Analysis of region-specific DNA methylation changes in age-related macular degeneration: evaluating a role for the major risk factor smoking, and folate metabolism genes

A thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Master of Philosophy

By

Sean Daniel Duffy

May 2021

Table of Contents

List of Figures 5

List of Tables 8

Abbreviations 10

Abstract 12

Acknowledgements 14

1 Chapter 1: Introduction 15

1.1 Age-related macular degeneration 15

1.1.1 Anatomy of the human eye with relevance to AMD 16

1.1.2 Prevalence and risk factors for AMD 19

1.1.3 Smoking as a risk factor of AMD 20

1.1.4 Diet as a risk factor of AMD 21

1.1.5 Current treatment modalities for AMD 22

1.2 A role for epigenetics in AMD 25

1.2.1 DNA methylation – mechanisms and function 26

1.2.2 Modification of DNA methylation 27

1.2.3 DNA methylation is tissue specific 29

1.2.4 Whole blood tissue vs RPE in investigating AMD epigenetics 29

1.2.5 Comparing the methylomes of blood and RPE 31

1.2.6 DNA methylation as a target for disease therapies 31

1.2.7 DNA methylation in AMD – evidence from the literature 33

1.3 Differential methylation of AMD RPE by Porter *et al* 35

1.3.1 General Transcription Factor IIH Subunit 4 37

1.3.2 SKI Proto-oncogene 37

1.3.3 Tenascin XB 38

1.4 Modifiable AMD risk factors and DNA methylation 39

1.4.1 Effects of smoking on the AMD epigenome 39

1.4.2 The folate pathway and AMD 39

1.4.3 The folate pathway in DNA methylation 41

1.4.4 Association of homocysteine and AMD 43

1.5 Project aims and objectives 44

1.6 Introduction to methods used 45

1.6.1 Using bisulphite pyrosequencing to observe methylation changes 45

1.6.2 DNA methylation arrays to study DNA methylation 48

2 Chapter 2: Materials and Methods 49
2.1 Sample collection, grading and DNA extraction 49
2.2 Ethics statement 50
2.3 Agarose gel electrophoresis 50
2.4 PCR primer optimisation 51
2.5 Polymerase chain reaction 52
2.6 Bisulphite DNA conversion 53
2.7 Methylation standard curve generation 54
2.8 Bisulphite pyrosequencing 54
2.9 Statistical analysis 55
2.10 Bioinformatic analyses (Hewitt *et al* 2017, Joehanes *et al* 2016) 56
2.11 Interrogation of mRNA sequencing data (Kim *et al* 2018) 57
2.12 Transcription factor binding prediction analysis 58
3 Chapter 3: Results I: Data mining and bioinformatic analyses of differentially methylated and expressed genes in AMD RPE (*SKI*, *GTF2H4* and *TNXB*) and genes implicated in folate metabolism 59
3.1 Introduction to chapter 3 59
3.2 Results 61
3.2.1 RNA Sequencing data shows differential expression of *SKI* and *TNXB*, as well as folate pathway genes *AHCY* and *MTR* in AMD RPE. 61
3.2.2 Univariate analyses of CpG loci β-values at loci associated with genes involved in the TFIIH complex and folate and methionine biosynthesis pathways genes in Porter et al. 2019 reveals candidate target genes for further analysis 66
3.2.3 Methylation of *GTF2H4*, *SKI*, *TNXB*, and folate pathway genes MTR and MAT2A show similar baseline methylation levels in blood and RPE. 72
3.2.4 Smoking sensitive methylation is observed at CpG loci within the differentially methylated genes in AMD *SKI* and *TNXB* 74
3.2.5 Transcription factor binding site predictions at selected CpG sites 78
3.3 Discussion for chapter 3 80
4 Chapter 4: Results II: Bisulphite pyrosequencing of genes of interest in whole blood-derived DNA of patients with and without AMD 84
4.1 Introduction to chapter 4 84
4.2 Results 85

4.2.1 Whole blood DNA samples show little signs of degradation 85
4.2.2 Bisulphite pyrosequencing assays detect DNA methylation changes at the candidate DNA loci 87
4.2.3 Significantly differential methylation is observed between control and AMD in genomic regions associated with *SKI* and *GTF2H4* in whole blood 89
4.2.4 Bisulphite pyrosequencing of folate pathway genes: significant increase in methylation was observed within *MAT2A* in whole blood AMD 96
4.2.5 Validation cohort studies show differential methylation of *SKI* in AMD whole blood 103
4.2.6 Combined cohort analysis of *SKI*, *GTF2H4* and *MAT2A* reveals differential methylation in *GTF2H4* and *MAT2A* methylation (p = <0.0001) in AMD whole blood and a significant effect of smoking status on methylation levels within *SKI, GTF2H4* and *MAT2A*. 108
4.3 Discussion for chapter 4 114
5 Chapter 5: Conclusions and future perspectives 119
6 Chapter 6: References 127
7 Chapter 7: Appendix 139

List of Figures

Figure 1.1	Anatomy of the eye and layers of the retina	17
Figure 1.2	Visual distinction between atrophic and neovascular AMD	18
Figure 1.3	Comparison of healthy and AMD retinal tissue	18
Figure 1.4	Structures of cytosine and 5-methylcytosine	26
Figure 1.5	Overview of the overlapping folate and methionine biosynthesis pathways	41
Figure 1.6	Overview of sodium bisulphite conversion workflow	46
Figure 3.1	RNA sequencing data for candidate gene *AHCY* obtained from Kim et al. (2018)	64
Figure 3.2	RNA sequencing data for candidate gene *MTR* obtained from Kim et al. (2018)	65
Figure 3.3	Histone modification data at identified smoking sensitive CpG sites within SKI	78
Figure 4.1	Example of DNA degradation analysis	86
Figure 4.2	Age analysis of AMD and control whole blood samples for further study	87
Figure 4.3	Methylation standard curves for target candidate genes	88
Figure 4.4	Bisulphite pyrosequencing of *SKI* in CTRL and AMD whole blood samples	90
Figure 4.5	Bisulphite pyrosequencing of *GTF2H4* in CTRL and AMD whole blood samples	90
Figure 4.6	Bisulphite pyrosequencing of *TNXB* in CTRL and AMD whole blood samples	91
Figure 4.7	Bisulphite pyrosequencing of *SKI* in CTRL and AMD whole blood samples, stratified by smoking status	92
Figure 4.8	Bisulphite pyrosequencing of *GTF2H4* in CTRL and AMD whole blood samples, stratified by smoking status	93
Figure 4.9	Bisulphite pyrosequencing of *TNXB* in CTRL and AMD whole blood samples, stratified by smoking status	94
Figure 4.10	Bisulphite pyrosequencing of *SKI* CpG position 4 in CTRL and AMD whole blood samples, stratified by smoking status	96

Figure 4.11	Bisulphite pyrosequencing of *AHCY* in CTRL and AMD whole blood samples	97
Figure 4.12	Bisulphite pyrosequencing of *MTR* in CTRL and AMD whole blood samples	98
Figure 4.13	Bisulphite pyrosequencing of *MAT2A* in CTRL and AMD whole blood samples	98
Figure 4.14	Bisulphite pyrosequencing of *AHCY* in CTRL and AMD whole blood samples, stratified by smoking status	100
Figure 4.15	Bisulphite pyrosequencing of *MTR* in CTRL and AMD whole blood samples, stratified by smoking status	101
Figure 4.16	Bisulphite pyrosequencing of *MAT2A* in CTRL and AMD whole blood samples, stratified by smoking status	102
Figure 4.17	Bisulphite pyrosequencing validation of *SKI* in CTRL and AMD whole blood samples	103
Figure 4.18	Bisulphite pyrosequencing validation of *GTF2H4* in CTRL and AMD whole blood samples	104
Figure 4.19	Bisulphite pyrosequencing validation of *MAT2A* in CTRL and AMD whole blood samples	104
Figure 4.20	Bisulphite pyrosequencing validation of *SKI* in CTRL and AMD whole blood samples, stratified by smoking status	105
Figure 4.21	Bisulphite pyrosequencing validation of *GTF2H4* in CTRL and AMD whole blood samples, stratified by smoking status	106
Figure 4.22	Bisulphite pyrosequencing validation of *MAT2A* in CTRL and AMD whole blood samples, stratified by smoking status	107
Figure 4.23	Combined cohort bisulphite pyrosequencing results in CTRL and AMD whole blood samples	109
Figure 4.24	Combined cohort bisulphite pyrosequencing results of *SKI* in CTRL and AMD whole blood samples, stratified by smoking status	110
Figure 4.25	Combined cohort bisulphite pyrosequencing results of *GTF2H4* in CTRL and AMD whole blood samples, stratified by smoking status	111
Figure 4.26	Combined cohort bisulphite pyrosequencing results of *MAT2A* in CTRL and AMD whole blood samples, stratified by smoking status	112

Figure 7.1 *GTF2H4* RNA sequencing data from Kim et al (2017) 147
Figure 7.2 *SKI* RNA sequencing data from Kim et al (2017) 148
Figure 7.3 *TNXB* RNA sequencing data from Kim et al (2017) 149
Figure 7.4 *MAT2A* RNA sequencing data from Kim et al (2017) 150
Figure 7.5 *GTF2H5* RNA sequencing data from Kim et al (2017) 151
Figure 7.6 *GTF2H1* RNA sequencing data from Kim et al (2017) 152
Figure 7.7 *GTF2H2* RNA sequencing data from Kim et al (2017) 153
Figure 7.8 *GTF2H3* RNA sequencing data from Kim et al (2017) 154
Figure 7.9 *ERCC2* RNA sequencing data from Kim et al (2017) 155
Figure 7.10 *ERCC3* RNA sequencing data from Kim et al (2017) 156
Figure 7.11 *DNMT1* RNA sequencing data from Kim et al (2017) 157
Figure 7.12 *DNMT3A* RNA sequencing data from Kim et al (2017) 158
Figure 7.13 *DNMT3B* RNA sequencing data from Kim et al (2017) 159
Figure 7.14 *SHMT1* RNA sequencing data from Kim et al (2017) 160
Figure 7.15 *DHFR* RNA sequencing data from Kim et al (2017) 161
Figure 7.16 *DHFRL1* RNA sequencing data from Kim et al (2017) 162
Figure 7.17 *CBS* RNA sequencing data from Kim et al (2017) 163
Figure 7.18 *MTHFR* RNA sequencing data from Kim et al (2017) 164

List of Tables

Table 1.1	Minnesota Grading System for AMD classification in post-mortem tissue AMD, adapted from AREDS classification system	16
Table 1.2	Summary of literature exploring DNA methylation changes of genes in AMD	35
Table 1.3	Identification of DML in AMD RPE by Porter *et al* (2019)	36
Table 1.4	Identification of DMR in AMD RPE by Porter *et al* (2019)	36
Table 2.1	Sample information overview for cohort 1 blood samples (Netherlands)	49
Table 2.2	Sample information overview for cohort 2 blood samples (Southampton)	49
Table 2.3	Sample information overview for combined cohort 1 and 2 blood samples	50
Table 2.4	Sequence of primers used in PCR and bisulphite pyrosequencing reactions	52
Table 2.5	Polymerase chain reaction thermocycling conditions	53
Table 2.6	Target gene PCR annealing temperatures	53
Table 2.7	Patient information for samples assessed from Hewitt et al (2017) study	56
Table 2.8	Patient demographics from Joehanes et al (2016) study	57
Table 2.9	Patient information for samples used in Kim et al (2018) study	58
Table 3.1	mRNA expression of genes of interest in AMD by Kim *et al* (2018)	63
Table 3.2:	Univariate analyses of Porter et al. 450k genome-wide DNA methylation array data at CpG loci associated with TFIIH complex genes.	67
Table 3.3:	Univariate analyses of Porter et al. 450k genome-wide DNA methylation array data at CpG loci associated with folate pathway genes	67
Table 3.4:	CpG loci identified as significant following univariate analysis of Porter *et al.* (2019) data	68
Table 3.5:	Summary of univariate analysis of Porter et al data at TFIIH complex genes	69

Table 3.6: Summary of univariate analyses of Porter et al. data for folate pathway genes 70

Table 3.7: Summary of CpG loci of interest for further study in AMD whole blood 72

Table 3.8: Mean methylation levels of CpG sites of interest in blood and RPE / choroid tissue reported in Hewitt *et al*. (2017) 74

Table 3.9: Mean methylation β-levels of CpG sites of interest within *SKI* measured in whole blood and RPE using bisulphite pyrosequencing 74

Table 3.10: “Smoking sensitive” CpG loci observed within *SKI* in Joehanes et al. (2016) in "former" smokers 76

Table 3.11: “Smoking sensitive” sites observed within *TNXB* in Joehanes et al. (2016) 77

Table 3.12: Transcription factor binding predictions obtained from the TRANSFAC database 79

Table 4.1: There is clear correlation between predicted and observed methylation from methylation standard curves 88

Table 7.1 Individual sample information for Cohort 1 (Netherlands) 192 blood DNA samples 139

Table 7.2 Individual sample information for Cohort 2 (Southampton) 96 blood DNA samples 144

Abbreviations

AMD	Age-related macular degeneration
AREDS	Age-related eye disease study
BER	Base-excision repair
CFH	Complement factor H
CNV	Choroidal neovascularisation
DHA	Docosahexaenoic acid
DHF	Dihydrofolate
DML	Differentially methylated locus / loci
DMR	Differentially methylated region
DNMT	DNA methyltransferase
ECM	Extracellular matrix
EWAS	Epigenome-wide association study
G2 AMD	Grade 2 / intermediate AMD
G3 AMD	Grade 3 / advanced AMD
GA	Geographic atrophy
hESC	Human embryonic stem cell
HLA	Human leukocyte antigen
MBD	Methyl-CpG-binding domain
MGS	Minnesota grading system
miRNA	MicroRNA
NER	Nucleotide excision repair
PBMC	Peripheral blood mononuclear cell
PCR	Polymerase chain reaction
PRCS	Peripheral RPE / choroid / sclera
ROS	Reactive oxygen species
RPE	Retinal pigment epithelium
SAH	S-adenosyl homocysteine
SAM	S-adenosyl methionine
SCLC	Small cell lung cancer
TGF-β	Transforming growth factor β

THF	Tetrahydrofolate
TSS	Transcription start site
UTR	Untranslated region
VEGF	Vascular endothelial growth factor

Abstract

Age-related macular degeneration (AMD) is a degenerative disease of the central retina and the foremost cause of blindness, predicted to affect up to 288 million people globally by the year 2040. Besides age, important risk factors include smoking and a diet low in folic acid, with oral supplementation shown to reduce AMD risk in epidemiological studies. Both smoking and dietary folate play crucial roles in DNA methylation. In AMD, differential methylation and expression of the genes *SKI*, *GTF2H4* and *TNXB* were recently observed in AMD retinal pigment epithelium. However, collection of retinal pigment epithelium tissue post-mortem presents major challenges for large-scale epigenetic studies.

My thesis explores the use of whole blood as a surrogate tissue for retinal pigment epithelium at observed DNA methylation change regions noted in Porter *et al*. (2019) at selected candidate genes, including *SKI*, *GTF2H4* and *TNXB.* The aim was to evaluate the potential to use DNA methylation changes in whole blood as a potential biomarker for early/intermediate AMD. I also investigated the possible impact of smoking on DNA methylation in AMD at target loci including *SKI, GTF2H4* and *TNXB* in blood-derived DNA, as well as differential methylation of genes involved in folate metabolism, with the aim of evaluating a role for major AMD environmental risk factors on DNA methylation. To this end, bioinformatic analyses were carried out to identify candidate folate pathway genes for analysis. Secondly, a first-line assessment of predicted reproducibility of baseline DNA methylation levels at target CpG regions across ocular tissue and whole blood was carried out. Results revealed that methylation levels of selected CpG loci of interest were comparable between retinal pigment epithelium and whole blood. A refined list of candidate genes based on published associations with smoking and AMD was then generated for study.

Further results generated using bisulphite pyrosequencing of patient whole blood revealed differential methylation in genomic regions within *SKI* and *GTF2H4* in AMD. Significantly increased methylation levels were also observed in the *MAT2A* gene in AMD whole blood. Importantly, *MAT2A* is involved in folate metabolism,

responsible for the biosynthesis of S-adenosyl methionine, the universal methyl donor, suggesting a role for differential methylation of a folate pathway gene in AMD. The CpG loci within *SKI* and *MAT2A* were also found to be significantly differentially methylated in AMD smokers, suggesting smoking-induced DNA methylation changes as a potential driver of AMD development.

Overall, this thesis forms the basis for future studies into the whole blood epigenome of AMD providing insights into a role for environmental risk factors influencing DNA methylation in AMD.

Acknowledgements

Firstly, I would like to thank my supervisor, Dr Louise Porter, for allowing me to work on the project and for her support and assistance throughout. I am very grateful to have been given this opportunity to learn so much about epigenetics, ageing and eye and vision science, and to have had the guidance of a supervisor who has offered unwavering support over the course of the project.

Thank you to Prof Luminita Paraoan for help and assistance especially in getting started with the project, and for guidance with regards to proper experimental technique and sound reporting of results. Thank you to Dr Lakis Liloglou, from whom I am incredibly grateful to have learned so much about pyrosequencing, and for always maintaining enthusiasm and a desire to teach, despite my seemingly never-ending stream of questions.

Thank you to Neil Saptarshi, who has been a pleasure to work alongside over the last few years, has helped me keep a level head and has provided plenty of helpful advice, both in and out of the lab. Thank you to all my friends and colleagues in the Institute of Ageing and Chronic Disease and Institute of Translational Medicine, for making these days that much brighter. And thank you to all of my friends outside of science, for always being ready for a pint and for politely asking how Uni is going, even if you come to regret the stream of nonsense that follows from me in response.

Thank you to my family, my mum, dad, brother, sister-in-law, and to Joshua, my nephew and latest arrival into the Duffy family, for always being there for me. I hope I can continue to make you all proud.

Finally, thank you to Beth – your love and support over these past few years has kept a smile on my face as we braved postgraduate study together. Most importantly, thank you for not complaining when, for almost two years, seemingly 99% of text messages I sent contained the word “pyrosequencing”.

Chapter 1: Introduction

1.1 Age-related Macular Degeneration

Age-related macular degeneration (AMD) is a degenerative disease of the central retina and the leading cause of irreversible loss of vision in the western world (1). The prevalence of AMD is increasing over time, with cases predicted to increase to 288 million by 2040 due to an increasingly ageing population worldwide (2). The primary symptom of AMD is central vision loss, either rapid (in the case of neovascular AMD) or gradual (non-neovascular AMD) in one or both eyes, leading to blurred vision, difficulties with reading, and difficulties recognising faces (3).

AMD is classified using the Age-Related Eye Disease Study (AREDS) classification, categorising AMD into early, intermediate or advanced stages. Early and intermediate stage AMD cases are classified based on the presence and size of sub-retinal deposits of lipids and proteins known as drusen, while in the advanced stage of the disease two forms, termed “neovascular” and “atrophic” AMD, are observed: “neovascular” AMD is characterised by choroidal neovascularisation (CNV) in the subretinal/intraretinal layers of the macular retina, with subsequent bleeding and/or fluid leakage causing rapid loss of central vision. “Atrophic” AMD involves geographic atrophy (GA), categorised by progressive degeneration of the retinal pigment epithelium (RPE) and photoreceptors (4, 5). Consensus of AMD prevalence studies suggests that atrophic AMD is as many as 1.7 times more prevalent than neovascular AMD, with an approximated annual incidence of 160,000 new atrophic cases in over-50 white Americans alone (6, 7). The Minnesota Grading System (MGS) was introduced following the AREDS study to evaluate AMD donor tissue using colour fundus photographs, and to allow classification of these donor eyes according to the AREDS grading system (Table 1.1) (8).

Table 1.1: Minnesota Grading System for AMD classification in post-mortem tissue AMD, adapted from AREDS classification system.

MGS AMD Level	AREDS Identifiers	AMD Classification
1	• Drusen maximum size < 63µm • Drusen total area < 125µm	No AMD
2	• Drusen max size (µm): 63 ≤ X < 125, or • Drusen total area ≥ 125µm, or • RPE abnormalities consistent with AMD	Early AMD
3	• Drusen maximum size ≥ 125µm, or • Drusen max size ≥ 63µm, total area > 180µm AND type identified as soft, distinct, or • Drusen max size ≥ 63µm, total area > 660µm AND type identified as soft, indistinct, or • Geographic atrophy (atrophy > 180µm) at grid, but not at centre of macula	Intermediate AMD
4	• Geographic atrophy at centre of subfield, with at least questionable involvement at centre of macula, or • Evidence of choroidal neovascularisation	Advanced AMD

MGS: Minnesota Grading System

1.1.1 Anatomy of the human eye with relevance to AMD

The human eye is a complex organ comprising of many different components, each with a specific role contributing to visual function (9). AMD is a disease of the posterior segment of the eye (the part of the eye located behind the lens) affecting the retina, retinal pigment epithelium and its basement membrane (Bruch's membrane), the associated sub-RPE vasculature, the choriocapillaris.

The retina and associated tissues is where AMD manifests. It is comprised of several distinct layers, that include the photoreceptor layer, with rod and cone photoreceptors. Rods are particularly light-sensitive and are primarily responsible for peripheral vision as well as vision in low-light conditions, while cones are responsible for detailed central vision, and are predominantly clustered near the centre of the retina, in an area known as the macula, which is where AMD changes

concentrate (10). Beyond the photoreceptor layer, lies the RPE, a single layer of post-mitotic cells responsible for the supply of nutrients to photoreceptors, and Bruch's membrane, a five-layered extracellular matrix separating the retina from the choroid. It is situated between the inner collagenous layer of Bruch's membrane and the RPE basement membrane and is the site where drusen accumulate (Figure 1.1) (11, 12). Accumulation of drusen at the level of Bruch's membrane in the between the subretinal pigment epithelial layer leads to metabolic stress within the RPE, possibly through interference with one of the primary RPE functions of fluid efflux from the RPE through Bruch's membrane, and acts as a hallmark of early AMD (13). The advanced stages are classified by the presence of larger and more numerous drusen, as well as specific changes to the macular region leading to the aforementioned "neovascular" and "atrophic" forms (14). Neovascular AMD is driven by over expression of vascular endothelial growth factor (VEGF) leading to the development of new blood vessels in the subretinal macula, while atrophic AMD is driven by larger drusen deposits and resulting RPE and photoreceptor cell death (Figures 1.2, 1.3).

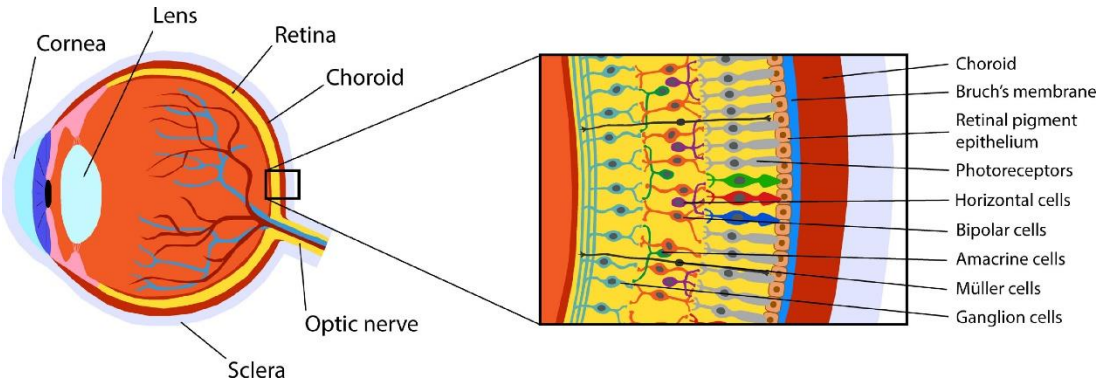

Figure 1.1: Anatomy of the eye and layers of the retina. Image taken from "Bio-inspired human in vitro outer retinal models: Bruch's membrane and its cellular interactions" by Murphy et al (15)

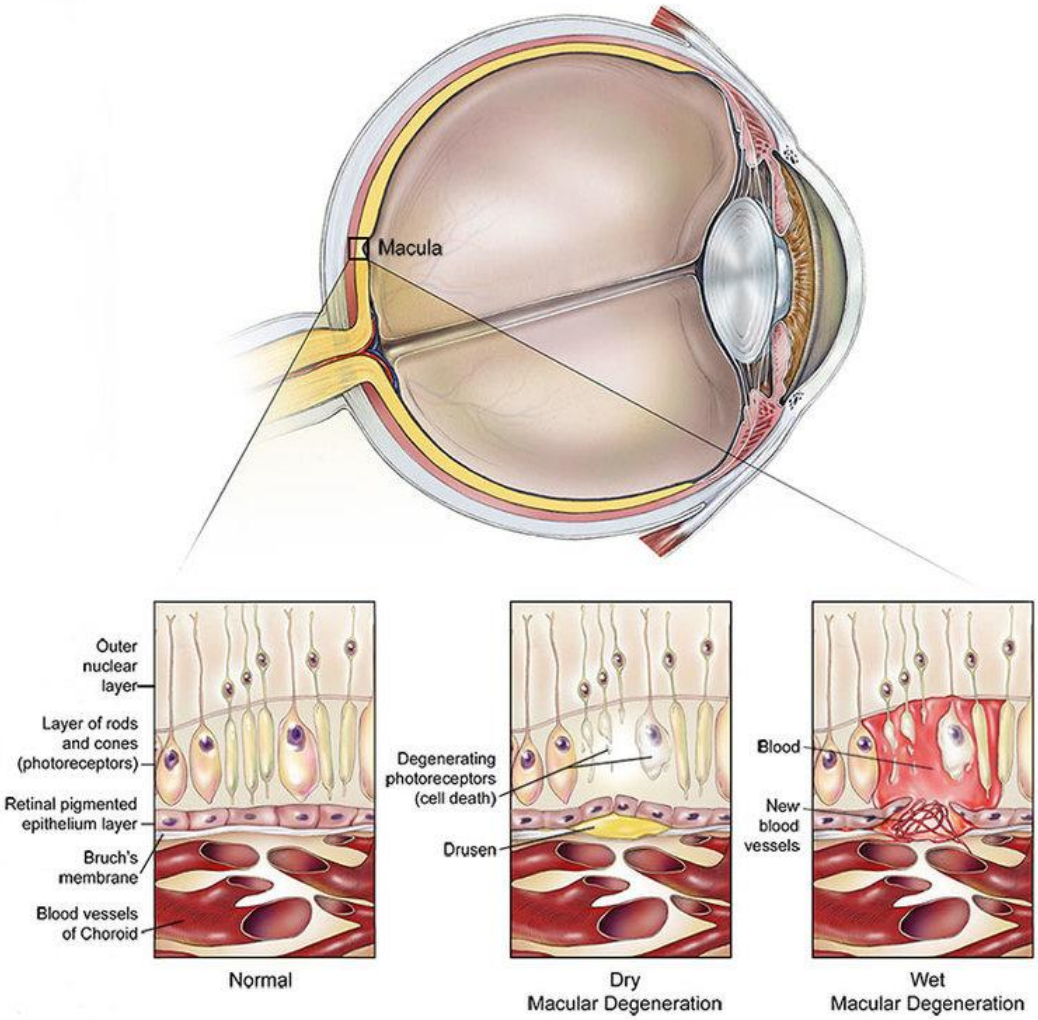

Figure 1.2: Visual distinction between atrophic and neovascular AMD. Image taken from www.brightfocus.org

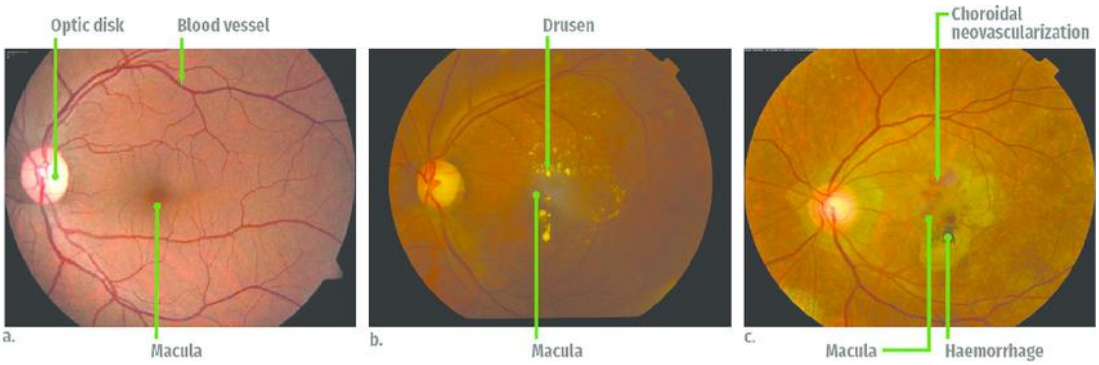

Figure 1.3: Comparison of healthy and AMD retinal tissue. Left: healthy; Centre: dry AMD; Right: Neovascular AMD. Fundus photographs obtained from Acharya *et al*. (2016). (16)

1.1.2 Prevalence and risk factors for AMD

Prevalence of AMD is higher in white European populations compared to Asian or African populations, across all forms and stages of the disease (2), though gender does not appear to be a principal risk factor despite some studies observing increased reports of CNV in women with AMD (17, 18). Age is the primary risk factor of AMD, with those over 70 years of age displaying symptoms of AMD at a significantly higher rate than younger cohorts, and AMD prevalence reported to increase four-fold per decade beyond this point (18, 19). However, a number of genetic and environmental risk factors have been associated with AMD development and progression with the major genetic risk loci located on chromosomes 1 and 10 associated with complement factor H (CFH) gene (chr1) and *ARMS2/HTRA1* (chr 10) (20). Genes involved in the complement system, a cellular system involved in the innate immune inflammatory responses, have been most frequently associated with AMD, specifically a genetic polymorphism associated with the CFH coding gene. Multiple studies have identified a significant association between AMD development and the Y402H tyrosine to histidine substitution in *CFH*. Furthermore, drusen from AMD patients contain a number of components of the complement pathway, including CFH and components C3 and C5 (21-25). More recent studies have identified a substantial number of further genetic variants associated with advanced AMD, although despite these new insights in the field of AMD genetics, the impacts of these genetic variations on biological function are not clearly defined (26-29).

Modifiable environmental risk factors for AMD include dietary intake of fat, omega-3 fatty acids and antioxidants, as well as smoking, helping to define the complex relationship between AMD development and the environment. Smoking is one of the major consistently reported modifiable risk factors for AMD. When compared to never-smokers, there is evidence from several studies to suggest that both current- and former-smokers are at greater risk of AMD, with the risk of developing AMD being two-fold or greater in current regular smokers compared to non-smokers (30). Importantly, the risk of AMD has been found to be reduced in former

smokers compared to current smokers, suggesting the potential for the reversibility or slowing of progression of AMD through cessation of smoking (31).

Tobacco use is one of the foremost lifestyle choices responsible for ill health, with the WHO estimating the number of tobacco users to have reached 1.3 billion worldwide, and tobacco smoking being the primary route of tobacco exposure (32). Responsible for an approximated five million deaths worldwide each year, tobacco smoking is reported as one of the leading preventable causes of death in adults (33). While the number of smokers in the UK has declined in recent years, smoking remains one of the primary causes of cancers, particularly lung cancer, as well as cardiovascular and respiratory disease, infertility, and pregnancy complications such as increased risk of miscarriage (34-37).

Stopping smoking is considered a healthy lifestyle choice and is widely recommended, with former smokers benefitting from greater health prospects compared to current smokers, for instance through lowering of the characteristic high blood pressure of smokers, as well as through decreased risk of stroke and increased general longevity (38-40). Evidence can be observed supporting the reversibility of genetic expression changes brought about by smoking (41), although this is not true in all cases – the expression of numerous smoking related genes associated with lung cancer are still found to be dysregulated even decades after cessation of smoking, and are reported to be potentially responsible for the elevated risk of lung cancer among former smokers compared to never-smokers (42). While many current smokers are aware of the risks of continued tobacco use and express a desire to quit, successful occurrences of long-term abstinence do not reflect this, owing to the difficulty to commit to permanently quitting smoking (43). Nonetheless, cessation of smoking where possible is one of the primary positive lifestyle changes advised to mitigate the risk of AMD.

1.1.3 Smoking as a risk factor of AMD

Smoking is one of the major consistently reported modifiable risk factors in AMD. When compared to never-smokers, there is evidence from several studies that both current- and former-smokers are at greater risk of AMD. A study by Khan *et al*

(2006) comprising of 435 AMD cases and 280 controls from the UK between 2001 and 2003 demonstrated the effect of cigarette smoking on the prevalence of AMD, with odds ratios for AMD, GA and CNV being highest in current smokers, while those who had stopped smoking for 20 years or more displayed odds ratios comparable to those of never-smokers, highlighting the importance of stopping smoking as a method of reducing the risk of AMD (44). Evidence for increased occurrence of oxidative stress in smokers has also been observed, highlighting the potential for smoking to be a driver of oxidative stress-induced RPE damage and AMD progression (45).

Three major studies on the effects of cigarette smoking on AMD based in Europe (Rotterdam Study), North America (Beaver Dam Eye Study) and Australia (Blue Mountains Eye Study) totalling 14752 participants all conclude that there is an association between smoking and the risk of AMD, with smokers being at higher risk of developing large drusen than non-smokers. A pooled study of these three analyses by Smith *et al* (2001) concluded that cigarette smoking was the prime modifiable risk factor of AMD in these studies, being the only risk factor consistently associated with AMD within the studies excluding age (46-49).

Seddon *et al* (2011) demonstrated in a study of 840 individual monozygotic twins that the heavier-smoking twin was more likely to have the more advanced stage of AMD, as well as larger drusen area, than the lighter-smoking twin. This suggests an association between modifiable environmental risk factors and AMD on an identical genetic background (50).

1.1.4 Diet as a risk factor of AMD

After smoking, diet is considered the largest modifiable risk factor of AMD. The AREDS study concluded that the consumption of dietary antioxidants, including carotenoids (such as lutein and zeaxanthin), commonly found in fruits and vegetables, as well as supplementation of zinc and folic acid was responsible for decreased risk of CNV and GA progression, as well as a decreased risk of developing intermediate or large drusen (51). Antioxidants regulate the levels of reactive oxygen species (ROS) which, when allowed to accumulate, can lead to oxidative

stress-induced cellular damage and apoptosis. The observed protective effect of antioxidants in AMD is thought to occur due to the role of antioxidants as ROS scavengers, reducing ROS-induced oxidative stress on retinal and RPE function (52), with studies demonstrating the ability of antioxidants to reduce oxidative damage of H_2O_2, a known ROS, in RPE cells (53). This hypothesis is supplemented by the observation that mice (aged 2-18 months) lacking the Nuclear factor erythroid 2-related factor 2 (*NRF2*) gene, known as a master regulator of the response to oxidative stress, were found to develop AMD-like retinal pathology (54, 55).

Similarly, dietary fats, including saturated, trans-unsaturated and poly-unsaturated fatty acids have long been associated with AMD progression to varying extents – several studies describe the relationship between dietary saturated, monounsaturated and trans-unsaturated fatty acid intake and increased risk of AMD (56-58). Conversely, the protective nature of dietary omega-3 fatty acids is well documented. Docosahexaenoic acid (DHA), a long-chain omega-3 fatty acid obtained mainly from oily fish, is concentrated in the retina and plays a role in photoreceptor differentiation during early retinal development, as well as subsequent retinal function (59). DHA has been reported to have a protective effect in early AMD, including in mouse models where a higher diet of omega-3 fatty acids resulted in a slower progression of retinal lesions (60, 61). Likely as a result of such studies, a recent survey of optometrists in Australia and New Zealand revealed that 79% of respondents recommended omega-3 fatty acid supplements to their patients in an effort to promote eye health in cases of AMD (62).

1.1.5 Current treatment modalities for AMD

To date, there is no effective treatment for the advanced atrophic form of AMD. Inhibitors of the complement pathway, such as C3 and C5, have been trialled for use in GA treatment, but have proven unsuccessful – the C3 inhibitor Compstatin showed early promise in drusen diffusion, however was aborted during Phase II trials due to lack of replication of these results, while AMD trials of the C5 inhibitor Eculizamab demonstrated no effect on GA progression, despite facilitating significant downregulation of C5 activity (63).

Standard treatment for neovascular AMD involves the use of regular or periodic (monthly, bi-monthly, as required protocols) intravitreal anti-VEGF injections, such as Lucentis (Ranibizumab) and Avastin (Bevacizumab) (64). VEGF-A is a known regulator of angiogenesis, and multiple animal models have demonstrated both the relationship between VEGF overexpression and CNV development, as well as between the application of anti-VEGF therapies and a reduction in further development of CNV and reduced leakage observed in eyes with established CNV (65-67). This antiangiogenic effect was also observed in subsequent clinical trials studies in humans (68). In any case, it is important to note that while anti-VEGF therapies demonstrate an effective means of slowing AMD progression in patients with CNV, reducing the amount of scarring resulting from haemorrhage, these therapies do not treat the atrophic consequences and subsequent progression to significant macular atrophy with gradual loss of central vision despite previous successful anti-VEGF treatment. An association has been observed between anti-VEGF treatment for neovascular AMD and increased GA development (69).

In addition, despite the relative success of intravitreal anti-VEGF treatments for neovascular AMD, the nature of these repeated monthly injections means that treatment may be required frequently for many years by a trained specialist and represent a significant burden to patients and healthcare providers. Moreover, repeated injections over a prolonged period increase the risk of the adverse effects including ocular pain and inflammation, as well as endophthalmitis, which itself can lead to a loss of vision (70). A novel idea to alleviate these issues is the proposal of a single intravitreal injection of a gene delivery vector such as an adenovirus, which would allow an antiangiogenetic protein to be expressed in the host to counter the neovascularisation observed in wet AMD patients (71).

Other suggested alternative approaches have targeted the RPE, a primary site of disease pathogenesis. One potential approach to AMD treatment would be reparation of the RPE layer, for instance through cell transplantation, and/or surgical removal of CNV in AMD eyes. Examples of the latter however have revealed no significant improvement in visual acuity following surgery, with mechanical disturbance of the RPE layer (72). More recently, through developments in stem cell

technology, human embryonic stem cells (hESCs) have been differentiated into RPE cells and used to alleviate visual loss in mouse models with reported near-normal levels of visual function following transplantation. This has since been studied in small-scale human trials with hESC-differentiated RPE cells being transplanted after successful surgical removal of the neovascular membrane in three wet AMD patients (73, 74). While visual function was improved following RPE transplantation in humans, the positive effects were variable between patients, with all three requiring follow-up surgeries to remove silicon oil as well as cataracts forming post-transplantation. Clearly, further improvements to transplantation methodologies are required before this can be considered a plausible treatment strategy for AMD, or as an alternative to anti-VEGF therapy in the treatment of neovascular AMD specifically.

Of major concern is the notion that in the "late" stages of AMD, where GA and / or CNV is observed, the disease may be too advanced for novel treatment strategies to be effectively assessed, with high levels of irreparable retinal and RPE cellular damage and death. Earlier administration of a treatment may be required to assess efficacy in clinical trials.

In conclusion, therapeutic measures to treat AMD are limited to the treatment of advanced neovascular AMD. Following a diagnosis of early or intermediate AMD (in the context of visual decline and identification of drusen and RPE pigmentary changes), dietary and lifestyle changes are recommended as methods of prevention of further AMD progression (75). As a result of these combined studies, dietary intervention to include a diet rich in fruits, vegetables and fish should be considered a crucial step in the treatment regimen of early AMD patients, in addition to antioxidant supplementation. Furthermore, studies demonstrating the reduced prevalence of AMD in former smokers compared to current smokers highlight the importance of stopping smoking as a method of reducing the risk of AMD and reducing the risk of CNV in particular – cigarette smoking is reported to significantly increase the risk of progression from atrophic AMD to neovascular AMD, with this progression observed to occur at an earlier age in smokers (76, 77).

1.2 A role for epigenetics in AMD

The terms “epigenetics” or "epigenetic changes" refer to modifications to DNA that to do not directly alter the genetic code itself, and result in changes in DNA accessibility to transcription factors, for example by conformational changes or chemical modifications to DNA bases, which can in turn result in alteration of gene expression (78). The principal forms of epigenetic regulation include chemical modifications to DNA (methylation for example), chemical modifications to histones (including methylation, acetylation, ubiquitination, sumoylation), chromatin conformation changes mediated by chromatin remodelling enzymes, and microRNA (miRNA) regulation (79).

As epigenetic processes can be regulated by environmental factors, it can be argued that epigenetic change reflects a regulatory interface between the genome and the environment, enabling adaptive or pathogenic changes in gene expression.

Histones are a family of proteins found in eukaryotic cell nuclei that, through electrostatic forces between their basic positive charge and the negative charge of DNA, fold and condense DNA into chromatin. Post-translational modifications to these histones (including methylation, acetylation, ubiquitination, etc.) can impact their interaction with DNA and therefore elicit gene expression changes through alteration of chromatin structure. This is particularly observed with several post-translational modifications at particular amino acid residues, forming the "histone code". Modifications of histone H3 for example, such as mono- and tri-methylation of lysine 4 (H3K4me1, H3K4me3), as well as acetylation of lysine 27 (H3K27ac), are well documented markers of active promoter and enhancer regions of DNA (80, 81). Conversely, H3K9 methylation is a known marker of transcriptional silencing and heterochromatin structure, a condensed chromatin structure inactive for transcription (82).

miRNAs are short, non-coding RNA molecules averaging 22 nucleotides in length involved in gene expression regulation by binding primarily to the 3’ untranslated region (3’ UTR) of mRNA, resulting in mRNA silencing either by facilitating their degradation or by blocking their translation. miRNA regulation has been

documented to impact a wide variety of biological processes, including apoptosis, immune response, and the cell cycle (83-85). For the purposes of this project, particular attention will be given to the third observed mechanism of epigenetic change – DNA methylation.

1.2.1 DNA methylation – mechanisms and function

DNA methylation is a chemical change that occurs at cytosine residues of eukaryotic DNA, primarily at the 5' position of cytosine, to from 5'-methylcytosine. In eukaryotic cells, DNA methylation of cytosine residues occurs almost exclusively on cytosine residues at CpG dinucleotide sites, areas where a cytosine residue is immediately followed by a guanine residue, linked by a phosphate bond, in the 5' → 3' direction (Figure 1.3) (86).

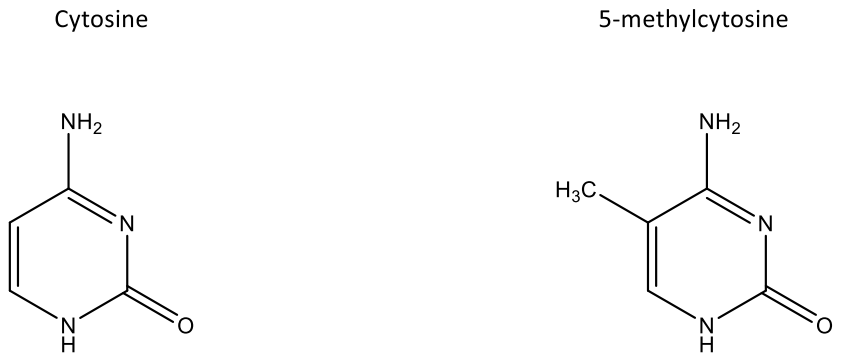

Figure 1.4: Structures of cytosine and 5-methylcytosine

The process of DNA methylation is regulated by a family of enzymes known as DNA methyltransferases (DNMTs). In humans, the catalytically active DNMTs responsible for DNA methylation are DNMT1, DNMT3a and DNMT3b, with s-adenosylmethionine (SAM), a product of the metabolism of methionine, acting as the universal methyl donor (87).

CpG dinucleotides are often found concentrated in areas of DNA known as CpG islands, defined as genomic regions of >200bp with GC content >50% and "observed to expected" CpG ratio of >0.6. Commonly occurring near promoter regions at the transcriptional start site (TSS) of genes, CpG islands are largely unmethylated, and play an important role in regulation of gene expression; methylation of CpG islands

at the promoter region of genes is strongly associated with transcriptional gene silencing (88).

Methylation of cytosine residues within gene promoters generally results in repression of transcription through several pathways, such as transcription factor binding inhibition due to the DNA structure alteration following CpG methylation (89). Additionally, recruited methyl-CpG-binding domain (MBD) family of proteins such as MeCP2, following DNA methylation, in turn associate with a number of separate elements including histone deacetylases, methyltransferases and chromatin remodelers, generating a repressive chromatin state (90).

1.2.2 Modification of DNA methylation

In humans, *de novo* DNA methylation during embryonic development is facilitated by DNMT3a and DNMT3b, with DNMT1 being responsible for subsequent DNA methylation maintenance (91). Following fertilisation, paternal and maternal genomes undergo a wave of demethylation, erasing inherited methylation patterns, before *de novo* methylation establishes a new methylation pattern. *De novo* DNA methylation is primarily observed at this developmental stage, before being largely suppressed following differentiation in the majority of cases (92).

DNA methylation modifications outside developmental processes occur by a variety of factors and in both physiological and pathological conditions. An erosion of DNA methylation patterns over time occurs in ageing for example, a phenomenon termed "epigenetic drift". An Example of this relationship between ageing and DNA methylation was observed in monozygotic twin pairs, where younger twins (3 years old) displayed largely indistinguishable genome-wide methylation patterns (methylomes), while older twins (50 years old) displayed significantly higher levels of variation in methylation levels measured by global genomic 5-methylcytosine content (93). Similarly, a study by Christensen *et al* (2009) employed DNA methylation arrays and bisulphite pyrosequencing techniques to interrogate 1413 distinct CpG loci across 10 different human tissue types. They observed a significant increase in methylation of CpG probes located within genomic regions of CpG island status with increased age, in addition to a significant reduction in methylation of

loci outside genomic regions of CpG island status, in accordance with CpG island definitions previously outlined (94). In addition, a significant correlation has been observed between increasing age and decreased methylation of the transposable elements Alu and long interspersed nuclear element 1 (LINE-1), a relationship which is seen to be exacerbated in later follow-up studies of the same patient samples (95). Typically transcriptionally inactive, retrotransposons such as Alu and LINE-1 are reported to make up an approximated 40% of the human genome, making them prime candidates for the representation of levels of global DNA methylation (96).

An environmental factor responsible for alterations in DNA methylation is diet – specifically, the dietary intake of the range of cofactors and substrates which facilitate DNA methylation. DNMT mediated DNA methylation requires sufficient levels of SAM to act as the universal methyl donor, SAM being a component of the methionine cycle which involves a range of additional dietary nutrients including folate, betaine, and methionine. Diets which are found to be lacking in one or more of these components could theoretically impact the levels of SAM available to be incorporated into the DNA methylation process, and therefore lead to changes in the human epigenome (97).

Several diseases are widely known to be associated with abnormal DNA methylation. Principally, many cancers are found to display atypical CpG methylation at specific loci, particularly tumour suppressors and oncogenes, in breast cancer, leukaemia, and small cell lung cancer (SCLC) (98-101). In addition, monozygotic twin studies have demonstrated discrepancies in epigenetic profiles in patients with autoimmune diseases such as rheumatoid arthritis on an identical genetic background, suggesting a role for epigenetics in the development of such diseases – this is of particular interest given the separate study detailing the relationship between increased methylation of the human leukocyte antigen (HLA) gene and reduced risk of development of rheumatoid arthritis (102, 103).

Although there are many links associating abnormal DNA methylation to disease, the methylation process is reversible, with conversion of 5’-methylcytosine back to its original state possible through two distinct pathways, termed “active” and

“passive” demethylation. Active DNA demethylation is mediated by ten eleven translocation (TET) protein activity, which binds to DNA selectively at CpG cytosine residues, facilitating the conversion of 5’-methylcytosine to a 5’-hydroxymethylcytosine intermediate via oxidation, before further TET-mediated conversion to 5’-formylcytosine and subsequently 5’-carboxylcytosine, which is finally replaced with an unmethylated cytosine residue through base-excision repair (BER). In contrast, passive DNA methylation refers to the gradual dilution of methylated cytosine residues through repeated DNA replication, due to dysfunctional activity of enzymes responsible for methylation such as DNMT1 (104, 105).

1.2.3 DNA methylation is tissue specific

Owing to the ability of DNA methylation to regulate gene expression and therefore influence cellular processes, DNA methylation profiles are not equal across all tissues, with each tissue type displaying a distinct epigenetic signature, reflecting their individual functions (106). It has been demonstrated that variation in tissue DNA methylation occurs at a particularly high level at gene enhancer regions, exemplifying the idea that differential DNA methylation between tissues is due to the need for varying levels of gene activity dependent upon tissue function (107). Recently developed algorithms using whole genome bisulphite sequencing have supported these findings, revealing unique methylation profiles within nine distinct cell types across blood, brain, and breast tissues (108).

1.2.4 Whole blood tissue vs RPE in investigating AMD epigenetics

As a primary site of disease pathogenesis, it is understandable and expected that many patient studies of AMD involve the use of RPE tissue to maximise physiological relevance to the disease. With regard to epigenetic studies in particular, the selection of tissue type for investigation is of prime importance, owing to the tissue specificity of DNA methylation, and therefore conclusions drawn from epigenetic studies in one tissue cannot necessarily be readily applied to another (109).

While AMD studies in RPE are indeed physiologically relevant, the use of RPE tissue also presents a series of complications. RPE tissue samples for DNA methylation and expression studies must be collected post-mortem, with critical importance placed on minimising the interval time between patient death and tissue preservation, due to the potential for alterations in gene expression resulting from long post-mortem intervals. Optimal post-mortem times of three to six hours for ensuring mRNA stability have been recently reported, however longer post-mortem times can be deemed acceptable for genetic expression studies (110). Specifically, a post-mortem index of 72 hours has been proposed as a maximum threshold for maintenance of sample quality for DNA methylation studies, with correlations observed between longer post-mortem periods and increased variance in DNA methylation facilitated by a greater degree of measurable DNA degradation (111). The need for RPE isolation from deceased patients can put strain on the sample numbers available for study, potentially drastically limiting the power and validity of conclusions drawn from such studies, especially in epigenetics studies where clinically significant DNA differences can be relatively small in magnitude, requiring larger sample sizes to validate these differences (112).

A proposed alternative tissue for study of AMD epigenetics is whole blood, due to relatively non-invasive methods for collection when compared to RPE, as well as a dramatically increased available sample size to include living AMD patients. This would allow blood samples to be taken from patients at many different stages of disease progression, including early stage i.e. immediately upon early AMD diagnosis, as a means of obtaining large sample numbers of whole blood tissue to study epigenetic changes as a result of AMD progression from early to late stage. Blood samples have already been utilised as surrogate tissues in the assessment of the epigenetics of a number of diseases including schizophrenia, Crohn's disease and psoriasis, with whole blood studies revealing differential DNA methylation in these diseases without the potential difficulties in obtaining tissues from the primary sites of pathogenesis in each case (113-115).

1.2.5 Comparing the methylomes of blood and RPE

As previously stated, epigenetic profiles can vary between tissue type. A recent study by Hewitt *et al.* (2016) described the comparison of the DNA methylome between ocular tissue (RPE / choroid, optic nerve and retina) and matched whole blood for eight individuals (all male, tissues obtained within 12 hours post-mortem) using an Illumina 450K DNA methylation array (116). Analysis of this study revealed a general trend of similarity between the methylomes of ocular tissues and whole blood, with >255,000 CpG loci found to exhibit similar methylation levels between all tissues, and only approximately an additional 16,000 loci found to have comparable methylation between only the ocular tissues. The study by Oliver *et al* (2015) which revealed differential methylation of the *ARMS2* gene in AMD patient blood samples also noted genome-wide concordance between the methylomes of blood and retinal samples as obtained through use of a 450K methylation array. Similarly, significantly differential methylation was observed in both blood and primary tumour tissue of patients with retinoblastoma, with almost 1000 CpG loci determined to be retinoblastoma-specific present in both the blood and tumour tissue cohorts when compared to healthy control groups per tissue (117). With whole blood samples having been successfully used in studies to determine significantly differential methylation in disease and with demonstrable overlap between the methylomes of blood and ocular tissues at many loci, analysis of DNA methylation in blood may be of great benefit in identifying suitable biomarkers for AMD.

1.2.6 DNA methylation as a target for disease therapies

The abundance of associations between alterations to DNA methylation and disease make DNA methylation an attractive prospect for therapeutic intervention in the development of several diseases particularly as a small number of key molecules, delivered to the correct targets, could potentially be utilised in a wide number of different pathologies displaying specific pathogenic DNA methylation modifications. In many cancers for example, increased promoter DNA methylation gives rise to transcriptional silencing of tumour-suppressor genes, a known hallmark of carcinogenesis (118). Due to the potential for reversibility of epigenetic

modifications, in contrast to direct modifications to the genetic code, targeting the mechanisms through which DNA methylation occurs is a viable and attractive therapeutic strategy for future consideration.

DNMT enzymes, the mediators of global DNA methylation, were the targets of early anti-cancer therapeutic strategies utilising nucleoside analogues. These analogues, including 5-azacytidine (Vidaza), worked via inhibition of DNMT enzymes, reducing the level of DNA binding to DNMT1, and facilitating DNMT1 proteasomal degradation. This resulted in an increase in "passive" gradual demethylation – inactivation or absence of DNMT1 results in decreased capability of methylation of DNA following cell division, which is exacerbated with each subsequent replication cycle (119). 5-azacytidine is still used as a first-line therapy in the treatment of myelodysplastic syndrome (MDS), however administration of high doses of 5-azacytidine leads to cytotoxicity and DNA damage, with safe therapeutic hypomethylation activity most evident at lower drug doses (120). Given the cytotoxic potential of 5-azacytidine, and their lack of specificity for a particular methyltransferase enzyme, further studies into the use of DNMT inhibitors have been undertaken to find more suitable therapeutics for various cancers. More recently, the drug Hydralazine (Apresoline), while primarily used as an antihypertensive, has been shown to selectively inhibit DNMT1 activity, with Hydralazine treatment seen to significantly reduce tumour volume in mice in breast cancer studies (121, 122).

While DNMT activity can be evidently altered via analogues of cytosine, the other fundamental component of DNMT activity, the universal methyl donor SAM, has also been investigated in terms of therapeutic strategies targeting DNA methylation. Recent medicinal chemistry studies investigating the inhibitory effect of DNMT activity by analogues of SAM have determined that replacement of the adenine group with a quinazoline group, when covalently attached to a quinoline group, was found to inhibit both DNMT1 and DNMT3a activity with low concentrations of 4µM and 46µM respectively (123). As analogues for both substrates necessary for DNA methylation have proven to be viable drug targets,

this reinforces the idea that epigenetics, and specifically DNA methylation, can be an attractive therapeutic target for disease treatment.

1.2.7 DNA methylation in AMD – evidence from the literature

While the genetic basis of AMD is well understood, environmental factors are also associated with AMD development, though the specific mechanisms by which these factors interact to cause AMD is not clear. As epigenetics is regarded as a mechanism by which the environment may influence genetic expression, DNA methylation changes may, at least in part, play a role in mediating the two major AMD risk factors, smoking and diet poor in vitamins and folate, influencing AMD development.

Based on this assumption, several studies have been undertaken to elucidate the possible role of DNA methylation in AMD development. Methylation array and bisulphite pyrosequencing studies have been utilised, revealing significant reductions in mRNA levels of glutathione s-transferase (GST) enzymes *GSTM1* and *GSTM5* in AMD RPE / choroid, with associated hypermethylation of the *GSTM1* promoter region, suggesting DNA methylation-mediated transcriptional silencing (124). These GST enzymes are of major importance in cellular detoxification, including in the quenching of ROS and therefore the mitigation of oxidative stress-induced cell damage, a proposed mechanism of retinal damage and AMD development (125).

A Genome-wide methylation analysis in AMD blood samples by Oliver *et al* (2015) revealed no sites of DNA methylation change achieving genome-wide significance, however upon further interrogation of the data with regards to known AMD risk loci, a small but significant decrease in promoter methylation of the age-related maculopathy susceptibility gene 2 (*ARMS2*) was identified in AMD blood samples, particularly in those samples displaying CNV (126). While the raw levels of methylation change between control and AMD patients at this locus were small, the identification of significantly differential DNA methylation at a gene previously associated with AMD suggests a potential mechanistic link between abnormal DNA methylation and AMD progression through gene expression change.

In a separate study, analysis of peripheral blood mononuclear cell (PBMC) DNA of monozygotic and dizygotic twins, with one AMD twin per pair, identified hypomethylation of the promoter region of interleukin 17 receptor C (*IL17RC*) in the AMD twin, a finding which was subsequently replicated with 7 pairs of non-twin siblings, with each pair consisting of one sibling with large drusen and one sibling with none (127). Interestingly, *IL17RC* binds interleukin 17A (IL17A), which has been identified as being cytotoxic to ARPE-19 cells, an RPE cell line, through activation of the pro-apoptotic caspase-3 and caspase-9 enzymes (128). As AMD progression is generally categorised by, among other factors, RPE loss, this finding is particularly noteworthy. Further to this, PBMC DNA from separate patients, including 96 non-AMD controls, 95 CNV and 107 GA patients, was assessed to validate the above findings. Again, significant hypomethylation of the *IL17RC* promoter was observed for AMD patients compared to non-AMD controls, while no significantly differential methylation was observed between the CNV and GA cohorts. This finding was supplemented by qPCR studies which revealed a significant increase in *IL17RC* expression in both peripheral blood as well as the macula of AMD patients compared to non-AMD controls (127). The observations of significant *IL17RC* promoter demethylation compounded with significant upregulation of *IL17RC* gene expression in multiple tissues suggest a potential mechanism of promoter DNA demethylation-mediated transcriptional un-silencing of *IL17RC*, again underpinning the potential role of DNA methylation in the development of AMD.

However, in response to these findings, a separate study investigating the claims that *IL17RC* promoter methylation may be useful as a clinical biomarker of AMD in blood was undertaken which presented contrasting results. Interrogation of multiple cohorts of AMD and non-AMD control peripheral whole blood samples resulted in no significant methylation difference between the groups at the DNA region previously reported by Wei *et al*. Further validation was attempted through bisulphite pyrosequencing of two additional distinct cohorts. While similar levels of methylation were observed between the two cohorts, the results of one of these two validation cohorts in fact displayed a small, yet significant ($p=<0.001$) increase in methylation in the AMD group compared to control. A final validation cohort of

retinal and RPE donor samples was subjected to bisulphite pyrosequencing with no significant methylation change in the *IL17RC* promoter observed between AMD and non-AMD samples (129), however, despite the comparable methylation levels to the whole blood samples used in previous cohorts, the low sample numbers used (retina: 9 AMD, 6 control; RPE: 3 AMD, 3 control) suggests that further validation may be necessary to conclude that the altered gene expression of *IL17RC* reported by Wei *et al* (2012) in ocular tissue should not be considered as a potential biomarker of AMD.

These studies, while outlining a number of candidate genes and loci in which altered DNA methylation can be potentially associated with AMD, suffer from the limitations previously outlined – namely, the lack of genome-wide significance in loci observed; the small numbers of ocular tissue samples of relevance to AMD, such as RPE; and a general reliance on whole blood DNA samples to draw conclusions regarding the effects of DNA methylation on the onset and development of a disease of the central retina without complementary data from RPE / retina samples. The results of these studies are summarised in Table 1.2.

Table 1.2: Summary of literature exploring DNA methylation changes of genes in AMD

Study	Gene	Observation in AMD	Tissue
Hunter *et al*. (2012)	*GSTM1*	Hypermethylation of promoter	RPE/Choroid
Oliver *et al*. (2015)	*ARMS2*	Hypomethylation of promoter	Blood
Wei *et al.* (2012)	*IL17RC*	Hypomethylation of promoter	Blood
Oliver *et al.* (2013)	*IL17RC*	No methylation change	Blood

1.3 Differential methylation of AMD RPE observed by Porter *et al*

A recent study by Porter et al identified differential DNA methylation at a number of novel genetic loci via genome-wide DNA methylation profiling in human donor RPE cells in AMD (130). DNA methylation levels were measured using an Illumina Infinium HumanMethylation450 BeadChip Array, allowing for a number of differentially methylated loci (DML) and differentially methylated region (DMR) sites to be identified (Tables 1.3, 1.4), the most significant of which were further

validated using bisulphite pyrosequencing. The study found significant differential methylation and gene expression between AMD and control patient RPE, in both the microarray and the bisulphite pyrosequencing validation, at regions within the genes *GTF2H4*, *SKI* and *TNXB*. As promising candidates for potential therapeutic targeting in AMD, these genes warrant further investigation.

Table 1.3: Identification of DML in AMD RPE by Porter *et al*. (2019).

Gene	Identifier	P value	Difference in β-value	Chr	Position	Probe location
FAIM2	cg18486102	1.08E-12	0.20	12	50297777	TSS 200
SKI	cg18934822	1.18E-09	0.11	1	2191402	Body
	cg26139512	2.08E-08	0.33	15	60290666	Intragenic
GTF2H4	cg22508626	7.03E-07	0.11	6	30879905	Body
RIC3	cg01560972	2.03E-06	0.33	11	8190837	TSS 1500
EDEM2	cg04838987	3.98E-06	0.18	20	33734406	Body
BDNF	cg11241206	4.44E-06	0.30	11	27723128	TSS 1500
EIF2AK3	cg26347887	5.16E-06	0.28	2	88927196	TSS 200

Differential methylation and expression of *SKI* (cg18934822) and *GTF2H4* (cg22508626) was subsequently validated by Porter *et al.* via bisulphite pyrosequencing and RT-qPCR respectively.

Table 1.4: Identification of DMR in AMD RPE by Porter *et al*. (2019).

Gene	CpGs	Chr	Coordinates	Min FDR	DMR length (bp)	Location
FAIM2	7	12	50297477-50297945	8.39E-16	469	Promoter
SKI	6	1	2190850-2191658	2.92E-08	809	Intron 1
	3	17	14201680-14201938	3.60E-07	259	Intragenic
TNXB	15	6	32063835-32064258	6.30E-06	424	Exon 3

Differential methylation and expression of *TNXB* was subsequently validated by Porter *et al.* via bisulphite pyrosequencing and RT-qPCR respectively.

1.3.1 General Transcription Factor IIH Subunit 4

General Transcription Factor IIH Subunit 4 (*GTF2H4*) is a core component of the TFIIH complex, involved in initiation of transcription, transition from initiation to elongation, and transcription-coupled nucleotide excision repair (TC-NER) of DNA (131). TC-NER acts on transcriptionally active strands of transcribed genes, at a higher rate than the non-transcribed strands and transcriptionally silent DNA, initiated by stalled RNA polymerase at a site of DNA damage (132). In TC-NER, stalled RNA polymerase serves as the DNA damage signal, differing from the mechanism of global genomic NER (GG-NER) which repairs transcribed and non-transcribed strands of active and inactive genes throughout the genome, is not dependent upon transcription, and is facilitated by damage sensing proteins such as DNA damage binding (DDB) and XPC-Rad23B complexes, which constantly scan the genome and recognise helix distortions for repair (133).

The TFIIH complex is comprised of seven subunits: *ERCC2* (XPD), *ERCC3* (XPB), *GTF2H1* (p62), *GTF2H2* (p44), *GTF2H3* (p34), *GTF2H4* (p52) and *GTF2H5* (p8). Mutations in these subunits have previously been linked to a number of diseases, principally trichothiodystrophy, Cockayne syndrome, and xeroderma pigmentosum (134-136). Single nucleotide polymorphisms (SNPs) associated with *GTF2H4* in particular have been linked with lung cancer, multiple sclerosis, and HPV persistence, as a potential precursor to cervical cancer (137-140). As differential expression of *GTF2H4* has been observed previously in AMD RPE (141), the significantly increased differential methylation and associated decreased gene expression of *GTF2H4* in AMD RPE observed by Porter et al suggests a potential role of TC-NER pathways in AMD development, with transcriptional silencing of *GTF2H4*, mediated by epigenetic mechanisms, potentially contributing to impaired DNA repair and subsequent loss of RPE (142).

1.3.2 SKI Proto-oncogene

SKI, coding for the SKI proto-oncogene protein, acts as a negative regulator of transforming growth factor beta (TGF-β) signalling. Of particular note, TGF-β signalling has been reported to be involved in RPE cell migration, oxidative stress-

induced RPE senescence, and modulation of the complement pathway. These mechanisms give insight into the potential functional role of *SKI* in AMD development, with reduction of TGF-β signalling seen as a potential preventative strategy in early AMD (143-146).

SKI associates with SMAD proteins SMAD2/3 to counteract the activation of gene expression mediated by TGF-β. The *SKI* interacting domains of SMAD2/3 reside within the mad homology 2 (MH2) regions, which are highly conserved across SMAD proteins and contain the SMAD2/3 transcriptional activation domains (147). This suggests the potential role of *SKI* in blocking TGF-β activity by binding to activation domains and blocking co-activators and co-binding transcription factors or basal transcription machinery (148).

Increased expression of TGF-β has been previously observed in mice subjected to laser-induced CNV compared to control groups during CNV development (149). Administration of TGF-β inhibitors resulted in a reduction of CNV lesion size compared to groups treated with a PBS control. TGF-β facilitates the phosphorylation of SMAD proteins and the formation of a SMAD complex, which is found to translocate into the nucleus to act as a transcription factor through DNA binding at the "SMAD binding element" (SBE) within target gene promoters to modulate gene transcription via recruitment of chromatin activators or repressors (150). The role of *SKI* as a negative regulator of TGF-β signalling leads to the hypothesis that differential *SKI* methylation and reduction in mRNA expression observed in AMD RPE by Porter *et al* may lead to increased TGF-β signalling, increased SMAD phosphorylation and therefore increased risk of AMD development.

1.3.3 Tenascin XB

Tenascin-X, coded for by the Tenascin XB (*TNXB*) gene, is involved in extracellular matrix maturation and wound healing, with mutations in *TNXB* being associated with the connective tissue disorder Ehlers-Danlos syndrome (151). In AMD, drusen deposits occur in Bruch's membrane, the ECM of the RPE, as a hallmark of early AMD. *TNXB* has been previously associated with AMD, with SNPs within *TNXB*

suggested to convey increased risk of neovascular AMD (152, 153). In this context, the discovery of differential methylation and expression of *TNXB* in AMD RPE is of particular interest.

1.4 Modifiable AMD risk factors and DNA methylation

1.4.1 Effects of smoking on the AMD epigenome

In addition to being a known risk factor for development of AMD, cigarette smoking has been shown to have a direct effect on DNA methylation. In a study by Koks *et al* (2015), the *GPR15* gene displayed significantly higher expression levels (log fold change 2.61) in both male and female smokers than in non-smokers, with DNA methylation analysis confirming a statistically significant hypomethylation of the *GPR15* gene in smokers (154). Similarly, Fragou *et al* (2019) demonstrated the association between smoking status and the methylation profile of patients with Barrett's oesophagus and oesophageal cancer and highlighted *TNXB*, among other genes as being significantly differentially methylated in smokers when compared to non-smokers (155). The changes in both methylation and expression of genes in smokers compared to non-smokers demonstrates the ability for the environment to have a direct influence over the methylome in disease pathogenesis.

In a large-scale study of 15,907 blood samples of current, former and never smokers, Joehanes *et al* (2016) described the impact of smoking on the human methylome, identifying 2623 CpG sites that were significantly differentially methylated between current and never smokers (156). The results of this study highlighted cg05603985, a CpG site within the *SKI* gene, as being one of the most significantly differentially methylated between current and never smokers. Porter *et al* also identified a genomic region within *SKI* that was both significantly differentially methylated and expressed in AMD RPE compared to control tissue.

1.4.2 The folate pathway and AMD

Folate (vitamin B9) is involved in a number of critical cellular processes including the synthesis and methylation of DNA. Classed as an essential nutrient, folate can

be incorporated into the diet through consumption of folate-rich foods, such as leafy green vegetables, nuts and seeds, and seafood, or through supplements of folic acid, a manufactured compound that is converted into folate in the body (157). Largely due to its role in DNA synthesis, folate demands increase during periods of growth, such as during pregnancy and adolescence (158).

The folate pathway, and closely associated methionine-homocysteine cycle, is a complex pathway of processes involving a multitude of enzymes and cofactors to facilitate the synthesis and methylation of DNA. The folate biosynthesis pathway, through the process known as one-carbon metabolism, involves dietary folate and folic acid first undergoing reduction to dihydrofolate (DHF), and subsequent reduction to tetrahydrofolate (THF) via dihydrofolate reductase (*DHFR*). One-carbon methyl addition to THF via serine hydroxymethyltransferase (*SHMT*) enzymes produces 5,10-methylene-THF, which in turn is converted to 5-methyltetrahydrofolate via methylene tetrahydrofolate reductase (*MTHFR*). Alternatively, in the presence of deoxyuridine monophosphate (dUMP), thymidylate synthase (*TS*) mediated conversion of 5,10-methylene-THF to deoxythymidine monophosphate (dTMP) can occur, also generating DHF as a by-product which can reinsert back into the folate pathway to mediate further THF synthesis (159). The generation of dTMP through the folate pathway underpins the importance of dietary folate in the regulation of DNA and RNA synthesis.

It is at this point that the folate pathway and methionine-homocysteine cycle become intertwined – the enzyme methionine synthase (MS), encoded by the *MTR* gene, catalyses the synthesis of methionine from the amino acid homocysteine, which itself is not utilised in protein synthesis. This conversion takes place in the presence of the cofactor cobalamin (vitamin B12), and utilises 5-methyl THF as a methyl donor for the reaction, regenerating THF as a result and restarting the folate pathway (160).

Following methionine synthesis, addition of ATP results in the conversion of methionine to SAM, the universal methyl donor, mediated by methionine adenosyltransferase (MAT) enzymes, encoded by the *MAT2A* gene. Use of SAM as a substrate by DNMT enzymes produces s-adenosyl homocysteine (SAH) as a result,

which in turn is converted to the amino acid homocysteine via adenosyl homocysteinase (*AHCY*), enabling homocysteine to enter back into the methionine synthesis pathway as above or be used as a substrate for cystathione synthesis, along with serine, mediated by the enzyme cystathione β-synthase (*CBS*) and the cofactor pyridoxine (vitamin B6). In this cystathione pathway, cystathione γ-lyase (*CTH*) catalyses the breakdown of cystathione into cysteine, revealing a role of the methionine and folate pathways in protein synthesis (Figure 1.5) (161, 162).

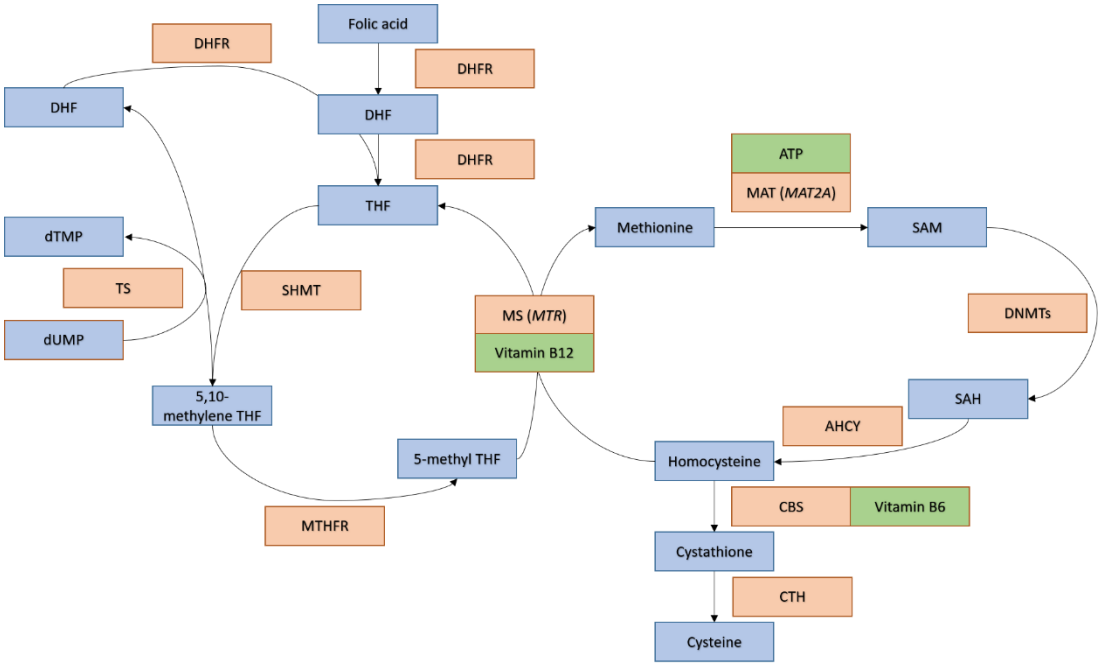

Figure 1.5: Overview of the overlapping folate and methionine biosynthesis pathways. Red boxes represent enzymes, green boxes represent cofactors.

1.4.3 The folate pathway in DNA methylation

As the folate / methionine pathways are responsible for the regulation of a number of cellular processes and involve a great many substrates and cofactors in balance, there are many factors to consider when assessing the potential associations between dietary folate, DNA methylation, and AMD.

The folate / methionine pathway facilitates, among many other reactions, the biosynthesis of SAM. As SAM is involved in DNA methylation as a substrate for DNMT enzymes, the importance of the interplay between the folate and methionine cycles in the regulatory mechanisms of DNA methylation is clear. A study investigating the association of dietary folic acid intake and global DNA

methylation found that folic acid supplementation of 400µg/day resulted in an increase in serum folate concentration of 81% compared to a placebo group, with global DNA methylation increasing by 31% in supplemented patient leukocytes (163). This was measured by a methylation acceptance assay, where genomic DNA is subjected to enzymatic addition of radiolabelled methyl groups via SssI methylase, methylating all currently unmethylated cytosine residues – as methyl acceptance will be inversely proportional to the pre-existing level of methylation, global DNA methylation can thus be measured (164). Further studies have demonstrated the effects of this same dose of folic acid supplementation on DNA methylation levels in late-stage pregnancies, observing increased methylation of a DMR upstream of *ZFP57*, the master regulatory gene of genomic imprinting, compared to a placebo control group. This same study additionally detailed that altered methylation at this particular DMR had a direct effect on transcriptional activity of *ZFP57*, suggesting a relationship between dietary folate intake, DNA methylation, and changes in gene expression (165).

However, high dietary folate intake has been suggested to be somewhat of a double-edged sword, with mice studies reporting that a 10-fold higher supplementation of folic acid in the treated versus control group resulted in reduced 5-methyltetrahydrofolate levels, as well as reduced *MTHFR* expression and activity. It has been reported that mutations in the *MTHFR* gene may be responsible for reductions in global DNA methylation, suggesting a link between the folate-mediated expression of this particular gene and the capability for DNA methylation (166). It has been suggested that high levels of folic acid supplementation may affect the ratio between SAM and SAH in the folate cycle causing feedback inhibition, which may explain the counterintuitive findings of decreased global DNA methylation with particularly high serum levels of folate (167).

As the relationship between the folate and methionine cycles has been well established, it is understandable that a reduction in dietary folate intake can theoretically be responsible for the alteration of a multitude of downstream factors – principally among these is the downregulation of THF synthesis, and therefore the resulting downregulation of THF mediated synthesis of methionine from

homocysteine and potential increase in intracellular levels of homocysteine. Owing to the reversibility of the reaction for the formation of homocysteine from SAH, a high serum concentration of homocysteine observed in the condition hyperhomocysteinemia was found to be responsible for similarly increased levels of SAH. This disruption of the SAM to SAH ratio, a key balance required for DNMT-mediated DNA methylation, is proposed to lead to a decrease in available methyl donation from SAM and as a result lead to global DNA hypomethylation (168). These claims are supported in an animal study by Jiang *et al* (2007), who report that high methionine diet fed rats were found, through induction of hyperhomocysteinemia, to have significantly decreased methylation of B1 repetitive elements (169). Again, the close association of homocysteine with the folate pathway, and the ability of increased dietary intake of folate to reduce serum homocysteine levels through conversion to either methionine or cystathione, underpins the importance of dietary folate in regulation of global DNA methylation.

1.4.4 Association of homocysteine and AMD

The association between homocysteine and AMD is less well defined (170), with a proposed association between hyperhomocysteinemia and induction of oxidative stress through interactions with NADPH oxidase and generation of ROS, which could facilitate the development of AMD through the proposed mechanism of ROS-mediated retinal damage (171, 172). A number of studies have reported elevated levels of serum homocysteine in AMD patients (173, 174), particularly in the neovascular form, suggesting hyperhomocysteinemia may play a role in the development of CNV in more advanced cases of the disease (175-181). However, a number of contrasting studies observed no significant elevation of serum homocysteine in AMD patients compared to non-AMD controls (182, 183).

Nonetheless, given the role of homocysteine in the complex cycles involved in facilitating methyl donors for DNA methylation, its potential reported associations with AMD development, and the observed regulatory effect of dietary folate intake (itself a recognised protective factor in AMD) on homocysteine levels, there exists a potential link between serum homocysteine levels and epigenetic mechanisms in the development of AMD. In this regard, investigation of DNA methylation levels of

the genes involved in the regulation of homocysteine in AMD may provide insight into the association between homocysteine and AMD. Genes involved include *AHCY* (coding for adenosyl homocysteinease, responsible for the formation of homocysteine from SAH), *MTR* (coding for methionine synthase, responsible for the conversion of homocysteine to methionine) and *MAT2A* (coding for methionine adenosyltransferase, responsible for the conversion of methionine to the universal methyl donor SAM).

1.5 Project aims and objectives

Smoking is a known risk factor for AMD with the potential to influence DNA methylation and therefore gene expression. However, the effects of smoking status on the methylation levels of differentially methylated genes in AMD has not been assessed. After validation of overlapping baseline methylation levels of the differentially methylated loci and regions within the genes *SKI*, *GTF2H4*, *TNXB* identified by Porter *et al*, levels of DNA methylation of AMD patient whole blood in smokers and non-smokers will be investigated to evaluate the impact of smoking on the epigenetics of AMD.

The folate pathway is of great interest in AMD, given the association between folic acid supplementation and reduced risk of AMD. Furthermore, high serum levels of homocysteine, a key component in the regulation of DNA methylation, are reported to play a potential role in the development of AMD, making the genes associated with the folate and methionine-homocysteine biosynthesis pathways attractive targets for further study with respect to a role for DNA methylation changes in AMD.

In brief, the major aims of the project are:

1) To validate the use of whole blood DNA as a proxy to RPE-derived DNA to identify differential methylation in the select differentially methylated genetic loci identified by Porter *et al.* within the genes *SKI*, *GTF2H4*, and *TNXB*.

2) To evaluate the impact of smoking, as a known major risk factor of AMD, on levels of DNA methylation at genetic loci of interest within *SKI*, *GTF2H4*, *TNXB* in AMD
3) To evaluate DNA methylation levels of genes associated with the folate and methionine-homocysteine biosynthesis cycles in blood-derived DNA of patients with AMD versus controls.

1.6 Introduction to methods used

1.6.1 Utilising bisulphite pyrosequencing to observe methylation changes

One of the primary experimental methods of observation of DNA methylation levels and changes is through bisulphite pyrosequencing of DNA fragments. Briefly, a fragment of target DNA to be analysed is outlined, and CpG sites within the fragment are identified. The target DNA sample is then subjected to sodium bisulphite conversion, a process where unmethylated cytosine residues are selectively converted to uracil by deamination, while 5'-methylcytosine residues resist the chemical change and remain unmodified. Complete bisulphite conversion of study DNA is essential for reliable DNA methylation results (184). Once the DNA sample has undergone complete bisulphite conversion, a pair of forward and reverse primers specific to the gene / region of interest are designed, with one amplicon biotin labelled, and polymerase chain reaction (PCR) amplification is performed on the sample. Bisulphite conversion results in unchanged 5'-methylcytosine residues being replicated in the PCR product, while uracil residues are converted to thymine residues in the PCR product (Figure 1.6), meaning that there is subsequently a distinct, measurable difference in genetic code between methylated and unmethylated DNA samples (185).

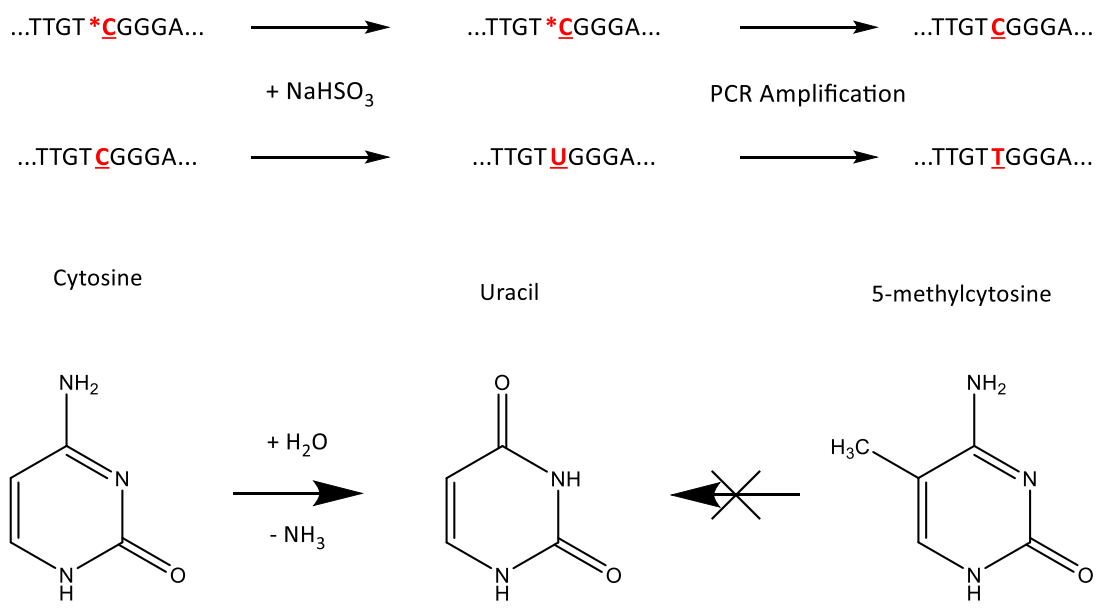

Figure 1.6: Overview of sodium bisulphite conversion workflow. *C represents 5-methylcytosine

The pyrosequencing reaction itself allows methylated cytosine residues in a given sample to be tracked and ultimately visualised. The method follows the principle of "sequencing by synthesis" due to the use of a DNA polymerase to facilitate nucleotide incorporation. Briefly, the bisulphite converted DNA is mixed with streptavidin beads where binding of the beads to the DNA sample occurs due to the biotinylated nature of the PCR product. Using a vacuum pump, the biotinylated product is isolated and subsequently denatured using sodium hydroxide (NaOH), with only the biotinylated single-stranded DNA (ssDNA) product remaining. A previously designed sequencing primer, specific to the sequence of this biotinylated ssDNA product, anneals to the sample and allows sequencing to begin on a nucleotide-by-nucleotide basis within the pyrosequencer (186).

The reagent mixture incorporated in pyrosequencing reactions, and its associated chemistry facilitate the visualisation of DNA methylation, with the mixture including a number of enzymes (DNA polymerase, ATP sulfurylase, luciferase and apyrase) and substrates (adenosine 5' phosphosulfate and luciferin). Following the dispensation and successful addition of the correct nucleotide to the ssDNA sample

via DNA polymerase (via a programmed sequence in the pyrosequencing software), inorganic pyrophosphate (PPi) is released, which reacts with adenosine 5' phosphosulfate (APS) to produce adenosine triphosphate (ATP), catalysed by ATP sulfurylase. This ATP in turn reacts with luciferin to produce oxyluciferin, and therefore visible light, catalysed by luciferase, as well as the by-products PPi and adenosine monophosphate (AMP). The enzyme apyrase degrades any free remaining nucleotides as well as ATP, to ensure that no reactive elements remain, whereby the next programmed dispensation begins, and the cycle is repeated until full synthesis of the target template has been achieved, with the pyrosequencer able to detect any visible light from oxyluciferin synthesis (187).

During the pyrosequencing run, the presence of a CpG site will prompt the instrument into addressing the two possible scenarios – one where 5'-methylcytosine is present and has been preserved in PCR amplification, and one where the unmethylated cytosine has been converted to uracil through bisulphite treatment and subsequently to thymine upon PCR amplification. At each given CpG site in the DNA sample, the relevant nucleotides for each situation will be dispensed consecutively, with the luciferin levels of each measured to give a ratio of the level of methylated / unmethylated cytosine residues in the original sample at the given CpG site, and therefore the degree of DNA methylation present.

The ability to quantitively determine the proportion of methylation for each cytosine residue present in a given sample makes pyrosequencing an invaluable tool in observing DNA methylation, for example in measuring epigenetic differences in particular genes between varying populations, in measuring epigenetic changes as a result of drug treatment, or in detecting epigenetic biomarkers of disease. However, a major limitation to the use of bisulphite pyrosequencing in the observation of DNA methylation changes is the size of template DNA required – pyrosequencing is primarily used for shorter target sequences of a maximum of approximately 200bp (188). While this approach can be advantageous e.g. for fast DNA methylation screening of large numbers of isolated small target sequences, or examining a small number of previously identified sequences of interest in a larger

number of patients, genome-wide DNA methylation studies such as epigenome-wide association studies (EWAS) require an alternative approach.

1.6.2 DNA methylation arrays to study DNA methylation

In contrast to bisulphite pyrosequencing, DNA methylation arrays, such as the Illumina methylation microarrays, allow for interrogation of vast numbers of CpG sites across the entire genome (189). This approach can be advantageous for example when attempting to identify candidate CpG sites or genes of interest for further study which display significantly differential methylation between distinct populations.

Briefly, genomic DNA is subjected to bisulphite conversion as previously described and purification steps utilised to provide purified DNA fragments, which are subsequently applied to the methylation array chip. Each chip, which can accommodate 12 distinct DNA samples, contains beads attached to probes specific to CpG loci (the number of which are dependent on the technology used), with separate probes for each locus corresponding to methylated and unmethylated cytosine residues, the unmethylated residues having been converted to thymine through bisulphite conversion and subsequent amplification. Again, in a similar process to that used for bisulphite pyrosequencing, DNA fragments are denatured to ssDNA and allowed to anneal to the corresponding probe dependent upon methylation status. Single-base extension takes place along each probe, with fluorescent reagent staining allowing for the proportion of DNA methylation at each probe to be observed (190).

The use of both bisulphite pyrosequencing and methylation microarray technology is vital in characterising DNA methylation levels at specific loci and genomic regions, and represent ideal technologies in combination for case-control studies monitoring of DNA methylation changes in a disease context versus a control.

Chapter 2: Materials and Methods

2.1 Sample collection, grading and DNA extraction

Peripheral whole blood samples and gDNA was extracted from individuals within the Netherlands and Southampton Case-Control Cohort (referred to as the "Southampton Cohort") using the salting-out procedure (191). Demographic information is provided including age (years), gender (male or female) and smoking status (Smoker or Non-Smoker) of both of these cohorts, as well as the combined cohort (Tables 2.1, 2.2, 2.3). Samples were obtained from individuals phenotyped according to the Age-Related Eye Disease Study (AREDS) classification. Samples from patients exhibiting advanced AMD (grade 4) were excluded from the study. From the Netherlands cohort, a total of 192 patient samples including 96 normal, 48 AREDS grade 2 (early AMD) and 48 AREDS grade 3 (intermediate AMD) (AMD total n=96) were selected. From the Southampton cohort, a total of 96 patient samples including 48 normal, 23 AREDS grade 2 and 25 AREDS grade 3 (AMD total n=48) were selected. More detailed information of individual sample information of each cohort is listed in the appendix (Tables 7.1, 7.2).

Table 2.1: Sample information overview for cohort 1 blood samples (Netherlands)

AREDS Grade	Samples (male)	No. of smokers	Age (mean ± SD)
0 CTRL	96 (47)	10 (10.4%)	73.08 ± 7.77
2 Early AMD	48 (21)	8 (16.7%)	72.23 ± 6.70
3 Intermediate AMD	48 (20)	4 (8.3%)	75.04 ± 7.98

Table 2.2: Sample information overview for cohort 2 blood samples (Southampton)

AREDS Grade	Samples (male)	No. of smokers	Age (mean ± SD)
0 CTRL	48 (24)	24 (50.0%)	77.54 ± 7.03
2 Early AMD	23 (8)	11 (47.8%)	73.78 ± 7.56
3 Intermediate AMD	25 (5)	13 (52.0%)	83.20 ± 7.08

Table 2.3: Sample information overview for combined cohort 1 and 2 blood samples

AREDS Grade	Samples (male)	No. of smokers	Age (mean ± SD)
0 CTRL	144 (71)	34 (23.6%)	74.57 ± 7.80
2 Early AMD	71 (29)	19 (26.8%)	72.73 ± 6.97
3 Intermediate AMD	73 (25)	17 (23.3%)	77.84 ± 8.97

2.2 Ethics statement

All procedures in this study adhered to the tenets of the Declaration of Helsinki. This project was approved by the University of Liverpool's institutional ethical review boards (University of Liverpool Central University Research Ethics Committee for Physical Interventions, ref. 2326, date of approval 31 January 2018) and the University of Pennsylvania Institutional Review Board (IRB). Guidelines established in the Human Tissue Act of 2008 (UK) were followed. Informed written consent was obtained from individuals who donated their blood samples for the purposes of this research.

2.3 Agarose gel electrophoresis

1% agarose gel electrophoresis was used to assess levels of DNA degradation prior to bisulphite pyrosequencing. A 1% gel was made by boiling 200ml of 0.5x TBE buffer (see appendix) and 2g of agarose (Fisher BioReagents, Pittsburgh, United States) in a conical flask by microwaving with stirring for 5x30 seconds, after which SafeView Nucleic Acid Stain (3ul, NBS Biologicals, Huntingdon, United Kingdom) was added and the gel poured into a cassette and allowed to cool.

For the purposes of standard PCR product assessment by agarose gel electrophoresis a 2% gel was produced in the manner outlined above using 4g of agarose in 200ml 0.5x TBE buffer.

5ul of PCR product, or 100 ng of DNA (for DNA degradation analysis), and a 100bp ladder (for PCR products, 5ul, NEB, Ipswich, United States) or a 1kb ladder (for DNA degradation analysis, 5ul, NEB, Ipswich, United States) were added to relevant wells, with the gel being run at 100V for 45 minutes. Images were taken using an

EC3 Bioimaging System (UVP, Upland, United States) using UVP imaging software. No-DNA negative controls were used in all cases.

2.4 PCR primer optimisation

Lyophilised primers were resuspended using TE buffer (10mM Tris, 0.1mM EDTA, pH 8.0) to give a final concentration of 100pmol/ul. Initial estimation of the annealing temperature for PCR was carried out based on the primer base sequences using the Tm calculator function of the GeneAmp PCR 9700 Thermal Cycler (Applied Biosystems, Bedford, United States), and 6 PCR reactions were performed in parallel at a range of temperatures up to 5°C away from the estimated annealing temperature using a Veriti 96-well Thermal Cycler (Applied Biosystems, Bedford, United States). PCR Optimisation reactions were carried out on healthy control PBMC DNA. Primers were obtained from Eurofins with the following sequences and experimentally optimised annealing temperatures.

Table 2.4: Sequence of primers used in PCR and bisulphite pyrosequencing reactions

Gene	Primer	Sequence (5' → 3')	Biotin Labelled?
SKI	FWD	TGT GGA GTG TTT TTA GTT TTT GAT	
	REV	TTC CAC ACC TAA ATA TCT ATC CAA	YES
	SEQ	TTT GAT TGT TAT AAA GAA AG	
GTF2H4	FWD	GGG TAT GTG GTA TTA GTT AGG TT	YES
	REV	CTC TTC CAA AAA CCC TCT A	
	SEQ	TTC CAA AAA CCC TCT ATC C	
TNXB	FWD	GGT TAT TTA GGG GAA GAT TGT AGT A	
	REV	AAC CTT CCC AAC AAA TAC ACA	YES
	SEQ	GGT TAG TGT TTA GAT GGG	
AHCY	FWD	GTT GTA TAA GGT TAG ATT AAG GTG	YES
	REV	CCC TCT CAT TAA TTA ATA ATA ACT C	
	SEQ	AAT AAC TCT CAT CCT CAC TAT A	
MTR	FWD	TGA TAG GAT GGT AGT AAT TAG AAT	
	REV	CCC CAT ATA ATA ATT CAA AAT	YES
	SEQ	TTT TTA GGG TTG TAG TAG G	
MAT2A	FWD	GAG TTG TTA TTG TTT GGG TT	YES
	REV	CCC TAA AAA TAT AAC TTT ACT CCT C	
	SEQ	AAA TAT AAC TTT ACT CCT CC	

"Seq" denotes sequencing primer structure used in bisulphite pyrosequencing reactions only. For each set of primers, either the forward or reverse primer was biotin labelled as denoted.

2.5 Polymerase chain reaction

For PCR studies, PyroMark (Qiagen, Hilden, Germany) reagents were used. 3ul of DNA was added to PyroMark master mix (2x, 15µl), CoralLoad dye (10x, 3ul), forward and reverse primer mix (1µl) and water (8µl) and amplified using a Veriti 96-well Thermal Cycler (Applied Biosystems, Bedford, United States). Thermocycling conditions were as follows:

Table 2.5: Polymerase chain reaction thermocycling conditions

Stage	Temperature	Time	Cycles
Enzyme activation	95°C	10min	1
Denaturing	94°C	30s	40
Annealing temp	X°C (See **Table 2.3**)	30s	
Extension	72°C	30s	
Final extension	72°C	10min	1

Thermocycling conditions were designed to provide sufficient DNA amplification with PCR reactions being evaluated via agarose gel electrophoresis.

Table 2.6: Target gene PCR annealing temperatures

Gene	Annealing Temp (°C)
SKI	53
GTF2H4	51
TNXB	58
AHCY	54
MTR	51
MAT2A	53

Optimal annealing temperatures were evaluated using parallel PCR reactions ranging up to 5°C away from the estimated annealing temperature as calculated by the manufacturer in each case.

2.6 Bisulphite DNA conversion

Concentration of blood DNA samples was assessed via nanodrop. Subsequently, 1ug DNA per sample was aliquoted into a separate 96-well plate, before being dried using a miVac DNA concentrator (GeneVac, Ipswich, United Kingdom) and resuspended in ddH_2O (20ul). Bisulphite DNA conversion was carried out using the EZ-96 DNA Methylation Gold kit (Zymo, Irvine, United States) as per the manufacturer's instructions. Briefly, CT conversion reagent (130ul) was added to each sample before incubation at 64°C for 2.5 hours. Following this, the samples were washed with the provided wash buffer (200ul) utilising a vacuum manifold before addition of desulphonation buffer (200ul) to facilitate conversion of

unmethylated cytosine to uracil. Samples were then left to incubate at room temperature for 20 mins. Samples were subsequently further washed with wash buffer (2x 200ul) before drying completely using a centrifuge (3000 x g, 5 mins). Following this, elution buffer (50ul, Zymo, Irvine, United States) was added with further incubation at room temperature for 5 mins, followed by elution under centrifuge (3000 x g, 5 mins). The resulting samples were later used for PCR and subsequent pyrosequencing studies.

2.7 Methylation standard curve generation

A DNA methylation standard curve was produced to assess the efficiency of the primers in detection of methylation changes across the range of methylation present in the DNA samples for that genomic region. To this effect, PBMC DNA (50ul, 100ng/ul) was incubated with CpG MethylTransferase (sssI methylase, 1ul, 20 units, NEB, Ipswich, United States), SAM (1ul, NEB, Ipswich, United States), methyltransferase reaction buffer (10x, 10ul, NEB, Ipswich, United States) and ddH_2O (38ul) for 2 hours at 37^oC, before addition of further SAM (1ul), and the reaction left at 37^oC overnight to facilitate complete CpG methylation of DNA. Serial dilutions of this 100% artificially methylated PBMC DNA were then created (80%, 40%, 20%, 10%, 5%, 2.5%, 0%) by diluting the 100% artificially methylated blood sample with original untreated PBMC DNA sample. Dilution with original untreated PMBC DNA rather than H2O is performed to ensure equal concentration of DNA in each reaction, as methylation assessment is sensitive to DNA concentration. In this context, the DNA methylation standard curve can assess the methylation efficiency of the primers across the methylation range of the samples (the 0% additional artificial methylation sample may have residual methylation, indicative of the baseline methylation level of the genomic region). All samples were subsequently subjected to bisulphite conversion treatment, prior to bisulphite pyrosequencing.

2.8 Bisulphite pyrosequencing

Reactions were set up on 96-well plates with 25ul of PCR product. 50ul of PyroMark binding buffer (Qiagen, Hilden, Germany) and 23ul of ddH_2O were mixed with 2ul of streptavidin beads (NEB, Ipswich, United States), and the plate sealed and shaken

on a plate shaker for 10-20 mins, to allow the streptavidin beads to bind to biotin-labelled strands. Sequencing primer (100µM, Eurofins, Luxembourg) stocks were diluted 1:10 with H_2O to a final concentration of 10µM. 1.5ul of sequencing primer (10µM) and 43.5ul of PyroMark annealing buffer (Qiagen, Hilden, Germany) were then added to each well of a 96-well pyrosequencing plate.

The PyroMark Q96 Vacuum Workstation (Qiagen, Hilden, Germany) was set up by filling the appropriate trays with ddH_2O, 70% ethanol, denaturation buffer (0.2M NaOH), and wash buffer (10mM tris acetate). After 10 min of plate shaking, the pyrosequencing station vacuum was engaged and vacuum filter probes were applied to the plate containing the streptavidin beads, until all sample had been passed through the vacuum and the streptavidin beads adhered to the probes. The streptavidin bead-bound filter probes were then placed into the tray containing 70% ethanol for 10 seconds. Probes were subsequently transferred to the tray containing 0.2M NaOH for 10 seconds, then to 10mM tris acetate for 10 seconds, before disengaging the vacuum. The probes were then transferred to the pyrosequencing plate containing the sequencing primer and annealing buffer mixture and scratched along the bottom of the wells. The probes were then finally transferred to a tray containing ddH_2O and the vacuum engaged for cleaning. The pyrosequencing plate was placed onto a heating block pre-heated at 80°C for 2 min, before being allowed to cool for a further 2 min and being transferred to the PyroMark Q96 Pyrosequencer (Qiagen, Hilden, Germany). A PyroMark Q96 cartridge (Qiagen, Hilden, Germany) was prepared by adding enzyme mixture, substrate mixture, and nucleotide mixtures from the PyroMark Gold Q96 Reagents kit (Qiagen, Hilden, Germany) to the relevant wells before placing inside the PyroMark Q96 instrument. Pyrosequencing assays were designed using PyroMark Assay Design software, with pyrosequencing data analysed using PyroMark Q96 software version 2.5.8.15 (Qiagen, Hilden, Germany).

2.9 Statistical analysis

All statistical analysis was performed using GraphPad Prism version 5. The Shapiro-Wilk normality test was applied to statistical analyses to identify normal or non-normal distribution of data, with $p = >0.05$ denoting normal distribution. Where

normal distribution of data was identified, an unpaired t-test was used to compare the means of two independent groups. Where non-normal distribution of data was identified, the Mann-Whitney U test was used to compare the means of two independent groups, while the Kruskal-Wallis test with Dunn's multiple comparison post-test was utilised for comparison of the means of more than two independent groups. In the results, the statistical test used is outlined in each case. Data was considered statistically significant where $p = \leq 0.05$.

2.10 Bioinformatic analyses (Hewitt *et al* 2017, Joehanes *et al* 2016)

Methylation data from Hewitt et al (2017) comparing the methylomes of blood and ocular tissue for matched donor samples was interrogated to determine the mean methylation of CpG loci within genes of interest in whole blood and RPE tissue (n = 7 samples per tissue). As a Shaprio-Wilk normality test determined normal distribution of data in each case, a two-tailed unpaired t-test was used to determine differential methylation, with $p = <0.05$ considered statistically significant. Sample information is outlined in Table 2.7.

Methylation data from Joehanes et al (2016) involving 15,907 patients was interrogated to identify smoking-sensitive CpG sites within genes of interest in AMD within their study. Identified CpG sites were subsequently mapped using UCSC genome browser and assessed for co-localisation with histone modification markers. Sample demographics are outlined in Table 2.8.

Table 2.7: Patient information for samples assessed from Hewitt et al (2017) study

Sample ID	Age at death (years)	Preservation time interval (hours)
3684	37	3.1
3675	56	9.3
3631	58	8.2
3685	63	11.3
3701	62	6.0
3689	67	9.1
3677	76	10.1

8 patients were used in the study by Hewitt et al, however both blood and ocular tissue was only obtained for 7 of the patients, so the 8th sample was omitted for the purposes of this thesis.

Table 2.8: Patient demographics from Joehanes et al (2016) study

Smoking status	Samples (% male)	BMI (mean ± SD)	Age (mean ± SD)
Never smoker	6956 (31.7%)	28.6 ± 5.3	61.2 ± 9.7
Former smoker	6518 (55.6%)	28.7 ± 5.0	64.8 ± 8.2
Current smoker	2433 (46.3%)	27.3 ± 5.4	57.7 ± 7.7

BMI = body mass index in kg/m^2.

2.11 Interrogation of mRNA sequencing data (Kim *et al* 2018)

mRNA sequencing data from Kim *et al* (2018) was interrogated to observe transcriptional changes in peripheral RPE / choroid / sclera at genes of interest to AMD. For each gene reported, UCSC genome browser data from 8 healthy control and 8 AMD samples was used to calculate average exon-exon junction read numbers for both control and AMD populations by dividing the sum of junction read numbers in the primary transcript variant (where applicable) by the number of junctions. Sample information is provided in Table 2.9.

Table 2.9: Patient information for samples used in Kim et al (2018) study

Sample ID	Age (years)	Gender	Disease status / phenotype
Normal #1	85	Male	Normal
Normal #2	84	Male	Normal
Normal #3	92	Female	Normal
Normal #4	86	Female	Normal
Normal #5	83	Male	Normal
Normal #6	83	Male	Normal
Normal #7	84	Male	Normal
Normal #8	83	Female	Normal
AMD #1	69	Male	Early AMD
AMD #2	85	Female	Early AMD
AMD #3	95	Male	Late dry AMD
AMD #4	87	Male	Late wet AMD
AMD #5	83	Female	Late dry AMD
AMD #6	86	Female	Dry AMD
AMD #7	86	Female	Late dry AMD
AMD #8	86	Female	Early AMD

Samples were subsequently interrogated for exon-exon junction read numbers using UCSC genome browser.

2.12 Transcription factor binding prediction analysis

Transcription factor binding prediction was carried out using AliBaba2.1 and PROMO software using the TRANSFAC database. A DNA sequence comprised of 50bp up/downstream of each selected CpG site was entered into each program with transcription factor binding motifs obtained from the TRANSFAC database mapped to matching motifs in the DNA sequence. Transcription factors for which binding motifs overlapped with a CpG site or the immediate surrounding area where noted with comparisons drawn between the results of the 2 databases. CpG sites for examination included the previous sites of interest within *SKI*, *GTF2H4* and *TNXB* as identified by Porter *et al* (2019), as well as additional sites of interest within *AHCY*, *MTR* and *MAT2A* as identified from the bioinformatics analysis outlined in chapter 3 of this thesis.

Chapter 3: Results I: Data mining and bioinformatic analyses of differentially methylated and expressed genes in AMD RPE (*SKI, GTF2H4 and TNXB*) and genes implicated in folate metabolism

3.1 Introduction to chapter 3

Porter *et al*. (2019) identified several genetic loci and regions with significantly different DNA methylation and expression levels in AMD RPE, principally at loci within the genes *GTF2H4, SKI* and *TNXB*. The aims of chapter 3 were the following: firstly, to generate a list of candidate genes including folate metabolism pathway genes with suggested roles in AMD uncovered by data mining for further investigation in blood of patients with AMD; secondly, to investigate baseline methylation levels in blood of differentially methylated and expressed selected candidate genes; thirdly, to identify any overlap with smoking sensitive CpG loci via data mining; and finally, to identify whether the regions of interest were sites of predicted TF binding.

This chapter focusses on the bioinformatic analyses performed to provide baseline data for laboratory studies. Firstly, I analysed published open access RNA sequencing data from Kim *et al*. (192) with a particular focus on genes of interest in AMD as outlined by Porter *et al*. (2019), as well as additional genes involved in the TFIIH complex and the folate pathway. This data was obtained from RPE/choroid of 8 AMD and 8 control patients and assessed at gene transcripts of interest using UCSC genome browser, to enquire whether transcriptional changes at these genes were observed in AMD. In view of the epidemiological association between dietary folate intake and AMD risk, and the direct role of folate in DNA methylation processes, I sought to generate a list of candidate folate metabolism pathway genes with suggested roles in AMD uncovered by data mining.

Secondly, a study of the comparison between the epigenomes of blood and ocular tissue (116) was assessed, allowing the candidate loci of interest in AMD outlined in the above study to be evaluated for analysis in whole blood. Interrogation of this

study allowed the methylation levels of matched patient blood and RPE samples to be compared at these specific loci, with subsequent bisulphite pyrosequencing validation performed on eight healthy control whole blood samples (all non-smokers, four male and four female) from samples shipped on dry ice from colleagues at Radboud University, The Netherlands (see Table 7.1), to confirm comparable baseline methylation levels in *SKI* and *GTF2H4* between blood and RPE using data from Porter *et al* (2019). DNA was extracted using the methods outlined in chapter 2.1, with samples being shipped following approval by the University of Liverpool's institutional ethical review boards (University of Liverpool Central University Research Ethics Committee for Physical Interventions, ref. 2326), with informed written consent being obtained from donors.

Thirdly, identified differentially methylated (DM) regions within *GTF2H4*, *SKI*, and *TNXB* were cross-referenced to DM loci identified in a large-scale study of the "smokers methylome" (156). This study assessed genome-wide differential methylation in smokers, never smokers and ex-smokers using the 450k Illumina Human Methylation BeadChip array, the same array used in Porter *et al* (2019). This cross-referencing enabled a first stage characterisation of a potential association between smoking sensitive CpG loci and differentially methylated genes associated with AMD.

Finally, transcription factor binding site prediction software was utilised to identify whether the loci of interest were predicted TF binding sites, because DNA methylation changes at sites of TF binding may affect TF binding affinity (89). Following the above analyses, a final list of genes for further study in AMD using whole blood as a surrogate tissue at these defined loci was outlined for examination by bisulphite pyrosequencing.

3.2 Results

3.2.1 RNA Sequencing data shows differential expression of *SKI* and *TNXB*, as well as folate pathway genes *AHCY* and *MTR* in AMD RPE

A transcriptome study in AMD using human ocular tissue samples from peripheral RPE/choroid/sclera (PRCS) by Kim *et al*. (2018) revealed differential expression of both coding and non-coding RNA in AMD (192). I analysed this data to assess the differential transcription of genes of interest *SKI, GTF2H4 and TNXB* and other genes of interest including all components of the TFIIH complex that includes *GTF2H*4, as well as genes of importance in the folate and methionine biosynthesis pathways. Kim *et al* aligned RNA-seq reads to the hg19 human genome assembly by using STAR (Spliced Transcript Alignment to a Reference) from ENSEMBL, with data being normalised, prior to quantification, using PORT (pipeline of RNA-sequencing transformations), and pseudogenes with unreliable alignments being filtered out. Normalised data was quantified by identifying all RNA-seq reads consistent with an ENSEMBL splice form for each gene in question, with differential expression of RNA-seq data between AMD and CTRL sample populations being identified by evaluating the difference in read numbers between each population. To assess differential expression between AMD and CTRL populations in genes of interest in AMD and the folate pathway, I interrogated this data further on a gene-by-gene basis to identify significantly differential read numbers between populations.

Analysis of this data revealed several points of interest: while a similar transcription profile for *GTF2H4* was observed between both control and AMD PCRS samples, an increase in AMD mRNA read numbers of exon-exon junctions within both *SKI* and *TNXB* transcripts compared to control was identified (see appendix Figures 7.1-7.3).

With regards to components of the TFIIH complex, many subunits showed little variation in RNA expression in this data: *GTF2H1* and *ERCC2* control samples displayed only slightly higher mRNA reads than AMD, with *GTF2H2* AMD reads slightly higher than control. The mRNA expression of *GTF2H3* and *ERCC3* appeared similar across both sample populations. However, analysis of *GTF2H5*, a gene

associated with trichothiodystrophy, revealed a transcript with higher mRNA expression, noticeably increased in AMD samples compared to control.

Thereafter, I interrogated the data from Kim *et al*. (2018) for mRNA expression changes in genes relevant to the folate and methionine biosynthesis pathways. While no major differences in expression of *DNMT3A*, *DNMT3B* and *SHMT1* and dihydrofolate reductase-like 1 (*DHFRL1)* were observed between AMD and control samples, several differentially expressed folate pathway genes in AMD tissue were identified. *DNMT1, DHFR, CBS, MTR* and *MAT2A* were found to be generally more highly expressed in AMD PRCS, while *MTHFR* and *AHCY* displayed fewer mRNA read counts in AMD. Examples of the transcriptional differences detailed in this study are observed in Table 3.1 and Figures 3.1 (*AHCY*) and 3.2 (*MTR*) below (also see appendix Figures 7.1-7.18 for Kim *et al* mRNA expression for all further genes outlined in Table 3.1).

Table 3.1: mRNA expression of genes of interest in AMD by Kim et al (2018)

Gene	Average CTRL read numbers	Average AMD read numbers	Change in AMD	Normality test?	p-value
SKI	96.85	253.57	↑	Yes	0.0333*
GTF2H1	540.36	517.86	↓	Yes	0.6419
GTF2H2	160.58	201.00	↑	Yes	0.5313
GTF2H3	258.25	268.27		Yes	0.9400
GTF2H4	61.62	59.62		Yes	0.9337
GTF2H5	333.50	449.50	↑	No	0.7813
ERCC2	143.14	126.56	↓	Yes	0.3320
ERCC3	291.64	307.00		Yes	0.5733
TNXB	152.26	175.00	↑	No	0.0441*
DNMT1	103.92	126.74	↑	Yes	0.0080**
DNMT3A	103.47	108.38		Yes	0.4346
DNMT3B	6.55	5.85		Yes	0.2384
DHFR	118.17	157.33	↑	Yes	0.0478*
DHFRL1	40.00	43.00		N/A	N/A
AHCY	738.22	663.89	↓	Yes	0.2012
MTR	300.61	377.94	↑	Yes	0.0006***
SHMT1	375.45	384.18		Yes	0.9041
CBS	252.88	292.56	↑	Yes	0.3680
MAT2A	703.50	1063.63	↑	Yes	0.0205*
MTHFR	212.27	161.72	↓	Yes	0.0459*

"Read numbers" refer to the average observed reads of exon-exon junctions within each gene, calculated as total junction reads divided by number of junctions per gene, for the primary transcript variant of each gene where applicable. "Normality test" indicates whether the data passed the Shapiro-Wilk normality test – normal data was assessed using an unpaired T-test, non-normal data was assessed using a Mann-Whitney U test, with significance level set at $p<0.05$. For DHFRL1, only one read value was present in both control and AMD, meaning no statistics could be performed. (* = $p\leq0.05$, ** = $p\leq0.01$, *** = $p\leq0.001$, **** = $p<0.0001$)

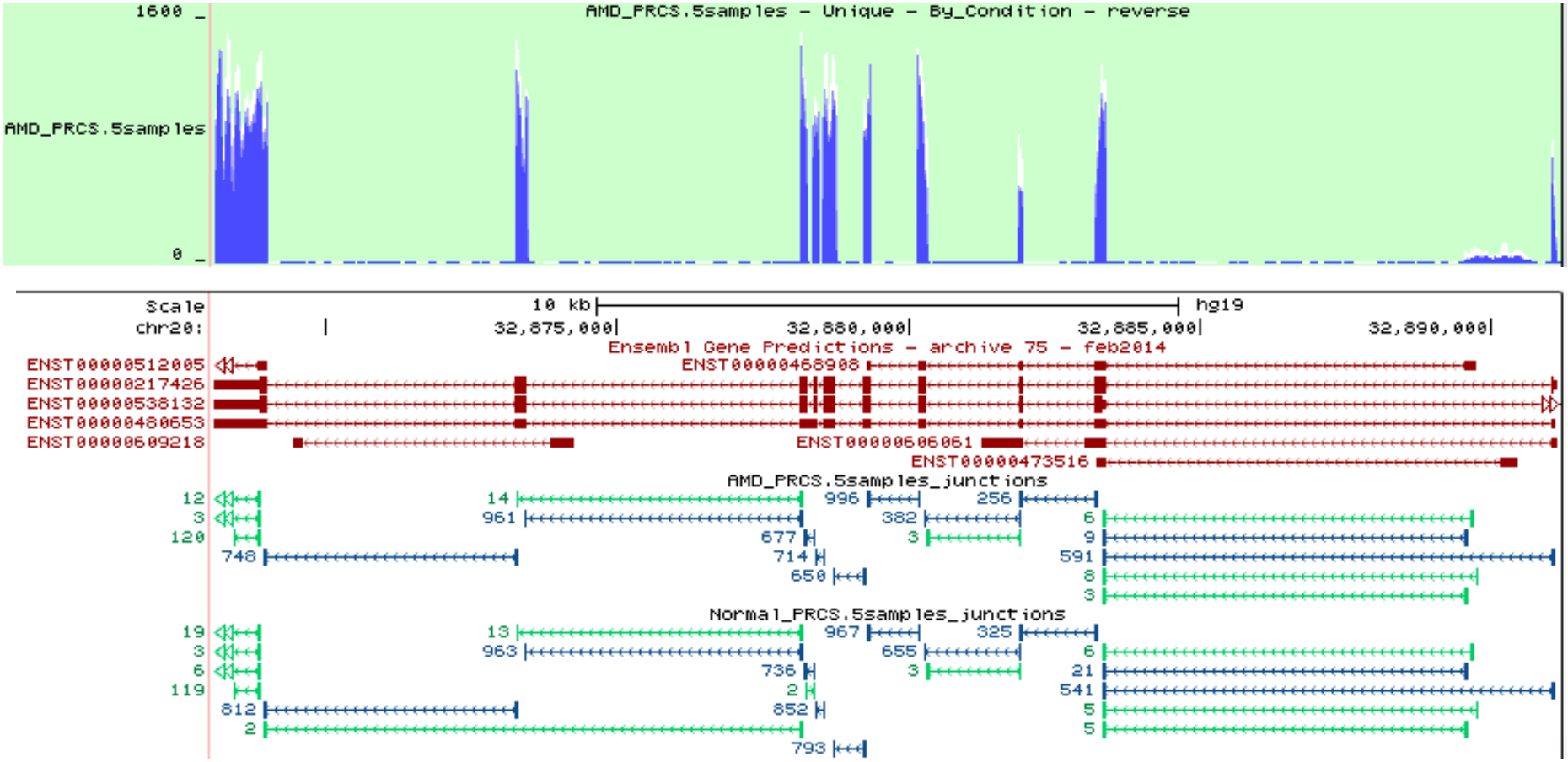

Figure 3.1: RNA sequencing data of *AHCY* obtained from Kim et al. (2018) viewed in UCSC genome browser. The study by Kim et al reports lower expression of *AHCY* in AMD PRCS compared to CTRL PRCS. Top panel: Representation of exon mRNA read numbers in AMD (blue) laid over CTRL (white). Middle panel: Representation of exons and exon-exon junctions of *AHCY* transcripts (red). Bottom panel: Representation of exon-exon junction mRNA read numbers in AMD (above) and CTRL (below) PRCS samples.

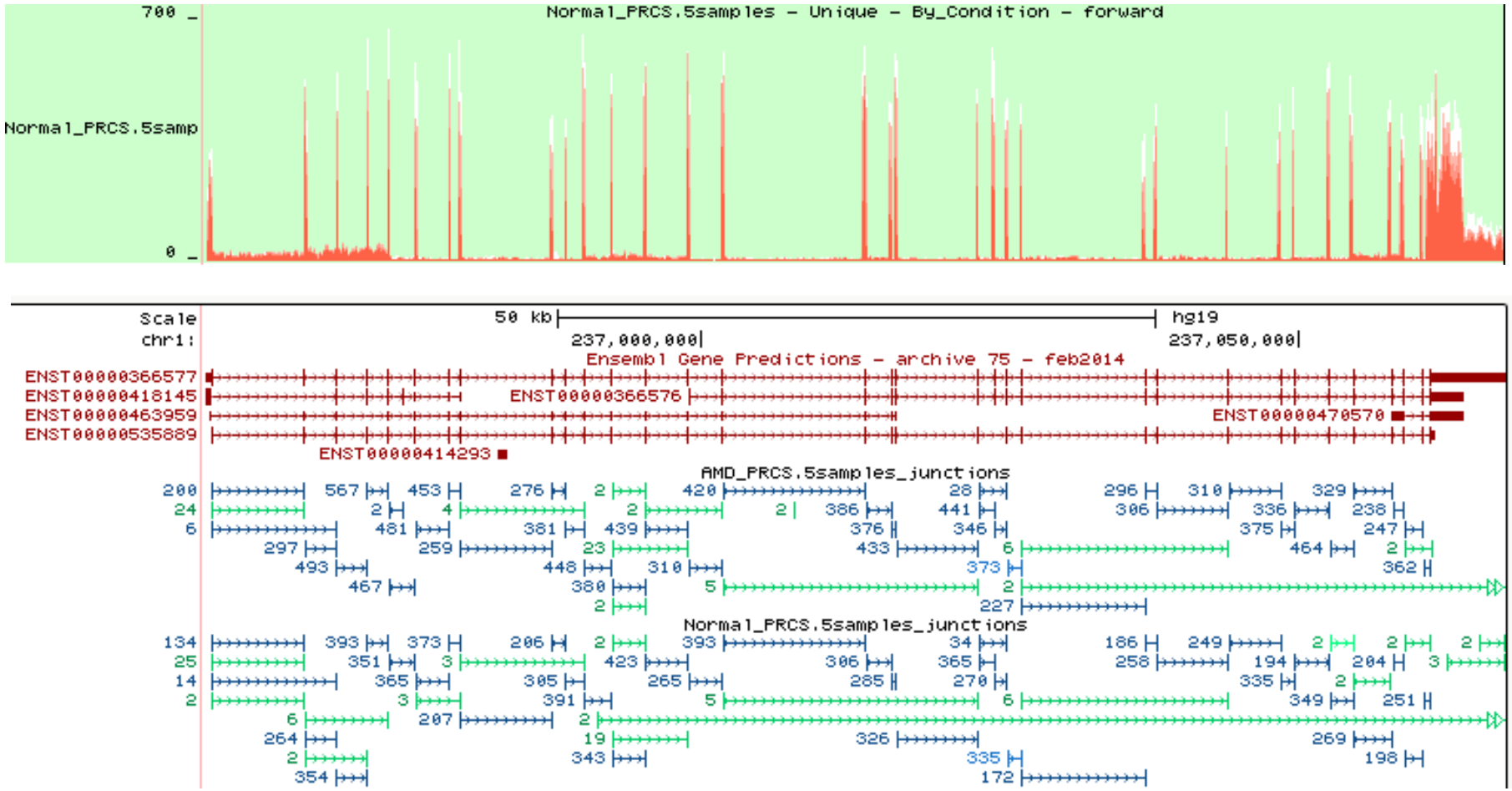

Figure 3.2: RNA sequencing data of *MTR* obtained from Kim et al. (2018) viewed in UCSC genome browser. The study by Kim et al reports higher expression of *MTR* in AMD PRCS compared to CTRL PRCS. Top panel: Representation of exon mRNA read numbers in CTRL (orange) laid over AMD (white). Middle panel: Representation of exons and exon-exon junctions of *MTR* transcripts (red). Bottom panel: Representation of exon junction mRNA read numbers in AMD (above) and CTRL (below) PRCS samples.

3.2.2 Univariate analyses of CpG loci β-values at loci associated with genes involved in the TFIIH complex and folate and methionine biosynthesis pathways genes in Porter et al. 2019 reveals candidate target genes for further analysis

Following the observation of differential mRNA expression of genes in AMD RPE, I then re-examined the DNA methylation data from Porter *et al*. (2019) conducting univariate analyses of DNA methylation β levels at CpG loci associated with genes involved in the TFIIH complex and folate and methionine biosynthesis pathways genes. Within the TFIIH complex, the genes *ERCC2*, *ERCC3*, *GTF2H1*, *GTF2H2*, *GTF2H3*, *GTF2H4* and *GTF2H5* were assessed, as well as the folate pathway genes *AHCY*, *MTR*, *DHFR*, *MAT2A*, *DNMT1*, *MTHFR* and *CBS*.

All CpG sites mapped to the genes were identified and cross-referenced to the methylation array data from Porter *et al*. (2019) to determine differential methylation levels in AMD RPE and control at these loci. The data was interrogated at each specific target CpG site using the mean methylation β values observed across all 25 AMD and 19 healthy control RPE tissues (Porter *et al*.) and a Mann-Whitney U test was performed in each case (Tables 3.2 – 3.4). Many of the data sets assessed failed the Shapiro-Wilk normality test ($p > 0.05$), meaning a non-parametric test was justified. The full list of CpG sites for interrogation within each gene were obtained from the Illumina 450k Methylation BeadChip Array manifest (193).

Significantly differential methylation at each individual CpG locus was determined as $p<0.05$. Genes involved in the folate / methionine-biosynthesis pathway that displayed differential mRNA transcription in Kim *et al*. (2018) were prioritised for further investigation. Therefore, CpG sites within *DNMT3A*, *DNMT3B*, *SHMT1* and *DHFRL1* were not investigated further.

Table 3.2: Univariate analyses of Porter et al. 450k genome-wide DNA methylation array data at CpG loci associated with TFIIH complex genes

Gene	Number of CpG sites associated with gene identified in Illumina 450k manifest	Number of CpG sites associated with gene in Illumina 450k manifest identified by Porter *et al*. (2019)	Univariate analysis results showing number of CpG loci associated with the gene showing significant methylation difference in Porter *et al*. (2019)
ERCC2	25	25	0
ERCC3	14	14	2
GTF2H1	16	16	0
GTF2H2	9	0	0
GTF2H3	18	17	1
GTF2H4	78	75	3
GTF2H5	19	17	0

All CpG sites documented in the Illumina 450K methylation array manifest (193) were checked against data from Porter *et al* (2019). A Mann-Whitney U test to identify significant methylation difference between mean beta value of CTRL (n = 19) and mean beta value of AMD (n = 25) at each individual CpG locus was performed. Significance was classed as $p < 0.05$.

Table 3.3: Univariate analyses of Porter et al. 450k genome-wide DNA methylation array data at CpG loci associated with folate pathway genes

Gene	CpG sites identified in manifest	CpG sites from manifest identified by Porter et al	CpG sites identified as significant by Porter et al
AHCY	27	25	1
MTR	22	21	2
DHFR	17	13	1
MAT2A	16	15	0
DNMT1	22	20	1
MTHFR	23	23	0
CBS	44	41	1

All CpG sites documented in the Illumina 450K methylation array manifest (193) were checked against data from Porter et al, and subsequently assessed using a Mann-Whitney U test to identify significance between CTRL (n = 19) and AMD (n = 25) methylation levels. Significance was classed as $p < 0.05$.

Table 3.4: CpG loci identified as significant following univariate analysis of Porter *et al.* (2019) data

Gene	Identifier	P value (Mann-Whitney U test)
ERCC3	cg07455051	0.0492
ERCC3	cg26107076	0.0028
GTF2H3	cg03156109	0.0330
GTF2H4	cg09474938	0.0080
GTF2H4	cg13472000	0.0492
GTF2H4	cg22508626	0.0003
AHCY	cg02138331	0.0006
MTR	cg17297354	0.0440
MTR	cg25552893	0.0371
DHFR	cg16161425	0.0244
DNMT1	cg15043801	0.0440
CBS	cg27024932	0.0293

With regards to the TFIIH complex, except for one CpG locus within *GTF2H4* (cg18934822) the CpG loci outlined in Table 3.4 did not reach genome-wide significance (Porter *et al*. 2019). Nonetheless, these loci were shortlisted as potential candidates for further study because they are integral components of the transcription dependant DNA repair complex that included *GTF2H4*. In addition, CpG loci within folate pathway genes approaching significance were also included for further investigation, including cg14520600 (*MTR*), cg09363733 (*MAT2A*), cg01817286 (*CBS*) and cg14171527 (*CBS*) (Tables 3.5, 3.6)

Table 3.5: Summary of univariate analysis of Porter et al data at TFIIH complex genes

Gene	Identifier	Mean meth β value CTRL	Mean meth β value AMD	P-value	Significance	Change in AMD	Chr	Location	Notes	≥10% change?
GTF2H3	cg03156109	0.03757	0.03310	0.0330	*	↓	12	124,118,386	5' UTR (approx. 20bp upstream of initiator methionine)	YES
GTF2H4	cg09474938	0.06083	0.05197	0.0080	**	↓	6	30,875,843	Approx. 150bp upstream of TSS	YES
	cg13472000	0.9099	0.9173	0.0492	*	↑	6	30,880,454	Intron 11	
	cg22508626	0.7453	0.8583	0.0003	***	↑	6	30,879,905	Porter et al CpG site (exon 10)	YES
ERCC3	cg07455051	0.03304	0.02938	0.0492	*	↓	2	128,051,853	Approx. 100bp upstream of TSS	YES
	cg26107076	0.04670	0.04013	0.0028	**	↓	2	128,051,925	Approx. 200bp upstream of TSS	YES

All CpG loci of interest were mapped using UCSC Genome Browser to determine proximity to genes. Methylation β levels were assessed to identify if a change of ≥10% of the CTRL mean value was observed in AMD. All statistical analyses were performed using Mann-Whitney U test with significance level set at $p<0.05$. (* = $p\leq0.05$, ** = $p\leq0.01$, *** = $p\leq0.001$, **** = $p<0.0001$)

Table 3.6: Summary of univariate analyses of Porter et al. data for folate pathway genes

Gene	Identifier	Mean meth β value CTRL	Mean meth β value AMD	P-value	Significance	Change in AMD	Chr	Location	Notes	≥10% change?
AHCY	cg02138331	0.9210	0.8981	0.0006	***	↓	20	32,893,975	Approx. 3kbp upstream of TSS	
MTR	cg14520600	0.07190	0.07982	0.0520		↑	1	236,958,447	Approx. 100bp upstream of TSS	YES
MTR	cg17297354	0.8346	0.8517	0.0440	*	↑	1	237,056,641	Intron 29 (co-localises with H3K4Me1 histone mark ENCODE data)	
MTR	cg25552893	0.1424	0.1237	0.0371	*	↓	1	236,958,163	Approx. 400bp upstream of TSS	YES
DHFR	cg16161425	0.06396	0.06930	0.0244	*	↑	5	79,950,956	Approx. 150bp upstream of TSS	
MAT2A	cg09363733	0.2418	0.2142	0.1355		↓	2	85,767,277	Intron 1 (co-localises with H3K4Me3 / H3K27Ac histone marks ENCODE data)	YES
DNMT1	cg15043801	0.03660	0.03266	0.0440	*	↓	19	10,305,911	Approx. 150bp upstream of TSS (co-localises with H3K4Me1, H3K4Me3 and H3K27Ac histone mark ENCODE data)	YES
CBS	cg01817286	0.7444	0.7288	0.0795		↓	21	44,480,336	Intron 12 (co-localises with H3K4Me1 histone mark)	
CBS	cg14171527	0.8075	0.7918	0.0795		↓	21	44,473,840	Exon 17	
CBS	cg27024932	0.1010	0.08944	0.0293	*	↓	21	44,495,796	Intron 1 (co-localises with histone marks H3K4Me1 and H3K27Ac from ENCODE data)	YES

All CpG loci of interest were mapped using UCSC Genome Browser to determine proximity to genes and presence of histone modifications. Methylation β levels were assessed to identify if a change of ≥10% of the CTRL mean value was observed in AMD. All statistical analyses performed using Mann-Whitney U test. (* = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

Based upon the information gained, priority targets were determined for assessment in blood of patients with AMD alongside the CpG loci identified by Porter et al. Of components of the TFIIH complex, cg22508626 within *GTF2H4* (Porter *et al*. 2019) displayed the greatest degree of differential methylation in AMD RPE and was retained for further study (Table 3.5). CpG loci observed within *GTF2H3* and *ERCC3*, all within close proximity to the transcriptional start site (TSS) of their respective genes, were excluded from further study due to low levels of methylation observed across all patient RPE samples at these loci, further reduced in AMD RPE in each case, suggesting that these loci are unlikely to be involved in transcriptional gene silencing, being almost completely unmethylated in both healthy and disease tissues.

While several CpG loci reached significance on univariate analyses in the folate pathway genes, the most significant by far was cg02138331 associated with *AHCY* (Table 3.6), where a significant reduction in methylation levels of a generally highly methylated locus upstream of the TSS was observed in AMD RPE tissue, suggesting that *AHCY* transcriptional activity may be increased in AMD. Of note, a marked decrease in transcriptional activity of *AHCY* in AMD PRCS samples was demonstrated by Kim *et al*. (2018) by RNA sequencing. In this context and given the role of *AHCY* in the biosynthesis of homocysteine, this locus was chosen for further interrogation. Similarly, reduced levels of methylation of cg25552893 associated with *MTR* was observed in AMD RPE, again upstream of the TSS. Differential transcription levels of *MTR* were also shown by RNA sequencing in AMD PRCS samples by Kim *et al*. (2018). In this context and given the function of *MTR* this CpG locus was also selected for further investigation as *MTR* facilitates the metabolism of homocysteine to methionine. Finally, while not significant, lower methylation levels were observed at cg09363733 within intron 1 of *MAT2A* in AMD RPE, at a site significantly enriched for H3K4Me3 and H3K27Ac, markers of transcriptional enhancer regions. As *MAT2A* also displayed differential mRNA transcription in the study by Kim *et al*. (2018) and is responsible for the biosynthesis of the universal methyl donor SAM (of paramount importance in global DNA methylation), cg09363733 was selected as the third and final additional CpG locus, alongside

those observed by Porter et al, for further investigation in the blood of AMD donors. Other candidate CpG loci observed either did not reach significance on univariate analyses, were present at genomic locations of lesser relevance, or were largely unmethylated between both sample groups, and were therefore not deemed priority targets. The final candidate loci are outlined in Table 3.7.

Table 3.7: Summary of CpG loci of interest for further study in AMD whole blood

Gene	Identifier	Chromosome	Location
GTF2H4	cg22508626	6	30,879,905
SKI	cg18934822	1	2,191,402
TNXB	DMR	6	32,063,865 – 32,064,058
AHCY	cg02138331	20	32,893,975
MTR	cg25552893	1	236,958,163
MAT2A	cg09363733	2	85,767,277

Each of the loci and surrounding CpG sites were subsequently assessed using bisulphite pyrosequencing in blood of age matched AMD donors and controls.

3.2.3 Methylation of *GTF2H4*, *SKI*, *TNXB*, and folate pathway genes MTR and MAT2A show similar baseline methylation levels in blood and RPE

Hewitt *et al*. (2017) compared whole genome methylation levels across ocular tissues (RPE / choroid, optic nerve and retina) and matched whole blood for eight individuals (all male, tissues obtained within 12 hours post-mortem) using the Illumina Infinium 450K array (116). Analysis of this study revealed a general trend of similarity between the methylomes of RPE / whole blood, with 255 586 probes (59%) demonstrating similar levels of methylation between all tissues (defined as either <20% or >80% methylation observed across all tissues). Interrogation of this open access data at the differentially methylated CpG loci reported by Porter *et al*. was performed. As RPE / choroid tissue was not available for one of the patients, the remaining seven were evaluated for differences between blood and RPE tissue at selected CpG loci and regions. Probes for the differentially methylated CpG sites within *SKI* (cg18934822) and *AHCY* (cg02138331), as well as one site within the

TNXB DMR (cg01137207) were not represented in the Hewitt *et al*. (2017) study, meaning that unfortunately a comparison of methylation using this method between RPE and matched whole blood at these sites could not be made. However, no significant methylation differences between RPE and whole blood in matched patient samples represented in Hewitt *et al*. were observed at CpG sites within the genes *GTF2H4* (cg22508626, p=0.4779) and *TNXB* (cg26266427 p=0.3811, cg10365886 p=0.9619, cg07524919 p=0.7466, cg00872984 p=0.5883, cg27387193 p=0.7943, cg17662683 p=0.7646). Similarly, CpG sites within the folate pathway genes *MTR* (cg25552893, p=0.1475) and *MAT2A* (cg09363733, p=0.4682) showed no methylation difference between tissue types. These results are outlined in Table 3.8 below. Application of the Shapiro-Wilk normality test demonstrated normal data distribution in all cases, with a two-tailed unpaired t-test therefore used to determine significance. No significant difference in methylation levels across tissues was noted at these loci.

To validate these findings, bisulphite pyrosequencing of eight healthy whole blood samples was carried out for the CpG sites of interest in both *SKI* and *GTF2H4* and compared to the pyrosequencing results of eight healthy RPE samples obtained from the Porter et al. study. Unfortunately, as bisulphite pyrosequencing had not been carried out across the DMR region for *TNXB* in the Porter *et al*. study, a comparison was not made between blood and RPE methylation for these CpG sites. Of note, similar and overlapping levels of methylation were observed between blood and RPE tissue at CpG loci assessed within *SKI* and *GTF2H4*, with no significant difference in methylation status identified (Table 3.9).

Table 3.8: Mean methylation levels of CpG sites of interest in blood and RPE / choroid tissue reported in Hewitt *et al*. (2017)

		Hewitt et al. (2017) – Mean methylation β-values per tissue type		
Gene	**Identifier**	**Blood**	**RPE / Choroid**	**P value**
SKI	cg18934822	N/A	N/A	
GTF2H4	cg22508626	0.9439	0.9523	0.4779
TNXB	cg01137207	N/A	N/A	
TNXB	cg26266427	0.7055	0.6596	0.3811
TNXB	cg10365886	0.7522	0.7545	0.9619
TNXB	cg07524919	0.6083	0.6356	0.7466
TNXB	cg00872984	0.5459	0.5815	0.5883
TNXB	cg27387193	0.5459	0.5620	0.7943
TNXB	cg17662683	0.3877	0.3684	0.7646
AHCY	cg02138331	N/A	N/A	
MTR	cg25552893	0.1773	0.1285	0.1475
MAT2A	cg09363733	0.2280	0.2519	0.4682

Statistical analysis was performed using a two-tailed unpaired t-test (n = 7 samples per tissue type) with significance level set at $p<0.05$.

Table 3.9: Mean methylation β-levels of CpG sites of interest within *SKI* measured in whole blood and RPE using bisulphite pyrosequencing

		Bisulphite pyrosequencing – Mean methylation β-values per tissue		
Gene	**Identifier**	**Blood**	**RPE**	**P value**
SKI	cg18934822	0.9703	0.9661	0.7170
GTF2H4	cg22508626	0.9636	0.9583	0.6854

Statistical analysis was performed using a two-tailed unpaired t-test (n = 8 per tissue) with significance level set at $p<0.05$.

3.2.4 Smoking sensitive methylation is observed at CpG loci within the differentially methylated genes in AMD *SKI* and *TNXB*

Smoking is a major risk factor for the development of AMD. A recent genome-wide DNA methylation study by Joehanes *et al.* (2016) performed on whole blood samples from 15907 patients (2433 current, 6518 former and 6956 never smokers) identified 2623 CpG sites across 1405 genes that were significantly differentially methylated between current and never smokers (156). I evaluated this data to

identify the prevalence of smoking-sensitive CpG sites within the AMD genes of interest *SKI, GTF2H4, TNXB* (Porter *et al*. 2019). Of these 2623 smoking-sensitive differentially methylated CpG sites, 19 were mapped to *SKI*, 24 to *TNXB*, and 1 to *GTF2H4*. Equal distribution of the smoking sensitive CpG loci identified by Joehanes *et al* would result in an average of 1.87 loci per gene (0.07% of total loci observed). As the proportion of total sites mapped to both *SKI* (0.72%) and *TNXB* (0.91%) are greater than this average, this would suggest an enrichment of smoking sensitive loci within these genes, while no such enrichment was observed for *GTF2H4* (0.038%).

For *SKI* and *TNXB*, 4 differentially methylated CpG sites within each gene were also found to be significantly differentially methylated between former and never smokers (*SKI*: cg01979157, cg05603985, cg09469355, cg08884752; *TNXB*: cg16225663, cg15697476, cg00188055, cg19108771). In Joehanes *et al* (2016), "current" smokers were defined as individuals who had smoked at least 1 cigarette a day for at least 12 months prior to sampling and "former" smokers were defined as those who smoked at least 1 cigarette a day in the past, but had stopped at least 12 months prior to sampling (i.e. at least 12 months "smoke-free"). Tables 3.10 and 3.11 contain details of smoking sensitive CpG sites within *SKI* / *TNXB.*

For each CpG site found to be significantly differentially methylated between smokers and non-smokers within *GTF2H4*, *SKI* and *TNXB*, including the Porter et al (2019) AMD specific CpG loci cg22508626 (*GTF2H4*), cg18934822 (*SKI*), cg01137207, cg26266427, cg10365886, cg07524919, cg00872984, cg27387193 and cg17662683 (*TNXB*), I interrogated the genomic region for evidence of histone modifications, including H3K4me1, H3K4me3 and H3K27ac using UCSC Genome Browser and data collated from the ENCODE project (194) to generate associations through histone markers with putative active promoter and enhancer regions in the vicinity of each site. Within *SKI* and *TNXB*, many CpG loci were found to occur in the presence of H3K4me1 and H3K27ac, suggesting that these sites of significantly differential methylation in smokers occur at putative enhancer regions (Tables 3.10, 3.11; Figure 3.3).

Table 3.10: "Smoking sensitive" CpG loci observed within *SKI* in Joehanes et al. (2016) in "former" smokers

CpG Site	Identifier	Chr	Intron / Exon	Location	Value	Former smokers?	H3K4me1	H3K4me3	H3K27ac
A	cg01979157	1	Exon 1	2,161,013	-0.0081	✓	✓	✓	✓
B	cg05603985	1	Exon 1	2,161,049	-0.0122	✓	✓	✓	✓
C	cg09469355	1	Intron 1	2,161,886	-0.0166	✓	✓	✓	✓
D	cg08884752	1	Intron 1	2,162,001	-0.0171	✓	✓	✓	✓
E	cg08640824	1	Intron 1	2,162,506	-0.0103		✓	✓	✓
F	cg01787285	1	Intron 1	2,162,682	-0.0065		✓	✓	✓
G	cg16417118	1	Intron 1	2,162,931	-0.0074		✓	✓	✓
H	cg19404444	1	Intron 1	2,164,602	-0.0045		✓		
I	cg12611488	1	Intron 1	2,169,282	0.0061		✓		
J	cg16704920	1	Intron 1	2,169,407	0.0043		✓		
Porter et al	cg18934822	1	Intron 1	2,191,402			✓		
K	cg01949002	1	Intron 1	2,203,589	-0.0044		✓		
L	cg15007228	1	Intron 1	2,210,335	-0.0041		✓		✓
M	cg03466562	1	Intron 1	2,220,317	-0.0046		✓		
N	cg16902087	1	Intron 1	2,220,497	-0.0071		✓		
O	cg08271229	1	Intron 1	2,222,674	0.0038		✓	✓	✓
P	cg15935121	1	Intron 1	2,230,601	0.0068		✓		
Q	cg26332488	1	Intron 1	2,230,640	0.0077		✓		
R	cg26017930	1	Intron 1	2,232,166	0.0014		✓		✓
S	cg07425568	1	Intron 5	2,236,650	0.0030				

"Value" indicates the direction and magnitude of methylation change in β-value observed in smokers compared to non-smokers. The column "former smoker" indicates whether the change is also present in ex-smokers. Co-localisation with histone modification information was gained using data from the ENCODE project within UCSC Genome Browser.

Table 3.11: "Smoking sensitive" sites observed within *TNXB* in Joehanes et al. (2016)

CpG Site	Identifier	Chr	Intron / Exon	Location	Value	Former smokers?	H3K4me1	H3K4me3	H3K27ac
A	cg12112110	6	Intron 43	32,009,422	-0.0064				
B	cg13409674	6	Intron 30	32,014,476	0.0073		✓		✓
C	cg02272968	6	Exon 30	32,015,773	-0.0078		✓		
D	cg05473289	6	Exon 29	32,016,257	-0.0039		✓		✓
E	cg24376689	6	Exon 29	32,016,336	-0.0057		✓		✓
F	cg09474017	6	Exon 29	32,016,368	-0.0066		✓		✓
G	cg03216698	6	Intron 24	32,021,500	-0.0062				
H	cg19071976	6	Intron 21	32,026,656	-0.0125		✓		
I	cg16225663	6	Intron 21	32,026,797	-0.0113	✓	✓		
J	cg16678426	6	Intron 21	32,026,891	-0.0074		✓		
K	cg10463626	6	Intron 6	32,054,094	-0.0028		✓		
L	cg06580770	6	Intron 6	32,054,790	-0.0051		✓		
M	cg02352370	6	Intron 6	32,054,845	-0.0046		✓		✓
N	cg11823230	6	Intron 6	32,055,385	-0.0032		✓		✓
O	cg05483184	6	Intron 6	32,055,393	-0.0030		✓		✓
P	cg21460606	6	Intron 6	32,055,402	-0.0016		✓		
Q	cg15697476	6	Intron 6	32,055,474	-0.0039	✓			
R	cg17001689	6	Intron 6	32,055,525	-0.0014				
S	cg02740950	6	Intron 6	32,055,534	-0.0034				
T	cg18178010	6	Intron 6	32,055,629	-0.0011				
Porter et al	DMR	6	Exon 3	32,063,865 – 32,064,058					
U	cg00188055	6	Intron 1	32,076,079	0.0064	✓	✓		
V	cg14630933	6	Intron 1	32,076,093	0.0107		✓		
W	cg09872737	6	Intron 1	32,076,270	0.0059		✓	✓	✓
X	cg19108771	6	Intron 1	32,076,337	0.0104	✓	✓	✓	✓

"Value" indicates the direction and magnitude of methylation change in β value observed in smokers compared to non-smokers. The column "former smoker" indicates whether the change is also present in ex-smokers. Co-localisation with histone modification information was gained using data from the ENCODE project within UCSC Genome Browser.

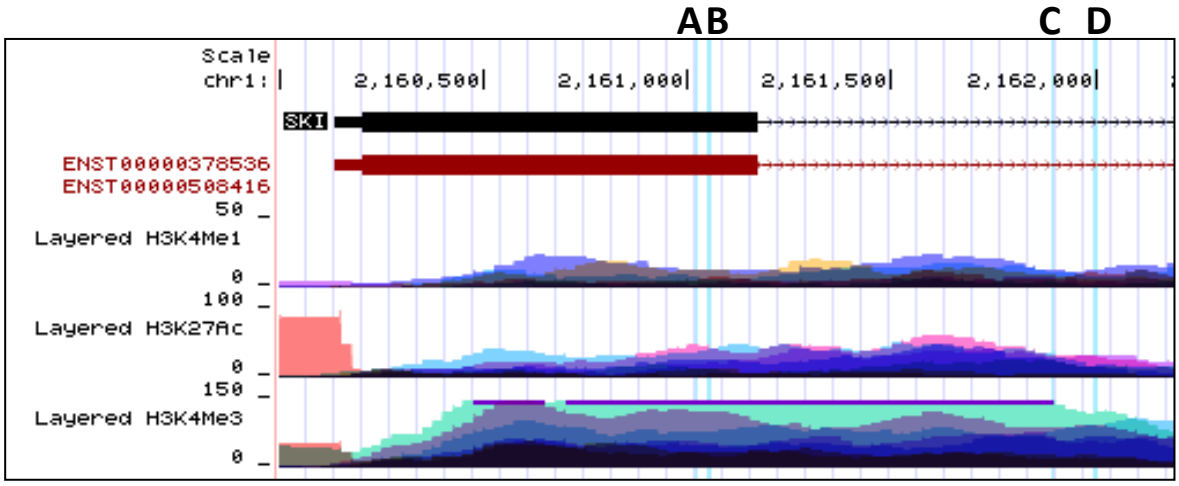

Figure 3.3: Histone modification data from USCS genome browser at smoking sensitive CpG sites within SKI (cg01979157, cg05603985, cg09469355, cg08884752**) observed by Joehanes et al.** Differentially methylated CpG sites in smokers observed within *SKI* exon 1 and intron 1 occur at sites of histone modifications, including H3K4me1, H3K4me3 and H3K27ac. A-D represent CpG loci outlined in Table 3.3, A: cg01979157; B: cg05603985; C: cg09469355; D: cg08884752 respectively.

3.2.5 Transcription factor binding site predictions at selected CpG sites

With 6 candidate target genes in mind for further investigation of differential methylation in whole blood of AMD patients, I employed transcription factor binding site prediction software packages to assess the predicted *in silico* presence of transcription factor binding sites in the vicinity of cg22508626 (*GTF2H4*), cg18934822 (*SKI*), cg02138331 (*AHCY*), cg25552893 (*MTR*), cg09363733 (*MAT2A*) and the CpG loci within the *TNXB* DMR (Porter *et al*. 2019).

I used AliBaba2.1 and PROMO for the analysis of transcription factor binding predictions at these loci. AliBaba2.1 is a prediction programme based on TRANSFAC, a database of eukaryotic transcription factors, and generates transcription factor binding predictions based on experimental evidence, matching potential binding motifs to a given DNA sequence. PROMO also utilises the TRANSFAC database to create weight matrices for transcription factor binding predictions. The DNA sequence encompassing each of the selected target loci / DMR (50bp either side of CpG site / DMR) was assessed using both programs, with

transcription factors predicted to bind at or immediately adjacent to each CpG locus documented. For the *TNXB* DMR, each of the 7 loci within the DMR was assessed separately. Results are outlined in Table 3.12.

Table 3.12: Transcription factor binding predictions obtained from the TRANSFAC database

Gene	Identifier	TFs Identified – AliBaba2.1	TFs Identified - PROMO	Overlapping sites
GTF2H4	cg22508626	C/EBPα	IRF-1 C/EBPα AP-2α	C/EBPα
SKI	cg18934822	IRF-1	IRF-1 E2F-1	IRF-1
TNXB	cg01137207	N/A	N/A	N/A
TNXB	cg26266427	N/A	N/A	N/A
TNXB	cg10365886	N/A	E2F-1 C/EBPβ C/EBPα	N/A
TNXB	cg07524919	SP-1	GR-α TFII-I	N/A
TNXB	cg00872984	C/EBPα AP-1	c-Jun GR-α	N/A
TNXB	cg27387193	SP-1 USF YY1	NF-kappaB1	N/A
TNXB	cg17662683	SP1	AhR/Arnt AP-2α AR	N/A
AHCY	cg02138331	GATA-1 Oct-1	GATA-1	GATA-1
MTR	cg25552893	N/A	E2F-1 FOXP3 p53	N/A
MAT2A	cg09363733	N/A	E2F-1 TFII-I	N/A

While Table 3.12 merely outlines predictions of transcription factor binding and not experimentally confirmed evidence for each gene, interesting observations can be made based on the nature of the transcription factors predicted to be associated with the genomic regions at and immediately surrounding these sites of differential methylation in AMD RPE.

With regards to TF binding site predictions, despite both drawing from the TRANSFAC database, there was a fairly small degree of overlap of predicted transcription factor binding observed between the two software platforms used. Common to both however was the predicted binding of interferon regulatory factor-1 (IRF-1) at the site of significantly differential methylation in *SKI*, also predicted in *GTF2H4* by the PROMO software. Also of note is the recurring prediction of the binding of retinoblastoma-associated protein 1 (E2F-1) at sites of differential DNA methylation in AMD RPE, identified at the CpG loci within *SKI*, *TNXB*, *MTR* and *MAT2A.*

3.3 Discussion for chapter 3

Of note, differential mRNA transcription in *SKI* and *TNXB* was shown in AMD RPE by Kim et al. and Porter et al., suggesting a potential role of differential methylation and expression of these genes in AMD. The identification of increased mRNA expression in control vs AMD PRCS samples of genes implicated in folate and methionine metabolism, including *AHCY*, as well as the decreased expression in control vs AMD PRCS samples of *MTR*, is particularly noteworthy. *AHCY* facilitates the biosynthesis of homocysteine, while *MTR* facilitates the metabolism of homocysteine to methionine. It is of interest that AMD patients appear to display reduced expression of a gene which functions to upregulate homocysteine, as well as an increased expression of a gene which facilitates the breakdown of homocysteine, given the proposed link between increased serum homocysteine levels and AMD progression. As such, these genes warrant further investigation.

Univariate analyses carried out on the data obtained from Porter et al. refined candidates to six target CpG loci/regions of interest. *SKI* (cg18934822) and *TNXB* (DMR ch6 32,063,865 – 32,064,058) were found to be differentially methylated and expressed in AMD RPE (Porter et al, 2019) and found to be closely related with "smoking-sensitive" CpG loci. *GTF2H4* (cg22508626) was differentially methylated and expressed in AMD RPE; *AHCY* and *MTR* were found to be differentially expressed in AMD RPE as observed by Kim *et al* (2018). Finally, given the role of

SAM in DNA methylation as the universal methyl donor, *MAT2A* was chosen as a further representative of the folate pathway due to the observation of lower methylation levels in AMD RPE at cg09363733, within intron 1 of *MAT2A*, a site with particularly high enrichment of enhancer markers H3K4me3 and H3K27ac.

One of the principal aims of the project was to determine the validity of using whole blood tissue as a proxy in epigenetic studies on AMD. I therefore examined Hewitt et al. (2017) data (116) that compared the methylomes of whole blood and ocular tissues in matched patient samples and revealed significant overlap between baseline methylation levels in blood and RPE at CpG loci identified by Porter *et al.* as well as at CpG loci within the folate pathway genes of interest *MTR* (cg25552893) and *MAT2A* (cg09363733). In addition, six of the seven CpG loci within the *TNXB* DMR identified by Porter et al. showed similar methylation levels between whole blood and RPE/choroid in Hewitt et al.

In vitro validation of similar baseline methylation levels for *SKI* and *GTF2H4* by bisulphite pyrosequencing revealed comparable methylation between whole blood and RPE tissue, supporting the use of whole blood as a proxy for studying AMD epigenetics at selected loci and regions.

Thereafter, I utilised data from studies of smoking-related DNA methylation generated in whole blood to investigate a role for smoking sensitivity at AMD associated differentially methylated loci and regions. Joehanes et al. (2016) revealed 2623 CpG sites across the genome as “smoking-sensitive” sites. While an impact of smoking on DNA methylation is therefore apparent, the impact of smoking on AMD risk through an interaction with DNA methylation is not clear. Analysis of Joehanes et al. data focussing on *SKI*, *GTF2H4* and *TNXB*, differentially methylated genes identified in AMD RPE (Porter *et al*) provided some insights. Smoking sensitive loci in the vicinity of *SKI*, *GTF2H4* and *TNXB* included cg05603985, within exon 1 of *SKI* (the third most significantly differentially methylated CpG locus between current and non-smokers), 30kb from cg18934822 within intron 1 of *SKI,* associated with AMD. A further 18 smoking-sensitive loci were associated with *SKI*, suggesting a possible association between smoking status and *SKI* transcriptional regulation.

With regards to TF binding site predictions IRF-1 was predicted to bind at the site of significantly differential methylation in *SKI* (cg18934822), and *GTF2H4* (cg22508626) by the PROMO software. IRF-1 is reported to be responsible for the induction of CFH expression in human RPE tissue, a known component of drusen in AMD patients, strengthening the suggested role of differential methylation of cg18934822 in *SKI* and cg22508626 in *GTF2H4* in AMD development (195).

E2F-1 is predicted to bind at sites of differential DNA methylation in AMD RPE, identified within *SKI*, *TNXB*, *MTR* and *MAT2A.* It has been recently reported that E2F-1 is involved in the protection of RPE cells against oxidative stress-induced damage, known to play a role in the pathogenesis of AMD, making the suggestion of E2F-1 binding sites at areas of known differential methylation in AMD, while only predictive, a particularly noteworthy finding (196).

Binding of transcription factor II-I (TFII-I) is predicted at both cg07524919 within the *TNXB* DMR, as well as at cg09363733 within *MAT2A*. TFII-I binding to the 3' UTR region of the *ARMS2* gene, a locus representing one of the most well documented associations with AMD genetic risk (197), has been associated with upregulation of the serine protease HTRA1 and subsequent generation of a CNV-like phenotype in mice through upregulation of VEGF (198). As HTRA1 is a known inhibitor of TGF-β signalling, this prediction could also suggest further implications for the role of differential methylation of *SKI* and the TGF-β signalling pathway in AMD development. Furthermore, at both cg22528626 in *GTF2H4* and cg17662683 within the *TNXB* DMR, biding of transcription factor AP-2-alpha (TFAP2α) is predicted. TFAP2α is a transcription factor associated with eye development, mutations in which are associated with branchio-oculo-facial syndromes (BOFS) (199).

The TF binding site prediction software packages PROMO and AliBaba2.1 were used together due to their shared utilisation of the TRANSFAC database, potentially allowing for increased reliability of any observed overlapping predictions. This would give the potential advantage of identifying specific CpG sites located at binding sites of transcription factors with relevance to AMD or ocular disease in general. However, the relatively small degree of overlap observed between the two platforms may be explained by differences in their prediction methods – while both

platforms employ weigh matrices matched to TRANSFAC data, one focus of AliBaba2.1 is the construction of multiple matrices specifically to identify overlapping TF binding sites (200), while PROMO utilises a higher number of matrices constructed computationally to improve reliability, and allow greater probability of any sequence alignment due to random chance being filtered out by the software, minimising false positives (201). While both draw from the TRANSFAC database, differences in the number and construction of the weight matrices used for prediction may yield different predictions, additionally with predictions identified using AliBaba2.1 being subsequently filtered out by the more robust matrices constructed using PROMO.

For more accurate prediction of transcription factor binding sites, more recent platforms with unique prediction methods could be utilised. For example, CiiiDER draws from the TRANSFAC database similar to the software packages used in this study, however also performs an “enrichment” step comparing TF binding site predictions to the distribution of transcription factors in relevant background sequences, to improve accuracy of true TF binding site identification (202). Alternatively, TFBSPred performs binding site prediction by drawing from DNAse I hypersensitivity data from ENCODE, allowing open chromatin regions to be identified for sites of interest and claiming to eliminate false discoveries by utilising novel modelling strategies instead of the weight matrix-based methods of older models such as PROMO and AliBaba2.1 (203). Incorporation of these more recent software packages in the future may allow for novel transcription factor binding sites to be identified, and with potentially more reliability and overlap than obtained with the packages used in this study, possibly further clarifying the role of these particular CpG sites and genes in the onset or development of AMD.

Chapter 4: Results II: Bisulphite pyrosequencing of genes of interest in whole blood-derived DNA from patients with and without AMD

4.1 Introduction to chapter 4

Following the identification of 6 candidate genes with promising associations with AMD for DNA methylation studies in AMD whole blood, further experiments were undertaken to achieve the remaining aims of the project.

This chapter focuses on the optimisation and use of PCR and bisulphite pyrosequencing techniques to evaluate methylation levels at target CpG loci in AMD patients, and to determine the impact of smoking status on DNA methylation at these loci. To achieve this, 188 AMD and 188 healthy control whole blood samples were shipped on dry ice from colleagues at Radboud University, The Netherlands following approval by the University of Liverpool's institutional ethical review boards (University of Liverpool Central University Research Ethics Committee for Physical Interventions, ref. 2326), with informed written consent being obtained from donors. Of these samples, 96 of each were selected for further study, further details of which can be found in Tables 2.1 and 7.1. To ensure validity of the pyrosequencing studies, these samples were first assessed for quality using agarose gel electrophoresis to assure no DNA degradation had taken place. Following the successful determination of intact DNA samples for further use, primers for regions of interest within each gene were designed and optimised for PCR amplification.

To confirm that bisulphite pyrosequencing could detect changes in methylation within these regions, and to verify that designed primers and pyrosequencing assays were fit for purpose, fragments corresponding to each genetic region were subjected to enzymatic methylation to generate a methylation standard curve in each case. Bisulphite pyrosequencing was then carried out on the entire cohort of whole blood samples for each genomic target region, stratified by disease state and smoking status, with statistical analysis performed to evaluate differential methylation.

To validate the findings from initial pyrosequencing studies, a validation cohort of AMD and healthy control blood samples, stratified by smoking status, was shipped on dry ice from collaborators at the University of Southampton comprising 48 AMD and 48 healthy control samples with DNA having been extracted using the aforementioned methods, again with informed written consent being obtained from donors. Further information regarding the demographics of these samples can be found in Tables 2.2 and 7.2. This validation cohort was used for further bisulphite pyrosequencing studies of the genes demonstrating smoking-sensitive CpG loci in the discovery set.

4.2 Results

4.2.1 Whole blood DNA samples show little signs of degradation

Whole blood samples obtained from the Netherlands (188 control: 36 smokers, 152 non-smokers; 187 AMD: 27 smokers, 160 non-smokers) were assessed for evidence of DNA degradation before further use. AMD samples were categorised as either AREDS grade 2 (early AMD) or grade 3 (intermediate AMD). For each sample, DNA concentration was determined via nanodrop, with 100ng DNA assessed via agarose gel electrophoresis at 100V for 45min. While most samples showed no signs of DNA degradation, for a small number of samples either weak / absent bands or signs of smearing were observed, suggested at least a degree of degradation. Such samples were removed from the study, with subsequent processes utilising only whole blood samples for which no degradation could be observed (Figure 4.1).

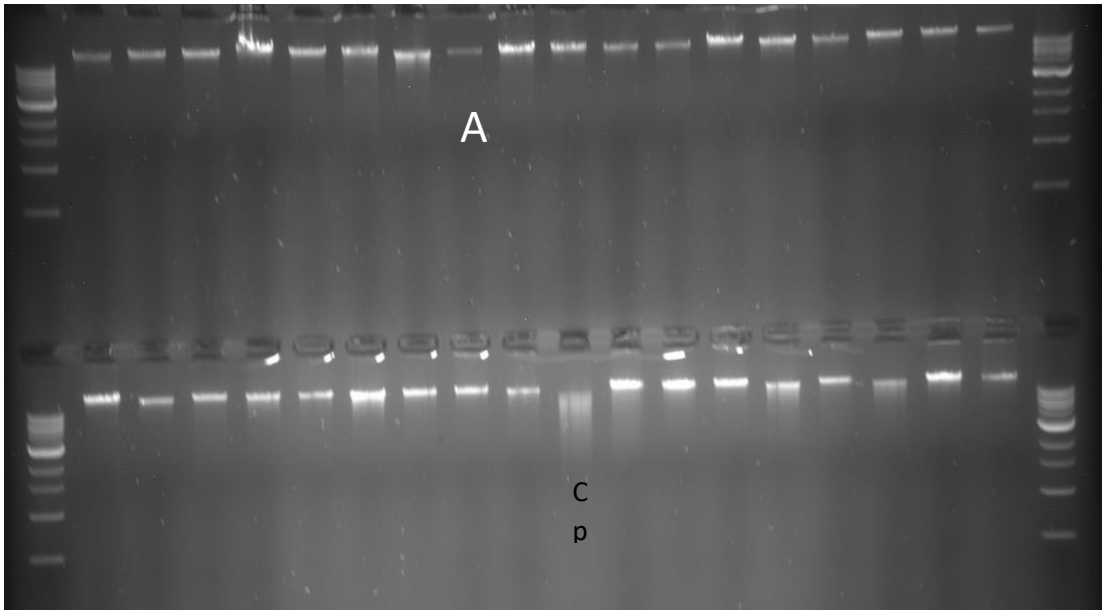

Figure 4.1: Example of DNA degradation analysis. An example of the results obtained from agarose gel electrophoresis of Netherlands cohort whole blood DNA samples. Evidence of weak bands (A) or smearing (B) were taken as signs of DNA degradation, with such samples being excluded from the project.

Once a list of suitable DNA samples had been collated, the decision was made to continue the bisulphite pyrosequencing studies using 96 control and 96 AMD (48 grade 2, 48 grade 3) samples. After exclusion of samples showing potential DNA degradation, the remaining 192 samples were chosen based on the observed DNA concentrations from nanodrop evaluation, with priority given to the most concentrated DNA samples to maximise the amount of DNA available for further study. To ensure no age bias between sample populations (control, grade 2, grade 3), samples were age matched between each population. Statistical analysis shows no significant difference in the mean age of each sample population (Figure 4.2).

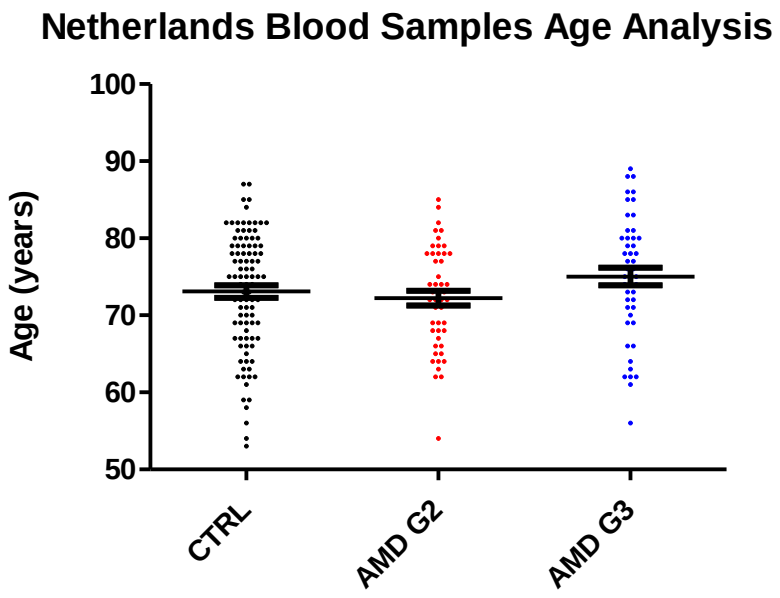

Figure 4.2: Age analysis of Netherlands sample cohort proposed for further study. A Kruskal-Wallis test determined no statistical significance in age between the sample populations (p = 0.1712). Average age ± SEM (years): CTRL = 73.1 ± 0.8; AMD G2 = 72.2 ± 1; AMD G3 = 75 ± 1.1

4.2.2 Bisulphite pyrosequencing assays detect DNA methylation changes at the candidate DNA loci

Pyrosequencing assays designed for each candidate gene were assessed with a methylation standard curve. Briefly, healthy PBMC DNA samples obtained from Dr Lakis Liloglou at the University of Liverpool were enzymatically methylated using *sssI* methylase and a serial dilution of DNA methylation was created (as described in chapter 2: materials and methods). These DNA dilutions were subsequently bisulphite converted and amplified via PCR for each gene of interest before bisulphite pyrosequencing, with the degree of methylation calculated for each dilution per target gene (Figure 4.3, Table 4.1).

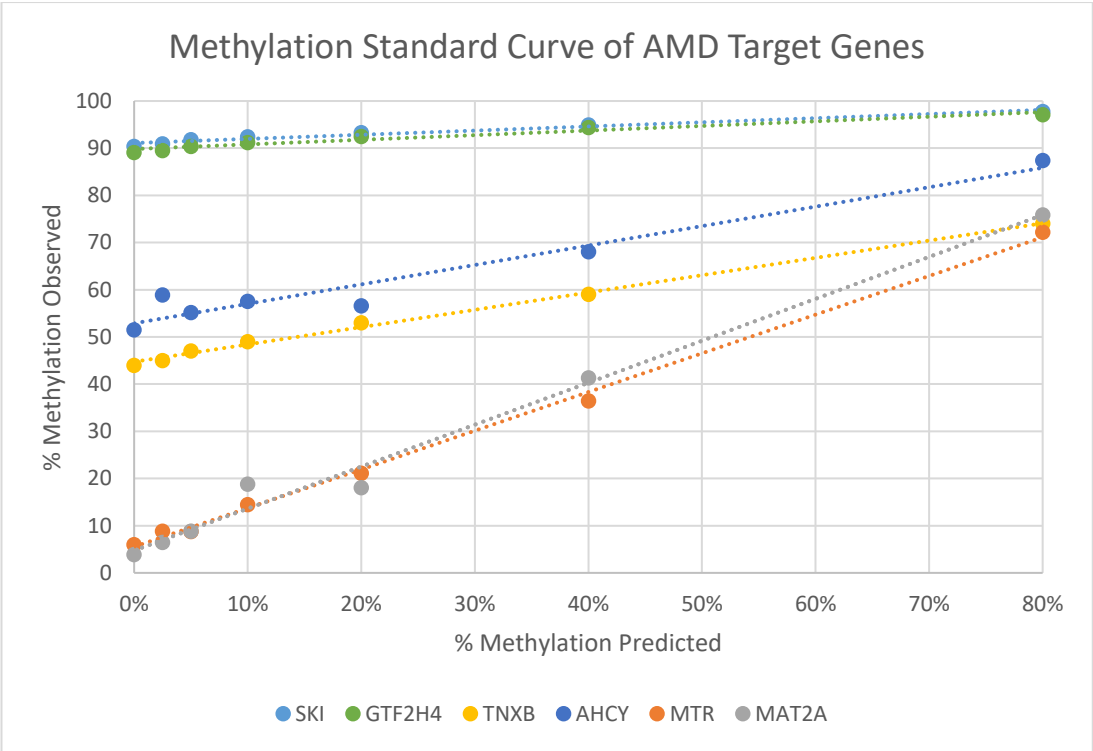

Figure 4.3: Methylation standard curves were successfully produced for each candidate gene. Methylation standard curves measure observed methylation by pyrosequencing vs predicted methylation from the proportion of enzymatically methylated DNA in the sample. The observed methylation at the 0% prediction value (no enzymatically methylated DNA) gives an indication of the baseline methylation of each CpG locus in healthy PBMC DNA.

Table 4.1: There is clear correlation between predicted and observed methylation from methylation standard curves

Gene	SKI	GTF2H4	TNXB	AHCY	MTR	MAT2A
R^2 Coefficient	0.9673	0.9544	0.9962	0.9426	0.9975	0.9878

R^2 values of ≈1 suggest that the pyrosequencing assays designed are fit for purpose, as differential methylation levels can be clearly observed at all CpG loci examined.

For each gene, the methylation standard curve produced showed a clear trend of increased observed DNA methylation levels with increased proportion of enzymatically methylated DNA present. The strong correlation observed suggests that the pyrosequencing assays and primers designed for each gene were valid and could therefore be used to study the smoking sensitivity on at select CpG regions on a larger scale using whole blood.

4.2.3 Significantly differential methylation is observed between control and AMD in genomic regions associated with *SKI* and *GTF2H4* in whole blood

The genomic regions associated with *SKI*, *GTF2H4* and *TNXB* differentially methylated in AMD RPE (Porter *et al*. 2019) were subjected to bisulphite conversion and subsequent pyrosequencing in control and AMD whole blood. Methylation levels for each pyrosequenced fragment were reported as the average value of all CpG sites encompassed within the fragment (*SKI* = 4 CpGs, *GTF2H4* = 4 CpGs, *TNXB* = 9 CpGs), however all CpG sites within each fragment were also examined independently to determine any specific differentially methylated loci.

The following results were obtained from bisulphite pyrosequencing of 96 healthy control whole blood samples, as well as 48 grade 2 (early) and 48 grade 3 (intermediate) AMD whole blood samples. Application of the Shapiro-Wilk normality test showed non-normal distribution of data, meaning subsequent statistical analyses were performed using either the Mann-Whitney test or Kruskal-Wallis test (with Dunn's multiple comparison post-test) where specified with significance level set at $p<0.05$.

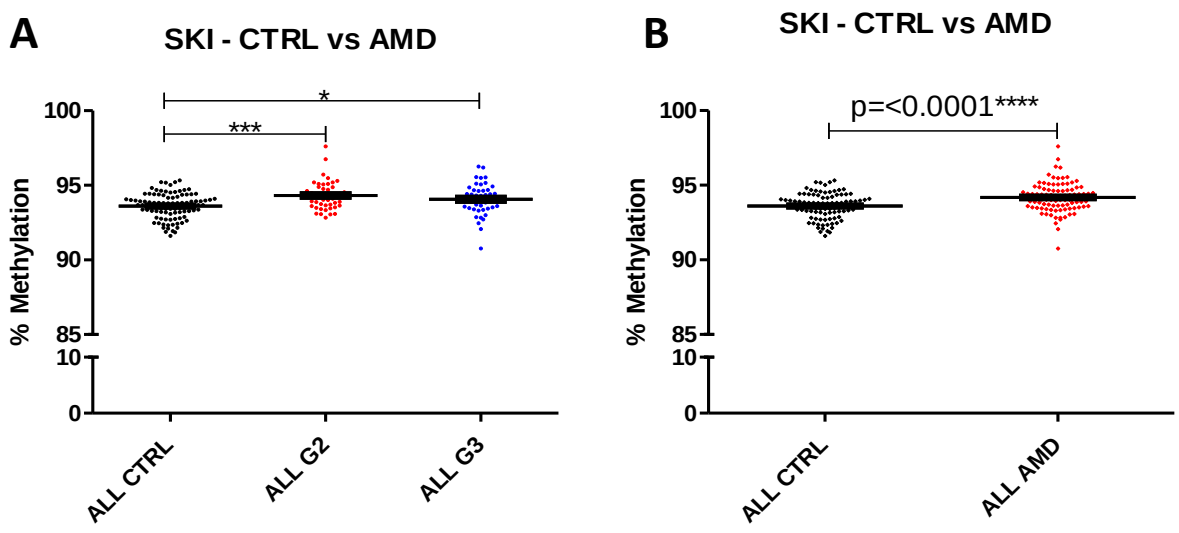

Figure 4.4: Bisulphite pyrosequencing of *SKI* in CTRL and AMD whole blood. A) Kruskal-Wallis test of 96 CTRL vs 48 G2 (grade 2 / early AMD) vs 48 G3 (grade 3 / intermediate AMD) samples. B) Mann-Whitney test of 96 CTRL vs 96 AMD samples. Main bars represent mean ± SEM. (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

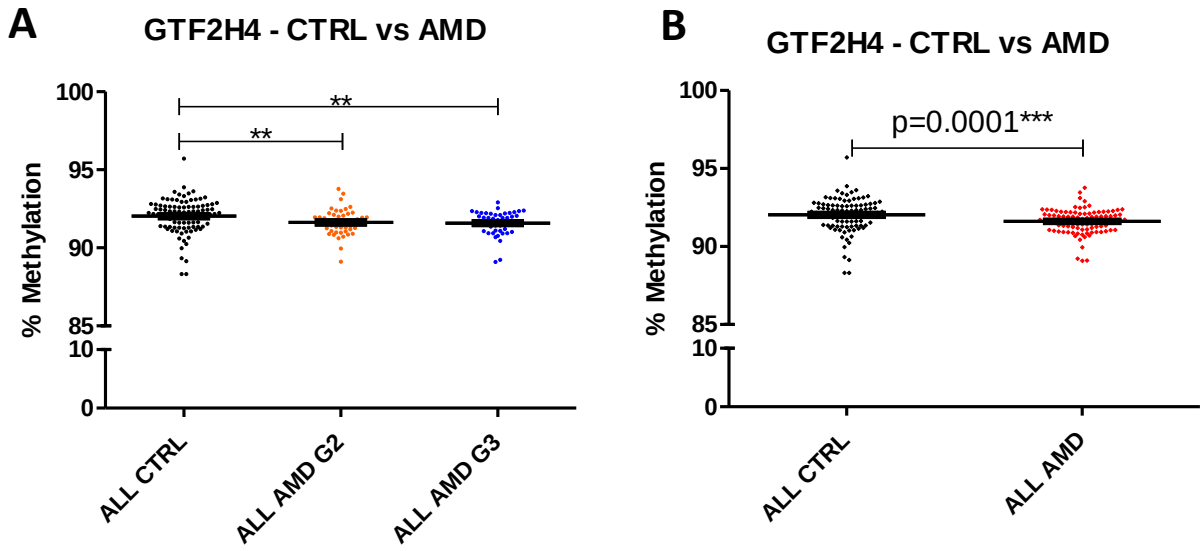

Figure 4.5: Bisulphite pyrosequencing of *GTF2H4* in CTRL and AMD whole blood. A) Kruskal-Wallis test of 96 CTRL vs 48 G2 (grade 2 / early AMD) vs 48 G3 (grade 3 / intermediate AMD) samples. B) Mann-Whitney test of 96 CTRL vs 96 AMD samples. Main bars represent mean ± SEM. (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

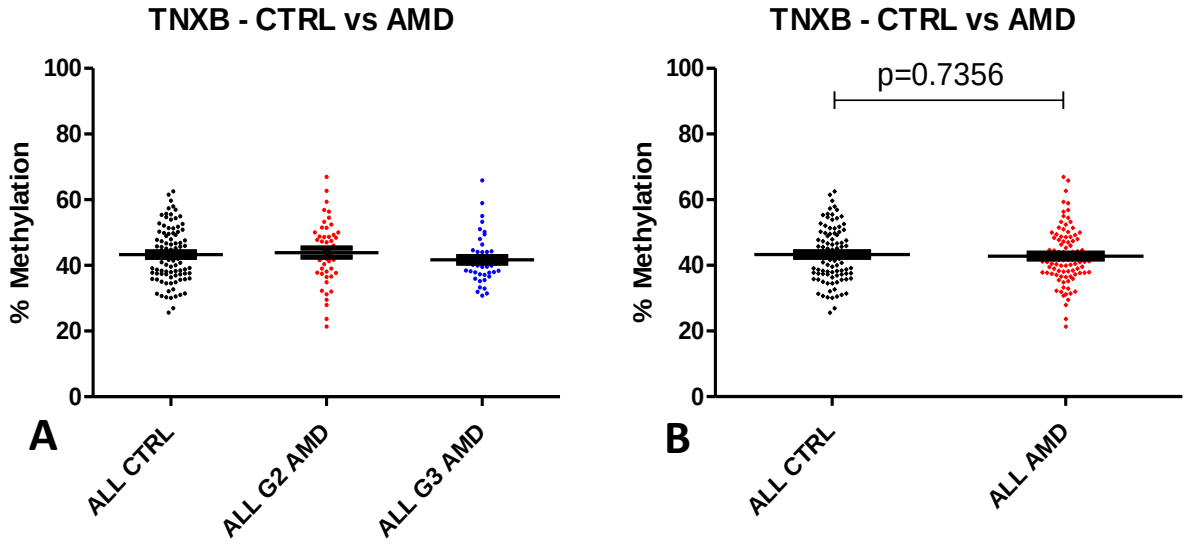

Figure 4.6: Bisulphite pyrosequencing of *TNXB* in CTRL and AMD whole blood. A) Kruskal-Wallis test of 96 CTRL vs 48 G2 (grade 2 / early AMD) vs 48 G3 (grade 3 / intermediate AMD) samples. B) Mann-Whitney test of 96 CTRL vs 96 AMD samples. Main bars represent mean ± SEM. (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

Significantly differential methylation was observed between control and collated AMD whole blood in *SKI* and *GTF2H4* (p = <0.0001, p = 0.0001 respectively), as well as between control and both grade 2 and grade 3 AMD whole blood in each case (Figures 4.4, 4.5, 4.6). In addition, for *SKI* and *GTF2H4* genomic regions, most of the differential methylation in AMD whole blood can be identified in the early phase of AMD development (grade 2), with no significant methylation changes observed between early AMD and intermediate AMD (grade 3). However, no differential methylation was observed between any two groups across the *TNXB* DMR.

These genes were also assessed for methylation changes associated with smoking status – samples were stratified by both disease state and smoking status to allow analysis of the effects of smoking on DNA methylation within *SKI*, *GTF2H4* and *TNXB* in control (86 non-smokers, 10 smokers) and AMD (84 non-smokers, 12 smokers) whole blood (Figures 4.7, 4.8, 4.9).

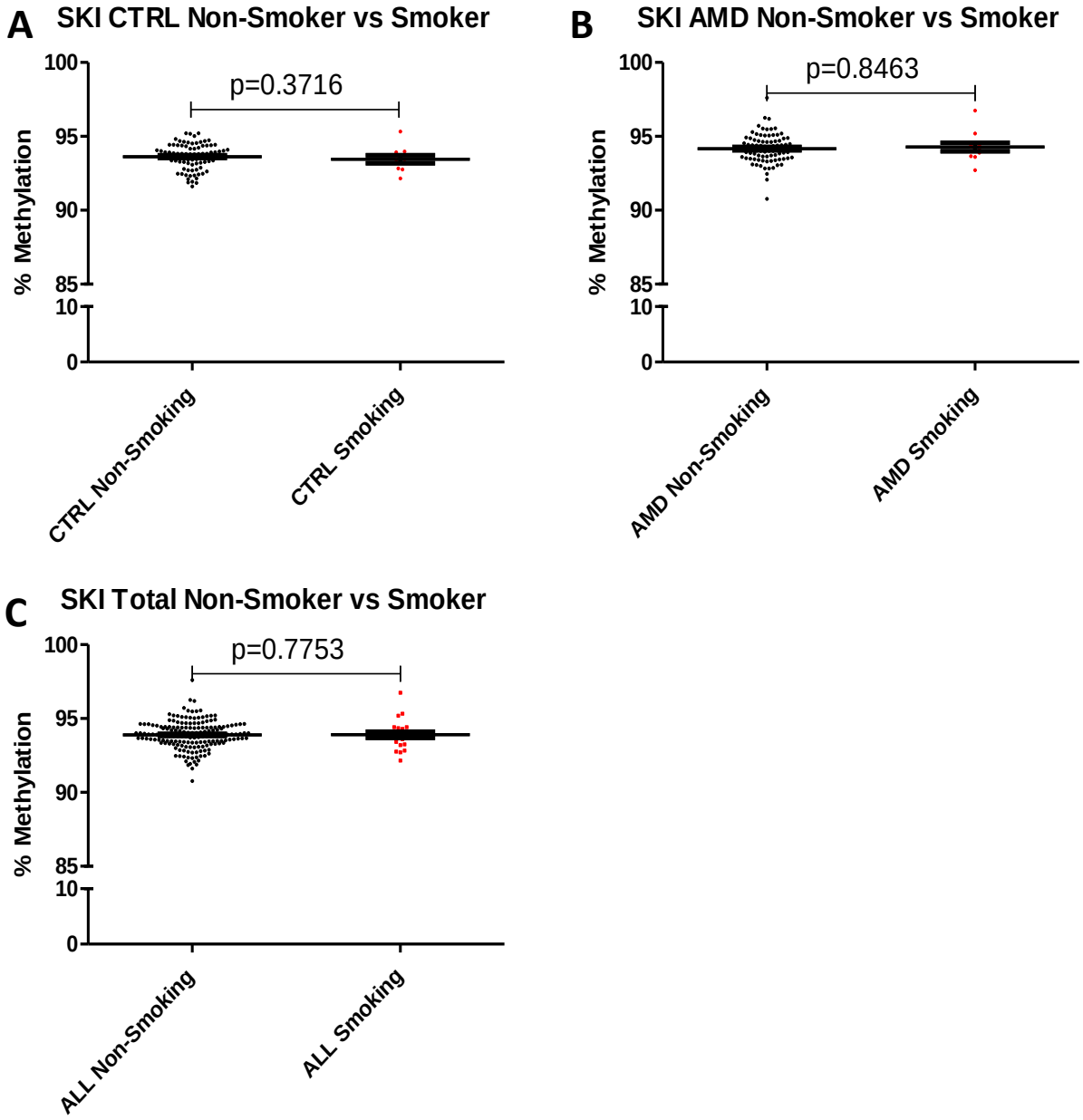

Figure 4.7: Bisulphite pyrosequencing of *SKI* in CTRL and AMD whole blood, stratified by smoking status. A) Mann-Whitney test of 86 Non-smoking CTRL vs 10 smoking CTRL samples. B) Mann-Whitney test of 84 Non-smoking AMD vs 12 smoking AMD samples. C) Mann-Whitney test of 170 combined non-smoking and 22 combined smoking samples. Main bars represent mean ± SEM. (ns = $p>0.05$, * = $p\leq0.05$, ** = $p\leq0.01$, *** = $p\leq0.001$, **** = $p<0.0001$)

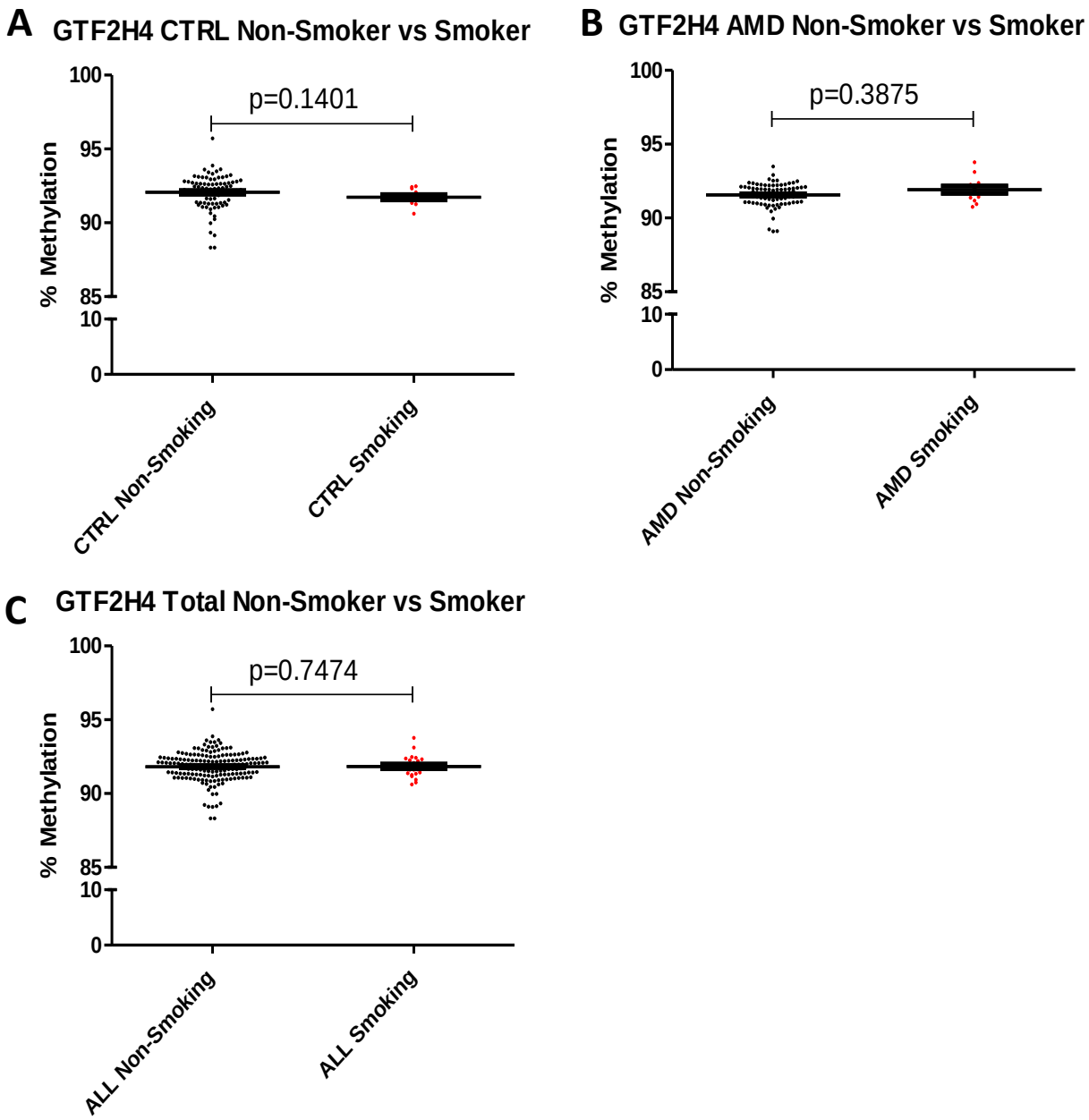

Figure 4.8: Bisulphite pyrosequencing of *GTF2H4* in CTRL and AMD whole blood, stratified by smoking status. A) Mann-Whitney test of 86 Non-smoking CTRL vs 10 smoking CTRL samples. B) Mann-Whitney test of 84 Non-smoking AMD vs 12 smoking AMD samples. C) Mann-Whitney test of 170 combined non-smoking and 22 combined smoking samples. Main bars represent mean ± SEM. (ns = $p>0.05$, * = $p\leq0.05$, ** = $p\leq0.01$, *** = $p\leq0.001$, **** = $p<0.0001$)

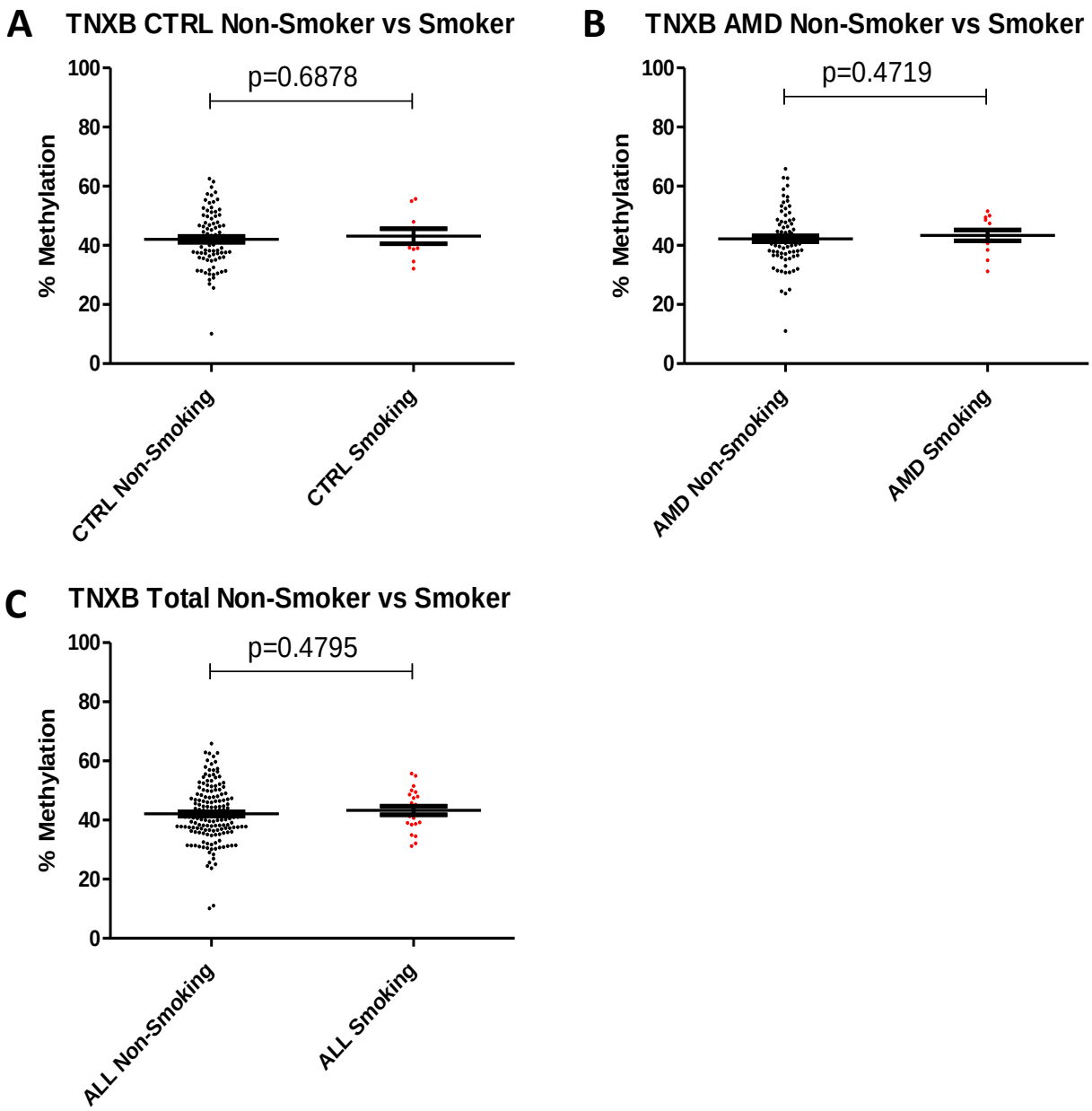

Figure 4.9: Bisulphite pyrosequencing of *TNXB* in CTRL and AMD whole blood, stratified by smoking status. A) Mann-Whitney test of 86 Non-smoking CTRL vs 10 smoking CTRL samples. B) Mann-Whitney test of 84 Non-smoking AMD vs 12 smoking AMD samples. C) Mann-Whitney test of 170 combined non-smoking and 22 combined smoking samples. Main bars represent mean ± SEM. (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

While no significantly differential methylation associated with smoking was observed for any of the above gene fragments overall, further investigation into the individual CpG sites for each pyrosequenced fragment revealed a “smoking-sensitive” locus within *SKI*. Significantly differential methylation was observed at CpG position 4 within the *SKI* fragment between control non-smokers and control smokers ($p = 0.0154$), as well as between both control smokers and AMD smokers ($p = <0.0001$) and also control non-smokers and AMD non-smokers ($p = 0.0022$). These results suggest that while smoking has an impact on DNA methylation at this locus, AMD appears to affect DNA methylation at the same locus regardless of smoking status (Figure 4.10).

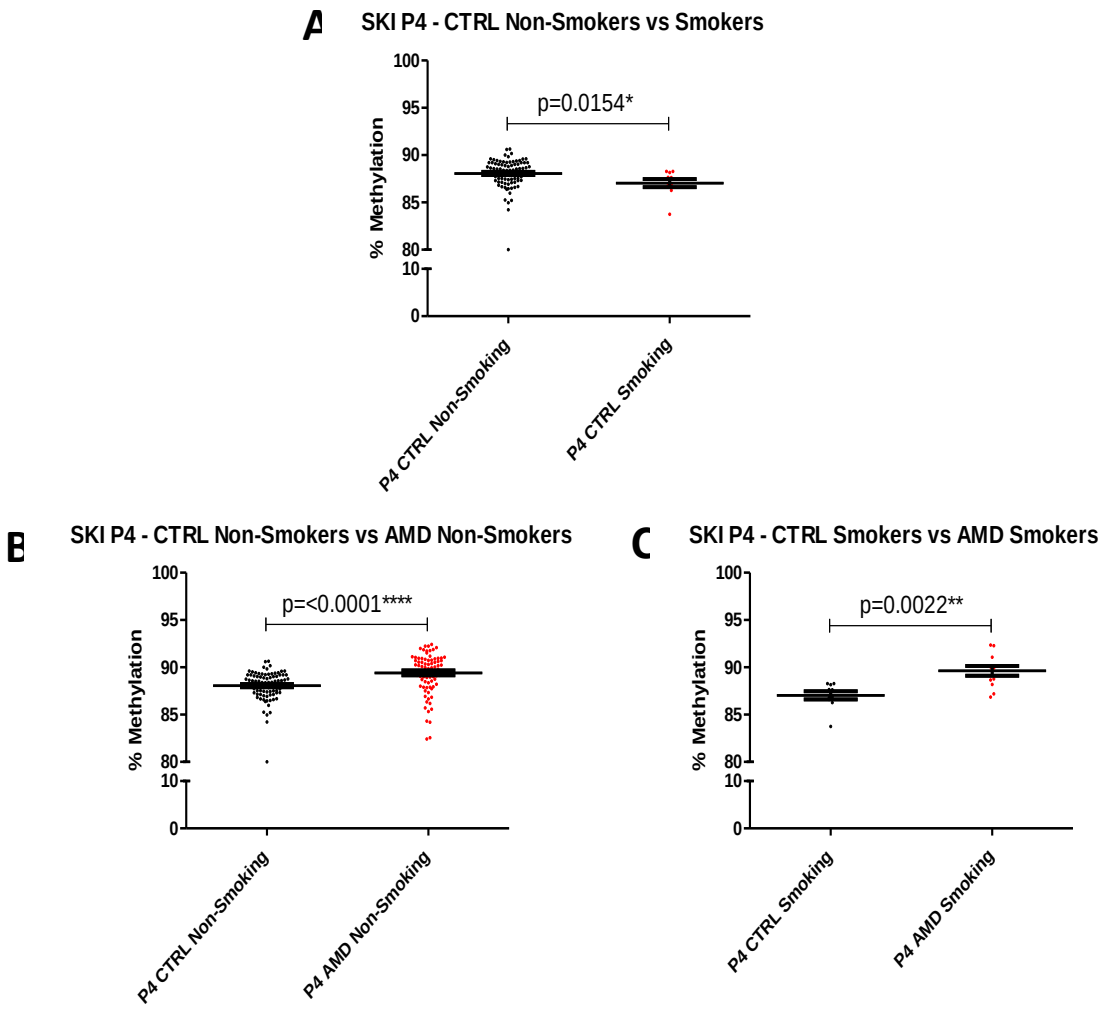

Figure 4.10: Bisulphite pyrosequencing of *SKI* CpG position 4 in CTRL and AMD whole blood, stratified by smoking status. A) Mann-Whitney test of 86 Non-smoking CTRL vs 10 smoking CTRL samples. B) Mann-Whitney test of 84 Non-smoking AMD vs 12 smoking AMD samples. C) Mann-Whitney test of 170 combined non-smoking and 22 combined smoking samples. Main bars represent mean ± SEM. (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

4.2.4 Bisulphite pyrosequencing of folate pathway genes: significant increase in methylation was observed within *MAT2A* in whole blood AMD

To supplement the above findings of significantly differential methylation of *SKI* and *GTF2H4* in AMD whole blood, bisulphite pyrosequencing was utilised to assess the methylation of the folate pathway genes *AHCY*, *MTR* and *MAT2A* using the same whole blood control and AMD patient samples. Methylation levels for each pyrosequenced fragment were again reported as the average value of all CpG sites

encompassed within the fragment (*AHCY* = 2 CpGs, *MTR* = 4 CpGs, *MAT2A* = 5 CpGs), with all CpG sites within each fragment also independently examined to determine any specific differentially methylated loci. The same whole blood samples previously used to interrogate genes identified by Porter et al. were used, with all statistical analysis performed using either the Mann-Whitney or Kruskal-Wallis tests as specified (Figures 4.11, 4.12, 4.13).

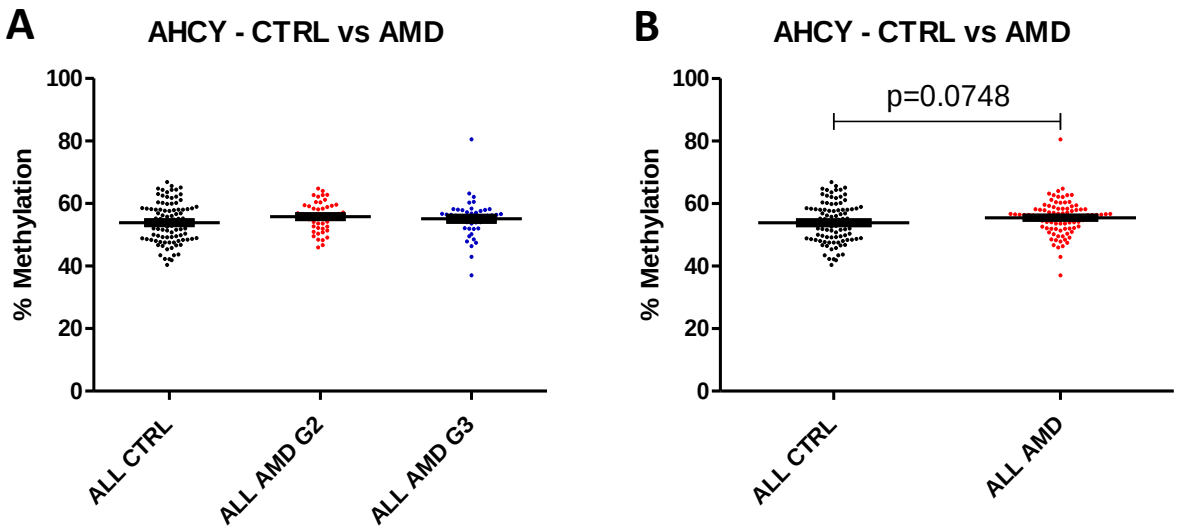

Figure 4.11: Bisulphite pyrosequencing of *AHCY* in CTRL and AMD whole blood. A) Kruskal-Wallis test of 96 CTRL vs 48 G2 (grade 2 / early AMD) vs 48 G3 (grade 3 / intermediate AMD) samples. B) Mann-Whitney test of 96 CTRL vs 96 AMD samples. Main bars represent mean ± SEM. (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

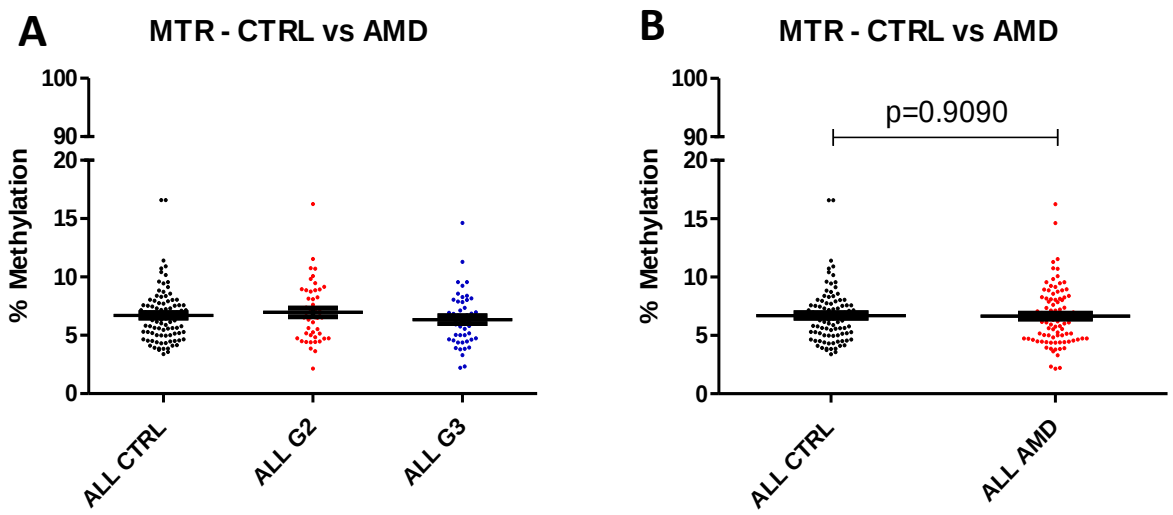

Figure 4.12: Bisulphite pyrosequencing of *MTR* in CTRL and AMD whole blood. A) Kruskal-Wallis test of 96 CTRL vs 48 G2(grade 2 / early AMD) vs 48 G3 (grade 3 / intermediate AMD) samples. B) Mann-Whitney test of 96 CTRL vs 96 AMD samples. Main bars represent mean ± SEM. (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

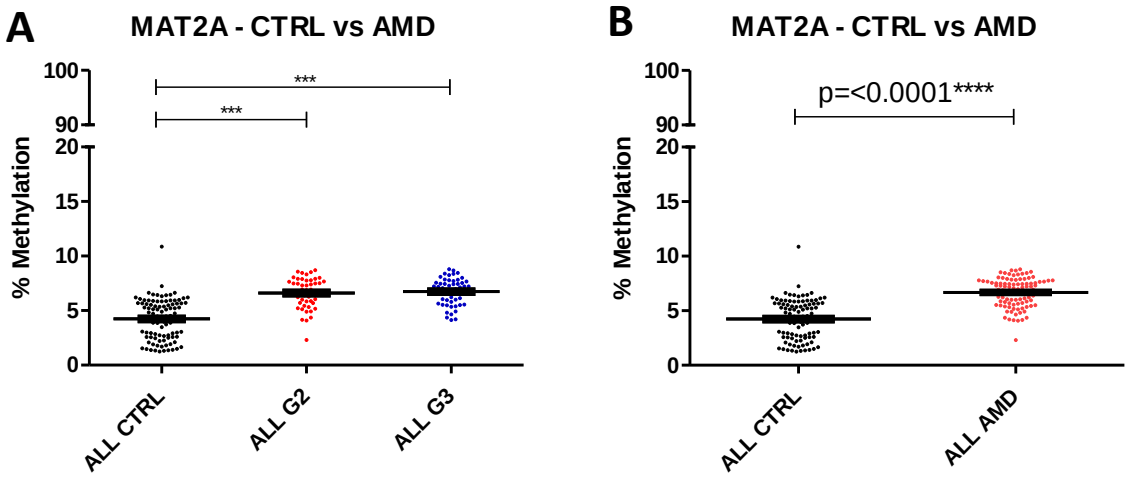

Figure 4.13: Bisulphite pyrosequencing of *MAT2A* in CTRL and AMD whole blood. A) Kruskal-Wallis test of 96 CTRL vs 48 G2(grade 2 / early AMD) vs 48 G3 (grade 3 / intermediate AMD) samples. B) Mann-Whitney test of 96 CTRL vs 96 AMD samples. Main bars represent mean ± SEM. (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

While no significantly differential methylation was observed between control and AMD whole blood in either *AHCY* or *MTR*, a significant increase in methylation was discovered within *MAT2A* in whole blood AMD ($p = <0.0001$). As noted for *SKI* and *GTF2H4*, no significant difference in methylation was discovered between G2 and G3 AMD samples within *MAT2A*, with most of the differential methylation in AMD whole blood identified in the early phase of AMD development (grade 2).

The pyrosequencing fragments surrounding the CpG sites of interest within *AHCY*, *MTR* and *MAT2A* were also assessed to identify differential methylation in smokers using the same whole blood samples. While no significant difference in methylation was observed between smokers and non-smokers in the pyrosequenced regions of *AHCY* and *MTR*, pyrosequencing of the *MAT2A* fragment revealed a general trend of increased methylation in smokers, with a significant increase in methylation observed in control smokers compared to control non-smokers ($p = 0.0168$) (Figures 4.14, 4.15, 4.16).

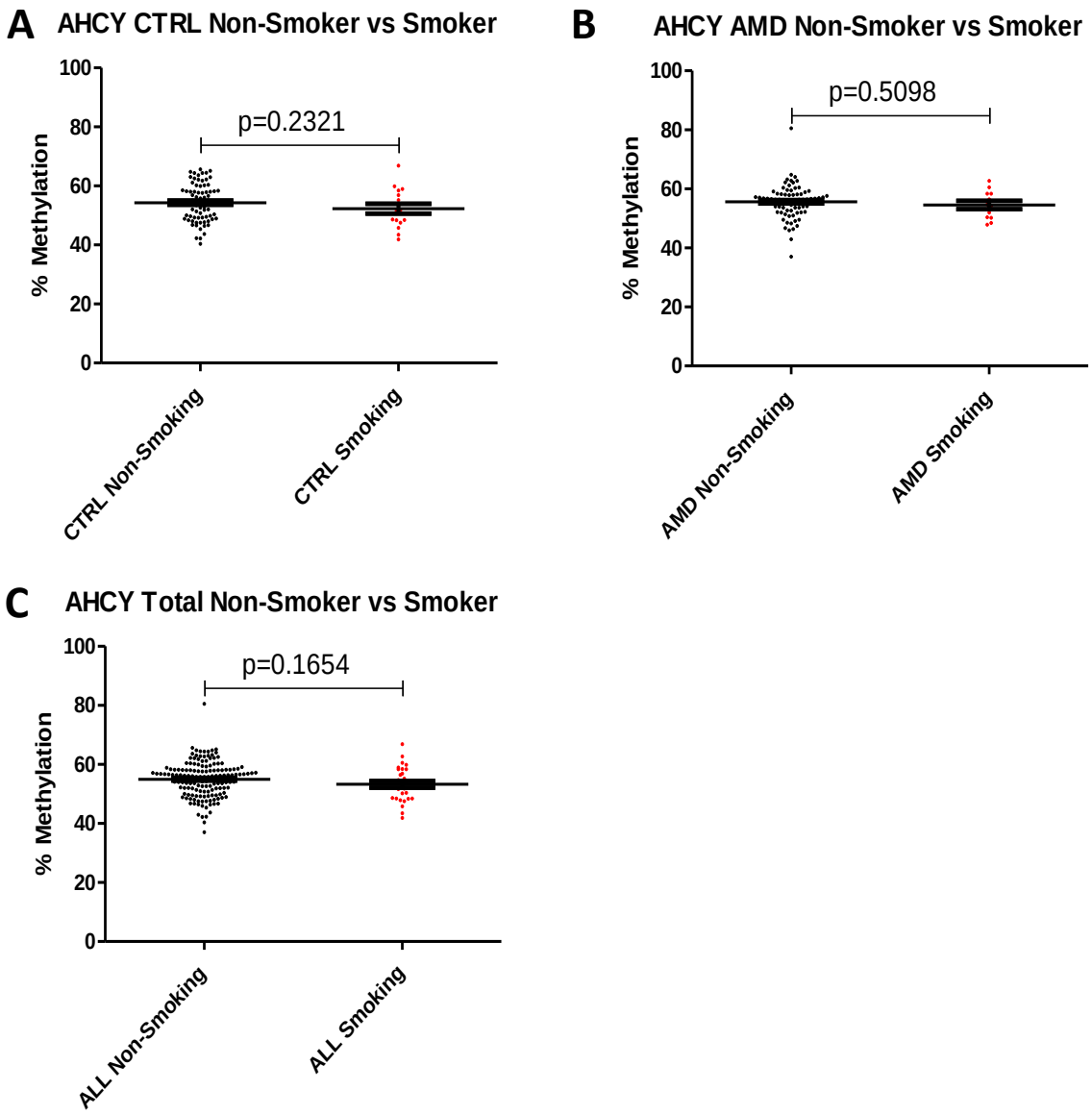

Figure 4.14: Bisulphite pyrosequencing of *AHCY* in CTRL and AMD whole blood, stratified by smoking status. A) Mann-Whitney test of 86 Non-smoking CTRL vs 10 smoking CTRL samples. B) Mann-Whitney test of 84 Non-smoking AMD vs 12 smoking AMD samples. C) Mann-Whitney test of 170 combined non-smoking and 22 combined smoking samples. Main bars represent mean ± SEM. (ns = $p>0.05$, * = $p\leq0.05$, ** = $p\leq0.01$, *** = $p\leq0.001$, **** = $p<0.0001$)

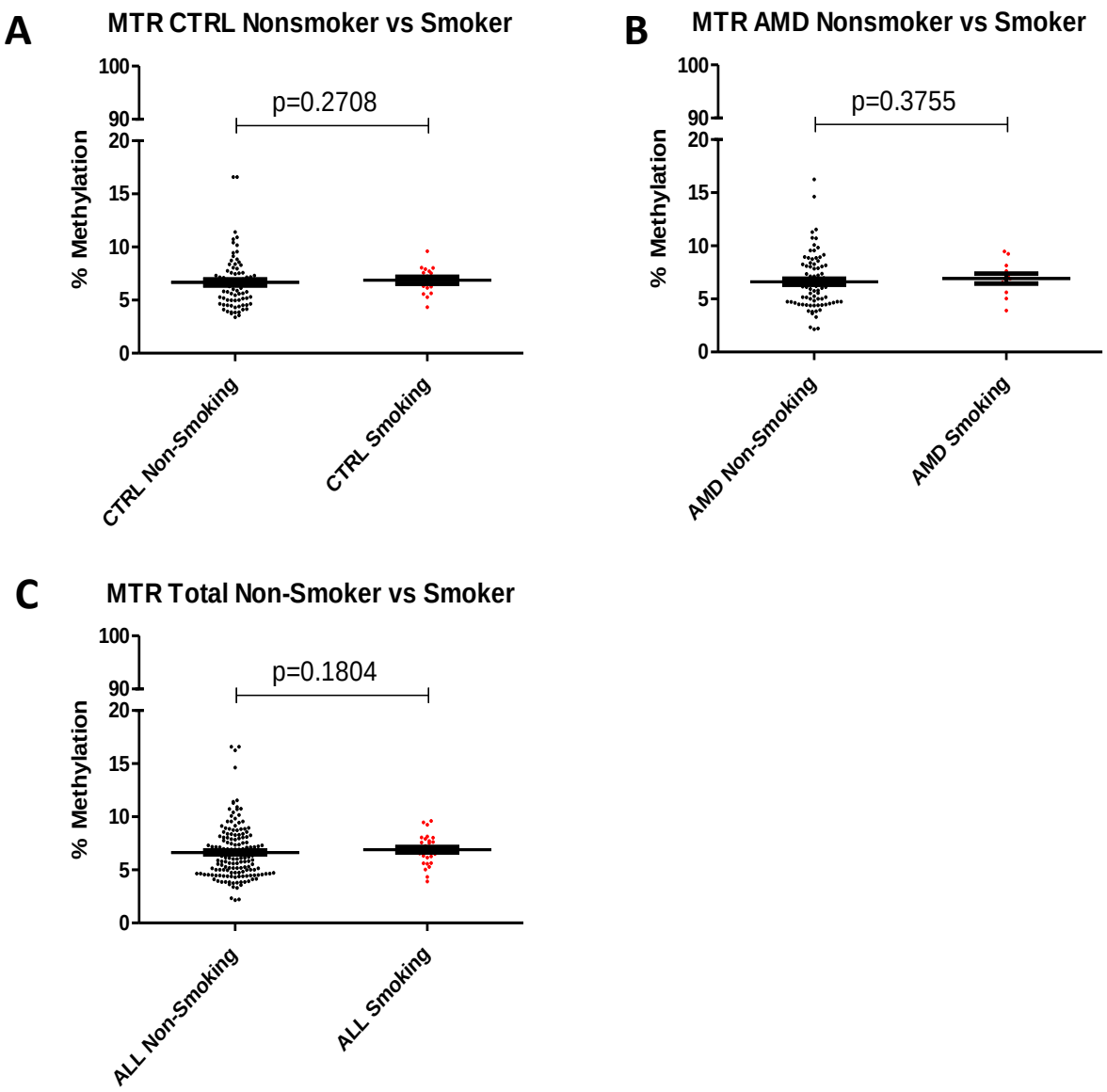

Figure 4.15: Bisulphite pyrosequencing of *MTR* in CTRL and AMD whole blood, stratified by smoking status. A) Mann-Whitney test of 86 Non-smoking CTRL vs 10 smoking CTRL samples. B) Mann-Whitney test of 84 Non-smoking AMD vs 12 smoking AMD samples. C) Mann-Whitney test of 170 combined non-smoking and 22 combined smoking samples. Main bars represent mean ± SEM. (ns = $p>0.05$, * = $p\leq0.05$, ** = $p\leq0.01$, *** = $p\leq0.001$, **** = $p<0.0001$)

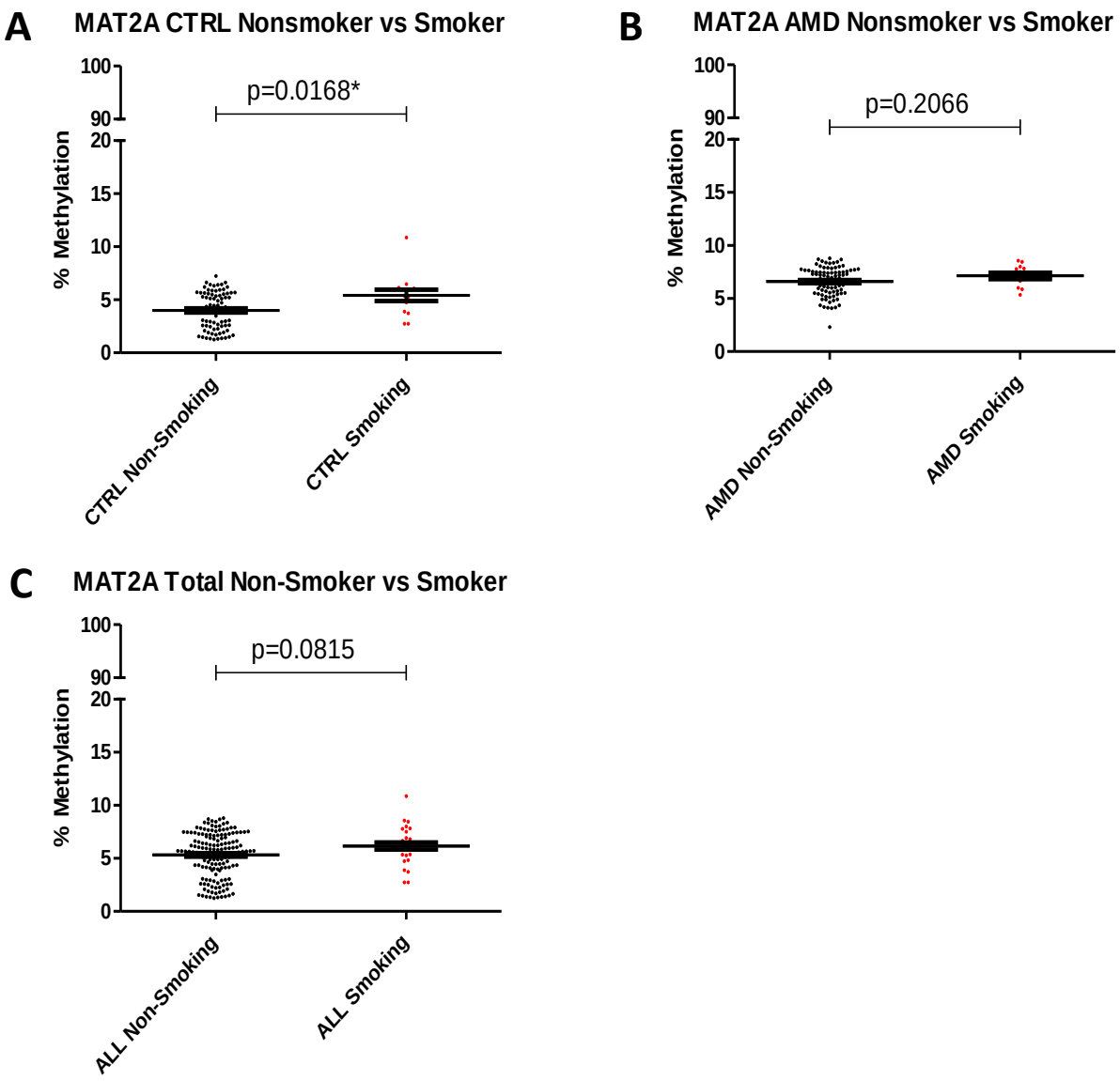

Figure 4.16: Bisulphite pyrosequencing of *MAT2A* in CTRL and AMD whole blood, stratified by smoking status. A) Mann-Whitney test of 86 Non-smoking CTRL vs 10 smoking CTRL samples. B) Mann-Whitney test of 84 Non-smoking AMD vs 12 smoking AMD samples. C) Mann-Whitney test of 170 combined non-smoking and 22 combined smoking samples. Main bars represent mean ± SEM. (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

4.2.5 Validation cohort studies show differential methylation of *SKI* in AMD whole blood

To validate differential methylation in AMD whole blood in *SKI, GTF2H4* and *MAT2A* and further address an impact of smoking on whole blood methylation at *MAT2A* a separate cohort of control and AMD patient whole blood DNA samples were obtained from colleagues at the University of Southampton. While significantly differential methylation was observed between smokers and non-smokers within *SKI* and *MAT2A*, results were hampered by low number of smokers in the discovery cohort. As such, the Southampton validation cohort was balanced for smoking status consisting of 48 healthy controls (24 smokers, 24 non-smokers), 23 G2 AMD (11 smokers, 12 non-smokers) and 25 G3 AMD (13 smokers, 12 non-smokers). As the pyrosequenced regions within *TNXB*, *AHCY* and *MTR* revealed no significantly differential methylation in the original whole blood cohort between either control and AMD patients or smokers and non-smokers, these genes were not studied further. The 96 validation cohort samples were then subjected to bisulphite conversion and pyrosequencing at the regions of interest within *SKI*, *GTF2H4* and *MAT2A* as before (Figures 4.17 to 4.22).

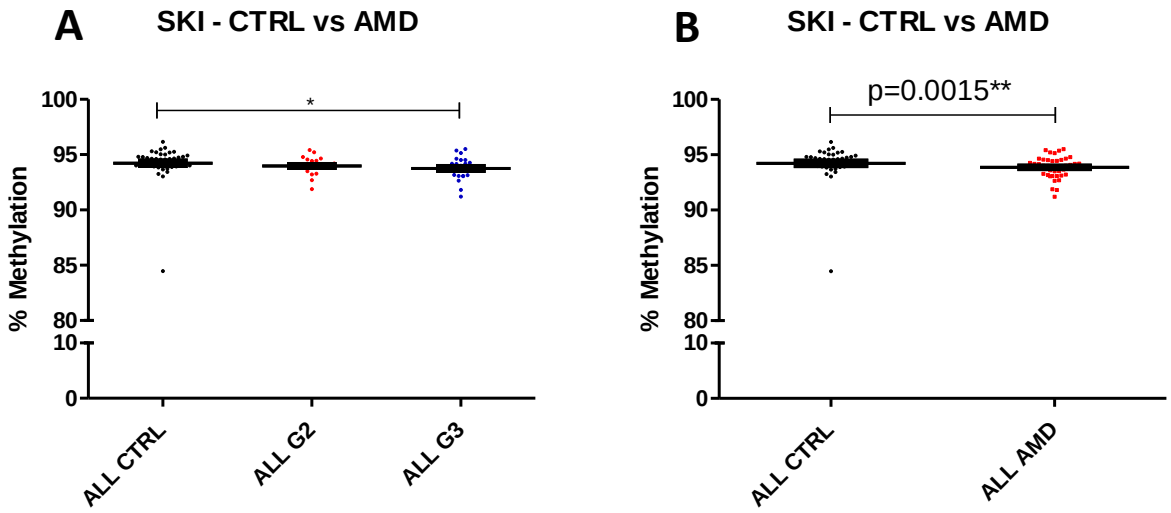

Figure 4.17: Bisulphite pyrosequencing validation of *SKI* in CTRL and AMD whole blood. A) Kruskal-Wallis test of 96 CTRL vs 48 G2(grade 2 / early AMD) vs 48 G3 (grade 3 / intermediate AMD) samples. B) Mann-Whitney test of 96 CTRL vs 96 AMD samples. Main bars represent mean ± SEM. (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

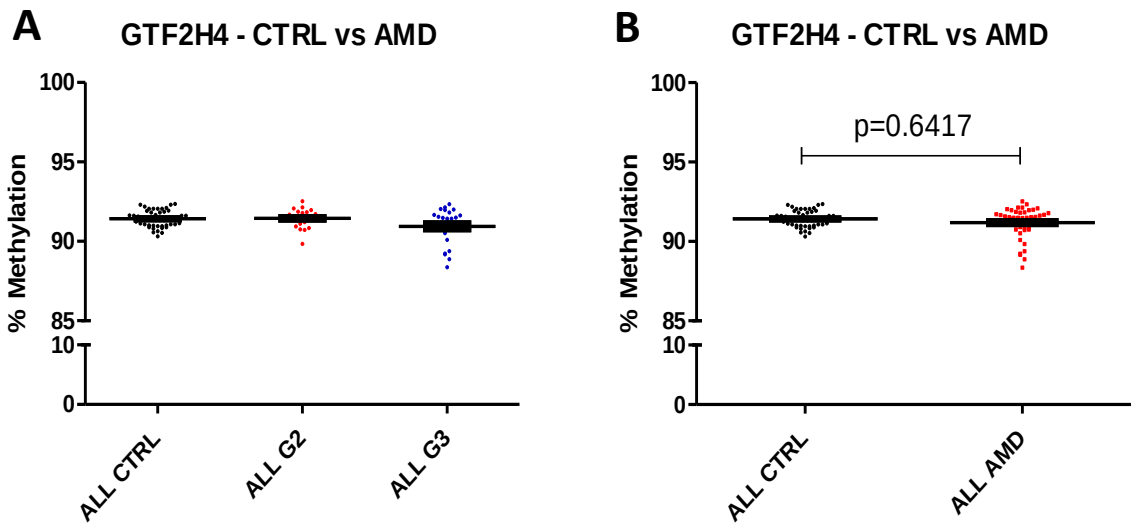

Figure 4.18: Bisulphite pyrosequencing validation of *GTF2H4* in CTRL and AMD whole blood. A) Kruskal-Wallis test of 96 CTRL vs 48 G2(grade 2 / early AMD) vs 48 G3 (grade 3 / intermediate AMD) samples. B) Mann-Whitney test of 96 CTRL vs 96 AMD samples. Main bars represent mean ± SEM. (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

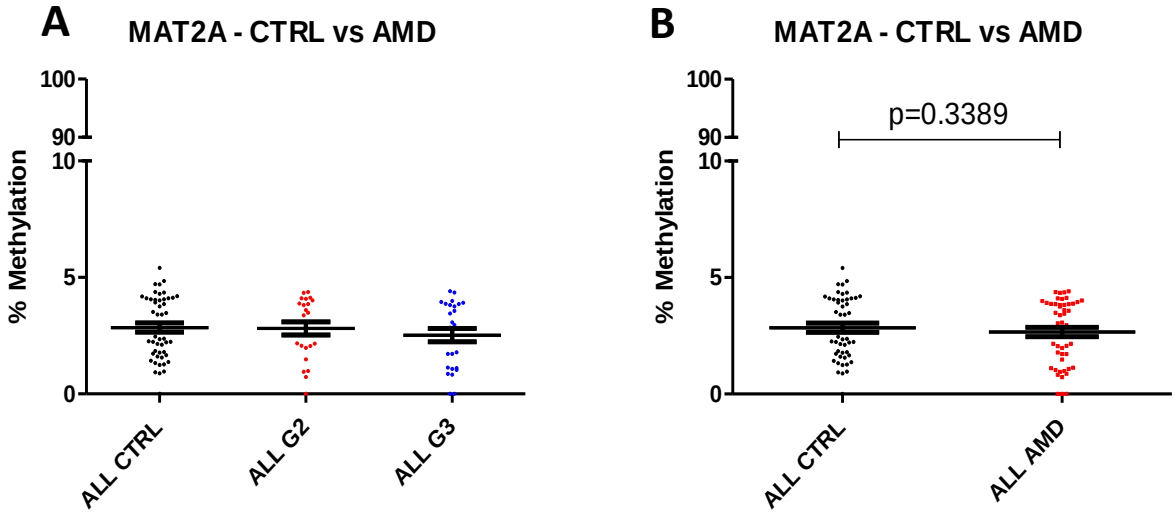

Figure 4.19: Bisulphite pyrosequencing validation of *MAT2A* in CTRL and AMD whole blood. A) Kruskal-Wallis test of 96 CTRL vs 48 G2(grade 2 / early AMD) vs 48 G3 (grade 3 / intermediate AMD) samples. B) Mann-Whitney test of 96 CTRL vs 96 AMD samples. Main bars represent mean ± SEM. (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

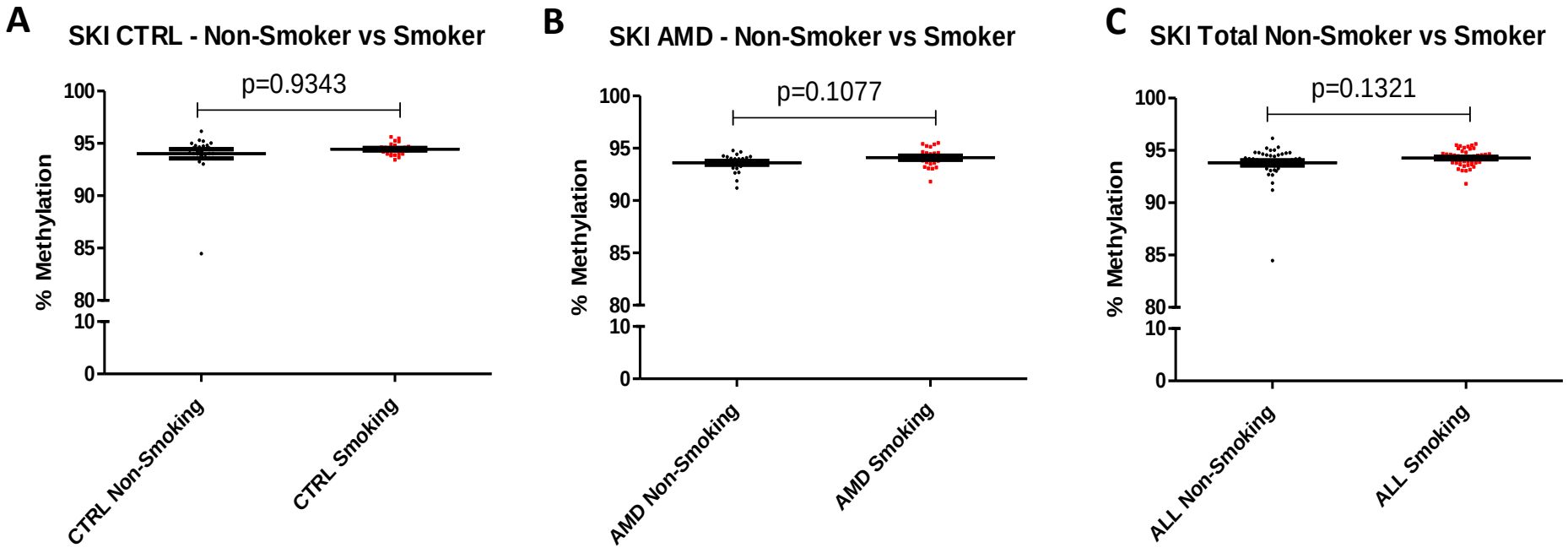

Figure 4.20: Bisulphite pyrosequencing validation of *SKI* in CTRL and AMD whole blood, stratified by smoking status. A) Mann-Whitney test of 24 Non-smoking CTRL vs 24 smoking CTRL samples. B) Mann-Whitney test of 24 Non-smoking AMD vs 24 smoking AMD samples. C) Mann-Whitney test of 48 combined non-smoking and 48 combined smoking samples. Main bars represent mean ± SEM. (ns = $p>0.05$, * = $p\leq0.05$, ** = $p\leq0.01$, *** = $p\leq0.001$, **** = $p<0.0001$)

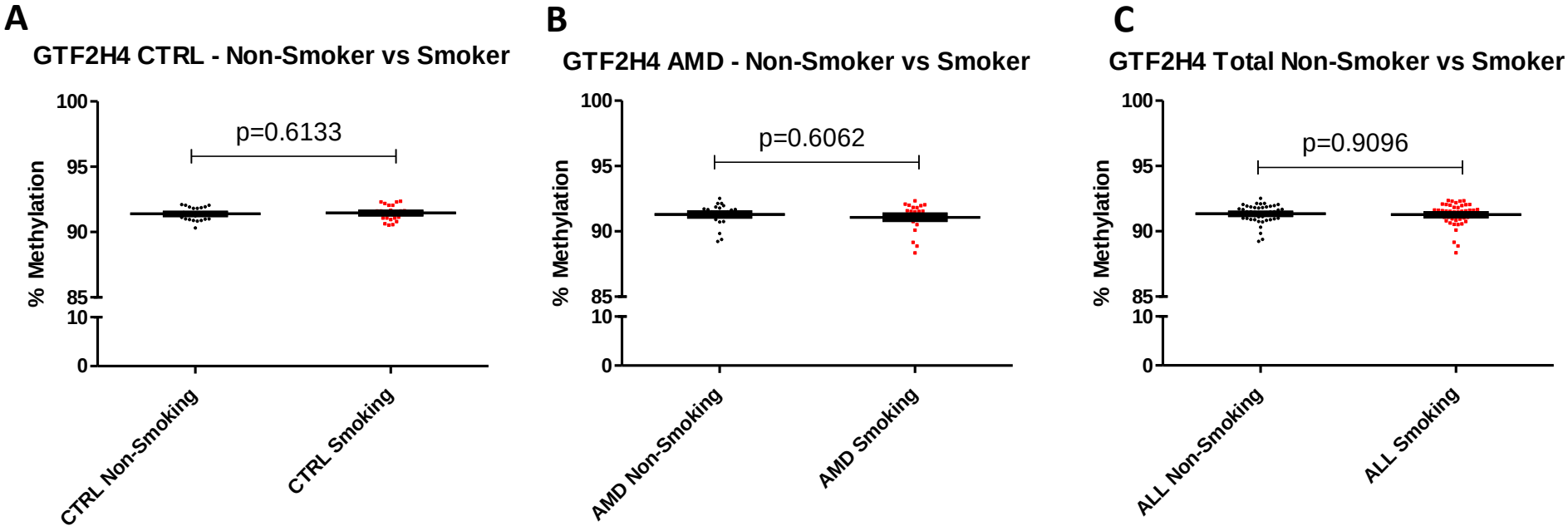

Figure 4.21: Bisulphite pyrosequencing validation of *GTF2H4* in CTRL and AMD whole blood, stratified by smoking status. A) Mann-Whitney test of 24 Non-smoking CTRL vs 24 smoking CTRL samples. B) Mann-Whitney test of 24 Non-smoking AMD vs 24 smoking AMD samples. C) Mann-Whitney test of 48 combined non-smoking and 48 combined smoking samples. Main bars represent mean ± SEM. (ns = $p>0.05$, * = $p\leq0.05$, ** = $p\leq0.01$, *** = $p\leq0.001$, **** = $p<0.0001$)

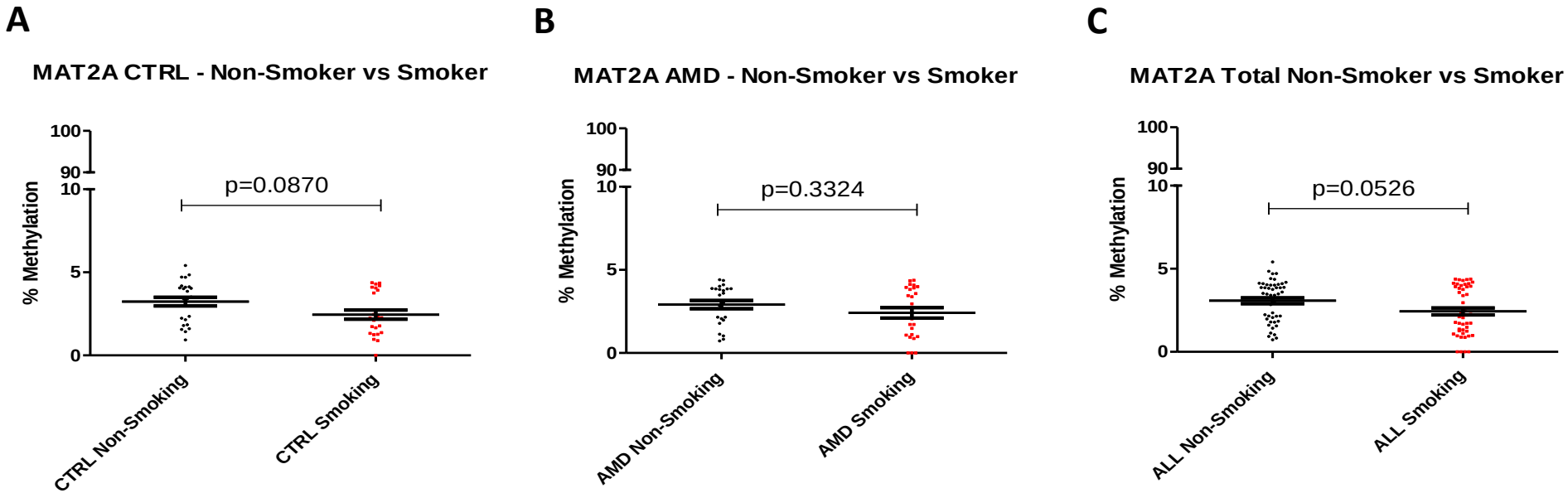

Figure 4.22: Bisulphite pyrosequencing validation of *MAT2A* in CTRL and AMD whole blood, stratified by smoking status. A) Mann-Whitney test of 24 Non-smoking CTRL vs 24 smoking CTRL samples. B) Mann-Whitney test of 24 Non-smoking AMD vs 24 smoking AMD samples. C) Mann-Whitney test of 48 combined non-smoking and 48 combined smoking samples. Main bars represent mean ± SEM. (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

Bisulphite pyrosequencing of the validation cohort revealed little differential methylation between control and AMD patient whole blood samples, with no significant difference in methylation within any genes between smokers and non-smokers in either the control or AMD sample groups. A decrease in AMD methylation within the *SKI* pyrosequencing fragment was observed (p = 0.0015), however the change in raw methylation was only slight.

4.2.6 Combined cohort analysis of *SKI*, *GTF2H4* and *MAT2A* reveals differential methylation in *GTF2H4* and *MAT2A* methylation (p = <0.0001) in AMD whole blood and a significant effect of smoking status on methylation levels within *SKI, GTF2H4* and *MAT2A*

Due to the observation of largely similar methylation levels within each gene across both sample cohorts, and given that both cohorts were of European origin, the decision was made to combine the results of both cohorts to see if any significantly differential methylation could be observed due to either disease or smoking status across a larger sample size. The resulting combined cohort comprised 144 control (34 smokers, 110 non-smokers) and 144 AMD (36 smokers, 108 non-smokers) whole blood samples. This cohort was reanalysed and stratified for both disease and smoking status per gene (Figures 4.23 to 4.26).

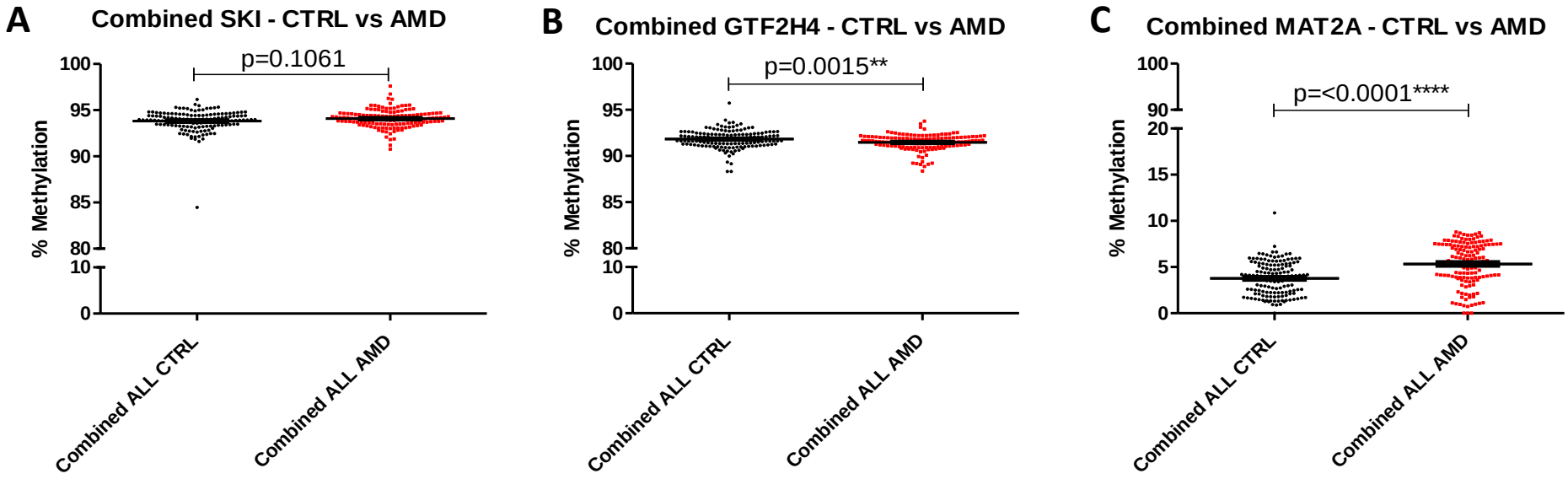

Figure 4.23: Combined cohort bisulphite pyrosequencing results in CTRL and AMD whole blood. A) Mann-Whitney test of 144 CTRL vs 144 AMD samples in *SKI*. B) Mann-Whitney test of 144 CTRL vs 144 AMD samples in *GTF2H4* C) Mann-Whitney test of 144 CTRL vs 144 AMD samples in *MAT2A*. Main bars represent mean ± SEM. (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

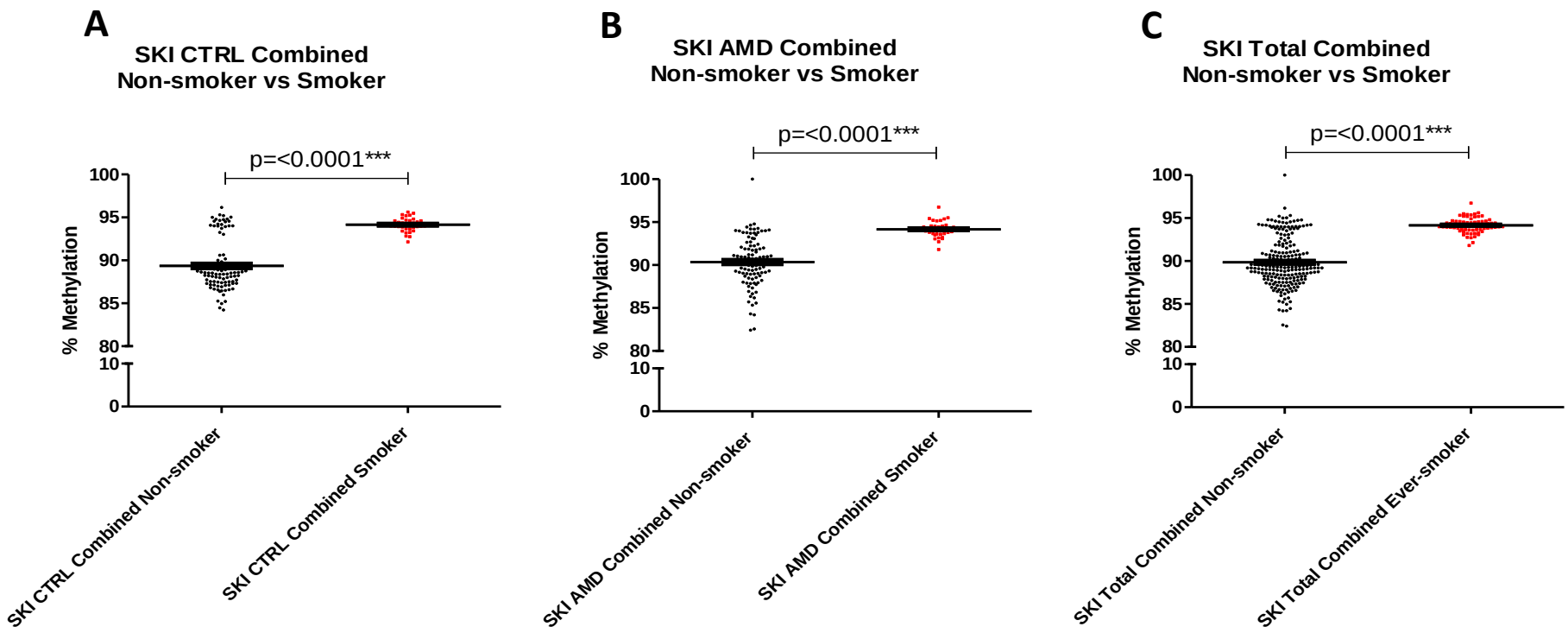

Figure 4.24: Combined cohort bisulphite pyrosequencing results of *SKI* in CTRL and AMD whole blood, stratified by smoking status. A) Mann-Whitney test of 110 Non-smoking CTRL vs 34 smoking CTRL samples. B) Mann-Whitney test of 108 Non-smoking AMD vs 36 smoking AMD samples. C) Mann-Whitney test of 218 combined non-smoking and 70 combined smoking samples. Main bars represent mean ± SEM. (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

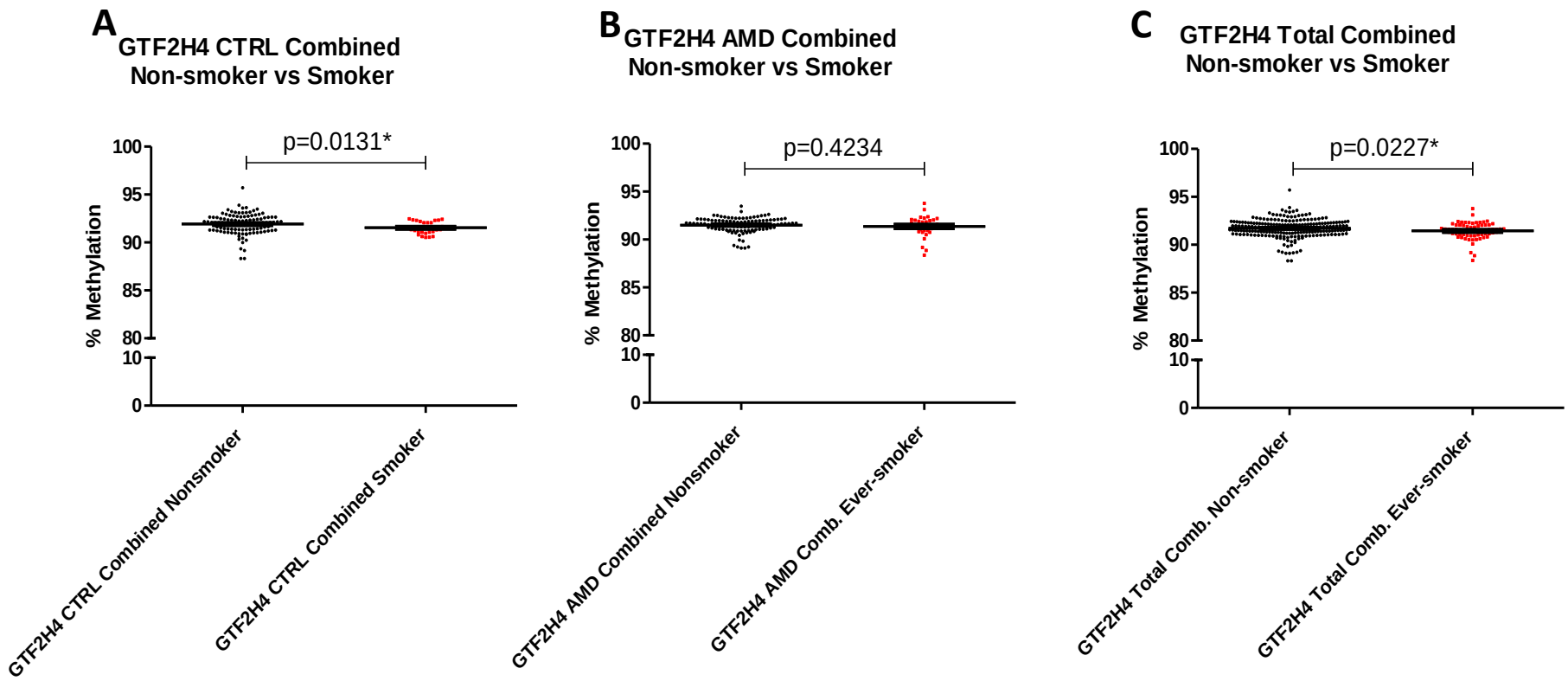

Figure 4.25: Combined cohort bisulphite pyrosequencing results of *GTF2H4* in CTRL and AMD whole blood, stratified by smoking status. A) Mann-Whitney test of 110 Non-smoking CTRL vs 34 smoking CTRL samples. B) Mann-Whitney test of 108 Non-smoking AMD vs 36 smoking AMD samples. C) Mann-Whitney test of 218 combined non-smoking and 70 combined smoking samples. Main bars represent mean ± SEM. (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

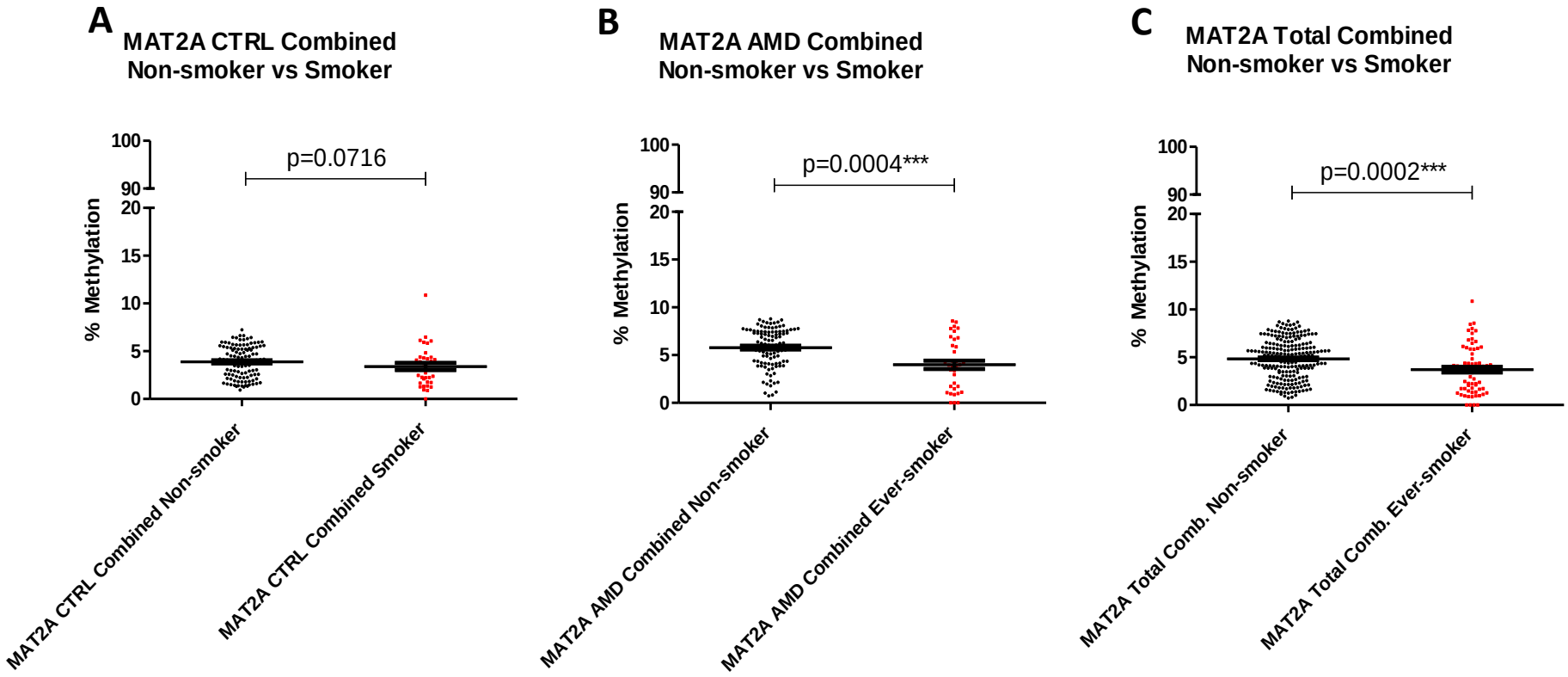

Figure 4.26: Combined cohort bisulphite pyrosequencing results of *MAT2A* in CTRL and AMD whole blood, stratified by smoking status. A) Mann-Whitney test of 110 Non-smoking CTRL vs 34 smoking CTRL samples. B) Mann-Whitney test of 108 Non-smoking AMD vs 36 smoking AMD samples. C) Mann-Whitney test of 218 combined non-smoking and 70 combined smoking samples. Main bars represent mean ± SEM. (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p<0.0001)

Analysis of the combined cohort revealed several significant regions of differential methylation. While a slight non-significant increase in methylation was observed in *SKI* in AMD ($p = 0.1061$), significant decreases in *GTF2H4* methylation ($p = 0.0015$) and increases in *MAT2A* methylation ($p = <0.0001$) were observed in AMD whole blood. All 3 genes displayed the same direction of methylation change in the combined cohort as was reported in the initial Netherlands cohort.

Furthermore, significant differential methylation can be observed within each gene for at least one subgroup when stratified by smoking status. The combined cohort analysis suggests that smoking has a significant effect increasing methylation levels within the intron 1 region of *SKI* ($p = <0.0001$), and a significant effect lowering methylation levels within *GTF2H4* ($p = 0.0227$) and *MAT2A* ($p = 0.0002$). In the case of *SKI*, the lack of significant differential methylation in AMD compared to control samples suggests that smoking had a much greater effect on determining levels of methylation in at this genomic region in whole blood than disease state.

Retrospective analysis of the power levels of the data obtained in this combined cohort was performed using a standardised alpha level of 0.05 while evaluating the mean and standard deviations of each population in each case – when evaluating CTRL vs AMD methylation differences in the combined cohort using 144 samples in each population, the calculated power levels of the results in *SKI*, *GTF2H4* and *MAT2A* were 57.1%, 90.7% and 100% respectively. For combined Non-smokers (218 samples) vs smokers (70 samples), the calculated power levels of the results in *SKI*, *GTF2H4* and *MAT2A* were 100%, 57.1% and 93.6% respectively. To achieve a power level of 80% for future studies in cases where the power was observed to be below 80%, it was calculated that approximately double the number of samples would be required in each case while maintaining the weighting used in this study (294 CTRL vs 294 AMD for *SKI*, 447 non-smokers vs 143 smokers for *GTF2H4*).

4.3 Discussion for chapter 4

Firstly, I justified the use of whole blood DNA as a proxy for interrogating DNA methylation levels in AMD at select genomic regions of interest. Secondly, after removal of DNA samples showing signs of DNA degradation (204), the resulting study cohort of 96 control and 96 AMD patient samples exemplifies the major advantage of whole blood use for AMD epigenetics study over RPE tissue – the ease of obtaining greater number of samples for analysis. With the constraints of the COVID-19 pandemic, the necessity to consider easily accessible yet validated surrogate tissues for DNA methylation studies in AMD is heightened, as human ocular tissue collections have been halted in many locations.

Generation of methylation standard curves by bisulphite pyrosequencing suggested that the assays designed could be used to interrogate genomic regions of interest. Baseline methylation levels observed in healthy PBMC DNA in each case, and the similarity between this baseline and the methylation values observed for healthy control whole blood samples in the pyrosequencing studies for each gene, suggested whole blood samples could be used as a surrogate for RPE samples at the target loci investigated.

Bisulphite pyrosequencing in whole blood samples of AMD and control patients of AMD-associated genomic regions within *SKI*, *GTF2H4* and *TNXB* (Porter *et al*. 2019) revealed significant methylation differences within *SKI* and *GTF2H4* ($p = <0.0001$, $p = 0.0001$ respectively). Interestingly, no significant difference in methylation was found between early (G2) and intermediate (G3) AMD whole blood samples for any of the six genomic fragments assessed. Most of the difference in methylation observed occurred between control and early AMD suggesting that DNA methylation changes at these loci appear at the early stages of AMD development.

In contrast to the findings by Porter *et al*. in RPE, no significant difference in DNA methylation was observed within *TNXB* between control and AMD whole blood samples ($p = 0.7356$). However, the Porter *et al*. study did not include bisulphite pyrosequencing validation of the differential RPE methylation at the *TNXB* DMR, observed from the Illumina 450K methylation array. While methylation arrays such

as the Illumina 450K array can interrogate a vast number of CpG sites in one experiment, bisulphite pyrosequencing relies on the design of valid primers to isolate target DNA fragments, including multiple CpG loci. Pyrosequencing of the *TNXB* fragment involved the inclusion of nine CpG loci, resulting in a longer fragment than used for the assessment of *SKI* and *GTF2H4,* which included four loci each. Pyrosequencing is designed for the assessment of smaller DNA fragments. On longer pyrosequencing runs, smaller visible light peak signals are emitted with efficient visualisation deteriorating beyond approximately 70bp (205), resulting in potential inconsistencies in quantification of DNA methylation observed as longer pyrosequencing runs progress (206, 207). In the case of the fragment containing the *TNXB* DMR, signal strength was found to drop towards the end of the pyrosequencing run – this could possibly be alleviated in future by the implementation of two unique, shorter yet overlapping pyrosequencing fragments.

The observation of significantly differential methylation in AMD whole blood within *SKI* and *GTF2H4* reinforce the proposed idea from Porter *et al*. of differential methylation at these genes playing a potential role in early and intermediate AMD. The increase in *SKI* intron 1 methylation in AMD was of particular interest given the recent observation that methylation levels within the first intron of genes may be inversely proportional to gene expression perhaps due to the proximity of the first intron to the promoter region in some cases. The study by Anastasiadi *et al* (2018) describes the identification of more DMRs in the first intron than in any other genetic feature, potentially suggesting a role of *SKI* intron 1 methylation in regulation of gene expression in AMD, despite being approximately 30kb downstream of the *SKI* TSS (208). As a negative regulator of TGF-β signalling, increased methylation and subsequent decreased expression of *SKI* may be responsible for the development of AMD through increased expression of TGF-β. Given the proposed role of TGF-β in oxidative stress-induced RPE cell senescence and modulation of the complement pathway, both recognised hallmarks of AMD development, transcriptional silencing of *SKI*, previously observed by Porter et al in AMD RPE, could lead to an increase in TGF-β levels and subsequent complement C3 deposition within RPE tissue (209). Given my previous finding of predicted binding

at the *SKI* CpG locus of IRF-1, reported to be responsible for the induction of CFH expression in human RPE tissue, these results are particularly noteworthy (195).

In contrast to Porter *et al.* (2019) where an increase in the *GTF2H4* exonic CpG region methylation was observed in the RPE and found to be associated with reduced expression of *GTF2H4*, in whole blood a significant decrease in *GTF2H4* methylation in AMD was observed. Differences in tissue types may explain this variance. Of note, a putative IRF-1 binding site also resides within the pyrosequencing *GTF2H4* genomic region associated with cg22508626. To date, the Y402H genetic variant in *CFH* remains the strongest genetic risk locus associated with AMD with CFH expressed in drusen. Methylation modifications in the region of cg22508626 associated with *GTF2H4* could alter binding capabilities of IRF-1 with subsequent transcriptional effects on CFH, however this hypothesis would require experimental assessment.

Joehanes *et al*. (2016) recently demonstrated a large number of differentially methylated CpG sites within *SKI* intron 1 in smokers. The observation from the pyrosequencing studies above of significantly lower methylation in control smokers than control non-smokers at CpG position 4 within the *SKI* pyrosequencing fragment ($p = 0.0154$), also located within intron 1, supports the idea of *SKI* intron 1 being associated with smoking-sensitive DNA genomic regions with regards to methylation. While differential methylation was not observed in smokers within *GTF2H4* or *TNXB* in the Netherlands cohort, this section of the study unfortunately suffered from a markedly smaller smoking cohort compared to non-smokers, with only 22 smokers in total (10 control, 12 AMD). To be able to determine the effects of smoking on DNA methylation more accurately in genes of interest in AMD, a greater number of smokers was needed for further study.

Bisulphite pyrosequencing of whole blood samples was also used to address the third major aim of the project – to evaluate epigenetic changes in genes associated with the folate and methionine-homocysteine biosynthesis pathways, given the association of these intertwining pathways with both global DNA methylation and AMD. As high serum levels of homocysteine are reported to play a role in AMD

development, I focussed my attention on DNA methylation within *AHCY* and *MTR*, proteins responsible for the formation and metabolism of homocysteine. While a general trend of increased methylation in AMD whole blood was observed in the pyrosequenced *AHCY* fragment, no significant methylation change was detected ($p = 0.0748$). This contrasted with the methylation change found in the DNA methylation array at this locus, with the CpG site of interest displaying lower methylation levels ($p = 0.0006$) in AMD RPE than controls. Differences may be due to tissue type, as the difference in mean raw methylation levels observed in healthy controls in RPE (92.10%) and whole blood (53.89%), highlights a potential issue when studying AMD epigenetics in whole blood, as evidently not all genes share a similar methylome between tissues. Baseline methylation data for *AHCY* was not available in the Hewitt *et al*. (2017) study for a preliminary comparison between whole blood and ocular tissue methylation to be made and the difference in RPE / blood AHCY methylation only became apparent when the pyrosequencing was performed. While a comparison between blood / RPE methylation of AHCY could not be made from the data by Hewitt *et al* (2017), the significant demethylation of cg02138331 (Porter *et al.* univariate analysis) made this locus an attractive prospect for further investigation. However, the difference in raw methylation levels observed between blood and RPE following bisulphite pyrosequencing in whole blood highlights the need to assess genomic regions for inter-tissue baseline methylation variability on a case-by-case basis before surrogate tissues can be used with confidence.

Bisulphite pyrosequencing also demonstrated a highly significant change in the whole blood methylation of *MAT2A*, responsible for the biosynthesis of the universal methyl donor SAM. *MAT2A* methylation was significantly increased in AMD whole blood ($p = <0.0001$), with increased methylation also observed in smokers, in particular non-AMD smokers ($p = 0.0168$). Given the findings from Hewitt *et al.* (2017) of similar methylation levels at cg09363733 between healthy blood and ocular tissues, this suggests that whole blood tissue can be effectively used to observe methylation changes in AMD at this locus. Of note, Oliver *et al*.

(2015) identified a number of differentially methylated CpG loci in AMD whole blood, the majority of which demonstrated a reduction in DNA methylation (126).

Due to the small size of the European ethnicity cohorts from the Netherlands and Southampton, limited numbers of smokers in each cohort, and identical methylation trends in individual experimental batches, both cohorts were combined to assess a total of 144 healthy control and 144 AMD whole blood samples. Following the combined analysis of both cohorts, significant differential methylation was maintained in AMD whole blood within *GTFH24* ($p = 0.0015$) and *MAT2A* ($p = <0.0001$) as evidenced in Figure 4.23, suggesting a potential role for DNA methylation within both genes in the development of AMD. While the absolute level of methylation decrease in *GTF2H4* was observed to be small (mean ± SEM for CTRL = 91.83 ± 0.08, AMD = 91.47 ± 0.07), the significance of this decrease highlights the need for subsequent functional studies to evaluate the potential biological relevance an epigenetic change of this scale. However, the relative increase between mean methylation of CTRL and AMD samples within *MAT2A* (mean ± SEM for CTRL = 3.77 ± 0.15, AMD = 5.34 ± 0.19) of 42% suggests a potentially impactful biologically relevant change, especially given the nature of *MAT2A* as the universal methyl donor involved in global DNA methylation – again, further functional studies would be required to confirm this. Combined analysis of both cohorts also displayed significantly differential methylation within all three genes in smokers compared to non-smokers (*SKI* $p = <0.0001$, *GTF2H4* $p = 0.0227$, *MAT2A* $p = 0.0002$), with a clear suggestion from this data that smoking influences DNA methylation at these loci. Given the observed differences in methylation levels of AMD patient blood samples compared to control samples at these smoking-sensitive genomic regions within *SKI*, *GTF2H4* and *MAT2A*, it can therefore be suggested that smoking is a major factor in the modulation of DNA methylation at loci previously associated with AMD, and also at a locus associated with a principal dietary risk factor of AMD.

Chapter 5: Conclusions and Future Perspectives

The work in this thesis describes the evaluation of whole blood tissue as a surrogate for RPE in the study of DNA methylation in AMD at loci previously associated with AMD in RPE. In addition, I describe the evaluation of the impact of smoking on whole blood DNA methylation at CpG loci previously reported as being differentially methylated and expressed in AMD RPE. I also investigate and contribute novel insights on a role for DNA methylation of folate and methionine-homocysteine biosynthesis pathway genes in AMD development. To counteract the difficulties observed in obtaining patient RPE tissue for large-scale AMD studies, this work reveals the potential for whole blood tissue use in future studies of AMD epigenetics, with the prospect of being able to perform longitudinal studies and assess the effect of drugs and environmental factors in patients with AMD, without waiting for the post-mortem period, with all its added retrieval time constraints which have been further exacerbated due to COVID restrictions. I outline a number of genetic loci found to be significantly differentially methylated in the whole blood of both smokers and AMD patients.

The first set of results presented in this thesis (Chapter 3: Results I) outlines a variety of data mining approaches used to evaluate the potential for whole blood tissue to be used in this study and to evaluate a number of potential candidate loci for further study of the epigenetic impact of smoking and the folate pathway on AMD development. Comparison of the methylomes of whole blood and ocular tissues demonstrated considerable overlap, with CpG loci recently associated with AMD by Porter *et al.* (2019) showing comparable methylation between tissue types in healthy control patients at specific loci including cg18934822 (*SKI*), cg22508626 (*GTF2H4*) and six of the seven loci within a *TNXB* DMR. As this comparable methylation was validated by subsequent bisulphite pyrosequencing in control PBMC DNA, the use of whole blood tissue to study the epigenetics of these loci of interest was justified.

As cigarette smoking is a recognised major risk factor of AMD, as well as a known factor of aberrant DNA methylation, it was theorised that further investigation into the proposed association between smoking and the epigenetics of AMD could lead to greater insights into the role of smoking on AMD development. Genes associated with AMD by Porter *et al*. were found to be largely "smoking-sensitive" in a recent large-scale study of the epigenetic "signature" of smokers by Joehanes *et al*, strengthening the proposal that smoking may play a major role in AMD development. With the identification of the presence of histone markers such as H3K4Me1, HEK4Me3 and H3K27Ac, common indicators of transcriptional promoter and enhancer regions, at many of these smoking-sensitive CpG loci, the potential for cigarette smoking to influence levels of gene transcription through alteration of DNA methylation and chromatin accessibility cannot be dismissed. Given that these smoking-sensitive genes were found to also be differentially expressed in AMD RPE by Porter *et al*, this discovery is of particular interest in the investigation of smoking on AMD development. However, to further evaluate this association, additional data sets could be interrogated in a similar manner and cross-referenced to identify any particularly well-documented genetic links between DNA methylation and association with both smoking and AMD. For example, Christiansen *et al* identified 952 smoking-sensitive CpG sites across 500 genes in a study involving blood samples of 1407 patients (210) – analysis of this data alongside both that of Joehanes *et al* and the data presented in this study could identify further smoking sensitive genes of relevance to AMD, or potentially further support the claims made in this study, in turn further prompting subsequent functional studies to identify the biological relevance of differential methylation in smokers with respect to AMD.

As folic acid supplementation has been linked to reduced risk of AMD (51) and the folate pathway is implicated in methionine-homocysteine biosynthesis and the generation of the universal DNA methyl donor SAM, there is potential for the folate pathway to exert a protective effect in AMD through mechanisms involving modulation of DNA methylation. In particular, high levels of serum homocysteine have been repeatedly reported in AMD patients (173, 174), genes that regulate

homocysteine levels are of interest for further study into their potential role in AMD development.

To further address these ideas in AMD whole blood, a series of candidate CpG loci were outlined, based on a number of factors: evidence of overlap of CpG blood and RPE methylation, evidence of smoking-sensitivity, evidence of the presence of histone markers, and evidence of differential CpG methylation and gene transcription in AMD RPE. With these factors in mind, and following a reinterrogation of the study by Porter et al to reveal further sites of differential methylation (albeit below the threshold of genome-wide significance), a final list of 6 candidate CpG loci was constructed.

Further scrutinization utilising the TRANSFAC database identified predictions of transcription factor binding at these sites of differential methylation leading to a number of interesting predictions, for example the recurrence of predicted binding for IRF-1, TFII-I and retinoblastoma-associated protein EF2-1 at nucleotide motifs encompassing the CpG sites of interest. Given the role of DNA methylation in disruption of transcription factor binding, the overlap between binding sites of transcription factors of reported relevance to AMD and CpG sites found to be differentially methylated in AMD could have profound implications on the role of aberrant DNA methylation, modulated by smoking and folic acid supplementation, on transcription factor binding and development of AMD. Future evaluation of transcription factor binding at these sites, for instance via electrophoretic mobility shift assay (EMSA), could prove invaluable in clarifying these findings.

The second set of results in this thesis (Chapter 4: Results II) outlined the series of pyrosequencing studies conducted to assess an impact of smoking on DNA methylation at select genomic regions and assess DNA methylation levels of key components of the folate pathway in whole blood samples from patients with AMD. Following the subsequent bisulphite pyrosequencing of all 6 genes of interest in AMD, statistical analysis was performed. Application of the Shapiro-Wilk normality test identified non-normal distribution of the methylation data ($p = <0.05$). The Mann-Whitney test was therefore applied to all studies of two groups (CTRL vs all AMD, smoker vs non-smoker), while the Kruskal-Wallis test was applied

to studies of more than two groups (CTRL vs G2 AMD vs G3 AMD), to negate the potential error introduced from multiple Mann-Whitney tests. Dunn's multiple comparison post-test allowed significance to be identified between pairs of sample populations following overall significant difference of results observed by the Kruskal-Wallis test. Statistical analysis in this manner allowed for a minimisation of potential statistical error by using as few tests as possible, but maximum coverage of the data by allowing the methylation levels of all sample populations stratified by disease state and smoking status to be effectively compared within each gene.

Key findings included significantly reduced methylation levels of cg22508626 within *GTF2H4* ($p = 0.0015$) and significantly increased methylation levels of cg09363733 within *MAT2A* ($p = <0.0001$) in AMD whole blood in a combined cohort consisting of 144 control and 144 AMD samples. While an increase in DNA methylation at cg18934822 within *SKI* was observed, the result was not significant ($p = 0.1061$). However, significantly increased methylation was observed within *SKI* ($p = <0.0001$), and significantly reduced methylation was observed in *GTF2H4* ($p = 0.0227$) and *MAT2A* ($p = 0.0002$) in smokers compared to non-smokers. Furthermore, these findings were preserved within *SKI* ($p = <0.0001$) and *MAT2A* ($p = 0.0004$) when comparing only AMD smokers and non-smokers, and in *SKI* ($p = <0.0001$) and *GTF2H4* ($p = 0.0131$) when comparing only control smokers and non-smokers with direction of methylation change consistent across each genomic region for all smoking related studies in this combined whole blood cohort.

Given the differing directions of methylation change between genes, follow-up qPCR studies evaluating the implications of the observed differential methylation on the expression of these genes in AMD whole blood would prove particularly interesting. Gene expression methods utilising RNA such as qPCR and northern blot studies would, however, likely require more strict conditions regarding whole blood sample collection than was necessary in this study. Of note, PAXgene blood collection tubes are required for the storage of blood samples to avoid the potentially significant levels of RNA degradation observed with samples stored in conventional collection tubes (211). Additional samples specifically collected and stored in PAXgene collection tubes would therefore be required to further assess

the impact of the differential methylation observed in this study on genetic expression.

The identification of significantly differential methylation in *GTF2H4* in both this thesis and the Porter *et al.* study suggests that cg22508626 within *GTF2H4* has potential to be a successful biomarker of AMD in whole blood pending further investigation. The development of such biomarkers in whole blood could potentially facilitate further studies into the role of epigenetics in AMD, using far less invasive methods of sample collection and, in the case of RPE tissue, negating the need to obtain samples post-mortem within a short time frame. In particular, this finding suggests a potentially greater role of *GTF2H4* and TFIIH-mediated DNA damage repair in the pathogenesis of AMD than previously assumed.

The observation of significantly differential methylation between AMD smokers and non-smokers within *SKI* and *MAT2A*, as well as between control smokers and non-smokers within *SKI* and *GTF2H4*, suggests an impact of cigarette smoking on DNA methylation independent of disease status, at loci associated with AMD or AMD risk factors. However, the discovery of significantly increased methylation in the region of cg18934822 within *SKI* across all smoking stratified studies of the combined cohort is of major interest, given the identification by Joehanes et al of cg05603985 within *SKI* being one of the most significant "smoking-sensitive" loci observed in their study, 30kb upstream of cg18934822. A potential enhancement of transcriptional silencing of *SKI*, mediated by smoking-induced DNA hypermethylation, could be a key factor in AMD development, particularly given the relative proximity of cg18934822 to the TSS, being approximately 32kb downstream of the TSS. Reduction in *SKI* transcript levels, as shown previously by Porter *et al*. in AMD RPE, may lead to an increase in TGF-β levels within RPE tissue. TGF-β has been previously implicated in oxidative stress-induced RPE senescence and modulation of the complement pathway, both recognised aspects of AMD pathogenesis, with TGF-β inhibitors reported to display antiangiogenic effects in animal studies by reducing CNV lesion size in an AMD-like phenotype (149) These results suggest that smoking may play a major role in upregulation of TGF-β and AMD development through increased methylation at this intron 1 locus associated

with *SKI*. To further evaluate the role of *SKI* in AMD and the effects of smoking, future work could include investigation of TGF-β expression in smokers and non-smokers with AMD using whole blood and RPE tissue for example. Such studies could provide greater insight into the precise mechanisms by which *SKI* and TGF-β interact in AMD, and by which smoking contributes to AMD development.

The studies conducted in this thesis were, however, bound by several limitations, meaning absolute conclusions cannot necessarily be drawn from the results presented. A distinct lack of literature covering the overlap between the epigenomes of matched blood and ocular tissue, specifically the RPE, means that far more work is required before a comprehensive list of genes with comparable methylomes can be constructed to broaden the study to other non-validated loci. While the study by Hewitt *et al.* (2017) suggests validation of the use of whole blood to study the epigenetics of AMD at the specific loci documented in this thesis, additional studies could be collated with the data observed by Hewitt *et al*. in a large-scale meta-analysis, whereby genes / CpG loci of similar methylation status across both tissues could be cross-referenced to identify genes of interest in AMD. While the observation of similar whole blood and RPE methylation of *SKI*, *GTF2H4* and *TNXB* was clearly advantageous in the work presented in this thesis, additional studies would lend more confidence to the use of whole blood tissue as a surrogate for AMD epigenetics studies in a broader context. However, the identification of significantly differential methylation of *SKI*, *GTF2H4* and *MAT2A* in AMD smoker whole blood DNA suggests that follow up studies in this tissue, for instance a 850K Epic methylation array of control and AMD patient whole blood stratified by smoking status, could potentially identify a far greater number of smoking-sensitive genes of relevance to AMD, further clarifying the role of smoking in AMD epigenetics, as well as potentially identifying new candidate loci as potential AMD biomarkers in whole blood.

Although whole blood tissue is far easier to obtain than RPE tissue post-mortem, due to the short duration of this project, only 144 total control and AMD whole blood samples could be assessed per gene by bisulphite pyrosequencing. These combined cohorts also only included a total of 70 smokers, with an imbalance

between smoking and non-smoking samples in the Netherlands cohort. However, these results suggest a significant role for smoking on DNA methylation levels at the genomic loci associated with *SKI*, *GTF2H4* and *MAT2A* and in AMD. Follow-up studies may focus on larger-scale analysis of these genes with greater numbers of smokers, potentially strengthening the claims made in this thesis and prompting further studies into the mechanisms by which smoking influences DNA methylation within these genes, and how this may contribute to AMD development. As no significant methylation change was found between G2 and G3 AMD samples within any of the genes assessed throughout this study, yet significant methylation changes between control and both G2 and G3 AMD samples were numerous, this raises the question of the disease stage at which differential methylation can be first observed. These results suggest the DNA methylation modifications are relatively small between early and intermediate AMD, compared to no disease and early disease, at least at the loci investigated. As my results may demonstrate epigenetic changes in the early stages of AMD development at key genomic loci of interest, this could potentially form the basis for the development of future therapeutic strategies targeting DNA methylation in AMD before progression to the advanced stages and development of GA or CNV.

While the project aimed to assess the epigenetics of genes involved in the folate and methionine-homocysteine biosynthesis pathways in AMD, of the three relevant genes assessed differential methylation was only observed within *MAT2A*. While this finding was highly significant, no evidence of differential methylation within *AHCY* or *MTR* was observed, meaning few conclusions can be drawn from this study regarding the impact of aberrant DNA methylation on homocysteine in AMD. However, while *MTR* facilitates the breakdown of homocysteine and thus can contribute to alterations in serum homocysteine levels, *CBS* and *CTH* facilitate the conversion of homocysteine into cystathione and cystathione into cysteine respectively. CpG sites within these genes were not assessed, meaning future bisulphite pyrosequencing of these genes in AMD whole blood may clarify a role of DNA methylation on homocysteine metabolism pathway genes in AMD. As the interlinked folate and methionine-homocysteine biosynthesis pathways involve a

large number of genes acting in tandem, analysis of a far larger number than is present in this study would be necessary to confidently unravel a role of these pathways on DNA methylation in AMD. Furthermore, while these results implicate altered DNA methylation of *MAT2A* as a potential cause or consequence of AMD development, large-scale clinical studies assessing the impact of regular administration of folic acid supplementation on the methylation of these genes in AMD would help to more fully evaluate the possibly intricate role of the folate pathway in the epigenetics of AMD, as well as the proposed protective nature of folic acid supplementation on disease progression.

In conclusion, the results outlined in this thesis have formed the basis for the further study of AMD DNA methylation using whole blood tissue, with a number of CpG loci previously reported to be differentially methylated in AMD RPE also found to be differentially methylated in AMD whole blood. Smoking-sensitive methylation observed within *SKI*, *GTF2H4* and *MAT2A* suggests a possible epigenetic mechanism by which smoking can impact AMD development. In particular, significantly differential methylation of *MAT2A* in AMD whole blood may provide insight between the interplay between AMD, DNA methylation, and the folate pathway. As such, these results help to explain the impact of environmental effects on changes in the AMD methylome.

Chapter 6: References

1. Gheorghe A, Mahdi L, Musat O. AGE-RELATED MACULAR DEGENERATION. Rom J Ophthalmol. 2015;59(2):74-7.
2. Wong WL, Su X, Li X, Cheung CM, Klein R, Cheng CY, et al. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. The Lancet Global health. 2014;2(2):e106-16.
3. Ambati J, Fowler BJ. Mechanisms of age-related macular degeneration. Neuron. 2012;75(1):26-39.
4. Olsen TW, Feng X. The Minnesota Grading System of eye bank eyes for age-related macular degeneration. Invest Ophthalmol Vis Sci. 2004;45(12):4484-90.
5. The Age-Related Eye Disease Study system for classifying age-related macular degeneration from stereoscopic color fundus photographs: the Age-Related Eye Disease Study Report Number 6. Am J Ophthalmol. 2001;132(5):668-81.
6. Rudnicka AR, Kapetanakis VV, Jarrar Z, Wathern AK, Wormald R, Fletcher AE, et al. Incidence of Late-Stage Age-Related Macular Degeneration in American Whites: Systematic Review and Meta-analysis. American Journal of Ophthalmology. 2015;160(1):85-93.e3.
7. Wilde C, Poostchi A, Mehta RL, MacNab HK, Hillman JG, Vernon SA, et al. Prevalence of age-related macular degeneration in an elderly UK Caucasian population—The Bridlington Eye Assessment Project: a cross-sectional study. Eye. 2017;31(7):1042-50.
8. Olsen TW, Liao A, Robinson HS, Palejwala NV, Sprehe N. The Nine-Step Minnesota Grading System for Eyebank Eyes With Age Related Macular Degeneration: A Systematic Approach to Study Disease Stages. Invest Ophthalmol Vis Sci. 2017;58(12):5497-506.
9. Willoughby CE, Ponzin D, Ferrari S, Lobo A, Landau K, Omidi Y. Anatomy and physiology of the human eye: effects of mucopolysaccharidoses disease on structure and function – a review. Clinical & Experimental Ophthalmology. 2010;38(s1):2-11.
10. Gupta MP, Herzlich AA, Sauer T, Chan CC. Retinal Anatomy and Pathology. Developments in ophthalmology. 2016;55:7-17.
11. Strauss O. The Retinal Pigment Epithelium in Visual Function. Physiological Reviews. 2005;85(3):845-81.
12. Spraul CW, Grossniklaus HE. Characteristics of Drusen and Bruch's membrane in postmortem eyes with age-related macular degeneration. Archives of ophthalmology (Chicago, Ill : 1960). 1997;115(2):267-73.
13. Bowes Rickman C, Farsiu S, Toth CA, Klingeborn M. Dry age-related macular degeneration: mechanisms, therapeutic targets, and imaging. Invest Ophthalmol Vis Sci. 2013;54(14):ORSF68-ORSF80.
14. Ferris FL, 3rd, Wilkinson CP, Bird A, Chakravarthy U, Chew E, Csaky K, et al. Clinical classification of age-related macular degeneration. Ophthalmology. 2013;120(4):844-51.
15. Murphy AR, Truong YB, O'Brien CM, Glattauer V. Bio-inspired human in vitro outer retinal models: Bruch's membrane and its cellular interactions. Acta Biomaterialia. 2020;104:1-16.
16. Acharya UR, Mookiah MRK, Koh JEW, Tan JH, Noronha K, Bhandary S, et al. Novel risk index for the identification of age-related macular degeneration using radon transform and DWT features. Computers in Biology and Medicine. 2016;73:131–40.
17. Zetterberg M. Age-related eye disease and gender. Maturitas. 2016;83:19-26.
18. Rudnicka AR, Jarrar Z, Wormald R, Cook DG, Fletcher A, Owen CG. Age and gender variations in age-related macular degeneration prevalence in populations of European ancestry: a meta-analysis. Ophthalmology. 2012;119(3):571-80.

19. Klein R, Klein BEK, Linton KLP. Prevalence of Age-related Maculopathy: The Beaver Dam Eye Study. Ophthalmology. 2020;127(4, Supplement):S122-S32.
20. Huang L, Meng Q, Zhang C, Sun Y, Bai Y, Li S, et al. Gene–gene interaction of CFH, ARMS2, and ARMS2/HTRA1 on the risk of neovascular age-related macular degeneration and polypoidal choroidal vasculopathy in Chinese population. Eye. 2015;29(5):691-8.
21. Haines JL, Hauser MA, Schmidt S, Scott WK, Olson LM, Gallins P, et al. Complement Factor H Variant Increases the Risk of Age-Related Macular Degeneration. Science. 2005;308(5720):419.
22. Edwards AO, Ritter R, Abel KJ, Manning A, Panhuysen C, Farrer LA. Complement Factor H Polymorphism and Age-Related Macular Degeneration. Science. 2005;308(5720):421.
23. Hageman GS, Anderson DH, Johnson LV, Hancox LS, Taiber AJ, Hardisty LI, et al. A common haplotype in the complement regulatory gene factor H (HF1/CFH) predisposes individuals to age-related macular degeneration. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(20):7227.
24. Clark Simon J, Bishop Paul N, Day Anthony J. Complement factor H and age-related macular degeneration: the role of glycosaminoglycan recognition in disease pathology. Biochemical Society Transactions. 2010;38(5):1342-8.
25. Nozaki M, Raisler BJ, Sakurai E, Sarma JV, Barnum SR, Lambris JD, et al. Drusen complement components C3a and C5a promote choroidal neovascularization. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(7):2328-33.
26. Arakawa S, Takahashi A, Ashikawa K, Hosono N, Aoi T, Yasuda M, et al. Genome-wide association study identifies two susceptibility loci for exudative age-related macular degeneration in the Japanese population. Nature genetics. 2011;43(10):1001-4.
27. Fritsche LG, Igl W, Bailey JNC, Grassmann F, Sengupta S, Bragg-Gresham JL, et al. A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. Nature genetics. 2016;48(2):134-43.
28. Fritsche LG, Chen W, Schu M, Yaspan BL, Yu Y, Thorleifsson G, et al. Seven new loci associated with age-related macular degeneration. Nature genetics. 2013;45(4):433-9, 9e1-2.
29. Seddon JM, Yu Y, Miller EC, Reynolds R, Tan PL, Gowrisankar S, et al. Rare variants in CFI, C3 and C9 are associated with high risk of advanced age-related macular degeneration. Nature genetics. 2013;45(11):1366-70.
30. Thornton J, Edwards R, Mitchell P, Harrison RA, Buchan I, Kelly SP. Smoking and age-related macular degeneration: a review of association. Eye. 2005;19(9):935-44.
31. Tomany SC, Wang JJ, Van Leeuwen R, Klein R, Mitchell P, Vingerling JR, et al. Risk factors for incident age-related macular degeneration: pooled findings from 3 continents. Ophthalmology. 2004;111(7):1280-7.
32. Fu H, Feng D, Tang S, He Z, Xiang Y, Wu T, et al. Prevalence of Tobacco Smoking and Determinants of Success in Quitting Smoking among Patients with Chronic Diseases: A Cross-Sectional Study in Rural Western China. Int J Environ Res Public Health. 2017;14(2):167.
33. Proctor RN. The global smoking epidemic: a history and status report. Clinical lung cancer. 2004;5(6):371-6.
34. Furrukh M. Tobacco Smoking and Lung Cancer: Perception-changing facts. Sultan Qaboos Univ Med J. 2013;13(3):345-58.
35. Gallucci G, Tartarone A, Lerose R, Lalinga AV, Capobianco AM. Cardiovascular risk of smoking and benefits of smoking cessation. J Thorac Dis. 2020;12(7):3866-76.
36. Harlev A, Agarwal A, Gunes SO, Shetty A, du Plessis SS. Smoking and Male Infertility: An Evidence-Based Review. World J Mens Health. 2015;33(3):143-60.

37. Pineles BL, Park E, Samet JM. Systematic review and meta-analysis of miscarriage and maternal exposure to tobacco smoke during pregnancy. American journal of epidemiology. 2014;179(7):807-23.
38. Taylor DH, Jr., Hasselblad V, Henley SJ, Thun MJ, Sloan FA. Benefits of smoking cessation for longevity. Am J Public Health. 2002;92(6):990-6.
39. Takami T, Saito Y. Effects of smoking cessation on central blood pressure and arterial stiffness. Vasc Health Risk Manag. 2011;7:633-8.
40. Wannamethee SG, Shaper AG, Whincup PH, Walker M. Smoking cessation and the risk of stroke in middle-aged men. Jama. 1995;274(2):155-60.
41. Beane J, Sebastiani P, Liu G, Brody JS, Lenburg ME, Spira A. Reversible and permanent effects of tobacco smoke exposure on airway epithelial gene expression. Genome Biol. 2007;8(9):R201-R.
42. Bossé Y, Postma DS, Sin DD, Lamontagne M, Couture C, Gaudreault N, et al. Molecular signature of smoking in human lung tissues. Cancer research. 2012;72(15):3753-63.
43. Reid RD, Pritchard G, Walker K, Aitken D, Mullen K-A, Pipe AL. Managing smoking cessation. CMAJ. 2016;188(17-18):E484-E92.
44. Khan JC, Thurlby DA, Shahid H, Clayton DG, Yates JRW, Bradley M, et al. Smoking and age related macular degeneration: the number of pack years of cigarette smoking is a major determinant of risk for both geographic atrophy and choroidal neovascularisation. Br J Ophthalmol. 2006;90(1):75-80.
45. Rahman I, Morrison D, Donaldson K, MacNee W. Systemic oxidative stress in asthma, COPD, and smokers. American journal of respiratory and critical care medicine. 1996;154(4 Pt 1):1055-60.
46. Tan JS, Mitchell P, Kifley A, Flood V, Smith W, Wang JJ. Smoking and the long-term incidence of age-related macular degeneration: the Blue Mountains Eye Study. Archives of ophthalmology (Chicago, Ill : 1960). 2007;125(8):1089-95.
47. Klein R, Klein BE, Moss SE. Relation of smoking to the incidence of age-related maculopathy. The Beaver Dam Eye Study. American journal of epidemiology. 1998;147(2):103-10.
48. Vingerling JR, Hofman A, Grobbee DE, de Jong PT. Age-related macular degeneration and smoking. The Rotterdam Study. Archives of ophthalmology (Chicago, Ill : 1960). 1996;114(10):1193-6.
49. Smith W, Assink J, Klein R, Mitchell P, Klaver CC, Klein BE, et al. Risk factors for age-related macular degeneration: Pooled findings from three continents. Ophthalmology. 2001;108(4):697-704.
50. Seddon JM, Reynolds R, Shah HR, Rosner B. Smoking, dietary betaine, methionine, and vitamin D in monozygotic twins with discordant macular degeneration: epigenetic implications. Ophthalmology. 2011;118(7):1386-94.
51. A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E, beta carotene, and zinc for age-related macular degeneration and vision loss: AREDS report no. 8. Archives of ophthalmology (Chicago, Ill : 1960). 2001;119(10):1417-36.
52. Plafker SM, O'Mealey GB, Szweda LI. Mechanisms for countering oxidative stress and damage in retinal pigment epithelium. Int Rev Cell Mol Biol. 2012;298:135-77.
53. Lu L, Hackett SF, Mincey A, Lai H, Campochiaro PA. Effects of different types of oxidative stress in RPE cells. Journal of Cellular Physiology. 2006;206(1):119-25.
54. Zhao Z, Chen Y, Wang J, Sternberg P, Freeman ML, Grossniklaus HE, et al. Age-related retinopathy in NRF2-deficient mice. PLoS One. 2011;6(4):e19456.
55. Bellezza I, Mierla AL, Minelli A. Nrf2 and NF-κB and Their Concerted Modulation in Cancer Pathogenesis and Progression. Cancers. 2010;2(2):483-97.

56. Seddon JM, Rosner B, Sperduto RD, Yannuzzi L, Haller JA, Blair NP, et al. Dietary fat and risk for advanced age-related macular degeneration. Archives of ophthalmology (Chicago, Ill : 1960). 2001;119(8):1191-9.
57. Cho E, Hung S, Willett WC, Spiegelman D, Rimm EB, Seddon JM, et al. Prospective study of dietary fat and the risk of age-related macular degeneration. Am J Clin Nutr. 2001;73(2):209-18.
58. Delcourt C, Carrière I, Cristol JP, Lacroux A, Gerber M. Dietary fat and the risk of age-related maculopathy: the POLANUT study. European journal of clinical nutrition. 2007;61(11):1341-4.
59. van Leeuwen EM, Emri E, Merle BMJ, Colijn JM, Kersten E, Cougnard-Gregoire A, et al. A new perspective on lipid research in age-related macular degeneration. Progress in retinal and eye research. 2018;67:56-86.
60. Tan JSL, Wang JJ, Flood V, Mitchell P. Dietary Fatty Acids and the 10-Year Incidence of Age-Related Macular Degeneration: The Blue Mountains Eye Study. Archives of Ophthalmology. 2009;127(5):656-65.
61. Tuo J, Ross RJ, Herzlich AA, Shen D, Ding X, Zhou M, et al. A high omega-3 fatty acid diet reduces retinal lesions in a murine model of macular degeneration. The American journal of pathology. 2009;175(2):799-807.
62. Zhang AC, Singh S, Craig JP, Downie LE. Omega-3 Fatty Acids and Eye Health: Opinions and Self-Reported Practice Behaviors of Optometrists in Australia and New Zealand. Nutrients. 2020;12(4).
63. Wu J, Sun X. Complement system and age-related macular degeneration: drugs and challenges. Drug Des Devel Ther. 2019;13:2413-25.
64. Moreno TA, Kim SJ. Ranibizumab (Lucentis) versus Bevacizumab (Avastin) for the Treatment of Age-Related Macular Degeneration: An Economic Disparity of Eye Health. Seminars in ophthalmology. 2016;31(4):378-84.
65. Gu L, Chen H, Tuo J, Gao X, Chen L. Inhibition of experimental choroidal neovascularization in mice by anti-VEGFA/VEGFR2 or non-specific siRNA. Experimental eye research. 2010;91(3):433-9.
66. Xin X, Zhu Y, Xi R, Hao Y. Establishing a mouse model of choroidal neovascularization to study the therapeutic effect of levotinib and its mechanism. Saudi Journal of Biological Sciences. 2020;27(9):2491-7.
67. Kwak N, Okamoto N, Wood JM, Campochiaro PA. VEGF Is Major Stimulator in Model of Choroidal Neovascularization. Invest Ophthalmol Vis Sci. 2000;41(10):3158-64.
68. Brown DM, Regillo CD. Anti-VEGF agents in the treatment of neovascular age-related macular degeneration: applying clinical trial results to the treatment of everyday patients. Am J Ophthalmol. 2007;144(4):627-37.
69. Gemenetzi M, Lotery AJ, Patel PJ. Risk of geographic atrophy in age-related macular degeneration patients treated with intravitreal anti-VEGF agents. Eye. 2017;31(1):1-9.
70. Sachdeva MM, Moshiri A, Leder HA, Scott AW. Endophthalmitis following intravitreal injection of anti-VEGF agents: long-term outcomes and the identification of unusual micro-organisms. J Ophthalmic Inflamm Infect. 2016;6(1):2-.
71. Guimaraes TACd, Georgiou M, Bainbridge JWB, Michaelides M. Gene therapy for neovascular age-related macular degeneration: rationale, clinical trials and future directions. British Journal of Ophthalmology. 2021;105(2):151.
72. Hawkins BS, Bressler NM, Miskala PH, Bressler SB, Holekamp NM, Marsh MJ, et al. Surgery for subfoveal choroidal neovascularization in age-related macular degeneration: ophthalmic findings: SST report no. 11. Ophthalmology. 2004;111(11):1967-80.
73. Lu B, Malcuit C, Wang S, Girman S, Francis P, Lemieux L, et al. Long-term safety and function of RPE from human embryonic stem cells in preclinical models of macular degeneration. Stem cells (Dayton, Ohio). 2009;27(9):2126-35.

74. Liu Y, Xu HW, Wang L, Li SY, Zhao CJ, Hao J, et al. Human embryonic stem cell-derived retinal pigment epithelium transplants as a potential treatment for wet age-related macular degeneration. Cell discovery. 2018;4:50.
75. Carneiro Â, Andrade JP. Erratum to "Nutritional and Lifestyle Interventions for Age-Related Macular Degeneration: A Review". Oxid Med Cell Longev. 2017;2017:2435963-.
76. Chakravarthy U, Augood C, Bentham GC, de Jong PT, Rahu M, Seland J, et al. Cigarette smoking and age-related macular degeneration in the EUREYE Study. Ophthalmology. 2007;114(6):1157-63.
77. Mitchell P, Wang JJ, Smith W, Leeder SR. Smoking and the 5-Year Incidence of Age-Related Maculopathy: The Blue Mountains Eye Study. Archives of Ophthalmology. 2002;120(10):1357-63.
78. Weinhold B. Epigenetics: the science of change. Environ Health Perspect. 2006;114(3):A160-A7.
79. Handy DE, Castro R, Loscalzo J. Epigenetic modifications: basic mechanisms and role in cardiovascular disease. Circulation. 2011;123(19):2145-56.
80. Calo E, Wysocka J. Modification of enhancer chromatin: what, how, and why? Mol Cell. 2013;49(5):825-37.
81. Sharifi-Zarchi A, Gerovska D, Adachi K, Totonchi M, Pezeshk H, Taft RJ, et al. DNA methylation regulates discrimination of enhancers from promoters through a H3K4me1-H3K4me3 seesaw mechanism. BMC Genomics. 2017;18(1):964.
82. Hyun K, Jeon J, Park K, Kim J. Writing, erasing and reading histone lysine methylations. Experimental & Molecular Medicine. 2017;49(4):e324-e.
83. Chuang JC, Jones PA. Epigenetics and microRNAs. Pediatric research. 2007;61(5 Pt 2):24r-9r.
84. Mehta A, Baltimore D. MicroRNAs as regulatory elements in immune system logic. Nature Reviews Immunology. 2016;16(5):279-94.
85. Pileczki V, Cojocneanu-Petric R, Maralani M, Neagoe IB, Sandulescu R. MicroRNAs as regulators of apoptosis mechanisms in cancer. Clujul Med. 2016;89(1):50-5.
86. Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology. 2013;38(1):23-38.
87. Jin B, Robertson KD. DNA methyltransferases, DNA damage repair, and cancer. Adv Exp Med Biol. 2013;754:3-29.
88. Teodoridis JM, Strathdee G, Brown R. Epigenetic silencing mediated by CpG island methylation: potential as a therapeutic target and as a biomarker. Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy. 2004;7(4-5):267-78.
89. Kribelbauer JF, Lu X-J, Rohs R, Mann RS, Bussemaker HJ. Toward a Mechanistic Understanding of DNA Methylation Readout by Transcription Factors. Journal of Molecular Biology. 2020;432(6):1801-15.
90. Fuks F, Hurd PJ, Wolf D, Nan X, Bird AP, Kouzarides T. The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. The Journal of biological chemistry. 2003;278(6):4035-40.
91. Li E, Zhang Y. DNA methylation in mammals. Cold Spring Harb Perspect Biol. 2014;6(5):a019133-a.
92. Chen T, Ueda Y, Dodge JE, Wang Z, Li E. Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. Mol Cell Biol. 2003;23(16):5594-605.
93. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, et al. Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci U S A. 2005;102(30):10604-9.

94. Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrensch MR, Wiemels JL, et al. Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. PLoS Genet. 2009;5(8):e1000602.
95. Bollati V, Schwartz J, Wright R, Litonjua A, Tarantini L, Suh H, et al. Decline in genomic DNA methylation through aging in a cohort of elderly subjects. Mechanisms of ageing and development. 2009;130(4):234-9.
96. Sae-Lee C, Biasi JD, Robinson N, Barrow TM, Mathers JC, Koutsidis G, et al. DNA methylation patterns of LINE-1 and Alu for pre-symptomatic dementia in type 2 diabetes. PLOS ONE. 2020;15(6):e0234578.
97. Zhang N. Role of methionine on epigenetic modification of DNA methylation and gene expression in animals. Animal Nutrition. 2018;4(1):11-6.
98. de Almeida BP, Apolónio JD, Binnie A, Castelo-Branco P. Roadmap of DNA methylation in breast cancer identifies novel prognostic biomarkers. BMC Cancer. 2019;19(1):219.
99. Yang X, Wong MPM, Ng RK. Aberrant DNA Methylation in Acute Myeloid Leukemia and Its Clinical Implications. Int J Mol Sci. 2019;20(18):4576.
100. Jiang H, Ou Z, He Y, Yu M, Wu S, Li G, et al. DNA methylation markers in the diagnosis and prognosis of common leukemias. Signal Transduction and Targeted Therapy. 2020;5(1):3.
101. Krushkal J, Silvers T, Reinhold WC, Sonkin D, Vural S, Connelly J, et al. Epigenome-wide DNA methylation analysis of small cell lung cancer cell lines suggests potential chemotherapy targets. Clinical Epigenetics. 2020;12(1):93.
102. Svendsen AJ, Gervin K, Lyle R, Christiansen L, Kyvik K, Junker P, et al. Differentially Methylated DNA Regions in Monozygotic Twin Pairs Discordant for Rheumatoid Arthritis: An Epigenome-Wide Study. 2016;7(510).
103. Liu Y, Aryee MJ, Padyukov L, Fallin MD, Hesselberg E, Runarsson A, et al. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. Nature Biotechnology. 2013;31(2):142-7.
104. Wu X, Zhang Y. TET-mediated active DNA demethylation: mechanism, function and beyond. Nature Reviews Genetics. 2017;18(9):517-34.
105. An J, Rao A, Ko M. TET family dioxygenases and DNA demethylation in stem cells and cancers. Experimental & Molecular Medicine. 2017;49(4):e323-e.
106. Armstrong DA, Lesseur C, Conradt E, Lester BM, Marsit CJ. Global and gene-specific DNA methylation across multiple tissues in early infancy: implications for children's health research. FASEB J. 2014;28(5):2088-97.
107. Garg P, Joshi RS, Watson C, Sharp AJ. A survey of inter-individual variation in DNA methylation identifies environmentally responsive co-regulated networks of epigenetic variation in the human genome. PLoS Genet. 2018;14(10):e1007707-e.
108. Zhang B, Zhou Y, Lin N, Lowdon RF, Hong C, Nagarajan RP, et al. Functional DNA methylation differences between tissues, cell types, and across individuals discovered using the M&M algorithm. Genome Res. 2013;23(9):1522-40.
109. Slieker RC, Relton CL, Gaunt TR, Slagboom PE, Heijmans BT. Age-related DNA methylation changes are tissue-specific with ELOVL2 promoter methylation as exception. Epigenetics & Chromatin. 2018;11(1):25.
110. Kallestad L, Blackshaw S, Khalil AM, Palczewski K. Tissue- and Species-Specific Patterns of RNA metabolism in Post-Mortem Mammalian Retina and Retinal Pigment Epithelium. Scientific Reports. 2019;9(1):14821.
111. Rhein M, Hagemeier L, Klintschar M, Muschler M, Bleich S, Frieling H. DNA methylation results depend on DNA integrity-role of post mortem interval. Front Genet. 2015;6:182-.

112. Nayak BK. Understanding the relevance of sample size calculation. Indian J Ophthalmol. 2010;58(6):469-70.
113. Gavin DP, Sharma RP. Histone modifications, DNA methylation, and schizophrenia. Neurosci Biobehav Rev. 2010;34(6):882-8.
114. Nimmo ER, Prendergast JG, Aldhous MC, Kennedy NA, Henderson P, Drummond HE, et al. Genome-wide methylation profiling in Crohn's disease identifies altered epigenetic regulation of key host defense mechanisms including the Th17 pathway. Inflammatory bowel diseases. 2012;18(5):889-99.
115. Gervin K, Vigeland MD, Mattingsdal M, Hammerø M, Nygård H, Olsen AO, et al. DNA methylation and gene expression changes in monozygotic twins discordant for psoriasis: identification of epigenetically dysregulated genes. PLoS Genet. 2012;8(1):e1002454-e.
116. Hewitt AW, Januar V, Sexton-Oates A, Joo JE, Franchina M, Wang JJ, et al. DNA methylation landscape of ocular tissue relative to matched peripheral blood. Scientific Reports. 2017;7(1):46330.
117. Berdasco M, Gómez A, Rubio MJ, Català-Mora J, Zanón-Moreno V, Lopez M, et al. DNA Methylomes Reveal Biological Networks Involved in Human Eye Development, Functions and Associated Disorders. Scientific Reports. 2017;7(1):11762.
118. Brooks J, Cairns P, Zeleniuch-Jacquotte A. Promoter methylation and the detection of breast cancer. Cancer Causes Control. 2009;20(9):1539-50.
119. Kuo HK, Griffith JD, Kreuzer KN. 5-Azacytidine induced methyltransferase-DNA adducts block DNA replication in vivo. Cancer research. 2007;67(17):8248-54.
120. Qin T, Jelinek J, Si J, Shu J, Issa JP. Mechanisms of resistance to 5-aza-2'-deoxycytidine in human cancer cell lines. Blood. 2009;113(3):659-67.
121. Segura-Pacheco B, Perez-Cardenas E, Taja-Chayeb L, Chavez-Blanco A, Revilla-Vazquez A, Benitez-Bribiesca L, et al. Global DNA hypermethylation-associated cancer chemotherapy resistance and its reversion with the demethylating agent hydralazine. Journal of translational medicine. 2006;4:32.
122. Andrade FdO, Nguyen NM, Warri A, Hilakivi-Clarke L. Reversal of increased mammary tumorigenesis by valproic acid and hydralazine in offspring of dams fed high fat diet during pregnancy. Scientific Reports. 2019;9(1):20271.
123. Halby L, Menon Y, Rilova E, Pechalrieu D, Masson V, Faux C, et al. Rational Design of Bisubstrate-Type Analogues as Inhibitors of DNA Methyltransferases in Cancer Cells. Journal of Medicinal Chemistry. 2017;60(11):4665-79.
124. Hunter A, Spechler PA, Cwanger A, Song Y, Zhang Z, Ying G-S, et al. DNA methylation is associated with altered gene expression in AMD. Invest Ophthalmol Vis Sci. 2012;53(4):2089-105.
125. Kim Y, Cha SJ, Choi H-J, Kim K. Omega Class Glutathione S-Transferase: Antioxidant Enzyme in Pathogenesis of Neurodegenerative Diseases. Oxid Med Cell Longev. 2017;2017:5049532.
126. Oliver VF, Jaffe AE, Song J, Wang G, Zhang P, Branham KE, et al. Differential DNA methylation identified in the blood and retina of AMD patients. Epigenetics. 2015;10(8):698-707.
127. Wei L, Liu B, Tuo J, Shen D, Chen P, Li Z, et al. Hypomethylation of the IL17RC promoter associates with age-related macular degeneration. Cell Rep. 2012;2(5):1151-8.
128. Ardeljan D, Wang Y, Park S, Shen D, Chu XK, Yu C-R, et al. Interleukin-17 Retinotoxicity Is Prevented by Gene Transfer of a Soluble Interleukin-17 Receptor Acting as a Cytokine Blocker: Implications for Age-Related Macular Degeneration. PLOS ONE. 2014;9(4):e95900.

129. Oliver VF, Franchina M, Jaffe AE, Branham KE, Othman M, Heckenlively JR, et al. Hypomethylation of the IL17RC promoter in peripheral blood leukocytes is not a hallmark of age-related macular degeneration. Cell Rep. 2013;5(6):1527-35.
130. Porter LF, Saptarshi N, Fang Y, Rathi S, den Hollander AI, de Jong EK, et al. Whole-genome methylation profiling of the retinal pigment epithelium of individuals with age-related macular degeneration reveals differential methylation of the SKI, GTF2H4, and TNXB genes. Clin Epigenetics. 2019;11(1):6.
131. Rimel JK, Taatjes DJ. The essential and multifunctional TFIIH complex. Protein Sci. 2018;27(6):1018-37.
132. Vermeulen W, Fousteri M. Mammalian transcription-coupled excision repair. Cold Spring Harb Perspect Biol. 2013;5(8):a012625-a.
133. Schärer OD. Nucleotide excision repair in eukaryotes. Cold Spring Harb Perspect Biol. 2013;5(10):a012609-a.
134. Vélez-Cruz R, Zadorin AS, Coin F, Egly JM. Sirt1 suppresses RNA synthesis after UV irradiation in combined xeroderma pigmentosum group D/Cockayne syndrome (XP-D/CS) cells. Proc Natl Acad Sci U S A. 2013;110(3):E212-20.
135. Orioli D, Compe E, Nardo T, Mura M, Giraudon C, Botta E, et al. XPD mutations in trichothiodystrophy hamper collagen VI expression and reveal a role of TFIIH in transcription derepression. Human Molecular Genetics. 2013;22(6):1061-73.
136. Naegeli H, Sugasawa K. The xeroderma pigmentosum pathway: Decision tree analysis of DNA quality. DNA Repair. 2011;10(7):673-83.
137. Buch SC, Diergaarde B, Nukui T, Day RS, Siegfried JM, Romkes M, et al. Genetic variability in DNA repair and cell cycle control pathway genes and risk of smoking-related lung cancer. Molecular carcinogenesis. 2012;51 Suppl 1(Suppl 1):E11-20.
138. Briggs FB, Goldstein BA, McCauley JL, Zuvich RL, De Jager PL, Rioux JD, et al. Variation within DNA repair pathway genes and risk of multiple sclerosis. American journal of epidemiology. 2010;172(2):217-24.
139. Wang M, Liu H, Liu Z, Yi X, Bickeboller H, Hung RJ, et al. Genetic variant in DNA repair gene GTF2H4 is associated with lung cancer risk: a large-scale analysis of six published GWAS datasets in the TRICL consortium. Carcinogenesis. 2016;37(9):888-96.
140. Wang SS, Gonzalez P, Yu K, Porras C, Li Q, Safaeian M, et al. Common genetic variants and risk for HPV persistence and progression to cervical cancer. PloS one. 2010;5(1):e8667-e.
141. Newman AM, Gallo NB, Hancox LS, Miller NJ, Radeke CM, Maloney MA, et al. Systems-level analysis of age-related macular degeneration reveals global biomarkers and phenotype-specific functional networks. Genome Med. 2012;4(2):16-.
142. Szaflik JP, Janik-Papis K, Synowiec E, Ksiazek D, Zaras M, Wozniak K, et al. DNA damage and repair in age-related macular degeneration. Mutation research. 2009;669(1-2):169-76.
143. Tecalco-Cruz AC, Ríos-López DG, Vázquez-Victorio G, Rosales-Alvarez RE, Macías-Silva M. Transcriptional cofactors Ski and SnoN are major regulators of the TGF-β/Smad signaling pathway in health and disease. Signal Transduction and Targeted Therapy. 2018;3(1):15.
144. Yu AL, Fuchshofer R, Kook D, Kampik A, Bloemendal H, Welge-Lüssen U. Subtoxic Oxidative Stress Induces Senescence in Retinal Pigment Epithelial Cells via TGF-β Release. Invest Ophthalmol Vis Sci. 2009;50(2):926-35.
145. Mitsuhiro MR, Eguchi S, Yamashita H. Regulation mechanisms of retinal pigment epithelial cell migration by the TGF-beta superfamily. Acta ophthalmologica Scandinavica. 2003;81(6):630-8.

146. Goswami MT, Reka AK, Kurapati H, Kaza V, Chen J, Standiford TJ, et al. Regulation of complement-dependent cytotoxicity by TGF-β-induced epithelial–mesenchymal transition. Oncogene. 2016;35(15):1888-98.
147. Luo K, Stroschein SL, Wang W, Chen D, Martens E, Zhou S, et al. The Ski oncoprotein interacts with the Smad proteins to repress TGFbeta signaling. Genes Dev. 1999;13(17):2196-206.
148. Deheuninck J, Luo K. Ski and SnoN, potent negative regulators of TGF-beta signaling. Cell Res. 2009;19(1):47-57.
149. Wang X, Ma W, Han S, Meng Z, Zhao L, Yin Y, et al. TGF-β participates choroid neovascularization through Smad2/3-VEGF/TNF-α signaling in mice with Laser-induced wet age-related macular degeneration. Scientific reports. 2017;7(1):9672-.
150. Tosi GM, Orlandini M, Galvagni F. The Controversial Role of TGF-β in Neovascular Age-Related Macular Degeneration Pathogenesis. Int J Mol Sci. 2018;19(11):3363.
151. Kaufman CS, Butler MG. Mutation in TNXB gene causes moderate to severe Ehlers-Danlos syndrome. World J Med Genet. 2016;6(2):17-21.
152. Ye Z, Shuai P, Zhai Y, Li F, Jiang L, Lu F, et al. Associations of 6p21.3 Region with Age-related Macular Degeneration and Polypoidal Choroidal Vasculopathy. Scientific Reports. 2016;6(1):20914.
153. Cipriani V, Leung H-T, Plagnol V, Bunce C, Khan JC, Shahid H, et al. Genome-wide association study of age-related macular degeneration identifies associated variants in the TNXB–FKBPL–NOTCH4 region of chromosome 6p21.3. Human Molecular Genetics. 2012;21(18):4138-50.
154. Kõks G, Uudelepp M-L, Limbach M, Peterson P, Reimann E, Kõks S. Smoking-Induced Expression of the GPR15 Gene Indicates Its Potential Role in Chronic Inflammatory Pathologies. The American Journal of Pathology. 2015;185(11):2898-906.
155. Fragou D, Pakkidi E, Aschner M, Samanidou V, Kovatsi L. Smoking and DNA methylation: Correlation of methylation with smoking behavior and association with diseases and fetus development following prenatal exposure. Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association. 2019;129:312-27.
156. Joehanes R, Just AC, Marioni RE, Pilling LC, Reynolds LM, Mandaviya PR, et al. Epigenetic Signatures of Cigarette Smoking. Circulation Cardiovascular genetics. 2016;9(5):436-47.
157. Delchier N, Herbig A-L, Rychlik M, Renard CMGC. Folates in Fruits and Vegetables: Contents, Processing, and Stability. Comprehensive Reviews in Food Science and Food Safety. 2016;15(3):506-28.
158. Milman N. Intestinal absorption of folic acid - new physiologic & molecular aspects. Indian J Med Res. 2012;136(5):725-8.
159. Froese DS, Fowler B, Baumgartner MR. Vitamin B(12) , folate, and the methionine remethylation cycle-biochemistry, pathways, and regulation. Journal of inherited metabolic disease. 2019;42(4):673-85.
160. Miller AL. The methionine-homocysteine cycle and its effects on cognitive diseases. Alternative medicine review : a journal of clinical therapeutic. 2003;8(1):7-19.
161. Chittiboyina S, Chen Z, Chiorean EG, Kamendulis LM, Hocevar BA. The role of the folate pathway in pancreatic cancer risk. PLoS One. 2018;13(2):e0193298.
162. Zheng Y, Cantley LC. Toward a better understanding of folate metabolism in health and disease. Journal of Experimental Medicine. 2018;216(2):253-66.
163. Pufulete M, Al-Ghnaniem R, Khushal A, Appleby P, Harris N, Gout S, et al. Effect of folic acid supplementation on genomic DNA methylation in patients with colorectal adenoma. Gut. 2005;54(5):648-53.

164. Nephew KP, Balch C, Skalnik DG. Methyl Group Acceptance Assay for the Determination of Global DNA Methylation Levels. In: Tost J, editor. DNA Methylation: Methods and Protocols. Totowa, NJ: Humana Press; 2009. p. 35-41.
165. Irwin RE, Thursby S-J, Ondičová M, Pentieva K, McNulty H, Richmond RC, et al. A randomized controlled trial of folic acid intervention in pregnancy highlights a putative methylation-regulated control element at ZFP57. Clinical Epigenetics. 2019;11(1):31.
166. Christensen KE, Mikael LG, Leung K-Y, Lévesque N, Deng L, Wu Q, et al. High folic acid consumption leads to pseudo-MTHFR deficiency, altered lipid metabolism, and liver injury in mice. The American Journal of Clinical Nutrition. 2015;101(3):646-58.
167. Oh KH, Lee JH, Kim KN, Chang N. Effects of folic acid intake on plasma folate, vitamin B12 and homocysteine levels, hepatic SAM/SAH ratio and the expression of cerebral myelin basic protein in pregnant and lactating rats. The FASEB Journal. 2009;23(S1):904.5-.5.
168. Mandaviya PR, Stolk L, Heil SG. Homocysteine and DNA methylation: A review of animal and human literature. Molecular Genetics and Metabolism. 2014;113(4):243-52.
169. Jiang Y, Sun T, Xiong J, Cao J, Li G, Wang S. Hyperhomocysteinemia-mediated DNA Hypomethylation and its Potential Epigenetic Role in Rats. Acta Biochimica et Biophysica Sinica. 2007;39(9):657-67.
170. Huang P, Wang F, Sah BK, Jiang J, Ni Z, Wang J, et al. Homocysteine and the risk of age-related macular degeneration: a systematic review and meta-analysis. Scientific reports. 2015;5:10585-.
171. Edirimanne VE, Woo CW, Siow YL, Pierce GN, Xie JY, O K. Homocysteine stimulates NADPH oxidase-mediated superoxide production leading to endothelial dysfunction in rats. Canadian journal of physiology and pharmacology. 2007;85(12):1236-47.
172. Tyagi N, Sedoris KC, Steed M, Ovechkin AV, Moshal KS, Tyagi SC. Mechanisms of homocysteine-induced oxidative stress. American Journal of Physiology-Heart and Circulatory Physiology. 2005;289(6):H2649-H56.
173. Kamburoglu G, Gumus K, Kadayifcilar S, Eldem B. Plasma homocysteine, vitamin B12 and folate levels in age-related macular degeneration. Graefe's archive for clinical and experimental ophthalmology = Albrecht von Graefes Archiv fur klinische und experimentelle Ophthalmologie. 2006;244(5):565-9.
174. Seddon JM, Gensler G, Klein ML, Milton RC. Evaluation of Plasma Homocysteine and Risk of Age-Related Macular Degeneration. American Journal of Ophthalmology. 2006;141(1):201-3.
175. Axer-Siegel R, Bourla D, Ehrlich R, Dotan G, Benjamini Y, Gavendo S, et al. Association of neovascular age-related macular degeneration and hyperhomocysteinemia. Am J Ophthalmol. 2004;137(1):84-9.
176. Nowak M, Swietochowska E, Wielkoszyński T, Marek B, Kos-Kudła B, Szapska B, et al. Homocysteine, vitamin B12, and folic acid in age-related macular degeneration. European journal of ophthalmology. 2005;15(6):764-7.
177. Coral K, Raman R, Rathi S, Rajesh M, Sulochana KN, Angayarkanni N, et al. Plasma homocysteine and total thiol content in patients with exudative age-related macular degeneration. Eye (London, England). 2006;20(2):203-7.
178. Ates O, Azizi S, Alp HH, Kiziltunc A, Beydemir S, Cinici E, et al. Decreased serum paraoxonase 1 activity and increased serum homocysteine and malondialdehyde levels in age-related macular degeneration. The Tohoku journal of experimental medicine. 2009;217(1):17-22.
179. Javadzadeh A, Ghorbanihaghjo A, Bahreini E, Rashtchizadeh N, Argani H, Alizadeh S. Plasma oxidized LDL and thiol-containing molecules in patients with exudative age-related macular degeneration. Molecular vision. 2010;16:2578-84.

180. Ghosh S, Saha M, Das D. A study on plasma homocysteine level in age-related macular degeneration. Nepalese journal of ophthalmology : a biannual peer-reviewed academic journal of the Nepal Ophthalmic Society : NEPJOPH. 2013;5(2):195-200.
181. Mulero J, Manresa N, Zafrilla P, Losada M. Markers of cardiovascular risk in elderly patients with age-related macular degeneration. Clinical hemorheology and microcirculation. 2014;58(3):447-53.
182. Wang JJ, Ross RJ, Tuo J, Burlutsky G, Tan AG, Chan CC, et al. The LOC387715 polymorphism, inflammatory markers, smoking, and age-related macular degeneration. A population-based case-control study. Ophthalmology. 2008;115(4):693-9.
183. Obeid R, Ninios K, Loew U, Gatzioufas Z, Hoffmann S, Seitz B, et al. Aqueous humor glycation marker and plasma homocysteine in macular degeneration. Clinical chemistry and laboratory medicine. 2013;51(3):657-63.
184. Owa C, Poulin M, Yan L, Shioda T. Technical adequacy of bisulfite sequencing and pyrosequencing for detection of mitochondrial DNA methylation: Sources and avoidance of false-positive detection. PloS one. 2018;13(2):e0192722-e.
185. Darst RP, Pardo CE, Ai L, Brown KD, Kladde MP. Bisulfite sequencing of DNA. Curr Protoc Mol Biol. 2010;Chapter 7:Unit-7.9.17.
186. Harrington CT, Lin EI, Olson MT, Eshleman JR. Fundamentals of pyrosequencing. Archives of pathology & laboratory medicine. 2013;137(9):1296-303.
187. Delaney C, Garg SK, Yung R. Analysis of DNA Methylation by Pyrosequencing. Methods Mol Biol. 2015;1343:249-64.
188. Marsh S. Chapter 8 - Pyrosequencing. In: Patrinos GP, Ansorge WJ, editors. Molecular Diagnostics (Second Edition). San Diego: Academic Press; 2010. p. 107-16.
189. Huang Y-W, Huang THM, Wang L-S. Profiling DNA methylomes from microarray to genome-scale sequencing. Technol Cancer Res Treat. 2010;9(2):139-47.
190. Kurdyukov S, Bullock M. DNA Methylation Analysis: Choosing the Right Method. Biology (Basel). 2016;5(1):3.
191. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16(3):1215-.
192. Kim EJ, Grant GR, Bowman AS, Haider N, Gudiseva HV, Chavali VRM. Complete Transcriptome Profiling of Normal and Age-Related Macular Degeneration Eye Tissues Reveals Dysregulation of Anti-Sense Transcription. Scientific Reports. 2018;8(1):3040.
193. manifest IKBM. https://emea.support.illumina.com/array/array_kits/infinium_humanmethylation450_beadchip_kit/downloads.html [
194. Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, et al. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57-74.
195. Amadi-Obi A, Yu C-R, Dambuza I, Kim S-H, Marrero B, Egwuagu C. Interleukin 27 Induces the Expression of Complement Factor H (CFH) in the Retina. PloS one. 2012;7:e45801.
196. Gong C, Qiao L, Feng R, Xu Q, Zhang Y, Fang Z, et al. IL-6-induced acetylation of E2F1 aggravates oxidative damage of retinal pigment epithelial cell line. Experimental eye research. 2020;200:108219.
197. Chakravarthy U, McKay GJ, de Jong PT, Rahu M, Seland J, Soubrane G, et al. ARMS2 increases the risk of early and late age-related macular degeneration in the European Eye Study. Ophthalmology. 2013;120(2):342-8.
198. Pan Y, Iejima D, Nakayama M, Suga A, Noda T, Kaur I, et al. Binding of Gtf2i-β/δ transcription factors to the ARMS2 gene leads to increased circulating HTRA1 in AMD patients and in vitro. Journal of Biological Chemistry. 2021;296:100456.
199. Gestri G, Osborne RJ, Wyatt AW, Gerrelli D, Gribble S, Stewart H, et al. Reduced TFAP2A function causes variable optic fissure closure and retinal defects and sensitizes eye

development to mutations in other morphogenetic regulators. Hum Genet. 2009;126(6):791-803.
200. Grabe N. AliBaba2: Context specific identification of transcription factor binding sites. In silico biology. 2002;2:S1-15.
201. Farré D, Roset R, Huerta M, Adsuara JE, Roselló L, Albà MM, et al. Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. Nucleic Acids Res. 2003;31(13):3651-3.
202. Gearing LJ, Cumming HE, Chapman R, Finkel AM, Woodhouse IB, Luu K, et al. CiiiDER: A tool for predicting and analysing transcription factor binding sites. PLOS ONE. 2019;14(9):e0215495.
203. Zogopoulos VL, Spaho K, Ntouka C, Lappas GA, Kyranis I, Bagos PG, et al. TFBSPred: A functional transcription factor binding site prediction webtool for humans and mice. Int J Epigen. 2021;1(4):9.
204. Swango KL, Timken MD, Chong MD, Buoncristiani MR. A quantitative PCR assay for the assessment of DNA degradation in forensic samples. Forensic Science International. 2006;158(1):14-26.
205. Ramon D, Braden M, Adams S, Marincola FM, Wang L. Pyrosequencing™ : A one-step method for high resolution HLA typing. Journal of translational medicine. 2003;1(1):9.
206. Gharizadeh B, Nordström T, Ahmadian A, Ronaghi M, Nyrén P. Long-read pyrosequencing using pure 2'-deoxyadenosine-5'-O'-(1-thiotriphosphate) Sp-isomer. Analytical biochemistry. 2002;301(1):82-90.
207. Marziali A, Akeson M. New DNA Sequencing Methods. Annual Review of Biomedical Engineering. 2001;3(1):195-223.
208. Anastasiadi D, Esteve-Codina A, Piferrer F. Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species. Epigenetics & Chromatin. 2018;11(1):37.
209. Li Y, Song D, Song Y, Zhao L, Wolkow N, Tobias JW, et al. Iron-induced Local Complement Component 3 (C3) Up-regulation via Non-canonical Transforming Growth Factor (TGF)-β Signaling in the Retinal Pigment Epithelium. The Journal of biological chemistry. 2015;290(19):11918-34.
210. Christiansen C, Castillo-Fernandez JE, Domingo-Relloso A, Zhao W, El-Sayed Moustafa JS, Tsai PC, et al. Novel DNA methylation signatures of tobacco smoking with trans-ethnic effects. Clinical Epigenetics. 2021;13(1):36.
211. Chai V, Vassilakos A, Lee Y, Wright JA, Young AH. Optimization of the PAXgene blood RNA extraction system for gene expression analysis of clinical samples. Journal of clinical laboratory analysis. 2005;19(5):182-8.

Chapter 7: Appendix

Table 7.1: Individual sample information for Cohort 1 (Netherlands) 192 blood DNA samples

Sample ID	Age (years)	Gender	Ever Smoked (Yes / No)	AREDS
53023	64	F	N	CTRL
53314	72	F	N	CTRL
53557	69	M	Y	CTRL
53022	82	M	N	CTRL
53506	59	M	N	CTRL
53589	74	F	N	CTRL
53333	75	M	N	CTRL
53318	77	F	N	CTRL
53028	72	M	N	CTRL
53253	69	F	N	CTRL
53508	53	F	Y	CTRL
50937	62	M	N	CTRL
53337	75	F	N	CTRL
53074	67	M	N	CTRL
52935	68	F	Y	CTRL
53510	67	M	N	CTRL
48005	66	M	N	CTRL
53491	82	F	N	CTRL
50748	82	F	N	CTRL
53039	66	F	N	CTRL
52964	87	M	N	CTRL
53091	70	F	N	CTRL
53521	59	F	Y	CTRL
53096	70	M	N	CTRL
50246	69	F	N	CTRL
53531	63	F	N	CTRL
53391	62	M	N	CTRL
52940	79	F	N	CTRL
53150	71	M	Y	CTRL
53532	54	M	N	CTRL
50743	73	M	N	CTRL
53554	64	F	N	CTRL
53018	76	F	N	CTRL
45756	80	F	N	CTRL
53189	81	F	N	CTRL
53160	81	M	N	CTRL
53306	84	F	N	CTRL
46646	78	F	N	CTRL
53317	70	F	N	CTRL
53328	75	F	N	CTRL

53035	78	M	Y	CTRL
53257	77	F	N	CTRL
53320	85	F	N	CTRL
50614	71	M	N	CTRL
52965	80	F	N	CTRL
53347	75	M	N	CTRL
50604	66	F	Y	CTRL
47396	79	F	N	CTRL
50809	78	M	N	CTRL
50595	69	M	N	CTRL
50624	79	M	N	CTRL
53112	79	F	N	CTRL
53237	65	F	N	CTRL
53049	80	F	N	CTRL
53158	77	F	N	CTRL
50617	71	F	N	CTRL
53075	76	F	N	CTRL
48038	69	M	N	CTRL
52948	75	M	N	CTRL
53005	72	F	N	CTRL
52986	75	F	N	CTRL
53147	61	M	N	CTRL
47772	72	M	N	CTRL
48013	67	F	N	CTRL
37670	76	F	N	CTRL
44060	85	M	N	CTRL
53170	73	F	N	CTRL
50768	81	F	N	CTRL
50254	82	M	N	CTRL
53343	74	M	N	CTRL
53279	77	M	N	CTRL
47670	82	F	N	CTRL
50596	78	M	N	CTRL
50282	79	M	N	CTRL
53019	77	M	N	CTRL
48011	80	F	N	CTRL
53474	62	F	N	CTRL
53156	87	F	N	CTRL
50611	81	F	N	CTRL
47416	75	M	Y	CTRL
50603	82	M	N	CTRL
47432	82	F	N	CTRL
53662	56	M	N	CTRL
50616	80	M	N	CTRL

47999	67	M	N	CTRL
53110	82	F	N	CTRL
53577	63	M	N	CTRL
47383	78	M	N	CTRL
53559	64	M	N	CTRL
50250	74	M	N	CTRL
47443	78	M	N	CTRL
47359	72	M	N	CTRL
50634	62	M	N	CTRL
48025	79	F	N	CTRL
50574	58	M	Y	CTRL
47405	67	M	Y	CTRL
53516	65	M	Y	2
44610	74	F	N	2
50708	78	F	Y	2
53624	63	F	N	2
53360	78	F	Y	2
53644	65	F	N	2
50832	71	M	N	2
53653	64	M	N	2
53294	67	F	N	2
50837	78	F	N	2
53390	72	M	N	2
53085	74	F	N	2
52917	72	M	N	2
50655	80	F	N	2
47824	73	F	N	2
48001	73	M	N	2
50816	81	F	N	2
53660	77	F	N	2
53668	64	F	N	2
53426	78	M	N	2
50631	74	M	N	2
50779	69	M	N	2
48015	73	F	N	2
50792	78	F	N	2
50784	79	F	Y	2
48009	85	M	N	2
56921	68	F	N	2
53311	72	M	N	2
50706	79	F	N	2
56562	68	F	N	2
53647	64	F	N	2
50807	66	M	N	2

53262	68	F	N	2
50866	82	M	N	2
56615	54	M	N	2
47398	84	F	N	2
53451	69	M	N	2
53431	72	F	N	2
50965	81	M	Y	2
20808	62	M	N	2
56642	77	M	Y	2
50747	79	M	N	2
56643	71	F	Y	2
51100	62	M	N	2
50838	69	F	N	2
50763	74	F	N	2
53640	66	M	Y	2
53252	75	F	N	2
50947	83	M	Y	3
37606	75	F	N	3
40462	80	M	N	3
56516	86	F	N	3
46478	77	F	N	3
13112	78	F	N	3
42811	74	M	N	3
37574	89	M	N	3
56433	80	F	N	3
35807	88	F	N	3
50878	76	F	N	3
40126	66	F	Y	3
50782	64	M	N	3
40903	73	F	N	3
50783	69	M	N	3
56460	85	M	N	3
50785	75	F	N	3
50672	88	F	N	3
50824	78	M	N	3
56523	74	F	N	3
44774	79	F	N	3
50833	80	F	N	3
56425	78	F	N	3
50777	62	F	N	3
50845	85	M	N	3
56517	71	F	N	3
50849	69	M	N	3
56453	73	M	N	3

42436	70	M	N	3
56585	72	F	N	3
56422	71	M	N	3
49520	81	M	N	3
50935	80	F	N	3
56586	77	M	N	3
56492	76	M	N	3
53013	79	M	N	3
56627	81	F	N	3
39810	62	M	N	3
50681	83	M	N	3
56519	61	F	Y	3
53309	72	F	N	3
50831	62	M	N	3
56555	86	M	N	3
41798	63	F	N	3
50853	56	F	Y	3
56594	74	F	N	3
44836	66	F	N	3
51583	75	F	N	3

Table 7.2: Individual sample information for Cohort 2 (Southampton) 96 blood DNA samples

Sample ID	Age (years)	Gender	Ever Smoked (Yes / No)	AREDS
SC007	78	M	Y	CTRL
SC035	75	M	Y	CTRL
SC054	67	F	Y	CTRL
SC068	81	M	Y	CTRL
SC086	73	M	Y	CTRL
SC090	65	M	Y	CTRL
SC092	72	M	Y	CTRL
SC095	76	M	Y	CTRL
SC109	71	M	Y	CTRL
SC136	71	F	Y	CTRL
SC137	68	F	Y	CTRL
SC152	78	M	Y	CTRL
SC184	78	F	Y	CTRL
SC187	63	M	Y	CTRL
SC205	67	F	Y	CTRL
SC230	70	M	Y	CTRL
SC241	65	M	Y	CTRL
SC245	81	M	Y	CTRL
SC261	78	F	Y	CTRL
SC272	65	M	Y	CTRL
SC291	67	M	Y	CTRL
SC323	82	M	Y	CTRL
SC340	77	F	Y	CTRL
SC366	67	F	Y	CTRL
SC034	80	M	N	CTRL
SC044	85	F	N	CTRL
SC050	78	F	N	CTRL
SC060	81	F	N	CTRL
SC065	88	M	N	CTRL
SC098	88	M	N	CTRL
SC105	85	F	N	CTRL
SC122	84	F	N	CTRL
SC129	81	F	N	CTRL
SC130	80	F	N	CTRL
SC141	82	F	N	CTRL
SC149	80	F	N	CTRL
SC157	83	F	N	CTRL
SC195	81	F	N	CTRL
SC209	79	M	N	CTRL
SC214	83	F	N	CTRL
SC217	86	M	N	CTRL
SC228	81	M	N	CTRL

SC253	84	F	N	CTRL
SC282	83	M	N	CTRL
SC287	81	F	N	CTRL
SC321	84	F	N	CTRL
SC339	89	F	N	CTRL
SC360	81	M	N	CTRL
SA031	73	F	N	2
SA039	87	F	N	2
SA065	73	M	N	2
SA068	75	F	N	2
SA076	75	M	N	2
SA108	65	F	N	2
SA109	68	F	N	2
SA131	95	F	N	2
SA168	83	F	N	2
SA191	63	M	N	2
SA235	78	F	N	2
SA315	73	F	N	2
SA009	68	F	Y	2
SA044	72	F	Y	2
SA097	71	F	Y	2
SA118	76	M	Y	2
SA159	74	F	Y	2
SA244	63	M	Y	2
SA258	75	F	Y	2
SA378	77	M	Y	2
SA387	63	M	Y	2
SA406	75	M	Y	2
SA413	75	F	Y	2
SA012	91	F	N	3
SA046	88	F	N	3
SA105	83	F	N	3
SA120	79	F	N	3
SA121	82	F	N	3
SA187	81	F	N	3
SA204	81	F	N	3
SA246	88	F	N	3
SA261	78	F	N	3
SA292	92	F	N	3
SA306	81	M	N	3
SA348	88	F	N	3
SA022	81	M	Y	3
SA086	61	F	Y	3
SA088	89	F	Y	3

SA136	95	F	Y	3
SA151	80	F	Y	3
SA167	94	M	Y	3
SA192	83	F	Y	3
SA273	80	F	Y	3
SA291	83	F	Y	3
SA308	87	M	Y	3
SA375	74	F	Y	3
SA379	78	F	Y	3
SA415	83	M	Y	3

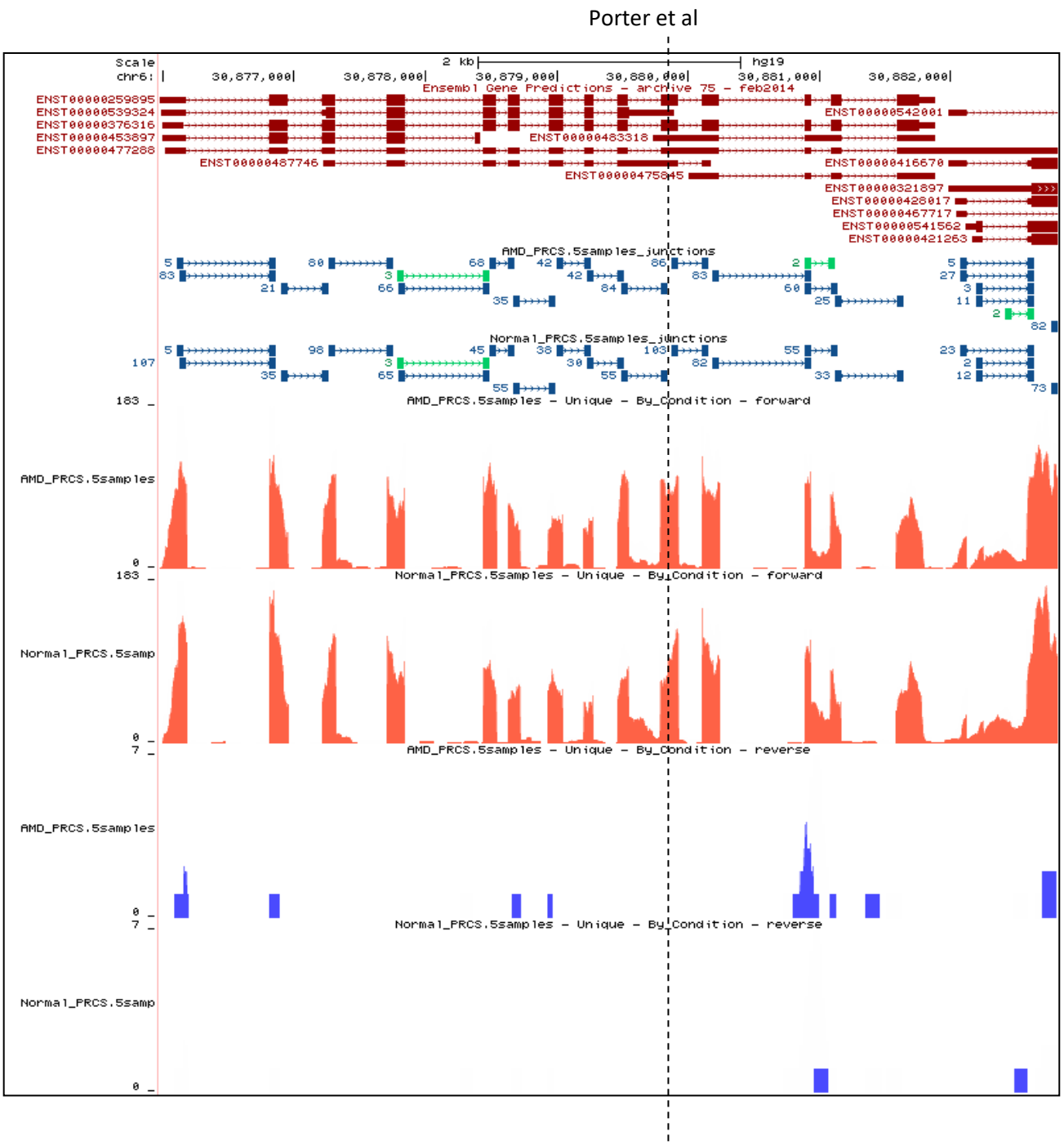

Figure 7.1: *GTF2H4* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. Dashed black line defines location of Porter et al CpG site of interest. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue.

Porter et al

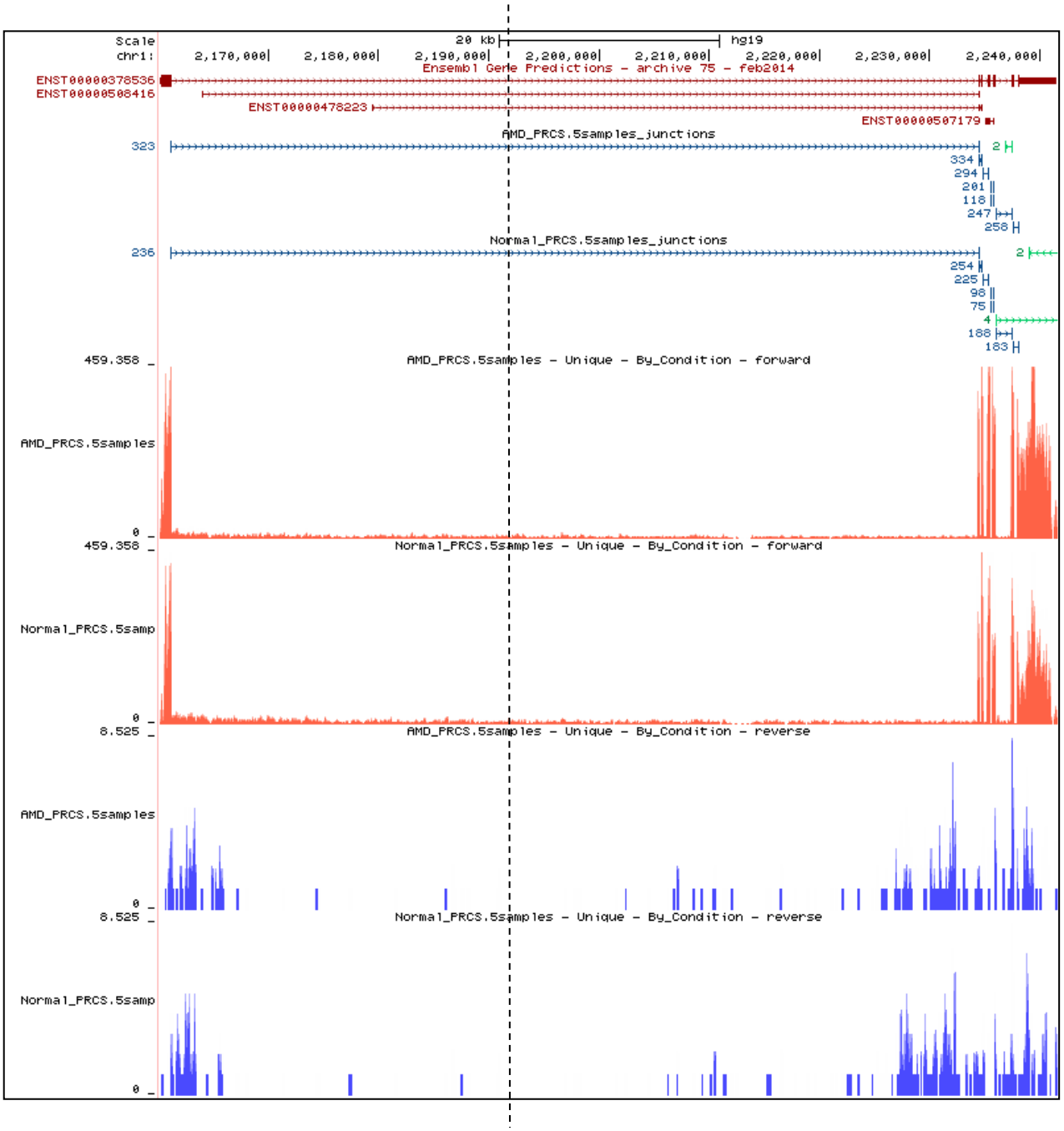

Figure 7.2: *SKI* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. Dashed black line defines location of Porter et al CpG site of interest. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue.

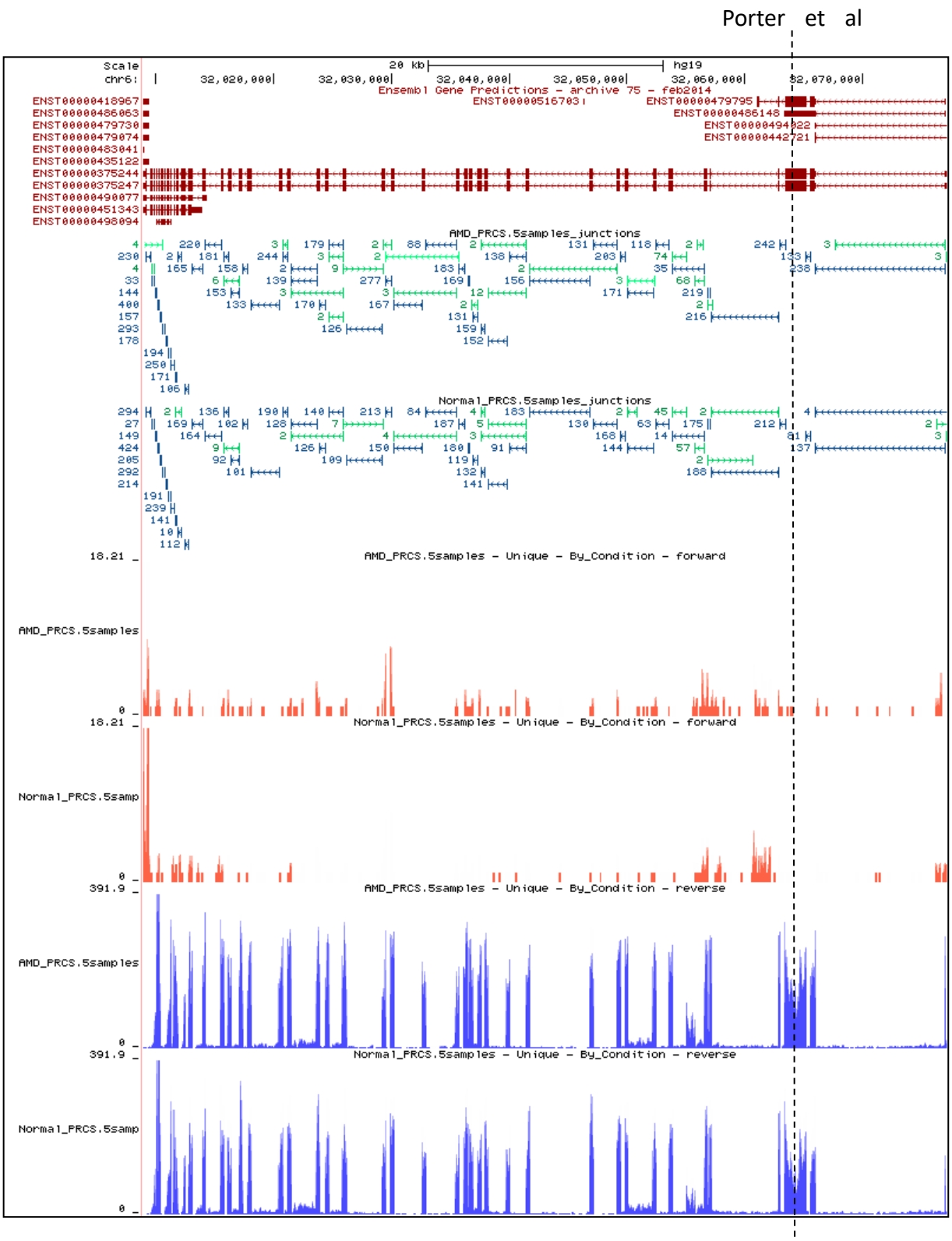

Figure 7.3: *TNXB* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. Dashed black line defines location of Porter et al DMR site of interest. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue

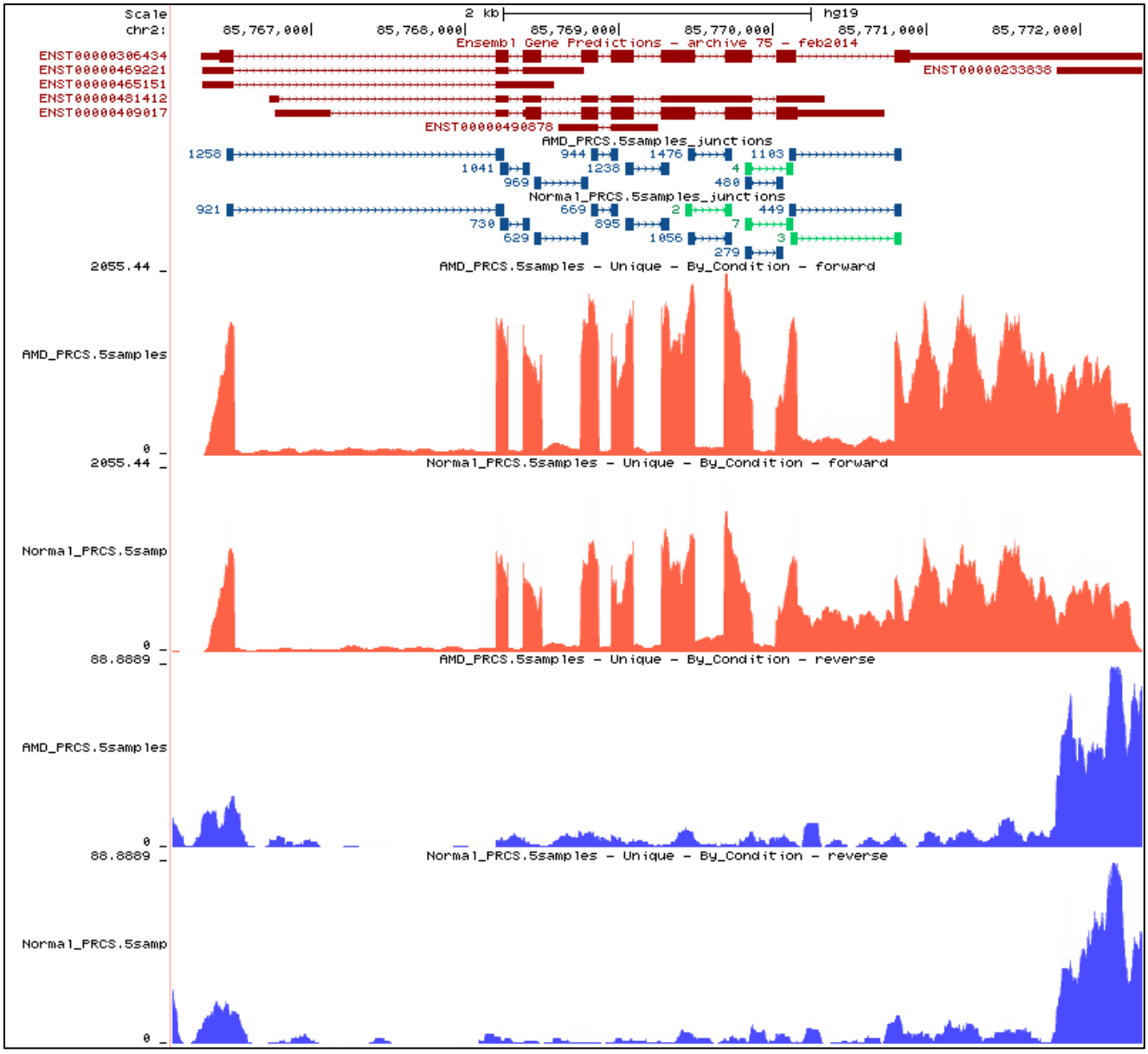

Figure 7.4: *MAT2A* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue.

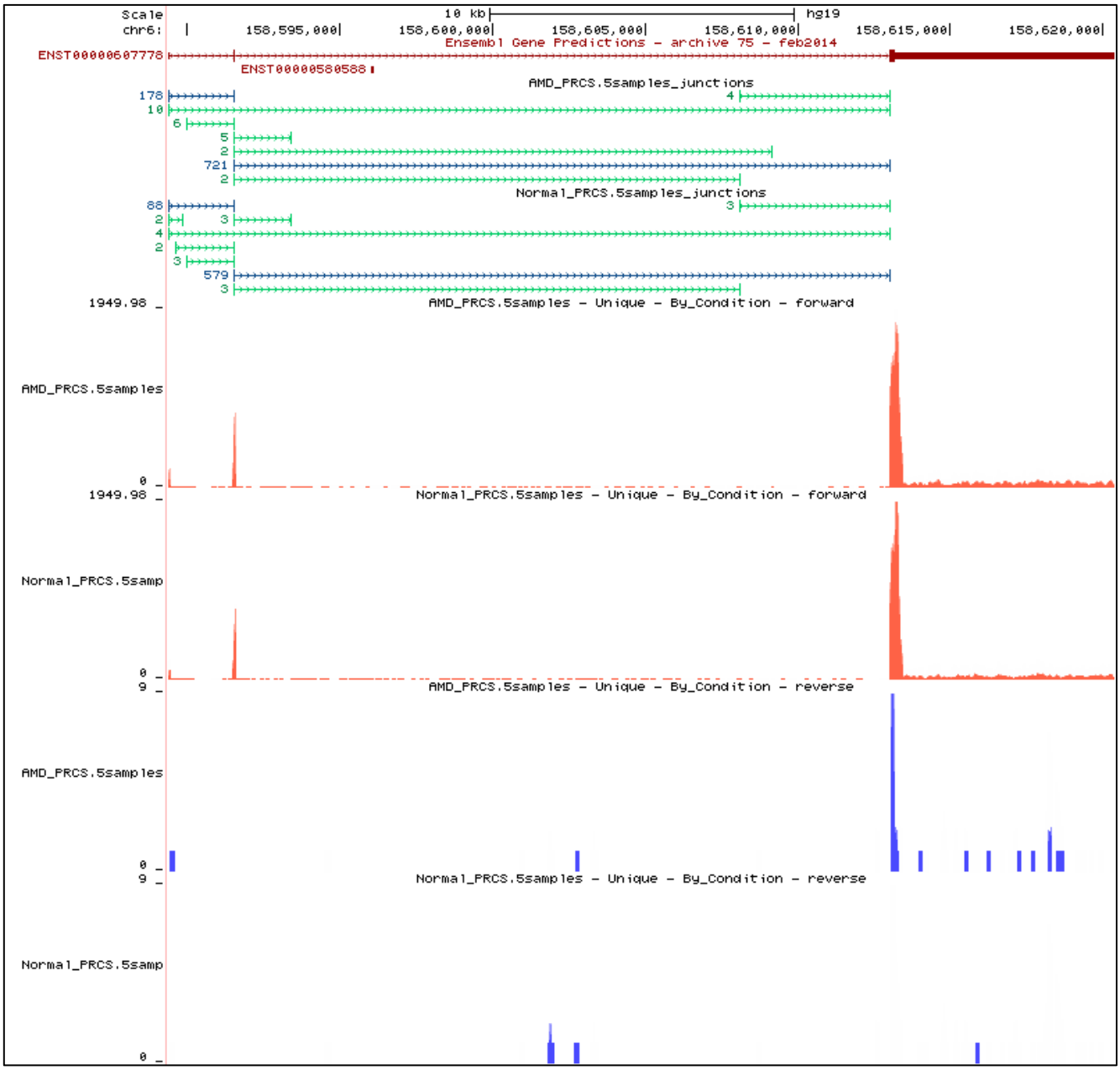

Figure 7.5: *GTF2H5* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue.

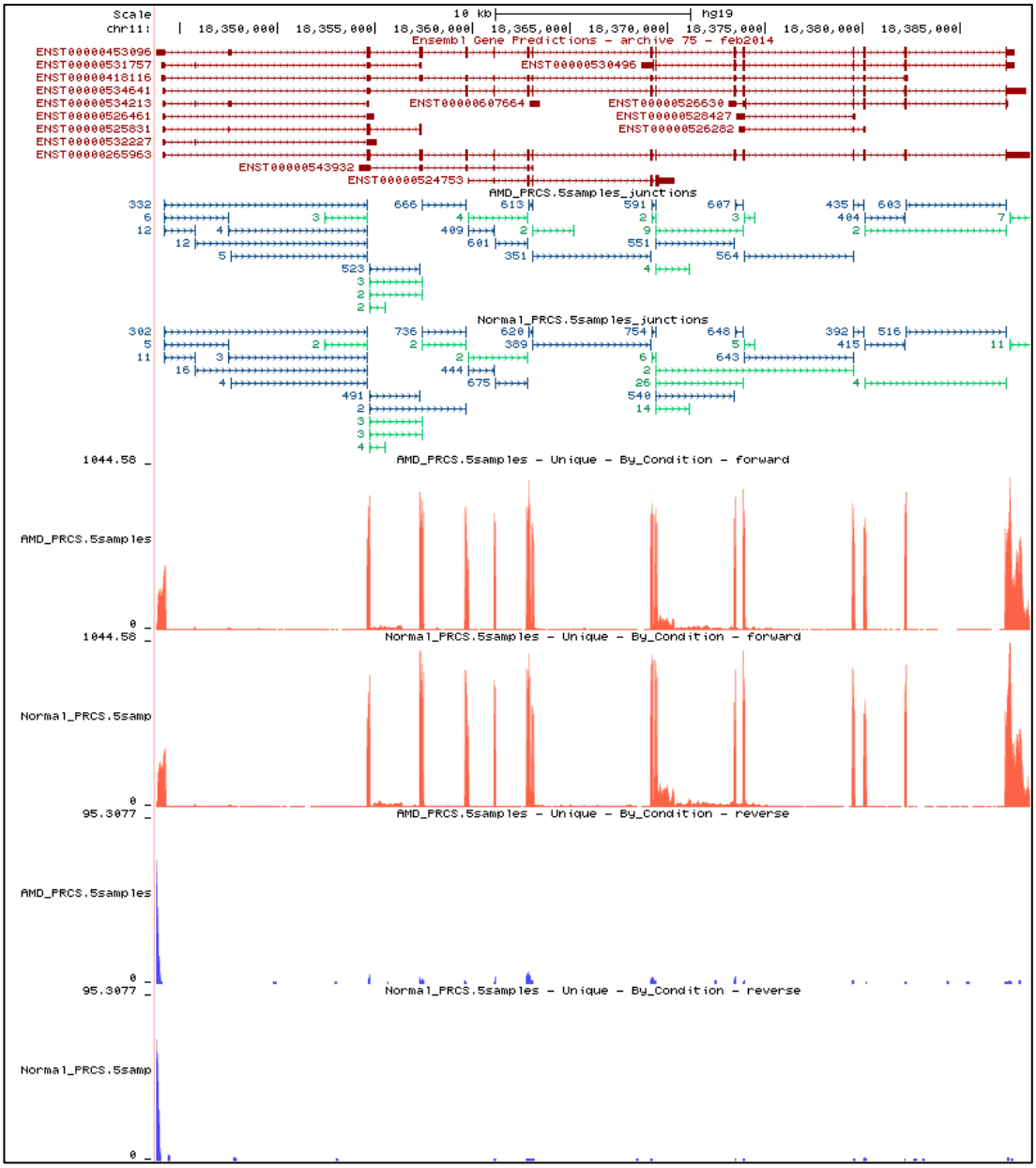

Figure 7.6: *GTF2H1* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue.

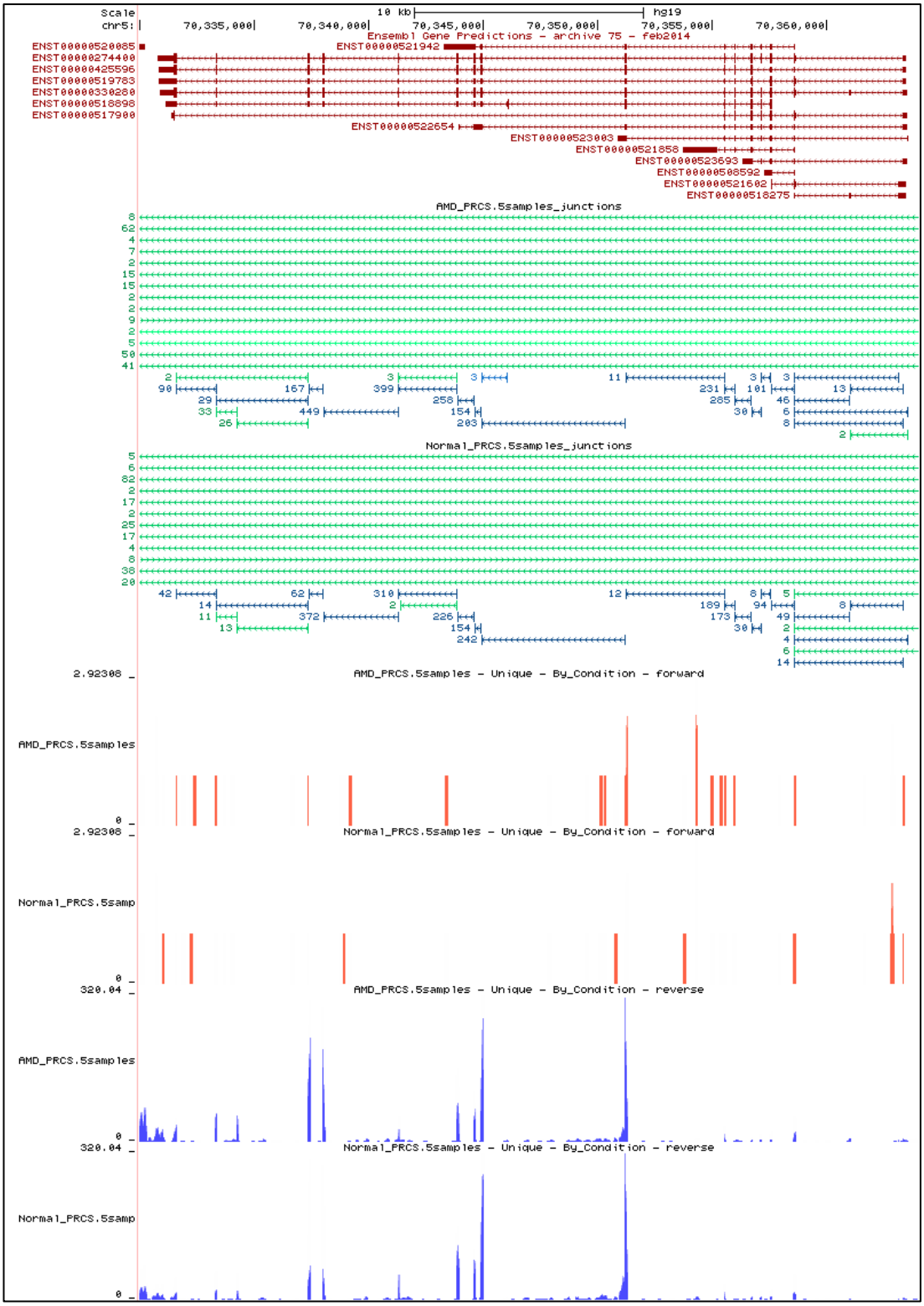

Figure 7.7: *GTF2H2* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue.

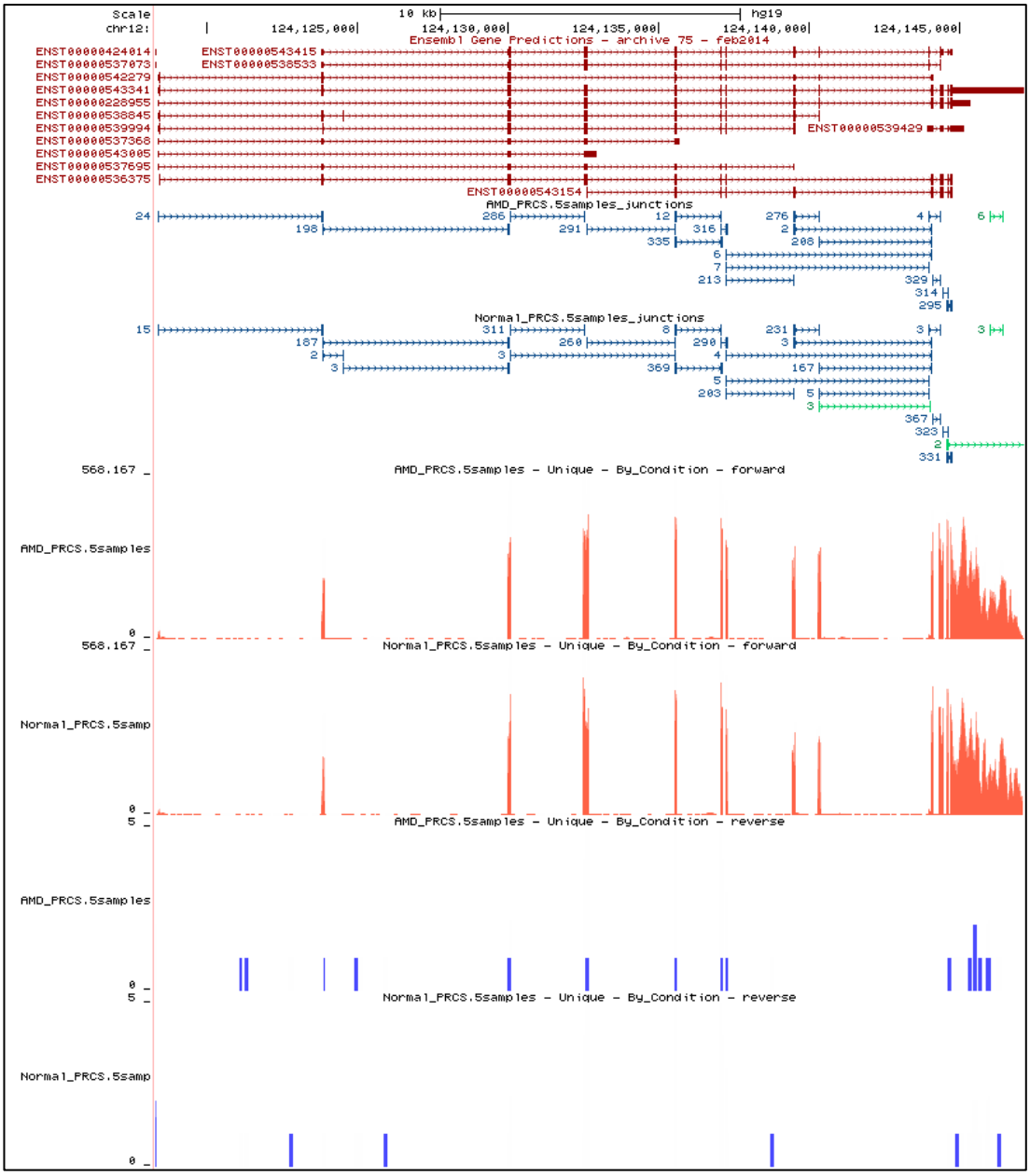

Figure 7.8: *GTF2H3* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue.

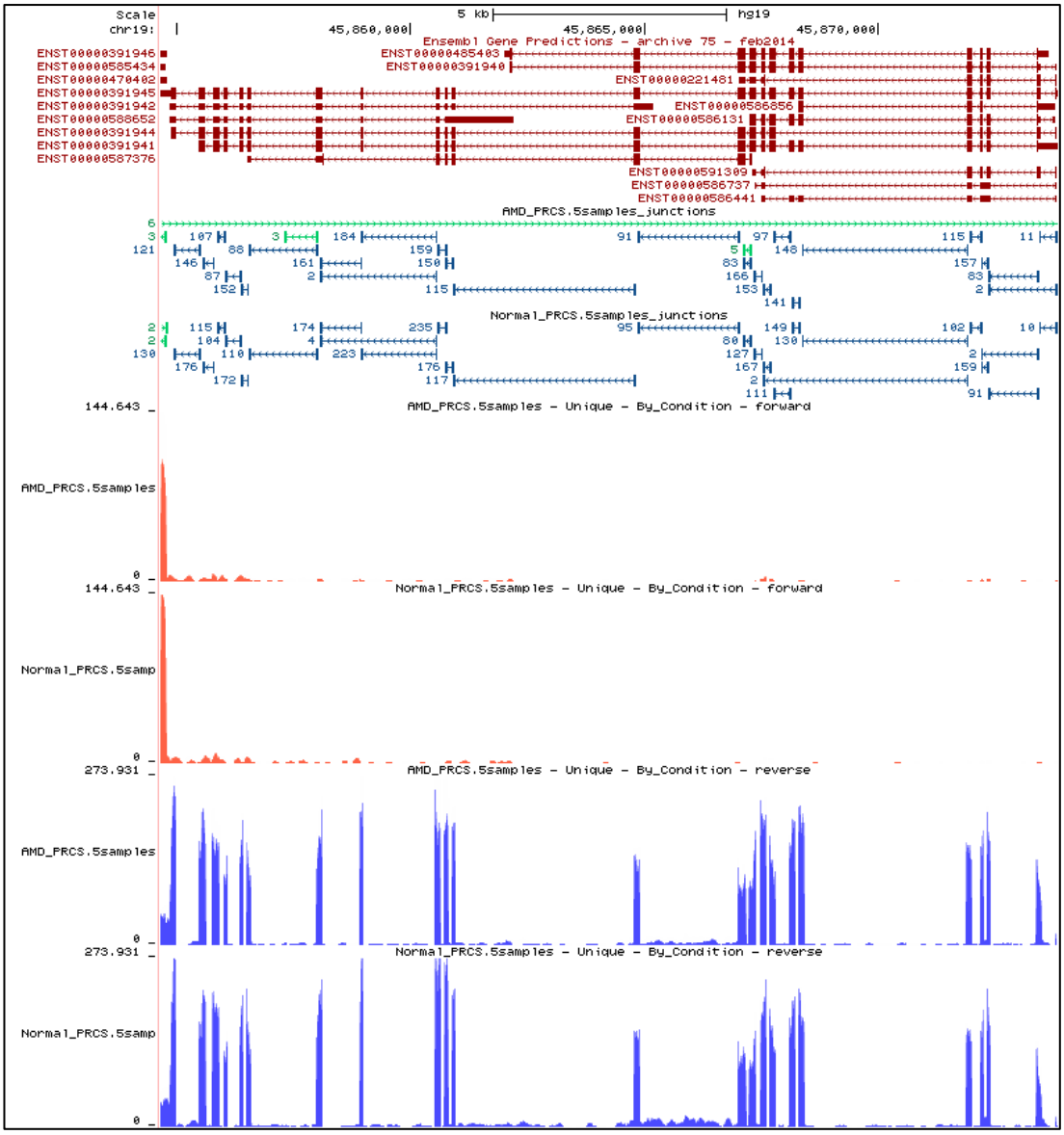

Figure 7.9: *ERCC2* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue.

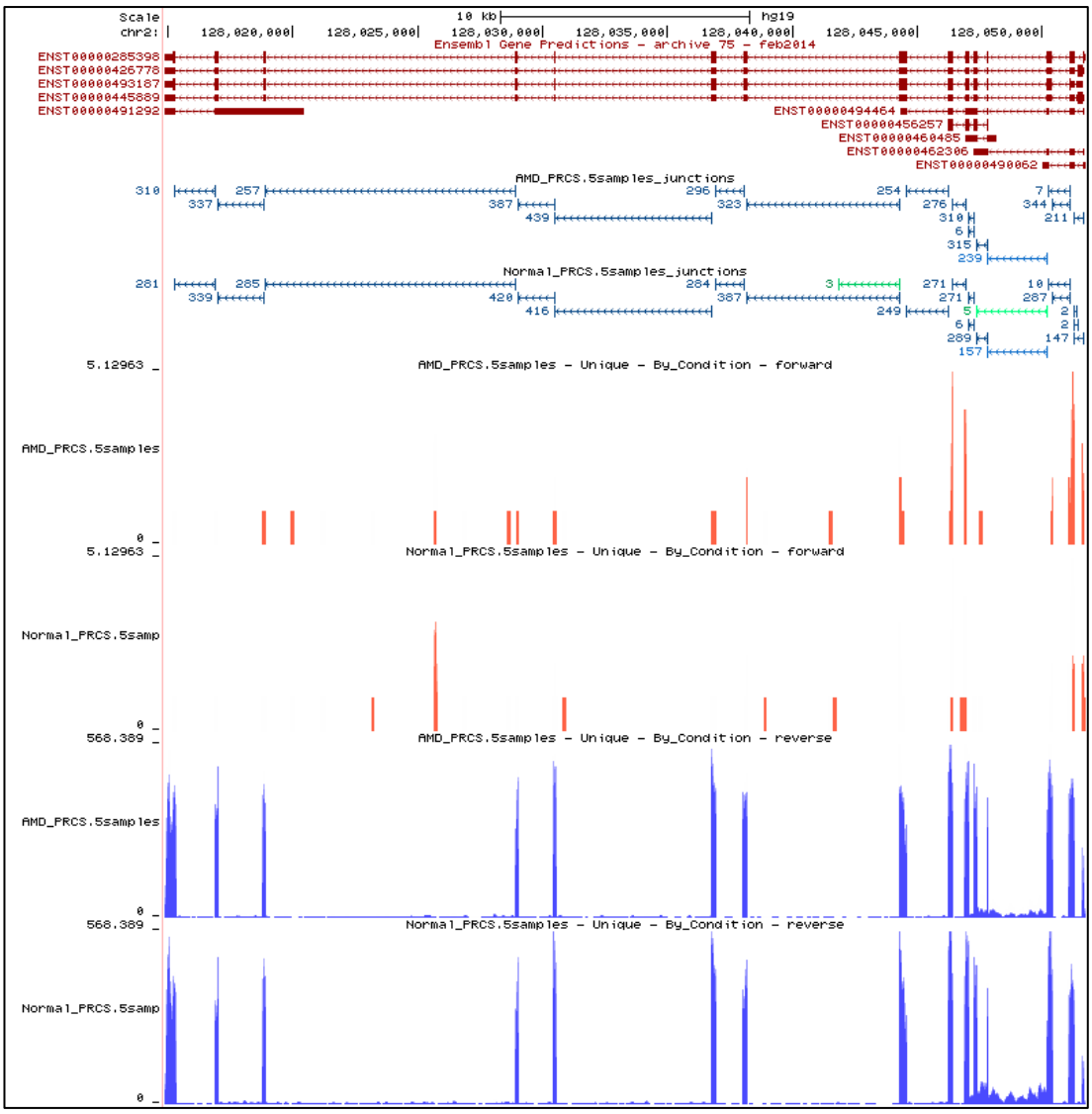

Figure 7.10: *ERCC3* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue.

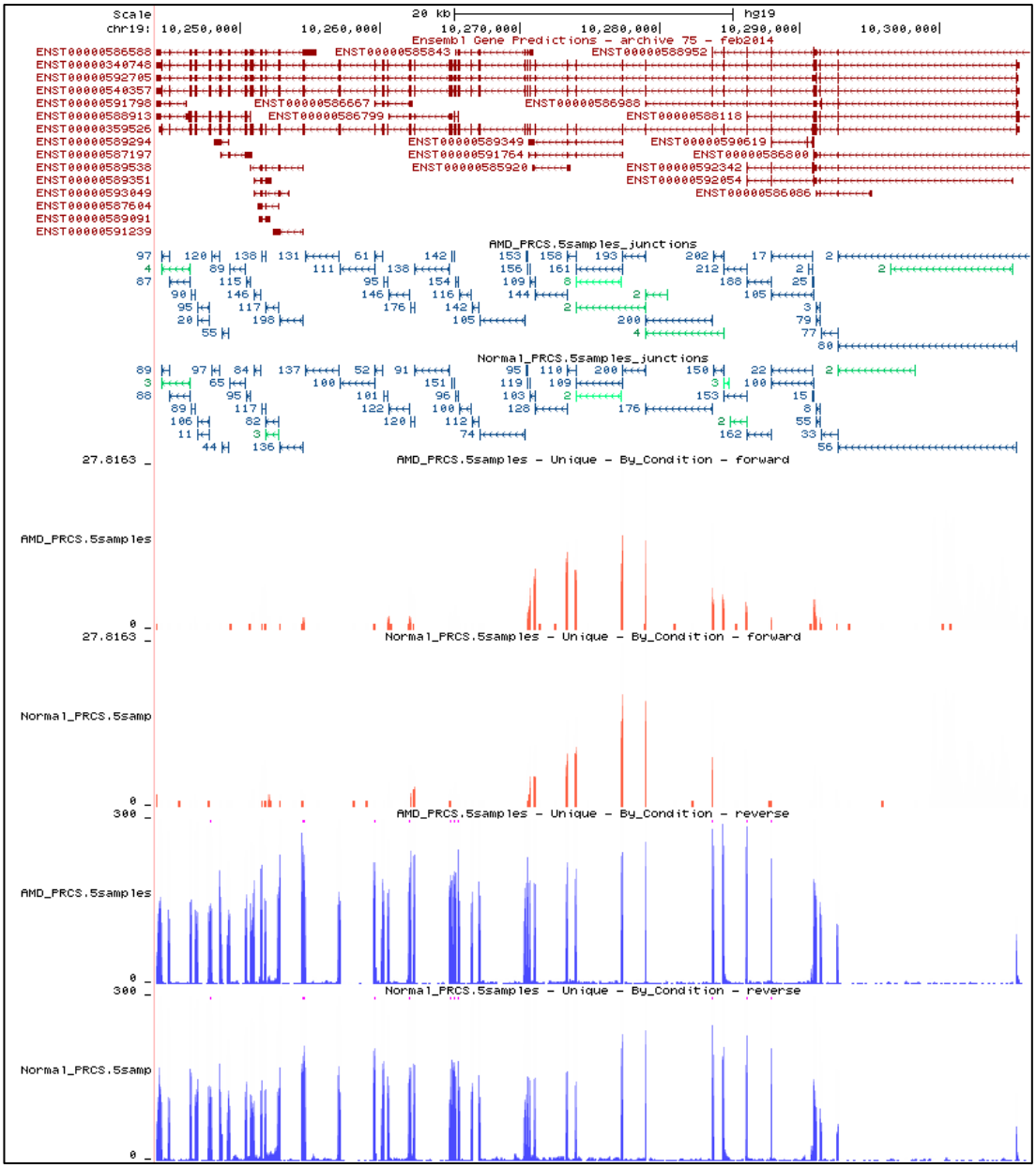

Figure 7.11: *DNMT1* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue.

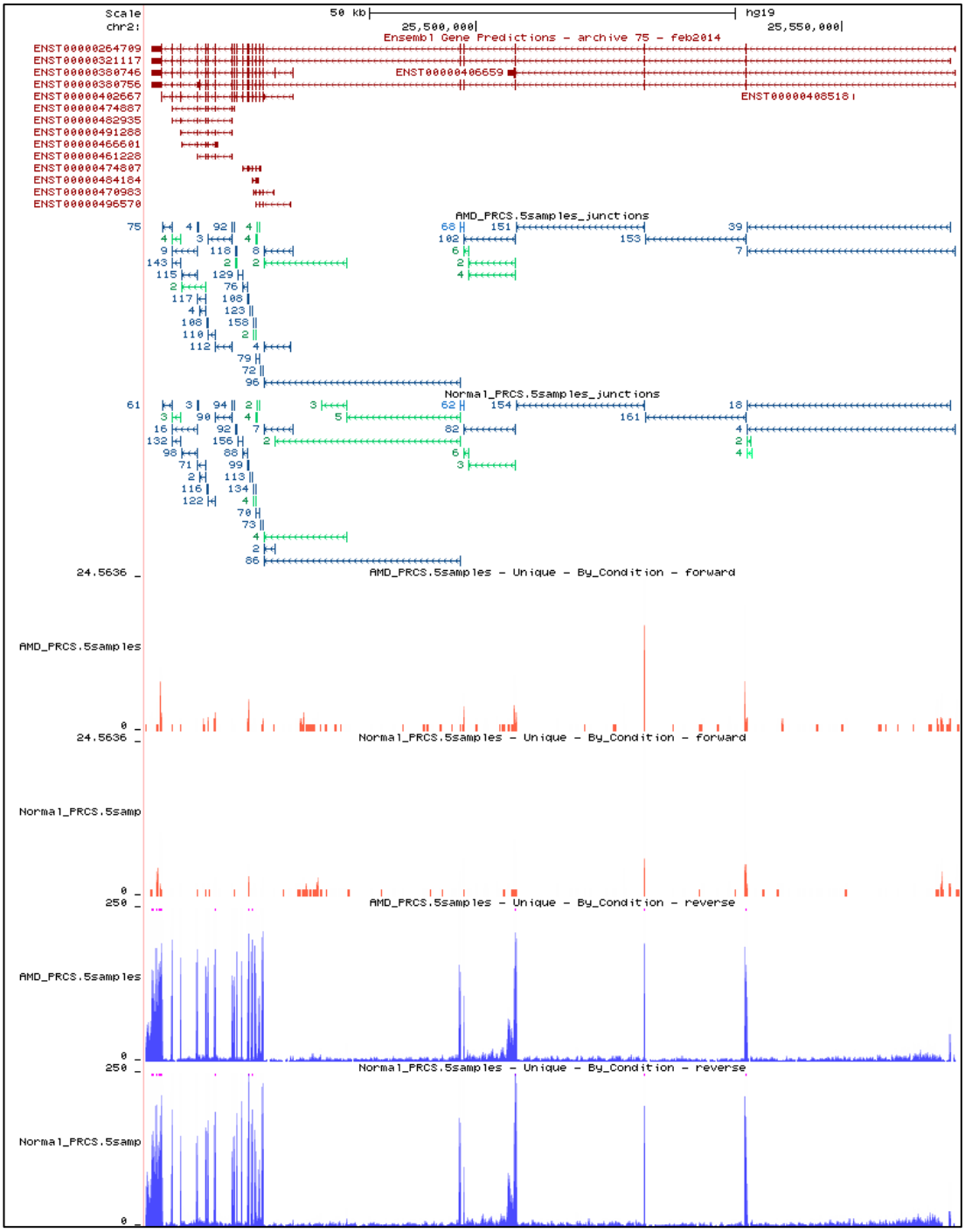

Figure 7.12: *DNMT3A* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue.

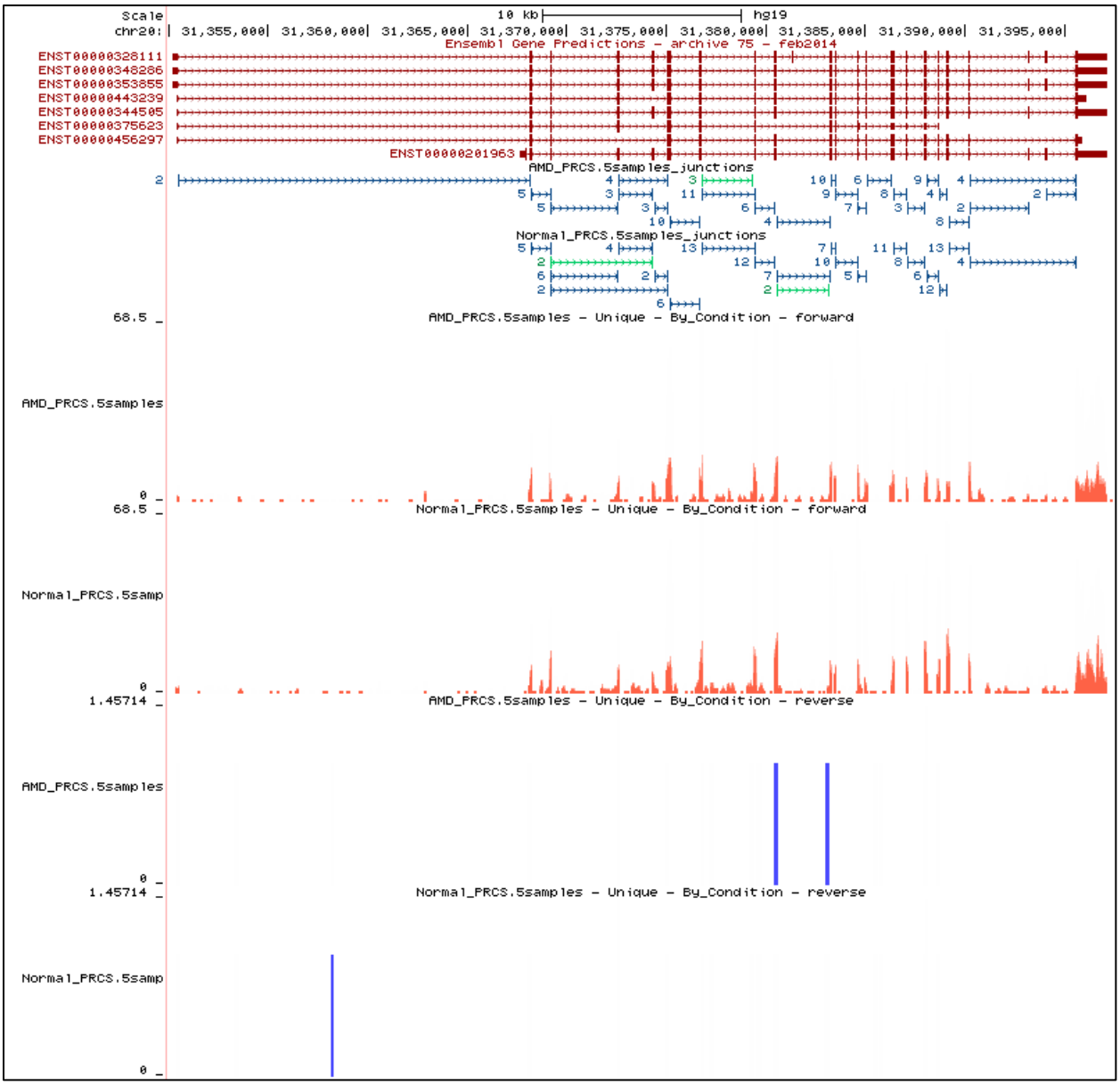

Figure 7.13: *DNMT3B* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue.

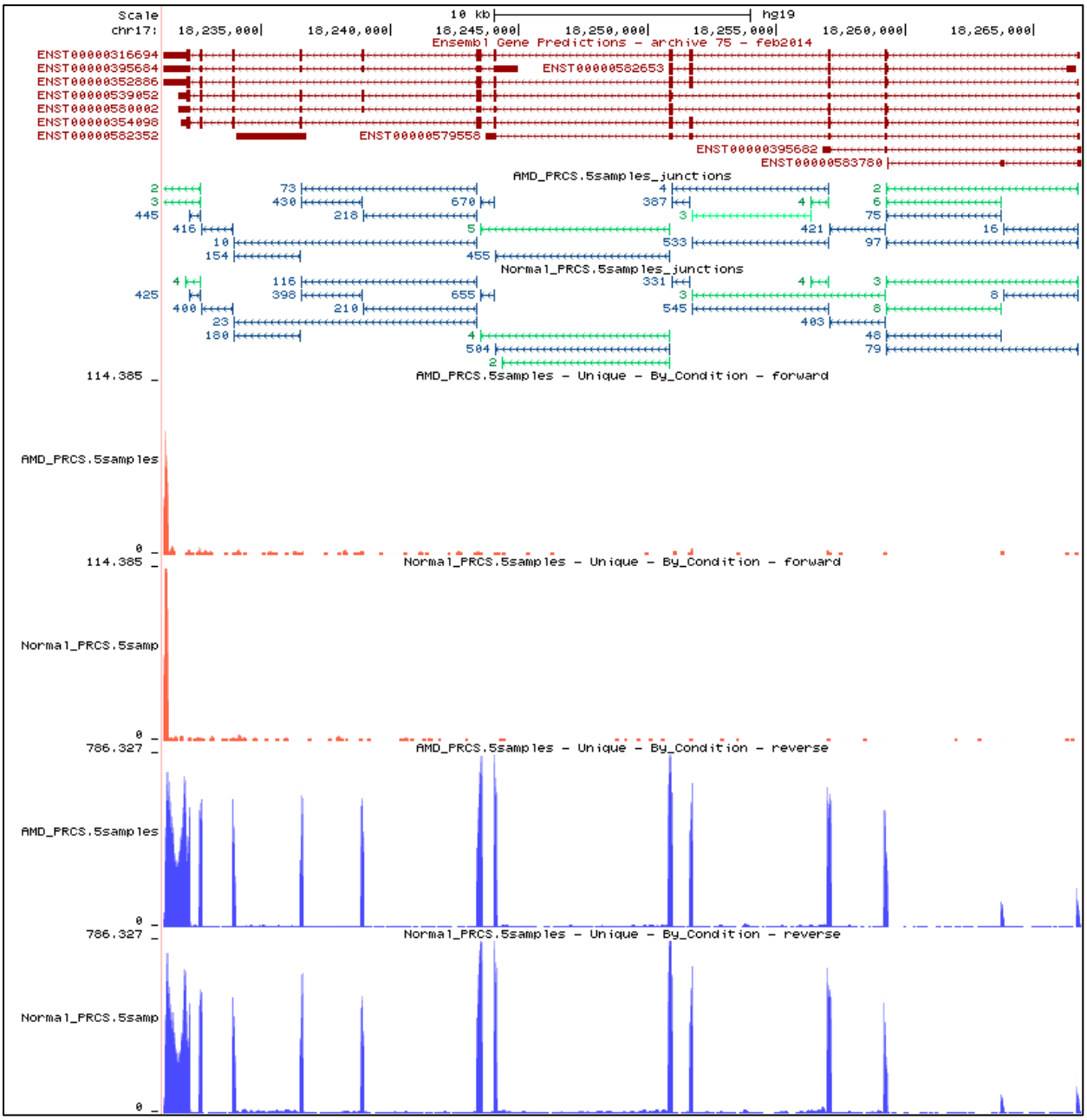

Figure 7.14: *SHMT1* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue.

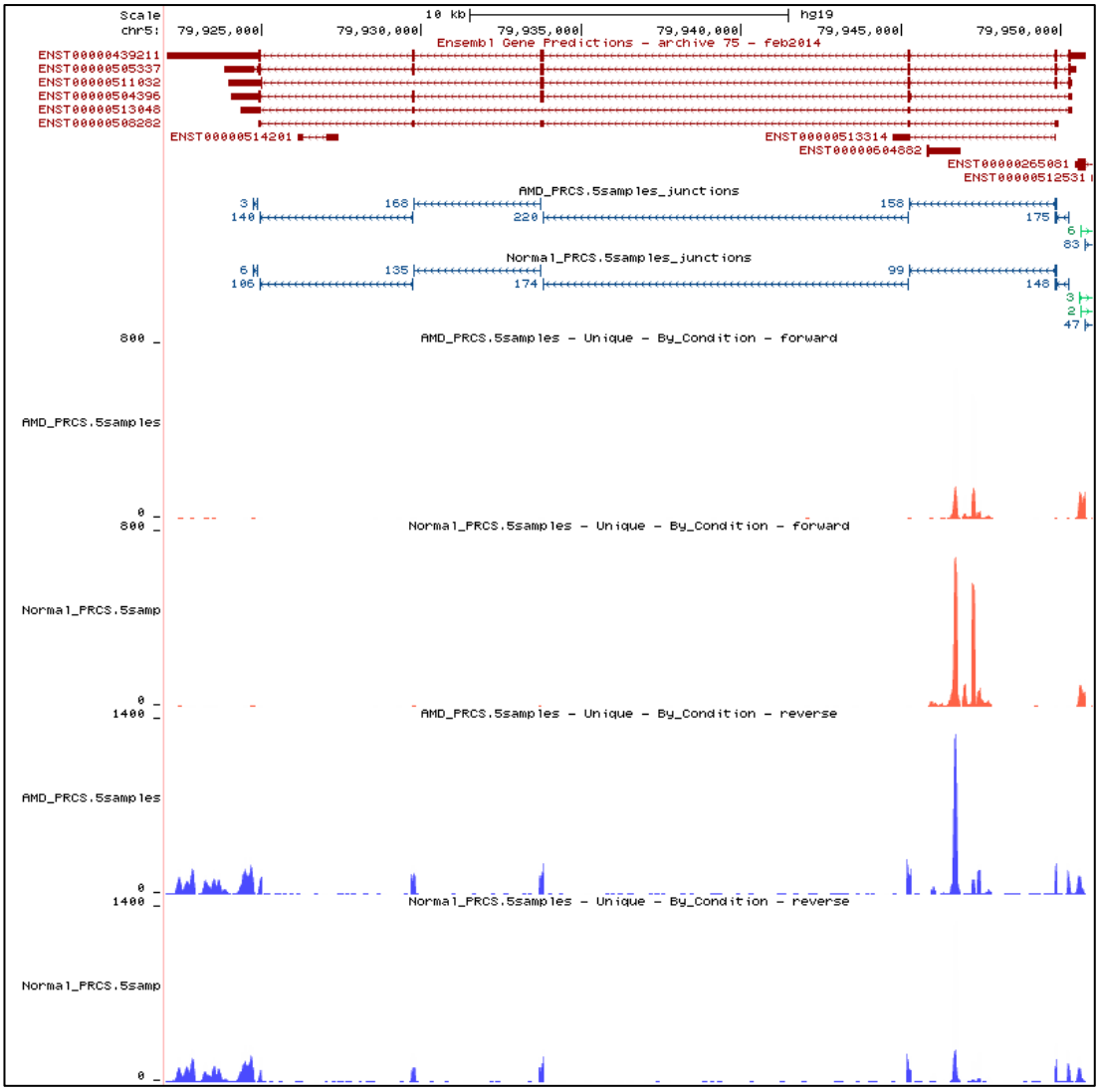

Figure 7.15: *DHFR* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue.

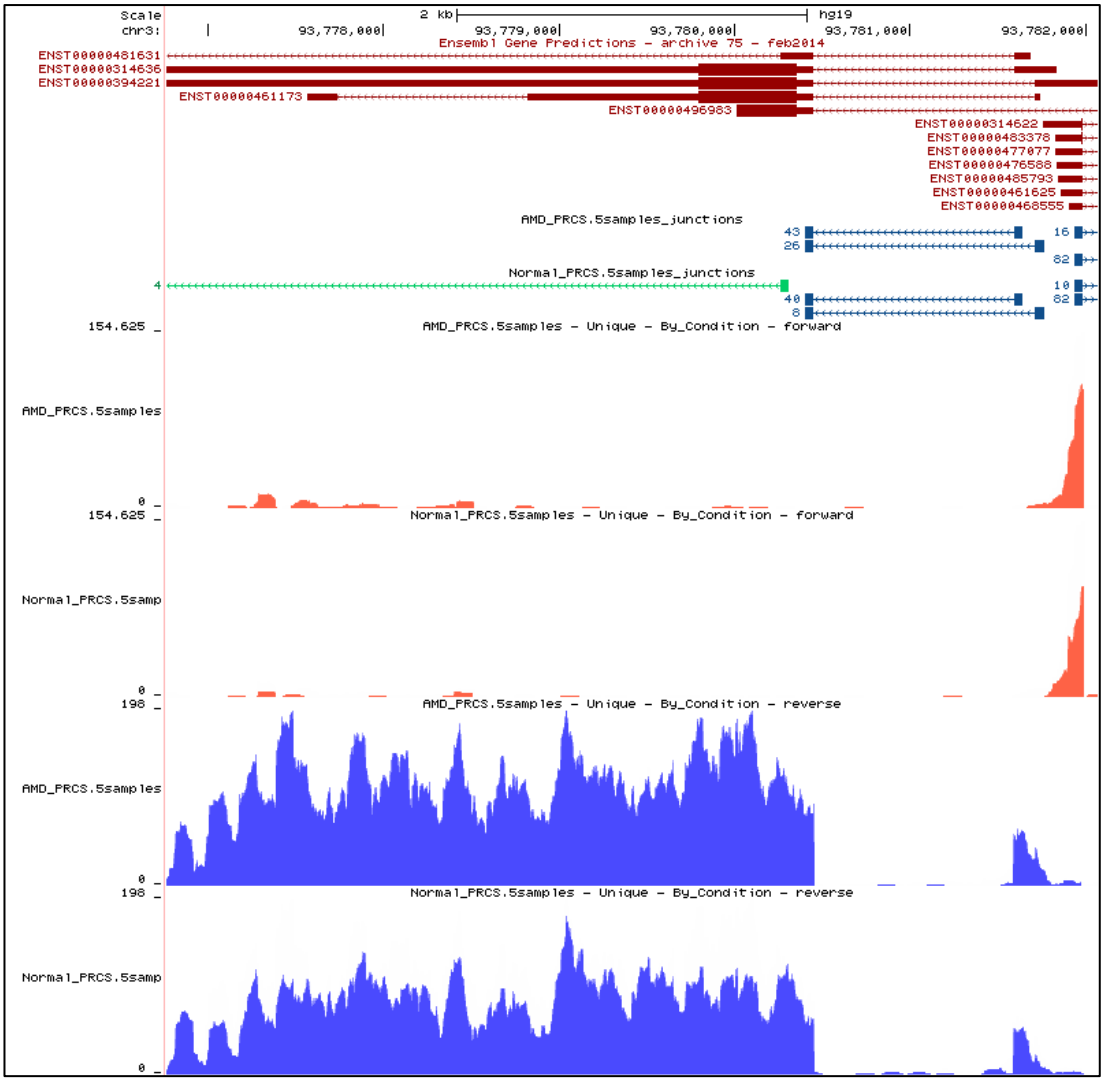

Figure 7.16: *DHFRL1* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue.

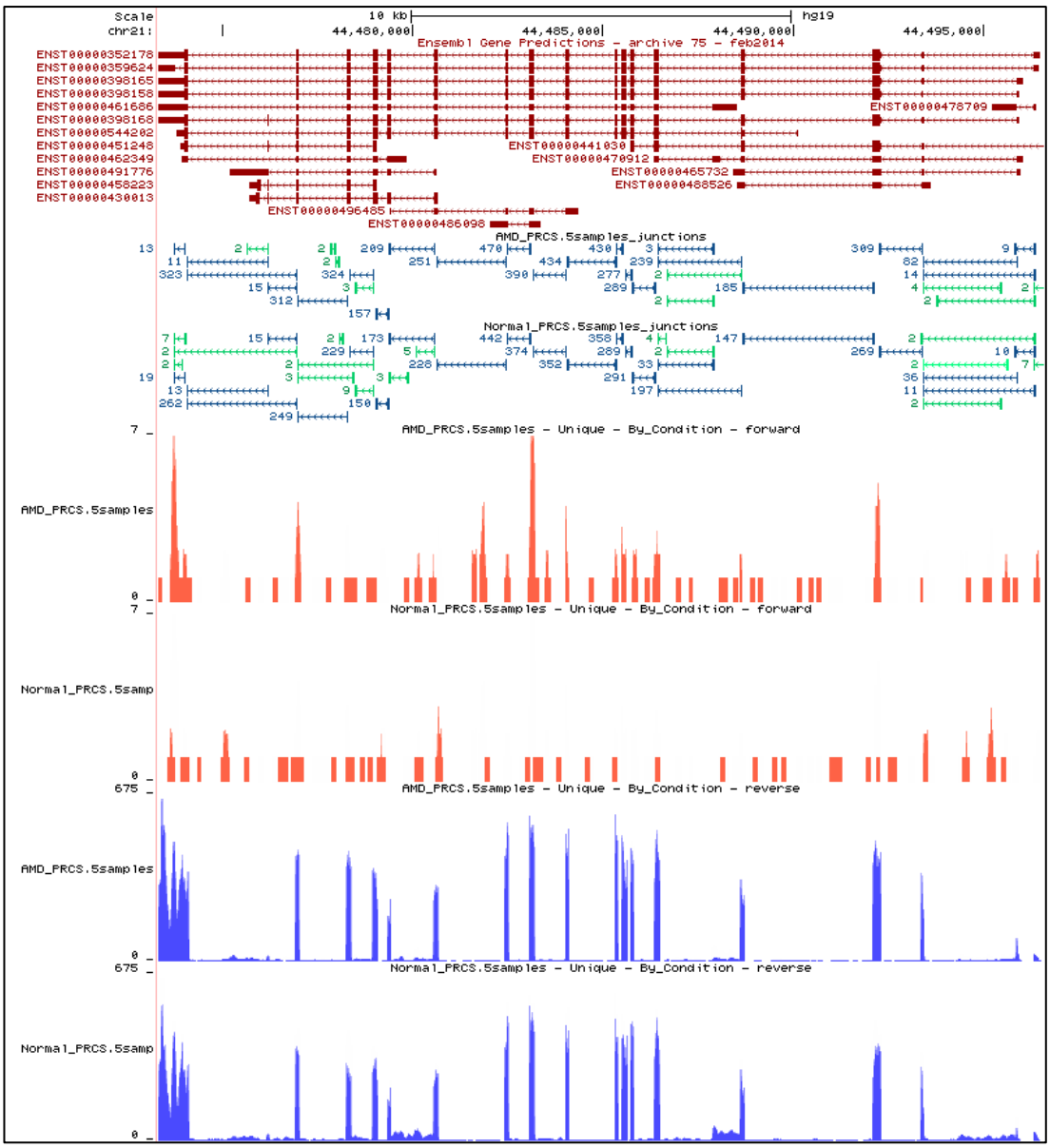

Figure 7.17: *CBS* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue.

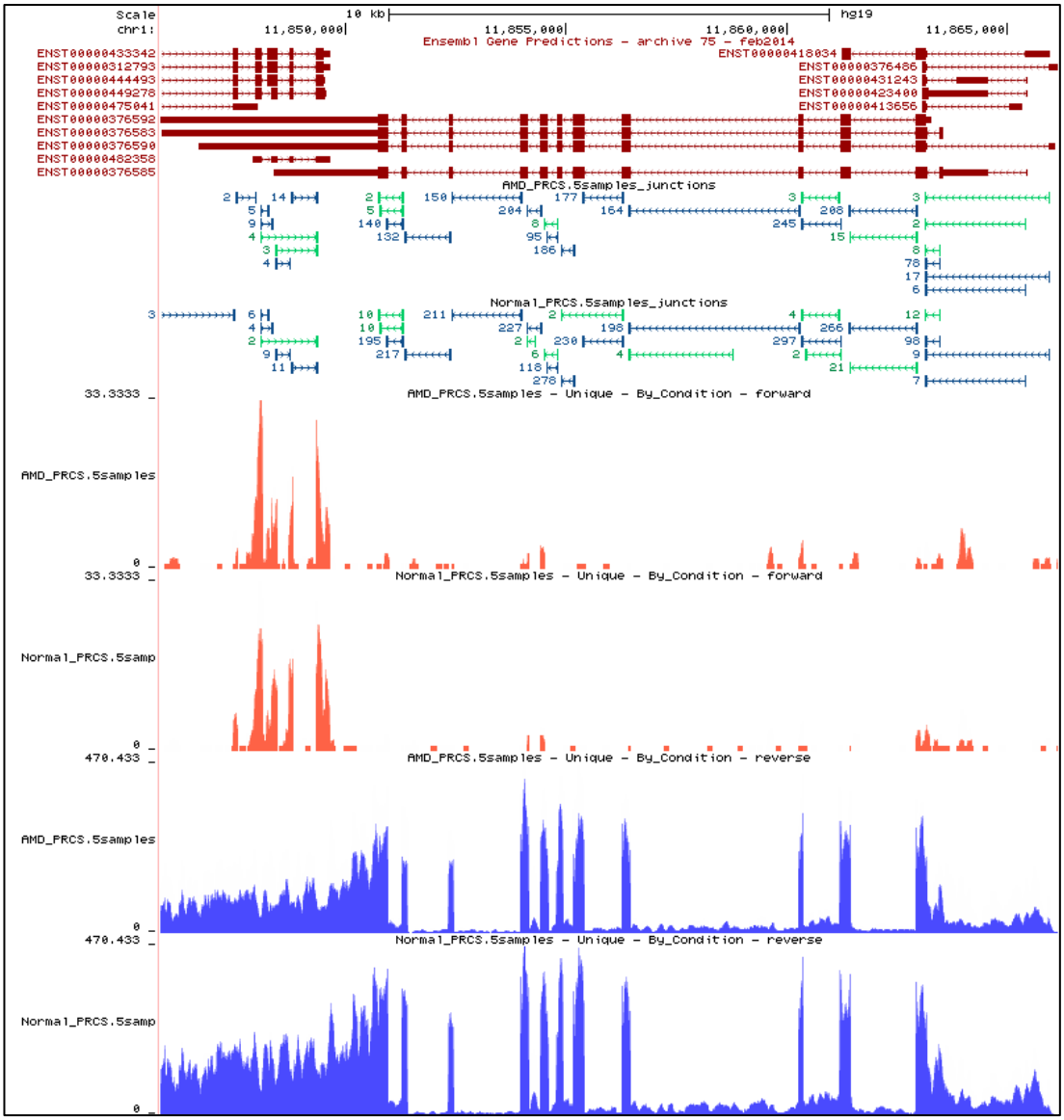

Figure 7.18: *MTHFR* RNA sequencing data from Kim et al (2017). Alternative gene transcripts derived from ENSEMBL are displayed in red. RNA sequencing data from Kim et al (2017) includes forward and reverse strands in both control and AMD PRCS (peripheral RPE / choroid / sclera) tissue.

Recipes and Reagents

10x stock TBE recipe used: 108g tris base, 55g boric acid, 40ml 0.5M EDTA Ph 8.0, ddH_2O to 1L