

University of Nottingham
UK | CHINA | MALAYSIA

Assessing DNA methylation epialleles for human cognitive dysfunction and brain disease

Nicholas Jude Rielly
BSc Biological Sciences,
MSc Microbiology and Immunology,

Thesis submitted to The University of Nottingham
for the degree of Doctor of Philosophy

September 2019

Dedicated with love to Paul, Chrys, and Tegan

Declaration

I declare that all work presented in this thesis has been produced by myself and has not been submitted for any other degree. Every effort has been made to acknowledge collaborations and the sources of previous research have been cited throughout.

Nicholas Jude Rielly

Acknowledgements

First and foremost, I must thank my supervisor, Dr Helen Miranda Knight, without whom this project would have been near impossible. She has guided not only the research but my academic career and nurtured my love of science. She continually surprises me with ingenuity in the face of adverse circumstances and has been a brilliant role model.

Professor John Armour has similarly had a huge impact on this project providing insight into the most complex aspects of this study. He has provided invaluable knowledge and his calm and rational approach to his work has not only made this project easier but has been something to aspire to. I am grateful for his advice and support.

I could not write these acknowledgements without thanking everyone in the Knight lab. Miles Flitton, Maria Koromina, Braulio Martinez De La Cruz, Eleanor Bellows, and Maria Isabell Haig. All of your support and kindness turned work into pleasure, and I will miss you all sorely when we part ways. I also must thank my MSc students, Rhian Warman, and Dominika Valent Raffajová, who were not only a pleasure to work with but were excellent in the lab. We changed in the Knight lab together, growing from inexperienced students into aspiring scientists. I know they will all have hugely successful careers, and I wish them the very best in their lives.

There are countless others whom which have supported this project and provided invaluable insight into new research and technologies. Daniel Ellis, Jessica Foraker, Chang En Yu, Matthew Carlisle, Djordje Gveric, and Rita San-Bento. Without these people, this project would not have become what it is now.

I would not have had this opportunity without the Rosetrees Trust and co-donors who funded the PhD and stipend for this project. Their continued generosity allowed me to explore multiple questions in this project, and they facilitate many more projects just like it.

I must also thank my fantastic partner Garreth Brown, who has always supported and believed in me. He made this project so much easier and was patient with me when I was struggling. He is one of the kindest people I know, and I hope one day I can return the favour.

Finally, I want to thank my parents, Paul and Chrys, and my sister Tegan, who have never doubted me for a second, even when I did. They have cheered me on ever since I expressed an interest in science as a little boy. They are the wind in my sails, and I am incredibly grateful.

Abstract

In 2016, neurological disorders were the leading cause of disability adjusted life years (number of years lost to illness or disability) and the second leading cause of death. Extensive research has been performed on associations between genetic variants and risk for neurodegenerative diseases. Now, epigenomic processes such as DNA methylation, the addition of a methyl group to the 5th carbon of cytosine (5mC) and its oxidised derivatives (5hmC, 5fC, and 5CaC), are being increasingly studied. DNA methylation contributes to changes in cognition and ageing and has been associated with several neurodegenerative diseases such as Parkinson's disease and Alzheimer's disease. However, little research has focused the effect methylation at these sites may have on the regulation of gene expression.

In this study 5mC, 5hmC, and 5fC/CaC DNA methylation levels were measured at the complementary antisense CpG of the *BDNF* Val66Met rs6265 variant, CpG sites at *APOE* rs7412 and rs429358 risk variants which determine APOE ε isoforms (ε2 - ε4), and DNMT3L rs7354779 (R278G) which are known risk variants for cognitive diseases. Methylation was quantified through bisulphite, oxibisulphite, and methylation assisted bisulphite sequencing and compared across multiple tissues, brain regions and across mild cognitive impairment, Alzheimer's disease, dementia with Lewy bodies, and Parkinson's disease as well as compared with sense and antisense *BDNF* transcript expression.

Analysis of methylation at the antisense *BDNF* Val66Met CpG (AS-rs6265) showed significantly higher levels of 5mC compared to 5hmC methylation in C homozygotes ($p = 1\times10^{-3}$). Tissue as well as tissue and genotype had significant associations with 5mC and 5hmC methylation, respectively, in the Imperial samples. Whole brain and hippocampus tissues, as well as Parkinson's disease and Alzheimer's disease analyses of C homozygotes all showed patterns where 5mC was higher than 5hmC levels. 5fC/CaC methylation was observed in all tissues and disease states. Samples from individuals with mild cognitive impairment or from the cerebellum showed higher

(but non-significant) levels of AS-rs6265 5fC/CaC than 5mC methylation, and AS-rs6265 5fC/CaC methylation was bimodal in distribution.

BDNF sense mRNA transcript expression correlated strongly and significantly with *BDNF* antisense transcript AS-204 expression across all samples while an alternative antisense transcript (AS-201) showed a negative correlation with both transcripts suggesting that AS-201 may be acting as a cis-regulatory element for *BDNF* sense mRNA. Expression analysis indicated significantly higher expression of BDNF antisense AS-201 transcript and lower expression of *BDNF* sense mRNA and BDNF antisense AS-204 in dementia with Lewy bodies samples compared to control samples ($p = 2.4\times10^{-2}$). *BDNF* mRNA showed a significant decrease in expression with AS-rs6265 5mC methylation in dementia with Lewy bodies ($p = 3\times10^{-2}$) and mild cognitive impairment samples ($p = 0.04$) and 5hmC levels were found to significantly correlate with decreased AS-201 transcript abundance in cerebellum ($p = 2\times10^{-3}$) and significantly correlate with increased mRNA abundance in C homozygotes ($p = 0.03$), respectively.

5mC methylation was non-significantly higher than 5hmC methylation across both *APOE* ε2-ε4 rs7412 and rs429358 variant sites and also observed in Parkinson's disease, dementia with Lewy bodies, and Alzheimer's disease samples at the rs7412 variant site. 5fC/CaC was observationally higher than 5mC in cerebellum and dementia with Lewy bodies samples at the rs429358 epiallele. When analysing by ε isoform, 5hmC methylation showed evidence of allele specific methylation and 5hmC methylation was highest in ε3/ε4 isoform individuals, which may suggest that methylation is being regulated and could influence disease risk.

This study indicates that methylation at risk variant sites shows brain region and disease specific methylation patterns and may, through cis-acting regulatory mechanisms, spatially and temporally regulate expression of proteins in disease. Furthermore, it provides new mechanistic insight into antisense strand methylation and ratios of 5mC to 5hmC in ASM.

Contents

Declaration i
Acknowledgements ii
Abstract iv
Contents vi
List of tables xi
List of Figures xiii
Abbreviations xvi
Chapter 1: Thesis Introduction 1
1.1 The Brain and cognition 2
1.2 Brain Disease 3
1.2.1 Schizophrenia 4
1.2.2 Autism spectrum disorders 5
1.2.3 Mild Cognitive Impairment 6
1.2.4 Alzheimer's Disease 6
1.2.5 Synucleinopathies 9
1.2.6 Parkinson's Disease 9
1.2.7 Dementia with Lewy Bodies 13
1.3 Epigenetics, DNA methylation and epialleles 14
1.3.1 DNA 5mC methylation 15
1.3.2 Dysregulation of 5mC methylation, methylation turnover, and disease 18
1.3.3 5hmC and other oxidised forms of methylation 21
1.3.4 5hmC, 5fC, and 5CaC dysregulation and disease 27
1.3.5 Technology to investigate methylation 32
1.4 Candidate DNA methylation epialleles 37
1.4.1 Apolipoprotein E and the rs7412 and rs429358 ε isoform variants 37
1.4.2 Brain Derived Neurotrophic Factor and rs6265 42
1.4.3 DNA Methyltransferase 3 Like and rs7354779 49
1.5 Introduction to thesis studies 51
Chapter 2: Materials and Methods 53
2.1 Materials and methods 54
2.1.1 Samples 54

2.1.2 Reagents 55
2.2 In Vitro Protocols and materials 63
2.2.1 Probe and primer designs 63
2.2.2 DNA & RNA extraction from snap frozen tissues 73
2.2.3 Amplification, Quantification and Quality control 74
2.2.4 DNA conversion to assess Methylation 80
2.2.5 Preparation of samples and sequencing 85
2.2.6 Digital Droplet PCR (DD-PCR) 87
2.3 Data Analysis 89
2.3.1 In vitro data analysis 89
2.3.2 Statistical analysis of methylation and gene expression data 90
2.3.3 Statistical analysis of Digital Droplet PCR expression and pyrosequencing data 94
2.3.4 Analysis of cognitive performance using UK10K Data 95
2.3.5 Bioinformatic tools 99
Chapter 3: Optimisation of techniques to examine DNA methylation in blood and brain 100
3.1 Workflow of the optimisation of techniques and DNA extraction 101
3.1.1 Extraction of genomic material from snap frozen tissues 103
3.2 Optimisation of primers for sequencing 104
3.3 Optimisation of (oxi)bisulphite conversion for the analysis of 5(h)mC 105
3.3.1 Kits to assess bisulphite and oxibisulphite conversion 105
3.3.2 Bisulphite and oxibisulphite conversion to assess 5mC/5hmC 106
3.4 Validating conversion efficiency 106
3.4.1 Methods to assess conversion efficiency. 106
3.4.2 Validating oxibisulphite conversion 107
3.4.3 Optimisation of 5fC and 5CaC Analysis 109
3.4.4 CEGX TrueMethyl Array fC analysis of 5fC methylation 109
3.4.5 Methylation assisted bisulphite conversion and sequencing 111
3.5 Optimising the quantification of methylation abundance 115
3.5.1 Quantifying methylation through Sanger and pyrosequencing 115
3.6 Optimisation of DNMT3L risk allele primer sets and DNMT3L analysis 116

3.7 Discussion 119
3.7.1 Bisulphite conversion of DNA to quantify methylation 119
3.7.2 Sequencing methods 121
3.7.3 Quantification of methylation at DNMT3L rs7354779 122
3.7.4 Future developments 123
3.7.5 Conclusions 123
Chapter 4: DNA methylation analysis of the antisense *BDNF* rs6265 epiallele 124
4.1 Workflow pipeline for the analysis of BDNF 125
4.2 Association study of the BDNF rs6265 variant in individuals with psychiatric disease 127
4.3 BDNF rs6265 complementary antisense CpG allele in individuals with neurodegenerative diseases 129
4.4 Pyrosequencing methylation analysis 130
4.4.1 5mC and 5hmC methylation at AS-rs6265 by genotype 131
4.4.2 Tissue, Disease, genotype status and AS-rs6265 5mC/5hmC methylation 136
4.4.3 Tissue, Disease, genotype status and AS-rs6265 5fC/CaC methylation 140
4.5 Discussion 145
4.5.1 BDNF rs6265 genotype association 145
4.5.2 5mC and 5hmC methylation across antisense rs6265 carrier status 146
4.5.3 5mC and 5hmC methylation across tissues and cognitive disease 147
4.5.4 5fC/CaC methylation analyses 151
4.5.5 Limitations and future studies 154
4.5.6 Conclusion 156
Chapter 5: *BDNF* rs6265 epiallele methylation and *BDNF* transcript abundance 158
5.1 Workflow pipeline for the analysis of BDNF 159
5.2 *BDNF* transcript expression and correlation within tissues and disease states 159
5.2.1 Transcript expression correlations 160
5.2.2 Comparison of BDNF transcript abundance across brain regions 163

5.2.3 Comparison of BDNF transcript abundance across disease phenotypes....167
5.2.4 Transcript expression between disease states within each brain region....169
5.2.5 Comparison of transcript abundance across brain regions within each disease state....173
5.3 *BDNF* transcript abundance, and 5mC, 5hmC, 5fC/CaC methylation at antisense *BDNF* rs6265....186
5.3.1 5mC and 5hmC methylation at AS-rs6265 and BDNF transcript expression....186
5.3.2 5(h)mC methylation, BDNF transcript abundance examined by genotype, brain region and disease phenotype, with age and gender as covariates....188
5.3.3 5fC/CaC methylation and BDNF transcript expression....193
5.3.4 5fC/CaC methylation, transcript expression, tissue type and diagnoses, with and without age and gender as covariates....195
5.4 Discussion....196
5.4.1 BDNF transcript abundance and correlations within brain tissues and disease states....196
5.4.2 Comparing BDNF transcript abundance and 5mC, 5hmC and 5fC/CaC methylation at antisense rs6265....202
5.4.3 Limitations and future research....207
5.4.4 Conclusion....208
Chapter 6: *APOE* rs7412 and rs42958 epiallele methylation analysis....210
6.1 Sequencing genotype and methylation analysis of *APOE* variants....211
6.1.1 Genotypes at APOE rs7412 and rs429358....214
6.2 5mC and 5hmC methylation analyses at *APOE* rs7412 and rs429358....214
6.2.1 5mC and 5hmC methylation compared to Carrier status....215
6.2.2 5mC and 5hmC methylation, compared to genotype and tissue types....220
6.2.3 5mC and 5hmC methylation, compared to genotype and disease states....222
6.3 5mC, 5hmC and 5fC/CaC methylation analyses....224
6.3.1 APOE rs7412 5mC, 5hmC, and 5fC/CaC methylation, across tissue types....224
6.3.2 APOE rs429358 5mC, 5hmC and 5fC/CaC methylation, and tissue types....227

6.3.3 APOE rs7412 5mC, 5hmC, and 5fC/CaC methylation across disease states 230
6.3.4 APOE rs429358 5mC, 5hmC, and 5fC/CaC methylation, across disease states 232
6.4 Comparison of ε Isoform 5mC and 5hmC methylation levels 234
6.4.1 Methylation levels per APOE ε isoform 234
6.4.2 5mC and 5hmC methylation split by disease state and ε isoform 236
6.4.3 Combined 5mC or 5hmC methylation across both rs7412 and rs429358 epialleles 238
6.5 Discussion 240
6.5.1 ε isoform analyses 240
6.5.2 5mC and 5hmC methylation, APOE risk genotype, tissue and disease 240
6.5.3 5mC, 5hmC, and 5fC/CaC methylation, tissue and disease 244
6.5.4 5mC, 5hmC methylation and ε Isoform analysis 246
6.5.5 Limitations and future research 248
6.5.6 Conclusion 249
Chapter 7: General Discussion 251
7.1 Methylation and the effects of ASM and CRE in disease 252
7.1.1 Methylation as a mechanism causing or as a result of disease 252
7.1.2 The role of genotype, ASM, and CRE 254
7.1.3 Methylation state ratios, and the significance of 5fC/CaC methylation 256
7.2 *BDNF* antisense methylation and mechanisms behind transcript expression 258
7.3 *APOE* variant analysis and proposed mechanisms of regulation 260
7.4 The field of epigenetics in disease, and future studies 262
7.5 *APOE* and *BDNF* as genes in cognition 264
References 266

List of tables

Table 1.1	The relationship between *APOE* rs7412 and rs429358 genotype, amino acid substitutions, resultant ε isoforms, and effects on risk for Alzheimer's disease	39
Table 2.1	Unaffected control DNA samples used to optimise DNA extraction, PCR amplification, DNA conversion, and bisulphite PCR amplification	56
Table 2.2	OPTIMA blood and brain samples used for analysis	57
Table 2.3	Parkinson's UK Imperial College Brain Bank snap frozen tissue Samples	58 -59
Table 2.4	Premade kits	60
Table 2.5	Reagents and polymerase systems used	61
Table 2.6	Solutions and reagents that were made in house with chemical constituents	62
Table 2.7	Primers used for PCR amplification	66-68
Table 2.8	Pyrosequencing primers	69-70
Table 2.9	The optimal pyrosequencing primer sets or combinations used to assess the genotype and methylation status of gDNA and converted DNA	71
Table 2.10	Digital Droplet probes used to assess the expression of target genes	72
Table 2.11	Standard thermal cycler running conditions used in all polymerase systems	75
Table 2.12	Optimal running conditions for primer sets in PCR amplification	76-77
Table 2.13	UK10K cohort genotype datasets used to examine BDNF rs6265 genotypes in association studies	96
Table 2.14	A list of bioinformatic tools and their associated websites	99
Table 3.1	Comparison of quantification of 5mC/5hmC methylation through Sanger sequencing (observed Sanger-seq), pyrosequencing (observed pyro-seq) and data reported by Schultz et. al.	118
Table 3.2	Bisulphite Sanger sequencing and semi-quantification of methylation at the *DNMT3L* rs7354779 variant using the ab1peakreporter tool	118
Table 4.1	Association of rs6265 minor allele with neuropsychiatric phenotypes	128
Table 4.2	*BDNF* rs6265 antisense complementary genotypes in disease and control samples	128
Table 4.3	Successful data points used in analysis of *BDNF* rs6265 antisense complementary genotypes, with 5(h)mC methylation, diagnoses, and tissues	133-134

Table 4.4	Linear mixed effects model analysis for disease, tissue, and genotype variables across different samples and 5mC or 5hmC methylation types	138
Table 4.5	Linear mixed effects model analysis for disease, tissue, and genotype variables across different samples and 5fC/CaC methylation	142
Table 5.1	*BDNF* rs6265 genotypes, gender, age, and diagnosis across samples	162
Table 5.2	Kruskal-Wallis tests across different brain regions investigating transcript expression in disease states	170
Table 5.3	Kruskal-Wallis tests across disease groups investigating transcript expression in tissue groups	174
Table 5.4	Kendal-Tau correlation analysis of transcript expression data and 5mC/5hmC methylation at *BDNF* AS-rs6265	187
Table 5.5	Kendal-Tau correlation analysis across methylation type, tissue groups, disease groups, and genotypes	191
Table 5.6	Comparison of age and gender distribution across Imperial Brain Bank samples	191
Table 6.1	*APOE* rs7412 and rs429358 genotypes, and ε isoforms for brain samples examined	212
Table 6.2	Major ε Isoforms with their associated rs7412 and rs429358 genotypes	212
Table 6.3	Samples used to assess *APOE* rs7412 5mC and 5hmC methylation across genotype, tissues and disease states	217
Table 6.4	Samples used to assess *APOE* rs429358 5mC and 5hmC methylation across genotype, tissues and disease states	218
Table 6.5	Samples used to assess *APOE* rs7412 5mC, 5hmC, and 5fC/CaC methylation across tissues and disease states	225
Table 6.6	Samples used to assess *APOE* rs429358 5mC, 5hmC, and 5fC/CaC methylation across tissues and disease states	228

List of Figures

Figure 1.1	Pathological progression of Parkinson's disease	11
Figure 1.2	Active demethylation cycle compared to the passive demethylation cycle	23
Figure 1.3	Summary of all conversions techniques to assess methylation types at base resolution	36
Figure 1.4	Structure of *BDNF* referencing all transcripts and anti-sense transcripts	46
Figure 1.5	*BDNF* antisense vs sense transcript expression	50
Figure 3.1	Pipeline of methods and techniques used to assess DNA methylation and gene expression at risk epialleles	102
Figure 3.2	Bisulphite, oxibisulphite and MAB conversion methods	108
Figure 3.3	A gel scan of digested and undigested PCR products	108
Figure 3.4	Validation of methylation assisted bisulphite sequencing using λ phage DNA	114
Figure 4.1	Workflow for the analysis of *BDNF* rs6265 antisense complementary CpG methylation and *BDNF* transcript abundance	126
Figure 4.2	5mC or 5hmC methylation at AS-rs6265 in carrier and non-carrier groups	135
Figure 4.3	5mC and 5hmC methylation across AS-rs6265 carrier/non-carrier groups and tissues and disease states	139
Figure 4.4	5mC, 5hmC, and 5fC/CaC methylation across AS-rs6265 C homozygote tissues	143
Figure 4.5	5mC, 5hmC, and 5fC/CaC methylation at AS-rs6265 across C homozygote disease states	144
Figure 5.1	Diagram showing sense and antisense transcripts at *BDNF* variant rs6265 with a major G allele genotype	161
Figure 5.2	Correlation analysis of *BDNF* sense/mRNA, AS-201 and AS-204 transcript abundance	162
Figure 5.3	Differences between *BDNF* sense transcript expression across brain regions	164
Figure 5.4	Differences between *BDNF* AS-201 transcript expression across brain regions	165
Figure 5.5	Differences between *BDNF* AS-204 transcript expression across brain regions	166

Figure 5.6	Differences in expression in *BDNF* AS-201 transcripts across disease states.	168
Figure 5.7	*BDNF* AS-201 transcript abundance across disease states within the frontal cortex	171
Figure 5.8	*BDNF* AS-204 transcript abundance across disease states within the frontal cortex	172
Figure 5.9	Expression of *BDNF* AS-201 transcript in tissue types of unaffected control samples	177
Figure 5.10	Expression of *BDNF* sense transcript in tissue types of Mild Cognitive Impairment samples	178
Figure 5.11	Expression of *BDNF* AS-201 transcript in tissue types of Mild Cognitive Impairment samples	179
Figure 5.12	Expression of *BDNF* AS-204 transcript in tissue types of Mild Cognitive Impairment samples	180
Figure 5.13	Expression of *BDNF* sense transcript in tissue types of Parkinson's disease samples	182
Figure 5.14	Expression of *BDNF* AS-204 transcript in tissue types of Parkinson's disease samples	183
Figure 5.15	Expression of *BDNF* sense transcript in tissue types of dementia with Lewy bodies (DLB) samples	185
Figure 5.16	Relationship between 5mC and 5hmC methylation levels at the AS-rs6265 epiallele and sense, AS-201, and AS-204 transcript abundance	187
Figure 5.17	Significant correlations between transcript abundance, methylation, tissue types, disease states, and genotypes at AS-rs6265	192
Figure 5.18	Scatterplot matrix showing the relationship between data transcript abundance and 5fC/CaC levels at AS-rs6265	194
Figure 5.19	Correlations between AS-rs6265 5fC/CaC methylation and *BDNF* transcripts	194
Figure 6.1	The structure of *APOE* and the surrounding sequence of rs7412 and rs429358	213
Figure 6.2	5mC and 5hmC methylation at *APOE* rs7412 in heterozygotes (C/T) and C homozygotes (C/C)	219
Figure 6.3	5mC and 5hmC methylation across *APOE* rs429358 heterozygotes (T/C), and C homozygote (C/C)	219
Figure 6.4	5mC and 5hmC methylation across tissue groups and genotype at rs7412	221

Figure 6.5	5mC and 5hmC methylation across tissue groups and genotype at rs429358	221
Figure 6.6	5mC and 5hmC methylation across disease groups and genotype at rs7412	223
Figure 6.7	5mC and 5hmC methylation across disease groups and genotype at rs429358	223
Figure 6.8	5mC, 5hmC and 5caC/5fC methylation levels across tissues at rs7412	226
Figure 6.9	5mC, 5hmC and 5caC/5fC methylation levels across tissues at rs429358	229
Figure 6.10	5mC, 5hmC, 5fC/CaC methylation type in disease states at rs7412	231
Figure 6.11	5mC, 5hmC, 5fC/CaC methylation in disease states at rs4293589	233
Figure 6.12	5mC and 5hmC methylation at APOE rs7412 and rs429358 epialleles and ε isoforms	235
Figure 6.13	5mC and 5hmC methylation at rs7412 and rs429358 sites compared against ε isoform split by disease states	237
Figure 6.14	Combined median and ranges of 5mC or 5hmC methylation from both rs7412 and rs429358 epialleles split by ε isoform	239

Abbreviations

(O)BS	Bisulphite and Oxibisulphite conversion
˚C	Degrees Celsius
µl	Microlitre
µM	Micromolar
4mC	N4-methylcytosine
5(h)mC	5mC and 5hmC
5CaC	5-Carboxylcytosine
5fC	5-Formylcytosine
5gmC	β-Glucosyl-5-Hydroxymethylcytosine
5hmC	5-Hydroxymethylcytosine
5mC	5-Methylcytosine
AChE	Acetylcholinesterase
AD	Alzheimer's Disease
AID	Activation Induced Deaminases
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
APOBEC	Apolipoprotein B mRNA editing Enzyme, Catalytic polypeptide-like enzymes
APOE	Apolipoprotein E
APOER2	Apolipoprotein E Receptor 2
ARC	Cytoskeleton-associated protein
ASD	Autism Spectrum Disorder
ASM	Allele Specific Methylation
AS-rs6265	Antisense *BDNF* rs6265 CpG
ATP	Adenosine Triphosphate
BDNF	(mature) Brain Derived Neurotrophic Factor
BER	Base Excision Repair
bp	Base-pair
BS	Bisulphite
Ca^{2+}	Calcium ions
cAMP	Cyclic adenosine mono-phosphate
CaMAB-Seq	5CaC Methylation Assisted Bisulphite Sequencing
cAMPRE	cAMP regulatory element
CER	Cerebellum
CGI	CpG Islands
CHST3	Carbohydrate Sulphotransferase 3
CI	Confidence Interval
COMT	Catechol-O-Methyltransferase
CpG	Cytosine Guanine dyad
CRE	Cis-acting Regulatory Element
CREB	cAMP response element binding protein
CRF	Clinical Research Facility
CSF	Cerebrospinal Fluid
DA	Dopamine

DD-PCR	Digital-Droplet PCR
DLB	Dementia with Lewy Bodies
DMR	Differentially Methylated Region
DMSO	Dimethyl Sulphoxide
DNA	Deoxyribonucleic Acid
DNMT	DNA Methyltransferase
DNMT3L	DNA Methyltransferase 3 Like
dNTP	Deoxyribonucleotide Triphosphate
DR2	Dopamine Receptor like 2
dsDNA	double stranded DNA
EDC	1-ethyl-3-[3-dimethylaminopropyl]-carbodiimide hydrochloride
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-Linked Immunosorbent Assays
EOPD	Early Onset Parkinson's Disease
EPPK1	Epiplakin 1
EZH2	Zeste Homologue 2
ε	epsilon
FC	Frontal Cortex
FFPE	Formalin Fixed Paraffin Embedded
FGFR3	Fibroblast growth factor receptor 3
FTO	Fat mass and obesity-associated protein
GAB2	GRB2 Associated Binding Protein 2
GOF	Gain of Function
GWAS	Genome Wide Association Studies
H3K27me	Methylated histone H3 tail
HBA2	Haemoglobin A2
HDAC	Histone Deacetylases
HIP	Hippocampus
HPLC-UV	High Performance Liquid Chromatography Ultra-Violet
HPLC-MS	High Performance Liquid Chromatography Mass-Spectrometry
HTR6	5-Hydroxytryptamine Receptor 6
ICR	Imprinting Control Regions
IGF2R	Insulin Like Growth Factor 2 Receptor
IHC	Immunohistochemistry
IL-6	Interleukin-6
IQ	Intelligence Quotient
LC-MS/MS	Liquid Chromatography Mass-Spectrometry
LDLR	Low Density Lipoprotein Receptors
lncRNA	Long non-coding RNA
LOF	Loss of Function
LOI	Loss of Imprinting
LTP	Long Term Potentiation
M.SssI	CpG Methyltransferase
m^6A	RNA N^6-methyladenosine
m^6Am	RNA N^6,2'-O-dimethyladenosine

m^6dA	DNA N^6-methyladenosine
MAB-Seq	Methylation Assisted Bisulphite Sequencing
MAF	Minor Allele Frequency
MBD	Methyl-Binding Domain
MCF2L	MCF.2 Cell Line Derived Transforming Sequence Like
MCI	Mild Cognitive Impairment
MECP1/2	Methyl CpG binding protein 1/2
meQTL	Methylation Quantitative Trait Loci
mg	Milligrams
$MgCl_2$	Magnesium Chloride
mRNA	Sense/messenger RNA
NAV2	Neuron navigator 2
NECTIN2	Nectin cell adhesion molecule 2
NeuN+	Neuronal Nuclei positive
ng	Nanograms
ng/µl	Nanogram $Microlitre^{-1}$
NGF	Nerve Growth Factor
NIEL3	Nei Like DNA glycosylase 3
nm	Nanometres
NT3/4	Neurotrophin 3/4
NTs	Neurotrophins
OBS	Oxibisulphite
OD	Optical Density
OR	Odds Ratio
PAGE	Polyacrylamide gel electrophoresis
PARK 1-4	Parkin 1-4
PCR	Polymerase Chain Reaction
PD	Parkinson's Disease
PDE2A	Phosphodiesterase 2A
pNTs	proneurotrophins
PRC2	Polycomb Repressive Complex 2
PINK1/PARK6	PTEN-Induced Kinase 1
Pro-BDNF	BDNF Precursor Molecule
PSEN1/2	Presenilin
PWS	Prader-Willi Syndrome
QTL	Quantitative Trait Loci
RELN	Reelin
RNA	Ribonucleic Acid
ROS	Reactive Oxygen Species
SAM	S-Adenosyl Methionine
SCZ	Schizophrenia and Schizoaffective disorder
SD	Standard Deviation
SE	Standard Error
SMRT	Single-Molecule Real-Time sequencing
SNCA	α-Synuclein
SNCAIP	Synphilin-1 Alpha-Synuclein-Interacting Protein

SNRP	Small Nuclear Ribonucleoprotein Polypeptide N
SNV	Single Nucleotide Variant
SOX5	SRY-Box 5
SRA	Set & Ring finger Associated protein
ssDNA	Single Strand DNA
STP	Short Term Potentiation
TAB-Seq	TET Assisted Bisulphite Sequencing
TBP	TATA-box Binding Protein
TDG	Thymine DNA Glycosylase
TET1/2/3	Ten-Eleven Translocation methylcytosine dioxygenase 1/2/3
TLC	Thin layer Liquid Chromatography
Tm	Primer Melting Temperature
TNF-α	Tumour Necrosis Factor-α
TOMM40	Translocase of Outer Mitochondrial Membrane 40
TRD	Target Recognition Domain
Trk	Tropomyosine-Related Kinase receptors
RTKs	Receptor Tyrosine Kinase
TSS	Transcriptional Start Site
UCHL1	Ubiquitin carboxyl-terminal esterase L1
UEB	Universal extraction buffer
UHRF1/2	Ubiquitin-like, containing PHD and RING finger domains, 1
VLDLR	Very Low-Density Lipoprotein Receptors
Wnt	Wingless-integrated 1
w/w	Weight per weight concentration
βGT	β-Glucosyltransferase
λ	Lambda

Chapter 1: Thesis Introduction

1.1 The Brain and cognition

The brain is a complex organ comprised of many different cell types which in concert allows the processing of external and internal stimuli and facilitates responses such as changes in behaviour. Estimates indicate that there are approximately 100 billion cells within the human brain, many of which are poorly defined (S. B. Nelson, Sugino, & Hempel, 2006). However, the brain volume and weight decrease dramatically over the age of 40 with ageing, altering cell population composition (Allen & Barres, 2009; Noctor, Martínez-Cerdeño, & Kriegstein, 2007; Peters, 2006). Brain cells vary in phenotype, density, and function which allows the interpretation of sensory information and the storage and recall of memories (Lynch, 2004). As such, the brain exhibits regional specificity in its cell populations, connections and function enabling different regions to perform separate but connected functions facilitating quick parallel cognitive processing.

During memory formation, the brain perceives stimuli which are relayed through multiple pathways to the hippocampus where an engram, the unit of memory, is stored in neural connections (Tonegawa, Liu, Ramirez, & Redondo, 2015). Memories may be explicit, i.e. recall of detail, or implicit i.e. muscle memory (Lynch, 2004). These memories are formed, maintained and lost through physical and temporal alterations in cell signalling at synapses (Bannerman et al., 2012; Morris, Anderson, Lynch, & Baudry, 1986). This may occur through loss or generation of synaptic contacts called synaptic pruning and synaptogenesis respectively, as well as alterations in the frequency of potentiation (T. Xu et al., 2009; Yang, Pan, & Gan, 2009). The frequency of potentiation can vary, i.e. short (STP), and long term potentiation (LTP), both of which are utilised in the processes of memory formation and maintenance (Bannerman et al., 2012). In this way specific connections are gained, maintained and lost in order to alter the connective architecture of the brain. These processes shape and activate neural circuitry in a process known as synaptic plasticity.

Synaptic plasticity is thought to be facilitated by neurotrophins, such as Brain Derived Neurotrophic Factor (*BDNF*) (Binder & Scharfman, 2004). Neurotrophin release has been found to be proportional to neuronal activity (H. Kang & Schuman, 1995; D. O. Wang et al., 2009). Furthermore, neurotrophins are activity dependent synaptic modifiers which modify presynaptic neurotransmitter release, post-synaptic sensitivity, and synaptic morphology as a whole causing persistent changes in synaptic signalling (Schinder & Poo, 2000). In this way, neurotrophins can directly affect potentiation (STP and LTP) and in some cases, as with BDNF, are essential for adequate signalling and synaptic plasticity (Figurov, Pozzo-Miller, Olafsson, Wang, & Lu, 1996; Kovalchuk, Hanse, Kafitz, & Konnerth, 2002; Minichiello et al., 2002; Patterson et al., 1996).

Other genes also influence synaptic plasticity by facilitating neurite outgrowth (J. Kim, Yoon, Basak, & Kim, 2014; R. W. Mahley & Huang, 2012). Apolipoprotein E (*APOE*) is a gene involved in lipid and cholesterol transport and metabolism (Bu, 2009; R. Mahley, 1988) and is expressed in response to neuronal injuries to repair damage and promote neurite outgrowth in synaptic plasticity (R. W. Mahley & Huang, 2012; Nathan et al., 1994). *APOE* contains two well-characterised variant sites rs7412 and rs429358 (Babenko et al., 2018). These variants cause coding changes in the APOE protein which result in different APOE ε isoforms (Babenko et al., 2018; R. W. Mahley, Weisgraber, & Huang, 2009). Responses to neuronal injury and neurite outgrowth have been shown to be augmented by APOE, however the efficacy of this response is regulated by ε isoforms and specific ε isoforms have been heavily associated with cognitive diseases such as Alzheimer's disease (AD) through several proposed pathways (Lane et al., 2009; C. C. Liu, Liu, Kanekiyo, Xu, & Bu, 2013; R. Mahley, 1988; Tsuang et al., 2013).

1.2 Brain Disease

Dysfunction in neuronal circuitry can lead to brain disease phenotypes. Broadly, these brain disease phenotypes fall into two diagnostic fields: neurodegenerative disease such as Parkinson's disease, Alzheimer's disease and dementias; and

neuropsychiatric or neurodevelopmental diseases such as schizophrenia, bipolar disorder, and major depressive disorder (I. Chou & Chouard, 2008; Williams, 2002). However a growing body of research suggests commonalities in the symptoms and causes between these fields, and this has given rise to new models of diagnosis and transdiagnostic symptoms (Husain, 2017).

As medicine progresses and increases the average lifespan, neurological and psychiatric disease contributes to an increasing proportion of disability adjusted life years, i.e. years of life lost and years lived with a disability, and deaths globally (Feigin & Vos, 2019). In 2016, neurological disorders were the leading cause of disability adjusted life years and the second leading cause of death globally (Feigin & Vos, 2019). Of this, Alzheimer's disease and other dementias contributed 10.4% of neurological disability adjusted life years lost (Feigin & Vos, 2019). Many brain diseases are thought to have a heritable component and typically involve dysfunction in the brain through degradation of cells and changes in cellular signalling. This can be caused or driven by genetic mutations of large effect such as loss of function (LOF) or gain of function (GOF) mutations in key genes (Van Den Eeden et al., 2003). As the presence and burden of neurological diseases is likely to increase, it is essential that research aims to determine the molecular causes and potential therapies for these diseases.

1.2.1 Schizophrenia

Schizophrenia is polygenic mental illness that has environmental and developmental/genetic factors (Lewis & Lieberman, 2000; S. H. Schultz, North, & Shields, 2007). There is an average 1% incidence of schizophrenia worldwide, which increases with relatedness to individuals with schizophrenia, indicating incidence is influenced by genetic factors (Gottesman & Wolfgram, 1991). Environmental factors such as autoimmune, traumatic, or toxic insults, severe infection, and stress during gestation and childhood, are all associated with developing schizophrenia, indicating onset is affected by outside factors (Marcelis et al., 1998). Diagnostic criteria for schizophrenia are presentation of characteristic symptoms such as delusions,

hallucinations, disorganised speech and/or behaviour, catatonic behaviour, and negative symptoms (e.g. alogia, avolition)(Washington D.C.: American Psychiatric Association, 2000).

It has been determined that the psychotic symptoms in schizophrenia are at least in part due to the over-stimulation of dopaminergic transmission, and abnormal dopamine storage, transport, release, and reuptake are proximal causes to these symptoms (Laruelle, Abi-Dargham, Gil, Kegeles, & Innis, 1999; Laruelle et al., 1996). NMDA receptor hypofunction also acts to increase overall dopamine release through decreasing corticofugal inhibition of dopaminergic neurons (Kegeles et al., 2000). *BDNF* has also been associated with schizophrenia in mice where TrkB was deleted in parvalbumin-expressing fast-spiking interneurons, and glutamatergic and GABAergic signalling was impaired in the hippocampus resulting in schizophrenia like symptoms (K. Zheng et al., 2011). Genetic association of the Val66Met rs6265 variant with schizophrenia has been controversial, with some studies finding associations and some not (Chao, Kao, & Porton, 2008; Naoe et al., 2007; Zakharyan & Boyajyan, 2014).

1.2.2 Autism spectrum disorders

Autism spectrum disorders are behavioural disorders that are characterised by reduced ability in communication, restricted interests and repetitive behaviours (Washington D.C.: American Psychiatric Association, 2000). Animal models of ASD show abnormalities in connectivity in the cerebellum (Courchesne et al., 1994; Santangelo & Tsatsanis, 2005), frontal cortex (Santangelo & Tsatsanis, 2005; Zilbovicius et al., 1995), hippocampus (Piven, Bailey, Ranson, & Arndt, 1998; Santangelo & Tsatsanis, 2005), and amygdala (Abell et al., 1999; Santangelo & Tsatsanis, 2005). However the mechanisms driving these changes are undetermined (Santangelo & Tsatsanis, 2005). *BDNF* is known to be key in neuronal development (Anita E. Autry, 2012) however, different studies have observed both heightened (Meng et al., 2017) and depleted levels (Katoh-Semba et al., 2007) of serum BDNF in children with ASD compared to controls. Similarly, while BDNF may be contributing

in some way to the development of the ASD phenotype in childhood, the rs6265 Val66Met variant has not shown association with ASD (Meng et al., 2017; Yoo, Yang, Cho, Park, & Kim, 2014), however the variant was associated with "Restricted, Repetitive and Stereotyped patterns of behaviour", a canonical trait of ASD (J. G. Lee et al., 2009).

1.2.3 Mild Cognitive Impairment

Mild cognitive impairment (MCI) is a brain disease characterised by a decline in cognitive faculties such as memory, processing speed, as well as other metrics of general and fluid IQ (Petersen et al., 1999). It is associated with the early stages of other clinical brain diseases such as Parkinson's disease (PD), Alzheimer's disease (AD), and Lewy-Body dementia (DLB). However, it can also manifest through ageing and senescence and is defined as 'the increased dementia risk between ageing without cognitive impairment or clinical dementia', although 5 to 10% of all MCI cases convert into dementias (Stephan, Brayne, Savva, Matthews, & Medical Research Council Cognitive Function and Ageing Study, 2011). Unlike aforementioned brain diseases, MCI itself is not fatal and it is not considered to be a major disease that can be treated, but a symptomatic state associated with senescence and other brain disease states. MCI has been associated with, and may be in part due to, changes at synapses and a fluctuation in specific proteins such as BDNF. Some studies have indicated there is a 34% reduction in BDNF and a 21% reduction in Pro-BDNF expression in the brain in MCI individuals compared with controls (Peng, Wuu, Mufson, & Fahnestock, 2005). In contrast, other studies have also suggested BDNF serum levels have been shown to increase in individuals with MCI (Francesco Angelucci et al., 2010).

1.2.4 Alzheimer's Disease

Alzheimer's disease (AD) is a disease characterised by deterioration in memory, recall and other associated cognitive faculties. Its prevalence increases with age and it is associated with several genetic and environmental risk factors (Blennow, de Leon, & Zetterberg, 2006; Ikeda & Yamada, 2010). The largest risk factor for sporadic

Alzheimer's disease is the *APOE* ε4 isoforms generated by genotype at the rs7412 and rs429358 variant sites (Corder et al., 1993; Green et al., 2009). Genotype at the rs7412 and rs429358 variant sites results in different amino acids in the APOE protein which generates the ε protein isoforms. The minor T allele of rs7412 causes an Arg158Cys amino acid substitution in the primary *APOE* transcript, and the minor C allele at rs429358 causes a Cys112Arg amino acid substitution (Zannis et al., 1982), and these determine protein folding and ε isoform. Additionally, environmental risk factors such as smoking (Durazzo, Mattsson, & Weiner, 2014; Wallin et al., 2017) and comorbid conditions such as type II diabetes (Peila, Rodriguez, & Launer, 2002), increases in adiponectin and insulin, and decreases in leptin, that have typically been associated with hypercholesterolemia (Chakrabarti et al., 2014) are all associated with risk for sporadic Alzheimer's disease. Alzheimer's disease is the most common neurodegenerative disorder (Nussbaum & Ellis, 2003).

Supplementation of vitamins has been shown to be a protective factor in cognitive decline, potentially in the early stages of Alzheimer's disease (De Jager, Oulhaj, Jacoby, Refsum, & Smith, 2012). Studies have indicated that many Alzheimer's disease patients and those with MCI reported significantly lower levels of vitamin E in blood serum compared to health controls (Baldeiras et al., 2008; Browne, McGuinness, Woodside, & McKay, 2019; Doltra et al., 2013; Mangialasche et al., 2012; Mullan et al., 2017). However supplementation of vitamins remains a controversial topic, as some studies suggest no difference in vitamin E blood serum levels (Browne et al., 2019; Charlton, Rabinowitz, Geffen, & Dhansay, 2004; Mas et al., 2006; Von Arnim et al., 2012). It is possible that while vitamin E plasma levels may be a factor in or a marker for Alzheimer's disease and MCI, it is not causative in the development of the disease, or there are other more influential factors acting on plasma vitamin E levels. If and how environmental factors (and potentially methylation) influence disease states are ongoing research topics that require more study before this can be elucidated.

Alzheimer pathology is defined by the excessive production and accumulation of amyloid-β oligomers which aggregate to create plaques causing cell death initiating

within the cortical and hippocampal regions (Murphy, LeVine, & III, 2010; Nussbaum & Ellis, 2003). In addition, hyperphosphorylation of τ (Tau) microtubule binding protein aggregate in neurofibrillary tangles which cause cell death (Alonso, Zaidi, Novak, Grundke-Iqbal, & Iqbal, 2001; Schindowski, Belarbi, & Buée, 2008). The accumulation of amyloid-β plaques has not shown to be load dependent in that individuals with greater, more widespread, plaques may not necessarily have as severe impairment in cognitive faculties compared to someone with fewer plaques (Giannakopoulos et al., 2003). Subsequently some hypothesise that the aggregation of these plaques may be a symptom of the disease, but not the underlying cause of cell death in AD as τ fibril accumulation and cognitive decline continue even after the clearance of plaques in animal models (McDade & Bateman, 2017; L. Wang et al., 2016). This is further supported by research that indicates clearance of the plaques through introduction of a chimeric monoclonal antibody, Solanezumab, destroys plaques but does not relieve the patient of the disease and does not alter the rate of decline in human trials (The Lancet Neurology, 2017). It has also been postulated that while immunotherapies may slow or halt the decline in AD, the therapies come at too late a stage to be markedly beneficial and hence they cannot reverse the damage already done (Murphy et al., 2010). This highlights the importance of the diagnosis and treatment of AD and other cognitive diseases earlier in disease course.

As mentioned previously, an important gene in the development of AD is *APOE*. This gene encodes a protein which facilitates lipid and lipoprotein transport as well as amyloid clearance in healthy ε2 or ε3 *APOE* isoforms. Single Nucleotide Variants (SNVs) and point mutations within this gene alter the efficacy of clearance and hence the chances of developing AD in later life.

As AD progresses there is also a reported fluctuation in the amount of BDNF and Pro-BDNF expressed in the brain, which may contribute to a decrease in cognitive faculties through synaptic loss. Some research has reported Pro-BDNF and BDNF levels decreasing (J. G. Lee et al., 2009; Peng et al., 2005), while others have indicated BDNF increases in patients with AD (Francesco Angelucci et al., 2010; Faria et al., 2014). It has been speculated that the increase of BDNF in AD patients is in part due

to an immune response to amyloid-β plaque pathologies, and other sources of inflammation in the brain (Faria et al., 2014). However, this occurs in a region specific manner with BDNF levels increased in the temporal cortex in AD patients (J. Lee et al., 2005).

1.2.5 Synucleinopathies

α-synucleinopathies represent a group of neurological pathologies that result in the accumulation and aggregation of the protein α-synuclein, called Lewy bodies, in the cytoplasm of neuronal cells which cause neurotoxicity and eventual cell death (W. S. Kim, Kågedal, & Halliday, 2014). α-synucleinopathies include Parkinson's disease, dementia with Lewy bodies disease and some familial forms of Alzheimer's disease (Crews, Tsigelny, Hashimoto, & Masliah, 2009). In α-synucleinopathies, α-synuclein is misfolded and forms oligomers which aggregate in the cell-body cytoplasm of the neuronal cells (Lippa et al., 1998) and interact with the cell membrane and cause cell neuroinflammation, synaptic dysfunction, and cell death (Anichtchik, Calo, & Spillantini, 2013; W. S. Kim et al., 2014). This aggregate acts as a neurotoxin that has been shown to propagate in a prion-like manner with different aggregation states altering the speed of propagation and transition of aggregates between neurons (J. L. Guo et al., 2013; W. S. Kim et al., 2014).

1.2.6 Parkinson's Disease

Parkinson's disease (PD) is a neurological disease characterised by dysfunction in motor as well as non-motor circuitry (McGregor & Nelson, 2019). It typically manifests in those over the age of 60 and is more common in Hispanic and white-Caucasians comparative to those of Asian or Black/minority ethnic descent (Van Den Eeden et al., 2003). It is also 1.5 times more common in men compared to women (Van Den Eeden et al., 2003). Recently the full spectrum of PD phenotype is being recognised and the neuropsychiatric ramifications of PD. It was previously thought that Parkinson's disease (as many other neurological diseases) is separate and distinct from other cognitive diseases, but over half of PD patients often develop depression, anxiety, or sleep disorders, and a third of PD patients develop dementia

(Anderson, 2004; Langston, 2006; Schrempf, Brandt, Storch, & Reichmann, 2014). This indicates that while the mechanisms by which PD may occur are distinct, they share pathways and comorbidities with other cognitive diseases.

The phenotypical diagnostic criteria of PD include tremor, bradykinesia, and rigidity (K. Chou, 2018). PD is clinically diagnosed as the accumulation of pathological α-synuclein aggregates (Lewy bodies) in neurons which cause dysfunction and eventually cell death (Braak, Ghebremedhin, Rüb, Bratzke, & Del Tredici, 2004). PD has six stages of progression which may be pre-symptomatic or symptomatic in nature and progresses in a predictable pattern as only certain cell types are vulnerable to develop α-synuclein aggregates (Braak et al., 2004). Vulnerable cells typically are projection neurons with long and thin axonal projections often with poor or no myelination and progression normally starts at the dorsal motor nucleus of the vagal nerve (stage 1) (Braak & Del Tredici, 2004; Braak et al., 2004). From here, pathology progresses to the gain setting nuclei (stage 2), the substantia nigra and amygdala (stage 3), mesocortex and thalamus (stage 4), the neocortex (stage 5) and finally the primary and secondary tissues of the neocortex (stage 6) (Braak et al., 2004, 2003) A figure depicting the pathological progression of Parkinson's disease (Figure 1.1).

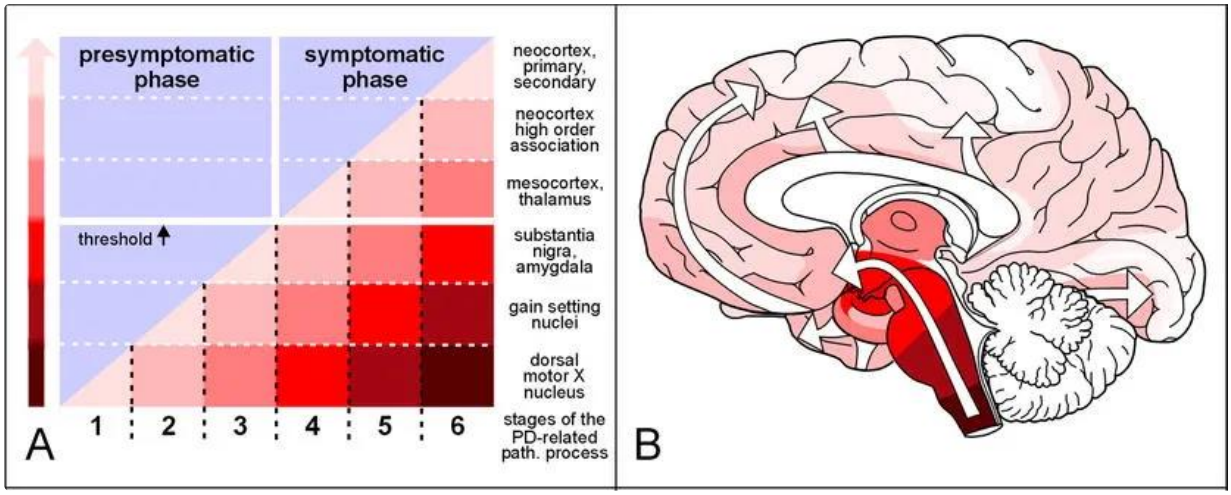


Figure 1.1 The pathological progression of Parkinson's disease. A) Progression from stages 1 to 6 starts in the presymptomatic phase until a threshold is met, after which the symptomatic phase begins. The intensity of the colour of each region is indicative of the severity of pathology in those regions as the disease progresses. B) A panel showing the progressing pathological progress in order. The shading corresponds to the regions shown in panel A when at stage 6. This figure was taken from (Braak et al., 2004).

The degeneration of substantia nigra and amygdala in Parkinson's disease leads to dopamine deficiencies, and case control studies have determined deficiencies as a factor associated with denervation in PD, independent of age (Scherman et al., 1989). Neurons of the substantia nigra secrete dopamine, and in Parkinson's disease these dopaminergic neurons degenerate and result in a dopamine deficiency (German, Manaye, Sonsalla, & Brooks, 1992). It is this denervation that results in the cardinal Parkinsonian traits observed. It is thought that dopamine acts to regulate and "tune" cortical excitation of striatal neurons, and lack of regulation develops as the Parkinson's phenotype (Nieoullon & Goff, 1992).

However some studies have indicated that while dopamine deficiency may be strongly associated with Parkinson's disease, replacement of lost dopamine in these phenotypes is associated with complications such as dyskinesias (involuntary muscle movements) (Ferrazzoli et al., 2016; Obeso, Rodríguez-Oroz, & Zamarbide, 2005), motor fluctuations ("wearing off" of beneficial effects) (Ferrazzoli et al., 2016; Kikuchi, 2007), and addiction-like behaviours (Ferrazzoli et al., 2016; Giovannoni, O'Sullivan, Turner, Manson, & Lees, 2000). It is therefore important that preventative measures to stop Parkinson's phenotype developing are investigated in tandem with restorative therapies such as treatment with L-DOPA.

Parkinson's disease appears to manifest more commonly in certain ethnicities and genders which suggests a strong genetic contribution to the expression of the disease. Mutations in the gene PTEN-induced kinase 1 (*PINK1* also known as *PARK6*) are suggested to result in mitochondrial dysfunction and excessive oxidative stress to neurons which causes cell death and has been associated with increased risk of familial PD (Klein & Westenberger, 2012). Through a different pathway, mutations in the α-synuclein gene (*SNCA*) cause an abnormal protein conformation which causes the α-synuclein to aggregate in bodies known as Lewy bodies (Polymeropoulos et al., 1996). This leads to dopaminergic cell death, and hence mutations in *SNCA* have been identified as causing genetic risk for PD (Polymeropoulos et al., 1996). Similarly, mutations in the *Parkin (PARK2)* gene are the most common cause for early onset PD (EOPD), with an estimated 77% of PD cases under 30 years of age result from *PARK2*

mutations (Lücking et al., 2000). Parkin proteins facilitate the ubiquitination of α-synuclein which in turn stops them from aggregating into Lewy bodies, and mutations in *Parkin* disrupt protein action and lead to α-synuclein aggregation, and cell death (Klein & Westenberger, 2012). PD has also tentatively been associated with mutations within the *APOE* gene. The risk for PD has been shown to increase with the different *APOE* Ɛ isoforms as with other synucleinopathies (Tsuang et al., 2013).

Recent research has also suggested that α-synuclein may be transported from the gut to the brain via the vagus nerve in models of sporadic forms of PD (S. Kim et al., 2019). Moreover, the dorsal motor nucleus of the vagal nerve is normally the starting region or stage 1 in PD progression (Braak et al., 2004) and presymptomatic or prodromal manifestations of the disease include gastrointestinal dysfunction which indicates that the gut may influence PD onset and progression in some way (Bridi & Hirth, 2018; Fitzgerald, Murphy, & Martinson, 2019). Several studies have found significant differences between the composition of gut microbiota in PD patients and healthy controls (Fitzgerald et al., 2019; Keshavarzian et al., 2015; Scheperjans et al., 2015; Scheperjans, Derkinderen, & Borghammer, 2018; Unger et al., 2016).

1.2.7 Dementia with Lewy Bodies

Dementia with Lewy bodies (DLB) also known as Lewy body dementia, is an α-synucleinopathy and form of dementia that closely resembles PD and AD. It is also associated with the accumulation of α-synuclein aggregates (Lewy bodies) in the brain (Jellinger & Korczyn, 2018; W. S. Kim et al., 2014). It is primarily differentiated from other synucleinopathies by fluctuations in cognition whereby patients enter phases of unresponsiveness and return to normal attention after a period of time (Mckeith et al., 2017). It is thought that these fluctuations in cognition are caused by thalamic damage and cholinergic imbalances (Delli Pizzi et al., 2015). DLB also typically has at least one cardinal Parkinsonism trait during progression, which may include resting tremor, rigidness or bradykinesia that increase severity over the years (Ferman et al., 2011; W. S. Kim et al., 2014; Mckeith et al., 2017; Weisman et al., 2007). However, a quarter of DLB patients may never develop Parkinsonism traits

(W. S. Kim et al., 2014; Mckeith et al., 2017) and similar to PD, DLB patients may experience visual hallucinations (Jellinger & Korczyn, 2018). Mutations in *APOE* have been associated with risk to develop DLB (Bras et al., 2014; Lane et al., 2009) and like AD and PD, *APOE* Ɛ isoforms have been associated with the risk of developing DLB (Tsuang et al., 2013).

1.3 Epigenetics, DNA methylation and epialleles

Research has shown that many psychopathologies are influenced by genetic and environmental components (Burmeister, McInnis, & Zöllner, 2008; Strathearn, 2009). One way in which the environment may influence phenotype is through regulating transcription or translation (Dolinoy, Huang, & Jirtle, 2007; Eccleston, DeWitt, Gunter, Marte, & Nath, 2007; Seong, Li, Shimizu, Nakamura, & Ishii, 2011; H. Wang et al., 2012). Epigenetics is a term which refers to heritable processes which regulate gene function or expression that does not alter the underlying DNA base sequence (Russo, Martienssen, & Riggs, 1996). Epigenetic processes have been associated with biological ageing, gene expression and associated phenotypes (Horvath, 2013; Law & Jacobsen, 2010). Typically, the most studied epigenetic modifications to date are histone modifications and DNA methylation.

Histones are post-translationally modified which alters gene expression through coiling of chromatin and nucleosomes which in turn affects the abundance of transcript and by extension expression (Duncan, Gluckman, & Dearden, 2014; Marmorstein & Trievel, 2009; Peterson & Laniel, 2004; C. Wu, Wong, & Elgin, 1979). Histone modifications are known to modulate the accessibility and expression of DNA through the folding and coiling of DNA (Duncan et al., 2014; Marmorstein & Trievel, 2009; Peterson & Laniel, 2004). Acetylation of histones facilitates transcription, and deacetylation of histones further recruits DNA methylation and reduces transcription (Bannister & Kouzarides, 2011; Clayton et al., 2006; Xhemalce, Dawson, & Bannister, 2011). In this way, histone modifications may store temporally separated signalling events that modulate the expression and phenotype of cells (Badeaux & Shi, 2013).

1.3.1 DNA 5mC methylation

DNA modifications in mammals typically take the form of methylation of cytosines at a cytosine-guanine dyad (CpG site) along DNA sequences. The most common DNA methylation state is 5-methylcytosine (5mC) where DNA methyltransferases (DNMTs) methylate the 5th carbon of the cytosine using S-Adenosyl methionine (SAM) as a substrate (Holliday & Pugh, 1975; Riggs, 1975). CpG methylation is thought to comprise 99.98% of methylation in human cell lines (Lister et al., 2009). More rarely, methylation may also occur at non-CpG sites such as CpH, CHG, and CHH sites where H may be cytosine, thymine or adenosine bases. Non-CpG methylation has been shown to be site specific and conserved in human embryonic stem cells indicating that it has a function in development (Ramsahoye et al., 2000; Woodcock, Crowther, & Diver, 1987).

De novo DNA methyltransferases, DNMT3a and DNMT3b methylate unmodified cytosines are known as methylation "writers" (Gowher, Liebert, Hermann, Xu, & Jeltsch, 2005; Li & Zhang, 2014). They are catalysed by DNMT3L which binds to their catalytic domains with its C-terminal domain, increasing activity by 15 fold (Gowher et al., 2005). The newly methylated cytosine is stable until it is demethylated through oxidation, DNA repair, or the DNA is replicated (Li & Zhang, 2014). During replication, only one strand will retain the original methyl mark while the other strand is unmethylated meaning the newly synthesised DNA is hemimethylated (Li & Zhang, 2014). DNA methyltransferase 1 (DNMT1) maintains the methylation by methylating the complementary CpG sites in hemimethylated DNA and has a preference to binding hemimethylated or unmethylated DNA (Holliday & Pugh, 1975; Li, Bestor, & Jaenisch, 1992; Okano, Xie, & Li, 1998; Robertson & Wolffe, 2000). If DNA methylation is not conserved through DNMT1 mediated maintenance it leads to loss of methylation over time in a process known as passive demethylation which may be fatal in certain stages of development (Li et al., 1992; Li & Zhang, 2014).

5mC is suggested to be repressive to gene expression (Deaton & Bird, 2011). 5mC motifs are bound by several proteins that facilitate transcriptional repression of DNA,

such as the Methyl-Binding Domain (MBD) proteins, Kaiso and Kaiso-like proteins, and Set & Ring Finger Associated (SRA) proteins (Defossez & Stancheva, 2011; Sasai, Nakao, & Defossez, 2010). These proteins are known as methylation "readers". Moreover, methylation of promoters and exonic/intronic regions surrounding CpG islands is associated with transcriptional repression of those genes (Deaton & Bird, 2011). In MBD protein repression, CpG methylation is bound by the Methyl CpG binding (MECP1, MECP2) proteins and facilitates transcriptional repression through the recruitment of Sin3 and Histone deacetylase (Boyes & Bird, 1991; Jones et al., 1998; Nan et al., 1998; Robertson & Wolffe, 2000). Sin3 is a corepressor that is bound by MECP1/2 protein, and that facilitates the deacetylation, chromatin compaction, and transcriptional silencing of a gene segment (Jones et al., 1998; Nan et al., 1998; Robertson & Wolffe, 2000). While 5mC CpG methylation is mainly repressive, studies have indicated that it may also facilitate transcriptional expression for some genes during development and stem cell differentiation (D. H. Yu et al., 2013).

While there is extensive research on CpG methylation of the sense strand, there is substantially little research investigating methylation of the antisense CpGs. As discussed above, DNMT1 acts to 5mC methylate CpG sites after strand replication to avoid passive dilution of methylation, which may be fatal (Holliday & Pugh, 1975; Li et al., 1992; Li & Zhang, 2014; Okano et al., 1998; Robertson & Wolffe, 2000). This suggests that the majority of complementary antisense CpG sites are methylated where their sense CpG counterparts are methylated. Antisense expression can lead to sense strand transcriptional or translational repression, by facilitating methylation of sense strands and histones (Barman, Reddy, & Bhaumik, 2019; Tufarelli et al., 2003), or forming RNA duplexes and acting as cis-regulatory elements (CREs) (Modarresi et al., 2012; Pruunsild, Kazantseva, Aid, Palm, & Timmusk, 2007).

It is known that CREs are crucial in the activity dependent expression of *BDNF* transcripts and that some of them are natural anti-sense transcripts of the *BDNF* gene (W. G. Chen et al., 2003; Modarresi et al., 2012; Varela, Roberts, Andaloussi, & Wood, 2012). Moreover, sense-antisense duplexes not only inhibit translation of the mRNA and sequester the transcript but may also increase the sensitivity of sense transcript

to degradation unless sequestered (W. G. Chen et al., 2003; Modarresi et al., 2012; J. Shi, 2015). In this way, transcriptional regulation of antisense transcripts through methylation may, in turn, regulate the methylation and transcription of sense transcripts.

Knowing that CpG 5mC methylation has a transcriptionally modular effect, this suggests a role for 5mC in transcriptional dysregulation during development and disease. It is known that methylation is key in development and developmental changes in methylation ensure totipotency (Baubec et al., 2015; Pujadas & Feinberg, 2012). Vertebrate gametes carry their own specific methylation patterns which are different from somatic cells and the sperm methylome has been shown to be inherited by the embryo in zebrafish and mice in genetic imprinting (Jiang et al., 2013; Smallwood et al., 2011). In addition, genes have been shown to be differentially methylated depending on their parental origins repressing the maternal or paternal alleles in genetic imprinting (Ferguson-Smith, Sasaki, Cattanach, & Surani, 1993; Li, Beard, & Jaenisch, 1993). This indicates that the gamete DNA methylome is crucial for embryo development. However, human embryos undergo genome wide demethylation prior development into a foetus (H. Guo et al., 2014; Z. D. Smith et al., 2014). As mentioned previously, mice embryos that have knock-out or repressed levels of DNMTs do not develop to term (Razin & Shemer, 1995) indicating that methylation is essential to embryo development and cell-fate in mice and potentially most mammals (Baubec et al., 2015).

DNA methylation has long been suggested to underlie changes in memory and cognitive performance in a general and pathological capacity (Day & Sweatt, 2010, 2011b; Flitton et al., 2019; Griffith & Mahler, 1969). 5mC has been proposed as a molecular “ticket” by which a memory engram can be stored by the cell within the DNA (Day & Sweatt, 2010; Griffith & Mahler, 1969). In this way information can be passed between generations of cells, retaining the methylation and engram long term (Day & Sweatt, 2010). Further to this, inhibition of the DNMT protein blocks memory formation in contextual fear conditioning studies in rat models which indicates that methylation is required for memory formation and consolidation (Day

& Sweatt, 2010; Feng et al., 2010; Lubin, Roth, & Sweatt, 2008; Miller et al., 2010). Inhibition of DNMT in the hippocampus has also been shown to cause hypomethylation and impaired memory recall (Lubin et al., 2008) as has DNMT inhibition in the prefrontal cortex (Miller et al., 2010). This indicates that engram formation, maintenance, and recall have different neuronal circuitry (Day & Sweatt, 2011b). Both active methylation and demethylation of specific genes are required for engram formation and retention (Day & Sweatt, 2010). General DNMT inhibition results in a decrease in methylation at *BDNF* and other genes associated with cognitive plasticity like *RELN* (Levenson et al., 2006; Miller & Sweatt, 2007) the protein of which inhibits LTP and therefore cognitive and synaptic plasticity.

DNA methylation has also been implicated in the regulation of alternative splice variants for genes through the repression of one transcript and the enhancement of another (Nilsen & Graveley, 2010). DNA methylation differentially expresses *BDNF* transcripts through methylation across *BDNF1* promoters in mice, typically repressing expression with 5mC methylation (Levenson et al., 2006). As *BDNF* has been shown to be important in hippocampal LTP and memory formation in mammals, regulation of *BDNF* with methylation may impact these mechanisms (Michael F. Egan et al., 2003; Schinder & Poo, 2000) Additionally, genetic imprinting, whereby maternal or paternal copies of genes are regulated and expressed differently, regulates the expression of imprinted parental genes in part through DNA methylation (Day & Sweatt, 2011a). Therefore, DNA methylation may affect memory storage and consolidation and plasticity in multiple separate pathways, and variations in the levels of methylation may alter memory formation and cognition.

1.3.2 Dysregulation of 5mC methylation, methylation turnover, and disease

Dysregulation of methylation in the form of mutations in key genes involved in methylation (*MECP2*, *DNMTs*, and Ten-eleven translocation methylcytosine dioxygenase; *TET*) may lead to disease phenotypes such as cancer, metabolic disease, immune disease, and cognitive diseases. The same cellular plasticity afforded by DNA methylation may enable cancerous growth through the dysregulation of methylation

and expression of oncogenes and tumour suppressor genes (Pujadas & Feinberg, 2012). Hypomethylation and loss of 5mC is aberrant across the genomes of cancerous cells and may primarily be attributed to demethylation of repeat regions of the human genome which leads to genomic instability (Robertson, 2005; Yoder, Walsh, & Bestor, 1997). In addition, hypermethylation in the CpG islands of tumour suppressor genes results in transcriptional repression and cancer formation (Herman & Baylin, 2003; Issa, 2004; Robertson, 2005). Furthermore, disruption of genomic imprinting (loss of imprinting; LOI) through changes in methylation can cause alterations in allele specific expression leading to genetic disease and cancers (Feinberg, Cui, & Ohlsson, 2002; Robertson, 2005).

DNA methylation has been linked to cognition and ageing, where quantification of methylation at specific CpG sites can be a good indication of the age of a cell, tissue or individual (Horvath, 2013; Z. Jin & Liu, 2018) and cognitive impairment associated ageing may therefore be an epigenetic process (Ciccarone, Tagliatesta, Caiafa, & Zampieri, 2018). Moreover, methylation and cognitive impairment have been associated with Huntington's disease (Horvath et al., 2016), Parkinson's disease (Horvath & Ritz, 2015), and Alzheimer's disease (Z. Jin & Liu, 2018; Levine, Lu, Bennett, & Horvath, 2015). Brain tissues show unique methylation profiles and many forms of dysregulation of methylation may have an impact on cognitive faculties (Bradley-Whitman & Lovell, 2013; Horvath, 2013; Lister & Mukamel, 2015; Masliah, Dumaop, Galasko, & Desplats, 2013).

One of the most established cognitive disease phenotypes caused by dysregulation in methylation or disruption of genomic imprinting is Rett syndrome, a neurological disorder of which 90% of cases, often with the most severe phenotypes, are attributed to loss of function mutations within *MECP2*. Mutations in *MECP2* in turn, inhibit methylation facilitated transcriptional repression (Amir et al., 1999; Liyanage & Rastegar, 2014; Rousseaud, Delépine, Nectoux, Billuart, & Bienvenu, 2015; Weaving et al., 2003). This indicates that the disruption of epigenetic regulation facilitated through methylation by *MECP2* is deleterious if not fatal in some phenotypes. A second methylation related cognitive condition is Prader-Willi

syndrome (PWS). PWS is associated with mental retardation, obesity, hyperphagia, and behavioural issues in infants (Goldstone, 2004). It arises due to microdeletions in the imprinting control regions (ICRs) of the *small nuclear ribonucleoprotein polypeptide N* (*SNRP*) gene (Nicholls & Knepper, 2001; Sutcliffe et al., 1994). These deletions alter the parentally specific methylation of the ICR by removing and creating CpG sites across the ICR, which leads to transcript dysregulation and PWS phenotype (Nicholls & Knepper, 2001; Robertson, 2005).

Because of the unique role DNA methylation appears to play in tissue specific gene expression, DNA methylation has now been linked with many common complex cognitive diseases. However, determining causation or interpretation of findings have been difficult. For instance, Alzheimer's disease studies have shown that there is broad genome wide hypermethylation in murine cerebral endothelial cell lines with AD associated mutations which increase amyloid-β (K.-L. Chen et al., 2009; Sanchez-Mut & Gräff, 2015; Sung, Choi, Ahn Jo, Oh, & Ahn, 2011). However, similar AD associated mutations in mice result in genome wide hypomethylation and result in similar phenotypes (Cong, Jia, Qin, Ren, & Sun, 2014; Sanchez-Mut & Gräff, 2015). These differences may be due to variation in methylation on an individual, tissue, cell type, disease, age, and gender basis.

This is further illustrated by contradicting findings on humans, where quantifying 5mC from human brain samples has indicated regional variations in methylation. For instance, studies report 5mC hypomethylation in the entorhinal cortex (Mastroeni et al., 2010) and hippocampus (Chouliaras et al., 2013) of AD patients compared to control individuals. Other studies reported no difference in 5mC methylation in entorhinal cortex (Lashley et al., 2015) and hypermethylation in the frontal cortex (Coppieters et al., 2014) and hippocampus (Bradley-Whitman & Lovell, 2013) of AD individuals. This highlights the difficulty in determining the function of DNA methylation in these pathologies as many case populations appear to have contrasting levels of methylation, due to a multitude of factors.

Not only is the regulation of methylation crucial for heritable diseases but it is also important in non-heritable diseases such as viral infections. The high frequency of repeating elements in the human genome are in part transposable viral sequences that have been silenced using methylation (de Koning, Gu, Castoe, Batzer, & Pollock, 2011). In this way DNA methylation and its regulation are crucial to many aspects of human development and conditions.

1.3.3 5hmC and other oxidised forms of methylation

Other oxidised DNA methylation forms are less well documented than the canonical 5mC CpG methylation motif. 5mC methylation can be oxidised by Ten-eleven Translocation (TET) enzymes and converted into 5-hydroxymethylcytosine (5hmC)(Ito et al., 2010; Mamta Tahiliani, 2009). These TET enzymes are known as methylation "erasers". 5hmC is further oxidised by TET enzymes and converted into 5-formylcytosine (5fC) and then further into 5-carboxylcytosine (5CaC) (Hahn, Szabó, & Pfeifer, 2014; Ito et al., 2011). Either 5fC or 5CaC can be glycosylated and excised by Thymine DNA Glycosylase (TDG) in base excision repair (BER) (Hahn et al., 2014; Neri, Incarnato, & Oliviero, 2015; S. C. Wu & Zhang, 2010) back to unmodified cytosine, in a process known as active demethylation. As mentioned previously passive demethylation or dilution of methylation occurs when DNA methylation is not conserved through DNMT1 mediated maintenance (Li et al., 1992; Li & Zhang, 2014) Figure 1.2 illustrates the active methylation cycle as well as passive demethylation.

Alternatively 5mC and 5hmC methylated cytosines may also be demethylated into unmodified cytosines through deamination by activation induced deaminases (AID) and apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like enzymes (APOBEC) which deaminate 5mC/5hmC (Morgan, Dean, Coker, Reik, & Petersen-Mahrt, 2004) (Figure 1.2). However some studies suggest that AID/APOBEC cytosine deaminases have a preference to unmodified cytosine over 5mC methylation and rarely deaminate 5hmC methylation at all (Nabel et al., 2012). Finally, 5hmC methylation can be removed by DNMT3A/3B when S-adenosyl methionine (SAM) is

absent (C.-C. Chen, Wang, & Shen, 2012; Liutkevičiūtė, Lukinavičius, Masevičius, Daujotytė, & Klimašauskas, 2009). Through these processes, unmethylated cytosines can be methylated, oxidised into 3 separate states and finally converted back to unmodified cytosine in active demethylation.

It was previously thought that the other intermediate stages of methylation (5hmC, 5fC, and 5CaC) were transient states as part of a pathway for demethylation. But recent findings have provided strong evidence that 5hmC and 5fC/CaC may have a more specific and permanent role in the regulation of transcription (Hahn et al., 2013, 2014). 5hmC is known to be a stable modification that is abundant in the brain (comprising 0.6% of total nucleotides compared to 5mC comprising 1.6% in Purkinje cells, which is a high ratio of 5hmC:5mC compared to other tissues) and is found almost exclusively at CpG sites (Hahn et al., 2014; Kriaucionis & Heintz, 2009; Lister et al., 2013). 5hmC also inhibits DNMT1 maintenance methyltransferases, resulting in the passive 5mC demethylation of replicating DNA strands and may be important in facilitating global demethylation events found in development (Hahn et al., 2014; Ji, Lin, Song, & Wang, 2014; Valinluck & Sowers, 2007).

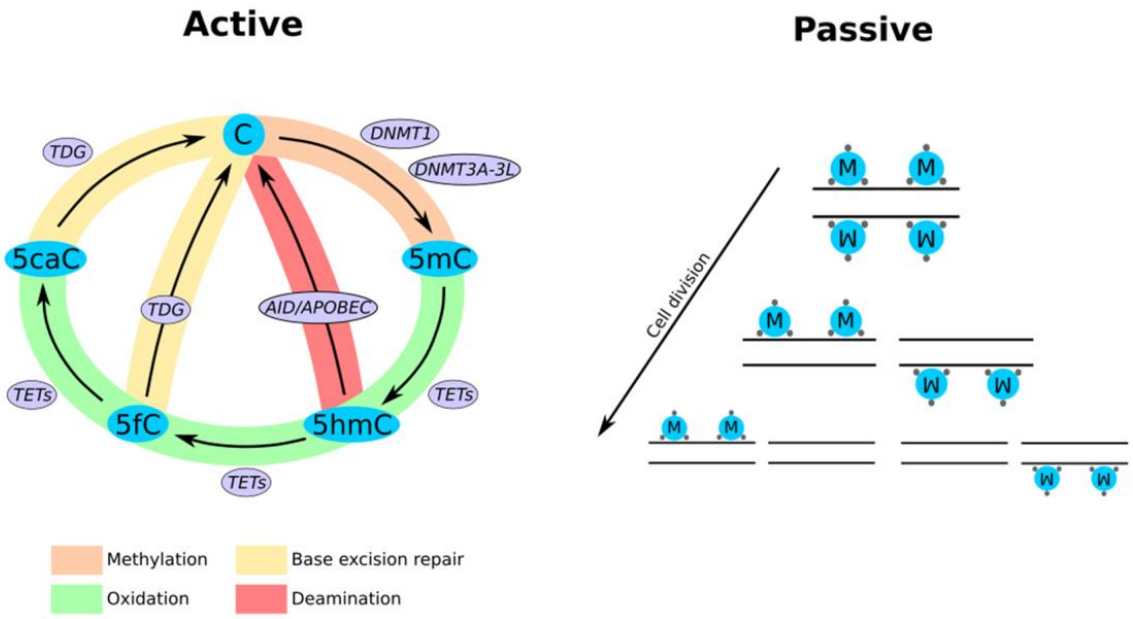


Figure 1.2 Active demethylation cycle compared to the passive demethylation cycle. In the active demethylation cycle unmodified cytosines (C) are methylated by DNMTs into 5mC. These can be oxidised by TET enzymes into derivative methylation states (5hmC, 5fC, 5CaC). 5hmC may be demethylated through deamination (AID/APOBEC), while 5fC/CaC may be demethylated through TDG and the base excision repair pathway. Passive methylation shows the passive dilution of methylation (shown as an M in blue) on a single strand over several generations of cell division. Figure was taken from (Flitton, 2018).

Methylation is primarily complementary on both strands, and additionally most methyl-binding domain proteins bind to both complementary methylated CpG sites on both strands, and "sit" in the DNA major groove (Clouaire & Stancheva, 2008; Hendrich & Bird, 1998). However, the ratio of different types of methylation at these complementary CpG sites is currently understudied. For instance, it is known that 5hmC shows an anti-sense strand preference (M. Yu et al., 2012), and this may change with cell type (Mooijman, Dey, Boisset, Crosetto, & Van Oudenaarden, 2016), and in non-human animals such as mice (C. X. Song et al., 2011). When asymmetrical methylation is measured in some studies, this could in part be due to bisulphite sequencing technologies that are unable to quantify oxidised forms of 5mC such as 5hmC, 5fC, and 5CaC. It is also possible that quantification using technologies that are unable to differentiate between 5mC and 5hmC are overestimating the levels of methylation complementarity at these sites.

Canonically 5mC methylation of promoters and gene surrounding CpG islands is associated with transcriptional repression of those genes (Deaton & Bird, 2011). Methylation of a promoter would typically reduce sense strand or coding transcript expression, and may do this through alternative pathways such as increased anti-sense CRE expression. However few studies have observed the effects of exon strand specific methylation on the levels of transcript expression, and how this contributes to gene regulation. Exon methylation (regardless of strand symmetry) may represses transcription in a strand specific manner- which in turn may associate with the differences observed in anti-sense and sense strand expression. This study aims in part to quantify the ratio of methylation types at key epialleles on the antisense strand to determine if they are associated with strand specific expression of transcripts. This ratio on different strands could be used as a potential biomarker, allowing for quantification of methylation, genotype associated with disease, and state of gene transcription.

5hmC is accumulated in the brain over life span (C. X. Song et al., 2011; Szulwach et al., 2011). 5hmC has also been shown to be bound by the methylation readers Ubiquitin-like, containing PHD and RING finger domains, 1/2 (UHRF1/2) and *MECP2*,

which suggests that 5hmC can interact with proteins and has a function in the brain (Hahn et al., 2013; Mellen, Ayata, Dewell, Kriaucionis, & Heintz, 2012; C. X. Song et al., 2011; Spruijt et al., 2013; Szulwach et al., 2011). While UHRF1 alone preferentially binds to hemi-5mC methylated DNA (Bostick et al., 2007; Otani et al., 2013), UHRF1 and UHRF2 complexes have a preference for hemi-5hmC methylation (Frauer et al., 2011; Vaughan et al., 2018; T. Zhou et al., 2014).

However, 5hmC methylation has also been shown to be inversely correlated with MECP2 as the binding of MECP2 with 5mC hinders oxidation of 5mC into 5hmC, so the more MECP2 is present, typically the less 5hmC is observed (D. Q. Shi, Ali, Tang, & Yang, 2017; Szulwach et al., 2011). Conversion of 5mC to 5hmC methylation through TET enzymes can also affect MBD proteins such as MECP2 binding to DNA (Day & Sweatt, 2011b; Mamta Tahiliani, 2009). 5hmC has also been shown to not be bound by other MBD proteins such as MBD1, MBD2, and MBD4 (S.-G. Jin, Kadam, & Pfeifer, 2010). This indicates that further research may be required before we know the functional purpose of reading 5hmC methylation and the relationship between 5hmC and *MECP2* other MBD proteins.

5hmC methylation appears to alter its modulation of gene expression dependent on cell type (Nestor et al., 2012). It is enriched in active genes and it is inversely correlated with 5mC frequency (Mellen et al., 2012). However, absolute levels of 5hmC are primarily influenced by the cell type (with gene transcription being a secondary factor), and the same genes with the same level of transcription in different cell types may have over a 20 fold difference in 5hmC levels (Nestor et al., 2012). 5hmC is enriched in the intragenic regions of brain specific and neuronal differentiation associated genes which are active, which suggests it enhances transcriptional expression for specific genes (Hahn et al., 2014; S.-G. Jin, Wu, Li, & Pfeifer, 2011; Mellen et al., 2012; C. X. Song et al., 2011). Although 5hmC in gene bodies has also been associated with enhanced transcription of DNA (S.-G. Jin et al., 2011; C. X. Song et al., 2011), 5hmC enrichment at the transcriptional start site (TSS) has been associated with repressed gene transcription (Lister et al., 2013; Wen et al.,

2014; M. Yu et al., 2012). This highlights a location specific effect of 5hmC methylation in modulating expression.

Due to the technological constraints in detection, comparatively little is known about the 5fC and 5CaC methylation states. 5fC and 5CaC are present in the brain in quantities between 10 and 100 times lower than that of 5hmC (Ito et al., 2011; S. Liu et al., 2013) which makes full quantification of the 5fC/CaC methylome difficult. However, as the interactions and functions of the 5mC and 5hmC methylation marks become clearer, more research is aiming to elucidate the roles of 5fC/5CaC. Indeed, mass-spectrometry has suggested that 5fC and 5CaC show far broader protein binding affinity compared to 5hmC, binding with more proteins despite being detected many times lower than 5hmC levels (Ito et al., 2011; Iurlaro et al., 2013; S. G. Jin et al., 2016; Spruijt et al., 2013). 5fC and 5CaC were previously thought of as transient forms of methylation because they were low in quantity. However, studies have shown that TET1 and TET2 enzymes preferentially oxidise 5mC methylation over 5hmC, 5fC and 5CaC (L. Hu et al., 2015). Further, once 5hmC has been established it is not common for it to be further oxidised to 5fC or 5CaC (L. Hu et al., 2015; D. Q. Shi et al., 2017).

Both 5fC and 5CaC reduce polymerase elongation and therefore transcription of DNA (Kellinger et al., 2012) which indicates that they may have a transcriptionally repressive effect similar to that of 5mC. In addition to this, both 5fC and 5CaC also block the activity of the DNMT1 maintenance methylase, therefore facilitating passive demethylation of DNA similar to the effect of 5hmC methylation (Ji et al., 2014). This inhibition of 5mC maintenance may in part be due to 5fC and 5CaC possessing the ability to block complementary methylation at the opposite strand CpG site by altering the orientation of the DNMT1 target recognition domain (TRD) and inhibiting its ability to methylate (Ji et al., 2014). However, contra to this, 5fC methylation has been shown to increase DNMT3A methylase activity at complementary CpG sites on the opposite strand as well as directly adjacent CpG sites (Ji et al., 2014). This suggests that not only does 5fC/CaC methylation show different affinities and functions with different methylation binding readers but also

that 5fC and 5CaC methylation show distinct properties separate from those of 5mC and 5hmC.

5fC has been shown to prime gene enhancer and regulatory elements for activation in mouse embryonic stem cells and is enriched at "poised" enhancers, i.e. gene enhancer regions which are ready to activate (C. X. Song et al., 2013). To further support transcription, accumulation of 5fC in TDG knock out cells recruits the transcriptional coactivator p300 and increases transcription (C. X. Song et al., 2013). This is in contrast to studies that suggested that 5fC reduced polymerase elongation which may show that 5fC not only has its own specific functions with readers but also has a gene regions specific effect by encouraging transcription at enhancer regions and reducing it at the gene body.

In addition to being their own oxidised form of methylation, 5fC and 5CaC can still function as intermediary forms of methylation in the process of BER and demethylation. In this way, 5fC and 5CaC may define which 5hmC methylated CpGs are stable and which are not (Shen et al., 2013; C. X. Song & He, 2013; C. X. Song et al., 2013). This is supported by the fact that while only ~30% of the hydroxy-methylome has 5fC methylation, this accounts for over 90% of all 5fC in mouse embryonic stem cells (Shen et al., 2013; C. X. Song & He, 2013; C. X. Song et al., 2013). This may indicate that 10% of the 5fC methylation observed is non-transient and may serve a function and that a portion of the hydroxy-methylome is oxidised by TET enzymes to maintain 5fC/CaC methylation at enhancer sites to facilitate "poised" state enhancers for genes (Shen et al., 2013; C. X. Song & He, 2013; C. X. Song et al., 2013).

1.3.4 5hmC, 5fC, and 5CaC dysregulation and disease

Similar to 5mC, the total 5hmC levels in brain tissue from individuals with Alzheimer's disease are variable and inconsistent, and one individual may have a higher/lower amount of 5hmC compared to another, despite analysing the same brain region and having the same AD phenotype. For instance the mid frontal and temporal gyri

(Coppieters et al., 2014) and hippocampus (Chouliaras et al., 2013) of Alzheimer's disease brain show higher 5hmC levels than that of control samples from the same regions (Sanchez-Mut & Gräff, 2015). However other studies have shown that the 5hmC methylation levels in the entorhinal cortex and cerebellum (Condliffe et al., 2014), as well as in the hippocampus (Chouliaras et al., 2013) from Alzheimer's individuals are lower than that of healthy control brain samples (Sanchez-Mut & Gräff, 2015). Many studies have shown a correlation between the total amount of 5mC and 5hmC methylation observed in brain regions and markers for Alzheimer's disease such as amyloid-β plaque and neurofibrillary tangle loads in patient tissue samples and cell lines (Chouliaras et al., 2013; Coppieters et al., 2014; Irier & Jin, 2012; Levine et al., 2015; Sung et al., 2011). These studies show that changes in DNA methylation in the brain are responding to (or potentially leading to) a dementia phenotype and dysregulation in the 5mC and 5hmC methylation may be a driving factor. But change in methylation while indicative of dysregulation and disease appears to be distinct and shows variation between individuals (Iwamoto et al., 2011; Lister et al., 2013), and further research will be required before methylation may be considered as a potential marker of disease in AD, PD and DLB diagnoses.

In addition to Alzheimer's disease, Parkinson's disease appears to involve epigenetic regulation. For example, Synphilin-1 Alpha-Synuclein-Interacting Protein (*SNCAIP*) has been shown to be hypermethylated in the cerebral cortex of patients with sporadic Parkinson's disease, when quantified in bisulphite sequencing (Dashtipour et al., 2017). The protein product of *SNCAIP* is Synphilin-1 which is both a constituent of Lewy bodies (Casadei et al., 2014) and interacts with α-synuclein in α-synucleinopathies to inhibit its degradation (Alvarez-Castelao & Castaño, 2011). Additional studies observed hypomethylation in the promoters of *SNCA* and *PARK2* genes in early onset Parkinson's disease patients when compared to controls (Eryilmaz et al., 2017). Similarly, a study has identified differentially methylated CpGs in the blood and saliva of Parkinson's disease patients in genes associated with the immune system (*ABCB9*, *CLORF200*, *AZU1*, all hypomethylated), Wnt pathway signalling (*PARK2*, *LRRK2*, both hypomethylated), neuronal projection (*CEP63*, hypermethylated), mitochondrial function (*LARS2*, *MIRI977*, *DDAH2*, all

hypomethylated), and others (Chuang et al., 2017). However, additional research would be required before a causal relationship could be determined.

As 5fC and 5CaC methylation were only recently discovered in 2011 (Ito et al., 2011), there is less research available on the relationship between these methylation forms and disease. However research has suggested that DNA extracted from the hippocampus and parahippocampal gyrus in both pre-clinical and late-stage Alzheimer's disease patients showed a significant enrichment of TET1, a genome wide enrichment of 5mC and 5hmC methylation, and a depletion of global 5fC and 5CaC methylation compared to controls (Bradley-Whitman & Lovell, 2013). As the roles of all the methylation forms have not been elucidated, it is debatable whether the changes in methylation are symptomatic or causative of dementia phenotypes.

The TET3 oxidation enzyme involved in the oxidation of 5mC to its oxidised forms, has 3 major isoforms, of which the full length isoform (TET3FL) possesses an N-terminal CXXC domain which has a very high affinity for 5CaC (S. G. Jin et al., 2016). TET3FL shows a very restricted localisation to TSS for key genes in neuronal cells including genes associated with mRNA processing, lysosome function, and key genes of BER (S. G. Jin et al., 2016). It is also thought to perform mass oxidation of 5mC to 5CaC to reduce methylation of the TSS at crucial genes (S. G. Jin et al., 2016). Several diseases are associated with dysfunction of *TET3* including neuronal ceroid lipofuscinoses (neurodegenerative diseases characterised by lysosomal dysfunction)(Shacka, 2012), Parkinson's disease (Gan-Or, Dion, & Rouleau, 2015; Schneider & Zhang, 2010; Schöndorf et al., 2014; Sidransky et al., 2009) and general degradation of cognitive faculties during ageing (Cuervo & Dice, 2000). In this way, TET3FL may exert a protective effect against general cognitive ageing, and the worsening of Parkinson's and lysosomal diseases by ensuring the TSS of such essential genes are not 5mC methylated and therefore repressed (S. G. Jin et al., 2016). Similarly, dysregulation of methylation states and mutations in methylation associated genes (TET, DNMT) have the potential to cause disease.

Methylation effects expression of genes and therefore phenotype, through repression of transcription, as enacted by methylation readers, such as MBD, Kaiso, Kaiso-like, and SRA proteins (Deaton & Bird, 2011; Defossez & Stancheva, 2011; Sasai et al., 2010). However, cell type, disease, and other environmental influences on the body may also show to affect methylation in certain gene regions. An individual's gender, age, diet, and sample cell type all influence the methylome observed (Whyte, Ellis, Brown, & Kenna, 2019).

Age related changes are observed across tissue types, with development involving several genome wide methylation and demethylation events, and most intra-individual differences in methylation not caused by cell type, are correlated with age (Gervin et al., 2016). Lifestyle also plays a large role in methylation, with global 5mC levels associated with serum glucose levels, and global 5hmC levels associated with BMI, waist circumference, cholesterol, and triglycerides (Nicoletti et al., 2016), with BMI influencing global methylation in response to adiposity (Wahl et al., 2017). Smoking has also been shown to reversibly deplete methylation at specific CpG sites across the genome, with recovery of methylation returning between 12 weeks to 10 years after cessation (Tsaprouni et al., 2014; Whyte et al., 2019). In this way, methylation is a mechanism that regulates transcription and therefore phenotype, but transcription and environmental effects may also regulate methylation levels.

As stated above, strand specific methylation while understudied is potentially important in expression. This is because while methylation is normally strand-complementary, it is known that hemi-methylation (methylation of only one strand, or unequal/uncomplimentary methylation between strands) is possible at genomic sites. For instance, hemi-methylation was reported to be key in genomic imprinting and regulation of genes in the 4C subpopulation of tetraploid cortical neurons of mice (Patiño-Parrado, Gómez-Jiménez, López-Sánchez, & Frade, 2017). Hemi-methylation of CpG sites with 5mC/5hmC or 5fC/5CaC reported differential repression of genes in isogenic HeLa derived cell lines, indicating that not only is the CpG site location and strand of methylation important in repression, but also the type of methylation (Kitsera et al., 2017). While this study cannot investigate methylation of both strands

at the rs6265 CpG, it is the first study to investigate 5mC, 5hmC, 5fC/CaC antisense methylation as a potential effect on the regulation of transcription.

I theorise that strand specific methylation effects the level of strand specific expression- which in the context of *BDNF* may contribute to the expression of the coding transcript through R-loop mediated inhibition of promoter methylation (Ginno, Lott, Christensen, Korf, & Chédin, 2012). In this fashion, modulated levels of methylation at the antisense rs6265 epiallele would result in altered levels of *BDNF* sense transcript. Changes in antisense transcript abundance may act to either increase or decrease sense transcript by several mechanisms. It is possible that as antisense transcript levels decrease, so would the sense transcript:

1. Heighted AS-rs6265 5mC methylation (or potentially decreased 5hmC methylation- known to enhance transcription in some scenarios (D. Q. Shi et al., 2017)) would decrease the level of anti-sense transcripts, which in turn decreases the amount of *BDNF* sense transcripts, as there would be no antisense transcript R-loop mediated repression of de novo methylation at the BDNF promoter. This type of cis-acting mechanism has been reported as a key mechanism of breast cancer in the *HOX* gene (Barman et al., 2019; Guil & Esteller, 2012; Gupta et al., 2010).

2. Alternatively, it is possible that the antisense transcript acts to stabilise some isoforms of the sense transcript, as seen in the stabilisation of β secretase 1 (*BACE1*) sense transcripts, through formation of a RNA-duplex with the antisense transcript, and disruption of this mechanism contributes to the Alzheimer's disease phenotype (Barman et al., 2019; Faghihi et al., 2010).

It is also possible that the *BDNF* antisense transcript negatively correlates with *BDNF* sense transcript, and that increased 5mC methylation of AS-rs6265 would act to increase overall *BDNF* sense transcript levels. A mechanism by which this could happen may be through transcription of the antisense transcript to repress transcription of the sense strand by 5mC methylating CpG islands. This mechanism is

seen in the repression of *HBA2* in α-thalassemia, and is similar to X inactivation mediated repression (Tufarelli et al., 2003). Similarly, *IGF2R* is also transcriptionally repressed through transcription of antisense AIRN in a lncRNA mediated mechanism, interfering with RNA polymerase II recruitment when AIRN overlapped with the IGF2R promoter sequence(Barman et al., 2019; Latos et al., 2012).

I further suggest that the types of methylation observed at this site directly influence the ratios and overall levels of transcripts expressed, which in turn will have a direct effect on phenotypes. As mentioned previously, not only is it the CpG site and strand, but also the type of methylation present that dictates the repressive/enhancing effects of methylation (Kitsera et al., 2017). I theorise that methylation, expression, and disease correlate at these CpG sites.

1.3.5 Technology to investigate methylation

Since its discovery, the quantification of DNA methylation has been an ongoing pursuit in epigenetic research. Two approaches for the quantification of methylation states at base resolution exist. One, conversion sequencing methods which typically bisulphite convert DNA before sequencing and comparison to normal DNA sequencing to determine which bases were methylated. Two, direct quantification methods that directly assess methylation at DNA segments. Both approaches are outlined below.

The most common protocol to quantify 5mC methylation at base resolution in DNA is through a conversion sequencing method and the bisulphite conversion of DNA. Bisulphite conversion treats ssDNA with sodium-bisulphite to deaminate 5fC/5CaC/unmodified cytosines into uracil which are read as thymine in sequencing. This allows for the quantification of 5mC and/or 5hmC methylation, but cannot differentiate between the two. Bisulphite mapping was then further developed to quantify 5mC methylation using oxidative bisulphite (oxibisulphite) conversion of DNA (whereby an oxidant is added to the bisulphite treated DNA, which will only read 5mC as cytosines in sequencing and 5hmC will be read as thymine) in tandem with

bisulphite sequencing and using subtraction-based mapping to differentiate 5mC and 5hmC methylation. This is illustrated in Figure 1.3.

Finally, more recently additional methods are being developed to quantify 5fC and 5CaC methylation. Methylation assisted bisulphite sequence (MAB-Seq) uses the *M.SssI* methylase to 5mC methylate any unmodified cytosines at CpG sites prior to bisulphite conversion and sequencing. The sequenced DNA can then be used in further subtraction-based mapping with bisulphite and oxibisulphite sequencing to determine which CpG sites are 5fC/CaC methylated. 5CaC methylation assisted bisulphite sequencing (CaMAB-Seq) can also be used to differentiate 5fC and 5CaC further through subtraction-based mapping by treating the DNA with sodium borohydride ($NaBH_4$) after methylation by *M.SssI* enzyme and prior bisulphite conversion (H. Wu & Zhang, 2015). (Figure 1.3)

While these main forms of conversion can differentiate and quantify methylation through subtraction-based mapping, there are alternative forms of conversion. For instance, TET assisted bisulphite sequencing (TAB-Seq) directly maps 5hmC methylation by glycosylating DNA with β-glucosyltransferase (βGT) which will glycosylate 5hmC into β-glucosyl-5-hydroxymethylcytosine (5gmC). The 5gmC is protected against TET enzymes, which will oxidise all other forms of methylation to 5CaC. The DNA is then bisulphite converted and sequenced, mapping 5hmC methylation as cytosines and all other forms of methylation as thymine (M. Yu et al., 2012).

Similarly, chemical modification assisted bisulphite sequencing for the analysis of 5CaC (CAB-Seq) and 5fC (fCAB-Seq) may be used as alternative forms of conversion to map these methylation marks through subtraction-based mapping. CAB-Seq uses 1-ethyl-3-[3-dimethylaminopropyl]-carbodiimide hydrochloride (EDC) to catalyse amide bond formation between amine and 5CaC. Modified 5CaC is then protected against deamination in bisulphite treatment and is compared in sequencing against standard bisulphite converted DNA to determine 5CaC location and levels (X. Lu et al., 2013). fCAB-Seq works in a very similar fashion, protecting 5fC against

deamination in bisulphite conversion by treatment with O-ethylhydroxylamine (EtONH2), and sequencing. This sequence is compared against standard bisulphite sequencing to determine which sites were 5fC methylated (C. X. Song et al., 2013). In this way, each methylation mark may be quantified through conversion and sequencing (H. Wu, Wu, & Zhang, 2016; H. Wu & Zhang, 2015). A summary of all the conversion sequencing based methods is presented in Figure 1.3.

Conversion sequencing-based methods typically quantify DNA methylation through Sanger sequencing or pyrosequencing. Sanger sequencing based methods involve cloning converted DNA and sequencing the clones in parallel. The sequences can be used in semi-quantification to determine the amount of methylation observed at a CpG as a ratio of cytosines to thymines at sites of interest. Pyrosequencing uses a direct quantification method by comparing the ratio of cytosines to thymines across all sequences. These methods typically target one region or gene of interest. However, converted libraries can be used to PCR amplify converted DNA amplicons for sequencing which when combined will allow for whole genome quantification of methylation. Alternatively, high-throughput investigation of the most commonly investigated CpGs may be performed by bead arrays/microarrays. If methylation is detected in the region of interest it is then possible to further quantify methylation using pyrosequencing (Kurdyukov & Bullock, 2016).

While conversion and sequencing methods are the most common commercially available methods for quantifying methylation, there are alternative forms of quantification. These include restriction enzyme digestion-based methods followed by PCR/qPRC, methyl-sensitive cut counting, high performance liquid chromatography ultra-violet (HPLC-UV), liquid chromatography mass-spectrometry (LC-MS/MS), and enzyme-linked immunosorbent assays (ELISA) (Kurdyukov & Bullock, 2016). These methods directly quantify DNA methylation without conversion. However, they may have varying ability in throughput and sensitivity. The “gold-standard” benchmarks for these methods are HPLC-UV and LC-MS/MS which while very sensitive to methylation, are usually expensive and low throughput.

In the past, methods to quantify oxidised methylation beyond 5hmC methylation were lacking and technologies to overcome this are still being pursued (C. X. Song et al., 2013). Recently, hybrid methods of conversion and tagging are being developed such as fC-Seal. fC-Seal maps 5fC across the genome by chemical reduction with biotin tagging. Similar to TAB-Seq 5hmC is first glycosylated with unmodified glucose using βGT. 5fC is then reduced to 5hmC using $NaBH_4$ and labelled with an azide-modified glucose. Azide modified glucose is pulled-down using biotin tagging, cleaved and sequenced. (C. X. Song et al., 2013). This is shown in Figure 1.3 as normal and modified Pvu-Seal-Seq. In this way, techniques have been developed to quantify each methylation state through high throughput and conversion-based methods, or the low-throughput gold standard; LC-MS/MS.

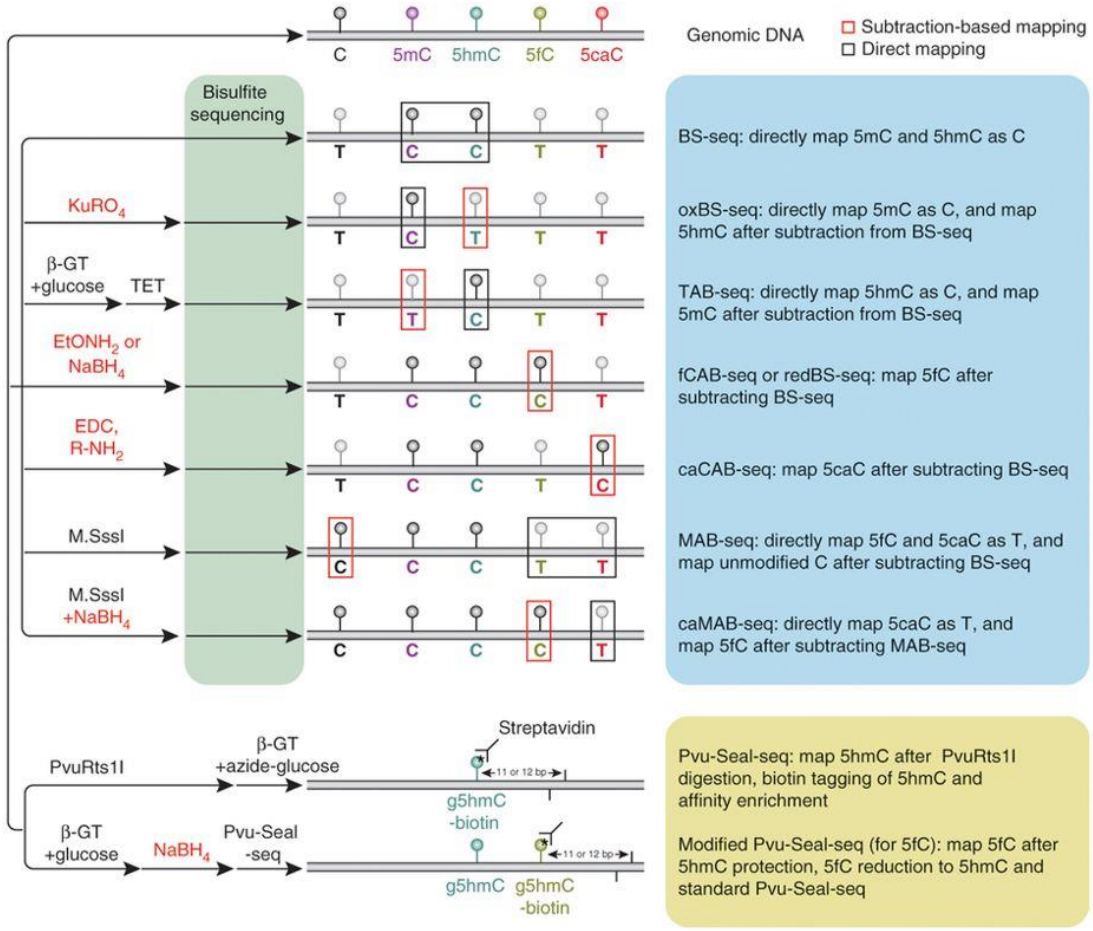


Figure 1.3 Summary of all conversions techniques to assess methylation types at base resolution. All treatments are followed by bisulphite conversion and sequencing to interpret base calls. BS-seq; Bisulphite sequencing, oxBS-seq; oxidative bisulphite sequencing, TAB-seq; TET assisted bisulphite sequencing, fCAB; 5fC chemical modification assisted bisulphite sequencing, caCAB-seq; 5CaC chemical assisted bisulphite sequencing, MAB-seq; methylation assisted bisulphite sequencing, caMAB-Seq; 5CaC methylation assisted bisulphite sequencing. Figure is taken from (H. Wu & Zhang, 2015).

1.4 Candidate DNA methylation epialleles

Genetic factors which contribute to neurodegenerative, neurodevelopmental and neuropsychiatric diseases have been extensively investigated. Typically, research has focused on the identification of common variants identified through Genome Wide Association Studies (GWAS) and indeed many cognitive and developmental disease phenotypes show associated for the same risk variants (Pizzo et al., 2019).

While many "common" variants have been studied previously, new research is investigating the effect of DNA methylation occurring at candidate risk sites and have coined the term epialleles. Risk epialleles are single nucleotide variant sites where one of the alleles has the potential to generate or disrupt a cytosine-guanine dinucleotide (CpG) site where DNA methylation can occur. This CpG site is capable of supporting cytosine DNA methylation in mammals. This means that the allele specific effects generated by the variant may not only be due to the base substitution but also whether DNA methylation is present at the CpG site. Three candidate epialleles for cognition across two genes will be the main focus of this thesis.

1.4.1 Apolipoprotein E and the rs7412 and rs429358 ε isoform variants

As mentioned previously, Apolipoprotein E (*APOE*) is a gene that when transcribed and translated produces APOE protein that processes and transports lipid and cholesterol (Bu, 2009; R. Mahley, 1988; R. W. Mahley & Rall, 2000). It transports cellular lipids and cholesterol by binding to cell surface APOE receptors. APOE is found most abundantly in the liver and brain (Bu, 2009; R. Mahley, 1988; R. W. Mahley & Rall, 2000). Receptors of APOE are low density lipoprotein receptors (LDLR) which include very low density lipoprotein receptors (VLDLR) and apolipoprotein E receptor 2 (APOER2) both of which are also receptors for the reelin protein, involved in cell to cell interactions and synaptic plasticity (Herz & Chen, 2006).

APOE possesses two coding variant sites associated with disease risk. rs7412 (Ch19:44,908,822 in Hg38) and rs429358 (Ch19:44,908,684 in Hg38) are located in *APOE* exon 4. The minor allele frequencies for rs7412 and rs429358 are 8% and 15%

of the population, respectively (R. A. Gibbs et al., 2015). A combination of these variant genotypes determines the *APOE* ε isoform expressed, e.g. ε2, ε3, or ε4, with ε3 being the "normal" or most common ε phenotype globally (Farrer et al., 1997). Possession of the ε4 allele has been significantly associated with Alzheimer's disease and ε4 is the largest risk factor for the disease accounting for 40-60% of the genetic variability observed in Alzheimer's disease (Farrer et al., 1997; Hatters, Peters-Libeu, & Weisgraber, 2006). Moreover, ε isoforms show load burden, where ε4 homozygotes possess the highest risk for AD and also report the highest levels of amyloid-β plaques (Hashimoto et al., 2012). Possession of the ε2 isoform, however, has been shown to be protective against Alzheimer's disease (Farrer et al., 1997).

As stated previously, the rs7412 and rs429358 variant genotypes determine the amino acids in the APOE protein and hence the ε protein isoforms. The minor T allele of rs7412 causes an Arg158Cys amino acid substitution, and the minor C allele at rs429358 causes a Cys112Arg amino acid substitution (Zannis et al., 1982). The typical ε3 isoform contains cysteine and arginine at positions 112 and 158, respectively, while the ε2 isoform has two cysteines and the ε4 isoform has two arginines, respectively (Hatters et al., 2006; Weisgraber, Rall, & Mahley, 1981). In the ε4 phenotype, possession of arginine at amino positions 112 and 158 results in a N- and C- terminal domain interaction (L. M. Dong & Weisgraber, 1996; Hatters et al., 2006). In this interaction an APOE side chain containing arginine at position 61 changes orientation, causing the Arg61 residue to interact with Glu255 and changing the proteins properties and preference to interact (L. M. Dong & Weisgraber, 1996; Hatters et al., 2006). Summaries of the different genotypes, amino acids, isoforms, and corresponding risk for Alzheimer's disease for APOE rs7412 and rs429358 is presented in Table 1.1. A further diagram (Figure 6.1) is presented in chapter 6 detailing the CpG sites surrounding APOE, the structure of the exons in the gene, how the mutation works in phase, etc.

rs7412		rs429358			
Genotype	AA158	Genotype	AA112	Isoform	Effects
C/T	Arg/Cys	T/C	Cys/Arg	ε2/ε4	Increased risk. Synonymous with ε1/ε3
T/T	Cys	T/T	Cys	ε2/ε2	Protective effect
C/T	Arg/Cys	T/T	Cys	ε2/ε3	Partial protective effect
C/C	Arg	T/T	Cys	ε3/ε3	Standard likelihood of AD
C/C	Arg	T/C	Cys/Arg	ε3/ε4	Increased risk
C/C	Arg	C/C	Arg	ε4/ε4	Maximum risk

Table 1.1 The relationship between *APOE* rs7412 and rs429358 genotype, amino acid substitutions, resultant ε isoforms, and effects on risk for Alzheimer's disease (AD). AA158 and AA112 denote the amino acids present at position 158 and 112 in APOE, as determined by the genotype at rs7412 and rs429358, respectively. Arg refers to arginine, and Cys refers to cysteine.

In addition to this, ε4 isoforms are less stable compared to the ε2 and ε3 counterparts being more susceptible to thermal and chemical degradation (Acharya et al., 2002; Hatters et al., 2006; Julie A. Morrow et al., 2000). Furthermore, ε2 and ε3 isoforms show no or few stable folding state intermediates, while the ε4 isoform has multiple folding state intermediates that do not operate as ε2 and ε3 do which may account for the deleterious effects of ε4 isoforms (Acharya et al., 2002; Hatters et al., 2006; Julie A. Morrow et al., 2000). These properties of ε4, which act through protein changes, have been proposed as the mechanism affording risk for disease phenotypes.

The ε4 isoform has also been implicated in the accumulation of amyloid-β plaques and τ neurofibrillary tangles in the progression of cognitive disease (Herz & Chen, 2006). While it is putatively accepted that disruption in amyloid-β and to a lesser extend τ fibril is the cause of Alzheimer's phenotypes, why or how clearance of these molecules fails is unknown (Selkoe, 2012). Moreover, a function of APOE is the clearance of amyloid-$\beta_{1\text{-}40}$ peptide through interaction with LDL receptor related protein-1 (Shibata et al., 2000). Reelin is a protein which facilitates synaptic plasticity, neuronal migration, microtubule stabilisation and τ hyperphosphorylation and is essential in brain development (Herz & Chen, 2006). APOE and reelin complete for the same VLDLR and APOER2 receptors, and it is possible that the instability and shift in interaction preferences present in the APOE ε4 isoform are also contributing to the disruption of the reelin pathway leading to the accumulation of τ fibrils and the neurodegenerative phenotypes observed (Herz & Chen, 2006).

In addition to this, APOE has been found in amyloid-β plaques and τ tangles (Namba, Tomonaga, Kawasaki, Otomo, & Ikeda, 1991; Tanzi & Bertram, 2001) and neuronal ε4 isoforms have been shown to be especially susceptible to proteolytic cleavage in the cell compared to the ε2 or ε3 isoforms, which produces bioactive C terminal truncated fragments (Brecht et al., 2004; Y. Huang et al., 2001; Tolar et al., 1999). These fragments accumulate in Alzheimer's brains in tangle-like cellular inclusions and cause neurodegeneration. The proteolysis of ε4 isoforms occurs preferentially in brain regions which are most adversely affected by Alzheimer's disease (Brecht et al.,

2004; Y. Huang et al., 2001). This highlights the multiple pathways in which the ε4 isoform of APOE may be causing or a result of Alzheimer's disease.

Some APOE isoforms have also been associated with risk for "pure" synucleinopathies and may be contributing to neurodegeneration outside of the A-β plaque pathway (Tsuang et al., 2013). In Parkinson's and Dementia with Lewy bodies disease patients with low amyloid-β plaque burden, the *APOE* ε4 isoform was reported significantly represented indicating that it may be operating outside of this pathway (Tsuang et al., 2013). However, as ε4 isoforms have reduced binding affinity and action when processing lipids and cholesterols, it may well also operate to increase amyloid-β plaque burden through the dysregulation of cholesterol and lipids (R. Mahley, 1988). In this way the cumulative phenotypic effect of APOE isoforms may vary between person to person, resulting in different forms and severities of cognitive disease.

One explanation for the variability in physiological symptoms seen between individuals with the same APOE isoforms and diagnoses might be DNA methylation regulating expression. For instance, the two *APOE* variant sites (rs7412 and rs429358) may also be the cytosine of a CpG dinucleotide (with the major and minor alleles, respectively), that have been significantly linked with the levels of amyloid-β burden present (Chibnik et al., 2015). *APOE* 5mC methylation has been indicated as a potential cause or marker for many diseases. *APOE* has shown to be hypomethylated in dementia with Lewy bodies when quantified in bisulphite sequencing (Tulloch et al., 2018), mild cognitive impairment (Mancera-Páez et al., 2019) and Alzheimer's disease diagnoses (Foraker et al., 2015). *APOE* 5mC methylation has also been associated with delayed memory recall in healthy individuals (J. Liu et al., 2018).

Moreover, as mentioned earlier, the methylation of a single CpG site has been reported to affect promoter activity in HeLa cell lines, indicating that one CpG site may be predictive for the expression of a gene or region (Kitsera et al., 2017). Similar to this study, the different types of methylation were investigated at the hemi-methylated CpG, and found to exert a differential effect on promoter repression

(Kitsera et al., 2017). Additionally, the rs67899457 epiallele (a schizophrenia risk variant) and its surrounding CpG sites in *KLC1* reported heightened 5(h)mC methylation with the deleterious form of the variant in human samples, and these differences in methylation may contribute to the mechanism by which the deleterious and healthy haplotypes contribute to the manifestation of schizophrenia (Alfimova, Kondratyev, Golov, & Golimbet, 2020).

Individual CpG sites have also been individually associated with life expectancy, and the process of ageing induces hypomethylation at these sites in humans (Y. Zhang, Hapala, Brenner, & Wagner, 2017). Another study investigated the effects of any single CpGs methylation on autism, however no single CpG site's methylation level could be determined to significantly associate with autism after correction for multiple testing (Ellis, Gupta, Moes, West, & Arking, 2017). It is possible therefore, that while some diseases have associations with the methylation of a single variant, others may not. It is also possible that many of these studies do not have sufficient power after corrections for multiple testing, due to the large number of samples required to counter-act the power reducing effects from corrections.

Additionally, possession of the ε4 allele is coinherited with the *SYMPK* (a gene encoding a nuclear protein involved in polyadenylation and gene expression) rs4420638 variant minor G allele, indicating that the rs7412 and rs429358 variant CpGs in *APOE* may be in linkage disequilibrium (J. Liu et al., 2018). However *APOE* methylation does correlate with age, meaning that the link between methylation and disease may be due in part with ageing, and not due to the progression of disease (Ma et al., 2015).

1.4.2 Brain Derived Neurotrophic Factor and rs6265

Brain Derived Neurotrophic Factor (BDNF) is a neurotrophic protein which is abundant in the amygdala, hippocampus and cerebral cortex. It is essential for the survival and development of neurons in the central and peripheral nervous system, playing key roles in LTP, STP, synaptic plasticity, axonal sprouting, dendritic

arborisation and differentiation of neurons (Budni, Bellettini-Santos, Mina, Garcez, & Zugno, 2015; Figurov et al., 1996; Kovalchuk et al., 2002; Murer, Yan, & Raisman-Vozari, 2001; Patterson et al., 1996; Tyler, Alonso, Bramham, & Pozzo-Miller, 2002). BDNF binds to protein-kinase neurotrophin receptors and tropomyosin-related kinase receptors (Trk) (Budni et al., 2015). Deficiencies in BDNF during development and through life are associated with learning deficits and poor spatial learning and memory in rodents (Petzold, Psotta, Brigadski, Endres, & Lessmann, 2015). However, an increase of serum BDNF (Francesco Angelucci et al., 2010; Faria et al., 2014) and hippocampus and parietal cortex BDNF levels is reported in MCI and AD patients when compared to controls (Durany et al., 2000). In AD, the increase is mainly in the hippocampus, cerebral cortex, and amygdala (Ingelsson & Hyman, 2002; J. Lee et al., 2005). This indicates the specificity of BDNF in its roles when acting within different regions of the brain.

BDNF originates as pro-BDNF, a precursor molecule that undergoes proteolytic cleavage to become mature BDNF, which is transported in vesicles along neuronal projections to different parts of the brain (Binder & Scharfman, 2004; Pang et al., 2004; Schindowski et al., 2008). Proteolysis of pro-BDNF to BDNF is facilitated in a neuronal activity dependent manner and reduction of neuronal activity reduces BDNF available (Binder & Scharfman, 2004; Borodinova & Salozhin, 2017). Cleavage of pro-BDNF into BDNF due to synaptic activation is essential for hippocampal plasticity (Pang et al., 2004). Where synapses are competing in axonal turnover, cleavage of pro-BDNF into BDNF stabilises the synapse while lack of cleavage and accumulation of pro-BDNF at the synapse causes axonal retraction in synaptic pruning (Je et al., 2012).

Accumulation of pro-BDNF in the presence of p75NTR and sortilin can initiate cell death (Friedman, 2000; Teng et al., 2005). Here, where a sympathetic neuron may express both p75NTR and sortilin, pro-BDNF acts as a apoptotic ligand, by binding with p75NTR (sortilin acts as a co-ligand) and inducing caspase 9, 3, and 6 activity, resulting in a DNAse cascade and inducing apoptosis (Teng et al., 2005; X. Wang et al., 2001). In this way, activity dependent conversion of pro-BDNF to BDNF regulates

synaptic plasticity, but failure of this mechanism may also induce apoptosis (Pang et al., 2004; Teng et al., 2005).

BDNF has 17 sense transcript variants and 12 anti-sense transcript variants derived from 11 exons (Pruunsild et al., 2007). Figure 1.4 illustrates the different *BDNF* sense and anti-sense transcript and corresponding exon locations. *BDNF* variant rs6265 (Ch11:27,658,369 in Hg38), also known as the Val66Met polymorphism, is located in *BDNF* exon IX and has been associated with a variety of neuropsychiatric disorders (Hing, Sathyaputri, & Potash, 2018). Possession of the minor allele at rs6265 results in a valine to methionine substitution at the 66th amino acid in BDNF. The major allele creates a CpG site while the minor allele creates a CpA site on the sense strand. On the antisense strand the major allele still creates a CpG (antisense major allele is C, generating a complementary CpG), however, the minor allele creates a TpG site.

Where normal "val-BDNF" would be secreted in a constitutive and activity dependent manner (Binder & Scharfman, 2004), mutant "met-BDNF" shows far less activity dependent secretion and fails to localise at secretory granules and synapses (Michael F. Egan et al., 2003). Reduced activity dependent secretion is associated with reduced synaptic and hippocampal plasticity and may induce apoptosis under specific conditions (Pang et al., 2004; Teng et al., 2005).

BDNF is also thought to play a key role in ageing, memory, and in pathologies such as synucleinopathies. *BDNF* has been identified as a key gene involved in synaptic plasticity, learning, and epigenetic regulation of *BDNF* is modulated in the process of memory formation, and proposed to be modulated in maintenance and retrieval (Day & Sweatt, 2011b; Lubin et al., 2008; Tyler et al., 2002). Contextual fear conditioning in rats results increasing histone H3 acetylation and phosphoacetylation at *BDNF* promoter 4, and decreasing histone H4 acetylation at *BDNF* promoter 2, in area CA1 of hippocampus (Lubin et al., 2008). This regulates BDNF transcript expression by increasing the expression of BDNF IV, and decreasing expression of BDNF II isoforms, and overall increasing the amount of BDNF transcripts being expressed in the hippocampus (Lubin et al., 2008). Methylation was also reported to

control exon specific expression of BDNF transcripts, with a significant increase in BDNF IV also being associated with a significant decrease in *BDNF* exon IV promoter methylation in the hippocampus CA1 region (Lubin et al., 2008). In this way DNA methylation of *BDNF* contributes to memory formation.

BDNF acts through the TrkB signalling pathway which in turn affects the MAP kinase pathway, enhancing excitory transmission (Budni et al., 2015; Kovalchuk et al., 2002; Minichiello et al., 2002). During ageing, the pituitary gland changes in its endocrine function, which is partly responsible for the ageing process (Veldhuis, 2013). In rat models, it has been shown that the number of Trk B receptor subunits (mRNA and protein) is impaired with age which in turn reduces the modulatory effect BDNF has on the pituitary (Rage, Silhol, Binamé, Arancibia, & Tapia-Arancibia, 2007).

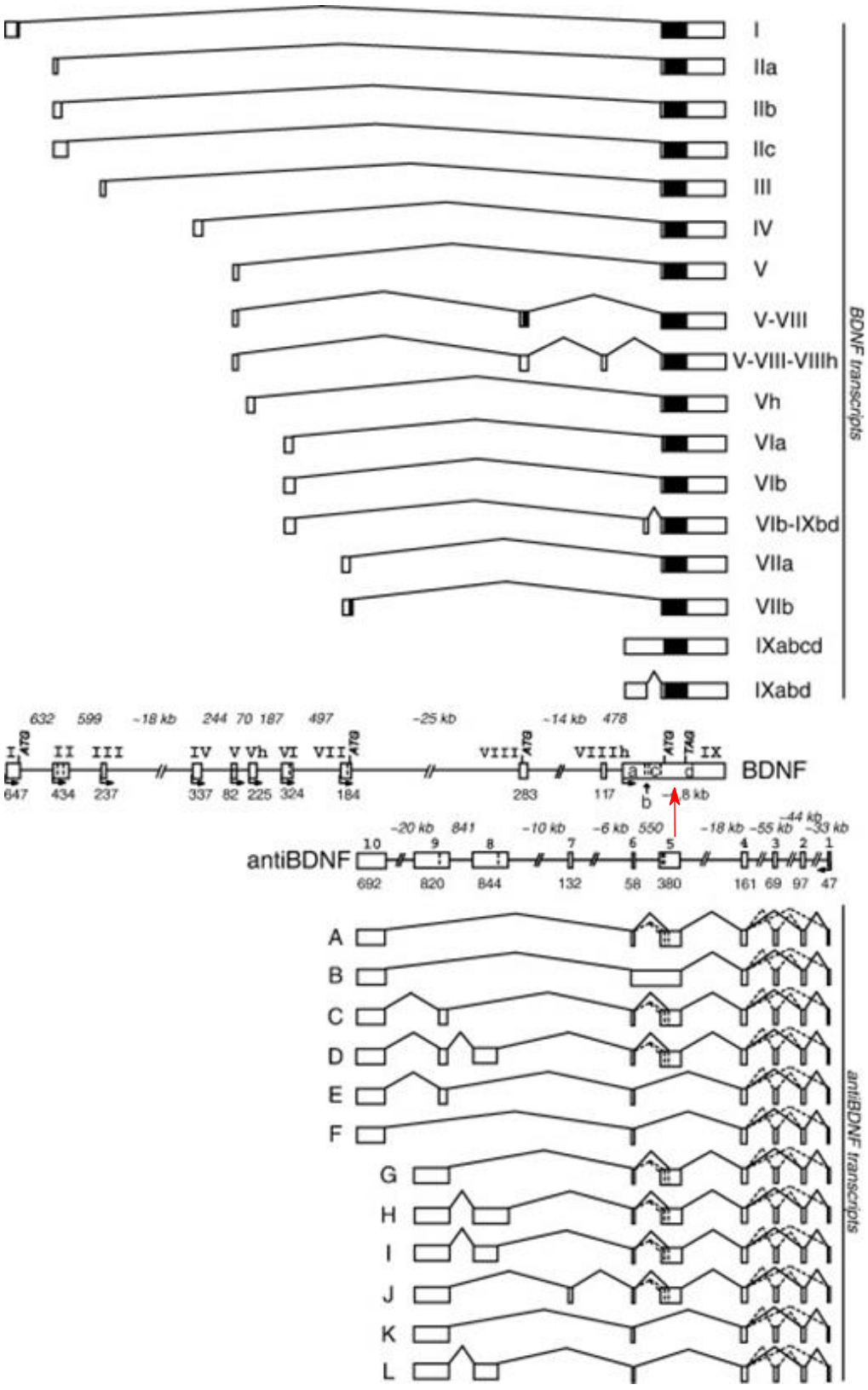


Figure 1.4 Structure of *BDNF* referencing all transcripts (top) and anti-sense transcripts (bottom). rs6265 in exon IX is detailed by a red arrow. Exons are represented as boxes and introns are represented as lines. Shading indicates translated regions in exons. Numbers presented above *BDNF* structure represent intron length in base pairs, numbers below exons represent exon length. *BDNF* rs6265 is located in exon IX. Image is taken from (Pruunsild et al., 2007).

In AD, the regions that are most affected by the disease (temporal cortex and hippocampus) are depleted of BDNF while the cerebellar, frontal and parietal cortices had no apparent change in BDNF (Hock, Heese, Hulette, Rosenberg, & Otten, 2000; J. Lee et al., 2005). Amyloid-β plaques induce deacetylation of histone H3 by accumulating histone deacetylases (HDAC) 2 and 5, repressing expression of BDNF, and drugs that reverse this deacetylation reduce the cell death at amyloid-β plaque sites (Yueling Wang, Jia, & Ma, 2018). Of interest, the regions showing the highest levels of BDNF often have the highest levels of amyloid-β plaques in the disease (Ferrer et al., 1999; Ingelsson & Hyman, 2002) while BDNF reduction was equally observed in Tau neurofibrillary tangle bearing and non-bearing neurons (Ferrer et al., 1999). This increase in BDNF surrounding amyloid-β plaques may be due to the activation of glial cells in the immune response to the plaques (Ferrer et al., 1999) and BDNF mRNA has been shown to increase in expression in monocytes when exposed to Tumour necrosis factor-α (TNF-α) and interleukin-6 (IL-6)(Schulte-Herbrüggen et al., 2005) which would be released in response to plaques.

While many disease states have been associated with the over or under expression of antisense and sense *BDNF* as a response or causal change, some cognitive disease states are associated with the methylation of *BDNF*. Significantly higher *BDNF* promoter methylation levels have been shown to be associated with multiple diseases such as depression (Fuchikami et al., 2011; H.-J. Kang et al., 2015), bipolar disorder (D'Addario et al., 2012), schizophrenia (Ikegame et al., 2013), and anxiety (Chagnon, Potvin, Hudon, & Préville, 2015) primarily in promoters I and IV (Zheleznyakova, Cao, & Schiöth, 2016). As DNA methylation of a gene region may modulate the isoform specific and total gene expression levels, changes in methylation of *BDNF* may up or down-regulate expression of BDNF which may contribute to disease phenotypes. As the major allele of rs6265 forms a CpG site, it supports that the rs6265 minor allele is associated with defective met-BDNF protein and lacks methylation at rs6265. This would also be reflected on the antisense rs6265 CpG site; as methylation is possible in the major allele CpG, but not the minor allele TpG. A lack of methylation in rs6265 minor alleles on both the sense and antisense strands may contribute to the pathology seen in met-BDNF.

Despite the fact that the rs6265 major allele can be methylated, it is often minor allele rs6265 carriers in transgenic mice that show repression in *BDNF* expression (Dulka et al., 2016; Ursini et al., 2016). This indicates that possession of the *BDNF* minor allele, while not methylated, still may lead to depletion of BDNF post-translationally as possession of the minor allele inhibits pro-BDNF cleavage, and reduces the amount of mature BDNF available to the cell (Dulka et al., 2016; Ursini et al., 2016). Methylation of the rs6265 in major allele homozygotes is also associated with impaired cognition and Schizophrenia (Ursini et al., 2016). However, these phenotypes are associated with an absence of methylation in rs6265 heterozygotes (Ursini et al., 2016). This suggests that not only does methylation operate in a context dependent way, but that methylation may be indicative of phenotypes.

Antisense *BDNF* is a long non-coding RNA (lncRNA) and is known to be a natural repressor of *BDNF* sense transcripts through CRE mediated repression (Modarresi et al., 2012). Here, The antisense transcript recruits the histone methyltransferase enhancer of Zeste homologue 2 (*EZH2*), a component of the polycomb repressive complex 2 (PCR2), to the BDNF promoter and K27 methylates the histone H3 tail (H3K27me) repressing transcription (Modarresi et al., 2012; Pruunsild et al., 2007; Varela et al., 2012)(Figure 1.5). As mentioned previously, sense-antisense duplexes may also form to inhibit translation of the mRNA and sequester the transcript, reducing the amount of mRNA available for translation into protein (W. G. Chen et al., 2003; Modarresi et al., 2012; Pruunsild et al., 2007; J. Shi, 2015). Overexpression of antisense *BDNF* transcripts has been associated with increased cellular apoptosis, reduced *BDNF* mRNA, neurodegenerative and neurodevelopmental disorders (B Lu, Nagappan, & Lu, 2014; Bai Lu, Nagappan, Guan, Nathan, & Wren, 2013; J.-H. Song, Yu, & Tan, 2015; Wei, Luo, Zou, & Wu, 2018).

In this way, while the Val66Met variant BDNF is associated with disease, methylation of both the sense and anti-sense transcripts may be acting as a secondary level of regulation in the expression of transcripts and hence contributing to phenotype. Many studies have investigated the effects of methylation of the sense strand on

mRNA expression, and studies have examined antisense RNA regulation on sense transcription (Tufarelli et al., 2003). However, no studies have quantified antisense-methylation at clinical variant sites and investigated the effects on the abundance of *BDNF* transcripts across different brain diseases.

1.4.3 DNA Methyltransferase 3 Like and rs7354779

The DNA Methyltransferase 3 Like (*DNMT3L*) gene encodes for a methyltransferase enzyme that is key in the methylation of CpG sites. DNMT3L catalyses DNMT3A and DNMT3B and enhances their methylase ability by 15 times facilitating d*e novo* methylation of unmethylated CpG sites (Chedin, Lieber, & Hsieh, 2002; Goll & Bestor, 2005; Jia, Jurkowska, Zhang, Jeltsch, & Cheng, 2007; Kareta, Botello, Ennis, Chou, & Chédin, 2006). As methylation is important in the expression of genes and the regulation of cellular and tissue phenotypes, DNA methyltransferase function is essential to these mechanisms.

DNMT3L houses a coding variant, rs7354779, where the wild type T allele may be substituted for a C allele. The minor allele is found in approximately 22% of the population (E. A. Consortium et al., 2015). This T to C change results in a codon and amino acid change from arginine (AGG) to glycine (GGG) at position 278 of the DNMT3L active domain (Haggarty et al., 2010). The major T allele is significantly associated with a higher childhood intelligence score and the minor C allele is associated with an increased risk of being below average intelligence (Haggarty et al., 2010). While the mechanisms by which this change in intelligence manifest are not known, it is known that the minor C allele of the rs7354779 variant is associated with genome wide significant hypomethylation as determined by bisulphite sequencing (El-Maarri et al., 2009). The C minor allele changes the amino and protein structure of DNMTL which compromises its ability to catalyse *de novo* methylation by DNMT3A (El-Maarri et al., 2009) and resulted in hypomethylation primarily at promoter CpG islands.

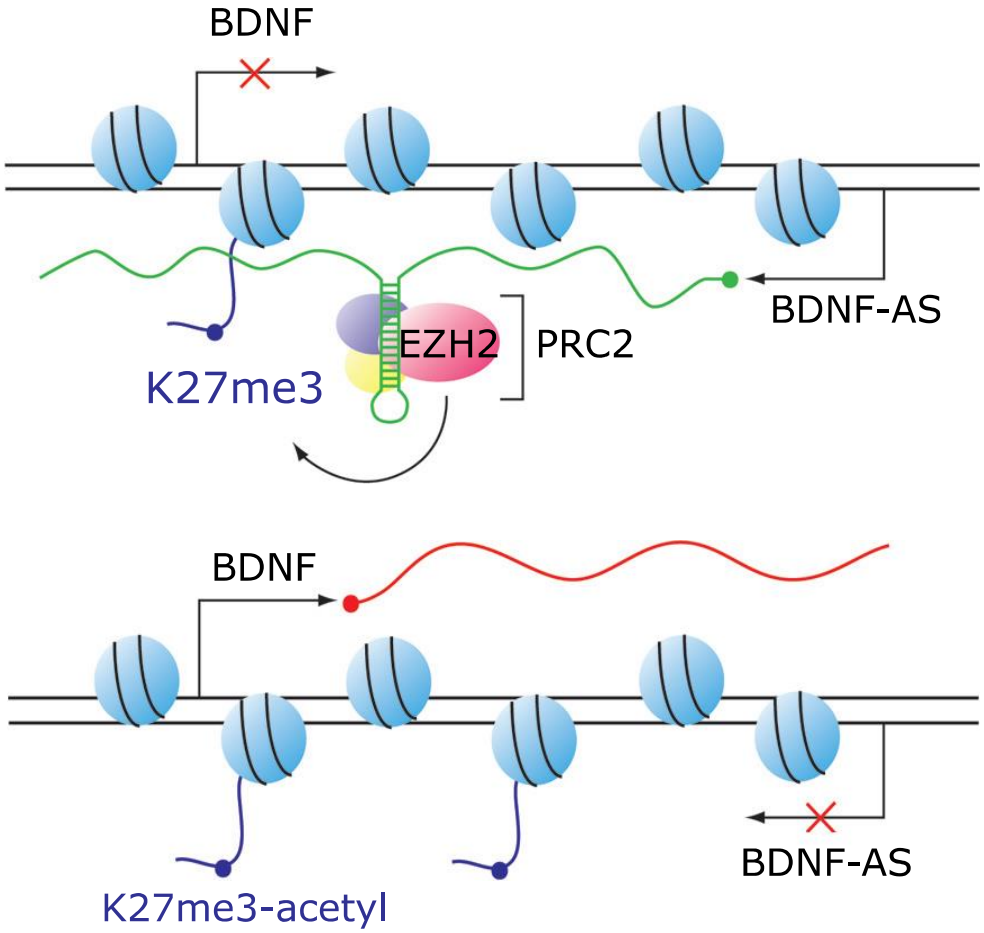


Figure 1.5 *BDNF* antisense vs sense transcript expression. The above panel displays *BDNF* antisense mediated suppression of the *BDNF* sense transcript, through histone methylation. The lower panel displays how reduced expression of *BDNF* antisense transcript increases *BDNF* sense transcript expression, through lack of repression. Image modified and taken from (Varela et al., 2012).

1.5 Introduction to thesis studies

Comparing 5mC and 5hmC methylation status at candidate epiallele loci associated with risk for neurological disease has great potential diagnostically. It is hypothesised that through the analysis of brain tissue from individuals with neurodegenerative diseases, there will be a correlation between genotype, methylation state at epialleles and disease phenotype. While many studies have quantified methylation on the sense strand, no studies have investigated 5mC, 5hmC, 5fC/CaC methylation on the antisense strand. As discussed earlier, single CpG sites and strand specific hemi-methylation have been proven to influence transcript expression and correlate with phenotypes. The three following aims will be explored in order to elucidate the role of DNA methylation at candidate risk epialleles within *DNMT3L*, antisense *BDNF* and *APOE* in cognitive disorders.

Aim 1. To optimise DNA conversion techniques to examine DNA methylation states in blood and across brain regions at candidate loci

Chapter 3	Optimisation of DNA conversion and the quantification of 5mC, 5hmC and 5fC/CaC methylation.
Chapter 3	Comparison of DNMT3L rs7354779 genotype and 5mC and 5hmC methylation levels

Aim 2. To compare genetic and methylation variation at the *BDNF* risk allele across brain regions and diseases

Chapter 4	Test for an association between genotype of *BDNF* rs6265 with Autism Spectrum Disorder (ASD) and Schizophrenia (SCZ) disease phenotypes
Chapter 4	Assess 5mC and 5hmC and 5fC/CaC methylation at the antisense *BDNF* rs6265 complementary CpG across brain regions and correlate with genotype and mild cognitive impairment, Parkinson's, dementia with Lewy bodies, and Alzheimer's disease phenotypes.

Chapter 5 Assess 5mC, 5hmC and 5fC/CaC methylation at antisense *BDNF* rs6265 complementary CpG across brain regions and mild cognitive impairment, Parkinson's, dementia with Lewy bodies, and Alzheimer's disease states and correlate with BDNF transcript expression.

Aim 3. Examine 5mC, 5hmC, and 5fC/CaC variation at *APOE* risk variants across brain regions and diseases

Chapter 6 Assess 5mC, 5hmC, and 5fC/CaC methylation levels at (sense) APOE ε isoform risk alleles rs7412 and rs429358 across brain tissues and mild cognitive impairment, Parkinson's, dementia with Lewy bodies, and Alzheimer's disease states.

Chapter 2: Materials and Methods

2.1 Materials and methods

2.1.1 Samples

Control DNA Samples

To optimise the methylation assays, DNA extracted from blood (S1-S6, donated by Dr Sally Pollard), Lymphoblastoid Cell Lines (AF103, AF105, AF111, TT296, and QQ0088), and sperm (N25) were used (Donated by Prof. John Armour). Formalin fixed, paraffin embedded (FFPE) mounted brain sections were obtained from the Nottingham Health Science Biobank (NHSB, Project number and ethics approval number: ACP000101) and Human Foetal Brain Total RNA (Clontech) was used as a positive control for cDNA synthesis and Digital Droplet PCR. Storage and use of these resources were in compliance with the EUTCD and HTA regulations of 2007. Non-affected, control samples are listed in detail in Table 2.1.

OPTIMA DNA samples

Samples from the Oxford Project to Investigate Memory and Ageing (OPTIMA)(D. Smith, 2009) were donated by Professor David Smith (lead researcher for the OPTIMA study) and Dr Donald Warden. DNA samples extracted from blood and snap frozen whole brain tissue were used for the preliminary analysis of genotype and methylation and split into Control, MCI, and AD phenotypes. Table 2.2 presents details of these samples. All samples had been previously genotyped at the rs7412 and rs429358 *APOE* variant sites which encodes for the *APOE* risk epsilon.

Imperial University Brain bank Snap Frozen Tissue Samples

Snap frozen brain tissue samples were obtained from Imperial College London Parkinson's UK Brain Bank (Research Tissue Bank reference number; 08/MRE09/31+5, research ethic committee approval number; 09/H606/70) from 3 brain regions; prefrontal cortex, hippocampus and cerebellum. Samples were taken from control, Mild Cognitive Impairment (MCI), Parkinson's Disease (PD), and Dementia with Lewy bodies (DLB) patients post-mortem. Details are provided in Table 2.3. Storage and use of these tissues was in compliance with the EUTCD and HTA regulations of 2007.

2.1.2 Reagents

Pre-made reagents and kits purchased, as well as bespoke solutions and reagents made in house, are listed in Tables 2.4, 2.5, and 2.6. Chemicals and generic solutions were in general purchased from Sigma-Aldrich and QIAGEN unless stated otherwise.

Sample ID	Species	Medium	Tissue	Gender	Diagnosis	Age	Source
S1	Human	DNA	Blood	-	None	-	Dr Sally Pollard
S2	Human	DNA	Blood	-	None	-	Dr Sally Pollard
S3	Human	DNA	Blood	-	None	-	Dr Sally Pollard
S4	Human	DNA	Blood	-	None	-	Dr Sally Pollard
S5	Human	DNA	Blood	-	None	-	Dr Sally Pollard
S6	Human	DNA	Blood	-	None	-	Dr Sally Pollard
AF103	Human	DNA	Lymphoblastoid Cell Lines	Female	None	32	Prof. John Armour
AF105	Human	DNA	Lymphoblastoid Cell Lines	Male	None	34	Prof. John Armour
AF111	Human	DNA	Lymphoblastoid Cell Lines	Female	None	68	Prof. John Armour
TT296	Human	DNA	Lymphoblastoid Cell Lines	Female	None (Myotonic Dystrophy)	34	Prof. John Armour
QQ0088	Human	DNA	Lymphoblastoid Cell Lines	Male	None (Myotonic Dystrophy)	17	Prof. John Armour
N25	Human	DNA	Sperm	Male	None	41	Prof. John Armour
NP22	Human	FFPE Section	Frontal Gyrus	Female	None	99	NHSB
NP22	Human	FFPE Section	Hippocampus	Female	None	99	NHSB
NP22	Human	FFPE Section	Parahippocampus	Female	None	99	NHSB
NP22	Human	FFPE Section	Cerebellum	Female	None	99	NHSB
Foetal RNA	Human	RNA	Whole Brain	N/A	-	20-33 Weeks	Clontech (UK)

Table 2.1 Unaffected control DNA samples used to optimise DNA extraction, PCR amplification, DNA conversion, and bisulphite PCR amplification. Sample ID, species, the medium (either extracted DNA or FFPE mounted sections), the origin tissue, gender, diagnosis, age, and sample source are listed if information was known. No samples had (relevant) diagnoses, allowing them to be used as control samples.

Sample ID	Tissue	Age	Gender	Diagnosis	APOE ε Isoform
OPTIMA 111	Blood	84	Male	Alzheimer's Disease	ε3/ε3
OPTIMA 111	Brain	84	Male	Alzheimer's Disease	ε3/ε3
OPTIMA 108	Blood	-	Female	Alzheimer's Disease	ε3/ε4
OPTIMA 108	Brain	-	Female	Alzheimer's Disease	ε3/ε4
OPTIMA 199	Blood	70	Female	Alzheimer's Disease	ε3/ε4
OPTIMA 199	Brain	70	Female	Alzheimer's Disease	ε3/ε4
OPTIMA 253	Blood	79	Female	Alzheimer's Disease	ε3/ε4
OPTIMA 253	Brain	79	Female	Alzheimer's Disease	ε3/ε4
OPTIMA 195	Blood	74	Male	Alzheimer's Disease	ε3/ε4
OPTIMA 195	Brain	74	Male	Alzheimer's Disease	ε3/ε4
OPTIMA 215	Blood	77	Female	Alzheimer's Disease	ε4/ε4
OPTIMA 215	Brain	77	Female	Alzheimer's Disease	ε4/ε4
OPTIMA 250	Blood	69	Female	Alzheimer's Disease	ε3/ε4
OPTIMA 250	Brain	69	Female	Alzheimer's Disease	ε3/ε4
OPTIMA 297	Blood	65	Female	Alzheimer's Disease	ε3/ε3
OPTIMA 297	Brain	65	Female	Alzheimer's Disease	ε3/ε3
OPTIMA 47	Blood	82	Female	Alzheimer's Disease	ε3/ε4
OPTIMA 47	Brain	82	Female	Alzheimer's Disease	ε3/ε4
OPTIMA 58	Blood	88	Male	Alzheimer's Disease	ε3/ε4
OPTIMA 58	Brain	88	Male	Alzheimer's Disease	ε3/ε4
OPTIMA 179	Blood	70	Female	Alzheimer's Disease	ε4/ε4
OPTIMA 179	Brain	70	Female	Alzheimer's Disease	ε4/ε4

Table 2.2 OPTIMA blood and brain samples used for analysis. Sample age at sampling, gender, diagnosis, APOE ε Isoform, and tissue are listed. APOE genotype information was made available by OPTIMA.

Sample ID	Tissue	Gender	Diagnosis	Age	APOE ε Isoform
PD687/A2A3	Prefrontal Cortex	Male	Parkinson's Disease	86	ε3/ε3
PD687/P2E1	Hippocampus	Male	Parkinson's Disease	86	ε3/ε3
PD687/CLB1	Cerebellum	Male	Parkinson's Disease	86	ε3/ε3
PD793/A2A7	Prefrontal Cortex	Female	Parkinson's Disease	75	ε3/ε3
PD793/P1F1	Hippocampus	Female	Parkinson's Disease	75	ε3/ε3
PD793/CLA1	Cerebellum	Female	Parkinson's Disease	75	ε3/ε3
PD898/A2A9	Prefrontal Cortex	Male	Parkinson's Disease	84	ε3/ε3
PD898/P2F10	Hippocampus	Male	Parkinson's Disease	84	ε3/ε3
PD898/CLB8	Cerebellum	Male	Parkinson's Disease	84	ε3/ε3
PD115/A2A2	Prefrontal Cortex	Male	Dementia with Lewy bodies	63	ε3/ε3
PD115/P2D3	Hippocampus	Male	Dementia with Lewy bodies	63	ε3/ε3
PD115/CLA2	Cerebellum	Male	Dementia with Lewy bodies	63	ε3/ε3
PD120/A202	Prefrontal Cortex	Female	Dementia with Lewy bodies	83	ε3/ε3
PD120/P1D3	Hippocampus	Female	Dementia with Lewy bodies	83	ε3/ε3
PD120/CLC1	Cerebellum	Female	Dementia with Lewy bodies	83	ε3/ε3
PD332/A2A2	Prefrontal Cortex	Male	Dementia with Lewy bodies	77	ε3/ε4
PD332/P2D2	Hippocampus	Male	Dementia with Lewy bodies	77	ε3/ε4
PD332/CLB4	Cerebellum	Male	Dementia with Lewy bodies	77	ε3/ε4
PDC007/A4B5	Prefrontal Cortex	Female	MCI	104	ε3/ε3
PDC007/P1D3	Hippocampus	Female	MCI	104	ε3/ε3
PDC007/CLB1	Cerebellum	Female	MCI	104	ε3/ε3
PDC111/A4A7	Prefrontal Cortex	Female	MCI	88	ε2/ε3
PDC111/P2D7	Hippocampus	Female	MCI	88	ε2/ε3
PDC111/CLA3	Cerebellum	Female	MCI	88	ε2/ε3
PD715/A5A6	Prefrontal Cortex	Male	MCI	86	ε3/ε3
PD715/P1G1	Hippocampus	Male	MCI	86	ε3/ε3

PD715/CLA5	Cerebellum	Male	MCI	86	ε3/ε3
C045/P2E5	Hippocampus	Male	MCI	77	ε3/ε3
C045/CLC5	Cerebellum	Male	MCI	77	ε3/ε3
PD065/A4C3	Prefrontal Cortex	Female	MCI	84	ε3/ε4
PD065/CLB4	Cerebellum	Female	MCI	84	ε3/ε4
PD092/A4A5	Prefrontal Cortex	Male	None	79	ε3/ε3
PD092/P1F1	Hippocampus	Male	None	79	ε3/ε3
PD092/CLB4B4	Cerebellum	Male	None	79	ε3/ε3
PDC110/A5L8	Prefrontal Cortex	Female	None	94	ε2/ε4
PDC110/P2E6	Hippocampus	Female	None	94	ε2/ε4
PDC110/C1B5	Cerebellum	Female	None	94	ε2/ε4
PDC112/A4A6	Prefrontal Cortex	Male	None	87	ε3/ε3
PDC112/P2F9	Hippocampus	Male	None	87	ε3/ε3
PDC112/CLB3	Cerebellum	Male	None	87	ε3/ε3

Table 2.3 Parkinson's UK Imperial College Brain Bank snap frozen tissue samples. Sample gender, age at sampling, diagnosis, tissue and APOE ε Isoform are detailed.

Assays	Kit/Reagent	Supplier
Extraction of DNA from FFPE tissue	QIAGEN QIAmp FFPE DNA extraction Kit	QIAGEN (UK)
Extraction of DNA & RNA from Snap Frozen Tissue	AllPrep DNA/RNA Mini Kit	QIAGEN (UK)
Bisulphite Conversion	CpGenome Turbo Bisulphite Modification Kit	Merck Millipore (UK)
cDNA Synthesis	RevertAid First Strand cDNA Synthesis Kit	Thermo Fisher Scientific (UK)
Bisulphite and Oxibisulphite conversion	CEGX TrueMethyl 6 Kit	Cambridge Epigenetix (UK)
Bisulphite and Oxibisulphite conversion	CEGX TrueMethyl Whole Genome Kit	Cambridge Epigenetix (UK)
Purifying PCR products/cloning prep	QIAquick PCR purification kit	QIAGEN (UK)
Extraction of DNA from FFPE tissue	QIAamp MinElute Columns	QIAGEN (UK)
E. Coli Cloning with Plasmid vectors	StrataClone Blunt Ends cloning kit	Agilent Technologies (UK)
Extraction of DNA from *E. Coli*	QIAmp Spin Miniprep Kit	QIAGEN (UK)
Extraction of RNA from Snap Frozen Tissue	mirVana miRNA extraction Kit	Thermo Fisher Scientific (UK)
Qubit Quantification	Qubit ssDNA Assay Kit	Thermo Fisher Scientific (UK)
Qubit Quantification	Qubit dsDNA Assay Kit	Thermo Fisher Scientific (UK)
cDNA Synthesis	RevertAid First Strand cDNA Synthesis Kit	New England Biolabs (UK)

Table 2.4 Premade kits. The assay which the kits were used for and the supplier is shown.

Method	Reagent	Supplier
PCR amplification	LongAmp *Taq* DNA Polymerase & Reaction Buffer	New England Biolabs
PCR amplification	Q5 High Fidelity DNA Polymerase & Reaction Buffer	New England Biolabs
PCR amplification	Hot Start Q5 High Fidelity DNA Polymerase & Reaction Buffer	New England Biolabs
PCR amplification	Invitrogen Platinum *Taq* Polymerase & Reaction Buffer	Thermo Fisher Scientific
PCR amplification	KAPA HiFi Uracil+ DNA Polymerase & Reaction Buffer	Kapa Biosystems
PCR amplification	One *Taq* Hot Start DNA Polymerase & Reaction Buffer	New England Biolabs
PCR amplification	AmpliTaq 360 Gold Master Mix	Thermo Fisher Scientific
PCR amplification	VeraSeq Ultra DNA Polymerase	Enzymatics
PCR amplification	dNTP solutions	New England Biolabs
Gel Electrophoresis	Agarose	Thermo Fisher Scientific
Gel Electrophoresis	Ethidium Bromide	Thermo Fisher Scientific
Gel Electrophoresis	100-1kb DNA ladder	New England Biolabs
Digital Droplet PCR	Foetal Brain RNA	Clontech (USA)
Digital Droplet PCR	DD-PCR Supermix for Probes	Bio-Rad (UK)
PCR amplification	ExoSAP-IT PCR product Cleanup	Thermo Fisher Scientific
E. Coli Cloning with Plasmid vectors	Ampicillin	Thermo Fisher Scientific
Methylation Assisted Bisulphite Conversion	CpG Methyltransferase (*M. SssI*) Enzyme & Reaction Buffer	New England Biolabs
Methylation Assisted Bisulphite Conversion	S-adenosylmethionine (SAM)	New England Biolabs
E. Coli Cloning with Plasmid vectors	S-Gal Blended High-Salt Agar	Thermo Fisher Scientific
E. Coli Cloning with Plasmid vectors	High-Salt LB Broth Agar	Thermo Fisher Scientific
Snap-Frozen Tissue DNA/RNA extraction	IsoPhenol-Chloroform	Thermo Fisher Scientific
cDNA synthesis, DD-PCR positive Control	Human Foetal Brain Total RNA	Clontech (USA)
FFPE Tissue Extraction	Proteinase K	New England Biolabs
FFPE Tissue Extraction	Hexadecane	Sigma Aldrich
Validation of bisulphite conversion	Methylated Human DNA Standard & hMLH1 primer set	Zymo Research
Validation of bisulphite conversion	Unmethylated Human DNA Standard & hMLH1 primer set	Zymo Research

Table 2.5 Reagents and polymerase systems. The assays in which the reagents were used and the supplier of the reagents are shown.

Solution/ Reagent	Method	Constituents
TAE Buffer	Gel Electrophoresis	242g Tris base, 57.1ml glacial acetic acid, 0.5M EDTA, made to 1L with ultrapure water. Makes 50x stock
TE Buffer (pH 8.0)	Eluting DNA	1M Tris-HCl (pH 8), 0.5M EDTA, Ultra-pure water
Universal Extraction Buffer	Extracting DNA from FFPE Tissue	50:1mM mix of Tris-HCl pH 8 to EDTA, 0.05% volume SDS
Gel Loading Dye	Gel Electrophoresis	Bromophenol Blue (0.03%), Glycerol (60%), 60mM EDTA
PCR Ready Lysis Buffer	DNA/RNA extraction from FFPE Tissue	40mM Tris, 1mM EDTA, 0.5% Tween-20
FFPE extraction Buffer	DNA/RNA extraction from FFPE Tissue	1M Tris-HCl (pH 8), 0.5M EDTA, 0.5ml 10% SDS, Ultra-pure water

Table 2.6 Solutions and reagents made in house with chemical constituents. The methods in which these solutions/reagents were used have also been listed.

2.2 In Vitro Protocols and materials

2.2.1 Probe and primer designs

Sanger sequencing primers

Bisulphite and normal primer design were performed using Zymo Research's Bisulphite Primer Seeker 12S, the Li Lab's MethPrimer and University of Massachusetts Primer3 design resources. Gene regions containing the epialleles of interest were taken from the University of California Santa Cruz (UCSC) Genome Browser and used to design primers. Primers were also designed to work in blunt-end transfection and cloning through vectors avoiding GG starts and topoisomerase I-binding sites if they were intended to be used in cloning. Where bisulphite APOE target alleles could not be assessed through novel primer designs, bisulphite ready primer sets for APOE designed by the Yu lab were produced with permission (Foraker et al., 2015). All primers were desalt purified and suspended in ultra-pure water. A table detailing all Sanger sequencing primer designs is shown below (Table 2.7).

Pyrosequencing Primer Design

Bisulphite and normal primers were designed using the *QIAGEN PyroMark Assay Design software 2.0* and involved defining design parameters to reduce primer mismatching and distant annealing temperatures. The sense strand of *APOE* and *DNMT3L* and the antisense strand of *BDNF* were selected for amplification. Selection of target gene regions was followed by multiple iterations of designs. Only primers that scored higher than 85 in *silico* prediction for success were tested in pyrosequencing. Pyrosequencing primer sets contain 3 individual primers; forward, reverse and sequencing primers. Either the forward or reverse primer of each set will be biotinylated at the 5' end (typically the reverse primer). Forward and reverse primers targeted the region to be amplified, while the sequencing primer specifically targeted the sequence containing the epiallele for genotype status or methylation analysis. Where bisulphite *APOE* target alleles could not be assessed through novel primer designs, bisulphite ready primer sets for *APOE* designed by the Yu lab were used with permission. Primers were synthesised by Sigma Aldrich, were PAGE purified, and suspended in ultra-pure water. A full list of all pyrosequencing primers

and their designs is provided in Table 2.8. A list of the optimised pyrosequencing primer sets, primer combinations and optimised melting temperatures is presented in Table 2.9.

Digital Droplet PCR probe selection

Digital Droplet PCR (DD-PCR) probes were selected to target the most common transcripts of BDNF sense and anti-sense mRNA. Transcripts of interest were identified using the Ensembl and Havana Genome Browsers. Commonly expressed transcripts in the brain regions being investigated were targeted for analysis. Quality checks were performed on isoform candidates confirming coding region between Ensembl and RefSeq (CCDS), TSL grade 1 indicating mRNA sequence has been confirmed multiple times, and APPRIS indicating that it is the principal isoform of this gene. The Genotype-Tissue Expression Project was used to observe differential transcript expression in different brain regions and cell types. No probes were found to bind or cross-react to ensure transcript targets were isolated in quantification of abundance.

To amplify the sense strand transcripts of interest, probe set Hs02718934_s1 was selected with a VIC dye reporter, which targets the two most common *BDNF* transcripts expressed in cerebellum, hippocampus and frontal cortex brain tissue; ENST00000395981 (BDNF II) & ENST0000039580.6 (BDNF V). The probe also targets expressed alternative sense *BDNF* transcripts; BDNF transcripts I, III, IV, VI, VII, VIII, IX, X, XI, XII, XIII, XIV, XVI, XVII, and XVIII, however, these are not expressed as much as *BDNF* II and V. To amplify the most common anti-sense transcripts in the same tissues, probe set Hs01010228_m1 was selected with a FAM dye reporter which targets transcripts NR_033314 (*BDNF* AS-204) and NR_002832.2 (*BDNS* AS-201).

AS-201 and AS-204 were quantified using the same probe and were differentiated by the identification of two distinct abundance populations in quantification. Comparing anti-sense transcript abundance with current expression database profiles differentiated AS-201 and AS-204 from each other and negative droplet data (Gte. Consortium, 2015). As a reference/control, probe set Hs00427620_m1 was used

which amplified TATA-box binding protein (*TBP*) transcript ENST00000392092.6 with a VIC dye reporter to be used simultaneously with the *BDNF* probes. The *TBP* probe set was selected for its relatively uniform expression level in tissues of interest. All probes were designed and synthesised by Thermo Fisher Scientific. A summary of DD-PCR probes used in this study can be found in Table 2.10.

Sanger sequencing primer	Methyl or No-Methyl	Target	Sequence (5' - 3')	Melting temperature (Tm ˚C)	Manufacturer
APOE_Norm_PCR1_F	No	*APOE*	CTGATGGACGAGACCATGAAGGAGTTGAAGG	60	Sigma-Aldrich
APOE_PCR1_F	No	*APOE*	GTTTGGAYGAGGTGAAGGAGTAGGTG	60	Sigma-Aldrich
APOE_PCR1_R	No	*APOE*	CTAATACCCACRACAACCTACACCTTCTCCACC	60	Sigma-Aldrich
APOE3'_PCR2_F	No	*APOE*	CCTGGCAGTGTACCAGGC	60	Sigma-Aldrich
APOE3'_PCR2_R	No	*APOE*	GCTGGGGCTTAGAGGAAATC	60	Sigma-Aldrich
APOE_PCR2_F	No	*APOE*	GYGTTAAGTTGGAGGAGTAGGTTTAG	61	Sigma-Aldrich
APOE_PCR2_R	No	*APOE*	CTGCATGTCTTCCACCAGGG	61	Sigma-Aldrich
APOE_PCR3_F	No	*APOE*	CTCCCACTGTGCGACACC	60	Sigma-Aldrich
APOE3'_Bisul_PCR1_F	Yes	*APOE*	GGTTTTTTTGGTTTATTTTTATTT	59	Sigma-Aldrich
APOE3'_Bisul_PCR1_R	Yes	*APOE*	CTATTCCACCAAAAACCCCA	59	Sigma-Aldrich
BS_APOE_PCR1_F	Yes	*APOE*	TATGAAGGAGTTGAAGGTTTATAAAT	58	Sigma-Aldrich
BS_APOE_PCR1_R	Yes	*APOE*	TATTAAACTAAAATCCACCCCAAAA	58	Sigma-Aldrich
BS_APOE_PCR2_R	Yes	*APOE*	AACTTAAAAAAAATCACAAAAAAAA	56	Sigma-Aldrich
BS_APOE_PCR3_F	Yes	*APOE*	TTGATGGAYGAGATTATGAAGGAGTTGAAGG	57	Sigma-Aldrich
BS_APOE_PCR3_R	Yes	*APOE*	ACCTACAAACRTATCTACTAAACCTACTCCTCCA AC	57	Sigma-Aldrich
BS_APOE_PCR4_F	Yes	*APOE*	GGGTAYGGTTGTTTAAGGAGTTGTAG	60	Sigma-Aldrich
BS_APOE_PCR4_R	Yes	*APOE*	ACTAACCRACCAAAAAACCCACAATAAC	60	Sigma-Aldrich
BS APOE Yu rs429358-F	Yes	*APOE rs429358*	AGGTTTATAAATCGGAATTGGAGGAATAAT	64	Sigma-Aldrich
BS APOE Yu rs429358-R	Yes	*APOE rs429358*	ATGATTTGTAGAAGCGTTTGGTAGTGT	64	Sigma-Aldrich
BS APOE Yu rs7412-F	Yes	*APOE rs7412*	AGGTTTATAAATAGGAATTGGAGGAATAAT	63	Sigma-Aldrich
BS APOE Yu rs7412-R	Yes	*APOE rs7412*	AACCCCCACCTAATACACTACCAA	63	Sigma-Aldrich
AS_BDNF_PCR1_F	No	*BDNF*	AACCTTGACCCTGCAGAATG	64	Sigma-Aldrich
AS_BDNF_PCR1_R	No	*BDNF*	ATGGGATTGCACTTGGTCTC	64	Sigma-Aldrich

AS_BS_BDNF_PCR1_F	Yes	*BDNF*	GGGAAGAAGGAGAATTTGAAGTTATA	62	Sigma-Aldrich
AS_BS_BDNF_PCR1_R	Yes	*BDNF*	TCCAATACCTTTTATCTATACCCCTAC	62	Sigma-Aldrich
AS_BS_BDNF_PCR2_R	Yes	*BDNF*	TTTTATCTATACCCCTACAACCTTCTTT	62	Sigma-Aldrich
AS_BS_BDNF_PCR3_F	Yes	*BDNF*	TTGGGTYGAATTTTTTGGTTTTTATTTAATAG	62	Sigma-Aldrich
AS_BS_BDNF_PCR3_R	Yes	*BDNF*	CTATCTTATTTCTACTTTCTCCCTACAATTCCACC	62	Sigma-Aldrich
DNMT3L_F	No	*DNMT3L*	GGGTTGCCTGTCTGTCTCTT	61	Sigma-Aldrich
DNMT3L_R	No	*DNMT3L*	GAGCCTGGGGTGTAGACTG	61	Sigma-Aldrich
BS_DNMT3L_PCR1_F	Yes	*DNMT3L*	TTTGTGTGTGGTGTATGTGTATAGTG	59	Sigma-Aldrich
BS_ DNMT3L _PCR1_R	Yes	*DNMT3L*	AAAAAACAAAAACAAAAAAATCCCT	59	Sigma-Aldrich
BS_ DNMT3L _PCR2_R	Yes	*DNMT3L*	AAAACAAAAAAATCCCTCAAACC	59	Sigma-Aldrich
BS_ DNMT3L _PCR3_F	Yes	*DNMT3L*	TTTTGTGTGTGGTGTATGTGTATAGT	58	Sigma-Aldrich
BS_DNMT3L_PCR4_F	Yes	*DNMT3L*	TTTTTAYGTGGGTAGGGGTTAGAGGG	62	Sigma-Aldrich
BS_DNMT3L_PCR4_F	Yes	*DNMT3L*	CTAAACAACRTCCCCATAACCAACCCCCC	62	Sigma-Aldrich
Rat_Prolactin_F	No	*Rat Prolactin*	GCTTCTGAGCAATGACACCA	55	Sigma-Aldrich
Rat_Prolactin_R	No	*Rat Prolactin*	ATTCCAGGAGTGCACCAAAC	55	Sigma-Aldrich
λ_Phage_F	No	λ phage genetic segment	GGAGAAACAGGGTGTGGAAA	60	In House
λ_Phage_R	No	λ phage genetic segment	GCAGAAAACAGCCGCATT	60	In House
Nested BS_λ_Phage_PCR1_F	Yes	λ phage genetic segment	TTATAAGGTGGATGGTAATTTTTATAGTT	60	Sigma-Aldrich
Nested BS_λ_Phage_PCR1_R	Yes	λ phage genetic segment	ACAATAACTAAAACAATTATTAATTAA	60	Sigma-Aldrich
BS_λ_Phage_PCR2_F	Yes	λ phage genetic segment	GGAGATATTGTAGTTTAGGATGGAAGTAA	60	Sigma-Aldrich
Universal Methylated Human DNA Control Primers	Yes	Methylated Control Segment	-	59	Cambridge Bioscience

Bisulphite Converted Universal Methylated DNA Primers	Yes	Methylated Control Segment	-	59	Cambridge Bioscience
T7P Promoter Sequencing Primer	Yes	Plasmid Amplicon	-	57	Designed In house
T7T Promoter Sequencing Primer	Yes	Plasmid Amplicon	-	57	Designed In house
Digestion Control Primers F	Yes	Digestion control	-	60	CEGX
Digestion Control Primers R	Yes	Digestion control	-	60	CEGX

Table 2.7 Primers used for PCR amplification. Primers were used for either Sanger Sequencing to genotype samples (Methyl or no-Methyl: No) or to assess methylation (Methyl or No-Methyl: Yes). Details of the target template are also provided as well as the 5' to 3' oligonucleotide sequence (with exception to pre-made primers designed by manufacturers). Details of the optimal annealing temperature (Tm˚C) and Manufacturers are also provided.

Pyrosequencing Primer	Methyl or No-Methyl	Target	Sequence (5' - 3')	Melting temperature (Tm ˚C)
APOE_PM_Norm_F	No	*APOE rs7412*	CTCCGCGATGCCGATGAC	60.8
APOE_PM_Norm_R	No	*APOE rs7412*	[Btn]CCCCGGCCTGGTACACTG	62.9
APOE_PM_Norm_S	No	*APOE rs7412*	CGATGACCTGCAGAAG	50.6
APOE PM Norm rs429358 -F	No	*APOE rs429358*	GCTGATGGACGAGACCATGAA	60
APOE PM Norm rs429358 -R	No	*APOE rs429358*	[Btn]AGGCGCTTCTGCAGGTCA	60
APOE PM Norm rs429358 -S	No	*APOE rs429358*	CTGCAGGCGGCGCAG	60
APOE PM BS rs7412-F	Yes	*APOE rs7412*	GGGTAAGGTTGTTTAAGGAGTTGTAG	50
APOE PM BS rs7412-S	Yes	*APOE rs7412*	GATGTAGATGATTTGTAGA	50
APOE_Yu_BS_rs7412-F (Yu:AE4_CGI_11-38_F)	Yes	*APOE rs7412/rs429358*	AGGTTTATAAATAGGAATTGGAGGAATAAT	50
APOE Yu BS rs7412-R (Yu: AE4_PyPCR_BSP_CpG38_R)	Yes	*APOE rs7412/rs429358*	[Btn]AACCCCCACCTAATACACTACCAA	50
APOE Yu BS rs7412-S (Yu: AE4_PySeq_BSP_CpG32_S)	Yes	*APOE rs7412/rs429358*	GGGTGAGTTTAGTTTTTTATTTG	50
APOE_Yu_BS_rs429358_S	Yes	*APOE rs429358*	GTTGGGAGAGGATATGGA	57
APOE_Yu_PM_BS_1_R	Yes	*APOE rs7412*	[Btn]ACACTACCAAACCCTTCTACAAATCAT	63.7
AS_BDNF_P3_BS_F	Yes	*BDNF rs6265*	TATTGAGTATTATTTTGGACG	56
AS_BDNF_P3_BS_R	Yes	*BDNF rs6265*	[Btn]GGTGGAATTGTAGGGAGAAAGTAGAAATAAGATA	56
AS_BDNF_P3_BS_S	Yes	*BDNF rs6265*	TTGGGTCGAATTTTTTGGTTTTTATTTAATAG	56
AS_BDNF_PM_Norm_F	No	*BDNF rs6265*	GGAGGTTTTAAAGGTATTTGATTATTGAG	63.3
AS_BDNF_PM_Norm_R	No	*BDNF rs6265*	[Btn]CTACCCAAATATACCAACCCATAAAACTCT	66.7
AS_BDNF_PM_Norm_S	No	*BDNF rs6265*	GAATTTTTTGGTTTTTATTTAATAG	52.7
AS_BDNF_PM_BS_1_S	Yes	*BDNF rs6265*	ATTTTTTGGTTTTTATTTAATAGTT	51
AS_BDNF_PM_BS_Long_F	Yes	*BDNF rs6265*	GGTAGAATTTTTTGGTTTTTATTTAATAGT	59.8

AS_BDNF_PM_BS_Long_R	Yes	*BDNF rs6265*	[Btn]ACCCTACAAAATAACCTAAAATTACAATC	62.1
AS_PM BDNF F2	No	*BDNF rs6265*	[Btn]TGGGCCGAACTTTCTGGT	60
AS_PM BDNF R2	No	*BDNF rs6265*	GGACTCTGGAGAGCGTGAATG	60
AS_PM BDNF S2	No	*BDNF rs6265*	CATCATTGGCTGACACT	60

Table 2.8 Pyrosequencing primers. Unlike Sanger sequencing primers, there are 3 primers per set. Primer sets were designed to determine genotype (Bisulphite ready: No) or quantify methylation (Yes). The oligonucleotide sequence is shown (5' to 3') with [Btn] indicating biotinylation. Target genes and epialleles are given, along with the optimal melting temperature for each primer individually. Overall primer annealing temperature for each set is determined in full by the PyroMark Q24 instrument.

Target	Methylation specific primers	Primer set/ combination	Optimised Tm (˚C)
APOE rs7412	No	APOE_PM_Norm_F APOE_PM_Norm_R APOE_PM_Norm_S	60
APOE rs7412	Yes	APOE_PM_BS_rs7412_F APOE Yu BS rs7412_R APOE_PM_BS_rs7412_S	56
APOE rs429358	No	APOE_PM_Norm_rs429358 _F APOE_PM_Norm_rs429358 _R APOE_PM_Norm_rs429358 _S	60
APOE rs429358	Yes	APOE_Yu_BS_rs7412_F APOE_Yu_BS_rs7412_R APOE_Yu_BS_rs429358_S	56
AS_BDNF rs6265	No	AS_PM_BDNF_F2 AS_PM_BDNF_R2 AS_PM_BDNF_S2	60
AS_BDNF rs6265	Yes	AS_BDNF_PM_Norm_F AS_BDNF_PM_Norm_R AS_BDNF_PM_BS1_S	56

Table 2.9 The optimal pyrosequencing primer sets or combinations used to assess the genotype and methylation status of gDNA and converted DNA. The target genes and epialleles are shown along with the optimal primer sets/combinations whether the primer sets were designed to assess genotype or methylation status as well as the optimised Tm for each of the primer sets during PCR amplification. Methylation specific primers denotes whether the primer set has been designed to amplify (O)BS converted DNA (Yes) or unconverted genomic DNA (No). *AS_BDNF* refers to the *BDNF* antisense strand.

Target Gene	Target Transcripts	Probe	Reporter system
BDNF Sense	ENST00000395981, ENST0000039580.6	Hs02718934_s1	VIC
BDNF Anti-Sense	NR_033314, NR_002832.2	Hs01010228_m1	FAM
TBP	ENST00000392092.6	Hs00427620_m1	VIC

Table 2.10 Digital Droplet probes used to assess expression of target genes. Target genes (and the transcript "direction"), along with the commonly expressed transcripts being analysed, the probe used to target the transcripts, and the reporter system used for quantification during fluorescence.

2.2.2 DNA & RNA extraction from snap frozen tissues

DNA and RNA were extracted from snap frozen tissue by using either a hand-held rotor homogeniser or Fast Prep 24 Ceramic Bead Mill Homogeniser. Tissue was transferred to sterile Eppendorf tubes and AllPrep DNA/RNA Mini Lysis Buffer added. Samples were kept on dry ice at all times. Homogenates were stored at -80˚C long.

Homogenates/lysates underwent DNA or RNA extraction using the AllPrep DNA/RNA Mini Kit which simultaneously extracted DNA and RNA. Following the kit protocol, homogenates were defrosted and held on ice for 10 minutes, vortexed and combined with 600µl-1ml IsoPhenol-Chloroform. Solutions were then further vortexed and centrifuged at 10,000g for 5 minutes. The aqueous top phase of the solution was then used in the standard AllPrep DNA/RNA Mini protocol. In brief, the AllPrep DNA/RNA extraction entailed adding RTL Buffest to the lysate, centrifuging at max speed for 3 minutes, extracting the supernatant and transferring it to the AllPrep DNA Spin Column. This column was further centrifuged at 8000g for 30 seconds and the flow through collected for the RNA extraction workflow while the DNA remained in the column.

For DNA purification, Buffer AW1 was added to the column and spun at 8000g for 15 seconds, discarding the flow through. Buffer AW2 was further added to the column and centrifuged at full speed for 2 minutes, discarding flow through. Elution Buffer was applied the membrane of the column and incubated at room temperature for 1 minute before being centrifuged at full speed for 1 minute to obtain eluate. This last elution step was repeated if higher volumes were required.

For RNA purification, the flow through from the AllPrep DNA Spin Column was mixed with 70% Ethanol, and pipetted to mix solution. Solution was transferred to a RNeasy Spin Column and centrifuged at 8000g for 15 seconds and the flow through discarded. Buffer RW1 was added to the column and spun at 8000g for 15 seconds, discarding flow through. Buffer RPE was added to the column, spun at 8000g for 2 minutes and flow through discarded. The collection tube was replaced and the

column spun at full speed for 1 minute. The collection tube was replaced again and RNase-free water added to the column membrane, before centrifuging at 10,000g for 1 minute. This elution step was repeated if higher volumes were required.

2.2.3 Amplification, Quantification and Quality control

PCR Amplification and Gel Electrophoresis

PCR amplification was performed according to the protocol of the polymerase used. Typically, a master mix of polymerase, reaction buffer, dNTP, nuclease free water was made and 5ng template DNA and 0.2 µM of forward and reverse primers were added. Running conditions and solution composition varied dependent on the amplification being performed and quality and efficacy of the template DNA and primers. Amplification was performed on an Applied Biosystems Veriti 96 Well Thermal Cycler according to the protocol of the polymerase and primer properties (Table 2.11).

Most polymerase systems required some optimisation. Typically, a master-mix comprising of appropriate quantities of polymerase, buffer, dNTP, and forward & reverse primers was prepared in a standard Eppendorf tube and then split between the number of tubes intending to be run. DNA was added last (with exception to sensitive polymerases such as Platinum *Taq* which took precedent) and made up to 25µl with ultrapure water. Reduced DNA input was used wherever possible to conserve DNA (down to 5ng input per well). Final volumes for PCR runs were reduced from 25µl to 12.5µl by halving the input of all reagents. This reduced consumption of DNA and *Taq*, while not compromising success of amplification. Many of the polymerase systems were optimised by using accessory buffers, such as the GC enhancer buffer (Q5 HiFi polymerase, One *Taq* polymerase), Kb extension buffer (Platinum *Taq*), etc. Additional $MgCl_2$, DMSO, and EDTA were added in optimisation and these changed from primer to primer. Final primer sets to analyse sites, along with optimal Tm˚C, polymerase systems, and running conditions can be found in (Table 2.12).

Polymerase System	Standard Running Conditions	
LongAmp *Taq* DNA Polymerase & Reaction Buffer	94°C (30s) 94°C (15-30s), [Tm˚C] (15-60s), 65˚C (50s/kb) 72˚C (10 mins), 4˚C Hold	30 cycles
Q5 High Fidelity DNA Polymerase & Reaction Buffer	98°C (30s) 98°C (5-10s), [Tm˚C] (10-30s), 72˚C (25s/kb) 72˚C (2 mins), 4˚C Hold	25-35 cycles
Hot Start Q5 High Fidelity DNA Polymerase & Reaction Buffer	98°C (30s) 98°C (5-10s), [Tm˚C] (10-30s), 72˚C (25s/kb) 72˚C (2 mins), 4˚C Hold	25-35 cycles
Invitrogen Platinum *Taq* Polymerase & Reaction Buffer	94°C (2mins) 94°C (30s), [Tm˚C] (30s), 72˚C (60s/kb) 72˚C (10 mins), 4˚C Hold	25-35 cycles
KAPA HiFi Uracil+ DNA Polymerase & Reaction Buffer	98°C (45s) 98°C (315s), [Tm˚C] (30s), 72˚C (30s) 72˚C (1 min), 4˚C Hold	30 cycles
One*Taq* Hot Start DNA Polymerase & Reaction Buffer	94°C (30s) 94°C (15-30s), [Tm˚C] (15-60s), 68˚C (60s/kb) 68˚C (5 mins), 4˚C Hold	30 cycles
AmpliTaq 360 Gold Master Mix	95°C (5-10mins) 95°C (15-30s), [Tm˚C] (30s), 72˚C (60s/kb) 72˚C (7 mins), 4˚C Hold	25-40 cycles
VeraSeq Ultra DNA Polymerase	98°C (30s) 98°C (5-10s), [Tm˚C] (10-30s), 72˚C (30s/kb) 72˚C (10 mins), 4˚C Hold	15-35 cycles

Table 2.11 Standard thermal cycler running conditions used in all polymerase systems. These running conditions vary based on the size of the amplicon being amplified, and the addition of other reagents such as GC Enhancing Buffer and $MgCl_2$. [Tm˚C] Indicates the primer pair annealing temperature, which will vary dependent on the primers being used. Optimal polymerase systems and running conditions are discussed further in the optimisation chapter.

Target	Methyl or No-Methyl	Primer set	Polymerase System	Additional reagents	Optimised Running Conditions		Amplicon size
APOE; rs7412, rs429358	No	APOE_PCR2_F/R	One *Taq* Hot Start DNA Polymerase & Reaction Buffer	15% volume GC Enhancer Buffer	94°C (30s) 94°C (30s), 60˚C (30s), 68˚C (1 minute) 68˚C (5 minutes), 4˚C Hold	30 cycles	~650bp
APOE; rs7412, rs429358	Yes	APOE_PCR2_F/ APOE_PCR1_R	Invitrogen Platinum *Taq* Polymerase & Reaction Buffer	$MgCl_2$	94°C (2 minutes) 94°C (5 minutes), 60-64˚C (30s), 72˚C (1 minute) 72˚C (10 minutes), 4˚C Hold	40 cycles	~200bp
AS_BDNF rs6265	No	BDNF_PCR1_F/R	LongAmp *Taq* DNA Polymerase & Reaction Buffer	-	94°C (30s) 94°C (1 minute), 62˚C (15s), 65˚C (1 minute) 72˚C (10 minutes), 4˚C Hold	30 cycles	~700bp
AS_BDNF rs6265	Yes	BS_BDNF_PCR3_F/R	KAPA HiFi Uracil+ DNA Polymerase & Reaction Buffer	-	98°C (45s) 98°C (15s), 60-62˚C (30s), 72˚C (30s) 72˚C (1 min), 4˚C Hold	30 cycles	~300bp
DNMT3L	No	DNMT3L_F/R	LongAmp *Taq* DNA Polymerase & Reaction Buffer	-	94°C (30s) 94°C (1 minute), 58˚C (30s), 65˚C (1 minute) 72˚C (10 minutes), 4˚C Hold	30 cycles	~200bp
DNMT3L	Yes	BS_DNMT3L_PCR4_F/R	LongAmp *Taq* DNA	-	94°C (30s) 94°C (15s), 62˚C (15s), 65˚C (1 minute)	30 cycles	~200bp

			Polymerase & Reaction Buffer		72˚C (7 minutes), 4˚C Hold		
Rat Prolactin	No	Rat_Prolactin_F/R	LongAmp *Taq* DNA Polymerase & Reaction Buffer	-	94°C (30s) 94°C (1 minute), 55˚C (30s), 65˚C (1 minute) 72˚C (10 minutes), 4˚C Hold	30 cycles	~400bp
λ phage genetic segment	No	λ_Phage_F/R	LongAmp *Taq* DNA Polymerase & Reaction Buffer	-	94°C (30s) 94°C (1 minute), 60˚C (30s), 65˚C (1 minute) 72˚C (10 minutes), 4˚C Hold	30 cycles	~450bp
λ phage genetic segment	Yes	BS_λ_Phage_PCR1_F/ BS_λ_Phage_PCR2_R	KAPA HiFi Uracil+ DNA Polymerase & Reaction Buffer	-	98°C (45s) 98°C (15s), 60˚C (30s), 72˚C (30s) 72˚C (1 min), 4˚C Hold	30 cycles	~300bp

Table 2.12 Optimal running conditions for primer sets in PCR amplification. The gene target, optimal primer set/combination, optimal Polymerase system, additional reagents required, and expected primer amplicon size are given for each amplification. Methyl or No methyl refers to whether the primer set are to be used on converted DNA or gDNA, respectively.

The processes of PCR amplification, analysis and purification were optimised for each primer set. Amplified product was quantified using Nanodrop or for extremely low concentration samples through Qubit quantification.

Nanodrop Quantification

Nanodrop quantification was performed on a Nanodrop 2000c UV/IV Spectrophotometer. 1µl of the sample to be quantified loaded. Reading the sample typically gave concentration (ng/µl), absorbance at 260nm, and absorbance at 280nm which was used to determine purity and contaminants.

Qubit Quantification

Qubit quantification performed used the DeNovix Flourometric Analyser and the Qubit ssDNA/dsDNA Assay Kits. The dsDNA kit was used for extracted and purified DNA and the ssDNA kit used for purified converted DNA. A working solution of Qubit ssDNA/dsDNA Reagent combined with Qubit ssDNA/dsDNA Buffer at 1:200 was made for each ssDNA/dsDNA standard and sample being quantified. Qubit ssDNA/dsDNA Standard 1 & 2 were combined with working solution in separate tubes. Purified genomic material was added to sample tubes and mixed. The final volume for all tubes was made to 200µl with ultrapure water. Tubes were incubated at room temperature for 2 minutes before calibrating the fluorometer using Standard 1 followed by Standard 2. Samples were then read on the fluorometer and quantified.

ExoSAP-IT purification and Gel electrophoresis

Amplified PCR products were purified with ExoSAP-IT. In brief, ExoSAP-IT purification combined 5µl of DNA amplicon with 2µl ExoSAP-IT and incubated in a thermal cycler at 37˚C for 15 minutes, followed by 80˚C for 15 minutes. Purified products were then taken forwards for Gel electrophoresis. 1-5% Agarose gels were made by combining agarose with TAE buffer and heated using a microwave until agarose had dissolved. Ethidium Bromide was used as the UV visualisation DNA intercalator. DNA ladder, 100bp to 1kb, depending on expected product size, and purified PCR product was loaded into gels with loading dye and typically run for 5 minutes at 40V followed by

45 minutes at 80V. Lower voltages were applied for low weight fragments as required. Gels were imaged on a UV visualiser (Bio-rad, UK) and observed products compared to expected product size for amplification.

Cloning

Cloning was initially performed to semi-quantitatively assess methylation at variant sites and to determine by proxy whether oxibisulphite conversion was successful. (O)BS converted human DNA was amplified by PCR and inserted into plasmids using blunt end vectors. Transfected colonies were grown and the plasmids sequenced. This increased the number of technical repeats for each sample, which allowed for semi-quantitative analysis through the comparison of each of the colony variant sites. Final methylation levels obtained through semi-quantitative analysis were determined as either the ratio of C to T bases at the variant sites in between the colonies' sequences, or the average of the ratios of the C to T peak heights at the variant sites between colonies (additional details in section 2.3.1).

(O)BS converted DNA samples were amplified using BS-BDNF primers and purified using ExoSAP-IT. Purified product was used to create a ligation reaction and cloning mixture which followed the standard StrataClone blunt ends cloning kit protocol. In brief, the purified amplicon was mixed with Cloning Buffer and Vector Mix amp. This mixture was incubated for 5 minutes at room temperature and then held on ice. 1µl of the mixture was mixed with an aliquot of StrataClone SoloPack Competent Cells and gently mixed. This cell mixture was incubated on ice for 20 minutes, followed by heatshock at 42˚C for 45 seconds and immediately returning to ice for 2 minutes. The mixture was transferred to a prewarmed aliquot of LB broth and incubated at 37˚C for 2 hours with agitation. S-Gal blended high-salt agar plates with 0.1% ampicillin and 2% X-gal were cast in sterile conditions and seeded with transformed cells. 5ul, 10ul or 100ul transformed bacteria were used to seed plates. Plates were left to grow over night at 37˚C and five transformed colonies from each plate taken and seeded into 2ml of high-salt LB broth with 0.1% ampicillin and left to grow over night at 37˚C with agitation. Resultant cultures were centrifuged at 10,000rpm (2 minutes) and purified using the QIAmp Spin Miniprep Kit.

In brief, the QIAmp Spin Miniprep kit resuspended the bacterial pellets in Buffer P1 and P2 in an Eppendorf tube and mixing. Buffer N3 was added, the tube was mixed and centrifuged at 13,000rpm for 10 minutes. The supernatant was moved to a QIAprep spin column and centrifuged at maximum speed for 30 seconds discarding the flow through. The spin column was washed with Buffer PB discarding flow through. The column was further washed with Buffer PE discarding flow through. The column was centrifuged at maximum speed for 60 seconds to dry the column membrane, the spin column was placed in a clean Eppendorf, and Buffer EB/Ultra-pure water applied to the membrane. The tube was incubated at room temperature for 1 minute then centrifuged at maximum speed for one minute. Flow through was stored at -20˚C long term. The plasmid solution yielded from the QIAmp Spin Miniprep kit was then quantified and sanger-sequenced using the T7 primer set with in house services.

2.2.4 DNA conversion to assess Methylation

Bisulphite conversion to assess 5(h)mC

In order to assess 5(h)mC methylation, The Millipore CpGenome Turbo Bisulphite Modification kit and the Bisulphite workflow for the CEGX TrueMethyl kits were used for bisulphite conversion of DNA. Both kits relied on strand separation of DNA followed by the deamination and sulphonation of unmethylated cytosine-guanine dinucleotides (CpG sites) into uracil, while 5(h)mC methylated CpG sites were protected. 5fC/CaC methylation is also decarbonylated and deaminated into uracil allowing for specific quantification of 5(h)mC methylation only.

In brief, the CpGenome Turbo method denatured DNA by incubating the DNA with NaOH at 37˚C for 10 minutes followed by bisulphite treatment with DNA modification Reagent at 70˚C for 40 minutes. Converted DNA was then desulphonated and purified, combining DNA and DNA Binding Buffer in a modified DNA Purification Column and centrifuging (11,000g for 1 minute). DNA was washed in the column with DNA Wash Buffer and desulphonated with Desulphonation Solution before a final

wash with DNA Wash Buffer. Converted DNA was eluted in DNA Elution Buffer heated to 70˚C. The Eluate was stored at -20˚C.

The CEGX TrueMethyl kit's bisulphite workflow entailed adding Digestion and Spike-In controls (2% w/w; Digestion control for the oxibisulphite pipeline and methylated λ-phage spike in control for bisulphite/MAB conversion pipelines) to the DNA. The solution was then mixed with Magnetic Bead Binding Solution 1 and incubated for 20 minutes at room temperature. DNA attached to the beads which were precipitated in a Magnetic Bead Rack and washed 3 times with 80% Acetonitrile. The beads were resuspended in Denaturing Solution and heated to 37 ˚C for 5 minutes to denature the DNA. Bisulphite Reagent Solution was added to the DNA, the mixture was repeatedly denatured (95˚C, 5 minutes) and then incubated (60˚C, 20 minutes) for 3 cycles in a thermocycler. The solution was centrifuged at 14,000g for 10 minutes at which point a qualitative quality control step was performed by observing the colour of the solution. The supernatant was combined with Magnetic Bead Binding Solution 2, and vortexed. This solution was precipitated in the Magnetic Bead Rack, supernatant aspirated and beads washed with 70% Ethanol before being resuspended in Desulphonation Solution. The Beads were then washed 3 more times with 70% Ethanol, left to dry, and resuspended in ultra-pure water. This mixture was precipitated in the Magnetic Bead Rack and the converted DNA supernatant taken. Converted DNA was stored at -20˚C.

Quantitative quality control assessing conversion efficiency of bisulphite converted samples was determined by amplifying methylated & converted λ-phage spike-in control through PCR. This was done using BS λ Phage PCR1 primers, followed by gel electrophoresis and sanger sequencing. Where bisulphite conversion has been successful, sequencing showed retention of cytosines at CpG sites but indicate thymines where non-CpG cytosines previously were reported. Retention of non-CpG cytosines in converted sequences indicated a failure in conversion.

Oxibisulphite Conversion to assess 5mC

Oxibiuslphite conversion was performed using the CEGX TrueMethyl HydroxyMethyl Workflow which oxidised ssDNA prior to deamination and sulphonation in bisulphite conversion. 5hmC methylation was oxidised to 5fC, while 5mC and unmethylated CpGs remained unchanged. Subtraction-based mapping could then be used with bisulphite converted DNA to determine levels of 5mC and 5hmC methylation individually. After conversion, converted ssDNA was purified and eluted in spin columns into ultra-pure water.

The TrueMethy HydroxyMethyl workflow was identical to that of the TrueMethyl bisulphite workflow (described above) however, after the DNA was denatured and prior to the Bisulphite Reagent Solution being added, oxidant solution was added. The oxidant solution was vortexed with the ssDNA solution, centrifuged and incubated at 40˚C for 10 minutes. The solution was then centrifuged at 14,000g for 10 minutes and a qualitative quality check performed based on the colour of the supernatant. The supernatant was then taken forwards to the bisulphite conversion workflow.

In addition to the qualitative control described above, quality control of converted samples was performed using *Taqαl* mediated digestion of spike in controls within converted samples. In brief, PCR amplification was performed on converted DNA sample (with digestion control), polymerase and buffer (Kapa HiFi U+ Polymerase), dNTP, and forward & reverse Digestion Control Primers. PCR amplification conditions were as follows: 95˚C initial denaturation for 2 minutes, 40 cycles of [95˚C for 30 seconds, 60˚C for 30 seconds, 72˚C for 15 seconds], followed by a final extension step of 70˚C for 5 minutes. The resultant amplified product was digested by incubating with *Taqαl* Digestion Enzyme, BSA, and *Taqαl* Digestion Buffer, and incubating at 65˚C for 18 hours followed by 80˚C for 20 minutes. Positive and negative controls were also digested using Cutting Control DNA, and by excluding *Taqαl* enzyme, respectively. Digested products were validated by agarose gel electrophoresis. In successful conversion, oxibisulphite converted DNA would not digest, however

bisulphite converted DNA would digest, leading to two distinct molecular weight bands (approximately 200bp and just above 100bp, respectively).

Where above enzymatic treatments with *M.SssI* and *Taq$^{\alpha}$I* were not successful or showed low efficiency, the QIAquick PCR purification kit was used in order to remove any further contaminants that may have been inhibiting reactions. In summary, DNA/amplicon was combined with PB Buffer and pH adjusted using the colour control. This mixture was transferred to QIAquick column and centrifuged at 8,000rpm for 1 minute. PE Buffer was applied to the column and run through via centrifuge. Flow through was discarded and the column centrifuged at maximum speed for 1 minute. Elution Buffer is added and incubated for 1 minute. The column is further centrifuged at maximum speed for 1 minutes and the eluate collected. Eluate was stored at -20˚C long term.

Reductive Bisulphite conversion to assess 5fC

The CEGX TrueMethyl Array fC workflow was also used to assess 5fC DNA methylation. Similar to the HydroxyMethyl workflow, the Array fC workflow incubates the ssDNA with a reductant solution prior to deamination and sulphonation in bisulphite conversion. This reduced 5fC methylated CpGs to 5hmC which was protected (along with nascent 5(h)mC) in bisulphite conversion, while unmethylated CpGs and 5CaC methylation were decarbonylated and deaminated into uracil. This was combined with bisulphite data in subtraction-based mapping to assess 5fC methylation.

The TrueMethyl Array fC workflow was performed by adding a FormylC Control to the DNA. The DNA was mixed with Magnetic Bead Binding Solution 1 and incubated for 20 minutes at room temperature. The solution was precipitated in the Magnetic Bead Rack (5 minutes) and washed 3 times with 80% Acetonitrile. The beads were left to dry, resuspended in Denaturing Solution, and incubated at 37˚C for 30 minutes. The Beads were then pelleted using the magnetic bead rack, the supernatant transferred to new tubes held on ice, before the Reductant Solution was added, mixed and incubated in darkness for 1 hour. The Bisulphite Reagent solution

was then added and the mixture was denatured (95˚C for 5 minutes) then incubated (60˚C for 20 minutes) repeatedly for 6 cycles in a thermocycler. After incubation/denaturing Magnetic Bead Binding Solution 2 was added and incubated at room temperature for 20 minutes. The beads were pelleted in the Magnetic Bead Rack for 30 minutes (or until solution was clear), supernatant aspirated, and beads washed with 70% Ethanol. Beads were resuspended in Desulphonation Buffer and pelleted by magnetic rack once more for 5 minutes. Two more 70% Ethanol washes were performed and the pellets air dried for 15 minutes. Elution buffer or ultra-pure water were added to the beads and incubated at room temperature for 20 minutes followed by pelleting in the Magnetic Bead Rack. Supernatant was transferred to a new tube and stored at -20˚C long term.

Methylation Assisted Bisulphite (MAB) conversion to assess 5fC/CaC

MAB conversion relied on bisulphite conversion and allowed for the assessment of 5fC/CaC DNA methylation. MAB conversion treated dsDNA with a CpG Methyltransferase *M.SssI* and S-Adenosyl Methionine (SAM) prior to bisulphite conversion. This methylated any previously unmodified/unmethylated CpG sites in the DNA to 5mC methylated CpG sites. Therefore, only 5fC and 5CaC methylation was converted into uracil during oxibisulphite/bisulphite conversion. This was again combined in subtraction-based mapping to assess 5fC and 5CaC methylation collectively.

CpG 5mC methylation of DNA was performed by diluting SAM stock to 1600µM, and combining with nuclease free water, NEBuffer 2, DNA, unmethylated λ-phage spike in control, and *M.SssI* methylase. This solution was mixed via pipette and incubated at 37˚C (one hour). The reaction was terminated by incubating at 65˚C for 20 minutes. The methylated DNA was then taken forward to the bisulphite conversion workflow.

Quality control assessing conversion efficiency of MAB converted samples was performed by the amplification of (CpG 5mC methylated) λ-phage spike in control using bisulphite λ-phage primers. This amplicon was processed following standard amplification, gel electrophoresis, sanger sequencing with reverse BS λ phage primer

and comparing amplicon with known λ-phage genome. Where CpG 5mC methylation by *M.SssI* and bisulphite conversion was successful, no CpG cytosines would be reported as thymines. Where CpG methylation by *M.SssI* failed, CpG sites would be indicated as thymines instead of cytosines. Where bisulphite conversion failed, non-CpG cytosines would not be converted to thymines.

Generation of λ Phage Spike in controls

Bisulphite converted λ phage spike-in controls were generated by amplifying the λ phage genome with the λ phage primer set. This amplicon was tested by performing agarose gel electrophoresis, and then purified using ExoSAP-IT and quantified through Nanodrop. The purified amplicon was then methylated using the *M.SssI* enzyme. A small amount of the methylated amplicon was then bisulphite converted and the further used as template in bisulphite PCR, using the nested BS λ phage primer set. The resultant bisulphite converted amplicon is once again tested through gel electrophoresis, ExoSAP-IT purified, and sent for sequencing using the reverse BS λ phage primer. Unconverted methylated amplicon was eligible for use as a λ phage spike-in control for bisulphite/oxibisulphite conversion and quantification of 5mC and 5hmC. Unconverted and unmethylated amplicon was used as a λ phage spike-in control for MAB-Seq quantification of 5fC/CaC.

2.2.5 Preparation of samples and sequencing

Sanger Sequencing

Sanger sequencing was performed to assess either genotype at variant sites, or semi-quantitative methylation analysis using cloning to analyse a sample multiple times by an inhouse service or via Eurofins & SourceBioscience sequencing services. ExoSAP-IT purified PCR product was diluted to appropriate concentrations dependent on expected PCR product size. The amplified product and an aliquot of the sequencing primer used in PCR amplification of the product was submitted to internal facilities or sequencing company.

For all sequencing services, sequencing reactions were prepared (comprised of Big Dye, Sequencing Buffer, diluted amplicon, primer, and HPLC grade water) and run on a thermocycler (25 cycles of; 96˚C for 30 seconds, 50˚C for 15 seconds, 60˚C for 4 minutes) in 96 well PCR plate. Sequencing products were then combined with precipitation mix (NaOAc pH 4.6, pure ethanol and HPLC water), and centrifuged at 3000 RCF (45 mins). The plate was inverted, placed on a towel and spun at 300 RCF for 1 minute. 70% Ethanol was added to each sequencing well and spun at 3000 RCF for 10 minutes. This was followed by an additional inverted spin before being left at room temperature for 10 minutes. The DNA was then reconstituted in formamide buffer for capillary electrophoresis. The solution was heated to 98˚C for 2 minutes and then held on ice for 2 minutes before being centrifuged at 2000 RCF for 1 minute. The plate was then placed in the genetic analyser and processed through capillary electrophoresis.

Sequencing reads were output as ABI format chromatograms. These files were interpreted using programs such as *ab1peakreporter* or *FinchTV*. Sequences were quality checked visually by peak size and also aligned to expected sequences using Nucleotide Blast (blastn) and Clustal Omega. Quality control of samples up until this point was performed using *fastQC* software (Babraham Bioinformatics) and the *ab1peakreporter* tool.

Pyrosequencing

Pyrosequencing analysis to assess methylation was performed by staff at the Wellcome Trust Clinical Research Facility (WTCRF) in Edinburgh using a PyroMark Q24 MDx instrument. Genotyping for the majority of samples was not performed by pyrosequencing. However, initially genomic primers and gDNA were sent to the facility to optimise the process as well as converted DNA and bisulphite primers.

In preparation for sequencing, the converted DNA was amplified by PCR using the pyrosequencing primers. PCR conditions followed the recommended QIAGEN protocol which combined QIAGEN CoralLoad PCR Buffer/PCR Buffer, $MgCl_2$, dNTP, forward and reverse primers, *Taq* Polymerase, template DNA and distilled water.

Running conditions were as follows: 95˚C for 15 minutes initial activation; 45 cycles of [94˚C for 30 seconds, [annealing temperature 5˚C below T_m of Primers] for 30 seconds, 72˚C for 30 seconds]; final 72˚C extension step for 10 minutes (from the QIAGEN PCR Handbook). After PCR, amplicons were run on an agarose gel to confirm amplification of specific product. If the gel indicated poor/non-specific product, additional optimisation steps were performed, e.g. varying concentrations of template DNA and primers, adding additional $MgCl_2$ and extending the number of cycles used in amplification.

PCR amplicons were bound to streptavidin coated sepharose beads by mixing the amplicon, Sepharose Bead Solution, and Binding Buffer together with high purity water in a PCR plate. The plate was sealed and agitated at 1400rpm for 10 minutes. Sequencing primer diluted in Annealing Buffer was prepared and dispensed into PyroMark Q24 sequencing plate wells. The DNA strands were then separated and the samples released into the Q24 plate through vacuum. The PCR plate was placed in a vacuum pump to sediment the sepharose beads onto Filter Probes, and aspirate all solution. The Filter Probes were then washed in 70% Ethanol for 5 seconds, Denaturing Solution for 5 seconds, Wash Buffer for 10 seconds, and finally drained of fluid before the beads were transferred to the PyroMark Q24 plate and into the sequencing primer/Annealing Buffer mixture. The Q24 plate was then heated at 80˚C for 2 minutes and cooled at room temperature for 5 minutes. The Q24 plate and a PyroMark Gold Q24 Reagent Cartridge were loaded into the PyroMark Q24 MDx instrument and sequencing performed. Allele Quantification (AQ) or CpG analysis reports targeting the regions of interest were outputted as PDFs.

2.2.6 Digital Droplet PCR (DD-PCR)

cDNA synthesis & DD-PCR

DD-PCR was performed on *BDNF* transcripts to analyse transcript expression in different clinical samples. cDNA synthesis was performed using the RevertAid First Strand cDNA Synthesis Kit. In summary, 1µg purified extracted RNA was mixed with Protoscript II Buffer, DDT, Protoscript II Reverse Transcriptase, dNTP, RNase

inhibitors, primers and nuclease-free water. This mixture was incubated at 42˚C for 1 hour followed by 65˚C for 20 minutes. RevertAid Random primers (Random Primer Mix also required a 25˚C incubation for 5 minutes prior to the 42˚C incubation) and oligo dT primers were used in amplification. To assess conversion efficacy, *Minus Reverse Transcriptase* was used as a negative control, and Foetal RNA was used as a positive control for each conversion. cDNA was also tested for DNA contamination by performing a PCR using generic GAPDH primers.

After cDNA synthesis was complete, the cDNA was combined with forward and reverse probes, DD-PCR Supermix for Probes and ultra-pure water in a master mix. Serial dilutions of the probe were analysed to determine the best concentration to observe differences in expression. Dilutions and controls were transferred to a Droplet Generator 8 well Cartridge, Droplet Generation Oil added and droplets generated using a QX100 Droplet Generator. Samples were transferred to a new 96 well plate and foil heat sealed. Plates were run in a thermal cycler with the following conditions: 95˚C for 10 minutes initial incubation; [94˚C for 30 seconds, 59˚C for 1 minute] repeated for 41 cycles, and a final 98˚C incubation for 10 minutes. Plates were then read on QX100 Droplet Reader instrument. Threshold values were calculated automatically by the instrument (using *TBP* probes as a reference; see section 2.2.1). Where threshold values could not be automatically calculated, they were manually set by calibrating for the negative and positive controls. Minus reverse transcriptase was used as a negative control as it showed no amplification/expression and foetal RNA was used as a positive control in high concentrations to conserve clinical RNA. Expression data were expressed relative to *TBP* expression and were exported to *.csv* for analysis.

2.3 Data Analysis

2.3.1 In vitro data analysis

Sanger Sequencing genotyping

Genotyping was originally performed in both Sanger and Pyrosequencing to confirm that no discrepancies in results were present. Once this was confirmed all further genotyping was performed by Sanger sequencing. Sanger sequencing analysis of genotype across sites of interest was performed manually by comparing sequencing ABI files. Variant sites were identified in chromatograms and inspected for peak height.

Analysis of methylation through Sanger Sequencing

DNA methylation was analysed using Sanger sequencing data through semi-quantitative means, with or without cloning. Using cloning, methylated amplicons for each sample were cloned, extracted from bacteria and sequenced from multiple colonies. The base calls at the variant sites for these sequences were compared and a ratio of methylated (C calls) vs unmethylated (T calls) were made and used to make a percentage of methylation at the site. This increases the reliability of the methylation analysis compared to when investigating a single sequence by increasing technical repeats.

When investigating sequences without cloning, trace files from Sanger-sequencing of (O)BS converted template from cloned sequences, were analysed using Thermo Fischer's *ab1PeakReporter* tool. In brief the *ab1PeakReporter* tool used files that were converted into sequence [.csv] files and analysed them for 5(h)mC epialleles. Semi-quantitative methylation analysis was performed by inferring methylation levels through peak height ratios at epialleles. Bisulphite (and other) conversions result in uracil bases in the sequence which were interpreted in sequencing as thymine. The filtered ratio of detectable C:T peaks over the locus of interest as well as a 7 base frame (1 codon before and after the locus of interest as reference) were compared and given an associated phred score (indicative of sequence quality) to give an average of methylation over loci investigated.

Analysis of methylation through pyrosequencing

Using pyrosequencing data for each site of interest, methylation levels were determined as a percentage based on the number of transcripts that had reported cytosine or thymine bases. Pyrosequencing has a reported 5% error margin so data reporting below 5% methylation were considered unreliable. Pyrosequencing data were assembled as a *.csv* file containing the reported methylation level and type at the variant site of interest, and quality report level generated by the PyroMark Q24 Software. Runs that were deemed "abnormal" or "failed" were excluded from data analysis or repeated where possible.

To calculate actual levels of methylation at each epiallele subtraction-based mapping was used. 5mC abundance values were the reported methylation or C levels in (O)BS converted DNA samples. 5hmC levels were calculated by [BS – OBS] levels. As MAB conversion and sequencing indicate unmethylated cytosines, 5mC, and 5hmC methylation were reported as cytosines, and 5fC/5CaC methylation was reported as thymines. 5fC/CaC could only be examined in homozygous C genotypes due to the selective nature of the CpG methyltransferase used in MAB-seq requiring cytosines. Hence if a non-cytosine allele is present the methyltransferase will not methylate the site and it will falsely report 50 to 100% 5fC/CaC methylation, for heterozygotes and homozygous non-cytosine alleles, respectively.

For some analyses, methylation was quantified at T homozygotes to investigate how reliable quantification was and determine any false positives. Because of the nature of subtraction-based mapping, all methylation forms (including unmodified cytosines) must add to 100%. If methylation data added up to over 100% or below 0% the data were excluded.

2.3.2 Statistical analysis of methylation and gene expression data

Statistical analysis of Pyrosequencing data

Statistical analysis of pyrosequencing data was performed by combining methylation levels into variables both individually and as total methylation observed. Because

5fC/CaC methylation data can only be observed at homozygous C allele genotypes, 5(h)mC data (all genotypes) and 5(h)mC/5fC/CaC data (Homozygous C genotypes) were processed separately. Genotype, minor allele carrier status, tissue type, and clinical diagnosis variables were matched along with methylation levels in a *.csv* file.

The relationship between 5mC and 5hmC methylation across genotype as a single variable was analysed using ANOVA for the *BDNF* and *APOE* analyses. The ANOVA compared 5mC or 5hmC methylation across genotypes at a single variant site.

Initially statistical analysis for the tissue and disease variables (also including genotype) was attempted using parametric and non-parametric ANOVAs, however due to the non-factorial and unmatched design of the data, many variables were non-independent. All other analyses for *BDNF* tissue and disease variables was performed by Linear Mixed Effects Model (LMEM), as a derivate form of General Linear Models (GLM) (Crawley, 2013). This allowed for the analysis of non-independent variables such as disease, tissue, and genotype across different models. *APOE* data were not suitable for ANOVA or LMEM analysis as the sample sizes were too small, resulting in factor reduction and low power. Analyses for *BDNF* were performed in R studio using the lme4 package, and entailed the following;

1. A “complete” linear model with all variables and interaction terms was designed.
2. The complete model was compared against models where one or more variable/interaction term had been removed, via ANOVA. Normality and homoskedasticity testing were performed on the residuals (obtained through ANOVA) to ensure distributions were appropriate for the model.
3. The models with the highest significance in the X^2 comparisons were taken forwards, and step two repeated until a “minimum adequate” model was reached. The X^2 statistic and P value for all variables were recorded to determine the effect of the variable in the model.

However, as there was strong collinearity between the Disease and Genotype data, the LMEM analysis often resulted in factorial reduction by the model. This is primarily

due to the Alzheimer's disease samples, of which almost all have a C/C genotype. To overcome this issue, 3 separate models were designed to assess the effect of the variables as best as possible:

1. A "complete" model with no interaction terms, or tissue variables: *lmer(m5hmC~Disease+rs6265+(1|Samples), data=BDNFrs6265)*
2. A model investigating Tissue and Disease terms and interactions, but excluding Alzheimer's disease samples: *lmer(m5mC~Disease+rs6265+Tissue+Tissue:Disease+Tissue:rs6265+(1|Samples), data=Imperial)*
3. A model investigating the Alzheimer's disease samples only, using Tissue and Genotype terms: *lmer(m5mC~Tissue+rs6265+(1|Samples), data=OPTIMA)*

Each of these models were made for 5mC, 5hmC, or 5fC/CaC analysis, to compare variables within methylation types, and determine if methylation changes with disease, tissue, or genotype. Unfortunately, there was not enough statistical power to include age and gender covariates into the analysis. An example of the R studio script used to create the minimum adequate model and derive the effect of variables is shown below:

- model1<-lmer(m5mC~Disease+rs6265+(1|Samples), data=BDNFrs6265)
- summary(model1)
- model2<-update(model1,~.-Disease)
- model3<-update(model1,~.-rs6265)
- anova(model1,model2)
- anova(model1,model3)
- par(mfrow=c(1,3))|qqnorm(residuals(model1))
- hist(residuals(model1))
- plot(fitted(model1),residuals(model1))
- qqnorm(residuals(model2))
- hist(residuals(model2))
- plot(fitted(model2),residuals(model2))
- qqnorm(residuals(model3))
- hist(residuals(model3))

- plot(fitted(model3),residuals(model3))

Because analysis of genotypes with methylation (excluding other factors such as diagnosis, tissue, age or gender) reported matched, factorial design, ANOVAs were used to compare methylation between genotypes.

Digital Droplet PCR analysis

As *TBP* probes were used as a reference for expression, all *BDNF* expression data were expressed relative to expression of *TBP*. *Csv* files containing the transcript expression from different samples were matched to pyrosequencing quantitative methylation data and phenotypical data for the same samples (tissue, clinical diagnosis, age of death, gender) and imported into SPSS for analysis.

Initially, *BDNF* expression was compared between transcripts and then with tissue, disease types, and genotype. Normality (Shapiro-Wilk) and homogeneity of variance (Levene's) tests were used to determine the distribution of the expression data and appropriate statistical testing used. ANOVA and Kruskal-Wallis tests were used for parametric and non-parametric distributions, respectively. Post Hoc analysis with Dunn's was performed to account for multiple testing, as group sizes were equal.

Transcript expression data were then compared to 5mC, 5hmC, and 5fC levels individually, and between 5mC and 5hmC methylation levels, tissue type, disease type, and genotype. Comparisons used Pearson's Correlation Coefficient or Kendall-Tau correlation analysis for parametric and non-parametric data, respectively. Covariate analysis was also performed using ANCOVA or Spearman's non-parametric partial correlation analysis (see below) for parametric and non-parametrically distributed data, respectively. Graphs and graphics were generated using GraphPad Prism.

Spearman's non-parametric partial correlation analysis script

Kendal-Tau correlation could not compare multiple comparisons and Quade's analyses could only analyse one categorical and one continuous variable. For this

reason, covariate analyses to observe the effects of age and gender covariates were performed using Spearman's non-parametric partial correlation analysis. Analyses were performed in SPSS using the following script:

- NONPAR CORR Meth Transc Age Gender
  /MISSING = LISTWISE
  /MATRIX OUT(*).
- RECODE rowtype_ ('RHO'='CORR').
- PARTIAL CORR Meth Transc BY Age/Gender
  /significance = twotail
  /MISSING = LISTWISE
  /MATRIX IN(*).

Where "Meth" and "Transc" refer to the methylation type (5mC, 5hmC, or 5fC/CaC) and transcript (Sense, AS-201, or AS-204) being analysed, respectively. "Age/Gender" refers to age, gender or both variables being analysed together.

2.3.3 Statistical analysis of Digital Droplet PCR expression and pyrosequencing data

Digital Droplet PCR expression *.csv* files coded each sample with one sense and two anti-sense *BDNF* transcripts as variables. Data were stored as relative expression to expression of the *TBP* housekeeping gene at 3 concentration levels; "minimum", "actual", and "maximum". These concentration values were included to give a relative point of transcript concentration in order to quantify expression in samples. For analysis only the "actual" expression values were used. Expression data for each sample were standardised ([*BDNF* transcript expression value]/[*TBP* expression value]). Normality tests were performed, and Ln(X+10) transformation was performed to improve gaussian distribution and retain data for analyses. Statistical Analyses were performed in GraphPad Prism and SPSS.

2.3.4 Analysis of cognitive performance using UK10K Data

Association analysis between Asperger's, autism spectrum disorders, schizophrenia and schizoaffective disorders and the *BDNF* rs6265 genotype were performed using UK10K data sets acquired from the Wellcome Trust and King's College London UK10K Project. All access to data was given with permission from the UK10K project. Genotype data were computed using Next Generation Sequencing on genome build hg19. A list of the UK10K data sets accessed for the *BDNF* analyses are shown in Table 2.13.

The UK10K genotype datasets were used to perform an association study of the *BDNF* rs6265 variant. All cohorts listed in Table 2.13 were accessed with permission. The VCF sequencing files containing the target variant *BDNF* rs6265 (chr11: 27,679,916 (hg19)) were downloaded and decrypted from the European Genome-Phenome Archive. Genotypic and phenotypic data for the UK10K_COHORT datasets were extracted using the following Linux BASH script.

- java –jar EgaDemoClient.jar –p username password –rf filename –re decryptionkey –label request_filename
  # Where "filename" is the EGAF file name, "decryptionkey" is the unique identifier you use for secure encryption/decryption.

- java –jar EgaDemoClient.jar –p username password –dr request_filename

- java –jar EgaDemoClient.jar –p username password –dc path/to/file –dck decryptionkey

Due to the size and nature of the resultant VCF files, the files were then "cut" to extract a 1kb region surrounding relevant site of interest from the genotype data and convert them into *.tab* files using VCFtools in Linux BASH script:

Dataset	N	Phenotype	Sequence Type	Dataset ID	EGA Project and ethics approval ID
UK10K_NEURO_ASD_ GALLAGHER	75	ASD	Exome	EGAD00001000436	EGAS00001000112
UK10K_NEURO_ASD_SKUSE	334	ASD	Exome	EGAD00001000614	EGAS00001000114
UK10K_NEURO_IOP_COLLIER	172	SCZ	Exome	EGAD00001000442	EGAS00001000121
UK10K_NEURO_MUIR	334	ASD, SCZ, ND, ID	Exome	EGAD00001000443	EGAS00001000122
UK10K_NEURO_EDINBURGH	233	SCZ	Exome	EGAD00001000438	EGAS00001000117
UK10K_NEURO_ABERDEEN	387	SCZ	Exome	EGAD00001000433	EGAS00001000109
UK10K_NEURO_GURLING	48	SCZ	Exome	EGAD00001000440	EGAS00001000225
UK10K_COHORT_IMGSAC	113	ASD	Exome	EGAD00001000441	EGAS00001000120
UK10K_COHORT_MGAS	97	ASD	Exome	EGAD00001000613	EGAS00001000113
UK10K_NEURO_UKSCZ	552	SCZ	Exome	EGAD00001000430	EGAS00001000106
UK10K_NEURO_FI	84	ASD	Exome	EGAD00001000435	EGAS00001000110
UK10K_NEURO_FSZ	119	SCZ	Exome	EGAD00001000615	EGAS00001000111
UK10K_NEURO_FSZNK	283	SCZ	Exome	EGAD00001000439	EGAS00001000123
UK10K_OBESITY_TWINSUK	403	Control	Exome	EGAD00001000431	EGAS00001000242
UK10K_COHORT_TWINSUK	1854	Control	Whole Genome	EGAD00001000194	EGAS00001000108

Table 2.13 UK10K cohort genotype datasets used to examine BDNF rs6265 genotypes in association studies. The dataset name, number of individuals (N), and Phenotype (non-clinical shown as Control, clinical shown as; ASD for autism spectrum disorder and Asperger's disease, and SCZ is schizophrenia and schizoaffective disorders) are shown. The sequences targeted (Exomic or Whole genome) and the Data Set ID's, and EGA Project and ethics approval ID are also provided.

- gunzip /path/to/VCFfile
  #Where “path/to/VCFfile” is the directory pathway to the file downloaded above

- cd /users/downloads/vcftools

- ./bin/vcftools –vcf /users/braingenomics/downloads/filename.vcf –chr 11 –from-bp startbp –to-bp endbp –out /users/desktop/newfilename.vcf –recode
  #Where “filename.vcf” is the file downloaded above, and “startbp” and “endbp” refer to the starting base-pair and ending base-pair number of the region being targeted for extraction

- Gzip /users/braingenomics/desktop/newfilename.vcf.recode.vcf

- Export
  PERL5LIB=/users/braingenomics/downloads/vcftools/lib/perl5/site.perl | zcat < /users/desktop/newfilename.vcf.recode.vcf.gz

- ./perl/vcf-to-tab > /users/desktop

.tab files were converted into *.xlsx* files. For the TWINSUK dataset genotype and phenotype data were matched in Excel using the VLOOKUP function.

The UK10K cohort data were then grouped based on diagnosis; ASD, SCZ, and Controls. The UK10K_NEURO_MUIR data were divided into subgroups to ensure only relevant diagnosis were put into each group, all other groups only had one diagnosis. Genotypes for each group of cohorts were split into sense strand (5’ to 3’) *BDNF* rs6265 major and minor alleles (G or A, respectively) and by diagnosis (Control or ASD/SCZ). Pearson’s Chi-squared analyses were performed to assess allele frequency using Rstudio:

- a = [a], b = [b], c = [c], d = [d], | R1 = c([a], [b]) | R2 = c([c], [d]) | Matriz = matrix(data = c(R1, R2), nrow = 2, ncol = 2) | Matriz
  #Where [a], [b], [c], and [d], refer to the frequencies of; Non-control T alleles, Non-control C alleles, Control T alleles, and Control C alleles, respectively.

- chisq.test(Matriz, correct = FALSE)

Resultant output contained the x^2 statistic and associated p-value. Additional commands were used to get Odds Ratio and Upper/Lower Confidence limits;

- p1=a/(a+b) | p2=c/(c+d) | OR = p1/p2 | SE.OR = sqrt((1/a+1/b+1/c+1/d)) | LCL = OR * exp(-1.96*SE.OR) | UCL = OR * exp(1.96*SE.OR)
  #Allows for the calculation of OR and Upper/Lower confidence limits using predefined terms above

- OR | LCL | UCL

2.3.5 Bioinformatic tools

Resource	Website address
Ab1 Peak Reporter	http://apps.lifetechnologies.com/ab1peakreporter.
Bisulphite Primer Seeker	https://www.zymoresearch.com/pages/bisulphite-primer-seeker
MethPrimer	http://www.urogene.org/methprimer/
Primer3	http://primer3.ut.ee/
UCSC Genome Browser	https://genome-euro.ucsc.edu/
Ensemble	http://www.ensembl.org/index.html
FinchTV	http://www.geospiza.com
Blast	http://ncbi,nlm.nih.gov/BLAST/
Clustal Omega	http://www.clustal.org/
FastQC	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Table 2.13 A list of bioinformatic tools used in this study. The resource and its corresponding web address are shown.

Chapter 3: Optimisation of techniques to examine DNA methylation in blood and brain

Acknowledgements

Several people assisted in the optimisation of the methods and protocols. Daniel Ellis (QIAGEN) assisted with the optimisation and design of the early iterations of pyrosequencing primers designed to target *BDNF*. Jonathan Rowlinson and Dr Teresa Pereira Raposo, University of Nottingham, aided in the optimisation of homogenisation and extraction of DNA from FFPE and snap-frozen tissues respectively. Maria Isabelle Haig, University of Nottingham, assisted in cloning of bisulphite DNA in semi-quantification. Rhian Warman, University of Nottingham, contributed to the optimisation of Sanger-sequencing primers to target *DNMT3L*.

3.1 Workflow of the optimisation of techniques and DNA extraction

Between 2015 and 2019, many emergent commercial technologies such as the CEGX TrueMethyl kits became available for the study of DNA methylation sites. However, although available, the optimisation of such methods using specific biological tissue types and examining specific genomic loci still required development. This chapter will describe the processes used to optimise new methods to quantify methylation at target epialleles from rat and human blood and brain tissues. A diagram showing the workflow pipeline for this project is shown in Figure 3.1.

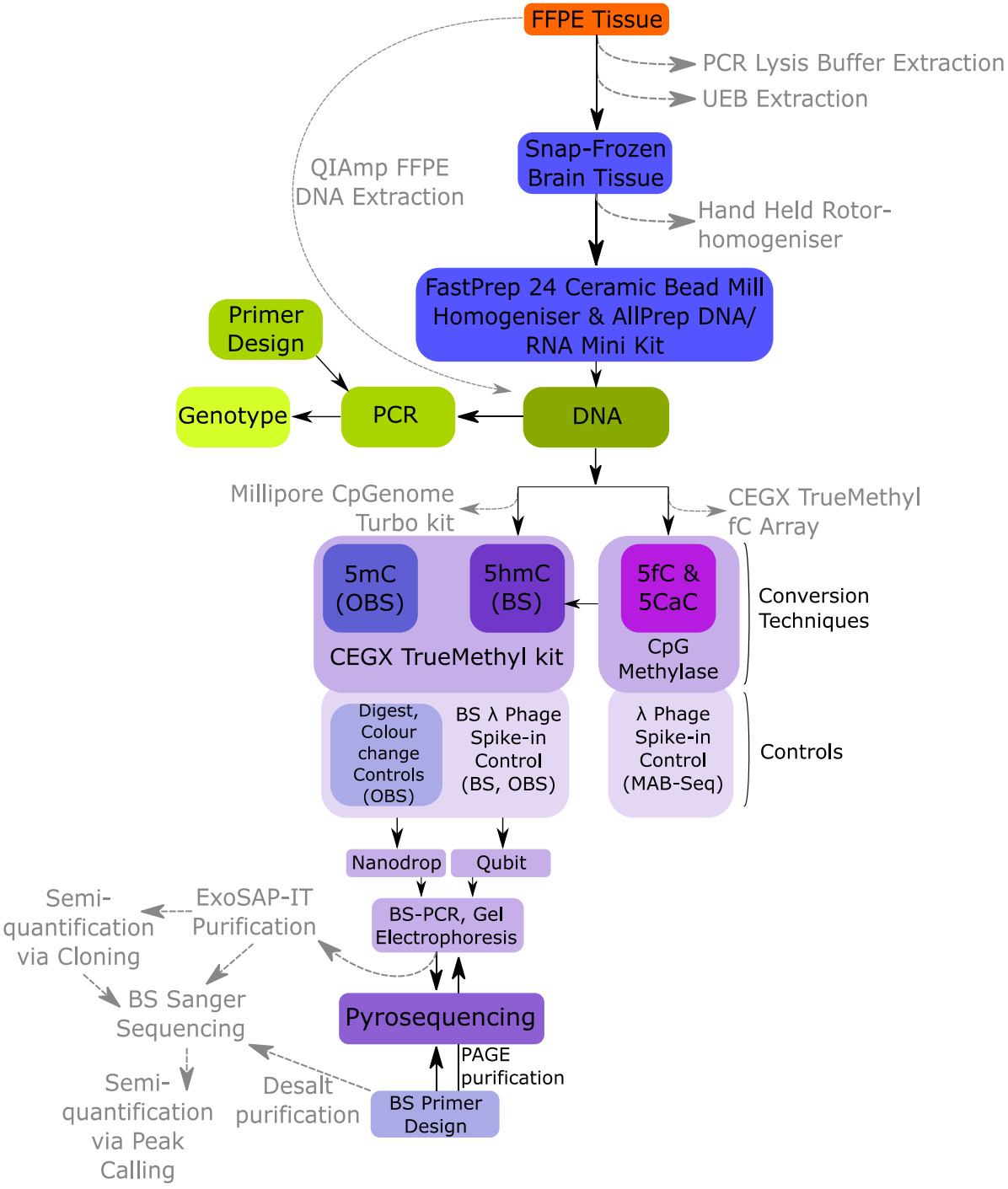


Figure 3.1 Pipeline of methods and techniques used to assess DNA methylation at risk epialleles. Dashed/grey lines or text indicate methods that were attempted but not continued. Extraction methods are detailed in orange/blue. The DNA pipelines are shown in green. Methods for the analysis of methylation are detailed in purple along with techniques and associated control assays. The goal of this pipeline was to analyse 5mC and 5hmC in all risk allele heterozygous and homozygous genotypes or 5mC, 5hmC, and 5fC/CaC methylation in C homozygotes. UEB refers to universal extraction buffer, FFPE refers to formalin fixed, paraffin embedded tissue, BS refers to bisulphite, OBS refers to oxibisulphite, MAB-Seq refers to methylation assisted bisulphite sequencing, PAGE refers to polyacrylamide gel electrophoresis.

3.1.1 Extraction of genomic material from snap frozen tissues

DNA and RNA extraction using the AllPrep DNA/RNA Mini kit

Extraction of DNA and RNA was first performed using 46mg of fresh frozen rat brain by homogenising tissue, and incubating with phenol-chloroform on ice. The homogenate was then taken forwards to the AllPrep extraction method, extending the final incubation of elution buffer by 5 minutes. To study the effects of homogenisation on DNA yields pre-sliced tissue by cryosection equipment was used. Tissue slices were vortexed in RTL Buffer until they had partially dissolved. DNA concentration was found to be 60ng/µl ($260/280_{nm}$ = 1.91nm, $260/230_{nm}$ = 1.59nm) and the OD values were close to the ideal 1.80nm absorbance. RNA concentrations extracted were substantially higher at 940.8ng/µl ($260/280_{nm}$ = 2.15nm, $260/230_{nm}$ = 2.06nm). The rat DNA was PCR amplified using the rat prolactin primers and showed specific product bands with high intensity fluorescence and little non-specific product, indicating DNA extracted through this method was suitable for conversion to assess methylation.

Tissue and Cell type variations in genomic material

Using the adjusted AllPrep protocols, extractions were performed on the Imperial brain bank Parkinson's disease (PD) clinical sample tissue sections from cerebellum, hippocampus and frontal cortex. The resultant yields varied with DNA concentrations ranging from 23.4ng/µl to 259.5ng/µl, and purity was high with all samples $260/280_{nm}$ around 1.8nm. DNA $260/230_{nm}$ values were all below 1.0nm, which indicated contaminants absorbing at the 230nm range which was most likely phenol. Despite the variation in concentration, all samples processed yielded sufficient DNA for genotyping and conversion. RNA concentrations similarly varied with the lowest concentration at 1.7ng/µl and highest at 44.7ng/µl. Similarly, $260/280_{nm}$ values were all around 2.0nm while $260/230_{nm}$ values indicated some phenol contamination. The higher concentration yields were sufficient for cDNA synthesis. However, the lower concentrations produced insufficient material.

The same protocol was used to extract from the Imperial brain bank dementia with Lewy bodies (DLB) clinical samples. Yields obtained from these samples were substantially higher than that of the Parkinson's disease samples with DNA concentrations ranging from 139 to 665ng/µl. The purity of the samples was slightly lower than that of the PD samples with $260/280_{nm}$ values between 1.64 to 1.85nm. DNA $260/230_{nm}$ values once again indicated phenol contamination. Amplification of both the PD and DLB samples DNA with AS *BDNF* primers showed specific amplification with little non-specific product. RNA values were similarly high in concentration (between 110 and 384.1ng/µl) with $260/280_{nm}$ values all around 1.8nm, an indication of slight phenol contamination. These yields supported the possibility that the variation in concentrations (as well as reported purity) appears to be sample and tissue specific. Both PD and DLB samples bisulphite (BS) converted well and post-conversion amplification with antisense (AS) BS *BDNF* primers indicated specific amplicons.

3.2 Optimisation of primers for sequencing

Sanger sequencing primer optimisation followed the standard methods with many redesigns performed to navigate poly A regions, GC repeats, nested design, as well as modifying salt concentration and melting temperatures. Some regions such as *BDNF* loci were comparatively easier to amplify using standard primers than others such as *APOE* which is reflected in the number of redesigns of primer sets in difficult regions, i.e. one primer pair was designed to capture genotype at AS-*BDNF* rs6265 variant, while 5 primer sets were designed to capture genotype at sense *APOE* rs7412 & rs429358 variants.

While pyrosequencing is relatively common, the use of pyrosequencing to quantify methylation at base-resolution is not as common. Because of this, optimisation was essential to quantifying methylation in the samples. Pyrosequencing primer optimisation was automated using the PyroMark Assay Design Software (QIAGEN). Inputting the base sequence surrounding the variant of interest was sufficient for workable primers covering the AS-*BDNF* rs6265 complementary CpG. However, the

APOE variant sites rs7412 and rs429358 proved more difficult in analysis due to high GC regions and homopolymers. The Yu lab designed methyl primer sets in order to assess the genotype and methylation status of the APOE CpG island, which contained the two variants of interest (Foraker et al., 2015). With permission, these primers were used in analysis of *APOE*. Additional assistance in primer design was given by QIAGEN representatives who also instructed on PCR optimisation to aid in pyrosequencing of samples. While designs were being optimised, alternative purification of the primers such as PAGE purified primers as opposed to desalt purified primers were also generated in order to assist the pyrosequencing chemistry which improved sequencing substantially.

3.3 Optimisation of (oxi)bisulphite conversion for the analysis of 5(h)mC

3.3.1 Kits to assess bisulphite and oxibisulphite conversion

Using oxibisulphite (OBS), and bisulphite (BS) conversion, 5mC and 5hmC were investigated. Bisulphite conversion assessed 5mC/5hmC and could not differentiate between the two. Oxibisulphite conversion assessed 5mC methylation, and subtraction-based mapping with bisulphite sequencing was used to determine 5hmC levels. In addition, different kits as well as different conversion workflows, such as the Millipore vs CEGX bisulphite conversion kits were followed. Figure 3.2 details two main forms of conversion used in this project after optimisation. Standard bisulphite sequencing maps 5mC and 5hmC methylation by converting unmodified cytosines into uracil with bisulphite treatment. Uracil bases are then read as thymine in sequencing, so any retained cytosines adjacent to a guanine in sequencing are 5mC or 5hmC methylated.

Oxibisulphite sequencing only maps 5mC directly by oxidising DNA prior to bisulphite treatment. Here, oxidation converts 5hmC into 5fC which is converted into uracil in bisulphite treatment and read as thymine in sequencing. Only 5mC which is oxidised to 5hmC will still be unconverted in bisulphite treatment, meaning that retained cytosines adjacent to guanines in sequencing are 5mC methylated. When combined

with bisulphite sequencing, oxibisulphite sequencing can differentiate between 5mC and 5hmC using subtraction-based mapping.

3.3.2 Bisulphite and oxibisulphite conversion to assess 5mC/5hmC

The Millipore CpGenome Turbo bisulphite modification kit was initially used to bisulphite convert samples to examine 5mC/5hmC methylation. Since the kit did not have any controls or methods to assess conversion efficiency these were designed (see Section 3.4). The kit although successful at conversions, showed variable yields and quality in conversion. Because of its inconsistency and long conversion times the CEGX conversion kits were preferable and used for conversion of clinical samples.

The CEGX TrueMethyl kit enabled 5mC and 5hmC methylation to be compared as it used BS and OBS pathways in parallel. An advantage of the CEGX TrueMethyl kits were the ability to process one sample in the bisulphite and oxibisulphite workflows simultaneously. Key observations/changes were made to the manufacturer's protocols in optimisation: The input DNA amount was increased above 400ng which increased total converted DNA yielded. The final elution incubation was extended by 10 minutes to increase yields marginally. Pelleting steps were still effective and reduced to 5 minutes (with the exception of the final pelleting step) to reduce time taken for the protocol. Typically, the workflow took 8 hours, however, the bisulphite conversion procedure could be left overnight. Splitting the workflow over two days allowed for tube & reagent preparation, as well as fewer errors. Contrary to kit protocol each bisulphite aliquot is sufficient to modify 11 samples (both bisulphite and oxibisulphite). This allowed the kit to process 128 (O)BS conversions or 64 samples, instead of the suggested 24 samples.

3.4 Validating conversion efficiency

3.4.1 Methods to assess conversion efficiency.

Bisulphite conversion of DNA using the CpGenome and TrueMethyl modification kits were initially validated using a methylated human DNA standard as a spike in control

(2% w/w). To validate conversion of the DNA, the spike in was amplified using KAPA HiFi *Taq* and supplied hMLH1 primers (Tm = 59˚C). However, due to difficulties in optimisation and evidence of inconsistencies being generated by the primer set when amplifying converted human DNA, λ phage spike-in controls were designed and used.

For the generation of CpG methylated λ phage spike-in controls, see Section 2.2.5. CpG methylated λ phage spike-in controls were spiked into the DNA at 2% w/w, and amplified after DNA conversion using the nested BS λ phage primer set. Sanger sequencing of amplicon confirmed conversion efficiency. Where conversion was successful all CpG sites remained unchanged, but all non-CpG site cytosines were converted into uracil and read as thymine in sequencing. The remaining non-CpG cytosines indicated incomplete conversion. The primer amplicon contained 20 CpG sites and 49 non-CpG site cytosines, thus each non-CpG cytosine represents close to 2% of potential conversions. Any conversion less than 94% effective (3 or more non-CpG cytosines unconverted) was repeated. Typically, one spike in was used per group of samples being converted in order to determine efficiency of conversion.

3.4.2 Validating oxibisulphite conversion

The oxibisulphite workflow of the TrueMethyl conversion kit contained qualitative controls to indicate failure in conversion as indicated by the colour of the solution after the oxidant solution has been added. The kit also provides a digest control which was spiked-in along with the λ phage spike-in control (2% w/w). The digestion control contains a 5'-TCGA-3' site that is specifically 5hmC methylated. In BS conversion this site is retained as a cytosine, however in OBS conversion the site is converted to uracil; giving 5'-TTGA-3'. The control was amplified using KAPA HiFi *Taq* and supplied Digestion Control Primers (Tm = 60˚C). The amplicon is incubated with *Taq$^{\alpha}$I* restriction enzyme (65˚C) which digests CpG sites in 5'-TCGA-3', and hence will only digest BS converted digestion control and will fail to digest OBS converted digestion amplicon. An example of digest amplicon gel is presented in Figure 3.3.

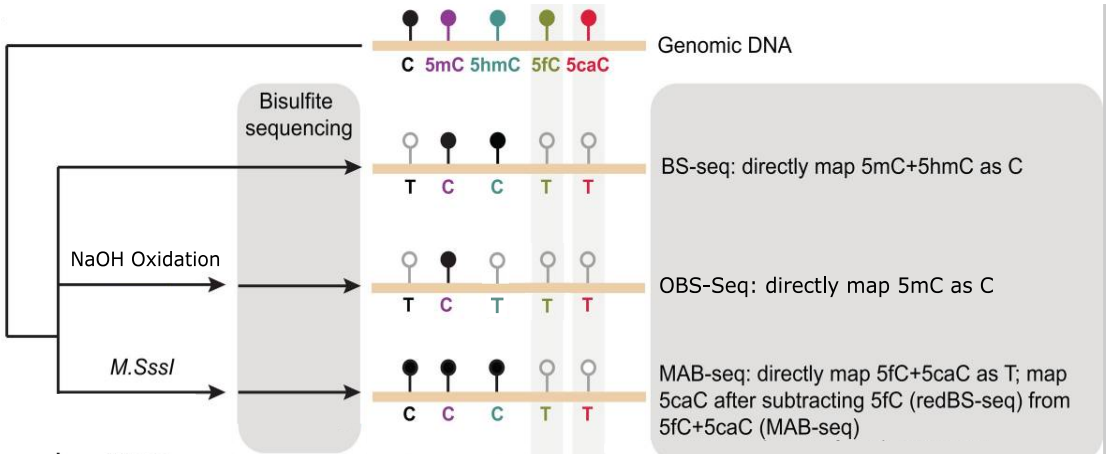


Figure 3.2 Bisulphite, oxibisulphite and MAB conversion methods. Standard bisulphite sequencing maps 5mC and 5hmC methylation. Oxibisulphite sequencing only maps 5mC directly and when combined with bisulphite sequencing can differentiate between 5mC and 5hmC using subtraction-based mapping. 5mC CpG methylating DNA with *M.SssI* enzyme prior to bisulphite conversion in MAB conversion will map 5fC & 5CaC as thymines during sequencing. This can be used in subtraction-based mapping with bisulphite and oxibisulphite sequencing. Diagram is modified from (H. Wu et al., 2016).

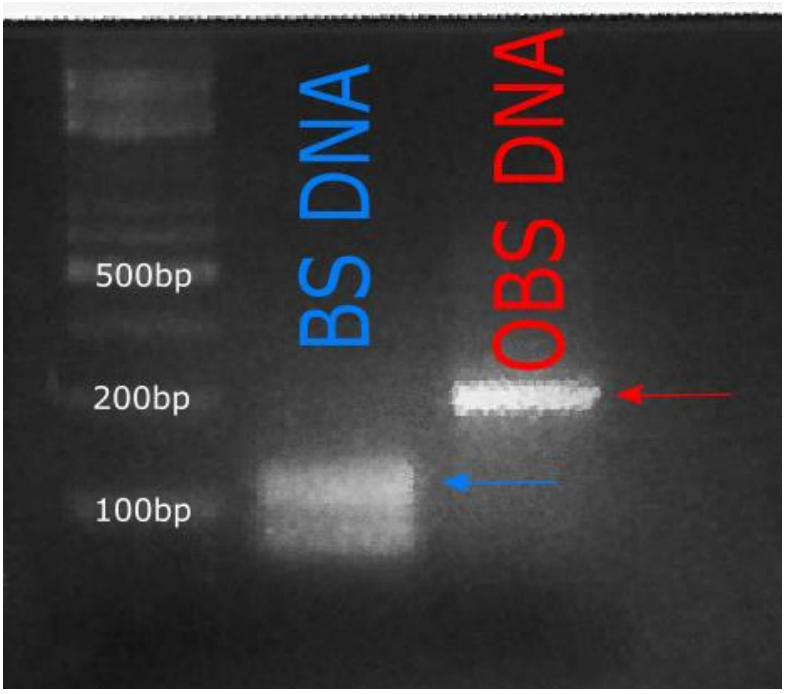


Figure 3.3. A gel scan of digested and undigested PCR products. 100bp ladder was used and shows a 100bp, 200bp and 500bp molecular sizes. Bisulphite and oxibisulphite converted digest control DNA amplified by digest primers are shown in blue and red, respectively. Both amplicons were treated with *Taqαl* digestion enzymes. Oxibisulphite DNA has failed to digest the 200bp product while bisulphite DNA has been digested and produced bands at around 100bp, indicating successful conversion of DNA.

Digestion protocol was optimised by making standard improvements to support enzyme activity. Initial failings in digests were determined to be due to contaminants in the DNA generated by extraction and conversion (phenol, sulphates, alcohol, excess salts etc). This issue was resolved through extra purifications such as a double ExoSAP-IT incubation. Where enzyme reactions were still failing and Nanodrop indicated the persistence of contaminants the QIAquick PCR Purification kit was used to purify the DNA prior to incubation with *Taq$^{\alpha}$I*.

3.4.3 Optimisation of 5fC and 5CaC Analysis

The decision to analyse 5fC/5CaC was made during the project and two main workflows were investigated. The CEGX TrueMethyl Array fC workflow was suggested as a non-commercial kit by CEGX technicians. In addition, published protocols to assess 5fC/5CaC used methylation assisted bisulphite (MAB) conversion (Neri, Incarnato, Krepelova, Parlato, & Oliviero, 2016). In this method, previously unmethylated CpG dyads in DNA are 5mC methylated with *M.SssI* methyltransferase prior to bisulphite conversion. Only 5fC and 5CaC methylation are converted into thymines during bisulphite treatment, meaning conversion into TpG dyads from CpG dyads in sequencing indicates 5fC/CaC methylation. This can be used in subtraction-based mapping with bisulphite sequencing to determine 5fC/CaC methylation at CpG sites. However, this method is unable to distinguish between the 5fC and 5CaC states (Neri et al., 2016; H. Wu et al., 2016). Both of these protocols were performed in tandem to assess 5fC/5CaC at methylated alleles.

3.4.4 CEGX TrueMethyl Array fC analysis of 5fC methylation

Analysis of 5fC was initially attempted using the CEGX TrueMethyl Array fC workflow. At the date of writing, this kit has not been commercially released. Analysis with this kit was undertaken in part to help the optimisation of the workflow, and was explicitly stated as an experimental protocol by CEGX during communications. An aliquot of the reductant solution required in the fC array was received along with a FormylC spike-in control and primer set which allowed for the analysis of 5fC using the CEGX TrueMethyl kits. Because the fC array methodology had not been

commercially released there were issues in conversion as the kit had not been optimised previously.

Changes made to the manufacturer's kit protocol were similar to the CEGX TrueMethyl oxibisulphite workflows and primarily focused on improving efficiency by reducing incubation times. The main issues identified in the conversion process arose with the reductant solution. Similar to the oxibisulphite workflow, the Array fC workflow treated the DNA with a reductant solution prior to the bisulphite conversion (redbisulphite conversion). However, where the oxidant solution of the OBS workflow was sensitive to alcohol, carbon dioxide, and acidic compounds, the reductant solution in the Array fC workflow was sensitive to all of these factors as well as light. Because of this, small alterations in timing and the environment affected the efficiency of conversion measurably. Moreover, similar to (O)BS conversions, the Array fC workflow typically would leave contaminants in eluted DNA as well as denature and shear fragments, which made PCR amplification of converted product difficult.

While concentrations observed in initial conversions were similar to that of previously observed OBS yields, whole brain DNA appeared to yield less in conversion than other samples. Moreover, the purity was variable with samples having an OD range of 0.7nm to 1.39nm indicating contaminants. Both the converted DNA and the FormylC control failed to amplify in PCR when using KAPA HiFi *Taq* and BS *BDNF*, or FormylC primers, respectively, after purification with ExoSAP-IT. A further purification attempt was made using the QIAquick PCR Purification kit, and once again amplified with KAPA *Taq* with no results.

Following advice from CEGX, PCR amplification with BS *BDNF* primers was attempted with VeraSeq polymerase. The BS *BDNF* primers were found to work optimally with the VeraSeq polymerase at 9% DMSO, with GC enhancer and double polymerase levels. However, when the array fC samples were PCR amplified using the VeraSeq polymerase/KAPA *Taq* and BS BDNF primers, the VeraSeq failed to amplify the DNA.

Subsequently, to avoid excessive consumption of DNA and the fC array reductant solution, the alternative MAB conversion method was assessed.

3.4.5 Methylation assisted bisulphite conversion and sequencing

As mentioned in chapter 2, the *M.SssI* enzyme is specific to CpG dyads and 5mC methylation by *M.SssI* outside of these motifs is extremely low (around 1.3% for CpH dyads)(Renbaum et al., 1990; H. Wu, Wu, Shen, & Zhang, 2014). Because of this, analysis of 5fC/CaC methylation using MAB conversion could only be performed at CpG sites generated by C or G genotypes. For example, the antisense complementary site of *BDNF* Val66Met rs6265 may be C or T in allele. If an individual is heterozygous (C/T genotype) or homozygous T (T/T genotype) at the antisense rs6265 site this halves or eliminates presence of CpGs, severely reducing methylation efficiency of *M.SssI*. Weighting the methylation reported based on the presence of CpGs was considered but not implemented due to the error carried forward from pyrosequencing and conversion potentially being increased due to the increased weight of C homozygous samples. In small sample sizes this could skew the representation of data by amplifying error and variation from C homozygotes twice as much as heterozygotes.

M.SssI methylase was used to 5mC CpG methylate cytosines in DNA (detailed in Section 2.2.5). The MAB conversion method was initially attempted using control whole brain DNA, to exclude variability generated by the extraction of DNA. Converted DNA showed good concentration despite having some contaminants (95.5ng/µl, 260/280nm = 2.22). Amplification of methylated and bisulphite converted λ spike-in controls with BS λ primers showed inefficient methylation (7/20 CpG sites reporting thymines and hence only 65% efficient conversion). The MAB conversion was repeated using a 2 hour final elution incubation on 4 control whole brain DNA samples which showed far better concentrations (3 to 6ng/µl). Amplification of these converted samples with KAPA *Taq* and AS BS BDNF primers showed specific amplicons and these were Sanger sequenced. Sequences indicated that 3 of 4 samples were efficiently converted (above 95% conversion) and that also

there was partial 5fC/CaC methylation at the *BDNF* rs6265 variant site as quantified by comparing peak ratios via the *ab1peakreporter* tool.

DNA extracted from 17 snap-frozen clinical tissue samples was converted. Methylation and conversion followed standard protocol. Amplification and sequencing of the λ phage spike in control indicated approximately 100% bisulphite conversion efficiency. However, it also indicated partial methylation by *M.SssI* with all CpG sites reporting cytosines except for 2 sites at the beginning of the sequence showing C/T heterozygous status when looking at C/T peak height (90% methylation efficiency).

In order to improve *M.SssI* methylation efficiency, λ phage amplicon (from λ phage primer amplification) was used to optimise current methylation protocol. Fresh stocks of SAM and *M.SssI* were used and incubated as per standard protocol with the λ phage amplicon. A standard mix was incubated as well as a mix containing twice as much SAM as standard, and both incubated for 4 hours before stopping the reaction. This was performed as SAM is the limiting factor for methylation efficiency in most reactions. After methylation, the amplicons were bisulphite converted, purified and sequenced. Once again conversion was approximately 100% efficient, and methylation by *M.SssI* was around 90% efficient in both cases with the only partial methylation occurring at the beginning of the sequence.

It was noted in the previous two λ phage sequencing reactions that all inefficient methylation occurred at the beginning of the sequence where the base-call quality was poor (phred score of 10-20). It was not determined as to why CpG methylation would be poorer at the start of sequences and was assumed to be due to poorer sequencing quality at the beginning of sequences. Once the poor sequence regions were removed from methylation efficiency calculations (excluding data where phred score < 30, as these regions would not be targeted in the quantification of methylation in pyrosequencing) the methylation efficiency increased above 95% and was deemed adequate for MAB sequencing. As all epiallele targets are typically 50-70bp away from the start of the sequence and with phred scores above 30, it was

unlikely that the methylation by *M.SssI* would poorly affect quantification at these sites. An example of the amplified sequence for the methylation assisted bisulphite sequencing conversion of λ phage DNA spiked into PD120 frontal cortex DNA is presented in in Figure 3.4.

Figure 3.4 Panel A compares the alignment of methylated λ phage DNA spiked into the PD120 frontal cortex DNA during bisulphite conversion with that of reference unconverted λ phage DNA. While the reference unmethylated and unconverted DNA retains non-CpG cytosines, the λ phage spike-in in PD120 only shows cytosines in CpGs, where all non-CpG cytosines have been read as thymines. Figure 3.4 Panel B shows the peak call file for the sequencing of the same sample. As the Phred score increases for the peak calls, the sequences show increasingly similar peak calls between the sample as the reference sequence. After approximately 130 bases the BS converted λ phage DNA spike-in DNA shows identical base calls to the reference sequence with the exception to non-CpG cytosines which had all been converted to thymines. All but one CpG site after base 130 in the converted spike in sample shows methylation by *M.SssI* and successful conversion.

DNA extracted from all clinical snap-frozen tissues were subsequently methylated and converted using the standard MAB conversion protocol. Amplification and sequencing of λ phage spike in showed both 100% methylation and conversion efficiency and these samples were AS BS *BDNF* primer amplified to assess 5fC/CaC methylation at rs6265 prior pyrosequencing. As the quantification of 5fC/CaC methylation was successful using the MAB-Seq method, the CEGX array fC methodology was not pursued. This however meant that 5fC and 5CaC methylation could not be distinguished between and that only cytosine homozygotes could be investigated in analysis, as CaMAB-Seq conversion using $NaBH_4$ was not known to us at this time. Therefore, 5fC and 5CaC were grouped for comparison in the subsequent analysis, presented in chapters 4 to 6.

A)

```
Ref_Lam CCATAAGGTGGATGGCAACCCCTACAGCCATCTTCCGGATGACGTCCGGGAGACACTGCA 60
120FC_L ------------------------------NNNNNNNNNNNNNNNNGGGANANNTTGTA 29
                                                      **** *   ** *

Ref_Lam GTCCCGGATGGACGCAACCCGCCAGATGTTTGCGCAGAAGGTGTCGGCATATACCGGCCT 120
120FC_L GTTTNNNATGGANGTAATTCGTTAGATGTTTGCGTAGAAGGTGTCGGTATATATNGGTTT 89
        **     ***** * **  **  ********** ********** *****  **  *

Ref_Lam GTCCGTGCAGGTTGTGCTGGATACCGAGGCTGCAGTGTACAGCGGTCAGGAGGCCATTGA 180
120FC_L GTTCGTGTAGGTTGTGTTGGATATCGAGGTTGTAGTGTATAGCGGTTAGGAGGTTATTGA 149
        ** **** ******** ****** ***** ** ****** ****** ******  *****

Ref_Lam TGCCGGACTGGCTGATGAACTTGTTAACAGCACCGATGCGATCACCGTCATGCGTGATGC 240
120FC_L TGTCGGATTGGTTGATGAATTTGTTAATAGTATCGATGCGATTATCGTTATGCGTGATGT 209
        ** **** *** ******* ******* ** * ********* * *** **********

Ref_Lam ACTGGATGCACGTAAATCCCGTCTCTCAGGAGGGCGAATGACCAAAGAGACTCAATCAAC 300
120FC_L ATTGGATGTACGTAAATTTCGTTTTTTAGGAGGGCGAATGATTAAAGAGATTTAATTAAT 269
        * ****** ********  *** * * **************  ******* * *** **

Ref_Lam AACTGTTTCAGCCACTGCT    319
120FC_L AATT---------------    273
        ** *
```

B)

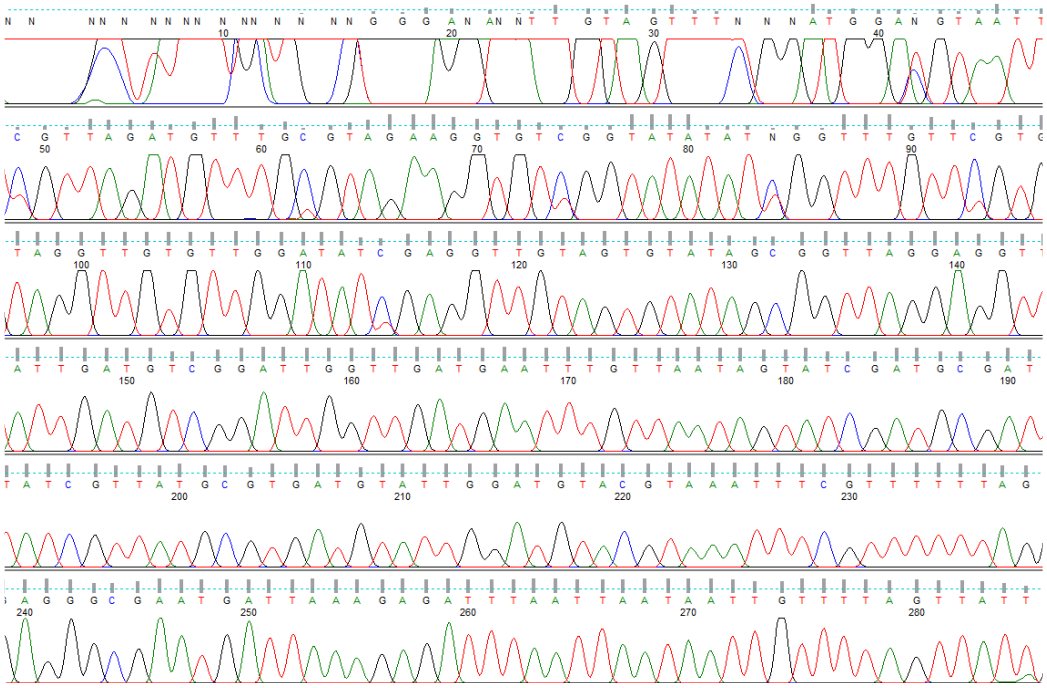


Figure 3.4 Validation of methylation assisted bisulphite sequencing using λ phage DNA. A) Aligned sequences of the reference unconverted λ amplicon (Ref_Lam) (Neri et al., 2016) and the BS converted λ amplicon DNA used as a spike in (120FC_L). The asterisks below the sequence indicate both sequences show the same base calls. CpG dyads have been underlined. B) Peak-call image for the sequencing of BS converted λ amplicon DNA used as a spike in after methylation and conversion. As with Sanger sequencing, peaks represent the call likelihood for each base. Base calls are shown above peaks (A, C, G, T or N). Phred scores are shown as grey bars above each base call.

3.5 Optimising the quantification of methylation abundance

3.5.1 Quantifying methylation through Sanger and pyrosequencing

Methylation was originally quantified through cloning of an amplicon and comparing Sanger sequencing peak height ratio. DNA extracted from blood from 3 clinical Alzheimer's samples (OPTIMA 111, 199, and 253) were bisulphite and oxibisulphite converted and amplified using AS BS *BDNF* primers. These samples were cloned and Sanger sequenced and semi quantified using the *ab1peakreporter* tool, only using base calls with a phred score of over 30. This tool is capable of estimating the level of methylation observed (C:T ratio) calibrating values by using the peak height of the codons either side of the site of interest as reference.

Levels of methylation obtained through *ab1peakreporter* tool analysis indicated substantial variations in the relative level of 5mC and 5hmC methylation, both between and within samples and tissues. Differences between the values varied by as much as 5% in bisulphite conversion, and 33% in oxibisulphite conversion, within the same sample and tissue type. Moreover, because subtraction-based mapping was required to calculate the levels of 5mC to 5hmC, the inconsistencies between runs created cumulative error. In addition to this, as methylation values were averaged to compare methylation between tissues and samples outlying values and low phred scores influenced the data.

Bisulphite converted non-clinical DNA (samples AF111 and S6) were also amplified for the AS-BDNF amplicon and cloned generating 3 clones each as a comparison to pyrosequencing data for the same samples using AS *BDNF* PM Norm/AS *BDNF* BS PM1 primer combinations. In addition, methylation data generated and reported by Schultz et al. was used as a further comparison to the Sanger sequencing peak calls (M. D. Schultz et al., 2015). These data were generated by mapping bisulphite (5mC and/or 5hmC) methylation using Sanger sequencing across the genome from snap-frozen tissues from multiple organs. To assist with the interpretation, an additional non-clinical variant site (rs779034534) which was covered by both the Sanger and pyrosequencing BS BDNF primers was also included. The external data set was

examined for both the antisense rs6265 and rs779034534 variant complementary CpGs, and alignment was performed using the hg18 genome build. A comparison of methylation levels as quantified through bisulphite sequencing for both sites is shown in Table 3.1.

5mC and 5hmC methylation values quantified by Sanger sequencing and pyrosequencing at the rs6265 variant site averaged at 77.4% and 89%, respectively and were relatively similar. In contrast, the Schultz data set reported 100% methylation at AS-rs6265. However, when comparing the values at the non-clinical AS-rs779034534 variant site, the pyrosequencing methylation value (27.5%) was closer to the Schultz methylation value (0%) than the Sanger sequencing value (96.5%). Indeed, the Sanger sequencing methylation value at the AS-rs779034534 site was higher than that at the AS-rs6265 site. These findings indicated that semi-quantification by Sanger sequencing generated a higher level of variability compared to pyrosequencing, as pyrosequencing only has a 5% margin of error generated per run (correspondence with Dr. Ellis, QIAGEN). Subsequent quantification of methylation at *BDNF* and *APOE* risk alleles in brain samples examining was performed by pyrosequencing.

3.6 Optimisation of DNMT3L risk allele primer sets and DNMT3L analysis

The *DNMT3L* rs7354779 missense variant has previously been reported to be associated with change in cognition (Flitton et al., 2019). As an extra validation of methods and a comparison to *BDNF* and *APOE* variant sites, primers targeting genomic and converted DNA capturing the *DNMT3L* rs7354779 region were designed (Table 2.7). The rs7354779 variant site is known to form a CpG when individuals possess the minor C allele.

In total 23 individuals were genotyped and compared for 5mC and 5hmC levels. All quantification was performed by Sanger sequencing peak calling. The ratios of homozygous T (T/T), heterozygous (C/T), and homozygous C (C/C) individuals were

7:13:3 respectively, or 30% T/T, 57% C/T, and 13% C/C genotypes. Semi-quantification of 5mC and 5hmC methylation is shown in Table 3.2.

Oxibisulphite sequencing was not successful for all samples. This meant we could not differentiate between 5mC and 5hmC methylation in some samples. Several of the 5mC and 5hmC percentages generated through subtraction-based mapping with oxibisulphite sequencing showed negative numbers which is an artefact of variation in the sequencing. Similarly, samples S1, S2, and OPTIMA 250 (blood) showed methylation despite being T allele homozygotes which therefore have no CpG sites at the variant site. Because of the exclusion of the minor allele cytosine it is also not possible for the variant to create a CpH site, which means no methylation should have been detected. Hence, the methylation observed at the variant site in the T allele homozygous samples indicates the margin of error for semi-quantification of methylation using sanger sequencing (minimum 16% margin of error).

Despite the limitations of the Sanger sequencing quantification method, 5mC and 5hmC methylation was observed within the C allele homozygous individual (AF111) at rs7354779, and is above the 16% margin of error generated in Sanger sequencing (in bold in Table 3.2). This is the first report that this allele is methylated and provides an alternative explanation involving transcription regulation rather than protein amino acid changes, of how the allele may contribute to changes in cognition. While the sample number was too low to perform any statistical analysis, both genotypes appear to show varying levels of 5mC/5hmC methylation.

Variant rs#	Schultz Meth count	Schultz Read depth	% Meth	Observed Sanger-seq	Observed PyroSeq
AS rs6265	8	8	100	89.0%	77.4%
AS rs779034534	0	9	0	96.5%	27.5%

Table 3.1 Comparison of quantification of 5mC/5hmC methylation through Sanger sequencing (observed Sanger-seq), pyrosequencing (observed pyro-seq) and data reported by Schultz et. al. (% meth) at antisense *BDNF* rs6265 and rs779034534 variant sites. Read depth and reported methylation calls indicate the Sanger sequencing calls reported in the Schultz data set.

Sample ID	Genotype	BS methylation (5mC/5hmC) (%)	5mC (%)	5hmC (%)
S01	T/T	10.74	15.25	-5.19
S02	T/T	15.74	2.77	12.97
BR250	T/T	11.65	-	-
AF103	C/T	28.07	-	-
AF105	C/T	72.3	-	-
QQ88	C/T	2.17	-	-
S08	C/T	12.38	-	-
S09	C/T	15.2	16.77	-1.56
TT296	C/T	5.07	22	-16.93
AF111	C/C	30.31	9.55	**20.76**
S11	C/C	12.09	23.2	-11.11
TT297	C/C	1.1	-	-

Table 3.2 Bisulphite Sanger sequencing and semi-quantification of methylation at the *DNMT3L* rs7354779 variant using the ab1peakreporter tool. Sample genotypes at *DNMT3L* rs7354779 variant site, the level of 5mC/5hmC methylation observed through bisulphite sequencing (BS methylation (5mC/5hmC) (%)) and the ratio of 5mC to 5hmC methylation calculated by performing subtraction-based mapping using oxibisulphite sequencing (5mC, 5hmC (%), respectively).

3.7 Discussion

This chapter has described measures taken in order to successfully optimise the quantification of 5mC, 5hmC and 5fC/5caC methylation in DNA isolated from blood and brain tissue. Many techniques needed to be optimised to ensure methylation data were reliable and accurate. To this end, where many processes were unable to provide results of sufficient quality or repeatably quantify methylation they were replaced with new methods.

Extraction of genomic materials

Extraction from snap frozen tissue presented its own set of issues and challenges for optimisation. These initial low yields attained might be due to the sequestering of genomic material in vesicles or protein structures. The adjusted protocol using the phenol-chloroform addition was effective in improving yields and may have acted to separate lipids from water soluble cell contents like DNA (Fan & Gulley, 2001). This is especially important as the lipid content found within human brain is high (forming 36% in grey matter to 81% in Myelin tissues) (O'Brien & Sampson, 1965).

3.7.1 Bisulphite conversion of DNA to quantify methylation

Bisulphite conversion methods to assess methylation in DNA were a large part of this thesis and required much time to optimise. Bisulphite conversion can shear DNA, lead to strand degradation, and is especially damaging to pyridamine bases (Ruga & Jia, 2019; Yi, Long, Cheng, & Huang, 2017). Because of the nature of bisulphite conversion, steps could not be altered extensively as DNA must be denatured, bisulphite treated, and desulphonated to analyse methylation. Nevertheless, conversion using the CEGX TrueMethyl kit showed a higher consistency in the quality and quantity of converted DNA yielded than the Millipore CpGenome Turbo bisulphite modification kit. Moreover a greater number of improved kits have been released that show numerous benefits dependent on the tissue and inputs being used (Leontiou et al., 2015).

MAB conversion enabled comparison of 5fC/CaC, but uses bisulphite conversion after treatment with *M.SssI*. Therefore, any issues in bisulphite conversion or methylation led to lower MAB converted DNA yields. Some of the inconsistency attributed to the Millipore kit may have been due to the need to weigh hygroscopic materials and adjust pH of solutions. These additional steps, led to more opportunities for error and increased amount of time required for conversion. The CEGX kit avoided this issue as the reagents had already been weighed and compartmentalised meaning only a few solutions needed to be prepared prior to conversion. Moreover, the CEGX method allowed for easy oxibisulphite conversion and came with its own controls to ensure high efficiency conversion of DNA.

5fC methylation was initially targeted using the CEGX TrueMethyl 5fC Conversion Array. However, as there were many issues within the optimisation due to the sensitivity of the reagent, low yields of converted DNA of poor quality were generated. Given more time and resources this method could be further investigated and optimised in order to specifically quantify the 5fC methylation levels. The methylation Assisted Bisulphite (MAB) conversion (Neri et al., 2016) which uses a *M.SssI* methylase and SAM, was successfully optimised. A drawback from using this method was that it is time consuming, could only be validated by sequencing at the end of bisulphite conversion and was unable to differentiate between 5fC and 5CaC methylation types. Because the analysis pipeline depends on methylation, bisulphite conversion, amplification, and sequencing before success and efficiency could be determined, there were many processes that could fail in conversion without being detected until much later.

In addition to the Millipore and CEGX conversion kits, newer more advance kits have been released for the quantification of methylation from samples. For instance the Epigentek Methylamp Whole Cell Bisulphite Modification Kit directly converts gDNA from whole cells or tissues, eliminating the need for DNA extraction ("Methylamp Whole Cell Bisulfite Modification Kit | EpiGentek," 2019). Other companies have released commercial kits to quantify 5fC methylation (Abcam 5-formylcytosine DNA quantification Kit) which are optimised and may be used in parallel with bisulphite

conversion("5-formylcytosine DNA Quantification Kit (ab156905) | Abcam," 2019). High throughput methods for the quantitative assessment of multiple methylation sites is also commercially available ("Infinium MethylationEPIC Kit | Methylation profiling array for EWAS," 2019). In the future, use of these conversion kits and quantification techniques will make the processing of many more samples easier.

3.7.2 Sequencing methods

Sanger sequencing was initially used to genotype and quantify methylation within converted samples but peak calling proved variable and with a high margin of error. Furthermore, throughout the study (discussed further in chapters 4 to 6) it was observed that the most accurate technique to quantify methylation levels is pyrosequencing. While pyrosequencing relies on the same method of quantification as Sanger sequencing, i.e. signal intensities generated by the incorporation of dNTP, it is far more sensitive (5% level of error in resolution), consistent in findings, and can consume as little as 5ng input transcript (Kurdyukov & Bullock, 2016). This is important as methylation levels are variable and quantification of methylation type and level may be essential for future diagnostics. Moreover, the cost and speed of pyrosequencing will very likely decrease in the future.

Furthermore, when combined with Next Generation Sequencing methods e.g. Roche 454, Ion Torrent sequencing, quantification of multiple methylation types across large sections of the genome over multiple samples will be possible. New techniques are also still being developed such as chromatography assisted mass spectrometry (LC-MS/MS) that will allow for the quantification of all methylation types at base resolution with no conversion (Chowdhury, Cho, & Irudayaraj, 2017). This would not only allow for the determination of the causative effects of methylation in phenotype, but also may allow us to better identify key biomarkers for multiple diseases, and potentially medications to counter the effects of aberrant methylation.

Controls to assess DNA conversion

In addition to the digestion control provided by the CEGX kit, a λ phage spike in control was designed and used to validate all conversions. Use of λ phage spike-in is cheap to analyse and has previously been used in methylation analysis with success (Leontiou et al., 2015; Shirane et al., 2013).The λ phage spike-in control provided a substantial advantage to the other controls provided as it could validate all forms of conversion (bisulphite, oxibisulphite and MAB conversion). This control procedure was routinely performed using Sanger sequencing to save on costs and time. However, pyrosequencing could be performed providing a more accurate measure of methylase activity quantification. This would show the level of 5mC methylation at each CpG site which would increase the quantification of the methylase action by 20 times.

During the MAB conversion process, many sequences showed a lack of methylase activity at the beginning of the sequenced λ phage control DNA spike in sequences. For the purposes of this study these start sequences below 30 phred score were ignored. However, this may represent a bias in the methylase activity or be an artefact of sequencing. It would be interesting to investigate in future studies whether this is due to poor sequencing or failure of the methylase at the start of the sequences.

3.7.3 Quantification of methylation at DNMT3L rs7354779

This study has observed 5mC and 5hmC methylation at the *DNMT3L* rs7354779 risk epiallele within blood DNA of healthy phenotype individuals. Previous studies have indicated that the rs7354779 C allele is associated with poor childhood intelligence, and also disrupts the protein interaction between *DNMT3L* and *DNMT3A*, which is required to form complexes for 5mC methylation (Flitton et al., 2019; Haggarty et al., 2010). These models have focused on the effect of the variant on protein changes, however 5mC and 5hmC methylation are known to modulate transcriptional expression, either by repressing or enhancing transcriptional expression, respectively (Hahn et al., 2013; S.-G. Jin et al., 2011; D. Q. Shi et al., 2017). This novel finding of

5mC and 5hmC methylation at the *DNMT3L* rs7354779 variant should be studied further by quantifying methylation in diseased and healthy controls. However unfortunately, due to monetary and time constraints of the project, development of pyrosequencing primers for this variant site were not possible, and additional investigation into the levels of methylation within disease, tissues and genotypes of the *DNMT3L* variant were not possible. Further research to quantify 5mC, 5hmC, 5fC, and 5CaC methylation at the variant through pyrosequencing in clinical and healthy samples in the future may elucidate the association between methylation and cognition.

3.7.4 Future developments

In the future, MAB conversion and sequencing using $NaBH_4$ could be used to further reduce methylated cytosines in order to differentiate 5fC and 5CaC methylated cytosines in quantification. (Booth, Marsico, Bachman, Beraldi, & Balasubramanian, 2014; H. Wu et al., 2016). Using subtraction-based mapping it would be possible to quantify all four methylation type levels at any CpG site. Although this technology has been optimised by different manufacturing companies, the sensitivity of the reductant itself may pose issues. Moreover, methylation by methylases could be optimised further top try and achieve close to 100% methylation across phage DNA. This may be pursued in future work further investigating the levels of methylation at these epialleles using much larger clinical sample numbers.

3.7.5 Conclusions

In this chapter, the optimisation of methylation methods in order to reliably and accurately quantify 5mC, 5hmC and 5fC/5caC methylation at target risk epialleles was presented. The results demonstrated successful conversion to assess 5mC, 5hmC and 5CaC or 5fC states enabling *BDNF* rs6265 and *APOE* rs7412 & rs429358 epialleles to examined as discussed in chapters 4-6. Furthermore, the *DNMT3L* rs7354779 risk allele was shown to be 5mC and 5hmC methylated (Table 3.2) providing further evidence that genetic risk for disease could be in part a DNA methylation processes regulating transcription of disease associated proteins.

Chapter 4: DNA methylation analysis of the antisense *BDNF* rs6265 epiallele

Acknowledgements

The analysis of methylation at the antisense *BDNF* rs6265 variant site would not have been possible without the assistance and advice from a number of people. Dr Miles Flitton, University of Nottingham, helped write the code which enabled the extraction of data analysed in the BDNF association study. Maria Isabelle Haig, the University of Nottingham, gave advice on the optimisation of the Sanger sequencing primers used to target *BDNF*. Finally, Dominika Valent Raffajová, University of Nottingham, assisted in the preparation to genotype samples.

4.1 Workflow pipeline for the analysis of BDNF

Studies have indicated that *BDNF* and its associated proteins play a key role in the molecular processes underlying cognition and that impaired BDNF function putatively contributes to cognitive and neurological diseases (Day & Sweatt, 2011b; Petzold et al., 2015). The coding variant and epiallele rs6265 results in a valine to methionine (Val66Met) coding substitution and has commonly been associated with a number of brain conditions (Binder & Scharfman, 2004; Michael F. Egan et al., 2003; L. Liu et al., 2008). This chapter will first discuss the results obtained in an association study of the *BDNF* rs6265 variant genotypes with neuropsychiatric phenotypes, followed by genotype and methylation analysis of the *BDNF* rs6265 antisense complementary CpG. A diagram showing the *BDNF* research pipeline is presented in Figure 4.1

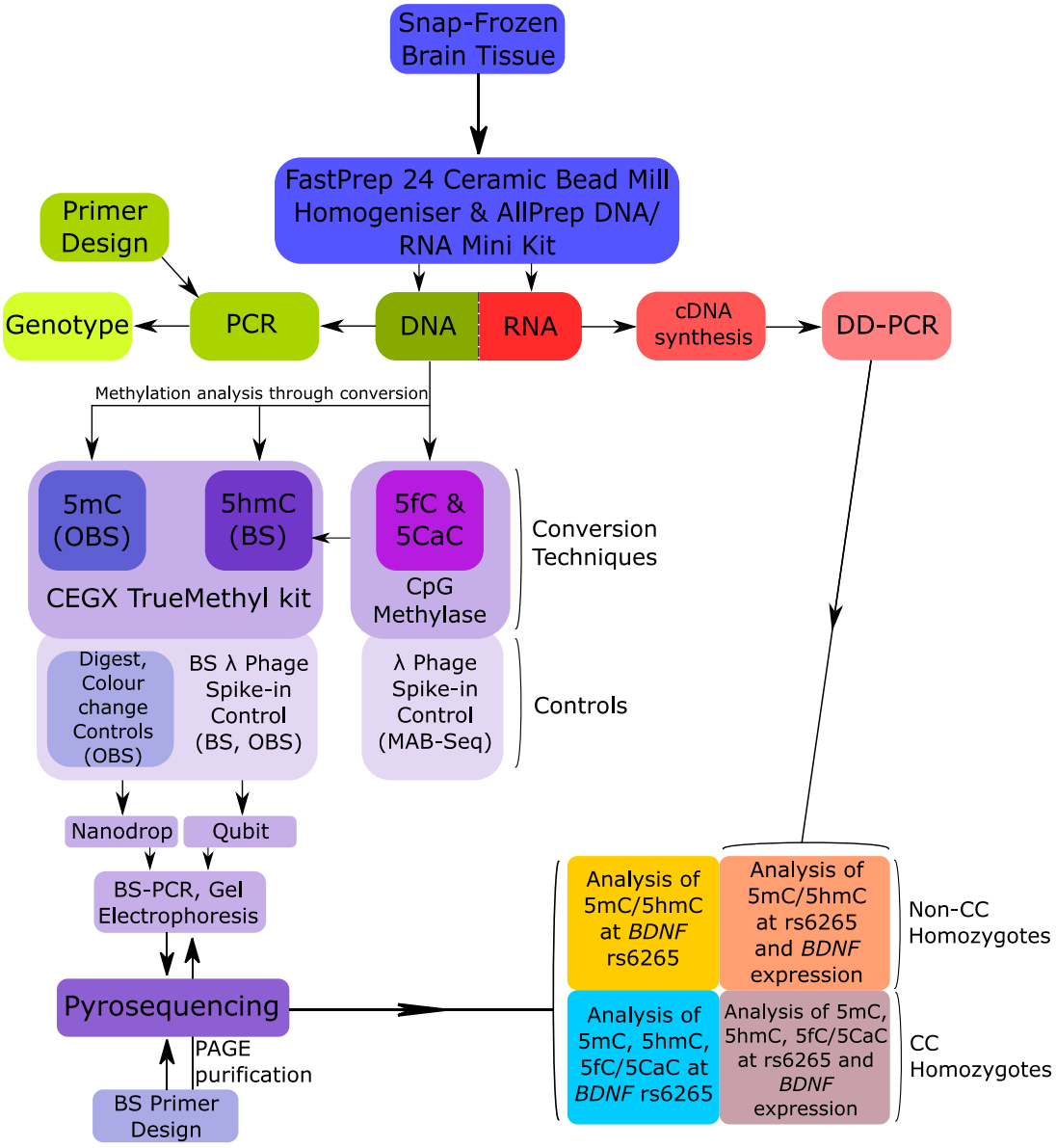


Figure 4.1 Workflow for the analysis of *BDNF* rs6265 antisense complementary CpG methylation (chapter 4) and *BDNF* transcript abundance (chapter 5). The DNA/RNA pipelines are shown in red and green respectively. Methods for the analysis of methylation are detailed in purple. The end goals were the analysis of methylation (either 5mC and 5hmC or 5mC, 5hmC, and 5fC/CaC) at rs6265 antisense CpG in minor allele genotypes (i.e. heterozygous and T homozygous individuals, shown in yellow) or major allele C homozygous individuals (shown in light blue), respectively. Digital Droplet PCR expression data (DD-PCR, shown in pink) for *BDNF* were combined with methylation data derived from pyrosequencing (purple) and this will be discussed in chapter 5. UEB refers to universal extraction buffer, FFPE refers to formalin fixed, paraffin embedded tissue, BS refers to bisulphite, OBS refers to oxibisulphite, MAB-Seq refers to methylation assisted bisulphite sequencing, PAGE refers to polyacrylamide gel electrophoresis.

4.2 Association study of the BDNF rs6265 variant in individuals with psychiatric disease.

Data from the UK10K next generation sequencing (NGS) cohort data were used to perform an association study of *BDNF* sense strand rs6265 variant. Individuals diagnosed with Autism Spectrum Disorder (ASD) comprised of autism and Asperger's diagnoses and unaffected control individuals were examined for an association with the rs6265 'A' minor allele (66Met). The total sample number analysed was 2512 individuals. Exome sequencing revealed 723 ASD (29% of samples) (observed minor allele frequency from UK10K data (OMAF) = 0.19) diagnosed individuals and 1789 control (71% of samples) (OMAF = 0.14) individuals carried the rs6265 minor allele. Chi-squared analysis showed that there was a significant association between genotype at the *BDNF* rs6265 variant site and ASD diagnosis, $X^2(1) = 23.09$, $p < 1\times10^{-3}$. The odds ratio (OR) of 1.49 and a confidence interval of 1.26-1.75 indicated that carriers of the minor A allele (66Met) are more likely to have an ASD diagnosis compared to controls. This is consistent with previous findings (Yoo et al., 2014).

Individuals diagnosed with schizophrenia and schizoaffective disorder (SCZ) were also investigated. The total sample number for analysis was 3671 individuals. Exome sequencing revealed 1882 SCZ (OMAF = 0.16, 51% of samples) individuals and 1789 controls (OMAF = 0.14, 49% of samples) carried the rs6265 minor A allele. Analysis of SCZ revealed a significant association between rs6265 minor allele and SCZ, $X^2(1) = 10.72$, $p = 1\times10^{-3}$, which indicated significant differences between observed and expected frequencies. Moreover, similar to the findings for ASD, the OR of 1.24 (1.09-1.42) indicated that carriers of the A allele are more likely to have a SCZ diagnosis compared to controls. Both the chi-squared statistics for SCZ and ASD diagnosis show that the rs6265 minor allele is associated with increased risk for psychiatric disease. Unfortunately, owing to a limited number of samples and access to genotype data, further investigation of the rs6265 risk allele association in cohorts with neurodegenerative phenotypes such as MCI, AD and PD, was not possible. The MAF, Chi-square, and p values of the rs6265 genotype and ASD/SCZ diagnosis analysis are shown in Table 4.1.

Diagnosis	Number of Cases (MAF)	Number of Controls (MAF)	X^2	P	OR (CI)
ASD	723 (0.19)	1789 (0.14)	23.09	$<1\times10^{-3}$	1.49 (1.26-1.75)
SCZ	1882 (0.16)	1789 (0.14)	10.72	1×10^{-3}	1.24 (1.09-1.42)

Table 4.1 Association of rs6265 minor allele with neuropsychiatric phenotypes (ASD and SCZ). Counts per cohort, MAF (shown in brackets below counts), Pearson's X^2 value, P-value, Odds ratio (OR) and 95% confidence interval (CI) are shown. The lower threshold for significance is $<1\times10^{-3}$.

Sample ID	Age	Gender	Diagnosis	Antisense *BDNF* rs6265 Genotype
OPTIMA 111	84	Male	Alzheimer's Disease	C/C
OPTIMA 108	-	Female	Alzheimer's Disease	C/C
OPTIMA 199	70	Female	Alzheimer's Disease	C/C
OPTIMA 253	79	Female	Alzheimer's Disease	C/C
OPTIMA 195	74	Male	Alzheimer's Disease	C/C
OPTIMA 215	77	Female	Alzheimer's Disease	C/C
OPTIMA 250	69	Female	Alzheimer's Disease	C/C
OPTIMA 297	65	Female	Alzheimer's Disease	C/T
OPTIMA 47	82	Female	Alzheimer's Disease	C/C
OPTIMA 58	88	Male	Alzheimer's Disease	C/C
OPTIMA 179	70	Female	Alzheimer's Disease	C/C
PD687	86	Male	Parkinson's Disease	C/C
PD793	75	Female	Parkinson's Disease	C/C
PD898	84	Male	Parkinson's Disease	C/C
PD115	63	Male	Dementia with Lewy bodies	C/T
PD120	83	Female	Dementia with Lewy bodies	C/T
PD332	77	Male	Dementia with Lewy bodies	C/T
PDC007	104	Female	MCI	C/C
PDC111	88	Female	MCI	C/T
PD715	86	Male	MCI	C/C
C045	77	Male	MCI	C/C
PD065	84	Female	MCI	C/T
PD092	79	Male	Control	C/T
PDC110	94	Female	Control	C/T
PDC112	87	Male	Control	T/T

Table 4.2 *BDNF* rs6265 antisense complementary genotypes in disease and control samples. Individual age at sampling, gender, and diagnosis are given in addition to genotype status, where C is the major allele, and T is the minor allele. MCI denotes mild cognitive impairment.

4.3 BDNF rs6265 complementary antisense CpG allele in individuals with neurodegenerative diseases

To perform analysis of DNA methylation levels at the *BDNF* rs6265 antisense complementary CpG variant site, genotypes were first determined. Eleven samples with an Alzheimer's disease diagnosis (the OPTIMA cohort samples) had DNA which was isolated from homogenised whole brain and blood samples. DNA from mild cognitive impairment, Alzheimer's disease, Parkinson's disease and dementia with Lewy bodies samples, were extracted from snap frozen cerebellum, hippocampus and frontal cortex brain regions. Genotypes of samples were organised by carrier status of the T minor allele at antisense *BDNF* rs6265 (Table 4.2). This allele is the complementary base of the rs6265 sense 'A' minor allele which showed association with neuropsychiatric diseases. Because of the nature of CpG methylation, possession of the minor T allele at antisense rs6265 created thymine guanine dyad (TpG) sites which cannot be methylated.

Of the 25 individuals genotyped, 16 individuals (64% of samples) were homozygous for the major C allele, 8 individuals (32% of samples) were heterozygous individuals, and one individual (4%) was a minor T allele homozygote. The rs6265 antisense MAF for these samples was 0.2. While this MAF was not the same as the MAF observed from the UK10K samples above (MAF = 0.14), this is the same expected genotype frequencies for antisense rs6265 genotype in individuals of European ancestry with homozygous C allele individuals composing 64%, heterozygotes composing 32%, and homozygous T allele individuals composing 4% of the population, with the T allele MAF at 0.2 (R. A. Gibbs et al., 2015). It should however be noted that the MAF expected for the UK population is 0.16 (R. A. Gibbs et al., 2015). In this way, while the clinical samples are not a true reflection of MAF expected in a normal population in the UK (MAF = 0.16) or the observed MAF in the UK10K cohort (MAF = 0.14), the samples do reflect the expected EU MAF (MAF = 0.2). As the genetic ancestry of the samples is unknown, it is possible that the observed MAF for these samples could in part be due to ancestry originating from the EU.

The frequency of genotypes by disease status was unequal. Of the 11 Alzheimer's disease samples, all but one sample were C homozygotes and the exception was heterozygous (Table 4.2). All Parkinson's disease samples were C homozygotes and all dementia with Lewy bodies samples were heterozygotes. The mild cognitive impairment (MCI) samples had a roughly equal split where 3 samples were C homozygotes and 2 samples were heterozygotes. Two control samples were heterozygotes and one was T homozygous, which is the only T homozygote present.

Most clinical groups presented mixed ratios of gender and age. However, the majority of groups showed a female bias. Indeed 14 of the 25 samples were women. While the other clinical groups have roughly equal gender ratios, Alzheimer's disease showed a large inequality in the ratio of women to men (8:3, respectively). Because of the small numbers available in this study, neither age nor gender could be accounted for in analysis.

4.4 Pyrosequencing methylation analysis

Methylation was quantified through pyrosequencing bisulphite, oxibisulphite and methylation assisted bisulphite (MAB) converted samples. Because 5fC/CaC methylation determined through MAB conversion can only be determined at C homozygous genotypes, analyses of 5mC/5hmC and 5fC/CaC methylation were separated into individual pipelines. Linear mixed effects models were used to analyse how disease, tissue, genotype and their interactions may be contributing to the levels of 5mC, 5hmC, or 5fC/CaC methylation observed.

It has been suggested that genotype, disease, and tissue may all be contributing to the observed levels of methylation. For instance the rs6265 minor allele is known to be associated with abnormal brain architecture and neurotransmission, which has also been associated with disease and changing methylation levels (Chao et al., 2008). 5mC and 5hmC methylation have been reported to vary in levels dependant on tissue and genes (Globisch et al., 2010; C. X. Song et al., 2011). Similarly, disease states are known to contribute to the modulation of methylation, such as global

hypomethylation events in the brain of dementia with Lewy bodies patients (Desplats et al., 2011). The mechanisms by which these variables may be contributing to methylation levels, and whether these variables are a result of or play a causative role in modulating methylation levels is currently disputed, and will be discussed at length later in this chapter.

4.4.1 5mC and 5hmC methylation at AS-rs6265 by genotype

Of the 62 data points sequenced from all available samples, disease states and tissues which passed pyrosequencing, data for 51 methylation data points were available for 5mC and 5hmC analysis which showed no erroneous results in sequencing (below quantification threshold of 5%) or subtraction-based mapping (reporting results below 0% or above 100%) (Data points successfully used are presented in Table 4.3). This led to unmatched group sizes and non-factorial design for analyses.

5mC and 5hmC methylation at the antisense *BDNF* rs6265 complementary CpG epiallele (hereafter referred to as AS-rs6265) was first analysed by regarding heterozygous (C/T) individuals as carriers of the Met66 allele due to base complementarity (Watson & Crick, 1953), and homozygous C (C/C) genotypes were regarded as non-carriers or Val66 alleles. Due to the nature of methylation and bisulphite sequencing, T homozygotes were excluded from analysis. Of the 37 non-carrier methylation data points, 7 samples gave erroneous results and were removed for analysis leaving 30 data points. Of the 25 carrier data points, four proved erroneous, or had a T/T genotype and were removed leaving 21 for analysis. The data points used in 5(h)mC analysis are shown in Table 4.3. No successful pyrosequencing data were obtained for Whole Brain DNA from Carriers. Some disease sub-groups had the same genotypes across all samples resulting in a lack of non-carrier control, non-carrier DLB, and carrier PD samples. Similarly, because the carrier AD group only had one sample it could not be included in analyses.

Because analysis of methylation only with genotype resulted in a factorial and non-collinear design, ANOVAs were performed on the normally distributed and

homoscedastic data, however LMEM analysis was performed when including other variables such as disease and tissue types, to determine the effect of each variable. Two-way (paired) ANOVAs indicated a significant difference in methylation levels between genotype groups and across methylation types (5mC and 5hmC) ($F(3,104) = 4.53$, $p = 5\times10^{-3}$), and Welch's ANOVA ($F(3, 54.33) = 5.21$, $p = 3\times10^{-3}$). This indicated that significant difference in methylation levels were not due to heterogeneity in variance.

To determine which groups showed significant differences in methylation levels, Games-Howell post-hoc analysis was performed. Only one significant interaction between non-carrier 5mC and 5hmC methylation groups was observed (Figure 4.2). It showed that 5mC levels observed at C homozygote genotypes was significantly greater (mean ± SE; 41.36 ± 3.45, $p = 1\times10^{-3}$) than C homozygote 5hmC levels (24.70 ± 2.43). This supports literature that suggests that 5mC is more abundant than 5hmC methylation in the exons of embryonic stem cells (M. Yu et al., 2012). Moreover, heterozygous carriers had more similar levels of 5mC and 5hmC methylation, compared to C homozygotes. This may suggest that 5mC and 5hmC methylation is not less in individuals with only one potential CpG site available for methylation (heterozygotes). As 5hmC is also known to be enriched in brain tissues (Hahn et al., 2014), it is possible that the high levels of 5hmC observed in heterozygotes are due to 20 of the 21 available data points originating from brain tissues (Table 4.3).

Sample	Brain Bank	Tissue	Diagnosis	rs6265 Genotype
OPT 111	OPTIMA	Whole Brain	Alzheimer's Disease	C/C
OPT 199	OPTIMA	Blood	Alzheimer's Disease	C/C
OPT 199	OPTIMA	Whole Brain	Alzheimer's Disease	C/C
OPT 253	OPTIMA	Blood	Alzheimer's Disease	C/C
OPT 253	OPTIMA	Whole Brain	Alzheimer's Disease	C/C
OPT 195	OPTIMA	Whole Brain	Alzheimer's Disease	C/C
OPT 215	OPTIMA	Blood	Alzheimer's Disease	C/C
OPT 215	OPTIMA	Whole Brain	Alzheimer's Disease	C/C
OPT 250	OPTIMA	Whole Brain	Alzheimer's Disease	C/C
OPT 297	OPTIMA	Blood	Alzheimer's Disease	C/T
OPT 47	OPTIMA	Blood	Alzheimer's Disease	C/C
OPT 47	OPTIMA	Whole Brain	Alzheimer's Disease	C/C
OPT 58	OPTIMA	Blood	Alzheimer's Disease	C/C
OPT 58	OPTIMA	Whole Brain	Alzheimer's Disease	C/C
OPT 179	OPTIMA	Whole Brain	Alzheimer's Disease	C/C
PD687	Imperial	Frontal Cortex	Parkinson's Disease	C/C
PD687	Imperial	Hippocampus	Parkinson's Disease	C/C
PD687	Imperial	Cerebellum	Parkinson's Disease	C/C
PD793	Imperial	Frontal Cortex	Parkinson's Disease	C/C
PD793	Imperial	Hippocampus	Parkinson's Disease	C/C
PD793	Imperial	Cerebellum	Parkinson's Disease	C/C
PD898	Imperial	Frontal Cortex	Parkinson's Disease	C/C
PD898	Imperial	Hippocampus	Parkinson's Disease	C/C
PD898	Imperial	Cerebellum	Parkinson's Disease	C/C
PD115	Imperial	Frontal Cortex	Dementia with Lewy bodies	C/T
PD115	Imperial	Hippocampus	Dementia with Lewy bodies	C/T
PD115	Imperial	Cerebellum	Dementia with Lewy bodies	C/T
PD120	Imperial	Frontal Cortex	Dementia with Lewy bodies	C/T
PD120	Imperial	Hippocampus	Dementia with Lewy bodies	C/T
PD120	Imperial	Cerebellum	Dementia with Lewy bodies	C/T
PD332	Imperial	Frontal Cortex	Dementia with Lewy bodies	C/T
PD332	Imperial	Hippocampus	Dementia with Lewy bodies	C/T
PD332	Imperial	Cerebellum	Dementia with Lewy bodies	C/T
PDC007	Imperial	Frontal Cortex	Mild Cognitive Impairment	C/C
PDC007	Imperial	Hippocampus	Mild Cognitive Impairment	C/C
PDC007	Imperial	Cerebellum	Mild Cognitive Impairment	C/C
PDC111	Imperial	Frontal Cortex	Mild Cognitive Impairment	C/T
PDC111	Imperial	Hippocampus	Mild Cognitive Impairment	C/T
PDC111	Imperial	Cerebellum	Mild Cognitive Impairment	C/T
PD715	Imperial	Frontal Cortex	Mild Cognitive Impairment	C/C
PD715	Imperial	Cerebellum	Mild Cognitive Impairment	C/C
C045	Imperial	Hippocampus	Mild Cognitive Impairment	C/C
C045	Imperial	Cerebellum	Mild Cognitive Impairment	C/C
PD065	Imperial	Frontal Cortex	Mild Cognitive Impairment	C/T

PD065	Imperial	Cerebellum	Mild Cognitive Impairment	C/T
PD092	Imperial	Frontal Cortex	Control	C/T
PD092	Imperial	Hippocampus	Control	C/T
PD092	Imperial	Cerebellum	Control	C/T
PDC110	Imperial	Frontal Cortex	Control	C/T
PDC110	Imperial	Hippocampus	Control	C/T
PDC110	Imperial	Cerebellum	Control	C/T

Table 4.3 Successful data points used in analysis of *BDNF* rs6265 antisense complementary genotypes, with 5(h)mC methylation, diagnoses, and tissues. Individual diagnoses are given in addition to genotype status, where C is the major allele, and T is the minor allele. Brain bank refers to whether the sample was derived from tissue from the OPTIMA cohort or Imperial brain bank.

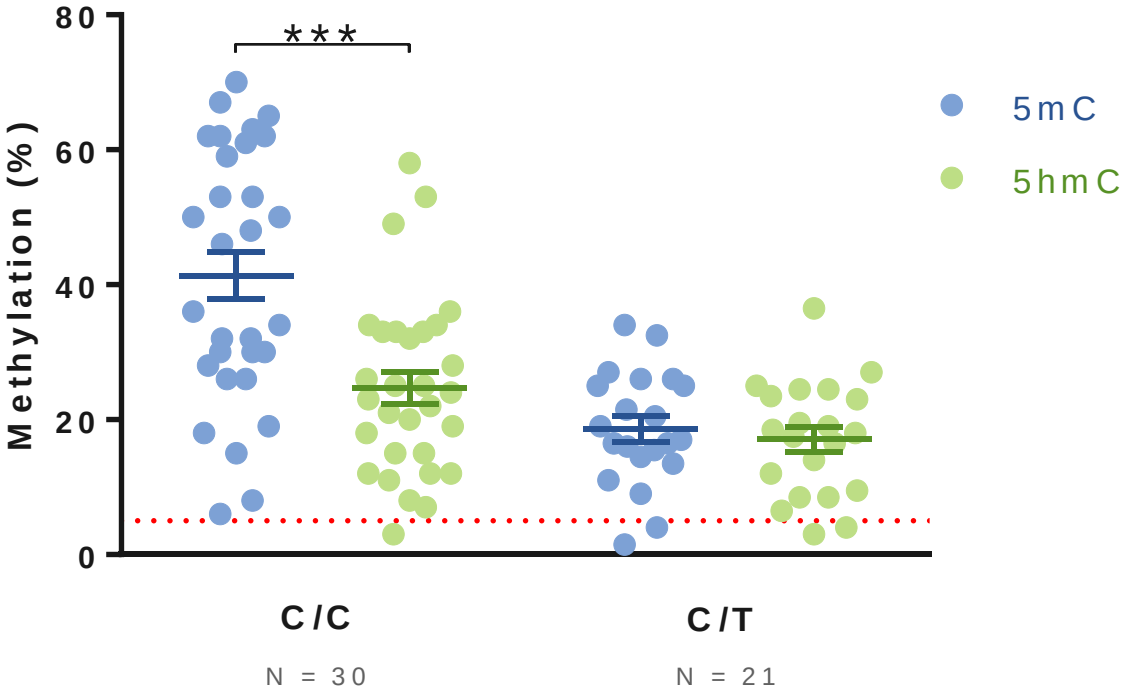


Figure 4.2 5mC or 5hmC methylation at AS-rs6265 in carrier and non-carrier groups. 5mC is shown in blue, 5hmC is shown in green. Long bars represent means, and short bars represent standard error around the mean. C homozygous genotype (N = 30), while C/T heterozygotes (N = 21). Red-line indicates 5% margin of error generated in pyrosequencing. Significance for Games-Howell post hoc is shown where (*** $p \leq 1 \times 10^{-3}$).

4.4.2 Tissue, Disease, genotype status and AS-rs6265 5mC/5hmC methylation

As stated in the methodology and above, because many data points came from the same individual, which often shared disease states and had different tissue types, the effect of each variable was determined using LMEM analysis. Analysis of 5mC or 5hmC methylation was performed by creating 3 different linear mixed effects models, to investigate effect of disease and genotype variables across all samples; effect of disease, genotype, and tissue variables across the Imperial brain bank samples; and finally, to investigate the effect of tissue and genotype in OPTIMA samples. A summary of the effect of each variable when creating the minimum adequate model for both 5mC and 5hmC methylation may be seen in Table 4.4. 5mC and 5hmC methylation levels across genotypes, tissues, and disease states may be seen in Figure 4.3. The 5mC, 5hmC and 5fC/CaC methylation data are presented in Figures 4.4 and 4.5 later in the chapter.

LMEM analysis of 5mC methylation

As Table 4.4 indicates, LMEM analysis of disease and genotype with 5mC methylation across all samples did not report significant results. From this, we can determine that when using both Imperial and OPTIMA samples, neither disease ($X^2(4) = 1.75$, $p = 0.78$) nor genotype ($X^2(1) = 0.11$, $p = 0.74$) significantly influenced the levels of 5mC methylation observed.

LMEM analysis of 5mC within the Imperial samples indicated that tissue ($X^2(2) = 8.33$, $p = 0.02$) was the main and only significant variable affecting the levels of 5mC methylation observed. When creating the minimum adequate model, the disease ($X^2(3) = 1.43$, $p = 0.70$) and tissue:disease ($X^2(6) = 3.28$, $p = 0.77$) interaction also indicated contribution to the levels of 5mC methylation observed, however these were not significant.

The analysis of OPTIMA samples did not report significant results. LMEM analysis indicated that neither tissue ($X^2(1) = 0.65$, $p = 0.42$) nor genotype ($X^2(1) = 0.06$, $p = 0.80$) significantly contributed to the levels of 5mC methylation. Genotype had far

less of an effect and was less significant compared to tissue, but this may be a reflection of the lack of C/T heterozygotes observed in the Alzheimer's population.

LMEM analysis of 5hmC methylation

Results for the LMEM analysis of 5hmC were similar to that of 5mC (Table 4.4). Analysis of 5hmC in Imperial and OPTIMA samples once again reported that neither disease ($X^2(4) = 1.77$, $p = 0.78$) or genotype ($X^2(1) = 2.40$, $p = 0.12$) contributed significantly to the levels of 5hmC observed across samples (Table 4.4). However, unlike analysis of 5mC, the effect of genotype in this model trended to significance and reported a larger X^2 statistic than that of disease.

LMEM analysis of 5hmC across the Imperial samples indicated that both genotype ($X^2(1) = 3.94$, $p = 0.05$) and tissue ($X^2(2) = 6.19$, $p = 0.05$) contributed to the level of 5hmC methylation observed, with the tissue variable reporting a higher X^2 statistic. Analysis of the OPTIMA samples once again indicated that 5hmC levels are not influenced by tissue ($X^2(1) = 2.22$, $p = 0.14$) or genotype ($X^2(1) = 1.39$, $p = 0.24$) variables (Table 4.4).

Methylation being analysed	Data set	Term	X^2	DF	P value
5mC	Imperial and OPTIMA	Disease	1.75	4	0.78
5mC	Imperial and OPTIMA	Genotype	0.11	1	0.74
5mC	Imperial	Tissue:rs6265	2.00	2	0.37
5mC	Imperial	Genotype	0.50	1	0.48
5mC	Imperial	Tissue:Disease	3.28	6	0.77
5mC	Imperial	Disease	1.43	3	0.70
5mC	Imperial	Tissue	8.33	2	0.02*
5mC	OPTIMA	Tissue	0.65	1	0.42
5mC	OPTIMA	Genotype	0.06	1	0.80
5hmC	Imperial and OPTIMA	Disease	1.77	4	0.78
5hmC	Imperial and OPTIMA	Genotype	2.40	1	0.12
5hmC	Imperial	Tissue:Disease	5.89	6	0.44
5hmC	Imperial	Disease	2.91	3	0.40
5hmC	Imperial	Tissue:rs6265	0.08	2	0.96
5hmC	Imperial	Genotype	3.94	1	0.05*
5hmC	Imperial	Tissue	6.19	2	0.05*
5hmC	OPTIMA	Tissue	2.22	1	0.14
5hmC	OPTIMA	Genotype	1.39	1	0.24

Table 4.4 Linear mixed effects model analysis for disease, tissue, and genotype variables across different samples and 5mC or 5hmC methylation types. The imperial data set contained frontal cortex, hippocampus, and cerebellum brain tissues, as well as dementia with Lewy bodies, Parkinson's disease, mild cognitive impairment, and control diagnoses. The OPTIMA data set contained Alzheimer's diagnoses across blood and whole brain tissues. Term refers to the variable being investigated, while terms separated by a colon (e.g. Tissue:Disease) indicate interaction terms between the variables. X^2 refers to the ANOVA derived Chi^2 statistic, DF refers to degrees of freedom, and P value indicates significance. (* $p \leq 0.05$).

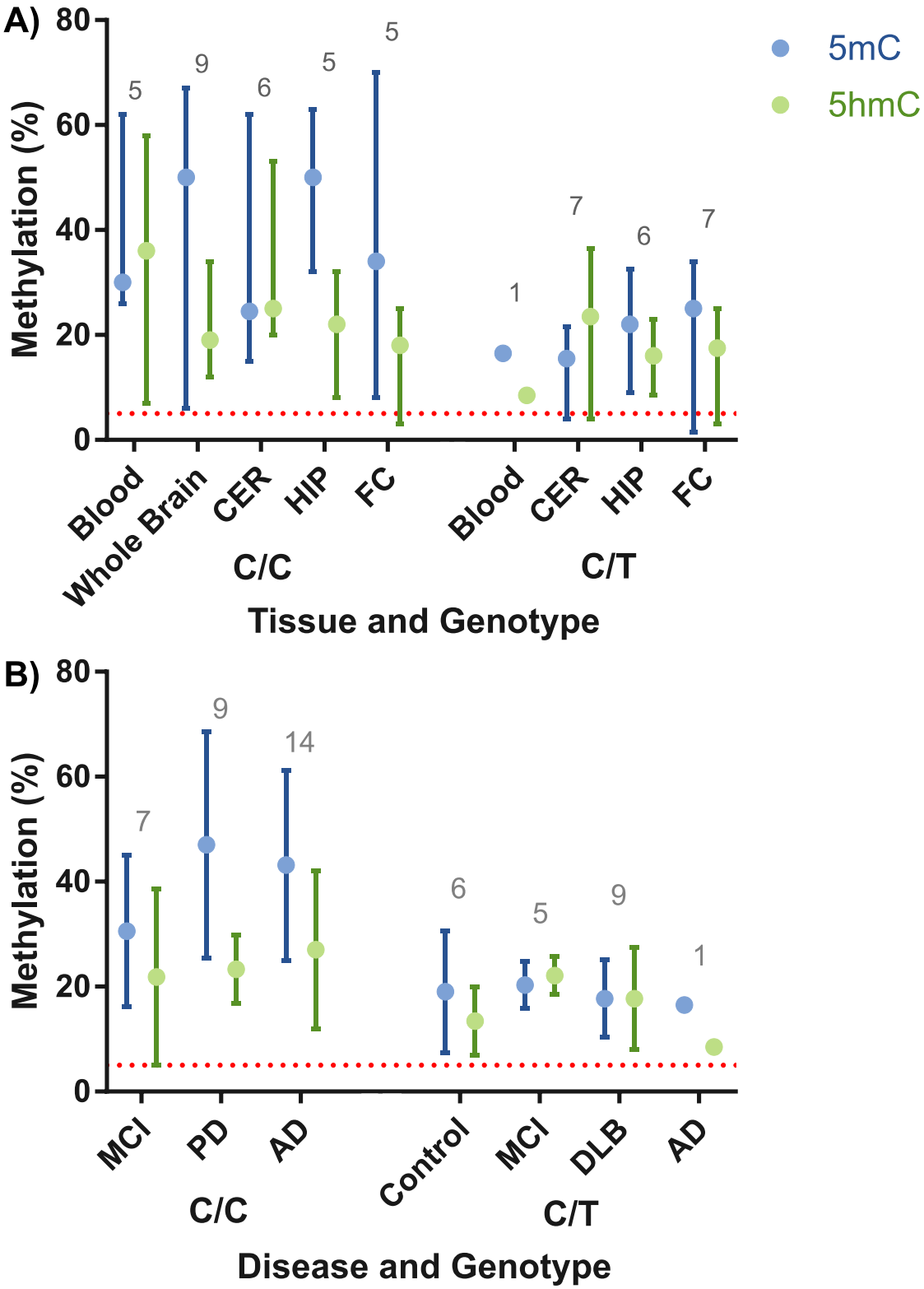


Figure 4.3 Median percentages and ranges of 5mC and 5hmC methylation at AS-rs6265 in C homozygous and heterozygous genotypes across tissue/disease states. 5mC is shown in blue, 5hmC is shown in green for both panels. Panel A displays methylation across tissues, where CER, HIP and FC refer to cerebellum, hippocampus, and frontal cortex, respectively. Panel B shows methylation across disease states, where MCI, PD, DLB, and AD, refer to mild cognitive impairment, Parkinson's disease, dementia with Lewy bodies, and Alzheimer's disease, respectively. The number of data points for each group (N) is displayed above the methylation levels.

4.4.3 Tissue, Disease, genotype status and AS-rs6265 5fC/CaC methylation

5fC/CaC methylation data generated by the MAB-Seq method were compared with 5mC/5hmC methylation data in analysis. Due to the nature of MAB-Seq, only C homozygotes could be analysed for 5fC/CaC quantification. Of the 30 C homozygous C non-carrier data points available, 28 data points were available for 5fC/CaC methylation analysis (Table 4.3). No tissue groups were excluded.

As above, 5fC/CaC methylation analysis was performed by creating 3 Linear Mixed Effects Models, and removing variables from the models to create the minimum adequate model and predict the effect of each variable on 5fC/CaC methylation. Because only C homozygotes could be used, genotype could not be assessed as a variable. The models investigated the effect of disease and tissue variables across all samples; effect of disease and tissue variables across the Imperial brain bank samples; and finally, the effect of tissue in the OPTIMA samples. A summary of the effect of each variable when creating the minimum adequate model for 5fC/CaC methylation may be seen in Table 4.5.

LMEM Analysis of 5fC/CaC methylation

Unlike the analysis of 5mC and 5hmC methylation, analysis of 5fC/CaC methylation indicated that the most significant terms were interactions between the tissue and disease variable (Tissue:Disease). This is observable in the analysis of the Imperial and OPTIMA data whereby the only significant term is the interaction between tissue and disease ($X^2(2) = 9.36$, $p = 0.01$). Both tissue ($X^2(3) = 2.90$, $p = 0.41$) and disease ($X^2(1) = 0.94$, $p = 0.33$) variables were not significant on their own.

This effect may be driven by the Imperial samples as a similar effect can be seen in the analysis of 5fC/CaC in these samples. Here, once again the interaction of tissue and disease ($X^2(2) = 10.59$, $p = 0.01$) was significant, while tissue ($X^2(2) = 1.92$, $p = 0.38$) and disease ($X^2(1) = 0.89$, $p = 0.35$) terms alone, were not. Unfortunately, because only one disease was present in the OPTIMA samples (Alzheimer's disease) only the tissue effect could be analysed ($X^2(1) = 0.99$, $p = 0.32$).

While not shown here, all models for both the 5mC, 5hmC, and 5fC/CaC analysis had the derivative residuals used to determine normal distribution and homoskedasticity. All models reported normal distribution and homoskedasticity. The lack of significance observed in the LMEM analysis across both the Imperial and OPTIMA data sets in methylation analysis may be in part to the non-factorial design of the study. Because only OPTIMA samples have Alzheimer's diagnoses and blood/whole brain tissues, this will directly affect the data available when investigating for differences between variables and methylation.

To visualise the relationship between 5mC methylation and tissues/disease, samples were divided into subgroups based on tissue type (blood, whole brain, frontal cortex (FC), hippocampus (HIP) and cerebellum (CER)), diagnosis (Control, mild cognitive impairment (MCI), Parkinson's disease (PD), dementia with Lewy bodies (DLB), and Alzheimer's disease (AD)), and carrier status (carrier vs. non-carrier). Medians and range of methylation levels are presented in panel A of Figure 4.4 and Figure 4.5, for tissue and disease groups, respectively. The averages of the methylation types (5mC, 5hmC and 5fC/CaC) are presented in panel B of Figure 4.4 and Figure 4.5 for tissue and disease groups, respectively.

The ratios indicate that the cerebellum shows a higher level of 5hmC when compared against 5mC methylation levels and other tissue type levels. The highest reported average levels of 5fC/CaC methylation were in the cerebellum (35%) and frontal cortex (36%), although 5fC/CaC was the least abundant form of methylation in all tissues. While both Parkinson's disease and Alzheimer's disease states showed similar distributions of methylation with the majority of methylation being 5mC, MCI showed a 5fC/CaC dominant distribution of methylation. All methylation types in all disease states showed large ranges. Quantification of 5fC/CaC methylation indicates that AS-rs6265 is an epiallele that is 5mC, 5hmC and 5fC/CaC methylated.

Methylation being analysed	Data set	Term	X^2	DF	P value
5fC/CaC	Imperial and OPTIMA	Tissue:Disease	9.36	2	0.01*
5fC/CaC	Imperial and OPTIMA	Tissue	2.9	3	0.41
5fC/CaC	Imperial and OPTIMA	Disease	0.94	1	0.33
5fC/CaC	Imperial	Tissue:Disease	10.59	2	0.01*
5fC/CaC	Imperial	Tissue	1.92	2	0.38
5fC/CaC	Imperial	Disease	0.89	1	0.35
5fC/CaC	OPTIMA	Tissue	0.99	1	0.32

Table 4.5 Linear mixed effects model analysis for disease, tissue, and genotype variables across different samples and 5fC/CaC methylation. The imperial data set contained frontal cortex, hippocampus, and cerebellum brain tissues, as well as dementia with Lewy bodies, Parkinson's disease, mild cognitive impairment, and control diagnoses. The OPTIMA data set contained Alzheimer's diagnoses across blood and whole brain tissues. Term refers to the variable being investigated, while terms separated by a colon (e.g. Tissue:Disease) indicate interaction terms between the variables. X^2 refers to the ANOVA derived Chi^2 statistic, DF refers to degrees of freedom, and P value indicates significance. (* $p \leq 0.05$).

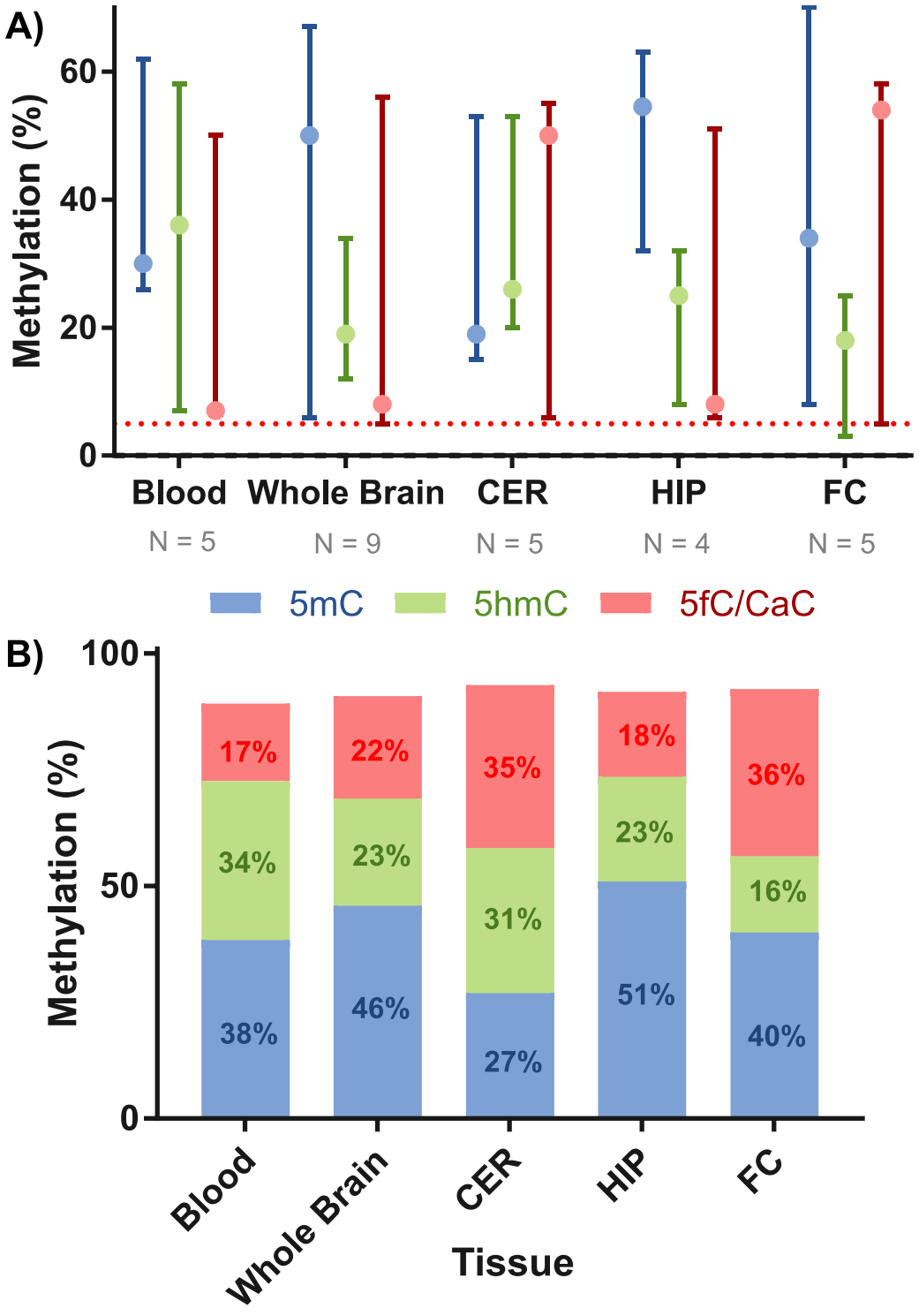


Figure 4.4 5mC, 5hmC, and 5fC/CaC methylation across AS-rs6265 C homozygote tissues. A) Median percentages and ranges of 5mC (blue), 5hmC (green), and 5fC/CaC (red). Tissues are blood, whole brain, cerebellum (CER), frontal cortex (FC) and hippocampus (HIP). Red dotted line indicates the 5% margin of error reported by pyrosequencing. Sample numbers (N) are shown below the tissue labels on the X-axis. B) Means of 5mC, 5hmC, and 5fC/CaC methylation types, and unmodified cytosine across the different tissues. Mean percentages are shown on the corresponding bar colour under each tissue, where percentages have been rounded to the nearest integer.

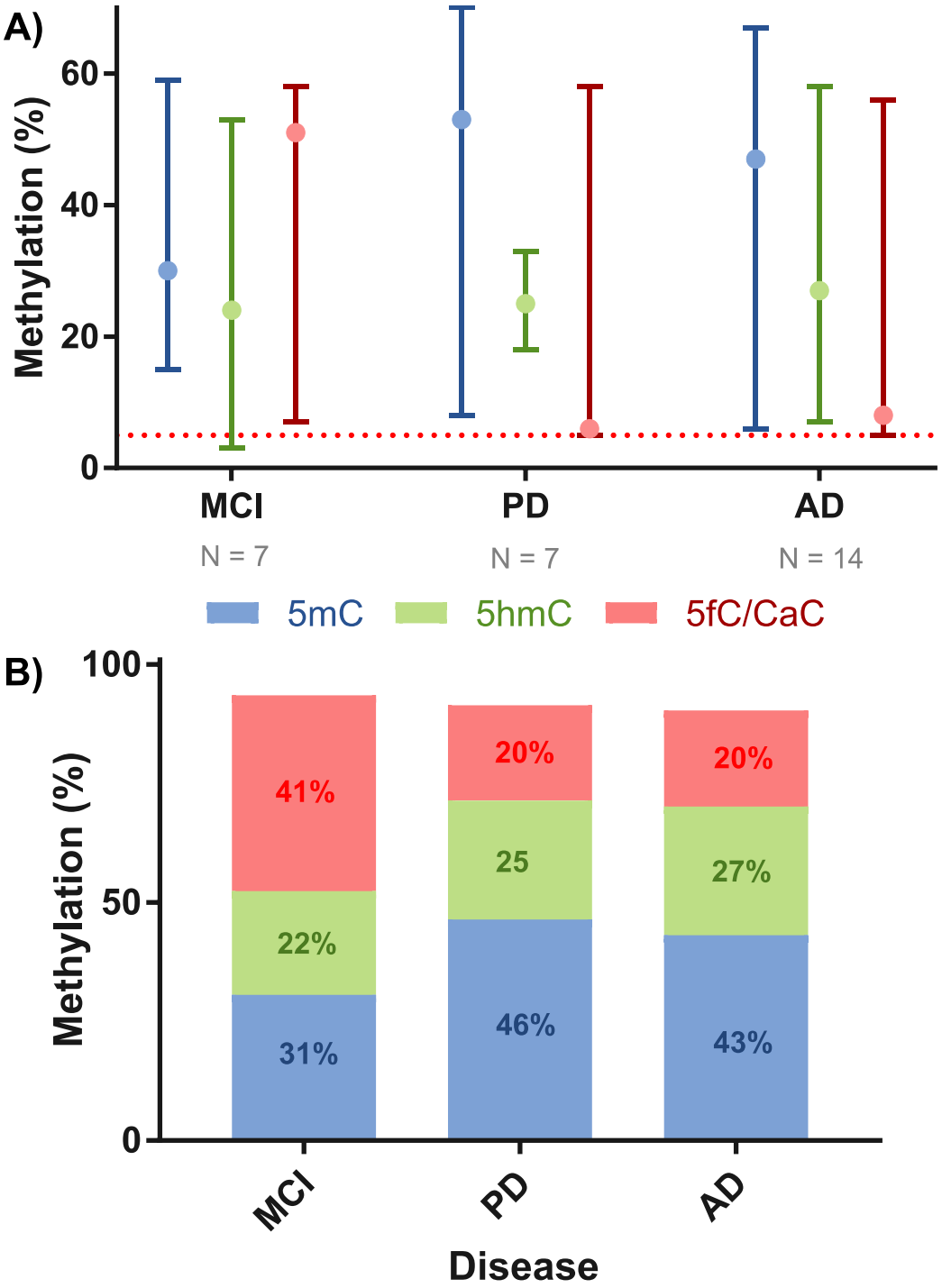


Figure 4.5 5mC, 5hmC, and 5fC/CaC methylation at AS-rs6265 across disease states in C homozygotes. A) median percentages and range bars for 5mC (blue), 5hmC (green), and 5fC/CaC (red) methylation. Methylation medians are shown across tissue groups (control, mild cognitive impairment (MCI), Parkinson's disease (PD), dementia with Lewy bodies (DLB), Alzheimer's disease (AD)). Red dotted line indicates the 5% margin of error reported by pyrosequencing. Sample numbers (N) are shown below the tissue labels on the X-axis. B) the mean percentages of different methylation types (5mC, 5hmC, 5fC/CaC) are represented as the same colours as is panel A, across different diagnoses (represented as in panel A). Mean percentages are shown on the corresponding bar colour under each disease, where percentages have been rounded to the nearest integer.

4.5 Discussion

4.5.1 BDNF rs6265 genotype association

The significant association found between the sense rs6265 minor A allele (66Met) and Schizophrenia and Schizoaffective disorders evident in this current study has previously been reported (F Angelucci, Brenè, & Mathé, 2005; Chao et al., 2008; Shoval & Weizman, 2005). Research suggests that the rs6265 variant may affect brain architecture and neurotransmission and could contribute through such mechanisms to the expression of SCZ by lowering the age of onset (Chao et al., 2008; Michael F. Egan et al., 2003; Hariri et al., 2003). Patients with SCZ were also found to have significantly lower levels of BDNF in blood serum compared to controls and this was associated with carrying the rs6265 minor A allele (Zakharyan & Boyajyan, 2014). This reduction of BDNF may contribute to the SCZ phenotype through the dysregulation of axonal and dendritic differentiation in development (Matsuda et al., 2009).

The findings of the association analysis performed in this study confirm an association between the rs6265 Met66 minor A allele and Autism Spectrum Disorders including autism and Asperger's diagnoses (ASD) phenotypes. However, previous studies of the rs6265 A minor allele in ASD cohorts have shown inconsistent findings. Studies have previously connected ASD phenotypes with BDNF hyperactivity in early life, however this is not connected to the rs6265 variant (Tsai, 2005). Moreover, as the rs6265 minor allele is thought to alter the activity-dependent expression of BDNF, the constitutive release of BDNF is unaffected (Michael F. Egan et al., 2003). Literature has suggested that ASD phenotypes show differential levels of BDNF in blood serum (Miyazaki et al., 2004; K. B. Nelson et al., 2001). However carrier status at rs6265 has not been shown to be significantly associated with ASD (Meng et al., 2017). In addition, association studies have begun to determine if the rs6265 variant may be enabling ASD phenotypes by altering BDNF protein abundance *in vivo* (Yoo et al., 2014). In family based associations no association was determined between the rs6265 minor allele and autism (Yoo et al., 2014). However quantitative transmission disequilibrium tests identified associations between the rs6265 minor allele and quantitative traits such as restricted, repetitive

and stereotyped patterns of behaviour, which are key symptoms of ASD (Yoo et al., 2014). Our findings support the association between risk for ASD and SCZ pathologies and carrying the A minor allele of rs6265. However how the rs6265 minor allele may be contributing to ASD phenotypes is unknown at this time.

4.5.2 5mC and 5hmC methylation across antisense rs6265 carrier status

The hypothesis of how this genetic change could contribute to disease is currently more focused on changes in the amino acid (Val to Met) and consequences on short term plasticity and learning (McHughen et al., 2010; Murer et al., 2001; Zuccato & Cattaneo, 2009). However *BDNF* rs6265 methylation is known to play a role in disease (Ursini et al., 2016). Previous studies have identified methylation through bisulphite sequencing at the sense rs6265 CpG, where the cytosine methylated is the neighbouring cytosine of rs6265 (Ursini et al., 2016). In the current study, a significant difference was observed between the 5mC and 5hmC levels at AS-rs6265 in non-carriers (C homozygotes) but not carriers of the T minor allele (heterozygotes). It is known that 99% of CpG sites in human cell lines are methylated on both strands (Lister et al., 2009). However, no research has endeavoured to quantify antisense methylation at key sites. At the point of writing, this study is believed to be the only quantification of 5mC and 5hmC methylation in both C homozygotes and T allele carriers at AS-rs6265 across brain diseases using DNA from human brain.

While there was a significant difference between AS-rs6265 5mC and 5hmC methylation in non-carriers, this was not present in carriers of the rs6265 minor allele and this may be due to the loss of a CpG site in AS-rs6265. This effect, known as allele specific methylation (ASM), is reported to be up to 37% bisulphite sequencing quantified methylation at heterozygous genotype sites (Shoemaker, Deng, Wang, & Zhang, 2010). The variation in 5mC/5hmC methylation at the C/T heterozygous genotypes in the current study is greater and partial or “fuzzy” methylation (whereby CpG sites have intermediate levels of methylation; 25% - 75%) is more common at heterozygous sites (Shoemaker et al., 2010).

4.5.3 5mC and 5hmC methylation across tissues and cognitive disease

In the last 10 years there has been growing evidence that the oxidised 5hmC methylation state is more abundant than the originally identified 5mC state in specific tissues (Wen et al., 2014). Linear mixed effects model analysis of the levels of 5mC and 5hmC methylation across tissue types in different groups indicated there were three significant variables contributing to differences in observed methylation. Firstly, tissue type significantly contributed to the levels of 5mC methylation observed in the Imperial samples ($p = 0.02$). Similarly, both tissues ($p = 0.05$) and genotype ($p = 0.05$) contributed to the differential levels of 5hmC observed in the Imperial samples, however the interaction between tissues and genotype was not significant ($p = 0.96$) indicating that they individually contribute more influence in methylation. The observation of significant effects of genotype on 5hmC methylation further support the theorised action of ASM in these samples. Further analysis with methylation types and transcript expression in chapter 5 may further elucidate the action of methylation types in different tissues and disease states.

While it is not possible to determine significant differences between specific tissues and genotypes in the samples using LMEM analysis, some observational interpretations were performed (Figure 4.3 Panel A). Median 5mC methylation was observed to be higher than 5hmC across most brain regions of AS-rs6265 non-carriers, and this supports the findings of the methylation across genotype analysis. 5hmC methylation is known to be enriched in the mouse genome in neurons, Purkinje cells, and the brain as a whole compared to other tissues such as blood (Globisch et al., 2010; Kriaucionis & Heintz, 2009; C. X. Song et al., 2011). However, these studies only looked at sense strand 5hmC methylation in mouse cerebellum or used HPLC-MS or thin layer chromatography (TLC) measurements of 5hmC, which are not strand specific. Therefore, while this study did not identify higher levels of 5hmC compared to 5mC in brain tissues, this may be due to species or strand specific methylation (Patiño-Parrado et al., 2017).

AS-rs6265 blood methylation showed relatively high levels of 5hmC compared to 5mC in non-carriers, while all non-carrier brain regions reported higher levels of 5mC compared to 5hmC. This increase in 5mC compared to 5hmC in non-carriers was consistently shown across the brain regions, and contrasts with other research that reports that 5hmC is typically enriched in the brain, and 5mC methylation is higher in blood. These results also allude to a tissue specific profile of methylation states which may relate to cell specific profiles as many brain tissues show different cell type compositions (Hawrylycz et al., 2012).

An increased amount of AS-rs6265 5mC/5hmC methylation in the frontal cortex of disease samples was also predicted. This is because previous studies have suggested an increase in methylation quantified through bisulphite sequencing in *BDNF* regions in cognitive and neuropsychiatric diseases such as bipolar disorder (Rao, Keleshian, Klein, & Rapoport, 2012). For instance the promoter region of BDNF was found to be hypermethylated in Alzheimer's disease and bipolar disorder phenotypes through bisulphite quantification of methylation (Rao et al., 2012). Additionally the promoter I and IV of *BDNF* have been shown through bisulphite sequencing to be significantly more methylated in amnestic MCI phenotypes that progress onto Alzheimer's disease, compared to those who do not develop AD (Xie et al., 2017). The increased methylation is associated with a reduction in BDNF expression as well as hypermethylation of the CREB promoter region, which may exacerbate reduced BDNF levels (Rao et al., 2012; Xie et al., 2017).

The cerebellum tissue region showed higher 5hmC than 5mC methylation or 5hmC roughly equal to 5mC methylation in heterozygotes and C homozygotes, respectively. It is possible that this was driven by the Parkinson's disease sample (due to the non-factorial design of the study) as it has been shown that there is a genome-wide significantly higher level of 5hmC in the cerebellum of PD patients compared to controls (Stöger, Scaife, Shephard, & Chakrabarti, 2017). It may also be driven by the reported enrichment of 5hmC methylation in the Purkinje cells of the cerebellum with age (F. C. Zhou, Resendiz, Lo, & Chen, 2016). This study's findings also indicated frontal cortex and hippocampus tissue regions observationally possessed higher

levels of 5mC compared to 5hmC in both carriers and non-carriers. While none of these observations were significant in difference, additional testing would be valuable to confirm the patterns observed between 5mC and 5hmC methylation, tissues, and disease.

5mC and 5hmC methylation in cognitive disease

Research suggests 5hmC methylation is thought to act as a cis-regulatory element (CRE) promoting and repressing gene expression and altering chromatin conformation (D. Q. Shi et al., 2017). Whole genome disruption in the level of 5hmC or the ratio of 5mC to 5hmC has been associated with Alzheimer's disease (Chouliaras et al., 2013), schizophrenia (Matrisciano et al., 2013), Parkinson's disease (Stöger et al., 2017), and other neurological disorders (E. Dong, Gavin, Chen, & Davis, 2012; D. Q. Shi et al., 2017). In addition to this, in mice 5hmC levels may vary and can increase or decrease expression of transcripts dependent on its location in genes (Hahn et al., 2013; S.-G. Jin et al., 2011). For instance activation of *SOX5*, *NAV2*, *UCHL1*, *PDE2A* and *ELK3* genes in mice were associated with increased intragenic 5hmC methylation at these gene sites (Hahn et al., 2013). Furthermore, 5hmC enrichment in gene bodies was significantly associated with *MCF2L*, *CHST3*, *EPPK1* and more gene expression levels (S.-G. Jin et al., 2011). Therefore, it would be expected that significantly different patterns of 5hmC methylation would be seen between control and disease samples.

In the current study, disease states did not significantly contribute to the levels of 5mC or 5hmC methylation observed, in either the Imperial or OPTIMA cohorts of samples (Table 4.5). Methylation patterns in the diseases were similar to that of the tissue findings in that AS-rs6265 5mC was observationally higher in level compared to 5hmC methylation (Figure 4.3 Panel B). Namely, Parkinson's disease and Alzheimer's disease diagnoses within non-carrier samples showed higher levels of 5mC compared to 5hmC, which may indicate disruption in transcription and therefore disease (Chouliaras et al., 2013; D. Q. Shi et al., 2017; Stöger et al., 2017). The high level of AS-rs6265 5mC to 5hmC in Parkinson's disease samples went against literature suggesting there is higher 5hmC levels in the cerebellum of Parkinson's

disease patients (Stöger et al., 2017). Moreover, as the only Parkinson's disease samples were in non-carriers, we could not determine if this was ASM or partial/fuzzy methylation. In addition to this, all control samples were AS-rs6265 carriers meaning there were no non-carrier controls to compare against the Parkinson's disease 5hmC levels.

Carrier dementia with Lewy bodies individuals showed similar levels of AS-rs6265 5mC and 5hmC compared to carrier controls (Figure 4.3 Panel B). These findings also contrast with current literature that suggested that global hypomethylation events occur in the brain samples of dementia with Lewy bodies patients (Desplats et al., 2011). This is facilitated by the sequestration of DNMT1 in the cytoplasm of dementia with Lewy bodies patients, due to the interaction of DNMT1 and α-synuclein (Desplats et al., 2011). The disease analysis performed in this present study suggest that the median levels of AS-rs6265 5mC and 5hmC are higher in carrier dementia with Lewy bodies than carrier control samples.

Higher levels of AS-rs6265 5hmC were observed in carrier compared to non-carrier 5hmC samples across disease states (Figure 4.3 Panel B). Furthermore, while non-carrier mild cognitive impairment AS-rs6265 methylation levels were similar to that of carrier control samples, carrier mild cognitive impairment samples showed higher levels of 5mC and 5hmC. As carriers have half the number of potential CpG sites to methylate than non-carriers, this may be ASM. Previous studies have identified multiple differentially methylated regions (DMRs) associated with mild cognitive impairment phenotypes in blood (Chouliaras et al., 2018; Park et al., 2018), and these results from brain DNA may be reflecting a similar pattern.

Literature suggests that enhanced methylation of *BDNF* promoters contribute in the development of amnestic mild cognitive impairment to Alzheimer's disease (Xie et al., 2017). Moreover rs6265 homozygous T carriers show a 23% increase in chance of developing amnestic mild cognitive impairment, and methylation of *BDNF* promoter regions also increases the chance to develop Alzheimer's disease from mild cognitive impairment (Xie et al., 2017). Higher levels of AS-rs6265 methylation in carriers in the

current study may be contributing to the development of mild cognitive impairment in a genotype in an ASM driven manner. Alzheimer's disease diagnoses observationally showed higher levels of AS-rs6265 5mC compared to 5hmC in non-carriers which is known to be associated with Alzheimer's disease phenotype (Chouliaras et al., 2013). The findings may support studies that show that *BDNF* is 5mC enriched in the frontal cortex of Alzheimer's disease patients (Levine et al., 2015; Rao et al., 2012).

4.5.4 5fC/CaC methylation analyses

5fC and 5CaC were previously thought to be temporary states in base excision repair, but it is now known that 5fC can be permanent and alter transcription and expression of genes (Bachman et al., 2015; Raiber et al., 2014). At the time of writing, this study is the first to show that the BDNF AS-rs6265 epiallele is 5fC and/or 5CaC methylated in brain tissue. LMEM analysis indicated that the interaction of tissue and disease variables (Tissue:Disease) significantly contributed to the levels of 5fC/CaC methylation observed in both the Imperial samples ($p = 0.01$), and the Imperial and OPTIMA samples analysed together ($p = 0.01$) (Table 4.5). This indicates that while the specific disease state in specific tissues affected the 5fC/CaC level, both tissue and disease individually, did not. Because the significant difference generated by the interaction of tissue and disease were observed in all the samples as a whole, and the imperial samples, this may suggest that the Imperial samples are driving the significant effect observed when analysing all samples.

While the observed levels of 5fC/CaC in this study were low in most states, they were still higher than other studies quantifications of sense strand 5fC/CaC (Y. Chen, Damayanti, Irudayaraj, Dunn, & Zhou, 2014; D. Q. Shi et al., 2017). The distribution of 5fC/CaC methylation in most samples was bimodal with either very high or low levels of 5fC/CaC methylation, but rarely an intermediate value. This is illustrated in the large range bars and low medians in most tissue and disease analyses (Figures 4.4 & 4.5). In mouse and zebrafish tissues methylation has been shown to be highly

species specific and also indicated a bimodal distribution (C. Zhang, Hoshida, & Sadler, 2016).

5fC/CaC methylation across tissues

5fC/CaC methylation is putatively detected in the brain in smaller concentrations (around 1% of 5hmC levels) compared to 5mC and 5hmC and exist in neurons throughout adulthood (Y. Chen et al., 2014; D. Q. Shi et al., 2017). 5fC/CaC methylation of an artificial promoter region CpG has also been shown to possess a thymine-DNA glycosylase (TDG) mediated transcriptionally repressive effect (Kitsera et al., 2017). Therefore, it would be expected that 5fC/CaC methylation of the *BDNF* AS-rs6265 epiallele would result in transcriptional repression of AS-BDNF.

AS-rs6265 5fC/CaC levels observed were on average similar to concentrations of 5hmC (as seen in whole brain, cerebellum, and hippocampus), if not higher than 5hmC levels (as in the frontal cortex) (Figure 4.4 Panel B). It should be noted, however, that the reported percentages are means which are limited in their representation of bimodal data. These results may allude to a tissue or cell specific profile as many brain tissues show different cell type compositions (Hawrylycz et al., 2012).

5fC/CaC methylation in cognitive disease

Little current literature is available on 5fC and 5CaC methylation at CpG resolution in pathological brain samples. Most literature suggests 5hmC is more concentrated than 5fC/CaC in most tissues (Y. Chen et al., 2014; D. Q. Shi et al., 2017), and the methylation patterns observed in both Parkinson's disease and Alzheimer's disease samples support this. However, mild cognitive impairment samples showed a high level of 5fC/CaC methylation compared to other disease states, but as this finding was non-significant and there were no controls to compare, nothing can be determined from this observation.

Current literature suggests brain is enriched genome-wide for 5hmC compared to blood and Parkinson's disease diagnoses are enriched for 5hmC compared to controls (Stöger et al., 2017). This may in part explain why the range of 5hmC methylation in Parkinson's disease samples was higher compared to other disease states, however this is once again non-significant and cannot be compared with controls. DNA methylation is known to play a role in and may be a valuable biomarker for cognitive dysfunction (Chouliaras et al., 2018) and more research must be performed to understand the role of these less studied forms of methylation in disease.

Antisense methylation

While previous studies have observed methylation through bisulphite sequencing at rs6265, the AS-rs6265 CpG has never been investigated (Ursini et al., 2016). While these results have broadly agreed with other research, such as consistently higher levels of 5mC than 5hmC/5fC/5CaC at CpGs, there have also been larger differences. For instance, 5hmC was not observed as being enriched across the genome in most brain regions (Szulwach et al., 2011). This may indicate that while the sense strand reports enrichment for 5hmC in brain regions, the antisense strand does not. While literature has indicated that almost all methylated CpGs show complementary methylation at the opposite strand CpG, the composition of these methylation types has not been investigated.

Additionally anti-sense transcripts are known to lead to gene silencing mediated by methylation (Tufarelli et al., 2003). In a rare case of α-thalassemia, a truncated form of the Haemoglobin gene, *HBA2*, became methylated at its CpG islands through the transcription of anti-sense RNA, which led to transcriptional silencing (Tufarelli et al., 2003). This effect was reproducible even when using different promoter regions which indicates that it is a general mechanism, capable of happening at other points in the genome (Tufarelli et al., 2003). This is similar to X inactivation mediated repression of gene transcription, facilitated through the expression of anti-sense RNA which leads to the methylation of gene promoter CpG islands, such as with *IGF2R* (Tufarelli et al., 2003). The mechanism by which anti-sense RNA is exerting this effect

on methylation is unknown, but understanding this effect will be important in providing insight into genotype, methylation, and disease dynamics.

4.5.5 Limitations and future studies

The major limitations found in this study were derived from the small number and unmatched and non-factorial design of the samples used. Subsequently linear mixed effect models were used to determine if any variable (genotype, disease, tissue, or their interactions) were significantly contributing to differences in methylation levels in the Imperial, OPTIMA, or Imperial & OPTIMA cohorts. Furthermore, statistically differentiating the effects of specific diseases and tissues within analysis was not possible, and only observational analysis could be performed. For instance, the different patterns observed between non-carrier Parkinson's disease 5hmC methylation and carrier mild cognitive impairment 5hmC methylation, may be due to the disease state, AS-rs6265 genotype, or because of tissue specific methylation patterns. For future studies increasing sample size and sequencing samples until all samples have been adequately sequenced would remove many of these limitations, and allow for a better study design, and allowing for parametric testing.

It should also be kept in mind that these methylation data are from one variant locus CpG site; AS-rs6265. Such methylation profiles may be very different when assessing other genomic sites and across the whole genome. This study has observed trends in methylation across diseased and control tissue, but this one variant site may not be representative of all the trends observed. Methylation biomarkers are valuable in the progression of personalised medicine, especially in the progression of pathologies associated with ageing. This highlights the need for larger sample sizes to determine if the trends observed in this study are due to sample variation or indicative of the mechanisms of pathology.

Future studies may wish to repeat this study with larger numbers, using liquid chromatography mass-spectrometry to quantify methylation. As this method is more sensitive in the quantification of methylation, levels of methylation below 5% (as is

common with 5fC/CaC) will be better quantified. Additionally, use of additional reagents in MAB-Seq such as $NaBH_4$ would allow for the differentiation between 5fC and 5CaC methylation in quantification. This would allow for the quantification of all cytosine DNA methylation at base resolution across different tissues and brain regions. Finally, quantifying methylation in the same samples on the *BDNF* sense strand would allow for comparison on methylation types and levels in the complementary CpG sites.

As yet, this field is in its infancy and 5hmC, 5fC, and 5CaC methylation at specific risk epialleles in the context of surrounding methylation, expression, and their effect on disease, are under researched. Further quantification of methylation in blood along with brain tissue regions would be of importance in diagnostic medicine. While understanding the methylation patterns in the brain is key in elucidating mechanisms behind disease phenotypes, unless this pathogenic methylation phenotype can be identified from blood, diagnosis using methylation patterns would be very invasive if brain tissue were required from living individuals. Additionally, quantification of methylation from cerebral spinal fluid may act as an intermediary in diagnosis, especially if differences in methylation used in diagnosis are small.

The bases surrounding the CpG dinucleotide may also influence the level of methylation and the potential variation (and bimodality) of methylation levels (L. Zhang, Gu, Yang, Tang, & Gao, 2017). For instance, CpG sites that are preceded by a guanine or cytosine show less variation in the level of methylation present at the CpG site, which is conserved in a cell, tissue and species specific manner (L. Zhang et al., 2017). If this is the case, future studies could analyse the surrounding base sequence to better quantify true differential methylation between bimodally distributed methylation levels. This may be valuable in the identification of differentially methylated loci in the characterisation of risk factors for disease.

4.5.6 Conclusion

In summary, the findings described in this chapter support research that suggests there is a significant association between cognitive disease states such as ASD and SCZ and the possession of the rs6265 Met66 minor allele. 5mC methylation was shown to be significantly higher than 5hmC methylation in AS-rs6265 non-carriers. Tissue and tissue & genotype variables significantly contributed to the observed levels of 5mC and 5hmC methylation, respectively, in the Imperial cohort. The interaction between tissue and disease significantly contributed to the levels of 5fC/CaC observed in the Imperial, and Imperial & OPTIMA cohorts. Different patterns of 5mC, and 5hmC methylation at *BDNF* AS-rs6265 were observed between different genotypes, tissue regions, and disease phenotypes. Additional research is required to determine if regulation of DNA methylation could contribute to the biological processes underlying brain diseases. The mechanisms by which the observed methylation may contribute to disease through regulation will be explored further in chapter 5.

5mC was typically the dominant form of methylation in all analyses, and 5mC methylation was significantly higher than 5hmC methylation in AS-rs6265 C homozygotes. Contrary to most research, an enrichment of 5mC methylation was observed in the brain and while non-significant, was higher than 5hmC in whole brain and hippocampus tissue region analyses, as well as Parkinson's disease and Alzheimer's disease analyses of C homozygotes. Our results do not support the current view that there are global hypomethylation events during dementia with Lewy bodies progression. 5hmC methylation was found to be higher in T allele carriers compared to 5hmC methylation in C homozygotes, despite having half the amount of CpG sites.

5fC/CaC methylation was observed at the AS-rs6265 variant site, which is the first time it has been quantified at this epiallele. 5fC/CaC methylation showed a bimodal distribution which is in keeping with other research, however was detected at higher levels compared to other research (Y. Chen et al., 2014; D. Q. Shi et al., 2017). While

non-significant the pattern of methylation observed in this study may indicate disease and genotype specific methylation profiles at the AS-rs6265 variant site could be valuable as a potential biomarker for diagnostic medicines and therapies to slow disease progression.

Chapter 5: *BDNF* rs6265 epiallele methylation and *BDNF* transcript abundance

Acknowledgements

The analysis of methylation at the BDNF variant site would not have been possible without the assistance and advice from a number of people. Maria Isabelle-Haig, University of Nottingham, gave advice on the optimisation of the Sanger sequencing primers used to target *BDNF*. Dominika Valent Raffajová, University of Nottingham, assisted in the preparation to genotype samples, and along with Dr Andrew Benest, University of Nottingham, aided in the optimisation and analysis of DD-PCR for the study of transcriptional expression of *BDNF*.

5.1 Workflow pipeline for the analysis of BDNF

As discussed in chapter 4, BDNF proteins play a key role in cognition and impaired BDNF function putatively contributes to cognitive and neurological diseases (Day & Sweatt, 2011b; Petzold et al., 2015). In addition to the rs6265 variant's genotype, and the resultant valine to methionine substitution in translated proteins, local changes in expression of *BDNF* may also be regulating phenotype (Pruunsild et al., 2007).

It is possible that DNA methylation of the AS-rs6265 major allele may regulate expression of the primary *BDNF* transcripts. This chapter will first discuss the results obtained by examining *BDNF* transcript abundance with AS-rs6265 genotype, tissue, and disease states, followed by correlation analysis between methylation at antisense *BDNF* rs6265, AS-rs6265 genotype, tissue, and disease states. A diagram showing the *BDNF* research workflow pipeline is presented in Figure 4.1.

5.2 *BDNF* transcript expression and correlation within tissues and disease states

Digital droplet PCR (DD-PCR) was performed in order to determine transcript abundance of the most highly expressed *BDNF* transcripts in brain tissue using the Imperial brain bank unaffected-control, MCI, PD and DLB samples (N = 14, 80 data points total) (Table 5.1). Each sample had tissue from three brain regions (frontal cortex, hippocampus and cerebellum). Four BDNF transcripts were targeted by two *BDNF* probes (see Section 2.2). The first probe targeted the two most abundant sense strand transcripts in brain (ENST00000395981 & ENST0000039580.6, or *BDNF*

transcripts 205 & 204, respectively) which were grouped as a single population and referred to as *BDNF* "sense" or "mRNA" transcript hereafter. The second probe targeted two different anti-sense transcripts (NR_033314 and NR_002832.2) known as AS-201 and AS-204, respectively. A diagram showing details on rs6265 and the transcripts may be seen in Figure 5.1.

The *BDNF* sense and anti-sense transcripts were targeted with VIC and FAM dye reporters respectively, to distinguish transcripts. AS-201 and AS-204 were quantified using the same probe, and were differentiated by the identification of two distinct populations in quantification. Comparing anti-sense transcript abundance with current expression database profiles differentiated AS-201 and AS-204 from each other and negative droplet data (Gte. Consortium, 2015). All transcript abundance was expressed relative to the abundance of TATA-box binding protein (*TBP*) housekeeping gene transcript ENST00000392092.6. The *BDNF* rs6265 variant is located within *BDNF* exon IX or AS-*BDNF* exon V and is present in all of these transcripts and hence methylation of the rs6265 epiallele could have a direct effect upon their expression.

5.2.1 Transcript expression correlations

Spearman correlation coefficient was performed to determine if transcript showed high correlations in their expression profiles. No correlations between sense and antisense AS-201 transcripts (Figure 5.2 Panel A) or the AS-201 and AS-204 transcripts (Figure 5.2 Panel B) were evident. In contrast, there was a strong correlation present between the sense and AS-204 transcripts ($r^2 = 0.88$, $P < 1\times10^{-3}$) across cerebellum, hippocampus and frontal cortex tissues (Figure 5.2 Panel C).

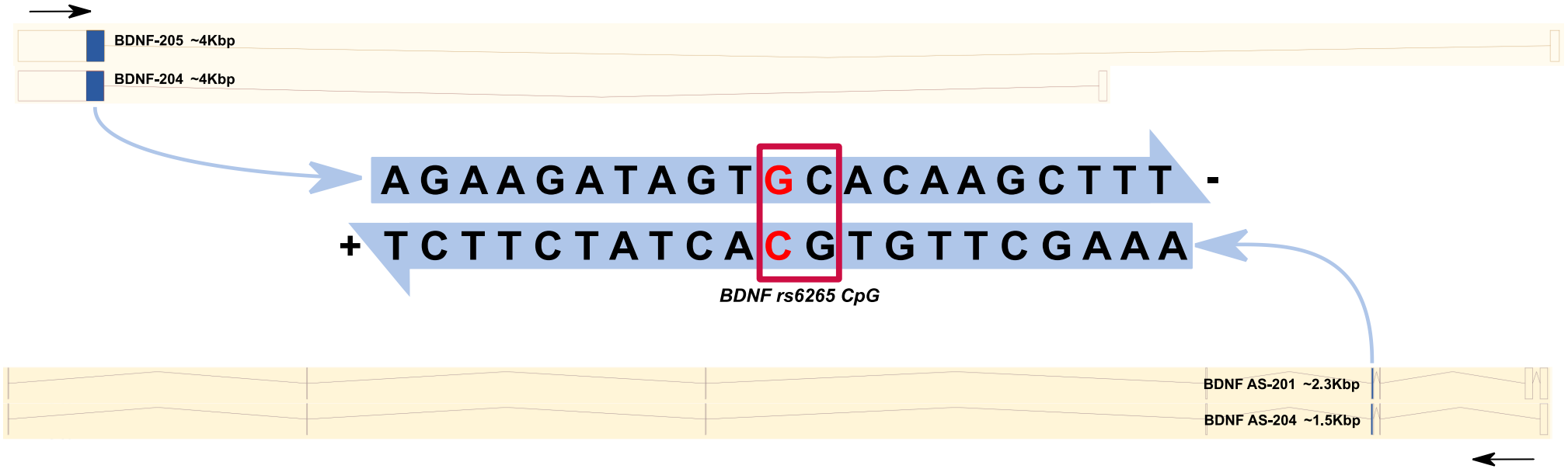


Figure 5.1 Sense and antisense transcripts at *BDNF* variant rs6265 with a major G allele genotype. In this study BDNF mRNA transcripts 204 and 205 are grouped as "sense" or "mRNA" transcript and are shown at the top, and as a sequence with "-". Both sense strand transcripts are approximately 4Kbp in length. Antisense transcripts 201 & 204 (AS-201 and AS-204) are shown on the bottom and as a sequence with "+". AS-201 is approximately 2.3Kbp, while AS-204 is approximately 1.5Kbp in length. *BDNF* rs6265 is displayed in red, (major allele is G on sense strand, C on antisense strand) and the rs6265 CpG has been highlighted by a red box. Possession of the minor A/T allele at the rs6265 variant disrupts the CpG. The *BDNF* rs6265 is located in sense strand exon IX, and antisense strand exon V. All exons containing the rs6265 variant are highlighted in blue for each transcript.

Sample ID	Age	Gender	Diagnosis	*BDNF* rs6265 Genotype
PD687	86	Male	Parkinson's Disease	C/C
PD793	75	Female	Parkinson's Disease	C/C
PD898	84	Male	Parkinson's Disease	C/C
PD115	63	Male	Dementia with Lewy bodies	C/T
PD120	83	Female	Dementia with Lewy bodies	C/T
PD332	77	Male	Dementia with Lewy bodies	C/T
PDC007	104	Female	MCI	C/C
PDC111	88	Female	MCI	C/T
PD715	86	Male	MCI	C/C
C045	77	Male	MCI	C/C
PD065	84	Female	MCI	C/T
PD092	79	Male	Control	C/T
PDC110	94	Female	Control	C/T
PDC112	87	Male	Control	T/T

Table 5.1 *BDNF* rs6265 genotypes, gender, age, and diagnosis across samples. Individual age at sampling, gender, and diagnosis are given in addition to genotype status. MCI denotes mild cognitive impairment.

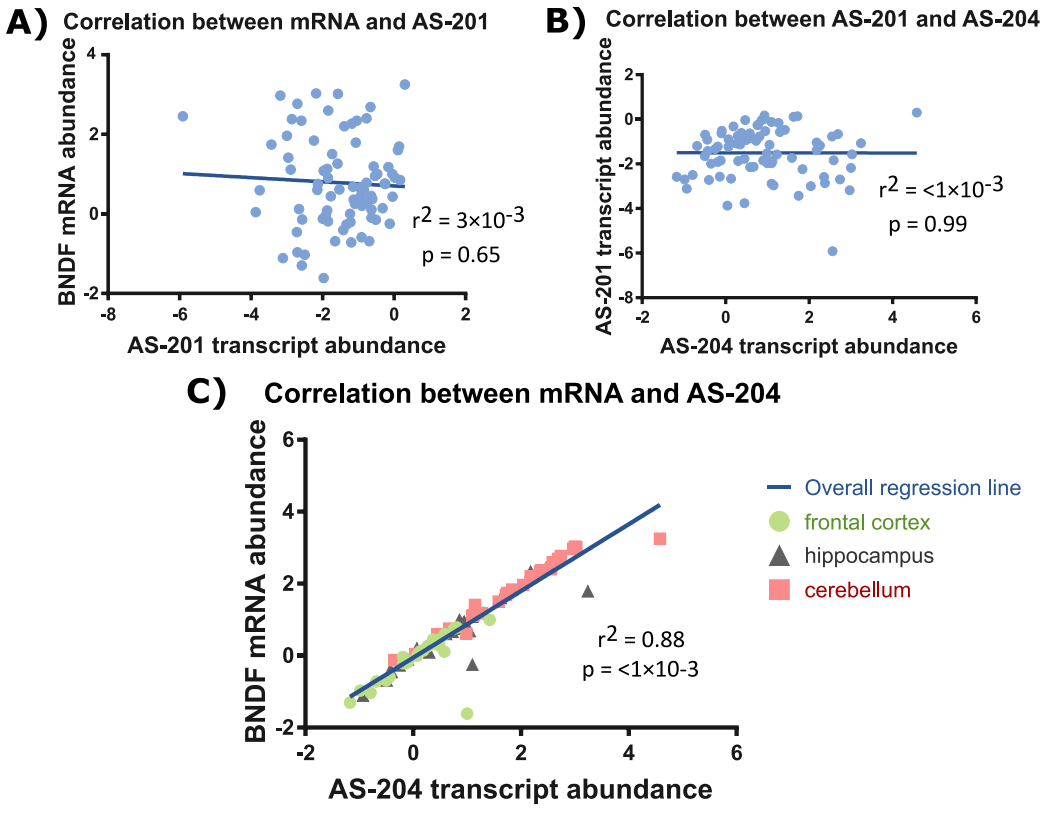


Figure 5.2 Correlation analysis of sense/mRNA, AS-201 and AS-204 transcript abundance. A) Spearman's correlation between sense/mRNA and AS-201 transcripts. B) Spearman's correlation between AS-201 and AS-204 transcripts. C) Spearman's correlation between sense/mRNA and AS-204 transcripts individual brain region tissues. Regression lines are displayed as blue lines for all graphs where r^2 is the correlation coefficient, significance is denoted by p.

5.2.2 Comparison of BDNF transcript abundance across brain regions

BDNF transcript abundance was analysed for differential expression across the cerebellum, hippocampus, and frontal cortex regions of the brain. Kruskal-Wallis analyses of the Ln(x+10) transformed sense, AS-201, and AS-204 transcript abundance across different brain tissues indicated significant differences in transcript expression. The results are displayed for sense, AS-201 and AS-204 transcripts in Figure 5.3, Figure 5.4, and Figure 5.5, respectively. Significant differences were identified between the cerebellum, frontal cortex and hippocampus for the sense (Kruskal-Wallis $X^2(2) = 31.96$, $p < 1\times10^{-3}$); the AS-201 transcript ($X^2(2) = 14.88$, $p = 1\times10^{-3}$); and the AS-204 ($X^2(2) = 28.24$, $p < 1\times10^{-3}$) transcripts. Dunn's post hoc analysis indicated all but one pairwise comparison as significant for all the transcripts. The exception was the pairwise comparison of hippocampus and frontal cortex in the antisense AS-201 transcript analysis ($p = 0.45$).

The highest observed transcript abundance for all three transcripts was in the cerebellum. The distribution of transcript abundance in cerebellum for both sense and AS-204 transcripts were higher than either hippocampus or frontal cortex. However AS-201 transcript, while showing the highest transcript abundance in cerebellum also reported the lowest transcript abundance in cerebellum i.e. the broadest range of expression was in cerebellum. Moreover, AS-201 expression in frontal cortex exhibited the highest range of AS-201 abundance. These results indicated that transcript expression may be regulated in a brain region specific manner. This also suggested that not only brain region but phenotype may influence the expression of these transcripts, as transcripts may be both the highest and lowest expressed in a single tissue, depending on the sample.

A)

K-W X^2= 31.96	DF = 2	p < 1×10^{-3}
Brain region	Median	IQR
Cerebellum	1.79	1.72
Hippocampus	0.72	1.25
Frontal Cortex	0.12	1.05

B)

Tissues being compared		Sig.
HIP	FC	1.7×10^{-2}
CER	FC	$<1\times10^{-3}$
CER	HIP	$<1\times10^{-3}$

C) **BDNF mRNA analysis across the brain regions**

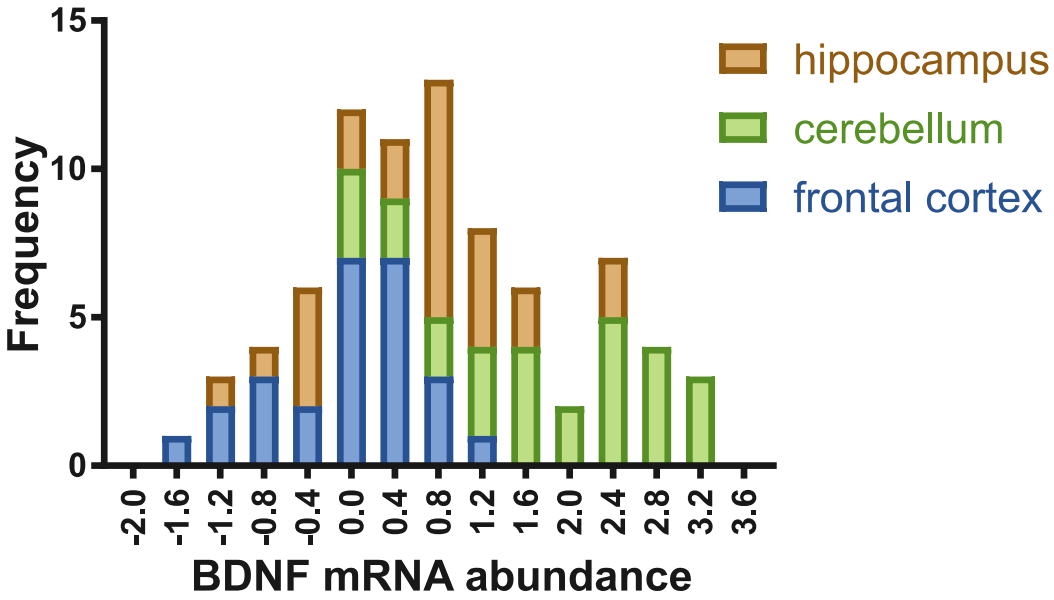


Figure 5.3 Differences between sense transcript expression in the cerebellum, frontal cortex, and hippocampus. A) Kruskal-Wallis analysis comparing sense transcript abundance across brain tissues, results, where Chi-squared statistic (X^2), Degrees of Freedom (DF) and Significance ($p < 1\times10^{-3}$) are shown. Additionally, median and IQR are given for all brain regions. B) Dunn's post hoc pairwise multiple comparison analyses for each pair of tissues (hippocampus (HIP), frontal cortex (FC) and cerebellum (CER)). C) Stacked histogram of the expression data for all tissues. The x-axis shows the relative transcript abundance to *TBP* abundance. Frequency shows how many samples were expressing that level of abundance.

A)

K-W X^2= 14.88	DF = 2	p = 1×10^{-3}
Brain region	Median	IQR
Cerebellum	-2.16	1.40
Hippocampus	-0.99	1.00
Frontal Cortex	-1.01	1.30

B)

Tissues being compared		Sig.
HIP	FC	0.45
CER	FC	4×10^{-3}
CER	HIP	<1×10^{-3}

C)

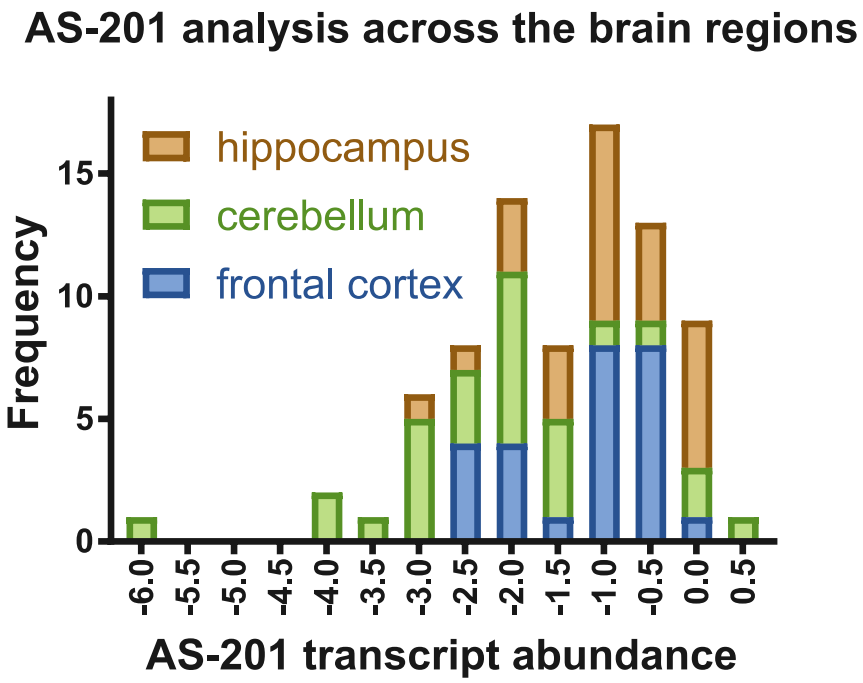


Figure 5.4 Differences between AS-201 transcript expression across the cerebellum, frontal cortex, and hippocampus. A) Kruskal-Wallis results, where Chi-squared statistic (X^2), Degrees of Freedom (DF) and Significance (p = 1×10^{-3}) are shown. Additionally, median and IQR are given for all brain regions. B) Dunn's post hoc pairwise multiple comparison analyses for each pair of tissues (hippocampus (HIP), frontal cortex (FC) and cerebellum (CER)). C) Stacked histogram of the expression data for all tissues. The x-axis shows the relative standardised transformed transcript abundance to *TBP* abundance. Frequency shows how many samples were expressing that level of abundance.

A)

K-W X^2= 28.24	DF = 2	p < 1×10^{-3}
Brain region	Median	IQR
Cerebellum	1.80	1.56
Hippocampus	0.88	1.11
Frontal Cortex	0.23	0.77

B)

Tissues being compared		Sig.
HIP	FC	1.7×10^{-2}
CER	FC	<1×10^{-3}
CER	HIP	<4×10^{-3}

C) **AS-204 analysis across the brain regions**

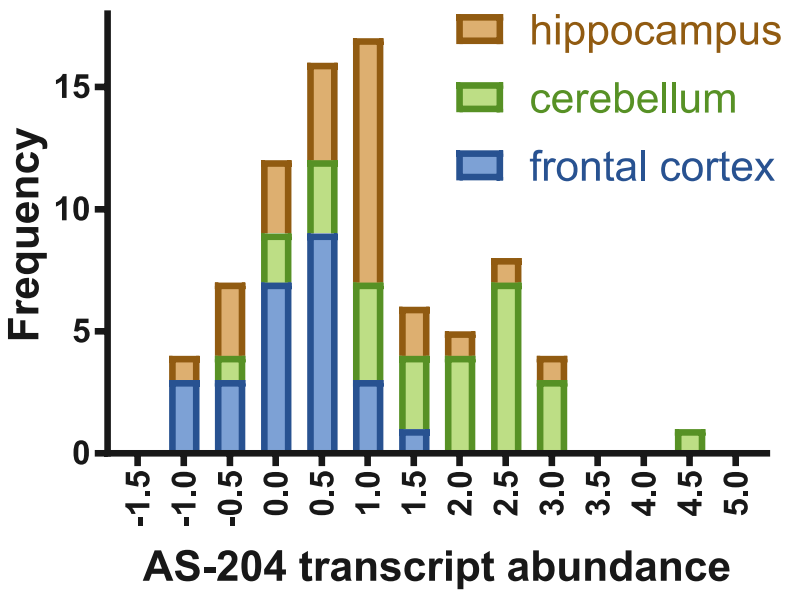


Figure 5.5 Differences between AS-204 transcript expression across the cerebellum, frontal cortex, and hippocampus. A) Kruskal-Wallis results, where Chi-squared statistic (X^2), Degrees of Freedom (DF) and Significance (p < 1×10^{-3}) are shown. Additionally, median and IQR are given for all brain regions. B) Dunn's post hoc pairwise multiple comparison analyses for each pair of tissues (hippocampus (HIP), frontal cortex (FC) and cerebellum (CER)). C) Stacked histogram of the expression data for all tissues. The x-axis shows the relative standardised transformed transcript abundance. Frequency shows how many samples were expressing that level of abundance.

5.2.3 Comparison of BDNF transcript abundance across disease phenotypes

Kruskal-Wallis analysis to compare differential gene expression across the disease states indicated both the sense ($X^2(3) = 1.83$, $p = 0.61$) and AS-20 4 ($X^2(3) = 4.24$, $p = 0.24$) transcripts did not show significant differences in transcript abundance between the disease phenotypes. However, the antisense transcript AS-201 did show significant differential transcript expression between disease states ($X^2(3) = 13.65$, $p = 3\times10^{-3}$).

Dunn's post hoc analysis of the AS-201 transcript expression indicated that there were significant differences between MCI and PD ($p = 6\times10^{-3}$), MCI and DLB ($p = 2\times10^{-3}$), and control and DLB ($p = 2.4\times10^{-2}$) (Figure 5.6). These comparisons revealed that healthy control and MCI individuals had significantly less expression of the AS-201 transcript compared to DLB and PD samples. This could suggest that in PD and DLB, the *BDNF* AS-201 antisense transcript abundance is dysregulated and the higher abundance of the AS-201 antisense transcript could interact and regulate both the amount of sense and AS-204 antisense transcript.

A)

K-W X^2= 13.65	DF = 3	p = 3×10^{-3}
Disease	Median	IQR
Control	-1.38	1.79
MCI	-1.92	1.43
PD	-0.88	1.13
DLB	-0.70	1.53

B)

Disease states being compared		P
MCI	Control	0.51
MCI	PD	6×10^{-3}
MCI	DLB	2×10^{-3}
Control	PD	5.7×10^{-2}
Control	DLB	2.4×10^{-2}
PD	DLB	0.72

C)

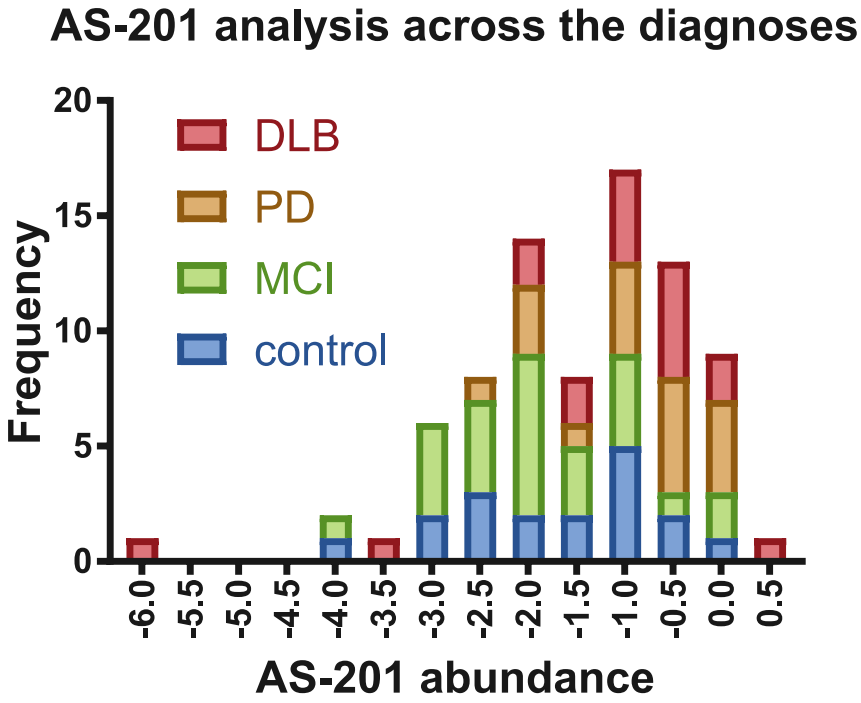


Figure 5.6 Differences in expression of AS-201 transcripts across disease states. Disease states are control, mild cognitive impairment (MCI), Parkinson's disease (PD) and dementia with Lewy bodies Disease (DLB). A) Kruskal-Wallis results, where Chi-squared statistic (X^2), Degrees of Freedom (DF) and Significance (p = 0.24) are shown. Additionally, median and IQR are given for all disease states. B) Dunn's post-hoc pairwise multiple comparisons for each disease state. The significance of each comparison is shown (P). C) Stacked histogram of the expression data for all diseases. AS-201 abundance shows the relative transcript abundance. Frequency shows how many samples were expressing that level of abundance.

5.2.4 Transcript expression between disease states within each brain region

A Kruskal-Wallis analysis was conducted to determine if there was differential expression between disease states when looking at each brain region independently (Table 5.2). Kruskal-Wallis analyses indicated that there were no significant differences observed between disease states within any brain region for the sense strand transcript abundance. Cerebellum (K-W $X^2(3) = 3.88$, $p = 0.27$), hippocampus ($X^2(3) = 0.63$, $p = 0.89$), and frontal cortex ($X^2(3) = 3.13$, $p = 0.37$) failed to show significant differences. However, the AS-201 transcript showed significant differences between disease states in the frontal cortex ($X^2(3) = 9.98$, $p = 1.9\times10^{-2}$) but not for the hippocampus ($X^2(3) = 0.63$, $p = 0.89$) or cerebellum ($X^2(3) = 7.77$, $p = 5.1\times10^{-2}$). AS-204 transcript expression analyses within tissues were similar to those for AS-201; where frontal cortex ($X^2(3) = 8.22$, $p = 4.2\times10^{-2}$) exhibited a significant difference between disease states, however cerebellum ($X^2(3) = 4.06$, $p = 0.26$) and hippocampus ($X^2(3) = 0.43$, $p = 0.94$) did not.

To determine which disease states were significantly different within frontal cortex for the AS-201 and AS-204 transcript, Dunn's post-hoc analyses were conducted (Figures 5.6, and 5.7). A significant difference between disease states for both anti-sense transcripts abundance (AS-201 and AS-204) within the frontal cortex was observed. The AS-201 transcripts showed significant differences between MCI and PD ($p = 1.1\times10^{-2}$), as well as between MCI and DLB ($p = 6\times10^{-3}$) (Figure 5.7 Panel B). Transcript AS-204 also showed significant differences between disease states within the frontal cortex when analysed by Dunn's post hoc test (Figure 5.8 Panel B). Control disease state AS-204 transcript expression within the frontal cortex was significantly different from that of PD ($p = 2.6\times10^{-2}$) and DLB ($p = 3.5\times10^{-2}$).

Transcript	Tissue	X^2	DF	P
Sense Strand	Cerebellum	3.88	3	0.27
	Hippocampus	0.63	3	0.89
	Frontal Cortex	3.13	3	0.37
AS-201	Cerebellum	7.77	3	5.1×10^{-2}
	Hippocampus	0.63	3	0.89
	Frontal Cortex	9.98	3	1.9×10^{-2}
AS-204	Cerebellum	4.06	3	0.26
	Hippocampus	0.43	3	0.94
	Frontal Cortex	8.22	3	4.2×10^{-2}

Table 5.2 Kruskal-Wallis tests across different brain regions to determine if there are significant differences between transcript expression in disease states. The transcripts and tissues being analysed, Kruskal-Wallis statistic (X^2), Degrees of Freedom (DF) and Significance of analysis (P) are indicated.

A)

K-W X^2= 9.98	DF = 3	p = 1.9×10^{-2}
Disease	Median	IQR
Control	-1.10	1.39
MCI	-2.00	1.23
PD	-0.78	0.47
DLB	-0.47	1.21

B)

Disease states being compared		P
MCI	Control	0.25
MCI	PD	0.01
MCI	DLB	6×10^{-3}
Control	PD	0.20
Control	DLB	0.14
PD	DLB	0.85

C)

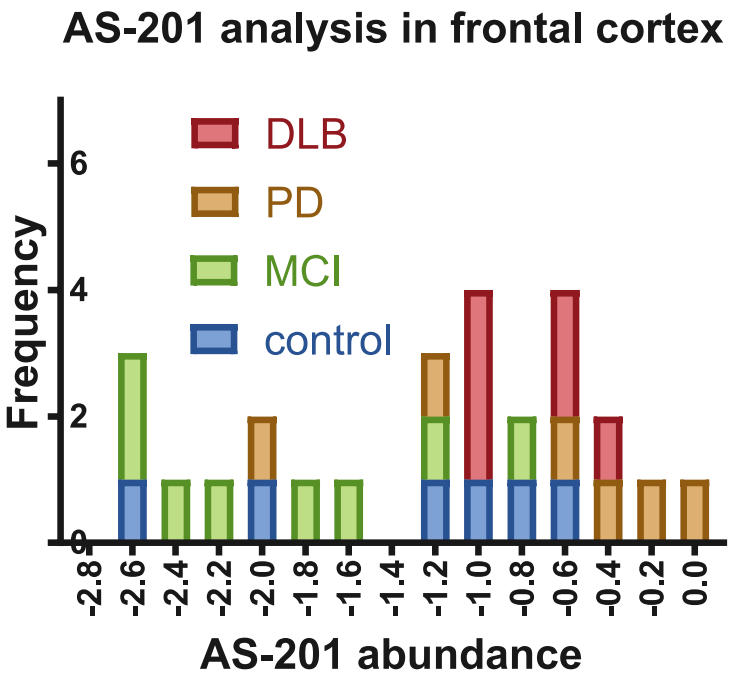


Figure 5.7 AS-201 transcript abundance across disease states within the frontal cortex. Disease states are control, mild cognitive impairment (MCI), Parkinson's disease (PD) and dementia with Lewy bodies (DLB). A) Kruskal-Wallis results within the frontal cortex, where Chi-squared statistic (X^2), degrees of freedom (DF) and p value are shown. Additionally, median and IQR are given for all disease states. B) Dunn's post-hoc pairwise multiple comparisons for each disease state analysed within frontal cortex. The significance of each comparison is shown (P). C) Stacked histogram of the expression data for all diseases in frontal cortex. X-axis shows the relative transcript abundance. Frequency shows how many samples were expressing that level of abundance.

A)

K-W X^2= 8.22	DF = 3	p = 4.2×10^{-2}
Disease	Median	IQR
Control	-0.32	1.10
MCI	-0.16	1.19
PD	0.49	0.62
DLB	0.43	0.66

B)

Disease states being compared		P
MCI	Control	0.65
MCI	PD	0.05
MCI	DLB	0.07
Control	PD	0.03
Control	DLB	0.04
PD	DLB	0.91

C)

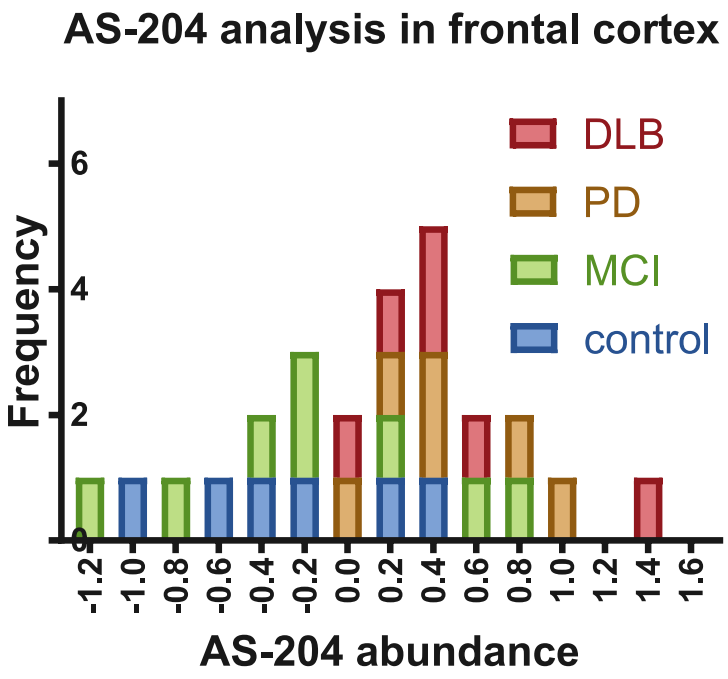


Figure 5.8 AS-204 transcript abundance across disease states within the frontal cortex. Disease states are control, mild cognitive impairment (MCI), Parkinson's disease (PD) and dementia with Lewy bodies (DLB). A) Kruskal-Wallis results within the frontal cortex, where Chi-squared statistic (X^2), degrees of freedom (DF) and p value are shown. Additionally, median and IQR are given for all disease states. B) Dunn's post-hoc pairwise multiple comparisons for each disease state analysed within frontal cortex. The significance of each comparison is shown (P). C) Stacked histogram of the expression data for all diseases in frontal cortex. X-axis shows the relative transcript abundance. Frequency shows how many samples were expressing that level of abundance.

5.2.5 Comparison of transcript abundance across brain regions within each disease state

Analyses were performed to investigate the expression of transcripts between different tissues within disease groupings (Table 5.3). Kruskal-Wallis analysis indicated the sense transcript showed the most significant differences in transcript abundance between disease states. Sense MCI ($X^2(2) = 15.39$, $p < 1\times10^{-3}$), PD ($X^2(2) = 6.35$, $p = 4.2\times10^{-2}$), and DLB ($X^2(2) = 8.78$, $p = 1.2\times10^{-2}$) all showed significant differences between brain regions. There were no significant differences in sense transcript abundance between any tissue type within control diagnoses samples ($X^2(2) = 4.67$, $p = 9.7\times10^{-2}$).

The AS-201 and AS-204 analyses each showed two significant differences. For example, AS-201 showed that the control ($X^2(2) = 10.25$, $p = 6\times10^{-3}$) and MCI ($X^2(2) = 6.44$, $p = 4\times10^{-2}$) states had a significant difference between brain region expression. Similarly, AS-204 transcript indicated MCI ($X^2(2) = 13.57$, $p = 1\times10^{-3}$), and PD ($X^2(2) = 7.45$, $p = 2.4\times10^{-2}$) had significant differences in transcript abundance across brain regions (Table 5.3). Dunn's post hoc analyses were used to determine differences between the brain regions' transcript expression within the disease states with significant differences.

Transcript	Disease	X^2	DF	P
Sense Strand	Control	4.67	2	0.097
	MCI	15.39	2	<0.001
	PD	6.35	2	0.042
	DLB	8.78	2	0.012
AS-201	Control	10.25	2	0.006
	MCI	6.44	2	0.04
	PD	0.36	2	0.834
	DLB	1.77	2	0.414
AS-204	Control	4.99	2	0.082
	MCI	13.57	2	0.001
	PD	7.45	2	0.024
	DLB	5.70	2	0.058

Table 5.3 Kruskal-Wallis tests across disease groups to determine significant differences in transcript expression between tissue groups. The disease states being tested are control, mild cognitive impairment (MCI), Parkinson's disease (PD), and dementia with Lewy bodies (DLB). The transcripts and disease states being analysed, Kruskal-Wallis statistic (X^2), degrees of freedom (DF) and p value (P) are indicated.

Comparison of AS-201 transcript expression in unaffected-control samples across brain regions

Dunn's post hoc tests to investigate differences in each disease state indicated that the antisense AS-201 transcript showed increased expression across all brain regions in all disease states compared to control controls, ($X^2(2)= 10.25$, $p = 6\times10^{-3}$) (Figure 5.9). The strongest decrease in relative expression occurred in the cerebellum (median of -2.78) as compared to the frontal cortex (-1.06) and hippocampus (-1.11), which is why cerebellar expression of AS-201 was significantly lower than both hippocampus ($p = 4\times10^{-3}$) and frontal cortex ($p = 9\times10^{-3}$).

Comparison of transcript expression in MCI samples across brain regions

The sense ($X^2(2) = 15.39$, $p < 1\times10^{-3}$), AS-201 ($X^2(2) = 6.44$, $p = 0.04$), and AS-204 ($X^2(2) = 13.57$, $p = 1\times10^{-3}$) transcripts all showed significant differences in abundance between brain region in MCI samples (Table 5.3). The results of Dunn's post hoc analysis of the sense, AS-201, and AS-204 transcript abundance within cerebellum, hippocampus and frontal cortex tissue from individuals with MCI are presented in Figures 5.9, 5.10, and 5.11, respectively.

As with previous analyses, sense and AS-204 transcripts showed similar distribution and expression levels to each other within the MCI samples (Figures 5.10 and 5.12) and were found to be highly expressed in cerebellum (median of 2.26 and 2.27, respectively). The sense transcript expression was significantly higher in cerebellum compared to hippocampus ($p = 2.1\times10^{-2}$) and frontal cortex ($p < 1\times10^{-3}$) tissues of MCI samples (Figure 5.10 panel B). However, the AS-204 transcript expression while significantly higher in cerebellum compared to frontal cortex ($p < 1\times10^{-3}$), was not significantly higher than hippocampus in MCI samples, as seen with the sense transcript (Figure 5.12 panel B).

The antisense AS-201 transcript expression was comparatively lower than that of the sense or AS-204 transcripts in all brain regions within MCI individuals (Figure 5.11 Panel C). The cerebellum, hippocampus and frontal cortex all showed negative relative median expression of AS-201 (-2.42, -1.01, and -2.00, respectively) (Figure

5.11 Panel A). Moreover, post hoc analysis revealed MCI AS-201 expression was significantly lower in cerebellum compared to hippocampus ($p = 1.1\times10^{-2}$), but not to frontal cortex (Figure 5.11 Panel B). In summary, within the MCI samples, the sense and AS-204 transcripts showed high abundance in the cerebellum compared to the frontal cortex whereas the AS-201 transcript showed lower expression in the cerebellum.

A)

K-W X^2= 10.25	DF = 2	p = 6×10^{-3}
Tissue	Median	IQR
Cerebellum	-2.78	0.74
Hippocampus	-1.11	0.67
Frontal Cortex	-1.06	1.39

B)

Tissue states being compared		P
HIP	FC	0.75
CER	FC	9×10^{-3}
CER	HIP	4×10^{-3}

C)

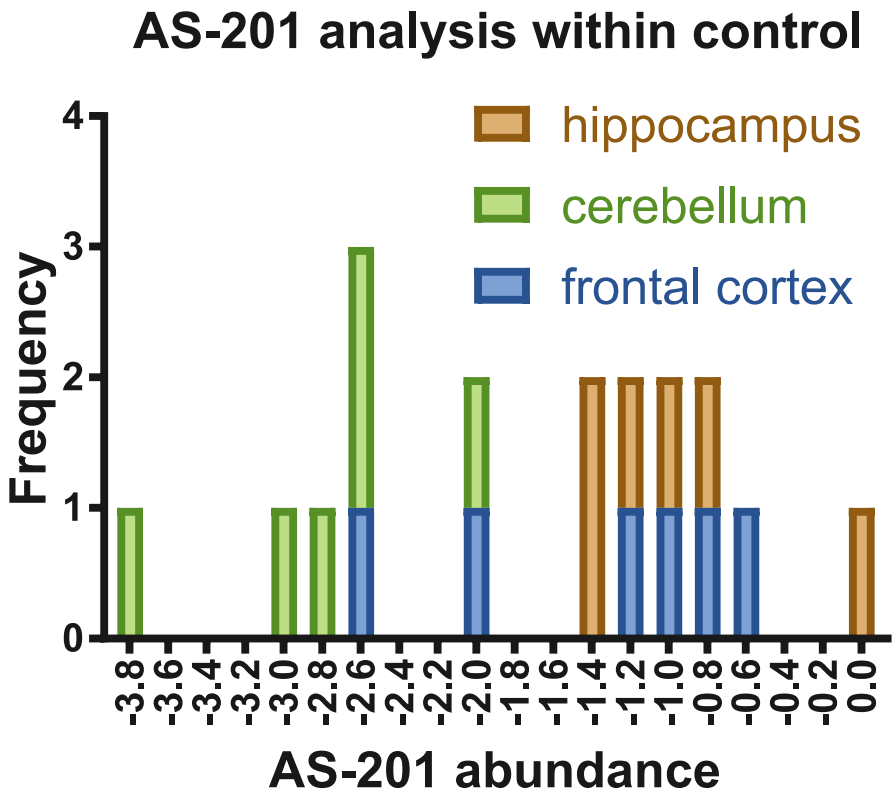


Figure 5.9 Expression of AS-201 transcript in tissue types of control samples. Tissues are cerebellum, hippocampus, and frontal cortex. A) Kruskal-Wallis summary statistic (X^2), degrees of freedom (DF), and p value. Each tissue types median and Inter-quartile range (IQR) are also shown. B) Significant findings of the Dunn's post hoc test for multiple comparisons between the tissue types where cerebellum (CER), hippocampus (HIP), and frontal cortex (FC) are shown. C) Frequency distribution in AS-201 transcript abundance across tissue types. X-axis shows the relative transcript abundance. Frequency shows how many samples were expressing that level of abundance.

A)

K-W X^2= 15.39	DF = 2	p < 1×10^{-3}
Tissue	Median	IQR
Cerebellum	2.26	1.33
Hippocampus	0.83	0.82
Frontal Cortex	-0.20	1.42

B)

Tissue states being compared		P
HIP	FC	0.14
CER	FC	$<1\times10^{-3}$
CER	HIP	2.1×10^{-2}

C)

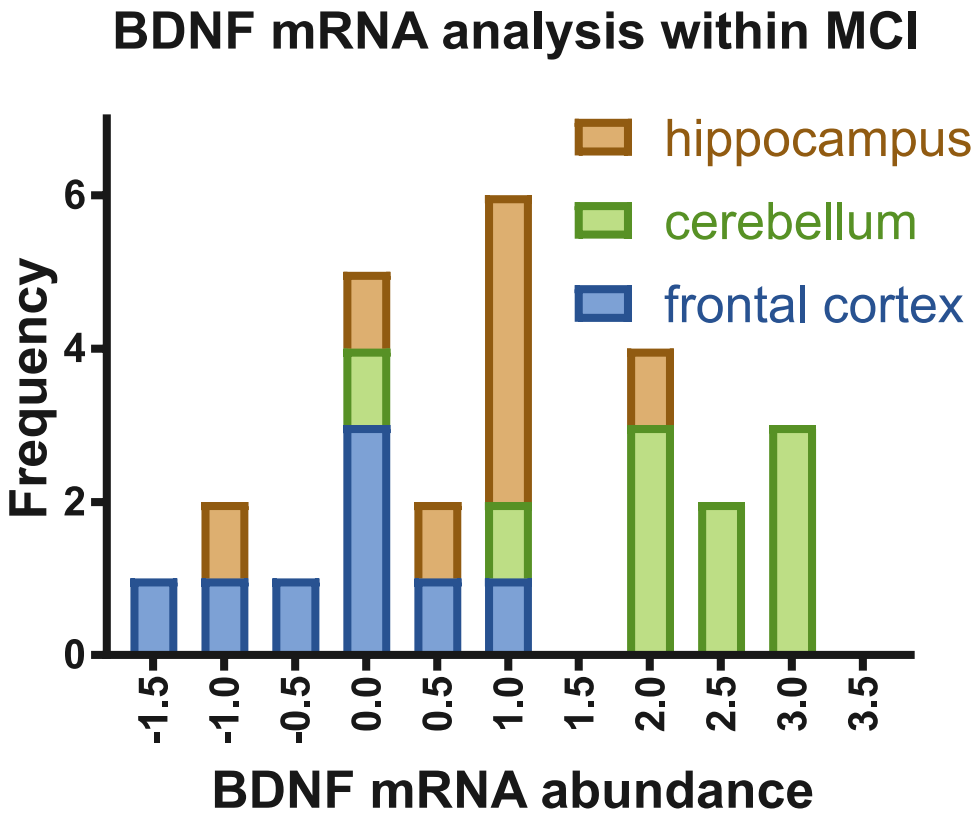


Figure 5.10 Expression of sense transcript in tissue types of Mild Cognitive Impairment samples. Tissues are cerebellum, hippocampus, and frontal cortex. A) Kruskal-Wallis summary statistic (X^2), degrees of freedom (DF), and p value. Each tissue types median and Inter-quartile range (IQR) are also shown. B) Significant findings of the Dunn's post hoc test for multiple comparisons between the tissue types where cerebellum (CER), hippocampus (HIP), and frontal cortex (FC) are shown. C) Frequency distribution in sense transcript abundance across tissue types. X-axis shows the relative transcript abundance. Frequency shows how many samples were expressing that level of abundance.

A)

K-W X^2= 6.44	DF = 2	p = 0.04
Tissue	Median	IQR
Cerebellum	-2.42	1.27
Hippocampus	-1.01	1.63
Frontal Cortex	-2.00	1.24

B)

Tissue states being compared		P
HIP	FC	0.21
CER	FC	0.22
CER	HIP	0.01

C)

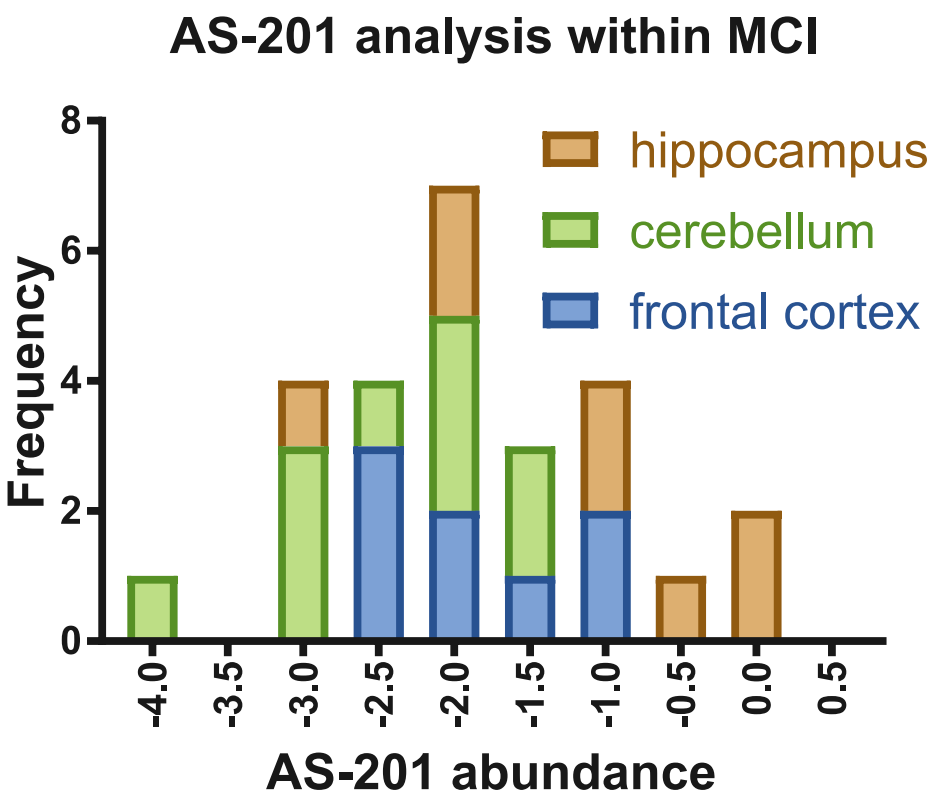


Figure 5.11 Expression of AS-201 transcript in tissue types of Mild Cognitive Impairment samples. Tissues are cerebellum, hippocampus, and frontal cortex. A) Kruskal-Wallis summary statistic (X^2), degrees of freedom (DF), and p value. Each tissue types median and Inter-quartile range (IQR) are also shown. B) Significant findings of the Dunn's post hoc test for multiple comparisons between the tissue types where cerebellum (CER), hippocampus (HIP), and frontal cortex (FC) are shown. C) Frequency distribution in AS-201 transcript abundance across tissue types. X-axis shows the relative transcript abundance. Frequency shows how many samples were expressing that level of abundance.

A)

K-W X^2= 23.57	DF = 2	p = 1×10^{-3}
Tissue	Median	IQR
Cerebellum	2.27	1.33
Hippocampus	0.86	0.65
Frontal Cortex	-0.16	1.19

B)

Tissue states being compared		P
HIP	FC	7.2×10^{-2}
CER	FC	$<1\times10^{-3}$
CER	HIP	7.4×10^{-2}

C)

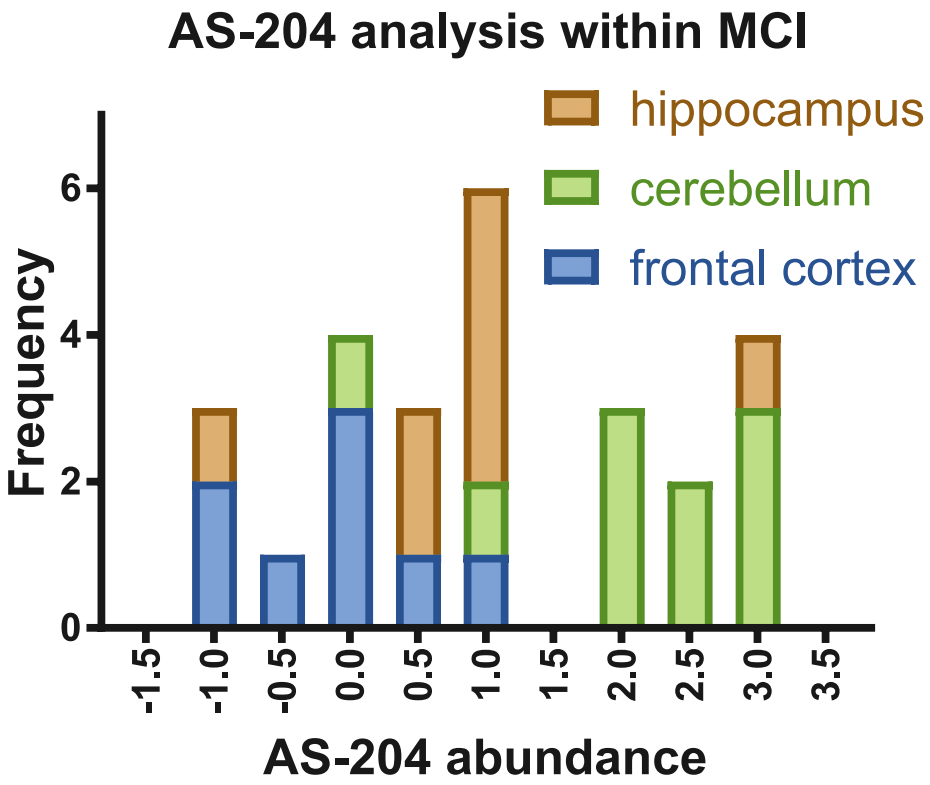


Figure 5.12 Expression of AS-204 transcript in tissue types of Mild Cognitive Impairment samples. Tissues are cerebellum, hippocampus, and frontal cortex. A) Kruskal-Wallis summary statistic (X^2), degrees of freedom (DF), and p value. Each tissue types median and Inter-quartile range (IQR) are also shown. B) Significant findings of the Dunn's post hoc test for multiple comparisons between the tissue types where cerebellum (CER), hippocampus (HIP), and frontal cortex (FC) are shown. C) Frequency distribution in AS-204 transcript abundance across tissue types. X-axis shows the relative transcript abundance. Frequency shows how many samples were expressing that level of abundance.

Comparison of transcript expression within PD samples across brain regions.

Kruskal-Wallis analysis revealed both the sense ($X^2(2) = 6.35$, $p = 4.2\times10^{-2}$), and AS-204 ($X^2(2) = 7.45$, $p = 2.4\times10^{-2}$) transcripts showed significant differences across brain regions within the Parkinson's disease (PD) diagnosis samples. However, the AS-201 transcript did not ($X^2(2) = 0.36$, $p < 8.3\times10^{-1}$) (Table 5.3). The PD sample expression profiles indicated similarities in distribution range between sense and AS-204 transcripts.

However, the sense transcript expression was relatively low in frontal cortex (Figures 5.13 and 5.14 Panel C). Both sense and AS-204 transcripts displayed a higher level of relative expression in cerebellum (median of 1.34 and 1.43, respectively) followed by hippocampus (0.52 and 0.64, respectively) and finally frontal cortex (0.26 and 0.42, respectively) (Figures 5.13 and 5.14 Panel A). Post hoc analyses indicated that like the control and MCI samples, both sense and AS-204 transcripts abundance was significantly higher in cerebellum tissue compared to hippocampus (sense $p = 2.7\times10^{-2}$ and AS-204 $p = 0.03$) and frontal cortex (sense $p = 5\times10^{-3}$ and AS-204 $p = 0.01$) in samples with PD (Figure 5.13 and 5.14, panel B).

A)

K-W X^2= 6.35	DF = 2	p = 4.2×10^{-2}
Tissue	Median	IQR
Cerebellum	1.34	0.83
Hippocampus	0.52	1.13
Frontal Cortex	0.26	0.92

B)

Tissue states being compared		P
HIP	FC	0.55
CER	FC	5×10^{-3}
CER	HIP	2.7×10^{-2}

C)

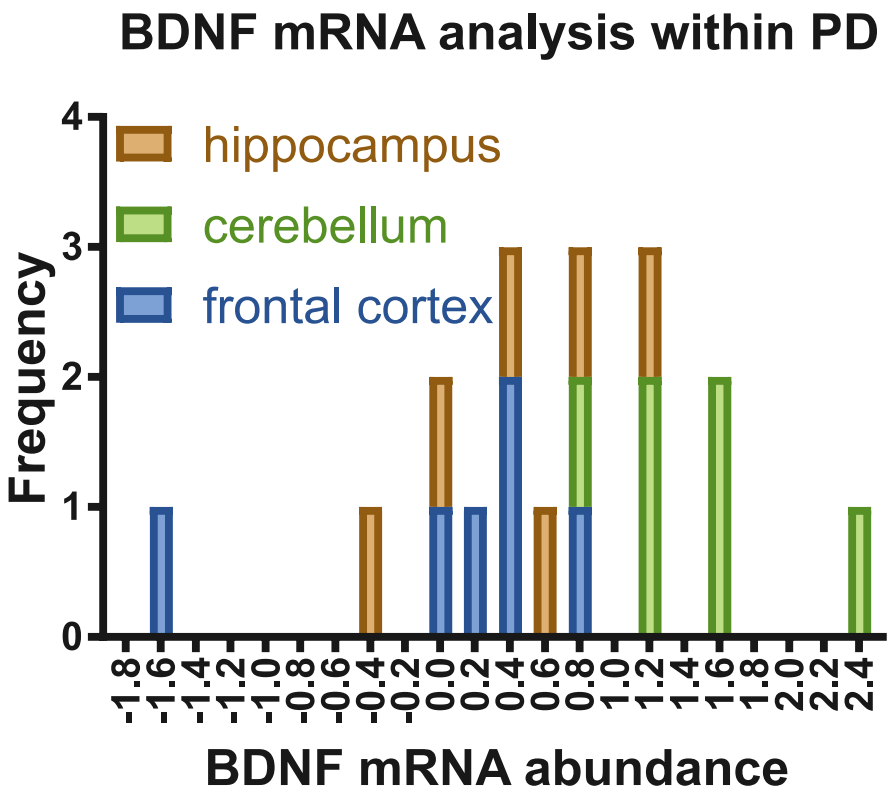


Figure 5.13 Expression of sense transcript in tissue types of Parkinson's disease samples. Tissues are cerebellum, hippocampus, and frontal cortex. A) Kruskal-Wallis summary statistic (X^2), degrees of freedom (DF), and p value. Each tissue types median and Inter-quartile range (IQR) are also shown. B) Significant findings of the Dunn's post hoc test for multiple comparisons between the tissue types where cerebellum (CER), hippocampus (HIP), and frontal cortex (FC) are shown. C) Frequency distribution in sense transcript abundance across tissue types. X-axis shows the relative transcript abundance. Frequency shows how many samples were expressing that level of abundance.

A)

K-W X^2= 7.45	DF = 2	p = 2.4×10^{-2}
Tissue	Median	IQR
Cerebellum	1.43	0.92
Hippocampus	0.64	1.29
Frontal Cortex	0.42	0.66

B)

Tissue states being compared		P
HIP	FC	0.79
CER	FC	0.01
CER	HIP	0.03

C)

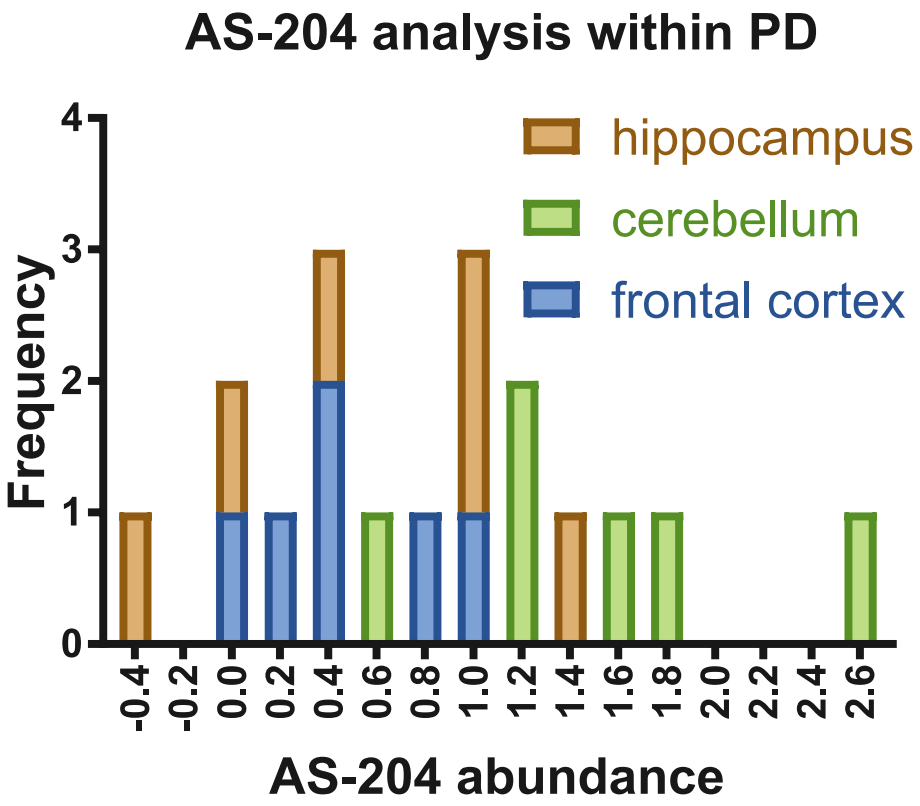


Figure 5.14 Expression of AS-204 transcript in tissue types of Parkinson's disease samples. Tissues are cerebellum, hippocampus, and frontal cortex. A) Kruskal-Wallis summary statistic (X^2), degrees of freedom (DF), and p value. Each tissue types median and Inter-quartile range (IQR) are also shown. B) Significant findings of the Dunn's post hoc test for multiple comparisons between the tissue types where cerebellum (CER), hippocampus (HIP), and frontal cortex (FC) are shown. C) Frequency distribution in AS-204 transcript abundance across tissue types. X-axis shows the relative transcript abundance. Frequency shows how many samples were expressing that level of abundance.

Comparison of transcripts within DLB samples across brain regions

In Kruskal-Wallis analysis the sense ($X^2(2) = 8.78$, $p = 1.2\times10^{-2}$) transcript showed a significant difference within the tissue types of dementia with Lewy bodies (DLB) diagnosis samples. However, the AS-201 ($X^2(2) = 1.77$, $p < 4.1\times10^{-1}$) and AS-204 ($X^2(2) = 5.69$, $p < 5.8\times10^{-2}$) antisense transcripts did not show any significant differences between tissues within the DLB samples. The sense transcript abundance followed patterns seen previously with highest expression in cerebellum (median of 2.10), followed by hippocampus (0.75) and finally frontal cortex (0.27) (Figure 5.15 Panel A). Expression of sense transcript was significantly higher in cerebellum compared to hippocampus ($p = 0.07$) and frontal cortex ($p = 0.02$) tissues in DLB diagnoses samples.

In summation, the *BDNF* transcripts investigated in this study showed unique relationships with brain regions and across disease states. Sense and AS-204 transcripts appear to correlate strongly, while AS-201 does not correlate significantly with either transcript. Expression for sense and AS-204 transcripts is typically highest in the cerebellum, followed by hippocampus and frontal cortex. The AS-201 transcript showed significant changes in expression dependent on disease state as evident in MCI and unaffected control samples. Both antisense transcripts also showed changes in expression in the frontal cortex, where expression of AS-201 was significantly lower in the frontal cortex of MCI samples compared to PD and DLB disease states. Further, AS-204 was significantly reduced in expression in the frontal cortex of MCI and control samples, when compared to PD samples. This illustrates the more severe disease phenotypes expressing higher levels of antisense transcripts.

A)

$X^2 = 8.78$	DF = 2	$p = 1.2 \times 10^{-2}$
Tissue	Median	IQR
Cerebellum	2.10	2.26
Hippocampus	0.75	1.70
Frontal Cortex	0.27	0.55

B)

Tissue states being compared		P
HIP	FC	0.52
CER	FC	0.02
CER	HIP	0.07

C)

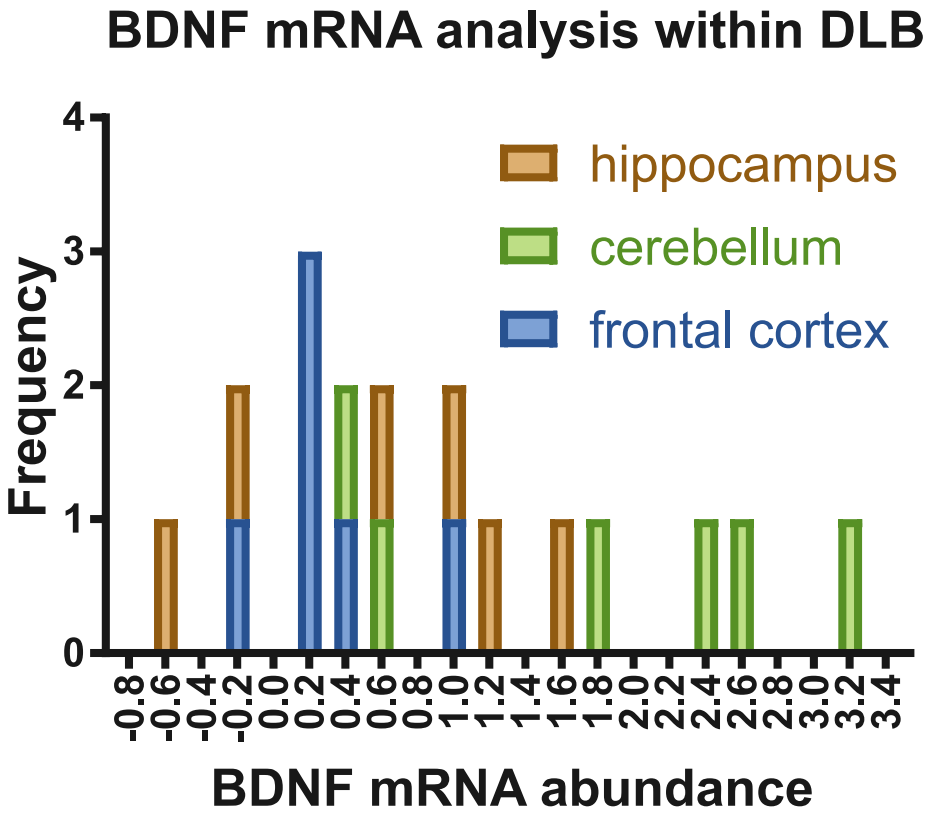


Figure 5.15 Expression of sense transcript in tissue types of dementia with Lewy bodies (DLB) samples. Tissues are cerebellum, hippocampus, and frontal cortex. A) Kruskal-Wallis summary statistic (X^2), degrees of freedom (DF), and p value. Each tissue types median and Inter-quartile range (IQR) are also shown. B) Significant findings of the Dunn's post hoc test for multiple comparisons between the tissue types where cerebellum (CER), hippocampus (HIP), and frontal cortex (FC) are shown. C) Frequency distribution in sense transcript abundance across tissue types. X-axis shows the relative transcript abundance. Frequency shows how many samples were expressing that level of abundance.

5.3 *BDNF* transcript abundance, and 5mC, 5hmC, 5fC/CaC methylation at antisense *BDNF* rs6265

The *BDNF* sense, antisense AS-201, and AS-204 transcript's expression was examined for correlations with 5mC, 5hmC, and 5fC/CaC methylation levels at the *BDNF* AS-rs6265 epiallele across tissue regions and disease phenotypes. Of the Imperial brain bank samples, only the T allele homozygous individual (PDC112) was excluded.

5.3.1 5mC and 5hmC methylation at AS-rs6265 and BDNF transcript expression

Thirteen individuals each with duplicate expression data for each of the three brain regions were examined. One sample (6 methylation data points) was eliminated as being an outlier (greater or lower than 1.5 × IQR) which gave a total of 72 expression data points which were matched to the 36 methylation data points. The data were then split into subgroups, i.e. into AS-*BDNF* rs6265 C/T heterozygotes (carriers) and C homozygotes (non-carriers), tissue region (cerebellum, hippocampus and frontal cortex), and finally disease states (control, mild cognitive impairment, Parkinson's disease, and dementia with Lewy bodies).

Correlation analysis between methylation and BDNF transcript expression profiles.
Table 5.4 and Figure 5.16 shows the Kendal-Tau correlations between 5mC/5hmC methylation at AS-rs6265 and *BDNF* transcript expression. Sense and AS-204 transcript expression were similar and showed similar methylation-expression profiles, negatively correlating with 5mC methylation and positively correlating with 5hmC methylation. However, no correlations were statistically significant for sense and AS-204 transcripts across both methylation types.

In contrast, the antisense AS-201 transcript abundance was found to be significantly positively correlated with 5mC methylation (K-T = 0.17, $p = 3.5\times10^{-2}$), and significantly negatively correlated with 5hmC methylation (K-T = -0.25, $p = 3\times10^{-3}$), i.e. low levels of 5hmC methylation were correlated with higher AS-201 transcript abundance.

Methylation	Transcript	N	Correlation coefficient	P
5mC	Sense	72	-0.14	8.7×10^{-2}
	AS-201	72	0.17	3.5×10^{-2}
	AS-204	72	-0.09	0.27
5hmC	Sense	72	0.08	0.34
	AS-201	72	-0.25	3×10^{-3}
	AS-204	72	0.09	0.26

Table 5.4 Kendal-Tau correlation analysis of transcript expression data and 5mC/5hmC methylation at *BDNF* AS-rs6265. Methylation state, transcript, sample number (N), correlation coefficient, and P value are indicated.

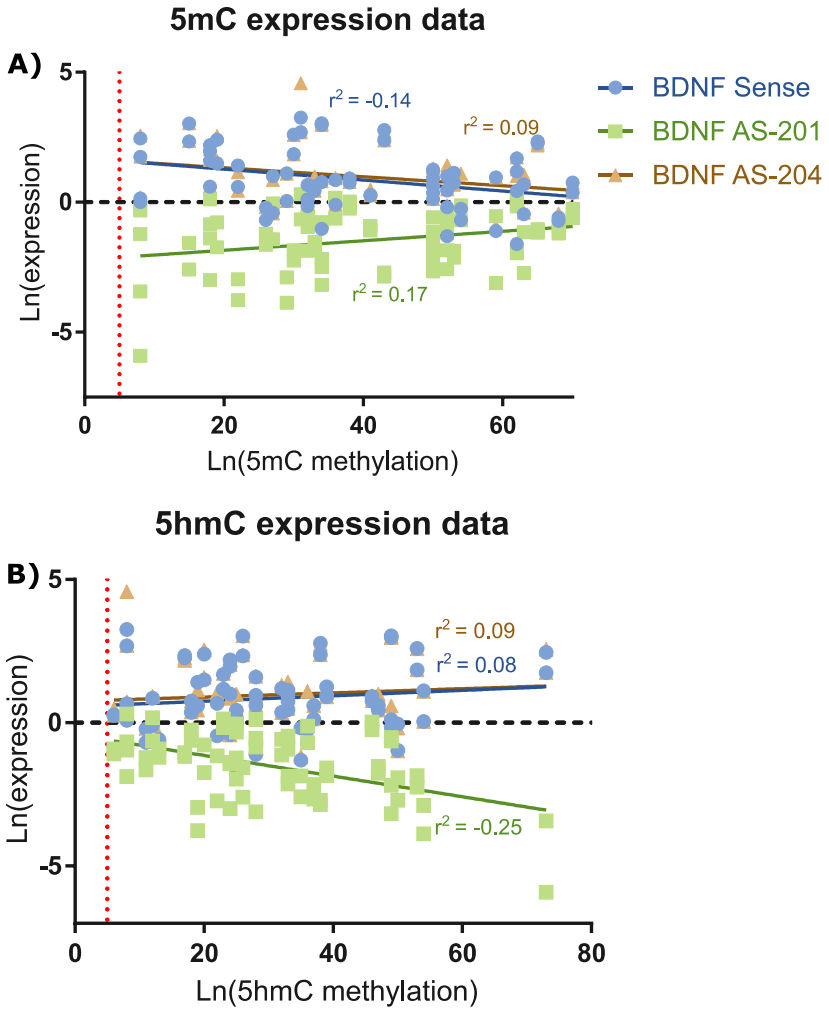


Figure 5.16 Relationship between 5mC and 5hmC methylation levels at the AS-rs6265 epiallele and sense, AS-201, and AS-204 transcript abundance. Sense, AS-201 and AS-204 transcripts are presented in blue, green and orange, respectively. Lines through data points represent linear regression. Ln transformed methylation data are shown on the x-axis, and Ln transformed transcript expression data are shown on the y-axis. Red dotted line indicates the 5% error threshold generated in the quantification of methylation by pyrosequencing. Black dotted line represents 0 relative expression. r^2 values represent correlations from Kendal-Tau analysis.

5.3.2 5(h)mC methylation, BDNF transcript abundance examined by genotype, brain region and disease phenotype, with age and gender as covariates

To further investigate the potential effect of AS-rs6265 CpG methylation on transcript expression across the brain regions and between disease phenotypes, data were grouped by brain region, by disease, as well as genotype (AS-rs6265 carriers; C/T heterozygotes, or non-carriers; C homozygotes). Correlations were examined by comparing methylation states (5mC or 5hmC), transcript expression (Sense, AS-201, AS-204) and either genotype, brain region or disease phenotype. The significant results of the Kendal-Tau correlations are presented in Table 5.5 and illustrated in Figure 5.17.

First, assessment of the disease phenotypes revealed that both sense and antisense AS-204 transcript abundance showed a significant negative correlation with 5mC methylation within MCI samples (sense, K-T = -0.31, $p = 4.1\times10^{-2}$; AS-204 K-T = -0.32, $p = 3.2\times10^{-2}$, respectively). These correlations were stronger and more significant than the negative correlations made between 5mC methylation and sense/AS-204 transcript abundance previously (Table 5.4). The sense transcript showed a significant negative correlation with 5mC methylation in DLB disease samples (K-T = -0.38, $p = 3.3\times10^{-2}$). However, unlike previous patterns seen between sense and AS-204 transcripts, the AS-204 transcript did not significantly negatively correlate with 5mC methylation in DLB samples (not shown; K-T = -0.30, $p = 9.3\times10^{-2}$).

When assessing differences between brain regions, AS-201 expression showed a significant and strong negative correlation with 5hmC methylation in the cerebellum (K-T = -0.43, $p = 2\times10^{-3}$) (Table 5.5). This was similar to the significant negative correlation observed between AS-201 transcript expression and 5hmC methylation previously (K-T = -0.25, $p = 3\times10^{-3}$) (Table 5.4, Figure 5.17 Panel B). This may indicate that the negative correlation is primarily driven by a decrease in AS-201 transcript expression in the cerebellum.

Likewise, when examining differences between the genotypes, the sense and AS-204 transcript abundance showed positive correlations with 5hmC methylation in non-carrier (C allele homozygotes) genotypes (K-T = 0.27, $p = 3.4\times10^{-2}$; and K-T = 0.29, $p = 2.1\times10^{-2}$, respectively) (Figure 5.17). Moreover AS-201 transcript expression negatively correlated with 5hmC methylation in carrier (heterozygous) genotypes (K-T = -0.33, $p = 3\times10^{-3}$). The negative correlation observed between 5hmC and AS-201 abundance previously (Table 5.4) is similar to the negative correlation between 5hmC and AS-201 expression in heterozygotes, as while the correlation is weaker (-0.25 compared to -0.33) it is just as significant.

In addition, transcript AS-201 abundance correlated negatively with 5hmC methylation in individuals with DLB diagnoses (K-T = -0.34, $p = 5.7\times10^{-2}$). This correlation only trended to significance however it demonstrates that 5hmC and AS-201 abundance are negatively correlated in disease states. These results collectively suggest a genotype, methylation, and transcript specific interaction between the transcripts abundance (Figure 5.17).

Additional analyses were performed using age and gender as covariates. Distributions of age and gender are presented in Table 5.6 which shows that while the samples were approximately equal in division of gender (38 samples from men, 34 samples from women) the ranges and average ages of men and women were very different with the age range of men lower than that of women (mean age of 79.6, youngest sample at 63, oldest at 86 years). Male samples included the youngest samples used in the study (63 years) while the youngest female samples used are over 10 years older (75 years). This demonstrates a clear separate distribution in ages between genders which may be affecting the correlations between methylation levels and expression.

Analyses of correlations between AS-rs62565 methylation and *BDNF* transcript abundance with age, gender, and age and gender as covariates was performed using the Spearman's non-parametric partial correlation. Most Spearman's partial correlation analyses showed that there was a stronger correlation between

expression and methylation levels and increased significance when age and gender were included as covariates. However, only seven correlations were significant with the inclusion of age and/or gender as covariates, and all significant correlations were between 5hmC methylation and transcript abundance.

The positive correlation between sense transcript abundance and 5hmC methylation became significant with the inclusion of gender (CC = 0.27, $p = 1.8\times10^{-2}$), as well as age and gender (CC = 0.23, $p = 4.8\times10^{-2}$), as covariates. This suggests that the driving covariate contributing to the change in significance is gender, as analysis with age as a covariate does not improve significance.

AS-201 transcript showed a significant negative correlation with 5hmC methylation (CC = -0.24, $p = 3\times10^{-3}$), which became stronger in correlation and more significant with the inclusion of age (CC = -0.35, $p = 2\times10^{-3}$), gender (CC = -0.29, $p = 1\times10^{-2}$), and age and gender (CC = -0.37, $p = 1\times10^{-3}$), as covariates. Both covariates appear to contribute to the significant negative correlation observed, however the age covariate both strengthens the correlation and shows the largest change in significance. The positive correlation between AS-204 transcript and 5hmC methylation became significant with gender (CC = 0.32, $p = 5\times10^{-3}$), and age and gender (CC = 0.27, $p = 2\times10^{-2}$) as covariates.

These results suggest that sense and AS-204 transcript abundance positively correlate with AS-rs6265 5hmC methylation and this interaction is affected by the gender of the individual, while AS-201 transcript abundance negatively correlates with 5hmC methylation and is affected by age as a covariate.

Methylation	Tis, Dis, Geno	Transcript	N	Correlation coefficient	P
5mC	MCI	Sense	24	-0.31	4.1×10^{-2}
5mC	MCI	AS-204	24	-0.32	3.2×10^{-2}
5mC	DLB	Sense	18	-0.38	3.3×10^{-2}
5hmC	CER	AS-201	26	-0.43	2×10^{-3}
5hmC	C/C	Sense	32	0.27	3.4×10^{-2}
5hmC	C/C	AS-204	32	0.29	2.1×10^{-2}
5hmC	C/T	AS-201	40	-0.33	3×10^{-3}

Table 5.5 Significant Kendal-Tau correlation analyses. Methylation type (Methylation), tissue groups, disease groups, and genotypes are presented (Tis, Dis, Geno). Transcript expression (Transcript), Sample number (N), Correlation coefficient, and Significance (P) are shown. C/C denotes homozygous C genotype and C/T denotes heterozygous genotype. MCI denotes mild cognitive impairment; DLB denotes dementia with Lewy bodies; CER denotes cerebellum.

Gender	N	Mean	Lower age range	Upper age range
Male	38	79.6	63	86
Female	34	88.2	75	104
Total	72			

Table 5.6 Comparison of age and gender distribution across Imperial Brain Bank data points in the sample population. The number of data points (N) for each gender as well as the mean age, and lowest and highest ages for each gender.

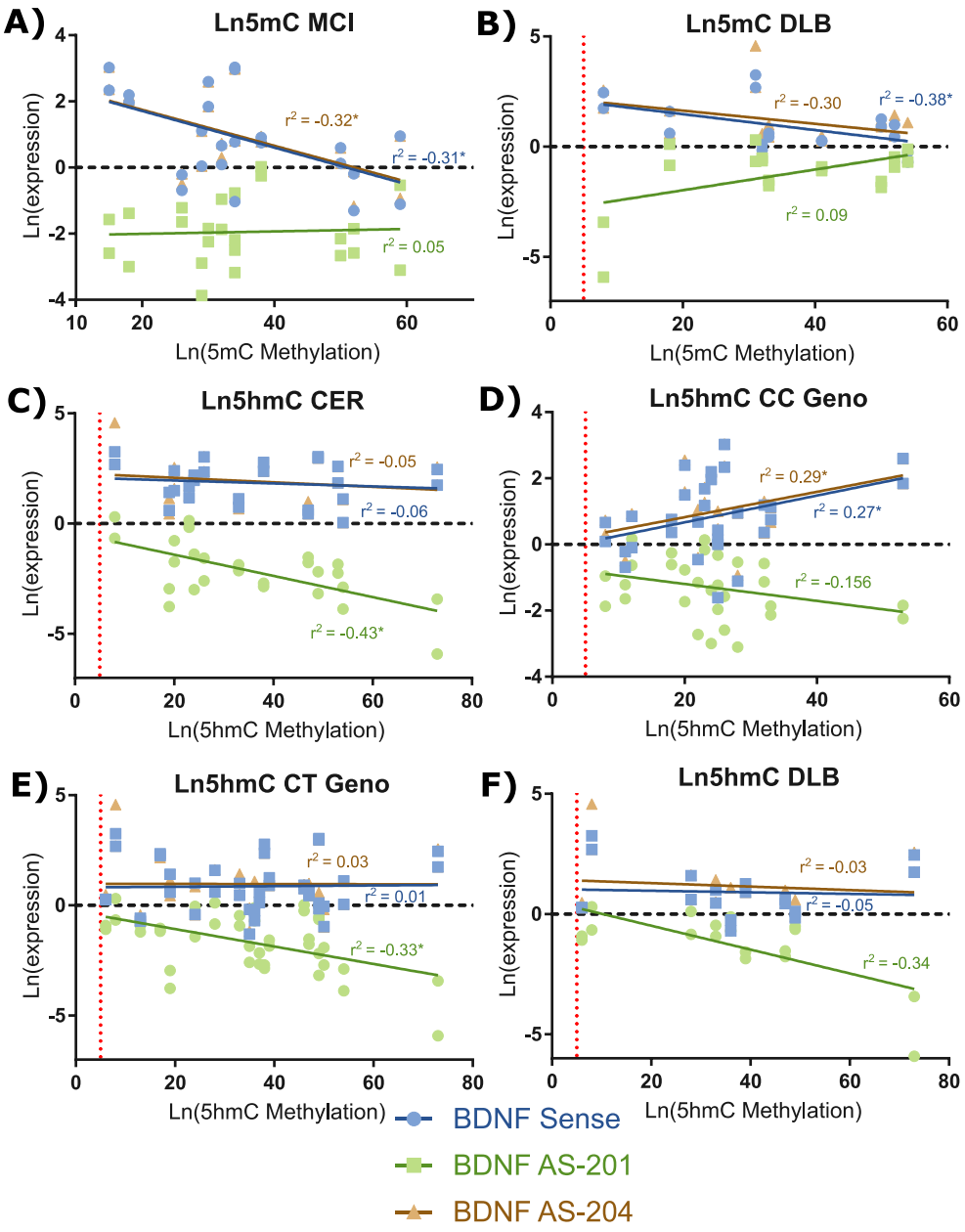


Figure 5.17 Significant correlations between transcript abundance, methylation, tissue types, disease states, and genotypes at AS-rs6265. Transcripts are sense, AS-201, and AS-204, methylation states are 5mC or 5hmC, genotypes are carriers (C/T) or non-carriers (C/C), tissue types are cerebellum, hippocampus, and frontal cortex, and disease states are control, mild cognitive impairment, Parkinson's disease, and dementia with Lewy bodies. A) The correlation between 5mC methylation and transcript abundance in MCI disease samples. B) The correlation between transcript abundance and 5mC in DLB diagnoses samples. C) The correlation between transcript abundance and 5hmC in cerebellum tissue. D) The correlation between transcript abundance and 5hmC in non-carrier genotypes. E) The correlation between transcript abundance and 5hmC in carrier genotypes. F) The correlation between transcript abundance and 5hmC in DLB diagnoses samples. In all panels, blue lines/points indicate sense transcripts, green indicates AS-201, and orange indicates AS-204. Black dotted lines refer to 0 relative expression and no change in expression. Red dotted lines indicate the 5% error threshold for methylation quantification. r^2 values represent correlations from Kendal-Tau analysis, * denotes significant correlations. Lines through data points represent linear regression.

5.3.3 5fC/CaC methylation and BDNF transcript expression

Correlation analysis was also performed for 5fC/CaC methylation at AS-rs6265 and *BDNF* transcript abundance for C homozygous genotypes. There were 34 C homozygous data points available. However, two data points were lost due to failed pyrosequencing and four data points were removed in quality control, hence 28 data points were used for 5fC/CaC analysis.

Figure 5.18 presents a scatterplot matrix of the data for AS-rs6265 5fC/CaC methylation and transcript abundance. All correlations between 5fC/CaC methylation and *BDNF* transcripts were weak in strength and none were not found to be statistically significant. 5fC/CaC methylation was also bimodally distributed when correlated with the transcripts, where 5fC/CaC methylation was either "low" or "high". Furthermore, although the correlations were not significant, as previously observed the sense and AS-204 transcript profiles were similar to each other and differed from the AS-201 profile (Figure 5.19). At the point of writing, this is the first time that antisense 5fC/CaC methylation and transcript abundance has been examined for correlations as a means to explain the potential role of oxidised methylation states in regulating BDNF abundance.

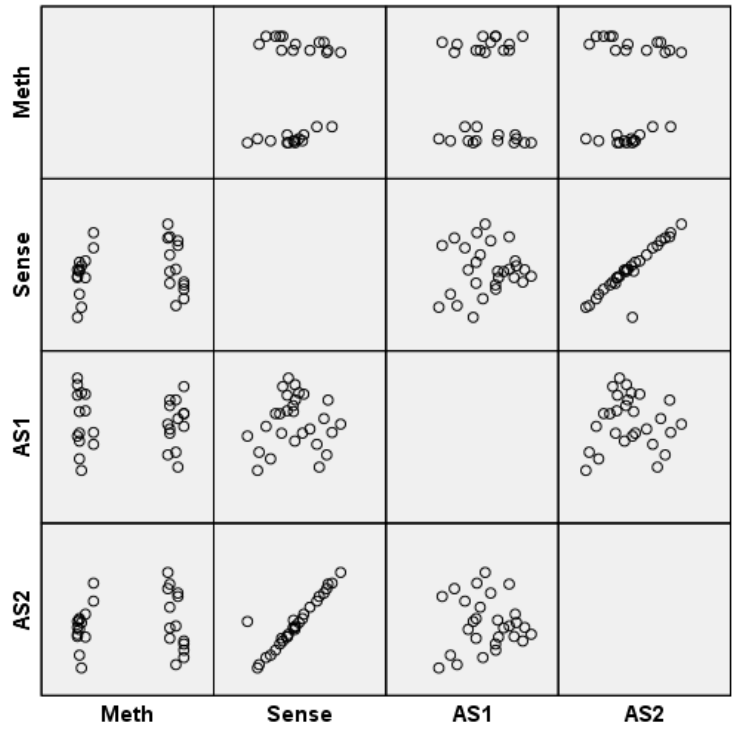


Figure 5.18 Scatterplot matrix showing the relationship between data transcript abundance and 5fC/CaC levels at AS-rs6265. Meth denotes 5fC/CaC methylation levels, sense, AS1 and AS2 denote sense, AS-201 and AS-204 transcripts, respectively.

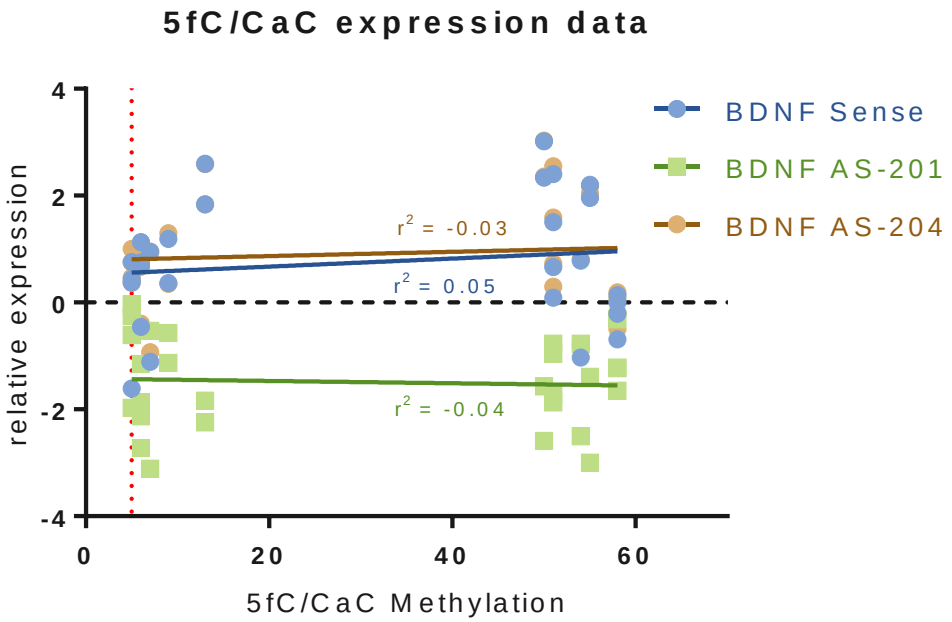


Figure 5.19 Correlations between AS-rs6265 5fC/CaC methylation and BDNF transcripts. Sense (blue), AS-201 (green), and AS-204 (orange) transcript abundance are presented. Lines through data points represent linear regression lines. Red dotted line indicates the 5% error threshold generated in the quantification of methylation by pyrosequencing. Black dotted line represents 0 or no change in relative expression. r^2 values represent correlations from Kendal-Tau analysis

5.3.4 5fC/CaC methylation, transcript expression, tissue type and diagnoses, with and without age and gender as covariates

Samples were subsequently grouped by brain region or disease phenotype status for all C homozygote samples. Unfortunately, only MCI and PD diagnoses remained in disease variable analysis as these were the only C homozygous diagnoses individuals (Table 5.1). Kendal-Tau analyses showed no significant correlations present between any transcript expression and 5fC/CaC methylation regardless of grouping by tissue or disease states.

Spearman's non-parametric partial correlation analyses were performed between 5fC/CaC methylation and the *BDNF* transcript abundance with age and gender as covariates. Including age and gender as covariates did not significantly change any correlations.

In summary, correlation analyses with transcripts and with methylation revealed several significant correlations. Sense mRNA expression and AS-204 expression appear to correlate strongly in hippocampus, cerebellum, and frontal cortex brain regions. *BDNF* mRNA and AS-204 transcript expression was significantly higher in the cerebellum compared to the frontal cortex and hippocampus, and this was true in mild cognitive impairment and Parkinson's disease samples, as well as when not accounting for disease. Additionally, mRNA expression was significantly higher in cerebellum than frontal cortex in dementia with Lewy body samples. AS-204 expression in the frontal cortex was also found to be significantly lower in healthy control samples compared to Parkinson's disease and dementia with Lewy body samples.

AS-201 reported a different expression profile, where expression of AS-201 was significantly decreased in cerebellum compared to hippocampus or frontal cortex samples. AS-201 was significantly increased Parkinson's disease and dementia Lewy bodies compared to mild cognitive impairment samples, and was also significantly increased in dementia with Lewy body compared to healthy control samples. This

significantly lower expression of AS-201 in MCI samples was observed specifically in the frontal cortex, but the effect was strong enough to be observed when not including brain regions as a variable.

AS-201 expression significantly correlated with 5mC methylation, while also reporting a significant negative correlation with 5hmC at antisense rs6265. When also including disease, tissue and genotype in analysis mRNA and AS-204 correlations also became significant. mRNA and AS-204 analysis reported significant negative correlations with 5mC methylation in MCI samples, and significant positive correlations with 5hmC methylation in homozygous C genotype samples. mRNA also reported a significant negative correlation with 5mC in DLB samples, however this was not observed in the AS-204 analysis. Finally, AS-201 expression reported a significant negative correlation with 5hmC methylation in cerebellum and C/T heterozygotes.

5.4 Discussion

5.4.1 BDNF transcript abundance and correlations within brain tissues and disease states

In this chapter, the transcript abundance of three *BDNF* transcript groups expressed in brain; the sense, AS-204, and AS-201 transcripts were compared. The sense and AS-204 transcripts showed high abundance and correlated strongly with each other, whereas AS-201 transcript expression was lower and did not correlate with the other two transcripts. Transcription of *BDNF* sense transcripts or coding mRNA enables the translation of functional BDNF proteins. The anti-sense transcripts (AS-201 and AS-204) are long non-coding elements that putatively act to regulate sense strand transcription and thereby regulate the translation into BDNF proteins (Modarresi et al., 2012; X. Zheng et al., 2016). The findings may indicate a regulatory role or interaction between the Sense/AS-204 and the AS-201 transcripts where either antisense transcript sequesters mRNA by creating an antisense-sense RNA duplex (Modarresi et al., 2012; Pruunsild et al., 2007). lncRNA regulated sequestration of mRNA is most likely performed by the AS-201 transcript as if the AS-204 transcript

acted to sequester mRNA via duplex, a proportional decrease in mRNA would be expected with AS-204; however, this is not the case and mRNA and AS-204 transcripts correlate in expression.

I propose that AS-201 acts as a lncRNA CRE, potentially as an R-loop mediated form of sequestration, or as an active recruiter of DNMTs to methylate the *BDNF* promoter and repress BDNF mRNA transcription, explaining why it has no correlation with mRNA and AS-204, and why it has opposite reactions to methylation observed at AS-rs6265. I report potential ASM in the form of significant negative correlations between mRNA and AS-204 with 5mC methylation in C homozygotes, and between AS-201 and 5hmC in heterozygotes at AS-rs6265, indicating that different genotypes and different transcripts harbour and respond to different types of methylation in different ways.

AS-204 correlated strongly with BDNF mRNA expression in this study, and previous work has shown that some *BDNF* anti-sense transcripts correlate with *BDNF* sense strand expression in the blood patients with multiple sclerosis (Gharzi et al., 2018), and indicate that correlations in transcripts may be modulated by regulatory mechanisms. What is more, while the anti-sense and sense *BDNF* were correlated in multiple sclerosis phenotypes, this was not true for control samples which indicates that this is a disease phenotype effect (Gharzi et al., 2018). Alternatively, antisense transcripts may not be a trans regulatory element but a long-non coding RNA (lncRNA) which acts upon other cis sites to repress transcription further downstream or upstream of the BDNF locus, as seen in the regulation of polycomb mediated repression and regulation of development (Dimitrova et al., 2014; Rapicavoli, Poth, Zhu, & Blackshaw, 2011).

The role of AS-204 transcripts may also be acting through an alternative regulatory process. While typically anti-sense transcripts act to repress transcription, some transcripts can repress elements that would otherwise repress transcription or translation resulting in an overall increase in transcription activation. One mechanism by which this may occur is through R-loop formation mediated

protection from methylation. CG rich regions may be transcribed to generate R-loops that inhibit *de novo* 5mC methylation of surrounding regions and stop transcriptional repression (Ginno et al., 2012). This is one explanation as to why the abundance of AS-204 transcript is similar to the amount of sense transcript, i.e. as AS-204 is expressed this may interact with and inhibit transcriptional repression of the sense transcript.

I theorise therefore, that the AS-204 transcript expression correlates with the *BDNF* mRNA transcript as 5mC methylation of AS-rs6265 for this transcript inhibits AS-204 transcription and therefore R-loop mediated repression of de novo methylation at the BDNF promoter, which in turn would decrease *BDNF* mRNA transcription. Alternatively, AS-204 may act as an intermediate stabiliser of the main BDNF mRNA transcripts by forming an RNA duplex in transportation, and repression of AS-204 through methylation decreases available transcript to stabilise *BDNF* mRNA, leading to reciprocal decrease in expression.

BDNF transcript expression across brain regions

BDNF transcript expression across the brain regions showed a brain region specific pattern. All transcripts (sense, AS-201, and AS-204) showed differential expression between brain regions except AS-201 expression in frontal cortex and hippocampus, which were not significantly different. This is supported by current research that suggests that the *BDNF* transcripts are expressed in a region specific manner (Pruunsild et al., 2007).

The highest levels of expression in sense and AS-204, and the lowest for AS-201, were observed in the cerebellum. This is in keeping with current expression profiles which show that the main *BDNF* sense transcript is most abundant in the cerebellum (Gte. Consortium, 2015). Although reference *BDNF* expression data (Gte. Consortium, 2015) were generated from healthy non-affected tissue and it is possible that disease samples in this study contributed to discrepancies from published data. High frontal cortex and hippocampus expression of anti-sense transcripts is consistent with published expression profiles and could, as noted above, be explained by AS-201

being a CRE (Gte. Consortium, 2015). Reference expression data does not support the low level of AS-201 transcript expression observed in the cerebellum. However, higher expression of the sense and AS-204 transcripts may be due to the low levels of AS-201 if it is acting as a repressive transcript.

The difference in transcript expression observed across brain regions may also be due to the composition of cells in the regions investigated. Cerebellum tissue showed a smaller range of cell types and are mostly composed of small neuronal granular cells and white matter glial cells (Hawrylycz et al., 2012). In contrast, frontal cortex and hippocampus regions have many different types of neuronal cells ranging from small inhibitory interneurons to large excitatory neurons and pyramidal cells as well as glial cells (Hawrylycz et al., 2012). The pathology associated with the disease phenotypes investigated in this study may also be contributing to the differential transcript expression. Previous studies have indicated that a similar level of AS-BDNF transcripts expression was detected throughout the brain regardless of region (Modarresi et al., 2012). As this was performed on control brain samples this may indicate a disease phenotype effect on the transcript levels detected in brain samples within this study.

Furthermore, cognitive pathological phenotypes are associated with disrupted neuronal integrity and an overall reduction in total neurons (Son, Shim, Kim, Ha, & Han, 2012). Total number of neurons in cognitive disease such as Alzheimer's disease and Parkinson's disease decrease in response to autophagy (Son et al., 2012). Moreover, α-synucleopathies such as Parkinson's disease and dementia with Lewy bodies disease accrue α-synuclein aggregates in pathology, which are known to disrupt cellular function and cause cell death (Jellinger & Korczyn, 2018; W. S. Kim et al., 2014; Spillantini et al., 1997). It is possible therefore that the disrupted phenotypes seen in these cognitive diseases may be due to altered transcript abundance. α-synucleopathies are known to sequester DNMT1 available in the cytoplasm and cause global hypomethylation which will have repercussions on the regulation of transcript abundance (Desplats et al., 2011). Risk loci associated with neuropsychiatric diseases such as autism, schizophrenia, Parkinson's disease and

Alzheimer's disease are also depleted for 5hmC in excitatory neurons and depleted for 5mC methylation in inhibitory neurons (Kozlenkov et al., 2018, 2016). This indicates distinct methylation profiles for cell type and disease, which may be modulating transcriptional abundance (Kozlenkov et al., 2018, 2016).

BDNF transcript expression across pathologies

BDNF is upregulated in expression as an immediate-early response to cellular insults, and BDNF levels increase in neurological tissue following injury (Hughes et al., 1999; Murer et al., 2001). However BDNF levels are also known to decrease with stress (Altar, 1999; Murer et al., 2001), age and gender (Lommatzsch et al., 2005). These factors could have influenced transcript expression levels observed in the current study. In contrast to differences found in transcript expression across brain regions, disease tissue samples only showed differential expression for the AS-201 transcript.

A significant upregulation of AS-201 abundance was observed in dementia with Lewy bodies samples compared to control samples. AS-201 transcript was also found to be expressed at significantly lower levels in mild cognitive impairment compared to Parkinson's disease and dementia with Lewy bodies samples. Together, this may indicate a load effect where more severe diseases show a higher upregulation of the AS-201 transcript compared to MCI and control samples. This also could suggest that AS-201 could consequently influence higher or lower repression of *BDNF* sense mRNA (and possible AS-204) dependent on the severity of pathology. If AS-201 is acting as a CRE it may be expected to observe the highest AS-201 expression in control, then Parkinson's disease, dementia with Lewy bodies, and finally mild cognitive impairment. While MCI is consistent with the CRE model, i.e. low AS-201 expression and high sense expression, the control samples are not. This indicates that other factors may be influencing the expression of *BDNF* sense transcript.

Transcript expression across disease states within tissue types, and tissue types within disease states

When investigating the effect of disease phenotypes within each brain region on transcript expression, AS-201 and AS-204 expression were observed to be

differentially expressed within the frontal cortex of all the diseased samples. This could be important in contributing to pathologies such as dementia, Alzheimer's disease, and mild cognitive impairment which are known to adversely affect the regulation of frontal cortex tissue.

The AS-201 transcript was significantly upregulated in Parkinson's disease and dementia with Lewy bodies compared to mild cognitive impairment samples but not control samples in frontal cortex. Mild cognitive impairment, Parkinson's disease and dementia with Lewy bodies are all associated with decreased BDNF levels in the brain (Borba et al., 2016; Ventriglia et al., 2013; von Bohlen und Halbach, Minichiello, & Unsicker, 2005). However α-synucleopathies such as Parkinson's disease and dementia with Lewy bodies, may be associated with reduced TrkB/C neurotrophic receptors while mild cognitive impairment is not (von Bohlen und Halbach et al., 2005). Therefore, although *BDNF* may contribute to the phenotypes a range of neurodegenerative disorders including cognitive impairment, they may not be through the same regulatory or physiological mechanisms.

The analysis of methylation and transcript expression in brain regions within disease states and disease states within brain regions provided similar results. Sense transcripts showed significant upregulation in abundance in the cerebellum of mild cognitive impairment, Parkinson's disease and dementia with Lewy bodies diagnoses. These findings support current expression profiles where cerebellum shows increased levels of *BDNF* mRNA, and this indicates that disease phenotypes in these tissues may not alter this dramatically (Gte. Consortium, 2015). These results suggest that transcript expression may have been altered in response to disease.

An increased expression of *BDNF s*ense transcript is reported to be protective against cognitive decline in later life through increased levels of the BDNF protein (Buchman et al., 2016). It is also known that anti-sense non-coding transcripts like AS-201 or AS-204 may act as interfering RNA that reduces overall mRNA available for translation into protein (Modarresi et al., 2012; Pruunsild et al., 2007). AS-201 transcript showed significantly lower expression in the cerebellum of control and mild cognitive

impairment samples. In this way, it is possible that the upregulation of AS-201 seen in disease states may be reducing the amount of *BDNF* sense mRNA available explaining why there is a reduced amount of sense transcript abundance disease states like Parkinson's disease and dementia with Lewy bodies and their tissues.

5.4.2 Comparing BDNF transcript abundance and 5mC, 5hmC and 5fC/CaC methylation at antisense rs6265

5mC, 5hmC AS-rs6265 methylation and correlation with transcript abundance

AS-201 abundance was significantly positively and negatively correlated with 5mC and 5hmC methylation at AS-rs6265, respectively. Previous research in rats has indicated that *BDNF* DNA methylation in the hippocampus as determined through bisulphite sequencing was associated with increased or decrease in abundance of exon-specific *BDNF* transcripts (Lubin et al., 2008). Our findings support *BDNF* transcript specific expression facilitated by *BDNF* AS-rs6265 5mC and 5hmC methylation. It also indicates that transcript expression mediated by methylation is facilitated by antisense methylation at key sites. However, whether AS-rs6265 methylation is causing the transcriptional abundance change alone, or whether methylation is representative of surrounding differentially methylated region (DMR) methylation, is unknown.

5mC is known to be a transcriptional repressor and is associated with decreased gene expression (Holliday & Pugh, 1975; Riggs, 1975). However, AS-201 expression was found to increase with 5mC methylation of the AS-rs6265 CpG which is opposite to the typical repressive effects of 5mC methylation on transcript expression. It is known that anti-sense transcripts can regulate sense transcript expression by exerting cis-regulatory effect. For example, Acetylcholinesterase (*AChE*) anti-sense transcripts have been reported to negatively correlate with the expression of *AChE* sense transcripts in hepatoma cells, and silencing *AChE* anti-sense transcripts increased *AChE* sense transcript expression (Xi et al., 2014). Moreover, a reduction in antisense *AChE* abundance increased H3K4 and decreased H3K9 histone methylation over the *AChE* promoter region, indicating alterations in transcript abundance may

influence epigenetic regulation (Xi et al., 2014). While this research was in carcinoma cells, *BDNF* may be being regulated by the same mechanisms as *AChE*.

In addition, Haemoglobin A2 (*HBA2*) anti-sense transcripts cause 5mC methylation of the genes regulatory promoter CpG island which causes repression of the genes mRNA in patients with α-thalassemia (Tufarelli et al., 2003). In this way *HBA2* anti-sense transcript directs 5mC methylation and represses expression of the gene in an imprinting and X inactivation independent manner. It is therefore possible that a similar effect is occurring in *BDNF* as while AS-rs6265 5mC increases and sense transcripts decrease, AS-201 transcript increases. If this were the case, it is possible that the 5mC methylation observed at AS-rs6265 is representative of the increase 5mC methylation observed across *BDNF*, which may be caused in part by AS-201 expression.

As the rs6265 variant is located within gene body and towards the 3' end of the coding region, AS-rs6265 5hmC methylation may be an indicator of increased transcriptional expression (S.-G. Jin et al., 2011) and our findings support this prediction. The presence of 5hmC may be capable of initiating demethylation to increase transcription (Junjie U Guo et al., 2014) and this is reflected in the abundance of sense and AS-204 increasing with 5hmC methylation at AS-rs6265 and the strong correlation between 5hmC methylation with these transcripts. This supports 5mC/5hmC regulation of transcription regardless of the strand in which CpG methylation is occurring.

BDNF transcript abundance, 5mC and 5hmC methylation, in tissues and disease states

Both sense and AS-204 showed significantly negative correlation between transcript expression and 5mC methylation at the AS-rs6265 CpG in mild cognitive impairment samples which supports our previous findings suggesting both of these transcripts were down regulated in frontal cortex in mild cognitive impairment samples (Figure 5.17). This suggests a link between sense and AS-204 transcript repression and AS-rs6265 5mC with the mild cognitive impairment phenotype. The sense transcript also exhibited significantly decreased expression in dementia with Lewy bodies samples

correlating with AS-rs6265 5mC methylation. These findings may indicate a disease specific effect in the regulation of all transcripts when DNA is 5mC methylated, irrespective of strand of CpG.

AS-201 transcript expression was significantly downregulated when AS-rs6265 was 5hmC methylated in both cerebellum and in heterozygous genotype samples (Figure 5.17). The downregulation of AS-201 transcript in C/T genotypes may be allele specific methylation (Shoemaker et al., 2010), but may also be driven by disease phenotypes.

As discussed in chapter 4, ASM has been associated with allele specific expression in a phenotype and tissue specific manner (Yan, Yuan, Velculescu, Vogelstein, & Kinzler, 2002). Kinzler et. al. showed that lymphoblastoid DNA taken from families which were *p73* single nucleotide polymorphism 629T/C heterozygotes showed up to a 4.3 fold higher expression for one allele over the other (Yan et al., 2002). Moreover, up to 48% of partial or "fuzzy" methylation at CpG sites occurs in an allele specific manner (Yan et al., 2002). In this way the modification in transcript expression may be due to the genotype at rs6265 (and therefore AS-rs6265), and/or the allele specific methylation available at the site. As methylation may favour one allele over the other at *BDNF* rs6265, the resultant bias in allele transcript expression may be a reflection of the methylation states there. Additionally, if 5hmC methylation was present, this may account for higher levels of expression, however not all 5hmC is stable long term (Hahn et al., 2014; Lin, Chen, & Hsu, 2017).

Dementia with Lewy bodies phenotype influences the expression of sense transcript and its correlation with methylation, and these data support a link between DLB phenotypes, methylation and transcript specific expression. Moreover, as the effect is persistent over transcripts which do not typically correlate it is a strong and conserved effect. Future research into the effect of methylation at *BDNF* rs6265 with the transcript expression would help to determine the full effect of DLB phenotypes at this epiallele.

5fC/CaC methylation and BDNF transcript abundance

Little information is available on base resolution whole genome 5fC/CaC methylation compared with 5mC and 5hmC methylation. This is due to the fact that technologies used to assess 5fC/CaC methylation are still being developed. However, Ito and colleagues have shown that 5fC/CaC is typically present at orders of magnitude lower levels than 5mC or 5hmC, and would typically be below the 5% margin of error in the quantification of 5fC/CaC by pyrosequencing on the sense strand (Ito et al., 2011; S. Liu et al., 2013). We found that the level of 5fC/CaC at the rs6265 was bimodal in nature with typically below 20% or above 50% of the AS-rs6265 site 5fC/CaC methylated.

Recent studies suggest that 5fC and 5CaC may act in a transcriptionally repressive way by interfering with RNA polymerase II transcription rate and substrates (Kellinger et al., 2012). However 5hmC, 5fC and 5CaC are all oxidised products of the DNA demethylation pathway (Ito et al., 2011), which brings into question whether the DNA modifications are actively maintained or a by-product of base excision repair into cytosine (or even if they can be actively maintained) (C. X. Song & He, 2013). Such mechanisms could affect how transient these modifications are and the likelihood of detection in that the higher rate of turnover and the more transient the methylation, the less likely the modification could be detected at sufficient levels by pyrosequencing. (C. X. Song, Diao, Brunger, & Quake, 2016). This may, in part, explain why all correlations observed between 5fC/CaC methylation and transcript expression were found to be weak. Additionally, changes in methylation may be a response to disease phenotypes, and methylation modifications may be unable to reverse back to “healthy” non-pathological levels after degenerative disease phenotypes such as Alzheimer’s disease, Parkinson’s disease, dementia with Lewy bodies have developed.

Additionally, 5fC and 5CaC have many more proteins capable of interacting or reading the methylation motifs compared to 5hmC in mouse embryonic stem cells, for example thymine DNA glycolase (*TDG*), p53 (*TP53*), and Nei like DNA glycosylase 3 (*NIEL3*) (Spruijt et al., 2013). This suggests that these modifications have a specific

function (Spruijt et al., 2013). Hence, methylation observed may be a result of many complex interactions and may be more likely to affect expression of one transcript over others.

Antisense rs6265 methylation

It is thought that this study is the first instance of analysis examining complementary CpG methylation at a risk allele and its effects on transcript expression. Our findings suggest that the complementary CpG methylation affects transcript expression in the same way the sense CpG methylation would; where 5mC is repressive and 5hmC enhances transcript specific expression. This study links methylation of the AS-rs6265 CpG with AS-201/4 transcript expression. As AS-204 transcript expression correlates strongly with strand expression, the AS-rs6265 CpG methylation may be a proxy measurement for expression. Whether this is due to complementary CpG methylation at rs6265 is unknown.

DNA methylation regulation of neighbouring genomic loci

Using existing data for 4C chromatin conformation capture which maps regions of DNA-DNA interactions, BDNF has been shown to interact with neighbouring genes such as Methyltransferase Like protein 15, *METTL15,* in multiple brain regions such as the hippocampus and frontal gyrus (Yanli Wang et al., 2018). This gene encodes a *de novo* methylase which creates N4-methylcytosine (4mC) methylated ribosomal RNA, using S-adenosyl-L-methionine (SAM) as a cofactor ("UniProt: a worldwide hub of protein knowledge," 2019). This leaves open the possibility that methylation of *BDNF* could regulate METTL15 expression which in turn influences rRNA production.

Similar to the association between the *BDNF* rs6265 minor allele, a decline in working memory and increased risk and severity of schizophrenia, Alzheimer's disease, and Parkinson's disease (M. F. Egan et al., 2001; R. Huang, Huang, Cathcart, Smith, & Poduslo, 2007; Ursini et al., 2016; Van Der Kolk et al., 2015) have been associated with the Val allele of variant Val108/158Met (rs4680) within Catechol-O-methyltransferase (*COMT*). This is facilitated by the Val allele while accounts for a

four-fold variation in COMT activity and increases dopamine catabolism (M. F. Egan et al., 2001).

Moreover, the minor Val allele introduces a CpG site similar to that of BDNF rs6265, and the Val/Val genotype as well as the methylation at the *COMT* rs4680 variant was significantly associated with homicidal behaviour in schizophrenia diagnoses (Y. Hu et al., 2018). This highlights the connection between variant epialleles, regulation of methylation, changes in cognition, and working memory. Further research into the effects of DNA methylation regulation at these genes and variant sites and BDNF gene expression would be warranted.

5.4.3 Limitations and future research

This study had several limitations. Small group sizes meant that the distribution of genotypes in this study were not a full representation of the population, and poor-quality pyrosequencing data often meant that insufficient power was available for significance to be determined. While age and gender covariates were included in analyses, other cofactors that are known to affect *BDNF* expression such as intake of vitamins, medication, and stress were not quantified. Future studies may aim to incorporate these factors. Furthermore, quantification of methylation only targeted the antisense strand in pyrosequencing. To account for this, future study should aim to quantify methylation on both the sense and anti-sense strands.

Moreover, further experiments which aim to clone BDNF into transgenic mice to determine the action of anti-sense BDNF expression through silencing and quantification of methylation would be required to determine if AS-201 is exerting an epigenetic pressure on sense strand expression. This could be further explored through the quantification of BDNF protein in blood to determine if changes in BDNF level occur because of DNA methylation, transcription, post transcriptional effects or post-translational effects. Additionally, analysis of the complementary CpG methylation will allow for the corroboration between strand specific methylation types and levels on transcript abundance.

Research targeting cell populations e.g. glial cells through cell sorting, would allow for the comparison of methylation and expression in cell types. NGS sequencing is a more sensitive and accurate method to quantify modified bases (Lander et al., 2001; Mardis, 2008) and would allow for more accurate quantification of methylation and the differentiation between 5fC and 5CaC methylation. Technologies such as melting-curve analysis are being developed to genotype *BDNF* haplotypes and variant sites more cheaply and quickly than before (Sánchez-Romero, Dorado, Guarino, & Llerena, 2009). The use of single-molecule real-time (SMRT) sequencing has the ability to detect all known forms of DNA methylation without conversion or additional (potentially damaging) treatments (Flusberg et al., 2010) which have reduced the success rate of our pyrosequencing. In the future these technologies will open up more accurate, fast paced, and complete analysis of methylation which will allow us to better understand the roles of these methyl motifs in cognition and disease.

5.4.4 Conclusion

BDNF sense and AS-204 transcript abundance were found to highly correlate whereas the antisense AS-201 transcript showed a different expression pattern. AS-201 did not correlate with Sense/AS-204, while both sense and AS-204 transcripts correlated. Either AS transcript may be a cis-regulatory element in the repression of *BDNF* sense transcript. While this cannot be determined in this study, the close correlation between the mRNA and AS-204 transcripts suggest that the AS-201 transcript may be regulating mRNA expression, and possible AS-204 by proxy. In disease analyses of transcript expression, dementia with Lewy bodies diagnoses displayed significantly upregulated expression of AS-201 and lower *BDNF* mRNA levels compared to controls, which may be a promising diagnostic if refined by future studies. Parkinson's disease and dementia with Lewy bodies diagnoses showed the highest upregulation of AS-201 transcript, which may indicate that there is a burden effect whereby more severe disease states show the highest AS-201 expression, and potentially resulting in the repression of *BDNF.* This difference may be driven by pathologies.

5mC and 5hmC methylation of AS-rs6265 was found to affect *BDNF* expression in a transcript specific manner. *BDNF* mRNA and AS-204 transcripts showed a significant decrease in expression when AS-rs6265 was 5mC methylated in severe disease states like dementia with Lewy bodies, while AS-201 expression increased. Homozygous C genotypes significantly increased expression of mRNA and AS-204 transcripts while decreasing that of AS-201, possibly due to ASM. These results indicate that methylation is occurring in a tissue, disease, genotype specific manner and methylation levels and transcript expression are affected by all of these variables. This may further support AS-201 as a CRE, and literature has also suggested that 5hmC is a transcriptional enhancer, and 5mC is a transcriptional repressor. It is possible that the *BDNF* AS-201 transcript may be acting similarly to the *HBA2* anti-sense transcript that facilitates the 5mC methylation of regulatory promoters which effectively repress sense strand expression (Tufarelli et al., 2003).

No significant correlations between 5fC/CaC methylation in Homozygous C individuals and *BDNF* transcript abundance were observed. However, this study observed bimodally distributed 5fC/CaC methylation at an antisense CpG which unlike 5mC or 5hmC methylation showed no correlations with sense or antisense transcripts, which is a novel finding. This research lays the foundation for future studies which can apply new technologies to enable a more accurate and comprehensive analysis of oxidised DNA methylation modifications at risk epialleles and role in gene regulation in cognitive diseases.

Chapter 6: *APOE* rs7412 and rs42958 epiallele methylation analysis

Acknowledgements

Daniel Ellis from QIAGEN aided in the optimisation and design of pyrosequencing primers designed to target both *APOE* variant sites. Dr Chang-En Yu, University of Washington, advised on the designing of primers to target bisulphite converted rs429358. William Hawkins, Edinburgh CRF, performed the pyrosequencing and all associated PCR amplifications and optimisations with pyrosequencing primers. Dominika Valent Raffajová, University of Nottingham, aided in genotyping of samples.

Preface

APOE rs7412 and rs429358 determine the ε isoform of APOE (Nickerson et al., 2000) and are the largest risk for late onset Alzheimer's disease (Corder et al., 1993; Green et al., 2009). While methylation has been quantified at rs7412 and rs429358 in Alzheimer's disease through bisulphite sequencing, which does not differentiate between 5mC and 5hmC methylation (Foraker et al., 2015), 5fC/CaC methylation have not been quantified, and the association between methylation at the variant sites, tissues, and other disease states such as Parkinson's disease and dementia with Lewy bodies, has not been examined. This chapter discusses the quantification of methylation at both the *APOE* rs7412 and rs429358 variant epialleles and observed the effect of tissue, disease status, and genotype, as well as age and gender covariates on methylation levels observed.

6.1 Sequencing genotype and methylation analysis of *APOE* variants

Snap frozen brain tissue samples from the OPTIMA cohort and Imperial Brain Bank were genotyped by Sanger Sequencing for the *Apolipoprotein E* (*APOE*) risk alleles rs7412 and rs429358. Demographic details, disease status and the results of the genotyping to determine APOE ε isoforms are presented in Table 6.1.

Sample ID	Age	Gender	Disease	*APOE* rs7412	*APOE* rs429358	ε Isoform
OPTIMA 111	84	Male	Alzheimer's Disease	C/C	T/T	ε3/ε3
OPTIMA 108	-	Female	Alzheimer's Disease	C/C	T/C	ε3/ε4
OPTIMA 199	70	Female	Alzheimer's Disease	C/C	T/C	ε3/ε4
OPTIMA 253	79	Female	Alzheimer's Disease	C/C	T/C	ε3/ε4
OPTIMA 195	74	Male	Alzheimer's Disease	C/C	T/C	ε3/ε4
OPTIMA 215	77	Female	Alzheimer's Disease	C/C	C/C	ε4/ε4
OPTIMA 250	69	Female	Alzheimer's Disease	C/C	T/C	ε3/ε4
OPTIMA 297	65	Female	Alzheimer's Disease	C/C	T/T	ε3/ε3
OPTIMA 47	82	Female	Alzheimer's Disease	C/C	T/C	ε3/ε4
OPTIMA 58	88	Male	Alzheimer's Disease	C/C	T/C	ε3/ε4
OPTIMA 179	70	Female	Alzheimer's Disease	C/C	C/C	ε4/ε4
PD687	86	Male	Parkinson's Disease	C/C	T/T	ε3/ε3
PD793	75	Female	Parkinson's Disease	C/C	T/T	ε3/ε3
PD898	84	Male	Parkinson's Disease	C/T	T/T	ε2/ε3
PD115	63	Male	DLB	C/C	C/C	ε4/ε4
PD120	83	Female	DLB	C/C	C/C	ε4/ε4
PD332	77	Male	DLB	C/C	T/C	ε3/ε4
PDC007	104	Female	MCI	C/T	T/T	ε2/ε3
PDC111	88	Female	MCI	C/T	T/C	ε2/ε4 (ε1/ε3)
PD715	86	Male	MCI	C/C	T/C	ε3/ε4
C045	77	Male	MCI	C/C	C/C	ε4/ε4
PD065	84	Female	MCI	C/C	C/C	ε4/ε4
PD092	79	Male	Control	C/C	C/C	ε4/ε4
PDC110	94	Female	Control	C/T	T/C	ε2/ε4 (ε1/ε3)
PDC112	87	Male	Control	C/C	T/T	ε3/ε3

Table 6.1 *APOE* rs7412 and rs429358 genotypes, and ε isoforms for brain samples examined. Individual age at sampling, gender, and diagnosis are presented. Mild cognitive impairment (MCI), and dementia with Lewy bodies (DLB).

rs7412	rs429358	Isoform	Effect
C/T	T/C	ε2/ε4	Increased risk. Indistinguishable from ε1/ε3
T/T	T/T	ε2/ε2	Protective effect
C/T	T/T	ε2/ε3	Partial protective effect
C/C	T/T	ε3/ε3	Standard likelihood of AD
C/C	T/C	ε3/ε4	Increased risk
C/C	C/C	ε4/ε4	Maximum risk

Table 6.2 Major ε Isoforms with their associated rs7412 and rs429358 genotypes. Effects denotes whether the isoform increases or increases the chances of an Alzheimer's disease (AD).

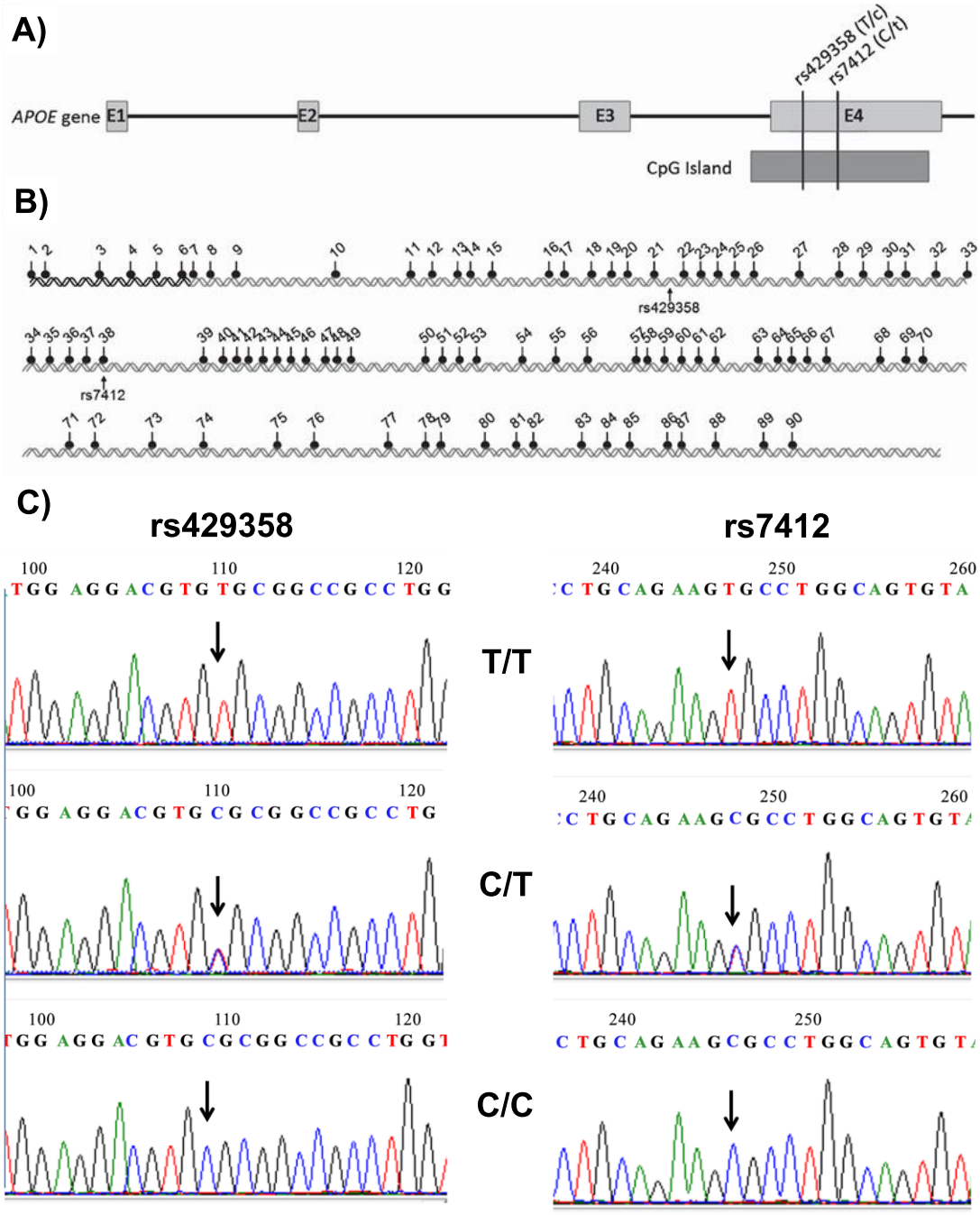


Figure 6.1 The structure of *APOE* and the surrounding sequence of rs7412 and rs429358. A) Location of the rs7412 and rs42358 variant sites in exon 4 of *APOE*. B) Area surrounding the rs7412 and rs429358 variants. Each point represents a CpG site, starting at the beginning of the *APOE* exon 4 DMR1. C) Examples of Sanger sequencing peak-calling files calling different genotypes at both the rs7412 and rs429358 sites. T is the major allele for rs429358, while C is the major allele for rs7412. Panel A and B were taken from (Foraker et al., 2015), panel C was taken from (C.-H. Chen, 2016).

6.1.1 Genotypes at APOE rs7412 and rs429358

The APOE rs7412 and rs429358 alleles and ε Isoforms

The *APOE* rs7412 variant is a coding variant that causes an Arg158Cys amino acid change, while rs429358 causes a Cys112Arg amino acid change in the primary APOE variant (Zannis et al., 1982). *APOE* genotypes, ε isoforms and their predicted risk effect are presented in Table 6.2. Figure 6.1 shows the region surrounding *APOE* rs7412 and rs429358 variant sites, exon location, and CpG sites in exon 4. Of the 25 individuals genotyped, 20 (80%) were C major allele homozygotes, 5 (20%) were heterozygous (C/T), and no samples were T homozygous for rs7412. The frequencies of C homozygous and heterozygous samples were similar to the reported frequencies observed in the 1000 genome project (R. A. Gibbs et al., 2015). Of the 25 individuals genotyped for rs429358, 9 (36%) were T major allele homozygotes, 10 (40%) were heterozygotes and 6 (24%) of the samples were C minor allele homozygotes at rs429358. Our samples reported a depletion of T homozygotes and an enrichment of C minor allele homozygotes (R. A. Gibbs et al., 2015). The current study's samples show a bias towards higher risk ε isoforms as the risk genotype is typically found more frequently in population with forms of dementia (Lane et al., 2009; McDowell, 2001). Within the Alzheimer's diagnoses, the ε4 isoform was common (50%), with an equal number of ε3 isoform and no ε2 isoforms. Overall the samples have an over-representation of ε4 and an under-representation of ε3 in both groups compared to a population control frequencies (R. A. Gibbs et al., 2015).

6.2 5mC and 5hmC methylation analyses at *APOE* rs7412 and rs429358

Unfortunately, statistical analysis was not possible for the *APOE* rs7412 and rs429358 variants beyond analysis of 5mC and 5hmC methylation across genotypes. This is because the small sample sizes made LMEM analysis unreliable, due to low power and factor reduction. Instead, the data were presented in graphs displaying either 5mC/5hmC across C/T heterozygote and C homozygote individuals, or 5mC, 5hmC, 5fC/CaC methylation across C homozygote individuals. All samples were grouped by brain region, disease, and genotype separately at each variant site. A joint graph

showing both risk variants and indicative of the isoforms is presented at the end of the chapter.

6.2.1 5mC and 5hmC methylation compared to Carrier status

5mC/5hmC methylation at APOE rs7412 by genotype carrier status

Samples were grouped by C/T heterozygous or C homozygous genotype status. Of the potential 62 5mC and 5hmC methylation data points available for analysis, 22 data points were excluded due to poor pyrosequencing quality. Therefore 40 data points remained for analyses. These samples are shown in Table 6.3 to show how tissues and disease states may overlap. No diagnostic or brain region subgroupings were excluded owing to missing data and each diagnostic group had at least one sample for each available tissue group.

5mC or 5hmC methylation levels were compared between the C homozygotes and C/T heterozygotes. Kruskal-Wallis analysis reported significant differences between the 5mC methylation levels reported in rs7412 carrier and non-carrier genotypes (K-W $X^2(1) = 7.85$, $p = 5\times10^{-3}$). This is presented in Figure 6.2. However, there was a similar pattern of 5mC and 5hmC methylation in both C homozygotes and C/T heterozygotes where 5mC levels were higher than 5hmC methylation levels (Figure 6.2). At the rs7412 allele both 5mC and 5hmC methylation was observed confirming this variant is a risk epiallele.

5mC/5hmC methylation at APOE rs429358 and genotype

Because rs429358 has a major T allele, the majority of samples possessed the T allele as T homozygotes or C/T heterozygotes. Methylation should not occur at non-cytosine containing dyads such as a TpGs, and of the potential 62 data points available for analysis 25 T homozygote data points were excluded. A further 15 data points were excluded due to poor pyrosequencing results, leaving 22 data points for analysis. These samples are shown in Table 6.4 to show how tissues and disease states may overlap. No disease or brain region groups were lost and each diagnostic group had at least one sample for each available tissue group.

At rs429358 both 5mC and 5hmC methylation were observed confirming this variant is also a risk epiallele. Kruskal-Wallis analysis indicated no significant differences between either 5mC or 5hmC methylation across genotype groups (K-W $X^2(1) = 2.43$, $p = 0.12$ for 5mC, K-W $X^2(1) = 2.38$, $p = 0.12$ for 5hmC methylation). This is presented in Figure 6.3.

Sample ID	Tissue	Disease	APOE rs7412
OPT 111	BL	AD	C/C
OPT 111	WB	AD	C/C
OPT 199	BL	AD	C/C
OPT 253	BL	AD	C/C
OPT 253	WB	AD	C/C
OPT 108	BL	AD	C/C
OPT 195	WB	AD	C/C
OPT 215	BL	AD	C/C
OPT 215	WB	AD	C/C
OPT 250	WB	AD	C/C
OPT 297	BL	AD	C/C
OPT 297	WB	AD	C/C
OPT 58	WB	AD	C/C
OPT 179	BL	AD	C/C
OPT 179	WB	AD	C/C
PD687	FC	PD	C/C
PD687	CER	PD	C/C
PD793	FC	PD	C/C
PD793	HIP	PD	C/C
PD793	CER	PD	C/C
PD120	FC	DLB	C/C
PD120	HIP	DLB	C/C
PD120	CER	DLB	C/C
PD332	FC	DLB	C/C
PDC111	FC	MCI	C/T
PDC111	HIP	MCI	C/T
PDC111	CER	MCI	C/T
PD715	FC	MCI	C/C
PD715	HIP	MCI	C/C
PD715	CER	MCI	C/C
C045	CER	MCI	C/C
PD065	FC	MCI	C/C
PD065	HIP	MCI	C/C
PD065	CER	MCI	C/C
PD092	HIP	CNTRL	C/C
PDC110	FC	CNTRL	C/T
PDC110	CER	CNTRL	C/T
PDC112	FC	CNTRL	C/C
PDC112	HIP	CNTRL	C/C
PDC112	CER	CNTRL	C/C

Table 6.3 Samples used to assess *APOE* rs7412 5mC and 5hmC methylation across genotype, tissues and disease states. Tissues are shown where BL, WB, FC, HIP and CER, refer to blood, whole brain, frontal cortex, hippocampus, and cerebellum, respectively. Disease states are shown where AD, PD, DLB, MCI and CNTRL, refer to Alzheimer's disease, Parkinson's disease, Dementia with Lewy Bodies, Mild Cognitive Impairment, and Control samples, respectively.

Sample ID	Tissue	Disease	APOE rs429358
OPT 108	BL	AD	C/T
OPT 179	WB	AD	C/C
OPT 195	WB	AD	C/T
OPT 199	BL	AD	C/T
OPT 215	BL	AD	C/C
OPT 215	WB	AD	C/C
OPT 250	WB	AD	C/T
OPT 253	BL	AD	C/T
OPT 58	WB	AD	C/T
PD115	FC	DLB	C/C
PD120	HIP	DLB	C/C
PD120	CER	DLB	C/C
PD332	FC	DLB	C/T
PD332	HIP	DLB	C/T
PD065	FC	MCI	C/C
PD065	CER	MCI	C/C
PDC111	CER	MCI	C/T
PD092	HIP	CNTRL	C/C
PD092	CER	CNTRL	C/C
PDC110	FC	CNTRL	C/T
PDC110	HIP	CNTRL	C/T
PDC110	CER	CNTRL	C/T

Table 6.4 Samples used to assess *APOE* rs429358 5mC and 5hmC methylation across genotype, tissues and disease states. Tissues are shown where BL, WB, FC, HIP and CER, refer to blood, whole brain, frontal cortex, hippocampus, and cerebellum, respectively. Disease states are shown where AD, PD, DLB, MCI and CNTRL, refer to Alzheimer's disease, Parkinson's disease, Dementia with Lewy Bodies, Mild Cognitive Impairment, and Control samples, respectively.

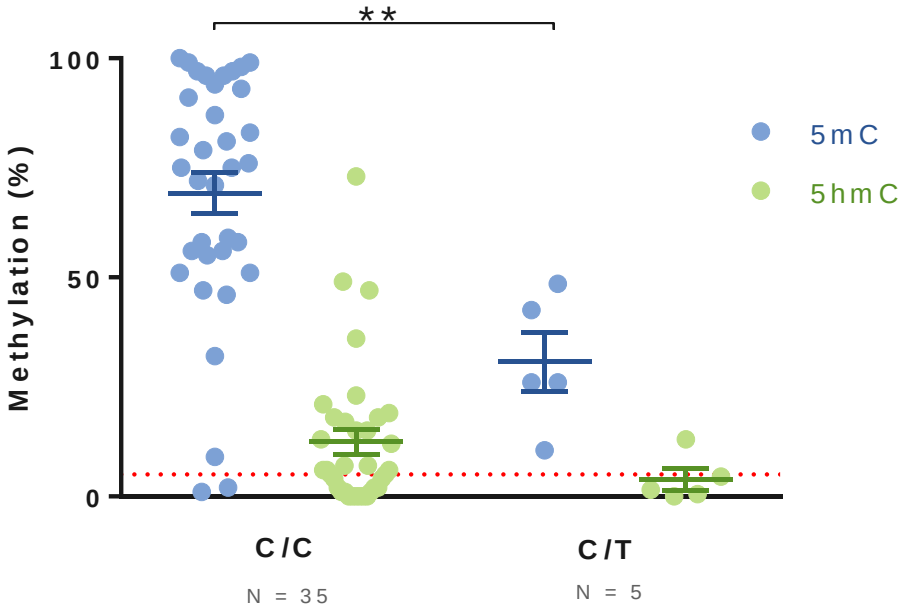


Figure 6.2 5mC and 5hmC methylation at *APOE* rs7412 in heterozygotes (C/T) and C homozygotes (C/C). 5mC and 5hmC methylation are presented in blue or green, respectively. Individual dots represent each sample methylation value, long bars represent means, and short bars represent standard error around the mean. Group numbers (N) are shown below carrier and non-carrier groups. 5% margin of error is indicated by the red dotted line. (* p≤0.05, ** p≤0.01, and *** p≤0.001).

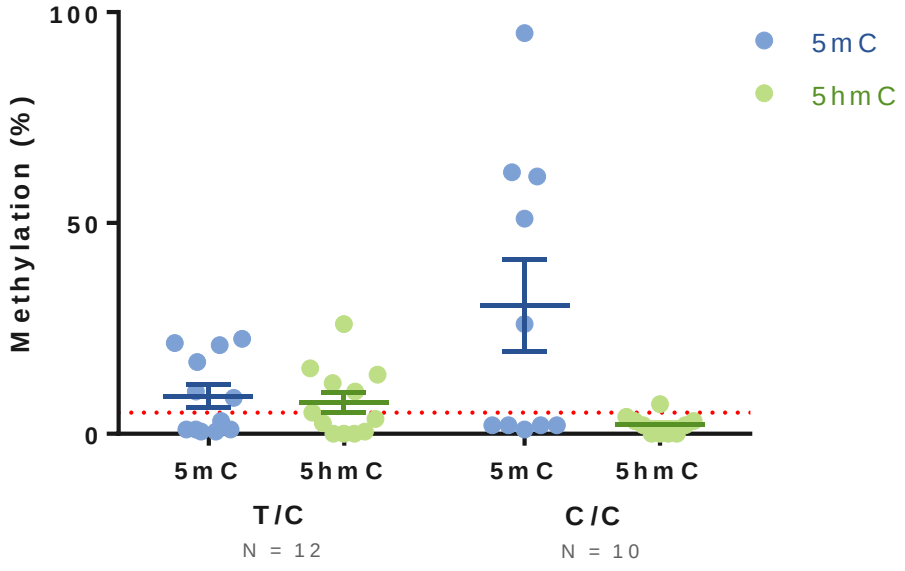


Figure 6.3 5mC and 5hmC methylation across *APOE* rs429358 heterozygotes (T/C), and C homozygote (C/C). 5mC and 5hmC methylation are presented in blue or green, respectively. Individual dots represent each sample methylation value, long bars represent means, and short bars represent standard error around the mean. Group numbers (N) are shown below genotype groups. 5% margin of error is indicated by the red dotted line.

6.2.2 5mC and 5hmC methylation, compared to genotype and tissue types

APOE rs7412 methylation, genotype, and tissue types

As there were no rs7412 C/T heterozygotes in the OPTIMA AD homogenised brain cohort, blood and whole brain samples were only present in C homozygotes (Table 6.3). Figure 6.4 shows the levels of 5mC and 5hmC methylation across tissue types and genotypes. Median levels of 5mC were higher than that of 5hmC within the same tissues and genotypes, for all groups. Some levels of 5hmC were close to or below the 5% limit for error as generated through pyrosequencing, such as cerebellum samples, and hippocampus samples in heterozygotes.

APOE rs429358 methylation, genotype and tissue types

Figure 6.5 reports the 5mC and 5hmC methylation levels across different tissues and genotypes, at the *APOE* variant site rs4293568. Unlike the observations for rs7412, the levels of methylation at the rs429358 variant were more variable. While median 5mC methylation was more commonly higher than that of 5hmC methylation, this was not the case for blood and whole brain samples in heterozygotes. Similarly, median levels of methylation in cerebellum samples, and in hippocampal samples from C homozygotes were below that of the 5% limit of quantification.

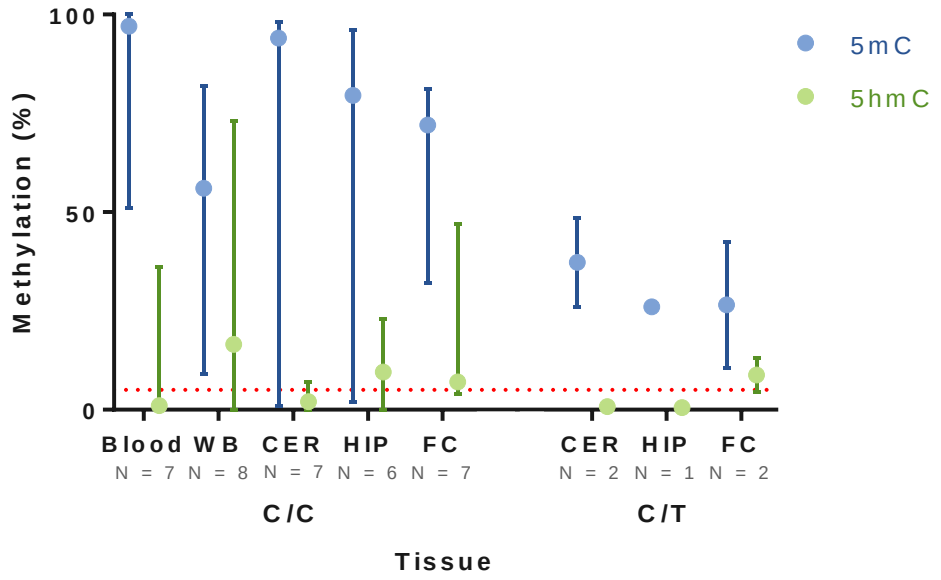


Figure 6.4 5mC and 5hmC methylation across tissue groups and genotype at rs7412. Median percentages and range for 5mC or 5hmC methylation levels are shown in blue, and green, respectively. Blood, whole brain (WB) cerebellum (CER), hippocampus (HIP), and frontal cortex (FC) tissue groups, and genotype are shown in the x-axis, with sample sizes (N) for each group. Genotype status is given below x-axis; either carrier (C/T) or non-carrier (C/C). 5% error margin for percentage of methylation is indicated by a red dotted line.

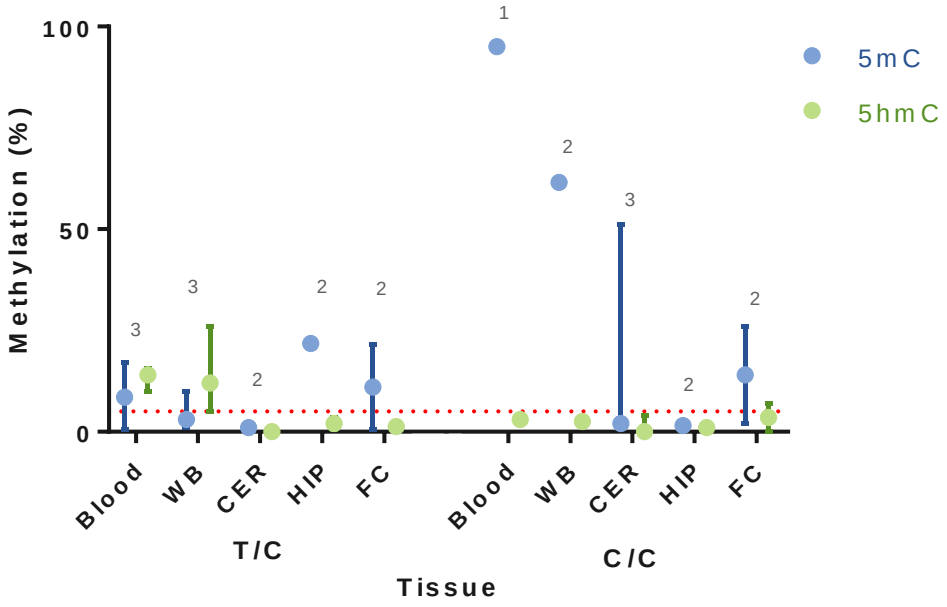


Figure 6.5 5mC and 5hmC methylation across tissue groups and genotype at rs429358. Median percentages and range for 5mC or 5hmC methylation levels are shown in blue, and green, respectively. Blood, whole brain (WB) cerebellum (CER), hippocampus (HIP), and frontal cortex (FC) tissue groups, and genotype are shown in the x-axis. Sample sizes (N) for each group are shown above the bars. Genotype status is given below x-axis; either heterozygote (C/T) or CC homozygote (C/C). 5% error margin for percentage of methylation is indicated by a red dotted line.

6.2.3 5mC and 5hmC methylation, compared to genotype and disease states

APOE rs7412 methylation, carrier status, and Disease states

As pyrosequencing failed for the Parkinson's disease rs7412 C/T heterozygotes, heterozygotes were present only as control and MCI data points (Table 6.3). As with 5mC and 5hmC methylation at rs7412 in tissues, methylation in disease states at rs7412 report higher levels of 5mC methylation compared to 5hmC methylation when in the same disease state and genotype status. Median 5mC was higher than 5hmC regardless of disease state or genotype (Figure 6.6). The range of 5hmC methylation in Alzheimer's diagnosis C homozygotes is observationally larger than the other diagnoses 5hmC range. This may be attributed to the high ranged of 5hmC methylation observed in the whole blood of C homozygotes at rs7412 (Figure 6.4). Due to small sample sizes there was insufficient power to determine if methylation levels were significantly different to that of other disease states or genotypes.

APOE rs429358, methylation, genotype and disease states

As the PD diagnosis samples were all T homozygotes, they were excluded (Table 6.4). The levels of 5mC and 5hmC methylation at rs429358 in different disease states and genotypes are similar to the levels observed when assessing methylation by tissues. A majority of samples have higher median 5mC than 5hmC within the disease states and genotypes, such as Alzheimer's disease and mild cognitive impairment in C homozygotes, as well as dementia with Lewy bodies in heterozygotes. The Alzheimer's disease group in heterozygotes report higher median levels and range of 5hmC than 5mC methylation. Since blood and whole brain tissues in this project only come from Alzheimer's diagnosis samples, the high levels of 5hmC observed in heterozygous Alzheimer's diagnosis are reflected in the higher levels of 5hmC in blood and whole brain tissue heterozygotes at rs429358. This is presented in Figure 6.7.

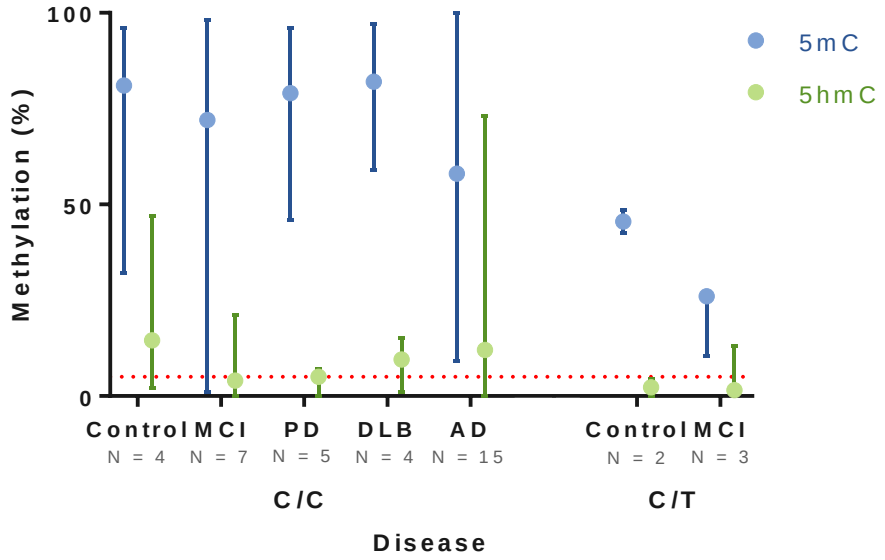


Figure 6.6 5mC and 5hmC methylation across disease groups and genotype at rs7412. Median percentages and range for 5mC or 5hmC methylation levels are shown in blue, and green, respectively. Control, mild cognitive impairment (MCI), Parkinson's Disease (PD), Dementia with Lewy bodies (DLB), and Alzheimer's Disease (AD) diagnoses groups are shown in the x-axis, with sample sizes (N) for each group. Genotype status is given below x-axis; either carrier (C/T) or non-carrier (C/C). 5% error margin for percentage of methylation is indicated by a red dotted line.

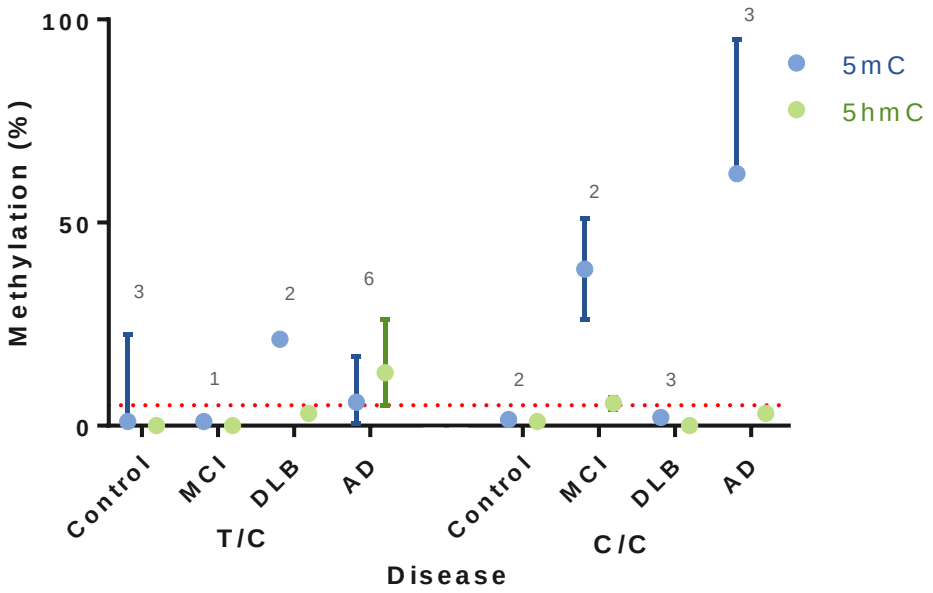


Figure 6.7 5mC and 5hmC methylation split by genotype, disease states at rs429358. Medians percentages and ranges for 5mC and 5hmC methylation are shown in blue and green, respectively. Control, mild cognitive impairment (MCI), Parkinson's disease (PD), dementia with Lewy bodies (DLB), and Alzheimer's disease (AD) diagnoses groups, and genotype are shown in the x-axis. Sample sizes (N) are shown above range bars. Genotype status is given below x-axis; either heterozygote (C/T) or CC homozygote (C/C). 5% error margin for percentage of methylation is indicated by a red dotted line.

6.3 5mC, 5hmC and 5fC/CaC methylation analyses

6.3.1 APOE rs7412 5mC, 5hmC, and 5fC/CaC methylation, across tissue types

Only C/C homozygotes could be used for the investigation of 5fC/CaC methylation. Of the 40 data points for which 5mC and 5hmC methylation data were available at rs7412, only 24 data points were both homozygous for the C allele and successfully analysed for 5fC/CaC methylation. While at least one of each tissue and disease state were present, many tissue groups did not have all disease states. This meant some of the effects observed may have been due to confounding effects of disease in specific brain regions or in blood. The samples used in 5fC/CaC methylation comparisons are shown in Table 6.5 and highlight overlaps in tissues/disease states.

5mC methylation was the most abundant form of methylation at the rs7412 site in all tissue types (Figure 6.8 Panel A & B). This is shown in the medians of 5mC in all tissue types being higher than that of 5hmC or 5fC/CaC methylation (Figure 6.8 Panel A). Whole blood reported a higher median level of 5hmC compared to other tissue types. Additionally, cerebellum reported higher ranges of 5mC and 5fC/CaC methylation. This may be due to cell populations within the cerebellum.

However, the ratio of average methylation changes in a tissue dependent manner where whole brain tissue showed a higher ratio of 5hmC to 5fC/CaC (23.9% : 3.3%), and cerebellum showed a lower ratio of 5hmC to 5fC/CaC (3% : 31.7%). Frontal cortex showed a higher absolute level of 5hmC and 5fC/CaC to 5mC, while blood showed a lower amount of 5hmC and 5fC. Cerebellum also shows the highest level of 5fC/CaC methylation (31.7%) and lowest 5hmC methylation (3%) compared to any other tissue.

Sample ID	Tissue	Disease
OPT 111	WB	ALZ
OPT 199	BL	ALZ
OPT 253	BL	ALZ
OPT 253	WB	ALZ
OPT 195	WB	ALZ
OPT 215	BL	ALZ
OPT 215	WB	ALZ
OPT 250	WB	ALZ
OPT 297	WB	ALZ
OPT 58	WB	ALZ
OPT 179	BL	ALZ
OPT 179	WB	ALZ
PD687	FC	PD
PD793	FC	PD
PD793	HIP	PD
PD793	CER	PD
PD120	FC	DLB
PD120	HIP	DLB
PD332	FC	DLB
PD715	FC	MCI
C045	CER	MCI
PD065	FC	MCI
PD065	CER	MCI
PDC112	HIP	CNTRL

Table 6.5 Samples used to assess *APOE* rs7412 5mC, 5hmC, and 5fC/CaC methylation across tissues and disease states. All samples used are C homozygous at rs7412. Tissues are shown where BL, WB, FC, HIP and CER, refer to blood, whole brain, frontal cortex, hippocampus, and cerebellum, respectively. Disease states are shown where AD, PD, DLB, MCI and CNTRL, refer to Alzheimer's disease, Parkinson's disease, Dementia with Lewy Bodies, Mild Cognitive Impairment, and Control samples, respectively.

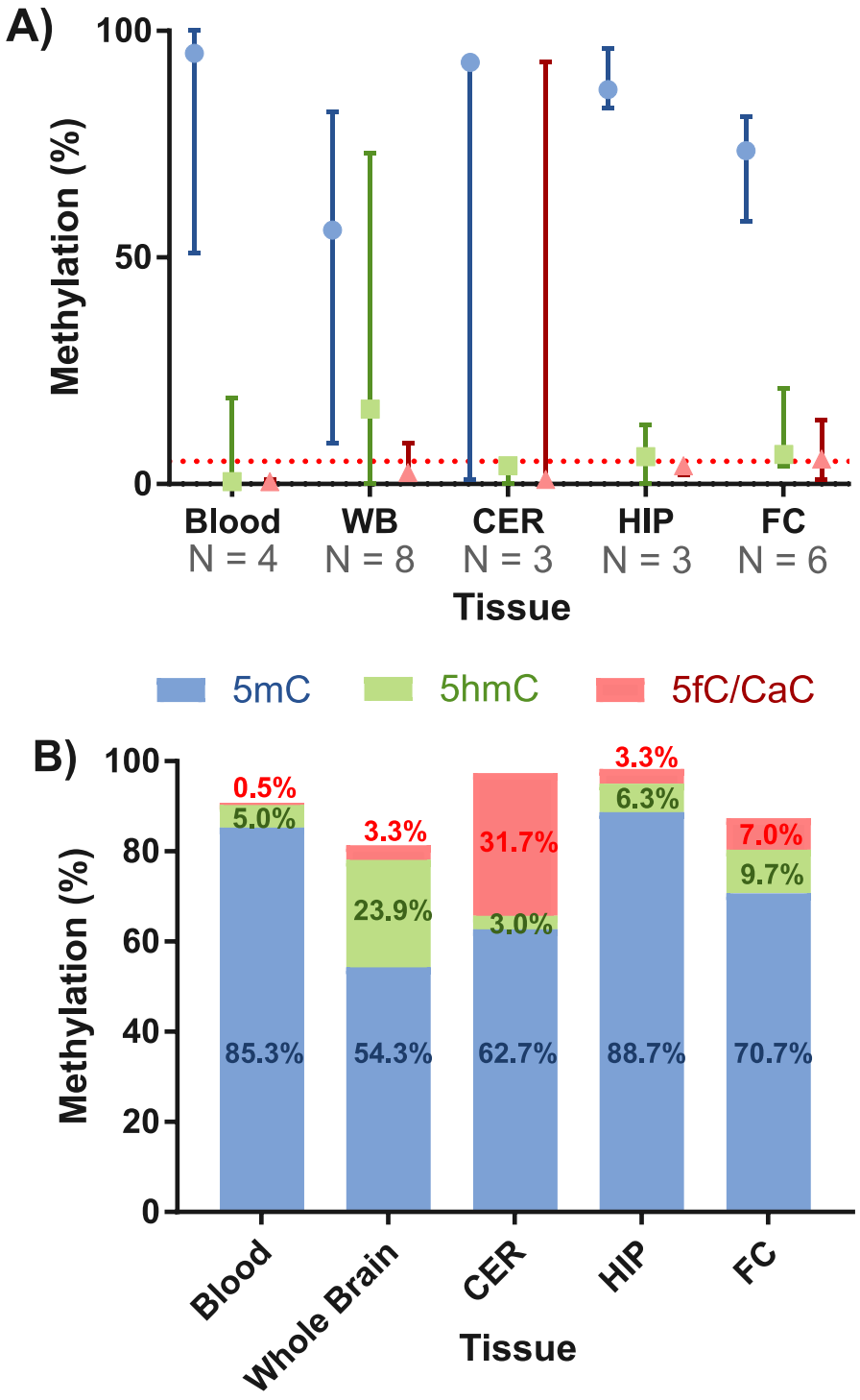


Figure 6.8 5mC, 5hmC and 5caC/5fC methylation levels across tissues at rs7412 A) Median percentages and ranges for 5mC, 5hmC, and 5fC/CaC methylation across tissue groups at rs7412, when using significant interactions resulting from inclusion of covariates. 5mC, 5hmC, and 5fC/CaC methylation are shown in blue, green, and red respectively. blood, whole brain (WB), cerebellum (CER), hippocampus (HIP), and frontal cortex (FC) tissues are shown on the x-axis, with sample sizes (N) for each group. 5% error margin for percentage of methylation is indicated by a red dotted line. B) Stacked bar-charts of the mean percentage methylation values for each type of methylation, within each tissue. Methylation and tissue types are represented as in Panel A. For each methylation type, a percentage is shown on each tissue bar showing the mean methylation level (rounded to 1 d.p.) in corresponding colours.

6.3.2 APOE rs429358 5mC, 5hmC and 5fC/CaC methylation, and tissue types

Because the major allele at rs429358 is Thymine, homozygous C genotypes are far fewer in frequency. Of the 22 data points for which 5mC and 5hmC methylation data were available, only 8 data points were both homozygous for the C allele and successfully analysed for 5fC/CaC methylation. No Parkinson's disease samples were present in the cohort and some disease and tissue groups (control, blood, hippocampus) had a data point size of N = 1. The samples used in 5fC/CaC methylation analysis at rs429358 are shown in Table 6.6 and highlight overlaps between tissues and disease states.

Although data point numbers were low, Figure 6.9 illustrates the wide variation observed in the level of methylation between tissue types. Both blood (median of 95.5%) and whole brain (61.5%) tissue showed high 5mC levels, while cerebellum, hippocampus and frontal cortex all showed 5fC/CaC to be abundant (98%, 98%, and 81.5%, respectively). Moreover, whole brain showed a large amount of unmethylated cytosines (34%) comparative to other tissue groups. Compared to previous 5mC and 5hmC analyses at rs429358, uncharacteristically low 5hmC levels were observed across all tissue types for homozygous C genotypes (the highest level at 3.5%), with cerebellum and hippocampus showing no 5hmC methylation or unmodified cytosines at all. While 5fC was a dominant form of methylation in brain regions when taking the mean percentage (81.5% to 98%), this may in part be due to the low sample numbers available at *APOE* rs429358.

Sample ID	Tissue	Disease
OPT 215	BL	ALZ
OPT 215	WB	ALZ
OPT 179	WB	ALZ
PD115	FC	DLB
PD120	HIP	DLB
PD120	CER	DLB
PD065	FC	MCI
PD092	CER	CNTRL

Table 6.6 Samples used to assess *APOE* rs429358 5mC, 5hmC, and 5fC/CaC methylation across tissues and disease states. All samples used are C homozygous at rs429358. Tissues are shown where BL, WB, FC, HIP and CER, refer to blood, whole brain, frontal cortex, hippocampus, and cerebellum, respectively. Disease states are shown where AD, DLB, MCI and CNTRL, refer to Alzheimer's disease, Dementia with Lewy Bodies, Mild Cognitive Impairment, and Control samples, respectively.

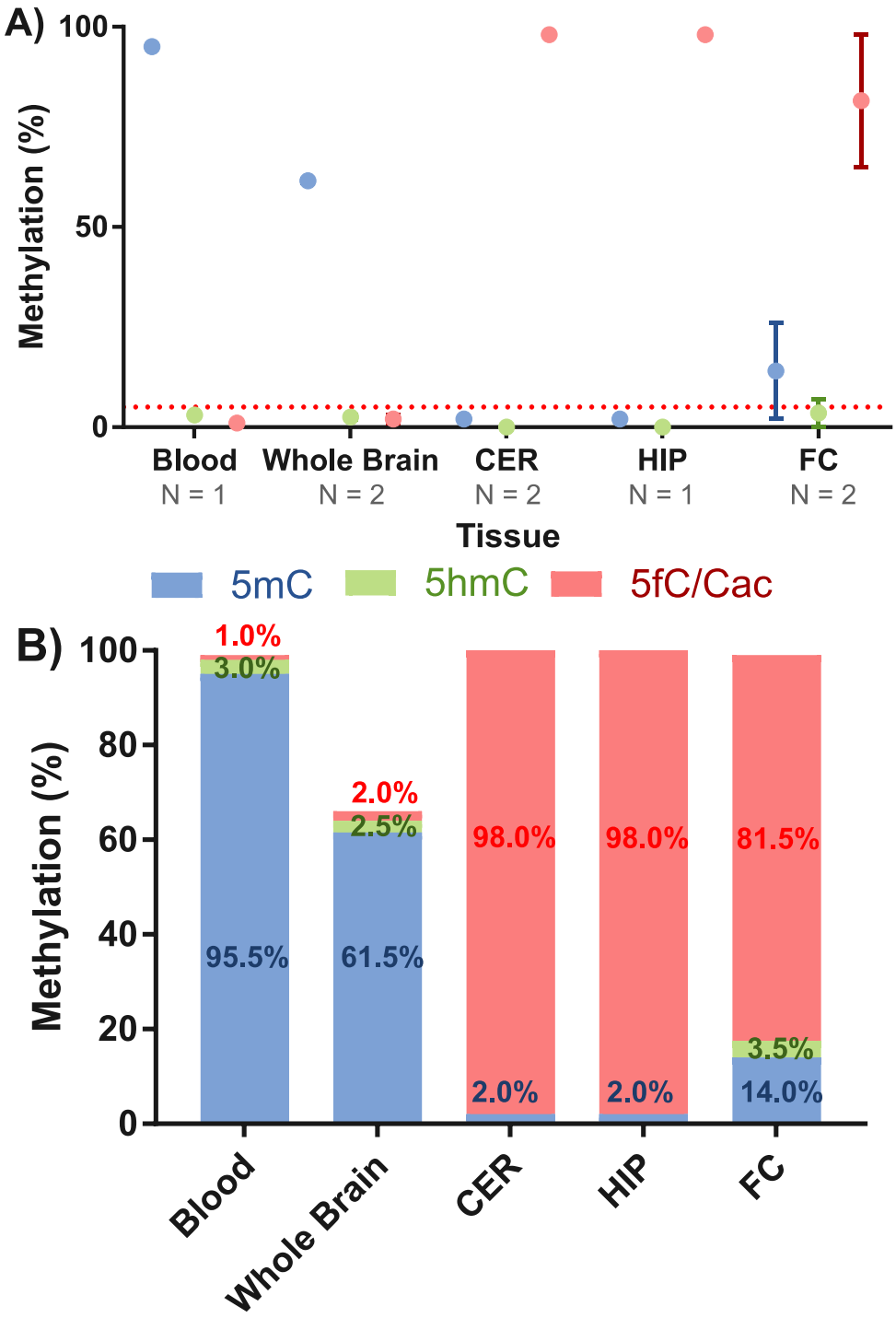


Figure 6.9 5mC, 5hmC and 5caC/5fC methylation levels across tissues at rs429358 A) Median percentages and ranges for 5mC, 5hmC, and 5fC/CaC methylation across tissue groups at rs429358, when using significant interactions resulting from inclusion of covariates. 5mC, 5hmC, 5fC/CaC methylation and unmodified cytosines are shown in blue, green, red, and grey, respectively. blood, whole brain (WB), cerebellum (CER), hippocampus (HIP), and frontal cortex (FC) tissues are shown on the x-axis, with sample sizes (N) for each group. 5% error margin for percentage of methylation is indicated by a red dotted line. B) Stacked bar-charts of the mean percentage methylation values for each type of methylation, within each tissue. Methylation and tissue types are represented as in Panel A. For each methylation type, a percentage is shown on each tissue bar showing the mean methylation level (rounded to 1 d.p.) in corresponding colours.

6.3.3 APOE rs7412 5mC, 5hmC, and 5fC/CaC methylation across disease states

The same samples used in the analysis of 5fC/CaC methylation at APOE rs7412 across tissues were used in the analysis of disease states (Table 6.5). Similar to the analysis of methylation across tissues, the most abundant form of methylation present across disease states was 5mC methylation. This is reflected in the high medians of 5mC for each disease state (Figure 6.10 Panel A).

Both 5hmC and 5fC/CaC methylation were comparatively lower in median abundance than that of 5mC methylation across all disease states. Some median 5hmC and 5fC/CaC methylation levels were observed to be close to or below the 5% threshold of methylation. MCI samples reported high median and ranges of 5fC/CaC methylation, compared to that of other samples. This may be a reflection of the tissue types of the MCI samples which are restricted to frontal cortex and cerebellum, and cerebellum has shown to have high ranges of 5fC/CaC (Figure 6.8 Panel A). This may also be true for the Alzheimer's disease samples reporting a high range of 5hmC methylation (Figure 6.10 Panel A) as the whole brain tissue exclusively found in Alzheimer's disease samples has previously reported high ranges of 5hmC (Figure 6.8 Panel A).

Overall, the mean levels of 5mC, 5hmC and 5fC/CaC methylation across disease states was comparable with that of methylation across tissues (Figure 6.8 Panel B, and Figure 6.10 Panel B). This is highlighted mainly as high mean 5mC methylation levels compared to other forms of methylation.

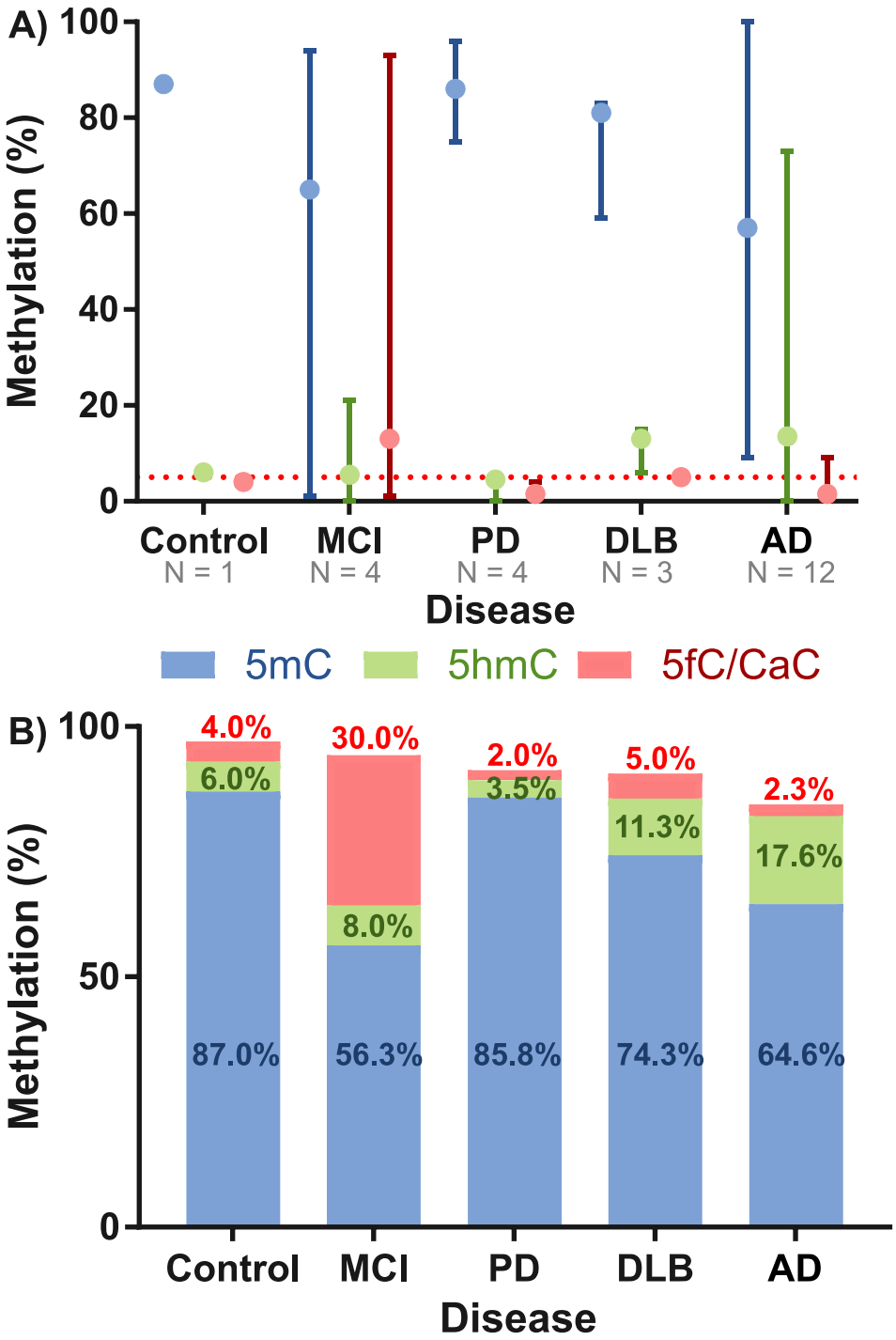


Figure 6.10 5mC, 5hmC, 5fC/CaC methylation type in disease states at rs7412 A) Median percentages and ranges for 5mC, 5hmC, and 5fC/CaC methylation levels in disease groups in homozygous C genotypes when using significant interactions resulting from inclusion of covariates. 5mC, 5hmC, 5fC/CaC methylation and unmodified cytosines are shown in blue, green, red, and grey, respectively. Disease states are controls, mild cognitive impairment (MCI), Parkinson's disease (PD), dementia with Lewy bodies (DLB), and Alzheimer's disease (AD) and are shown on the x-axis with sample sizes (N) for each group. 5% error margin for percentage of methylation is indicated by a red dotted line. B) Stacked Bar charts of the mean percentage methylation values for each type of methylation in each disease state. Methylation and disease states are as represented as in Panel A. For each methylation type, a percentage is shown on each tissue bar showing the mean methylation level (rounded to 1 d.p.) in corresponding colours.

6.3.4 APOE rs429358 5mC, 5hmC, and 5fC/CaC methylation, across disease states

The samples available for 5mC, 5hmC and 5fC/CaC methylation at the *APOE* rs429358 variant across disease states are shown in Table 6.6. As with analysis of methylation in tissues, analysis of 5fC/CaC methylation in disease states used samples which may not have all available tissues. For instance. While DLB diagnosis samples have been taken from frontal cortex, hippocampus and cerebellum, the MCI diagnosis was taken from the frontal cortex, and the control sample was taken from the cerebellum. No Parkinson's disease samples were available for analysis.

Analysis of methylation at the rs429358 variant across diseases revealed similar results seen in the analysis of tissue types. The Imperial samples (Dementia with Lewy bodies, Mild Cognitive Impairment, and control diagnoses) all reported high median levels of 5fC/CaC methylation compared to other methylation types (Figure 6.11), and this is reflected in the high levels of 5fC/CaC methylation observed in the brain regions in tissue analysis (Figure 6.9). However, the OPTIMA samples (Alzheimer's diagnosis) reported the canonically high median levels of 5mC (Figure 6.11) observed in the blood and brain tissue samples at rs429358 (Figure 6.9).

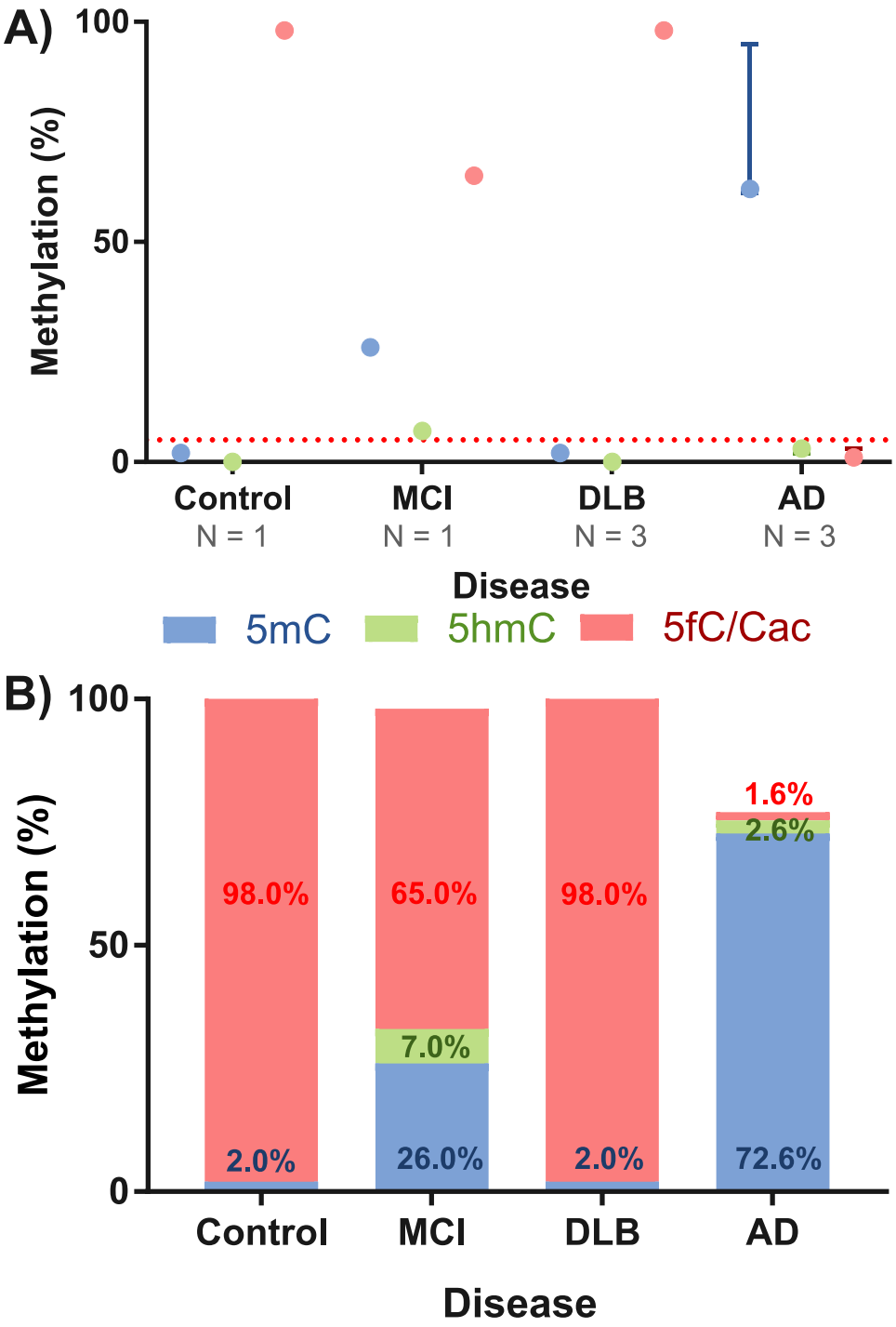


Figure 6.11 5mC, 5hmC, 5fC/CaC methylation in disease states at rs429358 A) Median percentages and ranges for 5mC, 5hmC, and 5fC/CaC methylation levels in disease groups in homozygous C genotypes when using significant interactions resulting from inclusion of covariates. 5mC, 5hmC, 5fC/CaC methylation and unmodified cytosines are shown in blue, green, red, and grey, respectively. Disease states are controls, mild cognitive impairment (MCI), Parkinson's disease (PD), dementia with Lewy bodies (DLB), and Alzheimer's disease (AD) and are shown on the x-axis with sample sizes (N) for each group. B) Stacked Bar charts of the mean percentage methylation values for each type of methylation in each disease state. Methylation and disease states are as represented as in Panel A. For each methylation type, a percentage is shown on each tissue bar showing the mean methylation level (rounded to 1 d.p.) in corresponding colours. 5% error margin for percentage of methylation is indicated by a red dotted line.

6.4 Comparison of ε Isoform 5mC and 5hmC methylation levels

Following individual site analyses, methylation was compared across APOE ε isoforms by combining the 5mC and 5hmC methylation levels for both the *APOE* rs7412 and rs429358 epialleles. Because multiple genotypes are required to analyse different isoforms, 5fC/CaC methylation analyses could not be investigated. Unfortunately, methylation could not be quantified at ε2/ε3 isoforms due to failed pyrosequencing in PD898 and PDC007 samples.

6.4.1 Methylation levels per APOE ε isoform

24 data points were eligible for comparison across both epialleles for 5mC and 5hmC methylation. These samples were then grouped by methylation type and *APOE* ε isoforms derived. `

As shown in figure 6.12 Panel A, there is a disproportionate level of methylation at some sites compared to the number of cytosines present. For instance, in ε2/ε4 and ε3/ε4 isoforms rs429358 is heterozygous. But while ε2/ε4 rs429358 shows both methylation types below 5%, ε3/ε4 rs429358 shows the highest level of 5hmC methylation at rs429358 and the highest range of 5mC. Once again this may be indicative of underlying factors such as disease or tissue or this may reflect sample sizes (N = 2 in ε2/ε4, N = 9 in ε3/ε4).

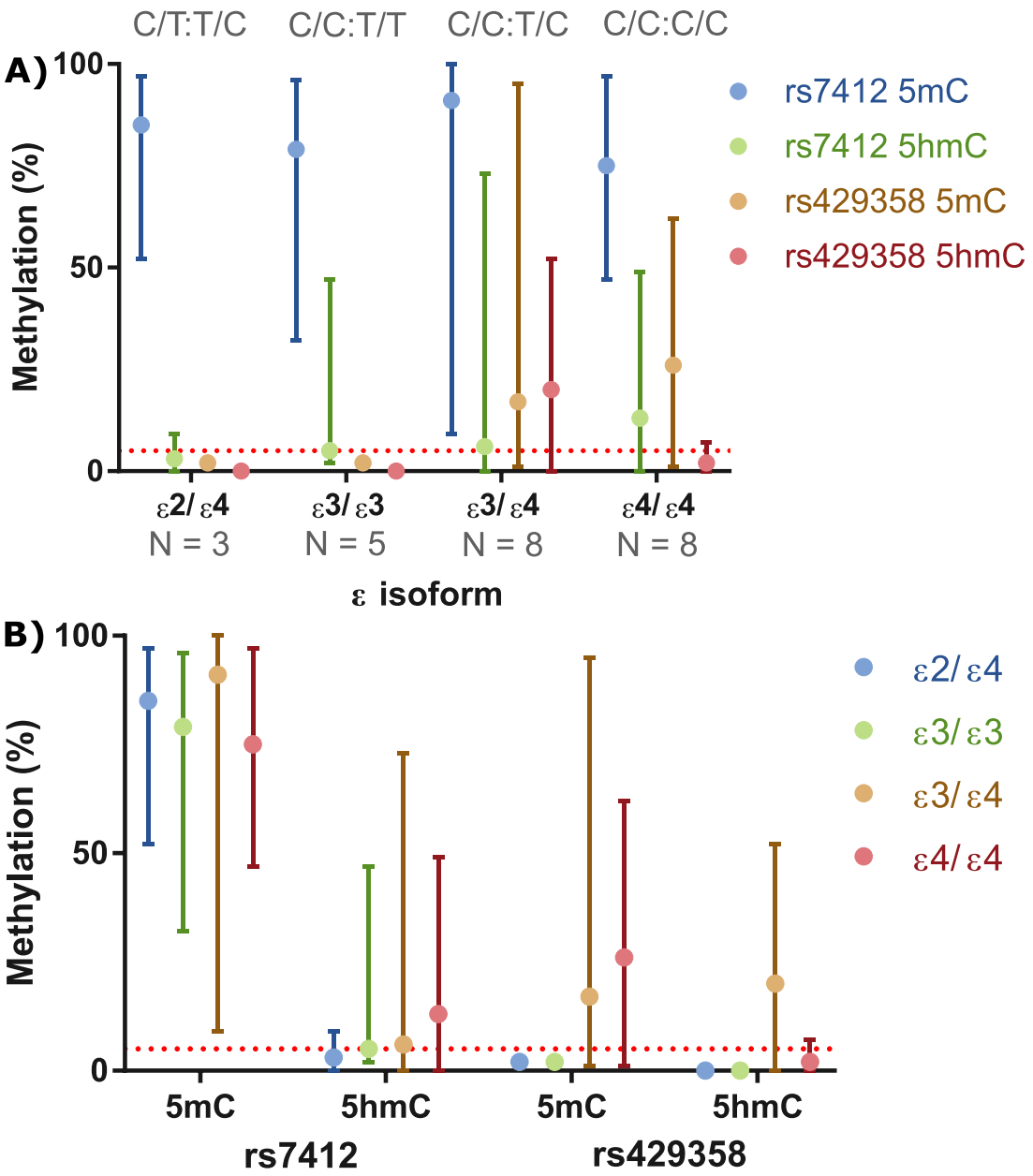


Figure 6.12 5mC and 5hmC methylation at APOE rs7412 and rs429358 epialleles and ε isoforms. A) median percentages and ranges of 5mC and 5hmC methylation levels at APOE rs7412 (blue and green, respectively) and rs429358 (yellow and red, respectively). ε isoforms are shown on the x-axis increasing in deleterious effect from left to right. Data point cohort sizes (N) are shown below the ε isoforms. The number of cytosines available for methylation at each ε isoform is displayed above Panel A, displayed as [genotype of rs7412]:[genotype of rs429358]. B) transposed mean percentages and ranges of Panel A across the ε isoforms at the rs7412 and rs429358 epialleles. Isoforms are shown from the least damaging isoform (ε2/ε4) in blue, the standard isoform ε3/ε3 in green, the deleterious isoform ε3/ε4 in yellow, and the most deleterious isoform (ε4/ε4) in pink/red. Both panels show the 5% margin of error for pyrosequencing as a red dotted line.

These differences between isoforms and methylation types can be observed more clearly by transposing the methylation data (Figure 6.12 Panel B). As there is a change from rs7412 heterozygous (ε2/ε4) to homozygous C (ε3/ε3, ε3/ε4, ε4/ε4) isoforms, the methylation levels at rs7412 would be expected to increase with the number of cytosines. However, the highest range and second highest median of 5mC methylation observed at rs7412 was in the ε2/ε4 where rs7412 is heterozygous. Unusually the highest level of 5hmC methylation at both epialleles was in the ε3/ε4 isoforms, which is further supported by the same isoform showing the lowest values of 5mC methylation. 5hmC methylation at rs429358 did not show typical increases of methylation with genotype and cytosine availability, however. This unpredictable methylation may further support the alteration of methylation profiles in a disease or tissue specific manner.

6.4.2 5mC and 5hmC methylation split by disease state and ε isoform

Methylation data were combined across epialleles and samples grouped based on disease states and ε isoform (Figure 6.13). Data suggest a trend whereby 5mC is the more abundant form of methylation regardless of disease state and ε isoform. Moreover, higher levels of 5hmC methylation typically occurred at heterozygous ε isoforms. Within the MCI and AD disease groups, the ε4/ε4 isoforms showed the highest median levels of 5mC methylation, which would be expected as the number of cytosines available for methylation at rs7412 and rs429358 increases with isoform deleteriousness. This trend is not observed in the control and DLB samples however, which report higher levels of 5mC in less deleterious isoforms.

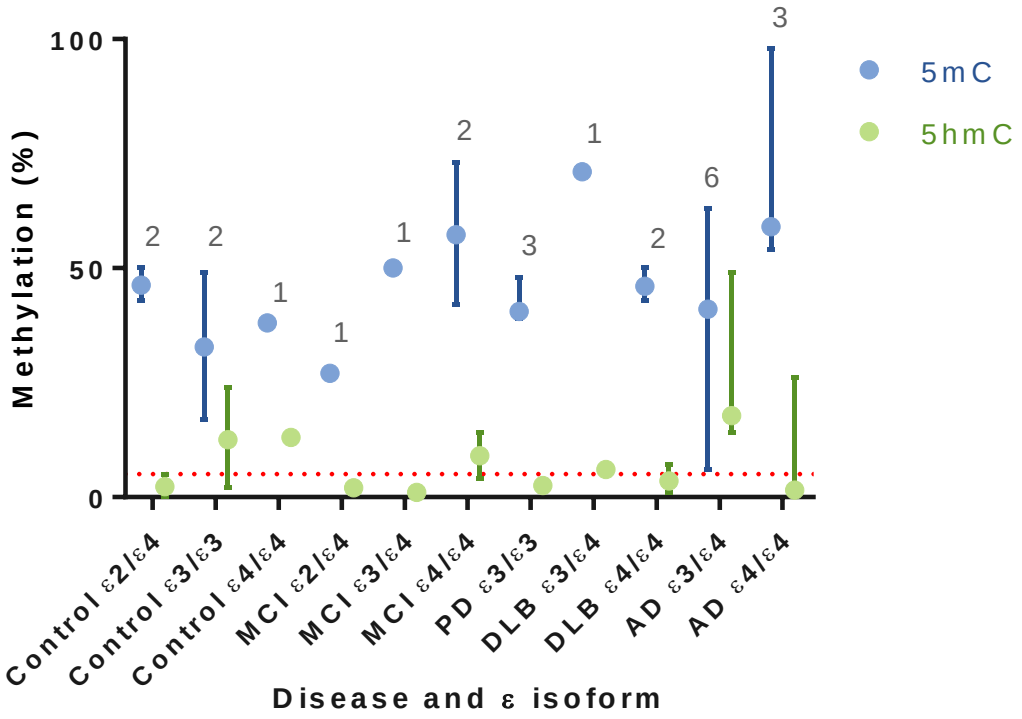


Figure 6.13 5mC and 5hmC methylation at rs7412 and rs429358 sites compared against ε isoform split by disease states. Medians and ranges of 5mC and 5hmC methylation, are presented in blue and green respectively. Disease states are controls, mild cognitive impairment (MCI), Parkinson's disease (PD), dementia with Lewy bodies (DLB), and Alzheimer's disease (AD). ε isoforms are ε2/ε4, ε3/ε3, ε3/ε4, and ε4/ε4. sample numbers (N) have been shown as grey numbers above disease isoform methylation data. 5% error threshold from pyrosequencing is shown as a red dotted line.

6.4.3 Combined 5mC or 5hmC methylation across both rs7412 and rs429358 epialleles

5mC or 5hmC methylation percentages were combined for each sample across both the *APOE* rs7412 and rs429358 epialleles and analysed together (Figure 6.14). This showed trends in 5mC and 5hmC methylation across both sites. 5mC was the most abundant form of methylation across all ε isoforms. There is a trend towards methylation of 5mC and 5hmC becoming more abundant with the increased risk isoforms (e3/e4 and e4/e4). Interestingly the ε3/ε4 isoforms reported the highest levels of 5mC and 5hmC methylation, despite having less cytosines compared to the ε4/ε4 isoforms. The disproportionate representation of 5mC at sites with less cytosines may indicate the amount of methylation observed is influenced by other factors such as disease, tissue, age.

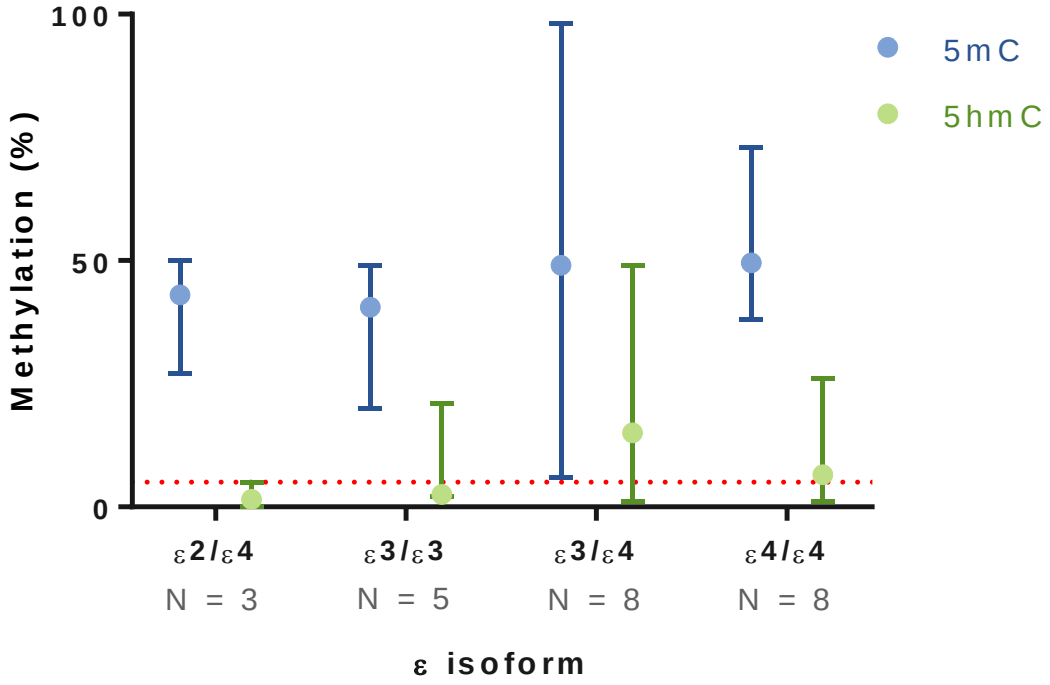


Figure 6.14 Combined median and ranges of 5mC or 5hmC methylation from both rs7412 and rs429358 epiallele sites and split by ε isoform. 5mC and 5hmC are presented in blue and green, respectively. Data point numbers (N) are shown below the isoforms. 5% error threshold from pyrosequencing is shown as a red dotted line.

6.5 Discussion

6.5.1 ε isoform analyses

Alzheimer's disease has been associated with the ε4 risk allele in a dose dependent manner (Corder et al., 1993; Strittmatter et al., 1993) and both ε3 and ε2 isoforms are known to be under-represented in Alzheimer's disease populations (Strittmatter et al., 1993). This is reflected in the current study's Alzheimer's disease samples. However, even 3.9% of protective ε2 homozygotes develop Alzheimer's disease (Benjamin et al., 1994; West, Rebeck, & Hyman, 1994) and other factors are influencing the risk of developing Alzheimer's disease in addition to ε isoform mechanisms.

Other cognitive diseases have shown mixed associations for disease risk with ε isoforms. Risk for dementia with Lewy bodies has been associated with ε4 isoform (Lane et al., 2009; Tsuang et al., 2013) and this is seen in the high number of DLB ε4/ε4 samples in this cohort. However, there has been no evidence linking the ε isoforms and Parkinson's disease (Federoff, Jimenez-Rolando, Nalls, & Singleton, 2012). ε4 isoform levels are also higher in the blood plasma of ε3/ε4 MCI samples than ε3/ε4 AD or control diagnoses (Fukumoto et al., 2003) and MCI ε4 carriers have been shown to have a faster rate of cognitive decline, memory deterioration, and brain atrophy (Farlow et al., 2004; Morgen et al., 2013).

6.5.2 5mC and 5hmC methylation, APOE risk genotype, tissue and disease

It is known that 5mC and 5hmC methylation may contribute to disease pathology e.g. Cancer, in a region and cell type specific manner (Phipps, Vickers, Taberlay, & Woodhouse, 2016). For instance, squamous cell carcinomas have been shown to have specific methylation patterns across the genome with some loci being hyper or hypomethylated in tumours (Rauch et al., 2008). Moreover while astrocytes typically have decreased 5mC and 5hmC methylation in AD, microglia do not appear to be differentially methylated compared to the microglia of controls (Phipps et al., 2016). 5mC and 5hmC methylation has also been shown to increase across the genome in

frontal cortex brain tissue in individuals with Alzheimer's disease or bipolar disorder (Coppieters et al., 2014; Rao et al., 2012). Of particular relevance, two studies have shown that APOE rs7412 C homozygotes generate ε3/ε3 and ε3/ε4 isoforms which increase the risk for disease through transcriptional regulation of *APOE* and facilitate higher levels of methylation, determined through bisulphite sequencing, across APOE CpG islands (CGIs) associated with an increasing number of available CpGs (Foraker et al., 2015; C. E. Yu et al., 2013).

This may be allele specific methylation (ASM), whereby different genotypes at a site determine the amount of methylation possible, and this is reflected in this study's combined epiallele analysis of 5mC and 5hmC (Figure 6.14) where as ε isoforms become more deleterious, methylation increases. The 3' CpG island exon containing rs7412 and rs429358 of *APOE* has been shown to exhibit transcriptional enhancer/silencer activity, and alters the amount and type of APOE transcripts expressed in a cell type, in a methylation and ε isoform specific manner (C. E. Yu et al., 2013). I propose that in this way, methylation at rs7412 and rs429358 may alter the transcriptional expression of *APOE*, and this may be linked to the frequency of cytosines available for methylation at these sites, genotype at these variants (ASM), and disease state of the samples. The mechanisms by which this may occur are discussed in this and the next chapter.

Allele-specific methylation across the human genome has been observed in other diseases and cell types, and is typically cis-acting (Meaburn, Schalkwyk, & Mill, 2010). This ASM may therefore be occurring in a disease phenotype specific manner, indicating that as the total number of cytosines increases, ε isoforms become more deleterious. Potabattula et. al. reported that that cis-acting ASM in imprinting was largely dependent on underlying variant haplotype, despite the small effect sizes (Potabattula et al., 2018). In this way, knowledge of a variant genotype, and allele-specific methylation levels may be indicative of disease states and may predict transcript expression in these states.

Similar to rs7412, rs429358 mean 5mC levels broadly appeared higher than 5hmC across genotypes (Figure 6.3). This was to be expected as 5mC methylation is more prominent at CpG sites compared to other methylation forms (Lister et al., 2009; Z. D. Smith & Meissner, 2013) and this is reflected in higher 5mC than 5hmC in rs7412 C homozygotes (Figure 6.2) and in blood and cerebellum of rs7412 C homozygotes (Figure 6.4). However, 5hmC methylation was highest in the blood and whole brain DNA of rs429358 heterozygotes (Figure 6.5). It is known that 5hmC may have up to 20-fold enrichment or depletion depending on tissue/cell types, as well as other factors (Coppieters et al., 2014; Nestor et al., 2012; Rao et al., 2012). Therefore, the high level of 5hmC observed in heterozygotes may be due to cell or tissue types. The overall change in total 5mC and 5hmC methylation may also be due to ASM (Meaburn et al., 2010).

Whole brain DNA reported the most similar levels of 5mC and 5hmC methylation at both rs7412 C homozygotes and rs429358 heterozygotes, which may indicate that the differences are preserved across multiple individuals and tissue types. As whole brain DNA is mixed DNA from multiple brain regions, the enrichment for 5hmC may be due to specific brain regions (S.-G. Jin et al., 2011). 5hmC methylation enrichment is also dependent on the genomic region, e.g. where gene body 5hmC is more common and typically is cis-acting in altering expression (Zhao et al., 2017). These differentially hydroxymethylated regions may also change with pathology (Zhao et al., 2017). Alzheimer's disease data have been shown to be differentially methylated depending on gene location and brain region, and this may in part be due to Alzheimer's disease present in these samples (Semick et al., 2019). This is reflected in more similar levels of 5hmC in Alzheimer's diagnoses at rs429358 heterozygotes (Figure 6.7). This further supports tissue or disease phenotype specific methylation profiles (Phipps et al., 2016; Rao et al., 2012).

As stated above, while possession of the ε4/ε4 phenotype is the largest genetic risk factor for Alzheimer's disease, it confers under 50% chance of developing late onset Alzheimer's disease (Liddell, Lovestone, & Owen, 2001). It is known that other factors, such as neuroinflammation and lipid metabolism also contribute to the risk

and severity of Alzheimer's disease (Y. S. Kim & Joh, 2006; Q. Liu & Zhang, 2014; McGeer & Rogers, 1992). Alternative mechanisms, such as allele specific methylation or other cis-acting regulatory elements (CRE), may also influence the likelihood of developing late onset Alzheimer's when the ε4/ε4 isoform is present (Tycko, 2010). The level of 5mC and 5hmC methylation correlate with hallmarks of Alzheimer's disease such as tau protein and amyloid-β plaques (Coppieters et al., 2014). 5mC methylation was observed as being higher than 5hmC methylation in heterozygous rs7412 Alzheimer's disease patients (Figure 6.6). However, whether this is a reflection of pathology or not is unknown.

rs7412 C homozygote Parkinson's disease data also showed higher levels of 5mC methylation compared to 5hmC methylation. Studies have shown that there is differential methylation between controls and Parkinson's disease diagnosis in a gene specific manner, where the α-synuclein (*SNCA*) gene is hypomethylated (Matsumoto et al., 2010; Miranda-Morales et al., 2017). This in turn reduces transcriptional repression of α-synuclein and contributes to the Parkinsonian phenotype (Matsumoto et al., 2010). While this is a good example of an epigenetic CRE, research has also shown that APOE ε isoform and blood dosage is not significantly different in Parkinson's disease, and it is unlikely that altered regulation of APOE isoform expression is contributing greatly to the Parkinsonian phenotype (Federoff et al., 2012). Therefore, the heightened level of 5mC compared to 5hmC in these groups may be a reflection of brain regions, genotype, or other factors.

Some observations in this study, such as methylation in dementia with Lewy bodies samples at rs429358 C homozygotes and heterozygotes, showed the same level of methylation across all samples despite having cohort sizes greater than one. These samples are most likely derived from 3 or 4 individuals across multiple tissue types which may suggest an inter-tissue pattern of methylation in dementia with Lewy bodies phenotypes. Studies have shown that dementia with Lewy bodies disease is associated with reduced 5mC methylation across *APOE* in a tissues specific manner between frontal lobe and cerebellum (Tulloch et al., 2018) and these findings would support tissue specific levels of 5mC and 5hmC across different brain tissues.

Similarly, the significantly higher level of 5mC in C homozygous DLB patients at rs7412 may be due to the 5mC methylation being a product of multiple regions.

Mild cognitive impairment reported variable 5mC methylation across both epialleles and genotypes. At rs7412 MCI 5mC methylation was high at C homozygotes but with a large range, while heterozygotes showed lower but less variable levels of methylation. rs429358 MCI data showed no or low 5mC methylation in heterozygotes and homozygotes, respectively. Current literature suggests 5mC methylation levels decrease globally within the ageing brain (Ciccarone et al., 2018; Wilson & Jones, 1983; Xiangru Xu, 2015). Changes in methylation alter cognition in rat models (Lubin et al., 2008; Miller & Sweatt, 2007) and different levels of 5mC in MCI diagnoses may be a result of epigenetic dysregulation in MCI phenotype, or multiple other factors.

6.5.3 5mC, 5hmC, and 5fC/CaC methylation, tissue and disease

In this study we observed higher 5mC methylation than 5fC/CaC methylation in DNA from whole brain, hippocampus and frontal cortex at *APOE* rs7412 (Figure 6.8). As 5fC/CaC are expected in 2-4 orders of magnitude less amounts compared to 5hmC in human brain, it is surprising that much 5fC/CaC methylation is detectable (S. Liu et al., 2013). Furthermore, our results support low levels of 5fC/CaC and site-specific levels of methylation, where 5fC/CaC appears enriched at rs429358 compared to rs7412, as demonstrated by higher levels of 5fC/CaC in cerebellum compared to 5mC at rs429358. Additionally, results support the bimodal distribution of 5fC/CaC previously seen in the *BDNF* analysis and supported by literature, where methylation is typically around 0% or 100% with few intermediate values (C. Zhang et al., 2016).

A heightened level of 5fC/CaC methylation in the cerebellum at *APOE* rs7412 and *APOE* rs429358 was also observed (Figures 6.7 & 6.8). Studies have shown by immunohistochemistry (IHC) methods that 5fC and 5CaC methylation is typically abundant in the cerebellum in Alzheimer's patients compared to other regions like hippocampus where 5fC/CaC methylation is low (Bradley-Whitman & Lovell, 2013).

Cerebral Purkinje cells are also known to increase or decrease 5fC/CaC methylation during differentiation of cerebral cells, so the observed methylation levels may be indicative of cell differentiation (Wheldon et al., 2014; F. C. Zhou et al., 2016).

High levels of 5fC/CaC in the hippocampus and frontal cortex have been reported previously in mouse models and in disease groups (Wheldon et al., 2014; F. C. Zhou et al., 2016) and this is observationally reflected in this study's 5fC/CaC analysis of rs429358, despite not being able to determine a significant difference due to low sample numbers. Immunofluorescent staining of mouse brain sections revealed that 5fC and 5CaC methylation are enriched in NeuN+ cell types of the cortex (Tao et al., 2018). Further, the hippocampus was enriched in 5fC and 5CaC compared to cortex and cerebellum of adult mice, and that 5CaC levels decrease with age in the development of neuronal stem cells (Tao et al., 2018). Furthermore 5CaC has been observed to be highly abundant in immunostaining of human hippocampus FFPE sections, and this was especially true for regions which had high population of stem cells (Haig, Flitton, Ruzov, & Knight, n.d.). The findings of this study reflect methylation patterns in a cell population and variant site-specific manner and support enrichment of 5fC/CaC methylation in the hippocampus.

While it is known that 5fC/CaC methylation is enriched in promoter regions and repeat elements, little is known as to the mechanisms they may be involved with in cognitive diseases outside of healthy tissue or Alzheimer's disease (Bradley-Whitman & Lovell, 2013; Shen et al., 2013; C. X. Song et al., 2013). Our results indicated similar 5fC/CaC methylation as seen in tissue analysis where 5fC/CaC was lower than 5mC at rs7412 in Parkinson's disease, Alzheimer's disease, and dementia with Lewy bodies. Both dementia with Lewy bodies and Parkinson's disease are α-synucleinopathies and it is interesting that the trends in methylation observed in these diseases at the rs7412 site were similar in pattern to Alzheimer's disease in that 5mC was higher than 5fC/CaC. Furthermore, both Alzheimer's disease and dementia with Lewy bodies disease show comorbidity of many overlapping symptoms (Chung et al., 2015). As these diseases are comorbid in many cases of dementia, this may show a similar methylation profile in these diseases (Bennett & Gordon, 2018). These results may

indicate a unique methylation profile for Alzheimer's disease that may be shared with Parkinson's disease and dementia with Lewy bodies disease, however as the trends in methylation are so similar to the trends observed in control samples, little can be determined at this time. Further analysis with more samples to determine smaller changes in methylation between control and disease samples are required.

The *APOE* rs429358 analysis of 5fC/CaC methylation across disease groups indicated that 5fC/CaC methylation was higher than 5mC and 5hmC in dementia with Lewy bodies diagnoses at rs429358. Methylation as determined by bisulphite sequencing has been shown to be significantly depleted in APOE across two loci which have a high frequency of CpG sites (known as CpG islands I and II) in patients with dementia with Lewy bodies diagnosis (Tulloch et al., 2018). However other genes were determined in bisulphite sequencing to be hypermethylated in dementia with Lewy bodies patients, such as the fibroblast growth factor receptor 3 (*FGFR3*) gene, and this hypermethylation is suggested to be directly correlated to the severity of α-synuclein pathology (Tsuchida et al., 2018). These results suggest that dementia with Lewy bodies shows a disease specific methylation profile that may differ at base resolution but how the 5fC/CaC enrichment of *APOE* rs429358 may be contributing to this pathology is unknown.

6.5.4 5mC, 5hmC methylation and ε Isoform analysis

Much of the variation observed between the APOE epialleles and methylation types may be attributed to ε isoforms determined by the genotypes at the rs7412 and rs429358 epialleles. Our findings exhibited observationally higher levels of 5mC methylation compared to 5hmC methylation in ε3/ε3 samples at both rs429358 and rs7412, and this may be due to ASM.

Our findings indicate that the epialleles are disproportionately methylated when compared to the amount of cytosines available for methylation. That is, total methylation does not appear to increase linearly with cytosine availability, as if methylation increased linearly with cytosine availability, the highest levels of

methylation would be expected at C homozygotes (as they have the largest number of sites that may be methylated), however this was not the case for all types of methylation. 5hmC methylation is detected at its highest at heterozygotes (only one site that may be methylated), and this may be an example of ASM. This is further supported by very different levels of 5fC/CaC methylation in control samples between both sites. This may show that methylation is not necessarily a product of genotype or CpG availability and may occur in distinctive profiles due to other factors. Further testing with larger numbers would determine if this trend is significant, or if the observed trend is due to variability in the samples used.

The site-specific differences in methylation may be explained by methylation quantitative trait loci (meQTL). That is the concept that genotypes at specific loci can be associated with different methylation patterns at sites in linkage disequilibrium (A. K. Smith et al., 2014). Such meQTLs showed the largest amount of overlap in brain regions, and genetic variation in APOE was associated with differential methylation in the 5-Hydroxytryptamine Receptor 6 (*HTR6*) gene in the frontal cortex and temporal cortex (A. K. Smith et al., 2014). *APOE* rs7412 has been shown to be a quantitative trait locus (QTL) associated with Alzheimer's disease that is co-inherited with Apolipoprotein C1 (*APOC1*) variant rs4420638 (Bertram, McQueen, Mullin, Blacker, & Tanzi, 2007). This means the G allele of rs4420638 is inherited with the rs7412 deleterious ε4 isoform allele and would show linkage disequilibrium with *APOE*. There may be other variant sites worth investigating such as *APOC1* rs4420638, that are affecting not only the genotype but level of methylation. However, whether the methylation observed is causative in disease phenotypes is yet to be discerned.

It is known that Alzheimer's patients show a reduction in the amount of available APOE in the cerebrospinal fluid (CSF) which correlates with deleterious ε isoforms (Lehtimäki et al., 1995; Tai et al., 2013; Talwar et al., 2016; Tapiola et al., 2000). However, whether this reduction in APOE is due to a reduction in transcriptional repression or binding of APOE protein to amyloid-β plaques is still debatable (Strittmatter et al., 1993; Tai et al., 2013). Increased total 5mC methylation observed

at rs7412 and rs429358 in the ε4 isoform individuals as observed in the present study may therefore be acting as a mediator for transcriptional expression and result in reduced APOE CSF levels typical of Alzheimer's disease (Coppieters et al., 2014). Differentially methylated regions have already been identified in Alzheimer's disease *APOE* (Foraker et al., 2015), and as the ε isoforms, rs7412, and rs429358 epialleles are heavily associated with the development of late onset Alzheimer's disease, identification of disease and isoform specific methylation profiles may be powerful as a diagnostic measure.

6.5.5 Limitations and future research

This study was limited primarily by small sample sizes for some tissue and diagnostic groups. These small and unmatched samples did not allow for the statistical analysis of methylation levels at the *APOE* variant sites across tissue and disease states, due to low power. Larger sample sizes in the future would provide the power to test significant differences where our results suggested differences were present observationally. In addition, larger sample sizes may allow for the identification and analysis of rare ε variants such as the ε1 isoform (Weisgraber or ε3r) (Murrell et al., 2006; Persico et al., 2004; Seripa et al., 2007), and increase the number of ε2 isoforms present for analysis.

Limitations in the quantification of methylation were mainly in issues generated from the 5% sensitivity threshold of pyrosequencing and subtraction-based mapping. Techniques such as chromatography assisted mass spectrometry (LC-MS/MS) are capable of quantifying all methylation types in high resolution without relying on the methylation assisted bisulphite sequencing subtraction based mapping (Chowdhury et al., 2017). Quantification by more sensitive means such as LC-MS/MS would allow for the quantification of methylation below the 5% error margin to confirm low levels of methylation versus depletion of methylation and would require no subtraction-based mapping.

Use of cell sorting to identify individual cell populations within tissue regions may elucidate cell population specific profiles of methylation and shed light onto individual cell types in disease state in the variation of methylation. Furthermore, quantification of transcript expression through digital droplet PCR may indicate the effects methylation have on protein abundance and how this may impact disease states in a brain region specific manner.

Amyloid-β plaque accumulation is heavily influenced by the Amyloid Precursor Protein (*APP*) and Presenilin (*PSEN1* & *PSEN2*) genes and their variant sites (Arber, Lovejoy, & Wray, 2017). Furthermore, there are gene mutations that influence the risk of developing Alzheimer's disease in *APOE* ε4 carriers, such as the rs2373115 variant in GRB2 Associated Binding Protein 2 (*GAB2*) (Babenko et al., 2018). In the future, these genes may also be analysed in conjunction with *APOE* variants for methylation to assess the effect of these genes on *APOE* related pathologies.

6.5.6 Conclusion

This study quantified 5mC, 5hmC and 5fC/CaC methylation at the *APOE* rs7412 and rs429358 sites in a blood and brain-tissue and disease-state specific manner. Our findings suggest that 5mC is the most abundant form of methylation at these sites and 5hmC is more abundant in brain tissue than blood.

This study determined that 5fC/CaC methylation at the rs7412 site appeared higher in cerebellum, while most other tissues showed comparatively lower levels. Similarly, 5fC/CaC methylation was lower at rs7412 in Parkinson's disease, dementia with Lewy bodies and Alzheimer's disease compared to the same site's 5mC levels. 5fC/CaC methylation at the rs429358 epiallele appeared enriched in all brain regions compared to blood and whole brain. Moreover, 5fC/CaC methylation was observationally higher than 5mC and 5hmC in dementia with Lewy bodies data at rs429358. These results together suggest that 5fC/CaC methylation may occur in a tissue and disease state specific manner and report high levels in cerebellum and dementia with Lewy bodies diagnoses.

Much research has been performed to quantify methylation by bisulphite conversion in the promoters of *APOE* to determine its influence on disease. However little research has quantified 5mC and 5hmC methylation at the rs7412 and rs429358 epialleles (Foraker et al., 2015; C. E. Yu et al., 2013) and no studies have established 5fC/CaC methylation at these variant sites. This study has highlighted allele specific 5mC, 5hmC and 5fC/CaC methylation patterns at the rs7412 and rs429358 epialleles between disease states and tissue types. However additional research is still required to elucidate the effect of different methylation types on phenotypes, not only to further understand the mechanisms driving cognitive disease states but also in the improvement of diagnostic medicines for brain disease.

Chapter 7: General Discussion

Although both *APOE* and *BDNF* have been strongly associated with risk for cognitive disease phenotypes, *APOE* is primarily associated with neurodegenerative conditions such as Alzheimer's disease and Parkinson's disease while *BDNF* has been linked with a broader spectrum of cognitive diseases. Traditionally both genes have been associated with risk for disease because of amino acid changes which putatively affects protein function. In this thesis, I argue for a role of DNA methylation in the regulation of these genes and their association with disease as an alternative mechanism that contributes to disease phenotypes. In this study, I investigated *APOE* risk rs7412 and rs429358 variants and the BDNF risk rs6265 missense variant on the antisense strand for variations in 5mC, 5hmC and 5fC/CaC DNA methylation.

I showed that at the *BDNF* AS-rs6265 site there were common patterns in 5mC and 5hmC methylation and expression in the context of tissue and disease states, and that methylation types investigated showed non-significant but differential methylation across the *APOE* risk alleles in different disease and tissue states. I theorise the *APOE* ε4 risk epsilon has the potential to influence APOE abundance though a DNA methylation regulatory mechanism and that *BDNF* rs6265 differential methylation may act as an indicator to surrounding gene methylation, in turn influencing specific transcript abundance and indicating disease by proxy. This discussion will elaborate on the mechanisms by which this may be possible.

7.1 Methylation and the effects of ASM and CRE in disease

7.1.1 Methylation as a mechanism causing or as a result of disease

Many studies have observed differential methylation states and levels and associated these differences with disease states, especially in brain disease. Throughout this thesis I have made the argument for methylation as a driving force contributing to disease phenotypes, but have also suggested that where this may not be true, methylation may still act as a biomarker for disease and would still be very useful in diagnostic medicine. The mechanisms by which these two possibilities may occur rely on the role of methylation in disease.

Methylation may act as a biomarker for disease where the underlying causes for the diseases result in differential methylation as a symptom of disease, but the methylation itself does not drive the deleterious phenotype. There have been many associations between disease states and altered methylation levels, in a cell and gene specific manner. A good example of this is methylation profiles in disease states in specific gene regions and cell types. For instance, 85 genes were identified that are significantly hypomethylated in Parkinson's disease patients compared to controls in at least one gene region (C. Wang, Chen, Yang, Zhang, & Wong, 2019). Similarly, *RARB*, *RASSF1*, and *GSTP1* genes report increased promoter methylation in the urine of patients with prostate cancers compared to individuals without prostate cancer (Bakavicius et al., 2019). In both of these cases, methylation levels have altered as an indirect effect of the disease state, but the change in methylation is not thought to be the driving causative factor behind the disease phenotype.

Methylation may drive disease where abnormal methylation of genes may result in altered transcription which in turn may manifest as a deleterious phenotype. For instance, fragile X syndrome is caused by CGG trinucleotide repeats in the 5' region of the *FMR1* gene, resulting in hypermethylation and gene silencing, stopping transcription and translation of the fragile X mental retardation protein (Penagarikano, Mulle, & Warren, 2007; Verkerk et al., 1991). Lack of this protein due to methylation mediated gene silencing results in severe mental deficiencies, and autism like behaviours (Penagarikano et al., 2007; Verkerk et al., 1991). While it could be argued that the CGG repeat mutation is driving the disease, it is the functional silencing of the gene that is the causative factor in this syndrome, and treatment of human cell lines with 5-azadeoxycytidine (a DNA methylation inhibitor which causes demethylation through dilution) and drugs that can cause histone hyperacetylation and histone decompaction reactivate transcription of the *FMR1* gene result in 2-5 fold increases in *FMR1* mRNA (Chiurazzi et al., 1999; Roberta et al., 2002).

An additional way aberrant methylation may contribute to disease is through loss of imprinting, which is critical to development. Both Prader-Willi syndrome is caused by loss of function across the 15q11–q13 region, in paternally expressed genes

(Goldstone, 2004; Robertson, 2005; Robertson & Wolffe, 2000). In Prader-Willi syndrome the loss of imprinting at the small nuclear riboprotein polypeptide N (*SNRPN*) 3' imprinting control region is partially due to de novo methylation of the paternal allele (Goldstone, 2004; Robertson, 2005). Similar to fragile X syndrome, treatment of lymphoblastoid cell lines derived from Prader-Willi syndrome patients with 5-azadeoxycytidine halted the loss of imprinting and allowed for the expression of *SNRPN* (Saitoh & Wada, 2000).

In this study, *BDNF* and *APOE* methylation were observed at key epialleles, however whether the methylation at these sites was causative or symptomatic of the disease states could not be determined. Future studies may endeavour to determine the effect of methylation in these diseases, perhaps by performing proteomic analysis of transcribed and translated transcripts in response to different levels of methylation in key genes, comparing these findings to transcript and protein levels and functions in patients with the disease. However, as Alzheimer's disease, Parkinson's disease, dementia with Lewy bodies, etc. are often complex diseases affecting different tissues, ages, and genes differently, it can be difficult to determine the pathways and mechanisms in which differences in genotype, methylation, transcription, translation, post-translational modification, distribution, and cell interaction may all be contributing to disease.

7.1.2 The role of genotype, ASM, and CRE

As epigenetics as an added level of transcriptional control "above" genetics, the genotype of an individual precludes the ability to have methylation, as well as the type and levels of methylation expected. This is known as allele specific methylation. Moreover, the majority of ASM mechanisms are thought to act in a cis-regulatory manner (i.e. are CREs), and interestingly ASM is commonly caused by cis-acting genetic variant states or epialleles, such as rs6265 (Tycko, 2010). Often variant genotypes can cause ASM which regulate expression in a cis-acting manner, and typically act by shifting the ratio of transcripts (Tycko, 2010). In this way, genotype, ASM and CRE are tightly inter-related and often involved in the same processes,

however this can make it difficult to differentiate which mechanisms are causative and contributing to drive phenotype.

This study observed ASM in the form of different levels and ratios of methylation types at different genotypes. Heterozygotes at *BDNF* rs6265 and *APOE* rs7412 and rs429358 reported more similar levels of 5mC and 5hmC compared to C homozygotes which consistently reported higher levels of 5mC compared to 5hmC. At *BDNF* rs6265 the Imperial cohort reported tissue, as well as tissue and genotype as significant factors affecting the observed levels of 5mC and 5hmC methylation, respectively. *APOE* appeared to show that the highest levels of methylation were observed at ε3/ε4 phenotypes, instead of ε4/ε4, although this was not significant. This indicates that genotype (ASM), as well as tissue types directly affected the level and type of methylation observed. This is reflected in the cerebellum samples whereby rs6265 observationally reported approximately equal levels of 5mC and 5hmC methylation (higher levels of 5hmC compared to other tissue types) in C homozygotes and in C/T heterozygotes reported higher levels of 5hmC than 5mC in the cerebellum.

The observed ASM at rs6265 also correlates with expression, as sense and AS-204 transcript expression significantly correlated with 5hmC methylation in C homozygous genotypes. However, a significant negative correlation was also observed between AS-201, and 5hmC methylation in C/T heterozygotes. This indicates that the methylation observed at rs6265 is cis-acting (a CRE) and this may contribute to phenotypes through the upregulation or repression of transcript. This may be the reason that there is a significant negative correlation between 5mC methylation and the sense, as well as the sense and AS-204 transcripts, in MCI and DLB samples, respectively.

There are several possible mechanisms by which ASM may act as a CRE and effect phenotype. 5mC methylation represses transcription through MECP2 mediated chromatin compaction and deacetylation (Boyes & Bird, 1991; Jones et al., 1998; Nan et al., 1998; Robertson & Wolffe, 2000). Alternatively, ASM that shifts the ratios of transcripts being transcribed may upregulate a CRE lncRNA that may sequester

mRNA in a RNA duplex (Modarresi et al., 2012; Pruunsild et al., 2007; Tycko, 2010). Additionally, methylation of a CpG site in a CGI can have a knock on effect by encouraging the methylation of surrounding CpGs in cancer cell lines- thus variant genotypes that generate novel CpGs and show ASM may play an active role in transcription, translation, and phenotype (Gebhard et al., 2010; Meaburn et al., 2010; Tycko, 2010). These mechanisms may alter the level of observed transcripts and therefore proteins, which may then alter phenotype and manifest in disease.

7.1.3 Methylation state ratios, and the significance of 5fC/CaC methylation

The levels and types of methylation observed are not merely a result of ASM however, there are other factors such as the gene, strand (M. Yu et al., 2012), surrounding CpG density (Gebhard et al., 2010), proximity to promoters or repeat sequences (Gebhard et al., 2010; Irizarry et al., 2009), and neighbouring bases to cytosines (Woodcock et al., 1987). 5mC and 5hmC show an sense and anti-sense strand preference, respectively (M. Yu et al., 2012), and this may change with cell type (Mooijman et al., 2016).

Unlike genomic imprinting, CRE ASM is also quantitative and heterogenous in nature meaning that the levels and ratios of methylation result in different isoforms being transcribed, but also that variation in methylation may derive from cell type, disease state, age, etc. (Meaburn et al., 2010; Schalkwyk et al., 2010; Shoemaker et al., 2010). The quantitative nature of ASM typically means that transcriptional expression is not biphasic, but normally an intermediate expression of transcripts mediated by allelic load, and this may be mediated by the type and relative amount of methylation present (Meaburn et al., 2010; Schalkwyk et al., 2010; Shoemaker et al., 2010). I believe that the higher levels of 5hmC to 5mC observed in heterozygotes, compared with C homozygotes in this study are indicative of ASM. I believe that higher 5hmC methylation is tied to higher mRNA transcription at rs6265 (as evidenced by the correlations observed in chapter 5) which demonstrates rs6265 as possessing cis-acting ASM.

Moreover, while this study has quantified methylation at the most common sites available for methylation (CpG sites) methylation may also occur in CpH, CHH, and CGH sites (Ramsahoye et al., 2000; Woodcock et al., 1987). CpH methylation is key in embryonic stem cell development and primarily occurs at CpA and CpT sites in human pluripotent stem cell lines and neuronal cells (J U Guo et al., 2014; Ramsahoye et al., 2000; Varley et al., 2013). CpH methylation (determined by bisulphite sequencing) has shown to be bound by MECP2 and is thought to be a transcriptional repressor similar to CpG 5mC methylation (J U Guo et al., 2014). In this way, CpH methylation reports factors that influence its methylation such as cell type (J U Guo et al., 2014), disease (Ellis et al., 2017), and shows a CRE function similar to that of CpG methylation. CpH methylation should also be incorporated for analysis in future studies to fully understand the mechanisms of ASM CRE in affecting phenotypes observed.

5fC/CaC methylation

5fC/CaC methylation has been reported to be essential in neuronal stem cell differentiation in mouse embryo development (Wheldon et al., 2014), and genome wide analysis identified TDG dependant accumulation of 5fC/CaC in gene regulatory elements to regulate this development (Shen et al., 2013; C. X. Song et al., 2013). 5fC/CaC are known to reduce transcription when mediated by T7 RNA polymerases in human cell lines (You, Ji, Dai, & Wang, 2014). Human neurons have also been shown to be differentially 5fC/CaC enriched in Alzheimer's disease compared to controls, indicating that it may be altered as a side effect/in response to the Alzheimer's phenotype, or used as a mechanism in the development of the Alzheimer's phenotype (Fetahu et al., 2019). Furthermore, within the cerebellum (Flitton, 2018) and within areas where adult neurogenesis occurs (the subventricular zone and dentate gyrus granular cell layer where high metabolism is needed), 5CaC was found to be present in stem cells (Haig et al., n.d.).

While no statements about 5fC/CaC methylation could be made in this study due to a lack of power there were some trends observed; such as higher levels of 5fC/CaC methylation in the brain regions, of which cerebellum typically had the higher

5fC/CaC methylation at most sites. *APOE* rs429358 reported very high 5fC/CaC methylation compared to the other sites, which may indicate that it was enriched at this site for a functional reason. While little is known about the 5fC/CaC methylation marks compared to the 5mC and 5hmC marks, 5fC/CaC is clearly being tightly regulated in different cell types (Wheldon et al., 2014), stages of development (Shen et al., 2013; C. X. Song et al., 2013), and disease states (Fetahu et al., 2019), and ongoing investigation to quantify these methylation types would prove to be valuable in the comprehension of how methylation, and ASM may play a role in regulating transcription (cis or trans) in the development of phenotype.

7.2 *BDNF* antisense methylation and mechanisms behind transcript expression

Methylation states are typically present at both sense and anti-sense complementary CpG sites due to DNMT1 mediated maintenance of methylation at opposite strand CpGs to avoid passive demethylation of DNA in replication (Li et al., 1992; Li & Zhang, 2014). As stated earlier, the amount and type of methylation present at a single CpG site has the ability to effect promoter activity and be indicative of a gene regions expression in human cell lines (Kitsera et al., 2017). 5mC methylation of a single promoter cAMP response element (cAMPRE) sequence (5'-TGACGTCA-3') can cause a 2 – 3 fold reduction in CREB transcription factor binding, and therefore cAMPRE dependant gene expression (Kitsera et al., 2017). Moreover, 5fC/CaC methylation of the same CpG in the CRE sequence initiated TDG mediated gene silencing, through BER which is known to inhibit transcription through the production of oxo-guanine, even when in non-transcribed and non-promoter regions (Allgayer, Kitsera, Bartelt, Epe, & Khobta, 2016; Kitsera et al., 2017; Lühnsdorf, Epe, & Khobta, 2014). This is relevant as the methylation measured at *BDNF* rs6265 (as well as *APOE* rs7412 and rs429358) is most probably indicative of methylation on the opposite strand at the same site, and has the ability to influence expression of different gene transcripts, dependant on the methylation types and levels present.

The Met66 variant is known to impede neuronal branching and reductions in dendritic length as well as impaired *BDNF* trafficking in neurons (Xin Xu, Garcia, Ewalt, Nason, & Pozzo-Miller, 2017). 5mC/5hmC methylation has been indicated at the

BDNF rs6265 adjacent cytosine on the sense strand and associated with hypoxia-related early life events and difference in cognitive performance in healthy individuals and individuals with Schizophrenia (Ursini et al., 2016). Moreover, increased 5mC at the rs6265 sense CpG was correlated with increased BDNF expression in serum but not in peripheral blood mononuclear cells (Ursini et al., 2016). However, nothing is known about the complementary anti-sense CpG methylation and how this potentially could facilitate increases in antisense transcript abundance, and therefore by proxy potentially regulate the abundance of the BDNF protein. At the point of writing, this is the first known research into anti-sense *BDNF* rs6265 CpG methylation.

Antisense *BDNF* lncRNA expression is known to suppress *BDNF* mRNA expression through cis-acting regulatory elements (CRE) mediated histone compaction (Modarresi et al., 2012; Zhong et al., 2017). Moreover, antisense lncRNA methylation in transcripts such as AS-*AFAP1* and AS-*SOX21* has been reported to significantly decrease in multiple cancer types and is associated with migration of cancerous cells and poorer prognosis (Bo et al., 2019; R. Wang et al., 2019). In this study, I report significant negative correlations between 5mC methylation at AS-rs6265 and sense (and antisense AS-204) transcript expression in individuals with MCI and DLB diagnoses, whereas the antisense AS-201 correlated weakly with AS-rs6265 5mC. I theorise that the *BDNF* sense and AS-204 transcripts are downregulated in expression mediated through increasing expression of AS-201 transcript through increases in AS-rs6265 5mC methylation, and this is mediated by an ASM CRE R-loop/promoter methylating mechanism. This supports evidence that some AS-*BDNF* transcripts like AS-201 transcripts are lncRNA CREs (Modarresi et al., 2012; Pruunsild et al., 2007). I propose that antisense CpG methylation at *BDNF* rs6265 acts as an additional regulator in gene mRNA expression, potentially through an ASM CRE mechanism, and this contributes to observed disease phenotypes which may be in combination with sense rs6265 methylation and the rs6265 amino acid protein changes.

Alternatively, AS-rs6265 5mC methylation may be a reflection of methylation at the sense rs6265 CpG which could also act to repress sense (and AS-204) transcript expression. It is known that CpG sites are methylated in a strand complementary fashion (J U Guo et al., 2014; Holliday & Pugh, 1975). However, no research has quantified the methylation types and levels present on complementary risk allele CpG sites, and this is yet to be determined. The resultant phenotype therefore may be a combination of both sense and antisense methylation operating through different mechanisms to regulate mRNA transcription, either directly or indirectly.

7.3 *APOE* variant analysis and proposed mechanisms of regulation

Methylation levels determined by bisulphite sequencing have been documented at the *APOE* rs7412 and rs429358 bases (Foraker et al., 2015). However, this study is the first to differentiate and quantify 5mC, 5hmC, and 5fC/CaC methylation at these sites in brain regions and tissues as well as multiple disease states. This study observed higher 5fC/CaC levels than 5mC at rs429358 in the cerebellum, and dementia with Lewy body samples which may suggest a connection between this methylation type, tissue and disease. This study also observed potential variant site-specific allele specific methylation (ASM) where rs429358 reported the highest level of 5hmC methylation in ε3/ε4 isoforms, which does not conform with typical methylation that would increase proportionally with cytosine availability. While this study could not determine significant findings, I theorise the *APOE* ε4 risk epsilon has the potential to influence APOE abundance though a DNA methylation regulatory mechanism; where the *APOE* variant sites may act as an indicator for surrounding methylation, or influence promoter activity directly which in turn effects APOE transcript abundance and therefore phenotype.

It is known that APOE transcripts are upregulated in the temporal lobe of Alzheimer's patients compared to controls (Twine, Janitz, Wilkins, & Janitz, 2011) and this may be in response to regulators of transcription such as DNA methylation or regulatory elements, independent of ε isoform and genotype. APOE rs7412 and rs429358 are thought to be a significant driving force in the activity exerted by the APOE 3' CGI, as

they not only change the CGI's CpG content but also alter the CpG density, which in turn affects binding ability of MBD proteins (C.-E. Yu & Foraker, 2015).

As stated above, a single CpG site has the ability to regulate a gene's transcript expression (Kitsera et al., 2017). The *APOE* 3' CGI containing the rs7412 and rs429358 variant sites is known to have CRE functional activity, and the rs7412 and rs429358 genotype, 3' exon CGI methylation levels, and cell types, have all been shown to influence *APOE* transcription (C. E. Yu et al., 2013). The negative correlation between the APOE 3' CGI and APOE transcript expression (including circular APOE and truncated APOE RNA isoforms) are thought to be mediated by MECP2, in a methylation dependant manner- whereby as methylation decreases at the *APOE* 3' CGI, so does MECP2 binding (E.-G. Lee et al., 2020; C. E. Yu et al., 2013). In this scenario, MECP2 binding of the *APOE* 3' CGI is thought to recruit and bind to the Sin3 corepressor and histone deacetylase to facilitate the deacetylation, chromatin compaction, and transcriptional silencing of APOE by acting on the *APOE* promoter (Boyes & Bird, 1991; Jones et al., 1998; Nan et al., 1998; Robertson & Wolffe, 2000).

I propose that this regulatory activity generated by the *APOE* 3' CGI is at least in part driven by the methylation of rs7412 and rs429358, whereby the type and level of methylation contribute to the observed levels of APOE mRNA. I also suggest that this mechanism may be ASM driven, where the methyl marks present may influence the type and amount of transcripts being transcribed. Several mechanisms are possible for the two epialleles to exert this effect. Firstly, rs7412 and rs429358 methylation may drive APOE transcriptional repression through the recruitment of MECP2 to the promoter of APOE. MECP2 would then represses APOE transcription through conventional deacetylation, chromatin compaction, or through the recruitment of DNMT1 to allow for methylation of the promoter (Boyes & Bird, 1991; Jones et al., 1998; Kimura & Shiota, 2003; Nan et al., 1998; Robertson & Wolffe, 2000). Methylation of the APOE promoter even at a single cAMPRE site would act to decrease *APOE* mRNA transcription (Kitsera et al., 2017).

Secondly, it is possible methylation of rs7412 and rs429358 results in antisense transcript expression (as seen in this studies BDNF results) such as lncRNA AS-*APOE* transcripts encoded by APOE exon IV, which have been identified as CREs to reduce APOE mRNA available for translation (Seitz et al., 2005; C.-E. Yu & Foraker, 2015). Additionally, APOE possesses multiple cis-element binding sites, such as TATA and GC boxes which regulate transcription in response to environmental factors (Chang et al., 1990; Paik et al., 1985; Theuns, 2000) and these binding sites may act as a location of AS-*APOE* lncRNA to bind and regulate *APOE* mRNA transcription.

This hypothesis could be tested by creating an artificial APOE construct in neuronal cell lines which have only been methylated at rs7412 and rs429358. Determining the expression of *APOE* based on rs7412, rs429358, and both sites methylation, would determine if these are the driving forces in regulation of transcription through MECP2 recruitment to the promoter, or other mechanisms. This would inform as to whether the methylation observed at rs7412 and rs429358 in current studies is driving the repression of *APOE* transcription, or is merely an indicator of the surrounding CGI methylation, which is driving repression. If the variant epialleles are shown to be driving repression of the APOE gene, further studies would be required to determine how much of the observed associations with disease are due to the genotype of the variants, the methylation, or a combination of the two (ASM).

7.4 The field of epigenetics in disease, and future studies

The results discussed in this thesis support other findings- such as differential DNA methylation being attributed to cis-acting ASM (J. R. Gibbs et al., 2010; Meaburn et al., 2010; D. Zhang et al., 2010), an enrichment for 5hmC methylation in brain tissues (Hahn et al., 2014; Kriaucionis & Heintz, 2009; Lister et al., 2013), association with 5mC with the downregulation of mRNA (Deaton & Bird, 2011), an association of 5hmC with the upregulation of mRNA (Wen et al., 2014), and cell specific methylation profiles in mammals (Mooijman et al., 2016). While this study could not determine significant differences due to lack of power, this is the first known study to quantify AS-rs6265 methylation and observe its effect on transcription, and observe 5fC/CaC

methylation at the *APOE* rs7412 and rs429358 epialleles. Observationally there appears to be higher levels of non-5mC methylation at APOE rs429358 compared to rs6265 and rs7412 which is a novel finding, however this may be due to smaller numbers, and this would need to be investigated in more depth in the future to determine if this is a significant difference. Additionally this study was able to distinguish between 5mC and 5hmC methylation at variant alleles which other studies investigating ASM have been unable to do (J. R. Gibbs et al., 2010; Meaburn et al., 2010; D. Zhang et al., 2010). This study therefore gives a novel insight into the ratios of 5mC to 5hmC at different genotypes- indicating that the ratio of 5mC:5hmC is closer in heterozygotes compared to C homozygotes which is far more 5mC dominant. This is a novel finding that may extend to other ASM sites.

Future studies leading on from this research may endeavour to repeat the study with large sample numbers, and now that the techniques for conversion have been optimised, repeatedly quantify methylation at these sites to generate multiple readings per sample and reduce variance. Additional technologies such as CaMAB-Seq would be able to differentiate 5fC and 5CaC and allow for quantification of all types of methylation (H. Wu & Zhang, 2015). Alternatively, quantification with HPLC-UV or LC-MS/MS would allow for more sensitive quantification below the 5% mark which may be of value when quantifying5fC and 5CaC, however these are lower throughput methods.

Another option for the quantification of methylation is the use of methylation arrays which would quantify methylation at hundreds to thousands of CpGs across the genome, and often provide free and accessible data sets for meta-analysis, for instance with the Pidsley et. al. study (Pidsley et al., 2016). While there was insufficient time in this study to access these data, many studies have used array data to analyse the rs6265 epiallele and combine them with fMRI data to show associations between *BDNF* methylation and hippocampal functional connectivity (Moser et al., 2015; Ursini et al., 2016; Wheater et al., 2020). These could be used in conjunction with more in-depth quantification of specific epialleles such as rs6265.

Additional studies may also want to quantify expression profiles for APOE in different disease states and across tissues, as was done in chapter 5. As APOE and BDNF are both key genes in cognition, they may be worth investigating in tandem to look at trans-acting mechanisms as well as cis acting mechanisms produced by ASM. Furthermore, other genes involved in cognition and methylation that possess variant epialleles may act to influence cognitive phenotypes through the ASM mechanisms discussed above. *DNMT3L* rs7354779 (C > T) is an epiallele which has been associated with childhood intelligence, but the mechanism by which alterations in methylation mediated by the *DNMT3L* mutation may be causing this are unknown (Haggarty et al., 2010). *SNCA* rs2583988 (C > T) is an epiallele which reported a significant association with Parkinson's disease and α-synucleopathies (Campêlo et al., 2017; Winkler et al., 2007) and would warrant investigation as a candidate epiallele for cognition.

7.5 *APOE* and *BDNF* as genes in cognition

In contrast to the antisense methylation investigated for BDNF, sense methylation was investigated at the *APOE* variants which is more relevant for the ε4 coding variants. APOE and BDNF have contrasting gene architecture and are on different chromosomes (chromosome 19, and 11, respectively). While *BDNF* has only one major clinical coding variant which has been consistently associated with risk for disease and changes in cognitive performance (Binder & Scharfman, 2004), *APOE* has two. Moreover, while the *BDNF* rs6265 variant may primarily affect transportation and activity dependent secretion of BDNF (Michael F. Egan et al., 2003), the *APOE* rs7412 and rs429358 variants are suggested to directly affect the ability of APOE to bind and function with other cellular machinery (L. M. Dong & Weisgraber, 1996).

APOE has also been shown to interact with *BDNF*, whereby ε isoforms reportedly differentially regulate the cleavage and secretion of pro-BDNF and BDNF in astrocytes (Sen, Nelson, & Alkon, 2017). Astrocytes treated with ε2 and ε3 APOE isoforms reported a 143.5 and a 38.4 fold increase in mature BDNF secretion after 24 hours, compared to those treated with ε4 APOE (Sen et al., 2017). Moreover, while

application of ε2 to astrocytes reduced the amount of secreted pro-BDNF, while application of ε3/ε4 APOE resulted in increased pro-BDNF secretion (Sen et al., 2017). This may explain the pathology generated by neurodegenerative diseases, as accumulation of pro-BDNF triggers the p57 cascade which results in apoptosis (Friedman, 2000). Furthermore, APOE ε4 isoforms and amyloid-β plaques have been shown to reduce the transcription of *BDNF* by recruiting histone deacetylases (HDACs) to *BDNF* exon IV, compacting chromatin, and impair transcription (Sen, Nelson, & Alkon, 2015). This is reflected in the reduction of BDNF in the hippocampus and frontal cortex of Alzheimer's patients (Siegel & Chauhan, 2000), as well as a reduction in the pro-BDNF and BDNF in preclinical stages of Alzheimer's disease (Peng et al., 2005). In this way the APOE protein is suggested to regulate the expression of BDNF protein. Therefore, it will be crucial to study the genes in parallel to understand how they regulate each other and contribute to pathology including regulation of gene transcription, maturation of pro-BDNF, and manifestation of disease phenotypes.

In summary, the effects of the *BDNF* and *APOE* variants on cognition and in neurodegenerative disease have been well documented. However, we still require a better understanding into the mechanisms linking allele risk and disease pathology. I propose that methylation at these risk epialleles on one or both strands contributes to disease through regulation of transcription. In this study, I show that CpGs comprising the antisense *BDNF* rs6265 variant and *APOE* rs7412 and rs429358 variants are 5mC, 5hmC, and 5fC/CaC methylated, and these methylation forms significantly correlate with transcript expression in tissue and disease groups. However further research into the effects of sense and antisense strand methylation on transcription, as well as interaction between different genes, and regulation through cis or trans acting regulatory elements is essential in understanding the aetiology of neurodegenerative diseases.

References

5-formylcytosine DNA Quantification Kit (ab156905) | Abcam. (2019). Retrieved June 12, 2019, from https://www.abcam.com/5-formylcytosine-dna-quantification-kit-colorimetric-ab156905.html

Abell, F., Krams, M., Ashburner, J., Passingham, R., Friston, K., Frackowiak, R., … Frith, U. (1999). The neuroanatomy of autism: A voxel-based whole brain analysis of structural scans. *NeuroReport*, *10*(8), 1647–1651. https://doi.org/10.1097/00001756-199906030-00005

Acharya, P., Segall, M. L., Zaiou, M., Morrow, J., Weisgraber, K. H., Phillips, M. C., … Snow, J. (2002). Comparison of the stabilities and unfolding pathways of human apolipoprotein E isoforms by differential scanning calorimetry and circular dichroism. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, *1584*(1), 9–19. https://doi.org/10.1016/S1388-1981(02)00263-9

Alfimova, M., Kondratyev, N., Golov, A., & Golimbet, V. (2020). Profiling haplotype specific CpG and CpH methylation within a schizophrenia GWAS locus on chromosome 14 in schizophrenia and healthy subjects. *Scientific Reports*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-61671-2

Allen, N. J., & Barres, B. A. (2009). Neuroscience: Glia — more than just brain glue. *Nature*, *457*(7230), 675–677. https://doi.org/10.1038/457675a

Allgayer, J., Kitsera, N., Bartelt, S., Epe, B., & Khobta, A. (2016). Widespread transcriptional gene inactivation initiated by a repair intermediate of 8-oxoguanine. *Nucleic Acids Research*, *44*(15), 7267–7280. https://doi.org/10.1093/nar/gkw473

Alonso, A. D. C., Zaidi, T., Novak, M., Grundke-Iqbal, I., & Iqbal, K. (2001). Hyperphosphorylation induces self-assembly of τ into tangles of paired helical filaments/straight filaments. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(12), 6923–6928. https://doi.org/10.1073/pnas.121119298

Altar, C. A. (1999). Neurotrophins and depression. *Trends in Pharmacological Sciences*, *20*(2), 59–62. https://doi.org/10.1016/S0165-6147(99)01309-7

Alvarez-Castelao, B., & Castaño, J. G. (2011). Synphilin-1 inhibits alpha-synuclein degradation by the proteasome. *Cellular and Molecular Life Sciences*, *68*(15), 2643–2654. https://doi.org/10.1007/s00018-010-0592-3

Amir, R. E., Van den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U., & Zoghbi, H. Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics*, *23*(2), 185–188. https://doi.org/10.1038/13810

Anderson, K. E. (2004). Behavioral disturbances in Parkinson's disease. *Dialogues in Clinical Neuroscience*, *6*(3), 323–332. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22033600

Angelucci, F, Brenè, S., & Mathé, A. A. (2005). BDNF in schizophrenia, depression and corresponding animal models. *Molecular Psychiatry*, *10*(4), 345–352. https://doi.org/10.1038/sj.mp.4001637

Angelucci, Francesco, Spalletta, G., di Iulio, F., Ciaramella, A., Salani, F., Colantoni, L., … Bossù, P. (2010). Alzheimer's disease (AD) and Mild Cognitive Impairment (MCI) patients are characterized by increased BDNF serum levels. *Current Alzheimer Research*, *7*(1), 15–20. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/20205668
Anichtchik, O., Calo, L., & Spillantini, M. G. (2013). Synaptic dysfunction in synucleinopathies. *CNS & Neurological Disorders Drug Targets*, *12*(8), 1094–1100. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24040824
Anita E. Autry, L. M. M. (2012). Brain-Derived Neurotrophic Factor and Neuropsychiatric Disorders. *Pharmacological Reviews*, *64*(2), 238.
Arber, C., Lovejoy, C., & Wray, S. (2017). Stem cell models of Alzheimer's disease: progress and challenges. *Alzheimer's Research & Therapy*, *9*(1), 42. https://doi.org/10.1186/s13195-017-0268-4
Babenko, V. N., Afonnikov, D. A., Ignatieva, E. V., Klimov, A. V., Gusev, F. E., & Rogaev, E. I. (2018). Haplotype analysis of APOE intragenic SNPs. *BMC Neuroscience*, *19*(S1), 16. https://doi.org/10.1186/s12868-018-0413-4
Bachman, M., Uribe-Lewis, S., Yang, X., Burgess, H. E., Iurlaro, M., Reik, W., … Balasubramanian, S. (2015). 5-Formylcytosine can be a stable DNA modification in mammals. *Nature Chemical Biology*, *11*(8), 555–557. https://doi.org/10.1038/nchembio.1848
Badeaux, A. I., & Shi, Y. (2013). Emerging roles for chromatin as a signal integration and storage platform. *Nature Reviews Molecular Cell Biology*, *14*(4), 211–224. https://doi.org/10.1038/nrm3545
Bakavicius, A., Daniunaite, K., Zukauskaite, K., Barisiene, M., Jarmalaite, S., & Jankevicius, F. (2019). Urinary DNA methylation biomarkers for prediction of prostate cancer upgrading and upstaging. *Clinical Epigenetics*, *11*(1), 115. https://doi.org/10.1186/s13148-019-0716-z
Baldeiras, I., Santana, I., Proença, M. T., Garrucho, M. H., Pascoal, R., Rodrigues, A., … Oliveira, C. R. (2008). Peripheral oxidative damage in mild cognitive impairment and mild Alzheimer's disease. *Journal of Alzheimer's Disease : JAD*, *15*(1), 117–128. https://doi.org/10.3233/jad-2008-15110
Bannerman, D. M., Bus, T., Taylor, A., Sanderson, D. J., Schwarz, I., Jensen, V., … Sprengel, R. (2012). Dissecting spatial knowledge from spatial choice by hippocampal NMDA receptor deletion. *Nature Neuroscience*, *15*(8), 1153–1159. https://doi.org/10.1038/nn.3166
Bannister, A. J., & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Research*, *21*(3), 381–395. https://doi.org/10.1038/cr.2011.22
Barman, P., Reddy, D., & Bhaumik, S. R. (2019). Mechanisms of Antisense Transcription Initiation with Implications in Gene Expression, Genomic Integrity and Disease Pathogenesis. *Non-Coding RNA*, *5*(1). https://doi.org/10.3390/ncrna5010011
Baubec, T., Colombo, D. F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A. R., … Schübeler, D. (2015). Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature*, *520*(7546), 243–247. https://doi.org/10.1038/nature14176
Benjamin, R., Leake, A., McArthur, F. K., Ince, P. G., Candy, J. M., Edwardson, J. A., … Bjertness, E. (1994). Protective effect of apoE epsilon 2 in Alzheimer's disease. *Lancet (London, England)*, *344*(8920), 473. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7914580
Bennett, D. A., & Gordon, B. A. (2018). Alzheimer disease biomarkers and

synucleinopathy. *Neurology*, *90*(12), 537–538. https://doi.org/10.1212/WNL.0000000000005171

Bertram, L., McQueen, M. B., Mullin, K., Blacker, D., & Tanzi, R. E. (2007). Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nature Genetics*, *39*(1), 17–23. https://doi.org/10.1038/ng1934

Binder, D. K., & Scharfman, H. E. (2004). Brain-derived neurotrophic factor. *Growth Factors (Chur, Switzerland)*, *22*(3), 123–131. https://doi.org/10.1080/08977190410001723308

Blennow, K., de Leon, M. J., & Zetterberg, H. (2006). Alzheimer's disease. *Lancet (London, England)*, *368*(9533), 387–403. https://doi.org/10.1016/S0140-6736(06)69113-7

Bo, H., Fan, L., Gong, Z., Liu, Z., Shi, L., Guo, C., … Xiong, F. (2019). Upregulation and hypomethylation of lncRNA AFAP1-AS1 predicts a poor prognosis and promotes the migration and invasion of cervical cancer. *Oncology Reports*. https://doi.org/10.3892/or.2019.7027

Booth, M. J., Marsico, G., Bachman, M., Beraldi, D., & Balasubramanian, S. (2014). Quantitative sequencing of 5-formylcytosine in DNA at single-base resolution. *Nature Chemistry*, *6*(5), 435–440. https://doi.org/10.1038/nchem.1893

Borba, E. M., Duarte, J. A., Bristot, G., Scotton, E., Camozzato, A. L., & Chaves, M. L. F. (2016). Brain-Derived Neurotrophic Factor Serum Levels and Hippocampal Volume in Mild Cognitive Impairment and Dementia due to Alzheimer Disease. *Dementia and Geriatric Cognitive Disorders Extra*, *6*(3), 559–567. https://doi.org/10.1159/000450601

Borodinova, A. A., & Salozhin, S. V. (2017). Differences in the Biological Functions of BDNF and proBDNF in the Central Nervous System. *Neuroscience and Behavioral Physiology*, *47*(3), 251–265. https://doi.org/10.1007/s11055-017-0391-5

Bostick, M., Kim, J. K., Estève, P.-O., Clark, A., Pradhan, S., & Jacobsen, S. E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science (New York, N.Y.)*, *317*(5845), 1760–1764. https://doi.org/10.1126/science.1147939

Boyes, J., & Bird, A. (1991). DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell*, *64*(6), 1123–1134. https://doi.org/10.1016/0092-8674(91)90267-3

Braak, H., & Del Tredici, K. (2004). Poor and protracted myelination as a contributory factor to neurodegenerative disorders. *Neurobiology of Aging*, *25*(1), 19–23. https://doi.org/10.1016/J.NEUROBIOLAGING.2003.04.001

Braak, H., Ghebremedhin, E., Rüb, U., Bratzke, H., & Del Tredici, K. (2004). Stages in the development of Parkinson's disease-related pathology. *Cell and Tissue Research*, *318*(1), 121–134. https://doi.org/10.1007/s00441-004-0956-9

Braak, H., Tredici, K. Del, Rüb, U., de Vos, R. A. ., Jansen Steur, E. N. ., & Braak, E. (2003). Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiology of Aging*, *24*(2), 197–211. https://doi.org/10.1016/S0197-4580(02)00065-9

Bradley-Whitman, M. A., & Lovell, M. A. (2013). Epigenetic changes in the progression of Alzheimer's disease. *Mechanisms of Ageing and Development*, *134*(10), 486–495. https://doi.org/10.1016/j.mad.2013.08.005

Bras, J., Guerreiro, R., Darwent, L., Parkkinen, L., Ansorge, O., Escott-Price, V., …

Hardy, J. (2014). Genetic analysis implicates APOE, SNCA and suggests lysosomal dysfunction in the etiology of dementia with Lewy bodies. *Human Molecular Genetics*, *23*(23), 6139–6146. https://doi.org/10.1093/hmg/ddu334

Brecht, W. J., Harris, F. M., Chang, S., Tesseur, I., Yu, G.-Q., Xu, Q., … Huang, Y. (2004). Neuron-Specific Apolipoprotein E4 Proteolysis Is Associated with Increased Tau Phosphorylation in Brains of Transgenic Mice. https://doi.org/10.1523/JNEUROSCI.4315-03.2004

Bridi, J. C., & Hirth, F. (2018). Mechanisms of α-Synuclein Induced Synaptopathy in Parkinson's Disease. *Frontiers in Neuroscience*, *12*, 80. https://doi.org/10.3389/fnins.2018.00080

Browne, D., McGuinness, B., Woodside, J. V., & McKay, G. J. (2019). Vitamin E and Alzheimer's disease: What do we know so far? *Clinical Interventions in Aging*. Dove Medical Press Ltd. https://doi.org/10.2147/CIA.S186760

Bu, G. (2009). Apolipoprotein E and its receptors in Alzheimer's disease: pathways, pathogenesis and therapy. *Nature Reviews. Neuroscience*, *10*(5), 333. https://doi.org/10.1038/NRN2620

Buchman, A. S., Yu, L., Boyle, P. A., Schneider, J. A., De Jager, P. L., & Bennett, D. A. (2016). Higher brain BDNF gene expression is associated with slower cognitive decline in older adults. *Neurology*, *86*(8), 735–741. https://doi.org/10.1212/WNL.0000000000002387

Budni, J., Bellettini-Santos, T., Mina, F., Garcez, M. L., & Zugno, A. I. (2015). The involvement of BDNF, NGF and GDNF in aging and Alzheimer's disease. *Aging and Disease*, *6*(5), 331–341. https://doi.org/10.14336/AD.2015.0825

Burmeister, M., McInnis, M. G., & Zöllner, S. (2008). Psychiatric genetics: progress amid controversy. *Nature Reviews Genetics*, *9*(7), 527–540. https://doi.org/10.1038/nrg2381

Campêlo, C. L. C., Cagni, F. C., Figueredo, D. de S., Oliveira, L. G., Silva-Neto, A. B., Macêdo, P. T., … Silva, R. H. (2017). Variants in SNCA Gene are associated with Parkinson's disease risk and cognitive symptoms in a Brazilian sample. *Frontiers in Aging Neuroscience*, *9*(JUN), 198. https://doi.org/10.3389/fnagi.2017.00198

Casadei, N., Pöhler, A.-M., Tomás-Zapico, C., Torres-Peraza, J., Schwedhelm, I., Witz, A., … Nuber, S. (2014). Overexpression of synphilin-1 promotes clearance of soluble and misfolded alpha-synuclein without restoring the motor phenotype in aged A30P transgenic mice. *Human Molecular Genetics*, *23*(3), 767–781. https://doi.org/10.1093/hmg/ddt467

Chagnon, Y. C., Potvin, O., Hudon, C., & Préville, M. (2015). DNA methylation and single nucleotide variants in the brain-derived neurotrophic factor (BDNF) and oxytocin receptor (OXTR) genes are associated with anxiety/depression in older women. *Frontiers in Genetics*, *6*, 230. https://doi.org/10.3389/fgene.2015.00230

Chakrabarti, S., Basu, D., Khemka, Vineet Kumar Bagchi, Debajit Bandyopadhyay, K., Bir, A., Chattopadhyay, M., & Biswas, A. (2014). Altered serum levels of adipokines and insulin in probable Alzheimer's disease. *Journal of Alzheimer's Disease : JAD*, *41*(2), 525–533. https://doi.org/10.3233/JAD-140006

Chang, D. J., Paik, Y. K., Leren, T. P., Walker, D. W., Howlett, G. J., & Taylor, J. M. (1990). Characterization of a human apolipoprotein E gene enhancer element and its associated protein factors. *The Journal of Biological Chemistry*, *265*(16),

9496–9504. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2188976
Chao, H. M., Kao, H.-T., & Porton, B. (2008). BDNF Val66Met variant and age of onset in schizophrenia. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics : The Official Publication of the International Society of Psychiatric Genetics*, *147B*(4), 505–506. https://doi.org/10.1002/ajmg.b.30619
Charlton, K. E., Rabinowitz, T. L., Geffen, L. N., & Dhansay, M. A. (2004). Lowered plasma vitamin C, but not vitamin E, concentrations in dementia patients. *The Journal of Nutrition, Health & Aging*, *8*(2), 99–107. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14978605
Chedin, F., Lieber, M. R., & Hsieh, C.-L. (2002). The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(26), 16916–16921. https://doi.org/10.1073/pnas.262443999
Chen, C.-C., Wang, K.-Y., & Shen, C.-K. J. (2012). The Mammalian *de Novo* DNA Methyltransferases DNMT3A and DNMT3B Are Also DNA 5-Hydroxymethylcytosine Dehydroxymethylases. *Journal of Biological Chemistry*, *287*(40), 33116–33121. https://doi.org/10.1074/jbc.C112.406975
Chen, C.-H. (2016). Development of a Melting Curve-Based Allele-Specific PCR of Apolipoprotein E (APOE) Genotyping Method for Genomic DNA, Guthrie Blood Spot, and Whole Blood. *PLOS ONE*, *11*(4), e0153593. https://doi.org/10.1371/journal.pone.0153593
Chen, K.-L., Wang, S. S.-S., Yang, Y.-Y., Yuan, R.-Y., Chen, R.-M., & Hu, C.-J. (2009). The epigenetic effects of amyloid-β1–40 on global DNA and neprilysin genes in murine cerebral endothelial cells. *Biochemical and Biophysical Research Communications*, *378*(1), 57–61. https://doi.org/10.1016/j.bbrc.2008.10.173
Chen, W. G., West, A. E., Tao, X., Corfas, G., Szentirmay, M. N., Sawadogo, M., … Greenberg, M. E. (2003). Upstream stimulatory factors are mediators of Ca2+-responsive transcription in neurons. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *23*(7), 2572–2581. https://doi.org/10.1523/jneurosci.4540-10.2011
Chen, Y., Damayanti, N. P., Irudayaraj, J., Dunn, K., & Zhou, F. C. (2014). Diversity of two forms of DNA methylation in the brain. *Frontiers in Genetics*, *5*, 46. https://doi.org/10.3389/fgene.2014.00046
Chibnik, L. B., Yu, L., Eaton, M. L., Srivastava, G., Schneider, J. A., Kellis, M., … De Jager, P. L. (2015). Alzheimer's loci: epigenetic associations and interaction with genetic factors. *Annals of Clinical and Translational Neurology*, *2*(6), 636–647. https://doi.org/10.1002/acn3.201
Chiurazzi, P., Pomponi, M. G., Pietrobono, R., Bakker, C. E., Neri, G., & Oostra, B. A. (1999). *Synergistic effect of histone hyperacetylation and DNA demethylation in the reactivation of the FMR1 gene*. *Human Molecular Genetics* (Vol. 8).
Chou, I., & Chouard, T. (2008). Neuropsychiatric disease. *Nature*, *455*(7215), 889–889. https://doi.org/10.1038/455889a
Chou, K. (2018). Clinical manifestations of Parkinson disease. Retrieved October 29, 2018, from https://www.uptodate.com/contents/clinical-manifestations-of-parkinson-disease
Chouliaras, L., Mastroeni, D., Delvaux, E., Grover, A., Kenis, G., Hof, P. R., … van den Hove, D. L. A. (2013). Consistent decrease in global DNA methylation and

hydroxymethylation in the hippocampus of Alzheimer's disease patients. *Neurobiology of Aging*, *34*(9), 2091–2099. https://doi.org/10.1016/j.neurobiolaging.2013.02.021

Chouliaras, L., Pishva, E., Haapakoski, R., Zsoldos, E., Mahmood, A., Filippini, N., … Ebmeier, K. P. (2018). Peripheral DNA methylation, cognitive decline and brain aging: pilot findings from the Whitehall II imaging study. *Epigenomics*, *10*(5), 585–595. https://doi.org/10.2217/epi-2017-0132

Chowdhury, B., Cho, I.-H., & Irudayaraj, J. (2017). Technical advances in global DNA methylation analysis in human cancers. *Journal of Biological Engineering*, *11*(1), 10. https://doi.org/10.1186/s13036-017-0052-9

Chuang, Y.-H., Paul, K. C., Bronstein, J. M., Bordelon, Y., Horvath, S., & Ritz, B. (2017). Parkinson's disease is associated with DNA methylation levels in human blood and saliva. *Genome Medicine*, *9*(1), 76. https://doi.org/10.1186/s13073-017-0466-5

Chung, E. J., Babulal, G. M., Monsell, S. E., Cairns, N. J., Roe, C. M., & Morris, J. C. (2015). Clinical Features of Alzheimer Disease With and Without Lewy Bodies. *JAMA Neurology*, *72*(7), 789–796. https://doi.org/10.1001/jamaneurol.2015.0606

Ciccarone, F., Tagliatesta, S., Caiafa, P., & Zampieri, M. (2018). DNA methylation dynamics in aging: how far are we from understanding the mechanisms? *Mechanisms of Ageing and Development*, *174*, 3–17. https://doi.org/10.1016/J.MAD.2017.12.002

Clayton, A. L., Hazzalin, C. A., Mahadevan, L. C., Allegra, P., Sterner, R., Clayton, D. F., … Bradbury, M. E. (2006). Enhanced Histone Acetylation and Transcription: A Dynamic Perspective. *Molecular Cell*, *23*(3), 289–296. https://doi.org/10.1016/j.molcel.2006.06.017

Clouaire, T., & Stancheva, I. (2008, May). Methyl-CpG binding proteins: Specialized transcriptional repressors or structural components of chromatin? *Cellular and Molecular Life Sciences*. Europe PMC Funders. https://doi.org/10.1007/s00018-008-7324-y

Condliffe, D., Wong, A., Troakes, C., Proitsi, P., Patel, Y., Chouliaras, L., … Lunnon, K. (2014). Cross-region reduction in 5-hydroxymethylcytosine in Alzheimer's disease brain. *Neurobiology of Aging*, *35*(8), 1850–1854. https://doi.org/10.1016/j.neurobiolaging.2014.02.002

Cong, L., Jia, J., Qin, W., Ren, Y., & Sun, Y. (2014). Genome-wide analysis of DNA methylation in an APP/PS1 mouse model of Alzheimer's disease. *Acta Neurologica Belgica*, *114*(3), 195–206. https://doi.org/10.1007/s13760-013-0267-6

Consortium, E. A., Lek, M., Karczewski, K., Minikel, E., Samocha, K., Banks, E., … MacArthur, D. (2015). *Analysis of protein-coding genetic variation in 60,706 humans*. *bioRxiv*. Cold Spring Harbor Labs Journals. https://doi.org/10.1101/030338

Consortium, Gte. (2015). The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science*, *348*(6235), 648–660.

Coppieters, N., Dieriks, B. V., Lill, C., Faull, R. L. M., Curtis, M. A., & Dragunow, M. (2014). Global changes in DNA methylation and hydroxymethylation in Alzheimer's disease human brain. *Neurobiology of Aging*, *35*(6), 1334–1344.

https://doi.org/10.1016/j.neurobiolaging.2013.11.031
Corder, E. H., Saunders, A. M., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., Small, G. W., … Pericak-Vance, M. A. (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science (New York, N.Y.)*, *261*(5123), 921–923. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8346443
Courchesne, E., Saitoh, O., Yeung-Courchesne, R., Press, G. A., Lincoln, A. J., Haas, R. H., & Schreibman, L. (1994). Abnormality of cerebellar vermian lobules VI and VII in patients with infantile autism: Identification of hypoplastic and hyperplastic subgroups with MR imaging. *American Journal of Roentgenology*, *162*(1), 123–130. https://doi.org/10.2214/ajr.162.1.8273650
Crawley, M. J. (2013). *The R Book* (Second). Wiley.
Crews, L., Tsigelny, I., Hashimoto, M., & Masliah, E. (2009). Role of synucleins in Alzheimer's disease. *Neurotoxicity Research*. https://doi.org/10.1007/s12640-009-9073-6
Cuervo, A. M., & Dice, J. F. (2000). When lysosomes get old. *Experimental Gerontology*, *35*(2), 119–131. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10767573
D'Addario, C., Dell'Osso, B., Palazzo, M. C., Benatti, B., Lietti, L., Cattaneo, E., … Altamura, A. C. (2012). Selective DNA Methylation of BDNF Promoter in Bipolar Disorder: Differences Among Patients with BDI and BDII. *Neuropsychopharmacology*, *37*(7), 1647–1655. https://doi.org/10.1038/npp.2012.10
Dashtipour, K., Tafreshi, A., Adler, C., Beach, T., Chen, X., Serrano, G., … Wang, C. (2017). Hypermethylation of Synphilin-1, Alpha-Synuclein-Interacting Protein (SNCAIP) Gene in the Cerebral Cortex of Patients with Sporadic Parkinson's Disease. *Brain Sciences*, *7*(7). https://doi.org/10.3390/brainsci7070074
Day, J. J., & Sweatt, J. D. (2010). DNA methylation and memory formation. *Nature Neuroscience*, *13*(11), 1319–1323. https://doi.org/10.1038/nn.2666
Day, J. J., & Sweatt, J. D. (2011a). Cognitive neuroepigenetics: A role for epigenetic mechanisms in learning and memory. *Neurobiology of Learning and Memory*, *96*(1), 2–12. https://doi.org/10.1016/j.nlm.2010.12.008
Day, J. J., & Sweatt, J. D. (2011b). Epigenetic Mechanisms in Cognition. *Neuron*, *70*(5), 813–829. https://doi.org/10.1016/j.neuron.2011.05.019
De Jager, C. A., Oulhaj, A., Jacoby, R., Refsum, H., & Smith, A. D. (2012). Cognitive and clinical outcomes of homocysteine-lowering B-vitamin treatment in mild cognitive impairment: A randomized controlled trial. *International Journal of Geriatric Psychiatry*. https://doi.org/10.1002/gps.2758
de Koning, A. P. J., Gu, W., Castoe, T. A., Batzer, M. A., & Pollock, D. D. (2011). Repetitive Elements May Comprise Over Two-Thirds of the Human Genome. *PLoS Genetics*, *7*(12), e1002384. https://doi.org/10.1371/journal.pgen.1002384
Deaton, A. M., & Bird, A. (2011). CpG islands and the regulation of transcription. *Genes & Development*, *25*(10), 1010–1022. https://doi.org/10.1101/gad.2037511
Defossez, P.-A., & Stancheva, I. (2011). Biological Functions of Methyl-CpG-Binding Proteins. In *Progress in molecular biology and translational science* (Vol. 101, pp. 377–398). https://doi.org/10.1016/B978-0-12-387685-0.00012-3

Delli Pizzi, S., Franciotti, R., Taylor, J.-P., Thomas, A., Tartaro, A., Onofrj, M., & Bonanni, L. (2015). Thalamic Involvement in Fluctuating Cognition in Dementia with Lewy Bodies: Magnetic Resonance Evidences. *Cerebral Cortex*, *25*(10), 3682–3689. https://doi.org/10.1093/cercor/bhu220

Desplats, P., Spencer, B., Coffee, E., Patel, P., Michael, S., Patrick, C., … Masliah, E. (2011). α-Synuclein Sequesters Dnmt1 from the Nucleus. *Journal of Biological Chemistry*, *286*(11), 9031–9037. https://doi.org/10.1074/jbc.C110.212589

Dimitrova, N., Zamudio, J. R., Jong, R. M., Soukup, D., Resnick, R., Sarma, K., … Jacks, T. (2014). LincRNA-p21 Activates p21 In cis to Promote Polycomb Target Gene Expression and to Enforce the G1/S Checkpoint. *Molecular Cell*, *54*(5), 777–790. https://doi.org/10.1016/j.molcel.2014.04.025

Dolinoy, D. C., Huang, D., & Jirtle, R. L. (2007). Maternal nutrient supplementation counteracts bisphenol A-induced DNA hypomethylation in early development. *Proceedings of the National Academy of Sciences*, *104*(32), 13056–13061. https://doi.org/10.1073/pnas.0703739104

Doltra, A., Dietrich, T., Schneeweis, C., Kelle, S., Doltra, A., Stawowy, P., & Fleck, E. (2013). Magnetic Resonance Imaging of Cardiovascular Fibrosis and Inflammation: From Clinical Practice to Animal Studies and Back Cardiovascular MRI View project Magnetic Resonance Imaging of Cardiovascular Fibrosis and Inflammation: From Clinical Practice to Ani. *BioMed Research International*, *676489*(10), 1–2. https://doi.org/10.1155/2013

Dong, E., Gavin, D. P., Chen, Y., & Davis, J. (2012). Upregulation of TET1 and downregulation of APOBEC3A and APOBEC3C in the parietal cortex of psychotic patients. *Translational Psychiatry*, *2*(9), e159–e159. https://doi.org/10.1038/tp.2012.86

Dong, L. M., & Weisgraber, K. H. (1996). Human apolipoprotein E4 domain interaction. Arginine 61 and glutamic acid 255 interact to direct the preference for very low density lipoproteins. *The Journal of Biological Chemistry*, *271*(32), 19053–19057. https://doi.org/10.1074/jbc.271.32.19053

Dulka, B. N., Ford, E. C., Lee, M. A., Donnell, N. J., Goode, T. D., Prosser, R., & Cooper, M. A. (2016). Proteolytic cleavage of proBDNF into mature BDNF in the basolateral amygdala is necessary for defeat-induced social avoidance. *Learning and Memory*, *23*(4), 156–160. https://doi.org/10.1101/lm.040253.115

Duncan, E. J., Gluckman, P. D., & Dearden, P. K. (2014). Epigenetics, plasticity, and evolution: How do we link epigenetic change to phenotype? *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, *322*(4), 208–220. https://doi.org/10.1002/jez.b.22571

Durany, Michel, Kurt, Cruz-Sánchez, Cervós-Navarro, & Riederer. (2000). Brain-derived neurotrophic factor and neurotrophin-3 levels in Alzheimer's disease brains. *International Journal of Developmental Neuroscience : The Official Journal of the International Society for Developmental Neuroscience*, *18*(8), 807–813. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11154850

Durazzo, T. C., Mattsson, N., & Weiner, M. W. (2014). Smoking and increased Alzheimer's disease risk: A review of potential mechanisms. *Alzheimer's and Dementia*, *10*(3 SUPPL.), S122. https://doi.org/10.1016/j.jalz.2014.04.009

Eccleston, A., DeWitt, N., Gunter, C., Marte, B., & Nath, D. (2007). Epigenetics. *Nature*, *447*(7143), 395–395. https://doi.org/10.1038/447395a

Egan, M. F., Goldberg, T. E., Kolachana, B. S., Callicott, J. H., Mazzanti, C. M., Straub, R. E., … Weinberger, D. R. (2001). Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proceedings of the National Academy of Sciences*, *98*(12), 6917–6922. https://doi.org/10.1073/pnas.111134598

Egan, Michael F., Kojima, M., Callicott, J. H., Goldberg, T. E., Kolachana, B. S., Bertolino, A., … Weinberger, D. R. (2003). The BDNF val66met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function. *Cell*, *112*(2), 257–269. https://doi.org/10.1016/S0092-8674(03)00035-7

El-Maarri, O., Kareta, M. S., Mikeska, T., Becker, T., Diaz-Lacava, a., Junen, J., … Chedin, F. (2009). A systematic search for DNA methyltransferase polymorphisms reveals a rare DNMT3L variant associated with subtelomeric hypomethylation. *Human Molecular Genetics*, *18*(10), 1755–1768. https://doi.org/10.1093/hmg/ddp088

Ellis, S. E., Gupta, S., Moes, A., West, A. B., & Arking, D. E. (2017). Exaggerated CpH methylation in the autism-affected brain. *Molecular Autism*, *8*(1), 6. https://doi.org/10.1186/s13229-017-0119-y

Eryilmaz, I. E., Cecener, G., Erer, S., Egeli, U., Tunca, B., Zarifoglu, M., … Kaleagasi, H. (2017). Epigenetic approach to early-onset Parkinson's disease: low methylation status of *SNCA* and *PARK2* promoter regions. *Neurological Research*, *39*(11), 965–972. https://doi.org/10.1080/01616412.2017.1368141

Faghihi, M. A., Zhang, M., Huang, J., Modarresi, F., Van der Brug, M. P., Nalls, M. A., … Wahlestedt, C. (2010). Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biology*, *11*(5), R56. https://doi.org/10.1186/gb-2010-11-5-r56

Fan, H., & Gulley, M. L. (2001). DNA Extraction from Fresh or Frozen Tissues. In *Molecular Pathology Protocols* (Vol. 49, pp. 5–10). New Jersey: Humana Press. https://doi.org/10.1385/1-59259-081-0:5

Faria, M. C., Gonçalves, G. S., Rocha, N. P., Moraes, E. N., Bicalho, M. A., Gualberto Cintra, M. T., … Sousa, L. P. (2014). Increased plasma levels of BDNF and inflammatory markers in Alzheimer's disease. *Journal of Psychiatric Research*, *53*, 166–172. https://doi.org/10.1016/j.jpsychires.2014.01.019

Farlow, M. R., He, Y., Tekin, S., Xu, J., Lane, R., & Charles, H. C. (2004). Impact of APOE in mild cognitive impairment. *Neurology*, *63*(10), 1898–1901. https://doi.org/10.1212/01.WNL.0000144279.21502.B7

Farrer, L. A., Cupples, L. A., Haines, J. L., Hyman, B., Kukull, W. A., Mayeux, R., … van Duijn, C. M. (1997). Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. *JAMA*, *278*(16), 1349–1356. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9343467

Federoff, M., Jimenez-Rolando, B., Nalls, M. A., & Singleton, A. B. (2012). A large study reveals no association between APOE and Parkinson's disease. *Neurobiology of Disease*, *46*(2), 389–392. https://doi.org/10.1016/j.nbd.2012.02.002

Feigin, V. L., & Vos, T. (2019). Global Burden of Neurological Disorders: From Global Burden of Disease Estimates to Actions. *Neuroepidemiology*, *52*(1–2), 1–2.

https://doi.org/10.1159/000495197
Feinberg, A. P., Cui, H., & Ohlsson, R. (2002). DNA methylation and genomic imprinting: insights from cancer into epigenetic mechanisms. *Seminars in Cancer Biology*, *12*(5), 389–398. https://doi.org/10.1016/S1044-579X(02)00059-7
Feng, J., Zhou, Y., Campbell, S. L., Le, T., Li, E., Sweatt, J. D., … Fan, G. (2010). Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. *Nature Neuroscience*, *13*(4), 423–430. https://doi.org/10.1038/nn.2514
Ferguson-Smith, A. C., Sasaki, H., Cattanach, B. M., & Surani, M. A. (1993). Parental-origin-specific epigenetic modification of the mouse H19 gene. *Nature*, *362*(6422), 751–755. https://doi.org/10.1038/362751a0
Ferman, T. J., Boeve, B. F., Smith, G. E., Lin, S.-C., Silber, M. H., Pedraza, O., … Dickson, D. W. (2011). Inclusion of RBD improves the diagnostic classification of dementia with Lewy bodies. *Neurology*, *77*(9), 875–882. https://doi.org/10.1212/WNL.0b013e31822c9148
Ferrazzoli, D., Carter, A., Ustun, F. S., Palamara, G., Ortelli, P., Maestri, R., … Frazzitta, G. (2016, June 14). Dopamine replacement therapy, learning and reward prediction in Parkinson's disease: Implications for rehabilitation. *Frontiers in Behavioral Neuroscience*. Frontiers Media S.A. https://doi.org/10.3389/fnbeh.2016.00121
Ferrer, I., Marín, C., Rey, M. J., Ribalta, T., Goutan, E., Blanco, R., … Martí, E. (1999). BDNF and Full-length and Truncated TrkB Expression in Alzheimer Disease. Implications in Therapeutic Strategies. *Journal of Neuropathology and Experimental Neurology*, *58*(7), 729–739. https://doi.org/10.1097/00005072-199907000-00007
Fetahu, I. S., Ma, D., Rabidou, K., Argueta, C., Smith, M., Liu, H., … Shi, Y. G. (2019). Epigenetic signatures of methylated DNA cytosine in Alzheimer's disease. *Science Advances*, *5*(8). https://doi.org/10.1126/sciadv.aaw2880
Figurov, A., Pozzo-Miller, L. D., Olafsson, P., Wang, T., & Lu, B. (1996). Regulation of synaptic responses to high-frequency stimulation and LTP by neurotrophins in the hippocampus. *Nature*, *381*(6584), 706–709. https://doi.org/10.1038/381706a0
Fitzgerald, E., Murphy, S., & Martinson, H. A. (2019). Alpha-Synuclein Pathology and the Role of the Microbiota in Parkinson's Disease. *Frontiers in Neuroscience*, *13*, 369. https://doi.org/10.3389/fnins.2019.00369
Flitton, M. (2018). Nucleotide methylation and neurocognitive phenotypes. *PhD Thesis*.
Flitton, M., Rielly, N., Warman, R., Warden, D., Smith, A. D., Macdonald, I. A., & Knight, H. M. (2019). Interaction of nutrition and genetics via DNMT3L-mediated DNA methylation determines cognitive decline. *Neurobiology of Aging*, *78*, 64–73. https://doi.org/10.1016/j.neurobiolaging.2019.02.001
Flusberg, B. A., Webster, D. R., Lee, J. H., Travers, K. J., Olivares, E. C., Clark, T. A., … Turner, S. W. (2010). Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nature Methods*, *7*(6), 461–465. https://doi.org/10.1038/nmeth.1459
Foraker, J., Millard, S. P., Leong, L., Thomson, Z., Chen, S., Keene, C. D., … Yu, C.-E.

(2015). The APOE Gene is Differentially Methylated in Alzheimer's Disease. *Journal of Alzheimer's Disease*, *48*(3), 745–755. https://doi.org/10.3233/JAD-143060

Frauer, C., Hoffmann, T., Bultmann, S., Casa, V., Cardoso, M. C., Antes, I., & Leonhardt, H. (2011). Recognition of 5-Hydroxymethylcytosine by the Uhrf1 SRA Domain. *PLoS ONE*, *6*(6), e21306. https://doi.org/10.1371/journal.pone.0021306

Friedman, W. J. (2000). Neurotrophins induce death of hippocampal neurons via the p75 receptor. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *20*(17), 6340–6346. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10964939

Fuchikami, M., Morinobu, S., Segawa, M., Okamoto, Y., Yamawaki, S., Ozaki, N., … Terao, T. (2011). DNA Methylation Profiles of the Brain-Derived Neurotrophic Factor (BDNF) Gene as a Potent Diagnostic Biomarker in Major Depression. *PLoS ONE*, *6*(8), e23881. https://doi.org/10.1371/journal.pone.0023881

Fukumoto, H., Ingelsson, M., Gårevik, N., Wahlund, L. O., Nukina, N., Yaguchi, Y., … Irizarry, M. C. (2003). APOE ε3/ε4 heterozygotes have an elevated proportion of apolipoprotein E4 in cerebrospinal fluid relative to plasma, independent of Alzheimer's disease diagnosis. *Experimental Neurology*, *183*(1), 249–253. https://doi.org/10.1016/S0014-4886(03)00088-8

Gan-Or, Z., Dion, P. A., & Rouleau, G. A. (2015). Genetic perspective on the role of the autophagy-lysosome pathway in Parkinson disease. *Autophagy*, *11*(9), 1443–1457. https://doi.org/10.1080/15548627.2015.1067364

Gebhard, C., Benner, C., Ehrich, M., Schwarzfischer, L., Schilling, E., Klug, M., … Rehli, M. (2010). General transcription factor binding at CpG islands in normal cells correlates with resistance to de novo DNA methylation in cancer cells. *Cancer Research*, *70*(4), 1398–1407. https://doi.org/10.1158/0008-5472.CAN-09-3406

German, D. C., Manaye, K. F., Sonsalla, P. K., & Brooks, B. A. (1992). Midbrain Dopaminergic Cell Loss in Parkinson's Disease and MPTP-Induced Parkinsonism: Sparing of Calbindin-D25k—Containing Cells. *Annals of the New York Academy of Sciences*, *648*(1), 42–62. https://doi.org/10.1111/j.1749-6632.1992.tb24523.x

Gervin, K., Andreassen, B. K., Hjorthaug, H. S., Carlsen, K. C. L., Carlsen, K. H., Undlien, D. E., … Munthe-Kaas, M. C. (2016). Intra-individual changes in DNA methylation not mediated by cell-type composition are correlated with aging during childhood. *Clinical Epigenetics*, *8*(1), 110. https://doi.org/10.1186/s13148-016-0277-3

Gharzi, V., Gangi, M., Sayad, A., Mazdeh, M., Arsang-Jang, S., & Taheri, M. (2018). Expression Analysis of BDNF Gene and BDNF-AS Long Noncoding RNA in Whole Blood Samples of Multiple Sclerosis Patients: Not Always a Negative Correlation between Them. *Iranian Journal of Allergy, Asthma, and Immunology*, *17*(6), 548–556. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/30644699

Giannakopoulos, P., Herrmann, F. R., Bussière, T., Bouras, C., Kövari, E., Perl, D. P., … Hof, P. R. (2003). Tangle and neuron numbers, but not amyloid load, predict cognitive status in Alzheimer's disease. *Neurology*, *60*(9), 1495–1500. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12743238

Gibbs, J. R., van der Brug, M. P., Hernandez, D. G., Traynor, B. J., Nalls, M. A., Lai, S.

L., … Singleton, A. B. (2010). Abundant quantitative trait loci exist for DNA methylation and gene expression in Human Brain. *PLoS Genetics*, *6*(5), 29. https://doi.org/10.1371/journal.pgen.1000952

Gibbs, R. A., Boerwinkle, E., Doddapaneni, H., Han, Y., Korchina, V., Kovar, C., … Rasheed, A. (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68–74. https://doi.org/10.1038/nature15393

Ginno, P. A., Lott, P. L., Christensen, H. C., Korf, I., & Chédin, F. (2012). R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Molecular Cell*, *45*(6), 814–825. https://doi.org/10.1016/j.molcel.2012.01.017

Giovannoni, G., O'Sullivan, J. D., Turner, K., Manson, A. J., & Lees, A. J. L. (2000). Hedonistic homeostatic dysregulation in patients with Parkinson's disease on dopamine replacement therapies. *Journal of Neurology Neurosurgery and Psychiatry*, *68*(4), 423–428. https://doi.org/10.1136/jnnp.68.4.423

Globisch, D., Münzel, M., Müller, M., Michalakis, S., Wagner, M., Koch, S., … Carell, T. (2010). Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PloS One*, *5*(12), e15367. https://doi.org/10.1371/journal.pone.0015367

Goldstone, A. P. (2004). Prader-Willi syndrome: advances in genetics, pathophysiology and treatment. *Trends in Endocrinology & Metabolism*, *15*(1), 12–20. https://doi.org/10.1016/J.TEM.2003.11.003

Goll, M. G., & Bestor, T. H. (2005). EUKARYOTIC CYTOSINE METHYLTRANSFERASES. *Annual Review of Biochemistry*, *74*(1), 481–514. https://doi.org/10.1146/annurev.biochem.74.010904.153721

Gottesman, I. I., & Wolfgram, D. L. (1991). *Schizophrenia genesis: The origins of madness*. New York: W. H. Freeman.

Gowher, H., Liebert, K., Hermann, A., Xu, G., & Jeltsch, A. (2005). Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. *The Journal of Biological Chemistry*, *280*(14), 13341–13348. https://doi.org/10.1074/jbc.M413412200

Green, R. C., Roberts, J. S., Cupples, L. A., Relkin, N. R., Whitehouse, P. J., Brown, T., … Farrer, L. A. (2009). Disclosure of *APOE* Genotype for Risk of Alzheimer's Disease. *New England Journal of Medicine*, *361*(3), 245–254. https://doi.org/10.1056/NEJMoa0809578

Griffith, J. S., & Mahler, H. R. (1969). DNA Ticketing Theory of Memory. *Nature*, *223*(5206), 580–582. https://doi.org/10.1038/223580a0

Guil, S., & Esteller, M. (2012, November). Cis-acting noncoding RNAs: Friends and foes. *Nature Structural and Molecular Biology*. https://doi.org/10.1038/nsmb.2428

Guo, H., Zhu, P., Yan, L., Li, R., Hu, B., Lian, Y., … Qiao, J. (2014). The DNA methylation landscape of human early embryos. *Nature*, *511*(7511), 606–610. https://doi.org/10.1038/nature13544

Guo, J. L., Covell, D. J., Daniels, J. P., Iba, M., Stieber, A., Zhang, B., … Lee, V. M. Y. (2013). Distinct α-Synuclein Strains Differentially Promote Tau Inclusions in Neurons. *Cell*, *154*(1), 103–117. https://doi.org/10.1016/j.cell.2013.05.057

Guo, J U, Su, Y., Shin, J. H., Shin, J., Li, H., Xie, B., … Song, H. (2014). Distribution, recognition and regulation of non-CpG methylation in the adult mammalian

brain. *Nat Neurosci*, *17*(2), 215–222. https://doi.org/10.1038/nn.3607

Guo, Junjie U, Szulwach, K. E., Su, Y., Li, Y., Yao, B., Xu, Z., … Song, H. (2014). Genome-wide antagonism between 5-hydroxymethylcytosine and DNA methylation in the adult mouse brain. *Frontiers in Biology*, *9*(1), 66–74. https://doi.org/10.1007/s11515-014-1295-1

Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., … Chang, H. Y. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*, *464*(7291), 1071–1076. https://doi.org/10.1038/nature08975

Haggarty, P., Hoad, G., Harris, S. E., Starr, J. M., Fox, H. C., Deary, I. J., & Whalley, L. J. (2010). Human Intelligence and Polymorphisms in the DNA Methyltransferase Genes Involved in Epigenetic Marking. *PLoS ONE*, *5*(6), e11329. https://doi.org/10.1371/journal.pone.0011329

Hahn, M. A., Qiu, R., Wu, X., Li, A. X., Zhang, H., Wang, J., … Lu, Q. (2013). Dynamics of 5-hydroxymethylcytosine and chromatin marks in Mammalian neurogenesis. *Cell Reports*, *3*(2), 291–300. https://doi.org/10.1016/j.celrep.2013.01.011

Hahn, M. A., Szabó, P. E., & Pfeifer, G. P. (2014). 5-Hydroxymethylcytosine: a stable or transient DNA modification? *Genomics*, *104*(5), 314–323. https://doi.org/10.1016/j.ygeno.2014.08.015

Haig, M., Flitton, M., Ruzov, A., & Knight, H. (n.d.). 5-carboxylcytosine 5caC abundance across human brain regions and neuronal populations. *Unpublished*.

Hariri, A. R., Goldberg, T. E., Mattay, V. S., Kolachana, B. S., Callicott, J. H., Egan, M. F., & Weinberger, D. R. (2003). Brain-derived neurotrophic factor val66met polymorphism affects human memory-related hippocampal activity and predicts memory performance. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *23*(17), 6690–6694. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12890761

Hashimoto, T., Serrano-Pozo, A., Hori, Y., Adams, K. W., Takeda, S., Banerji, A. O., … Hyman, B. T. (2012). Apolipoprotein E, Especially Apolipoprotein E4, Increases the Oligomerization of Amyloid Peptide. *Journal of Neuroscience*, *32*(43), 15181–15192. https://doi.org/10.1523/JNEUROSCI.1542-12.2012

Hatters, D. M., Peters-Libeu, C. A., & Weisgraber, K. H. (2006). Apolipoprotein E structure: insights into function. *Trends in Biochemical Sciences*, *31*(8), 445–454. https://doi.org/10.1016/J.TIBS.2006.06.008

Hawrylycz, M. J., Lein, E. S., Guillozet-Bongaarts, A. L., Shen, E. H., Ng, L., Miller, J. A., … Jones, A. R. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, *489*(7416), 391–399. https://doi.org/10.1038/nature11405

Hendrich, B., & Bird, A. (1998). Identification and Characterization of a Family of Mammalian Methyl-CpG Binding Proteins. *Molecular and Cellular Biology*, *18*(11), 6538–6547. https://doi.org/10.1128/mcb.18.11.6538

Herman, J. G., & Baylin, S. B. (2003). Gene Silencing in Cancer in Association with Promoter Hypermethylation. *New England Journal of Medicine*, *349*(21), 2042–2054. https://doi.org/10.1056/NEJMra023075

Herz, J., & Chen, Y. (2006). Reelin, lipoprotein receptors and synaptic plasticity. *Nature Reviews Neuroscience*, *7*(11), 850–859. https://doi.org/10.1038/nrn2009

Hing, B., Sathyaputri, L., & Potash, J. B. (2018). A comprehensive review of genetic and epigenetic mechanisms that regulate *BDNF* expression and function with relevance to major depressive disorder. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *177*(2), 143–167. https://doi.org/10.1002/ajmg.b.32616

Hock, C., Heese, K., Hulette, C., Rosenberg, C., & Otten, U. (2000). Region-Specific Neurotrophin Imbalances in Alzheimer Disease. *Archives of Neurology*, *57*(6), 846. https://doi.org/10.1001/archneur.57.6.846

Holliday, R., & Pugh, J. E. (1975). DNA modification mechanisms and gene activity during development. - PubMed - NCBI. *Science*, *187*(4173), 226–232. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/1111098

Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biology*, *14*(10), R115. https://doi.org/10.1186/gb-2013-14-10-r115

Horvath, S., Langfelder, P., Kwak, S., Aaronson, J., Rosinski, J., Vogt, T. F., … Yang, X. W. (2016). Huntington's disease accelerates epigenetic aging of human brain and disrupts DNA methylation levels. *Aging*, *8*(7), 1485–1512. https://doi.org/10.18632/aging.101005

Horvath, S., & Ritz, B. R. (2015). Increased epigenetic age and granulocyte counts in the blood of Parkinson's disease patients. *Aging*, *7*(12), 1130–1142. https://doi.org/10.18632/aging.100859

Hu, L., Lu, J., Cheng, J., Rao, Q., Li, Z., Hou, H., … Xu, Y. (2015). Structural insight into substrate preference for TET-mediated oxidation. *Nature*, *527*(7576), 118–122. https://doi.org/10.1038/nature15713

Hu, Y., Li, C., Wang, Y., Li, Q., Liu, Y., Liao, S., … Xu, H. (2018). Analysis of COMT Val158Met polymorphisms and methylation in Chinese male schizophrenia patients with homicidal behavior. *International Journal of Legal Medicine*, *132*(6), 1537–1544. https://doi.org/10.1007/s00414-018-1773-0

Huang, R., Huang, J., Cathcart, H., Smith, S., & Poduslo, S. E. (2007). Genetic variants in brain-derived neurotrophic factor associated with Alzheimer's disease. *Journal of Medical Genetics*, *44*(2), e66. https://doi.org/10.1136/jmg.2006.044883

Huang, Y., Qin Liu, X., Wyss-Coray, T., Brecht, W. J., Sanan, D. A., & Mahley, R. W. (2001). *Apolipoprotein E fragments present in Alzheimer's disease brains induce neurofibrillary tangle-like intracellular inclusions in neurons*. Retrieved from www.pnas.orgcgidoi10.1073pnas.151254698

Hughes, P. E., Alexi, T., Walton, M., Williams, C. E., Dragunow, M., G. Clark, R., & Gluckman, P. D. (1999). Activity and injury-dependent expression of inducible transcription factors, growth factors and apoptosis-related genes within the central nervous system. *Progress in Neurobiology*, *57*(4), 421–450. https://doi.org/10.1016/S0301-0082(98)00057-4

Husain, M. (2017). Transdiagnostic neurology: neuropsychiatric symptoms in neurodegenerative diseases. *Brain*, *140*(6), 1535–1536. https://doi.org/10.1093/brain/awx115

Ikeda, T., & Yamada, M. (2010). Risk factors for Alzheimer's disease. *Brain and Nerve = Shinkei Kenkyu No Shinpo*, *62*(7), 679–690. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20675872

Ikegame, T., Bundo, M., Murata, Y., Kasai, K., Kato, T., & Iwamoto, K. (2013). DNA

methylation of the BDNF gene and its relevance to psychiatric disorders. *Journal of Human Genetics*, *58*(7), 434–438. https://doi.org/10.1038/jhg.2013.65

Infinium MethylationEPIC Kit | Methylation profiling array for EWAS. (2019). Retrieved June 12, 2019, from https://www.illumina.com/products/by-type/microarray-kits/infinium-methylation-epic.html?langsel=/vg/

Ingelsson, M., & Hyman, B. T. (2002). Disordered proteins in dementia. *Annals of Medicine*, *34*(4), 259–271. https://doi.org/10.1080/078538902320322529

Irier, H. A., & Jin, P. (2012). Dynamics of DNA methylation in aging and Alzheimer's disease. *DNA and Cell Biology*, *31 Suppl 1*(Suppl 1), S42-8. https://doi.org/10.1089/dna.2011.1565

Irizarry, R. A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., … Webster, M. (2009). Genome-wide methylation analysis of human colon cancer reveals similar hypo-and hypermethylation at conserved tissue-specific CpG island shores. *Nature Genetics*, *41*(2), 178. https://doi.org/10.1038/ng.298.Genome-wide

Issa, J. P. (2004). CpG island methylator phenotype in cancer. *Nature Reviews Cancer*, *4*(12), 988–993. https://doi.org/10.1038/nrc1507

Ito, S., D'Alessio, A. C., Taranova, O. V, Hong, K., Sowers, L. C., & Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*, *466*(7310), 1129–1133. https://doi.org/10.1038/nature09303

Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A., … Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science (New York, N.Y.)*, *333*(6047), 1300–1303. https://doi.org/10.1126/science.1210597

Iurlaro, M., Ficz, G., Oxley, D., Raiber, E.-A., Bachman, M., Booth, M. J., … Reik, W. (2013). A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation. *Genome Biology*, *14*(10), R119. https://doi.org/10.1186/gb-2013-14-10-r119

Iwamoto, K., Bundo, M., Ueda, J., Oldham, M. C., Ukai, W., Hashimoto, E., … Kato, T. (2011). Neurons show distinctive DNA methylation profile and higher interindividual variations compared with non-neurons. *Genome Research*, *21*(5), 688–696. https://doi.org/10.1101/gr.112755.110

Je, H. S., Yang, F., Ji, Y., Nagappan, G., Hempstead, B. L., & Lu, B. (2012). Role of pro-brain-derived neurotrophic factor (proBDNF) to mature BDNF conversion in activity-dependent competition at developing neuromuscular synapses. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(39), 15924–15929. https://doi.org/10.1073/pnas.1207767109

Jellinger, K. A., & Korczyn, A. D. (2018). Are dementia with Lewy bodies and Parkinson's disease dementia the same disease? *BMC Medicine*, *16*(1), 34. https://doi.org/10.1186/s12916-018-1016-8

Ji, D., Lin, K., Song, J., & Wang, Y. (2014). Effects of Tet-induced oxidation products of 5-methylcytosine on Dnmt1- and DNMT3a-mediated cytosine methylation. *Molecular BioSystems*, *10*(7), 1749. https://doi.org/10.1039/c4mb00150h

Jia, D., Jurkowska, R. Z., Zhang, X., Jeltsch, A., & Cheng, X. (2007). Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature*, *449*(7159), 248–251. https://doi.org/10.1038/nature06146

Jiang, L., Zhang, J., Wang, J.-J., Wang, L., Zhang, L., Li, G., … Liu, J. (2013). Sperm, but Not Oocyte, DNA Methylome Is Inherited by Zebrafish Early Embryos. *Cell*, *153*(4), 773–784. https://doi.org/10.1016/J.CELL.2013.04.041

Jin, S.-G., Kadam, S., & Pfeifer, G. P. (2010). Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Research*, *38*(11), e125–e125. https://doi.org/10.1093/nar/gkq223

Jin, S.-G., Wu, X., Li, A. X., & Pfeifer, G. P. (2011). Genomic mapping of 5-hydroxymethylcytosine in the human brain. *Nucleic Acids Research*, *39*(12), 5015–5024. https://doi.org/10.1093/nar/gkr120

Jin, S. G., Zhang, Z.-M., Dunwell, T. L., Harter, M. R., Wu, X., Johnson, J., … Pfeifer, G. P. (2016). Tet3 Reads 5-Carboxylcytosine through Its CXXC Domain and Is a Potential Guardian against Neurodegeneration. *Cell Reports*, *14*(3), 493–505. https://doi.org/10.1016/j.celrep.2015.12.044

Jin, Z., & Liu, Y. (2018). DNA methylation in human diseases. *Genes & Diseases*, *5*(1), 1–8. https://doi.org/10.1016/j.gendis.2018.01.002

Jones, P. L., Jan Veenstra, G. C., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., … Wolffe, A. P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature Genetics*, *19*(2), 187–191. https://doi.org/10.1038/561

Julie A. Morrow, ‡,§, Mark L. Segall, ‖, Sissel Lund-Katz, ‖, Michael C. Phillips, ‖, Mark Knapp, ⊥, Bernhard Rupp, ⊥ and, & Karl H. Weisgraber*, ‡,§,#. (2000). Differences in Stability among the Human Apolipoprotein E Isoforms Determined by the Amino-Terminal Domain†. https://doi.org/10.1021/BI000099M

Kang, H.-J., Kim, J.-M., Bae, K.-Y., Kim, S.-W., Shin, I.-S., Kim, H.-R., … Yoon, J.-S. (2015). Longitudinal associations between BDNF promoter methylation and late-life depression. *Neurobiology of Aging*, *36*(4), 1764.e1-1764.e7. https://doi.org/10.1016/j.neurobiolaging.2014.12.035

Kang, H., & Schuman, E. (1995). Long-lasting neurotrophin-induced enhancement of synaptic transmission in the adult hippocampus. *Science*, *267*(5204), 1658–1662. https://doi.org/10.1126/science.7886457

Kareta, M. S., Botello, Z. M., Ennis, J. J., Chou, C., & Chédin, F. (2006). Reconstitution and mechanism of the stimulation of de novo methylation by human DNMT3L. *The Journal of Biological Chemistry*, *281*(36), 25893–25902. https://doi.org/10.1074/jbc.M603140200

Katoh-Semba, R., Wakako, R., Komori, T., Shigemi, H., Miyazaki, N., Ito, H., … Nakayama, A. (2007). Age-related changes in BDNF protein levels in human serum: differences between autism cases and normal controls. *International Journal of Developmental Neuroscience*, *25*(6), 367–372. https://doi.org/10.1016/J.IJDEVNEU.2007.07.002

Kegeles, L. S., Abi-Dargham, A., Zea-Ponce, Y., Rodenhiser-Hill, J., Mann, J. J., Van Heertum, R. L., … Laruelle, M. (2000). Modulation of amphetamine-induced striatal dopamine release by ketamine in humans: Implications for schizophrenia. *Biological Psychiatry*, *48*(7), 627–640. https://doi.org/10.1016/S0006-3223(00)00976-8

Kellinger, M. W., Song, C.-X., Chong, J., Lu, X.-Y., He, C., & Wang, D. (2012). 5-

formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription. *Nature Structural & Molecular Biology*, *19*(8), 831–833. https://doi.org/10.1038/nsmb.2346

Keshavarzian, A., Green, S. J., Engen, P. A., Voigt, R. M., Naqib, A., Forsyth, C. B., … Shannon, K. M. (2015). Colonic bacterial composition in Parkinson's disease. *Movement Disorders*, *30*(10), 1351–1360. https://doi.org/10.1002/mds.26307

Kikuchi, S. (2007). Motor fluctuations in Parkinson's disease. In *Journal of Neurology* (Vol. 254, pp. 32–40). Springer. https://doi.org/10.1007/s00415-007-5006-6

Kim, J., Yoon, H., Basak, J., & Kim, J. (2014). Apolipoprotein E in synaptic plasticity and Alzheimer's disease: potential cellular and molecular mechanisms. *Molecules and Cells*, *37*(11), 767–776. https://doi.org/10.14348/molcells.2014.0248

Kim, S., Kwon, S.-H., Kam, T.-I., Panicker, N., Karuppagounder, S. S., Lee, S., … Ko, H. S. (2019). Transneuronal Propagation of Pathologic α-Synuclein from the Gut to the Brain Models Parkinson's Disease. *Neuron*, *0*(0). https://doi.org/10.1016/j.neuron.2019.05.035

Kim, W. S., Kågedal, K., & Halliday, G. M. (2014). Alpha-synuclein biology in Lewy body diseases. *Alzheimer's Research & Therapy*, *6*(5–8), 73. https://doi.org/10.1186/s13195-014-0073-2

Kim, Y. S., & Joh, T. H. (2006, August 31). Microglia, major player in the brain inflammation: Their roles in the pathogenesis of Parkinson's disease. *Experimental and Molecular Medicine*. Korean Society of Med. Biochemistry and Mol. Biology. https://doi.org/10.1038/emm.2006.40

Kimura, H., & Shiota, K. (2003). Methyl-CpG-binding protein, MeCP2, is a target molecule for maintenance DNA methyltransferase, Dnmt1. *Journal of Biological Chemistry*, *278*(7), 4806–4812. https://doi.org/10.1074/jbc.M209923200

Kitsera, N., Allgayer, J., Parsa, E., Geier, N., Rossa, M., Carell, T., & Khobta, A. (2017). Functional impacts of 5-hydroxymethylcytosine, 5-formylcytosine, and 5-carboxycytosine at a single hemi-modified CpG dinucleotide in a gene promoter. *Nucleic Acids Research*, *45*(19), 11033–11042. https://doi.org/10.1093/nar/gkx718

Klein, C., & Westenberger, A. (2012). Genetics of Parkinson's disease. *Cold Spring Harbor Perspectives in Medicine*, *2*(1), a008888. https://doi.org/10.1101/cshperspect.a008888

Kovalchuk, Y., Hanse, E., Kafitz, K. W., & Konnerth, A. (2002). Postsynaptic Induction of BDNF-Mediated Long-Term Potentiation. *Science (New York, N.Y.)*, *295*(5560), 1729–1734. https://doi.org/10.1126/science.1067766

Kozlenkov, A., Li, J., Apontes, P., Hurd, Y. L., Byne, W. M., Koonin, E. V., … Dracheva, S. (2018). A unique role for DNA (hydroxy)methylation in epigenetic regulation of human inhibitory neurons. *Science Advances*, *4*(9), eaau6190. https://doi.org/10.1126/sciadv.aau6190

Kozlenkov, A., Wang, M., Roussos, P., Rudchenko, S., Barbu, M., Bibikova, M., … Dracheva, S. (2016). Substantial DNA methylation differences between two major neuronal subtypes in human brain. *Nucleic Acids Research*, *44*(6), 2593–2612. https://doi.org/10.1093/nar/gkv1304

Kriaucionis, S., & Heintz, N. (2009). The nuclear DNA base, 5-hydroxymethylcytosine is present in brain and enriched in Purkinje neurons. *Science (New York, N.Y.)*, *324*(5929), 929.

Kurdyukov, S., & Bullock, M. (2016). DNA Methylation Analysis: Choosing the Right Method. *Biology*, *5*(1). https://doi.org/10.3390/biology5010003

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., … Morgan, M. J. (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*(6822), 860–921. https://doi.org/10.1038/35057062

Lane, R., He, Y., Morris, C., Leverenz, J. B., Emre, M., & Ballard, C. (2009). BuChE-K and APOE ϵ4 allele frequencies in Lewy body dementias, and influence of genotype and hyperhomocysteinemia on cognitive decline. *Movement Disorders*, *24*(3), 392–400. https://doi.org/10.1002/mds.22357

Langston, J. W. (2006). The Parkinson's complex: parkinsonism is just the tip of the iceberg. *Annals of Neurology*, *59*(4), 591–596. https://doi.org/10.1002/ana.20834

Laruelle, M., Abi-Dargham, A., Gil, R., Kegeles, L., & Innis, R. (1999). Increased dopamine transmission in schizophrenia: relationship to illness phases. *Biological Psychiatry*, *46*(1), 56–72. https://doi.org/10.1016/S0006-3223(99)00067-0

Laruelle, M., Abi-Dargham, A., Van Dyck, C. H., Gil, R., D'Souza, C. D., Erdos, J., … Innis, R. B. (1996). Single photon emission computerized tomography imaging of amphetamine-induced dopamine release in drug-free schizophrenic subjects. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(17), 9235–9240. https://doi.org/10.1073/pnas.93.17.9235

Lashley, T., Gami, P., Valizadeh, N., Li, A., Revesz, T., & Balazs, R. (2015). Alterations in global DNA methylation and hydroxymethylation are not detected in Alzheimer's disease. *Neuropathology and Applied Neurobiology*, *41*(4), 497–506. https://doi.org/10.1111/nan.12183

Latos, P. A., Pauler, F. M., Koerner, M. V., Şenergin, H. B., Hudson, Q. J., Stocsits, R. R., … Barlow, D. P. (2012). Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. *Science*, *338*(6113), 1469–1472. https://doi.org/10.1126/science.1228110

Law, J. A., & Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics*, *11*(3), 204–220. https://doi.org/10.1038/nrg2719

Lee, E.-G., Tulloch, J., Chen, S., Leong, L., Saxton, A. D., Kraemer, B., … Yu, C.-E. (2020). Redefining transcriptional regulation of the APOE gene and its association with Alzheimer's disease. *PLOS ONE*, *15*(1), e0227667. https://doi.org/10.1371/journal.pone.0227667

Lee, J., Fukumoto, H., Orne, J., Klucken, J., Raju, S., Vanderburg, C. R., … Ingelsson, M. (2005). Decreased levels of BDNF protein in Alzheimer temporal cortex are independent of BDNF polymorphisms. *Experimental Neurology*, *194*(1), 91–96. https://doi.org/10.1016/J.EXPNEUROL.2005.01.026

Lee, J. G., Shin, B. S., You, Y. S., Kim, J. E., Yoon, S. W., Jeon, D. W., … Kim, Y. H. (2009). Decreased serum brain-derived neurotrophic factor levels in elderly korean with dementia. *Psychiatry Investigation*, *6*(4), 299–305. https://doi.org/10.4306/pi.2009.6.4.299

Lehtimäki, T., Pirttilä, T., Mehta, P., Wisniewski, H., Frey, H., & Nikkari, T. (1995). Apolipoprotein E (apoE) polymorphism and its influence on ApoE concentrations in the cerebrospinal fluid in Finnish patients with Alzheimer's

disease. *Human Genetics*, *95*(1). https://doi.org/10.1007/BF00225071

Leontiou, C. A., Hadjidaniel, M. D., Mina, P., Antoniou, P., Ioannides, M., & Patsalis, P. C. (2015). Bisulfite Conversion of DNA: Performance Comparison of Different Kits and Methylation Quantitation of Epigenetic Biomarkers that Have the Potential to Be Used in Non-Invasive Prenatal Testing. *PloS One*, *10*(8), e0135058. https://doi.org/10.1371/journal.pone.0135058

Levenson, J. M., Roth, T. L., Lubin, F. D., Miller, C. A., Huang, I.-C., Desai, P., … Sweatt, J. D. (2006). Evidence that DNA (cytosine-5) methyltransferase regulates synaptic plasticity in the hippocampus. *The Journal of Biological Chemistry*, *281*(23), 15763–15773. https://doi.org/10.1074/jbc.M511767200

Levine, M. E., Lu, A. T., Bennett, D. A., & Horvath, S. (2015). Epigenetic age of the pre-frontal cortex is associated with neuritic plaques, amyloid load, and Alzheimer's disease related cognitive functioning. *Aging*, *7*(12), 1198–1211. https://doi.org/10.18632/aging.100864

Lewis, D. A., & Lieberman, J. A. (2000). Catching up on schizophrenia: Natural history and neurobiology. *Neuron*. Cell Press. https://doi.org/10.1016/S0896-6273(00)00111-2

Li, E., Beard, C., & Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. *Nature*, *366*(6453), 362–365. https://doi.org/10.1038/366362a0

Li, E., Bestor, T. H., & Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, *69*(6), 915–926. https://doi.org/10.1016/0092-8674(92)90611-F

Li, E., & Zhang, Y. (2014). DNA methylation in mammals. *Cold Spring Harbor Perspectives in Biology*, *6*(5), a019133. https://doi.org/10.1101/cshperspect.a019133

Liddell, M. B., Lovestone, S., & Owen, M. J. (2001). Genetic risk of Alzheimer's disease: advising relatives. *British Journal of Psychiatry*, *178*(1), 7–11. https://doi.org/10.1192/bjp.178.1.7

Lin, I.-H., Chen, Y.-F., & Hsu, M.-T. (2017). Correlated 5-Hydroxymethylcytosine (5hmC) and Gene Expression Profiles Underpin Gene and Organ-Specific Epigenetic Regulation in Adult Mouse Brain and Liver. *PLOS ONE*, *12*(1), e0170779. https://doi.org/10.1371/journal.pone.0170779

Lippa, C. F., Fujiwara, H., Mann, D. M., Giasson, B., Baba, M., Schmidt, M. L., … Trojanowski, J. Q. (1998). Lewy bodies contain altered alpha-synuclein in brains of many familial Alzheimer's disease patients with mutations in presenilin and amyloid precursor protein genes. *The American Journal of Pathology*, *153*(5), 1365–1370. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9811326

Lister, R., & Mukamel, E. A. (2015). Turning over DNA methylation in the mind. *Frontiers in Neuroscience*, *9*, 252. https://doi.org/10.3389/fnins.2015.00252

Lister, R., Mukamel, E. A., Nery, J. R., Urich, M., Puddifoot, C. A., Johnson, N. D., … Ecker, J. R. (2013). Global epigenomic reconfiguration during mammalian brain development. *Science (New York, N.Y.)*, *341*(6146), 1237905. https://doi.org/10.1126/science.1237905

Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., … Ecker, J. R. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, *462*(7271), 315–322. https://doi.org/10.1038/nature08514

Liu, C. C., Liu, C. C., Kanekiyo, T., Xu, H., & Bu, G. (2013). Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. *Nature Reviews. Neurology*, *9*(2), 106–118. https://doi.org/10.1038/nrneurol.2012.263

Liu, J., Zhao, W., Ware, E. B., Turner, S. T., Mosley, T. H., & Smith, J. A. (2018). DNA methylation in the APOE genomic region is associated with cognitive function in African Americans. *BMC Medical Genomics*, *11*(1), 43. https://doi.org/10.1186/s12920-018-0363-9

Liu, L., Foroud, T., Xuei, X., Berrettini, W., Byerley, W., Coryell, W., … Nurnberger, J. I. (2008). Evidence of association between brain-derived neurotrophic factor gene and bipolar disorder. *Psychiatric Genetics*, *18*(6), 267–274. https://doi.org/10.1097/YPG.0b013e3283060f59

Liu, Q., & Zhang, J. (2014). Lipid metabolism in Alzheimer's disease. *Neurosci Bull*, *30*(2), 331–345. https://doi.org/10.1007/s12264-013-1410-3

Liu, S., Wang, J., Su, Y., Guerrero, C., Zeng, Y., Mitra, D., … Wang, Y. (2013). Quantitative assessment of Tet-induced oxidation products of 5-methylcytosine in cellular and tissue DNA. *Nucleic Acids Research*, *41*(13), 6421–6429. https://doi.org/10.1093/nar/gkt360

Liutkevičiūtė, Z., Lukinavičius, G., Masevičius, V., Daujotytė, D., & Klimašauskas, S. (2009). Cytosine-5-methyltransferases add aldehydes to DNA. *Nature Chemical Biology*, *5*(6), 400–402. https://doi.org/10.1038/nchembio.172

Liyanage, V. R. B., & Rastegar, M. (2014). Rett syndrome and MeCP2. *Neuromolecular Medicine*, *16*(2), 231–264. https://doi.org/10.1007/s12017-014-8295-9

Lommatzsch, M., Zingler, D., Schuhbaeck, K., Schloetcke, K., Zingler, C., Schuff-Werner, P., & Virchow, J. C. (2005). The impact of age, weight and gender on BDNF levels in human platelets and plasma. *Neurobiology of Aging*, *26*(1), 115–123. https://doi.org/10.1016/j.neurobiolaging.2004.03.002

Lu, B, Nagappan, G., & Lu, Y. (2014). BDNF and synaptic plasticity, cognitive function, and dysfunction. *Handbook of Experimental Pharmacology*, *220*, 223–250. https://doi.org/10.1007/978-3-642-45106-5_9

Lu, Bai, Nagappan, G., Guan, X., Nathan, P. J., & Wren, P. (2013). BDNF-based synaptic repair as a disease-modifying strategy for neurodegenerative diseases. *Nature Reviews. Neuroscience*, *14*(6), 401–416. https://doi.org/10.1038/nrn3505

Lu, X., Song, C.-X., Szulwach, K., Wang, Z., Weidenbacher, P., Jin, P., & He, C. (2013). Chemical modification-assisted bisulfite sequencing (CAB-Seq) for 5-carboxylcytosine detection in DNA. *Journal of the American Chemical Society*, *135*(25), 9315–9317. https://doi.org/10.1021/ja4044856

Lubin, F. D., Roth, T. L., & Sweatt, J. D. (2008). EPIGENETIC REGULATION OF BDNF GENE TRANSCRIPTION IN THE CONSOLIDATION OF FEAR MEMORY. *The Journal of Neuroscience*, *28*(42), 10576. https://doi.org/10.1523/JNEUROSCI.1786-08.2008

Lücking, C. B., Dürr, A., Bonifati, V., Vaughan, J., De Michele, G., Gasser, T., … Brice, A. (2000). Association between Early-Onset Parkinson's Disease and Mutations in the *Parkin* Gene. *New England Journal of Medicine*, *342*(21), 1560–1567. https://doi.org/10.1056/NEJM200005253422103

Lühnsdorf, B., Epe, B., & Khobta, A. (2014). Excision of uracil from transcribed DNA negatively affects gene expression. *Journal of Biological Chemistry*, *289*(32), 22008–22018. https://doi.org/10.1074/jbc.M113.521807

Lynch, M. A. (2004). Long-Term Potentiation and Memory. https://doi.org/10.1152/physrev.00014.2003.-One

Ma, Y., Smith, C. E., Lai, C.-Q., Irvin, M. R., Parnell, L. D., Lee, Y.-C., … Arnett, D. K. (2015). Genetic variants modify the effect of age on APOE methylation in the Genetics of Lipid Lowering Drugs and Diet Network study. *Aging Cell*, *14*(1), 49–59. https://doi.org/10.1111/acel.12293

Mahley, R. (1988). Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. *Science*, *240*(4852), 622–630. https://doi.org/10.1126/science.3283935

Mahley, R. W., & Huang, Y. (2012). Apolipoprotein e sets the stage: response to injury triggers neuropathology. *Neuron*, *76*(5), 871–885. https://doi.org/10.1016/j.neuron.2012.11.020

Mahley, R. W., & Rall, S. C. (2000). A POLIPOPROTEIN E: Far More Than a Lipid Transport Protein. *Annual Review of Genomics and Human Genetics*, *1*(1), 507–537. https://doi.org/10.1146/annurev.genom.1.1.507

Mahley, R. W., Weisgraber, K. H., & Huang, Y. (2009). Apolipoprotein E: structure determines function, from atherosclerosis to Alzheimer's disease to AIDS. *Journal of Lipid Research*, *50 Suppl*(Supplement), S183-8. https://doi.org/10.1194/jlr.R800069-JLR200

Mamta Tahiliani, K. P. K. Y. S. W. A. P. H. B. Y. B. S. A. L. M. I. D. R. L. L. A. A. R. (2009). Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science (New York, N.Y.)*, *324*(5929), 930.

Mancera-Páez, O., Estrada-Orozco, K., Mahecha, M., Cruz, F., Bonilla-Vargas, K., Sandoval, N., … Arboleda, H. (2019). Differential Methylation in APOE (Chr19; Exon Four; from 44,909,188 to 44,909,373/hg38) and Increased Apolipoprotein E Plasma Levels in Subjects with Mild Cognitive Impairment. *International Journal of Molecular Sciences*, *20*(6), 1394. https://doi.org/10.3390/ijms20061394

Mangialasche, F., Xu, W., Kivipelto, M., Costanzi, E., Ercolani, S., Pigliautile, M., … Mecocci, P. (2012). Tocopherols and tocotrienols plasma levels are associated with cognitive impairment. *Neurobiology of Aging*, *33*(10), 2282–2290. https://doi.org/10.1016/j.neurobiolaging.2011.11.019

Marcelis, M., van Os, J., Sham, P., Jones, P., Gilvarry, C., Cannon, M., … Murray, R. (1998). Obstetric complications and familial morbid risk of psychiatric disorders. *American Journal of Medical Genetics*, *81*(1), 29–36. https://doi.org/10.1002/(sici)1096-8628(19980207)81:1<29::aid-ajmg6>3.0.co;2-i

Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. *Trends in Genetics*, *24*(3), 133–141. https://doi.org/10.1016/J.TIG.2007.12.007

Marmorstein, R., & Trievel, R. C. (2009). Histone modifying enzymes: structures, mechanisms, and specificities. *Biochimica et Biophysica Acta*, *1789*(1), 58–68. https://doi.org/10.1016/j.bbagrm.2008.07.009

Mas, E., Dupuy, A. M., Artero, S., Portet, F., Cristol, J. P., Ritchie, K., & Touchon, J. (2006). Functional Vitamin E deficiency in ApoE4 patients with Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, *21*(3), 198–204. https://doi.org/10.1159/000090868

Masliah, E., Dumaop, W., Galasko, D., & Desplats, P. (2013). Distinctive patterns of DNA methylation associated with Parkinson disease. *Epigenetics*, *8*(10), 1030–1038. https://doi.org/10.4161/epi.25865

Mastroeni, D., Grover, A., Delvaux, E., Whiteside, C., Coleman, P. D., & Rogers, J. (2010). Epigenetic changes in Alzheimer's disease: Decrements in DNA methylation. *Neurobiology of Aging*, *31*(12), 2025–2037. https://doi.org/10.1016/j.neurobiolaging.2008.12.005

Matrisciano, F., Tueting, P., Dalal, I., Kadriu, B., Grayson, D. R., Davis, J. M., … Guidotti, A. (2013). Epigenetic modifications of GABAergic interneurons are associated with the schizophrenia-like phenotype induced by prenatal stress in mice. *Neuropharmacology*, *68*, 184–194. https://doi.org/10.1016/j.neuropharm.2012.04.013

Matsuda, N., Lu, H., Fukata, Y., Noritake, J., Gao, H., Mukherjee, S., … Poo, M.-M. (2009). Differential activity-dependent secretion of brain-derived neurotrophic factor from axon and dendrite. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(45), 14185–14198. https://doi.org/10.1523/JNEUROSCI.1863-09.2009

Matsumoto, L., Takuma, H., Tamaoka, A., Kurisaki, H., Date, H., Tsuji, S., & Iwata, A. (2010). CpG demethylation enhances alpha-synuclein expression and affects the pathogenesis of Parkinson's disease. *PLoS ONE*, *5*(11). https://doi.org/10.1371/journal.pone.0015522

McDade, E., & Bateman, R. J. (2017). Stop Alzheimer's before it starts. *Nature*, *547*(7662), 153–155. https://doi.org/10.1038/547153a

McDowell, I. (2001). Alzheimer's disease: insights from epidemiology. *Aging (Milan, Italy)*, *13*(3), 143–162. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11442298

McGeer, P. L., & Rogers, J. (1992). Anti–inflammatory agents as a therapeutic approach to Alzheimer's disease. *Neurology*, *42*(2), 447–448. https://doi.org/10.1212/wnl.42.2.447

McGregor, M. M., & Nelson, A. B. (2019). Circuit Mechanisms of Parkinson's Disease. *Neuron*, *101*(6), 1042–1056. https://doi.org/10.1016/j.neuron.2019.03.004

McHughen, S. A., Rodriguez, P. F., Kleim, J. A., Kleim, E. D., Crespo, L. M., Procaccio, V., & Cramer, S. C. (2010). BDNF Val66Met Polymorphism Influences Motor System Function in the Human Brain. *Cerebral Cortex*, *20*(5), 1254–1262. https://doi.org/10.1093/cercor/bhp189

Mckeith, I. G., Sci, M., Boeve, B. F., Dickson, D. W., Halliday, G., Taylor, J.-P., … Iranzo, A. (2017). Diagnosis and management of dementia with Lewy bodies Fourth consensus report of the DLB Consortium. https://doi.org/10.1212/WNL.0000000000004058

Meaburn, E. L., Schalkwyk, L. C., & Mill, J. (2010). Allele-specific methylation in the human genome: implications for genetic studies of complex disease. *Epigenetics*, *5*(7), 578–582. https://doi.org/10.4161/epi.5.7.12960

Mellen, M., Ayata, P., Dewell, S., Kriaucionis, S., & Heintz, N. (2012). MeCP2 binds to 5hmc enriched within active genes and accessible chromatin in the nervous system. *Cell*, *151*(7), 1417. https://doi.org/10.1016/J.CELL.2012.11.022

Meng, W. D., Sun, S.-J., Yang, J., Chu, R.-X., Tu, W., & Liu, Q. (2017). Elevated Serum Brain-Derived Neurotrophic Factor (BDNF) but not BDNF Gene Val66Met

Polymorphism Is Associated with Autism Spectrum Disorders. *Molecular Neurobiology*, *54*(2), 1167–1172. https://doi.org/10.1007/s12035-016-9721-9

Methylamp Whole Cell Bisulfite Modification Kit | EpiGentek. (2019). Retrieved June 12, 2019, from https://www.epigentek.com/catalog/methylamp-whole-cell-bisulfite-modification-kit-p-341.html

Miller, C. A., Gavin, C. F., White, J. A., Parrish, R. R., Honasoge, A., Yancey, C. R., … Sweatt, J. D. (2010). Cortical DNA methylation maintains remote memory. *Nature Neuroscience*, *13*(6), 664–666. https://doi.org/10.1038/nn.2560

Miller, C. A., & Sweatt, J. D. (2007). Covalent modification of DNA regulates memory formation. *Neuron*, *53*(6), 857–869. https://doi.org/10.1016/j.neuron.2007.02.022

Minichiello, L., Calella, A. M., Medina, D. L., Bonhoeffer, T., Klein, R., & Korte, M. (2002). Mechanism of TrkB-mediated hippocampal long-term potentiation. *Neuron*, *36*(1), 121–137. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12367511

Miranda-Morales, E., Meier, K., Sandoval-Carrillo, A., Salas-Pacheco, J., Vázquez-Cárdenas, P., & Arias-Carrión, O. (2017, July 18). Implications of DNA methylation in Parkinson's disease. *Frontiers in Molecular Neuroscience*. Frontiers Media S.A. https://doi.org/10.3389/fnmol.2017.00225

Miyazaki, K., Narita, N., Sakuta, R., Miyahara, T., Naruse, H., Okado, N., & Narita, M. (2004). Serum neurotrophin concentrations in autism and mental retardation: a pilot study. *Brain and Development*, *26*(5), 292–295. https://doi.org/10.1016/S0387-7604(03)00168-2

Modarresi, F., Faghihi, M. A., Lopez-Toledano, M. A., Fatemi, R. P., Magistri, M., Brothers, S. P., … Wahlestedt, C. (2012). Inhibition of natural antisense transcripts in vivo results in gene-specific transcriptional upregulation. *Nature Biotechnology*, *30*(5), 453–459. https://doi.org/10.1038/nbt.2158

Mooijman, D., Dey, S. S., Boisset, J. C., Crosetto, N., & Van Oudenaarden, A. (2016). Single-cell 5hmC sequencing reveals chromosome-wide cell-to-cell variability and enables lineage reconstruction. *Nature Biotechnology*, *34*(8), 852–856. https://doi.org/10.1038/nbt.3598

Morgan, H. D., Dean, W., Coker, H. A., Reik, W., & Petersen-Mahrt, S. K. (2004). Activation-induced Cytidine Deaminase Deaminates 5-Methylcytosine in DNA and Is Expressed in Pluripotent Tissues. *Journal of Biological Chemistry*, *279*(50), 52353–52360. https://doi.org/10.1074/jbc.M407695200

Morgen, K., Frölich, L., Tost, H., Plichta, M. M., Kölsch, H., Rakebrandt, F., … Meyer-Lindenberg, A. (2013). APOE-Dependent Phenotypes in Subjects with Mild Cognitive Impairment Converting to Alzheimer's Disease. *Journal of Alzheimer's Disease*, *37*(2), 389–401. https://doi.org/10.3233/JAD-130326

Morris, R. G. M., Anderson, E., Lynch, G. S., & Baudry, M. (1986). Selective impairment of learning and blockade of long-term potentiation by an N-methyl-D-aspartate receptor antagonist, AP5. *Nature*, *319*(6056), 774–776. https://doi.org/10.1038/319774a0

Moser, D. A., Paoloni-Giacobino, A., Stenz, L., Adouan, W., Manini, A., Suardi, F., … Schechter, D. S. (2015). BDNF Methylation and Maternal Brain Activity in a Violence-Related Sample. *PLOS ONE*, *10*(12), e0143427. https://doi.org/10.1371/journal.pone.0143427

Mullan, K., Williams, M. A., Cardwell, C. R., McGuinness, B., Passmore, P., Silvestri, G., … McKay, G. J. (2017). Serum concentrations of vitamin E and carotenoids are altered in Alzheimer's disease: A case-control study. *Alzheimer's and Dementia: Translational Research and Clinical Interventions*, *3*(3), 432–439. https://doi.org/10.1016/j.trci.2017.06.006

Murer, M. ., Yan, Q., & Raisman-Vozari, R. (2001). Brain-derived neurotrophic factor in the control human brain, and in Alzheimer's disease and Parkinson's disease. *Progress in Neurobiology*, *63*(1), 71–124. https://doi.org/10.1016/S0301-0082(00)00014-9

Murphy, M. P., LeVine, H., & III. (2010). Alzheimer's disease and the amyloid-beta peptide. *Journal of Alzheimer's Disease : JAD*, *19*(1), 311–323. https://doi.org/10.3233/JAD-2010-1221

Murrell, J. R., Price, B. M., Baiyewu, O., Gureje, O., Deeg, M., Hendrie, H., … Hall, K. (2006). The fourth *apolipoprotein E* haplotype found in the Yoruba of Ibadan. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *141B*(4), 426–427. https://doi.org/10.1002/ajmg.b.30295

Nabel, C. S., Jia, H., Ye, Y., Shen, L., Goldschmidt, H. L., Stivers, J. T., … Kohli, R. M. (2012). AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. *Nature Chemical Biology*, *8*(9), 751–758. https://doi.org/10.1038/nchembio.1042

Namba, Y., Tomonaga, M., Kawasaki, H., Otomo, E., & Ikeda, K. (1991). Apolipoprotein E immunoreactivity in cerebral amyloid deposits and neurofibrillary tangles in Alzheimer's disease and kuru plaque amyloid in Creutzfeldt-Jakob disease. *Brain Research*, *541*(1), 163–166. https://doi.org/10.1016/0006-8993(91)91092-f

Nan, X., Ng, H.-H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., & Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*, *393*(6683), 386–389. https://doi.org/10.1038/30764

Naoe, Y., Shinkai, T., Hori, H., Fukunaka, Y., Utsunomiya, K., Sakata, S., … Nakamura, J. (2007). No association between the brain-derived neurotrophic factor (BDNF) Val66Met polymorphism and schizophrenia in Asian populations: Evidence from a case-control study and meta-analysis. *Neuroscience Letters*, *415*(2), 108–112. https://doi.org/10.1016/j.neulet.2007.01.006

Nathan, B. P., Bellosta, S., Sanan, D. A., Weisgraber, K. H., Mahley, R. W., & Pitas, R. E. (1994). Differential effects of apolipoproteins E3 and E4 on neuronal growth in vitro. *Science (New York, N.Y.)*, *264*(5160), 850–852. https://doi.org/10.1126/science.8171342

Nelson, K. B., Grether, J. K., Croen, L. A., Dambrosia, J. M., Dickens, B. F., Jelliffe, L. L., … Phillips, T. M. (2001). Neuropeptides and neurotrophins in neonatal blood of children with autism or mental retardation. *Annals of Neurology*, *49*(5), 597–606. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11357950

Nelson, S. B., Sugino, K., & Hempel, C. M. (2006). The problem of neuronal cell types: a physiological genomics approach. *Trends in Neurosciences*, *29*(6), 339–345. https://doi.org/10.1016/J.TINS.2006.05.004

Neri, F., Incarnato, D., Krepelova, A., Parlato, C., & Oliviero, S. (2016). Methylation-assisted bisulfite sequencing to simultaneously map 5fC and 5caC on a genome-

wide scale for DNA demethylation analysis. *Nature Protocols*, *11*(7), 1191–1205. https://doi.org/10.1038/nprot.2016.063

Neri, F., Incarnato, D., & Oliviero, S. (2015). DNA methylation and demethylation dynamics. *Oncotarget*, *6*(33), 34049–34050. https://doi.org/10.18632/oncotarget.6039

Nestor, C. E., Ottaviano, R., Reddington, J., Sproul, D., Reinhardt, D., Dunican, D., … Meehan, R. R. (2012). Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Research*, *22*(3), 467–477. https://doi.org/10.1101/gr.126417.111

Nicholls, R. D., & Knepper, J. L. (2001). G ENOME O RGANIZATION , F UNCTION , AND I MPRINTING IN P RADER -W ILLI AND A NGELMAN S YNDROMES. *Annual Review of Genomics and Human Genetics*, *2*(1), 153–175. https://doi.org/10.1146/annurev.genom.2.1.153

Nickerson, D. A., Taylor, S. L., Fullerton, S. M., Weiss, K. M., Clark, A. G., Stengård, J. H., … Sing, C. F. (2000). Sequence Diversity and Large-Scale Typing of SNPs in the Human Apolipoprotein E Gene. *Genome Research*, *10*(10), 1532.

Nicoletti, C. F., Nonino, C. B., de Oliveira, B. A. P., Pinhel, M. A. de S., Mansego, M. L., Milagro, F. I., … Martinez, J. A. (2016). DNA Methylation and Hydroxymethylation Levels in Relation to Two Weight Loss Strategies: Energy-Restricted Diet or Bariatric Surgery. *Obesity Surgery*, *26*(3), 603–611. https://doi.org/10.1007/s11695-015-1802-8

Nieoullon, A., & Goff, L. K. (1992, October). Cellular interactions in the striatum involving neuronal systems using “classical” neurotransmitters: Possible functional implications. *Movement Disorders*. https://doi.org/10.1002/mds.870070404

Nilsen, T. W., & Graveley, B. R. (2010). Expansion of the eukaryotic proteome by alternative splicing. *Nature*, *463*(7280), 457–463. https://doi.org/10.1038/nature08909

Noctor, S. C., Martínez-Cerdeño, V., & Kriegstein, A. R. (2007). Contribution of Intermediate Progenitor Cells to Cortical Histogenesis. *Archives of Neurology*, *64*(5), 639. https://doi.org/10.1001/archneur.64.5.639

Nussbaum, R. L., & Ellis, C. E. (2003). Alzheimer’s Disease and Parkinson’s Disease. *New England Journal of Medicine*, *348*(14), 1356–1364. https://doi.org/10.1056/NEJM2003ra020003

O’Brien, J. S., & Sampson, E. L. (1965). Lipid composition of the normal human brain: gray matter, white matter, and myelin. *Journal of Lipid Research*, *6*(4), 537–544. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5865382

Obeso, J. A., Rodríguez-Oroz, M. C., & Zamarbide, I. (2005). Clinical features, pathophysiology and management of motor complications in parkinson’s disease. In *Principles of Treatment in Parkinson’s Disease* (pp. 99–115). Elsevier Inc. https://doi.org/10.1016/B978-0-7506-5428-9.50009-3

Okano, M., Xie, S., & Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature Genetics*, *19*(3), 219–220. https://doi.org/10.1038/890

Otani, J., Kimura, H., Sharif, J., Endo, T. A., Mishima, Y., Kawakami, T., … Tajima, S. (2013). Cell Cycle-Dependent Turnover of 5-Hydroxymethyl Cytosine in Mouse Embryonic Stem Cells. *PLoS ONE*, *8*(12), e82961.

https://doi.org/10.1371/journal.pone.0082961
Paik, Y. K., Chang, D. J., Reardon, C. A., Davies, G. E., Mahley, R. W., & Taylor, J. M. (1985). Nucleotide sequence and structure of the human apolipoprotein E gene. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.82.10.3445
Pang, P. T., Teng, H. K., Zaitsev, E., Woo, N. T., Sakata, K., Zhen, S., … Lu, B. (2004). Cleavage of proBDNF by tPA/Plasmin Is Essential for Long-Term Hippocampal Plasticity. *Science*, *306*(5695), 487–491. https://doi.org/10.1126/science.1100135
Park, M.-K., Lee, J.-C. J.-W., Lee, J.-C. J.-W., Hwang, S.-J., Roh, H. W., Hong, C. H., & Son, S. J. (2018). NLRP1 and NTN1, Deregulated Blood Differentially Methylated Regions in Mild Cognitive Impairment Patients. *Journal of Molecular Neuroscience*, *66*(4), 561–571. https://doi.org/10.1007/s12031-018-1180-5
Patiño-Parrado, I., Gómez-Jiménez, Á., López-Sánchez, N., & Frade, J. M. (2017). Strand-specific CpG hemimethylation, a novel epigenetic modification functional for genomic imprinting. *Nucleic Acids Research*, *45*(15), 8822–8834. https://doi.org/10.1093/nar/gkx518
Patterson, S. L., Abel, T., Deuel, T. A. ., Martin, K. C., Rose, J. C., & Kandel, E. R. (1996). Recombinant BDNF Rescues Deficits in Basal Synaptic Transmission and Hippocampal LTP in BDNF Knockout Mice. *Neuron*, *16*(6), 1137–1145. https://doi.org/10.1016/S0896-6273(00)80140-3
Peila, R., Rodriguez, B. L., & Launer, L. J. (2002). Type 2 diabetes, APOE gene, and the risk for dementia and related pathologies: The Honolulu-Asia Aging Study. *Diabetes*, *51*(4), 1256–1262. https://doi.org/10.2337/diabetes.51.4.1256
Penagarikano, O., Mulle, J. G., & Warren, S. T. (2007). The Pathophysiology of Fragile X Syndrome. *Annual Review of Genomics and Human Genetics*, *8*(1), 109–129. https://doi.org/10.1146/annurev.genom.8.080706.092249
Peng, S., Wuu, J., Mufson, E. J., & Fahnestock, M. (2005). Precursor form of brain-derived neurotrophic factor and mature brain-derived neurotrophic factor are decreased in the pre-clinical stages of Alzheimer's disease. *Journal of Neurochemistry*, *93*(6), 1412–1421. https://doi.org/10.1111/j.1471-4159.2005.03135.x
Persico, A. M., D'Agruma, L., Zelante, L., Militerni, R., Bravaccio, C., Schneider, C., … Keller, F. (2004). Enhanced APOE2 transmission rates in families with autistic probands. *Psychiatric Genetics*, *14*(2), 73–82. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15167692
Peters, R. (2006). Ageing and the brain. *Postgraduate Medical Journal*, *82*(964), 84–88. https://doi.org/10.1136/pgmj.2005.036665
Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology*. https://doi.org/10.1001/archneur.56.3.303
Peterson, C. L., & Laniel, M.-A. (2004). Histones and histone modifications. *Current Biology : CB*, *14*(14), R546-51. https://doi.org/10.1016/j.cub.2004.07.007
Petzold, A., Psotta, L., Brigadski, T., Endres, T., & Lessmann, V. (2015). Chronic BDNF deficiency leads to an age-dependent impairment in spatial learning. *Neurobiology of Learning and Memory*, *120*, 52–60. https://doi.org/10.1016/j.nlm.2015.02.009

Phipps, A. J., Vickers, J. C., Taberlay, P. C., & Woodhouse, A. (2016). Neurofilament-labeled pyramidal neurons and astrocytes are deficient in DNA methylation marks in Alzheimer's disease. *Neurobiology of Aging*, *45*, 30–42. https://doi.org/10.1016/j.neurobiolaging.2016.05.003

Pidsley, R., Zotenko, E., Peters, T. J., Lawrence, M. G., Risbridger, G. P., Molloy, P., … Clark, S. J. (2016). Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biology*, *17*(1), 208. https://doi.org/10.1186/s13059-016-1066-1

Piven, J., Bailey, J., Ranson, B. J., & Arndt, S. (1998). No difference in hippocampus volume detected on magnetic resonance imaging in autistic individuals. *Journal of Autism and Developmental Disorders*, *28*(2), 105–110. https://doi.org/10.1023/a:1026084430649

Pizzo, L., Jensen, M., Polyak, A., Rosenfeld, J. A., Mannik, K., Krishnan, A., … Girirajan, S. (2019). Rare variants in the genetic background modulate cognitive and developmental phenotypes in individuals carrying disease-associated variants. *Genetics in Medicine*, *21*(4), 816–825. https://doi.org/10.1038/s41436-018-0266-3

Polymeropoulos, M. H., Higgins, J. J., Golbe, L. I., Johnson, W. G., Ide, S. E., Di Iorio, G., … Duvoisin, R. C. (1996). Mapping of a gene for Parkinson's disease to chromosome 4q21-q23. *Science (New York, N.Y.)*, *274*(5290), 1197–1199. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8895469

Potabattula, R., Dittrich, M., Böck, J., Haertle, L., Müller, T., Hahn, T., … Haaf, T. (2018). Allele-specific methylation of imprinted genes in fetal cord blood is influenced by cis-acting genetic variants and parental factors. *Epigenomics*, *10*(10), 1315–1326. https://doi.org/10.2217/epi-2018-0059

Pruunsild, P., Kazantseva, A., Aid, T., Palm, K., & Timmusk, T. (2007). Dissecting the human BDNF locus: Bidirectional transcription, complex splicing, and multiple promoters. *Genomics*, *90*(3), 397–406. https://doi.org/10.1016/j.ygeno.2007.05.004

Pujadas, E., & Feinberg, A. P. (2012). Regulated Noise in the Epigenetic Landscape of Development and Disease. *Cell*, *148*(6), 1123–1131. https://doi.org/10.1016/J.CELL.2012.02.045

Rage, F., Silhol, M., Binamé, F., Arancibia, S., & Tapia-Arancibia, L. (2007). Effect of aging on the expression of BDNF and TrkB isoforms in rat pituitary. *Neurobiology of Aging*, *28*(7), 1088–1098. https://doi.org/10.1016/j.neurobiolaging.2006.05.013

Raiber, E.-A., Murat, P., Chirgadze, D. Y., Beraldi, D., Luisi, B. F., & Balasubramanian, S. (2014). 5-Formylcytosine alters the structure of the DNA double helix. *Nature Structural & Molecular Biology*, *22*(1), 44–49. https://doi.org/10.1038/nsmb.2936

Ramsahoye, B. H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A. P., & Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proceedings of the National Academy of Sciences*, *97*(10), 5237–5242. https://doi.org/10.1073/PNAS.97.10.5237

Rao, J. S., Keleshian, V. L., Klein, S., & Rapoport, S. I. (2012). Epigenetic modifications in frontal cortex from Alzheimer's disease and bipolar disorder patients. *Translational Psychiatry*, *2*(7), e132–e132. https://doi.org/10.1038/tp.2012.55

Rapicavoli, N. A., Poth, E. M., Zhu, H., & Blackshaw, S. (2011). The long noncoding RNA Six3OS acts in trans to regulate retinal development by modulating Six3 activity. *Neural Development*, *6*(1), 32. https://doi.org/10.1186/1749-8104-6-32

Rauch, T. A., Zhong, X., Wu, X., Wang, M., Kernstine, K. H., Wang, Z., … Pfeifer, G. P. (2008). High-resolution mapping of DNA hypermethylation and hypomethylation in lung cancer. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(1), 252–257. https://doi.org/10.1073/pnas.0710735105

Razin, A., & Shemer, R. (1995). DNA methylation in early development. *Human Molecular Genetics*, *4*(suppl_1), 1751–1755. https://doi.org/10.1093/hmg/4.suppl_1.1751

Renbaum, P., Abrahamove, D., Fainsod, A., Wilson, G. G., Rottem, S., & Razin, A. (1990). Cloning, characterization, and expression in Escherichia coli of the gene coding for the CpG DNA methylase from Spiroplasma sp. strain MQ1(M.SssI). *Nucleic Acids Research*, *18*(5), 1145. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC330428/

Riggs, A. D. (1975). X inactivation, differentiation, and DNA methylation. *Cytogenetic and Genome Research*, *14*(1), 9–25. https://doi.org/10.1159/000130315

Roberta, P., Maria, G. P., Elisabetta, T., Ben, O., Pietro, C., & Giovanni, N. (2002). Quantitative analysis of DNA demethylation and transcriptional reactivation of the FMR1 gene in fragile X cells treated with 5-azadeoxycytidine. *Nucleic Acids Research*, *30*(14), 3278–3285. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC135754/

Robertson, K. D. (2005). DNA methylation and human disease. *Nature Reviews Genetics*, *6*(8), 597–610. https://doi.org/10.1038/nrg1655

Robertson, K. D., & Wolffe, A. P. (2000). DNA methylation in health and disease. *Nature Reviews Genetics*, *1*(1), 11–19. https://doi.org/10.1038/35049533

Rousseaud, A., Delépine, C., Nectoux, J., Billuart, P., & Bienvenu, T. (2015). Differential Expression and Regulation of Brain-Derived Neurotrophic Factor (BDNF) mRNA Isoforms in Brain Cells from Mecp2308/y Mouse Model. *Journal of Molecular Neuroscience*, *56*(4), 758–767. https://doi.org/10.1007/s12031-014-0487-0

Ruga, S., & Jia, L. (2019). Bisulfite Conversion: Key Considerations from the Experts. Retrieved June 12, 2019, from https://epigenie.com/bulletproof-your-bisulfite-conversion-key-considerations-from-the-experts/

Russo, V. E. A. (Vincenzo E. A. ., Martienssen, R. A., & Riggs, A. D. (1996). *Epigenetic mechanisms of gene regulation*. Cold Spring Harbor Laboratory Press. Retrieved from http://agris.fao.org/agris-search/search.do?recordID=US201300304567

Saitoh, S., & Wada, T. (2000). Parent-of-origin specific histone acetylation and reactivation of a key imprinted gene locus in Prader-Willi syndrome. *American Journal of Human Genetics*, *66*(6), 1958–1962. https://doi.org/10.1086/302917

Sanchez-Mut, J. V, & Gräff, J. (2015). Epigenetic Alterations in Alzheimer's Disease. *Frontiers in Behavioral Neuroscience*, *9*, 347. https://doi.org/10.3389/fnbeh.2015.00347

Sánchez-Romero, M. A., Dorado, P., Guarino, E., & Llerena, A. (2009). Development of a new genotyping assay for detection of the *BDNF* Val66Met polymorphism using melting-curve analysis. *Pharmacogenomics*, *10*(6), 989–995.

https://doi.org/10.2217/pgs.09.44
Santangelo, S. L., & Tsatsanis, K. (2005). What is known about autism: Genes, brain, and behavior. *American Journal of PharmacoGenomics*. https://doi.org/10.2165/00129785-200505020-00001
Sasai, N., Nakao, M., & Defossez, P.-A. (2010). Sequence-specific recognition of methylated DNA by human zinc-finger proteins. *Nucleic Acids Research*, *38*(15), 5015–5022. https://doi.org/10.1093/nar/gkq280
Schalkwyk, L. C., Meaburn, E. L., Smith, R., Dempster, E. L., Jeffries, A. R., Davies, M. N., … Mill, J. (2010). Allelic Skewing of DNA Methylation Is Widespread across the Genome. *American Journal of Human Genetics*, *86*(2), 196–212. https://doi.org/10.1016/j.ajhg.2010.01.014
Scheperjans, F., Aho, V., Pereira, P. A. B., Koskinen, K., Paulin, L., Pekkonen, E., … Auvinen, P. (2015). Gut microbiota are related to Parkinson's disease and clinical phenotype. *Movement Disorders*, *30*(3), 350–358. https://doi.org/10.1002/mds.26069
Scheperjans, F., Derkinderen, P., & Borghammer, P. (2018). The Gut and Parkinson's Disease: Hype or Hope? *Journal of Parkinson's Disease*, *8*(s1), S31–S39. https://doi.org/10.3233/JPD-181477
Scherman, D., Desnos, C., Darchen, F., Pollak, P., Javoy-Agid, F., & Agid, Y. (1989). Striatal dopamine deficiency in parkinson's disease: Role of aging. *Annals of Neurology*, *26*(4), 551–557. https://doi.org/10.1002/ana.410260409
Schinder, A. F., & Poo, M. (2000). The neurotrophin hypothesis for synaptic plasticity. *Trends in Neurosciences*, *23*(12), 639–645. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11137155
Schindowski, K., Belarbi, K., & Buée, L. (2008). Neurotrophic factors in Alzheimer's disease: Role of axonal transport. In *Genes, Brain and Behavior*. https://doi.org/10.1111/j.1601-183X.2007.00378.x
Schneider, L., & Zhang, J. (2010). Lysosomal function in macromolecular homeostasis and bioenergetics in Parkinson's disease. *Molecular Neurodegeneration*, *5*(1), 14. https://doi.org/10.1186/1750-1326-5-14
Schöndorf, D. C., Aureli, M., McAllister, F. E., Hindley, C. J., Mayer, F., Schmid, B., … Deleidi, M. (2014). iPSC-derived neurons from GBA1-associated Parkinson's disease patients show autophagic defects and impaired calcium homeostasis. *Nature Communications*, *5*(1), 4028. https://doi.org/10.1038/ncomms5028
Schrempf, W., Brandt, M. D., Storch, A., & Reichmann, H. (2014). Sleep disorders in Parkinson's disease. *Journal of Parkinson's Disease*, *4*(2), 211–221. https://doi.org/10.3233/JPD-130301
Schulte-Herbrüggen, O., Nassenstein, C., Lommatzsch, M., Quarcoo, D., Renz, H., & Braun, A. (2005). Tumor necrosis factor-α and interleukin-6 regulate secretion of brain-derived neurotrophic factor in human monocytes. *Journal of Neuroimmunology*, *160*(1–2), 204–209. https://doi.org/10.1016/J.JNEUROIM.2004.10.026
Schultz, M. D., He, Y., Whitaker, J. W., Hariharan, M., Mukamel, E. A., Leung, D., … Ecker, J. R. (2015). Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature*, *523*(7559), 212–216. https://doi.org/10.1038/nature14465
Schultz, S. H., North, S. W., & Shields, C. G. (2007, June 15). Schizophrenia: A review.

American Family Physician.

Seitz, A., Gourevitch, D., Zhang, X.-M., Clark, L., Chen, P., Kragol, M., … Heber-Katz, E. (2005). Sense and antisense transcripts of the apolipoprotein E gene in normal and ApoE knockout mice, their expression after spinal cord injury and corresponding human transcripts. *Human Molecular Genetics*, *14*(18), 2661–2670. https://doi.org/10.1093/hmg/ddi296

Selkoe, D. J. (2012). Preventing Alzheimer's Disease. *Science*, *337*(6101), 1488–1492. https://doi.org/10.1126/SCIENCE.1228541

Semick, S. A., Bharadwaj, R. A., Collado-Torres, L., Tao, R., Shin, J. H., Deep-Soboslay, A., … Mattay, V. S. (2019). Integrated DNA methylation and gene expression profiling across multiple brain regions implicate novel genes in Alzheimer's disease. *Acta Neuropathologica*, *137*(4), 557–569. https://doi.org/10.1007/s00401-019-01966-5

Sen, A., Nelson, T. J., & Alkon, D. L. (2015). ApoE4 and Aβ Oligomers Reduce BDNF Expression via HDAC Nuclear Translocation. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *35*(19), 7538–7551. https://doi.org/10.1523/JNEUROSCI.0260-15.2015

Sen, A., Nelson, T. J., & Alkon, D. L. (2017). ApoE isoforms differentially regulates cleavage and secretion of BDNF. *Molecular Brain*. https://doi.org/10.1186/s13041-017-0301-3

Seong, K.-H., Li, D., Shimizu, H., Nakamura, R., & Ishii, S. (2011). Inheritance of Stress-Induced, ATF-2-Dependent Epigenetic Change. *Cell*, *145*(7), 1049–1061. https://doi.org/10.1016/j.cell.2011.05.029

Seripa, D., Matera, M. G., Daniele, A., Bizzarro, A., Rinaldi, M., Gravina, C., … Pilotto, A. (2007). The Missing ApoE Allele. *Annals of Human Genetics*, *71*(4), 496–500. https://doi.org/10.1111/j.1469-1809.2006.00344.x

Shacka, J. J. (2012). Mouse models of neuronal ceroid lipofuscinoses: Useful pre-clinical tools to delineate disease pathophysiology and validate therapeutics. *Brain Research Bulletin*, *88*(1), 43–57. https://doi.org/10.1016/j.brainresbull.2012.03.003

Shen, L., Wu, H., Diep, D., Yamaguchi, S., D'Alessio, A. C., Fung, H.-L., … Zhang, Y. (2013). Genome-wide Analysis Reveals TET- and TDG-Dependent 5-Methylcytosine Oxidation Dynamics. *Cell*, *153*(3), 692–706. https://doi.org/10.1016/j.cell.2013.04.002

Shi, D. Q., Ali, I., Tang, J., & Yang, W. C. (2017). New Insights into 5hmC DNA Modification: Generation, Distribution and Function. *Frontiers in Genetics*, *8*, 100. https://doi.org/10.3389/fgene.2017.00100

Shi, J. (2015). Regulatory networks between neurotrophins and miRNAs in brain diseases and cancers. *Acta Pharmacologica Sinica*, *36*(2), 149. https://doi.org/10.1038/APS.2014.135

Shibata, M., Yamada, S., Kumar, S. R., Calero, M., Bading, J., Frangione, B., … Zlokovic, B. V. (2000). Clearance of Alzheimer's amyloid-ss(1-40) peptide from brain by LDL receptor-related protein-1 at the blood-brain barrier. *The Journal of Clinical Investigation*, *106*(12), 1489–1499. https://doi.org/10.1172/JCI10498

Shirane, K., Toh, H., Kobayashi, H., Miura, F., Chiba, H., Ito, T., … Sasaki, H. (2013). Mouse oocyte methylomes at base resolution reveal genome-wide accumulation of non-CpG methylation and role of DNA methyltransferases. *PLoS*

Genetics, *9*(4), e1003439. https://doi.org/10.1371/journal.pgen.1003439

Shoemaker, R., Deng, J., Wang, W., & Zhang, K. (2010). Allele-specific methylation is prevalent and is contributed by CpG-SNPs in the human genome. *Genome Research*, *20*(7), 883–889. https://doi.org/10.1101/gr.104695.109

Shoval, G., & Weizman, A. (2005). The possible role of neurotrophins in the pathogenesis and therapy of schizophrenia. *European Neuropsychopharmacology : The Journal of the European College of Neuropsychopharmacology*, *15*(3), 319–329. https://doi.org/10.1016/j.euroneuro.2004.12.005

Sidransky, E., Nalls, M. A., Aasly, J. O., Aharon-Peretz, J., Annesi, G., Barbosa, E. R., … Ziegler, S. G. (2009). Multicenter Analysis of Glucocerebrosidase Mutations in Parkinson's Disease. *New England Journal of Medicine*, *361*(17), 1651–1661. https://doi.org/10.1056/NEJMoa0901281

Siegel, G. J., & Chauhan, N. B. (2000). Neurotrophic factors in Alzheimer's and Parkinson's disease brain. *Brain Research. Brain Research Reviews*, *33*(2–3), 199–227. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11011066

Smallwood, S. A., Tomizawa, S., Krueger, F., Ruf, N., Carli, N., Segonds-Pichon, A., … Kelsey, G. (2011). Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nature Genetics*, *43*(8), 811–814. https://doi.org/10.1038/ng.864

Smith, A. K., Kilaru, V., Kocak, M., Almli, L. M., Mercer, K. B., Ressler, K. J., … Conneely, K. N. (2014). Methylation quantitative trait loci (meQTLs) are consistently detected across ancestry, developmental stage, and tissue type. *BMC Genomics*, *15*(1), 145. https://doi.org/10.1186/1471-2164-15-145

Smith, D. (2009). Dispelling myths about Alzheimer's disease. *Oxford Medicine; Newsletter of the Oxford Medical Alumni*. Retrieved from https://www.ndcn.ox.ac.uk/research/centre-prevention-stroke-dementia/resources/optima-oxford-project-to-investigate-memory-and-ageing

Smith, Z. D., Chan, M. M., Humm, K. C., Karnik, R., Mekhoubad, S., Regev, A., … Meissner, A. (2014). DNA methylation dynamics of the human preimplantation embryo. *Nature*, *511*(7511), 611–615. https://doi.org/10.1038/nature13581

Smith, Z. D., & Meissner, A. (2013). DNA methylation: roles in mammalian development. *Nature Reviews. Genetics*, *14*(3), 204–220. https://doi.org/10.1038/nrg3354

Son, J. H., Shim, J. H., Kim, K.-H., Ha, J.-Y., & Han, J. Y. (2012). Neuronal autophagy and neurodegenerative diseases. *Experimental & Molecular Medicine*, *44*(2), 89. https://doi.org/10.3858/EMM.2012.44.2.031

Song, C. X., Diao, J., Brunger, A. T., & Quake, S. R. (2016). Simultaneous single-molecule epigenetic imaging of DNA methylation and hydroxymethylation. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(16), 4338–4343. https://doi.org/10.1073/pnas.1600223113

Song, C. X., & He, C. (2013). Potential functional roles of DNA demethylation intermediates. *Trends in Biochemical Sciences*, *38*(10), 480–484. https://doi.org/10.1016/j.tibs.2013.07.003

Song, C. X., Szulwach, K. E., Dai, Q., Fu, Y., Mao, S. Q., Lin, L., … He, C. (2013). Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell*, *153*(3), 678–691. https://doi.org/10.1016/j.cell.2013.04.001

Song, C. X., Szulwach, K. E., Fu, Y., Dai, Q., Yi, C., Li, X., … He, C. (2011). Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nature Biotechnology*, *29*(1), 68–72. https://doi.org/10.1038/nbt.1732

Song, J.-H., Yu, J.-T., & Tan, L. (2015). Brain-Derived Neurotrophic Factor in Alzheimer's Disease: Risk, Mechanisms, and Therapy. *Molecular Neurobiology*, *52*(3), 1477–1493. https://doi.org/10.1007/s12035-014-8958-4

Spillantini, M. G., Schmidt, M. L., Lee, V. M.-Y., Trojanowski, J. Q., Jakes, R., & Goedert, M. (1997). α-Synuclein in Lewy bodies. *Nature*, *388*(6645), 839–840. https://doi.org/10.1038/42166

Spruijt, C. G., Gnerlich, F., Smits, A. H., Pfaffeneder, T., Jansen, P. W. T. C., Bauer, C., … Vermeulen, M. (2013). Dynamic Readers for 5-(Hydroxy)Methylcytosine and Its Oxidized Derivatives. *Cell*, *152*(5), 1146–1159. https://doi.org/10.1016/J.CELL.2013.02.004

Stephan, B. C. M., Brayne, C., Savva, G. M., Matthews, F. E., & Medical Research Council Cognitive Function and Ageing Study. (2011). Occurrence of medical co-morbidity in mild cognitive impairment: implications for generalisation of MCI research. *Age and Ageing*, *40*(4), 501–507. https://doi.org/10.1093/ageing/afr057

Stöger, R., Scaife, P. J., Shephard, F., & Chakrabarti, L. (2017). Elevated 5hmC levels characterize DNA of the cerebellum in Parkinson's disease. *Npj Parkinson's Disease*, *3*(1), 6. https://doi.org/10.1038/s41531-017-0007-3

Strathearn, L. (2009). The elusive etiology of autism: nature and nurture? *Frontiers in Behavioral Neuroscience*, *3*, 11. https://doi.org/10.3389/neuro.08.011.2009

Strittmatter, W. J., Saunders, A. M., Schmechel, D., Pericak-Vance, M., Enghild, J., Salvesen, G. S., & Roses, A. D. (1993). Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(5), 1977–1981. https://doi.org/10.1073/pnas.90.5.1977

Sung, H. Y., Choi, E. N., Ahn Jo, S., Oh, S., & Ahn, J.-H. (2011). Amyloid protein-mediated differential DNA methylation status regulates gene expression in Alzheimer's disease model cell line. *Biochemical and Biophysical Research Communications*, *414*(4), 700–705. https://doi.org/10.1016/j.bbrc.2011.09.136

Sutcliffe, J. S., Nakao, M., Christian, S., Örstavik, K. H., Tommerup, N., Ledbetter, D. H., & Beaudet, A. L. (1994). Deletions of a differentially methylated CpG island at the SNRPN gene define a putative imprinting control region. *Nature Genetics*, *8*(1), 52–58. https://doi.org/10.1038/ng0994-52

Szulwach, K. E., Li, X., Li, Y., Song, C.-X., Wu, H., Dai, Q., … Jin, P. (2011). 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nature Neuroscience*, *14*(12), 1607–1616. https://doi.org/10.1038/nn.2959

Tai, L. M., Bilousova, T., Jungbauer, L., Roeske, S. K., Youmans, K. L., Yu, C., … Ladu, M. J. (2013). Levels of soluble apolipoprotein E/amyloid-β (Aβ) complex are reduced and oligomeric Aβ increased with APOE4 and Alzheimer disease in a transgenic mouse model and human samples. *The Journal of Biological Chemistry*, *288*(8), 5914–5926. https://doi.org/10.1074/jbc.M112.442103

Talwar, P., Sinha, J., Grover, S., Agarwal, R., Kushwaha, S., Srivastava, M. V. P., &

Kukreti, R. (2016). Meta-analysis of apolipoprotein E levels in the cerebrospinal fluid of patients with Alzheimer's disease. *Journal of the Neurological Sciences*, *360*, 179–187. https://doi.org/10.1016/J.JNS.2015.12.004

Tanzi, R. E., & Bertram, L. (2001). New frontiers in Alzheimer's disease genetics. *Neuron*, *32*(2), 181–184. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11683989

Tao, H., Xie, P., Cao, Y., Shu, L., Li, L., Chen, J., … Li, X. (2018). The Dynamic DNA Demethylation during Postnatal Neuronal Development and Neural Stem Cell Differentiation. *Stem Cells International*, *2018*, 1–10. https://doi.org/10.1155/2018/2186301

Tapiola, T., Pirttilä, T., Mehta, P. D., Alafuzofff, I., Lehtovirta, M., & Soininen, H. (2000). Relationship between apoE genotype and CSF beta-amyloid (1-42) and tau in patients with probable and definite Alzheimer's disease. *Neurobiology of Aging*, *21*(5), 735–740. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11016543

Teng, H. K., Teng, K. K., Lee, R., Wright, S., Tevar, S., Almeida, R. D., … Hempstead, B. L. (2005). ProBDNF Induces Neuronal Apoptosis via Activation of a Receptor Complex of p75NTR and Sortilin. *Journal of Neuroscience*, *25*(22), 5455–5463. https://doi.org/10.1523/JNEUROSCI.5123-04.2005

The Lancet Neurology. (2017). Solanezumab: too late in mild Alzheimer's disease? *The Lancet Neurology*, *16*(2), 97. https://doi.org/10.1016/S1474-4422(16)30395-7

Theuns, J. (2000). Transcriptional regulation of Alzheimer's disease genes: implications for susceptibility. *Human Molecular Genetics*. https://doi.org/10.1093/hmg/9.16.2383

Tolar, M., Keller, J. N., Chan, S., Mattson, M. P., Marques, M. A., Crutcher, K. A., … Huang, Y. (1999). Truncated apolipoprotein E (ApoE) causes increased intracellular calcium and may mediate ApoE neurotoxicity. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *19*(16), 7100–7110. https://doi.org/20026488

Tonegawa, S., Liu, X., Ramirez, S., & Redondo, R. (2015). Memory Engram Cells Have Come of Age. *Neuron*, *87*(5), 918–931. https://doi.org/10.1016/j.neuron.2015.08.002

Tsai, S.-J. (2005). Is autism caused by early hyperactivity of brain-derived neurotrophic factor? *Medical Hypotheses*, *65*(1), 79–82. https://doi.org/10.1016/J.MEHY.2005.01.034

Tsaprouni, L. G., Yang, T. P., Bell, J., Dick, K. J., Kanoni, S., Nisbet, J., … Deloukas, P. (2014). Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. *Epigenetics*, *9*(10), 1382–1396. https://doi.org/10.4161/15592294.2014.969637

Tsuang, D., Leverenz, J. B., Lopez, O. L., Hamilton, R. L., Bennett, D. A., Schneider, J. A., … Zabetian, C. P. (2013). APOE ϵ4 Increases Risk for Dementia in Pure Synucleinopathies. *JAMA Neurology*, *70*(2), 223. https://doi.org/10.1001/jamaneurol.2013.600

Tsuchida, T., Mano, T., Koshi-Mano, K., Bannai, T., Matsubara, T., Yamashita, S., … Iwata, A. (2018). Methylation changes and aberrant expression of FGFR3 in Lewy body disease neurons. *Brain Research*, *1697*, 59–66.

https://doi.org/10.1016/J.BRAINRES.2018.06.017
Tufarelli, C., Stanley, J. A. S., Garrick, D., Sharpe, J. A., Ayyub, H., Wood, W. G., & Higgs, D. R. (2003). Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nature Genetics*, *34*(2), 157–165. https://doi.org/10.1038/ng1157
Tulloch, J., Leong, L., Chen, S., Keene, C. D., Millard, S. P., Shutes-David, A., … Tsuang, D. (2018). APOE DNA methylation is altered in Lewy body dementia. *Alzheimer's & Dementia*, *14*(7), 889–894. https://doi.org/10.1016/J.JALZ.2018.02.005
Twine, N. A., Janitz, K., Wilkins, M. R., & Janitz, M. (2011). Whole transcriptome sequencing reveals gene expression and splicing differences in brain regions affected by Alzheimer's disease. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0016266
Tycko, B. (2010). Allele-specific DNA methylation: beyond imprinting. *Human Molecular Genetics*, *19*(R2), R210-20. https://doi.org/10.1093/hmg/ddq376
Tyler, W. J., Alonso, M., Bramham, C. R., & Pozzo-Miller, L. D. (2002). From Acquisition to Consolidation: On the Role of Brain-Derived Neurotrophic Factor Signaling in Hippocampal-Dependent Learning. *Learning & Memory*, *9*(5), 224–237. https://doi.org/10.1101/lm.51202
Unger, M. M., Spiegel, J., Dillmann, K.-U., Grundmann, D., Philippeit, H., Bürmann, J., … Schäfer, K.-H. (2016). Short chain fatty acids and gut microbiota differ between patients with Parkinson's disease and age-matched controls. *Parkinsonism & Related Disorders*, *32*, 66–72. https://doi.org/10.1016/j.parkreldis.2016.08.019
UniProt: a worldwide hub of protein knowledge. (2019). *Nucleic Acids Research*, *47*(D1), D506–D515. https://doi.org/10.1093/nar/gky1049
Ursini, G., Cavalleri, T., Fazio, L., Angrisano, T., Iacovelli, L., Porcelli, A., … Bertolino, A. (2016). BDNF rs6265 methylation and genotype interact on risk for schizophrenia. *Epigenetics*, *11*(1), 11–23. https://doi.org/10.1080/15592294.2015.1117736
Valinluck, V., & Sowers, L. C. (2007). Endogenous Cytosine Damage Products Alter the Site Selectivity of Human DNA Maintenance Methyltransferase DNMT1. *Cancer Research*, *67*(3), 946–950. https://doi.org/10.1158/0008-5472.CAN-06-3123
Van Den Eeden, S. K., Tanner, C. M., Bernstein, A. L., Fross, R. D., Leimpeter, A., Bloch, D. A., & Nelson, L. M. (2003). Incidence of Parkinson's disease: variation by age, gender, and race/ethnicity. *American Journal of Epidemiology*, *157*(11), 1015–1022. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12777365
Van Der Kolk, N. M., Speelman, A. D., Van Nimwegen, M., Kessels, R. P. C., Inthout, J., Hakobjan, M., … Van De Warrenburg, B. P. (2015). BDNF polymorphism associates with decline in set shifting in Parkinson's disease. https://doi.org/10.1016/j.neurobiolaging.2014.08.023
Varela, M. A., Roberts, T. C., Andaloussi, S. El, & Wood, M. J. (2012). Natural antisense makes sense for gene-specific a ctivation in brain. *Molecular Therapy - Nucleic Acids*, *1*(5), e24. https://doi.org/10.1038/mtna.2012.17
Varley, K. E., Gertz, J., Bowling, K. M., Parker, S. L., Reddy, T. E., Pauli-Behn, F., … Myers, R. M. (2013). Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Research*, *23*(3), 555–567. https://doi.org/10.1101/gr.147942.112

Vaughan, R. M., Dickson, B. M., Cornett, E. M., Harrison, J. S., Kuhlman, B., & Rothbart, S. B. (2018). Comparative biochemical analysis of UHRF proteins reveals molecular mechanisms that uncouple UHRF2 from DNA methylation maintenance. *Nucleic Acids Research*, *46*(9), 4405–4416. https://doi.org/10.1093/nar/gky151

Veldhuis, J. D. (2013). Changes in pituitary function with ageing and implications for patient care. *Nature Reviews. Endocrinology*, *9*(4), 205–215. https://doi.org/10.1038/nrendo.2013.38

Ventriglia, M., Zanardini, R., Bonomini, C., Zanetti, O., Volpe, D., Pasqualetti, P., … Bocchio-Chiavetto, L. (2013). Serum Brain-Derived Neurotrophic Factor Levels in Different Neurological Diseases. *BioMed Research International*, *2013*, 1–7. https://doi.org/10.1155/2013/901082

Verkerk, A. J. M. H., Pieretti, M., Sutcliffe, J. S., Fu, Y. H., Kuhl, D. P. A., Pizzuti, A., … Warran, S. T. (1991). Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell*, *65*(5), 905–914. https://doi.org/10.1016/0092-8674(91)90397-H

Von Arnim, C. A. F., Herbolsheimer, F., Nikolaus, T., Peter, R., Biesalski, H. K., Ludolph, A. C., … Nagel, G. (2012). Dietary antioxidants and dementia in a population-based case-control study among older people in South Germany. *Journal of Alzheimer's Disease*, *31*(4), 717–724. https://doi.org/10.3233/JAD-2012-120634

von Bohlen und Halbach, O., Minichiello, L., & Unsicker, K. (2005). Haploin-sufficiency for trkB and trkC receptors induces cell loss and accumulation of α-synuclein in the substantia nigra. *The FASEB Journal*, *19*(12), 1740–1742. https://doi.org/10.1096/fj.05-3845fje

Wahl, S., Drong, A., Lehne, B., Loh, M., Scott, W. R., Kunze, S., … Chambers, J. C. (2017). Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature*, *541*(7635), 81–86. https://doi.org/10.1038/nature20784

Wallin, C., Sholts, S. B., Österlund, N., Luo, J., Jarvet, J., Roos, P. M., … Wärmländer, S. K. T. S. (2017). Alzheimer's disease and cigarette smoke components: Effects of nicotine, PAHs, and Cd(II), Cr(III), Pb(II), Pb(IV) ions on amyloid-β peptide aggregation. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-13759-5

Wang, C., Chen, L., Yang, Y., Zhang, M., & Wong, G. (2019). Identification of potential blood biomarkers for Parkinson's disease by gene expression and DNA methylation data integration analysis. *Clinical Epigenetics*, *11*(1), 24. https://doi.org/10.1186/s13148-019-0621-5

Wang, D. O., Kim, S. M., Zhao, Y., Hwang, H., Miura, S. K., Sossin, W. S., & Martin, K. C. (2009). Synapse- and Stimulus-Specific Local Translation During Long-Term Neuronal Plasticity. *Science*, *324*(5934), 1536–1540. https://doi.org/10.1126/science.1173205

Wang, H., Maurano, M. T., Qu, H., Varley, K. E., Gertz, J., Pauli, F., … Stamatoyannopoulos, J. A. (2012). Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Research*, *22*(9), 1680–1688. https://doi.org/10.1101/gr.136101.111

Wang, L., Benzinger, T. L., Su, Y., Christensen, J., Friedrichsen, K., Aldea, P., … Ances,

B. M. (2016). Evaluation of Tau Imaging in Staging Alzheimer Disease and Revealing Interactions Between β-Amyloid and Tauopathy. *JAMA Neurology*, *73*(9), 1070. https://doi.org/10.1001/jamaneurol.2016.2078

Wang, R., Li, Y., Du, P., Zhang, X., Li, X., & Cheng, G. (2019). Hypomethylation of the lncRNA SOX21-AS1 has clinical prognostic value in cervical cancer. *Life Sciences*, 116708. https://doi.org/10.1016/j.lfs.2019.116708

Wang, X., Bauer, J. H., Li, Y., Shao, Z., Zetoune, F. S., Cattaneo, E., & Vincenz, C. (2001). Characterization of a p75NTR Apoptotic Signaling Pathway Using a Novel Cellular Model. *Journal of Biological Chemistry*, *276*(36), 33812–33820. https://doi.org/10.1074/jbc.M010548200

Wang, Yanli, Song, F., Zhang, B., Zhang, L., Xu, J., Kuang, D., … Yue, F. (2018). The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. *Genome Biology*, *19*(1), 151. https://doi.org/10.1186/s13059-018-1519-9

Wang, Yueling, Jia, A., & Ma, W. (2018). Dexmedetomidine attenuates the toxicity of β-amyloid on neurons and astrocytes by increasing BDNF production under the regulation of HDAC2 and HDAC5. *Molecular Medicine Reports*, *19*(1), 533–540. https://doi.org/10.3892/mmr.2018.9694

Washington D.C.: American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th editio).

Watson, J. D., & Crick, F. H. C. (1953). THE STRUCTURE OF DNA. *Cold Spring Harb Symp Quant Biol*, *18*, 123–131. https://doi.org/10.1101/SQB.1953.018.01.020

Weaving, L. S., Williamson, S. L., Bennetts, B., Davis, M., Ellaway, C. J., Leonard, H., … Christodoulou, J. (2003). Effects ofMECP2 mutation type, location and X-inactivation in modulating Rett syndrome phenotype. *American Journal of Medical Genetics*, *118A*(2), 103–114. https://doi.org/10.1002/ajmg.a.10053

Wei, C. W., Luo, T., Zou, S.-S., & Wu, A.-S. (2018). The Role of Long Noncoding RNAs in Central Nervous System and Neurodegenerative Diseases. *Frontiers in Behavioral Neuroscience*. https://doi.org/10.3389/fnbeh.2018.00175

Weisgraber, K. H., Rall, S. C. J., & Mahley, R. W. (1981). Human E apoprotein heterogeneity. Cysteine-arginine interchanges in the amino acid sequence of the apo-E isoforms. - PubMed - NCBI. *The Journal of Biologiccal Chemistry*, *256*(17), 9077–9083. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7263700?dopt=Abstract

Weisman, D., Cho, M., Taylor, C., Adame, A., Thal, L. J., & Hansen, L. A. (2007). In dementia with Lewy bodies, Braak stage determines phenotype, not Lewy body distribution. *Neurology*, *69*(4), 356–359. https://doi.org/10.1212/01.wnl.0000266626.64913.0f

Wen, L., Li, X. X., Yan, L., Tan, Y., Li, R. R., Zhao, Y., … Lu Wen, X. L. L. Y. Y. T. R. L. Y. Z. Y. W. J. X. Y. Z. C. S. M. Y. X. L. P. Z. X. L. Y. H. H. G. X. W. C. H. R. L. F. T. J. Q. (2014). Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain. *Genome Biology*, *15*(3), R49. https://doi.org/10.1186/gb-2014-15-3-r49

West, H. L., Rebeck, G. W., & Hyman, B. T. (1994). Frequency of the apolipoprotein E epsilon 2 allele is diminished in sporadic Alzheimer disease. *Neuroscience Letters*, *175*(1–2), 46–48. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7970208

Wheater, E. N. W., Stoye, D. Q., Cox, S. R., Wardlaw, J. M., Drake, A. J., Bastin, M. E., & Boardman, J. P. (2020, June 1). DNA methylation and brain structure and function across the life course: A systematic review. *Neuroscience and Biobehavioral Reviews*. Elsevier Ltd. https://doi.org/10.1016/j.neubiorev.2020.03.007

Wheldon, L. M., Abakir, A., Ferjentsik, Z., Dudnakova, T., Strohbuecker, S., Christie, D., … Ruzov, A. (2014). Transient Accumulation of 5-Carboxylcytosine Indicates Involvement of Active Demethylation in Lineage Specification of Neural Stem Cells. *Cell Reports*, *7*(5), 1353–1361. https://doi.org/10.1016/J.CELREP.2014.05.003

Whyte, J. M., Ellis, J. J., Brown, M. A., & Kenna, T. J. (2019, June 3). Best practices in DNA methylation: Lessons from inflammatory bowel disease, psoriasis and ankylosing spondylitis. *Arthritis Research and Therapy*. BioMed Central Ltd. https://doi.org/10.1186/s13075-019-1922-y

Williams, A. (2002). Defining neurodegenerative diseases. *BMJ (Clinical Research Ed.)*, *324*(7352), 1465–1466. https://doi.org/10.1136/bmj.324.7352.1465

Wilson, V. L., & Jones, P. A. (1983). DNA methylation decreases in aging but not in immortal cells. *Science (New York, N.Y.)*, *220*(4601), 1055–1057. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6844925

Winkler, S., Hagenah, J., Lincoln, S., Heckman, M., Haugarvoll, K., Lohmann-Hedrich, K., … Klein, C. (2007). α-synuclein and Parkinson disease susceptibility. *Neurology*, *69*(18), 1745–1750. https://doi.org/10.1212/01.wnl.0000275524.15125.f4

Woodcock, D. M., Crowther, P. J., & Diver, W. P. (1987). The majority of methylated deoxycytidines in human DNA are not in the CpG dinucleotide. *Biochemical and Biophysical Research Communications*, *145*(2), 888–894. https://doi.org/10.1016/0006-291X(87)91048-5

Wu, C., Wong, Y. C., & Elgin, S. C. (1979). The chromatin structure of specific genes: II. Disruption of chromatin structure during gene activity. *Cell*, *16*(4), 807–814. https://doi.org/10.1016/0092-8674(79)90096-5

Wu, H., Wu, X., Shen, L., & Zhang, Y. (2014). Single-base resolution analysis of active DNA demethylation using methylase-assisted bisulfite sequencing. *Nature Biotechnology*, *32*(12), 1231–1240. https://doi.org/10.1038/nbt.3073

Wu, H., Wu, X., & Zhang, Y. (2016). Base-resolution profiling of active DNA demethylation using MAB-seq and caMAB-seq. *Nature Protocols*, *11*(6), 1081–1100. https://doi.org/10.1038/nprot.2016.069

Wu, H., & Zhang, Y. (2015). Charting oxidized methylcytosines at base resolution. *Nature Structural & Molecular Biology*, *22*(9), 656–661. https://doi.org/10.1038/nsmb.3071

Wu, S. C., & Zhang, Y. (2010). Active DNA demethylation: many roads lead to Rome. *Nature Reviews Molecular Cell Biology*, *11*(9), 607–620. https://doi.org/10.1038/nrm2950

Xhemalce, B., Dawson, M. A., & Bannister, A. J. (2011). Histone Modifications. In *Encyclopedia of Molecular Cell Biology and Molecular Medicine*. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA. https://doi.org/10.1002/3527600906.mcb.201100004

Xi, Q., Gao, N., Zhang, X., Zhang, B., Ye, W., Wu, J., & Zhang, X. (2014). A natural

antisense transcript regulates acetylcholinesterase gene expression via epigenetic modification in Hepatocellular Carcinoma. *The International Journal of Biochemistry & Cell Biology*, *55*, 242–251. https://doi.org/10.1016/J.BIOCEL.2014.09.012

Xie, B., Xu, Y., Liu, Z., Liu, W., Jiang, L., Zhang, R., … Xu, S. (2017). Elevation of Peripheral BDNF Promoter Methylation Predicts Conversion from Amnestic Mild Cognitive Impairment to Alzheimer's Disease: A 5-Year Longitudinal Study. *Journal of Alzheimer's Disease*, *56*(1), 391–401. https://doi.org/10.3233/JAD-160954

Xu, T., Yu, X., Perlik, A. J., Tobin, W. F., Zweig, J. A., Tennant, K., … Zuo, Y. (2009). Rapid formation and selective stabilization of synapses for enduring motor memories. *Nature*, *462*(7275), 915–919. https://doi.org/10.1038/nature08389

Xu, Xiangru. (2015). DNA methylation and cognitive aging. *Oncotarget*, *6*(16), 13922–13932. https://doi.org/10.18632/oncotarget.4215

Xu, Xin, Garcia, J., Ewalt, R., Nason, S., & Pozzo-Miller, L. (2017). The BDNF val-66-met Polymorphism Affects Neuronal Morphology and Synaptic Transmission in Cultured Hippocampal Neurons from Rett Syndrome Mice. *Frontiers in Cellular Neuroscience*. https://doi.org/10.3389/fncel.2017.00203

Yan, H., Yuan, W., Velculescu, V. E., Vogelstein, B., & Kinzler, K. W. (2002). Allelic variation in human gene expression. *Science (New York, N.Y.)*, *297*(5584), 1143. https://doi.org/10.1126/science.1072545

Yang, G., Pan, F., & Gan, W.-B. (2009). Stably maintained dendritic spines are associated with lifelong memories. *Nature*, *462*(7275), 920–924. https://doi.org/10.1038/nature08577

Yi, S., Long, F., Cheng, J., & Huang, D. (2017). An optimized rapid bisulfite conversion method with high recovery of cell-free DNA. *BMC Molecular Biology*, *18*(1), 24. https://doi.org/10.1186/s12867-017-0101-4

Yoder, J. A., Walsh, C. P., & Bestor, T. H. (1997). Cytosine methylation and the ecology of intragenomic parasites. *Trends in Genetics*, *13*(8), 335–340. https://doi.org/10.1016/S0168-9525(97)01181-5

Yoo, H. J., Yang, S. Y., Cho, I. H., Park, M., & Kim, S. A. (2014). Polymorphisms of BDNF gene and autism spectrum disorders: family based association study with korean trios. *Psychiatry Investigation*, *11*(3), 319–324. https://doi.org/10.4306/pi.2014.11.3.319

You, C., Ji, D., Dai, X., & Wang, Y. (2014). Effects of Tet-mediated oxidation products of 5-methylcytosine on DNA transcription in vitro and in mammalian cells. *Scientific Reports*, *4*, 7052. https://doi.org/10.1038/srep07052

Yu, C.-E., & Foraker, J. (2015). Epigenetic considerations of the APOE gene. *Biomolecular Concepts*, *6*(1), 77–84. https://doi.org/10.1515/bmc-2014-0039

Yu, C. E., Cudaback, E., Foraker, J., Thomson, Z., Leong, L., Lutz, F., … Bekris, L. M. (2013). Epigenetic signature and enhancer activity of the human APOE gene. *Human Molecular Genetics*, *22*(24), 5036–5047. https://doi.org/10.1093/hmg/ddt354

Yu, D. H., Ware, C., Waterland, R. A., Zhang, J., Chen, M.-H., Gadkari, M., … Shen, L. (2013). Developmentally Programmed 3' CpG Island Methylation Confers Tissue- and Cell-Type-Specific Transcriptional Activation. *Molecular and Cellular Biology*, *33*(9), 1845–1858. https://doi.org/10.1128/MCB.01124-12

Yu, M., Hon, G. C., Szulwach, K. E., Song, C.-X., Zhang, L., Kim, A., … He, C. (2012). Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell*, *149*(6), 1368–1380. https://doi.org/10.1016/j.cell.2012.04.027

Zakharyan, R., & Boyajyan, A. (2014). Brain-derived neurotrophic factor blood levels are decreased in schizophrenia patients and associate with rs6265 genotypes. *Clinical Biochemistry*, *47*(12), 1052–1055. https://doi.org/10.1016/J.CLINBIOCHEM.2014.03.021

Zannis, V. I., Breslow, J. L., Utermann, G., Mahley, R. W., Weisgraber, K. H., Havel, R. J., … Blum, C. (1982). Proposed nomenclature of apoE isoproteins, apoE genotypes, and phenotypes. *Journal of Lipid Research*, *23*(6), 911–914. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7130859

Zhang, C., Hoshida, Y., & Sadler, K. C. (2016). Comparative Epigenomic Profiling of the DNA Methylome in Mouse and Zebrafish Uncovers High Interspecies Divergence. *Frontiers in Genetics*, *7*, 110. https://doi.org/10.3389/fgene.2016.00110

Zhang, D., Cheng, L., Badner, J. A., Chen, C., Chen, Q., Luo, W., … Liu, C. (2010). Genetic Control of Individual Differences in Gene-Specific Methylation in Human Brain. *American Journal of Human Genetics*, *86*(3), 411–419. https://doi.org/10.1016/j.ajhg.2010.02.005

Zhang, L., Gu, C., Yang, L., Tang, F., & Gao, Y. Q. (2017). The sequence preference of DNA methylation variation in mammalians. *PloS One*, *12*(10), e0186559. https://doi.org/10.1371/journal.pone.0186559

Zhang, Y., Hapala, J., Brenner, H., & Wagner, W. (2017, February 2). Individual CpG sites that are associated with age and life expectancy become hypomethylated upon aging. *Clinical Epigenetics*. Springer Verlag. https://doi.org/10.1186/s13148-017-0315-9

Zhao, J., Zhu, Y., Yang, J., Li, L., Wu, H., Jager, P. L. De, … Bennett, D. A. (2017). A genome-wide profiling of brain DNA hydroxymethylation in Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *13*(6), 674–688. https://doi.org/10.1016/J.JALZ.2016.10.004

Zheleznyakova, G. Y., Cao, H., & Schiöth, H. B. (2016). BDNF DNA methylation changes as a biomarker of psychiatric disorders: literature review and open access database analysis. *Behavioral and Brain Functions : BBF*, *12*(1), 17. https://doi.org/10.1186/s12993-016-0101-4

Zheng, K., An, J. J., Yang, F., Xu, W., Xu, Z. Q. D., Wu, J., … Lu, B. (2011). TrkB signaling in parvalbumin-positive interneurons is critical for gamma-band network synchronization in hippocampus. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(41), 17201–17206. https://doi.org/10.1073/pnas.1114241108

Zheng, X., Lin, C., Li, Y., Ye, J., Zhou, J., & Guo, P. (2016). Long noncoding RNA BDNF-AS regulates ketamine-induced neurotoxicity in neural stem cell derived neurons. *Biomedicine & Pharmacotherapy*, *82*, 722–728. https://doi.org/10.1016/j.biopha.2016.05.050

Zhong, J.-B., Li, X., Zhong, S.-M., Liu, J.-D., Chen, C.-B., & Wu, X.-Y. (2017). Knockdown of long noncoding antisense RNA brain-derived neurotrophic factor attenuates hypoxia/reoxygenation-induced nerve cell apoptosis through the BDNF-TrkB-PI3K/Akt signaling pathway. *Neuroreport*, *28*(14), 910–916.

https://doi.org/10.1097/WNR.0000000000000860
Zhou, F. C., Resendiz, M., Lo, C.-L., & Chen, Y. (2016). Cell-Wide DNA De-Methylation and Re-Methylation of Purkinje Neurons in the Developing Cerebellum. *PLOS ONE*, *11*(9), e0162063. https://doi.org/10.1371/journal.pone.0162063
Zhou, T., Xiong, J., Wang, M., Yang, N., Wong, J., Zhu, B., & Xu, R.-M. (2014). Structural Basis for Hydroxymethylcytosine Recognition by the SRA Domain of UHRF2. *Molecular Cell*, *54*(5), 879–886. https://doi.org/10.1016/J.MOLCEL.2014.04.003
Zilbovicius, M., Garreau, B., Samson, Y., Remy, P., Barthelemy, C., Syrota, A., & Lelord, G. (1995). Delayed maturation of the frontal cortex in childhood autism. *American Journal of Psychiatry*, *152*(2), 248–252. https://doi.org/10.1176/ajp.152.2.248
Zuccato, C., & Cattaneo, E. (2009). Brain-derived neurotrophic factor in neurodegenerative diseases. *Nature Reviews Neurology*, *5*(6), 311–322. https://doi.org/10.1038/nrneurol.2009.54