bisulfite-converted thymine from unmethylated cytosine. The mean coverages for the WGBS and WGS data are 20.6X (±8.8) and 23.2X (±5.9), respectively.

2.2.2 Conservation and divergence of cell-type-specific CG methylation

Due to the high rate of CG mutations associated with DNA methylation (Elango, Kim et al. 2008), only 9.6 million CG sites (out of 28 million total human CGs) are conserved in all three species (Figure A. 2) and these sites are biased toward hypomethylation (Figure A. 3). As expected, evolutionarily conserved CpG sites co-localize with CpG islands and exons (Kim, Elango et al. 2006, Mendizabal and Yi 2016) (Figure A. 4). In addition, evolutionarily conserved CpGs and human- specific CpGs are enriched in distinctive transcription factor binding motifs (Figure A. 5). Interestingly, HOX and FOX transcription factor families, among others, are significantly more often associated with human-specific CpGs than conserved CpGs (Table A. 4).

To avoid bias associated with CG conservation, we first identified differentially methylated regions (DMRs) that distinguish humans and chimpanzees using conserved sites (21 out of 25 million CGs analyzed), and subsequently added DNA methylation data from rhesus macaques to polarize direction of evolutionary change (Methods). In this analysis, we applied methods developed for the analysis of whole genome bisulfite sequencing data to identify species, cell-type and interaction effects on DNA methylation while taking into account variation due to sex, age, and bisulfite conversion rates (Methods).

Non-human primate methylomes of NeuN+ and OLIG2+ are highly distinct from each other and show clear clustering of cell types in each species (Figure A. 6), as in humans (Mendizabal, Berto et al. 2019). There are 56,532 CG DMRs (75.9 Mbp) between NeuN+ and OLIG2+ DNA that are conserved in all three species (Figure 2.1A). These conserved DMRs account for nearly 50% of all DMRs between NeuN+ and OLIG2+ in humans (Figure 2.1B). Consequently, a large portion of differential CG methylation

between NeuN+ and OLIG2+ DNA originated before the divergence of hominoids and catarrhine monkeys.

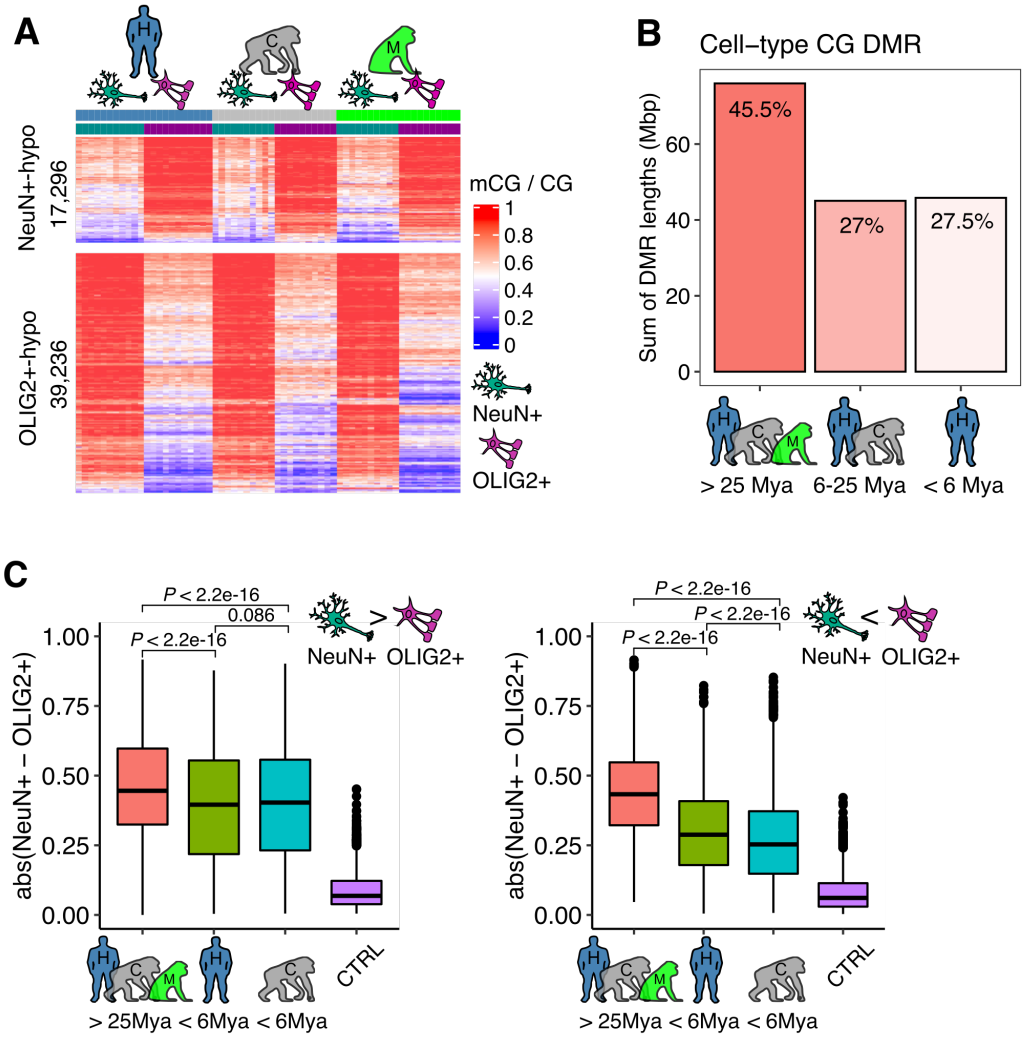

Figure 2.1 DNA methylomes of neurons (NeuN+) and oligodendrocytes (OLIG2+) are highly distinct and show conserved patterns across species. (A) CG methylation levels in neurons (left columns for each species) and oligodendrocytes (right columns for each species). A greater number of DMRs are hypermethylated in neurons (red, in the left columns) compared to oligodendrocytes (right columns). **(B)** Approximately half (45.5%) of CG DMRs differentially methylated between NeuN+ and OLIG2+ cells are conserved in all three species, with 27% conserved between humans and chimpanzees, and 27.5% specific to the human. **(C)** The absolute methylation difference of NeuN+ and OLIG2+ cells is highest for DMRs conserved in all three species (39,202 and 17,284 DMRs hypermethylated in neurons and oligodendrocytes, respectively) compared to those specific to humans (3,103 and 5,361 DMRs hypermethylated in neurons and

oligodendrocytes, respectively) or chimpanzees (4,370 and 2,989 DMRs hypermethylated in neurons and oligodendrocytes, respectively). Methylation differences between NeuN+ and OLIG2+ cells calculated from genomic regions serving as statistical control (CTRL), with a matched number of CG and G+C nucleotide contents, are also displayed. Statistical significance was computed using two-sided Mann-Whitney U-test. Box represents a range from the first quartile to the third quartile. The line in the box indicates the median value. The minima and maxima are within 1.5 times the distance between the first and third quartiles from box.

Enrichment tests utilizing cis-regulatory interactions based on long-range regulatory domains (McLean, Bristor et al. 2010) show that these regions are highly enriched in genes harboring functions specific to neurons and oligodendrocytes (Table A. 5). For example, we show one conserved DMR spanning the whole QKI locus. This gene, which is an RNA binding protein involved in myelination and oligodendrocyte differentiation (Ebersole, Chen et al. 1996), is covered entirely by a DMR in all three species so that it is hypomethylated in oligodendrocytes while hypermethylated in neurons. Gene expression data from matched samples (Berto, Mendizabal et al. 2019) shows that QKI is significantly up-regulated in oligodendrocytes compared to neurons in all three species ($P < 10^{-7}$ in all three species, Methods). This example illustrates that differential DNA methylation may facilitate cell-type specific regulation in human and non-human primate brains. Interestingly, the absolute methylation difference between neurons and oligodendrocytes was significantly more pronounced in the evolutionarily 'old' DMRs conserved in all three species compared to those recently evolved in human (Figure 2.1C).

2.2.3 Pronounced CG hypomethylation of human prefrontal cortex and human neuron-specific regulatory landscape

We found 23,703 CG DMRs (13.1Mbp) that experienced differential CG methylation since the divergence of humans and chimpanzees (Methods, Figure 2.2A and Figure 2.2B), distributed across different functional categories, including regions currently annotated as non-coding intergenic (Figure A. 7). These CG DMRs include 7,861 for which both cell types are differentially methylated between humans and chimpanzees (4,253 human-specific and 3,608 chimpanzee-specific CG DMRs, based on the comparison to macaques). The rest of the CG DMRs show DNA methylation changes in a cell-type-specific manner in each species (Figure 2.2A). Interestingly, CG DMRs were found more often than expected near previously identified brain mQTLs (Ng, White et al. 2017) (Figure A. 8), suggesting that some genomic regions might be more susceptible to genetic changes that affect DNA methylation. This is in line with the observation that the evolution of DNA methylation is associated with underlying genetic sequences (Yi 2017).

To provide insights into how DNA methylation changes at cell-type level have affected gene expression and other functional features, in the following we present results of DNA methylation analyses for each cell type, combining DMRs that are common in both cell types and DMRs that are cell-type-specific in each species (Methods). While most previous studies focused on neurons, recent studies have begun to unveil the functional and evolutionary importance of oligodendrocytes-specific changes (Berto, Mendizabal et al. 2019, Castelijns, Baak et al. 2020). Indeed, we identified a substantial number of human-derived hypomethylated DMRs specific to oligodendrocytes (Figure 2.2B).

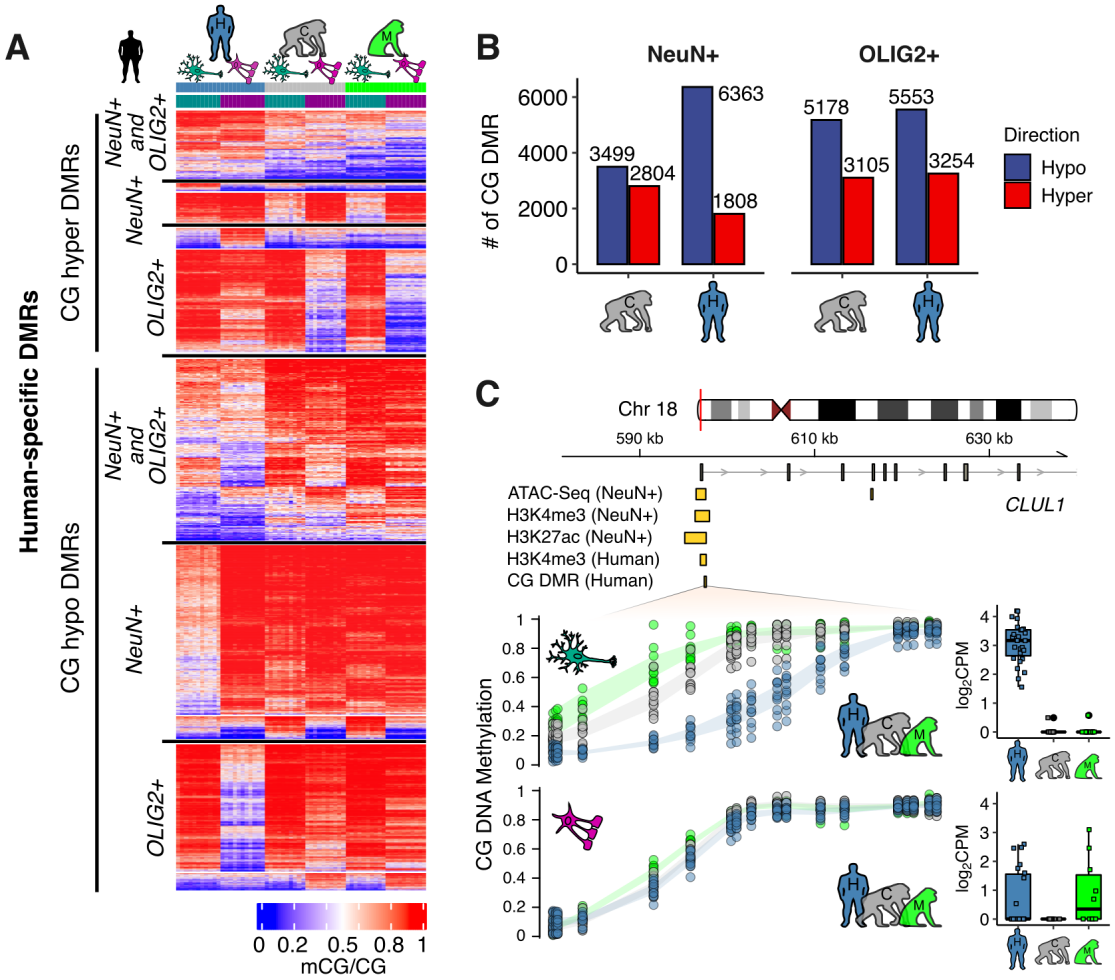

Figure 2.2 Evolutionary changes in CG methylation. (A) Heatmap representation of mean DNA methylation of all 23,703 human DMRs in the three species illustrates dramatic reduction of CG methylation in human prefrontal cortex, especially in neurons. **(B)** Numbers of DMRs in NeuN+ and OLIG2+ cells in human and chimpanzee frontal cortex. **(C)** An example of the relationship between human neuron-hypo CG DMR and other epigenetic marks in the CLUL1 locus, a gene widely expressed in the brain. This DMR overlaps with multiple other epigenetic marks of active chromatin in the human brain, including neuron-specific ATAC-Seq peak, neuron-specific H3K4me3 peak, neuron-specific H3K27ac peak. This DMR also overlaps with a human-specific brain H3K4me3 peak compared to chimpanzee and macaque. Box represents a range from the first quartile to the third quartile. The line in the box indicates the median value. The minima and maxima are within 1.5 times the distance between the first and third quartiles from box.

CG DMRs tend to show reduction of DNA methylation (hypomethylation) in human prefrontal cortex compared to chimpanzee in both cell-types (Figure 2.2B). Their enrichments in promoters and the 5' end of genes (Figure A. 7) suggest impacts on gene regulation, as hypomethylation near transcription start sites is significantly associated with up-regulation of gene expression (Schübeler 2015). Indeed, genes harboring human-hypomethylated CG DMRs are significantly enriched in human up-regulated genes compared to chimpanzees, in the same oligodendrocyte and neuron cell populations (Berto, Mendizabal et al. 2019) (Table A. 6). These results indicate widespread and significant contributions of recent CG hypomethylation to the transcriptional landscape of the human brain.

Human neurons in particular harbor a large number of hypomethylated CG DMRs compared to chimpanzee neurons (6,363 hypomethylated CG DMRs in human neurons versus 3,499 hypomethylated DMRs in chimpanzee neurons, OR = 2.82, P = 5.5 x 10^{-20}, chi-square test). Taking advantage of recent functional genomics data from human neurons, we show that human neuron-specific hypomethylated CG DMRs (referred to as 'neuron-hypo CG DMRs' henceforth) mark active regulatory regions of the neuronal genome (Figure A. 9). Specifically, a substantial portion of human neuron-hypo CG DMRs co-localize with brain-specific enhancers (Figure A. 10), as well as other recently characterized cell-type specific human brain epigenetic marks, including neuron-specific H3K27ac (fold-enrichment = 3.1, P < 0.01, permutation test), H3K4me3 (fold-enrichment = 8.5, P < 0.01), and ATAC-Seq (fold-enrichment = 8.2, P < 0.01) peaks (Fullard, Hauberg et al. 2018, Girdhar, Hoffman et al. 2018) (Figure A. 9). For example, we show a human-specific neuron-hypo CG DMR in a 5' region of the CLUL1 locus, which overlaps with

other epigenomic signatures of active chromatin marks observed in human neurons (Figure 2.2C). Even though its functional role is not resolved yet, previous studies showed that this gene is highly expressed across different brain regions (GTEx, Aguet et al. 2017). Using matched gene expression data, we show that this locus is up-regulated in a cell-type and lineage-specific manner in human neurons, consistent with the role of human-specific neuron CG hypomethylation.

In order to reveal the target genes of these epigenetically coordinated regulatory elements in human neurons, we integrated three-dimensional maps of chromatin contacts from the developing human cortex (Won, de la Torre-Ubieta et al. 2016). This analysis identified 213 enhancer-promoter pairs (Figure A. 11, fold-enrichment = 2.45, $P < 0.01$, permutation test), supporting physical chromatin interactions between spatially adjacent human neuron-hypo CG DMRs in human neuron nuclei (Table A. 7). Interestingly, genes affected by these enhancer-promoter interactions are enriched in functional categories including neuron differentiation and development (Table A. 8).

We also explored the co-occurrence of epigenetically identified regulatory elements with those emerging from DNA sequence analyses. Human hypo-methylated CG DMRs, while enriched for both conserved and human-specific CpGs, are significantly associated with binding motifs for three transcription factors, including two Forkhead box factors (FOXP1 and FOXK1) and the nuclear factor 1 C-type, NFIC. The presence of these motifs further associates with greater hypomethylation of the DMRs themselves, as well as with increased expression of downstream genes (Figure A. 12). Furthermore, non-coding human accelerated regions (ncHAR) significantly overlap with human-specific hypomethylated CG DMRs (Figure A. 11, fold-enrichment = 4.45, $P < 0.01$, permutation test). In contrast,

chimpanzee-specific hypomethylated CG DMRs did not show significant patterns (Figure A. 11). Notably, ncHARs also show an excess of three-dimensional interactions with distant human hypo CG DMRs, which include 7 experimentally validated human brain enhancer ncHARs (Capra, Erwin et al. 2013). In addition, human neuron-hypo CG DMRs frequently co-occur with human neuron-specific histone H3-trimethyl-lysine 4 (H3K4me3) modification (Shulha, Crisci et al. 2012) (fold-enrichment = 18.1, empirical P-value < 0.01, permutation test). Taken together, these results demonstrate the confluence of human-derived genetic and epigenetic innovations, and that CG hypomethylation of human neurons contributed to the active chromatin landscape of human prefrontal cortex in a cell-type specific manner.

2.2.4 Human neuron-specific CG methylation contributes additional risk to schizophrenia heritability

We have previously shown that genomic regions exhibiting differential CG methylation between neurons and oligodendrocytes are associated with increased risk for neuropsychiatric disorders, especially for schizophrenia (Mendizabal, Berto et al. 2019). Other studies have noted that sites of differential histone modification (Girdhar, Hoffman et al. 2018) or DNA methylation (Rizzardi, Hickey et al. 2019) between neurons and non-neurons significantly contribute to heritability for neuropsychiatric disorders. Our data can provide further insights into the evolution of genetic risk for neuropsychiatric disorders.

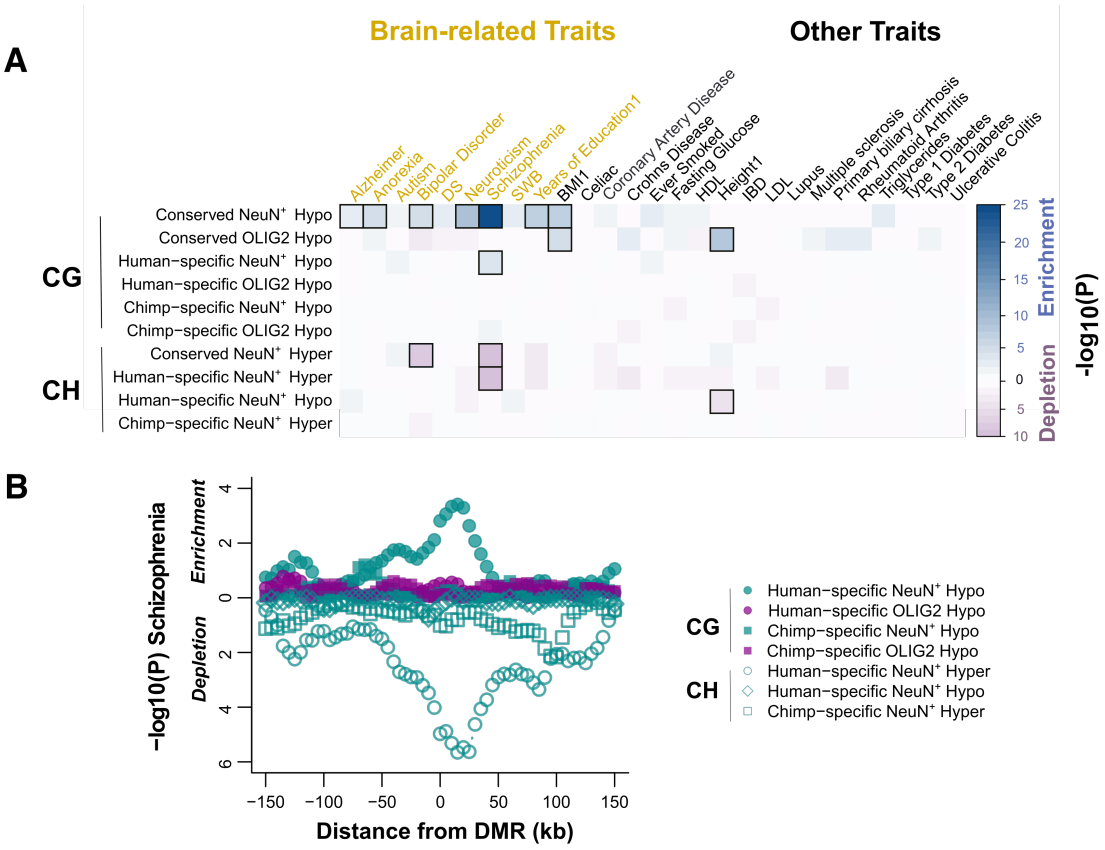

Figure 2.3 Evolutionarily derived DMRs contribute to brain disease susceptibility. (A) Significance levels for the enrichment for genetic heritability in different DMRs (+/- 25kb) and complex traits. Both conserved and human-specific neuronal DMRs are associated with schizophrenia. Enrichment with FDR < 0.05 are highlighted in squares. Notably, CG DMRs hypomethylated in NeuN+ cells compared to OLIG2+ cells in all three species (conserved NeuN+ hypo) are highly enriched in variants for several brain-related traits, and human-specific NeuN+ hypo shows enrichment in schizophrenia. **(B)** A sliding-windows analysis further demonstrates that the aforementioned signal for schizophrenia was centered at the DMRs and did not originate from extended adjacent regions. The Y-axis represents the P-values in sliding windows around DMRs classified by species (human or chimpanzee), cell-type (NeuN+ or OLIG2+), and cytosine context (mCG or mCH).

We used the stratified linkage disequilibrium score regression framework (Finucane, Bulik-Sullivan et al. 2015) to estimate the contribution of DMRs to the genetic heritability of various diseases and complex traits (Methods). We found a strong enrichment of risk

for schizophrenia and other brain-related traits at neuron-hypo CG DMRs that are evolutionarily conserved in the three species, while no signal was detected at OLIG2+ conserved DMRs (Figure 2.3A and Table A. 9). Non-brain polygenic traits such as height and body mass index (BMI) were also detected, consistent with the previously proposed role of the central nervous system in the genetic architecture of BMI (Finucane, Bulik-Sullivan et al. 2015). Moreover, human-specific neuron-hypo CG DMRs exhibited significant enrichment for schizophrenia heritability (Figure 2.3A), even though the degree of enrichment is lower than that for the conserved DMRs as suggested by down-sampling analyses (Figure A. 13). In contrast, chimpanzee neuron-hypo CG DMRs did not show significant enrichment for any human trait, while both conserved and human-specific DMRs in the non-CpG context (i.e., CH DMRs) show significant depletion for schizophrenia heritability (Figure 2.3A). Notably, the depletion signal was centered around the CH DMRs, whereas no other diseases (with the exception of bipolar disorder) nor chimpanzee-specific regions showed a significant trend (Figure 2.3B), implying that CH hypermethylated genomic regions are devoid of common DNA polymorphisms associated specifically with schizophrenia. Given that CH DMRs are enriched in inhibitory neuron markers, this observation may suggest that different neuron subclasses contribute disproportionately to schizophrenia phenotype (Finucane, Reshef et al. 2018, Hauberg, Creus-Muncunill et al. 2020).

2.3 DISCUSSION

Decades of research have solidified DNA methylation as a critical regulatory mechanism in human brains, including but not limited to brain development (Hon, Rajagopal et al. 2013, Lister, Mukamel et al. 2013, Price, Collado-Torres et al. 2019), cell-

type differentiation (Lister, Mukamel et al. 2013, Luo, Keown et al. 2017, Mendizabal, Berto et al. 2019), and disease susceptibility (Mendizabal, Berto et al. 2019, Rizzardi, Hickey et al. 2019). These processes are associated with cognitive and neurodevelopmental programs and neuropsychiatric disorders that are key to human uniqueness (Varki, Geschwind et al. 2008, Jakovcevski and Akbarian 2012). Despite such importance for genome regulation and human evolution, how DNA methylation and other epigenetic mechanisms have changed in human brains have not previously been characterized at the cell-type level. Reliable identification of human-specific epigenetic modifications at the cell-type level has been a limiting factor in previous studies due to the heterogeneity of brain tissue and the different relative cell compositions of different species. Here, we have presented comprehensive analyses of whole-genome methylomes of neurons and oligodendrocytes from humans, chimpanzees, and rhesus macaques, thus elucidating evolutionary changes of DNA methylation during human brain evolution with unprecedented cell-type resolution.

Previous studies have demonstrated an excess of CG hypomethylation in human prefrontal cortex compared to chimpanzee (Zeng, Konopka et al. 2012), mostly impacting noncoding regulatory regions of the human genome (Mendizabal, Shi et al. 2016). We find this to be the case for both neurons and oligodendrocytes, which could contribute to increased gene expression levels that have been reported in human brains (Caceres, Lachuer et al. 2003, Preuss, Caceres et al. 2004, Babbitt, Fedrigo et al. 2010, Sousa, Meyer et al. 2017, Berto, Mendizabal et al. 2019). Furthermore, these epigenomic innovations connect to potential underpinnings in genome evolution. For example, human derived hypomethylated CG DMRs are enriched for binding motifs for specific transcription

factors including FOXP1, a hub gene in human-specific transcriptional networks in the brain and which is implicated in several cognitive diseases in humans, including language, intellectual disability and autism (Konopka, Friedrich et al. 2012). In addition, non-coding human accelerated regions (ncHAR) are preferentially found in human-specific hypomethylated CG DMRs. These results begin to reveal the connections between genetic and epigenetic innovations of the human brain involving CG hypomethylation.

Our data also demonstrate that the majority of differential DNA methylation between neurons and oligodendrocytes has long been established before the divergence of apes and other catarrhine monkeys, echoing that a large portion of human brain regulatory programs have deep evolutionary roots (Vermunt, Tan et al. 2016). We further investigated the implication of this finding in the context of a complex neuropsychiatric disorder. We and others have previously shown that, in humans, epigenetic differences between neurons and non-neuronal cells are prevalent in non-coding regions and locate in regions that account for schizophrenia heritability (Girdhar, Hoffman et al. 2018, Mendizabal, Berto et al. 2019, Rizzardi, Hickey et al. 2019). Here we show that genomic regions with differential CG methylation between neurons and oligodendrocytes that contribute greatest to schizophrenia risk originated before the emergence of the catarrhine ancestor. It is known that genomic regions under strong and ancestral purifying selection (thus remain conserved) are enriched for disease genes and heritability (Domazet-Lošo and Tautz 2008, Finucane, Bulik-Sullivan et al. 2015, Hujoel, Gazal et al. 2019). For example, ancient enhancers and promoters have greater contributions to susceptibility to complex diseases compared to more recently evolved regulatory regions (Hujoel, Gazal et al. 2019). Our

results suggest that even though the phenotype of schizophrenia is highly specific to humans, the molecular and developmental mechanisms of this disease have deep phylogenetic roots. Moreover, human- brain specific CG hypomethylation provides additional significant genetic risk to schizophrenia, albeit a relatively small proportion. Therefore, recent, human brain specific epigenetic changes also contribute to schizophrenia pathology. These results advance our understanding of the relevance of conserved and derived regulatory mechanisms to the genetic and epigenetic architecture of complex diseases.

2.4 Methods

2.4.1 Sample acquisition, whole-genome sequencing, and whole-genome bisulfite sequencing

Information on samples used in this work was previously described in Berto et al. (Berto, Mendizabal et al. 2019). Briefly, adult human postmortem brain samples from Brodmann area 46 (BA46) were acquired from the National Institutes of Health NeuroBioBank (the Harvard Brain Tissue Resource Center, the Human Brain and Spinal Fluid Resource Center, VA West Los Angeles Healthcare Center, and the University of Miami Brain Endowment Bank) and the University of Texas Neuropsychiatry Research Program (Dallas Brain Collection). These samples included 25 and 22 NeuN+ and OLIG2+ specimens, respectively. Nonhuman primate tissue samples were obtained from Yerkes National Primate Research Center (macaque samples) and the National Chimpanzee Brain Resource (chimpanzee samples). For human samples, UT Southwestern Medical Center

Institutional Review Board has determined that as this research was conducted using post-mortem specimens, the project does not meet the definition of human subjects research and does not require IRB approval and oversight. Non-human primate samples were obtained from archival, post-mortem brain tissue opportunistically collected from subjects that died from natural causes, and following procedures approved by the Emory Institutional Animal Care and Use Committee and in accordance with federal and institutional guidelines for the humane care and use of experimental animals. No living great apes were used in this study. All non-human primate samples were obtained from homologous regions in chimpanzees (NeuN+ n = 11, OLIG2+ n = 11) and rhesus macaques (NeuN+ n = 15, OLIG2+ n = 13).

Nuclei isolation was performed as described previously (Berto, Mendizabal et al. 2019). Briefly, frozen postmortem brain was homogenized and subject to sucrose gradient and ultracentrifuge. The resulting nuclei pellet was then incubated with mouse NeuN and OLIG2 antibodies (alexa488 conjugated anti-NeuN (1:200), #MAB377X, Millipore, Billerica, MA and rabbit alexa555 conjugated anti-OLIG2 (1:75), #AB9610-AF555, Millipore). We then performed the fluorescence-activated nuclei sorting (FANS), followed by nucleic acid purification via the ZR-Duet DNA/RNA MiniPrep (Plus) kit (#D7003, Zymo Research, Irvine, CA).

2.4.2 Whole-genome bisulfite data processing

We followed the same data processing steps described in our previous work (Mendizabal, Berto et al. 2019). Briefly, extracted DNA was fragmented by S-series Focused-ultrasonicator (Covaris, Woburn, MA) using the “200 bp-target peak size

protocol". Fragmented DNA was then size selected (200-600 bp) with an Agencourt AMPure XP bead-based (#A63880, Beckman Coulter, Brea, CA), followed by the End repair step was performed with End-It DNA End-Repair Kit (#ER81050, Epicentre, Madison, WI) and A-tailing (#M0202, New England Biolabs, Ipswich, MA), and ligation of methylated adaptors (#511911, B100 Scientific, Austin, TX). The methylome libraries were diluted and loaded onto Illumina HiSeqX system for sequencing using 150bp paired-end reads. We performed quality and adapter trimming using TrimGalore v.0.4.1 (Babraham Institute) with default parameters. Reads were mapped first to PhiX genome (NC_001422.1) to remove the spike-in control and the remaining reads were subsequently mapped to the chimpanzee PanTro5 and macaque rheMac8 reference genomes using Bismark v 0.14.5 (Krueger and Andrews 2011) and bowtie v2.3.4 (Langmead and Salzberg 2012). After de-duplication, we obtained coverage for over 84% of the CpGs in the chimpanzee genome with an average read depth 19.32x, and over 91% of CpGs in the macaque genome with an average read depth of 21.61x. We calculated fractional methylation (ratio of the number of methylated cytosine reads to the total number of reads) levels at individual cytosines. Bisulfite conversion rates were estimated by mapping the reads to the lambda phage genome (NC_001416.1).

2.4.3 Whole-genome sequencing data processing

Quality and adapter trimming was performed using TrimGalore v.0.4.1 (Babraham Institute) with default parameters. Reads were mapped to the hg19, PanTro5 or rheMac8 reference genomes using BWA v0.7.4 (Li 2013) and duplicates were removed using picard v2.8.3 (https://broadinstitute.github.io/picard/index.html). We identified genetic polymorphisms from re-sequencing data following the GATK v4 best practices workflow

(McKenna, Hanna et al. 2010). For base recalibration, we used vcf files for known variants from dbSNP for chimpanzee and macaque from the following links:

ftp://ftp.ncbi.nlm.nih.gov/snp/organisms/chimpanzee_9598/VCF/

ftp://ftp.ncbi.nlm.nih.gov/snp/organisms/macaque_9544/VCF/

We applied hard filters for genotype calling with the following parameters:

--filterExpression "QD < 2.0 || FS > 60.0 || MQ < 40.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0"

For chimpanzee, we identified 10,980,856 variants with mean depth >24x. For macaque, we identified 30,001,119 variants with mean depth >24x. Since C>T and G>A polymorphisms at CpG sites can generate spurious differential methylation patterns, we removed polymorphic CpGs from downstream differential methylation analyses keeping a total of 26,024,877 and 24,740,404 non-polymorphic CpGs for chimpanzee and macaque genomes, respectively. For quality control of SNP calling, we performed principal component analyses using additional chimpanzee and bonobo samples from de Manuel et al. (De Manuel, Kuhlwilm et al. 2016) using 75,575 common SNPs from chromosome 20. As expected, our chimpanzee samples clustered with other chimpanzees and not with bonobos. We recapitulated the genetic ancestry of de Manuel et al. samples and identified most of our individuals as Western chimpanzees (*Pan troglodytes verus*) while one sample (sample ID Anja) clustered with Nigeria-Cameroon chimpanzees (*Pan troglodytes ellioti*).

2.4.4 Transcription factor motif enrichment analyses

We performed TF enrichment tests using the MEME suite's (Bailey, Boden et al. 2009) AME software and two HOCOMOCO v11 databases (Kulakovskiy, Vorontsov et al. 2016) of human TF motifs. We used 7 primates (human, chimpanzee, gorilla, orangutan, rhesus macaque, baboon, and gibbon) for which we have high quality genome sequences to identify cytosines that are conserved in all 7 primate species (n = 567,893) as 'conserved CpGs'. In comparison, 'variable CpGs' refer to CpGs that are specific to humans but not in other primates (n = 237,956). We identified TF motifs enriched at variable CpGs compared to conserved CpGs, as defined above. For this analysis, we added 20 bps to each side of each CpG given that the longest motif length in the database is 25bp. We compared the variable CpGs to control CpG sets as follows. We ran AME 100 times comparing the variable CpGs to a matched number of random CpG (defined as not overlapping with variable or conserved CpGs) using the following command:

ame --verbose 2 --oc variable_CpG.fa --scoring avg --method fisher --hit-lo-fraction 0.25 --evalue-report-threshold 10.0 --control control_CpG_1.fa HOCOMOCOv11_core_HUMAN_mono_meme_format.meme

Similarly, we also ran AME for conserved CpGs using 100 control CpG sets as background, as well as using the Full Homocomo v11 database.

We subsequently defined variable CpG-specific motif as those that satisfy both of the following conditions:

(frequency of enrichment in variable CpGs compared to control CpGs >0.95 in the 100 comparisons) AND (frequency of enrichment in conserved CpGs compared to control CpGs < 0.05)

In comparison, conserved CpG-specific motifs are those that satisfy both of the following conditions:

(frequency of enrichment in variable CpGs compared to control CpGs for >0.95) AND (frequency of enrichment in conserved CpGs compared to control CpGs < 0.05)

A total of 81 and 121 motifs were identified as variable CpG-specific and conserved CpG-specific in the core database, and 183 and 190 in full database, respectively (Table A. 4). The TF families with at least 5% difference between the two categories are shown in Figure A. 5.

We also applied MEME suite's AME software and two HOCOMOCO v11 databases to compare human-hypomethylated DMRs to chimpanzee-specific hypomethylated DMRs. We extended the DMRs 10bp to each side and run AME using the parameters as shown before. We found 3 TF motifs significantly associated with human hypomethylated DMRs, including two Forkhead box factors (FOXP1 and FOXK1) and the nuclear factor 1 C-type, NFIC. Identical results were obtained for both core and full datasets. 79% of human-hypomethylated DMRs showed a hit in any of the three TF motifs. A total of 1996 human-specifically hypomethylated DMRs associated with FOXP1 motif, 1906 DMRs with FOXK1 and 462 with NFIC motif. The DMRs with positive hits were highly shared among TFs, with around 80% shared between FOXP1 and FOXK1, and around 60% of NFIC binding-DMRs also bind the other two TFs. We compared the methylation levels of these DMRs and the associated gene expression patterns compared to other DMRs without enriched motifs (Figure A. 12).

2.4.5 RNA-Seq data

We used our previously generated matched samples of RNA-Seq datasets for human (without brain-related diseases), chimpanzee, and rhesus macaques from GSE108066, GSE107638, and GSE123936. The list of differentially expressed genes (DEGs) were also obtained from this previous work (Berto, Mendizabal et al. 2019).

2.4.6 Liftover of non-human primates cytosine positions to human genome

We lifted over the non-human primates' cytosine coordinates to human hg19 genome using UCSC batch liftover tool (panTro5ToHg19.over.chain.gz and rheMac8ToHg19.over.chain.gz for chimpanzee and rhesus macaque, respectively). For the CG DMR analysis, we did not perform three-way species analyses based on lifted over coordinates due to the rapid evolutionary loss of CG sites since the macaque split. Compared to around 21 million CG sites conserved between human and chimpanzee, only around 9.6 million CGs are conserved between human and macaque, whereas 13 million CGs in macaque show non-CG dinucleotides in human. To circumvent this issue, we first identified human-chimpanzee differentially methylated regions (DMRs) using conserved CGs and then used orthologous regions in the macaque rheMac8 genome to polarize the DMRs (see "Incorporation of Rhesus Macaque as an outgroup species" for additional details). We removed cytosines located in paralogous sequences in at least one species to avoid erroneous mapping (i.e. one-to-many or many-to-one mapping between species). For the CH methylation analysis, we used orthologous cytosines conserved among the three species.

2.4.7 Identification of CG differential methylation

We identified differentially methylated positions of 1) cell-types (NeuN+ vs. OLIG2+), 2) species (human vs. chimpanzee where both cell types show the same direction and magnitude of methylation differences between two species) and 3) cell-type-specific species changes (either cell-type exclusively shows DNA methylation difference between species) using DSS (ver. 2.3) Bioconductor package (Feng, Conneely et al. 2014). DSS handles variance across biological replicates and models read counts from WGBS experiments while accounting for additional biological factors. Specifically, we considered age (converted to three level categorical variable), sex, and conversion rates as covariates in the following model;

Fractional methylation ~ cell_type + species + species:cell_type + sex + age_class + conversion_rates

To remove low coverage loci, we only included sites with at least 5x coverage in 80% of individuals per species or cell-type. We used a false discovery rate (FDR) threshold of 5% to identify significant differentially methylated positions. For DMR identification, we considered a minimum length of 50bp with at least 4 significant differentially methylated positions. We removed cell-type DMRs and species DMRs that overlap with cell-type-specific species changes (i.e. interaction of cell-type and species effects) to remove redundant DMRs. We only considered the DMRs that show >10% of average methylation difference between human and chimpanzee for species DMR and >15% of average methylation difference between cell-types for cell-type DMR (please also see the section "Incorporation of Rhesus Macaque as an outgroup species" for detailed explanation of final set of DMRs).

Of note, as our differential methylation analyses were run under a multifactor design in DSS, the estimated coefficients in the regression were based on a generalized linear model framework using the arcsine link function to reduce dependence of variance on the fractional methylation levels (Park and Wu 2016). The distribution of the statistic is determined by the differences in methylation levels as well as by biological and technical factors such as read depth. The sign of the test statistic indicates the direction of methylation. However, the values of the test statistic cannot be directly interpreted as fractional methylation differences. For DMRs, the tool generates "areaStat" values which are defined as the sum of the test statistic of all CG sites within the DMR. To identify the stringent sets of DMRs we excluded DMRs if the average test statistics of corresponding CGs in the region (areaStat divided by the number of CGs) was below the test statistic corresponding to FDR = 0.05.

2.4.8 Incorporation of rhesus macaque as an outgroup species

We retrieved the corresponding genomic coordinates in rheMac8 using the Ensembl Primate EPO multiple sequence alignment (Zerbino, Achuthan et al. 2018). Read counts and methylation values of the CGs in corresponding regions were obtained from the macaque samples. Only CG sites with at least 5x coverage in 80% of the individuals per species were considered. The DMRs resulting from human and chimpanzee samples that had low alignment coverages with macaque (<50%) or included less than 4 CGs in macaque were considered "unclassified" DMRs. After adding macaque data, we fitted a beta regression model using the average methylation level of each individual accounting for the covariates indicated above. Among the cell-type DMRs resulting from human and chimpanzee samples, DMRs in which macaque showed cell-type changes in the same

direction and exhibited >15% fractional methylation difference were considered conserved cell-type DMRs.

We then used stringent criteria to categorize the species specificity of DMRs as human- or chimpanzee-specific. For example, a human-specific hypomethylated DMR should satisfy the following criteria: 1) the average fractional methylation of human is significantly lower than that of chimpanzee and macaque (FDR < 0.05), 2) the absolute methylation difference between human and macaque is greater than that between chimpanzee and macaque, 3) the proportion of the absolute methylation difference between human and macaque is greater than 5%, and 4) both of the two cell-types satisfy these criteria. Those DMRs that did not satisfy these criteria were considered "unclassified". We used the same logic to specify human-specific hypermethylated DMRs and chimpanzee-specific hypo- and hypermethylated DMRs. We also examined species-specific DMRs that show differential methylation between species but exclusively in one cell-type (i.e. either cell-type shows differential methylation patterns derived from either the human or chimpanzee lineage).

2.4.9 Lineage-specific accelerated non-coding regions

We used a set of human accelerated regions from Capra et al. (Capra, Erwin et al. 2013), which combined regions identified from independent studies (i.e. the 721 'Pollard HARs' from Lindblad-Toh et al. (Lindblad-Toh, Garber et al. 2011), the 1356 'ANC' regions from Bird et al. (Bird, Stranger et al. 2007), the 992 'HACNS' regions from Prabhakar et al. (Prabhakar, Noonan et al. 2006), and the 63 'Bush08' regions from Bush and Lahn (Bush and Lahn 2008)). Statistical significance and fold-enrichment for DMRs

were computed from the occurrences of DMRs for each feature compared to GC matched control region sets (n=100).

2.4.10 Hydroxymethylation

We used previously published methylome and hydroxymethylome maps at nucleotide resolution in the adult human brain (Wen, Li et al. 2014). The hmC and mC sites were defined in the original paper. We included the cytosines that are orthologous across the three species (n = 2,905,389). We compared the proportions of differentially methylated loci between 5-hydroxymethylcytosines (hmC) and 5-methylcytosines (mC). The proportions of the differentially methylated loci at hmC loci (4.2%) and mC loci (4.2%) showed no difference.

2.4.11 Contribution of DMRs to disease heritability using stratified LD score regression

To quantify the contribution of DMRs to the genetic risk of different traits and diseases, we performed stratified LD score regression analyses (Finucane, Bulik-Sullivan et al. 2015). This method estimates the percentage of heritability explained by a set of SNPs in a certain trait using GWAS summary statistics and computes the enrichment and significance by comparing the observed heritability to the expectation given the fraction of the genome considered. We used default parameters and excluded the MHC region as in Finucane et al. (Finucane, Bulik-Sullivan et al. 2015). Together with the DMR annotations, we also included the basal functional categories described in the original paper. The list of GWAS traits and references are listed in Table A. 10.

The stratified LD score regression method produces large standard errors when the annotation categories cover a small fraction of the genome. Since evolutionary DMRs are generally short (e.g. the median lengths of human CG DMR and CH DMR are 471 bps and 246 bps, respectively) we extended the DMR windows by 25kbp on both sides to improve the confidence intervals of the estimates as in other studies (Karczewski, Francioli et al. 2020). To ensure the GWAS signals were centered around the DMRs and not emerging from the extended regions, we further performed the stratified LD score regression in sliding windows 300kb around the DMRs with a window size of 20kb and step size of 5kb.

Conserved CG DMRs were more numerous and longer than human-specific ones, which could lead to increased statistical power on stratified LD score regression analyses. In order to directly compare the significance of conserved and human-specific DMR categories to schizophrenia heritability, we performed partitioned stratified LD score analyses using 100 random sub-samplings of conserved regions.

2.5 Acknowledgements

This chapter has been published in Nature Communications. Based on the journal's rights and permission policy (https://www.nature.com/nature-portfolio/editorial-policies), the contents of this published work can be included in my thesis dissertation.

This work was partially supported by the Asan Foundation (Biomedical Science Scholarship) to Hyeonsoo Jeong; the James S. McDonnell Foundation 21st Century Science Initiative in Understanding Human Cognition – Scholar Award and the Jon Heighten Scholar in Autism Research at UT Southwestern to Dr. Genevieve Konopka.; National Science Foundation (SBE-131719 and EF-2021635) to Dr. Soojin V. Yi; and the

NIMH (MH103517), to Dr. Todd M. Preuss, Dr. Genevieve Konopka, and Dr. Soojin V. Yi. The National Chimpanzee Brain Resource was supported by NINDS (R24NS092988). Macaque tissue collection and archiving was supported by the NIH National Center for Research Resources (P51RR165; superseded by the Office of Research Infrastructure Programs (OD P51OD11132)) to the Yerkes National Primate Research Center.

CHAPTER 3. EVOLUTIONARY DYNAMICS OF CH METHYLATION IN HUMAN NEURONS

DNA methylation shows strong dynamics during mammalian brain development, and the comparative study of cell-type specific DNA methylation patterns across primate brains is highly relevant to further the understanding of human brain evolution. Although DNA methylation at CG context is a dominant form of methyl-cytosines, DNA methylation at non-CG contexts is also abundant in a few cell types, especially enriched in neurons. Here we have elucidated DNA methylation changes in non-CG context, namely CH methylation, during recent human brain evolution. CH methylation has increased (hypermethylation) in neuronal gene bodies during human brain evolution, contributing to human-specific down-regulation of genes and co-expression modules. The effects of CH hypermethylation is particularly pronounced in early development and neuronal subtypes. Contents in this chapter have been published in *Nature Communications*, as Jeong et al. 2021 "Evolution of DNA methylation in the human brain".

3.1 Introduction

DNA methylation at non-CG contexts (CH methylation, where H = A, C, T) is relatively abundant in brains, where it is associated with postnatal neuronal maturation and cell-type specific transcriptional activity (Lister, Mukamel et al. 2013, Kozlenkov, Wang et al. 2016, Stroud, Su et al. 2017). Despite such importance, the evolutionary trajectories and significance of CH methylation during human brain evolution remain little understood.

In this work, we present comparative analyses of whole-genome CH DNA methylomes of humans, chimpanzees, and rhesus macaques. By integrating these data with matched transcriptome data from the same individuals (Berto, Mendizabal et al. 2019) and public data from studies of single-cell DNA methylation of human brains, we show that dramatic changes of CH methylation have occurred in neuronal gene bodies during human brain evolution, contributing to human-specific down-regulation of genes and co-expression modules. Our work extends the knowledge of the unique roles of CH methylation in human brain evolution and offers a new framework for investigating the role of the epigenome evolution in connecting the genome to brain development, function and diseases.

3.2 Results

3.2.1 Distinctive genomic methylation patterns in different contexts of cytosine methylation

We found that as in humans, non-human primate prefrontal cortex is highly methylated at CG sites, and NeuN+ DNA is more highly methylated than OLIG2+ DNA ($P < 10^{-10}$, two-sample K-S tests, Figure 3.1A). In comparison, CH methylation occurs in much lower frequencies than CG methylation, and is nearly exclusive to NeuN+ DNA in humans (Lister, Mukamel et al. 2013, Mendizabal, Berto et al. 2019) and non-human primates (Figure 3.1A). Interestingly, neurons of humans and chimpanzees have significantly more highly CH methylated sites than those of rhesus macaques and mice ($P = 4.3 \times 10^{-5}$, Kruskal-Wallis test), indicating that brain CH methylation may have increased

in human and ape brains. In turn, human brains show greater CH methylation compared to chimpanzee brains (P = 0.03, Mann-Whitney U test using proportions of mCH > 10%).

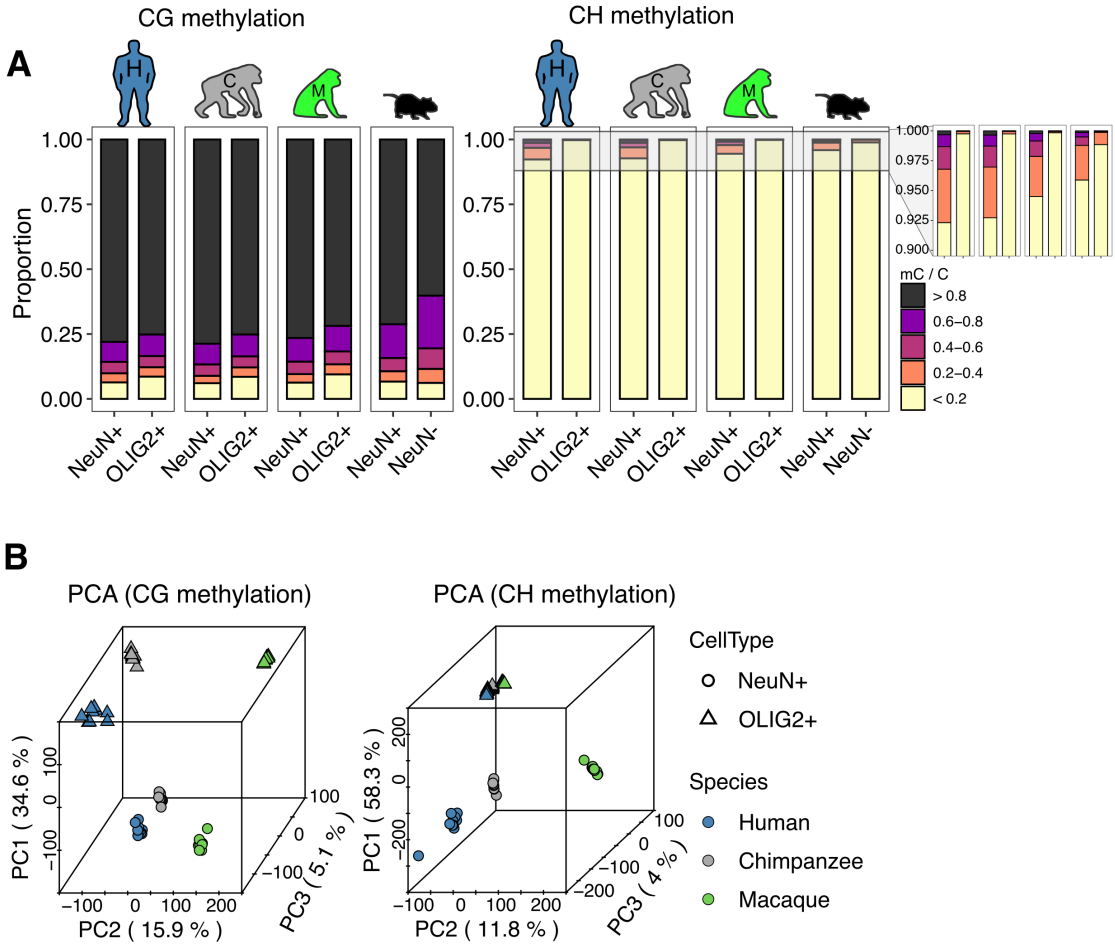

Figure 3.1 CG and CH methylation in NeuN+ and OLIG2+ in human and non-human primate prefrontal cortex. (A) The proportions of methylated CG and CH sites. Human and non-human primate neurons and oligodendrocytes are highly CG methylated. Human and non-human primate neurons show low levels of CH methylation and oligodendrocytes show even lower levels. CH methylation is highest in human neurons, followed by chimpanzees, rhesus macaques, and mice. (B) Principal component analysis of methylated cytosines in two contexts (CG and CH). The top two principal components (PCs), PC1 and PC2, distinguish cell type and species, respectively.

Principal component analyses demonstrate that cell type explains the largest amount of variation in both methylation contexts, followed by species (Figure 3.1B). Since OLIG2+ DNA is largely devoid of CH methylation, there is little separation of species for CH OLIG2+ (Figure 3.1B). As the genomic patterns and cellular distributions of CG and CH methylation are highly distinct from each other, we analyzed them separately.

3.2.2 Signature of evolutionarily recent CH hypermethylation in human neurons

CH methylation is limited to a few cell types in the body (Lister, Pelizzola et al. 2011, Ziller, Müller et al. 2011), and occurs at much lower frequency than CG methylation (Fig. 1a). Nucleotide substitution rates at CH sites and CH methylation do not have a significant correlation (Mugal and Ellegren 2011). Consequently, we were able to follow the evolutionary dynamics of CH methylation for the majority of CH positions. Among the 1.1 billion CH positions examined in the human genome, 716 million sites (71.2%) were found in the three species we examined (Methods). We found 51.9 million CH sites hypermethylated in NeuN+ compared to OLIG2+ DNA (FDR < 0.05). Among these, 23.6 million sites (45.5%) show NeuN+ DNA hypermethylation in all three species. Human and chimpanzee neurons share an additional 16.3 million (31.4%) CH hypermethylated sites not found in macaque (Figure 3.2A). Moreover, an additional 3.1 million CH sites gained methylation in the human neurons (Figure B. 1), which is a significant excess compared to the 2.2 million sites gained via CH methylation in the chimpanzee neurons (OR = 1.54, 95% CI 1.534 – 1.546, $P < 10^{-20}$, chi-square test). Thus, in contrast to the pronounced hypomethylation in the CG context, human neurons are predominantly hypermethylated compared to other primates (Figure 3.2B).

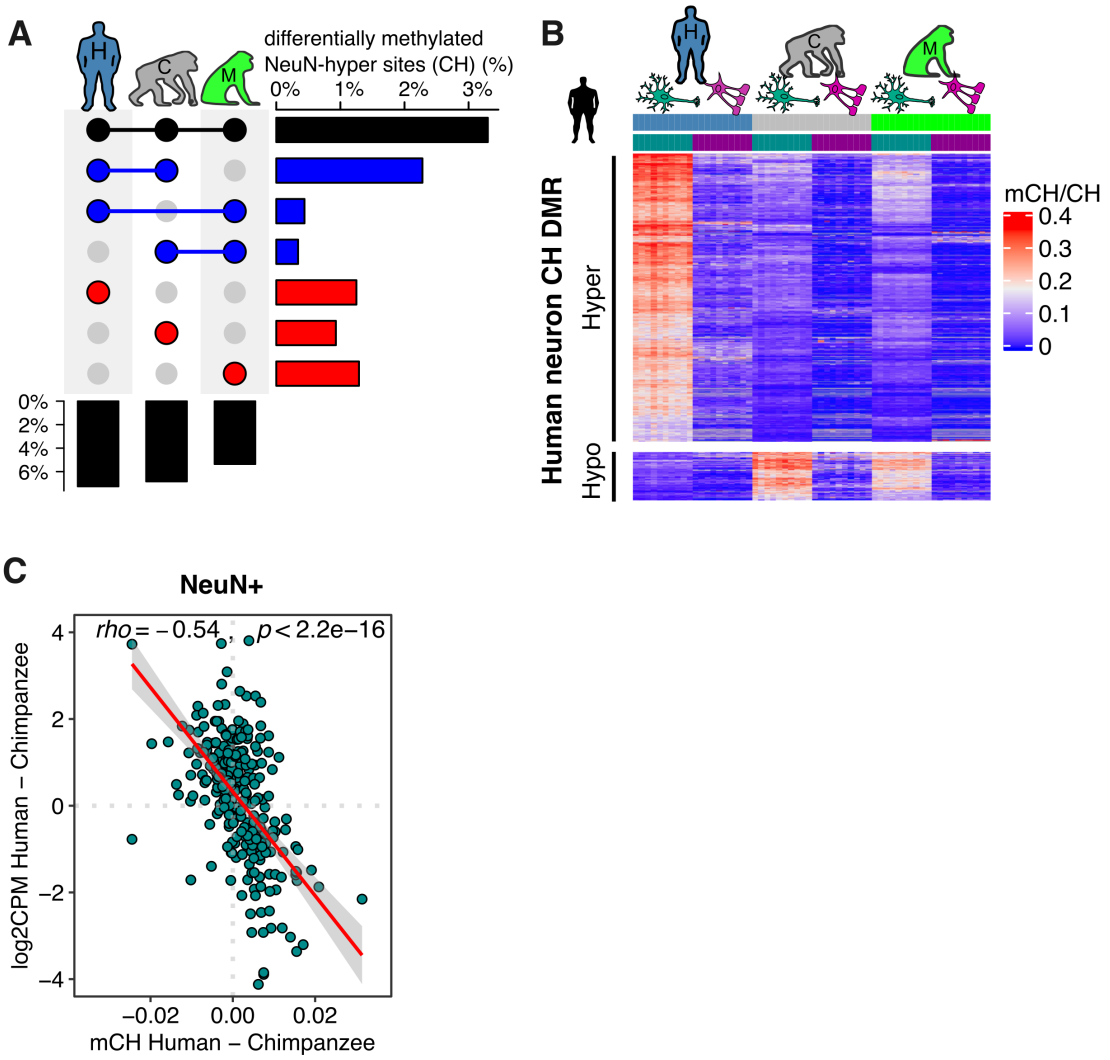

Figure 3.2 CH hypermethylation is significantly higher in human neurons compared to other primates. (A) Differences in the proportions of sites with neuronal CH methylation between species. (B) Mean methylation levels of human-specific CH DMRs demonstrate pronounced hypermethylation of human neurons. (C) CH methylation between humans and chimpanzees strongly predicts gene expression difference. The shaded band represents the 95% confidence interval for the fitted regression line.

CH methylation of gene bodies is one of the strongest predictors of repression of gene expression in humans and mice (Lister, Mukamel et al. 2013, Stroud, Su et al. 2017, Mendizabal, Berto et al. 2019, Rizzardi, Hickey et al. 2019). We find similarly strong repressive effects of genic CH methylation on gene expression in human and non-human primate neurons (Figure B. 2). Moreover, differential CH methylation between species is strongly negatively correlated with gene-expression differences between species, indicating that the change of CH methylation is a major determinant of neuronal transcriptional divergence (Figure 3.2C).

3.2.3 Distinctive evolutionary signatures of CG and CH methylation on the human neuronal transcriptome

We have demonstrated that DNA methylation at different cytosine contexts shows distinctive patterns during the recent evolutionary history of human brains. Specifically, the pronounced hypomethylation in CG context, associated with active cis-regulatory elements, contrasts with the repressive hypermethylation observed at CH sites in gene-bodies in human neurons. Given that both types of methylation correlate with gene expression (Lister, Mukamel et al. 2013, Schübeler 2015, Stroud, Su et al. 2017, Spainhour, Lim et al. 2019) (Figure B. 3), we analyzed their effects jointly using tools designed to measure independent effects of highly correlated variables (Kim and Yi 2007). These analyses point to significant and independent effects of both CG hypomethylation and CH hypermethylation (Table B. 1 and Table B. 2). Compared to chimpanzees, genes up-regulated in human neurons are more likely to have been impacted by CG hypomethylation at promoters, while those down-regulated are prone to genic CH hypermethylation (Figure B. 4). In line with these observations, coordinately up-regulated gene modules in human

neurons are enriched in promoter CG hypomethylation, whereas down-regulated modules are significantly enriched in CH hypermethylated gene bodies. These results illuminate contrasting yet additive effects of CG and CH during recent evolution of human neurons.

3.2.4 Developmental and cellular specificity of CH methylation

CH methylation is nearly absent in fetal brains and accumulates rapidly after birth (Lister, Mukamel et al. 2013). We thus hypothesized that the repressive impact of CH methylation might be more pronounced in early postnatal development, and subsequently examined gene expression data from bulk brain tissue during development (Zhu, Sousa et al. 2018). Indeed, genes bearing signatures of human-specific CH methylation accumulation (referred to as human CH DMR genes, Methods) are similarly expressed in human and macaque brains during prenatal growth but show reduced expression in humans following birth (Figure 3.3A). In contrast, chimpanzee CH DMR genes do not exhibit such a pattern (Figure 3.3A and Figure B. 5). We integrated our data with those from sorted neurons from individuals of different ages (Berto, Mendizabal et al. 2019, Price, Collado-Torres et al. 2019), to examine cell-type differences. Human CH DMR-genes showed lower expression in neurons than in non-neurons or oligodendrocytes in most developmental stages, and the reduction of neuronal expression was more evident in toddler and early teen data compared to data from adults (Figure B. 6).

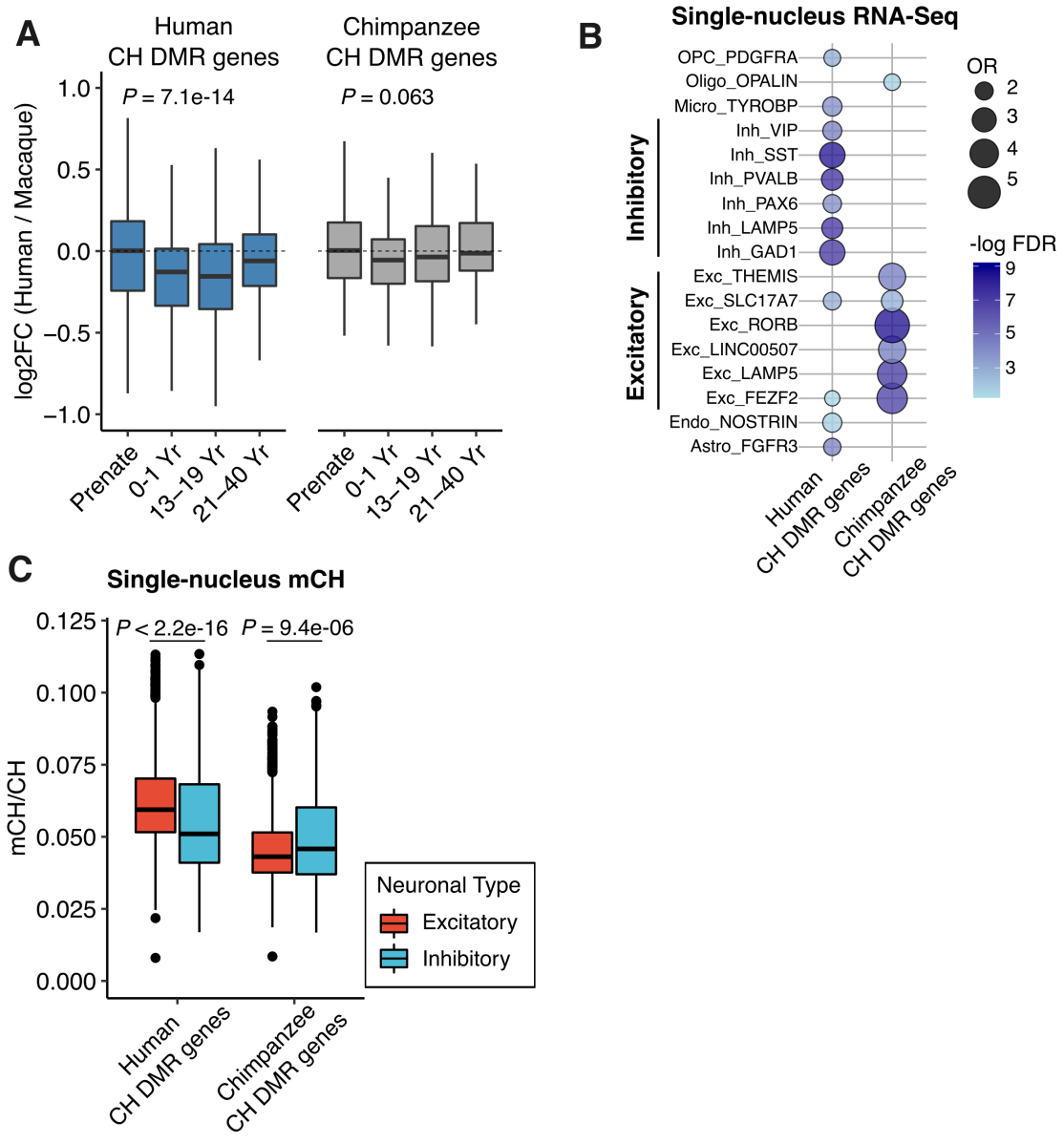

Figure 3.3 CH methylation has profound influence on developmental and cellular specificity. (A) Gene expression fold-change between human and macaque in CH DMR genes across developmental time points (Human CH DMR genes, n= 450 and Chimpanzee CH DMR genes, n=144). Macaque samples were age-matched to human developmental time points. Statistical significance was computed using Kruskal-Wallis test (two-sided). **(B)** Enrichment of human and chimpanzee CH DMR genes in specific cell-types. Human CH DMR genes are enriched in inhibitory neurons whereas chimpanzee CH DMR genes are enriched in excitatory neurons. In each gene set, genes expressed in at least 50% of the cells that are statistically significant (FDR < 0.05 and log2FC > 0.3) are included. Cell-type data are from human medial temporal gyrus (MTG). OR = Odds Ratio. **(C)** CH methylation of neuronal subtypes for CH DMR genes using methylation of single nuclei from the human frontal cortex. Human CH DMR genes are hypomethylated in inhibitory

neurons whereas chimpanzee CH DMR genes are hypomethylated in excitatory neurons (excitatory neurons, n = 1879 and inhibitory neurons, n = 861). Statistical significance was computed using two-sided Mann-Whitney U-test. Box represents a range from the first quartile to the third quartile.

Interestingly, human CH DMR genes are significantly enriched in gene sets representing inhibitory neurons, based on single-nucleus transcriptome data from the middle temporal gyrus (Hodge, Bakken et al. 2019) (Figure 3.3B), as well as those previously identified as markers of inhibitory neurons (Wonders and Anderson 2006, Lein, Hawrylycz et al. 2007, Luo, Keown et al. 2017) (FE = 5.3, $P < 0.0001$, permutation test). Moreover, these genes were more highly methylated in excitatory neurons than in inhibitory neurons in single-nucleus DNA methylation data from the same brain region (Luo, Keown et al. 2017) (Figure 3.3C and Figure B. 7). Integrating these observations, we hypothesize that human-specific CH methylation of inhibitory-neuron-specific genes may silence their expression in the genomes of excitatory neurons, thereby promoting functional specificity of neuron subtypes. Alternatively, there may have been a substantial shift of cell type composition in the human brain since the divergence from chimpanzees, to increase the ratio of excitatory to inhibitory neurons.

3.3 Discussion

We show that CH methylation is significantly higher in human and chimpanzee prefrontal cortex neurons compared to rhesus macaque and mice. Moreover, human prefrontal cortex neurons have higher CH methylation than chimpanzees and rhesus macaques. Although more data from brain tissues of a wider variety primates and other mammals are necessary to fully understand evolutionary dynamics of DNA methylation,

our observation suggests that CH methylation in the prefrontal cortex neurons has increased during the evolution of primates. CH methylation is highly negatively correlated with gene expression and is a strong predictor of gene-expression divergence between neurons of different species. Consequently, the evolutionary trajectory of increasing CH methylation during primate brain evolution may have contributed to shaping finer resolution transcriptional identities of cell types. In this regard, yet a further human-brain specific increase of CH methylation is intriguing. Based on joint analyses of CG and CH methylation, we show that these distinctive cytosine contexts both contribute additively to the human brain transcriptional program. Integrating our results with developmental bulk tissue data and single-cell functional genomics data from human brains, we show that the human-specific increase of CH methylation appears particularly important for early human brain development, and fine-tuning of neuron subtype cell identities.

Due to the limitation of bisulfite sequencing, our data cannot separate methylcytosines from hydroxymethylcytosines (hmCs), which might play distinctive roles in neuron subtypes (Kozlenkov, Li et al. 2018). While additional data are needed, currently available maps (Wen, Li et al. 2014) do not suggest a significant impact of hmC on the differential methylation patterns identified in this study.

3.4 Methods

3.4.1 Identification of CH differential methylation

Unlike CG methylation, >70% of cytosine positions were conserved among the three species. Thus, we used orthologous cytosines across the three species to infer differentially methylated positions. Because CH methylation is sensitive to bisulfite conversion rate (Warnecke, Stirzaker et al. 2002), we only used individuals with high bisulfite conversion rates (>99.5%). We down-sampled and matched sample size across the species to avoid any bias derived from the different sample sizes across groups (N=11 for each species and cell-type). We removed sites in which >50% of individuals in at least one group have fewer than 5 read counts.

For each CH site, we fitted a generalized linear model using the arcsine function to identify differentially methylated CH positions among species adjusting for other covariates (age, sex, and bisulfite conversion rate) using DSS. To fit our parsimonious approach, we also performed pair-wise analyses between species considering all combinations (i.e. human vs. chimpanzee, human vs. macaque, and chimpanzee vs. macaque). Benjamini–Hochberg correction (FDR) was used to perform multiple comparisons. We used the parsimonious approach to detect species-specific methylation changes with a cutoff of fractional methylation difference between species > 10% and FDR < 0.05. For example, human-specific CH methylated sites showed FDR < 0.05 from both human vs. chimpanzee and human vs. macaque comparisons and FDR > 0.05 from the chimpanzee vs. macaque comparison as well as a >10% difference of fractional methylation in humans compared to both chimpanzee and macaque fractional methylation levels.

To identify human-specific and chimpanzee-specific CH DMRs, we identified significantly differentially methylated regions between human and chimpanzee using the

differentially methylated positions generated from a human-chimpanzee comparison. We considered a minimum region of 50bp with at least 4 significant differentially methylated positions (FDR < 0.05) and covering >10 cytosines. Similarly, we used an average methylation difference of 10% as a cutoff. Using average methylation of macaque from corresponding regions, we detected human-specific and chimpanzee-specific CH DMRs using the following criteria. Human-specific CH DMRs are defined as DMRs that show a significant human-chimp difference with at least 4 differentially methylated positions as well as a methylation difference between human and macaque of >5% that is also greater than the methylation difference between chimpanzee and macaque. Similarly, chimp-specific CH DMRs are DMRs that satisfy the following criteria: a significant human-chimp difference with at least 4 differentially methylated positions and a methylation difference between chimpanzee and macaque of >5% that is also greater than methylation difference between human and macaque. To obtain regions in which both human and chimpanzee were differentially methylated compared to macaque, we checked the overlap between human-macaque CH DMRs and chimpanzee-macaque CH DMRs.

3.4.2 Identification of DMR genes

To identify differentially methylated genes, we extracted genes with at least one DMR within a 3kb window upstream and downstream of the gene body. To remove redundant genes among different categories of DMR genes, we used average gene body methylation as an additional indicator to assign genes into the DMR gene category using the following criteria. Human-specific hyper CH DMR-genes are defined as DMR genes that include at least one human-specific hyper CH DMR and show higher average gene body methylation compared to the average gene body methylation of chimpanzee and

macaque. Also, the absolute methylation difference between human and macaque should be greater than the methylation difference between chimpanzee and macaque.

3.4.3 CH methylation of neuronal subtypes

We examined methylation patterns of neuronal subtypes for CH DMR genes. Average gene body methylation of CH DMR genes was calculated for neuronal cells from 21 human neuronal subtypes (Luo, Keown et al. 2017). For the marker gene analysis of neuron subtypes, we used known excitatory and inhibitory neuron markers from Luo et al. 2017 (Luo, Keown et al. 2017). We included the marker genes that are orthologous to the three species. These include 20 excitatory neuron markers (SATB2, TYRO3, ARPP21, SLC17A7, TBR1, CAMK2A, ITPKA, ABI2, RASAL1, FOXP1, SLC8A2, SV2B, PTPRD, LTK, LINGO1, NRGN, NPAS4, KCNH3, BAIAP2, ARPP19) and 13 inhibitory neuron markers (ERBB4, GAD1, SLC6A1, CCNE1, EPHB6, KCNAB3, LPP, TBC1D9, DUSP10, KCNMB2, UBASH3B, MAF, ANK1).

CHAPTER 4. DYSREGULATION OF CELL-TYPE EPIGENETIC IDENTITY ASSOCIATED WITH AGING IN HUMAN BRAIN

4.1 Introduction

It has been observed for several decades that aging has a profound influence on DNA methylation. Studies dating back several decades have demonstrated that DNA methylation levels of specific CpGs were subject to age-associated changes (e.g., (Wilson Vincent and Jones Peter 1983, Hoal-van Helden and van Helden 1989, Richardson 2003)).

Technical advances in the last decade led to the development of relatively cost-friendly microarray methods to study DNA methylation of a large number of CpGs. Subsequently, DNA methylation profiling of large cohorts and from different tissue types followed (e.g., (Rakyan, Down et al. 2010, Hernandez, Nalls et al. 2011, Bell, Tsai et al. 2012, Heyn, Li et al. 2012, Numata, Ye et al. 2012, Hannum, Guinney et al. 2013)). To study the effect of aging on DNA methylation, these studies typically performed either correlation and/or linear regression analyses between DNA methylation and age, identifying numerous CpGs that showed significant variation with aging. The results of these studies solidified that aging has fundamental impacts on DNA methylation.

It was shown that DNA methylation of genetically identical monozygotic twins also diverges with aging (Fraga, Ballestar et al. 2005), indicating that aging-associated DNA methylation changes are not necessarily programmed in the genome. Rather, the term ‘epigenetic drift’ is often used to refer to changes of DNA methylation, and of other epigenetic marks that occur during aging (Cooney 1993, Egger, Liang et al. 2004,

Teschendorff, West et al. 2013). While there was some earlier disagreement over the nature of epigenetic drift regarding whether it involves a decrease or increase of DNA methylation, it became apparent that both patterns were prevalent. Promoters and CpG islands, which tend to have lower levels of DNA methylation (e.g., (Saxonov, Berg et al. 2006, Weber, Hellmann et al. 2007, Elango and Yi 2008)), are often subject to hypermethylation, while intergenic/repetitive regions with higher DNA methylation tend to experience hypomethylation (Teschendorff, West et al. 2013, Jones, Goodman et al. 2015, Sun and Yi 2015). These patterns support the idea that epigenetic drift might be due to gradual dysregulation of epigenetic maintenance over the lifespan, a pattern we will demonstrate more clearly in this work using nucleotide-resolution data of nearly all CpGs in the human genome.

Another exciting development in aging-associated DNA methylation research is the development and application of the so-called 'DNA methylation clocks' (Hannum, Guinney et al. 2013, Horvath 2013, Horvath and Raj 2018, Levine, Lu et al. 2018, Bell, Lowe et al. 2019, Mammalian Methylation, Lu et al. 2021). Briefly, these are subsets of CpGs whose DNA methylation can be used as predictors of age. They are often identified using supervised machine learning methods with DNA methylation values as independent variables and age as the dependent variable. DNA methylation clock studies require large cohorts, and each clock can be constructed using data from single tissue or multiple tissues, using either biological age or phenotypic age derived from multiple measures (Horvath and Raj 2018, Bell, Lowe et al. 2019). DNA methylation clocks are known to be remarkably robust and sometimes perform better than other traditional predictors of biological aging

(Horvath and Raj 2018). DNA methylation clocks have wide-ranging applicability for the study of human health and medicine.

These two aspects of aging-associated DNA methylation changes, namely epigenetic drift and DNA methylation clocks, open many questions and opportunities to study aging from the perspective of epigenetic programs over the lifespan (Jones, Goodman et al. 2015, Horvath and Raj 2018, Bell, Lowe et al. 2019). At the same time, there are several current deficiencies of knowledge that are critical to fully understanding and utilizing these patterns. For example, is the epigenetic drift a feature of all CpGs in the genome? How different or similar are epigenetic drift occurring in distinct tissues and cell types? What is the relationship between epigenetic clocks and epigenetic drift? Fundamentally, how do the epigenetic drift and epigenetic clocks relate to the underlying biological mechanisms of aging?

To address these questions, we need to extend the study of aging-associated DNA methylation changes to the whole genome, using methods developed to examine all genomic CpGs, such as the whole-genome bisulfite sequencing (WGBS). Ideally, these studies should be conducted in large cohorts, to enable the analyses of epigenetic clocks and a more refined analysis of phenotypic traits. Another key missing pieces of information in addressing these issues is understanding DNA methylation at cellular resolution (Horvath and Raj 2018, Bell, Lowe et al. 2019). As epigenomic studies begin to reveal tremendous heterogeneity of cellular epigenetic programs, it is necessary to evaluate how aging-associated DNA methylation changes occur in different cell populations.

Here, we present our analyses of extensive whole-genome bisulfite sequencing data sets of DNA methylation from neuronal nuclei, separated by fluorescence-activated nuclei sorting, from 77 postmortem brains of individuals ranging from neonate to 85 years old. We identify CpGs that show the most DNA methylation changes with aging in neurons, using a regression method specifically developed for the analysis of WGBS data. By doing so, we demonstrate genomic patterns of epigenetic drift in neurons. We further contrast these results with those of non-neuronal cells, including oligodendrocytes.

Our study first demonstrates a clear relationship between DNA methylation levels and age-associated DNA methylation change in the genome-wide scale for both cell populations. While the overall trend is consistent in the two cell populations, the specific CpGs that show aging-associated methylation changes are distinct, highlighting the cell-type specific nature of DNA methylation. Moreover, we show that specific DNA methylation landscapes of distinct cell types become less distinct with aging, as a consequence of epigenetic drift and the divergent DNA methylation landscapes of different cell types. In other words, aging is associated with dysregulation of cell type specific epigenetic identities. Given the emerging significance of cell type specificity in disease and development, dysregulation of cell type specific epigenetic identity may be a meaningful component of aging.

4.2 Results

4.2.1 Age is a major driver of DNA methylation change in the WGBS data

To gain insight into the aging programs in the human brain at cell type resolution, we examined DNA methylation of two major brain cell populations - neurons and oligodendrocytes, separated using fluorescence-activated nuclei sorting method. Specifically, we analyzed 127 whole-genome bisulfite sequencing (WGBS) data sets of neurons (NeuN+, N=77), oligodendrocytes (OLIG2+, N = 42) and non-neuronal cells (NeuN-, N = 8) from the dorsolateral prefrontal cortex, collected from two independent studies (Mendizabal, Berto et al. 2019, Price, Collado-Torres et al. 2019). These data are from individuals across a broad age span, ranging from 2.4 months to 85 years; mean = 43.5 years (Table C. 1).

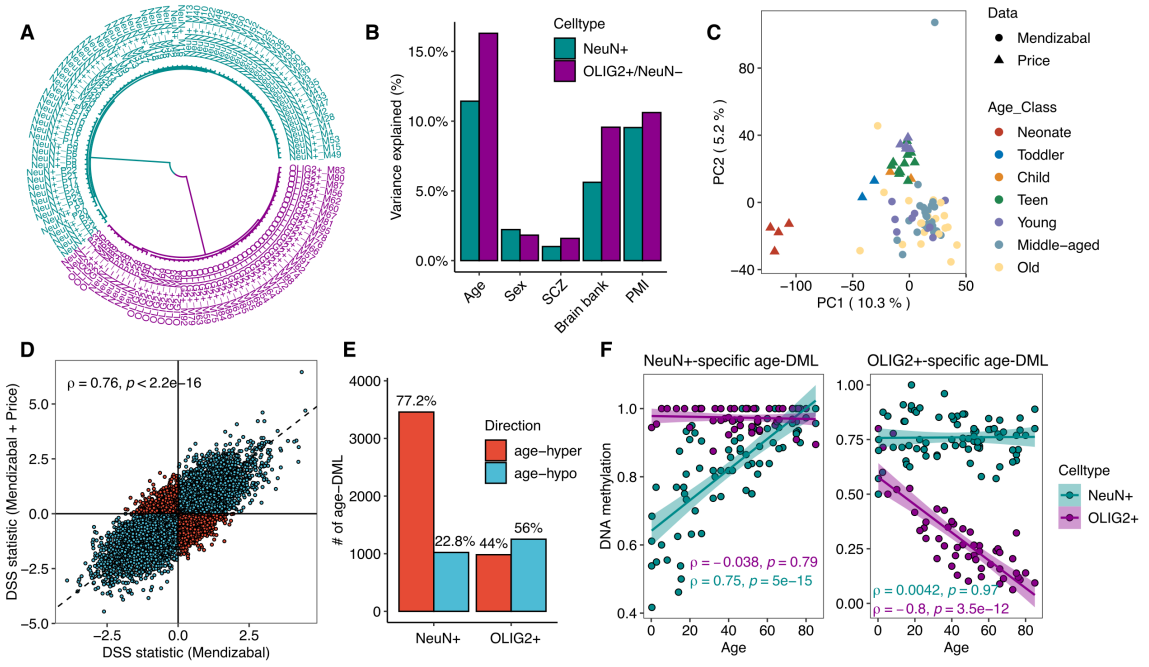

Figure 4.1 Age is a main determinant of DNA methylation variation. (A) Circular dendrogram plot resulting from hierarchical clustering of 127 whole-genome bisulfite sequencing (WGBS) data used in this study. Branches and nodes were colored based on the number of clusters setting to k=2. **(B)** The proportion of variation explained by each

variable controlled across 10 principal components (SCZ, schizophrenia; PMI, post-mortem interval). **(C)** Principal component analysis of whole-genome CpG methylation values. The first principal component (PC1) which explains 10.3% of the total variation of DNA methylation distinguishes the age groups. **(D)** Concordance of DSS statistics between results from the combined data set (Mendizabal and Price) and results from using only Mendizabal data set. CpG sites with disconcordant DSS statistics from two analyses were colored in red. **(E)** Barplot shows the numbers and percentage of significant age-associated CpG methylation changes (age-DML) in NeuN+ and OLIG2+. **(F)** Examples of NeuN+ or OLIG2+ specific age-DML. DNA methylation indicates fractional methylation of the CpG site.

To avoid erroneous methylation calls due to genetic polymorphism at cytosine bases, we first mapped all the matched whole genome sequencing data and excluded positions that were polymorphic at cytosines (Methods). Consequently, we were able to determine DNA methylation levels of 23.6 million CpGs from these data. Hierarchical clustering analysis indicated that NeuN+ samples from the two data sets cluster together, while OLIG2+ and NeuN- cluster together, reaffirming cell identities (Figure 4.1A).

In both data sets, age was the main determinant of variation of DNA methylation. Overall, age explained a greater amount of DNA methylation (11.4% and 16.3% for NeuN+ and OLIG2+, respectively) compared to other factors such as disease status (<1.6%) and sex (<2.2%, using CpGs in autosome) (Figure 4.1B). Using only Price et al. (Price, Collado-Torres et al. 2019) data set, age explained an even greater amount of DNA methylation variation (23.4% and 23.6% for NeuN+ and NeuN-, respectively) than in the total data set (Figure C. 1). It should be noted that samples from the Price et al. (Price, Collado-Torres et al. 2019) are biased toward relatively young age groups (neonate to early 20s) while those from Mendizabal et al. (Mendizabal, Berto et al. 2019) range from late 20s to 80s (Table C. 1). Given that DNA methylation dramatically changes during early

development (Price et al. 2019, other papers), the amount of DNA methylation explained by age in Price et al. data set would also include the effect of the shift of DNA methylation during early development. Accordingly, principal components and pairwise correlation coefficients showed that DNA methylation levels are influenced by both the difference between early age and the two studies (Figure 4.1C). In addition to the developmental shift of DNA methylation, the two data sets are generated from two different labs, thus subject to slightly different protocols, sequencing platforms, and other potential yet unknown differences. Also, phenotypic attributes associated with data are distinct between the two data sets, thus making it difficult to control for confounding factors while performing statistical tests. Consequently, even though cell identities and the major impact of age on DNA methylation are consistent between the two data sets, we treated the two data sets separately in the subsequent analyses.

4.2.2 CpGs that vary with aging are highly cell type specific

To characterize age-associated DNA methylation changes, we investigate aging-associated DNA methylation changes using data from Mendizabal et al. (2019). This data set contains samples from the aging lifespan, rather than developmental shift. We applied a generalized linear model framework developed to analyze WGBS data (Park and Wu 2016) and considered covariates including the postmortem interval, sex, disease status, and bisulfite conversion rates (similar to in Mendizabal et al. 2019, Methods).

Following these procedures, we identified 4,480 and 2,253 CpGs that show significant age-associated methylation changes in NeuN+ and OLIG2. These CpGs were detected using the cutoff of FDR < 0.1, which corresponds to $P< 3.18 \times 10^{-5}$ and $P<1.52 \times$

10^{-5} for NeuN+ and OLIG2+, respectively. They are henceforth referred to as 'age-differentially methylated loci' or simply 'age-DMLs'. We also performed the same analysis after combining Mendizabal et al. and Price et al. data sets. Test statistics from Mendizabal et al. and the combined data sets are strongly correlated (Spearman's rho = 0.76) (Figure 4.1D). For example, we observed that 73% of age-DMLs resulting from Mendizabal et al. are overlapped with DMLs resulting from the combined sets, suggesting consistent patterns of age-associated DNA methylation changes in the two data sets.

DNA methylation changes with age show distinct patterns between two cell types. In NeuN+, the majority of age-DMLs exhibit hypermethylation (77.2%, also referred to as 'age-hyper DMLs'). In contrast, age-DMLs in OLIG2+ are biased toward hypomethylation (55.6%, also referred to as 'age-hypo DMLs') (Figure 4.1E). Aging DMLs of the two cell types had only 16 overlapping CpGs, demonstrating highly distinctive sets of age-associated differentially methylated sites between NeuN+ and OLIG2+. Some examples of aging-DMLs are shown in Figure 4.1F.

The number of age-DMLs highly correlated with the size of chromosomes (rho = 0.9) (Figure C. 2), with a few outliers, notably chromosomes 1, 4, and 13. This pattern is not caused by clusters of adjacent CpGs with similar DNA methylation levels (i.e., (Lister, O'Malley et al. 2008, Huh, Yang et al. 2014)); when we re-examined the pattern after excluding age-DMLs within 1kbp of any other age-DMLs, the differences remained.

4.2.3 Genomic patterns of epigenetic drift

Previous studies have demonstrated that genome-wide DNA methylation changes with aging could be explained by the so-called epigenetic drift, or random changes of DNA

methylation. In particular, aging associated epigenetic drift may reflect decreased efficiency of DNA methylation maintenance, therefore leading to an increase or decrease of DNA methylation for sites that are initially lowly or highly methylated (Teschendorff, West et al. 2013, Jones, Goodman et al. 2015, Sun and Yi 2015).

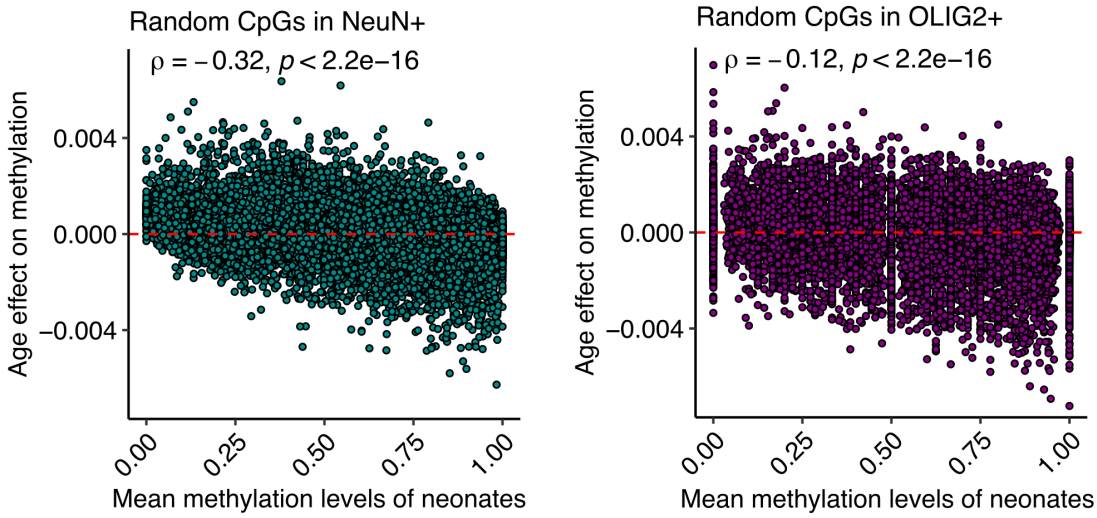

Figure 4.2 The initial methylation level is a major determinant of epigenetic drift. For each CpG site, we fitted a linear model to estimate the age effect on DNA methylation adjusted for other biological variables (post-mortem interval, sex, disease status, and bisulfite conversion rate). Y-axis indicates mean methylation levels of neonates for the corresponding CpG site. For an illustration purpose, only 10,000 randomly selected CpG sites were displayed in the plots.

Our data on DNA methylation on a nearly entire set of CpGs in the genome offer an unprecedented opportunity to examine this in a truly genome-wide fashion. We examined age-associated DNA methylation changes of CpG positions in relation to their DNA methylation states. We used an average methylation level of neonates as a proxy for the

putative initial methylation state and asked if the direction of epigenetic drift is related to the initial level of DNA methylation. We found that the initial methylation state and the direction of aging-associated change were significantly negatively correlated (Spearman's rho = -0.32, $P< 2.2X10^{-16}$, Figure 4.2). In other words, CpG positions that were lowly methylated in neonates tended to increase methylation with age, while those with heavy initial methylation tended to lose DNA methylation. It should be noted, however, that the relationship between these two variables is messy and by no means represent a perfect correlation. Nevertheless, the genome-wide trend is clear. When we used the mean methylation of individuals with ages less than 20 as a predictor, we found similarly significant patterns (Figure C. 3). These results indicate that the baseline level of DNA methylation is an important, and likely a major, determinant of age-associated DNA methylation change.

In mammalian and other vertebrate genomes, DNA methylation levels show clear bimodal distribution (Elango and Yi 2008). Given the direction of epigenetic drift we have demonstrated above, it is expected that the genome-wide epigenetic drift diminishes differences between the extreme ends of DNA methylation (since heavily methylated CpGs lose DNA methylation and lowly methylated CpGs gain DNA methylation). Indeed, when we compared DNA methylation distributions across the age groups in adults, we observed that the clear bimodality is more dispersed in old adults, although the distributions vary across individuals (Figure C. 4).

4.2.4 Connecting epigenetic drift with highly cell-type-specific aging differential DNA methylation

Having examined the genome-wide patterns of epigenetic drift, we questioned whether the observed age-DMLs in NeuN+ and OLIG2+ follow the genome-wide pattern of epigenetic drift. Distributions of DNA methylation in age-DMLs show cell-type specific patterns (Figure 4.3). For example, accumulation of DNA methylation (referred to as 'age-hyper' in Figure 4.3) in neurons shift initially lowly or intermediately methylated positions in NeuN+ (but highly methylated in OLIG2+) to a high level of DNA methylation (Figure 4.3A; see also Figure 4.3C for DNA methylation in OLIG2+ for the corresponding CpG sites). Loss of DNA methylation in neurons was observed for positions that are highly methylated NeuN+ while exhibiting a broad range of DNA methylation in oligodendrocytes (Figure 4.3A). On the other hand, accumulation of DNA methylation in the oligodendrocyte genomes mainly occurs in sites where early methylation levels were very low in both cell types, but increase in OLIG2+ only (Figure 4.3B; see also Figure 4.3D for DNA methylation in NeuN+ for the corresponding CpG sites).

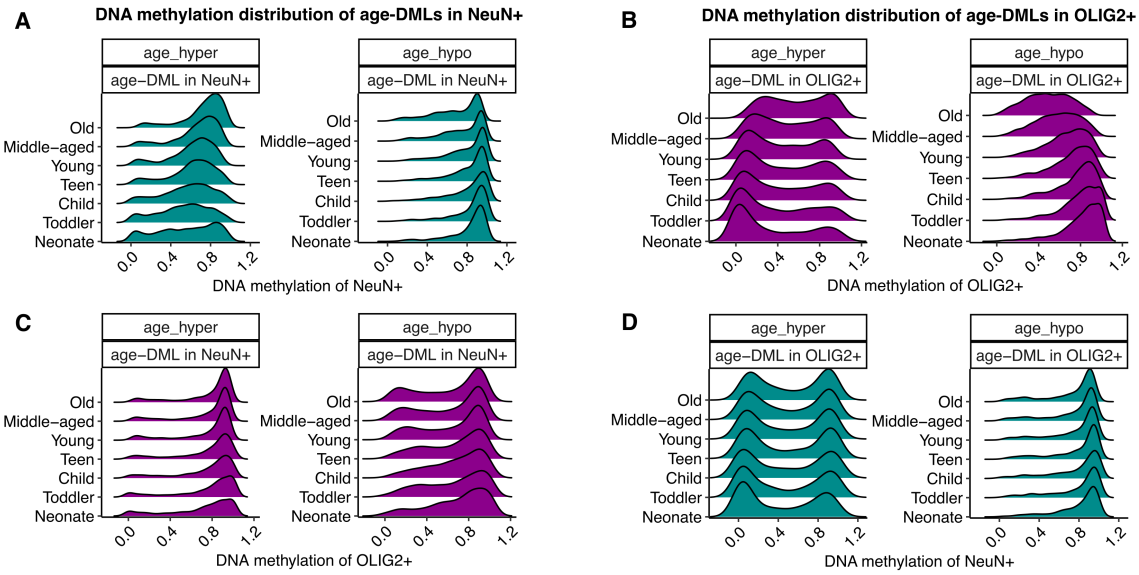

Figure 4.3 Accumulation and loss of DNA methylation with age show highly cell-type-specific patterns. (A) Trajectories of distributions of DNA methylation in NeuN+ age-DMLs. **(B)** Trajectories of distributions of DNA methylation in OLIG2+ age-DMLs. **(C)** and **(D)** DNA methylation distributions of the same CpG positions in the other cell type are also shown for control.

It should be noted that although the two brain cell types showed cell-type-specific patterns of genome-wide aging epigenetic drift, there is a weak but positive correlation of DNA methylation changes with aging between cell types (Figure C. 5). It is not surprising given that brain cell types contain both common and cell-type-specific methylated regions.

4.2.5 Dysregulation of cell-type identity is one potential mechanism of aging

Above we have demonstrated that age-associated changes of DNA methylation occur in distinctive CpGs in NeuN+ and OLIG2+, even though the overall pattern of epigenetic drift follows a common trend. These observations could be reconciled by noting that NeuN+ and OLIG2+ exhibit highly distinctive DNA methylation landscapes. For example,

we have previously shown that more than 20% of the human genome exhibits significant differential methylation between neuron and oligodendrocyte (Mendizabal, Berto et al. 2019) and the majority of these differences are evolutionarily conserved (Jeong, Mendizabal et al. 2021). CpGs that are highly differentiated between cell types tend to be more often identified as age-DMLs (Figure 4.4A). Concordantly, DNA methylation between the cell types shows more pronounced differences in aging-DMLs than randomly selected CpGs (Figure 4.4B). This pattern is stronger in OLIG2+ than in NeuN+.

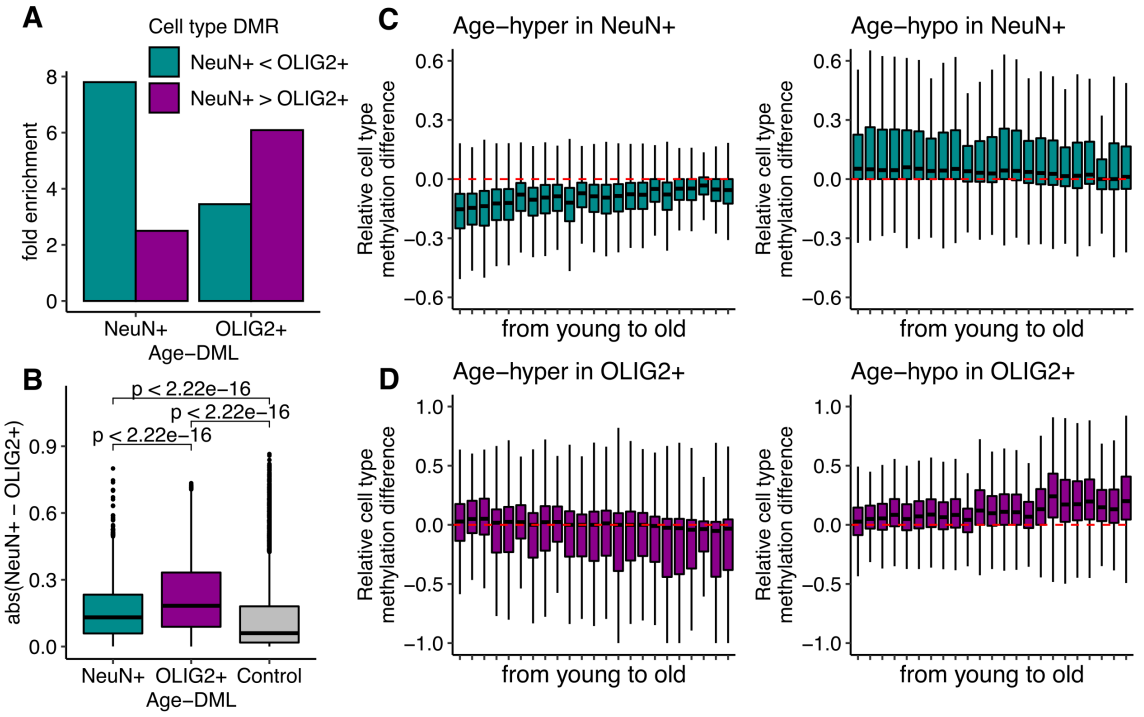

Figure 4.4 Age-associated DNA methylation changes contribute to the dysregulation of cell type identity. (A) Age-DML is highly enriched for the genomic regions differentially methylated between NeuN+ and OLIG2+ cells (cell type DMR). Fold enrichment analysis was performed based on the occurrences of NeuN+ age-DMLs for cell type DMR compared to random control sets (n=100). **(B)** The absolute CpG methylation difference between cell types is significantly higher for age-DMLs compared to control sets with matched G+C nucleotide contents. **(C)** and **(D)** Trajectories of cell type

methylation difference with age. Relative methylation difference between NeuN+ and OLIG2+ was calculated for the samples collected from the same brain tissue.

Since epigenetic drift pushes DNA methylation levels to intermediate levels, epigenetic drift can ultimately dim cell-type methylation differences. To test this hypothesis, for each CpG, we calculated the relative DNA methylation difference between neuron and oligodendrocyte collected from the same brain tissue sample. We then examined the cell-type methylation difference trajectories with age. The relative differences in DNA methylation between cell types tend to gradually decrease with age (Figure 4.4C). This trend is especially pronounced in NeuN+ age-hyper DML. However, we observed that epigenetic drift does not always act in the direction of decreasing DNA methylation differences between cell types. For example, OLIG2+ age-hypo DML showed a slight increase of cell type methylation difference between with age (Figure 4.4D).

Such observed aging-associated DNA methylation changes can alter epigenetic cell type identities. We examined DNA methylation cell type identity of different age groups to test this hypothesis. Specifically, we defined DNA methylation cell type identity as a point in a two-dimensional plane, where one axis is the methylation in neurons and the other axis is the methylation level in oligodendrocytes or non-neurons. We then examined how the DNA methylation cell type identities change with aging, by approximating epigenetic distances between age groups using the Euclidian distance between the above-defined cell type epigenetic identities (Methods for more details). Figure C. 6 demonstrates the results of this analysis. Epigenetic distance of cell type identity from the neonate is depicted for different age groups for age-DMLs. Epigenetic distances of the randomly

selected CpG positions are also shown for control. First, we observe that cell type epigenetic identity changes dramatically from neonates to toddlers, which is explained by the developmental shift of DNA methylation. Second, we show that cell type identity of age-DMLs was drifted further away from the initial epigenetic profile set up in the neonates. For example, the mean epigenetic distance of NeuN+ age-DMLs in old adults is 0.28, while those in control is 0.18 ($P < 2.2 \times 10^{-16}$, Mann-Whitney test). Finally, this trend is more pronounced in OLIG2+ age-DMLs.

4.2.6 Functional consequences of aging DMLs and disease enrichment

We examined functional implications of age-DMLs using genomic annotations, matched RNA-seq data, and comparisons with GWAS variants of different diseases. In both cell types, age-hyper DMLs were significantly enriched in promoters, suggesting their potential impact on the regulation of gene expression (Figure 4.5A). It is well known that promoter methylation is negatively correlated with its corresponding gene expression. Using gene expression data from the matched samples, we found that genes that accumulate DNA methylation in the promoter (i.e., harboring age-hyper DML) are expressed most strongly early in life (Figure 4.5B). In contrast, promoters with age-hypo tend to increase gene expression with age. In addition, the hypomethylating age DMLs exhibited strong enrichment in brain enhancers in the dorsolateral prefrontal cortex inferred from Roadmap epigenomics data.

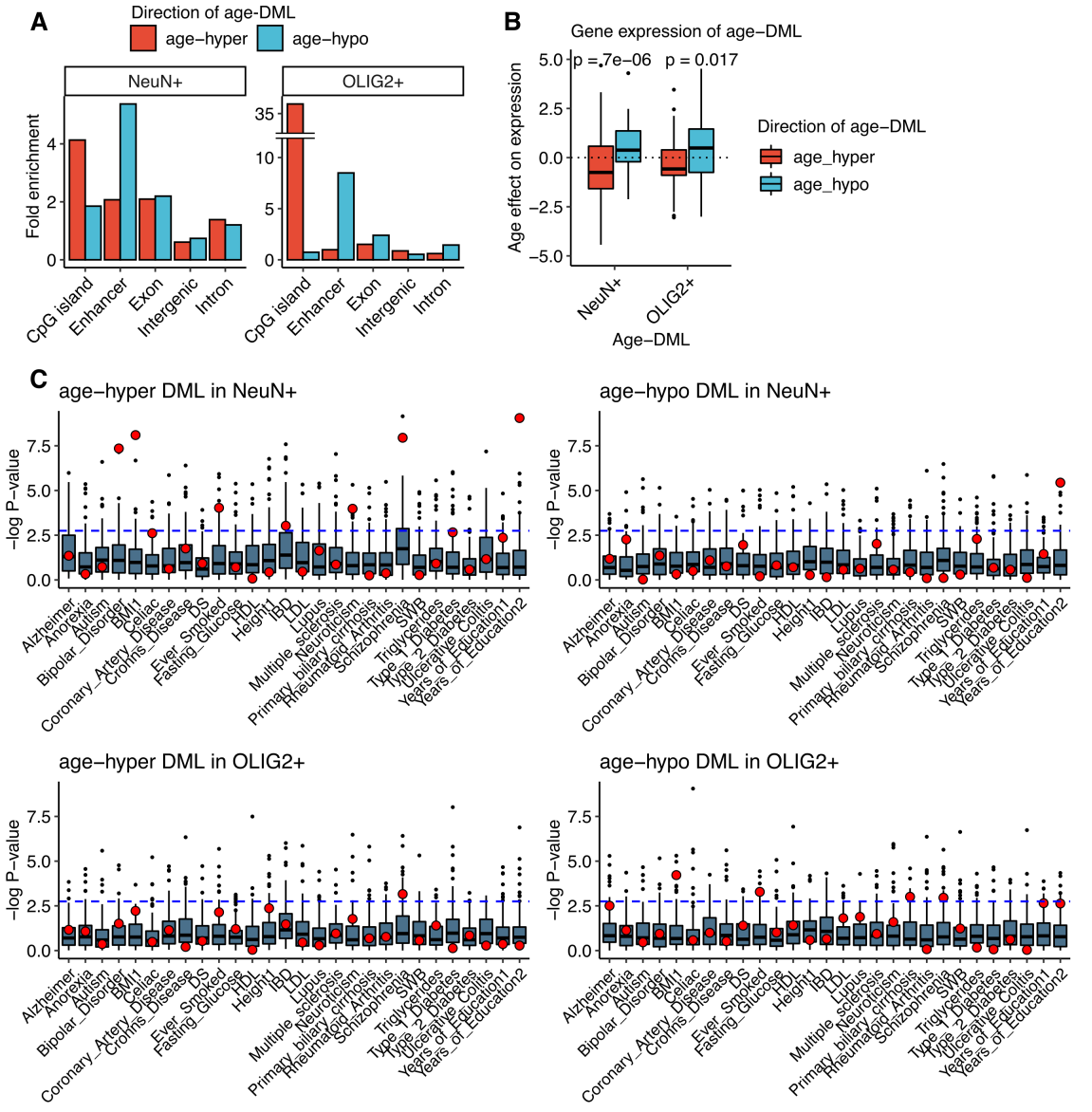

Figure 4.5 Functional implication and brain trait heritability of age-DMLs. (A) Distribution of Age-DML in different functional genomic regions. Age-hyper DMLs are highly enriched for promoter regions, while age-hypo DMLs are enriched for brain enhancers. Fold enrichment analysis was performed based on the occurrences of age-DMLs for each feature compared to random control sets (n=100). **(B)** The effect of age on gene expression for age-DML using the matched RNA-Seq data (Method). **(C)** Significance levels for genetic heritability in different age DML and complex traits. Red dots indicate the statistical significance of age-DML for the traits. Boxplots represent results from random control sets (Method).

We also find that age-DMLs are enriched in distinctive transcription factor (TF) binding motifs. Within the significant binding motifs was the nuclear factor 1 A-type (NFIA). Its expression drives neural stem cells toward astrocyte fate by regulating Gfap expression (Cebolla and Vallejo, 2006, Glasgow et al., 2014, Kang et al., 2012). Some age-DMLs occurred nearby, indicating the presence of genomic regions whose DNA methylation undergoes aging-associated changes (Figure C. 7). Interestingly, 21 OLIG2+ age-hypo DMLs are clustered in a window (chr1: 226.83Mbp – 226.92Mbp) residing on ITPKB locus and significantly overlapping with a significant GWAS risk locus of Parkinson's disease (rs16846351).

Given that cell type specificity is implicated in disease and development, we further examined cell-type aging epigenetic drift may contribute to genetic heritability associated with human diseases. We performed the stratified linkage disequilibrium score regression (Finucane, Bulik-Sullivan et al. 2015) to estimate the contribution of age-DMLs to disease and other complex traits using GWAS summary statistics (Figure 4.5C). Previously, we showed that differentially methylated regions between neurons and oligodendrocytes contribute to neuropsychiatric and neurodegenerative disorders (Mendizabal, Berto et al. 2019). To avoid a bias derived from the DNA methylation between neurons and oligodendrocytes, we compared the results with two control sets controlling for the cell-type methylation difference and GC ratios (Methods). We found a significant enrichment for various brain disorders (e.g., bipolar disorder, neuroticism, and schizophrenia) in NeuN+ age-hyper DMLs (Bonferroni adjusted P-value < 0.05). Educational attainment and body mass index (BMI) also showed significant enrichment.

4.2.7 Relationship with DNA methylation clocks

Recent studies of age predicter models from DNA methylation arrays from multiple studies and tissues showed that chronological ages could be accurately predicted (within four years) based on the DNA methylation of a few hundred CpG positions (Horvath 2013, Levine, Lu et al. 2018). Our data, while providing nucleotide-resolution DNA methylation data from the largest number of CpGs possible, lack the adequate sample size to construct a DNA methylation clock. Nevertheless, we examined whether the DNA methylation age predictor could estimate biological ages close to the known chronological ages of our WGBS samples. We performed the Horvath multi-tissue DNA methylation age clock (Horvath 2013), the most well-known age predicter.

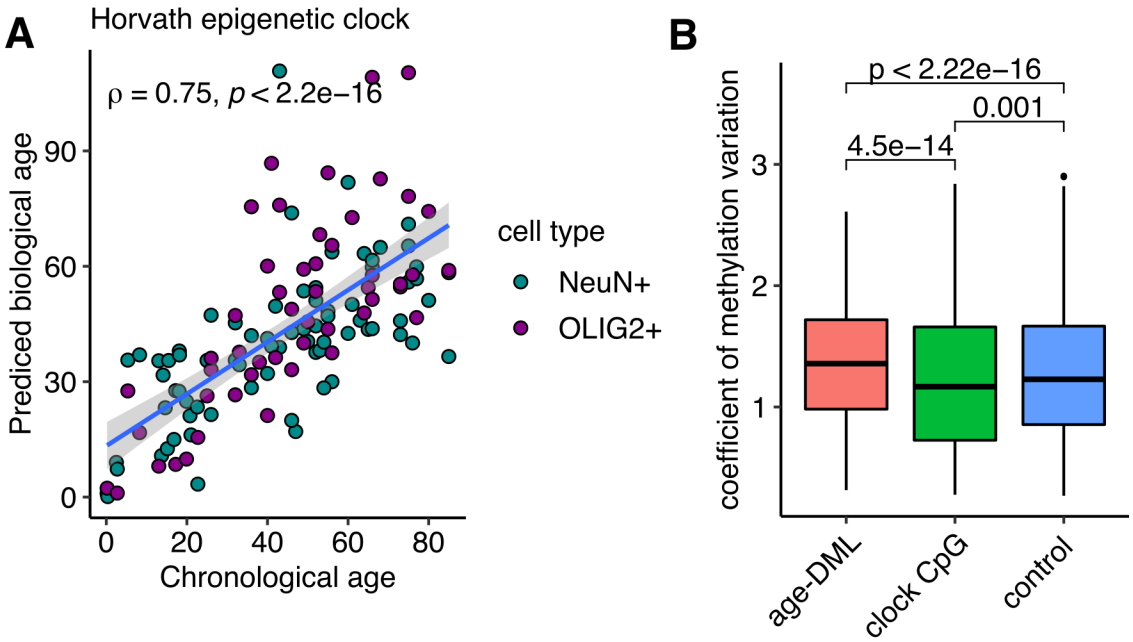

Figure 4.6 DNA methylation variation of multiple tissues in clock CpGs. (A) Horvath estimated DNA methylation age is highly correlated with a chronological age of WGBS samples. **(B)** The coefficient of methylation variation resulting from 10 different WGBS

tissues shows highly variable DNA methylation for age-DML and reduced DNA methylation variation for clock CpGs.

The Horvath clock accurately predicts methylation ages very similar to chronological ages in our WGBS samples (Figure 4.6A). The correlation between the estimated DNA methylation ages and the chronological ages is 0.76 and 0.72 for NeuN+ and OLIG2+, respectively, reaffirming that clock CpGs can accurately predict ages from multiple tissues. Given that we have demonstrated that age-DMLs are highly distinct between cell types, we hypothesized that clock CpGs might represent those that show less variability between tissues compared to age-DMLs. Indeed, the multi-tissue age predictor clock CpGs such as Horvath (Horvath 2013) and Levine (Levine, Lu et al. 2018) exhibited reduced DNA methylation variation in 10 tissues (Methods), while age-DML found in neurons and oligodendrocytes show significantly higher methylation variation compared to clock CpGs and randomly selected CpGs (Figure 4.6B and Figure C. 8).

4.3 Discussion

Our study is, as far as we are aware, the first to examine aging-associated DNA methylation changes at the whole-genome scale (23.6 million CpGs) from distinctive cell populations. By doing so, we first show that aging explains a significant amount of variation observed in DNA methylation. Across the two different data sets, aging explained more than 10% of the total variation in DNA methylation, which is greater than the estimated effect of sex in both data sets, and the effect of schizophrenia diagnosis in

Mendizabal et al.'s data. Moreover, we show that the majority of age-associated variation in DNA methylation can be explained by the effect of DNA methylation in early ages. Using DNA methylation levels at neonates as a proxy, nearly 8.1% of the total variation in genomic DNA methylation could be explained by the level of DNA methylation in the early stage.

The term 'epigenetic drift' has been used in the context of age-associated DNA methylation changes (Cooney 1993, Egger, Liang et al. 2004, Teschendorff, West et al. 2013, Sun and Yi 2015). The term drift itself explicitly refers to random changes, and epigenetic drift implicates stochastic change of DNA methylation. However, having demonstrated that the direction of DNA methylation change is significantly related to the DNA methylation level in early age, our study indicates that a significant amount of aging-associated DNA methylation changes is non-random.

Epigenetic drift was thought to occur due to the noise introduced by the imperfect DNA methylation machinery accumulated in every cell cycle (e.g., Cooney 1983). We show here that neurons show age-associated DNA methylation changes even though they are largely devoid of cell divisions. Therefore, epigenetic drift does not necessarily associate with cell division.

Emerging studies, including those from our group, have shown that DNA methylation and other epigenetic landscapes of different cell types are highly distinct, which in turn are enriched in positions implicated in diseases (Mendizabal et al. 2019, others). As we have demonstrated above, age-associated DNA methylation changes are dependent on DNA methylation levels in early stages. Consequently, CpGs that show DNA

methylation change with aging is likely to be highly divergent between different cell types, as we demonstrate here for between neurons and oligodendrocytes.

One of the most exciting recent development in aging research in general is the presence of Aging CpG clocks, which are subsets of CpGs can predict the biological and phenotypic ages with high accuracy. How can we reconcile aging CpG clocks with our findings that aging-associated DNA methylation changes occur in a highly cell type specific manner? It should be noted that our study and aging CpG clocks use fundamentally different approaches. The goal of aging CpG lock studies is to identify predictors of aging from the CpGs that are present in large studies, which are typically those included in widely used DNA methylation arrays. The specific CpGs within the clocks often do not show a higher correlation with age than other CpGs (Horvath 2013). Rather, these two types of CpGs represent two different aspects of aging-associated changes of DNA methylation. Nevertheless, interestingly, clock CpGs appear to have less variability of DNA methylation between tissues compared to randomly selected CpGs as well as aging-DMLs. Among the clock CpGs, those selected from a single tissue (Hannum clock) show a similar level of variability with the cell-type aging-DMLs.

4.4 Methods

4.4.1 Whole-genome bisulfite sequencing data processing

To investigate comprehensive brain cell-type DNA methylation changes with aging, we used our previously published WGBS data set of 53 NeuN+ and 42 OLIG2+, which

comprised samples from individuals of different ages (range from 25 yrs to 85 yrs). We also collected brain cell-type WGBS data from neonates, toddlers, and teens to examine DNA methylation trajectories across different age groups and infer the initial cell-type DNA methylation state.

WGBS reads were quality trimmed with TrimGalore v.0.4.1 (Babraham Institute) using a default setting. The trimmed reads were mapped to the human reference genome (hg19) using Bismark v.0.14.5 (Krueger and Andrews 2011) with Bowtie 2 mode. Duplicated reads were further processed and filtered by the deduplicate module in Bismark (deduplicate_bismark). We removed lowly mapped CpG positions (average mapped reads less than 5).

Principal component analysis of methylated cytosines was performed using randomly selected 2 million CpG sites. To retain informative CpG sites, we excluded the positions with the mean fractional methylation either less than 0.1 or greater than 0.9. To compute the proportion of the overall variance of DNA methylation explained by covariates, we regress each principal component on all covariates, take the coefficient of determination multiplied by the fraction of variance that the principal component captures, and sum over results from all principal components. To reduce computation cost, we used the top 10 principal components.

4.4.2 Identification of age-DML

For each CpG position, we fitted a generalized linear model using the arcsine link function to estimate the age effect. The model fitting is conducted utilizing the DSS (ver. 2.4) Bioconductor package (Feng, Conneely et al. 2014). We considered bisulfite

conversion rates, sex, postmortem intervals, and disease status as covariates. A hypothesis test of the age effect for each CpG site was performed using the Wald test using the estimated coefficient and standard error from the fitted model. False discovery rate (FDR) is computed using the Benjamini-Hochberg method.

The coefficients of age from the DSS analyses are estimated from the GLM framework, making it difficult to interpret the explicit meaning of methylation differences compared across CpG sites. Thus, we also fitted general linear models using fractional methylation to estimate biologically interpretable methylation level differences with aging.

4.4.3 RNA-Seq data processing

RNA-Seq data from the matched with WGBS samples were collected from previous studies (Mendizabal 2019, Price et al. 2019). Raw sequencing reads were quality trimmed with Trimmomatic (ver. 0.39) and then mapped to the human reference genome (hg19) using STAR (ver. 2.7) with the following options: --alignSJDBoverhangMin 1 --outFilterMismatchNmax 3 --outFilterMultimapNmax 10 --alignSJoverhangMin 10 --twopassMode Basic. We removed any secondary alignments and duplicated reads using Samtools (ver. 1.13) to ensure that only uniquely mapped reads were retained for further analyses. We calculated the gene expression using htseq (ver. 0.11.2) using intersection-strict mode by the exonic regions. We quantified protein-coding genes using the human Ensembl annotation (GRCh37.87).

4.4.4 Disease heritability using stratified LD score regression

To measure the contribution of age-DML to the genetic risk of disease and complex traits, we performed the stratified LD score regression analysis. We followed the same processing steps and used the list of GWAS traits described in the previous work (Jeong et al. 2021). Briefly, we examined extended genomic regions of age-DML (25 kbp on both sides of focal age-DML) to improve the confidence intervals of the estimates. As a statistical control, we performed partitioned stratified LD score analyses using CpG positions that are differentially methylated between neurons and oligodendrocytes. G+C nucleotide contents were matched based on the GC ratio of 1kbp window (+-500bp of age-DML).

4.4.5 Coefficient of methylation variation

To estimate the coefficient of variation of each CpG site, we processed WGBS data from 9 tissues (placenta, sperm, hair follicle, adrenal gland, liver, colon, ovary, embryonic stem cell, and b-cell). Because CpGs selected from methylation array data are often biased toward promoter regions, coefficient of variation of those CpGs are inflated (because mean methylation is close to 0). Thus, the coefficient of variation was calculated using the corrected equation.

$$\text{Coefficient of methylation variation} = \frac{\sqrt{\sum(x_i - \mu)^2}}{\sqrt{\mu(1-\mu)}}$$

where x_i denotes fractional methylation of i-th tissue and μ denotes the mean methylation value.

CHAPTER 5. DIFFERENTIAL EPIGENETIC AGING ASSOCIATED BY SOCIAL REARING EXPERIENCES DURING EARLY LIFE

5.1 Introduction

One of the most notable advances in DNA methylation research in the last decade is the development of aging CpG clocks, which are subsets of CpGs whose DNA methylation levels can predict chronological and phenotypic ages with high accuracy (Horvath 2013, Horvath and Raj 2018, Horvath, Zoller et al. 2021). These clocks are obtained via machine learning methods to identify reliable predictors of age. Initially developed from single tissues from DNA methylation arrays (Hannum, Guinney et al. 2013), aging CpG clocks are now being obtained from multiple tissues, DNA methylation arrays with a greater number of CpGs, and/or custom DNA methylation arrays (Horvath, Zoller et al. 2021). These clocks offer tremendous diversity of opportunities to study health and aging.

Aging is nearly ubiquitous in the tree of life. Molecular mechanisms of aging thus may share common evolutionary characteristics (Horvath, Zoller et al. 2021). It follows that aging CpG clocks can be also obtained for many species, provided that we can measure DNA methylation levels from a large number of individuals across the lifespan. Indeed, aging CpG clocks are actively being developed from many species (Bell, Lowe et al. 2019, Anderson, Johnston et al. 2021, Bors, Baker et al. 2021, Horvath, Zoller et al. 2021). These aging CpG clocks should provide information on common and divergent mechanisms of aging for different species, as well as candidate genes and pathways to understand species-

specific health and aging trajectories. For example, a study of wild baboons (*Papio cynocephalus*) in the Amboseli ecosystem of Kenya demonstrated that dominance rank in male baboons, but not female baboons, is strongly correlated with the acceleration of aging CpG clock (Anderson, Johnston et al. 2021). This result provides several hypotheses on how to understand aging in male and female baboons, which could yield information the conserved, universal mechanisms of aging.

Apart from the aging CpG clock generated in the Anderson et al. study, there are other aging CpG clocks from four other non-human primates, including the vervet monkey, rhesus macaque, marmoset, and the hamadryas baboon (*P. hamadryas*) (Horvath, Haghani et al. 2020). Aging CpG clocks from the hamadryas baboon hybrids showed high accuracy when used against the human data, emphasizing their potential significance in aging studies.

In this study, we utilized the Specific Pathogen Free (SPF) colony of the anubis baboons (*P. anubis*) at the MD Anderson Center. This colony is specifically developed for biomedical research, with medical and health records concurrently being developed. This specific colony setting provides another unique and interesting biological factor. To create a SPF breeding colony, newborn infant baboons must be separated from non-SPF females to prevent maternal transfer of all undesirable pathogens. When removed from the non-SPF females, the neonates are raised in a human nursery-setting ('nursery-raised' or NR baboons). There are also offspring born to existing SPF females who are not separated and are therefore raised by their biological mother in a nuclear family setting ('mother-raised' or MR baboons). Therefore, these groups of baboons experience different early rearing experience. Previous studies that examined the effect of similarly different rearing

experience discovered several characteristics that lasted into the adulthoods (Perris and Andersson 2000, Goel and Lee 2007, Lomanowska, Lovic et al. 2011).

The goal of this study is three-fold. First, we wanted to develop aging CpG clocks from the anubis baboons, which can be applied to the study of health and aging in this population. Second, we wanted to test if rearing experience during early life affect rates of epigenetic aging. Third, we wanted to further explore the relationships between epigenetic aging and differential DNA methylation with sex, aging and rearing experience.

5.2 Results

5.2.1 DNA methylation-based age estimator accurately predicts chronological age

To better understand the long-term effects of early-life social rearing experiences on epigenetics, we generated reduced-representation bisulfite sequencing (RRBS) data from blood-derived DNA samples of baboons (N = 140; 119 females and 22 males). These data were collected from two distinct groups of early social rearing experience. One group of individuals were raised by their biological mothers (mother-reared, MR), while the other group was raised in a human-nursery setting (nursery-reared, NR). Also, we collected the samples from across broad age span, ranging from 14 months to 19.3 years (mean = 9.6 years) to investigate age effects on epigenetic variation. RRBS reads were mapped to the baboon papAnu4 reference genome to quantify DNA methylation at base level resolution. Using whole-genome sequencing (WGS) data from 100 baboons, we excluded any nucleotide polymorphisms occurring at cytosine sites to avoid incorrect methylation calls

due to the technical limitation of distinguishing bisulfite converted thymine from unmethylated cytosine. After stringent filtering procedures (Methods), we obtained coverage for over 2.6 million CpGs in the baboon genome with an average read depth of >5.

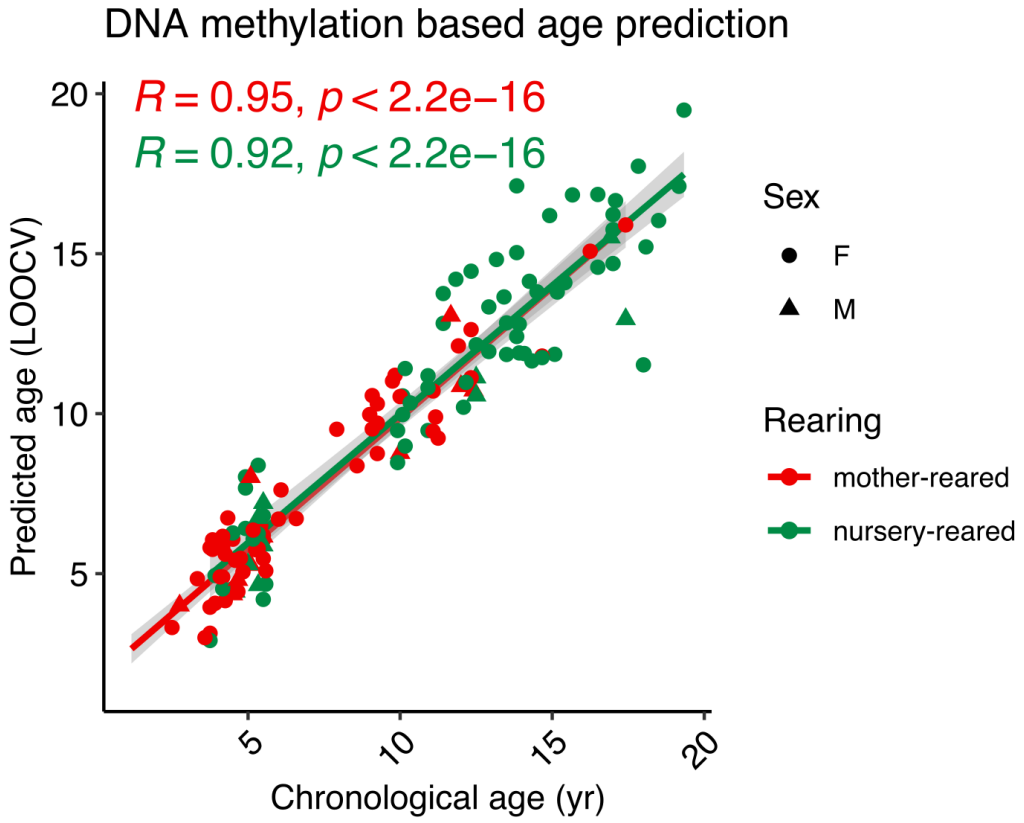

Figure 5.1 Age prediction using DNA methylation of baboon blood samples. Predicted age is estimated using the leave-one-out cross-validation approach (Methods). Solid lines indicate the fitted regression line.

DNA methylation can be used as a strong predictor to estimate the biological age of samples (Horvath 2013). Using our RRBS data, we generated DNA methylation-based baboon age predictor (Figure 5.1). The baboon epigenetic clock accurately predicts age using DNA methylation of 153 CpG sites (R=0.98 and mean error 2.1 years). Consistent with previous findings, the clock CpGs are more enriched in CpG islands than control CpGs that match GC-content (± 500 bp). We found that the clock CpGs consist of both age-hypermethylated (n=89) and age-hypomethylated sites (n=64).

We further examined whether the clock CpGs resulting from our study are located closely with other clock CpGs that are identified from an independent study (Anderson 2021). We found that 34 clock CpGs from our study (22.2%) are overlapping or closely located (< 500 bp) with the other clock CpGs (P-value < 0.01, permutation test). Notably, two epigenetic clock CpGs were generated from different baboon species (Papio Anubis and Papio cynocephalus) and different rearing environments (domesticated vs. wild). This result implies that epigenetic clock CpGs are a robust predictor of age and applicable for other species.

5.2.2 Age-associated methylation changes vary with early life social rearing experience

We performed principal component analyses (PCA) using DNA methylation of the CpG sites to explore the association of biological and environmental factors with DNA methylation. The first PC shows a clear trajectory by age (Figure 5.2A). This result is in agreement with the previous studies that age is a major determinant of DNA methylation changes. Interestingly, we found that there is a distinct separation of old individuals that has nursery-reared experience. However, there is no separation of individuals for rearing

experience when we observe young individuals. This result may indicate that there exists a long-term effect of the social rearing experience. Indeed, a significant amount of DNA methylation variation is explained by age, rearing experience, and their interaction effect (Figure 5.2B).

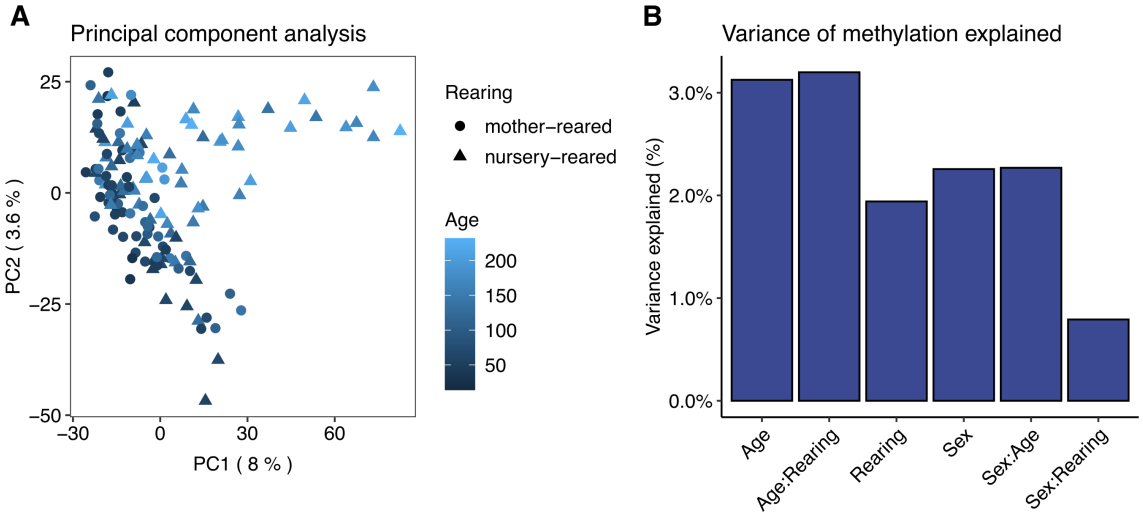

Figure 5.2 DNA methylation in baboon blood samples from RRBS data. (A) Principal component analysis of fractional methylation of CpG sites. Samples were colored and shaped by age and rearing experience, respectively. **(B)** DNA methylation variance explained by biological and environmental factors.

We compared DNA methylation levels from two rearing groups to identify methyl-CpG positions associated with differences in social rearing experience. For each CpG, we fitted a generalized linear model controlling for biological covariates (Methods). We identified 285 CpGs that show significant social rearing-associated methylation changes (henceforth, rearing-differentially methylated loci or rearing-DML). These CpGs were

detected using the cutoff of FDR < 0.1, which corresponds to $P < 6.28 \times 10^{-5}$. We also compared the relationship between DNA methylation and age for two social rearing experiences to identify age-DMLs associated with differences in social rearing experience. We identified 18 CpGs that show a significant interaction effect between age and social rearing experience.

5.3 Methods

5.3.1 Generation of DNA methylation data

Baboon blood samples were obtained from the MD Anderson Center. We used blood-derived DNA to generate DNA methylation data. All the samples were obtained following all relevant ethical regulations and institutional review boards of the MD Anderson Center.

We generated reduced-representation bisulfite sequencing (RRBS) data from 140 baboons (N = 140; 119 females and 22 males). These data are from samples across a broad age span, ranging from 14 months to 19.3 years (mean = 9.6 years). RRBS libraries were constructed using the NuGEN library preparation kit according to the manufacturer's protocol. The RRBS libraries were diluted and loaded onto Illumina HiSeq3000 system for sequencing using 57 bp single-end reads. We performed quality and adapter trimming using TrimGalore v.0.4.1 with a default setting (Babraham Bioinformatics). The libraries from the NuGEN kit use a 6-base barcode with an additional six random bases, which can be used for determining duplicate reads. We removed the additional adaptor sequences

added by the diversity adaptors using a custom python script provided by NuGEN Technologies (https://github.com/nugentechnologies/NuMetRRBS). The sequencing reads were mapped to the baboon papAnu4 reference genome using Bismark v 0.14 (Krueger and Andrews 2011). Duplicated reads were removed using the deduplicate module built in the Bismark software program.

Because genetic polymorphisms of thymine at CpG sites are not distinguishable from bisulfite-converted cytosines, we removed polymorphic CpGs from downstream analyses. Genetic variants collected from 100 baboons were downloaded from https://doi.org/10.5281/zenodo.2583266.

5.3.2 Elastic net regression model

To retain informative CpG sites, we removed CpGs with a mean methylation level either less than 0.1 or greater than 0.9. Also, we removed CpGs with a mean depth of coverage less than 5. We excluded CpG sites with missing data in any individuals. DNA methylation clock for baboons was built using elastic net regression. We used the R package glmnet to build the elastic net regression model (Engebretsen and Bohlin 2019). The optimal regularization parameter, lambda, was determined by 10-fold cross-validation on the data using cv.glmnet. We estimated the DNA methylation age of each individual using the leave-one-out cross-validation approach in which age is predicted for the sample using all samples but that sample as a training set.

5.3.3 Functional enrichment analysis

We investigated the genomic distribution of the clock CpG sites. Statistical significance and fold-enrichment were determined by the occurrences of the CpGs for each genomic feature compared with those in random control sets with matched numbers of the clock CpGs. The random sets were G+C content matched control CpGs (GC ratio of ±500 bp of the focal CpG). Genomic coordinates of functional genomic regions were downloaded from the UCSC genome browser.

5.3.4 Identification of differentially methylated CpGs

For each CpG site, we fitted a linear model to estimate the interaction effect of age and rearing experience. The generalized linear models were created with the DSS (ver. 2.4) Bioconductor package (Park and Wu 2016). We considered sex and bisulfite conversion rates as covariates. We conducted a hypothesis test of the interaction effect using the Wald test based on the estimated coefficient and standard error from the fitted model. We used the Benjamini-Hochberg method for multiple testing corrections.

CHAPTER 6. CONCLUSIONS

A long-standing hypothesis states that due to the high similarity at the protein level, regulatory genetic changes rather than protein sequence changes may explain most species-derived traits over evolutionary time (King and Wilson 1975). This implies that epigenetic functional changes implicated in the gene regulatory differences play a key role in understanding phenotypic variation within and between species. This thesis focuses on understanding the role of epigenetic variation in the genotype-to-phenotype relationship within and between human and non-human primate species and what are the mechanistic basis of the interplay between genome and epigenome.

The focus of the chapters 2 and 3 is the epigenetic evolution of the human brain evolution. The human brain underwent such a dramatic expansion with structural and molecular reorganization in a short period of time. The rapid expansion of the human brain is a fascinating example of evolutionary innovation. Humans are specifically susceptible to many neuropsychiatric and neurodegenerative diseases compared to other closely related non-human primates (Varki, Geschwind et al. 2008, Jakovcevski and Akbarian 2012, Mendizabal, Berto et al. 2019). Thus, some of these human brain-specific traits can be better understood in the context of their evolutionary origins. Previous studies have shown that dysregulation of DNA methylation process is associated with various diseases including cancer and neuropsychiatric disorders (Aran, Sabato et al. 2013, Mendizabal, Berto et al. 2019). Therefore, the characterization of DNA methylation modifications is critical for understanding disease vulnerability in humans as well as morphological and cognitive differences between human and non-human primates. In chapter 2 and 3, we

investigated genome-wide DNA methylation differences in the human brain comparing with non-human primate brains, by performing whole genome methylation sequencing from neurons and oligodendrocytes.

Consistent with the previous findings of an excess of hypomethylation in the human brain from bulk tissues, we found that human brains have overall reduced CG methylation compared to brains of non-human primates for both neurons and oligodendrocytes. The reduction of CG methylation in human brains contributed to increased gene expression levels and human brain-specific active regulatory landscape.

Interestingly, we show that the reduction of human hypomethylation is significantly more pronounced in neurons than in oligodendrocytes. This human neuron-hypo CG DMR resides on brain-specific enhancers as well as other cell-type specific human brain epigenetic marks. Also, we found that the neuron-specific human hypomethylated regions significantly contribute to genetic risk for schizophrenia. These results demonstrate that the human-specific neuron-hypomethylation contributes to the human-derived genetic and epigenetic innovations in a cell-type specific manner.

Although a majority of DNA methylation occurs at CpG sites, DNA methylation at non-CG sites (CH methylation) is relatively abundant in brains. Recent studies have shown that CH methylation plays an important role in gene expression regulation in early brain development. Despite such importance, knowledge is lacking on the evolutionary trajectories and significance of CH methylation during human brain evolution. In chapter 3, we focused our attention on the degree of CH methylation divergence of prefrontal cortex neurons between human and non-human primates. We found that cytosine

methylation in different contexts has played distinctive roles during human brain evolution. First, we discovered that CH methylation has increased along the evolution of human brains. The increase of CH methylation in human brains is associated with the definition of neuronal subtypes. This human-specific CH methylation changes are highly correlated (negatively) with neuronal gene expression divergence between human and chimpanzee. Thus, we hypothesize that evolutionary trajectory of human-derived hypermethylation of CH positions may contribute to fine-tuning of neuron subtype cell identities.

It has been known for several decades that aging has a significant influence on DNA methylation. Traditionally, DNA methylation profiling of age-associated changes has been studied using relatively cost-efficient methods such as DNA methylation arrays. Therefore, genomic patterns of DNA methylation with aging remain yet to be comprehensively characterized. In chapter 4, we examined aging-associated DNA methylation changes at the whole-genome scale and the cellular resolution. Our comprehensive analyses on DNA methylation at the nearly entire set of CpGs in the genome offer an unprecedented opportunity to estimate the age effect on DNA methylation in a truly genome-wide fashion. We showed that age explains more than 10% of the total variation observed in DNA methylation. Moreover, we show that DNA methylation level in early age is a significant determinant of the direction of DNA methylation changes with aging. This result indicates that a significant amount of epigenetic drift is a non-random process.

One of the exciting developments in the aging research field is the development of epigenetic clocks. Using a subset of CpGs, the epigenetic clock can predict biological ages with very high accuracy. Because molecular and physiological mechanisms of aging may share common evolutionary characteristics across closely related species, the development

of epigenetic clocks from non-human primates can be applied to the study of human health and aging. In chapter 5, we developed DNA methylation-based epigenetic clock using newly generated baboon RRBS data. Our epigenetic baboon age predictor clock accurately predicts DNA methylation ages very close to chronological ages in our RRBS samples (R = 0.93). More importantly, we showed the evidence of social rearing experience affecting rates of epigenetic aging.

In summary, our comprehensive analyses of DNA methylation profiles of human and non-human primates will expand our understanding of the epigenetic evolution of the human brain and aging epigenetic programs. We hope these novel findings stimulate interest in studies linking the epigenetic evolution of human brains to regulation and disease susceptibility.

APPENDIX A. SUPPLEMENTARY MATERIAL FOR CHAPTER 2

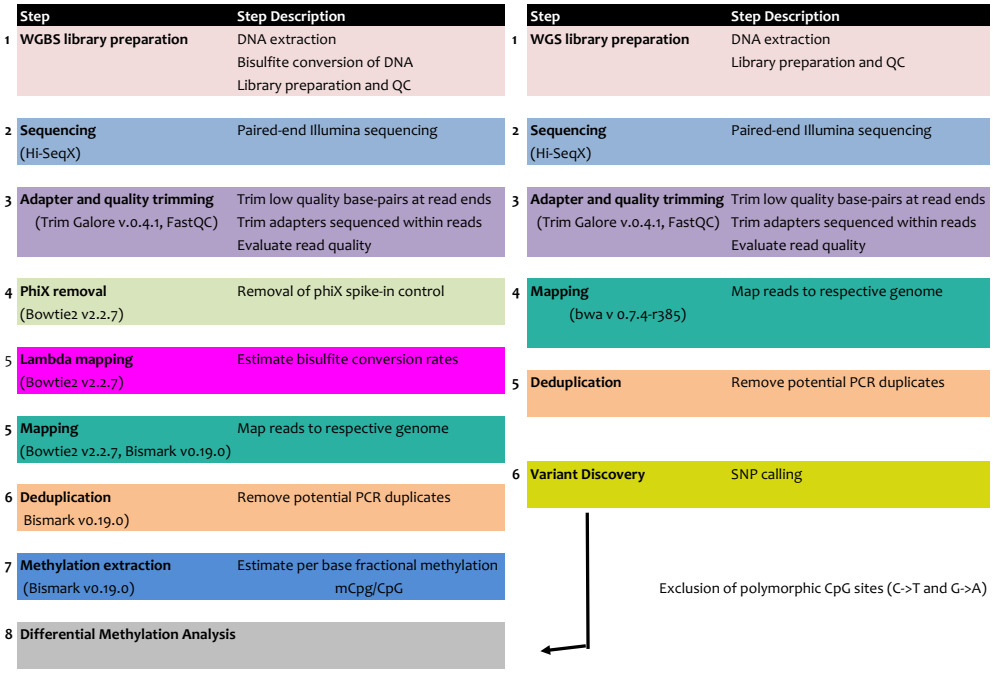

Figure A. 1. Overview of the workflow for WGBS and WGS data processing and differential methylation analyses.

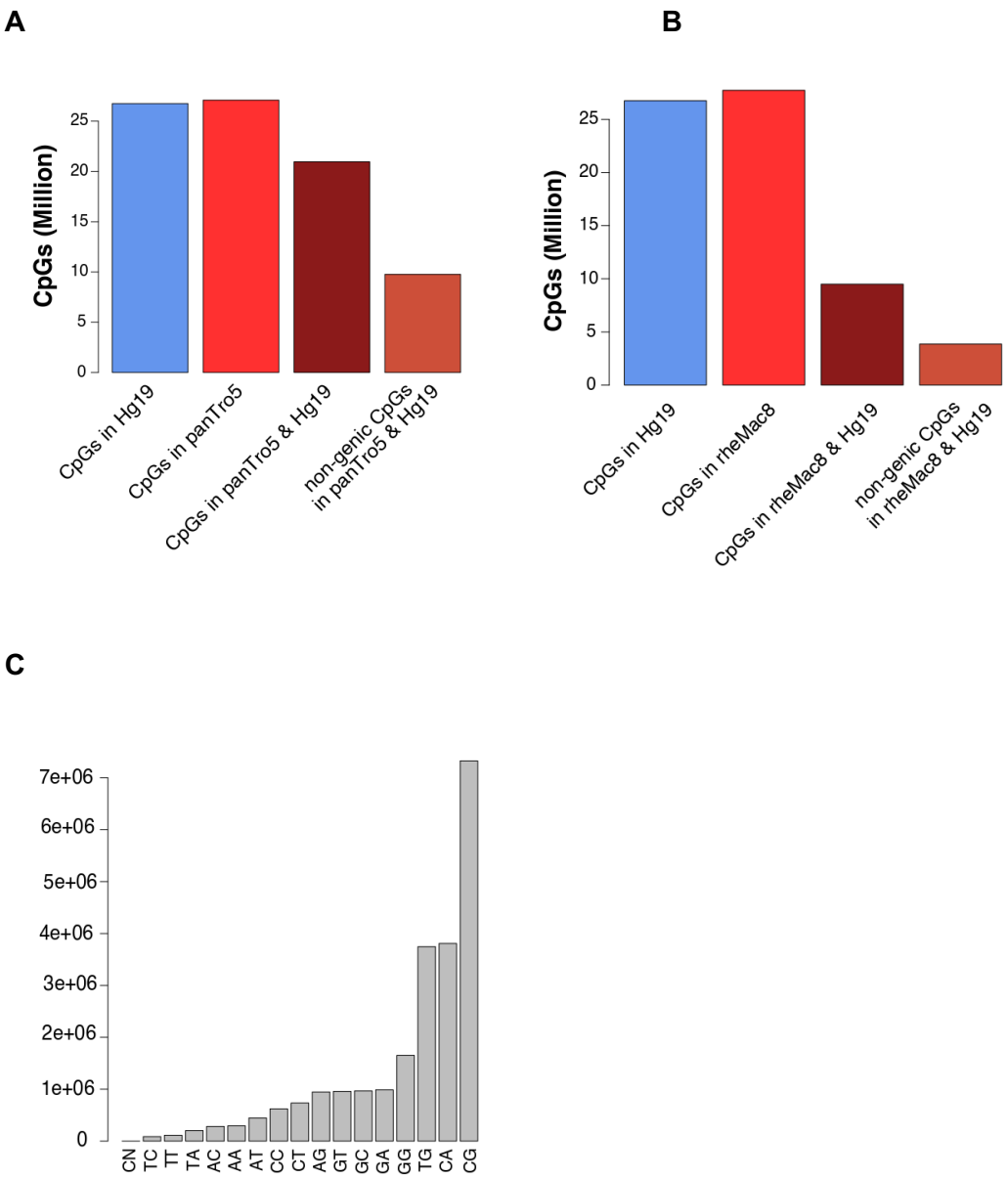

Figure A. 2. Loss of CG sites over evolutionary time. (A) Comparison of the total number of CpGs in either the human (hg19) or chimpanzee (panTro5) genomes with the total number of conserved CpGs in genic regions in both genomes or in non-genic regions in both genomes. (B) Comparison of the total number of CpGs in human (hg19) or macaque (rheMac8) genomes with the total number of conserved CpGs in genic regions in both genomes or in non-genic regions in both genomes. (C) Dinucleotide composition of CpGs of rheMac8 in the hg19 genome.

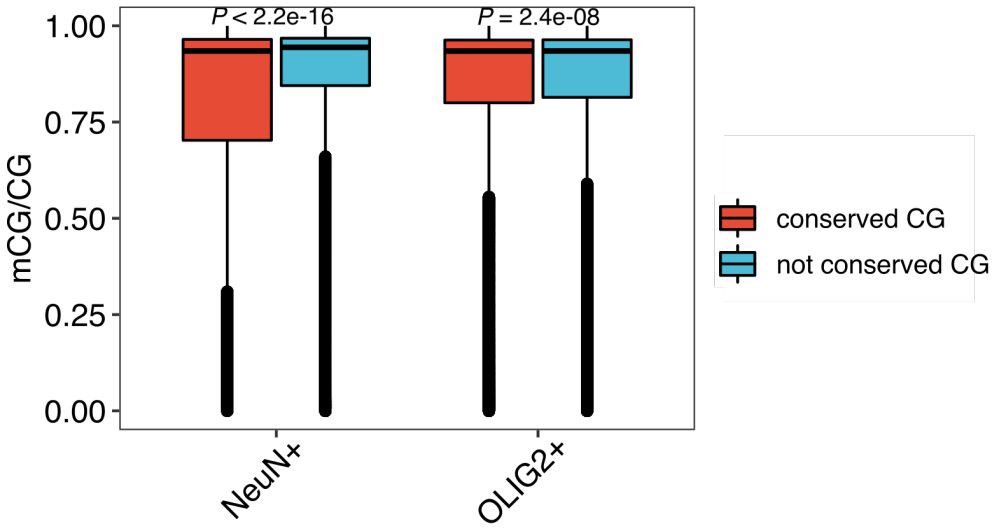

Figure A. 3. Comparison of fractional methylation between conserved CG sites and species-specific CG sites in humans. To illustrate, a subset of 100,000 sites were selected for each CG group. A one-tailed t-test was conducted to test whether conserved CG sites are biased toward hypomethylation. Box represents a range from the first quartile to the third quartile. The line in the box indicates the median value. The minima and maxima are within 1.5 times the distance between the first and third quartiles from box.

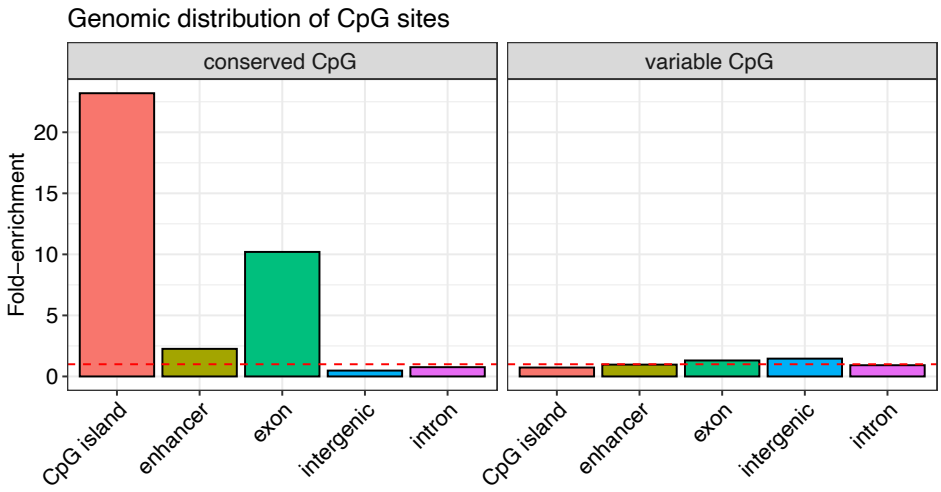

Figure A. 4. Distribution of evolutionarily conserved CpGs and variable CpGs in different functional genomic regions. Fold-enrichment was computed from the occurrences of the CpGs for each feature compared to random control sets (n=100). Red dashed lines indicate fold-enrichment values of 1.

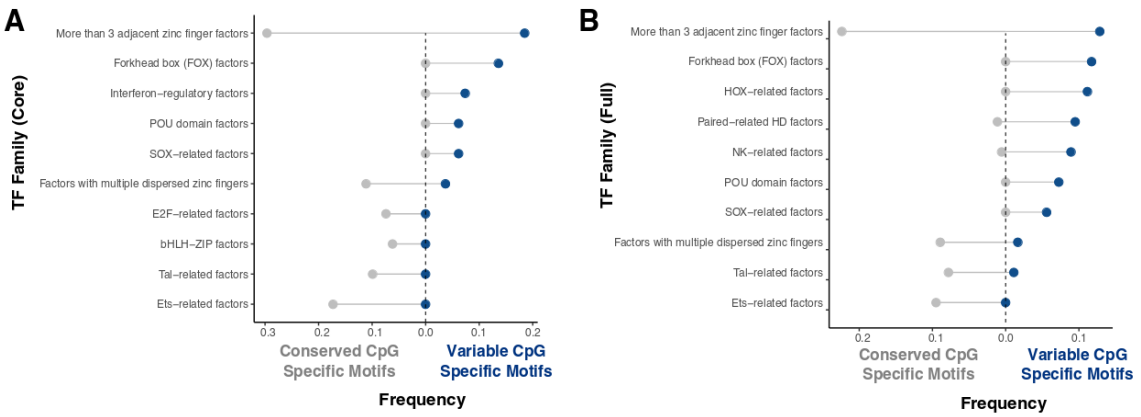

Figure A. 5. Transcription factor families with differential motif enrichment in conserved CpGs (gray dots) vs. human-specific-CpGs (variable CpGs, blue dots) in (A) core v11 and (B) full HOCOMOCO v11 databases.

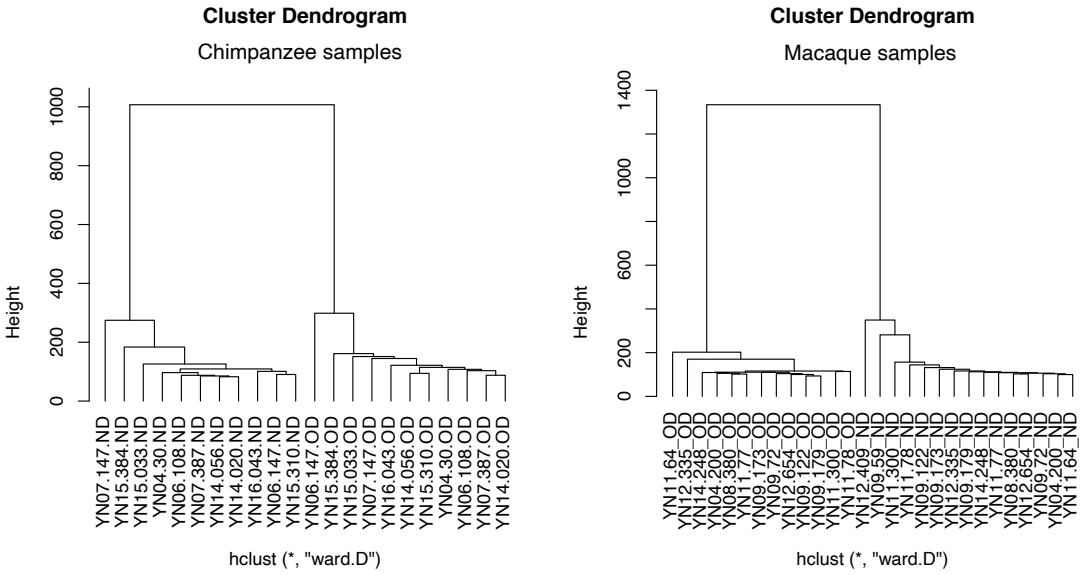

Figure A. 6. Hierarchical clustering of CG methylomes for chimpanzees and macaques. Names ending in ND represent neuronal (NeuN+) cell samples and names ending in OD represent oligodendrocyte (OLIG2+) cell samples.

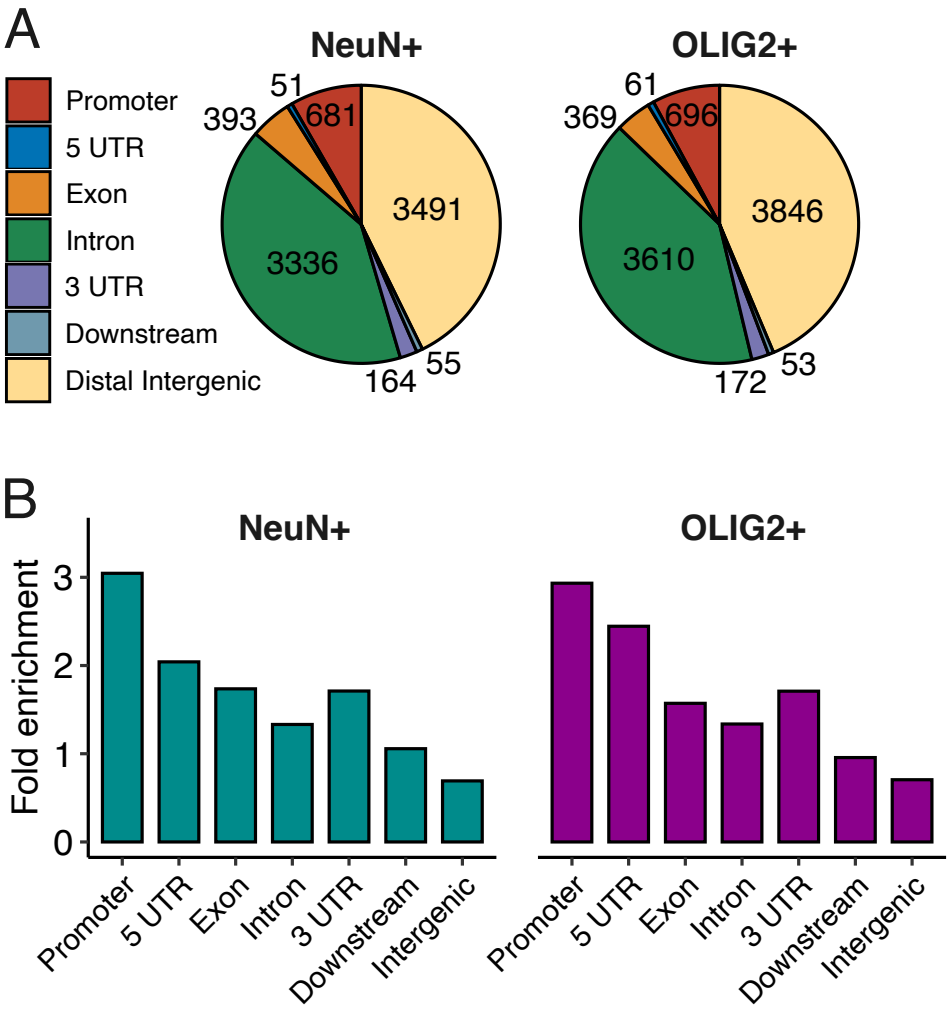

Figure A. 7. (A) Genomic locations of human DMRs show that most DMRs are within or nearby genes and (B) enriched in promoters and genic regions. Fold enrichment is computed by occurrences of DMRs in each genomic feature compared to GC matched control region sets (n=100).

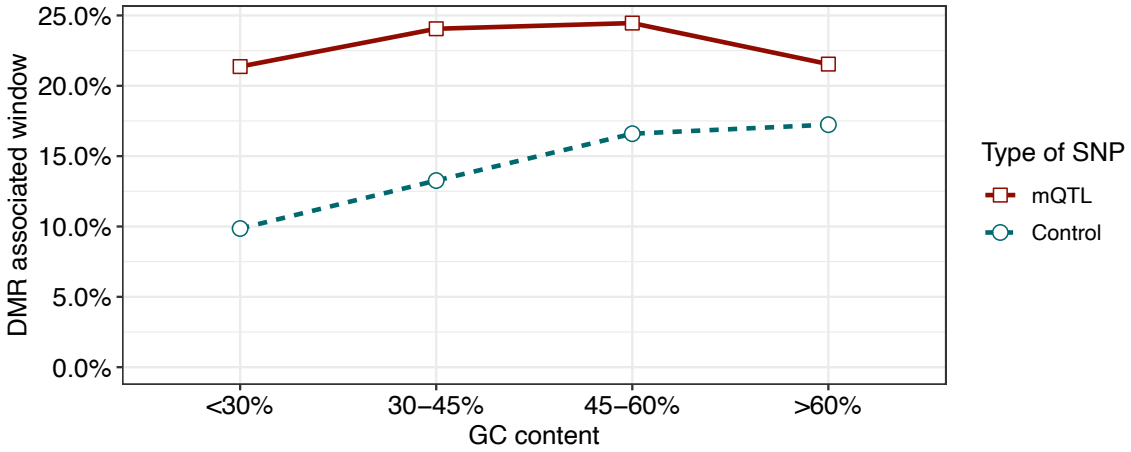

Figure A. 8. Genomic windows (200bps each) containing mQTLs are more often associated with DMRs than genomic windows containing SNPs matched for their minor allele frequency (MAF) (control) of the same size across different GC contents.

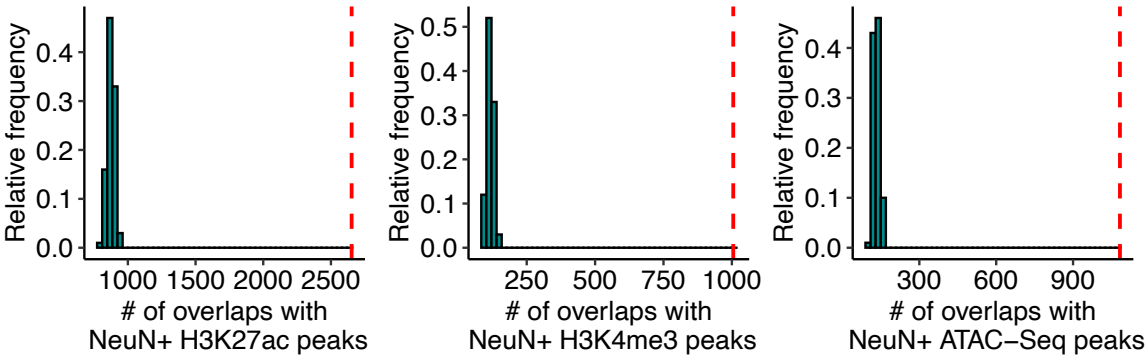

Figure A. 9. Genomic locations of human neuron-specific hypomethylated CG DMRs are enriched in cell-type specific human brain epigenetic marks. Fold enrichment is computed by observed numbers of DMRs (red dashed line) overlapping with each neuron-specific epigenetic mark compared to GC matched control region sets (n=100).

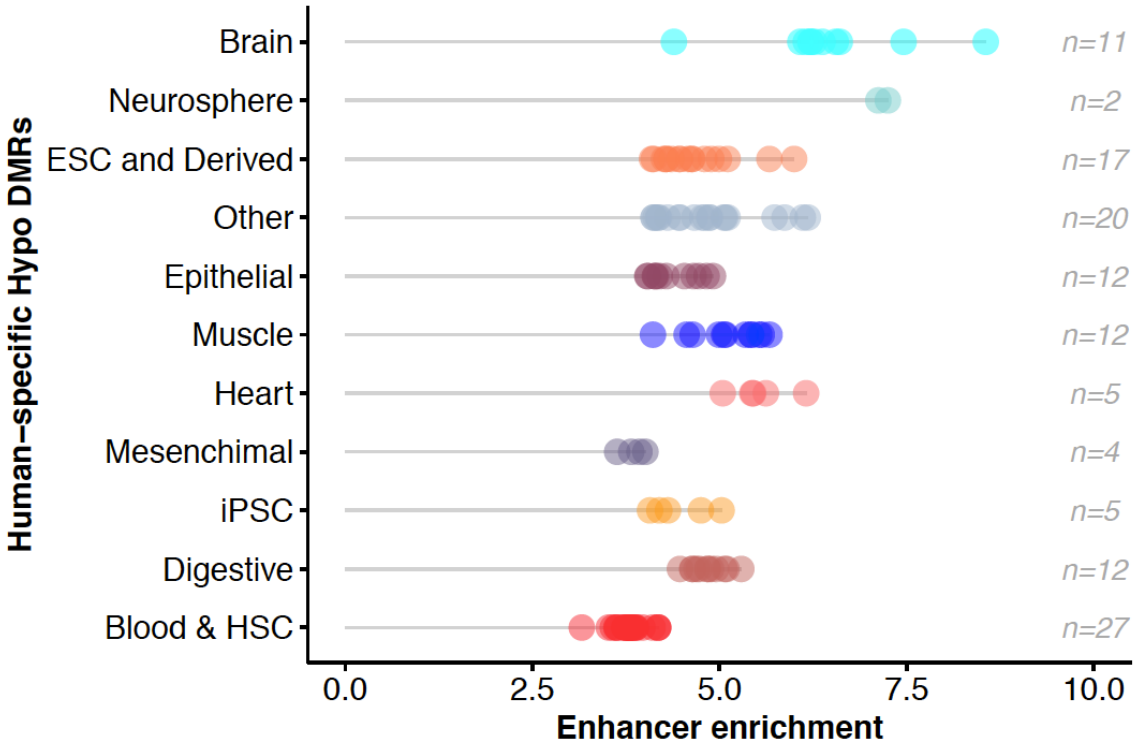

Figure A. 10. Enhancer enrichment at human-specific hypomethylated CG DMRs. 25 chromatin state-model maps based on 6 chromatin mark ChIP-Seq experiments (H3K4me3, H3K4me1, H3K36me3, H3K27me3, H3K9me3 and H3K27ac) were obtained from the Roadmap Epigenomics Project. Each dot represents the enrichment for enhancer-related states (TxReg, TxEnh5', TxEnh3', TxEnhW, EnhA1, EnhA2, EnhW1, EnhW2, and EnhAc) compared to 100 sets of GC-content matched control DMR sets for a given cell-type or tissue. The original 117 cell-type and tissue-types were grouped into 11 categories shown in the y-axes (total number of cell-type/tissues per group is indicated). Empirical $P<0.01$ for all enrichments.

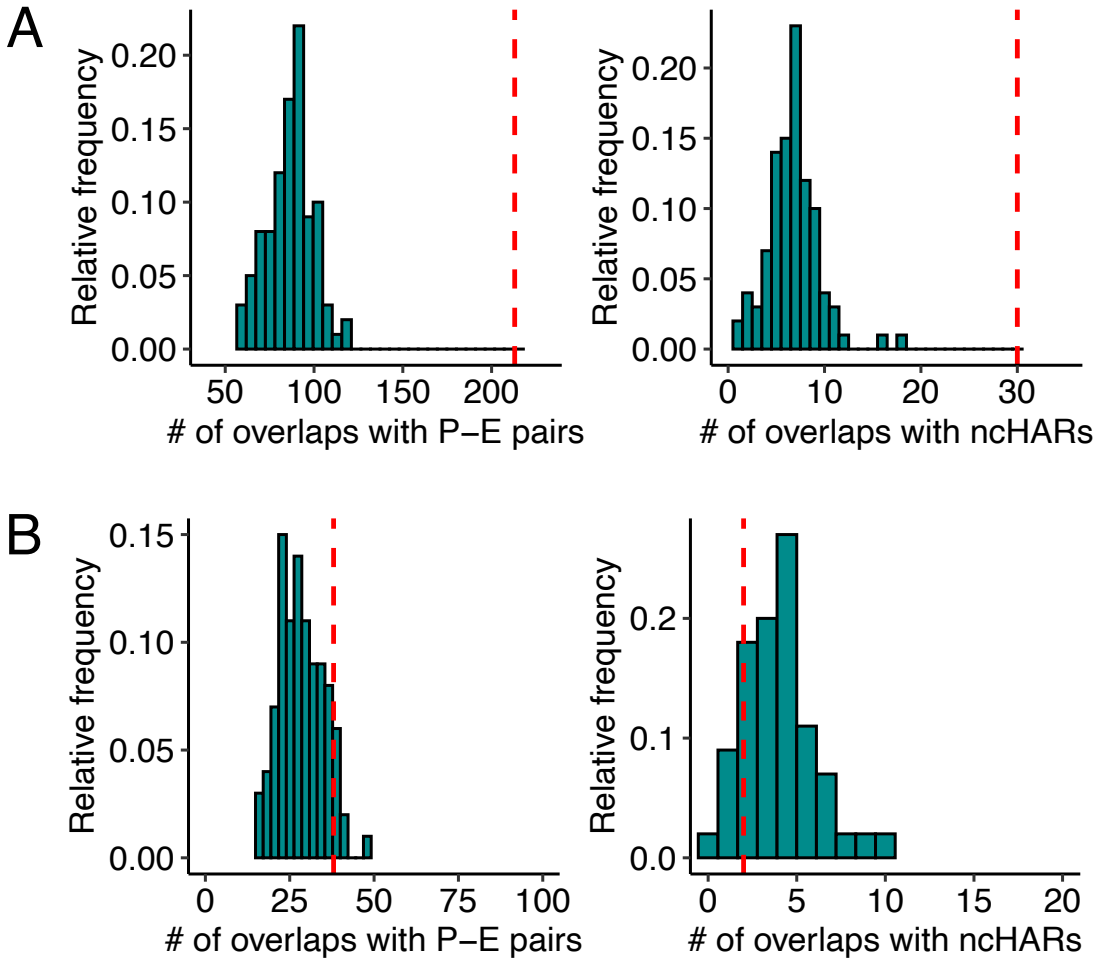

Figure A. 11. (A) Human neuron-hypo CG DMRs are significantly co-localized with enhancer-promoter (E-P) pairs and ncHARs (red dashed lines). Null distributions were plotted based on the GC matched control region sets (n=100) that overlap with enhancer-promoter pairs and ncHARs. (B) Chimpanzee DMRs that overlap with enhancer-promoter pairs and ncHARs were falling in the distribution based on the GC matched control region sets (n=100) that overlap with enhancer-promoter pairs and ncHARs.

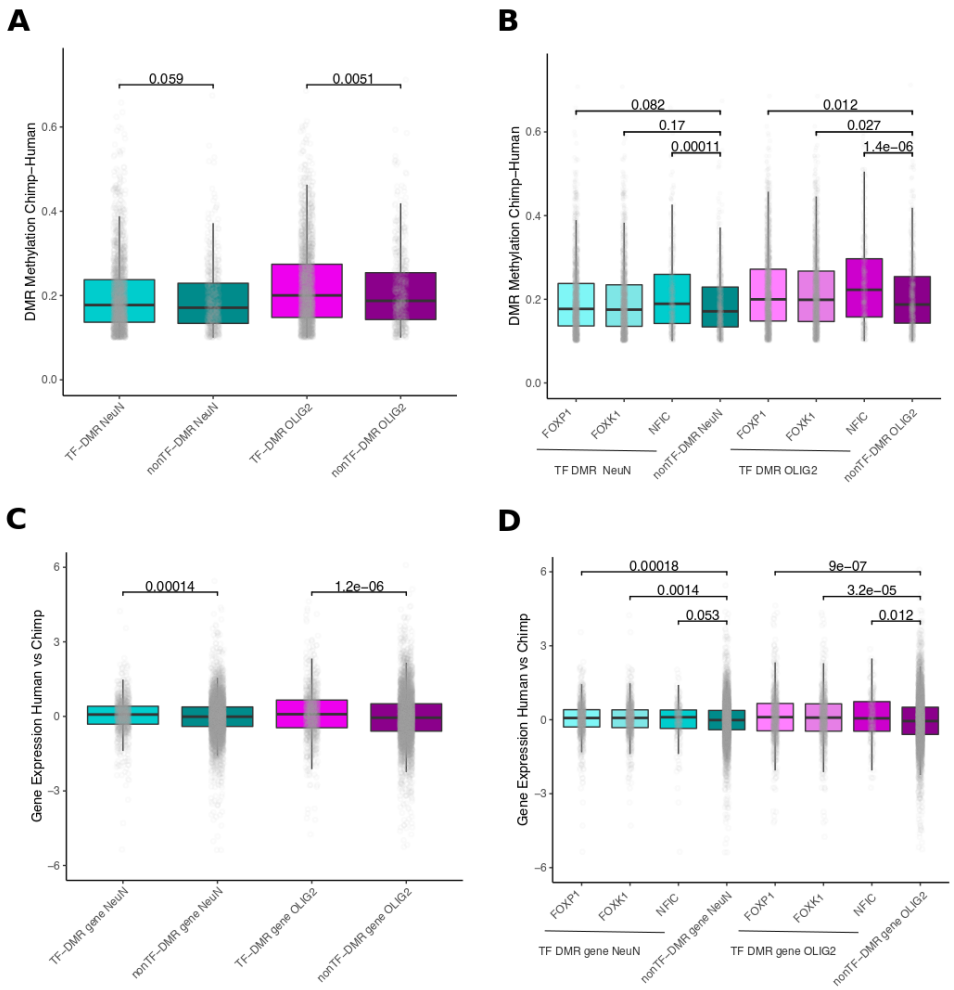

Figure A. 12. (A) Distribution of DNA methylation differences between human and chimpanzee brain cell-types at human-specific hypomethylated DMRs with and without enriched TF motifs. P-values for one-sided Wilcoxon signed-ranked test with alternative = greater (TF n=2439 vs non-TF n=652 for each cell-type). (B) Same as in panel A but separated by the specific TF enriched (FOXP1 n=1996, FOXK1 n=1906, NFIC n=462 and non-TF n=652 for each cell-type). (C) Distribution of gene expression differences between human and chimpanzee at human-specific hypomethylated DMRs with and without enriched TF motifs. P-values for one-sided Wilcoxon signed-ranked test with alternative = greater (NeuN+: TF n=1110 vs non-TF n=7262, and OLIG2+: TF n=1031 vs non-TF n=6529). (D) Same as in panel C but separated by the specific TF enriched (NeuN: FOXP1 n=957, FOXK1 n=901, NFIC n=248, non-TF n=7262, and OLIG2: FOXP1 n=892, FOXK1 n=833, NFIC n=229 and non-TF n=6529). Box represents a range from the first quartile to the third quartile. The line in the box indicates the median value. The minima and maxima are within 1.5 times the distance between the first and third quartiles from box.

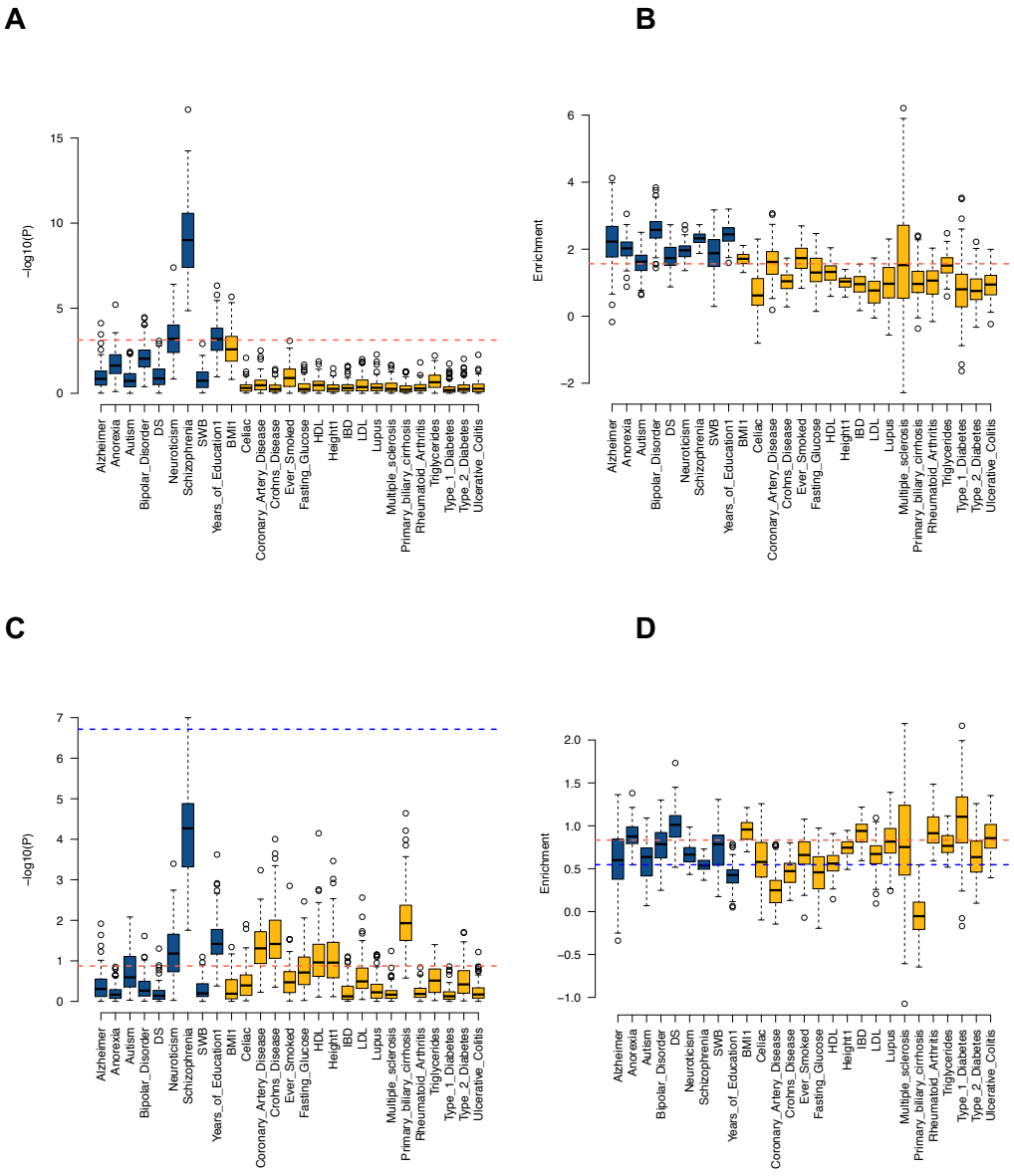

Figure A. 13. (A) and (B) Enrichment and P-values for schizophrenia heritability at 100 subsets of conserved NeuN+ Hypo DMRs that match the number and length of human-derived NeuN+ Hypo DMRs. Boxplots show the results of 100 conserved NeuN+ Hypo DMR subsets and the red lines indicate the observed values for human-derived NeuN+ Hypo DMRs. (C) and (D) Similar analyses for subsampling of human hyper CH DMRs to match the number and length distribution of chimpanzee hyper CH DMRs. Boxplots show the human hyper CH DMR subsets, red lines indicate observed chimpanzee hyper DMRs and blue lines indicate observed human hyper DMRs (full set). Box represents a range from the first quartile to the third quartile. The line in the box indicates the median value. The minima and maxima are within 1.5 times the distance between the first and third quartiles from box.

Table A. 1. List of covariates included in the WGBS analyses.

Sample ID	Species	Sex	Age class	Pmi	Cell type
X4615_Control_NeuN	Human	M	2	15	NeuN
X3611_Control_NeuN	Human	M	3	17.5	NeuN
X3602_Control_NeuN	Human	M	3	13.2	NeuN
X3590_Control_NeuN	Human	M	3	11.5	NeuN
X3586_Control_NeuN	Human	M	3	16	NeuN
X3545_Control_NeuN	Human	M	3	14	NeuN
X1541_Control_NeuN	Human	F	2	15.3	NeuN
X1539_Control_NeuN	Human	F	2	23	NeuN
X1538_Control_NeuN	Human	F	2	25	NeuN
X1537_Control_NeuN	Human	M	2	22	NeuN
X1536_Control_NeuN	Human	M	2	24	NeuN
X1535_Control_NeuN	Human	M	2	15.4	NeuN
X1534_Control_NeuN	Human	M	3	7.5	NeuN
X1533_Control_NeuN	Human	M	2	19.3	NeuN
X1532_Control_NeuN	Human	F	3	19	NeuN
X1531_Control_NeuN	Human	M	2	17.1	NeuN
X1527_Control_NeuN	Human	M	1	23	NeuN
X1525_Control_NeuN	Human	F	3	11	NeuN
X1524_Control_NeuN	Human	M	1	10	NeuN
Miami0001_Control_NeuN	Human	M	1	16.3	NeuN
AN16799_Control_NeuN	Human	M	2	14.68	NeuN
AN15240_Control_NeuN	Human	F	1	18.08	NeuN
AN10090_Control_NeuN	Human	M	2	13.12	NeuN
AN05483_Control_NeuN	Human	M	3	16.97	NeuN
AN03398_Control_NeuN	Human	F	3	12.1	NeuN
YN14.020_Chimp_NeuN	Chimpanzee	F	2	0.9	NeuN
YN15.033_Chimp_NeuN	Chimpanzee	F	3	0.9	NeuN
YN06.108_Chimp_NeuN	Chimpanzee	F	3	0.5	NeuN
YN07.147_Chimp_NeuN	Chimpanzee	M	1	3	NeuN
YN14.056_Chimp_NeuN	Chimpanzee	F	2	0.9	NeuN
YN07.387_Chimp_NeuN	Chimpanzee	M	3	2	NeuN
YN06.147_Chimp_NeuN	Chimpanzee	M	3	2.5	NeuN
YN15.384_Chimp_NeuN	Chimpanzee	F	3	0.9	NeuN
YN15.310_Chimp_NeuN	Chimpanzee	F	3	0.9	NeuN
YN04.30_Chimp_NeuN	Chimpanzee	M	2	0.5	NeuN

Table A. 1 (continued)

YN14.020_Chimp_NeuN	Chimpanzee	M	3	0.9	NeuN
YN04.200_Macaque_NeuN	Rhesus macaque	M	1	0.9	NeuN
YN08.380_Macaque_NeuN	Rhesus macaque	F	3	0.9	NeuN
YN09.122_Macaque_NeuN	Rhesus macaque	F	3	0.9	NeuN
YN09.173_Macaque_NeuN	Rhesus macaque	M	3	0.9	NeuN
YN09.179_Macaque_NeuN	Rhesus macaque	F	3	0.9	NeuN
YN09.59_Macaque_NeuN	Rhesus macaque	F	3	0.9	NeuN
YN09.72_Macaque_NeuN	Rhesus macaque	F	3	0.9	NeuN
YN11.300_Macaque_NeuN	Rhesus macaque	M	2	0.9	NeuN
YN11.64_Macaque_NeuN	Rhesus macaque	F	1	0.9	NeuN
YN11.77_Macaque_NeuN	Rhesus macaque	M	1	0.9	NeuN
YN11.78_Macaque_NeuN	Rhesus macaque	F	1	0.9	NeuN
YN12.335_Macaque_NeuN	Rhesus macaque	F	3	0.9	NeuN
YN12.409_Macaque_NeuN	Rhesus macaque	M	2	0.9	NeuN
YN12.654_Macaque_NeuN	Rhesus macaque	F	3	0.9	NeuN
YN14.248_Macaque_NeuN	Rhesus macaque	M	1	0.9	NeuN
X1524_Control_Olig2	Human	M	1	10	OLIG2
X1525_Control_Olig2	Human	F	3	11	OLIG2
X1527_Control_Olig2	Human	M	1	23	OLIG2
X1532_Control_Olig2	Human	F	3	19	OLIG2
X1536_Control_Olig2	Human	M	2	24	OLIG2
X1538_Control_Olig2	Human	F	2	25	OLIG2

Table A. 1 (continued)

X1539_Control_Olig2	Human	F	2	23	OLIG2
X1541_Control_Olig2	Human	F	2	15.3	OLIG2
X3545_Control_Olig2	Human	M	3	14	OLIG2
X3586_Control_Olig2	Human	M	3	16	OLIG2
X3590_Control_Olig2	Human	M	3	11.5	OLIG2
X3602_Control_Olig2	Human	M	3	13.2	OLIG2
X3611_Control_Olig2	Human	M	3	17.5	OLIG2
X4615_Control_Olig2	Human	M	2	15	OLIG2
AN03398_Control_Olig2	Human	F	3	12.1	OLIG2
AN05483_Control_Olig2	Human	M	3	16.97	OLIG2
AN10090_Control_Olig2	Human	M	2	13.12	OLIG2
AN15240_Control_Olig2	Human	F	1	18.08	OLIG2
AN16799_Control_Olig2	Human	M	2	14.68	OLIG2
Miami0001_Control_Olig2	Human	M	1	16.3	OLIG2
YN16.043_Chimp_Olig2	Chimpanzee	F	2	0.9	OLIG2
YN15.033_Chimp_Olig2	Chimpanzee	F	3	0.9	OLIG2
YN06.108_Chimp_Olig2	Chimpanzee	F	3	0.5	OLIG2
YN07.147_Chimp_Olig2	Chimpanzee	M	1	3	OLIG2
YN14.056_Chimp_Olig2	Chimpanzee	F	2	0.9	OLIG2
YN07.387_Chimp_Olig2	Chimpanzee	M	3	2	OLIG2
YN06.147_Chimp_Olig2	Chimpanzee	M	3	2.5	OLIG2
YN15.384_Chimp_Olig2	Chimpanzee	F	3	0.9	OLIG2
YN15.310_Chimp_Olig2	Chimpanzee	F	3	0.9	OLIG2
YN04.30_Chimp_Olig2	Chimpanzee	M	2	0.5	OLIG2
YN14.020_Chimp_Olig2	Chimpanzee	M	3	0.9	OLIG2
YN04.200-Macaque_Olig2	Rhesus macaque	M	1	0.9	OLIG2
YN08.380_Macaque_Olig2	Rhesus macaque	F	3	0.9	OLIG2
YN09.122_Macaque_NeuN	Rhesus macaque	F	3	0.9	OLIG2
YN09.173_Macaque_Olig2	Rhesus macaque	M	3	0.9	OLIG2
YN09.179_Macaque_Olig2	Rhesus macaque	F	3	0.9	OLIG2
YN09.72_Macaque_Olig2	Rhesus macaque	F	3	0.9	OLIG2

Table A. 1 (continued)

YN11.300_Macaque_Olig2	Rhesus macaque	M	2	0.9	OLIG2
YN11.64_Macaque_Olig2	Rhesus macaque	F	1	0.9	OLIG2
YN11.77_Macaque_Olig2	Rhesus macaque	M	1	0.9	OLIG2
YN11.78_Macaque_Olig2	Rhesus macaque	F	1	0.9	OLIG2
YN12.335_Macaque_Olig2	Rhesus macaque	F	3	0.9	OLIG2
YN12.654_Macaque_Olig2	Rhesus macaque	F	3	0.9	OLIG2
YN14.248_Macaque_Olig2	Rhesus macaque	M	1	0.9	OLIG2

Table A. 2. WGBS data processing and mapping statistics.

Sample	Species	Cell type	Mapping %	Mean DP
YN15-384_ND	Pan Troglodites	NeuN	72.31	8.29
YN15-033_ND	Pan Troglodites	NeuN	69.76	13.68
YN16-043_ND	Pan Troglodites	NeuN	71.73	27.39
YN06-108_ND	Pan Troglodites	NeuN	63.83	22.99
YN07-147_ND	Pan Troglodites	NeuN	40.90	3.38
YN04-30_ND	Pan Troglodites	NeuN	58.52	16.77
YN07-387_ND	Pan Troglodites	NeuN	71.23	25.86
YN06-147_ND	Pan Troglodites	NeuN	63.35	20.63
YN14-056_ND	Pan Troglodites	NeuN	74.38	31.99
YN15-310_ND	Pan Troglodites	NeuN	72.63	32.04
YN14-020_ND	Pan Troglodites	NeuN	70.56	26.14
YN15-384_OD	Pan Troglodites	OLIG2	73.87	9.01
YN15-033_OD	Pan Troglodites	OLIG2	75.56	10.28
YN16-043_OD	Pan Troglodites	OLIG2	77.31	19.66
YN06-108_OD	Pan Troglodites	OLIG2	78.17	17.28
YN07-147_OD	Pan Troglodites	OLIG2	76.32	9.10
YN04-30_OD	Pan Troglodites	OLIG2	72.63	16.45
YN07-387_OD	Pan Troglodites	OLIG2	75.97	21.45
YN06-147_OD	Pan Troglodites	OLIG2	41.70	2.58
YN14-056_OD	Pan Troglodites	OLIG2	74.85	28.26
YN15-310_OD	Pan Troglodites	OLIG2	71.94	31.98
YN14-020_OD	Pan Troglodites	OLIG2	74.11	29.79
YN08-380_ND	Rhesus Macaque	NeuN	74.12	27.29
YN09-122_ND	Rhesus Macaque	NeuN	70.94	17.55
YN09-179_ND	Rhesus Macaque	NeuN	64.36	19.22
YN09-59_ND	Rhesus Macaque	NeuN	49.30	10.58
YN11-300_ND	Rhesus Macaque	NeuN	51.29	8.99
YN11-77_ND	Rhesus Macaque	NeuN	73.88	22.89
YN11-78_ND	Rhesus Macaque	NeuN	67.74	14.04
YN12-335_ND	Rhesus Macaque	NeuN	63.16	16.47
YN12-409_ND	Rhesus Macaque	NeuN	27.18	2.09
YN12-654_ND	Rhesus Macaque	NeuN	70.79	24.13
YN04-200_ND	Rhesus Macaque	NeuN	72.26	23.35
YN09-173_ND	Rhesus Macaque	NeuN	72.93	17.73

Table A. 2 (continued)

YN09-72_ND	Rhesus Macaque	NeuN	75.46	25.50
YN11-64_ND	Rhesus Macaque	NeuN	71.00	23.52
YN14-248_ND	Rhesus Macaque	NeuN	68.37	22.95
YN08-380_OD	Rhesus Macaque	OLIG2	74.01	34.08
YN09-122_OD	Rhesus Macaque	OLIG2	69.45	30.31
YN09-179_OD	Rhesus Macaque	OLIG2	73.79	36.82
YN11-300_OD	Rhesus Macaque	OLIG2	68.19	24.54
YN11-77_OD	Rhesus Macaque	OLIG2	70.75	30.05
YN11-78_OD	Rhesus Macaque	OLIG2	71.88	27.19
YN12-335_OD	Rhesus Macaque	OLIG2	62.40	10.44
YN12-654_OD	Rhesus Macaque	OLIG2	71.08	30.20
YN04-200_OD	Rhesus Macaque	OLIG2	70.09	24.61
YN09-173_OD	Rhesus Macaque	OLIG2	70.35	24.75
YN09-72_OD	Rhesus Macaque	OLIG2	74.49	26.59
YN11-64_OD	Rhesus Macaque	OLIG2	71.2	5.81
YN14-248_OD	Rhesus Macaque	OLIG2	69.01	23.37

Table A. 3. WGS data processing and mapping statistics

Sample	Mapped reads	Mapping %	Mean DP
YN16-043	446678671	89.3%	24.1
YN15-033	436626713	95.9%	33.48
YN06-108	455929295	90.6%	23.58
YN07-147	441722229	89.6%	26.04
YN14-056	374973270	96.0%	21.86
YN07-387	201895989	99.7%	17.97
YN06-147	222137036	99.5%	19.35
YN15-384	441160742	95.7%	33.96
YN15-310	417208392	94.6%	24.95
YN04-30	421681615	91.4%	23.64
YN14-020	220850613	99.5%	18.06
YN04-200	211971044	99.1%	17.03
YN08-380	376925788	94.9%	29.75
YN09-59	180566422	99.1%	15.57
YN09-72	181442176	99.1%	15.58
YN09-122	376252657	89.7%	23.6
YN09-173	229848671	99.2%	18.12
YN09-179	354685331	95.0%	28.57
YN11-64	174683507	99.1%	14.81
YN11-77	376350631	96.6%	30.89
YN11-78	369949049	92.2%	23.04
YN11-300	403831657	92.3%	16.73
YN12-335	433762376	92.7%	29.76
YN12-409	387683227	88.5%	25.07
YN12-654	386504856	94.3%	29.98
YN14-248	202304006	99.1%	16.42

Table A. 4. List of variable CpG-specific and conserved CpG-specific transcription factor (TF) motifs.

Motif_class	TF	TF.family
Conserved_specific	FEV	Ets-related factors[3.5.2]
Conserved_specific	ETV1	Ets-related factors[3.5.2]
Conserved_specific	ELK4	Ets-related factors[3.5.2]
Conserved_specific	TGIF1	TALE-type homeo domain factors[3.1.4]
Conserved_specific	MYB	Myb/SANT domain factors[3.5.1]
Conserved_specific	NEUROD 2	Tal-related factors[1.2.3]
Conserved_specific	ETS1	Ets-related factors[3.5.2]
Conserved_specific	LYL1	Tal-related factors[1.2.3]
Conserved_specific	GABPA	Ets-related factors[3.5.2]
Conserved_specific	PTF1A	Tal-related factors[1.2.3]
Conserved_specific	ELF1	Ets-related factors[3.5.2]
Conserved_specific	ERG	Ets-related factors[3.5.2]
Conserved_specific	ELK1	Ets-related factors[3.5.2]
Conserved_specific	ATOH1	Tal-related factors[1.2.3]
Conserved_specific	ELF2	Ets-related factors[3.5.2]
Conserved_specific	BHLHA15	Tal-related factors[1.2.3]
Conserved_specific	OSR2	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	PBX1	TALE-type homeo domain factors[3.1.4]
Conserved_specific	RBPJ	CSL-related factors[6.1.4]
Conserved_specific	ETV5	Ets-related factors[3.5.2]
Conserved_specific	NR2C2	RXR-related receptors (NR2)[2.1.3]
Conserved_specific	ZBTB14	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	NEUROD 1	Tal-related factors[1.2.3]
Conserved_specific	ETV4	Ets-related factors[3.5.2]
Conserved_specific	MYOD1	MyoD / ASC-related factors[1.2.2]
Conserved_specific	TFAP4	bHLH-ZIP factors[1.2.6]
Conserved_specific	OLIG2	Tal-related factors[1.2.3]
Conserved_specific	MECP2	
Conserved_specific	MYF6	MyoD / ASC-related factors[1.2.2]
Conserved_specific	MYOG	MyoD / ASC-related factors[1.2.2]
Conserved_specific	NR1H4	Thyroid hormone receptor-related factors (NR1)[2.1.2]
Conserved_specific	ELF5	Ets-related factors[3.5.2]

Table A. 4 (continued)

Conserved_specific	PKNOX1	TALE-type homeo domain factors[3.1.4]
Conserved_specific	NKX2-5	NK-related factors[3.1.2]
Conserved_specific	ETV2	Ets-related factors[3.5.2]
Conserved_specific	PBX3	TALE-type homeo domain factors[3.1.4]
Conserved_specific	TCF12	E2A-related factors[1.2.1]
Conserved_specific	ZNF563	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	TFAP2A	AP-2[1.3.1]
Conserved_specific	ZBTB48	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	MZF1	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	FLI1	Ets-related factors[3.5.2]
Conserved_specific	E2F1	E2F-related factors[3.3.2]
Conserved_specific	TFAP2B	AP-2[1.3.1]
Conserved_specific	TCF4	E2A-related factors[1.2.1]
Conserved_specific	NR5A1	FTZ-F1-related receptors (NR5)[2.1.5]
Conserved_specific	NRF1	NRF[0.0.6]
Conserved_specific	MAFB	Maf-related factors[1.1.3]
Conserved_specific	EBF1	Early B-Cell Factor-related factors[6.1.5]
Conserved_specific	ZIC3	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	ZBTB18	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	TFDP1	E2F-related factors[3.3.2]
Conserved_specific	TFAP2C	AP-2[1.3.1]
Conserved_specific	CTCF	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	RFX2	RFX-related factors[3.3.3]
Conserved_specific	ZNF667	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	ASCL1	MyoD / ASC-related factors[1.2.2]
Conserved_specific	NHLH1	Tal-related factors[1.2.3]
Conserved_specific	RELB	NF-kappaB-related factors[6.1.1]
Conserved_specific	ZBTB33	Other factors with up to three adjacent zinc fingers[2.3.2]
Conserved_specific	E2F2	E2F-related factors[3.3.2]
Conserved_specific	ZNF341	Factors with multiple dispersed zinc fingers[2.3.4]
Conserved_specific	TCF3	E2A-related factors[1.2.1]
Conserved_specific	PPARG	Thyroid hormone receptor-related factors (NR1)[2.1.2]
Conserved_specific	RFX1	RFX-related factors[3.3.3]
Conserved_specific	MYC	bHLH-ZIP factors[1.2.6]
Conserved_specific	GRHL2	Grainyhead-related factors[6.7.1]

Table A. 4 (continued)

Conserved_specific	HNF4G	RXR-related receptors (NR2)[2.1.3]
Conserved_specific	ZNF547	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	NR5A2	FTZ-F1-related receptors (NR5)[2.1.5]
Conserved_specific	ZFP42	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	ZBTB7A	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	MBD2	
Conserved_specific	REST	Factors with multiple dispersed zinc fingers[2.3.4]
Conserved_specific	ATF6	CREB-related factors[1.1.7]
Conserved_specific	MYCN	bHLH-ZIP factors[1.2.6]
Conserved_specific	PPARA	Thyroid hormone receptor-related factors (NR1)[2.1.2]
Conserved_specific	EPAS1	PAS domain factors[1.2.5]
Conserved_specific	ZNF467	Factors with multiple dispersed zinc fingers[2.3.4]
Conserved_specific	MAX	bHLH-ZIP factors[1.2.6]
Conserved_specific	RFX5	RFX-related factors[3.3.3]
Conserved_specific	TAF1	TCF-7-related factors[4.1.3]
Conserved_specific	ARNT	PAS domain factors[1.2.5]
Conserved_specific	VEZF1	Factors with multiple dispersed zinc fingers[2.3.4]
Conserved_specific	HIF1A	PAS domain factors[1.2.5]
Conserved_specific	YY1	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	CTCFL	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	E2F6	E2F-related factors[3.3.2]
Conserved_specific	ATF2	Jun-related factors[1.1.1]
Conserved_specific	MXI1	bHLH-ZIP factors[1.2.6]
Conserved_specific	CREM	CREB-related factors[1.1.7]
Conserved_specific	NFE2	Jun-related factors[1.1.1]
Conserved_specific	INSM1	Factors with multiple dispersed zinc fingers[2.3.4]
Conserved_specific	ZFX	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	E2F3	E2F-related factors[3.3.2]
Conserved_specific	ZNF76	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	ZBTB6	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	THAP11	THAP-related factors[2.9.1]
Conserved_specific	ZNF281	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	NR2C1	RXR-related receptors (NR2)[2.1.3]
Conserved_specific	HIC1	Factors with multiple dispersed zinc fingers[2.3.4]
Conserved_specific	ZNF257	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	ZNF335	Factors with multiple dispersed zinc fingers[2.3.4]

Table A. 4 (continued)

Conserved_specific	ZNF143	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	NFKB1	NF-kappaB-related factors[6.1.1]
Conserved_specific	CLOCK	PAS domain factors[1.2.5]
Conserved_specific	HINFP	Factors with multiple dispersed zinc fingers[2.3.4]
Conserved_specific	NFIC	Nuclear factor 1[7.1.2]
Conserved_specific	T	Brachyury-related factors[6.5.1]
Conserved_specific	MAFK	Maf-related factors[1.1.3]
Conserved_specific	MAFF	Maf-related factors[1.1.3]
Conserved_specific	ZNF263	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	VDR	Thyroid hormone receptor-related factors (NR1)[2.1.2]
Conserved_specific	NFKB2	NF-kappaB-related factors[6.1.1]
Conserved_specific	MAFG	Maf-related factors[1.1.3]
Conserved_specific	ZNF549	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	E2F4	E2F-related factors[3.3.2]
Conserved_specific	PATZ1	Factors with multiple dispersed zinc fingers[2.3.4]
Conserved_specific	SNAI1	More than 3 adjacent zinc finger factors[2.3.3]
Conserved_specific	ESRRG	Steroid hormone receptors (NR3)[2.1.1]
Conserved_specific	GLI3	More than 3 adjacent zinc finger factors[2.3.3]
Variable_specific	SRY	SOX-related factors[4.1.1]
Variable_specific	IRF7	Interferon-regulatory factors[3.5.3]
Variable_specific	PRDM6	More than 3 adjacent zinc finger factors[2.3.3]
Variable_specific	SOX5	SOX-related factors[4.1.1]
Variable_specific	HOXB13	HOX-related factors[3.1.1]
Variable_specific	FOXK1	Forkhead box (FOX) factors[3.3.1]
Variable_specific	FOXP1	Forkhead box (FOX) factors[3.3.1]
Variable_specific	GATA3	GATA-type zinc fingers[2.2.1]
Variable_specific	TBP	TBP-related factors[8.1.1]
Variable_specific	FOXM1	Forkhead box (FOX) factors[3.3.1]
Variable_specific	POU1F1	POU domain factors[3.1.10]
Variable_specific	NFATC1	NFAT-related factors[6.1.3]
Variable_specific	HOXA10	HOX-related factors[3.1.1]
Variable_specific	SMARCA1	Myb/SANT domain factors[3.5.1]
Variable_specific	DBP	C/EBP-related[1.1.8]
Variable_specific	FOXO4	Forkhead box (FOX) factors[3.3.1]
Variable_specific	POU2F1	POU domain factors[3.1.10]
Variable_specific	FEZF1	More than 3 adjacent zinc finger factors[2.3.3]

Table A. 4 (continued)

Variable_specific	SOX2	SOX-related factors[4.1.1]
Variable_specific	ZNF350	More than 3 adjacent zinc finger factors[2.3.3]
Variable_specific	PBX2	TALE-type homeo domain factors[3.1.4]
Variable_specific	NKX6-1	NK-related factors[3.1.2]
Variable_specific	ZNF394	More than 3 adjacent zinc finger factors[2.3.3]
Variable_specific	FOXA1	Forkhead box (FOX) factors[3.3.1]
Variable_specific	FOXJ2	Forkhead box (FOX) factors[3.3.1]
Variable_specific	AR	Steroid hormone receptors (NR3)[2.1.1]
Variable_specific	SOX4	SOX-related factors[4.1.1]
Variable_specific	HNF1A	POU domain factors[3.1.10]
Variable_specific	ARID5B	ARID-related factors[3.7.1]
Variable_specific	FOXO1	Forkhead box (FOX) factors[3.3.1]
Variable_specific	FOXA3	Forkhead box (FOX) factors[3.3.1]
Variable_specific	LEF1	TCF-7-related factors[4.1.3]
Variable_specific	ALX1	Paired-related HD factors[3.1.3]
Variable_specific	STAT2	STAT factors[6.2.1]
Variable_specific	NKX3-2	NK-related factors[3.1.2]
Variable_specific	IRF3	Interferon-regulatory factors[3.5.3]
Variable_specific	IRF1	Interferon-regulatory factors[3.5.3]
Variable_specific	LHX3	HD-LIM factors[3.1.5]
Variable_specific	GATA6	GATA-type zinc fingers[2.2.1]
Variable_specific	HOXA13	HOX-related factors[3.1.1]
Variable_specific	ZNF354A	More than 3 adjacent zinc finger factors[2.3.3]
Variable_specific	MEF2B	Regulators of differentiation[5.1.1]
Variable_specific	MEF2C	Regulators of differentiation[5.1.1]
Variable_specific	NR2E3	RXR-related receptors (NR2)[2.1.3]
Variable_specific	AIRE	AIRE[5.3.1]
Variable_specific	HNF1B	POU domain factors[3.1.10]
Variable_specific	MEF2A	Regulators of differentiation[5.1.1]
Variable_specific	BATF	B-ATF-related factors[1.1.4]
Variable_specific	FOXQ1	Forkhead box (FOX) factors[3.3.1]
Variable_specific	MECOM	Factors with multiple dispersed zinc fingers[2.3.4]
Variable_specific	CEBPE	C/EBP-related[1.1.8]
Variable_specific	IRF8	Interferon-regulatory factors[3.5.3]
Variable_specific	FOXC1	Forkhead box (FOX) factors[3.3.1]
Variable_specific	IRF9	Interferon-regulatory factors[3.5.3]

Table A. 4 (continued)

Variable_specific	SRF	Responders to external signals (SRF/RLM1)[5.1.2]
Variable_specific	DUX4	Paired-related HD factors[3.1.3]
Variable_specific	ZNF8	Factors with multiple dispersed zinc fingers[2.3.4]
Variable_specific	ZIM3	More than 3 adjacent zinc finger factors[2.3.3]
Variable_specific	SOX17	SOX-related factors[4.1.1]
Variable_specific	IRF2	Interferon-regulatory factors[3.5.3]
Variable_specific	HLF	C/EBP-related[1.1.8]
Variable_specific	ZNF250	More than 3 adjacent zinc finger factors[2.3.3]
Variable_specific	ZNF146	More than 3 adjacent zinc finger factors[2.3.3]
Variable_specific	ZNF85	More than 3 adjacent zinc finger factors[2.3.3]
Variable_specific	ZNF260	More than 3 adjacent zinc finger factors[2.3.3]
Variable_specific	OTX2	Paired-related HD factors[3.1.3]
Variable_specific	MEF2D	Regulators of differentiation[5.1.1]
Variable_specific	ZFP82	More than 3 adjacent zinc finger factors[2.3.3]
Variable_specific	ZNF418	Factors with multiple dispersed zinc fingers[2.3.4]
Variable_specific	BATF3	B-ATF-related factors[1.1.4]
Variable_specific	FOXJ3	Forkhead box (FOX) factors[3.3.1]
Variable_specific	TEAD4	TEF-1-related factors[3.6.1]
Variable_specific	ZNF136	More than 3 adjacent zinc finger factors[2.3.3]
Variable_specific	NKX3-1	NK-related factors[3.1.2]
Variable_specific	ZFP28	More than 3 adjacent zinc finger factors[2.3.3]
Variable_specific	ZNF490	More than 3 adjacent zinc finger factors[2.3.3]
Variable_specific	POU3F2	POU domain factors[3.1.10]
Variable_specific	ZNF586	More than 3 adjacent zinc finger factors[2.3.3]
Variable_specific	ATF4	ATF-4-related factors[1.1.6]
Variable_specific	TEAD1	TEF-1-related factors[3.6.1]
Variable_specific	CEBPG	C/EBP-related[1.1.8]

Table A. 5. Gene ontology (GREAT) output for genes associated with conserved CG DMRs between NeuN+ and OLIG2+.

DMR type	# Term Name	Binom FDR	FE[1]
CG NeuN hypo DMR	regulation of synaptic plasticity	9.1E-78	2.27
CG NeuN hypo DMR	regulation of dendritic spine development	4.4E-45	2.36
CG NeuN hypo DMR	positive regulation of synaptic transmission	5.0E-43	2.04
CG NeuN hypo DMR	neuromuscular junction development	1.9E-37	2.83
CG NeuN hypo DMR	long-term synaptic potentiation	2.2E-37	2.65
CG NeuN hypo DMR	regulation of neurotransmitter receptor activity	1.4E-36	2.73
CG NeuN hypo DMR	regulation of vesicle fusion	3.9E-35	2.52
CG NeuN hypo DMR	regulation of glutamate receptor signaling pathway	1.9E-34	2.72
CG NeuN hypo DMR	ephrin receptor signaling pathway	1.2E-32	2.12
CG NeuN hypo DMR	peptidyl-threonine phosphorylation	2.5E-32	2.23
CG NeuN hypo DMR	cardiac conduction	6.0E-32	2.06
CG NeuN hypo DMR	activation of GTPase activity	8.5E-32	2.16
CG NeuN hypo DMR	regulation of cardiac conduction	1.1E-30	2.24
CG NeuN hypo DMR	glutamate secretion	2.0E-30	2.99
CG NeuN hypo DMR	regulation of alpha-amino-3-hydroxy-5-methyl-4-isoxazole propionate selective glutamate receptor activity	2.0E-29	3.38
CG NeuN hypo DMR	vesicle docking	4.0E-29	2.56
CG NeuN hypo DMR	mRNA splice site selection	5.8E-29	3.29

Table A. 5 (continued)

CG NeuN hypo DMR	peptidyl-threonine modification	1.1E-28	2.03
CG NeuN hypo DMR	regulation of dendritic spine morphogenesis	3.4E-28	2.52
CG NeuN hypo DMR	detection of calcium ion	5.7E-26	3.42
CG NeuN hypo DMR	postsynapse	6.8E-180	2.03
CG NeuN hypo DMR	asymmetric synapse	3.0E-137	2.33
CG NeuN hypo DMR	postsynaptic specialization	6.8E-136	2.35
CG NeuN hypo DMR	postsynaptic density	4.4E-135	2.35
CG NeuN hypo DMR	neuron to neuron synapse	6.5E-135	2.31
CG NeuN hypo DMR	cation channel complex	1.1E-85	2.15
CG NeuN hypo DMR	axon part	5.4E-62	2.06
CG NeuN hypo DMR	potassium channel complex	9.5E-46	2.11
CG NeuN hypo DMR	voltage-gated potassium channel complex	2.6E-45	2.13
CG NeuN hypo DMR	G-protein coupled receptor heterodimeric complex	3.2E-43	15.21
CG NeuN hypo DMR	voltage-gated calcium channel complex	8.5E-30	2.57
CG NeuN hypo DMR	calcium channel complex	8.7E-30	2.21
CG NeuN hypo DMR	GABA receptor complex	1.6E-29	3.59
CG NeuN hypo DMR	terminal bouton	4.2E-29	2.35
CG NeuN hypo DMR	ionotropic glutamate receptor complex	1.6E-28	2.19

Table A. 5 (continued)

CG NeuN hypo DMR	main axon	2.2E-28	2.15
CG NeuN hypo DMR	G-protein coupled receptor dimeric complex	1.0E-20	4.40
CG NeuN hypo DMR	axon initial segment	1.3E-20	3.32
CG NeuN hypo DMR	postsynaptic specialization membrane	1.7E-20	3.48
CG NeuN hypo DMR	dense body	5.2E-20	5.40
CG NeuN hypo DMR	calmodulin binding	1.1E-60	2.07
CG NeuN hypo DMR	voltage-gated cation channel activity	3.9E-51	2.05
CG NeuN hypo DMR	calcium ion transmembrane transporter activity	2.2E-48	2.06
CG NeuN hypo DMR	syntaxin-1 binding	8.0E-37	3.69
CG NeuN hypo DMR	G-protein coupled GABA receptor activity	1.1E-35	10.10
CG NeuN hypo DMR	pre-mRNA binding	6.2E-31	3.20
CG NeuN hypo DMR	GABA receptor activity	1.1E-30	3.33
CG NeuN hypo DMR	voltage-gated calcium channel activity	1.2E-30	2.60
CG NeuN hypo DMR	glutamate-gated calcium ion channel activity	2.3E-28	6.17
CG NeuN hypo DMR	high voltage-gated calcium channel activity	4.2E-28	3.89
CG NeuN hypo DMR	NMDA glutamate receptor activity	2.9E-24	4.68
CG NeuN hypo DMR	calcium-transporting ATPase activity	3.9E-24	4.13
CG NeuN hypo DMR	neurotrophin binding	1.9E-23	4.03

Table A. 5 (continued)

CG NeuN hypo DMR	calcium-dependent protein binding	2.7E-23	2.41
CG NeuN hypo DMR	neurotrophin receptor activity	6.2E-22	5.53
CG NeuN hypo DMR	calcium-dependent protein kinase activity	5.3E-21	4.61
CG NeuN hypo DMR	brain-derived neurotrophic factor binding	2.3E-20	5.20
CG NeuN hypo DMR	delayed rectifier potassium channel activity	1.1E-19	2.81
CG NeuN hypo DMR	adenylate cyclase binding	7.2E-19	3.05
CG NeuN hypo DMR	calcium-dependent protein serine/threonine kinase activity	2.2E-18	4.38
CG OLIG2 hypo DMR	myelin assembly	8.9E-19	2.08
CG OLIG2 hypo DMR	otic vesicle formation	3.4E-39	3.35
CG OLIG2 hypo DMR	otic vesicle morphogenesis	1.4E-37	3.00
CG OLIG2 hypo DMR	prostate gland morphogenetic growth	2.3E-36	3.30
CG OLIG2 hypo DMR	lacrimal gland development	3.1E-35	3.05
CG OLIG2 hypo DMR	ureter development	1.8E-34	2.52
CG OLIG2 hypo DMR	otic vesicle development	2.7E-32	2.49
CG OLIG2 hypo DMR	limb bud formation	4.3E-29	2.28
CG OLIG2 hypo DMR	forebrain ventricular zone progenitor cell division	1.6E-28	3.14
CG OLIG2 hypo DMR	neural crest cell fate specification	1.8E-28	4.35
CG OLIG2 hypo DMR	nephron tubule formation	8.4E-28	2.12

Table A. 5 (continued)

CG OLIG2 hypo DMR	regulation of epithelial cell proliferation involved in lung morphogenesis	1.6E-27	2.55
CG OLIG2 hypo DMR	embryonic camera-type eye morphogenesis	1.6E-26	2.02
CG OLIG2 hypo DMR	stem cell fate specification	1.7E-26	3.70
CG OLIG2 hypo DMR	primary sex determination	2.4E-25	3.75
CG OLIG2 hypo DMR	intrahepatic bile duct development	5.9E-25	5.03
CG OLIG2 hypo DMR	epithelial cell proliferation involved in prostatic bud elongation	5.9E-25	5.03
CG OLIG2 hypo DMR	regulation of cell proliferation involved in tissue homeostasis	5.9E-25	5.03
CG OLIG2 hypo DMR	metanephric nephron morphogenesis	8.5E-25	2.08
CG OLIG2 hypo DMR	neural crest cell fate commitment	1.4E-24	3.44
CG OLIG2 hypo DMR	ARC complex	5.2E-14	5.10
CG OLIG2 hypo DMR	guanyl-nucleotide exchange factor complex	2.8E-13	2.08
CG OLIG2 hypo DMR	RISC-loading complex	3.3E-13	2.91
CG OLIG2 hypo DMR	micro-ribonucleoprotein complex	2.1E-11	2.70
CG OLIG2 hypo DMR	RNAi effector complex	4.4E-11	2.25
CG OLIG2 hypo DMR	Derlin-1-VIMP complex	4.8E-11	5.88
CG OLIG2 hypo DMR	PCAF complex	8.5E-10	4.07
CG OLIG2 hypo DMR	XPC complex	6.6E-09	3.22
CG OLIG2 hypo DMR	hemidesmosome	1.5E-08	2.39

Table A. 5 (continued)

CG OLIG2 hypo DMR	beta-catenin-TCF7L2 complex	2.5E-08	2.34
CG OLIG2 hypo DMR	catenin-TCF7L2 complex	8.2E-08	2.26
CG OLIG2 hypo DMR	DBIRD complex	8.2E-08	3.26
CG OLIG2 hypo DMR	signal recognition particle receptor complex	1.9E-07	3.14
CG OLIG2 hypo DMR	CBM complex	3.5E-07	2.89
CG OLIG2 hypo DMR	VCP-NPL4-UFD1 AAA ATPase complex	4.6E-07	2.36
CG OLIG2 hypo DMR	condensed nuclear chromosome inner kinetochore	9.8E-07	7.84
CG OLIG2 hypo DMR	tumor necrosis factor receptor superfamily complex	1.0E-06	7.02
CG OLIG2 hypo DMR	mitotic checkpoint complex	1.0E-06	3.61
CG OLIG2 hypo DMR	bub1-bub3 complex	1.0E-06	3.61
CG OLIG2 hypo DMR	phosphopyruvate hydratase complex	1.5E-06	3.91
CG OLIG2 hypo DMR	RNA polymerase II repressing transcription factor binding	7.2E-19	2.03
CG OLIG2 hypo DMR	fibroblast growth factor-activated receptor activity	1.8E-16	3.39
CG OLIG2 hypo DMR	CD8 receptor binding	2.2E-16	5.86
CG OLIG2 hypo DMR	ribonuclease III activity	7.2E-15	3.84
CG OLIG2 hypo DMR	armadillo repeat domain binding	1.1E-14	2.54
CG OLIG2 hypo DMR	1-phosphatidylinositol-5-phosphate 4-kinase activity	1.9E-14	4.62
CG OLIG2 hypo DMR	type 5 metabotropic glutamate receptor binding	5.7E-14	3.01

Table A. 5 (continued)

CG OLIG2 hypo DMR	deoxyribonuclease I activity	1.6E-13	4.65
CG OLIG2 hypo DMR	1-phosphatidylinositol-4-phosphate 5-kinase activity	1.3E-11	3.29
CG OLIG2 hypo DMR	1-phosphatidylinositol-3-phosphate 4-kinase activity	3.3E-11	3.44
CG OLIG2 hypo DMR	methylated-DNA-[protein]-cysteine S-methyltransferase activity	5.6E-11	3.44
CG OLIG2 hypo DMR	pre-miRNA binding	1.1E-10	3.43
CG OLIG2 hypo DMR	pre-mRNA intronic binding	4.5E-10	2.48
CG OLIG2 hypo DMR	T cell receptor binding	5.3E-10	3.11
CG OLIG2 hypo DMR	tropomyosin binding	5.5E-10	2.27
CG OLIG2 hypo DMR	virion binding	9.3E-10	2.58
CG OLIG2 hypo DMR	rRNA (guanosine-2'-O-)-methyltransferase activity	2.0E-09	4.01
CG OLIG2 hypo DMR	DNA-methyltransferase activity	3.0E-09	2.39
CG OLIG2 hypo DMR	G-protein coupled glutamate receptor binding	1.8E-08	2.24
CG OLIG2 hypo DMR	CD4 receptor binding	1.9E-08	2.65

[1]FE: fold enrichment

Table A. 6. Association of differentially expressed genes (DEG) and DMR-genes.

DEG type	DMR-gene type	# of genes[1]	Odds ratio	P-value
Human-UP NeuN+	mCG Human-hyper NeuN+	9037;313;262;5	0.54	9.50E-01
Human-UP NeuN+	**mCG Human-hypo NeuN+**	**9037;313;717;37**	**1.52**	**9.00E-03**
Human-DOWN NeuN+	mCG Human-hyper NeuN+	9037;181;262;5	0.95	6.10E-01
Human-DOWN NeuN+	mCG Human-hypo NeuN+	9037;181;717;11	0.76	8.60E-01
Human-UP OLIG2+	mCG Human-hyper OLIG2+	9037;354;298;10	0.85	7.40E-01
Human-UP OLIG2+	**mCG Human-hypo OLIG2+**	**9037;354;552;31**	**1.46**	**2.70E-02**
Human-DOWN OLIG2+	mCG Human-hyper OLIG2+	9037;179;298;3	0.5	9.40E-01
Human-DOWN OLIG2+	mCG Human-hypo OLIG2+	9037;179;552;10	0.91	6.60E-01

[1] # of all orthologous genes expressed in NeuN+; # of DEGs; # of DMR-genes; # of overlap between DEGs with DMR-genes

Table A. 7. Three-dimensional chromatin Interaction profiles for human neuron CG DMRs.

Chromosome	DMR1_start	DMR1_end	DMR2_start	DMR2_end
chr8	145148281	145149216	145998048	145999189
chr9	140374163	140374609	140425003	140425420
chr17	5488727	5488946	6297077	6297463
chr2	120397080	120397368	120581182	120582332
chr18	61009566	61009665	60683377	60683494
chr6	159423639	159423743	160070649	160071041
chr4	185186763	185187403	185905643	185905908
chr12	31741600	31742027	32674286	32674774
chr4	1974091	1974618	2007413	2007773
chr12	122109709	122110258	122355246	122355636
chr11	60693440	60694121	60897649	60897782
chr8	142158244	142159039	141849213	141849299
chr8	144156247	144156462	144432099	144432433
chr3	107409990	107410835	108180570	108181050
chr15	93618135	93618635	94550656	94551123
chr6	710426	711826	1699131	1700235
chr8	41692955	41693075	41573648	41574533
chr20	30007668	30011036	29998900	30000624
chr2	208919627	208920218	208621549	208622136
chr16	48644359	48644730	48309672	48310323
chr16	12990780	12990909	13929128	13930648
chr3	126399360	126399896	127178020	127178134
chr6	169361157	169361978	169899046	169900347
chr8	67342861	67343760	66975924	66976744
chr2	68696763	68696957	69547747	69547844
chr2	68696763	68696957	69152321	69152841
chr11	60708025	60708478	60809013	60809344
chr16	48644359	48644730	48310977	48311513
chr4	113432450	113432637	113271457	113271652
chr20	29963100	29964215	29998900	30000624
chr11	69514746	69515349	68919875	68920140
chr18	77288526	77288614	77152688	77153550
chr2	9487885	9488285	9319444	9319552
chr2	9487885	9488285	9302585	9302974
chr6	110526893	110526966	110677371	110677726

Table A. 7 (continued)

chr11	574130	574273	696503	697163
chr8	11626450	11626687	11664930	11665230
chr5	173236055	173236120	173260376	173260821
chr18	3279931	3280035	3453836	3454017
chr8	142193554	142193720	141849213	141849299
chr15	100252193	100252905	99792722	99792809
chr20	29998900	30000624	29979118	29980313
chr20	29998900	30000624	29963100	29964215
chr4	113432450	113432637	113218384	113219113
chr11	62181904	62182027	62316367	62316430
chr18	3279931	3280035	3652876	3654558
chr9	546271	547720	546271	547720
chr8	142193554	142193720	141863750	141864595
chr12	116993376	116993629	117048319	117048384
chr6	27112594	27113055	27728086	27731326
chr2	202006735	202007150	201754731	201755116
chr20	29979118	29980313	29998900	30000624
chr15	86287198	86287302	86380318	86380711
chr4	7043263	7043392	6986436	6987164
chr14	91075377	91076483	90896401	90896783
chr14	91075377	91076483	90882977	90883566
chr21	43223019	43223402	42637335	42638459
chr2	10439495	10439753	10360711	10361108
chr2	10439495	10439753	10342647	10343058
chr7	121514511	121514732	121243330	121245795
chr11	598826	599155	598826	599155
chr20	29950946	29952785	30007668	30011036
chr11	93392145	93392818	92969873	92970428
chr16	87899783	87900324	87614309	87614552
chr20	29950946	29952785	29979118	29980313
chr3	127633556	127633768	127494063	127494368
chr3	71353312	71353997	70579578	70580143
chr8	142158244	142159039	141863750	141864595
chr4	6788323	6788647	6986436	6987164
chr20	29979118	29980313	29950946	29952785
chr20	29979118	29980313	29935052	29935123

Table A. 7 (continued)

chr8	142148178	142148470	141849213	141849299
chr9	135293208	135294865	135463667	135463754
chr2	2337254	2337759	3304537	3305034
chr13	114107155	114107280	113929508	113930151
chr9	140637039	140637617	140723560	140724155
chr11	92969873	92970428	93392145	93392818
chr5	4519547	4519879	3709040	3710552
chr8	144432099	144432433	144156247	144156462
chr6	27112594	27113055	27778414	27778541
chr6	110677371	110677726	110526893	110526966
chr4	3204685	3204843	3278747	3279277
chr4	3204685	3204843	3267372	3267657
chr16	3451820	3452364	3361385	3362750
chr16	48285998	48286127	48644359	48644730
chr3	16349338	16350044	16546090	16546888
chr18	77152688	77153550	77378864	77380085
chr18	77152688	77153550	77288526	77288614
chr1	11026448	11026635	10570639	10573205
chr1	36615863	36617947	36754509	36754863
chr11	66606019	66606263	66492913	66493554
chr8	142148178	142148470	142049147	142050235
chr11	117742251	117742667	117843375	117843895
chr5	173236055	173236120	173184454	173184824
chr8	41689805	41691302	41573648	41574533
chr2	120407892	120410678	120581182	120582332
chr8	9521379	9522029	9207998	9208423
chr4	3174179	3175520	3809147	3810264
chr11	62316367	62316430	62181904	62182027
chr16	51183282	51183378	51789366	51789674
chr5	139536374	139536525	139428483	139428585
chr11	68919875	68920140	69514746	69515349
chr14	103968154	103968949	104014168	104014266
chr11	125777710	125777873	126068708	126069272
chr6	16200639	16201079	15706076	15706585
chr1	204412084	204412402	204477275	204477744
chr10	116849576	116851817	116637263	116637625

Table A. 7 (continued)

chr18	3652876	3654558	3279931	3280035
chr4	152601154	152603165	152032769	152033047
chr3	124175533	124175633	123720218	123720576
chr4	6986436	6987164	6788323	6788647
chr6	159459522	159460416	160070649	160071041
chr6	159423639	159423743	159128888	159129250
chr15	99792722	99792809	100252193	100252905
chr6	52283764	52284433	52376482	52376647
chr15	101981827	101982177	102156691	102156911
chr12	49751948	49752319	49937369	49939010
chr1	36754509	36754863	36615863	36617947
chr9	135293208	135294865	135435912	135437844
chr5	139428483	139428585	139536374	139536525
chr9	115140681	115140853	115468241	115468325
chr13	112237303	112237489	113097314	113099228
chr8	1489854	1490077	1471821	1472040
chr17	1488595	1490976	1262954	1263885
chr2	217364686	217364951	216981251	216983091
chr18	3453836	3454017	3279931	3280035
chr4	909244	909785	830232	832167
chr4	909244	909785	909244	909785
chr4	113218384	113219113	113432450	113432637
chr4	113271457	113271652	113432450	113432637
chr6	159639020	159639293	160070649	160071041
chr1	36043652	36043901	36240504	36240784
chr4	1974091	1974618	1508772	1508925
chr20	29935052	29935123	29979118	29980313
chr20	29935052	29935123	29982661	29984776
chr3	108180570	108181050	107409990	107410835
chr10	12231172	12232484	13116696	13116777
chr9	98075703	98076254	98980833	98982341
chr9	13444503	13444652	14346519	14346674
chr2	3304537	3305034	2337254	2337759
chr2	9530868	9531082	9302585	9302974
chr2	9530868	9531082	9353287	9353484
chr6	169361157	169361978	170054064	170055846

Table A. 7 (continued)

chr8	142158244	142159039	142049147	142050235
chr5	67535222	67535514	66534239	66534504
chr1	154404242	154404455	154439458	154439542
chr8	38562302	38562375	38408369	38408924
chr16	4002026	4002442	3976480	3976729
chr5	168125769	168126016	169008697	169009297
chr16	1675018	1675252	1607684	1607791
chr11	66492913	66493554	66606019	66606263
chr1	8680303	8680895	8746926	8747163
chr17	1262954	1263885	1488595	1490976
chr14	104014168	104014266	103968154	103968949
chr4	6788323	6788647	6944314	6945708
chr5	134892901	134893026	134779092	134779952
chr11	125803029	125803296	126068708	126069272
chr5	173211905	173212814	173236055	173236120
chr8	11664930	11665230	11626450	11626687
chr2	216981251	216983091	217364686	217364951
chr14	21798680	21798806	21701119	21702057
chr9	27527070	27527513	27338161	27338736
chr7	130598497	130598908	130698629	130698697
chr12	122355246	122355636	122109709	122110258
chr6	28616985	28617483	28910952	28911221
chr13	27998114	27998202	28555201	28557055
chr16	1669023	1669244	1607684	1607791
chr13	44715484	44717469	44880709	44881113
chr9	14346519	14346674	13444503	13444652
chr16	3361385	3362750	3451820	3452364
chr9	130693708	130693848	130369975	130370458
chr2	10976424	10977657	11103608	11104445
chr6	170494794	170495159	170433199	170434077
chr6	134492177	134492300	135223946	135224369
chr6	167555853	167556037	167192345	167192689
chr3	52520163	52520423	52494991	52495352
chr9	140723560	140724155	140883593	140883876
chr6	16132062	16136539	15706076	15706585
chr6	16146602	16146948	15706076	15706585

Table A. 7 (continued)

chr7	104558936	104559408	103872030	103872393
chr13	21061903	21062662	21394314	21395209
chr5	173236055	173236120	173211905	173212814
chr17	161146	161901	868468	868830
chr20	30007668	30011036	29950946	29952785
chr20	29998900	30000624	29982661	29984776
chr8	145148281	145149216	145522376	145522552
chr9	139989779	139990557	140425003	140425420
chr9	140723560	140724155	140637039	140637617
chr3	37864732	37865546	37374546	37375216
chr21	46312390	46312868	45775435	45776847
chr8	11626450	11626687	11828094	11828386
chr7	104558936	104559408	103882407	103883182
chr3	52494991	52495352	52520163	52520423
chr6	27778414	27778541	27112594	27113055
chr5	137475360	137475717	137167320	137168403
chr12	122109709	122110258	122223048	122223163
chr5	34006884	34007296	34467975	34468475
chr15	101833362	101834236	102216098	102216850
chr15	101833362	101834236	102156691	102156911
chr6	34458536	34458635	34137456	34138344
chr4	3045531	3045901	3809147	3810264
chr13	45152779	45153144	45481055	45481432
chr2	240114989	240115861	240139250	240139380
chr20	29998900	30000624	30007668	30011036
chr4	2609495	2609934	3278747	3279277
chr3	128349611	128349731	127821216	127821862
chr12	49937369	49939010	49751948	49752319
chr4	6986436	6987164	7043263	7043392
chr5	169008697	169009297	168477935	168478206
chr9	130369975	130370458	130693708	130693848
chr18	3029220	3029814	3093161	3094168
chr18	3093161	3094168	3029220	3029814
chr2	201754731	201755116	202006735	202007150

Table A. 8. Gene ontology output (ShinyGO) for three-dimensional chromatin Interaction profiles for human neuron CG DMRs.

Enrichment FDR	**# of genes**	**Functional Category**
2.6E-02	45	Organelle organization
2.6E-02	22	Cytoskeleton organization
2.6E-02	24	Neurogenesis
2.6E-02	22	Neuron differentiation
2.6E-02	19	Cellular component morphogenesis
2.6E-02	29	Cell development
2.6E-02	19	Neuron development
2.6E-02	23	Generation of neurons
3.4E-02	32	Regulation of cellular component organization
3.5E-02	22	Plasma membrane bounded cell projection organization

Table A. 9. LD score regression results

DMR class	Disease	h2	FE	P-value	FDR
CG Conserved NeuN Hypo	Alzheimer	0.39	1.92	2.1E-03	**4.1E-02**
CG Conserved NeuN Hypo	Anorexia	0.34	1.69	6.2E-05	**1.7E-03**
CG Conserved NeuN Hypo	Autism	0.28	1.39	3.7E-02	2.5E-01
CG Conserved NeuN Hypo	Bipolar_Disorder	0.40	1.95	3.1E-05	**9.3E-04**
CG Conserved NeuN Hypo	DS	0.32	1.58	3.9E-03	6.7E-02
CG Conserved NeuN Hypo	Neuroticism	0.34	1.65	9.2E-09	**1.2E-06**
CG Conserved NeuN Hypo	Schizophrenia	0.39	1.94	1.4E-23	**3.6E-21**
CG Conserved NeuN Hypo	SWB	0.35	1.74	3.5E-03	6.3E-02
CG Conserved NeuN Hypo	Years_of_Education1	0.37	1.84	6.4E-07	**2.5E-05**
CG Conserved NeuN Hypo	BMI1	0.31	1.52	5.0E-07	**2.2E-05**
CG Conserved NeuN Hypo	Celiac	0.20	0.96	8.9E-01	9.9E-01
CG Conserved NeuN Hypo	Coronary_Artery_Disease	0.32	1.59	2.3E-02	2.0E-01
CG Conserved NeuN Hypo	Crohns_Disease	0.21	1.02	8.8E-01	9.9E-01
CG Conserved NeuN Hypo	Ever_Smoked	0.30	1.50	8.4E-03	9.9E-02
CG Conserved NeuN Hypo	Fasting_Glucose	0.33	1.61	3.1E-02	2.4E-01
CG Conserved NeuN Hypo	HDL	0.27	1.33	3.5E-02	2.5E-01
CG Conserved NeuN Hypo	Height1	0.22	1.09	3.6E-01	7.6E-01
CG Conserved NeuN Hypo	IBD	0.20	0.98	8.7E-01	9.9E-01
CG Conserved NeuN Hypo	LDL	0.20	1.00	9.9E-01	1.0E+00
CG Conserved NeuN Hypo	Lupus	0.26	1.26	3.1E-01	7.3E-01

Table A. 9 (continued)

CG Conserved NeuN Hypo	Multiple_sclerosis	0.30	1.49	5.0E-01	8.2E-01
CG Conserved NeuN Hypo	Primary_biliary_cirrhosis	0.25	1.22	3.9E-01	7.9E-01
CG Conserved NeuN Hypo	Rheumatoid_Arthritis	0.27	1.31	1.5E-01	6.2E-01
CG Conserved NeuN Hypo	Triglycerides	0.31	1.52	1.4E-02	1.4E-01
CG Conserved NeuN Hypo	Type_1_Diabetes	0.23	1.11	8.1E-01	9.6E-01
CG Conserved NeuN Hypo	Type_2_Diabetes	0.21	1.04	8.5E-01	9.9E-01
CG Conserved NeuN Hypo	Ulcerative_Colitis	0.18	0.88	5.3E-01	8.3E-01
CG Conserved OLIG2 Hypo	Alzheimer	0.42	0.94	7.8E-01	9.4E-01
CG Conserved OLIG2 Hypo	Anorexia	0.53	1.20	4.8E-02	3.1E-01
CG Conserved OLIG2 Hypo	Autism	0.49	1.10	3.5E-01	7.6E-01
CG Conserved OLIG2 Hypo	Bipolar_Disorder	0.29	0.66	4.9E-03	7.8E-02
CG Conserved OLIG2 Hypo	DS	0.52	1.16	1.1E-01	5.4E-01
CG Conserved OLIG2 Hypo	Neuroticism	0.51	1.14	1.6E-02	1.6E-01
CG Conserved OLIG2 Hypo	Schizophrenia	0.46	1.02	4.7E-01	8.1E-01
CG Conserved OLIG2 Hypo	SWB	0.49	1.09	5.3E-01	8.3E-01
CG Conserved OLIG2 Hypo	Years_of_Education1	0.43	0.96	6.2E-01	8.6E-01
CG Conserved OLIG2 Hypo	BMI1	0.55	1.24	1.4E-04	**3.5E-03**
CG Conserved OLIG2 Hypo	Celiac	0.45	1.02	8.8E-01	9.9E-01
CG Conserved OLIG2 Hypo	Coronary_Artery_Disease	0.49	1.10	4.8E-01	8.1E-01
CG Conserved OLIG2 Hypo	Crohns_Disease	0.55	1.23	9.8E-03	1.1E-01

Table A. 9 (continued)

CG Conserved OLIG2 Hypo	Ever_Smoked	0.51	1.16	1.4E-01	5.8E-01
CG Conserved OLIG2 Hypo	Fasting_Glucose	0.59	1.32	5.1E-02	3.1E-01
CG Conserved OLIG2 Hypo	HDL	0.51	1.14	9.2E-02	4.5E-01
CG Conserved OLIG2 Hypo	Height1	0.57	1.29	6.0E-08	**5.4E-06**
CG Conserved OLIG2 Hypo	IBD	0.50	1.12	1.6E-01	6.2E-01
CG Conserved OLIG2 Hypo	LDL	0.44	0.99	9.2E-01	9.9E-01
CG Conserved OLIG2 Hypo	Lupus	0.52	1.16	2.5E-01	6.9E-01
CG Conserved OLIG2 Hypo	Multiple_sclerosis	0.73	1.64	8.9E-02	4.4E-01
CG Conserved OLIG2 Hypo	Primary_biliary_cirrhosis	0.59	1.32	6.7E-03	9.2E-02
CG Conserved OLIG2 Hypo	Rheumatoid_Arthritis	0.58	1.30	6.5E-03	9.2E-02
CG Conserved OLIG2 Hypo	Triglycerides	0.47	1.06	4.7E-01	8.1E-01
CG Conserved OLIG2 Hypo	Type_1_Diabetes	0.45	1.01	9.6E-01	1.0E+00
CG Conserved OLIG2 Hypo	Type_2_Diabetes	0.58	1.31	2.6E-02	2.1E-01
CG Conserved OLIG2 Hypo	Ulcerative_Colitis	0.49	1.11	3.0E-01	7.3E-01
CG Human-specific NeuN Hypo	Alzheimer	0.05	0.98	9.8E-01	1.0E+00
CG Human-specific NeuN Hypo	Anorexia	0.07	1.36	3.5E-01	7.6E-01
CG Human-specific NeuN Hypo	Autism	0.09	1.84	4.5E-02	3.0E-01
CG Human-specific NeuN Hypo	Bipolar_Disorder	0.09	1.74	1.7E-01	6.3E-01

Table A. 9 (continued)

CG Human-specific NeuN Hypo	DS	0.05	0.99	9.8E-01	1.0E+00
CG Human-specific NeuN Hypo	Neuroticism	0.06	1.25	4.0E-01	7.9E-01
CG Human-specific NeuN Hypo	Schizophrenia	0.08	1.57	7.4E-04	**1.5E-02**
CG Human-specific NeuN Hypo	SWB	0.07	1.31	6.0E-01	8.6E-01
CG Human-specific NeuN Hypo	Years_of_Education1	0.05	1.01	9.9E-01	1.0E+00
CG Human-specific NeuN Hypo	BMI1	0.05	1.06	7.5E-01	9.4E-01
CG Human-specific NeuN Hypo	Celiac	0.05	0.97	9.6E-01	1.0E+00
CG Human-specific NeuN Hypo	Coronary_Artery_Disease	0.07	1.47	3.3E-01	7.4E-01
CG Human-specific NeuN Hypo	Crohns_Disease	0.04	0.77	4.5E-01	8.0E-01
CG Human-specific NeuN Hypo	Ever_Smoked	0.10	1.90	3.1E-02	2.4E-01
CG Human-specific NeuN Hypo	Fasting_Glucose	0.04	0.78	7.1E-01	9.1E-01
CG Human-specific NeuN Hypo	HDL	0.04	0.85	7.1E-01	9.1E-01
CG Human-specific NeuN Hypo	Height1	0.06	1.18	3.9E-01	7.9E-01

Table A. 9 (continued)

CG Human-specific NeuN Hypo	IBD	0.04	0.83	5.9E-01	8.6E-01
CG Human-specific NeuN Hypo	LDL	0.06	1.26	5.4E-01	8.3E-01
CG Human-specific NeuN Hypo	Lupus	0.05	1.00	1.0E+00	1.0E+00
CG Human-specific NeuN Hypo	Multiple_sclerosis	0.09	1.80	6.4E-01	8.6E-01
CG Human-specific NeuN Hypo	Primary_biliary_cirrhosis	0.09	1.86	1.2E-01	5.6E-01
CG Human-specific NeuN Hypo	Rheumatoid_Arthritis	0.07	1.35	5.0E-01	8.2E-01
CG Human-specific NeuN Hypo	Triglycerides	0.07	1.42	2.9E-01	7.3E-01
CG Human-specific NeuN Hypo	Type_1_Diabetes	0.09	1.79	4.6E-01	8.1E-01
CG Human-specific NeuN Hypo	Type_2_Diabetes	0.05	0.94	9.2E-01	9.9E-01
CG Human-specific NeuN Hypo	Ulcerative_Colitis	0.03	0.54	3.3E-01	7.4E-01
CG Human-specific OLIG2 Hypo	Alzheimer	0.00	0.10	2.5E-01	6.9E-01
CG Human-specific OLIG2 Hypo	Anorexia	0.06	1.40	3.4E-01	7.5E-01
CG Human-specific OLIG2 Hypo	Autism	0.05	1.26	5.4E-01	8.3E-01

Table A. 9 (continued)

CG Human-specific OLIG2 Hypo	Bipolar_Disorder	0.04	0.92	9.0E-01	9.9E-01
CG Human-specific OLIG2 Hypo	DS	0.06	1.40	4.1E-01	7.9E-01
CG Human-specific OLIG2 Hypo	Neuroticism	0.05	1.24	3.6E-01	7.6E-01
CG Human-specific OLIG2 Hypo	Schizophrenia	0.04	0.99	9.6E-01	1.0E+00
CG Human-specific OLIG2 Hypo	SWB	0.03	0.66	6.4E-01	8.6E-01
CG Human-specific OLIG2 Hypo	Years_of_Education1	0.05	1.33	4.2E-01	8.0E-01
CG Human-specific OLIG2 Hypo	BMI1	0.05	1.15	5.4E-01	8.3E-01
CG Human-specific OLIG2 Hypo	Celiac	0.06	1.59	5.1E-01	8.3E-01
CG Human-specific OLIG2 Hypo	Coronary_Artery_Disease	0.03	0.81	7.6E-01	9.4E-01
CG Human-specific OLIG2 Hypo	Crohns_Disease	0.03	0.77	5.6E-01	8.5E-01
CG Human-specific OLIG2 Hypo	Ever_Smoked	0.05	1.16	7.4E-01	9.4E-01
CG Human-specific OLIG2 Hypo	Fasting_Glucose	0.02	0.48	4.8E-01	8.1E-01
CG Human-specific OLIG2 Hypo	HDL	0.07	1.73	2.1E-01	6.7E-01

Table A. 9 (continued)

CG Human-specific OLIG2 Hypo	Height1	0.04	1.04	8.8E-01	9.9E-01
CG Human-specific OLIG2 Hypo	IBD	0.01	0.22	2.6E-02	2.1E-01
CG Human-specific OLIG2 Hypo	LDL	0.02	0.51	2.6E-01	6.9E-01
CG Human-specific OLIG2 Hypo	Lupus	0.03	0.67	6.1E-01	8.6E-01
CG Human-specific OLIG2 Hypo	Multiple_sclerosis	0.04	0.99	1.0E+00	1.0E+00
CG Human-specific OLIG2 Hypo	Primary_biliary_cirrhosis	0.08	1.96	1.8E-01	6.3E-01
CG Human-specific OLIG2 Hypo	Rheumatoid_Arthritis	0.03	0.64	4.8E-01	8.1E-01
CG Human-specific OLIG2 Hypo	Triglycerides	0.06	1.57	2.5E-01	6.9E-01
CG Human-specific OLIG2 Hypo	Type_1_Diabetes	0.07	1.80	4.7E-01	8.1E-01
CG Human-specific OLIG2 Hypo	Type_2_Diabetes	0.06	1.56	4.0E-01	7.9E-01
CG Human-specific OLIG2 Hypo	Ulcerative_Colitis	0.01	0.18	1.7E-01	6.3E-01
CG Chimp-specific NeuN Hypo	Alzheimer	0.06	2.25	2.5E-01	6.9E-01
CG Chimp-specific NeuN Hypo	Anorexia	0.03	1.26	6.4E-01	8.6E-01

Table A. 9 (continued)

CG Chimp-specific NeuN Hypo	Autism	0.05	1.99	1.3E-01	5.8E-01
CG Chimp-specific NeuN Hypo	Bipolar_Disorder	0.05	1.72	2.7E-01	7.0E-01
CG Chimp-specific NeuN Hypo	DS	0.01	0.50	4.2E-01	8.0E-01
CG Chimp-specific NeuN Hypo	Neuroticism	0.03	1.20	5.0E-01	8.2E-01
CG Chimp-specific NeuN Hypo	Schizophrenia	0.03	1.06	7.8E-01	9.4E-01
CG Chimp-specific NeuN Hypo	SWB	0.03	1.26	7.5E-01	9.4E-01
CG Chimp-specific NeuN Hypo	Years_of_Education1	0.03	0.96	9.3E-01	1.0E+00
CG Chimp-specific NeuN Hypo	BMI1	0.02	0.75	3.2E-01	7.3E-01
CG Chimp-specific NeuN Hypo	Celiac	0.06	2.05	3.0E-01	7.3E-01
CG Chimp-specific NeuN Hypo	Coronary_Artery_Disease	0.04	1.53	4.4E-01	8.0E-01
CG Chimp-specific NeuN Hypo	Crohns_Disease	0.04	1.50	3.7E-01	7.9E-01
CG Chimp-specific NeuN Hypo	Ever_Smoked	0.02	0.86	7.9E-01	9.5E-01
CG Chimp-specific NeuN Hypo	Fasting_Glucose	0.00	-0.13	1.2E-01	5.6E-01

Table A. 9 (continued)

CG Chimp-specific NeuN Hypo	HDL	0.04	1.57	4.0E-01	7.9E-01
CG Chimp-specific NeuN Hypo	Height1	0.02	0.83	4.1E-01	8.0E-01
CG Chimp-specific NeuN Hypo	IBD	0.04	1.46	4.3E-01	8.0E-01
CG Chimp-specific NeuN Hypo	LDL	0.00	0.00	3.4E-02	2.5E-01
CG Chimp-specific NeuN Hypo	Lupus	0.00	0.04	1.7E-01	6.3E-01
CG Chimp-specific NeuN Hypo	Multiple_sclerosis	0.04	1.62	7.7E-01	9.4E-01
CG Chimp-specific NeuN Hypo	Primary_biliary_cirrhosis	0.02	0.59	6.2E-01	8.6E-01
CG Chimp-specific NeuN Hypo	Rheumatoid_Arthritis	0.02	0.78	7.5E-01	9.4E-01
CG Chimp-specific NeuN Hypo	Triglycerides	0.03	1.15	7.7E-01	9.4E-01
CG Chimp-specific NeuN Hypo	Type_1_Diabetes	0.04	1.56	7.0E-01	9.1E-01
CG Chimp-specific NeuN Hypo	Type_2_Diabetes	0.05	1.69	3.2E-01	7.3E-01
CG Chimp-specific NeuN Hypo	Ulcerative_Colitis	0.06	2.14	2.2E-01	6.9E-01
CG Chimp-specific OLIG2 Hypo	Alzheimer	0.01	2.13	6.0E-01	8.6E-01

Table A. 9 (continued)

CG Chimp-specific OLIG2 Hypo	Anorexia	0.00	0.01	4.3E-01	8.0E-01
CG Chimp-specific OLIG2 Hypo	Autism	0.01	2.04	5.2E-01	8.3E-01
CG Chimp-specific OLIG2 Hypo	Bipolar_Disorder	0.01	3.02	3.1E-01	7.3E-01
CG Chimp-specific OLIG2 Hypo	DS	0.01	3.14	1.8E-01	6.3E-01
CG Chimp-specific OLIG2 Hypo	Neuroticism	0.01	2.28	1.8E-01	6.3E-01
CG Chimp-specific OLIG2 Hypo	Schizophrenia	0.01	2.12	6.5E-02	3.7E-01
CG Chimp-specific OLIG2 Hypo	SWB	0.01	2.20	5.7E-01	8.6E-01
CG Chimp-specific OLIG2 Hypo	Years_of_Education1	0.01	2.52	3.4E-01	7.5E-01
CG Chimp-specific OLIG2 Hypo	BMI1	0.01	1.85	2.3E-01	6.9E-01
CG Chimp-specific OLIG2 Hypo	Celiac	-0.01	-1.80	1.5E-01	6.0E-01
CG Chimp-specific OLIG2 Hypo	Coronary_Artery_Disease	-0.01	-1.80	1.8E-01	6.3E-01
CG Chimp-specific OLIG2 Hypo	Crohns_Disease	-0.01	-1.33	1.7E-02	1.7E-01
CG Chimp-specific OLIG2 Hypo	Ever_Smoked	0.00	0.94	9.7E-01	1.0E+00

Table A. 9 (continued)

CG Chimp-specific OLIG2 Hypo	Fasting_Glucose	0.02	3.91	2.0E-01	6.7E-01
CG Chimp-specific OLIG2 Hypo	HDL	0.00	0.81	8.6E-01	9.9E-01
CG Chimp-specific OLIG2 Hypo	Height1	0.00	0.42	4.0E-01	7.9E-01
CG Chimp-specific OLIG2 Hypo	IBD	0.00	-0.89	2.2E-02	2.0E-01
CG Chimp-specific OLIG2 Hypo	LDL	0.00	1.00	1.0E+00	1.0E+00
CG Chimp-specific OLIG2 Hypo	Lupus	0.01	2.98	6.2E-01	8.6E-01
CG Chimp-specific OLIG2 Hypo	Multiple_sclerosis	-0.03	-5.64	1.6E-01	6.2E-01
CG Chimp-specific OLIG2 Hypo	Primary_biliary_cirrhosis	0.01	1.23	9.0E-01	9.9E-01
CG Chimp-specific OLIG2 Hypo	Rheumatoid_Arthritis	0.00	-1.03	2.3E-01	6.9E-01
CG Chimp-specific OLIG2 Hypo	Triglycerides	0.01	3.09	1.6E-01	6.3E-01
CG Chimp-specific OLIG2 Hypo	Type_1_Diabetes	0.01	2.74	6.3E-01	8.6E-01
CG Chimp-specific OLIG2 Hypo	Type_2_Diabetes	0.01	1.44	8.5E-01	9.9E-01
CG Chimp-specific OLIG2 Hypo	Ulcerative_Colitis	0.00	0.61	7.6E-01	9.4E-01
CH NeuN Conserved Hyper	Alzheimer	0.71	1.02	8.3E-01	9.8E-01

Table A. 9 (continued)

CH NeuN Conserved Hyper	Anorexia	0.68	0.98	6.5E-01	8.6E-01
CH NeuN Conserved Hyper	Autism	0.75	1.08	7.9E-02	4.1E-01
CH NeuN Conserved Hyper	Bipolar_Disorder	0.48	0.69	1.9E-06	**6.3E-05**
CH NeuN Conserved Hyper	DS	0.70	1.00	9.9E-01	1.0E+00
CH NeuN Conserved Hyper	Neuroticism	0.69	0.99	7.5E-01	9.4E-01
CH NeuN Conserved Hyper	Schizophrenia	0.61	0.87	2.1E-07	**1.1E-05**
CH NeuN Conserved Hyper	SWB	0.63	0.90	2.2E-01	6.9E-01
CH NeuN Conserved Hyper	Years_of_Education1	0.61	0.87	8.4E-03	9.9E-02
CH NeuN Conserved Hyper	BMI1	0.71	1.01	6.2E-01	8.6E-01
CH NeuN Conserved Hyper	Celiac	0.75	1.08	2.6E-01	6.9E-01
CH NeuN Conserved Hyper	Coronary_Artery_Disease	0.61	0.87	8.1E-02	4.1E-01
CH NeuN Conserved Hyper	Crohns_Disease	0.75	1.07	1.3E-01	5.8E-01
CH NeuN Conserved Hyper	Ever_Smoked	0.65	0.94	2.8E-01	7.1E-01
CH NeuN Conserved Hyper	Fasting_Glucose	0.70	1.01	9.1E-01	9.9E-01
CH NeuN Conserved Hyper	HDL	0.70	1.01	8.6E-01	9.9E-01
CH NeuN Conserved Hyper	Height1	0.74	1.05	4.4E-02	3.0E-01
CH NeuN Conserved Hyper	IBD	0.73	1.04	3.3E-01	7.4E-01
CH NeuN Conserved Hyper	LDL	0.72	1.02	6.5E-01	8.6E-01
CH NeuN Conserved Hyper	Lupus	0.75	1.07	5.0E-01	8.2E-01
CH NeuN Conserved Hyper	Multiple_sclerosis	0.87	1.24	2.3E-01	6.9E-01

Table A. 9 (continued)

CH NeuN Conserved Hyper	Primary_biliary_cirrhosis	0.71	1.01	8.8E-01	9.9E-01
CH NeuN Conserved Hyper	Rheumatoid_Arthritis	0.76	1.09	2.1E-01	6.8E-01
CH NeuN Conserved Hyper	Triglycerides	0.66	0.95	2.5E-01	6.9E-01
CH NeuN Conserved Hyper	Type_1_Diabetes	0.80	1.14	4.1E-01	7.9E-01
CH NeuN Conserved Hyper	Type_2_Diabetes	0.74	1.05	4.4E-01	8.0E-01
CH NeuN Conserved Hyper	Ulcerative_Colitis	0.73	1.05	2.8E-01	7.2E-01
CH NeuN Human-specific Hypo	Alzheimer	0.07	2.94	7.0E-02	3.8E-01
CH NeuN Human-specific Hypo	Anorexia	0.05	1.82	1.7E-01	6.3E-01
CH NeuN Human-specific Hypo	Autism	0.04	1.58	4.4E-01	8.0E-01
CH NeuN Human-specific Hypo	Bipolar_Disorder	0.02	0.91	9.0E-01	9.9E-01
CH NeuN Human-specific Hypo	DS	0.03	1.08	9.0E-01	9.9E-01
CH NeuN Human-specific Hypo	Neuroticism	0.01	0.48	6.5E-02	3.7E-01
CH NeuN Human-specific Hypo	Schizophrenia	0.02	0.84	4.8E-01	8.1E-01
CH NeuN Human-specific Hypo	SWB	0.07	2.78	7.0E-02	3.8E-01
CH NeuN Human-specific Hypo	Years_of_Education1	0.01	0.35	1.8E-01	6.3E-01

Table A. 9 (continued)

CH NeuN Human-specific Hypo	BMI1	0.02	1.00	9.9E-01	1.0E+00
CH NeuN Human-specific Hypo	Celiac	0.01	0.54	6.3E-01	8.6E-01
CH NeuN Human-specific Hypo	Coronary_Artery_Disease	0.03	1.04	9.6E-01	1.0E+00
CH NeuN Human-specific Hypo	Crohns_Disease	0.04	1.61	1.9E-01	6.4E-01
CH NeuN Human-specific Hypo	Ever_Smoked	0.01	0.26	2.0E-01	6.6E-01
CH NeuN Human-specific Hypo	Fasting_Glucose	0.01	0.33	4.0E-01	7.9E-01
CH NeuN Human-specific Hypo	HDL	0.02	0.69	5.9E-01	8.6E-01
CH NeuN Human-specific Hypo	Height1	0.01	0.24	3.3E-04	**7.5E-03**
CH NeuN Human-specific Hypo	IBD	0.03	1.37	4.5E-01	8.0E-01
CH NeuN Human-specific Hypo	LDL	0.02	0.82	7.6E-01	9.4E-01
CH NeuN Human-specific Hypo	Lupus	0.01	0.48	4.4E-01	8.0E-01
CH NeuN Human-specific Hypo	Multiple_sclerosis	-0.03	-1.32	2.7E-01	7.0E-01
CH NeuN Human-specific Hypo	Primary_biliary_cirrhosis	0.04	1.47	5.7E-01	8.6E-01

Table A. 9 (continued)

CH NeuN Human-specific Hypo	Rheumatoid_Arthritis	0.03	1.06	9.4E-01	1.0E+00
CH NeuN Human-specific Hypo	Triglycerides	0.02	0.95	9.3E-01	1.0E+00
CH NeuN Human-specific Hypo	Type_1_Diabetes	0.00	-0.09	4.5E-01	8.0E-01
CH NeuN Human-specific Hypo	Type_2_Diabetes	0.04	1.56	5.0E-01	8.2E-01
CH NeuN Human-specific Hypo	Ulcerative_Colitis	0.01	0.52	4.0E-01	7.9E-01
CH NeuN Human-specific Hyper	Alzheimer	0.09	0.78	6.3E-01	8.6E-01
CH NeuN Human-specific Hyper	Anorexia	0.10	0.90	7.0E-01	9.1E-01
CH NeuN Human-specific Hyper	Autism	0.08	0.69	2.3E-01	6.9E-01
CH NeuN Human-specific Hyper	Bipolar_Disorder	0.08	0.71	3.1E-01	7.3E-01
CH NeuN Human-specific Hyper	DS	0.11	0.96	8.9E-01	9.9E-01
CH NeuN Human-specific Hyper	Neuroticism	0.08	0.72	6.4E-02	3.7E-01
CH NeuN Human-specific Hyper	Schizophrenia	0.06	0.55	1.9E-07	**1.1E-05**
CH NeuN Human-specific Hyper	SWB	0.09	0.78	5.7E-01	8.6E-01

Table A. 9 (continued)

CH NeuN Human-specific Hyper	Years_of_Education1	0.04	0.40	7.2E-03	9.2E-02
CH NeuN Human-specific Hyper	BMI1	0.11	0.99	9.5E-01	1.0E+00
CH NeuN Human-specific Hyper	Celiac	0.06	0.52	2.4E-01	6.9E-01
CH NeuN Human-specific Hyper	Coronary_Artery_Disease	0.03	0.29	1.9E-02	1.7E-01
CH NeuN Human-specific Hyper	Crohns_Disease	0.05	0.48	1.0E-02	1.1E-01
CH NeuN Human-specific Hyper	Ever_Smoked	0.08	0.69	2.6E-01	6.9E-01
CH NeuN Human-specific Hyper	Fasting_Glucose	0.05	0.42	7.2E-02	3.8E-01
CH NeuN Human-specific Hyper	HDL	0.06	0.58	6.1E-02	3.7E-01
CH NeuN Human-specific Hyper	Height1	0.08	0.74	3.3E-02	2.5E-01
CH NeuN Human-specific Hyper	IBD	0.10	0.85	4.4E-01	8.0E-01
CH NeuN Human-specific Hyper	LDL	0.07	0.61	1.1E-01	5.4E-01
CH NeuN Human-specific Hyper	Lupus	0.09	0.79	5.2E-01	8.3E-01
CH NeuN Human-specific Hyper	Multiple_sclerosis	0.12	1.07	9.4E-01	1.0E+00

Table A. 9 (continued)

CH NeuN Human-specific Hyper	Primary_biliary_cirrhosis	0.01	0.07	7.2E-03	9.2E-02
CH NeuN Human-specific Hyper	Rheumatoid_Arthritis	0.11	1.02	9.4E-01	1.0E+00
CH NeuN Human-specific Hyper	Triglycerides	0.09	0.80	2.6E-01	6.9E-01
CH NeuN Human-specific Hyper	Type_1_Diabetes	0.13	1.12	8.5E-01	9.9E-01
CH NeuN Human-specific Hyper	Type_2_Diabetes	0.07	0.63	2.6E-01	6.9E-01
CH NeuN Human-specific Hyper	Ulcerative_Colitis	0.09	0.81	5.1E-01	8.2E-01
CH NeuN Chimp-specific Hyper	Alzheimer	0.08	1.10	8.8E-01	9.9E-01
CH NeuN Chimp-specific Hyper	Anorexia	0.09	1.32	3.0E-01	7.3E-01
CH NeuN Chimp-specific Hyper	Autism	0.07	1.06	8.5E-01	9.9E-01
CH NeuN Chimp-specific Hyper	Bipolar_Disorder	0.02	0.32	5.1E-02	3.1E-01
CH NeuN Chimp-specific Hyper	DS	0.08	1.16	6.7E-01	8.9E-01
CH NeuN Chimp-specific Hyper	Neuroticism	0.07	1.05	8.0E-01	9.5E-01
CH NeuN Chimp-specific Hyper	Schizophrenia	0.06	0.83	1.3E-01	5.8E-01

Table A. 9 (continued)

CH NeuN Chimp-specific Hyper	SWB	0.06	0.86	8.0E-01	9.5E-01
CH NeuN Chimp-specific Hyper	Years_of_Education1	0.06	0.90	7.0E-01	9.1E-01
CH NeuN Chimp-specific Hyper	BMI1	0.08	1.20	4.3E-01	8.0E-01
CH NeuN Chimp-specific Hyper	Celiac	0.11	1.55	3.1E-01	7.3E-01
CH NeuN Chimp-specific Hyper	Coronary_Artery_Disease	0.08	1.13	7.8E-01	9.4E-01
CH NeuN Chimp-specific Hyper	Crohns_Disease	0.04	0.63	2.3E-01	6.9E-01
CH NeuN Chimp-specific Hyper	Ever_Smoked	0.08	1.21	5.9E-01	8.6E-01
CH NeuN Chimp-specific Hyper	Fasting_Glucose	0.03	0.37	1.9E-01	6.3E-01
CH NeuN Chimp-specific Hyper	HDL	0.07	1.09	7.8E-01	9.4E-01
CH NeuN Chimp-specific Hyper	Height1	0.09	1.27	1.4E-01	5.8E-01
CH NeuN Chimp-specific Hyper	IBD	0.05	0.70	3.0E-01	7.3E-01
CH NeuN Chimp-specific Hyper	LDL	0.11	1.59	2.1E-01	6.7E-01
CH NeuN Chimp-specific Hyper	Lupus	0.05	0.73	5.4E-01	8.3E-01

Table A. 9 (continued)

CH NeuN Chimp-specific Hyper	Multiple_sclerosis	0.13	1.92	4.4E-01	8.0E-01
CH NeuN Chimp-specific Hyper	Primary_biliary_cirrhosis	0.05	0.72	5.4E-01	8.3E-01
CH NeuN Chimp-specific Hyper	Rheumatoid_Arthritis	0.05	0.80	6.1E-01	8.6E-01
CH NeuN Chimp-specific Hyper	Triglycerides	0.06	0.89	6.8E-01	9.0E-01
CH NeuN Chimp-specific Hyper	Type_1_Diabetes	0.04	0.63	6.4E-01	8.6E-01
CH NeuN Chimp-specific Hyper	Type_2_Diabetes	0.08	1.21	6.4E-01	8.6E-01
CH NeuN Chimp-specific Hyper	Ulcerative_Colitis	0.06	0.81	6.0E-01	8.6E-01

Table A. 10. The list of GWAS traits and references used in LD score regression analyses.

Disease	**Reference**
BMI	Speliotes et al., 2010 Nat Genet
Height	Lango Allen et al., 2010 Nature
Anorexia	Boraska et al., 2014 Mol Psych
Bipolar Disorder	BIP Working Group of the PGC, 2011 Nat Genet
Coronary Artery Disease	Schunkert et al., 2011 Nat Genet
Crohn's Disease	Jostins et al., 2012 Nature
Ever Smoked	TAG Consortium, 2010 Nat Genet
Fasting Glucose	Manning et. al., 2012 Nat Genet
HDL	Teslovich et al., 2010 Nature
LDL	Teslovich et al., 2010 Nature
Rheumatoid Arthritis	Okada et al., 2014 Nature
Schizophrenia	SCZ Working Group of the PGC, 2014 Nature
Triglycerides	Teslovich et al., 2010 Nature
Type 2 Diabetes	Morris et al., 2012 Nat Genet
Ulcerative Colitis	Jostins et al., 2012 Nature
Years of Education	Rietveld et al., 2013 Science
IBD	Jostins et al., 2012 Nature
Autism Spectrum	PGC Cross-Disorder Group, 2013 Lancet
Alzheimer's Disease	Lambert et al., 2013 Nat Genet
Type 1 Diabetes	Bradfield et al., 2011 Plos Genet
Multiple Sclerosis	IMS Genetics Consortium, 2011 Nature
Celiac Disease	Dubois et al., 2010 Nat Genet
Primary Biliary Cirrhosis	Cordell et al., 2015 Nat Commun
Systemic Lupus Erythematosus	Bentham et al., 2015 Nat Genet
Subject well being	Okbay et al., 2016 Nat Genet
Depressive symptoms	Okbay et al., 2016 Nat Genet
Neuroticism	Okbay et al., 2016 Nat Genet

APPENDIX B. SUPPLEMENTARY MATERIAL FOR CHAPTER 3

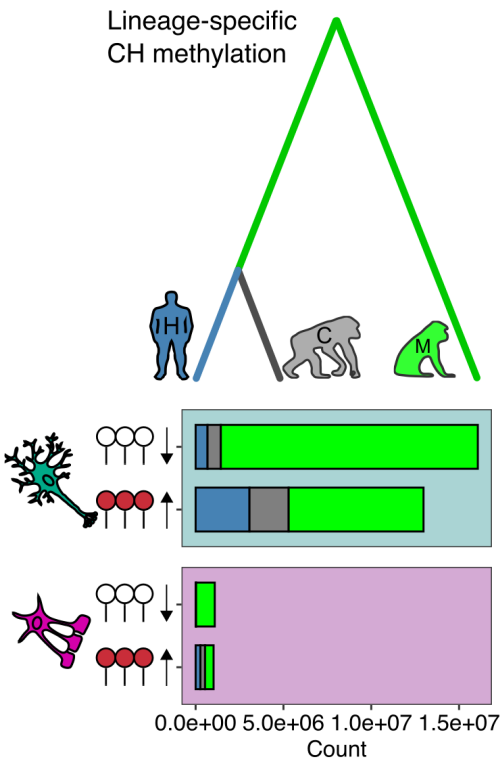

Figure B. 1. Evolutionary lineages and their divergence time among the three primate species. Human-specific methylation changes are depicted in blue and chimpanzee-specific methylation changes are shown in grey. Sites in which macaque show methylation divergence from human and chimpanzee but exhibit no difference between human and chimpanzee are denoted in green. Numbers of CH sites showing significant methylation changes on the three evolutionary lineages are plotted.

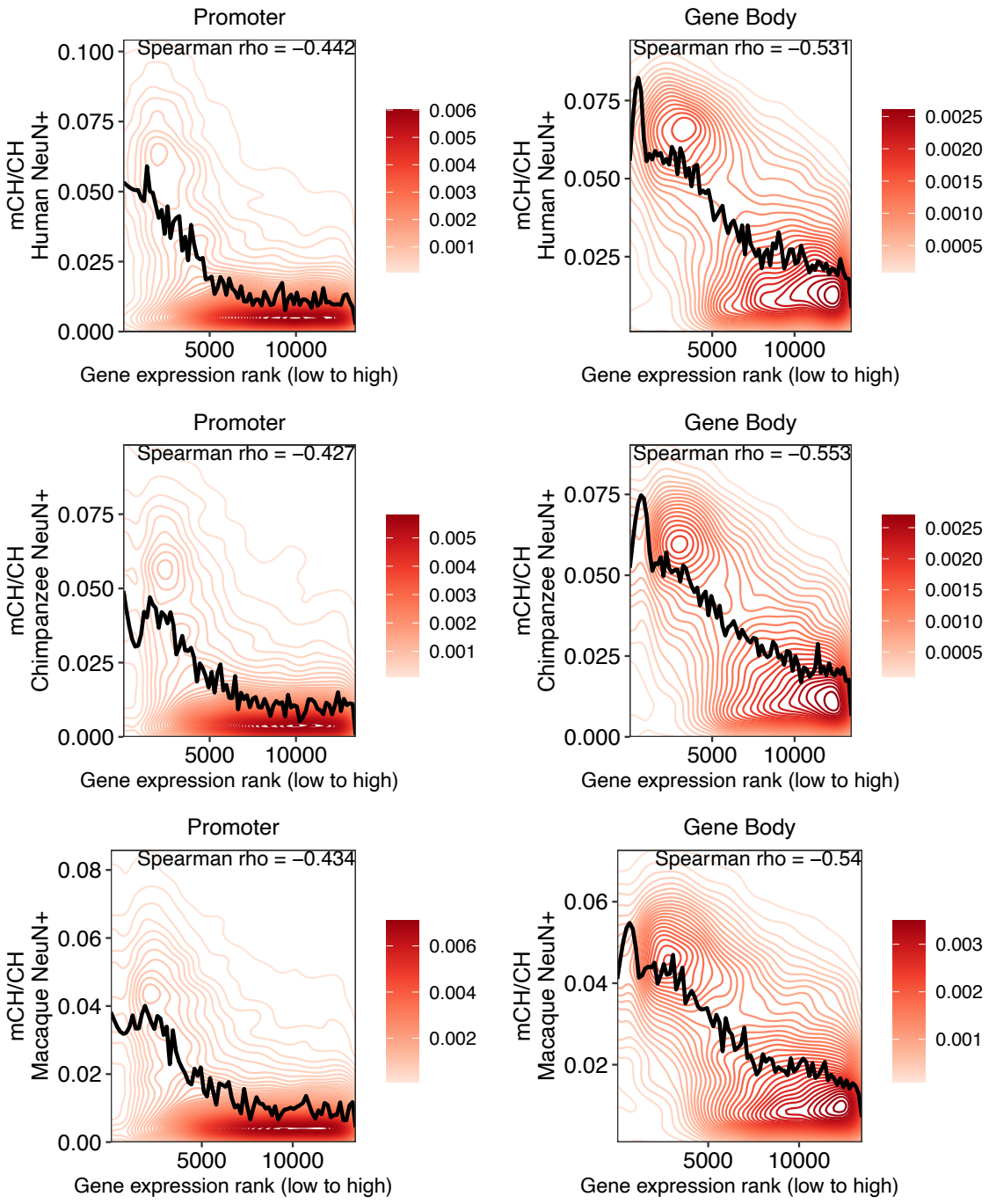

Figure B. 2. Correlation between gene expression and CH methylation in neuronal cells across all three primate species.

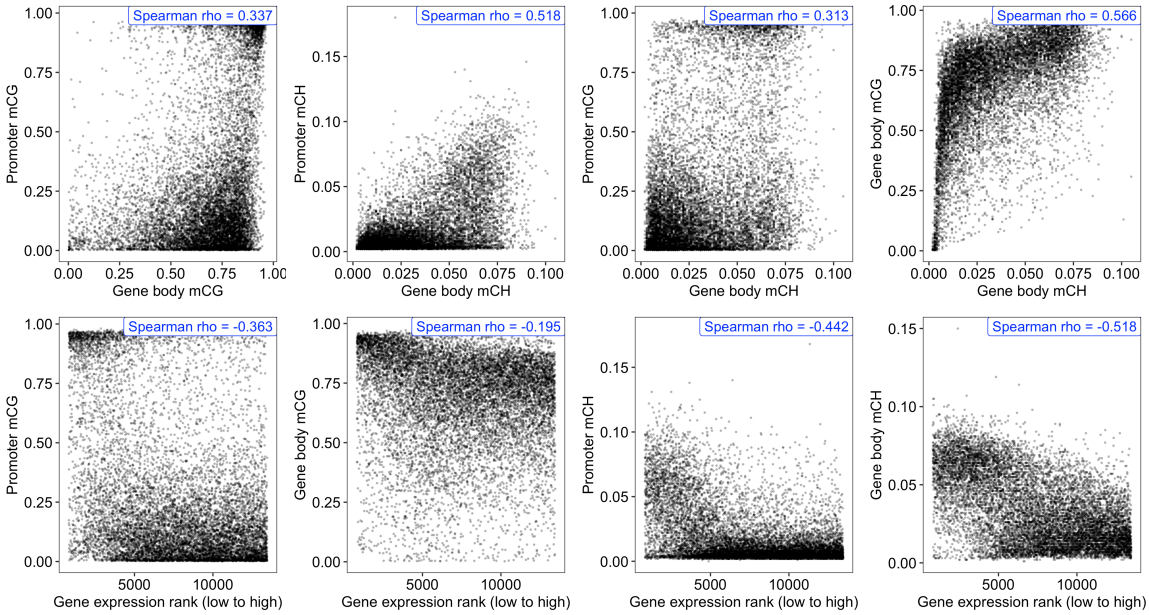

Figure B. 3. Correlation coefficient between different methylation contexts and between methylation and gene expression in human NeuN+.

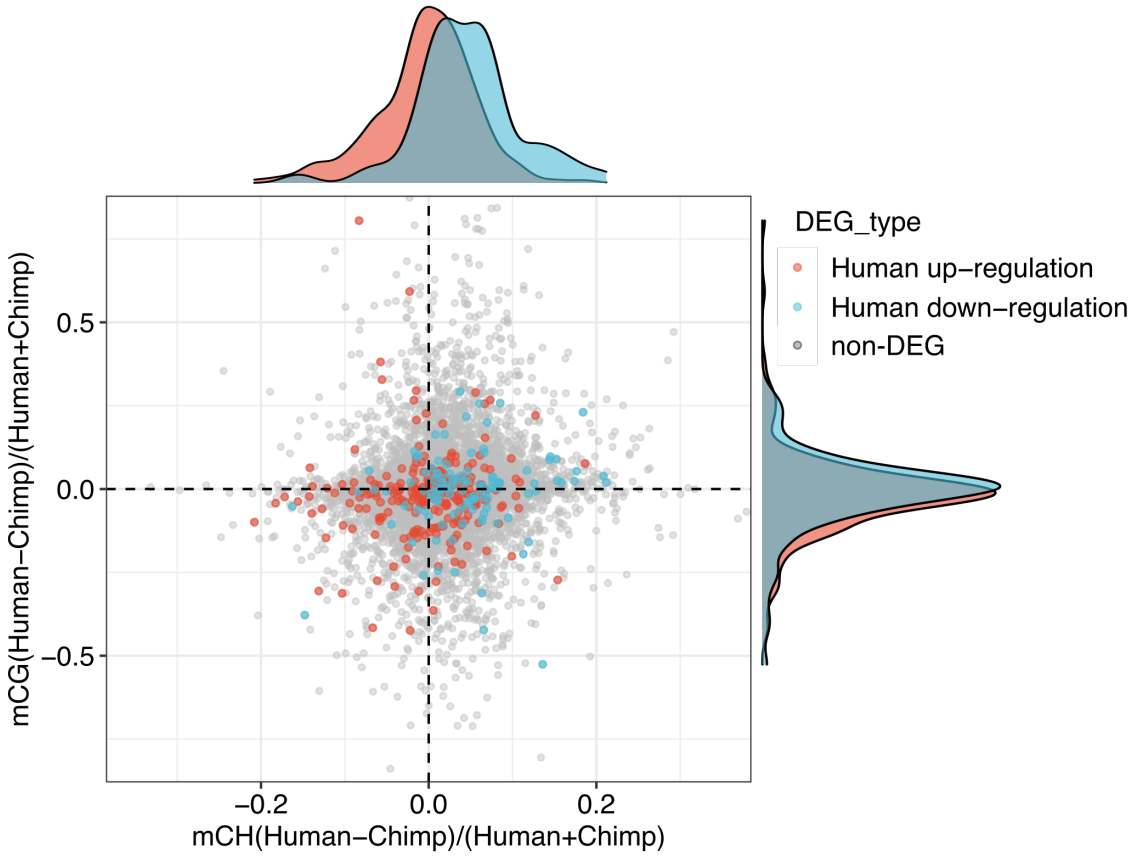

Figure B. 4. Relationship between CH gene body methylation (relative difference between species, X-axis) and CG promoter methylation (relative difference between species, Y-axis) in different gene types. DEG: differentially expressed gene list from Berto et al. 2019.

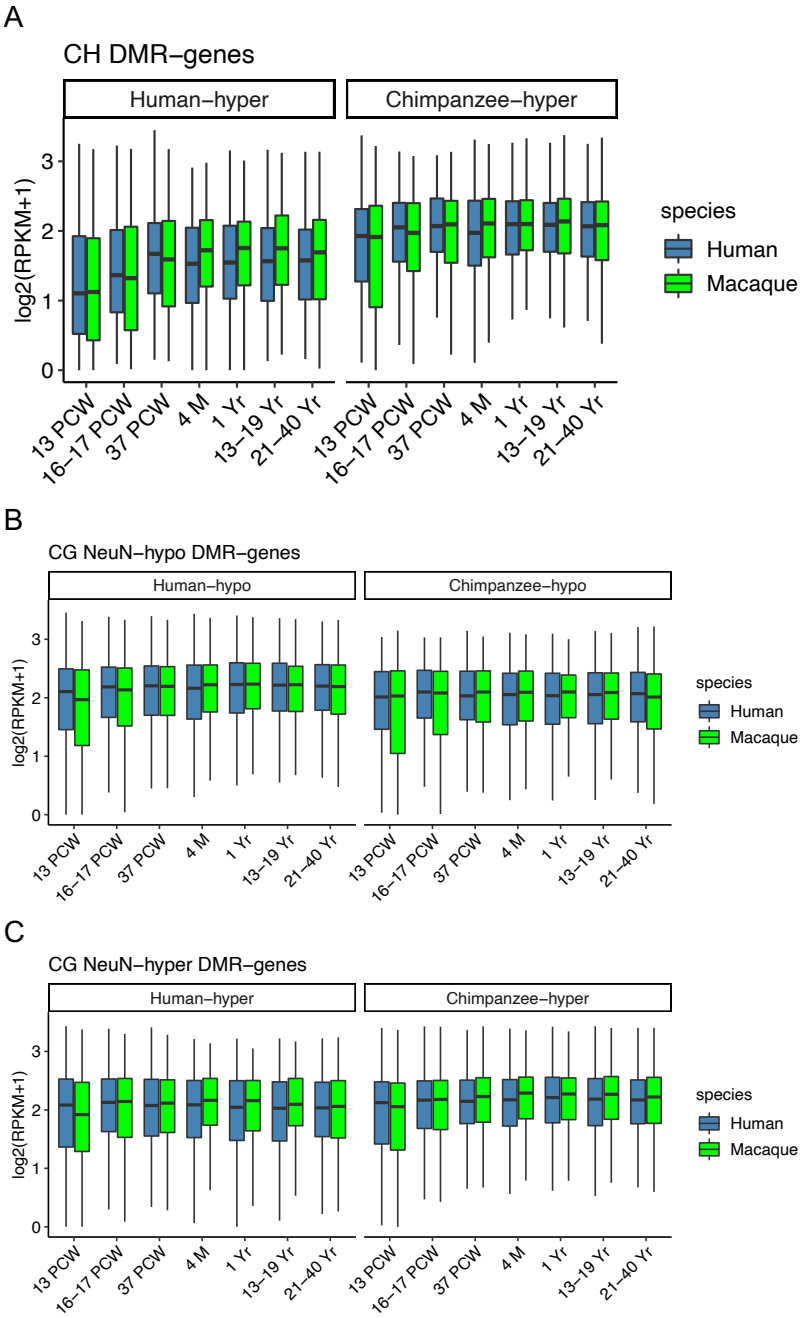

Figure B. 5. (A) Gene expression in human and macaque for CH DMR genes over developmental time points (Human CH DMR genes, n = 450 and Chimpanzee CH DMR genes, n =144). Macaque samples were age-matched to human developmental time points (PCW: Post conception week; M: Month; Yr: Year). Same as in panel A but (B) using CG DMR genes (hypo; Human DMR genes, n = 742 and Chimpanzee DMR genes, n = 201) and (C) using CG DMR genes (hyper; Human DMR genes, n = 223 and Chimpanzee DMR genes, n = 627). Box represents a range from the first quartile to the third quartile. The line in the box indicates the median value. The minima and maxima are within 1.5 times the distance between the first and third quartiles from box.

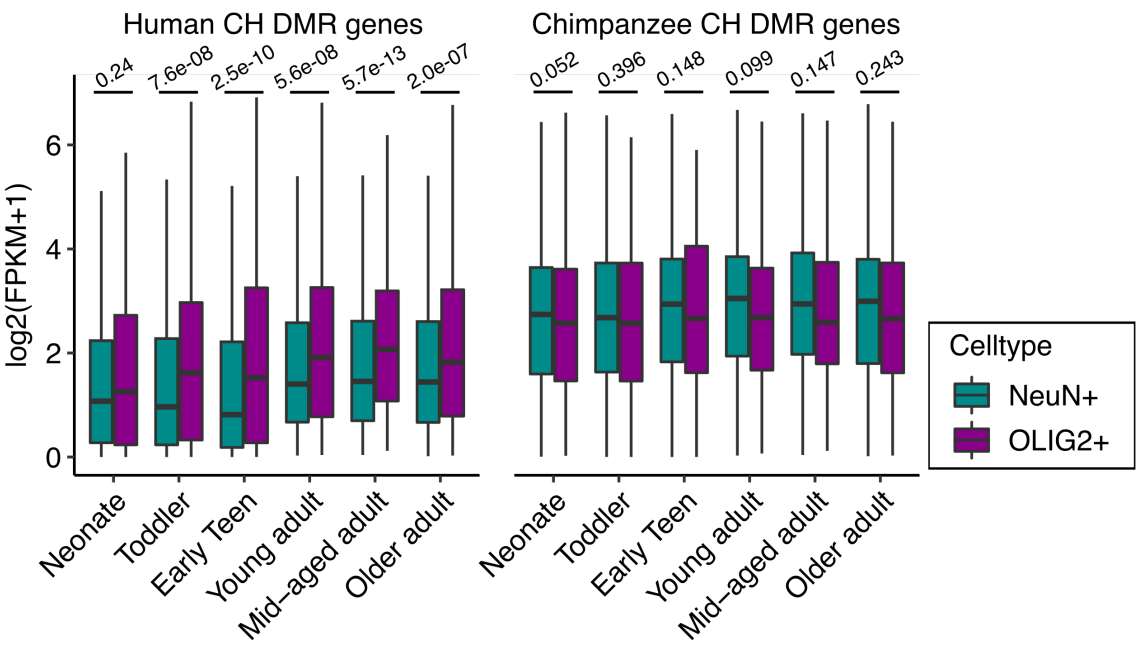

Figure B. 6. Gene expression of lineage-specific CH DMR genes (Human CH DMR genes, n = 590 and Chimpanzee CH DMR genes, n =159) in neuronal (NeuN+) cell samples and oligodendrocyte (OLIG2+) cell samples. Statistical significance was calculated using two-sided Mann-Whitney U-test. Box represents a range from the first quartile to the third quartile. The line in the box indicates the median value. The minima and maxima are within 1.5 times the distance between the first and third quartiles from box.

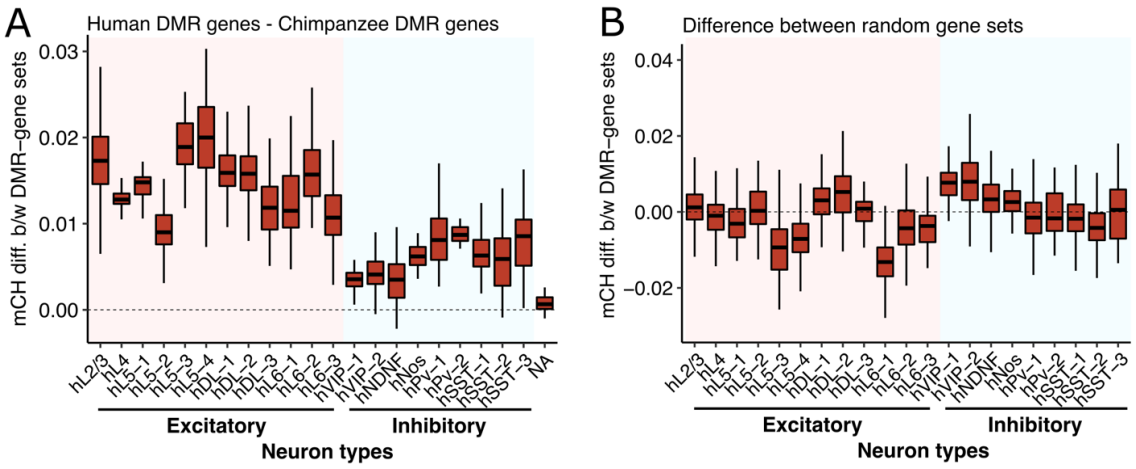

Figure B. 7. CH methylation of single cell methylomes. (A) Boxplots display average methylation differences between human CH DMR genes (n = 674) and Chimpanzee CH DMR genes (n = 173) of neuronal cells for neuronal subtypes. (B) As a control, we randomly selected the same number of genes. Box represents a range from the first quartile to the third quartile. The line in the box indicates the median value. The minima and maxima are within 1.5 times the distance between the first and third quartiles from box.

Table B. 1. Partial correlation analysis explaining correlation coefficient between methylation and expression account for effects from other methylation contexts

Predictors	Gene expression	
	Normal[1] (P-value)	**Partial[2] (P-value)**
mCG promoter	-0.36 ($P < 10^{-10}$)	-0.27 ($P < 10^{-10}$)
mCG gene body	-0.22 ($P < 10^{-10}$)	0.18 ($P < 10^{-10}$)
mCH gene body	-0.53 ($P < 10^{-10}$)	-0.48 ($P < 10^{-10}$)

[1]ordinary correlation coefficient (spearman) [2]partial correlation coefficient (spearman)

Table B. 2. Multiple linear regression models explaining variation of gene expression levels of humans and human-chimpanzee difference.

Predictors	**Estimate of *b***	**t-value**	**Significance**	
Human data alone				
Intercept	3.05	67.15	$< 10^{-15}$	
Promoter mCG	-1.46	-33.28	$< 10^{-15}$	
Gene body mCG	2.2	29.38	$< 10^{-15}$	
Gene body mCH	-44.14	-65.74	$< 10^{-5}$	
Adj-R^2				0.39
Human-chimpanzee difference				
Intercept	0.005	1.83	0.06	
Promoter mCG difference	***-0.17***	***-3.19***	***0.001***	
Gene body mCG difference	0.21	1.91	0.06	
Gene body mCH difference	***-11.72***	***-19.32***	***$< 10^{-10}$***	
Adj-R^2				0.14

APPENDIX C. SUPPLEMENTARY MATERIAL FOR CHAPTER 4

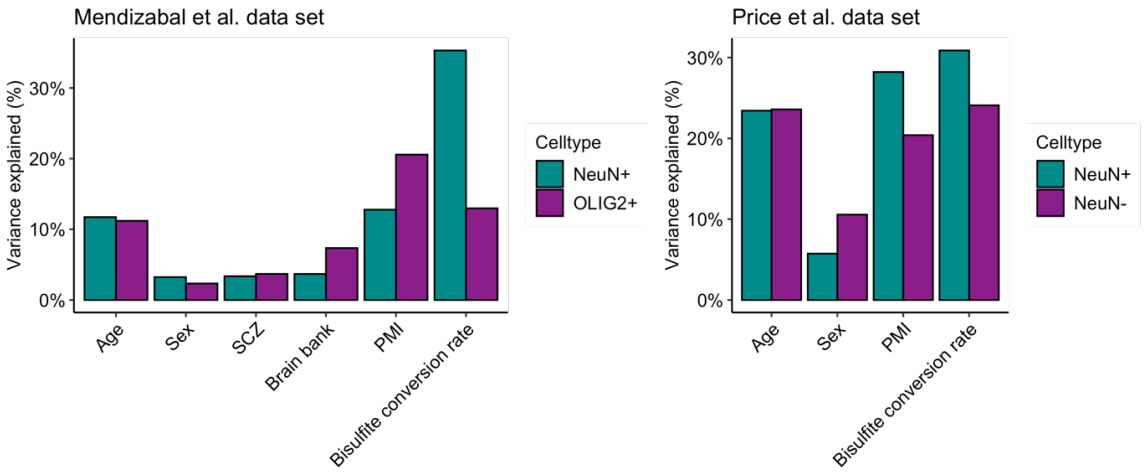

Figure C. 1. Proportion of variation explained by each variable controlled across 10 principal components (SCZ, schizophrenia; PMI, post-mortem interval). See also Figure 1B for results from the combined data sets.

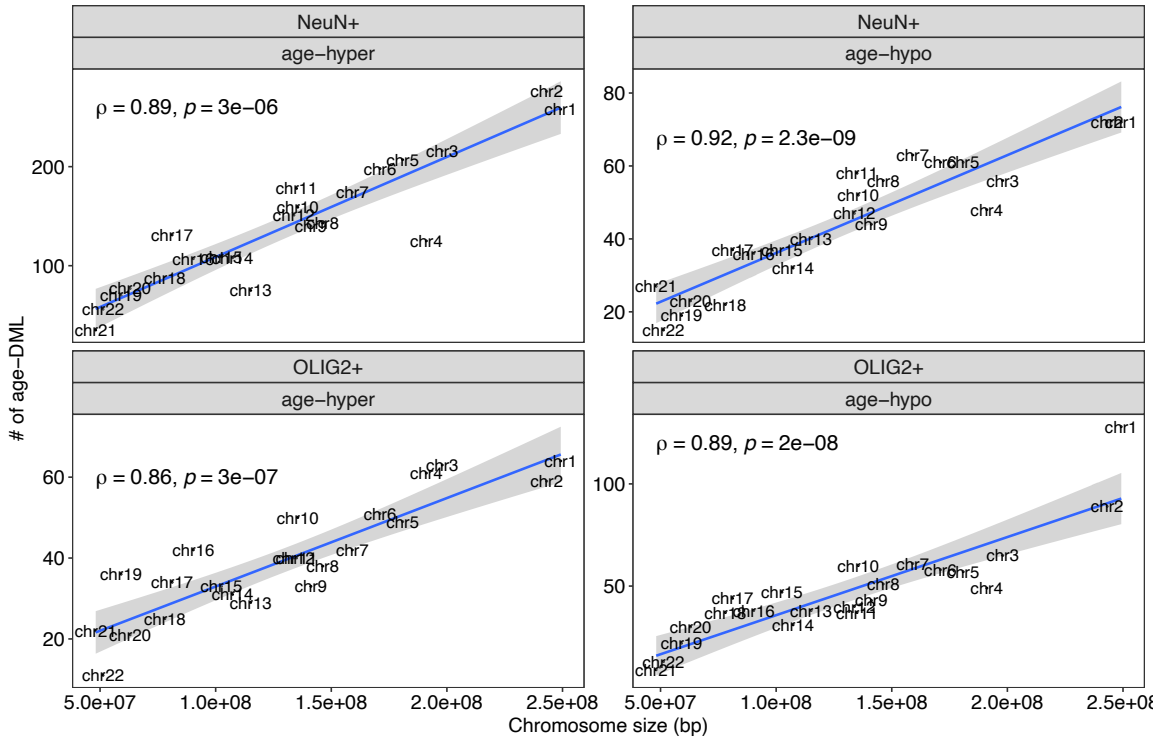

Figure C. 2. The number of age-DMLs occurred in each chromosome. Correlation coefficient is computed between the number of DMLs and the chromosome length.

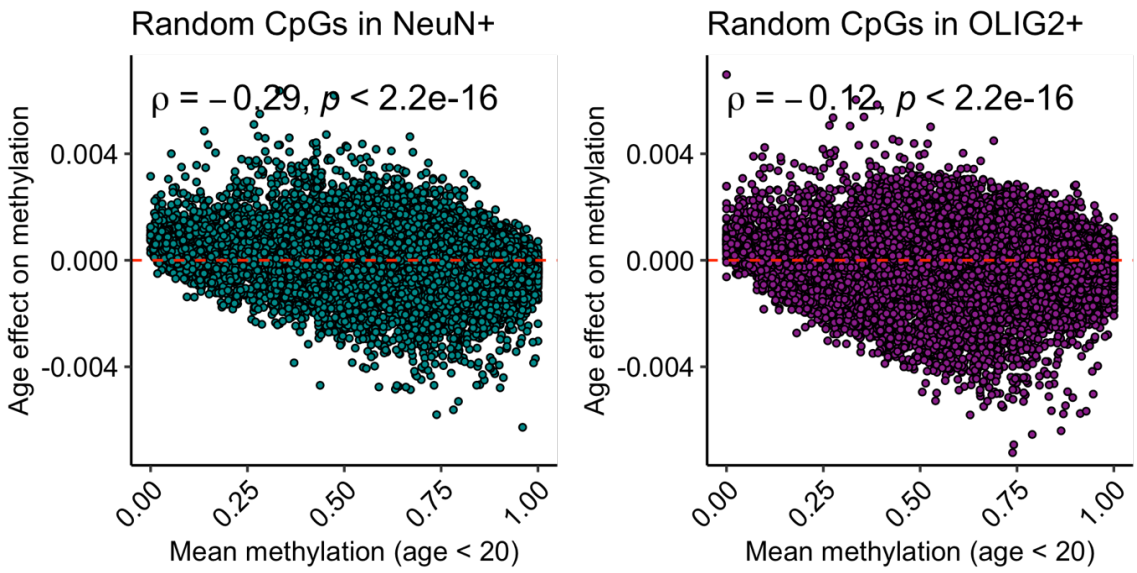

Figure C. 3. A relationship between mean methylation and the effect of age on methylation. For each CpG site, we fitted a linear model to estimate the age effect on DNA methylation adjusted for other biological variables (post-mortem interval, sex, disease status, and bisulfite conversion rate). Y-axis indicates mean methylation levels of samples with age < 20 for the corresponding CpG site. For the illustration purpose, only 10,000 randomly selected CpG sites were displayed in the plots.

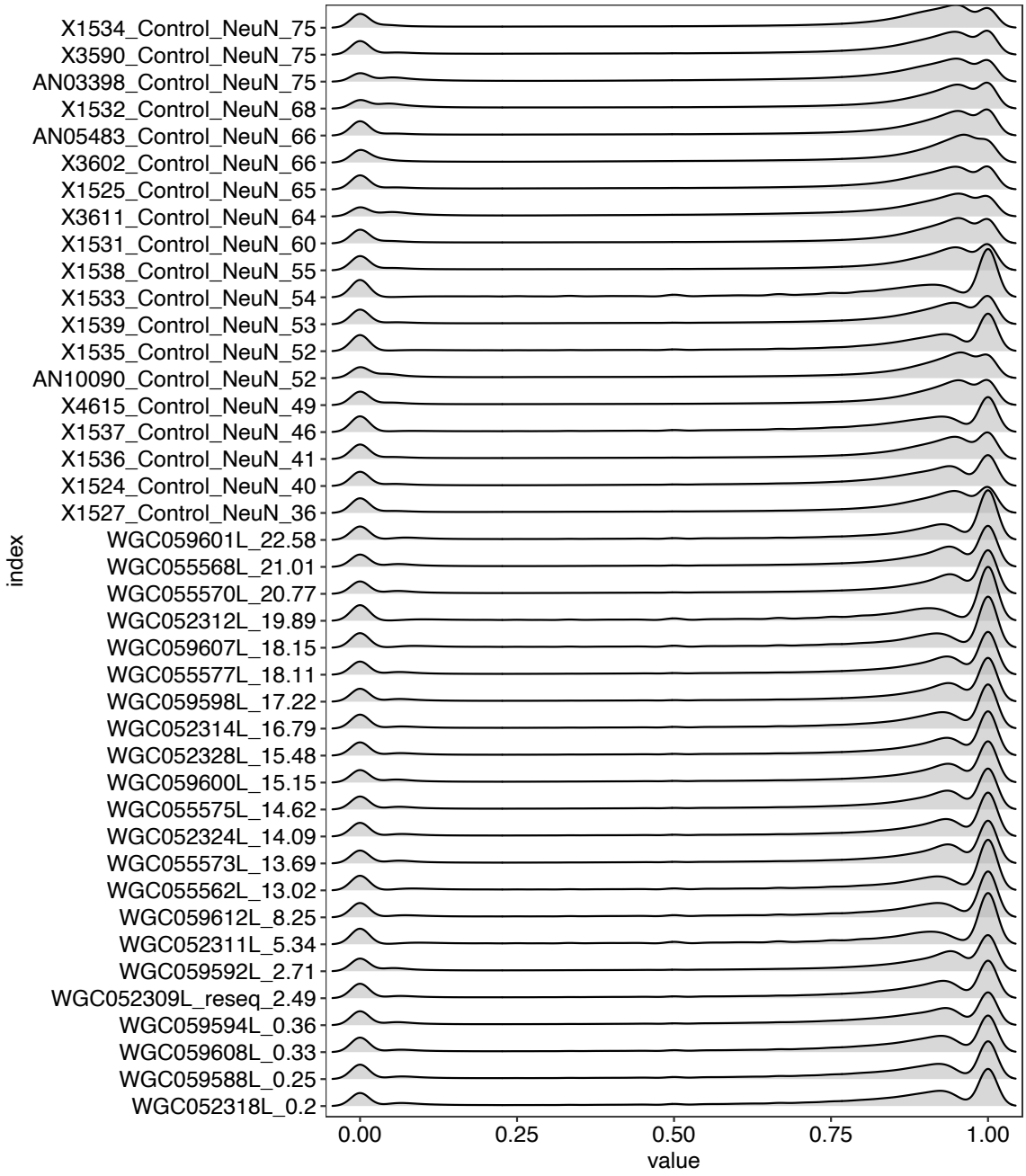

Figure C. 4. Distributions of whole-genome CpG methylation across age groups. NeuN+ samples were ordered by age (from bottom to top). For the illustration purpose, CpG sites with fractional methylation greater than 0.9 or less than 0.1 for more than 90% of samples were excluded. Sample index is denoted as 'sample name_age'

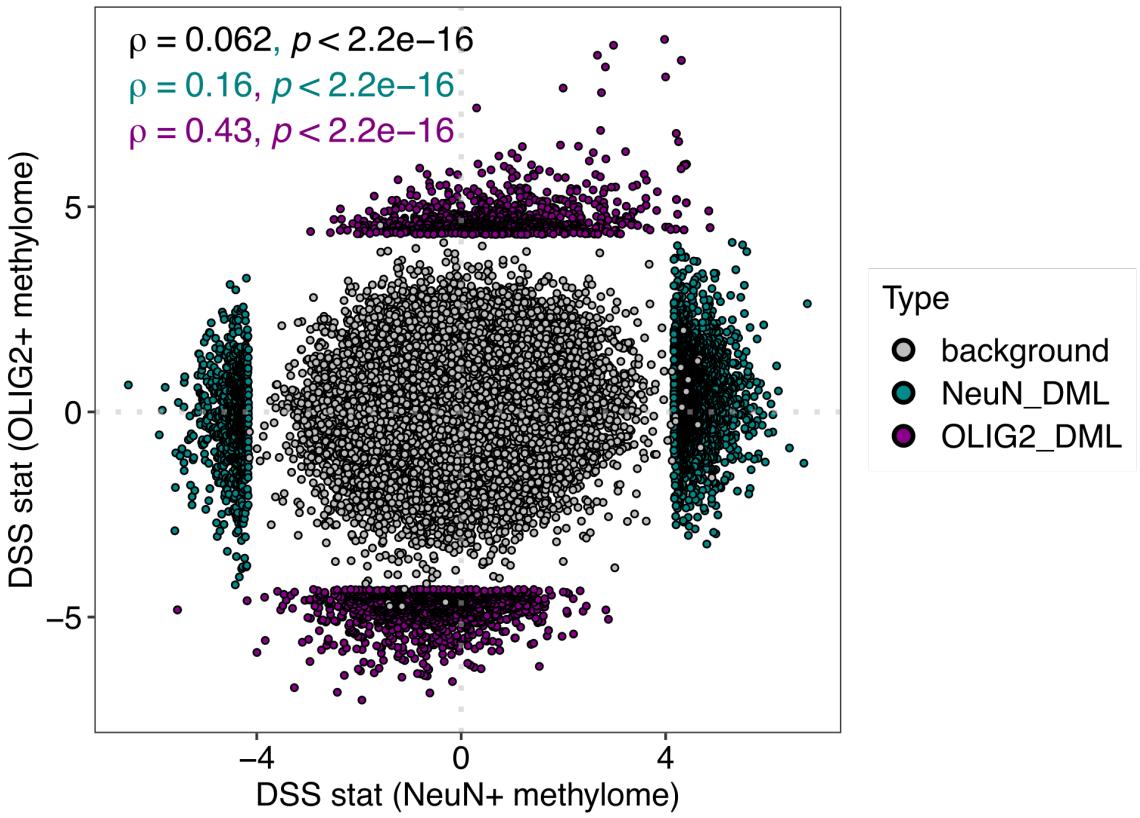

Figure C. 5. Comparison of DSS statistics of age effect between the cell types. NeuN+ age-DMLs and OLIG2+ age-DMLs are colored in cyan and magenta, respectively. Results from randomly selected 10,000 CpGs were also displayed (colored in grey).

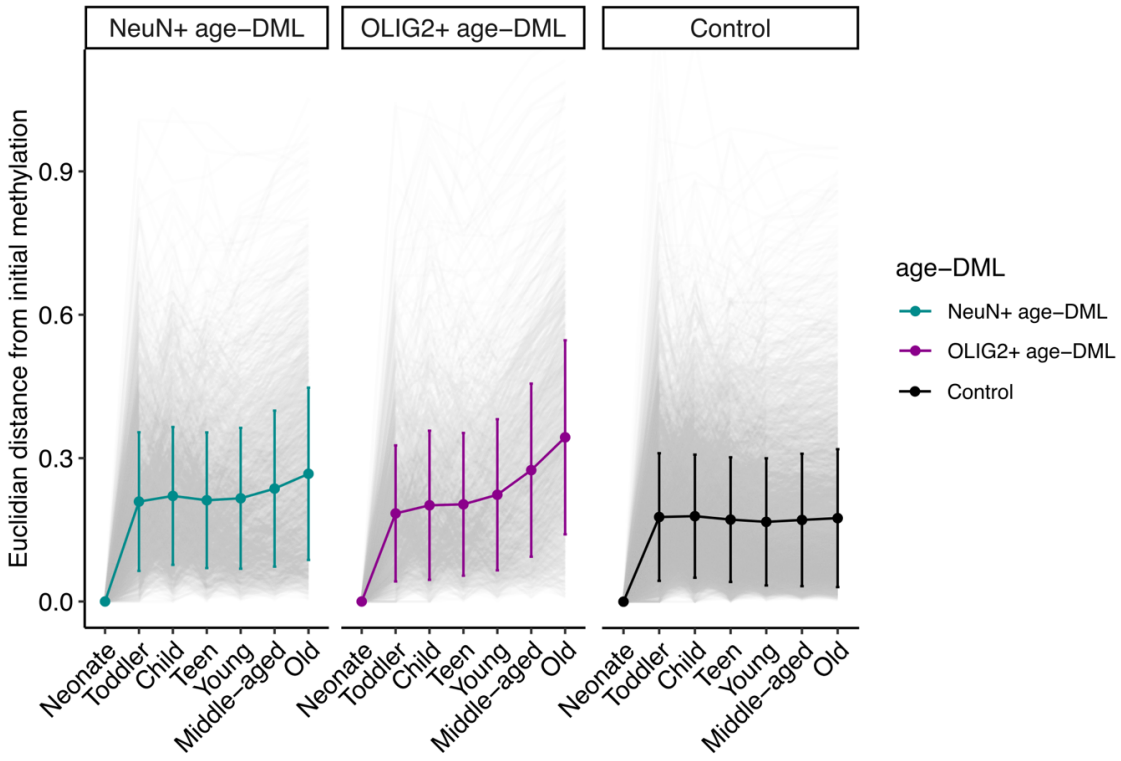

Figure C. 6. Dysregulation of cell identity with age measured by DNA methylation changes in neurons and oligodendrocytes in each age group from initial cell type DNA methylation. Epigenetic distance is calculated by Euclidian distance between points in a two-dimensional plane (DNA methylation of two cell types). Epigenetic distances of cell type identity from the neonate is depicted for different age groups for cell-type specific age-DMLs. Epigenetic distances of the randomly selected CpG positions are also shown for control.

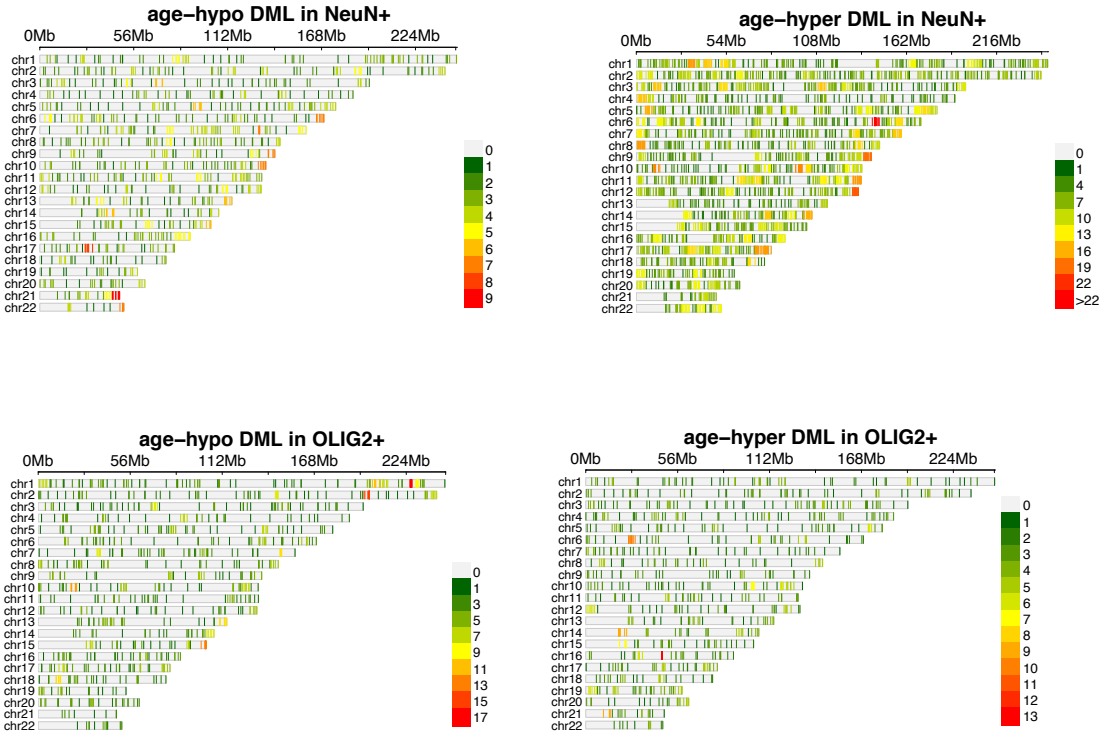

Figure C. 7. Distributions of age-DMLs across the genome. Each vertical line indicates age-DML and is colored based on the occurrence of age-DML in the 5Mbp bin window.

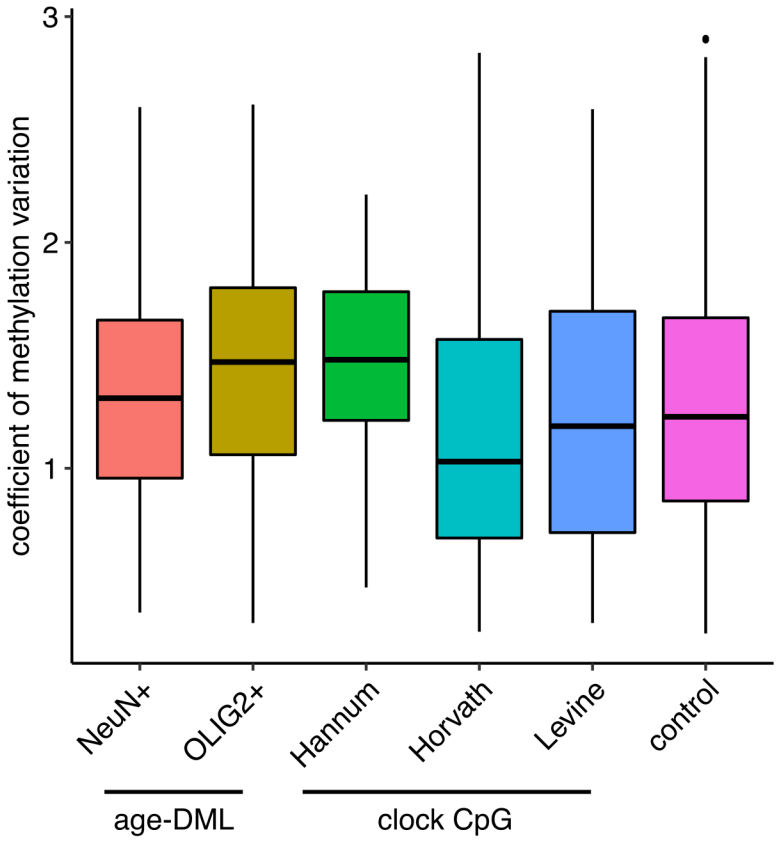

Figure C. 8. DNA methylation variation of multiple tissues in clock CpGs. Coefficient of methylation variation resulting from 10 different WGBS tissues shows highly variable DNA methylation for age-DML and reduced DNA methylation variation for clock CpGs.

Table C. 1. List of WGBS samples

Sample index	Disease status	Cell type	Sex	Age	Pmi	BS Conv Rate	Data
Miami0001_Control_NeuN	Control	NeuN+	M	25	16.3	0.942	Mendizabal
AN15240_Control_NeuN	Control	NeuN+	F	36	18.08	0.971	Mendizabal
X1507_Schizo_NeuN	Schizo	NeuN+	M	32	27	0.991	Mendizabal
AN09634_Schizo_NeuN	Schizo	NeuN+	M	26	16	0.997	Mendizabal
X1524_Control_NeuN	Control	NeuN+	M	40	10	0.997	Mendizabal
X1527_Control_NeuN	Control	NeuN+	M	36	23	0.998	Mendizabal
X4730_Schizo_NeuN	Schizo	NeuN+	F	32	12.3	0.998	Mendizabal
X1514_Schizo_NeuN	Schizo	NeuN+	M	40	9	0.998	Mendizabal
X1511_Schizo_NeuN	Schizo	NeuN+	M	26	24	0.998	Mendizabal
X1510_Schizo_NeuN	Schizo	NeuN+	M	33	20	0.998	Mendizabal
AN16799_Control_NeuN	Control	NeuN+	M	43	14.68	0.935	Mendizabal
X4504_Schizo_NeuN	Schizo	NeuN+	M	55	10.7	0.971	Mendizabal
X1541_Control_NeuN	Control	NeuN+	F	49	15.3	0.979	Mendizabal
AN10090_Control_NeuN	Control	NeuN+	M	52	13.2	0.992	Mendizabal

Table C. 1 (continued)

X1516_Schizo_NeuN	Schizo	NeuN+	M	47	25	0.996	Mendizabal
X4448_Schizo_NeuN	Schizo	NeuN+	M	46	21.7	0.996	Mendizabal
AN17799_Schizo_NeuN	Schizo	NeuN+	F	56	10.5	0.996	Mendizabal
X1531_Control_NeuN	Control	NeuN+	M	60	17.1	0.996	Mendizabal
X1505_Schizo_NeuN	Schizo	NeuN+	F	46	23	0.996	Mendizabal
X4615_Control_NeuN	Control	NeuN+	M	49	15	0.997	Mendizabal
X1538_Control_NeuN	Control	NeuN+	F	55	25	0.997	Mendizabal
X1521_Schizo_NeuN	Schizo	NeuN+	M	60	22.5	0.997	Mendizabal
X1518_Schizo_NeuN	Schizo	NeuN+	F	52	27.5	0.997	Mendizabal
X1512_Schizo_NeuN	Schizo	NeuN+	F	43	11	0.997	Mendizabal
X1539_Control_NeuN	Control	NeuN+	F	53	23	0.997	Mendizabal
X1523_Schizo_NeuN	Schizo	NeuN+	F	42	15.5	0.998	Mendizabal
X4395_Schizo_NeuN	Schizo	NeuN+	M	56	14.7	0.998	Mendizabal
X1536_Control_NeuN	Control	NeuN+	M	41	24	0.998	Mendizabal
X1515_Schizo_NeuN	Schizo	NeuN+	F	50	27.15	0.999	Mendizabal

Table C. 1 (continued)

X1533_Control_NeuN	Control	NeuN+	M	54	19.3	0.999	Mendizabal
X1519_Schizo_NeuN	Schizo	NeuN+	M	54	15.5	0.999	Mendizabal
X1535_Control_NeuN	Control	NeuN+	M	52	15.4	0.999	Mendizabal
X1537_Control_NeuN	Control	NeuN+	M	46	22	0.999	Mendizabal
X1509_Schizo_NeuN	Schizo	NeuN+	M	52	21	0.999	Mendizabal
AN03398_Control_NeuN	Control	NeuN+	F	75	12.1	0.990	Mendizabal
X3545_Control_NeuN	Control	NeuN+	M	80	14	0.945	Mendizabal
X4336_Schizo_NeuN	Schizo	NeuN+	F	85	11.5	0.976	Mendizabal
X3586_Control_NeuN	Control	NeuN+	M	76	16	0.979	Mendizabal
X1532_Control_NeuN	Control	NeuN+	F	68	19	0.983	Mendizabal
X3611_Control_NeuN	Control	NeuN+	M	64	17.5	0.985	Mendizabal
X4804_Schizo_NeuN	Schizo	NeuN+	M	61	28	0.995	Mendizabal
X3590_Control_NeuN	Control	NeuN+	M	75	11.5	0.996	Mendizabal
X4361_Schizo_NeuN	Schizo	NeuN+	F	77	14.7	0.997	Mendizabal
AN18099_Schizo_NeuN	Schizo	NeuN+	M	66	16.47	0.997	Mendizabal

Table C. 1 (continued)

X1522_Schizo_NeuN	Schizo	NeuN+	M	63	17	0.997	Mendizabal
X1534_Control_NeuN	Control	NeuN+	M	75	7.5	0.997	Mendizabal
X1525_Control_NeuN	Control	NeuN+	F	65	11	0.997	Mendizabal
X3602_Control_NeuN	Control	NeuN+	M	66	13.2	0.998	Mendizabal
X1513_Schizo_NeuN	Schizo	NeuN+	F	73	14	0.998	Mendizabal
X1517_Schizo_NeuN	Schizo	NeuN+	M	77	32	0.998	Mendizabal
X1506_Schizo_NeuN	Schizo	NeuN+	F	85	9	0.998	Mendizabal
X1508_Schizo_NeuN	Schizo	NeuN+	F	73	36	0.998	Mendizabal
AN05483_Control_NeuN	Control	NeuN+	M	66	16.97	0.999	Mendizabal
X1520_Schizo_Olig2	Schizo	OLIG2+	M	38	14	0.984	Mendizabal
Miami0001_Control_Olig2	Control	OLIG2+	M	25	16.3	0.987	Mendizabal
X1507_Schizo_Olig2	Schizo	OLIG2+	M	32	27	0.992	Mendizabal
AN15240_Control_Olig2	Control	OLIG2+	F	36	18.08	0.996	Mendizabal
X1514_Schizo_Olig2	Schizo	OLIG2+	M	40	9	0.997	Mendizabal
AN09634_Schizo_Olig2	Schizo	OLIG2+	M	26	16	0.997	Mendizabal

Table C. 1 (continued)

X1510_Schizo_Olig2	Schizo	OLIG2+	M	33	20	0.997	Mendizabal
X1511_Schizo_Olig2	Schizo	OLIG2+	M	26	24	0.997	Mendizabal
X1527_Control_Olig2	Control	OLIG2+	M	36	23	0.998	Mendizabal
X1524_Control_Olig2	Control	OLIG2+	M	40	10	0.998	Mendizabal
X4730_Schizo_Olig2	Schizo	OLIG2+	F	32	12.3	0.998	Mendizabal
X4448_Schizo_Olig2	Schizo	OLIG2+	M	46	21.7	0.978	Mendizabal
X1536_Control_Olig2	Control	OLIG2+	M	41	24	0.992	Mendizabal
X1505_Schizo_Olig2	Schizo	OLIG2+	F	46	23	0.995	Mendizabal
AN16799_Control_Olig2	Control	OLIG2+	M	43	14.68	0.997	Mendizabal
X1515_Schizo_Olig2	Schizo	OLIG2+	F	50	27.15	0.997	Mendizabal
X1518_Schizo_Olig2	Schizo	OLIG2+	F	52	27.5	0.997	Mendizabal
X1541_Control_Olig2	Control	OLIG2+	F	49	15.3	0.997	Mendizabal
X1512_Schizo_Olig2	Schizo	OLIG2+	F	43	11	0.997	Mendizabal
X1538_Control_Olig2	Control	OLIG2+	F	55	25	0.997	Mendizabal
X1539_Control_Olig2	Control	OLIG2+	F	53	23	0.997	Mendizabal

Table C. 1 (continued)

X4395_Schizo_Olig2	Schizo	OLIG2+	M	56	14.7	0.998	Mendizabal
X4504_Schizo_Olig2	Schizo	OLIG2+	M	55	10.7	0.998	Mendizabal
AN10090_Control_Olig2	Control	OLIG2+	M	52	13.2	0.998	Mendizabal
X1523_Schizo_Olig2	Schizo	OLIG2+	F	42	15.5	0.998	Mendizabal
X4615_Control_Olig2	Control	OLIG2+	M	49	15	0.998	Mendizabal
AN17799_Schizo_Olig2	Schizo	OLIG2+	F	56	10.5	0.998	Mendizabal
X3590_Control_Olig2	Control	OLIG2+	M	75	11.5	0.916	Mendizabal
X3611_Control_Olig2	Control	OLIG2+	M	64	17.5	0.947	Mendizabal
X1525_Control_Olig2	Control	OLIG2+	F	65	11	0.980	Mendizabal
X3545_Control_Olig2	Control	OLIG2+	M	80	14	0.988	Mendizabal
X4361_Schizo_Olig2	Schizo	OLIG2+	F	77	14.7	0.994	Mendizabal
X4336_Schizo_Olig2	Schizo	OLIG2+	F	85	11.5	0.995	Mendizabal
X3586_Control_Olig2	Control	OLIG2+	M	76	16	0.996	Mendizabal
X1513_Schizo_Olig2	Schizo	OLIG2+	F	73	14	0.996	Mendizabal
AN18099_Schizo_Olig2	Schizo	OLIG2+	M	66	16.47	0.996	Mendizabal

Table C. 1 (continued)

X4804_Schizo_Olig2	Schizo	OLIG2+	M	61	28	0.997	Mendizabal
X3602_Control_Olig2	Control	OLIG2+	M	66	13.2	0.997	Mendizabal
AN03398_Control_Olig2	Control	OLIG2+	F	75	12.1	0.997	Mendizabal
X1508_Schizo_Olig2	Schizo	OLIG2+	F	73	36	0.998	Mendizabal
AN05483_Control_Olig2	Control	OLIG2+	M	66	16.97	0.998	Mendizabal
X1532_Control_Olig2	Control	OLIG2+	F	68	19	0.998	Mendizabal
WGC052318L	Control	NeuN+	M	0.2	36	0.982	Price
WGC059588L	Control	NeuN+	M	0.25	36.5	0.983	Price
WGC059608L	Control	NeuN+	M	0.33	22	0.987	Price
WGC059594L	Control	NeuN+	M	0.36	16	0.986	Price
WGC052309L_reseq	Control	NeuN+	F	2.49	8	0.991	Price
WGC059592L	Control	NeuN+	F	2.71	44	0.986	Price
WGC052311L	Control	NeuN+	M	5.34	18	0.981	Price
WGC059612L	Control	NeuN+	M	8.25	30	0.986	Price
WGC055562L	Control	NeuN+	F	13.02	26	0.986	Price

Table C. 1 (continued)

WGC055573L	Control	NeuN+	M	13.69	18	0.985	Price
WGC052324L	Control	NeuN+	M	14.09	16	0.989	Price
WGC055575L	Control	NeuN+	M	14.62	10	0.985	Price
WGC059600L	Control	NeuN+	M	15.15	20	0.983	Price
WGC052328L	Control	NeuN+	M	15.48	14.5	0.989	Price
WGC052314L	Control	NeuN+	M	16.79	25	0.992	Price
WGC059598L	Control	NeuN+	F	17.22	19	0.986	Price
WGC059603L	Control	NeuN+	M	18.05	14.5	0.980	Price
WGC055577L	Control	NeuN+	M	18.11	6	0.985	Price
WGC059607L	Control	NeuN+	M	18.15	36.5	0.985	Price
WGC052312L	Control	NeuN+	M	19.89	21.5	0.982	Price
WGC055570L	Control	NeuN+	M	20.77	25.5	0.985	Price
WGC055568L	Control	NeuN+	M	21.01	28.5	0.990	Price
WGC059601L	Control	NeuN+	M	22.58	54.5	0.983	Price

Table C. 1 (continued)

WGC059604L	Control	NeuN+	M	22.71	38.5	0.979	Price
WGC052316L	Control	NeuN-	M	0.2	36	0.991	Price
WGC059613L	Control	NeuN-	F	2.71	44	0.983	Price
WGC059614L	Control	NeuN-	M	5.34	18	0.980	Price
WGC052317L	Control	NeuN-	M	8.25	30	0.991	Price
WGC055558L	Control	NeuN-	F	13.02	26	0.984	Price
WGC055559L	Control	NeuN-	F	17.22	19	0.984	Price
WGC059596L	Control	NeuN-	M	19.89	21.5	0.986	Price
WGC055561L	Control	NeuN-	M	22.71	38.5	0.986	Price

REFERENCES

Anderson, J. A., R. A. Johnston, A. J. Lea, F. A. Campos, T. N. Voyles, M. Y. Akinyi, S. C. Alberts, E. A. Archie and J. Tung (2021). "High social status males experience accelerated epigenetic aging in wild baboons." Elife **10**: e66128.

Aran, D., S. Sabato and A. Hellman (2013). "DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes." Genome biology **14**(3): 1-14.

Babbitt, C. C., O. Fedrigo, A. D. Pfefferle, A. P. Boyle, J. E. Horvath, T. S. Furey and G. A. Wray (2010). "Both Noncoding and Protein-Coding RNAs Contribute to Gene Expression Evolution in the Primate Brain." Genome Biology and Evolution **2**: 67-79.

Bailey, T. L., M. Boden, F. A. Buske, M. Frith, C. E. Grant, L. Clementi, J. Ren, W. W. Li and W. S. Noble (2009). "MEME Suite: tools for motif discovery and searching." Nucleic Acids Research **37**(suppl 2): W202-W208.

Bell, C. G., R. Lowe, P. D. Adams, A. A. Baccarelli, S. Beck, J. T. Bell, B. C. Christensen, V. N. Gladyshev, B. T. Heijmans, S. Horvath, T. Ideker, J.-P. J. Issa, K. T. Kelsey, R. E. Marioni, W. Reik, C. L. Relton, L. C. Schalkwyk, A. E. Teschendorff, W. Wagner, K. Zhang and V. K. Rakyan (2019). "DNA methylation aging clocks: challenges and recommendations." Genome Biology **20**(1): 249.

Bell, J. T., P. C. Tsai, T. P. Yang, R. Pidsley, J. Nisbet, D. Glass, M. Mangino, G. Zhai, F. Zhang, A. Valdes, S. Y. Shin, E. L. Dempster, R. M. Murray, E. Grundberg, A. K. Hedman, A. Nica, K. S. Small, T. C. Mu, E. T. Dermitzakis, M. I. McCarthy, J. Mill, T. D. Spector and P. Deloukas (2012). "Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population." PLoS Genet **8**(4): e1002629.

Benito-Kwiecinski, S., S. L. Giandomenico, M. Sutcliffe, E. S. Riis, P. Freire-Pritchett, I. Kelava, S. Wunderlich, U. Martin, G. A. Wray and K. McDole (2021). "An early cell shape transition drives evolutionary expansion of the human forebrain." Cell **184**(8): 2084-2102. e2019.

Berto, S., I. Mendizabal, N. Usui, K. Toriumi, P. Chatterjee, C. Douglas, C. A. Tamminga, T. M. Preuss, S. V. Yi and G. Konopka (2019). "Accelerated evolution of oligodendrocytes in the human brain." Proceedings of the National Academy of Sciences **116**(48): 24334.

Bird, C. P., B. E. Stranger, M. Liu, D. J. Thomas, C. E. Ingle, C. Beazley, W. Miller, M. E. Hurles and E. T. Dermitzakis (2007). "Fast-evolving noncoding sequences in the human genome." Genome biology **8**(6): R118.

Bors, E. K., C. S. Baker, P. R. Wade, K. B. O'Neill, K. E. Shelden, M. J. Thompson, Z. Fei, S. Jarman and S. Horvath (2021). "An epigenetic clock to estimate the age of living beluga whales." Evolutionary applications **14**(5): 1263-1273.

Bush, E. C. and B. T. Lahn (2008). "A genome-wide screen for noncoding elements important in primate evolution." BMC evolutionary biology **8**(1): 17.

Caceres, M., J. Lachuer, M. A. Zapala, J. C. Redmond, L. Kudo, D. H. Geschwind, D. J. Lockhart, T. M. Preuss and C. Barlow (2003). "Elevated gene expression levels distinguish human from non-human primate brains." Proc Natl Acad Sci U S A **100**(22): 13030-13035.

Capra, J. A., G. D. Erwin, G. McKinsey, J. L. Rubenstein and K. S. Pollard (2013). "Many human accelerated regions are developmental enhancers." Philosophical Transactions of the Royal Society B: Biological Sciences **368**(1632): 20130025.

Capra, J. A., G. D. Erwin, G. McKinsey, J. L. R. Rubenstein and K. S. Pollard (2013). "Many human accelerated regions are developmental enhancers." Philosophical Transactions of the Royal Society B: Biological Sciences **368**(1632): 20130025.

Castelijns, B., M. L. Baak, I. S. Timpanaro, C. R. Wiggers, M. W. Vermunt, P. Shang, I. Kondova, G. Geeven, V. Bianchi and W. de Laat (2020). "Hominin-specific regulatory elements selectively emerged in oligodendrocytes and are disrupted in autism patients." Nature communications **11**(1): 1-12.

Castelijns, B., M. L. Baak, I. S. Timpanaro, C. R. M. Wiggers, M. W. Vermunt, P. Shang, I. Kondova, G. Geeven, V. Bianchi, W. de Laat, N. Geijsen and M. P. Creyghton (2020). "Hominin-specific regulatory elements selectively emerged in oligodendrocytes and are disrupted in autism patients." Nature Communications **11**(1): 301.

Cooney, C. A. (1993). "Are somatic cells inherently deficient in methylation metabolism? A proposed mechanism for DNA methylation loss, senescence and aging." Growth, development, and aging : GDA **57**(4): 261-273.

De Manuel, M., M. Kuhlwilm, P. Frandsen, V. C. Sousa, T. Desai, J. Prado-Martinez, J. Hernandez-Rodriguez, I. Dupanloup, O. Lao and P. Hallast (2016). "Chimpanzee genomic diversity reveals ancient admixture with bonobos." Science **354**(6311): 477-481.

de Sousa, A. A., C. C. Sherwood, A. Schleicher, K. Amunts, C. E. MacLeod, P. R. Hof and K. Zilles (2010). "Comparative cytoarchitectural analyses of striate and extrastriate areas in hominoids." Cerebral Cortex **20**(4): 966-981.

DeFelipe, J., L. Alonso-Nanclares and J. I. Arellano (2002). "Microstructure of the neocortex: comparative aspects." Journal of neurocytology **31**(3): 299-316.

Dennis, M. Y., X. Nuttle, P. H. Sudmant, F. Antonacci, T. A. Graves, M. Nefedov, J. A. Rosenfeld, S. Sajjadian, M. Malig and H. Kotkiewicz (2012). "Evolution of human-specific neural SRGAP2 genes by incomplete segmental duplication." Cell **149**(4): 912-922.

Domazet-Lošo, T. and D. Tautz (2008). "An Ancient Evolutionary Origin of Genes Associated with Human Genetic Diseases." Molecular Biology and Evolution **25**(12): 2699-2707.

Du, A., A. M. Zipkin, K. G. Hatala, E. Renner, J. L. Baker, S. Bianchi, K. H. Bernal and B. A. Wood (2018). "Pattern and process in hominin brain size evolution are scale-dependent." Proceedings of the Royal Society B: Biological Sciences **285**(1873): 20172738.

Ebersole, T. A., Q. Chen, M. J. Justice and K. Artzt (1996). "The quaking gene product necessary in embryogenesis and myelination combines features of RNA binding and signal transduction proteins." Nature Genetics **12**(3): 260-265.

Egger, G., G. Liang, A. Aparicio and P. A. Jones (2004). "Epigenetics in human disease and prospects for epigenetic therapy." Nature **429**(6990): 457-463.

Elango, N., S.-H. Kim, N. C. S. Program, E. Vigoda and S. V. Yi (2008). "Mutations of different molecular origins exhibit contrasting patterns of regional substitution rate variation." PLoS Computational Biology **4**(2): e1000015.

Elango, N. and S. V. Yi (2008). "DNA methylation and structural and functional bimodality of vertebrate promoters." Mol. Biol. Evol. **25**: 1602-1608.

Engebretsen, S. and J. Bohlin (2019). "Statistical predictions with glmnet." Clinical epigenetics **11**(1): 1-3.

Feng, H., K. N. Conneely and H. Wu (2014). "A Bayesian hierarchical model to detect differentially methylated loci from single nucleotide resolution sequencing data." Nucleic acids research **42**(8): e69-e69.

Feng, J., H. Chang, E. Li and G. Fan (2005). "Dynamic expression of de novo DNA methyltransferases Dnmt3a and Dnmt3b in the central nervous system." Journal of neuroscience research **79**(6): 734-746.

Feng, J. and G. Fan (2009). "The role of DNA methylation in the central nervous system and neuropsychiatric disorders." International review of neurobiology **89**: 67-84.

Finucane, H. K., B. Bulik-Sullivan, A. Gusev, G. Trynka, Y. Reshef, P.-R. Loh, V. Anttila, H. Xu, C. Zang, K. Farh, S. Ripke, F. R. Day, S. Purcell, E. Stahl, S. Lindstrom, J. R. B. Perry, Y. Okada, S. Raychaudhuri, M. J. Daly, N. Patterson, B. M. Neale, A. L. Price, C. ReproGen, C. Schizophrenia Working Group of the Psychiatric Genomics and R. C. The (2015). "Partitioning heritability by functional annotation using genome-wide association summary statistics." Nature Genetics **47**(11): 1228-1235.

Finucane, H. K., Y. A. Reshef, V. Anttila, K. Slowikowski, A. Gusev, A. Byrnes, S. Gazal, P.-R. Loh, C. Lareau, N. Shoresh, G. Genovese, A. Saunders, E. Macosko, S. Pollack, C. Brainstorm, J. R. B. Perry, J. D. Buenrostro, B. E. Bernstein, S. Raychaudhuri, S. McCarroll, B. M. Neale and A. L. Price (2018). "Heritability enrichment of specifically

expressed genes identifies disease-relevant tissues and cell types." Nature genetics **50**(4): 621-629.

Fraga, M. F., E. Ballestar, M. F. Paz, S. Ropero, F. Setien, M. L. Ballestar, D. Heine-Suner, J. C. Cigudosa, M. Urioste, J. Benitez, M. Boix-Chornet, A. Sanchez-Aguilera, C. Ling, E. Carlsson, P. Poulsen, A. Vaag, Z. Stephan, T. D. Spector, Y. Z. Wu, C. Plass and M. Esteller (2005). "Epigenetic differences arise during the lifetime of monozygotic twins." Proc Natl Acad Sci U S A **102**: 10604-10609.

Fullard, J. F., M. E. Hauberg, J. Bendl, G. Egervari, M.-D. Cirnaru, S. M. Reach, J. Motl, M. E. Ehrlich, Y. L. Hurd and P. Roussos (2018). "An atlas of chromatin accessibility in the adult human brain." Genome Research **28**(8): 1243-1252.

Girdhar, K., G. E. Hoffman, Y. Jiang, L. Brown, M. Kundakovic, M. E. Hauberg, N. J. Francoeur, Y.-C. Wang, H. Shah, D. H. Kavanagh, E. Zharovsky, R. Jacobov, J. R. Wiseman, R. Park, J. S. Johnson, B. S. Kassim, L. Sloofman, E. Mattei, Z. Weng, S. K. Sieberts, M. A. Peters, B. T. Harris, B. K. Lipska, P. Sklar, P. Roussos and S. Akbarian (2018). "Cell-specific histone modification maps in the human frontal lobe link schizophrenia risk to the neuronal epigenome." Nature neuroscience **21**(8): 1126-1136.

Goel, A. and H.-K. Lee (2007). "Persistence of experience-induced homeostatic synaptic plasticity through adulthood in superficial layers of mouse visual cortex." Journal of Neuroscience **27**(25): 6692-6700.

GTEx, C., F. Aguet, A. A. Brown, S. E. Castel, J. R. Davis, Y. He, B. Jo, P. Mohammadi, Y. Park, P. Parsana, A. V. Segrè, B. J. Strober, Z. Zappala, B. B. Cummings, E. T. Gelfand, K. Hadley, K. H. Huang, M. Lek, X. Li, J. L. Nedzel, D. Y. Nguyen, M. S. Noble, T. J. Sullivan, T. Tukiainen, D. G. MacArthur, G. Getz, A. Addington, P. Guan, S. Koester, A. R. Little, N. C. Lockhart, H. M. Moore, A. Rao, J. P. Struewing, S. Volpi, L. E. Brigham, R. Hasz, M. Hunter, C. Johns, M. Johnson, G. Kopen, W. F. Leinweber, J. T. Lonsdale, A. McDonald, B. Mestichelli, K. Myer, B. Roe, M. Salvatore, S. Shad, J. A. Thomas, G. Walters, M. Washington, J. Wheeler, J. Bridge, B. A. Foster, B. M. Gillard, E. Karasik, R. Kumar, M. Miklos, M. T. Moser, S. D. Jewell, R. G. Montroy, D. C. Rohrer, D. Valley, D. C. Mash, D. A. Davis, L. Sobin, M. E. Barcus, P. A. Branton, N. S. Abell, B. Balliu, O. Delaneau, L. Frésard, E. R. Gamazon, D. Garrido-Martín, A. D. H. Gewirtz, G. Gliner, M. J. Gloudemans, B. Han, A. Z. He, F. Hormozdiari, X. Li, B. Liu, E. Y. Kang, I. C. McDowell, H. Ongen, J. J. Palowitch, C. B. Peterson, G. Quon, S. Ripke, A. Saha, A. A. Shabalin, T. C. Shimko, J. H. Sul, N. A. Teran, E. K. Tsang, H. Zhang, Y.-H. Zhou, C. D. Bustamante, N. J. Cox, R. Guigó, M. Kellis, M. I. McCarthy, D. F. Conrad, E. Eskin, G. Li, A. B. Nobel, C. Sabatti, B. E. Stranger, X. Wen, F. A. Wright, K. G. Ardlie, E. T. Dermitzakis, T. Lappalainen, F. Aguet, K. G. Ardlie, B. B. Cummings, E. T. Gelfand, G. Getz, K. Hadley, R. E. Handsaker, K. H. Huang, S. Kashin, K. J. Karczewski, M. Lek, X. Li, D. G. MacArthur, J. L. Nedzel, D. T. Nguyen, M. S. Noble, A. V. Segrè, C. A. Trowbridge, T. Tukiainen, N. S. Abell, B. Balliu, R. Barshir, O. Basha, A. Battle, G. K. Bogu, A. Brown, C. D. Brown, S. E. Castel, L. S. Chen, C. Chiang, D. F. Conrad, N. J. Cox, F. N. Damani, J. R. Davis, O. Delaneau, E. T. Dermitzakis, B. E. Engelhardt, E. Eskin, P. G. Ferreira, L. Frésard, E. R. Gamazon, D. Garrido-Martín, A. D. H. Gewirtz, G. Gliner,

M. J. Gloudemans, R. Guigo, I. M. Hall, B. Han, Y. He, F. Hormozdiari, C. Howald, H. Kyung Im, B. Jo, E. Yong Kang, Y. Kim, S. Kim-Hellmuth, T. Lappalainen, G. Li, X. Li, B. Liu, S. Mangul, M. I. McCarthy, I. C. McDowell, P. Mohammadi, J. Monlong, S. B. Montgomery, M. Muñoz-Aguirre, A. W. Ndungu, D. L. Nicolae, A. B. Nobel, M. Oliva, H. Ongen, J. J. Palowitch, N. Panousis, P. Papasaikas, Y. Park, P. Parsana, A. J. Payne, C. B. Peterson, J. Quan, F. Reverter, C. Sabatti, A. Saha, M. Sammeth, A. J. Scott, A. A. Shabalin, R. Sodaei, M. Stephens, B. E. Stranger, B. J. Strober, J. H. Sul, E. K. Tsang, S. Urbut, M. van de Bunt, G. Wang, X. Wen, F. A. Wright, H. S. Xi, E. Yeger-Lotem, Z. Zappala, J. B. Zaugg, Y.-H. Zhou, J. M. Akey, D. Bates, J. Chan, L. S. Chen, M. Claussnitzer, K. Demanelis, M. Diegel, J. A. Doherty, A. P. Feinberg, M. S. Fernando, J. Halow, K. D. Hansen, E. Haugen, P. F. Hickey, L. Hou, F. Jasmine, R. Jian, L. Jiang, A. Johnson, R. Kaul, M. Kellis, M. G. Kibriya, K. Lee, J. Billy Li, Q. Li, X. Li, J. Lin, S. Lin, S. Linder, C. Linke, Y. Liu, M. T. Maurano, B. Molinie, S. B. Montgomery, J. Nelson, F. J. Neri, M. Oliva, Y. Park, B. L. Pierce, N. J. Rinaldi, L. F. Rizzardi, R. Sandstrom, A. Skol, K. S. Smith, M. P. Snyder, J. Stamatoyannopoulos, B. E. Stranger, H. Tang, E. K. Tsang, L. Wang, M. Wang, N. Van Wittenberghe, F. Wu, R. Zhang, C. R. Nierras, P. A. Branton, L. J. Carithers, P. Guan, H. M. Moore, A. Rao, J. B. Vaught, S. E. Gould, N. C. Lockart, C. Martin, J. P. Struewing, S. Volpi, A. M. Addington, S. E. Koester, A. R. Little, L. E. Brigham, R. Hasz, M. Hunter, C. Johns, M. Johnson, G. Kopen, W. F. Leinweber, J. T. Lonsdale, A. McDonald, B. Mestichelli, K. Myer, B. Roe, M. Salvatore, S. Shad, J. A. Thomas, G. Walters, M. Washington, J. Wheeler, J. Bridge, B. A. Foster, B. M. Gillard, E. Karasik, R. Kumar, M. Miklos, M. T. Moser, S. D. Jewell, R. G. Montroy, D. C. Rohrer, D. R. Valley, D. A. Davis, D. C. Mash, A. H. Undale, A. M. Smith, D. E. Tabor, N. V. Roche, J. A. McLean, N. Vatanian, K. L. Robinson, L. Sobin, M. E. Barcus, K. M. Valentino, L. Qi, S. Hunter, P. Hariharan, S. Singh, K. S. Um, T. Matose, M. M. Tomaszewski, L. K. Barker, M. Mosavel, L. A. Siminoff, H. M. Traino, P. Flicek, T. Juettemann, M. Ruffier, D. Sheppard, K. Taylor, S. J. Trevanion, D. R. Zerbino, B. Craft, M. Goldman, M. Haeussler, W. J. Kent, C. M. Lee, B. Paten, K. R. Rosenbloom, J. Vivian and J. Zhu (2017). "Genetic effects on gene expression across human tissues." Nature **550**: 204.

Guerrier, S., J. Coutinho-Budd, T. Sassa, A. Gresset, N. V. Jordan, K. Chen, W.-L. Jin, A. Frost and F. Polleux (2009). "The F-BAR domain of srGAP2 induces membrane protrusions required for neuronal migration and morphogenesis." Cell **138**(5): 990-1004.

Hannum, G., J. Guinney, L. Zhao, L. Zhang, G. Hughes, S. Sadda, B. Klotzle, M. Bibikova, J. B. Fan, Y. Gao, R. Deconde, M. Chen, I. Rajapakse, S. Friend, T. Ideker and K. Zhang (2013). "Genome-wide methylation profiles reveal quantitative views of human aging rates." Mol Cell **49**: 359-367.

Hauberg, M. E., J. Creus-Muncunill, J. Bendl, A. Kozlenkov, B. Zeng, C. Corwin, S. Chowdhury, H. Kranz, Y. L. Hurd, M. Wegner, A. D. Børglum, S. Dracheva, M. E. Ehrlich, J. F. Fullard and P. Roussos (2020). "Common schizophrenia risk variants are enriched in open chromatin regions of human glutamatergic neurons." Nature communications **11**(1): 5581-5581.

Herculano-Houzel, S. (2012). "The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost." Proceedings of the National Academy of Sciences **109**(Supplement 1): 10661-10668.

Hernandez, D. G., M. A. Nalls, J. R. Gibbs, S. Arepalli, M. van der Brug, S. Chong, M. Moore, D. L. Longo, M. R. Cookson, B. J. Traynor and A. B. Singleton (2011). "Distinct DNA methylation changes highly correlated with chronological age in the human brain." Human Molecular Genetics **20**(6): 1164-1172.

Heyn, H., N. Li, H. J. Ferreira, S. Moran, D. G. Pisano, A. Gomez, J. Diez, J. V. Sanchez-Mut, F. Setien, F. J. Carmona, A. A. Puca, S. Sayols, M. A. Pujana, J. Serra-Musach, I. Iglesias-Platas, F. Formiga, A. F. Fernandez, M. F. Fraga, S. C. Heath, A. Valencia, I. G. Gut, J. Wang and M. Esteller (2012). "Distinct DNA methylomes of newborns and centenarians." Proceedings of the National Academy of Sciences **109**(26): 10522-10527.

Hoal-van Helden, E. G. and P. D. van Helden (1989). "Age-related methylation changes in DNA may reflect the proliferative potential of organs." Mutation Research/DNAging **219**(5–6): 263-266.

Hodge, R. D., T. E. Bakken, J. A. Miller, K. A. Smith, E. R. Barkan, L. T. Graybuck, J. L. Close, B. Long, N. Johansen, O. Penn, Z. Yao, J. Eggermont, T. Höllt, B. P. Levi, S. I. Shehata, B. Aevermann, A. Beller, D. Bertagnolli, K. Brouner, T. Casper, C. Cobbs, R. Dalley, N. Dee, S.-L. Ding, R. G. Ellenbogen, O. Fong, E. Garren, J. Goldy, R. P. Gwinn, D. Hirschstein, C. D. Keene, M. Keshk, A. L. Ko, K. Lathia, A. Mahfouz, Z. Maltzer, M. McGraw, T. N. Nguyen, J. Nyhus, J. G. Ojemann, A. Oldre, S. Parry, S. Reynolds, C. Rimorin, N. V. Shapovalova, S. Somasundaram, A. Szafer, E. R. Thomsen, M. Tieu, G. Quon, R. H. Scheuermann, R. Yuste, S. M. Sunkin, B. Lelieveldt, D. Feng, L. Ng, A. Bernard, M. Hawrylycz, J. W. Phillips, B. Tasic, H. Zeng, A. R. Jones, C. Koch and E. S. Lein (2019). "Conserved cell types with divergent features in human versus mouse cortex." Nature **573**(7772): 61-68.

Holliday, R. and J. E. Pugh (1975). "DNA modification mechanisms and gene activity during development." Science **187**(4173): 226-232.

Hon, G. C., N. Rajagopal, Y. Shen, D. F. McCleary, F. Yue, M. D. Dang and B. Ren (2013). "Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues." Nat Genet **45**(10): 1198-1206.

Horvath, S. (2013). "DNA methylation age of human tissues and cell types." Genome Biology **14**(10): 3156.

Horvath, S. (2013). "DNA methylation age of human tissues and cell types." Genome biology **14**(10): 1-20.

Horvath, S., A. Haghani, J. A. Zoller, J. Ernst, M. Pellegrini, A. J. Jasinska, J. A. Mattison, A. B. Salmon, K. Raj and S. Jenkins (2020). "DNA methylation study of age and sex in baboons and four other primates." bioRxiv.

Horvath, S. and K. Raj (2018). "DNA methylation-based biomarkers and the epigenetic clock theory of ageing." Nature Reviews Genetics **19**(6): 371-384.

Horvath, S., J. A. Zoller, A. Haghani, A. J. Jasinska, K. Raj, C. E. Breeze, J. Ernst, K. L. Vaughan and J. A. Mattison (2021). "Epigenetic clock and methylation studies in the rhesus macaque." GeroScience: 1-13.

Hubisz, M. J. and K. S. Pollard (2014). "Exploring the genesis and functions of Human Accelerated Regions sheds light on their role in human evolution." Current opinion in genetics & development **29**: 15-21.

Huh, I., X. Yang, T. Park and S. Yi (2014). "Bis-class: a new classification tool of methylation status using bayes classifier and local methylation information." BMC Genomics **15**(1): 608.

Hujoel, M. L. A., S. Gazal, F. Hormozdiari, B. van de Geijn and A. L. Price (2019). "Disease Heritability Enrichment of Regulatory Elements Is Concentrated in Elements with Ancient Sequence Age and Conserved Function across Species." The American Journal of Human Genetics **104**(4): 611-624.

Hutnick, L. K., X. Huang, T.-C. Loo, Z. Ma and G. Fan (2010). "Repression of retrotransposal elements in mouse embryonic stem cells is primarily mediated by a DNA methylation-independent mechanism." Journal of Biological Chemistry **285**(27): 21082-21091.

Jakovcevski, M. and S. Akbarian (2012). "Epigenetic mechanisms in neurological disease." Nature medicine **18**(8): 1194-1204.

Jeong, H., I. Mendizabal, S. Berto, P. Chatterjee, T. Layman, N. Usui, K. Toriumi, C. Douglas, D. Singh, I. Huh, T. M. Preuss, G. Konopka and S. V. Yi (2021). "Evolution of DNA methylation in the human brain." Nature Communications **12**(1): 2021.

Jones, M. J., S. J. Goodman and M. S. Kobor (2015). "DNA methylation and healthy human aging." Aging cell **14**(6): 924-932.

Kappelman, J. (1996). "The evolution of body mass and relative brain size in fossil hominids." Journal of Human Evolution **30**(3): 243-276.

Karczewski, K. J., L. C. Francioli, G. Tiao, B. B. Cummings, J. Alföldi, Q. Wang, R. L. Collins, K. M. Laricchia, A. Ganna, D. P. Birnbaum, L. D. Gauthier, H. Brand, M. Solomonson, N. A. Watts, D. Rhodes, M. Singer-Berk, E. M. England, E. G. Seaby, J. A. Kosmicki, R. K. Walters, K. Tashman, Y. Farjoun, E. Banks, T. Poterba, A. Wang, C. Seed, N. Whiffin, J. X. Chong, K. E. Samocha, E. Pierce-Hoffman, Z. Zappala, A. H. O'Donnell-Luria, E. V. Minikel, B. Weisburd, M. Lek, J. S. Ware, C. Vittal, I. M. Armean, L. Bergelson, K. Cibulskis, K. M. Connolly, M. Covarrubias, S. Donnelly, S. Ferriera, S. Gabriel, J. Gentry, N. Gupta, T. Jeandet, D. Kaplan, C. Llanwarne, R. Munshi, S. Novod, N. Petrillo, D. Roazen, V. Ruano-Rubio, A. Saltzman, M. Schleicher, J. Soto, K. Tibbetts, C. Tolonen, G. Wade, M. E. Talkowski, C. A. Aguilar Salinas, T. Ahmad, C. M. Albert,

D. Ardissino, G. Atzmon, J. Barnard, L. Beaugerie, E. J. Benjamin, M. Boehnke, L. L. Bonnycastle, E. P. Bottinger, D. W. Bowden, M. J. Bown, J. C. Chambers, J. C. Chan, D. Chasman, J. Cho, M. K. Chung, B. Cohen, A. Correa, D. Dabelea, M. J. Daly, D. Darbar, R. Duggirala, J. Dupuis, P. T. Ellinor, R. Elosua, J. Erdmann, T. Esko, M. Färkkilä, J. Florez, A. Franke, G. Getz, B. Glaser, S. J. Glatt, D. Goldstein, C. Gonzalez, L. Groop, C. Haiman, C. Hanis, M. Harms, M. Hiltunen, M. M. Holi, C. M. Hultman, M. Kallela, J. Kaprio, S. Kathiresan, B.-J. Kim, Y. J. Kim, G. Kirov, J. Kooner, S. Koskinen, H. M. Krumholz, S. Kugathasan, S. H. Kwak, M. Laakso, T. Lehtimäki, R. J. F. Loos, S. A. Lubitz, R. C. W. Ma, D. G. MacArthur, J. Marrugat, K. M. Mattila, S. McCarroll, M. I. McCarthy, D. McGovern, R. McPherson, J. B. Meigs, O. Melander, A. Metspalu, B. M. Neale, P. M. Nilsson, M. C. O'Donovan, D. Ongur, L. Orozco, M. J. Owen, C. N. A. Palmer, A. Palotie, K. S. Park, C. Pato, A. E. Pulver, N. Rahman, A. M. Remes, J. D. Rioux, S. Ripatti, D. M. Roden, D. Saleheen, V. Salomaa, N. J. Samani, J. Scharf, H. Schunkert, M. B. Shoemaker, P. Sklar, H. Soininen, H. Sokol, T. Spector, P. F. Sullivan, J. Suvisaari, E. S. Tai, Y. Y. Teo, T. Tiinamaija, M. Tsuang, D. Turner, T. Tusie-Luna, E. Vartiainen, J. S. Ware, H. Watkins, R. K. Weersma, M. Wessman, J. G. Wilson, R. J. Xavier, B. M. Neale, M. J. Daly, D. G. MacArthur and C. Genome Aggregation Database (2020). "The mutational constraint spectrum quantified from variation in 141,456 humans." Nature **581**(7809): 434-443.

Khrameeva, E., I. Kurochkin, D. Han, P. Guijarro, S. Kanton, M. Santel, Z. Qian, S. Rong, P. Mazin, M. Sabirov, M. Bulat, O. Efimova, A. Tkachev, S. Guo, C. C. Sherwood, J. G. Camp, S. Pääbo, B. Treutlein and P. Khaitovich (2020). "Single-cell-resolution transcriptome map of human, chimpanzee, bonobo, and macaque brains." Genome Research **30**(5): 776-789.

Kim, S.-H., N. Elango, C. W. Warden, E. Vigoda and S. Yi (2006). "Heterogeneous genomic molecular clocks in primates." PLoS Genetics **2**: e163.

Kim, S.-H. and S. Yi (2007). "Understanding relationship between sequence and functional evolution in yeast proteins." Genetica **131**: 151-156.

King, M.-C. and A. C. Wilson (1975). "Evolution at two levels in humans and chimpanzees." Science **188**(4184): 107-116.

Konopka, G., T. Friedrich, J. Davis-Turak, K. Winden, M. C. Oldham, F. Gao, L. Chen, G.-Z. Wang, T. M. Preuss and D. H. Geschwind (2012). "Human-specific transcriptional networks in the brain." Neuron **75**: 601-617.

Kozlenkov, A., J. Li, P. Apontes, Y. L. Hurd, W. M. Byne, E. V. Koonin, M. Wegner, E. A. Mukamel and S. Dracheva (2018). "A unique role for DNA (hydroxy)methylation in epigenetic regulation of human inhibitory neurons." Science Advances **4**(9): eaau6190.

Kozlenkov, A., M. Wang, P. Roussos, S. Rudchenko, M. Barbu, M. Bibikova, B. Klotzle, A. J. Dwork, B. Zhang, Y. L. Hurd, E. V. Koonin, M. Wegner and S. Dracheva (2016). "Substantial DNA methylation differences between two major neuronal subtypes in human brain." Nucleic acids research **44**(6): 2593-2612.

Krueger, F. and S. R. Andrews (2011). "Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications." bioinformatics **27**(11): 1571-1572.

Kulakovskiy, I. V., I. E. Vorontsov, I. S. Yevshin, A. V. Soboleva, A. S. Kasianov, H. Ashoor, W. Ba-alawi, V. B. Bajic, Y. A. Medvedeva, F. A. Kolpakov and V. J. Makeev (2016). "HOCOMOCO: expansion and enhancement of the collection of transcription factor binding sites models." Nucleic Acids Research **44**(D1): D116-D125.

Langmead, B. and S. L. Salzberg (2012). "Fast gapped-read alignment with Bowtie 2." Nature methods **9**(4): 357.

Lein, E. S., M. J. Hawrylycz, N. Ao, M. Ayres, A. Bensinger, A. Bernard, A. F. Boe, M. S. Boguski, K. S. Brockway, E. J. Byrnes, L. Chen, L. Chen, T.-M. Chen, M. Chi Chin, J. Chong, B. E. Crook, A. Czaplinska, C. N. Dang, S. Datta, N. R. Dee, A. L. Desaki, T. Desta, E. Diep, T. A. Dolbeare, M. J. Donelan, H.-W. Dong, J. G. Dougherty, B. J. Duncan, A. J. Ebbert, G. Eichele, L. K. Estin, C. Faber, B. A. Facer, R. Fields, S. R. Fischer, T. P. Fliss, C. Frensley, S. N. Gates, K. J. Glattfelder, K. R. Halverson, M. R. Hart, J. G. Hohmann, M. P. Howell, D. P. Jeung, R. A. Johnson, P. T. Karr, R. Kawal, J. M. Kidney, R. H. Knapik, C. L. Kuan, J. H. Lake, A. R. Laramee, K. D. Larsen, C. Lau, T. A. Lemon, A. J. Liang, Y. Liu, L. T. Luong, J. Michaels, J. J. Morgan, R. J. Morgan, M. T. Mortrud, N. F. Mosqueda, L. L. Ng, R. Ng, G. J. Orta, C. C. Overly, T. H. Pak, S. E. Parry, S. D. Pathak, O. C. Pearson, R. B. Puchalski, Z. L. Riley, H. R. Rockett, S. A. Rowland, J. J. Royall, M. J. Ruiz, N. R. Sarno, K. Schaffnit, N. V. Shapovalova, T. Sivisay, C. R. Slaughterbeck, S. C. Smith, K. A. Smith, B. I. Smith, A. J. Sodt, N. N. Stewart, K.-R. Stumpf, S. M. Sunkin, M. Sutram, A. Tam, C. D. Teemer, C. Thaller, C. L. Thompson, L. R. Varnam, A. Visel, R. M. Whitlock, P. E. Wohnoutka, C. K. Wolkey, V. Y. Wong, M. Wood, M. B. Yaylaoglu, R. C. Young, B. L. Youngstrom, X. Feng Yuan, B. Zhang, T. A. Zwingman and A. R. Jones (2007). "Genome-wide atlas of gene expression in the adult mouse brain." Nature **445**(7124): 168-176.

Levine, M. E., A. T. Lu, A. Quach, B. H. Chen, T. L. Assimes, S. Bandinelli, L. Hou, A. A. Baccarelli, J. D. Stewart, Y. Li, E. A. Whitsel, J. G. Wilson, A. P. Reiner, A. Aviv, K. Lohman, Y. Liu, L. Ferrucci and S. Horvath (2018). "An epigenetic biomarker of aging for lifespan and healthspan." Aging **10**(4): 573-591.

Li, H. (2013). "Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM." arXiv preprint arXiv:1303.3997.

Lindblad-Toh, K., M. Garber, O. Zuk, M. F. Lin, B. J. Parker, S. Washietl, P. Kheradpour, J. Ernst, G. Jordan and E. Mauceli (2011). "A high-resolution map of human evolutionary constraint using 29 mammals." Nature **478**(7370): 476-482.

Lister, R., E. A. Mukamel, J. R. Nery, M. Urich, C. A. Puddifoot, N. D. Johnson, J. Lucero, Y. Huang, A. J. Dwork and M. D. Schultz (2013). "Global epigenomic reconfiguration during mammalian brain development." Science **341**(6146).

Lister, R., E. A. Mukamel, J. R. Nery, M. Urich, C. A. Puddifoot, N. D. Johnson, J. Lucero, Y. Huang, A. J. Dwork, M. D. Schultz, M. Yu, J. Tonti-Filippini, H. Heyn, S. Hu, J. C.

Wu, A. Rao, M. Esteller, C. He, F. G. Haghighi, T. J. Sejnowski, M. M. Behrens and J. R. Ecker (2013). "Global Epigenomic Reconfiguration During Mammalian Brain Development." Science **341**(6146): 1237905.

Lister, R., R. C. O'Malley, J. Tonti-Filippini, B. D. Gregory, C. C. Berry, A. H. Millar and J. R. Ecker (2008). "Highly Integrated Single-Base Resolution Maps of the Epigenome in Arabidopsis." Cell **133**(3): 523-536.

Lister, R., M. Pelizzola, Y. S. Kida, R. D. Hawkins, J. R. Nery, G. Hon, J. Antosiewicz-Bourget, R. O'Malley, R. Castanon, S. Klugman, M. Downes, R. Yu, R. Stewart, B. Ren, J. A. Thomson, R. M. Evans and J. R. Ecker (2011). "Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells." Nature **471**(7336): 68-73.

Lomanowska, A. M., V. Lovic, M. J. Rankine, S. J. Mooney, T. E. Robinson and G. W. Kraemer (2011). "Inadequate early social experience increases the incentive salience of reward-related cues in adulthood." Behavioural brain research **220**(1): 91-99.

Lu, A. T., Z. Fei, A. Haghani, T. R. Robeck, J. A. Zoller, C. Z. Li, J. Zhang, J. Ablaeva, D. M. Adams and J. Almunia (2021). "Universal DNA methylation age across mammalian tissues." Biorxiv.

Luo, C., C. L. Keown, L. Kurihara, J. Zhou, Y. He, J. Li, R. Castanon, J. Lucero, J. R. Nery, J. P. Sandoval, B. Bui, T. J. Sejnowski, T. T. Harkins, E. A. Mukamel, M. M. Behrens and J. R. Ecker (2017). "Single-cell methylomes identify neuronal subtypes and regulatory elements in mammalian cortex." Science **357**(6351): 600.

Mammalian Methylation, C., A. T. Lu, Z. Fei, A. Haghani, T. R. Robeck, J. A. Zoller, C. Z. Li, J. Zhang, J. Ablaeva, D. M. Adams, J. Almunia, R. Ardehali, A. Arneson, C. S. Baker, K. Belov, P. Black, D. T. Blumstein, E. K. Bors, C. E. Breeze, R. T. Brooke, J. L. Brown, A. Caulton, J. M. Cavin, I. Chatzistamou, H. Chen, P. Chiavellini, O.-W. Choi, S. Clarke, J. DeYoung, C. Dold, C. K. Emmons, S. Emmrich, C. G. Faulkes, S. H. Ferguson, C. J. Finno, J.-M. Gaillard, E. Garde, V. N. Gladyshev, V. Gorbunova, R. G. Goya, M. J. Grant, E. N. Hales, M. B. Hanson, M. Haulena, A. N. Hogan, C. J. Hogg, T. A. Hore, A. J. Jasinska, G. Jones, E. Jourdain, O. Kashpur, H. Katcher, E. Katsumata, V. Kaza, H. Kiaris, M. S. Kobor, P. Kordowitzki, W. R. Koski, B. Larison, S.-G. Lee, Y. C. Lee, M. Lehmann, J.-F. Lemaitre, A. J. Levine, C. Li, X. Li, D. T. S. Lin, N. Macoretta, D. Maddox, C. O. Matkin, J. A. Mattison, J. Mergl, J. J. Meudt, K. Mozhui, A. Naderi, M. Nagy, P. Narayan, P. W. Nathanielsz, N. B. Nguyen, C. Niehrs, A. G. Ophir, E. A. Ostrander, P. O'Tierney Ginn, K. M. Parsons, K. C. Paul, M. Pellegrini, G. M. Pinho, J. Plassais, N. A. Prado, B. Rey, B. R. Ritz, J. Robbins, M. Rodriguez, J. Russell, E. Rydkina, L. L. Sailer, A. B. Salmon, A. Sanghavi, K. M. Schachtschneider, D. Schmitt, T. Schmitt, L. Schomacher, L. B. Schook, K. E. Sears, A. Seluanov, D. Shanmuganayagam, A. Shindyapina, K. Singh, I. Sinha, R. G. Snell, E. Soltanmaohammadi, M. L. Spangler, L. Staggs, K. J. Steinman, V. J. Sugrue, B. Szladovits, M. Takasugi, E. C. Teeling, M. J. Thompson, B. Van Bonn, S. C. Vernes, D. Villar, H. V. Vinters, M. C. Wallingford, N. Wang, R. K. Wayne, G. S. Wilkinson, C. K. Williams, R. W. Williams, X. W. Yang, B. G. Young, B. Zhang, Z. Zhang, P. Zhao, Y. Zhao, J. Zimmermann, W. Zhou, J. Ernst, K. Raj and S. Horvath (2021).

"Universal DNA methylation age across mammalian tissues." bioRxiv: 2021.2001.2018.426733.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel and M. Daly (2010). "The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data." Genome research **20**(9): 1297-1303.

McLean, C. Y., D. Bristor, M. Hiller, S. L. Clarke, B. T. Schaar, C. B. Lowe, A. M. Wenger and G. Bejerano (2010). "GREAT improves functional interpretation of cis-regulatory regions." Nature Biotechnology **28**(5): 495-501.

Mendizabal, I., S. Berto, N. Usui, K. Toriumi, P. Chatterjee, C. Douglas, I. Huh, H. Jeong, T. Layman and C. A. Tamminga (2019). "Cell type-specific epigenetic links to schizophrenia risk in the brain." Genome biology **20**(1): 1-21.

Mendizabal, I., S. Berto, N. Usui, K. Toriumi, P. Chatterjee, C. Douglas, I. Huh, H. Jeong, T. Layman, C. A. Tamminga, T. M. Preuss, G. Konopka and S. V. Yi (2019). "Cell type-specific epigenetic links to schizophrenia risk in the brain." Genome Biology **20**(1): 135.

Mendizabal, I., L. Shi, T. E. Keller, G. Konopka, T. M. Preuss, T.-F. Hsieh, E. Hu, Z. Zhang, B. Su and S. V. Yi (2016). "Comparative methylome analyses identify epigenetic regulatory loci of human brain evolution." Molecular biology and evolution **33**(11): 2947-2959.

Mendizabal, I. and S. V. Yi (2016). "Whole-genome bisulfite sequencing maps from multiple human tissues reveal novel CpG islands associated with tissue-specific regulation." Human Molecular Genetics **25**(1): 69-82.

Moore, L. D., T. Le and G. Fan (2013). "DNA methylation and its basic function." Neuropsychopharmacology **38**(1): 23-38.

Mugal, C. and H. Ellegren (2011). "Substitution rate variation at human CpG sites correlates with non-CpG divergence, methylation level and GC content." Genome Biology **12**(6): R58.

Ng, B., C. C. White, H.-U. Klein, S. K. Sieberts, C. McCabe, E. Patrick, J. Xu, L. Yu, C. Gaiteri, D. A. Bennett, S. Mostafavi and P. L. De Jager (2017). "An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome." Nature neuroscience **20**(10): 1418-1426.

Numata, S., T. Ye, T. M. Hyde, X. Guitart-Navarro, R. Tao, M. Wininger, C. Colantuoni, D. R. Weinberger, J. E. Kleinman and B. K. Lipska (2012). "DNA methylation signatures in development and aging of the human prefrontal cortex." American journal of human genetics **90**(2): 260-272.

Park, Y. and H. Wu (2016). "Differential methylation analysis for BS-seq data under general experimental design." Bioinformatics **32**(10): 1446-1453.

Perris, C. and P. Andersson (2000). "Experiences of parental rearing and patterns of attachment in adulthood." Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice **7**(4): 279-288.

Pollard, K. S., S. R. Salama, N. Lambert, M.-A. Lambot, S. Coppens, J. S. Pedersen, S. Katzman, B. King, C. Onodera and A. Siepel (2006). "An RNA gene expressed during cortical development evolved rapidly in humans." Nature **443**(7108): 167-172.

Prabhakar, S., J. P. Noonan, S. Pääbo and E. M. Rubin (2006). "Accelerated evolution of conserved noncoding sequences in humans." Science **314**(5800): 786-786.

Prescott, S. L., R. Srinivasan, M. C. Marchetto, I. Grishina, I. Narvaiza, L. Selleri, F. H. Gage, T. Swigut and J. Wysocka (2015). "Enhancer divergence and cis-regulatory evolution in the human and chimp neural crest." Cell **163**(1): 68-83.

Preuss, T. M. (2011). "The human brain: rewired and running hot." Annals of the New York Academy of Sciences **1225**(S1): E182-E191.

Preuss, T. M., M. Caceres, M. C. Oldham and D. H. Geschwind (2004). "Human brain evolution: insights from microarrays." Nat Rev Genet **5**(11): 850-860.

Price, A. J., L. Collado-Torres, N. A. Ivanov, W. Xia, E. E. Burke, J. H. Shin, R. Tao, L. Ma, Y. Jia, T. M. Hyde, J. E. Kleinman, D. R. Weinberger and A. E. Jaffe (2019). "Divergent neuronal DNA methylation patterns across human cortical development reveal critical periods and a unique role of CpH methylation." Genome Biology **20**(1): 196.

Rakic, P. (1995). "A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution." Trends in neurosciences **18**(9): 383-388.

Rakic, P. (2007). "The radial edifice of cortical architecture: from neuronal silhouettes to genetic engineering." Brain research reviews **55**(2): 204-219.

Rakyan, V. K., T. A. Down, S. Maslau, T. Andrew, T.-P. Yang, H. Beyan, P. Whittaker, O. T. McCann, S. Finer, A. M. Valdes, R. D. Leslie, P. Deloukas and T. D. Spector (2010). "Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains." Genome Research **20**(4): 434-439.

Razin, A. and A. D. Riggs (1980). "DNA methylation and gene function." Science **210**(4470): 604-610.

Richardson, B. (2003). "Impact of aging on DNA methylation." Ageing Res Rev **2**(3): 245-261.

Rizzardi, L. F., P. F. Hickey, V. Rodriguez DiBlasi, R. Tryggvadóttir, C. M. Callahan, A. Idrizi, K. D. Hansen and A. P. Feinberg (2019). "Neuronal brain-region-specific DNA methylation and chromatin accessibility are associated with neuropsychiatric trait heritability." Nature Neuroscience **22**(2): 307-316.

Saxonov, S., P. Berg and D. L. Brutlag (2006). "A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters." Proc. Nat. Acad. Sci. USA **103**(5): 1412-1417.

Schübeler, D. (2015). "Function and information content of DNA methylation." Nature **517**: 321.

Sherwood, C. C., F. Subiaul and T. W. Zawidzki (2008). "A natural history of the human mind: tracing evolutionary changes in brain and cognition." Journal of Anatomy **212**(4): 426-454.

Shulha, H. P., J. L. Crisci, D. Reshetov, J. S. Tushir, I. Cheung, R. Bharadwaj, H.-J. Chou, I. B. Houston, C. J. Peter and A. C. Mitchell (2012). "Human-specific histone methylation signatures at transcription start sites in prefrontal neurons." PLoS Biol **10**(11): e1001427.

Shulha, H. P., J. L. Crisci, D. Reshetov, J. S. Tushir, I. Cheung, R. Bharadwaj, H.-J. Chou, I. B. Houston, C. J. Peter, A. C. Mitchell, W.-D. Yao, R. H. Myers, J.-f. Chen, T. M. Preuss, E. I. Rogaev, J. D. Jensen, Z. Weng and S. Akbarian (2012). "Human-Specific Histone Methylation Signatures at Transcription Start Sites in Prefrontal Neurons." PLoS Biology **10**(11): e1001427.

Sousa, A. M. M., K. A. Meyer, G. Santpere, F. O. Gulden and N. Sestan (2017). "Evolution of the Human Nervous System Function, Structure, and Development." Cell **170**(2): 226-247.

Spainhour, J. C. G., H. S. Lim, S. V. Yi and P. Qiu (2019). "Correlation Patterns Between DNA Methylation and Gene Expression in The Cancer Genome Atlas." Cancer Informatics **18**: 1176935119828776.

Spiers, H., E. Hannon, L. C. Schalkwyk, R. Smith, C. C. Y. Wong, M. C. O,ÄôDonovan, N. J. Bray and J. Mill (2015). "Methylomic trajectories across human fetal brain development." Genome Research **25**(3): 338-352.

Stroud, H., S. C. Su, S. Hrvatin, A. W. Greben, W. Renthal, L. D. Boxer, M. A. Nagy, D. R. Hochbaum, B. Kinde and H. W. Gabel (2017). "Early-life gene expression in neurons modulates lasting epigenetic states." Cell **171**(5): 1151-1164. e1116.

Stroud, H., S. C. Su, S. Hrvatin, A. W. Greben, W. Renthal, L. D. Boxer, M. A. Nagy, D. R. Hochbaum, B. Kinde, H. W. Gabel and M. E. Greenberg (2017). "Early-Life Gene Expression in Neurons Modulates Lasting Epigenetic States." Cell **171**(5): 1151-1164.e1116.

Sudmant, P. H., J. O. Kitzman, F. Antonacci, C. Alkan, M. Malig, A. Tsalenko, N. Sampas, L. Bruhn, J. Shendure and G. Project (2010). "Diversity of human copy number variation and multicopy genes." Science **330**(6004): 641-646.

Sun, D. and S. V. Yi (2015). "Impacts of chromatin states and long-range genomic segments on aging and DNA methylation." PloS one **10**(6): e0128517.

Teschendorff, A. E., J. West and S. Beck (2013). "Age-associated epigenetic drift: implications, and a case of epigenetic thrift?" Human molecular genetics **22**(R1): R7-R15.

Vaillant, I. and J. Paszkowski (2007). "Role of histone and DNA methylation in gene regulation." Current opinion in plant biology **10**(5): 528-533.

Varki, A., D. H. Geschwind and E. E. Eichler (2008). "Human uniqueness: genome interactions with environment, behaviour and culture." Nature Reviews Genetics **9**(10): 749-763.

Varki, A., D. H. Geschwind and E. E. Eichler (2008). "Human uniqueness: genome interactions with environment, behaviour and culture." Nat Rev Genet **9**(10): 749-763.

Vermunt, M. W., S. C. Tan, B. Castelijns, G. Geeven, P. Reinink, E. De Bruijn, I. Kondova, S. Persengiev, R. Bontrop and E. Cuppen (2016). "Epigenomic annotation of gene regulatory alterations during evolution of the primate brain." Nature neuroscience **19**(3): 494-503.

Vermunt, M. W., S. C. Tan, B. Castelijns, G. Geeven, P. Reinink, E. de Bruijn, I. Kondova, S. Persengiev, R. Bontrop, E. Cuppen, W. de Laat, M. P. Creyghton and B. Netherlands Brain (2016). "Epigenomic annotation of gene regulatory alterations during evolution of the primate brain." Nature Neuroscience **19**(3): 494-503.

Wang, H.-Y., H.-C. Chien, N. Osada, K. Hashimoto, S. Sugano, T. Gojobori, C.-K. Chou, S.-F. Tsai, C.-I. Wu and C.-K. J. Shen (2007). "Rate of evolution in brain-expressed genes in humans and other primates." PLoS biology **5**(2): e13.

Warnecke, P. M., C. Stirzaker, J. Song, C. Grunau, J. R. Melki and S. J. Clark (2002). "Identification and resolution of artifacts in bisulfite sequencing." Methods **27**(2): 101-107.

Weber, M., I. Hellmann, M. B. Stadler, L. Ramos, S. Pääbo and M. Rebhan (2007). "Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome." Nat Genet. **39**.

Wen, L., X. Li, L. Yan, Y. Tan, R. Li, Y. Zhao, Y. Wang, J. Xie, Y. Zhang, C. Song, M. Yu, X. Liu, P. Zhu, X. Li, Y. Hou, H. Guo, X. Wu, C. He, R. Li, F. Tang and J. Qiao (2014). "Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain." Genome Biology **15**(3): R49.

Wilson Vincent, L. and A. Jones Peter (1983). "DNA Methylation Decreases in Aging But Not in Immortal Cells." Science **220**(4601): 1055-1057.

Won, H., L. de la Torre-Ubieta, J. L. Stein, N. N. Parikshak, J. Huang, C. K. Opland, M. J. Gandal, G. J. Sutton, F. Hormozdiari, D. Lu, C. Lee, E. Eskin, I. Voineagu, J. Ernst and D. H. Geschwind (2016). "Chromosome conformation elucidates regulatory relationships in developing human brain." Nature **538**(7626): 523-527.

Wonders, C. P. and S. A. Anderson (2006). "The origin and specification of cortical interneurons." Nature Reviews Neuroscience **7**(9): 687-696.

Yi, S. V. (2017). "Insights into Epigenome Evolution from Animal and Plant Methylomes." Genome Biology and Evolution **9**(11): 3189-3201.

Zeng, J., G. Konopka, B. G. Hunt, T. M. Preuss, D. Geschwind and S. V. Yi (2012). "Divergent whole-genome methylation maps of human and chimpanzee brains reveal epigenetic basis of human regulatory evolution." American journal of human genetics **91**(3): 455-465.

Zerbino, D. R., P. Achuthan, W. Akanni, M. R. Amode, D. Barrell, J. Bhai, K. Billis, C. Cummins, A. Gall and C. G. Girón (2018). "Ensembl 2018." Nucleic acids research **46**(D1): D754-D761.

Zhu, Y., A. M. M. Sousa, T. Gao, M. Skarica, M. Li, G. Santpere, P. Esteller-Cucala, D. Juan, L. Ferrández-Peral, F. O. Gulden, M. Yang, D. J. Miller, T. Marques-Bonet, Y. Imamura Kawasawa, H. Zhao and N. Sestan (2018). "Spatiotemporal transcriptomic divergence across human and macaque brain development." Science **362**(6420): eaat8077.

Ziller, M. J., F. Müller, J. Liao, Y. Zhang, H. Gu, C. Bock, P. Boyle, C. B. Epstein, B. E. Bernstein, T. Lengauer, A. Gnirke and A. Meissner (2011). "Genomic distribution and inter-sample variation of non-CpG methylation across human cell types." PLoS genetics **7**(12): e1002389-e1002389.