**Epigenetic modification of mitochondrial genes in Alzheimer's disease (AD)**

Submitted by

Matthew Adrian Michael Devall

to the University of Exeter

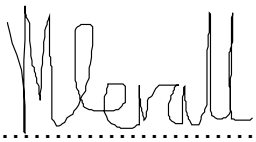
as a thesis for the degree of

Doctor of Philosophy in Medical Studies

in April 2017



I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature  .....

## Abstract

Alzheimer's disease is a chronic, neurodegenerative disease characterised by amyloid plaque accumulation, neurofibrillary tangles and eventual neuronal cell loss. The complex aetiology exhibited in late-onset Alzheimer's disease presents a considerable challenge in the field of genetics, with identified variants from genome-wide association studies collectively only explaining about a third of disease incidence. As such, new avenues are being explored to elucidate underlying mechanisms associated with disease onset and progression.

In 2014, the first epigenome-wide association studies in Alzheimer's disease were published, identifying several, novel differentially methylated loci in the nuclear genome in cortical brain samples, highlighting that epigenetic mechanisms may play a role in disease aetiology. Further, a growing body of evidence has implicated mitochondrial dysfunction as an early feature of disease pathogenesis. Despite this, few studies have investigated the role of mitochondrial DNA epigenetics in Alzheimer's disease. Indeed, the relatively nascent field of mitochondrial epigenetics has largely been restricted to candidate-based gene approaches to identify differential methylation associated with disease.

The main aim of this thesis was therefore to design an experimental and bioinformatic pipeline for the analysis of mitochondrial DNA methylation in post-mortem human brain tissue; first in healthy non-demented control donors, and

subsequently in individuals with Alzheimer's disease. Our work therefore represents the first epigenome wide studies of mitochondrial DNA methylation at single nucleotide resolution, providing a framework not only for mitochondrial DNA methylation in Alzheimer's disease, but also in a number of complex diseases characterised by mitochondrial dysfunction.

## Acknowledgements

"*I drink and I know things*" Tyrion Lannister.

Never have truer words been spoken!

I'm sure at this time, I should write much more sensible and deeper quotes that have inspired me through my PhD, quotes that I turn to when times are tough, but in truth, quotes don't inspire me. The people that I acknowledge below however have made my time in Exeter, and now Utrecht, something that I will never forget!

**First and foremost, I would like to thank my parents, my grandfather and my sisters for their continued love and support.** There's nothing more inspiring in this world than having a fantastic role model to look up to, and in my father, I have been blessed. Never have I met a man who has worked so hard to provide for their family and to care so wholeheartedly about their wellbeing. My best bud, it looks like we have made it across the finish line! In many ways, it still feels so recent that I left for Exeter and was left wondering, 'well now what?' But you have always been my centre, even if we don't always see eye to eye! I thank my mother for keeping me company on any walk that takes more than five minutes, if only my phone bill was as loving to me as you have been! The amount of times you must have wondered over the years if you could make it through the day without hearing me moan about one 'crisis' or another must be staggering! I thank my sisters Ellen, Charmayne and Ashleigh for always

being there when needed, in particular Ashleigh whose regular visits always makes sure that I was never more than a month or two without an oatcake! Leaving the country has been especially stressful at times, but your warm words and care have never been far from my heart.

**My warmest thanks go out to the epigenetics group and to my PhD supervisors, Katie, Jon and Mike.** Without their kind support and honest words, I would not be sitting here, writing this. When given the chance to reflect, I feel that the drunken stupidity, the annual midweek breakdowns, the lengthy 'meetings' in Mike's office spent discussing the world and the benefits of internet dating, have given me the strength, or sometimes just the ventilation needed to continue driving forward. I am very grateful to each of you for the generosity, understanding and confidence that you have shown me over the years and I hope I haven't been too much of a pain!

**In addition, a more personal thank you should really be extended to Katie!** I would like to thank you Katie from the bottom of my heart for all of the hard work you have put in throughout the last three and a half years. Not only has your support stopped me from falling to pieces on a professional level, but you have also been there on a personal level on more than one occasion, far exceeding your duties as a supervisor.

**To my best friend Myles,** I thank you for all that you have ever done for me, and for coming up for a question to ask at 'that induction'. I thank you for not

giving up on me when many others would have. When I first arrived, my social anxiety was quite easily at an all-time high and your continued persistence for me to join the group for a drink will always remain one of the most important social events in my life. From that Tuesday evening in Firehouse, I knew that I had found a friend, a brother, for life. Since then I have had so many good memories to treasure, whether it be Pizza Sundays in the apartment with Hannah, dragging you to Timepiece or getting lost in Prague, I have been lucky to know you. I don't think that I will ever forget the sheer fear of that Prague mishap and the embrace I was taken into when I finally found my way back, scarred by what I had seen! It seems odd to have to summarise a brotherhood in such a small space and I do not do you justice with this, but I love you my friend and I look forward to knowing you for so many years to come.

**Jenny,** my dearest friend and tea lover! I knew you were less quiet than you pretend to be from the first day that I met you! I am sorry that I am seemingly very poor with directions to Mosaic, I really had no excuse for that. I have missed our prosecco and movie nights in recent months. I cannot watch Forrest Gump now without thinking of you, nor Big Hero 6, nor Emperor's New….well you get the idea. With four sisters already, you wouldn't think that I would find another, but yet, here you are! In all of my time in Exeter, I have never met a sweeter, more honest individual who just understands, and yet still, surprisingly accepts me, and for that, thank you! Perhaps now that I have left, you may finally get chance to carry on with some work, absent of regular coffee breaks, over-exuberant hugs and 'surprising' dilemmas; because quite frankly, no one says 'what did you do now' with as much frequency and despair as you have to

me since you arrived. Thank you for always being there and for your continued advice and support.

**To Jeff, someone who has always had my back!** I don't think I've ever met anyone who jumps into the geopolitical landscape of the world or how the planet's current state of affair was brought about by an unbroken chain of events that was determined by the very first interactions to ever be had. I must admit, I'm usually a couple in by this time and so I probably missed quite a bit of what you actually strove for in your continued attempts to convert me to determinism, but it has always been fun to listen to nonetheless! Looking back on my time in this city, I really can't say that I would have come close to enjoying it without a spontaneous trip to the pub or a round of 'who loves Marvel the most', which I always inevitably win of course! Thank you for helping me get to this point, now keep away from my sister!

**Joe and Nyree, also known as my adopted parents!** Thank you for allowing me to (almost) conquer my unease around snakes, and to opening my eyes to the pagan culture. I very much hope that I can tag along to Edinburgh in the near future to see it in all its glory! Of course, I could just go to Glastonbury….Moving on from those F's and C's, thank you for showing me the true Devon, my final summer here has given me some fantastic experiences, I even got to spear fish, kind of, (thank god for sausages!). I'm very glad that in return I introduced you to Cards Against Humanity, and you plundered their reserves for every card that ever existed! I will miss our irregular coffee and lunch run (drives) and all of the wisdom that you have imparted upon me over

said times. Lord knows it must have been like banging your head against a brick wall at times! I look forward to seeing you both in Utrecht next year!

**Baz,** I still think that I'm scared of you! I'm pretty sure the first word you said to me would have been subtitled as an 'indiscernible grunt'. But as I stifle past this, I thank you for being there to help me troubleshoot, for being the only other person who seemingly like's football in Devon, (even if that does make you an Arsenal fan), and for the long discussions on Star Wars, Marvel, DC and general gaming. Importantly, Barry you will always hold the honour of being the person to introduce me to the Witcher, and whilst I've never cared much for cards, if Gwent was a real game, I would now be bankrupt! Thanks for everything!

**Janou,** it's quite incredible that the first time you were in Exeter, we rarely spoke. Thank God this changed the second time. Your warmth and generosity made my last few months in Exeter truly happy, despite even the terrible attempt at a German market. Going to tar barrels in Ottery St Mary remains one of my happiest evenings in my time here, despite the million and one people that were there. I just hope that no matter how lost we get in the future, we will just be one 'cacaw' away for each other.

**Bex,** I actually still think I'd be doing the MeDIP paper now if you hadn't have arrived when you did! It still makes me laugh when I think about the time you were so shocked that I called you a nerd, I mean Bex, come on! Thank you for

introducing me to board games, even if I must be silent whilst playing them!! Here was I thinking that all board games involved some exchange of money for properties around a board, how wrong I was. I have always known where I stand with you duck, (usually outside of punching distance) and that is something that I have always appreciated! Also, you make a mean bacon sandwich.

**Marta and Eleni,** you have both made my first months in Utrecht very interesting in your own ways. To move to a new city whilst writing a thesis and starting a new position left me in a place of fear and doubt. You have both been there to provide support, to enjoy in the good times, to keep my glass half full (with beer, or coffee (or both)) and to make me feel so very welcome. For this, I thank you all.

**Abhi,** ah Abhi, for without you I would about to become homeless! Is there anything more to say? Well yes, there is much. I'm very happy to have met you, to have been forced into the motherlands of nowhere (Cinemec) to watch one of the worst films of my life (Kong) and to enjoy my first shisha and first ziplining. WTFN. I really hope that we make sure to do our trip of Europe next year!

**Gio,** you've quickly become like a brother to me in recent months, largely because we spend about 8 hours (sleeping) apart before you have to endure my presence further! It has been great to have a gym buddy back, to go out for random drinks and be introduced to whole new genres of music. It isn't often

that you meet someone with so many different passions and such varied interests. A conversation with you is never anything other than entertaining. I wish you all the best of luck in the future for getting your own Ph.D, but more than that, I look forward to many trips to experience true Italian culture with you!

**Carlos:** you my friend, are an enigma! You introduced me to nightlife in Utrecht and Amsterdam, and no matter when (if) I ever read this again, I will probably still have your jeans! It's always amazing to meet someone with such a good heart who truly cares for others. I have really appreciated the last few months, introducing me to Brian and Miguel and am hoping to one day get my full initiation to 'the terror squad'!

**Stefania,** it's great to get to meet someone in the last couple of months of your thesis who is full of kindness and good spirit, who helps to get you away from the tortures of writing, and whose honesty and wisdom are so strong. In you Stefania, I met that person and whilst it's a shame our stays in Exeter didn't overlap for too long, you have made a true impression on me and I hope to see you in Utrecht very soon (and often!).

**Lastly Artemis,** for everything that you have done for me and for the courage that you have given me, I will be eternally grateful. No single person has ever had such a strong, positive impact on my life, that will never be forgotten.

# Table of Contents

## Table of Figures

## Table of Tables

## Publications Arising from this Thesis

Chapter 1: see Appendix 1 and 2.

**Devall, M.,** Mill, J. and Lunnon, K. (2014) The mitochondrial epigenome: a role in Alzheimer's disease? *Epigenomics,* **6**, 665-675.

***Devall, M.,*** *Roubroeks, J., Mill, J., Weedon, M. and Lunnon, K. (2016)* Epigenetic regulation of mitochondrial function in neurodegenerative disease: New insights from advances in genomic technologies. *Neuroscience letters.*

Chapter 3: see Appendix 3

**Devall, M.,** Smith, R.G., Jeffries, A.R., Hannon, E., Davies, M.N., Schalkwyk, L., Mill, J., Weedon, M. and Lunnon, K. (2016) Regional differences in mitochondrial DNA methylation in human post-mortem brain tissue. *Clinical epigenetics*, **In Press.**

Chapter 4: see Appendix 4

**Devall, M.,** Burrage, J., Caswell, R., Johnson, M., Troakes, C., Al-Saraj, S., Jeffries, A., Mill, J. and Lunnon, K. (2015) A comparison of mitochondrial DNA isolation methods in frozen post-mortem human brain tissue—applications for studies of mitochondrial genetics in brain disorders. *BioTechniques*, **59**, 241-246.

## Declarations

The samples used throughout this thesis were obtained from the Medical Research Council London Neurodegenerative Diseases Brain Bank (LBB), at the Institute of Psychiatry, Psychology & Neuroscience, King's College London (courtesy of Dr Claire Troakes and Dr Safa Al-Sarraj).

In **Chapters 4-6**, all laboratory work was carried out by myself under the guidance of Dr Katie Lunnon and Dr Joseph Burrage, with three exceptions: in **Chapter 4**, library preparation for sequencing of samples was performed by Dr Richard Caswell; secondly in **Chapters 4-6**, sequencing of pooled libraries was carried out by Dr Karen Moore and Audrey Farbos at the Exeter Sequencing Service and finally in **Chapter 6**, mitochondrial isolations were performed with the assistance of Miss Artemis Iatrou, a PhD student on placement with the group.

All bioinformatic and statistical analysis performed throughout **Chapters 3-6** was carried out by myself under the guidance of Dr Katie Lunnon, Dr Rebecca Smith and Prof. Michael Weedon, with two exceptions. First, **Figures 3.2** and **3.3** were generated by Dr Rebecca Smith. Second, initial Illumina Infinium450K Methylation array pre-processing and data normalisation in **Chapter 6** was carried out by Dr Katie Lunnon, Dr Rebecca Smith and Adam Smith, a PhD student in the group.

## List of Abbreviations

| | |
|---|---|
| 450K array | Illumina Infinium 450K Methylation BeadChip Array |
| 5-caC | 5-carboxylcytosine |
| 5-fC | 5-formylcytosine |
| 5-hmC | 5-hydroxymethylcytosine |
| 5-mc | 5-methylcytosine |
| AD | Alzheimer's disease |
| Aβ | Amyloid-beta |
| ALS | Amyotrophic lateral sclerosis |
| ATP | Adenosine triphosphate |
| BA | Brodmann area |
| BLAST | Basic Local Alignment Search Tool |
| BLD | Blood |
| BS | Bisulfite |
| BS-Seq | Bisulfite Sequencing |
| $Ca^{2+}$ | Calcium |
| CER | Cerebellum |
| CGIs | CpG Islands |
| chrM | Mitochondrial Chromosome |
| CTX | Cortex |
| Cybrids | Cytoplasmic Hybrids |
| DC | Differential Centrifugation |
| $DH_2O$ | Distilled Water |
| ddH2O | Double Distilled Water |
| DMPs | Differentially Methylated Positions |

| | |
|---|---|
| DMRs | Differentially Methylated Regions |
| DNMTs | DNA Methyltransferases |
| DS | Down's Syndrome |
| ECX | Entorhinal CTX |
| ER | Endoplasmic Reticulum |
| ERRBS | Enhanced Reduced Representation Bisufite Sequencing |
| ETC | Electron Transport Chain |
| EWAS | Epigenome-wide Association Studies |
| FACS | Fluorescence-Activated Cell Sorting |
| GWAS | Genome-Wide Association Studies |
| HD | Huntington's Disease |
| HDAC | Histone Deacetylases |
| hMeDIP-Seq | Hydroxymethylated DNA Immunoprecipitation Sequencing |
| HMT | Histone Methylases |
| IF | Immunofluorescence |
| IMM | Inner Mitochondrial Membrane |
| IMS | Inter-Membrane Space |
| IP | Immunoprecipitation |
| LA | α-Lipoic Acid |
| LBB | MRC London Neurodegenerative Disease Brain Bank |
| LC-ESI-MS | Liquid Chromatography-Electrospray Ionization Tandem Mass Spectrometry |
| LC-MS | Liquid Chromatography Mass Spectrometry |
| LPS | Lipopolysaccharide |
| MBD | Methyl-CpG Binding Domain Proteins |
| MeDIP-Seq | Methylated DNA Immunoprecipitation Sequencing |

| | |
|---|---|
| MELAS | Mitochondrial Encephalomyopathy With Lactic Acidosis And Stroke-Like Episodes |
| MM | Mixed Model |
| mPT | Mitochondrial Permeability Transition |
| mPTP | Mitochondrial Permeability Transition Pore |
| mtDNA | Mitochondrial DNA |
| mt-SNPs | Mitochondrial Single Nucleotide Polymorphisms |
| NASH | Non-Alcoholic Steatohepatitis |
| ncDNA | Nuclear DNA |
| NGS | Next Generation Sequencing |
| NUMTs | Nuclear-Mitochondrial Pseudogenes |
| OMM | Outer Mitochondrial Membrane |
| OXBS | Oxidative Bisulfite |
| OXPHOS | Oxidative Phosphorylation |
| PCA | Principal Component Analysis |
| PD | Parkinson's Disease |
| QQ | Quantile-Quantile |
| PCR | Polymerase Chain Reaction |
| qRT-PCR | Quantitative Real-Time Polymerase Chain Reaction |
| RLGS | Restriction Landmark Genomic Scanning |
| ROS | Reactive Oxygen Species |
| RPKM | Reads Per Kilobase Of Transcript Per Million Mapped Reads |
| RRBS | Reduced Representation Bisulphite Sequencing |
| rRNA | Ribosomal RNA |
| SAM | S-adenosyl L-methionine |
| SMRT | Single-Molecule Real-Time Sequencing |

| | |
|---|---|
| SS | Simple Steatosis |
| SNP | Single Nucleotide Polymorphism |
| STG | Superior Temporal Gyrus |
| TET | Ten-Eleven Translocase |
| TF | Transcription Factor |
| TFR | Transcription Factor Receptor |
| tRNAs | Transfer RNA |
| VDAC1 | Voltage Dependent Anion Channel 1 |
| WGBS-Seq | Whole Genome Shotgun Bisulphite Sequencing |
| ZMW | Zero-mode Waveguide |

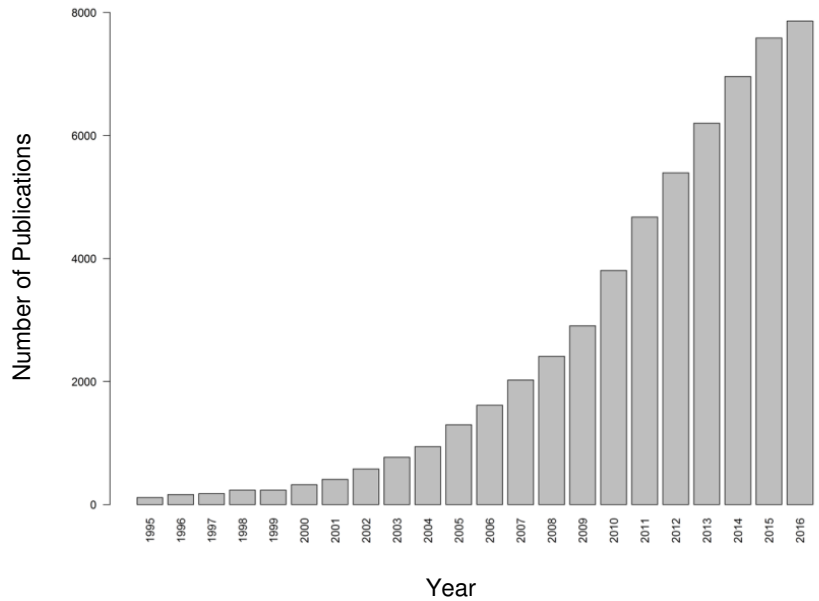
# Chapter 1. Introduction

***The work presented in this chapter is based on two published reviews; Devall et al., 2014 and Devall et al. 2016 (1,2), which can be found in Appendix 1 and 2, respectively.***

## 1.1. Epigenetics

Since initial studies in the field by Conrad Waddington (3), epigenetics has been increasingly investigated. Specifically, since the recent advances in genetic technology, there has been an exponential increase in the number of epigenetic studies published (**Figure 1.1**).

***Figure 1.1: Bar chart to show the number of publications per year in the field of epigenetics.*** *For each year, the keyword search term "epigenetic" was entered into Pubmed (https://www.ncbi.nlm.nih.gov/pubmed/) and the resulting number of publications for that year were plotted.*



Epigenetic processes mediate the reversible regulation of gene expression, occurring independently of DNA sequence and acting principally through chemical modifications to DNA and nucleosomal histone proteins. There are a number of epigenetic modifications, including phosphorylation, ubiquitination, SUMOylation, histone (de)acetylation and methylation of both DNA and protein. However, due to the relative stability of DNA methylation as a mark, and owing to the development of technologies capable of genome-wide DNA methylation profiling, this epigenetic modification is currently the most widely studied (1,4).

### *1.1.1. DNA Methylation*

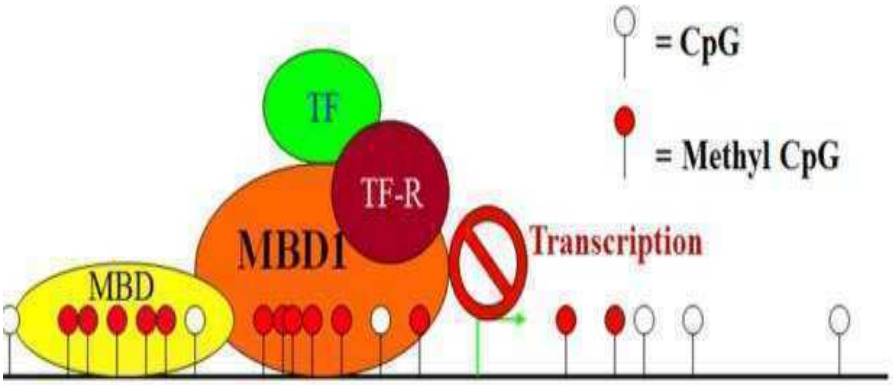
DNA methylation refers to the covalent addition of a methyl group to the fifth carbon of cytosine, resulting in the creation of 5-methylcystosine (5-mC). The majority of nuclear DNA (ncDNA) methylation in eukaryotes occurs in a CpG dinucleotide context, with 60-90% of total cytosines at these sites being methylated (5). A notable exception to the high proportion of methylated cytosines in a CpG context can be seen in CpG islands (CGIs), regions over 500bp with greater than 50% GC content. These regions are commonly associated with gene promoter regions and are generally unmethylated (6,7). DNA Methylation has been shown to be present in non-CpG contexts in somatic cells. However, modifications in this context generally appear to occur with low abundance and in specific cell types (8-10). It has been generally accepted that DNA methylation is associated with gene silencing (11,12); however, recent studies have found that this relationship is impacted by factors such as genic location of DNA methylation as well as the effect on the overlapping chromatin structure (13). For example, whilst methylation changes in regions of low-CG

content DNA at alternate promoters show a strong, inverse correlation with gene expression (14), the vast majority of CGIs within promoters are unmethylated, regardless of gene expression. Further, whilst methylation at intergenic regions has also been suggested to lead to decreased gene expression through regulation of DNA enhancers (15), gene body methylation has been shown to be positively correlated with expression in a number of studies (16).

#### *1.1.1.1. Mechanisms of DNA Methylation*

The transfer of a methyl group from the methyl donor S-adenosyl L-methionine (SAM) to create 5-mC is catalysed by a conserved group of enzymes called DNA methyltransferases (DNMTs) (17). Importantly, two classes of DNMTs exist in eukaryotic organisms. DNMT3A and DNMT3B are important for establishing *de novo* methylation and, as such, are highly expressed in cells during processes involving differentiation and development (18), whereas DNMT1 plays a fundamental role in the maintenance of hemi-methylated DNA during semi conservative replication and is ubiquitously expressed regardless of a cell's spatiotemporal state (18). Methylated DNA is then 'read' by a number of different protein classes, with the most commonly researched being the methyl-CpG binding domain proteins (MBD), which affect gene expression through their interactions with histone and chromatin modifying proteins, typically leading to a closed chromatin structure and reduced transcription factor binding (**Figure 1.2**) (19).

***Figure 1.2: An overview of DNA methylation machinery.*** *Methylated cytosines at a CpG dinucleotide recruit MBDs, which bind preferentially to single methylated cytosines and contain a transcriptional repressor domain, aiding in gene silencing. Figure taken from (20). Abbreviations: Methyl binding domain proteins (MBDs), Transcription factor (TF), Transcription Factor Receptors (TF-R).*




*1.1.1.2. Other Methylation Marks*

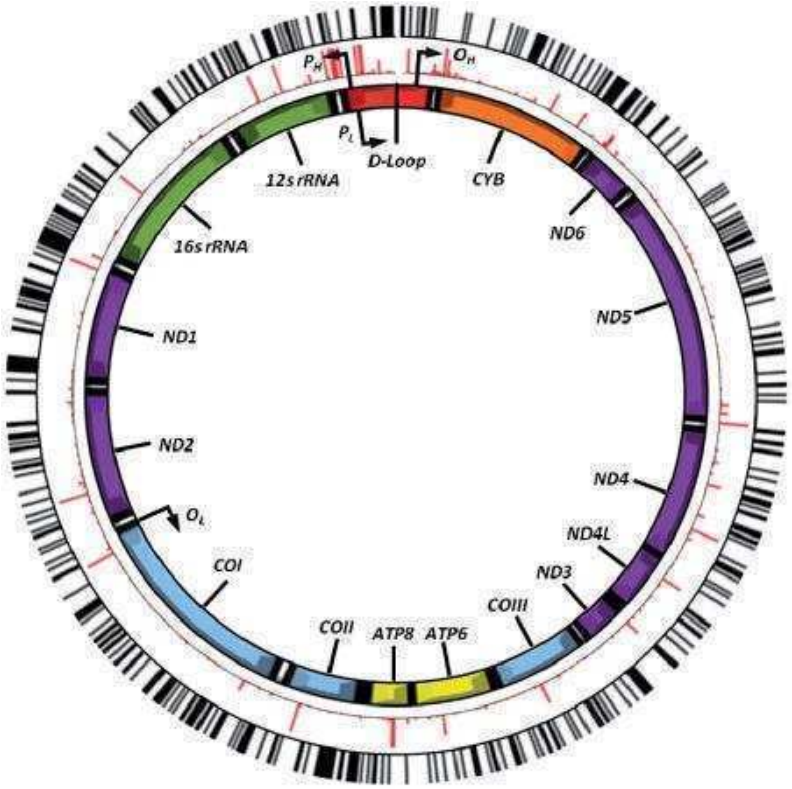
A number of technologies can be used to interrogate ncDNA methylation; the most commonly used in epigenetic studies are bisulfite (BS) based approaches (2). Briefly, this methodology allows for the distinction of methylated and unmethylated cytosines through the addition of BS. This converts unmethylated cytosines to uracil, which are then converted to thymine after a further desulphonation step. Methylated cytosines remain unaffected by this process and can therefore be distinguished based on their sequence and the base pairing of the now modified thymine. One limitation of using BS based sequencing or microarray based approaches are their inability to distinguish

between 5-mC and another DNA modification, namely 5-hydroxymethylcytosine, (5-hmC) (4). The potential role of 5-hmC as an independent epigenetic mark was initially met with controversy after a number of early contradictory findings (21). However, the discovery of Ten-eleven translocation 1 (TET1), which catalyses the oxidation of 5-mC into 5-hmC, has led to a number of studies investigating its role in the epigenome (22), particularly as 5-hmC appears to be enriched in the brain (23). Additional DNA modifications have also been identified, namely 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC), which are thought to be intermediates in the DNA methylation pathway, as levels of both modifications are far lower than 5-mC and 5-hmC, (see (24) for review).

## ***1.2. Mitochondria: The Powerhouse of the Cell***

Being the site of adenosine triphosphate (ATP) generation, the mitochondria provide the cell with the energy required to properly function; as such mitochondria are often described as 'the powerhouse of the cell'. Mitochondria are cylindrical organelles containing ~16.6kb of DNA (mtDNA) (25), which is separate to the nuclear genome and inherited in a maternal, non-Mendelian fashion. The mitochondrial genome consists of 37 genes, 13 of which encode for polypeptides required for the electron transport chain (ETC), in addition to two ribosomal RNAs (rRNAs) and 22 transfer RNAs (tRNAs) (**Figure 1.3**). MtDNA also exists as a multi-copy genome, with 2-10 copies of mtDNA per mitochondrion, and multiple mitochondria per cell (26).

***Figure 1.3: The structure of the mitochondrial genome**. Shown is the mitochondrial genome (ChrM), with the position of protein coding ETC genes and two rRNAs highlighted. Known mtDNA variants are shown in red and black lines highlight the predicted CpG sites. $P_H$ and $P_L$ represent the heavy and light strand promoter regions and $O_H$ and $O_L$ represent the origins of heavy-strand and light-strand replication respectively. Image taken from (27).*

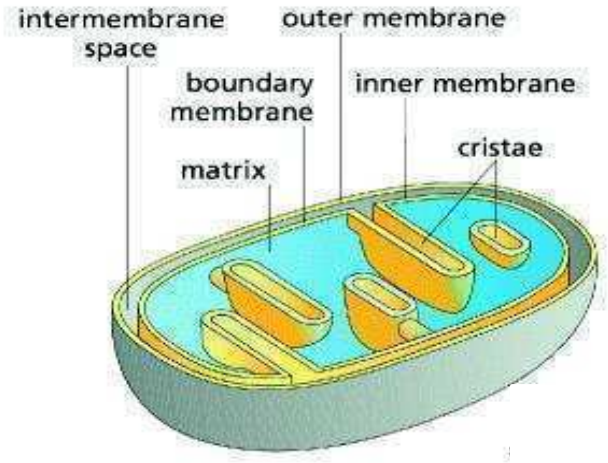



## ***1.3. Mitochondrial Structure***

The mitochondria are made of five distinct components, the cristae, matrix, inner mitochondrial membrane (IMM), inter-membrane space (IMS), and the

outer mitochondrial membrane (OMM), with the OMM acting to separate the IMS from the cytosol (28) (**Figure 1.4**).

***Figure 1.4: Diagram to show the inner compartments of the mitochondria**. The dual membrane structure of the mitochondria separates the matrix from the cytosol. Folds of the IMM form the mitochondrial cristae, which provide a greater surface area for aerobic respiration and thus, ATP generation to occur. Image taken from (29).*




Transport into the IMS, through the OMM is mediated by two pathways and is dependant upon molecular size. Small molecules less than 6kDA can pass through the OMM using porins, which are beta barrel proteins present on the membrane surface (28); whilst molecules greater than 6kDa can only enter after binding to translocase proteins on the OMM (30). Mitochondrial cristae are folds in the IMM which provide a greater surface area to facilitate the chemical reactions that occur during ATP production. Being the site of oxidative phosphorylation (OXPHOS), cristae shape has been found to have dramatic

effects on the kinetics of chemical reactions occurring in the ETC (31). The mitochondrial matrix plays a vital role in calcium ($Ca^{2+}$) buffering, which is required for correct neuronal transmission. Increased levels of $Ca^{2+}$ in the mitochondrial matrix leads to increased levels of Krebs cycle enzymes and increased ATP production. Further, mitochondria can also aid in the regulation of $Ca^{2+}$ levels, for example through the opening of the mitochondrial permeability transition pore (mPTP). As such, mitochondria are critical determinants of neuronal stability, not only through ATP production, but also through their ability to buffer calcium (32), with changes in correct mitochondrial function often accompanied by neuronal dysfunction and neurodegeneration (1).

### ***1.4. Mitochondria, mtDNA and their Relationship to Complex- and Age-Related Disorders***

With the exception of mature red blood cells, mitochondria are present in every cell within the human body. However, despite the importance of maintaining correct mitochondrial function for cellular respiration, mtDNA is subjected to approximately 100x the mutagenic stress of ncDNA (33). Reactive oxygen species (ROS) are a major source of mutagenic stress and an ever-growing body of evidence indicates a relationship between the production of ROS and subsequent mitochondrial damage, associated with some of the adverse effects of ageing. It is important to note that ROS play important roles as secondary messengers between mitochondria and the rest of the cell (retrograde signalling) and also from the cytosol to the mitochondria (anterograde signalling) (34). However, the mitochondrial theory of ageing states that, over

time, an increase in ROS creates a positive feedback loop, leading to oxidative damage in the mitochondria and further ROS production. In turn, these free radicals then contribute to the development of a number of degenerative diseases characterised by mitochondrial dysfunction. (35).

### *1.4.1. Mitochondria and Cell Death*

Importantly, changes in ROS levels and mitochondrial function have been associated with both apoptosis and necrosis by altering the permeability of the mitochondrial double membrane (36). Dramatic changes in IMM permeabilisation, the mitochondrial permeability transition (mPT), are caused by the opening of the mPTP, allowing solutes greater than 1.5kDa to pass through the IMM (37). If open for a short period, this change not only allows for the release of more Cytochrome C, a component of the ETC which acts to transfer electrons from Complex III – IV (38), but also leads to further increases in ROS production. A short-term release of cytochrome C leads to the activation of caspases, permeabilisation of the OMM and programmed cell death in a pathway mediated by BCL2 Associated X Protein (BAX) and BCL2 Antagonist/Killer 1 (BAK) (39). In some cases this transition persists and spreads, leading to further depolarisation of the mitochondria. This loss of mitochondrial membrane potential has been shown to lead to the complete loss of ATP generation, given that the ETC requires an electrochemical gradient to function (37). The absence of ATP generation leads to mitochondrial swelling, changes in ion homeostasis and eventual necrosis (36). A multitude of factors are associated with the opening of the mPTP and the widespread development of necrosis across cell populations, but the two most studied are oxidative

stress and excessive calcium accumulation in the mitochondrial matrix (40). Abnormalities in apoptosis and necrosis have been associated with mitochondrial dysfunction in a range of complex diseases (40-42); however, a full mechanistic understanding about the onset of this dysfunction is still unclear.

### *1.4.2. Mitophagy, Fission and Fusion*

In response to external stimuli, such as high levels of oxidative stress within dysfunctional mitochondria, a process of mitochondrial autophagy (mitophagy) may be evoked to degrade the organelle. A number of possible mitophagic pathways have been suggested. The most extensively characterised pathway includes the mitochondrial kinase PTEN Induced Putative Kinase 1 (PINK1) (43). Changes in mitochondrial permeability lead to an accumulation of this kinase on the mitochondria, phosphorylating ubiquitin and subsequently recruiting Parkin, a ligase. Parkin serves to add more ubiquitin chains to the mitochondria. Upon ubiquitination, chains are further modified via phosphorylation by PINK1, leading to the subsequent recruitment of autophagosome components and Microtubule-Associated Protein 1 Light Chain 3 (LC3) to engage the dysfunctional mitochondria in lysosomal degradation (43,44). As such, mitophagy acts to reduce the mitochondrial aspect of ageing by removing dysfunctional mitochondria. Interestingly, genetic changes in both PINK1 and PARKIN have been associated with Parkinson's disease (PD), an age-related disease (45).

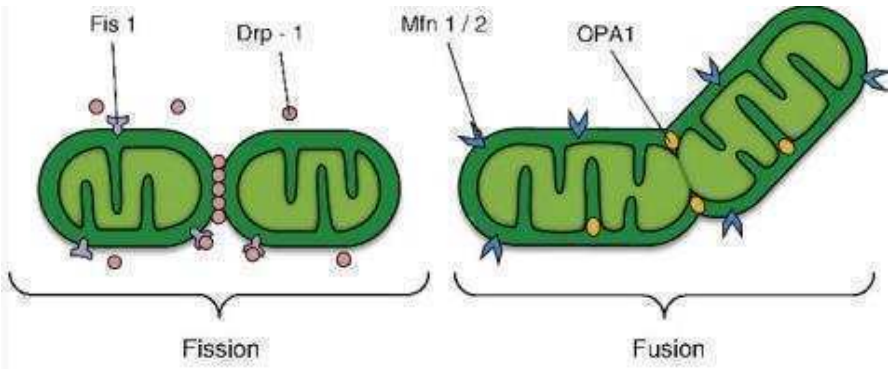
Mitochondrial quality is maintained not only through mitophagy, but also by a number of other processes, most notably mitochondrial fission and fusion. Essentially, extensively damaged mitochondria are fragmented through fission before eventual degradation by the process of mitophagy (46). Alternatively, during mitochondrial fusion, damaged mitochondria are fused with healthy mitochondria, thus reducing the overall percentage of the genetic mutational load within the mitochondria, stabilising mtDNA copy number and preserving mtDNA fidelity (47).

Mitochondrial fission is mediated by Dynamin-related Protein 1 (DRP1), a ubiquitously expressed protein, which is recruited from the cytosol and binds to Mitochondrial Fission 1 (FIS1), Mitochondrial Fission Factor (MFF), Mitochondrial Dynamics Protein of 49 kDa (MID49) and Mitochondrial Dynamics Protein of 51 kDa (MID51), which are proteins on the OMM (48,49). DRP1 is regulated by a number of epigenetic, post-translational modifications. Following binding to the OMM proteins, DRP1 oligomerises into a ring like structure, and upon constriction, divides mitochondria (49).

Mitochondrial fusion occurs at two levels, at the OMM, which is predominately mediated by Mitofusion 1 and Mitofusion 2, (MFN1 and MFN2 respectively), and subsequently at the IMM, which is mediated by Optic Atrophy 1 (OPA1) (**Figure 1.5**). In mouse embryonic cells, deletion of *Mfn1* or *Mfn2* leads to partial reduction in mitochondrial fusion, whilst a double deletion mouse model has complete loss of fusion (50). Therefore, current evidence suggests that a select number of ncDNA genes are directly responsible for adequate fission and fusion

of the mitochondrion and that dysregulation of these genes can have severe consequences on mitochondrial homeostasis.

***Figure 1.5: Overview of mitochondrial fusion and fission****. Despite their co-ordinated balance to control mitochondrial copy number, the two events are regulated by different proteins. The fusion of two or more mitochondria is regulated by three major proteins, MFN1 and MFN2 which reside at the OMM, and OPA1, residing in the IMM. The regulation of mitochondrial fission is regulated by several other proteins acting initially on the OMM. Most importantly, DRP1 is recruited by FIS1, which oligomerises and constricts around the mitochondria. Figure taken from (51).*




*1.4.3. Mitochondrial Genetics*

Parental transmission of mtDNA occurs in a non-Mendelian, maternal fashion. As opposed to ncDNA, mtDNA contains no histone structure and is instead organised in protein-DNA complexes named nucleoids. This lack of protective histones has been associated with the increased level of mtDNA mutations that occur throughout ageing.

Each mitochondrion contains between 2-10 copies of mtDNA. However, not all mtDNA in each mitochondrion share the same DNA sequence. Mutations that affect all copies are referred to as homoplasmic (52). However, mutations may also arise in a small proportion of the mtDNA molecules in each mitochondrion, creating a mosaic of different sequences known as mitochondrial heteroplasmy. Mitochondrial heteroplasmy has been linked to various mitochondrial diseases, (53), and poses a real issue when studying mitochondrial diseases. On a larger scale, mutations in mtDNA can be used to help group populations. Throughout evolution, collections of mutations in mtDNA may be conserved and passed on through maternal inheritance, thus allowing for the tracing of common ancestral lineages, these groups are referred to as haplogroups. Whilst many studies have investigated the risk associated with different mitochondrial haplogroups and complex diseases (54-56), investigations into individual mitochondrial single nucleotide polymorphisms (mt-SNPs) are limited, primarily as many genome-wide association studies (GWAS) often discard reads associated with the mitochondrial genome. This is in part due to the lack of genome-wide coverage that many single nucleotide polymorphism (SNP) microarrays offer (57) and that mitochondrial heteroplasmy may have large effects on current genotype calling algorithms, due to the possibility that there may be more than two minor alleles in highly heteroplasmic cells, requiring custom analysis pipelines for mtDNA variant analysis (58). Further, given the tissue-specific nature of heteroplasmy and sequence-specific mutational threshold required to drive mitochondrial dysfunction (59,60), and as many GWAS use easily accessible tissue such as blood, it is possible one may not observe peripheral changes in mtDNA genome stability contributing to disease in other tissues (2).

#### *1.4.3.1. Homology to the Nuclear Genome*

A major factor influencing mitochondrial genetic studies is the presence of nuclear-mitochondrial pseudogenes (*NUMTs*). *NUMT*s are regions of the mitochondrial genome that, over an evolutionary period of time, have translocated to the nuclear genome and therefore, share a high sequence homology with their mitochondrial paralogue. Failure to account for these regions has led to misinterpretations of genetic sequencing data (61,62), and is one pitfall of using publicly available data, although bioinformatic approaches to account for these regions are being developed (63,64). Further, the presence of *NUMTs* can cause major issues in genomic analyses using pre-sequencing enrichment methods such as custom capture or long-range polymerase chain reaction (PCR) as the likelihood of *NUMT* co-amplification, or even preferential amplification, increases due to the strong sequence similarity between the two segments of genome (65). Ideally to avoid misinterpretation of data, mitochondria should be isolated from the tissue of choice prior to sequencing (1) and nominated loci should be further assessed using programs such as the Basic Local Alignment Search Tool (BLAST), a program used to look for similarity between two regions of DNA (66).

## ***1.5. Mitochondrial Epigenetics***

The diverse range of mitochondrial functions coupled with their importance in the governance of cellular energy demands means that the expression of mitochondrial proteins requires sophisticated levels of fine-tuning. Although ~99% of mitochondrial proteins are encoded by the nuclear genome, differences in mtDNA methylation have been associated with a range of

complex diseases (1,2). Given the fact that mitochondria are not associated with histone proteins, it is unlikely that histone modifications will have a direct impact on mtDNA regulation. MtDNA methylation, on the other hand, may play an important role in regulating mitochondrial function, either independently, or via a complex mechanism involving nuclear-mitochondrial crosstalk (1).

### *1.5.1. MtDNA Methylation: The Controversial Mark*

Despite rapid progress in the field of nuclear epigenetics, the field of mitochondrial epigenetics has received little attention since initial, but contradictory studies published in the 1970's (67-69). These controversies continue, with one recent report concluding an absence of 'biologically significant levels' of mtDNA methylation in four regions of mtDNA analysed in human HEK293 cells and publicly available data (70). However, given the low sequencing depth of mtDNA in the publicly available data used in this study (94x), and the multi-copy nature of the mtDNA genome present in any given cell, it is possible that the true extent of mtDNA methylation was not determined. Despite this, DNMT1 and DNMT3B have been identified in the mitochondria of human and murine cell lines and may play an important role in methylation of the *mitochondrial Displacement loop* (*D-LOOP)* (71). The same study also identified TET1 and TET2 in the mitochondria, which alongside the presence of 5-hmC in the mitochondrial *D-LOOP*, strengthens evidence suggesting that methylation and demethylation could have important roles in defining mitochondrial function. Both 5-mC and 5-hmC have also been identified in a number of mammalian mitochondria from cell lines (72). Furthermore, 5-mC has been found to co-localise in motor neurons with the mitochondrial marker

Superoxide Dismutase 2 (SOD2) (73). In addition, the identification of Dnmt3a in mitochondria isolated from a mouse motor neuron cell line (73), as well as in mitochondria isolated from mouse brain and spinal cord (72), suggests a potential tissue-specificity for mtDNA methylation. As well as displaying signs of tissue specificity, mtDNA methylation also correlates well with other factors strongly associated with ncDNA methylation, such as age and sex in BS pyrosequencing studies (74).

### *1.5.2. MtDNA Methylation: A Role in Complex Disease*

Despite little being known about the physiological impact of variation in mtDNA methylation, some recent studies have shown that it may be associated with a variety of diseases. The majority of studies have focussed on diseases where mitochondrial dysfunction is known to be prevalent, for example in cancer (75) and also in Down's Syndrome (DS) (76). Particularly, for the purpose of this thesis, mitochondrial dysfunction and mtDNA methylation aberrations in DS cells are interesting given that these patients have an increased likelihood of presenting with Alzheimer's Disease (AD)-like phenotypes throughout ageing (77,78), due to possessing an extra copy of the amyloid precursor protein (*APP*) gene. An overview of studies of mtDNA epigenetics in disease is given in **Table 1.1**.

***Table 1.1: An overview of current studies of mitochondrial epigenetics in disease.*** *New abbreviations: Quantitative Real-Time PCR (qRT-PCR), simple steatosis (SS), non-alcoholic steatohepatitis (NASH), Liquid chromatography-electrospray ionization tandem mass spectrometry (LC-ESI-MS), Liquid chromatography mass spectrometry (LC-MS), Immunofluorescence (IF), Amyotrophic lateral sclerosis (ALS).*

| Research Question | Techniques | Main Findings | Ref |
|---|---|---|---|
| The effect of different environmental exposures (metal-rich particulate matter, air benzene levels and traffic derived elemental carbon levels) on mitochondria | Pyrosequencing<br><br>qRT-PCR | Increased exposure to particulate matter increases *MT-RNR1* and *MT-TF* gene methylation<br><br>Increased *MT-RNR1* methylation is associated with a significant increase in mtDNA copy number. | (79) |
| The effect of mtDNA methylation in the mitochondrial *D-LOOP* on gene expression in colorectal cancer cells. | Methylation-specific PCR<br><br>Western blotting | Significant hypomethylation in the *D-LOOP* of tumour cells is strongly associated with increased *MT-ND2* expression and mtDNA copy number. | (75) |
| The effect of methylation in the *D-LOOP* , *MT-ND6* and *MT-CO1* on disease progression in SS and NASH | Methylation-specific PCR<br><br>qRT-PCR | Increased *MT-ND6* methylation and DNMT1 and decreased MT-ND6 protein in NASH compared to SS. Physical activity reduced *MT-ND6* methylation in NASH. | (80) |
| The effect of decreased SAM on mtDNA methylation in DS lymphoblastoid cells | LC-ESI-MS<br><br>LC-MS/MS | Decreased SAM availability in DS lymphoblastoid cells reduces methyl uptake to mitochondria and leads to mtDNA hypomethylation. | (81) |
| The tissue specificity of DNMTs and 5-mC in the mitochondria in relation to ALS models. | IF<br><br>Pyrosequencing | Increased methylation at six cytosine sites in the *16S rRNA* gene in the spinal cord of an ALS mouse cell line.<br><br>Reduced levels of mtDNMT3a protein in skeletal muscle and spinal cord early in disease. | (82) |
| The effect of mtDNA methylation on mtDNA copy number in gastric cancer. | qRT-PCR<br><br>Pyrosequencing | Reduced mtDNA copy number levels in late clinicopathological stages.<br><br>Demethylation of mtDNA demethylation increases mtDNA copy number. | (83) |

*1.5.3. MtDNA Methylation: Regulation of Gene Expression*

The role of mtDNA genetic variation on underlying mitochondrial function has become increasingly studied in recent years. MtDNA haplogroups and genetic variation have been found to have pronounced effects on mitochondrial functions, leading to altered ETC functions and ROS levels, which are associated with increased breast cancer risk (84). However, given that a single mitochondrion can contain up to 10 copies of the mitochondrial genome (85), and as there are multiple mitochondria in any given cell, the extent of heteroplasmic mutations and their effect on mitochondrial function is far from clear (53). Further, relatively few of the studies investigating mtDNA methylation have related identified changes to alterations in gene expression.

One recent study showed significant hypermethylation of the *D-LOOP* is negatively associated with expression of three mitochondrial-encoded ETC genes, *Mitochondrially Encoded Cytochrome B* (*MT-CYB), Mitochondrially Encoded NADH: Ubiquinone Oxidoreductase Core Subunit 6* (*MT-ND6*) and *Mitochondrially Encoded Cytochrome C Oxidase II* (*MT-*CO2) in retinal microvessels derived from diabetic retinopathy donors (86). However, the group failed to find significant differences in the methylation of *MT-CYB* and concluded that, given the importance of the *D-LOOP* in replication and transcription of the mitochondrial genome, changes in *D-LOOP* methylation may result in transcriptional changes across the mitochondrial genome, potentially contributing to the pathogenesis of diabetic retinopathy (86).

However, a recent *in vivo* study showed *D-LOOP* methylation did not correlate with the expression of mitochondrial-encoded genes during inflammation (87). Briefly, it was found that lipopolysaccharide (LPS)-mediated inflammation in mice led to significant changes in mtDNA transcript levels, which were reversed upon treatment with α-Lipoic acid (LA), an agent with anti-inflammatory properties. However, despite these transcriptional changes, neither treatment altered *D-LOOP* methylation. Taken together, these two studies suggest that mitochondrial transcription may be driven by mtDNA methylation of the *D-LOOP*; however, this phenomenon is not the sole driver of mitochondrial gene expression. Further studies should potentially employ a more holistic approach to investigate the relationship between mtDNA methylation, mtDNA gene expression and potential nuclear-encoded effectors of mtDNA transcription. To better understand the relationship between mtDNA methylation and gene expression, the use of whole DNA methylome sequencing at single base resolution should be considered.

### *1.5.4. The Bi-directionality of Mitoepigenetics*

As research into the field of mitochondrial epigenetics gains momentum, studies have focused on a potential *trans*-acting role of mtDNA in the epigenetic regulation of ncDNA, whereby covalent modifications across the mtDNA genome may affect not only the expression of a gene in *cis*, but also have *trans*-acting effects on the transcription of genes in the nuclear genome. A recent study found DNA methylation levels in the nuclear-encoded, mitochondrial-specific transcription factor *DNA polymerase gamma A* (*POLGA*) inversely correlate with mtDNA copy number in pluripotent and multipotent cell types (88).

The removal of mtDNA (89), or changes in mtDNA haplogroups (90), have been found to be associated with differences in ncDNA methylation levels, suggesting an important interplay between the two genomes and a bi-directionality to mitoepigenetics (91).

Cytoplasmic hybrids (cybrids) are an important cell type for studying mitochondrial function as they contain identical nuclei to the parent cell but different mtDNA, allowing for a controlled investigation into the role of mtDNA variants. Using Restriction Landmark Genomic Scanning (RLGS) and $Rho^0$ cells, a form of cybrid cell line designed for investigating mtDNA depletion, one study found that mtDNA depletion significantly altered DNA methylation at CGIs in nuclear encoded genes (89), indicating that there are functional interactions between the two genomes. Re-introduction of wild-type mtDNA restored DNA methylation levels at some RLGS spots, suggesting that, at least for some genes, mitochondria may play a role in ncDNA methylation. Other cybrid studies have shown that mtDNA derived from haplogroup J had consistently higher levels of global ncDNA methylation than haplogroup H, suggesting that mtDNA variants may play an important role in influencing ncDNA methylation (92). This difference in DNA methylation was accompanied by significant decreases in the expression of six of the 11 nuclear-encoded genes assayed that are involved in DNA methylation and acetylation processes, as well as significant increases in the transcription of *Methionine Adenosyltrasferase 2B* (*MAT2B*) and *MBD4*, both of which are important for DNA methylation. Given the importance of these two processes in epigenomic regulation, this further illustrates how mtDNA genetic variation can impact on the nuclear genome. Another recent study

investigating the mitochondrial 3243A>G heteroplasmy, which causes several clinical phenotypes including mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes (MELAS), showed that changes in heteroplasmy levels led to large and widespread changes in mitochondrial and nuclear transcriptional regulation, including changes in transcripts involved in DNA methylation and histone acetylation. Interestingly, high levels of this heteroplasmic variant (50-90%) resulted in ncDNA gene expression changes that are similar to those seen in AD, PD and Huntington's disease (HD).

Each mitochondrion is comprised of approximately 1500 proteins; however, of these, only 13 are encoded by the mitochondrial genome. The remainder are encoded by the nuclear genome and imported into the mitochondria. A recent study found that over 600 of these genes have tissue-specific differentially methylated regions (DMRs), ultimately leading to changes in mitochondrial function, dependent upon tissue type (93). This suggests that there is an additional level of complexity to consider in the study of mitochondrial epigenetics, whereby epigenetic changes in one genome may affect transcriptional control in another in a tissue-specific manner. With alterations in ncDNA methylation in neurodegenerative disorders now well established (4,94-97), future studies may aim to study mtDNA and ncDNA epigenomes in parallel to gain an understanding of the complete cellular epigenetic landscape.

*1.5.5. Current Methodologies for MtDNA Methylation Analysis*

An important consideration in mtDNA methylation is defining the most appropriate method for which to study the epigenetic mark. A host of different platforms have been developed that are capable of assessing DNA methylation in the nuclear genome. However, given the relative nascence of the field of mtDNA epigenetics, the appropriateness of the wide range of genomic technologies available and their utility for this field has yet to be reviewed. Advances in genomic technologies has seen ncDNA methylation studies move from profiling methylation at a handful of candidate CpG sites to epigenome-wide association studies (EWAS). However, the current workhorse for such studies, the Illumina 450K methylation array, provides no coverage of the mitochondrial genome. Given that mtDNA methylation is now becoming increasingly investigated, an understanding of the technologies available to assess genome-wide changes in mitochondrial DNA methylation is of upmost importance. An overview of genomic technologies that could be utilised to measure mtDNA methylation is provided in **Table 1.2**.

To capture an accurate and precise representation of potential mtDNA methylation changes, a technique should ideally be able to detect DNA modifications with low error rate and at single nucleotide resolution. This is more pertinent to studies of mtDNA epigenetics than in studies of ncDNA methylation given the low levels of mtDNA methylation identified in recent studies (79,98). Currently, the only published genome-wide studies investigating mtDNA 5-mC and 5-hmC levels have utilised publicly-available methylated DNA immunoprecipitation sequencing (MeDIP-Seq) data (63,99). Whilst the study of

5-mC was able to identify spatio-temporal patterns of mtDNA methylation in a variety of cells and tissues, the coverage was relatively low (5x). Given sequencing and genomic biases, low levels of coverage may result in some bases not being sequenced adequately for analysis. Further, the resolution of MeDIP-Seq also remains a limitation. Despite high levels of concordance between MeDIP-Seq data, which averages the methylation values across a region, and other technologies that focus on single CpG sites (100), it is possible that a correlation may not be observed across every region of the mitochondrial genome. For example, averaging methylation in regions as small as 100bp in the mitochondrial genome could include averaging methylation across many gene boundaries of tRNAs in some regions and could greatly mask the true methylation status at specific regions within these genes.

***Table 1.2: An overview of sequencing technologies that could be utilised to study mtDNA methylation in neurodegenerative diseases.*** *New abbreviations: Hydroxymethylated DNA Immunoprecipitation sequencing (hMeDIP-Seq), Immunoprecipitation (IP), Reduced Representation Bisulphite Sequencing (RRBS), Enhanced Reduced Representation Bisulfite Sequencing (ERRBS), Oxidative Bisulphite (OXBS), Single-molecule real-time sequencing (SMRT), Zero-mode Waveguide (ZMW), Next-Generation Sequencing (NGS).*

| Method | General Overview | Advantages | Disadvantages |
|---|---|---|---|
| **MeDIP- & hMeDIP-Seq (NGS)** | MeDIP-Seq is an IP based method. In brief, purified genomic DNA is sheared by sonication to produce random fragments. These fragments are then denatured and immunoprecipitated, followed by PCR amplification. Using high-throughput sequencing at a depth of two Gb, around 70-80% of CpGs in the human genome can be identified (101). | 1. Staining with anti-5-hmC antibodies allows for 5-hmC Stroud et al. (102), which is not only present in the brain and significantly reduced in AD (103), but is also present in brain mitochondria (104).<br><br>2. Sequencing produces files typically less than 6GB (fastq format)<br><br>3. Cost-effective relative to most BS based approaches (105).<br><br>4. Only regions of interest are investigated. | 1. Investigation of both 5-mC and 5-hmC is costly and requires several µg mtDNA (102,106).<br><br>2. Resolution reduced to methylated windows. Analysis of single cytosine methylation sites is not possible (107).<br><br>3. Count based data means that reads must be normalised for CpG density and total read counts (108), although pipelines are now more capable of accounting for this (109). |

| Method | General Overview | Advantages | Disadvantages |
|---|---|---|---|
| **Custom Capture (NGS)** | Custom capture kits available allow for a custom design of a library for the enrichment of specific DNA target regions.<br><br>In brief, DNA undergoes standard NGS pipelines approaches of sonication, end repair, A-tailing and in this case, methyl-adaptor ligation before being bound by custom baits, bisulphite treated and amplified. Samples can then be run on a sequencer such as the Illumina HiSeq. | 1. Single base resolution for DNA methylation analysis<br><br>2. Customised library to capture 100% of mtDNA further enriches mtDNA, reducing risk of *NUMT* amplification.<br><br>3. Can use as little as 1µg mtDNA (Nimblegen).<br><br>4. Small target genome would likely generate small output files from sequencing.<br><br>5. Simultaneous analysis of CpG and non-CpG methylation.<br><br>6. Small size of the genome allows for ample space on a Mi-Seq lane for high (>1000x) coverage of over 100 samples at a time. | 1. Currently the process is not supported by Agilent.<br><br>2. Typically cost exceeds that of non-bisulphite experiments.<br><br>3. Reduced complexity of sequence due to bisulphite conversion (110). |
| **RRBS & ERRBS (NGS)** | RRBS makes use of MspI restriction enzymes for selective digestion of genomic DNA. This produces fragments of genomic regions enriched for CpG sites, which can be BS treated and sequenced irrespective of their methylation status (111). | 1. Requires less DNA input than WGS-BS, and most other techniques, with RRBS requiring ~10ng (112) and Enhanced RRBS requiring ~50ng (113)<br><br>2. Can provide coverage of CpG promoters and CpG islands at a fraction of the cost of WGS-BS at single base resolution.<br><br>3. Has been modified to allow for single cell epigenomic analysis (114,115) . | 1. Targets CpG islands, however the mitochondrial genome contains no CpG islands, therefore could give a very poor coverage of the mitochondrial methylome.<br><br>2. Reduced complexity of sequence due to bisulphite conversion (110). |

| Method | General Overview | Advantages | Disadvantages |
|---|---|---|---|
| **WGS-BS & OxBS-Seq (NGS)** | WGS-BS converts cytosines to uracil bases by BS treatment, After PCR amplification, 5-mC will be read as thymine in a sequencer, and can thus be distinguished from unmethylated sites. The reads can then be aligned to recreate the DNA sequence (116).<br><br>OXBS adaptation converts 5hmC base to 5fC before BS treatment. This enables a direct measurement of 5mC, and an indirect measurement of 5hmC at base-pair resolution (117). | 1. Single base resolution for DNA methylation analysis<br><br>2. Simultaneous analysis of CpG and non-CpG methylation<br><br>3. The development of OxBS-Seq allows for the identification of both 5-hmC and 5-mC when used in combination with standard BS-Seq (118,119).<br><br>4. OxBS-Seq can be further combined with RRBS to characterise 5-mC, 5-hmC and 5-fC (120). | 1. Sequences whole genome, including many repetitive, non-methylated AT-rich regions.<br><br>2. Vast quantities of data generated leads to massive (100GB+ files) following sequencing, leading to time consuming processing (121).<br><br>3. High cost of sequencing to improve sensitivity makes it unsuitable for large studies (110).<br><br>4. Reduced complexity of sequence due to bisulphite conversion (110) |

| Method | General Overview | Advantages | Disadvantages |
|---|---|---|---|
| **PacBio (Generation 2.5)** | SMRT by Pacific Biosciences uses a single DNA polymerase molecule attached to the bottom of a ZMW hole (tens of nanometers in diameter). By illuminating only the bottom 30nm of the ZMW with a laser, single nucleotide addition to the DNA can be measured. Each nucleotide fluoresces when bound to the DNA polymerase, which is detected by a camera before being cleaved off. Bases can be identified by corresponding fluorescent colours (122). | 1. SMRT avoids potential PCR bias by using pre-amplified DNA as an input for sequencing (123).<br><br>2. Optimised for circular genomes<br><br>3. Increased library complexity to BS based methods as DNA modifications are determined based upon changes to polymerase kinetics specific to modification present (124). Potentially, this may be developed to accurately characterise all methylation marks in one simultaneous run.<br><br>4. Longer read length could lead to improved coverage and accuracy for processes such as variant calling (125).<br><br>5. Ideal for de novo sequencing as capable of reads at a length of 8kbp (126)<br><br>6. Can detect strand specific patterns of 5hmC, after selective chemical labelling, in a high throughput manner (127) | 1. The RSII platform is capable of producing continuous long reads and circular consensus reads, at present, these reads are typically associated with high error rates (128). However, given the random nature of the errors, with sufficient depth SMRT can give an accurate read of the genome if sequenced at high coverage (129).<br><br>2. The cost and size of PacBio RSII, coupled with the need for specialist technical and bioinformatic support may limit much research to outsourcing data to specialist sequencing services. |

| Method | General Overview | Advantages | Disadvantages |
|---|---|---|---|
| **Nanopore Third Generation** | The nanopore technique makes use of either biological nanopores or solid state nanopores, which are embedded in a membrane immersed in salt solution. To create a flow of ions through the pore, an electrical current can be applied. A single stranded DNA molecule passing through the pore will create measurable changes in the intensity of the current, and different DNA bases could be distinguished by the degree and duration of modulation of the current (130). | 1. Although still in alpha testing, nanopore technologies such as Oxford's MinION™ allow for sequencing analysis to be done on a portable, benchtop device in real-time (131).<br><br>2. Identification of 5-hmC and 5mC<br><br>3. Systems like the Oxford MinION™ allows for the activation of hundreds of nanopores in parallel to reduce time of sequencing large genomes versus single pore technologies.<br><br>4. Sample preparation steps are reduced as no amplification or cloning is necessary, and use of enzymes is limited. This in turn reduces costs of sequencing (132). | 1. Observed per base error rates for applications such as DNA sequencing have reduced in recent years, but are still very high, recently reported at 30% (131).<br><br>2. Some nanopore technologies utilise the binding of MBD1 to 5mC, altering current flow through the Nanopore, to detect the methylation mark (133); MBD1 has not yet been identified in mitochondria.<br><br>3. Although nanopore sequencing could use < 1 µg of genomic DNA, an estimated 700 µg is needed due to the concentration-limited rate at which DNA molecules can be captured (132) |

Although this study was able to utilise publicly available data that had been generated to inexpensively assess ncDNA methylation, studies to specifically assess mtDNA methylation would ideally investigate DNA methylation at single base resolution. BS sequencing (BS-Seq) is another source of publicly available data that could be exploited to investigate mtDNA methylation; however, although this allows the interrogation of DNA methylation at single nucleotide resolution, data with an appropriate depth of sequencing is lacking due to the expense of sequencing the entire ~3,000 Mb human genome. An alternative, and more appropriate next generation sequencing (NGS), approach would specifically capture the entire mitochondrial genome prior to sequencing and due to its relatively small size, would allow for inexpensive sequencing at single base resolution analysis. Currently, two custom capture methylation technologies are available from Agilent and Nimblegen, and coverage can be tailored to suit the researcher's needs and allow for the interrogation of epigenetic variability within the mitochondrial genome. Despite costing considerably less than Whole genome BS sequencing (WGBS-Seq), a targeted capture of the mitochondrial genome is still relatively expensive and requires a higher input of DNA. This therefore, is a particular limitation in studies of brain tissue, as it would require the isolation of large quantities of mtDNA from small amounts of frozen post-mortem tissue (134).

Finally, the dependence of NGS technologies on PCR amplification prior to sequencing leads to a potential PCR bias and over-representation of some sequences. Given the tissue-specific nature of mitochondrial heteroplasmy, it is possible that these amplifications could influence findings. As such, the

development of new generations of sequencing systems such as the Pac-Bio RS II and Oxford Nanopore Technologies' MinION$^{TM}$ is an appealing prospect. Furthermore, despite both technologies being associated with a high individual base call error rate at present, the Pac-Bio's long read based deep sequencing has been used to accurately map complex regions of the human genome (135) as well as mitochondrial genomes in other species (136-139). Interestingly, given that both technologies use native DNA, both technologies provide the possibility of simultaneous genetic and epigenetic analyses, including a range of DNA modifications (124,140-143). The circular consensus reads of PacBio RS II could provide an interesting platform for the simultaneous identification of the four major epigenetic marks in real time at a high coverage. However, isolation of high quality mtDNA from the range of frozen tissues commonly used in nuclear epigenetics could prove to be difficult and would require further investigation. The continual development of more accurate calling algorithms and likely future cost reductions may therefore lead to these third-generation sequencing technologies being the optimal choice for providing specific, deep and targeted analyses of the mitochondrial (epi)genome

### 1.5.6. Caveats Associated with MtDNA Methylation Analysis

Despite the potential importance of mitochondrial DNA modifications in disease, there are a number of technical challenges specific to interrogating the mitochondrial epigenome that have hampered widespread studies to date. These issues can be broadly summarised as encompassing genetic issues and specificity issues, which are summarised briefly with potential solutions in **Table 1.3**.

***Table 1.3: A summary of the major issues and potential solutions in the field of mitochondrial epigenetics**. New abbreviations: Quality Control (QC).*

| Caveat | Potential Issues | Potential Solutions |
|---|---|---|
| NUMT | Wrongful determination of pseudogenes as mtDNA will affect the validity of results. | Isolate mitochondria before mtDNA extraction to avoid nuclear contamination of the mitochondria. |
| | | The use of QC and more specific primers designed with the consideration of NUMT amplification (144). |
| | | BLAST searches to identify known NUMTs and can be used to compare sequence similarity of results and NUMT to avoid misrepresentation. |
| Isolation of Mitochondria | The most common methods of mtDNA isolation/extraction (precipitation/ ultracentrifugation) are time consuming and labour intensive and may still leave ncDNA. | A range of kits are available for mitochondrial isolation from tissue, which have been used to isolate and store mtDNA for further analysis in recent mitochondrial epigenetic studies (72); (145). |
| | No standard isolation protocol for mitochondria from blood. | Other methods may be adapted and employed to preferentially bind to circular genomes such as the mitochondria instead of linear nuclear DNA (146). |
| | | Those wishing to analyse mtDNA methylation in blood samples, would have to develop an accurate and efficient protocol for isolating mtDNA prior to carrying out their study. |
| Heteroplasmy | Genetic mutations in mtDNA may have specific mtDNA signature. | Studies of differential methylation between different mitochondria and the functional effects that they present in AD. Percentage of heteroplasmy in the sample should be considered alongside results. |
| Differentiating between epigenetic marks | Not all methods (i.e. BS) can distinguish 5-mC from 5-hmC | Using oxidative bisulphite-sequencing (OXBS), 5-hmC is oxidised to 5-fC unlike 5-mC, allowing for a distinction to be made between the two marks at single-base resolution due to an initial lack of deamination by 5hmC (119). |

| Caveat | Potential Issues | Potential Solutions |
|---|---|---|
| Cell Specificity | Different brain regions have differential methylation patterns and different cell populations | Large samples sizes on specific brain subregions will improve the likelihood of determining methylation changes to a statistically significant degree if present. |
| | | The use of techniques such as FACS can potentially separate cell types such as glia and neurons to analyse samples in a cell-type specific manner. |
| Sensitivity | Reduced methylation levels in mitochondria Diluted signals | Comparative analysis of techniques should be used to determine the relative precision and accuracy of each and allow for the determination of their suitability to mitochondrial methylation studies. |

#### *1.5.6.1. Genetic Issues*

##### *1.5.6.1.1. Nuclear pseudogenes*

By far the greatest concern when analysing mtDNA methylation arises from regions of homology between the mitochondrial genome and *NUMTs*. These genes are nuclear paralogs of mtDNA which have been translocated and inserted into the nuclear genome during the evolution of both genomes (147). These insertions were thought to typically occur in non-coding regions; however, more evolutionary recent translocations have actually been shown to prefer integration into coding regions, thus leading to potential alterations in gene function with implications for disease (148). *NUMT*s are generally small and typically comprise ~0.1% of the nuclear genome (149). However in humans, it has been shown that some *NUMTs* can be as large as 14.7kb, representing a significant portion of the ~16.6kb human mitochondrial genome (150). As such, the presence of *NUMTs* can cause major issues in genomic

analyses using pre-sequencing enrichment methods such as custom capture or long-range PCR as the likelihood of *NUMT* co-amplification, or even preferential amplification, increases due to the strong sequence similarity between these segments of the two genomes (65). As such, this sequence similarity can lead to the miss-classification of *NUMT*s as mtDNA during analysis, and has led to a number of publications wrongly describing *NUMTs* as mtDNA (61,62). *NUMT* miss-classification has also been observed in AD genetic studies whereby amplification of the *NUMT* sequence has led to false heteroplasmies being reported (151,152). One potential solution is to isolate mitochondria prior to DNA extraction in an attempt to reduce the risk of contaminating the nuclear genome. However, despite extensive research being dedicated to mtDNA analysis, existing methods for mitochondrial isolation and mtDNA extraction via the use of fractional precipitation or gradient ultracentrifugation remain time consuming and labour intensive (153) and often leave residual nuclear DNA contamination following mitochondrial isolation (146).

#### *1.5.6.1.2. Variation in MtDNA: Haplogroups and Genetic and Epigenetic Heteroplasmy*

The multi-copy nature of the mitochondrial genome and the subsequent heteroplasmy that arises poses a real issue when studying mitochondrial diseases, as inter- and intra-individual heteroplasmic variation can confuse the association of a haplogroup with its corresponding phenotype (154). The importance of this issue is highlighted by a recent study demonstrating that mitochondrial heteroplasmy alters DNA methylation across the nuclear-encoded mitochondrial genes, *Transcription factor alpha mitochondrial* (*TFAM*) and *RNA*

*Polymerase Mitochondrial* (*POLMRT*) (155). Finally, if mtDNA methylation is altered across different mtDNA in the same mitochondrion, it could create an epigenetic mosaic within the mitochondrion, the cell and across the tissue, whereby each copy of mtDNA may possess its own methylation profile. If this 'methylomic heteroplasmy' were to occur, it could be very difficult to tease apart the effects of such a mosaic in functional studies.

On a larger scale, mutations in mtDNA can be used to identify haplogroups. Throughout evolution, mutations in mtDNA may be conserved and passed on through maternal inheritance, thus allowing for the tracing of common ancestral lineage by comparing haplogroups. Numerous studies have identified both contributory and protective effects of different haplogroups in diseases such as AD. For example, haplogroup K reduces the risk of developing sporadic AD onset in *Apolipoprotein ε4* (*APOE* ε4) carriers in an Italian population (156), but not in the Polish population (157). This presents an additional potential caveat in mitochondrial epigenetics, as mitochondrial haplogroups have been found to affect global levels of ncDNA methylation (90). As such, extra care should be taken to account for haplogroup variability in AD mitochondrial epigenetic studies, when possible.

#### *1.5.6.2. Specificity and Technical Issues*

The brain is a complicated, heterogeneous organ with numerous functionally-distinct subregions, each with their own individual composition of cell types. Unsurprisingly, there are clear tissue-specific epigenetic differences in ncDNA

across brain regions (158,159). There is an added level of complexity with respect to the mitochondrial epigenome because each mitochondrion contain between 2-10 copies of mtDNA and each cell contains varying levels of mitochondria; therefore the amount of mtDNA copies in each cell can vary between 100-10,000 dependent upon cell type. Further, a recent study using laser capture microdissection demonstrated that alterations in mitochondrial 5-hmC are seen with age in mouse cerebellar purkinje cells, which was not evident in whole cerebellar tissue (104), demonstrating the importance of cell-specific analyses in heterogeneous tissue, particularly when investigating functional impact. Currently, the most common method of measuring DNA methylation is via the conversion of DNA with sodium BS followed by subsequent sequencing analysis. However, these approaches are unable to distinguish between 5-mC and 5-hmC (160), an important limitation given recent studies confirmed the presence of 5-hmC in mitochondria in brain tissue (104). Studies have found that although both DNA modifications are present in the mitochondria, they occur at much lower levels compared to in ncDNA (72,161), and thus methods used for quantification may need to be more sensitive. Furthermore, variation in mitochondrial copy number may lead to the dilution of signals and reduce detection if the tissue is largely heterogeneous. Finally, it is important to note that 5-mC and 5-hmC in mitochondria have been found to predominantly occur at cytosines not preceding a guanine base (non-CpG) (71,162). Therefore the use of techniques that primarily investigate CpG methylation (69) must be called in to question to the extent by which they can depict the true epigenetic landscape of this genome.

Given the growing interest in studying epigenetic changes within mtDNA in a range of pathologies, there are several important considerations pertinent to this field of research. First, tissue-specific effects of DNA methylation mean that profiling a disease-relevant tissue is likely to be critical. Further, different cell types will have varying levels of mtDNA copy number and different DNA methylation levels dependent upon cellular requirements and localised environments (1). This is a potentially important issue in mtDNA methylation analysis given that a number of studies have implicated an association between mtDNA methylation status and mtDNA copy number (163) and that tissue homogenates contain a number of different cell types, many with varying levels of mtDNA copy numbers. One potential way to solve this problem would be to sort cell types from tissue homogenates using Fluorescence-activated cell sorting (FACS), however isolating sufficient levels of mitochondria for mtDNA methylation analyses using current technologies seems unlikely.

### *1.5.7. Other MtDNA modifications*

In a recent study, 5-fC and 5-caC, but not 5-hmC were found to lead to modest blocks in DNA transcription when mediated by either T7 RNA polymerase or by human RNA polymerase II in an *in vitro* transcription assay, which was also replicated in two human cell lines (164). Interestingly, T7 RNA polymerase shares a high degree of homology with human mitochondrial polymerases, and given the identification of 5-hmC and TET proteins in mitochondria (72,104), this study hypothesised a mechanism by which mtDNA demethylation products may be important for the regulation of mtDNA transcription. However, at present, no study has investigated 5-fC and 5-caC levels in mtDNA and further investigation

into the modulation of mtDNA methylation needs to be undertaken. Given the huge DNA input requirements for OXBS sequencing, which allows the quantification of 5-mC and 5-hmC independently, subsequent analysis of these other marks would likely require vast quantities of mtDNA, currently unavailable using non-PCR based mtDNA enrichment methods. However, improvements in the detection of these bases in newer technologies, such as PacBio RS II, could provide novel insights into their role in mtDNA epigenetic regulation of mitochondrial function.

## ***1.6. The Role of Mitochondria in the Brain***

The brain accounts for approximately 20% of the human body's energy expenditure during resting metabolism (165). As such, this complex organ requires continuous and vast quantities of ATP, ensuring that the correct maintenance of mitochondrial function is a constant priority during healthy ageing (166). Mitochondria are also implicated in many other important neurophysiological functions, for example synaptic mitochondria are believed to be involved in the regulation of neurotransmission by buffering extra intracellular $Ca^{2+}$ (167), making the mitochondria a vital organelle in the establishment of $Ca^{2+}$ homeostasis. Further, mitochondrial morphology and dynamics vary between synaptically immature and mature cortical neurons, with shorter mitochondrial lengths being observed in immature neurons, allowing for increased movement and greater ability to meet the high energy demands of immature neurons, providing a key role for this organelle in neuronal development (168). However, despite being the major site of ROS production, the mitochondria lack protective histones and defective mtDNA repair has been

suggested to play a role in a number of neurodegenerative diseases (169). As a by-product of ATP generation, ROS accumulation in areas of high ATP demand, such as post-mitotic neuronal cells, has been associated with neuronal loss (170). Mutations in mitochondrial-encoded genes have been shown to cause a number of maternally heritable, monogenic diseases. Most notably, MELAS syndrome, a disease affecting multiple organs which can lead to the development of a number of syndromes, ranging from muscular weakness, fatigue and stroke-like episodes to dementia, epilepsy and diabetes in later stages (171). As such, mtDNA variants have been shown to be an important driver of mitochondrial dysfunction and can have significant neuropathological effects (172).

## 1.7. Alzheimer's Disease

AD, the most common form of dementia accounting for over 60% of all cases, is a chronic, currently incurable, neurodegenerative disorder. Current estimates predict that over two million people in the United Kingdom will be affected by dementia by 2050 (173). The classic neuropathological hallmarks associated with AD include the formation of amyloid beta (Aβ) plaques and neurofibrillary tangles of hyperphosphorylated microtubule-associated protein Tau. Importantly, the continued spread and development of neurofibrillary tangles through the brain provides a standardised measure of AD pathology (174). Each significant stage of pathology development can be grouped into categories known as Braak stages, with Braak stages I and II primarily associated with tau tangle accumulation within the transentorhinal cortex. Later stages see the spread of tangles to the limbic regions (Braak Stages III and IV)

and gradually to more regions of the neocortex (Braak Stages V and VI) (174). This well documented development provides a strong post-mortem marker for accurate AD diagnosis (175). However, despite the pathological manifestation of AD being well documented, currently no anti-Aβ or anti-tau drug has passed Phase III trials (1). As such, a number of alternate hypotheses for the development of AD have been suggested, including the mitochondrial cascade hypothesis (176).

### *1.7.1. Mitochondrial Dysfunction in AD*

The mitochondria play a vital role in a variety of key biological functions, including apoptosis via caspase dependent and independent mechanisms (177), the regulation of calcium homeostasis (178,179) and the production of ROS (180). As such, the role of mitochondrial dysfunction has been implicated in AD pathogenesis (181,182) and forms the basis of the mitochondrial cascade hypothesis (183). Proposed by Swerdlow et al, this hypothesis states that an individual's genetic code will determine their basal mitochondrial function and that, throughout ageing; this function will decline due to a combination of genetic and environmental factors, determining an individual's time of disease onset (183).

Further evidence for a role of mitochondrial dysfunction in AD pathogenesis comes from a study demonstrating increased levels of mitochondrial gene expression and oxidative damage in a transgenic *APP* mutant mouse model of AD as a compensatory response to either Aβ or the APP mutant protein itself

(184). Further, various components of the mPTP have been shown to interact with Aβ in a number of murine models of AD as reviewed by Du and Yan (185). For example, one recent study found that, in *APP* transgenic mice, Aβ may act to upregulate Voltage Dependent Anion Channel 1 (VDAC1), a component of the mPTP, leading to mPTP blockade. (186). Interestingly, the authors also found that VDAC1 may interact with hyperphosphorylated tau, suggesting that VDAC1 interacts with both tau and Aβ, leading to mitochondrial dysfunction. An earlier study found that Aβ present in mitochondria interacts with Cyclophillin D (CYPD), another component of the mPTP, in cortical samples from both post-mortem AD patients and *APP* transgenic mice (187). In the mouse model, this was shown to lead to increased ROS production and neuronal cell death. Taken together, this illustrates a variety of mechanisms by which Aβ interacts with the mitochondria in AD, and demonstrates how mitochondrial dysfunction can lead to changes associated with AD, thus highlighting the need for continued research into the field.

### *1.7.2. The Role of Epigenetics in AD*

Epigenetic mechanisms orchestrate a diverse range of important neurobiological and cognitive processes in the brain. As such, epigenetic modifications to the nuclear genome have been widely hypothesised to play a role in many neurological disorders, including AD (4,188,189). Given that mitochondrial dysfunction is a prominent feature of AD, we recently hypothesised that epigenetic modifications to the mitochondrial genome could be important in disease progression and pathology (1).

Given the high heritability estimates for AD (190), considerable effort has focussed on understanding the role of genetic variation in disease aetiology, although more recently it has been hypothesised that epigenetic dysfunction may also be important (191). A number of studies have shown reduced global levels of the DNA modifications 5-mC and 5-hmC in AD brain (103,192-194) with only a handful of studies having looked at changes occurring at specific loci (reviewed in Lunnon and Mill, 2013 (191)). More recently, studies have investigated the effect of changes in 5-mC and 5-hmC levels within *Triggering Receptor Expressed On Myeloid Cells 2* (*TREM2*), a candidate AD risk loci from exome sequencing studies in AD (195,196). Recent methodological advances in microarray and genomic sequencing technologies have enabled researchers to undertake EWAS in AD brain, identifying several consistent DMRs associated with disease (197-199). Many of these DMRs are tissue-specific, restricted to regions of the brain associated with AD pathology, and correlate strongly with quantitative measures of neuropathology. As such, a strong case is being built for a role of epigenetics in the aetiology of AD. To date none of the published EWAS in AD brain have investigated differential DNA methylation in mtDNA, as the current workhorse for EWAS, the Illumina Infinium 450K Beadarray, has no coverage of ChrM.

### *1.7.3 MtDNA Methylation: A Role In AD?*

A recent nine year follow up study has shown that individuals with DNA methylation levels greater than a set threshold of 10% at ChrM: 932bp (*MT-RNR1*), were associated with a two-fold increased risk of mortality in a healthy control population study (74). This study suggests that increased mtDNA

methylation at specific sites may lead to genomic instability and dysregulation. As such, it is possible that low levels of mtDNA methylation are required for normal control of mtDNA expression and mitochondrial function. But, given the ubiquitous distribution of mitochondria across tissues, when methylation is increased passed a threshold, the consequences for the cell could be severe.

However to date, only two studies have investigated the role of mtDNA methylation in AD (161,200). This is potentially due to both the technological challenges associated with studying mtDNA methylation, and because many studies report low levels of mtDNA methylation (2); as such, both studies have been limited to candidate gene or global DNA methylation changes. The first study published identified a trend towards a significant increase in global levels of 5-hmC in the superior and middle temporal gyrus of late stage AD patients with respect to age-matched controls (161). The second identified increased levels of 5-mC in the *D-LOOP* between Braak Stages I to II and III to IV of AD entorhinal cortex (ECX) (200), but not in Braak Stage V and VI. Despite this, significant CG and non-CpG hypomethylation differences in *MT-ND1* during early AD stages were also observed. These changes were accompanied by an increase in *MT-ND1* expression levels in ECX in late-stage AD (stages V-VI). However, it is important to note that these changes in expression were modest and most importantly, were identified in a different cohort of cases. As such, to gain a better understanding of the role of mtDNA methylation in AD and its effect on mitochondrial gene expression, steps should be taken to investigate changes in larger, well matched cohorts using a genome-wide approach.

## 1.8. Conclusions

Despite a recent increase in publications in the field of mtDNA methylation, no study has investigated genome-wide mtDNA methylation differences at a high resolution to determine how these differences may contribute to AD pathology, nor any pathology characterised by mitochondrial dysfunction. This lack of an in-depth interrogation into the field of mtDNA methylation represents the current state of the field of mtDNA methylation, with most studies using candidate-based methods such as pyrosequencing or looking at global mtDNA methylation levels to associate the mark with disease. As such, for the field to progress, it is important to design an assay that will allow for accurate and precise measurement of mtDNA methylation across the genome and to apply this assay to future analysis. Further, the 'gold standard' of methylation analysis is BS-Seq; however the typical input requirements for this protocol, 1-3 µg, is in excess of what could typically be expected from DNA extraction from a small genome, unless tissue input quantities were exceedingly high. As such, a protocol for BS-Seq of low yields of mtDNA must also be developed. Given the role of mitochondrial dysfunction in complex diseases such as AD, the development of such a technology would allow for the interrogation of a range of relevant diseases in which mitochondrial dysfunction is also a component of the aetiology. Further, the vast majority of mitochondrial proteins are nuclear-encoded and imported into the mitochondria to aid in regulating a variety of mitochondrial functions. Despite this and the high levels of crosstalk observed between the nuclear and mitochondrial proteins, the two genomes are often investigated in isolation. As such, in the future, a more holistic approach to the investigation of mtDNA methylation should be considered.

## 1.9. Aims

1. To develop a pipeline to make use of current, publicly available genome-wide sequencing data, which will for allow the analysis of mtDNA methylation from whole genome datasets.
2. To optimise a protocol for reliable isolation of mtDNA from frozen, post-mortem human brain tissue suitable for downstream epigenetic analysis.
3. To design a targeted BS-Seq assay to investigate mtDNA methylation across the entire mitochondrial genome at single base pair resolution.
4. To use this assay to investigate whether mtDNA methylation in frozen, post-mortem brain shows sex-, age-, disease- and tissue-specific differences.
5. To use this assay to investigate the role of mtDNA methylation and its relationship to nuclear encoded mitochondrial gene methylation.

# Chapter 2. Materials and Methods

## 2.1 Introduction

This chapter will describe the general materials and methods used across multiple experiments in this thesis. Methods specific to only individual chapters will be discussed in detail in the methods within the relevant chapter itself.

## 2.2. Materials

### *2.2.1. DNA Extractions*

**Lysis Buffer for DNA Extraction:** 75mM NaCl, 10mM Tris-HCl, 25mM EDTA, 0.5% SDS

### *2.2.2. Agarose Gel Electrophoresis*

**SYTO 60:** 1µl stock SYTO 60 (Fisher Scientific; Cat No: 10194852), 99µl of distilled water ($dH_2O$). 10µl aliquots of this solution were then distributed into microcentrifuge tubes, where they were stored in a foil coated 50ml falcon flask at -20℃ until required.

**1x Tris-Borate-EDTA (TBE):** 100ml of 10x stock buffer (as provided), 900ml $dH_2O$.

*2.2.3. Miltenyi Biotec Mitochondrial Isolations*

**1x Protease Inhibition Buffer:** 1ml 10x Solution 2 (Miltenyi Biotec; Cat No: 130-097-340), 9ml double distilled water ($ddH_2O$), 1 protease inhibitor tablet (Sigma-Aldrich: P2308-100MG).

**1x Solution 2:** 1ml 10x Solution 2 (Miltenyi Biotec; Cat No: 130-097-340), 9ml $ddH_2O$.

**Extraction Buffer:** 40µl Solution1 (Miltenyi Biotec; Cat No: 130-097-340), 1ml 1x Solution 2 (Miltenyi Biotec; Cat No: 130-097-340).

**1x Separation Buffer**: 2.3ml 10x Separation Buffer (Miltenyi Biotec; 130-097-340), 20.7ml $ddH_2O$.

*2.2.4. MtDNA Extraction using Qiagen DNA Minikit*

**Buffer AW1:** 25ml of 96-100% ethanol, 19ml of the concentrated buffer (provided in the kit).

**Buffer AW2:** 30ml of 96-100% ethanol, 13ml of the concentrated buffer (provided in the kit).

*2.2.5. Bisulfite Treatment*

**EZ DNA Methylation-Gold Kit CT Conversion Reagent**: One vial of solid CT Conversion Reagent was resuspended in 900µl nuclease-free water, 300µl of M-Dilution Buffer and 50µl of M-Dissolving Buffer, after which the vial was subjected to frequent vortexing at room temperature for a minimum of 10 minutes.

## 2.3. Methods

*2.3.1. Total Genomic DNA Extraction*

DNA was isolated from frozen post-mortem human brain for use in a number of experiments presented in this thesis. The main objective when isolating DNA from such sources is to extract high molecular weight DNA, whilst removing contaminants such as proteins and salts. To do this, our group has developed an in-house, phenol-chloroform method (201), described below.

All extractions were carried out on 100mg samples of frozen, post-mortem human brain tissue. Briefly, tissue was added to 500µl of lysis buffer (see Section 2.2.1) in a 2ml microcentrifuge tube and lysed using a plastic pestle and motor (Sigma-Aldrich; Cat No: Z359947). After which, 2µl (10mg/ml) RNase-A (Sigma-Aldrich; Cat No: R6148-25ML) was added to each solution, which was then left to incubate at 37℃ for 2 hours. Following RNase-A treatment, 20µl of

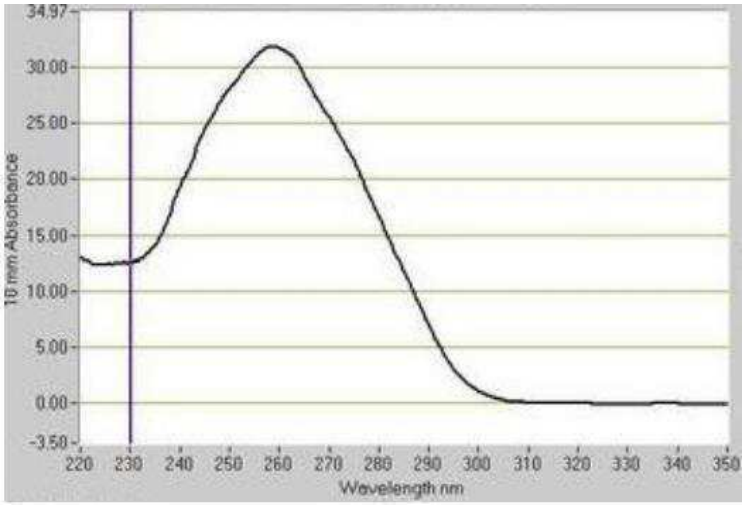
20mg/ml Proteinase K (Sigma-Aldrich; Cat No: P2308-100MG) was then added to each solution and left to incubate at 50℃ overnight in a water bath. During this incubation, in-house phase lock tubes were made by adding silicone grease (Sigma-Aldrich; Cat No: Z273554-1EA) to a fresh 2ml microcentrifuge tube to fill the inside of each lid. The microcentrifuge tubes were then centrifuged at 3500rpm for 10 minutes to pellet the grease. After the overnight incubation, 10µl of 20mg/ml Proteinase K (Sigma-Aldrich; Cat No: P2308-100MG) was then added to each sample, which were then returned to the water bath for 1 hour at 50℃. The samples were then heated to 65℃ for 30 minutes to heat inactivate the Proteinase K, after which samples were stored at 4℃ for 5 minutes to allow samples to cool to room temperature. Samples were then transferred into a phase lock tube and an equal volume of Phenol: Chloroform: Isoamyl Alcohol at a ratio of 25:24:1 (Thermo-Fisher; Cat No: 15593049) was added to each. Samples were inverted 10 times to mix thoroughly and were then centrifuged for 15 minutes at 13,000rpm. The top aqueous layer was then removed without disturbing the phase lock and added to a fresh microcentrifuge tube before an equal volume of chloroform was added to each sample, which were then inverted 10 times to mix, and centrifuged at 13,000rpm for 15 minutes. The top aqueous layer was then removed and this step was repeated one further time to improve purity. To precipitate the DNA, two volumes of 100% ice-cold ethanol was added to each sample and mixed by slowly inverting the tube 10 times, or until white, solid DNA precipitated out into aggregates. Samples were then left at -20℃ overnight to aid the precipitation. Samples were then centrifuged at 13,000rpm for 20 minutes, following which the supernatant was removed, with care not to dislodge the pellet. 1ml of 70% ethanol was added to each sample, which were then centrifuged at 13,000rpm for 5 minutes. The supernatant was

subsequently removed using a pipette and the DNA pellets were left to air dry for approximately 30 minutes. Samples were then resuspended in 200µl of $dH_2O$ and incubated at 37°C for 2 hours. Samples were then vortexed, and checked for purity using the Nanodrop ND8000 and Agarose Gel Electrophoresis.

### *2.3.2. Spectrophotometric Analysis Using Nanodrop D8000*

To quantify and check the purity of DNA samples, spectrophotometric analysis was performed. Briefly, DNA absorbs light at around 260nm, whilst protein contaminants absorb at 280nm and other contaminants, such as ethanol, absorb at around 230nm. A greater absorbance at 260nm, compared to 280nm or 230nm, is therefore an indicator of highly concentrated DNA. Ratios of 260/280nm and 260/230nm should lie between the ranges of 1.70-1.90 and 2.00-2.20 respectively as an indicator of good purity (**Figure 2.1**). For accurate quantification, 1µl of DNA was added to the Nanodrop ND8000 as per the manufacturer's instructions. Further details can be found at: http://www.nanodrop.com/Library/nd-8000-v2.2%20-users-manual-8.5%20x%2011.pdf.

***Figure 2.1: A typical spectrophotogram produced using Nanodrop ND8000****. The sample has a peak at 230nm, with a 260/230nm ratio between 2.00 and 2.20 and a 260/280nm ratio between 1.70 and 1.90, indicating DNA with good purity.*




*2.3.3. Agarose Gel Electrophoresis*

To allow for the separation and visualisation of DNA products, 3% agarose gels were made. Typically, 3g of Agarose powder (Sigma Aldrich; Cat No: A9539-500g) was added to 100ml of 1x TBE (Sigma Aldrich; Cat No: T4415-1L) (see Section 2.2.2.) in a conical flask**.** The mixture was then heated until all agarose was dissolved, after which 10 µl of SYTO 60 (see Section 2.2.2.) was added per 100ml of TBE and the solution was gently mixed. The solution was then poured into a gel cassette equipped with a comb and sealed at either end with laboratory tape. The cassette was then wrapped in paper towels and placed in the refrigerator at 4℃ for 15 minutes to allow sufficient time for setting. Once set, the cassette containing the gel was then removed from the refrigerator and the tape and combs were removed prior to placing in a gel tank. Up to 5 µl (or

100ng) of DNA was mixed with 1 µl of 6x Orange G (New England Biolabs; Cat No: B7022S). Each sample mixture was then loaded into an individual well. One well was left per row for the addition of a 100bp ladder (New England Biolabs; Cat No: N3231S) for ncDNA or 1Kb ladder (New England Biolabs; Cat No: N3232S) for mtDNA. The gel tank was set to run for 45 minutes at 120V. Following this period, gels were visualised on the Licor Odyssey CLx to assess DNA quality. For example, good quality, high molecular weight mtDNA will appear as a single band around 16.6kb whereas a single smear is indicative of degraded DNA or ncDNA contaminants.

### *2.3.4. Mitochondrial Isolation*

To reduce the risk of nuclear contamination in the analysis of mitochondrial methylation, several methods were tested for their efficiency of isolating mtDNA from frozen, post-mortem brain tissue (202), which will be discussed later in Chapter 4. However, the most efficient method, which was then utilised in Chapters 5 and 6, is detailed below.

With a few exceptions, mitochondria were isolated from tissue as per the manufacturer instructions. To reduce tissue degradation, centrifugation steps associated with the isolation of mitochondria were carried out at 4℃. All reagents were prepared fresh and kept on ice. Briefly, mitochondria were extracted from 200mg of human, post-mortem brain tissue using the Miltenyi Biotech mitochondria kit for humans (Miltenyi Biotech; Cat No: 130-097-340). First, the tissue was rinsed twice in 3-4ml ice cold 1x Solution 2 (see Section

2.2.3), The tissue was quickly transferred to a 35mm cell culture dish (over ice), containing a further 3-4ml of ice cold 1x Solution 2. Visible sections of fat were then resected and the tissue was cut into 1-2mm sections using a scalpel. Solution 2 was then aspirated off and the sample was then transferred to a 2ml microcentrifuge tube containing 1ml extraction buffer (see Section 2.2.3.) and incubated on ice for 30 minutes. After 30 minutes, this solution was then centrifuged at 300 x g for 5 minutes at 4℃ and the supernatant was discarded. 2ml ice cold Protease Inhibition Buffer (see Section 2.2.3.) was then added to resuspend the pellet. This suspension was then poured into a pre-cooled C Tube (Miltenyi Biotec; Cat No: 130-093-237) and an additional 2ml ice-cold Protease Inhibition Buffer was added. The C Tube was then closed and the tissue was gently homogenised using the gentleMACS dissociator (Miltenyi Biotec; Cat No: 130-093-235) by running gentleMACS Program h_mito_tissue_01. The C Tube was then centrifuged at 200 x g for 30 seconds at 4℃ to collect the homogenate at the bottom of the tube. A Pre-Separation 70um Filter (Miltenyi Biotec; Cat No: 130-095-823) was then placed on a pre-cooled 15ml conical tube over ice. The homogenate from the C Tube was then transferred to the reservoir of the filter, allowing for the homogenate to pass through. The filter was then washed twice with 1ml ice-cold Solution 3 (provided with the kit). The homogenate was then centrifuged at 500 x g for 5 minutes at 4℃ and the supernatant containing mitochondria was transferred to a 15ml conical tube. The pellet containing the nuclear contaminants was stored at -80℃ for later use as an unenriched mitochondrial fraction. Proceeding immediately, ice-cold 1x Separation Buffer (see Section 2.2.3.) was added to the supernatant to a total volume of 10ml. 100ul Anti-TOM22 MicroBeads (Miltenyi Biotec; Cat No: 130-094-532) were added to magnetically label the

mitochondria. The solution was then mixed well and incubated for 1 hour at 4℃ under continuous rotation using the MACS mix Tube Rotator, (Miltenyi Biotec; Cat No: 130-090-753) at a permanent run speed of 12rpm. Next, an LS Column (Miltenyi Biotec; Cat No: 130-042-401) was placed into the magnetic field of a MidiMACS separator (Miltenyi Biotec; Cat No: 130-042-302) and a Pre-Separation 30µm Filter (Miltenyi Biotec; Cat No: 130-041-407) was placed on top of the LS column. The column was then prepared by rinsing with 3ml of 1x Separation Buffer. Magnetically labelled mitochondria were added to the column in three aliquots of 3.3ml, adding only when the column reservoir was empty. After, the column was washed with three lots of 3ml of 1x Separation Buffer before being removed from the magnet and placed in a 15ml conical tube. 1.5ml of 1x Separation Buffer was immediately pipetted onto the column and the enriched mitochondria were flushed out by firmly pushing the plunger into the column. The mitochondria were then centrifuged at 13,000 x g for 2 minutes at 4℃. Following this, the supernatant was aspirated off and the mitochondrial pellet was resuspended in 100µl of Storage Buffer. It was then centrifuged again at 13,000 x g for 2 minutes at 4℃ to improve purity. The supernatant was aspirated off and the pellet was taken immediately to the next stage.

DNA was then isolated from the mitochondrial pellet using the QIAamp DNA Micro Kit (Qiagen; Cat No: 56304), as per the manufacturer's instructions. 180µl Buffer ATL was added to the mitochondrial pellet and allowed to equilibrate to room temperature. 20µl 10mg/ml Proteinase K, (Sigma-Aldrich; Cat No: P2308-100MG) was then added and mixed using a vortex for 15 seconds. The tube was then incubated at 56℃ for 3 hours. After incubation, 200µl Buffer AL was

added and mixed by pulse-vortexing for 15 seconds. Next, 200µl ethanol (96–100%) was added and mixed thoroughly by pulse-vortexing for 15 seconds. The solution was then incubated for 5 minutes at room temperature and then briefly centrifuged. The entire lysate was then transferred to a QIAamp MinElute column (in a 2 ml collection tube) and centrifuged at 6000 x g for 1 minute. The QIAamp MinElute column was then placed in a clean 2 ml collection tube. 500µl Buffer AW1 (see Section 2.2.4.) was then added to the QIAamp MinElute column, which was centrifuged at 6000 x g for 1 minute. The QIAamp MinElute column was then placed in a clean 2 ml collection tube. 500µl Buffer AW2 (see Section 2.2.4) was then added to the QIAamp MinElute column, which was centrifuged at 6000 x g for 1 minute. The QIAamp MinElute column was then placed in a clean 2 ml collection tube. Following this, the column was centrifuged at 20,000 x g for 3 minutes to dry the membrane completely. The QIAamp MinElute column was then placed in a clean 1.5 ml microcentrifuge tube and 50µl Buffer AE was applied to the centre of the membrane. The lid was then closed and the column was left to incubate at room temperature for 5 minutes before centrifugation at room temperature for 1 minute at 20,000 x g. Finally, this step was repeated with a further 50µl Buffer AE, to give a total elution volume of 100µl. DNA was then quantified as described in Sections 2.3.2 and 2.3.3.

### *2.3.5. MtDNA Copy Number Quantification*

To accurately determine the purity of mtDNA enrichments from mitochondrial isolation methods, a qRT-PCR was utilised to simultaneously assess the number of mtDNA and ncDNA copies within each sample. NcDNA and mtDNA

were extracted simultaneously using the method described in Section 2.3.4. Primers specific for a mitochondrial gene (with no known *NUMTs*), *MT-CYB,* and a nuclear housekeeping gene, *human beta 2-microglobulin,* (*hB2M)* were utilised to assess mitochondrial purity as previously described (203) see **Table 2.1**).

***Table 2.1: Primer sequences for mitochondrial copy number determination.***

| | ***hB2M*** | ***MT-CYB*** |
|---|---|---|
| Forward Sequence | TGCTGTCTCCATGTTTG ATGTATCT | CCTCATCCTAGCAATAATCCC CATCCTCCATATA |
| Reverse Sequence | TCTCTGCTCCCCACCTC TAAGT | ACTTGTCCAATGATGGTAAAA GG |

To determine the optimal conditions of the assay for use in our laboratory, a temperature gradient PCR was performed to determine the annealing temperature at five, 2℃ intervals between 50-60℃ for both primer sets using volumes described in **Table 2.2** and conditions described in **Table 2.3**. This revealed an optimum temperature at 60℃ for both primer sets. qRT-PCR reactions were then performed using the same volumes as described in **Table 2.2**, checking for a single peak in the melt curve assay, demonstrating the specificity of the primers designed. The specificity of the primers was also checked using gel electrophoresis (see Section 2.3.2).

***Table 2.2: Volumes added to each well for a 10µl PCR reaction.***

| Reagent | Volume (µl) |
|---|---|
| Forward primer (10µM) | 0.25 |
| Reverse primer (10µM) | 0.25 |
| EvaGreen (Solis Biodyne; Cat No: 08-26-00001) | 2 |
| $ddH_2O$ | 5.5 |
| DNA | 2 |
| **Total** | **10 µL** |

***Table 2.3: PCR conditions for gradient PCR.***

| Stage | | PCR temperature | Time |
|---|---|---|---|
| Hot start | | 95°C | 15 minutes |
| 30x Cycles | Denaturing | 95°C | 10 seconds |
| | Annealing | 50-60°C | 30 seconds |
| | Extension | 72°C | 1 minute |
| Final Elongation | | 72°C | 7 minutes |

We generated standard curves for each gene of interest, by scaling up the PCR volumes shown in **Table 2.2**, to 50µl and by increasing the number of PCR cycles to 50. After PCR we determined the yield of DNA using the Nanodrop, and using Avogadro's constant, the length of the amplicon and the yield of DNA, to determine the number of copies of the amplicon in the sample. The sample was then diluted to $10^{10}$ copies/µl (Standard 10) and a subsequent serial dilution was generated. Briefly, 10 µl of Standard 10 was added to 90 µl of nuclease-free $H_2O$. The solution was then vortexed for 30 seconds and centrifuged. This process of ten-fold dilutions was repeated for each standard from Standard 10 ($10^{10}$ copies/µl) to Standard 1 ($10^{1}$ copies/µl). The standard curves for both mtDNA and ncDNA were run in duplicate alongside duplicate DNA samples using the conditions described in **Tables 2.2** and **2.4** on a Taqman 7900HT (ThermoFisher; Cat No: 4329001).

Briefly, qRT-PCR relies on the exponential increase in specific DNA fragments during PCR to accurately quantify DNA. DNA is labelled with a fluorescent dye. Accordingly, the amount of fluorescence released during amplification is directly proportional to the concentration of the amplicon. To avoid the effects of background fluorescence, a Ct value is set by the computer software. This is deemed the threshold cycle, i.e. the number of cycles of amplification required for DNA to exceed background fluorescent levels. Ct values are therefore inversely proportional to both fluorescence emission and DNA concentration.

***Table 2.4: qRT-PCR conditions used for determination of mtDNA purity.***

| Stage | | PCR temperature | Time |
|---|---|---|---|
| Hot start | | 95 °C | 15 minutes |
| 30x Cycles | Denaturing | 95 °C | 10 seconds |
| | Annealing | 60 °C | 30 seconds |
| | Extension | 72 °C | 1 minute |
| Dissociation step | | 95 °C | 15 seconds |
| | | 60 °C | 15 seconds |

All analyses were performed in Microsoft excel; samples where duplicate variation exceeded a 0.5 Ct difference were excluded and re-run. Further, Ct variation between standards in the serial dilution was used to inspect PCR efficiencies. Any qRT-PCR efficiency not within 90-110% was also re-run. We then calculated the number of mtDNA copies/ncDNA copy by dividing the number of copies of *MT-CYB* by the number of copies of *hB2M* in each sample.

*2.3.6. Capture of MtDNA Methylome*

To assess mitochondrial methylation status across the mitochondrial genome, a custom targeted bisulphite sequencing assay was designed. This Sure Select assay (Agilent Technologies; Design ID: 0687721) consisted of 656 probes which covered the entire mitochondrial genome using the most stringent repeat masking parameters and a 5x tiling density. This was then utilised in Chapters 5 and 6.

This experiment was carried out in accordance with the manufacturer's specifications for 1µg DNA, with a few exceptions due to both the quantity of mtDNA available and the small size of the mitochondrial genome. Briefly, mtDNA was concentrated to a final volume of 30µl. Salmon sperm DNA (Sigma-Aldrich; Cat No: 15632-011) was then added to each sample as a carrier DNA to yield a total volume of 50µl (20ng/µl). Each sample was then placed on ice for 30 minutes and the DNA was then sheared using a Bioruptor (Diagenode; Cat No: B01020001) at 4℃ for 60 minutes for 30 second on/off cycles. After shearing, DNA quality was assessed and fragment size was determined using the Agilent 2200 TapeStation System (Agilent Technologies; Cat No: G2964AA) and DNA 1000 Screentape (Agilent Technologies; Cat No: 5067-5582), as per protocol guidelines. For samples with fragments larger than 100-175bp, DNA was re-sonicated. After DNA sonication, samples were kept on ice whilst End Repair master mix was prepared, as per the manufacturer's instructions. Briefly, 52µl of the master mix was added to each 48µl sample, which were then vortexed and briefly centrifuged, before being incubated at 20℃ for 30 minutes in a thermal cycler. 180µl of Agencourt AMPure XP beads

(Beckman Coulter; Cat No: A63881) were then added to each 100µl sample, ensuring that AMPure XP beads were equilibrated at room temperature for at least 30 minutes prior to the addition. Samples and beads were pipette mixed 10 times and incubated for 5 minutes at room temperature. The plate was then placed into a magnetic separation device and further incubated for 7-10 minutes until the sample cleared. Whilst the plate was in the magnetic stand, the cleared solution was carefully discarded and 200µl of 70% ethanol was added to each well containing beads. The beads were allowed to resettle, taking approximately 1 minute before removing the ethanol. This ethanol wash was then repeated. The wells were then sealed and the plate was briefly centrifuged before being placed on the magnetic stand to remove any residual ethanol. To ensure the complete removal of ethanol, the unsealed plate was then placed on a heat block at 37℃ for 5 minutes. After ethanol evaporation, 44µl of nuclease free water was added to each well, the plate was sealed, vortexed, briefly centrifuged and incubated at room temperature for 2 minutes. The plate was then placed in the magnetic stand and left for 3 minutes until the solution cleared. The cleared supernatant was then removed and placed into a fresh well and the beads were discarded.

The samples were then transferred to ice and the 3' ends of the DNA were adenylated. Briefly, 9µl of adenylation master mix, made per protocol guidelines, was added to each 41µl sample, and then briefly vortexed and centrifuged. The samples were then incubated in the thermal cycler at 37℃ for 30 minutes. Upon removal from the thermal cycler, the samples underwent another stage of AMPure clean up. 90µl of AMPure XP beads were then added

to each 50µl sample, with AMPure XP beads allowed to equilibrate to room temperature for at least 30 minutes prior to the addition. Samples and beads were pipette mixed 10 times and incubated for 5 minutes at room temperature. The plate was then placed into a magnetic separation device and further incubated for 5 minutes until the sample cleared. Whilst the plate was in the magnetic stand, the cleared solution was carefully discarded and 200µl of 70% ethanol was added to each well containing beads. The beads were allowed to resettle, taking approximately 1 minute, before removing the ethanol. This ethanol wash was then repeated. The wells were then sealed and the plate was briefly centrifuged before being placed on the magnetic stand to remove any residual ethanol. To ensure the complete removal of ethanol, the unsealed plate was then placed on a heat block at 37℃ for 2 minutes. After ethanol evaporation, 35µl of nuclease free water was added to each well, the plate was sealed, vortexed, briefly centrifuged and incubated at room temperature for 2 minutes. The plate was then placed in the magnetic stand and left for 3 minutes until the solution cleared. The cleared supernatant was then removed and placed into a fresh well and the beads were discarded.

The purified DNA was then ligated using a methylated adaptor ligation mix, prepared as per the protocol guidelines. Briefly, 16.5µl of the ligation mix was added to each 33.5µl DNA sample, which was then briefly vortexed and centrifuged before being incubated at 20℃ for 15 minutes in a thermal cycler. After ligation, DNA was purified using AMPure XP beads. 90µl of AMPure XP beads were then added to each 50µl sample (AMPure XP beads were allowed to equilibrate to room temperature for at least 30 minutes prior to the addition).

Samples and beads were pipette mixed 10 times and incubated for 5 minutes at room temperature. The plate was then placed into a magnetic separation device and further incubated for 5 minutes until the sample cleared. Whilst the plate was in the magnetic stand, the cleared solution was carefully discarded and 200µl of 70% ethanol was added to each well containing beads. The beads were allowed to resettle, for approximately 1 minute, before removing the ethanol. This ethanol wash was then repeated. The wells were then sealed and the plate was briefly centrifuged, before being placed on the magnetic stand to remove any residual ethanol. To ensure the complete removal of ethanol, the unsealed plate was then placed on a heat block at 37℃ for 2 minutes. After ethanol evaporation, 22µl of nuclease free water was added to each well, the plate was sealed, vortexed, briefly centrifuged and incubated at room temperature for 2 minutes. The plate was then placed in the magnetic stand and left for 3 minutes until the solution cleared. The cleared supernatant was then removed and placed into a fresh well and the beads were discarded. The DNA was then assessed for quality and quantity using the DNA D1000 Assay as previously described.

After initial testing within the group, it was determined that the previous method of hybridisation (January 2013), provided consistent results across experiments. As such this method was utilised for all experiments described in this thesis. Briefly, DNA was concentrated to a total volume of 3.4µl using a vacuum concentrator at 45℃. The samples were then briefly vortexed and centrifuged. Hybridization Buffer and Block Mix were made to specifications indicated on the protocol, however a 10% RNase block solution was made using 0.2µl RNase

Block and 1.8µl of nuclease free water (as recommended by Agilent technical support). Further, given the size of the genome, only 2µl of the custom capture library was added to a PCR plate containing 2µl RNase block solution and 3µl of nuclease free water (as recommended by Agilent technical support). In a separate plate, the 3.4µl samples were added to Row B. 5.6µl of freshly prepared Block Mix was then added to the samples, which were pipette mixed and the wells were sealed. The PCR plate containing samples and Block Mix was then added to a thermal cycler and heated to 95℃ for 5 minutes before being held at 65℃ for 5 minutes, with a heated lid set at 105℃. After 5 minutes, 7µl capture library mix was added to Row C of the plate, which was then sealed and left to incubate at 65℃ for 2 minutes. 13µl of Hybridisation buffer was added to the Sure Select capture library, which was then followed by the addition of the prepared library to the capture library. The contents of the wells were then pipette mixed 10 times before being sealed using two layers of adhesive tape and left to incubate at 65℃ for 24 hours.

Sure Select Wash Buffer 2 (provided with kit) was incubated at 65℃ in a water bath. Dynabeads MyOne Streptavidin T1 magnetic beads (ThermoFisher Scientific; Cat No: 65601) were then prepared prior to removal of the plate from the thermal cycler. Briefly, the magnetic beads were vortexed for 30 seconds and 50µl per sample was added to a plate containing 200µl of SureSelect binding buffer (provided with kit). The beads and buffer were then pipette mixed 10 times. The plate was then placed on a magnetic separation device and the solution was allowed to clear for 5 minutes before removing the supernatant. This washing was then repeated two more times and the beads were re-

suspended in binding buffer (provided with kit). The DNA extraction mixture was then immediately transferred into the plate containing the Streptavidin beads and incubated at room temperature for 30 minutes, during which time three 200µl aliquots per sample of Wash Buffer 2 were warmed to 65℃. After the DNA-bead hybridisation, the plate was briefly centrifuged and placed in a magnetic separator until the solution cleared, at which point the supernatant was removed. The beads were then re-suspended in Sure Select Wash Buffer 1 (provided with kit) and pipette mixed until fully re-suspended. The beads were then left at room temperature for 15 minutes, after which the plate was briefly centrifuged. Beads were then placed back into the magnetic separator and the supernatant was discarded after the solution cleared.

Beads were then re-suspended by pipette mixing in 200µl of Wash Buffer 2, which had been pre-warmed to 65℃. Wells were then capped and incubated at 65℃ for 10 minutes, after which they were placed in a magnetic separator and the supernatant was removed upon the solution clearing. This wash and incubation step was repeated twice more, ensuring that all Wash Buffer 2 was removed in the final wash. 20µl of Sure Select Elution Buffer (provided with kit) was then added to the beads, the wells were capped and the solution was vortex mixed for 5 seconds before being left to incubate at room temperature for 20 minutes. During this period, one vial of CT Conversion Reagent (Section 2.2.5.) from the EZ DNA Methylation Gold Kit (Zymo Research; Cat No: D5005) was prepared. Following the 20-minute incubation, the beads were placed on a magnetic separator and the supernatant was transferred into a new plate upon clearing. 130µl of CT Conversion Reagent was then added to each 20µl

captured library, which was then vortexed and briefly centrifuged. The plate was then placed in a thermal cycler and incubated at 64℃ for 2.5 hours.

Following incubation, 600µl of M-Binding Buffer (provided with kit) was added to each Zymo-Spin IC Column in collection tubes, followed by 150µl of BS-converted DNA. The columns were then capped, inverted five times and centrifuged at 13,000rpm. The flow through was then discarded and the column was placed back in the same collection tube. The column was then washed by adding 100µl of M-Wash Buffer (provided with kit), and centrifuged for 60 seconds at 13,000rpm. The flow through was discarded and the column was placed back into the same collection tube. 200µl of M-Desulphonation Buffer (provided with kit) was then added to each column, and they were then incubated for 20 minutes at room temperature before being centrifuged at 13,000rpm for 60 seconds. The flow through was then discarded and the column was placed back in the same collection tube. Each column was then subjected to a further wash step, by adding 200µl M-Wash Buffer, followed by another centrifugation step at 13,000rpm for 60 seconds. The column was then placed in a fresh 1.5ml microcentrifuge tube and incubated at room temperature for 2 minutes. 10µl of M-Elution Buffer (provided with kit) was then added to each column, and the samples were then incubated at room temperature for 3 minutes before being centrifuged at 13,000rpm for 60 seconds. The flow through was retained in the collection tube whilst a further 10µl of M-Elution Buffer was added to the column. Each column was again incubated at room temperature for 3 minutes prior to centrifugation at 13,000rpm for 60 seconds. Following desulphonation, samples were immediately taken to the next step:

PCR amplification of bisulphite-treated libraries. The PCR reaction mixture was prepared as per protocol guidelines, however given the small genome size and low DNA input, the number of PCR cycles were increased (see **Table 2.5**). AMPure XP beads were incubated at room temperature for 30 minutes. 180µl of homogenous AMPure XP beads were added to each sample and pipette mixed 10 times before being incubated at room temperature for 5 minutes. The plate containing the samples was then placed in a magnetic separator for 10 minutes, until the solution had cleared, at which point the supernatant was removed. 200µl of 70% ethanol was then added to each sample and left to incubate for 1 minute before the supernatant was removed. The ethanol wash was then repeated. At the end of the second wash the plate was placed in a heat block at 37℃ for 4 minutes until residual ethanol had evaporated. 21µl of nuclease-free water was then added to each sample, and samples were then pipette mixed and left to incubate for 2 minutes at room temperature, after which the plate was placed in the magnetic separator and incubated for a further 3 minutes. When the solution cleared, the supernatant was removed and placed in a fresh well.

***Table 2.5: Bisulfite PCR conditions for mtDNA methylation.***

| Segment | Number of Cycles | Temperature | Time |
|---|---|---|---|
| 1 | 1 | 95℃ | 2 mins |
| 2 | 14 | 95℃ | 30 seconds |
| | | 60℃ | 30 seconds |
| | | 72℃ | 30 seconds |
| 3 | 1 | 72℃ | 7 mins |
| 4 | 1 | 4℃ | Hold |

A PCR reaction mixture for sample indexing was then made as per the protocol guidelines. 25.5µl of the PCR Indexing Reaction Mix was added to each 19.5µl sample with 5µl of the appropriate Indexing Primer. This mixture was then subjected to PCR amplification using the Agilent Methyl-Seq protocol conditions before being analysed using D1000 High Sensitivity Screentape (Agilent Technologies; Cat No: 5067-5584) and pooled, as per protocol guidelines. Prior to being sent to the Exeter Sequencing Service (http://sequencing.exeter.ac.uk/), the sample pool was analysed using the D1000 High Sensitivity Screentape to determine concentration. The sample pool was then run on the Illumina HiSeq 2500, high Output format, using 100bp paired-end tags.

### *2.3.7. Determination of MtDNA Methylation State*

For sequencing alignments presented in Chapters 5 and 6, the most current version of the human mitochondrial genome, the revised Cambridge Reference Sequence, (rCRS), was downloaded from ensemble (ftp://ftp.ensembl.org/pub/release-83/fasta/homo_sapiens/dna/). Fastq files were unpackaged and indexed using bismark, a BS-specific alignment tool which utilises Bowtie and/or Bowtie 2 (204).

Following sequencing, paired-end fastq files were de-multiplexed and quality checked with FastQC, using default options. Paired-end fastq files were then trimmed for adaptors and for the first base pair from the 3' end of both reads, to improve later alignment mapping efficiencies using trim-galore.

Following trimming, files were aligned to rCRS under default settings. Following alignments, mapping efficiencies and sequencing depth were recorded and the Sequence Alignment/Map files were deduplicated using the shell function deduplicate_bismark before methylation levels in each cytosine context were determined using the shell function bismark_methylation_extractor, both can be found within the Bismark program (204). The percentage methylation at each methylated cytosine, for each sample, was then extracted and all further analyses were carried out in R statistical environment version 3.2.1 (205). For all analysis, a nominally significant threshold was defined at $P<0.05$. Bonferroni correction thresholds were defined as 0.05/N, with N being defined as the number of tests implemented. For most instances, this corresponds directly to the number of probes, or windows, present in each study, unless stated otherwise.

# Chapter 3. Investigating the use of Publicly Available Data for the Determination of MtDNA Methylation Patterns Across Different Tissue Types

***The work described in this chapter is currently in press at the journal "Clinical Epigenetics". A copy of the* accepted *article can be found in Appendix 3.***

## 3.1. Aims

1. To create a pipeline to account for the presence of *NUMT*s in total genomic DNA using publicly available MeDIP-Seq datasets.
2. To assess the tissue specificity of mtDNA methylation in a publicly available MeDIP-Seq dataset.

## 3.2. Introduction

Mitochondria are unique organelles in that they have their own, circular genome, approximately 16.6kb in size (25). mtDNA consists of 37 genes, 22 encoding for tRNAs, two for rRNAs and 13 encoding for proteins important in the ETC. However, mitochondria have an array of other important cellular roles such as $Ca^{2+}$ homeostasis (206) and neural stem cell differentiation (207). As such, abnormal mitochondrial function, dynamics and trafficking have been associated with a number of brain disorders including AD (208,209), schizophrenia (210), bipolar disorder (211) and major depressive disorder (212).

Epigenetic processes mediate the reversible regulation of gene expression, occurring independently of DNA sequence variation, acting principally through chemical modifications to DNA and nucleosomal histone proteins to orchestrate

a diverse range of important physiological functions. DNA methylation is the best characterised and most stable epigenetic modification, modulating the transcription of mammalian genomes. Further, as it can be robustly assessed using existing genomic DNA resources, it is also the focus of most human epidemiological epigenetic research to date (191). The most widely used method for epigenome-wide analysis of DNA methylation is the Illumina 450K methylation array and a number of studies have recently shown differential DNA methylation of ncDNA, between different tissue types (213-215) and also in a range of complex diseases, from brain disorders such as AD (198,199,216) and schizophrenia (217,218), to systemic diseases such as type 2 diabetes (219) and Crohn's disease (220). However, with no representation of the mitochondrial genome on this platform, as well as a lack of analysis on other genome-wide platforms, the role of mtDNA methylation has been largely neglected (1,221).

Since the identification of 5-mC in mitochondria, research into mtDNA methylation as an independent, and potentially relevant, mark has received more regular attention (72,73). However, most research is either focussed on low resolution, global DNA methylation, or candidate gene DNA methylation changes using techniques such as BS-pyrosequencing (1). These recent publications have indicated that differences in mtDNA methylation are present in a variety of different phenotypes (75,80-83,200) and may have potential utility as a biomarker (153). In addition, a recent study has explored the use of MeDIP-Seq to investigate changes in mtDNA methylation across 39 cell lines and tissues from publicly available data (63). At present, genome-wide

sequencing technologies have not yet been used to interrogate alterations in the mtDNA methylome across tissues in the same individuals.

A high proportion of current, publicly available, genome-wide DNA methylation data has been generated through the use of MeDIP-Seq, a method designed to interrogate genome-wide changes in methylation at a high throughput and for a low cost (100). However, given the presence of *NUMT*s, mitochondrial reads are often discarded from further analysis. The development of bioinformatic pipelines to investigate regions of differential mtDNA methylation from whole genome data would provide a novel way in which to interrogate the mtDNA methylome in publicly available data. Here, we control for the presence of *NUMTs* in a previously published MeDIP-Seq dataset, to investigate differential DNA methylation across the mitochondrial genome in human post-mortem brain samples.

## 3.3 Methods

### *3.3.1. Data Collection*

We utilised publicly available MeDIP-Seq data from Davies *et al* (158). In brief, this data was generated using 5µg fragmented gDNA, which, following end repair <A> base addition and adaptor ligation, was immunoprecipitated using an anti-5mC antibody (Diagenode, Liège, Belgium). MeDIP DNA was purified and then amplified using adaptor-mediated PCR, with DNA fragments between 220-320bp subjected to highly parallel 50bp paired-end sequencing on the Illumina

Hi-Seq platform. The paired-end, raw fastq files were provided by the authors and quality checked using FastQC. Sample information is provided in **Table 3.1**.

***Table 3.1: Demographic information for MeDIP study****. MeDIP-Seq data was available from post-mortem brain samples obtained from three individuals free of any neuropathology and neuropsychiatric disease. Data was available for five different regions of the cortex (CTX), Brodmann areas (BA) 8, 9 and 10, superior temporal gyrus (STG) and ECX, the cerebellum (CER) and pre-mortem blood (BLD). MeDIP-Seq data was available for all individuals from cortical and cerebellar samples, however blood MeDIP-Seq data was not available for individual 2. Data is freely available to download from http://epigenetics.iop.kcl.ac.uk/brain.*

| Individual | Age at Death (Years) | Age at Bloods Sampled (Years) | Post-Mortem Delay (Hours) | Sex |
|---|---|---|---|---|
| 1 | 82 | 79 | 43 | Female |
| 2 | 92 | N/A | 17 | Female |
| 3 | 78 | 78 | 10 | Male |

*3.3.2. Quality Control and NUMT exclusion*

Fastq files were subjected to adaptor and Phred score (q<20) trimming. In an attempt to remove any potential contamination of possible *NUMT*s, multiple alignments to the reference genome were undertaken. Briefly, paired fastq files were aligned to GRCH37 using BWA. Unique and mapped reads aligning to the mitochondria were then re-mapped to a custom GRCH37 reference without the mitochondrial chromosome. Reads not mapping to the custom reference were

then taken forward and realigned to the full GRCH37 reference to eliminate the possibility of homologous regions mapping falsely to the mitochondrial genome. All alignments were carried out using BWA mem and default settings.

Reads Per Kilobase of transcript per Million mapped reads (RPKM) values for each sample were calculated using the MEDIPS R package (222). Methylation was averaged across 100bp non-overlapping windows (default parameter setting in MEDIPS) and only windows with read counts >10 were considered for analysis. Due to the non-normal distribution of all cohorts, RPKM values were $log^2$ transformed before statistical analysis.

*3.3.3. Statistical analyses*

All analyses were performed in the R statistical environment version 3.2.1 (205). For our initial regression analysis , a nominally significant threshold of P<0.05, and a Bonferroni significant threshold of P<1.39E-04 were used**,** with the number of tests defined as five (number of cortical regions) multiplied by 72 (windows). Given the matched nature of this cohort, two tailed, paired t-tests were performed at each window along the mitochondrial genome to identify DMRs between individual cortical regions and CER.

To compare CTX to CER, we performed a multilevel mixed-effects model in the Lme4 package in R (223), using brain region as the random effect and individual as the fixed effect. A nominally significant threshold of P<0.05 and a bonferroni significant threshold of P<7.40E-04 were used. To assess the similarity of brain regions we used the R function “hclust” to cluster average

RPKM values for brain regions using Euclidean distance. We used the R function corrgram within the corrgram package (224) to order samples based upon the similarity of their principal components.

## 3.4 Results

### *3.4.1. MtDNA Methylation Patterns are Correlated between the CTX, CER and Blood.*

To date, no study has investigated differences in mtDNA methylation across different matched regions of human brain and blood samples. Our cohort (**Table 3.1**) consisted of MeDIP-Seq data from three individuals, free of any neuropathology and neuropsychiatric disease, for five different regions of the CTX: BA 8, 9 and 10, STG and ECX, the CER and pre-mortem blood (158).

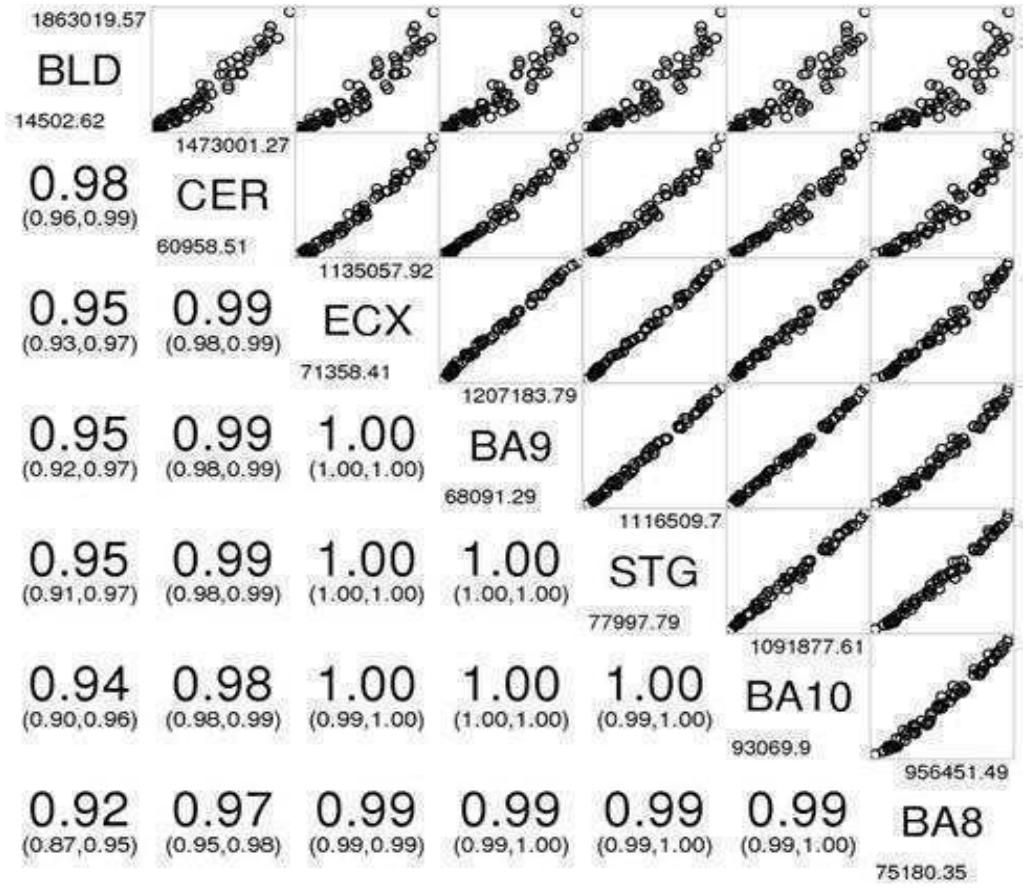
Given that MeDIP-Seq data has been generated from standardly extracted total genomic DNA, and thus contains a mixture of ncDNA and mtDNA (225), we initially controlled for regions of high sequence homology between the two genomes within our data by realigning mtDNA reads to a series of custom reference genomes using an in-house pipeline to specifically analyse mtDNA methylation, see Section 3.3.2.

Initially we were interested to investigate whether changes in mtDNA across the mitochondrial genome were highly correlated between different tissue types. Using principal component analysis (PCA) we found that mtDNA methylation

patterns are highly correlated between different cortical regions, ($r>0.99$, $p<2.2E-16$), with a slightly weaker correlation between CER and CTX ($r>0.97$, $p<2.2E-16$) (**Figure 3.1**).

Due to the small number of blood samples available, deriving a significance level for the correlations between CER and blood could not be made. Instead, in an attempt to explore the similarity between matched blood and CER samples, direction of differential methylation from the CTX was used. Here we found that 93.1% of the windows analysed in CER and blood had the same direction of methylation difference from the CTX, further suggesting a strong correlation between the two tissue types.

***Figure 3.1: Correlogram to show similarity between each tissue type.*** *Principal components of each sample were determined in R and the correlation (R) between these components were plotted to determine the similarity between each average tissue type. For each tissue, the mean and standard deviation was calculated and shown above and below the respective tissues.*



### *3.4.2. Differentially Methylated Regions of the Mitochondrial Genome can be Identified between Anatomically Distinct Cortical Regions and CER*

Having identified correlated mtDNA methylation patterns across different brain regions, we were interested to investigate whether we could identify DMRs in the mitochondrial genome between different regions of the CTX and CER. To identify such tissue-specific DMRs within the mitochondrial genome, paired t-tests were performed within matched cortical and CER samples at 100bp windows across the mitochondrial genome. In total, we identified 74 nominally significant DMRs ($P<0.05$) between the five individual cortical regions and the CER (**Table 3.2; Figure 3.2**). Of these DMRs, seven (**Table 3.2**, bold face) were found to be present across all prefrontal CTX areas (BA8, BA9, BA10). Furthermore, the direction of methylation difference was maintained in all Brodmann area regions, with three conserved regions of hypomethylation and four conserved regions of hypermethylation, with respect to CER. Furthermore, four of the seven conserved regions were adjacent to each other within the mitochondrial *D-LOOP* (16201-16600bp), a region associated with gene transcription and DNA replication.

***Table 3.2: List of DMRs identified between five anatomically discreet cortical regions and CER.*** *Shown is the location of the DMR within the mitochondrial genome (ChrM) (based on GENCODE), the gene(s) residing within the 100bp window, and P value from paired t-tests between each of the five cortical regions: BA8, BA9, BA10, ECX, STG compared to the CER. Results are displayed in order of genomic position. RPKM and corresponding P values are shown for windows if P<0.05. Key: - denotes data not significant (P>0.05); Results shown in bold represent those found to be present across all prefrontal CTX areas (BA8, BA9, BA10).*

| **Start of 100bp** | **Gene(s)** | **BA8** | | **BA9** | | **BA10** | | **ECX** | | **STG** | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **P Value** | **Δ RPKM** | **P Value** | **Δ RPKM** | **P Value** | **Δ RPKM** | **P Value** | **Δ RPKM** | **P Value** | **Δ RPKM** |
| 701 | *MT-RNR1* | - | - | - | - | - | - | 0.032 | 79156 | - | **-** |
| 801 | *MT-RNR1* | 0.043 | 94246 | 0.029 | -43592 | - | - | - | **-** | 0.018 | 95257 |
| **2301** | ***MT-RNR2*** | **0.025** | **-97351** | **0.032** | **-29509** | **0.043** | **-82048** | - | - | - | - |
| 2401 | *MT-RNR2* | 8.90E-03 | 70950 | - | - | - | - | - | - | - | - |

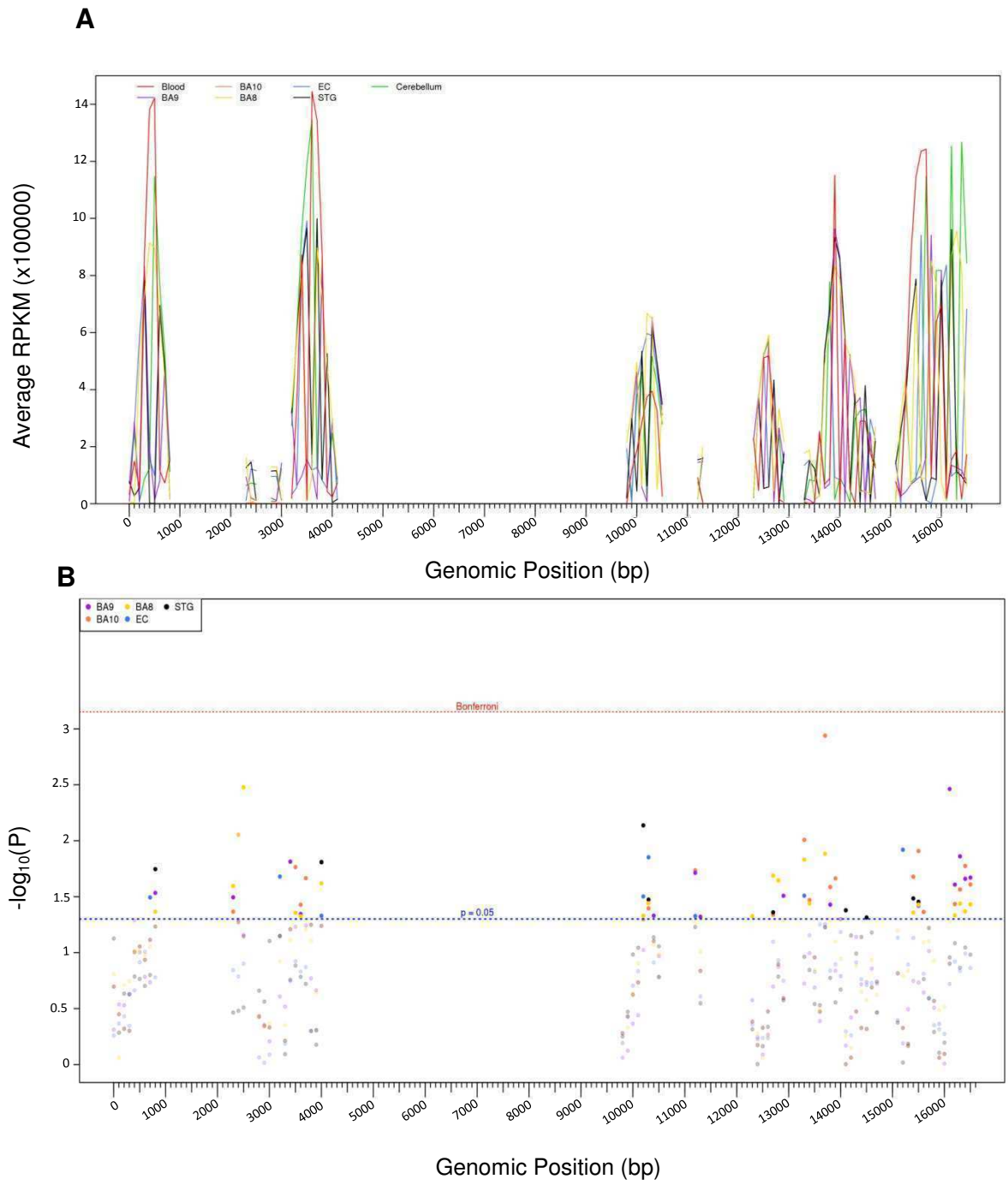
| Start of 100bp | Gene(s) | BA8 | | BA9 | | BA10 | | ECX | | STG | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | P Value | Δ RPKM | P Value | Δ RPKM | P Value | Δ RPKM | P Value | Δ RPKM | P Value | Δ RPKM |
| 2501 | *MT-RNR2* | 3.30E-03 | 51937 | - | - | - | - | - | - | - | - |
| 2901 | *MT-RNR2* | - | - | - | - | - | - | - | - | - | - |
| 3201 | *MT-RNR2 / MT-TL* | - | - | - | - | - | - | 0.021 | 242424 | - | - |
| 3401 | *MT-ND1* | - | - | 0.015 | 864685 | - | - | - | - | - | - |
| 3501 | *MT-ND1* | 0.041 | 1180914 | - | - | 0.017 | 226066 | - | - | - | - |
| **3601** | ***MT-ND1*** | **0.047** | **1250436** | **0.045** | **1228936** | **0.037** | **1183681** | - | - | - | - |
| 3701 | *MT-ND1* | - | - | - | - | 0.022 | 20818 | - | - | - | - |
| 4001 | *MT-ND1* | 0.024 | -39498 | 1.55E-02 | -38704 | - | - | 0.047 | -49753 | 0.016 | 243767 |
| 10201 | *MT-ND3* | 0.047 | -662126 | - | - | - | - | 0.031 | -591067 | 7.30E-03 | -57012 |

| Start of 100bp | Gene(s) | BA8 | | BA9 | | BA10 | | ECX | | STG | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | P Value | Δ RPKM | P Value | Δ RPKM | P Value | Δ RPKM | P Value | Δ RPKM | P Value | Δ RPKM |
| **10301** | ***MT-ND3 / MT-TR / MT-ND4L*** | **0.036** | **132719** | **0.036** | **135759** | **0.040** | **115364** | **0.014** | **73063** | **0.034** | **105106** |
| 10401 | *MT-ND4L* | - | - | 0.047 | -71246 | - | - | - | - | - | - |
| 10501 | *MT-ND4L* | - | - | - | - | - | - | - | - | - | - |
| 11201 | *MT-ND4* | - | - | 0.019 | -68903 | 0.018 | -79625 | 0.047 | 58742 | - | - |
| 11301 | *MT-ND4* | - | - | 0.048 | -139847 | - | - | - | - | - | - |
| 12301 | *MT-TL2 / MT-ND5* | 0.047 | 1787 | - | -197592 | - | - | - | - | - | - |
| 12701 | *MT-ND5* | 0.021 | -5702 | - | - | 0.046 | -388706 | - | - | 0.044 | -394978 |
| 12901 | *MT-ND5* | - | - | 0.031 | -130016 | - | - | - | - | - | - |

| Start of 100bp | Gene(s) | BA8 | | BA9 | | BA10 | | ECX | | STG | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | P Value | Δ RPKM | P Value | Δ RPKM | P Value | Δ RPKM | P Value | Δ RPKM | P Value | Δ RPKM |
| 13301 | *MT-ND5* | 0.015 | -175917 | - | - | 9.80E-03 | -144949 | 0.034 | -133694 | - | - |
| 13401 | *MT-ND5* | 0.036 | -104010 | - | - | 0.034 | -81264 | - | - | - | - |
| 13701 | *MT-ND5* | 0.013 | -422610 | - | - | 1.20E-03 | 3714 | - | - | - | - |
| 13801 | *MT-ND5* | - | - | 0.037 | 708761 | 0.026 | 123249 | - | - | - | - |
| 13901 | *MT-ND5* | - | - | - | - | 0.022 | -75118 | - | - | - | - |
| 14101 | *MT-ND5 / MT-ND6* | - | - | - | - | - | - | - | - | 0.042 | -534766 |
| 14501 | *MT-ND6* | - | - | - | - | - | - | - | - | 0.049 | -82767 |
| 15201 | *MT-CYB* | - | - | - | - | - | - | 0.012 | 254183 | - | - |
| 15301 | *MT-CYB* | - | - | - | - | - | - | - | - | - | - |

| Start of 100bp | Gene(s) | BA8 | | BA9 | | BA10 | | ECX | | STG | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | P Value | Δ RPKM | P Value | Δ RPKM | P Value | Δ RPKM | P Value | Δ RPKM | P Value | Δ RPKM |
| 15401 | *MT-CYB* | 0.044 | 10628 | - | - | 0.021 | -554527 | - | - | 0.033 | -554225 |
| 15501 | *MT-CYB* | 0.037 | -671707 | - | - | 0.012 | -685630 | 0.038 | 9301 | 0.035 | -695515 |
| 15601 | *MT-CYB* | - | - | - | - | 0.043 | -746270 | - | - | - | - |
| 16101 | *D-LOOP* | - | - | 3.40E-03 | 3014 | - | - | - | - | - | - |
| **16201** | ***D-LOOP*** | **0.046** | **380407** | **0.025** | **1117943** | **0.037** | **296890** | - | - | - | - |
| **16301** | ***D-LOOP*** | **0.037** | **-941720** | **0.014** | **-112453** | **0.027** | **-177146** | - | - | - | - |
| **16401** | ***D-LOOP*** | **0.043** | **444051** | **0.022** | **1151402** | **0.017** | **313156** | - | - | - | - |
| **16501** | ***D-LOOP*** | **0.037** | **784572** | **0.021** | **764994** | **0.025** | **836662** | - | - | - | - |

***Figure 3.2: DNA methylation differences are seen in the mitochondrial genome between brain regions and blood.*** *(**A**) Composite plot of mtDNA methylation for each brain region. Average RPKM for CER and each cortical region alongside matched blood samples were plotted per 100bp window. (**B**) For each 100bp, paired t-tests were performed between each cortical region and CER (bottom panel).*



### *3.4.3. A Number of Differentially Methylated Regions in MtDNA can be Observed between CTX and* CER

We were also interested to see whether total cortical tissue was significantly different to matched CER samples. Given the paired nature of the different anatomical regions of CTX, we used a multi-level model to compare total CTX to CER.

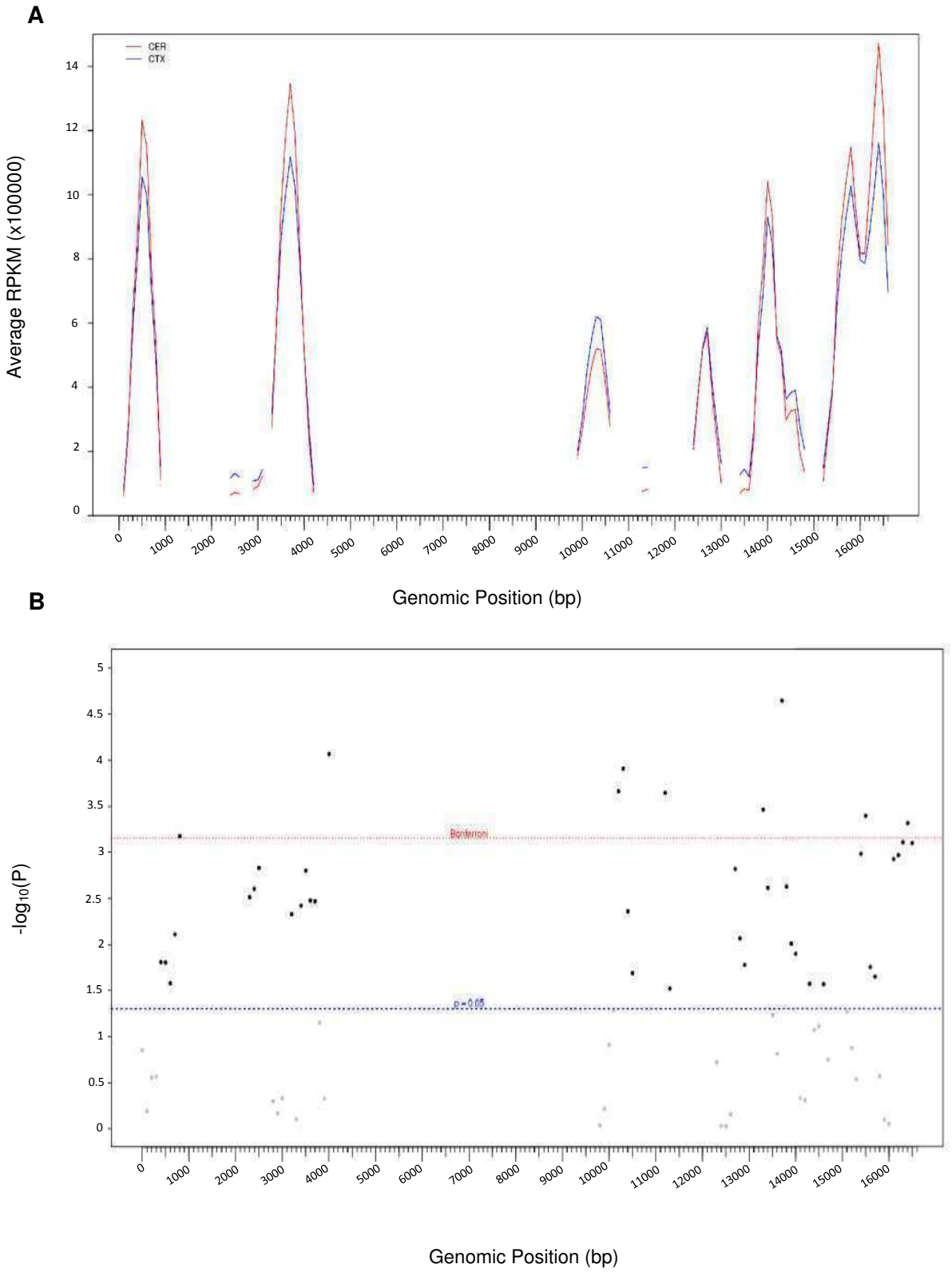
This analysis revealed 48 nominally significant (P<0.05) windows (**Table 3.3; Figure 3.3**), of which eight passed Bonferroni correction (**Table 3.3**, bold face). Interestingly, three of these eight were adjacent to each other, lying between 10301-10600bp and covering *Mitochondrially encoded NADH dehydrogenase 3 (MT-ND3) / Mitochondrially encoded NADH dehydrogenase 4 (MT-ND4L)* and *Mitochondrially encoded tRNA arginine* (*MT-TR*). We also saw a Bonferroni significant difference in DNA methylation in the *D-LOOP*, where we earlier noted DNA methylation changes across all three Brodmann area regions.

***Table 3.3: List of DMRs identified between total CTX and CER.*** *Shown is the location of the DMR within ChrM (based on GENCODE), the gene(s) residing within the 100bp window, and P value from a mixed effects model. Results are displayed in order of genomic position. RPKM and corresponding P values are shown for windows if P<0.05. Key; Bold denotes windows that reached our Bonferroni significant threshold of P<7.04E-04.*

| Start (bp) | Stop (bp) | Gene(s) | P-Value |
|---|---|---|---|
| 401 | 500 | *D-LOOP* | 7.99E-03 |
| 501 | 600 | *MT-TF* | 4.49E-03 |
| 601 | 700 | *MT-TF / MT-RNR1* | 1.16E-02 |
| 701 | 800 | *MT-RNR1* | 3.22E-03 |
| **801** | **900** | ***MT-RNR1*** | **1.91E-04** |
| 2301 | 2400 | *MT-RNR2* | 8.10E-03 |
| 2401 | 2500 | *MT-RNR2* | 7.50E-03 |
| 2501 | 2600 | *MT-RNR2* | 5.14E-03 |
| 3201 | 3300 | *MT-RNR2 / MT-TL* | 1.51E-03 |
| 3401 | 3500 | *MT-ND1* | 1.37E-02 |
| 3501 | 3600 | *MT-ND1* | 4.27E-03 |
| 3601 | 3700 | *MT-ND1* | 5.56E-03 |
| 3701 | 3800 | *MT-ND1* | 8.21E-03 |
| **4001** | **4100** | ***MT-ND1*** | **3.07E-06** |
| 10001 | 10100 | *MT-TG / MT-ND3* | 1.81E-02 |
| 10101 | 10200 | *MT-ND3* | 1.39E-02 |
| **10201** | **10300** | ***MT-ND3*** | **3.53E-04** |
| **10301** | **10400** | ***MT-ND3 / MT-TR / MT-ND4L*** | **1.19E-05** |
| **10401** | **10500** | ***MT-ND4L*** | **2.61E-04** |
| 10501 | 10600 | *MT-ND4L* | 9.05E-04 |
| **11201** | **11300** | ***MT-ND4*** | **8.86E-05** |
| 11301 | 11400 | *MT-ND4* | 1.65E-03 |
| 12701 | 12800 | *MT-ND5* | 1.81E-03 |
| 12801 | 12900 | *MT-ND5* | 1.05E-02 |
| 12901 | 13000 | *MT-ND5* | 8.23E-03 |
| 13301 | 13400 | *MT-ND5* | 1.44E-03 |
| 13401 | 13500 | *MT-ND5* | 9.13E-04 |
| 13501 | 13600 | *MT-ND5* | 1.61E-02 |
| 13601 | 13700 | *MT-ND5* | 3.89E-02 |
| **13701** | **13800** | ***MT-ND5*** | **2.77E-04** |

| Start (bp) | Stop (bp) | Gene(s) | P-Value |
|---|---|---|---|
| 13801 | 13900 | *MT-ND5* | 2.80E-03 |
| 13901 | 14000 | *MT-ND5* | 1.88E-02 |
| 14001 | 14100 | *MT-ND5* | 9.31E-03 |
| 14301 | 14400 | *MT-ND6* | 1.04E-02 |
| 14401 | 14500 | *MT-ND6* | 1.99E-02 |
| 14501 | 14600 | *MT-ND6* | 2.26E-02 |
| 14601 | 14700 | *MT-ND6 / MT-TE* | 3.82E-03 |
| 14701 | 14800 | *MT-TE / MT-CYB* | 1.92E-02 |
| 15101 | 15200 | *MT-CYB* | 3.30E-02 |
| 15401 | 15500 | *MT-CYB* | 8.52E-04 |
| 15501 | 15600 | *MT-CYB* | 7.43E-04 |
| 15601 | 15700 | *MT-CYB* | 1.16E-02 |
| 15701 | 15800 | *MT-CYB* | 2.24E-02 |
| 16101 | 16200 | *D-LOOP* | 2.23E-03 |
| **16201** | **16300** | ***D-LOOP*** | **5.02E-04** |
| 16301 | 16400 | *D-LOOP* | 1.84E-03 |
| 16401 | 16500 | *D-LOOP* | 1.12E-03 |
| 16501 | 16600 | *D-LOOP* | 2.20E-03 |

***Figure 3.3: DNA methylation differences are seen in the mitochondrial genome between the CTX and CER.*** *(**A**) Average RPKM values for CTX and CER at each 100bp window were plotted. (**B**) For each 100p window a mixed effects model was used to compare CTX to CER.*

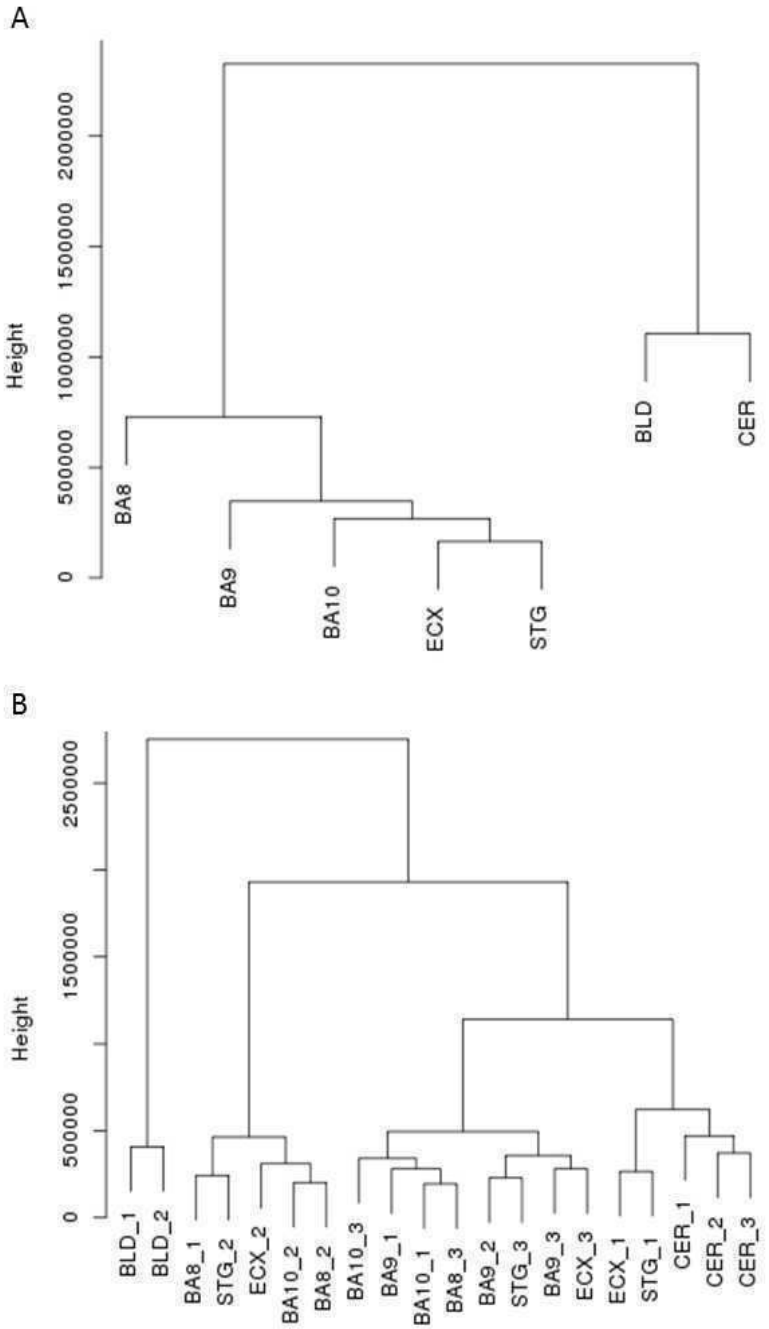


### *3.4.4. MtDNA Methylation Patterns Distinguish between Tissue Types*

Although we have shown that mtDNA methylation patterns are highly similar between distinct anatomical regions of the human brain and blood, we were also interested to identify whether mtDNA methylation patterns could distinguish between these tissue types. Through unsupervised hierarchical clustering we showed that average mtDNA methylation patterns can segregate these tissues (**Figure 3.4A)**.

Importantly, ncDNA methylation profiles in the same samples have also been previously shown to separate the CTX, CER and blood (158). Interestingly, when we performed unsupervised hierarchical clustering on the individual samples we found that, in most cases, intra-individual differences across tissue types are greater than inter-individual differences within each tissue-type, as the CTX, CER and blood samples clustered with their own tissue type respectively (**Figure 3.4B**).

***Figure 3.4: Hierarchical clustering of mtDNA methylation.*** *Unsupervised hierarchical clustering was performed on the average tissue methylation (**A**) and individual samples (**B**), which showed mtDNA methylation patterns could separate cortical, cerebellar and blood samples.*



## 3.5. Discussion

Publicly available epigenomic data provides a great resource for mitochondrial epigenetics, a field that is relatively nascent and has yet to be thoroughly investigated in a range of complex diseases. Here, we present evidence that mtDNA methylation patterns across mtDNA are brain region specific. However, data such as that presented here is confounded by a lack of isolation of mtDNA prior to antibody enrichment and sequencing. As such, the potential of including *NUMT*s in datasets derived from data generated using total genomic DNA could generate misleading results unless this is taken into account..

Here, we controlled for regions of high sequence homology between the nuclear and mitochondrial genomes. However, this approach is likely over-conservative and does lead to the generation of a somewhat truncated consensus sequence. MtDNA methylation has been shown to be cell line dependent in the past (63). Despite our stringent control for nuclear homology, a PCA of the mitochondrial epigenome was able to separate individuals belonging to the three main tissue types in this study: blood, CTX and CER, based on mtDNA methylation variation among tissue types. The tissue-specificity of this mark is further highlighted by the identification of eight DMRs that pass Bonferroni correction for multiple testing between total CTX and CER in our dataset.

Although overall DNA methylation levels were low in all tissues, it is worth noting that the study was performed on non BS-treated DNA. As such, the low percentage of mtDNA methylation is not a pitfall due to a lack of total BS-

treatment efficiency. One limitation of the current study is the lack of publicly available MeDIP-Seq datasets of matched cortical and CER tissue from other cohorts for validation purposes. Future work would aim to replicate our findings in additional study cohorts, and also to investigate the relationship between mitochondrial DNA methylation and gene expression.

Despite a number of nominally significant windows being identified between each individual cortical region and the CER, these did not pass either Bonferroni or Benjamani-Hochberg corrections for multiple testing. Nevertheless, the conservation of seven nominally significant windows across each Brodmann area is interesting to note. Four of these windows lie adjacent to each other and correspond to the mitochondrial *D-LOOP*, a region containing the only two mitochondrial promoters, which are typically associated with gene transcription and DNA replication.

One limitation of this study is owed to the use of antibody based enrichment, resulting in the analysis being limited to a window based approach. Despite this, studies of the nuclear genome have shown high correlation between window based approaches and, more sensitive, single site assays such as the Illumina 450K beadarray (100). However, given the small size of the mitochondrial genome, and that 23 of the 37 genes present in the genome are below 100bp in size, this window based approach may not be the most appropriate for future studies designed to specifically assess mtDNA methylation as it can result in a window intersecting two genes in the polycistronic transcript.

## 3.6. Conclusions

This method provides a conservative approach to determine mtDNA methylation across the genome for data previously generated using NGS approaches such as MeDIP-Seq. Its conservative nature reduces the risk of the inclusion of *NUMT*s in the final analysis of whole genome data, but may also lead to the inclusion of false-negatives as well as potential gaps in the reference sequence. As such, it is best suited to analysing previously generated whole genome data and is not a replacement for the isolation of mitochondrial DNA (225) prior to targeted methylation studies, which would be the optimal approach for investigating mitochondrial epigenomics. However, our method has allowed the identification of novel brain-region specific DMRs in a previously generated publicly available dataset. Furthermore, the identification of brain region specific mtDNA methylation patterns across the mitochondrial epigenome suggests the importance of a focussed, tissue-specific study design when investigating mtDNA methylation. As previously discussed, one caveat when utilising MeDIP-Seq data is the segregation of data into neighbouring windows, meaning that determining the exact corresponding gene of a DMR is difficult and, as such, future directed studies should aim to sequence the mitochondrial DNA methylome at single base resolution to address this.

# Chapter 4. A Comparison of Mitochondrial DNA Isolation Methods in Post-Mortem Archived Human Brain Tissue: Applications for Studies of Mitochondrial Genetics in Brain Disorders.

***The work presented in this chapter has been previously published (202) and can be found in Appendix 4.***

## 4.1. Aims:

1. To isolate mtDNA from frozen, post-mortem human brain tissue using methods commonly used to isolate mtDNA from fresh tissue and/or cultured cells.
2. To evaluate the effectiveness of these isolation methods using qRT-PCR, to compare the methods to each other and to total genomic DNA extracted using a phenol-chloroform method.

## 4.2. Introduction

Mitochondria generate ATP, regulate $Ca^{2+}$ homeostasis (178,179), mediate apoptosis (177), and produce ROS. Mitochondrial dysfunction has been implicated in a number of diseases, including in the pathogenesis of brain disorders such as AD (181,182,226). Mitochondria are unique mammalian organelles in that they contain their own genome; the mitochondrial genome is ~16.6kb of circular DNA (25), which is separate to ncDNA and inherited in a maternal, non-Mendelian fashion.

Because of its role in ROS production, mtDNA has a higher mutation rate than ncDNA (227). Mutations in mtDNA are relatively common, with at least one in

200 healthy humans harbouring a potentially pathogenic mtDNA mutation (228). Indeed more than 300 point mutations in mtDNA are associated with disease risk and pathology in MitoMAP (229). Interestingly, as each mitochondrion contains 2-10 copies of mtDNA and there are multiple mitochondria in any given cell, somatic mutations result in a mosaic of different mtDNA sequences within any given tissue. This phenomenon is known as mitochondrial heteroplasmy and is linked to various mitochondrial diseases (154). Such heterogeneity is a potential confounder in studies of mitochondrial diseases, because inter- and intra-individual heteroplasmic variation can confuse the association between a haplogroup and its corresponding phenotype. Therefore, unlike studies of ncDNA variation, it is important to use the specific tissue of interest for aetiological research.

Another interesting feature of the mitochondria is that over evolution, sequences of mtDNA have translocated to the nuclear genome. Traditional mitochondrial genetic research, and more recently studies of mitochondrial epigenetics, can be hampered by the presence of these *NUMTs* as they share a high homology with their mitochondrial paralogs (230,231). Given the interest in studying mtDNA genetic and epigenetic changes in the pathology of brain diseases characterised by mitochondrial dysfunction, it is imperative that *NUMT*s are correctly accounted for (1).

The specific isolation of mitochondria prior to downstream processing is vital to remove issues relating to *NUMT* contamination. For this purpose, a number of methods have been developed to specifically isolate mtDNA, although few of

these approaches have been specifically optimised for use on post-mortem, archived, frozen, human brain tissue, a major resource in many epidemiological studies. In fact, most studies investigating mtDNA use fresh animal tissue or cell lines. The insult of freezing tissue prior to isolation will potentially alter the effectiveness of these techniques and increase the risk of *NUMT* inclusion in downstream analysis.

In this study, we compared the effectiveness of five different mitochondrial isolation methods on human, post-mortem brain tissue using qRT-PCR to determine the optimal method for the specific enrichment of mtDNA. The methods tested have been previously successful in isolating pure mtDNA from fresh animal tissue/cell lines, and included protocols based on A) Percoll gradients, B) linear DNA digestion, C) differential centrifugation, D) rapid differential centrifugation using a commercial kit (Promokine) and E) magnetic isolation of mitochondria using anti-TOM22 antibodies (Miltenyi Biotech). We compared these approaches to DNA we had previously extracted (232), using a phenol-chloroform protocol, which isolates both ncDNA and mtDNA.

## 4.3. Methods

### *4.3.1. Method A*

Method A was a modification of the method by Sims and Anderson (233). Briefly, 150mg of tissue was dissociated using the gentleMACS dissociator (Miltenyi Biotech; Cat No: 130-093-235) and a mitochondrial extraction kit

(Miltenyi Biotech; Cat No: 130-097-340). After removal of the nuclear fraction, the supernatant was spun at 13,000 x g for 30 minutes at 4℃ to form a crude mitochondrial pellet. The pellet was homogenised in a 12% Percoll solution (Sigma Aldrich; Cat No: P1644-25ML) and added above two layers (26% and 40%) of Percoll solution. Samples were spun at 30,700 x g for 5 minutes at 4℃ with the lower band containing the enriched mitochondrial fraction. Each mitochondrial fraction was diluted in four volumes of isolation buffer (0.25M sucrose/10mM EDTA/30mM Tris-HCl, pH 7.5) and centrifuged at 16,700 x g for 10 minutes at 4℃ to form a loose mitochondrial fraction. The supernatant was discarded, and mtDNA extracted using a DNA Minikit (Qiagen; Cat No: 51304) (16).

*4.3.2. Method B*

Method B was based on the method by Zhou *et al* (146) that digests linear DNA but leaves circular DNA intact. In brief, 20µg genomic DNA (previously extracted using a phenol-chloroform protocol: see section 2.3.1) was treated with 4µl lambda exonuclease (5 U/µl) (New England Biolabs; Cat No: M0262S) and 12µl RecJf (30 U/µl) (New England Biolabs; Cat No: M0264S) in 400µl 1x lambda exonuclease buffer (New England Biolabs; Cat No: B0262S) at 37℃ for 16 hours. Samples were then incubated at 65℃ for 10 minutes to inactivate the enzymes and subsequently purified using a Qiagen DNA Minikit.

*4.3.3. Method C*

Method C was based on the method by Clayton and Shadel (234). Briefly, 100mg brain tissue was homogenized in 1ml chilled homogenisation buffer (0.25M sucrose/10mM EDTA/30mM Tris-HCl, pH 7.5). The homogenate was centrifuged at 1,000 x g for 15 minutes at 4℃ and the supernatant removed. The pellet was re-homogenised in 600µl chilled homogenisation buffer and spun at 1,000 x g for 10 minutes at 4℃. The supernatant was combined with the supernatant from the previous step and centrifuged at 12,000 x g for 30 minutes at 4℃ to pellet the mitochondria. MtDNA was extracted using a Qiagen DNA Minikit (18).

*4.3.4. Method D*

Method D was a modification of the method by Clayton and Shadel (234) and purchased as a commercial kit (Promokine; Cat No: PK-CA577-K280). Briefly, 100mg of brain tissue was homogenised with the reagents provided, as per the manufacturer's instructions (18).

*4.3.5. Method E*

The final isolation method used was adapted from a commercially available mitochondrial isolation kit and protocol from Miltenyi Biotec. Briefly, 200mg of tissue was dissociated using the gentleMACS dissociator (Miltenyi Biotech; Cat No: 130-093-235) and a mitochondrial extraction kit (Miltenyi Biotech; Cat No: 130-097-340) and centrifuged to pellet the nuclear contaminants. The supernatant was then removed and magnetically labelled anti-TOM22 antibody

beads (Miltenyi Biotech; Cat No: 130-094-532) were added. This mixture was then incubated at 4℃ for 1 hour, after which the mixture was processed through a magnetic column. Magnetically bound mitochondria were then removed from the column and mtDNA was extracted using a Qiagen DNA Minikit. Full details on this protocol can be found in section 2.3.4.

### *4.3.6. Method F*

Total DNA was extracted from human brain tissue using a standard phenol-chloroform extraction protocol routinely used in our group. This method isolates both ncDNA and mtDNA and this DNA was thus used to determine the average mtDNA/ncDNA ratio from non-enriched samples. Full protocol details can be found in section 2.3.1.

### *4.3.7. QRT-PCR of MtDNA Copy Number*

QRT-PCR was used to quantify the amount of mtDNA relative to ncDNA using primers specific to a mitochondrial gene (with no known *NUMTs*), and a nuclear housekeeping gene. Details on the assay and experimental conditions can be found in section 2.3.5.

### *4.3.8. Sequencing of MtDNA*

DNA samples were fragmented by sonication using a Bioruptor (Diagenode; Cat No: UCD-200) to an average size of ~240bp. Sequencing libraries were prepared using the NEXTflex Rapid DNA-Seq kit (Bioo Scientific; Cat No: 5144-

01) and ligated to pre-indexed adapters (NEXTflex-96 DNA Barcodes; Bioo Scientific; Cat No: 514101). Adapter-ligated DNA was amplified for 10 cycles using Herculase II Fusion DNA Polymerase (Agilent Technologies; Cat No: 600675) and NEXTflex PCR primer mix, then pooled for sequencing on an Illumina HiSeq2500 (100bp paired-end, rapid run mode). Raw reads were quality and adaptor trimmed using TrimGalore! (235) before being aligned to GRCH37. Only high quality (Phred<20) reads, uniquely mapping to the genome, were considered and total read counts were taken. The mapping percentages to this reference genome and the total number of reads mapped for enriched mitochondrial samples and the non-enriched standard were then compared.

## 4.4. Results

### *4.4.1. Method E Significantly Enriches for MtDNA*

Our data showed that method B (linear DNA digestion) gave the lowest purity (1,242 mtDNA copies/ncDNA copy), and method E (magnetic-microbeads) the highest purity (14,654 mtDNA copies/ncDNA copy) (**Figure 4.1; Table 4.1**).

***Figure 4.1: Enrichment of mtDNA relative to ncDNA.*** *In total five methods (Percoll (A), DNase digestion (B), differential centrifugation (C), rapid differential centrifugation (D) and magnetic microbeads (E) were compared to a non-enriched standard (phenol-chloroform (F)). Shown is the ratio of mtDNA/ncDNA (±SEM).*

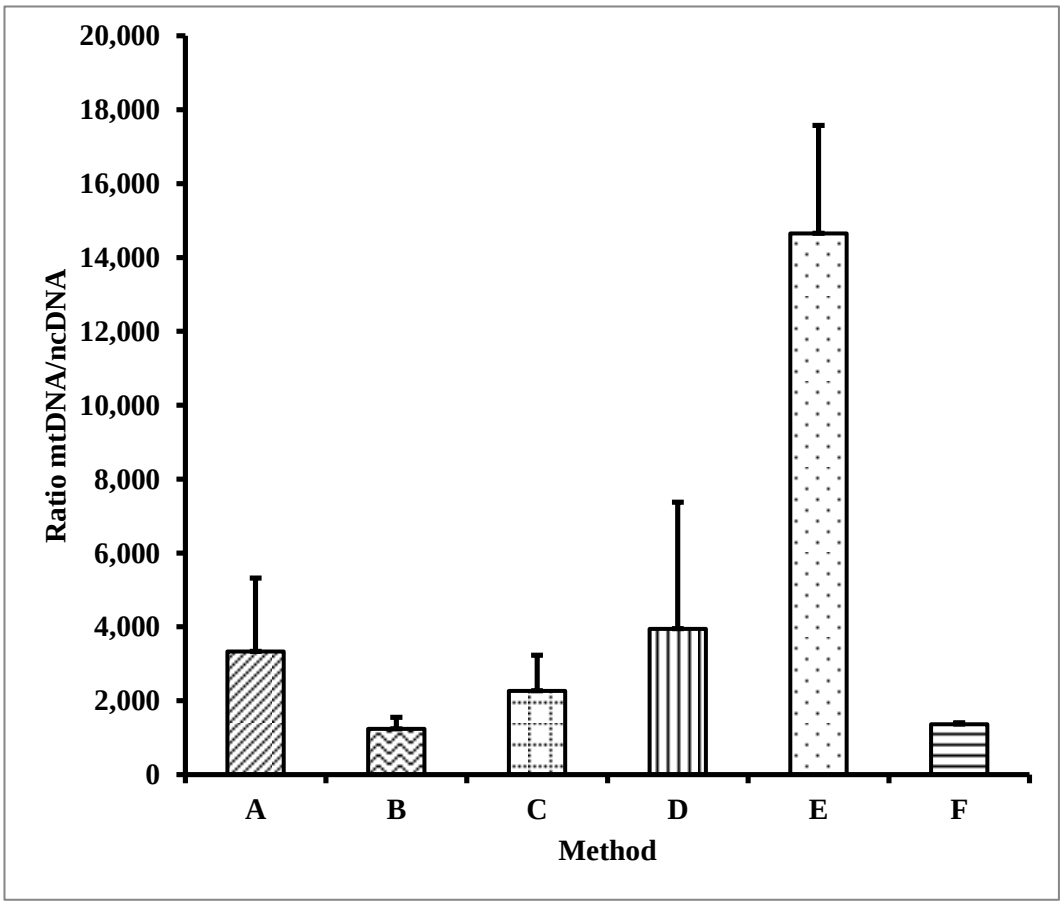


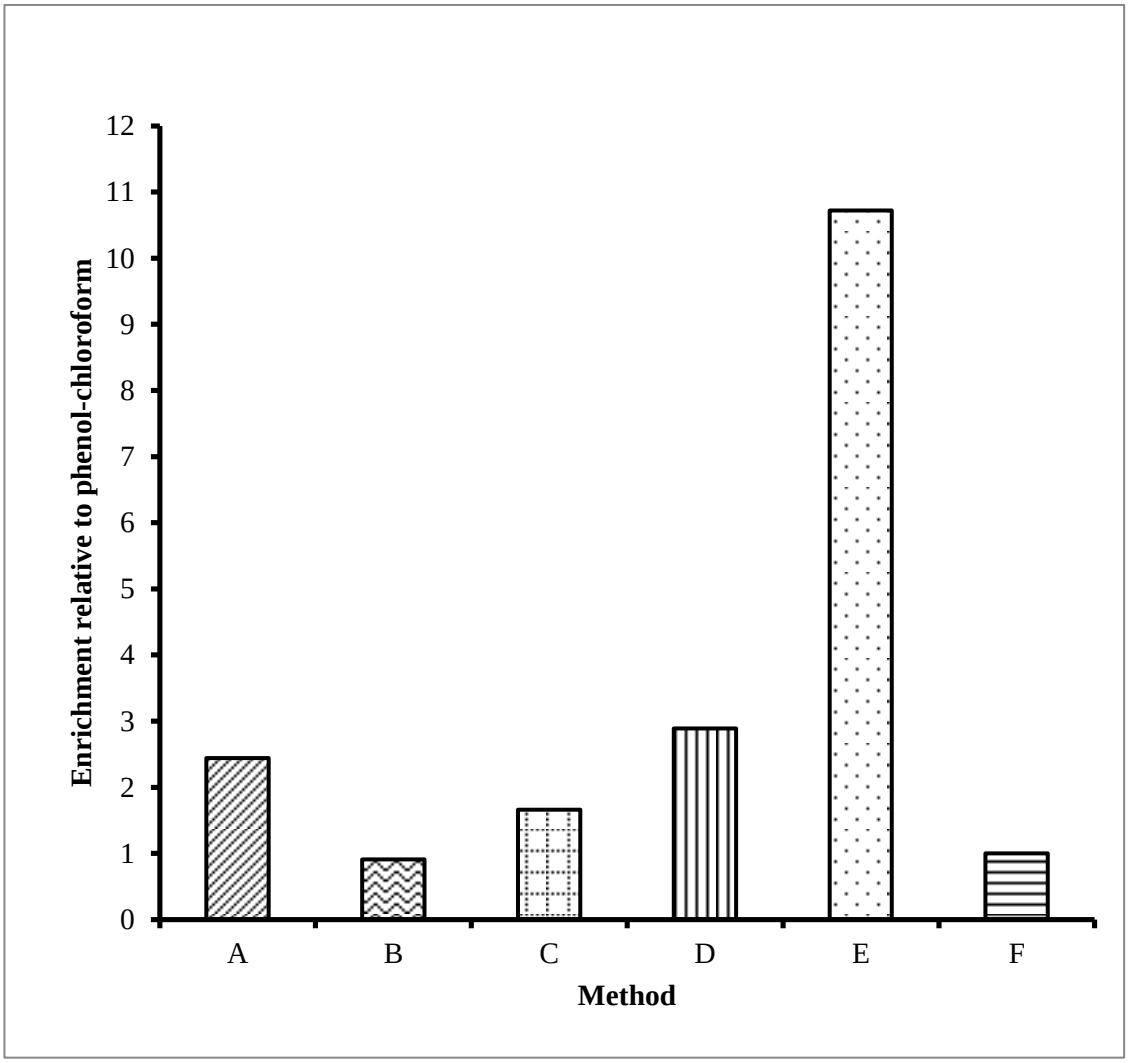
***Table 4.1: An overview of starting material for each isolation technique and resulting yield.*** *In brief we compared five methods of isolating mtDNA in post-mortem human brain tissue; (A) discontinuous Percoll gradient, (B) DNase digestion of linear DNA (RecJ$_f$), (C) differential centrifugation (DC), (D) rapid (commercial) mitochondrial isolation via differential centrifugation (provided by Promokine) (E) magnetic labelling and pull-down of mitochondria using an antibody to TOM22 (MACS kit provided by Miltenyi Biotec). We compared the yield and enrichment to a non-enriched standard (phenol-chloroform) using an unpaired two-tailed t-test.*

| **Method** | **A** | **B** | **C** | **D** | **E** | **F** |
|---|---|---|---|---|---|---|
| Quantity of starting Material | 150mg (tissue) | 20µg (DNA) | 100mg (tissue) | 100mg (tissue) | 200mg (tissue) | 100mg (tissue) |
| Average concentration (ng/µl) DNA collected (±SEM) | 162.4 (11.6) | 40.3 (10.1) | 23.7 (8.0) | 257 (138.4) | 32.4 (7.0) | N/A |
| Average yield (µg) DNA collected (±SEM) | 8.1 (0.58) | 2.0 (0.50) | 4.7 (1.60) | 13.7 (6.35) | 3.2 (0.70) | N/A |

| Method | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| Average copies of mtDNA (±SEM) | 41,530 (15,468) | 12,092,501 (7,742,804) | 28,886,594 (12,405,939) | 403,381 (194,557) | 5,321,960 (1,246,724) | 459,264 (751,075) |
| Number of samples | 4 | 6 | 4 | 3 | 5 | 55 |
| Average copies of ncDNA (±SEM) | 155 (85) | 8,768 (3,731) | 20,746 (11,256) | 720 (449) | 402 (86) | 4,011 (585) |
| Average ratio mtDNA/ncDNA (±SEM) | 3,337 (1,988) | 1,242 (309) | 2,270 (960) | 3,949 (3,424) | 14,654 (2,922) | 1,367 (35) |
| Fold Enrichment mtDNA/ncDNA relative to phenol-chloroform (P-value) | 2.44 (0.297) | 0.91 (0.725) | 1.66 (0.320) | 2.89 (0.342) | 10.72 ($1.88x10^{-3}$) | N/A |

Of particular interest to the study was the relative enrichment compared to a standard phenol-chloroform extraction (**Figure 4.2**).

***Figure 4.2: Enrichment of mtDNA relative to phenol: chloroform.*** *In total five methods (Percoll (A), DNase digestion (B), differential centrifugation (C), rapid differential centrifugation (D) and magnetic microbeads (E) were compared to a non-enriched standard (phenol-chloroform (F)).*



The only method to show no enrichment was Method B (linear DNA digestion). All other techniques showed a positive enrichment of mtDNA compared to Method F (phenol: chloroform). Method A (Percoll), Method C (differential centrifugation) and Method D (rapid differential centrifugation) all gave modest positive enrichments of 2.4-, 1.7- and 2.9- fold respectively. Although giving one of the lowest yields (3.2µg), the optimal method for enrichment relative to ncDNA was Method E (magnetic-microbeads), which gave a 10.7-fold enrichment, and was the only method to show significantly more copies of mtDNA/ncDNA copy compared to phenol: chloroform extraction (P=1.88E-03). Using this method we saw a significant enrichment of mtDNA/ncDNA compared to Methods A (P=0.019), B ($P=6.97x10^{-4}$) and C ($P=8.48x10^{-3}$).

### *4.4.2. Validation using NGS Technologies*

To validate our enrichment, two of the biological replicates from Method E were compared to a non-enriched standard from phenol: chloroform (Method F) extraction using NGS. For the non-enriched standard 1.1% of reads mapped to the mitochondrial genome, compared to an average of 18.7% (16.2% and 21.2% respectively) of reads with Method E, demonstrating an average 16.8 fold enrichment in mtDNA.

## 4.5. Discussion

A previous *in vitro* study by Hornig-Do *et al* has shown that magnetic-microbeads gives greater yields of mitochondria compared to both Percoll and differential centrifugation in an osteosarcoma cell line (236). However, although

increased yield of mitochondrial organelles was demonstrated, they showed the relative purity of mtDNA compared to ncDNA was similar between Percoll and magnetic-microbeads. In the context of genomic studies of mtDNA, where the exclusion of *NUMTs* is imperative, the relative enrichment of mtDNA is of far greater importance than the yield. Thus, although we saw a lower yield with magnetic-microbeads compared to the majority of methods tested, Method E gave the greatest purity. The reasons for the observed greater enrichment of mtDNA/ncDNA relative to Method A in our study compared to the analysis by Hornig-Do and colleagues potentially may include i) our use of qRT-PCR, which is more sensitive than western blot and ii) our use of frozen, archived samples, rather than fresh samples, which may lead to more fragile mitochondria that could degrade at higher centrifugation speeds.

## 4.6. Conclusion

To our knowledge this study represents the first to systematically compare and contrast methods for isolating mtDNA from small quantities of frozen, archived, post-mortem human brain. Our findings suggest that magnetic-microbeads provide a significant enrichment of mtDNA compared to any other method tested. This may be due to a number of reasons, for example the automated homogenisation of tissue in this protocol could provide a more consistent and gentle approach than other techniques and the use of magnetically labelled antibodies provides a specific capture of intact mitochondria, which may also contain less degraded mtDNA. We recommend that given the current interest in studying the mitochondrial genome in human brain, that the magnetic-

microbead method from Miltenyi Biotech is used prior to DNA extraction to minimise the inclusion of NUMTs in downstream analyses.

# Chapter 5. A Genome-wide Interrogation of MtDNA Methylation via NGS

## 5.1. Aims

1. To develop a methodology capable of interrogating the mitochondrial epigenome at single cytosine resolution.

2. To employ this methodology to investigate the associations between mtDNA methylation and each of age, sex and tissue type.

## 5.2. Introduction

Mitochondria are unique organelles in that they have their own genome, ~16.6kb in size. MtDNA consists of 37 genes, 22 encoding for tRNAs, two for rRNAs and 13 encoding for proteins in the ETC. The ETC consists of 97 genes; in addition to the 13 mitochondrial-encoded genes, an additional 84 are encoded by the nuclear genome and imported into the mitochondrion (237). ETC proteins are directly involved in the regulation of cellular respiration, generating the majority of ATP required for the process. However, mitochondria play a vital role in a variety of key biological functions, including apoptosis, as reviewed in (238), the regulation of $Ca^{2+}$ homeostasis (239,240) and the production of ROS (241).

Epigenetic processes mediate the reversible regulation of gene expression, occurring independently of DNA sequence variation and acting principally through chemical modifications to DNA and nucleosomal histone proteins. Of all the epigenetic marks, DNA methylation is the most stable and well-studied (4). The advent of cost-effective epigenome-wide technologies, such as the Illumina 450K array and its successor, the Illumina Infinium EPIC Beadchip Array, have

led to the identification of differentially methylated genes in a range of complex diseases, including AD (95,216,242,243), schizophrenia (244), cancer (245,246) and type I diabetes (247). However, one caveat of this approach is the complete absence of coverage of the mitochondrial genome. Recent studies have suggested a role for differential mtDNA methylation in numerous pathologies characterised by mitochondrial dysfunction, including cancer (248), amyotrophic lateral sclerosis (249), and AD (1,250). However, until recently studies have been limited to low-resolution immunohistochemical techniques (72) or candidate based approaches such as BS-pyrosequencing (251). To date, the only studies to investigate epigenome-wide mtDNA methylation have utilized publicly available MeDIP-Seq datasets (63,64,252). However, these studies are limited to a semi-quantitative calling of regional methylation in windows. Furthermore, the presence of *NUMT*s in the nuclear genome leads to the requirement for the removal of homologous regions from the analysis (63,64) and, as such, only a truncated sequence can be analysed accurately.

We have previously shown in Chapter 4 that a commercially available method from Miltenyi Biotec, which magnetically isolates mitochondria using anti-TOM22 antibodies, gives a significant enrichment of mtDNA to ncDNA (202). Here, we combine this method with a customised, targeted bisulphite sequencing approach to allow us to specifically investigate mtDNA methylation patterns in post-mortem brain samples at single nucleotide resolution along the entire 16.569kb genome.

## 5.3. Methods

### *5.3.1. Sample Demographics for Chapter 5*

200mg sections of brain tissue were obtained from seven donors archived in the LBB (http://www.kcl.ac.uk/iop/depts/cn/research/MRC-London-Neurodegenerative-Diseases-Brain-Bank/MRC-London-Neurodegenerative-Diseases-Brain-Bank.aspx). From each donor we sourced matched superior STG and CER brain tissue. The donors were selected based upon a similar age, distribution of sex across the group, low post-mortem intervals and no evidence of neurodegenerative disease. Sample demographics are provided in **Table 5.1.**

***Table 5.1: Demographic information for samples used in Chapter 5****. Brain tissue was collected from seven donors archived in the LBB. For each donor two brain regions, the STG and CER were collected. Samples selected for the study had a low-post-mortem interval (compared to other samples available), no evidence of neurodegenerative disease and a similar age and distribution of sex across the group.*

| Individual | Age at Death (Years) | Sex | Post-Mortem Interval (hours) |
|---|---|---|---|
| 1 | 82 | Female | 43 |
| 2 | 81 | Female | 17 |
| 3 | 80 | Female | 3 |
| 4 | 86 | Male | 6 |
| 5 | 87 | Female | 22 |
| 6 | 81 | Male | 18 |
| 7 | 77 | Male | 11 |

### *5.3.2. Mitochondrial Isolation*

Mitochondria were isolated from frozen, post-mortem brain tissue according to the protocol outlined in Section 2.3.4. MtDNA was further enriched through the use of a custom library designed to amplify the mitochondrial genome (Agilent Technologies, California, U.S.A.) as described in Section 2.3.6. To test for the effect of enrichment on mtDNA methylation levels, matched nuclear isolates from both STG and CER were kept and sequenced for individual 4. MtDNA methylation levels in nuclear isolates were then compared to mtDNA methylation levels of pre-enriched mitochondria for this individual.

### *5.3.3. Custom Capture of the Mitochondrial Epigenome*

To capture the mitochondrial epigenome, a custom library of RNA baits (Agilent, California, U.S.A.: Design ID 0687721) were designed to provide 100% coverage of the genome at 5x tiling density. Isolated mtDNA extracted from frozen, archived brain tissue was subjected to an adapted version of the Agilent 1µg Methyl-Seq protocol (described in detail in Section 2.3.6). After generating indexed libraries for each sample, samples were pooled in equimolar concentrations and sequenced using the Illumina HiSeq 2500 by Exeter Sequencing Service (Exeter, U.K.) to generate 16 fastq files, processed in house.

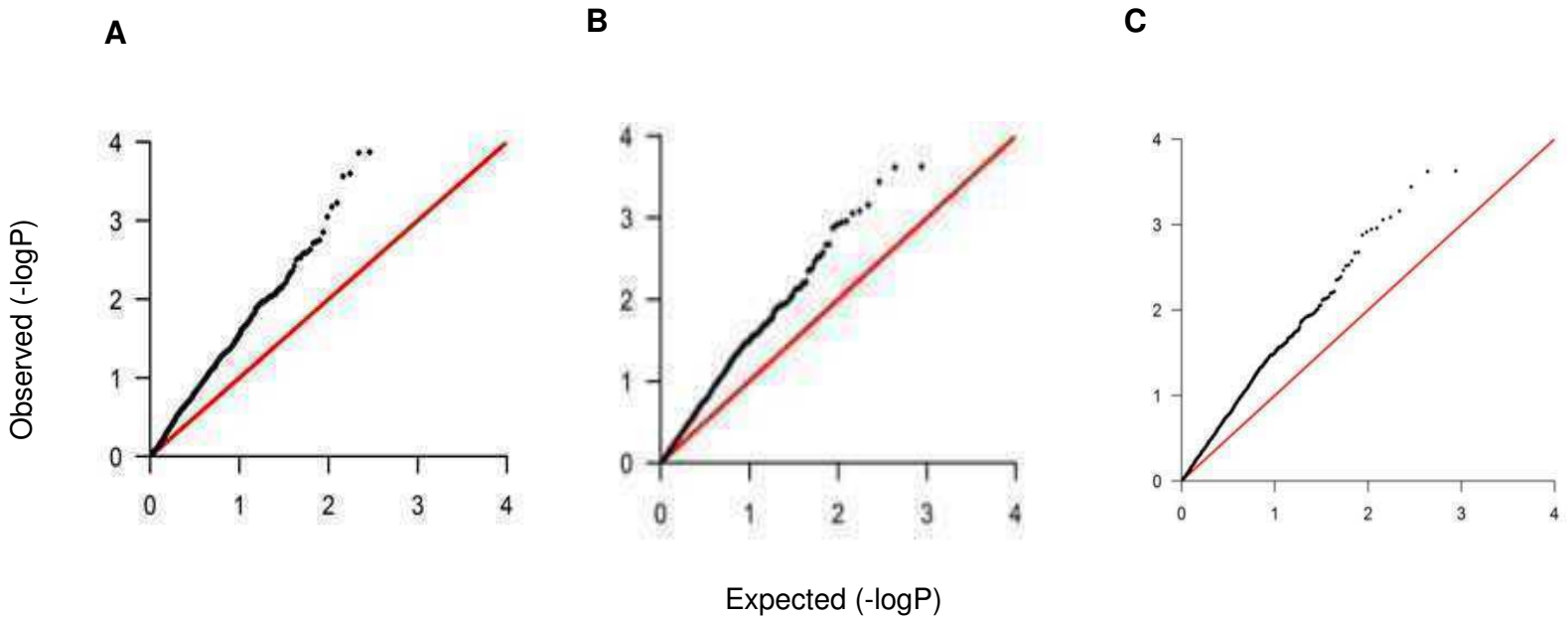
### *5.3.4. Analytical Workflow*

After sequencing, 100bp paired-end reads were de-multiplexed and, following quality assessment using FastQC, were trimmed for adaptor content using

TrimGalore (253). Using the same package, we also trimmed the first base-pair from the 3' end of both reads. Trimmed files were then aligned to the reference human mitochondrial genome, rCRS, through Bismark, using Bowtie 1 (204). Mapped reads were then de-duplicated using the deduplicate_bismark function, before CpG and non-CpG methylation was called using bismark_methylation_extractor function. All statistical analyses were carried out in R (v3.3.0) (254). Our initial analysis aimed to control for the matched nature of our different tissue samples by running a mixed effect model.

Briefly, this model used each of the covariates of interest (age, sex and tissue type) as fixed effects and the individual as the random effect. To determine the significance and effect size of each covariate in isolation at each methylated cytosine identified, a second (null) model was created, which didn't contain the covariate of interest. An analysis of variance (ANOVA) was then performed between the two models to identify significantly differentially methylated bases for each covariate. We also performed post-hoc analyses to determine the effect of sex and age within each brain region independently. This analysis consisted of two linear regressions, one for each tissue. Quantile-quantile (Q-Q) plots were used to check for P-value inflation in all models used for sex (**Figure 5.1.A**), age (**Figure 5.1.B.**) and for tissue (**Figure 5.1.C.**). Throughout all analyses, nominally significant P values were defined as $P<0.05$. A Bonferroni correction value of 5.734E-05 was determined. Q values were also determined using a Benjamani-Hochberg correction for multiple testing. To test for enrichment of trends found within our cohort, a Fisher's exact test was carried out using an in-built function within the R environment (254),

***Figure 5.1: Composite quantile-quantile (QQ) plots for P Value dispersion.*** *The expected P values are plotted against those observed in the data to check for P Value inflation in our multi-model level analysis. This analysis was performed for (A) sex, (B) age and (C) tissue type.*



## 5.4. Results

### *5.4.1. MtDNA Methylation Levels Differ Depending upon Level of Enrichment Prior to Sequencing*

On average, 723,765 reads mapped to the mitochondrial reference genome per sample, providing an average sequencing depth of 3019 times. To test whether methylation levels differ between enriched and non-enriched mtDNA, nuclear and mitochondrial isolates derived from matched STG and CER samples from one individual were processed at the same time and sequenced. On average, 58% of reads mapped to the mitochondrial genome, suggesting a large enrichment for mtDNA, exceeding levels previously reported through the use of antibody-based enrichment alone (202).

In total, 870 variably methylated cytosines sites, sites where varying calls of methylation are detected upon sequencing, were captured for all samples. To determine the effect of enrichment, the standard error for each site was calculated within each enriched sample. DNA methylation levels were then compared between the matched enriched and non-enriched samples. Sites with differences greater than the standard error for that tissue type were said to be enriched. In total, only 167 of 870 sites were unaffected by the enrichment for STG. However, 679 of 870 probes were unaffected by the enrichment process in CER.

*5.4.2. MtDNA Methylation Occurs with Distinct Prevalence at non-CpG sites*

NcDNA methylation occurs with a high prevalence in CpG contexts in mammalian cell lines and human brain tissue (255). However, a previous study has indicated a potential prevalence for non-CpG methylation in mtDNA (256), particularly in the *D-LOOP*. To investigate this further, we found that mtDNA methylation occurs at low levels and with a striking prevalence at non-CpG sites, with an average of 81% (S.D = 0.74) of all mtDNA methylation sites being found in a non-CpG context across the samples in our study (**Table 5.2**).

***Table 5.2: Table highlighting sequencing quality of samples.*** *For each sample, the mapping efficiency, total number of reads mapped, sequencing depth and percentage of non-CpG methylated cytosines were calculated. High levels of non-CpG methylated cytosines were found in all samples. NC corresponds to samples taken from the nuclear pellet of these isolations prior to mitochondrial isolation. Averages were only taken from mitochondrial isolates.*

| Sample | Mapping Efficiency (%) | Total Reads Mapped After Deduplication | Sequencing Depth | Total Methylated Cytosines in Non-CpG Context (%) |
|---|---|---|---|---|
| 1-STG | 44.60% | 364292 | 2198.64 | 80.68 |
| 1-CER | 48.20% | 364292 | 2198.64 | 80.91 |
| 2-STG | 60.10% | 250010 | 1508.90 | 81.62 |
| 2-CER | 53.80% | 1258516 | 7595.61 | 81.94 |
| 3-STG | 56.50% | 395091 | 2384.52 | 81.64 |
| 3-CER | 71.30% | 351541 | 2121.68 | 80.38 |
| 4-STG | 84.60% | 728332 | 4395.75 | 80.63 |
| 4NC-STG | 82.90% | 687317 | 4148.21 | 79.71 |
| 4-CER | 83.90% | 945709 | 5707.70 | 80.24 |
| 4NC-CER | 72.90% | 1085810 | 6553.26 | 79.19 |
| 5A-STG | 35.20% | 356999 | 2154.62 | 81.09 |
| 5B-CER | 27.70% | 191466 | 1155.57 | 81.64 |
| 6A-STG | 74.40% | 697378 | 4208.93 | 81.12 |
| 6B-CER | 74.50% | 516986 | 3120.20 | 81.66 |
| 7A-STG | 57.40% | 369870 | 2232.30 | 81.21 |
| 7B-CER | 37.00% | 213705 | 1289.79 | 80.88 |
| Average STG | 58.97% | 451710 | 2726.24 | 81.14 |
| Average CER | 56.63% | 548888 | 3312.74 | 81.09 |

*5.4.3. MtDNA Methylation Exhibits Sex-Specific Patterns Across the Genome*

Given that ncDNA methylation patterns can be strongly associated with sex (257), we were interested to see whether the mtDNA methylome exhibited sex-specific patterns. Our initial multi-level model identified 131 nominally significant differentially methylated positions (DMPs). Two sites, at 8722bp (*MT-ATP6*; $P=2.05x10^{-06}$) and 9055bp (*MT-ATP6;* $P=1.95x10^{-5}$) passed the strict Bonferroni correction for multiple testing **(Figure 5.2)**. In total five sites passed a less strict Benjamani-Hochberg correction for multiple testing (**Table 5.3**): 8722bp (*MT-ATP6*; Q=6.562E-03), 9055bp (*MT-ATP6*; Q=0.0297), 7019bp (*MT-CO1*; Q=0.0297), 499bp (*D-LOOP*; Q=0.0297) and 14559bp (*MT-ND6;* Q=0.0395).

To test whether one particular tissue was more responsible for driving these changes we ran two linear regression models, one in each tissue type, with age and sex as covariates. This allowed for an estimation of the effect of both covariates on each position. Our analysis identified a total of 97 nominally significant DMPs ($P<0.05$) for sex in the STG and 32 in the CER (**Table 5.3**). Of these, five were found to be significant in both tissues and shared the same direction of methylation between tissue types; 499bp (*D-LOOP*) and 9055bp (*Mitochondrially Encoded ATP Synthase 6; MT-ATP6*) were found to be hypermethylated in females whilst 8722bp (*MT-ATP6*), 14,269bp (*Mitochondrially Encoded NADH:Ubiquinone Oxidoreductase Core Subunit*; *MT-ND6*) and 15,761bp *(MT-CYB*) were found to be hypermethylated in males.

Further, our analysis revealed a tendency towards hypermethylation in females in both tissues. In total, 23 of the 33 and 89 of the 97 nominally significant differentially methylated positions (DMPs) were hypermethylated in CER and STG females respectively. To determine whether the direction of methylation change at these sites represented a significant enrichment a Fisher's exact test was performed on each brain region, demonstrating a significant enrichment in the STG (P=0.003; Odds Ratio=2.743) and a trend for enrichment in the CER (P=0.137; Odds Ratio=1.623).

***Figure 5.2: Composite boxplots to show sex-associated DMPs identified in both CER and STG, which survived Bonferroni correction in our mixed model, 8722bp (left) and 9055bp (right).***

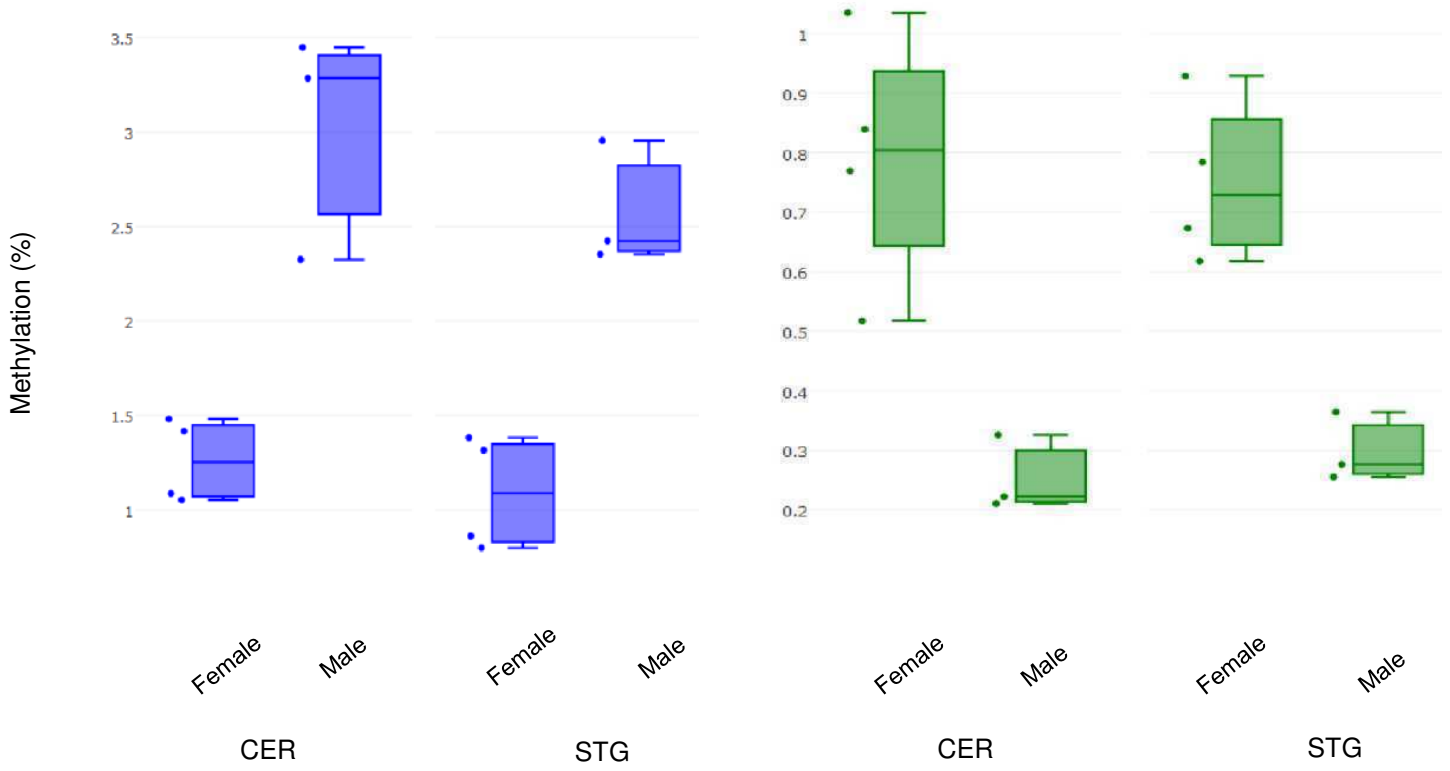


***Table 5.3: Nominally significant DMPs associated with sex****. In total 131 DMPs were nominally associated with sex in our multi-level model, 97 in STG and 32 in CER. Hypermethylation in females is indicated by a negative effect size value. Abbreviations: C1 (Complex I), C2 (Complex II), C3 (Complex III), C4 (Complex IV), C5 (Complex V), NCR (non-coding region), rRNA (ribosomal RNA), tRNA (transfer RNA), STG (superior temporal gyrus), CER (CER), MM (mixed model), N.S. (Not significant). Bold indicates DMPs present in both tissues.*

| Genomic Position (bp) | Gene | Complex | Effect Size MM | Significance Level (P) MM | Effect Size CER | Significance Level (P) CER | Effect Size STG | Significance Level (P) STG |
|---|---|---|---|---|---|---|---|---|
| 62 | *D-LOOP* | NCR | N.S | N.S | N.S | N.S | 0.147 | 4.580E-03 |
| 316 | *D-LOOP* | NCR | -1.003 | 7.478E-03 | N.S | N.S | N.S | N.S |
| 498 | *D-LOOP* | NCR | -7.345 | 0.010 | N.S | N.S | N.S | N.S |
| 499 | ***D-LOOP*** | **NCR** | 1.307 | 1.370E-04 | **1.098** | **0.011** | **1.516** | **0.011** |
| 526 | *D-LOOP* | NCR | N.S | N.S | N.S | N.S | -0.559 | 0.040 |
| 545 | *D-LOOP* | NCR | -3.002 | 2.531E-03 | -3.133 | 0.011 | N.S | N.S |
| 739 | *MT-RNR1* | rRNA | N.S | N.S | N.S | N.S | -0.408 | 0.031 |
| 785 | *MT-RNR1* | rRNA | N.S | N.S | 0.936 | 0.012 | N.S | N.S |
| 809 | *MT-RNR1* | rRNA | N.S | N.S | N.S | N.S | -0.446 | 0.039 |
| 900 | *MT-RNR1* | rRNA | N.S | N.S | N.S | N.S | -0.489 | 0.020 |
| 932 | *MT-RNR1* | rRNA | N.S | N.S | N.S | N.S | -0.901 | 0.025 |
| 1023 | *MT-RNR1* | rRNA | -0.926 | 0.011 | N.S | N.S | N.S | N.S |
| 1024 | *MT-RNR1* | rRNA | -0.251 | 5.312E-03 | N.S | N.S | -0.446 | 0.028 |
| 1146 | *MT-RNR1* | rRNA | N.S | N.S | N.S | N.S | -0.594 | 0.039 |
| 1177 | *MT-RNR1* | rRNA | -0.645 | 9.298E-03 | N.S | N.S | -1.217 | 6.000E-03 |
| 1217 | *MT-RNR1* | rRNA | -0.214 | 0.033 | -0.254 | 0.030 | N.S | N.S |
| 1262 | *MT-RNR1* | rRNA | -1.009 | 1.406E-03 | N.S | N.S | -1.484 | 0.025 |

| Genomic Position (bp) | Gene | Complex | Effect Size MM | Significance Level (P) MM | Effect Size CER | Significance Level (P) CER | Effect Size STG | Significance Level (P) STG |
|---|---|---|---|---|---|---|---|---|
| 1389 | *MT-RNR1* | rRNA | -0.105 | 0.049 | -0.165 | 0.047 | N.S | N.S |
| 1473 | *MT-RNR1* | rRNA | -0.877 | 2.567E-03 | -1.103 | 0.026 | N.S | N.S |
| 1475 | *MT-RNR1* | rRNA | N.S | N.S | -0.523 | 0.045 | N.S | N.S |
| 1489 | *MT-RNR1* | rRNA | -0.179 | 0.028 | N.S | N.S | N.S | N.S |
| 1561 | *MT-RNR1* | rRNA | -0.784 | 9.879E-03 | N.S | N.S | -1.042 | 0.020 |
| 1916 | *Mt-RNR2* | rRNA | -0.192 | 8.703E-03 | -0.174 | 0.026 | N.S | N.S |
| 2566 | *Mt-RNR2* | rRNA | 0.342 | 0.049 | N.S | N.S | 0.651 | 0.044 |
| 2643 | *Mt-RNR2* | rRNA | N.S | N.S | N.S | N.S | -0.608 | 0.023 |
| 2491 | *Mt-RNR2* | rRNA | -1.040 | 0.016 | N.S | N.S | N.S | N.S |
| 2809 | *Mt-RNR2* | rRNA | 0.640 | 0.021 | N.S | N.S | 0.894 | 0.038 |
| 2810 | *Mt-RNR2* | rRNA | -0.260 | 0.040 | N.S | N.S | N.S | N.S |
| 2878 | *Mt-RNR2* | rRNA | -0.206 | 0.010 | N.S | N.S | N.S | N.S |
| 2943 | *Mt-RNR2* | rRNA | -0.258 | 0.022 | N.S | N.S | N.S | N.S |
| 3035 | *Mt-RNR2* | rRNA | N.S | N.S | N.S | N.S | 0.888 | 9.000E-03 |
| 3047 | *Mt-RNR2* | rRNA | N.S | N.S | N.S | N.S | -0.345 | 0.045 |
| 3254 | *MT-TL1* | tRNA | -0.620 | 0.040 | N.S | N.S | N.S | N.S |
| 3351 | *MT-ND1* | C1 | -0.718 | 0.029 | N.S | N.S | -1.238 | 0.029 |
| 3407 | *MT-ND1* | C1 | N.S | N.S | N.S | N.S | -0.151 | 0.029 |
| 3436 | *MT-ND1* | C1 | N.S | N.S | N.S | N.S | -0.448 | 0.021 |
| 3453 | *MT-ND1* | C1 | N.S | N.S | N.S | N.S | -0.414 | 0.05 |
| 3531 | *MT-ND1* | C1 | -0.775 | 2.996E-03 | N.S | N.S | N.S | N.S |
| 3699 | *MT-ND1* | C1 | -0.446 | 1.764E-03 | N.S | N.S | N.S | N.S |
| 3922 | *MT-ND1* | C1 | N.S | N.S | N.S | N.S | -0.213 | 0.034 |
| 3946 | *MT-ND1* | C1 | -0.215 | 0.020 | N.S | N.S | N.S | N.S |
| 3952 | *MT-ND1* | C1 | -0.176 | 0.045 | N.S | N.S | -0.363 | 0.027 |
| 4051 | *MT-ND1* | C1 | -0.134 | 9.467E-03 | N.S | N.S | -0.204 | 0.031 |
| 4127 | *MT-ND1* | C1 | N.S | N.S | N.S | N.S | -0.096 | 0.044 |
| 4141 | *MT-ND1* | C1 | -1.231 | 0.041 | N.S | N.S | N.S | N.S |
| 4152 | *MT-ND1* | C1 | -0.778 | 0.047 | N.S | N.S | N.S | N.S |

| Genomic Position (bp) | Gene | Complex | Effect Size MM | Significance Level (P) MM | Effect Size CER | Significance Level (P) CER | Effect Size STG | Significance Level (P) STG |
|---|---|---|---|---|---|---|---|---|
| 4153 | *MT-ND1* | C1 | -0.227 | 8.578E-03 | N.S | N.S | -0.335 | 5.000E-03 |
| 4375 | *MT-TQ* | tRNA | -0.916 | 0.017 | N.S | N.S | N.S | N.S |
| 4426 | *MT-TM* | tRNA | -0.491 | 0.029 | N.S | N.S | -0.962 | 0.047 |
| 4463 | *MT-TM* | tRNA | -0.501 | 0.036 | N.S | N.S | -0.519 | 0.007 |
| 4490 | *MT-ND2* | C1 | N.S | N.S | 0.491 | 0.027 | N.S | N.S |
| 4531 | *MT-ND2* | C1 | N.S | N.S | N.S | N.S | N.S | N.S |
| 4995 | *MT-ND2* | C1 | N.S | N.S | N.S | N.S | -0.601 | 0.034 |
| 5164 | *MT-ND2* | C1 | N.S | N.S | N.S | N.S | -0.192 | 0.025 |
| 5237 | *MT-ND2* | C1 | -0.210 | 3.134E-03 | N.S | N.S | -0.303 | 0.012 |
| 5271 | *MT-ND2* | C1 | N.S | N.S | N.S | N.S | -0.267 | 0.039 |
| 5399 | *MT-ND2* | C1 | N.S | N.S | N.S | N.S | -1.557 | 0.023 |
| 5737 | *N* | NA | N.S | N.S | N.S | N.S | -0.534 | 0.016 |
| 5743 | *N* | NCR | N.S | N.S | N.S | N.S | -0.805 | 0.025 |
| 5910 | *MT-CO1* | C4 | -0.167 | 0.033 | N.S | N.S | -0.289 | 0.02 |
| 5913 | *MT-CO1* | C4 | N.S | N.S | N.S | N.S | -0.25 | 0.042 |
| 5972 | *MT-CO1* | C4 | N.S | N.S | -0.727 | 0.023 | N.S | N.S |
| 6020 | *MT-CO1* | C4 | -0.390 | 0.031 | N.S | N.S | N.S | N.S |
| 6069 | *MT-CO1* | C4 | -0.140 | 0.046 | N.S | N.S | N.S | N.S |
| 6128 | *MT-CO1* | C4 | -0.620 | 0.033 | N.S | N.S | N.S | N.S |
| 6164 | *MT-CO1* | C4 | -0.702 | 0.043 | N.S | N.S | -1.461 | 0.032 |
| 6181 | *MT-CO1* | C4 | -0.642 | 0.046 | N.S | N.S | N.S | N.S |
| 6190 | *MT-CO1* | C4 | -0.196 | 2.639E-03 | N.S | N.S | -0.246 | 0.047 |
| 6464 | *MT-CO1* | C4 | N.S | N.S | -0.42 | 2.33E-03 | N.S | N.S |
| 6563 | *MT-CO1* | C4 | -0.603 | 0.014 | N.S | N.S | N.S | N.S |
| 6572 | *MT-CO1* | C4 | -0.656 | 0.018 | N.S | N.S | -1.191 | 0.032 |
| 6824 | *MT-CO1* | C4 | -0.426 | 0.019 | N.S | N.S | N.S | N.S |
| 6876 | *MT-CO1* | C4 | -0.172 | 0.016 | N.S | N.S | -0.212 | 4.67E-03 |
| 7007 | *MT-CO1* | C4 | -0.679 | 0.013 | N.S | N.S | N.S | N.S |
| 7012 | *MT-CO1* | C4 | -0.643 | 0.033 | -0.562 | 3.95E-03 | N.S | N.S |
| 7019 | *MT-CO1* | C4 | -1.176 | 1.343E-04 | -1.612 | 7.13E-03 | N.S | N.S |

| Genomic Position (bp) | Gene | Complex | Effect Size MM | Significance Level (P) MM | Effect Size CER | Significance Level (P) CER | Effect Size STG | Significance Level (P) STG |
|---|---|---|---|---|---|---|---|---|
| 7161 | *MT-CO1* | C4 | N.S | N.S | N.S | N.S | -0.381 | 0.02 |
| 7332 | *MT-CO1* | C4 | N.S | N.S | N.S | N.S | -0.213 | 0.048 |
| 7599 | *MT-CO2* | C4 | -0.925 | 0.044 | -1.148 | 0.029 | N.S | N.S |
| 7757 | *MT-CO2* | C4 | N.S | N.S | -0.251 | 0.049 | N.S | N.S |
| 7792 | *MT-CO2* | C4 | -0.923 | 4.513E-03 | -0.882 | 4.89E-03 | N.S | N.S |
| 7813 | *MT-CO2* | C4 | -0.827 | 0.015 | N.S | N.S | N.S | N.S |
| 7849 | MT-CO2 | C4 | N.S | N.S | N.S | N.S | 0.551 | 0.016 |
| 7859 | MT-CO2 | C4 | N.S | N.S | N.S | N.S | -0.143 | 0.021 |
| 7910 | MT-CO2 | C4 | -0.108 | 0.016 | N.S | N.S | N.S | N.S |
| 7928 | MT-CO2 | C4 | -0.228 | 6.582E-03 | -0.229 | 1.45E-03 | N.S | N.S |
| 7978 | MT-CO2 | C4 | -0.663 | 3.750E-03 | N.S | N.S | N.S | N.S |
| 7979 | MT-CO2 | C4 | N.S | N.S | N.S | N.S | -0.346 | 0.012 |
| 8036 | MT-CO2 | C4 | N.S | N.S | -0.558 | 0.041 | N.S | N.S |
| 8140 | MT-CO2 | C4 | 0.495 | 0.032 | N.S | N.S | N.S | N.S |
| 8213 | MT-CO2 | C4 | N.S | N.S | N.S | N.S | -0.237 | 8.85E-03 |
| 8387 | MT-ATP8 | C5 | N.S | N.S | N.S | N.S | -0.439 | 0.019 |
| 8584 | MT-ATP6 | C5 | -0.727 | 0.046 | N.S | N.S | N.S | N.S |
| 8722 | **MT-ATP6** | **C5** | **-1.547** | **7.560E-06** | **-1.649** | **1.04E-03** | **-1.446** | **4.20E-03** |
| 8879 | MT-ATP6 | C5 | -0.179 | 0.011 | N.S | N.S | N.S | N.S |
| 8958 | MT-ATP6 | C5 | -1.089 | 9.311E-03 | N.S | N.S | -1.729 | 0.012 |
| 9009 | MT-ATP6 | C5 | -0.875 | 5.939E-04 | N.S | N.S | -1.045 | 0.038 |
| 9010 | MT-ATP6 | C5 | N.S | N.S | N.S | N.S | -0.153 | 0.033 |
| 9055 | **MT-ATP6** | **C5** | **0.519** | **7.934E-05** | **0.571** | **7.78E-03** | **0.467** | **6.38E-03** |
| 9195 | MT-ATP6 | C5 | N.S | N.S | N.S | N.S | -0.985 | 0.047 |
| 9381 | MT-CO3 | C4 | -0.736 | 0.019 | -1.163 | 0.021 | N.S | N.S |
| 9384 | MT-CO3 | C4 | -0.174 | 0.048 | N.S | N.S | N.S | N.S |
| 9449 | MT-CO3 | C4 | -0.575 | 0.022 | N.S | N.S | N.S | N.S |
| 9488 | MT-CO3 | C4 | -0.704 | 0.048 | N.S | N.S | N.S | N.S |
| 9774 | MT-CO3 | C4 | -0.148 | 7.120E-03 | N.S | N.S | -0.272 | 0.016 |
| 9786 | MT-CO3 | C4 | -0.295 | 0.020 | N.S | N.S | N.S | N.S |

| Genomic Position (bp) | Gene | Complex | Effect Size MM | Significance Level (P) MM | Effect Size CER | Significance Level (P) CER | Effect Size STG | Significance Level (P) STG |
|---|---|---|---|---|---|---|---|---|
| 9828 | MT-CO3 | C4 | -0.161 | 0.012 | N.S | N.S | -0.231 | 3.00E-03 |
| 10067 | MT-ND3 | C1 | N.S | N.S | N.S | N.S | -0.949 | 0.038 |
| 10142 | MT-ND3 | C1 | -0.976 | 8.609E-03 | N.S | N.S | -1.275 | 0.04 |
| 10143 | MT-ND3 | C1 | -0.288 | 7.821E-03 | N.S | N.S | -0.477 | 0.035 |
| 10176 | MT-ND3 | C1 | -0.345 | 0.021 | N.S | N.S | N.S | N.S |
| 10197 | MT-ND3 | C1 | -0.213 | 0.033 | N.S | N.S | N.S | N.S |
| 10200 | MT-ND3 | C1 | -0.750 | 0.038 | N.S | N.S | -1.396 | 0.043 |
| 10202 | MT-ND3 | C1 | -0.639 | 6.312E-03 | N.S | N.S | N.S | N.S |
| 10203 | MT-ND3 | C1 | N.S | N.S | N.S | N.S | -0.451 | 0.020 |
| 10436 | MT-TR | tRNA | -1.690 | 0.015 | N.S | N.S | N.S | N.S |
| 10586 | MT-ND4L | C1 | -0.277 | 0.011 | N.S | N.S | N.S | N.S |
| 11126 | MT-ND4 | C1 | -0.185 | 0.011 | N.S | N.S | -0.324 | 2.00E-03 |
| 11389 | MT-ND4 | C1 | -0.576 | 0.038 | N.S | N.S | N.S | N.S |
| 11422 | MT-ND4 | C1 | -0.689 | 0.032 | N.S | N.S | N.S | N.S |
| 11690 | MT-ND4 | C1 | -0.243 | 2.970E-03 | N.S | N.S | -0.39 | 0.024 |
| 11693 | MT-ND4 | C1 | N.S | N.S | N.S | N.S | -0.23 | 0.047 |
| 11716 | MT-ND4 | C1 | 0.461 | 0.038 | N.S | N.S | N.S | N.S |
| 11914 | MT-ND4 | C1 | 0.384 | 0.013 | N.S | N.S | N.S | N.S |
| 12053 | MT-ND4 | C1 | -1.502 | 4.841E-03 | N.S | N.S | N.S | N.S |
| 12054 | MT-ND4 | C1 | -0.159 | 0.011 | N.S | N.S | -0.265 | 0.033 |
| 12192 | MT-TH | tRNA | -0.152 | 0.047 | N.S | N.S | N.S | N.S |
| 12207 | MT-TS2 | tRNA | -0.254 | 0.024 | -0.207 | 0.029 | N.S | N.S |
| 12456 | MT-ND5 | C1 | -1.363 | 0.040 | N.S | N.S | N.S | N.S |
| 12457 | MT-ND5 | C1 | -0.125 | 0.029 | N.S | N.S | N.S | N.S |
| 12528 | MT-ND5 | C1 | N.S | N.S | N.S | N.S | N.S | N.S |
| 12813 | MT-ND5 | C1 | N.S | N.S | N.S | N.S | -0.423 | 0.034 |
| 12814 | MT-ND5 | C1 | -0.187 | 6.048E-03 | N.S | N.S | -0.303 | 0.049 |
| 12868 | MT-ND5 | C1 | -0.230 | 0.043 | N.S | N.S | -0.499 | 0.047 |
| 12871 | MT-ND5 | C1 | -0.148 | 0.045 | N.S | N.S | -0.279 | 0.047 |
| 13121 | MT-ND5 | C1 | N.S | N.S | N.S | N.S | -0.349 | 0.048 |

| Genomic Position (bp) | Gene | Complex | Effect Size MM | Significance Level (P) MM | Effect Size CER | Significance Level (P) CER | Effect Size STG | Significance Level (P) STG |
|---|---|---|---|---|---|---|---|---|
| 13210 | MT-ND5 | C1 | -0.268 | 6.665E-04 | N.S | N.S | -0.38 | 0.01 |
| 13288 | MT-ND5 | C1 | -0.216 | 0.037 | N.S | N.S | N.S | N.S |
| 13322 | MT-ND5 | C1 | -0.461 | 0.043 | N.S | N.S | N.S | N.S |
| 13336 | MT-ND5 | C1 | -0.244 | 0.045 | N.S | N.S | N.S | N.S |
| 13406 | MT-ND5 | C1 | N.S | N.S | N.S | N.S | -0.119 | 0.041 |
| 13530 | MT-ND5 | C1 | N.S | N.S | N.S | N.S | -0.846 | 0.037 |
| 13551 | MT-ND5 | C1 | -0.714 | 0.024 | N.S | N.S | -0.945 | 0.035 |
| 13596 | MT-ND5 | C1 | -0.556 | 0.050 | N.S | N.S | N.S | N.S |
| 13597 | MT-ND5 | C1 | -0.207 | 1.857E-03 | N.S | N.S | -0.271 | 0.033 |
| 13610 | MT-ND5 | C1 | -0.285 | 0.024 | N.S | N.S | N.S | N.S |
| 13643 | MT-ND5 | C1 | -0.566 | 0.028 | N.S | N.S | -0.814 | 0.047 |
| 13703 | MT-ND5 | C1 | -0.179 | 0.044 | N.S | N.S | N.S | N.S |
| 13758 | MT-ND5 | C1 | -0.858 | 0.028 | -1.098 | 7.63E-04 | N.S | N.S |
| 13915 | MT-ND5 | C1 | -0.284 | 8.943E-04 | N.S | N.S | -0.452 | 0.01 |
| 13940 | MT-ND5 | C1 | -0.214 | 0.023 | N.S | N.S | N.S | N.S |
| 14160 | MT-ND6 | C1 | -0.345 | 0.022 | N.S | N.S | -0.445 | 0.023 |
| 14226 | MT-ND6 | C1 | N.S | N.S | 0.168 | 0.039 | N.S | N.S |
| 14249 | **MT-ND6** | **C1** | **-0.326** | **1.939E-03** | **-0.214** | **0.023** | **-0.437** | **0.033** |
| 14384 | MT-ND6 | C1 | -0.234 | 0.013 | N.S | N.S | N.S | N.S |
| 14559 | MT-ND6 | C1 | -3.281 | 2.507E-04 | N.S | N.S | -4.35 | 0.002 |
| 14568 | MT-ND6 | C1 | -2.704 | 0.020 | N.S | N.S | N.S | N.S |
| 14697 | MT-TE | tRNA | -0.833 | 0.038 | N.S | N.S | N.S | N.S |
| 14680 | MT-TE | tRNA | N.S | N.S | N.S | N.S | -1.265 | 0.036 |
| 14698 | MT-TE | tRNA | N.S | N.S | N.S | N.S | -0.283 | 0.042 |
| 14720 | MT-TE | tRNA | -1.369 | 9.117E-03 | N.S | N.S | -1.943 | 0.016 |
| 14759 | MT-CYB | C3 | -0.879 | 0.012 | N.S | N.S | -1.115 | 0.046 |
| 14831 | MT-CYB | C3 | -0.187 | 0.017 | N.S | N.S | N.S | N.S |
| 14945 | MT-CYB | C3 | 0.089 | 0.045 | 0.147 | 0.037 | N.S | N.S |
| 14963 | MT-CYB | C3 | N.S | N.S | N.S | N.S | N.S | N.S |
| 14996 | MT-CYB | C3 | N.S | N.S | N.S | N.S | -0.236 | 0.028 |

| Genomic Position (bp) | Gene | Complex | Effect Size MM | Significance Level (P) MM | Effect Size CER | Significance Level (P) CER | Effect Size STG | Significance Level (P) STG |
|---|---|---|---|---|---|---|---|---|
| 15005 | MT-CYB | C3 | -0.171 | 0.042 | N.S | N.S | N.S | N.S |
| 15044 | MT-CYB | C3 | N.S | N.S | -0.468 | 0.043 | N.S | N.S |
| 15045 | MT-CYB | C3 | -0.218 | 0.027 | N.S | N.S | -0.248 | 0.037 |
| 15148 | MT-CYB | C3 | -0.263 | 0.011 | N.S | N.S | N.S | N.S |
| 15199 | MT-CYB | C3 | N.S | N.S | N.S | N.S | -0.726 | 0.049 |
| 15391 | MT-CYB | C3 | N.S | N.S | N.S | N.S | -0.956 | 0.046 |
| 15431 | MT-CYB | C3 | -0.225 | 7.557E-03 | N.S | N.S | -0.304 | 0.042 |
| 15500 | MT-CYB | C3 | -0.238 | 0.024 | N.S | N.S | N.S | N.S |
| 15590 | MT-CYB | C3 | -0.473 | 0.050 | N.S | N.S | -1.099 | 0.030 |
| 15591 | MT-CYB | C3 | -0.275 | 9.043E-03 | N.S | N.S | -0.45 | 9.39E-03 |
| 15595 | MT-CYB | C3 | -0.986 | 4.346E-03 | N.S | N.S | N.S | N.S |
| 15616 | MT-CYB | C3 | -0.663 | 0.013 | N.S | N.S | N.S | N.S |
| 15761 | **MT-CYB** | **C3** | **-0.374** | **2.732E-04** | **-0.491** | **1.15E-03** | **-0.257** | **0.028** |
| 15811 | MT-CYB | C3 | -1.576 | 2.328E-03 | -1.471 | 0.027 | N.S | N.S |
| 15812 | MT-CYB | C3 | N.S | N.S | 0.09 | 0.037 | N.S | N.S |
| 16083 | D-LOOP | NCR | -1.498 | 0.022 | N.S | N.S | N.S | N.S |
| 16129 | D-LOOP | NCR | N.S | N.S | N.S | N.S | -0.515 | 9.44E-03 |
| 16329 | D-LOOP | NCR | -0.159 | 0.048 | N.S | N.S | -0.303 | 0.012 |
| 16495 | D-LOOP | NCR | 0.547 | 7.115E-03 | N.S | N.S | 0.765 | 0.025 |

### *5.4.4. MtDNA Methylation Varies Significantly between Different Anatomical Regions of the Brain*

Given that we have recently identified tissue-specific DMRs in mtDNA across different brain regions using MeDIP-Seq (Chapter 3) (64), we were interested in investigating whether differences in mtDNA could be identified at single cytosine resolution between anatomically distinct brain regions. In our multi-level model, we identified 80 nominally significant ($P<0.05$) differences in mtDNA methylation across the genome that were associated with tissue type, of which 57 were found to be hypermethylated in CER relative to STG (**Table 5.4**).

***Table 5.4: Nominally significant DMPs associated with tissue type.*** *In total 80 tissue specific DMPs were identified between STG and CER, shown here in order of significance. Positive effect sizes represent hypermethylation in the CER compared to the superior temporal gyrus. Abbreviations: C1 (Complex I), C2 (Complex II), C3 (Complex III), C4 (Complex IV), C5 (Complex V), NCR (non-coding region), rRNA (ribosomal RNA), tRNA (transfer RNA).*

| Genomic Position (bp) | Gene | Complex | Effect Size | Significance Level (P) |
|---|---|---|---|---|
| 14846 | *MT-CYB* | C3 | 0.233 | 2.159E-04 |
| 6563 | *MT-CO1* | C4 | -0.996 | 2.242E-04 |
| 14945 | *MT-CYB* | C3 | 0.188 | 2.514E-04 |
| 3699 | *MT-ND1* | C1 | -0.503 | 3.343E-04 |
| 13198 | *MT-ND5* | C1 | 0.110 | 1.465E-03 |
| 4712 | *MT-ND2* | C1 | 1.631 | 1.902E-03 |
| 1216 | *MT-RNR1* | rRNA | -0.828 | 2.154E-03 |
| 5756 | *N* | NCR | 0.325 | 2.423E-03 |
| 8580 | *MT-ATP6* | C5 | 1.350 | 2.809E-03 |
| 9673 | *MT-CO3* | C4 | 0.209 | 3.283E-03 |
| 6824 | *MT-CO1* | C4 | 0.424 | 3.554E-03 |
| 10670 | *MT-ND4L* | C1 | -0.575 | 3.751E-03 |
| 6015 | *MT-CO1* | C4 | 0.434 | 4.413E-03 |
| 10202 | *MT-ND3* | C1 | 0.639 | 4.651E-03 |
| 12735 | *MT-ND5* | C1 | 0.965 | 4.685E-03 |
| 9776 | *MT-CO3* | C4 | 0.558 | 4.959E-03 |
| 16454 | *D-LOOP* | NCR | 0.657 | 5.208E-03 |
| 2718 | *MT-RNR2* | rRNA | 0.781 | 5.426E-03 |
| 13669 | *MT-ND5* | C1 | 0.229 | 5.518E-03 |
| 3406 | *MT-ND1* | C1 | -0.732 | 6.134E-03 |
| 11717 | *MT-ND4* | C1 | 0.260 | 6.446E-03 |
| 1537 | *MT-RNR1* | rRNA | -0.845 | 6.464E-03 |
| 10536 | *MT-ND4L* | C1 | -1.696 | 6.720E-03 |
| 1314 | *MT-RNR1* | rRNA | 1.056 | 7.333E-03 |
| 9773 | *MT-CO3* | C4 | 0.651 | 7.361E-03 |
| 9294 | *MT-CO3* | C4 | -0.233 | 8.445E-03 |
| 12876 | *MT-ND5* | C1 | -0.505 | 9.341E-03 |
| 6020 | *MT-CO1* | C4 | 0.323 | 0.010 |
| 7859 | *MT-CO2* | C4 | 0.152 | 0.011 |
| 15590 | *MT-CYB* | C3 | 0.629 | 0.011 |
| 13524 | *MT-ND5* | C1 | 0.498 | 0.012 |

| Genomic Position (bp) | Gene | Complex | Effect Size | Significance Level (P) |
|---|---|---|---|---|
| 11163 | *MT-ND4* | C1 | 0.134 | 0.012 |
| 1488 | *MT-RNR1* | rRNA | -1.085 | 0.013 |
| 12813 | *MT-ND5* | C1 | 0.371 | 0.014 |
| 3453 | *MT-ND1* | C1 | -0.553 | 0.014 |
| 4996 | *MT-ND2* | C1 | 0.208 | 0.015 |
| 6329 | *MT-CO1* | C4 | 0.497 | 0.015 |
| 8114 | *MT-CO2* | C4 | 0.290 | 0.015 |
| 2809 | *MT-RNR2* | rRNA | 0.657 | 0.016 |
| 6456 | *MT-CO1* | C4 | 0.184 | 0.017 |
| 6075 | *MT-CO1* | C4 | 0.099 | 0.018 |
| 6016 | *MT-CO1* | C4 | 0.086 | 0.020 |
| 1226 | *MT-RNR1* | rRNA | -0.435 | 0.020 |
| 33 | *D-LOOP* | NCR | -0.840 | 0.021 |
| 6053 | *MT-CO1* | C4 | 0.429 | 0.021 |
| 1597 | *MT-RNR1* | rRNA | -0.637 | 0.021 |
| 9828 | *MT-CO3* | C4 | 0.122 | 0.021 |
| 6128 | *MT-CO1* | C4 | 0.392 | 0.022 |
| 8544 | *MT-ATP8* | C5 | 0.553 | 0.023 |
| 3094 | *MT-RNR2* | rRNA | 0.601 | 0.023 |
| 740 | *MT-RNR1* | rRNA | -0.161 | 0.026 |
| 9752 | *MT-CO3* | C4 | 0.498 | 0.026 |
| 3889 | *MT-ND1* | C1 | -0.637 | 0.027 |
| 120 | *D-LOOP* | NCR | 0.647 | 0.028 |
| 3435 | *MT-ND1* | C1 | -0.464 | 0.028 |
| 15045 | *MT-CYB* | C3 | 0.199 | 0.029 |
| 187 | *D-LOOP* | NCR | 0.140 | 0.030 |
| 11184 | *MT-ND4* | C1 | 0.099 | 0.030 |
| 11861 | *MT-ND4* | C1 | 0.135 | 0.031 |
| 15005 | *MT-CYB* | C3 | 0.139 | 0.032 |
| 6808 | *MT-CO1* | C4 | -0.104 | 0.032 |
| 2566 | *MT-RNR2* | rRNA | 0.340 | 0.031 |
| 7130 | *MT-CO1* | C4 | -0.446 | 0.035 |
| 412 | *D-LOOP* | NCR | 0.199 | 0.037 |
| 15147 | *MT-CYB* | C3 | -0.460 | 0.038 |
| 900 | *MT-RNR1* | rRNA | 0.393 | 0.039 |
| 3127 | *MT-RNR2* | rRNA | 0.521 | 0.039 |
| 1323 | *MT-RNR1* | rRNA | 0.184 | 0.039 |
| 8722 | *MT-ATP6* | C5 | -0.286 | 0.039 |
| 1302 | *MT-RNR1* | rRNA | 0.461 | 0.040 |
| 11004 | *MT-ND4* | C1 | 0.169 | 0.040 |
| 5353 | *MT-ND2* | C1 | 0.184 | 0.041 |
| 10736 | *MT-ND4L* | C1 | 0.812 | 0.043 |
| 5053 | *MT-ND2* | C1 | -1.278 | 0.043 |
| 13579 | *MT-ND5* | C1 | 0.119 | 0.044 |

| Genomic Position (bp) | Gene | Complex | Effect Size | Significance Level (P) |
|---|---|---|---|---|
| 3946 | *MT-ND1* | C1 | 0.137 | 0.045 |
| 15059 | *MT-CYB* | C3 | 0.141 | 0.045 |
| 8958 | *MT-ATP6* | C5 | -0.739 | 0.047 |
| 3172 | *MT-RNR2* | rRNA | 1.149 | 0.049 |
| 14681 | *MT-TE* | tRNA | 0.184 | 0.049 |

### *5.4.5. MtDNA Methylation has a Tendency to Increase with Age Across the Mitochondrial Genome*

Given that alterations in mitochondrial function have been associated with age, as reviewed by Amigo et al. (258), we were interested to investigate whether age may alter mtDNA methylation patterns. Our initial multilevel model identified 128 nominally significant probes associated with age, 119 of which were positively correlated with increasing age. Using our post-hoc linear regression models defined earlier for CER and STG, we tested this association in both tissues separately. Despite the small age range of the samples (77-87, mean=82, S.D.=3.46), we were able to identify 59 and 71 nominally significant DMPs for age in CER and STG respectively (**Table 5.5**). There was a positive association with age for 40 nominally significant DMPs in CER and 69 nominally significant DMPs in STG. In total, only two DMPs were found to be nominally significant, with the same direction of methylation change (both hypermethylated) in both tissues: 1597bp (*MT-RNR1)* and 4463bp (*MT-TM)* (**Figure 5.3**). Given that, for the majority of age-associated DMPs we observed hypermethylation with increasing age, we performed a Fisher's exact test for enrichment on each individual tissue. This identified a significant enrichment for hypermethylated DMPs in association with increasing age in both the STG (P=5.074E-05; Odds Ratio = 8.439) and the CER (P=0.010; Odds Ratio=1.992).

***Table 5.5: Nominally significant DMPs associated with age.*** *DMPs were identified initially in the multi-level model, also shown are results from subsequent post-hoc linear regressions within each tissue seperately. Shown is the direction of effect and significance level for each DMP identified in the multi-level model and individual tissue regressions. Positive effect sizes show methylation levels positively correlating with age. Probes highlighted in bold indicate consistent DMPs in both significance and direction of effect in both CER and STG. Abbreviations: C1 (Complex I), C2 (Complex II), C3 (Complex III), C4 (Complex IV), C5 (Complex V), NCR (non-coding region), rRNA (ribosomal RNA), tRNA (transfer RNA), STG (superior temporal gyrus), CER (CER), MM (mixed model), N.S. (Not significant)*

| Genomic Position (bp) | Gene | Complex | Effect Size MM | Significance level (P Value) MM | Effect Size CER | Significance level (P Value) CER | Effect Size STG | Significance level (P Value) STG |
|---|---|---|---|---|---|---|---|---|
| 62 | *D-LOOP* | NCR | N.S | N.S | N.S | N.S | 0.011 | 0.044 |
| 80 | *D-LOOP* | NCR | N.S | N.S | -0.094 | 0.043 | N.S | N.S |
| 97 | *D-LOOP* | NCR | N.S | N.S | N.S | N.S | 0.040 | 0.013 |
| 186 | *D-LOOP* | NCR | -0.263 | 0.036 | N.S | N.S | N.S | N.S |
| 412 | *D-LOOP* | NCR | 0.056 | 1.327E-03 | N.S | N.S | 0.085 | 2.000E-03 |
| 498 | *D-LOOP* | NCR | -0.863 | 0.034 | N.S | N.S | N.S | N.S |
| 525 | *D-LOOP* | NCR | N.S | N.S | N.S | N.S | 0.443 | 0.032 |
| 545 | *D-LOOP* | NCR | 0.345 | 0.011 | 0.393 | 0.023 | N.S | N.S |
| 739 | *MT-RNR1* | rRNA | 0.143 | 1.223E-03 | N.S | N.S | 0.124 | 0.003 |
| 808 | *MT-RNR1* | rRNA | 0.110 | 3.033E-03 | N.S | N.S | 0.129 | 0.014 |
| 900 | *MT-RNR1* | rRNA | 0.073 | 0.018 | N.S | N.S | 0.095 | 9.31E-03 |
| 932 | *MT-RNR1* | rRNA | 0.085 | 0.036 | N.S | N.S | 0.173 | 0.012 |
| 935 | *MT-RNR1* | rRNA | N.S | N.S | N.S | N.S | 0.13 | 0.041 |

| Genomic Position (bp) | Gene | Complex | Effect Size MM | Significance level (P Value) MM | Effect Size CER | Significance level (P Value) CER | Effect Size STG | Significance level (P Value) STG |
|---|---|---|---|---|---|---|---|---|
| 1134 | *MT-RNR1* | rRNA | 0.049 | 4.091E-03 | 0.049 | 0.044 | N.S | N.S |
| 1146 | *MT-RNR1* | rRNA | N.S | N.S | N.S | N.S | 0.122 | 0.016 |
| 1217 | *MT-RNR1* | rRNA | N.S | N.S | -0.038 | 0.033 | N.S | N.S |
| 1226 | *MT-RNR1* | rRNA | 0.120 | 0.030 | N.S | N.S | N.S | N.S |
| 1227 | *MT-RNR1* | rRNA | 0.035 | 0.011 | N.S | N.S | N.S | N.S |
| 1323 | *MT-RNR1* | rRNA | -0.033 | 0.026 | N.S | N.S | N.S | N.S |
| 1561 | *MT-RNR1* | rRNA | 0.142 | 4.348E-03 | N.S | N.S | 0.173 | 0.015 |
| **1597** | ***MT-RNR1*** | **rRNA** | **0.226** | **2.396E-04** | **0.234** | **0.045** | **0.217** | **0.030** |
| 1750 | *Mt-RNR2* | rRNA | N.S | N.S | -0.091 | 6.779E-03 | N.S | N.S |
| 1916 | *Mt-RNR2* | rRNA | -0.025 | 0.021 | -0.033 | 0.013 | N.S | N.S |
| 2202 | *Mt-RNR2* | rRNA | 0.315 | 0.030 | N.S | N.S | N.S | N.S |
| 2203 | *Mt-RNR2* | rRNA | 0.045 | 0.034 | N.S | N.S | N.S | N.S |
| 2566 | *Mt-RNR2* | rRNA | -0.054 | 0.046 | N.S | N.S | N.S | N.S |
| 2824 | *Mt-RNR2* | rRNA | 0.114 | 0.019 | N.S | N.S | N.S | N.S |
| 2878 | *Mt-RNR2* | rRNA | 0.025 | 0.034 | N.S | N.S | N.S | N.S |
| 2988 | *Mt-RNR2* | rRNA | N.S | N.S | N.S | N.S | -0.173 | 8.29E-03 |
| 3078 | *Mt-RNR2* | rRNA | -0.369 | 0.011 | N.S | N.S | N.S | N.S |
| 3079 | *Mt-RNR2* | rRNA | 0.063 | 0.016 | N.S | N.S | N.S | N.S |
| 3247 | *MT-TL1* | tRNA | 0.231 | 0.034 | N.S | N.S | N.S | N.S |
| 3248 | *MT-TL1* | tRNA | 0.038 | 0.032 | N.S | N.S | N.S | N.S |
| 3254 | *MT-TL1* | tRNA | 0.168 | 2.662E-03 | 0.227 | 0.045 | N.S | N.S |
| 3351 | *MT-ND1* | C1 | 0.103 | 0.039 | N.S | N.S | N.S | N.S |
| 3352 | *MT-ND1* | C1 | 0.041 | 1.101E-03 | 0.039 | 0.030 | N.S | N.S |
| 3376 | *MT-ND1* | C1 | 0.041 | 0.025 | N.S | N.S | 0.063 | 0.022 |
| 3407 | *MT-ND1* | C1 | N.S | N.S | N.S | N.S | 0.024 | 0.026 |
| 3453 | *MT-ND1* | C1 | N.S | N.S | -0.143 | 0.039 | N.S | N.S |
| 3460 | *MT-ND1* | C1 | 0.052 | 0.013 | N.S | N.S | 0.085 | 0.034 |

| Genomic Position (bp) | Gene | Complex | Effect Size MM | Significance level (P Value) MM | Effect Size CER | Significance level (P Value) CER | Effect Size STG | Significance level (P Value) STG |
|---|---|---|---|---|---|---|---|---|
| 3496 | *MT-ND1* | C1 | N.S | N.S | -0.047 | 0.026 | N.S | N.S |
| 3526 | *MT-ND1* | C1 | 0.036 | 9.772E-03 | N.S | N.S | N.S | N.S |
| 3699 | *MT-ND1* | C1 | 0.044 | 0.020 | N.S | N.S | N.S | N.S |
| 3954 | *MT-ND1* | C1 | 0.125 | 0.012 | 0.172 | 0.048 | N.S | N.S |
| 3955 | *MT-ND1* | C1 | 0.034 | 8.938E-03 | N.S | N.S | N.S | N.S |
| 3967 | *MT-ND1* | C1 | 0.025 | 0.024 | 0.027 | 8.626E-04 | N.S | N.S |
| 4127 | *MT-ND1* | C1 | N.S | N.S | N.S | N.S | 0.038 | 1.82E-03 |
| 4141 | *MT-ND1* | C1 | 0.288 | 6.403E-03 | 0.316 | 0.026 | N.S | N.S |
| 4439 | *MT-TM* | tRNA | 0.162 | 2.134E-03 | N.S | N.S | N.S | N.S |
| 4463 | *MT-TM* | tRNA | 0.167 | 2.359E-04 | 0.234 | 0.019 | 0.101 | 3.08E-03 |
| 4490 | *MT-ND2* | C1 | 0.122 | 4.461E-03 | 0.141 | 3.000E-03 | N.S | N.S |
| 4540 | *MT-ND2* | C1 | 0.093 | 0.047 | N.S | N.S | N.S | N.S |
| 5054 | *MT-ND2* | C1 | 0.037 | 0.013 | N.S | N.S | N.S | N.S |
| 4848 | *MT-ND2* | C1 | N.S | N.S | N.S | N.S | 0.057 | 0.039 |
| 5111 | *MT-ND2* | C1 | 0.460 | 2.107E-03 | N.S | N.S | 0.759 | 0.027 |
| 5146 | *MT-ND2* | C1 | 0.204 | 0.043 | N.S | N.S | N.S | N.S |
| 5163 | *MT-ND2* | C1 | 0.329 | 6.088E-03 | 0.368 | 0.045 | N.S | N.S |
| 5164 | *MT-ND2* | C1 | N.S | N.S | N.S | N.S | 0.036 | 0.013 |
| 5237 | *MT-ND2* | C1 | 0.020 | 0.040 | N.S | N.S | 0.041 | 0.018 |
| 5244 | *MT-ND2* | C1 | 0.036 | 0.012 | N.S | N.S | N.S | N.S |
| 5271 | *MT-ND2* | C1 | 0.021 | 0.040 | N.S | N.S | 0.039 | 0.045 |
| 5459 | *MT-ND2* | C1 | 0.351 | 0.034 | 0.687 | 0.046 | N.S | N.S |
| 5470 | *MT-ND2* | C1 | 0.274 | 0.030 | 0.492 | 0.035 | N.S | N.S |
| 5737 | *N* | NCR | N.S | N.S | N.S | N.S | 0.074 | 0.023 |
| 5768 | *MT-TC* | tRNA | N.S | N.S | N.S | N.S | 0.066 | 0.03 |
| 5972 | *MT-CO1* | C4 | N.S | N.S | -0.118 | 0.020 | N.S | N.S |
| 5973 | *MT-CO1* | C4 | 0.022 | 0.013 | N.S | N.S | 0.035 | 0.040 |

| Genomic Position (bp) | Gene | Complex | Effect Size MM | Significance level (P Value) MM | Effect Size CER | Significance level (P Value) CER | Effect Size STG | Significance level (P Value) STG |
|---|---|---|---|---|---|---|---|---|
| 6016 | *MT-CO1* | C4 | 0.012 | 0.033 | N.S | N.S | N.S | N.S |
| 6020 | *MT-CO1* | C4 | 0.062 | 0.027 | N.S | N.S | N.S | N.S |
| 6165 | *MT-CO1* | C4 | 0.031 | 0.020 | 0.046 | 0.039 | N.S | N.S |
| 6263 | *MT-CO1* | C4 | N.S | N.S | 0.100 | 0.039 | N.S | N.S |
| 6464 | *MT-CO1* | C4 | 0.132 | 7.709E-03 | N.S | N.S | N.S | N.S |
| 6564 | *MT-CO1* | C4 | N.S | N.S | 0.140 | 1.204E-04 | N.S | N.S |
| 6572 | *MT-CO1* | C4 | 0.098 | 0.021 | 0.147 | 0.022 | N.S | N.S |
| 6689 | *MT-CO1* | C4 | 0.217 | 3.436E-03 | N.S | N.S | 0.249 | 0.034 |
| 6808 | *MT-CO1* | C4 | 0.015 | 0.049 | N.S | N.S | N.S | N.S |
| 6824 | *MT-CO1* | C4 | 0.067 | 0.018 | N.S | N.S | N.S | N.S |
| 6852 | *MT-CO1* | C4 | 0.040 | 0.031 | N.S | N.S | N.S | N.S |
| 6855 | *MT-CO1* | C4 | 0.022 | 0.047 | N.S | N.S | N.S | N.S |
| 6876 | *MT-CO1* | C4 | N.S | N.S | N.S | N.S | 0.024 | 0.014 |
| 7012 | *MT-CO1* | C4 | 0.126 | 0.012 | 0.178 | 2.534E-04 | N.S | N.S |
| 7160 | *MT-CO1* | C4 | 0.092 | 7.440E-03 | N.S | N.S | N.S | N.S |
| 7161 | *MT-CO1* | C4 | N.S | N.S | N.S | N.S | 0.050 | 0.035 |
| 7218 | *MT-CO1* | C4 | N.S | N.S | -0.175 | 0.010 | N.S | N.S |
| 7427 | *MT-CO1* | C4 | 0.097 | 0.048 | N.S | N.S | N.S | N.S |
| 7462 | *MT-TS1* | tRNA | 0.088 | 0.032 | 0.121 | 0.037 | N.S | N.S |
| 7600 | *MT-CO2* | C4 | 0.035 | 7.255E-03 | N.S | N.S | N.S | N.S |
| 7775 | *MT-CO2* | C4 | N.S | N.S | N.S | N.S | 0.015 | 0.019 |
| 7849 | *MT-CO2* | C4 | -0.089 | 0.038 | N.S | N.S | -0.108 | 6.92E-03 |
| 7859 | *MT-CO2* | C4 | N.S | N.S | N.S | N.S | 0.022 | 0.023 |
| 7928 | *MT-CO2* | C4 | 0.031 | 0.012 | 0.030 | 3.000E-03 | N.S | N.S |
| 7978 | *MT-CO2* | C4 | 0.077 | 0.020 | N.S | N.S | N.S | N.S |
| 7979 | *MT-CO2* | C4 | N.S | N.S | N.S | N.S | 0.045 | 0.020 |
| 7997 | *MT-CO2* | C4 | 0.015 | 0.048 | N.S | N.S | N.S | N.S |

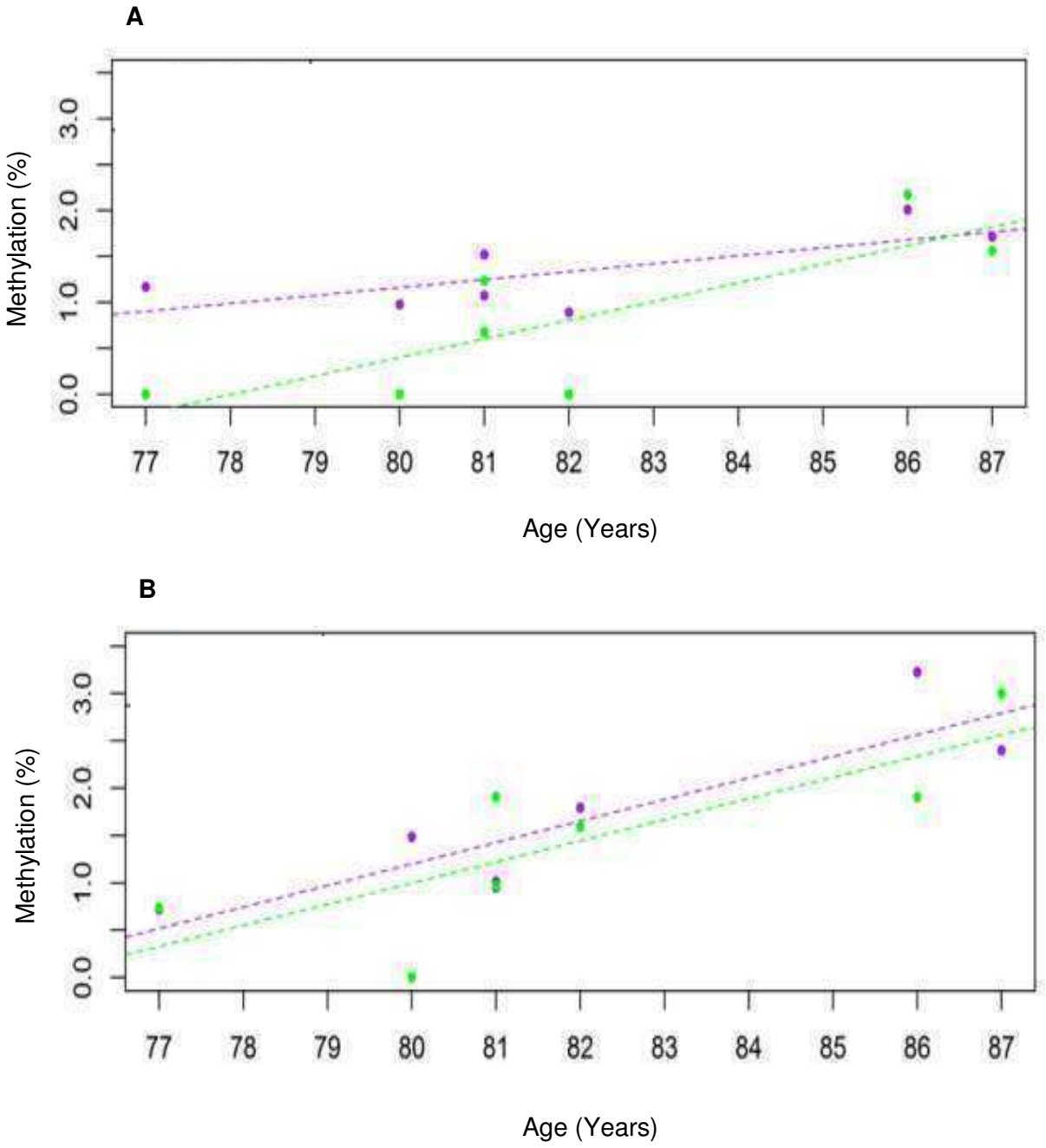
| Genomic Position (bp) | Gene | Complex | Effect Size MM | Significance level (P Value) MM | Effect Size CER | Significance level (P Value) CER | Effect Size STG | Significance level (P Value) STG |
|---|---|---|---|---|---|---|---|---|
| 8036 | *MT-CO2* | C4 | 0.103 | 3.611E-04 | 0.109 | 0.020 | N.S | N.S |
| 8113 | *MT-CO2* | C4 | 0.083 | 0.033 | N.S | N.S | N.S | N.S |
| 8117 | *MT-CO2* | C4 | 0.096 | 0.024 | N.S | N.S | N.S | N.S |
| 8140 | *MT-CO2* | C4 | N.S | N.S | 0.105 | 0.042 | N.S | N.S |
| 8141 | *MT-CO2* | C4 | N.S | N.S | N.S | N.S | 0.027 | 0.033 |
| 8147 | *MT-CO2* | C4 | 0.158 | 0.033 | N.S | N.S | N.S | N.S |
| 8151 | *MT-CO2* | C4 | 0.083 | 0.022 | N.S | N.S | N.S | N.S |
| 8165 | *MT-CO2* | C4 | 0.032 | 0.037 | N.S | N.S | N.S | N.S |
| 8165 | *MT-CO2* | C4 | N.S | N.S | N.S | N.S | 0.054 | 0.032 |
| 8545 | *MT-ATP8* | C5 | 0.033 | 0.018 | N.S | N.S | 0.058 | 0.022 |
| 8580 | *MT-ATP6* | C5 | 0.156 | 0.027 | N.S | N.S | 0.265 | 0.015 |
| 8648 | *MT-ATP6* | C5 | 0.025 | 0.047 | N.S | N.S | N.S | N.S |
| 8722 | *MT-ATP6* | C5 | -0.065 | 0.018 | -0.095 | 0.033 | N.S | N.S |
| 8879 | *MT-ATP6* | C5 | 0.025 | 0.019 | N.S | N.S | N.S | N.S |
| 8959 | *MT-ATP6* | C5 | N.S | N.S | -0.026 | 0.025 | N.S | N.S |
| 9010 | *MT-ATP6* | C5 | N.S | N.S | N.S | N.S | 0.032 | 0.013 |
| 9328 | *MT-CO3* | C4 | N.S | N.S | N.S | N.S | 0.085 | 0.043 |
| 9393 | *MT-CO3* | C4 | N.S | N.S | N.S | N.S | 0.090 | 0.033 |
| 9445 | *MT-CO3* | C4 | N.S | N.S | 0.032 | 0.040 | N.S | N.S |
| 9449 | *MT-CO3* | C4 | 0.077 | 0.044 | N.S | N.S | N.S | N.S |
| 9774 | *MT-CO3* | C4 | N.S | N.S | 0.021 | 2.434E-03 | N.S | N.S |
| 9777 | *MT-CO3* | C4 | 0.035 | 0.044 | N.S | N.S | N.S | N.S |
| 9918 | *MT-TG* | tRNA | 0.025 | 0.011 | N.S | N.S | N.S | N.S |
| 10067 | *MT-ND3* | C1 | 0.135 | 0.014 | N.S | N.S | N.S | N.S |
| 10182 | *MT-ND3* | C1 | 0.030 | 0.046 | N.S | N.S | N.S | N.S |
| 10202 | *MT-ND3* | C1 | 0.099 | 6.257E-03 | N.S | N.S | N.S | N.S |
| 10203 | *MT-ND3* | C1 | 0.040 | 0.044 | N.S | N.S | 0.062 | 0.029 |

| Genomic Position (bp) | Gene | Complex | Effect Size MM | Significance level (P Value) MM | Effect Size CER | Significance level (P Value) CER | Effect Size STG | Significance level (P Value) STG |
|---|---|---|---|---|---|---|---|---|
| 10436 | *MT-TR* | tRNA | -0.232 | 0.026 | -0.492 | 0.015 | N.S | N.S |
| 10676 | *MT-ND4L* | C1 | 0.179 | 1.136E-03 | 0.258 | 8.800E-03 | N.S | N.S |
| 10677 | *MT-ND4L* | C1 | 0.022 | 0.038 | N.S | N.S | N.S | N.S |
| 10684 | *MT-ND4L* | C1 | 0.111 | 0.037 | 0.221 | 0.019 | N.S | N.S |
| 10934 | *MT-ND4* | C1 | 0.068 | 0.040 | N.S | N.S | N.S | N.S |
| 11162 | *MT-ND4* | C1 | 0.024 | 7.039E-03 | N.S | N.S | 0.126 | 0.016 |
| 11184 | *MT-ND4* | C1 | 0.023 | 0.033 | N.S | N.S | 0.034 | 0.033 |
| 11192 | *MT-ND4* | C1 | N.S | N.S | N.S | N.S | 0.034 | 0.036 |
| 11389 | *MT-ND4* | C1 | N.S | N.S | N.S | N.S | 0.138 | 0.032 |
| 11455 | *MT-ND4* | C1 | 0.093 | 0.021 | N.S | N.S | 0.143 | 4.601E-03 |
| 11492 | *MT-ND4* | C1 | N.S | N.S | N.S | N.S | 0.225 | 0.037 |
| 11611 | *MT-ND4* | C1 | N.S | N.S | N.S | N.S | 0.033 | 0.029 |
| 11647 | *MT-ND4* | C1 | N.S | N.S | N.S | N.S | N.S | N.S |
| 11648 | *MT-ND4* | C1 | 0.079 | 0.027 | -0.028 | 1.194E-03 | N.S | N.S |
| 11689 | *MT-ND4* | C1 | 0.129 | 2.981E-03 | N.S | N.S | N.S | N.S |
| 11852 | *MT-ND4* | C1 | N.S | N.S | N.S | N.S | N.S | N.S |
| 11913 | *MT-ND4* | C1 | 0.161 | 0.012 | N.S | N.S | N.S | N.S |
| 12054 | *MT-ND4* | C1 | 0.021 | 0.023 | N.S | N.S | N.S | N.S |
| 12124 | *MT-ND4* | C1 | N.S | N.S | 0.334 | 0.018 | N.S | N.S |
| 12192 | *MT-TH* | tRNA | 0.049 | 6.945E-04 | 0.056 | 5.77E-03 | N.S | N.S |
| 12207 | *MT-TS2* | tRNA | N.S | N.S | -0.045 | 9.06E-03 | 0.059 | 0.028 |
| 12405 | *MT-ND5* | C1 | 0.320 | 0.042 | N.S | N.S | N.S | N.S |
| 12774 | *MT-ND5* | C1 | 0.116 | 0.044 | N.S | N.S | N.S | N.S |
| 12813 | *MT-ND5* | C1 | 0.051 | 0.028 | N.S | N.S | N.S | N.S |
| 12817 | *MT-ND5* | C1 | N.S | N.S | N.S | N.S | 0.057 | 0.028 |
| 12863 | *MT-ND5* | C1 | 0.024 | 0.027 | N.S | N.S | N.S | N.S |
| 12870 | *MT-ND5* | C1 | 0.065 | 0.046 | 0.096 | 0.010 | N.S | N.S |

| Genomic Position (bp) | Gene | Complex | Effect Size MM | Significance level (P Value) MM | Effect Size CER | Significance level (P Value) CER | Effect Size STG | Significance level (P Value) STG |
|---|---|---|---|---|---|---|---|---|
| 12876 | *MT-ND5* | C1 | 0.065 | 0.028 | N.S | N.S | N.S | N.S |
| 12888 | *MT-ND5* | C1 | 0.061 | 0.048 | 0.079 | 0.043 | N.S | N.S |
| 13210 | *MT-ND5* | C1 | 0.019 | 0.046 | N.S | N.S | 0.041 | 0.032 |
| 13530 | *MT-ND5* | C1 | N.S | N.S | N.S | N.S | 0.127 | 0.040 |
| 13668 | *MT-ND5* | C1 | N.S | N.S | N.S | N.S | 0.145 | 0.014 |
| 13758 | *MT-ND5* | C1 | N.S | N.S | -0.107 | 4.263E-03 | 0.218 | 0.041 |
| 13810 | *MT-ND5* | C1 | N.S | N.S | -0.045 | 0.037 | N.S | N.S |
| 13939 | *MT-ND5* | C1 | N.S | N.S | N.S | N.S | 0.301 | 0.045 |
| 13967 | *MT-ND5* | C1 | 0.193 | 0.025 | N.S | N.S | 0.279 | 0.036 |
| 13968 | *MT-ND5* | C1 | N.S | N.S | N.S | N.S | 0.034 | 0.038 |
| 14159 | *MT-ND6* | C1 | 0.528 | 0.035 | 1.009 | 0.010 | N.S | N.S |
| 14225 | *MT-ND6* | C1 | 0.460 | 8.214E-04 | N.S | N.S | 0.698 | 6.60E-03 |
| 14226 | *MT-ND6* | C1 | 0.027 | 0.026 | N.S | N.S | N.S | N.S |
| 14268 | *MT-ND6* | C1 | N.S | N.S | 0.026 | 0.019 | N.S | N.S |
| 14383 | *MT-ND6* | C1 | 0.258 | 8.788E-04 | N.S | N.S | 0.397 | 0.018 |
| 14458 | *MT-ND6* | C1 | 0.218 | 0.017 | 0.344 | 0.031 | N.S | N.S |
| 14459 | *MT-ND6* | C1 | 0.234 | 0.027 | N.S | N.S | N.S | N.S |
| 14568 | *MT-ND6* | C1 | 0.379 | 0.029 | N.S | N.S | N.S | N.S |
| 14698 | *MT-TE* | tRNA | 0.029 | 7.440E-03 | 0.024 | 0.023 | N.S | N.S |
| 14804 | *MT-CYB* | C3 | N.S | N.S | N.S | N.S | 0.047 | 0.048 |
| 14929 | *MT-CYB* | C3 | 0.078 | 0.027 | 0.107 | 0.046 | N.S | N.S |
| 14958 | *MT-CYB* | C3 | 0.026 | 0.042 | N.S | N.S | N.S | N.S |
| 14963 | *MT-CYB* | C3 | N.S | N.S | -0.013 | 0.010 | N.S | N.S |
| 14996 | *MT-CYB* | C3 | 0.023 | 0.024 | 0.041 | 1.191E-03 | N.S | N.S |
| 15040 | *MT-CYB* | C3 | N.S | N.S | N.S | N.S | 0.091 | 0.034 |
| 15045 | *MT-CYB* | C3 | N.S | N.S | N.S | N.S | 0.048 | 0.019 |
| 15092 | *MT-CYB* | C3 | 0.042 | 0.013 | 0.047 | 0.043 | N.S | N.S |

| Genomic Position (bp) | Gene | Complex | Effect Size MM | Significance level (P Value) MM | Effect Size CER | Significance level (P Value) CER | Effect Size STG | Significance level (P Value) STG |
|---|---|---|---|---|---|---|---|---|
| 15199 | *MT-CYB* | C3 | N.S | N.S | N.S | N.S | 0.118 | 0.043 |
| 15275 | *MT-CYB* | C3 | 0.073 | 0.020 | N.S | N.S | 0.113 | 0.039 |
| 15391 | *MT-CYB* | C3 | 0.170 | 0.017 | N.S | N.S | 0.169 | 0.031 |
| 15437 | *MT-CYB* | C3 | N.S | N.S | -0.038 | 0.033 | N.S | N.S |
| 15557 | *MT-CYB* | C3 | 0.040 | 0.011 | N.S | N.S | N.S | N.S |
| 15590 | *MT-CYB* | C3 | 0.084 | 0.027 | 0.090 | 0.035 | N.S | N.S |
| 15591 | *MT-CYB* | C3 | N.S | N.S | N.S | N.S | 0.057 | 0.018 |
| 15595 | *MT-CYB* | C3 | 0.094 | 0.040 | N.S | N.S | N.S | N.S |
| 15596 | *MT-CYB* | C3 | 0.025 | 8.795E-03 | N.S | N.S | 0.046 | 0.010 |
| 15761 | *MT-CYB* | C3 | N.S | N.S | -0.030 | 0.030 | N.S | N.S |
| 15811 | *MT-CYB* | C3 | N.S | N.S | 0.194 | 0.043 | N.S | N.S |
| 15812 | *MT-CYB* | C3 | N.S | N.S | -0.028 | 3.611E-03 | N.S | N.S |
| 15927 | *MT-TT* | tRNA | N.S | N.S | N.S | N.S | 0.070 | 0.045 |
| 16083 | *D-LOOP* | NCR | 0.208 | 0.034 | N.S | N.S | N.S | N.S |
| 16095 | *D-LOOP* | NCR | 0.212 | 0.010 | N.S | N.S | N.S | N.S |
| 16360 | *D-LOOP* | NCR | N.S | N.S | 0.184 | 0.023 | N.S | N.S |

***Figure 5.3: DNA methylation levels for significant DMPs associated with age present in both STG and CER.*** *Shown are DNA methylation levels for the two age-associated DMPs that were present in the STG (**A**) and CER (**B**), which resided at 4463bp (purple: R=0.720), and 1597bp (green: R=0.557).*

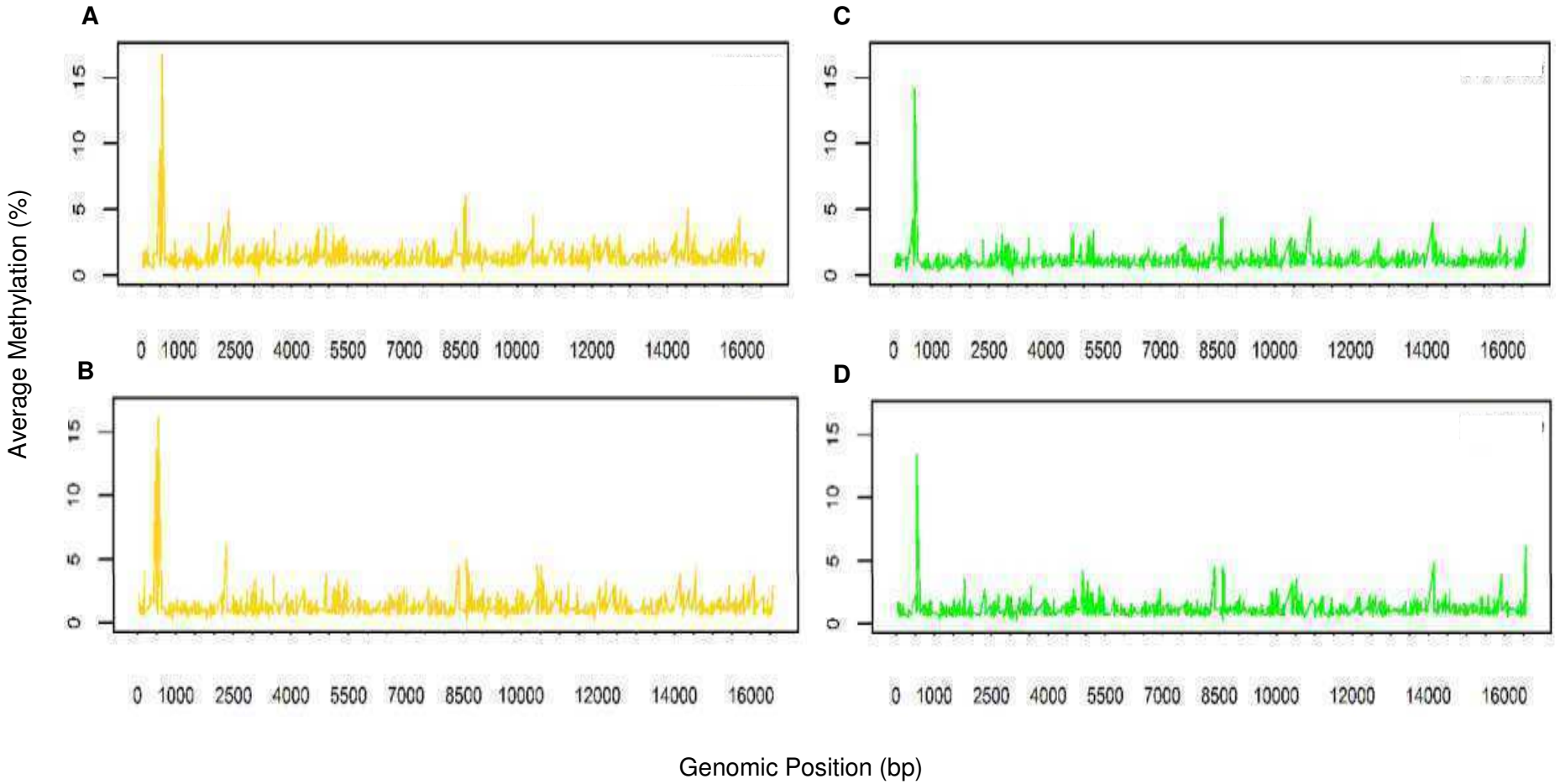


*5.4.6. Conservation of Methylation Levels across the Mitochondrial Genome*

Speculation over the validity of mtDNA methylation has been rife in the field of mitochondrial epigenetics, with studies indicating mtDNA methylation levels are non-biologically across specific regions of the genome (70). We were therefore interested to see how much variability could be observed between mtDNA methylation across the genome. We identified a conserved pattern of mtDNA methylation across the genome. By measuring the standard deviation at each site, we looked to determine the variation in methylation amongst individuals. In total, 496 sites had a standard deviation less than 0.5%, whilst only 104 sites were identified with a standard deviation greater than 1.0%. As such, a relatively conserved pattern of mtDNA methylation between individuals was identified across the mitochondrial genome (**Figure 5.4**).

By looking at methylation levels across each locus and each individual, we identified a total of 576 sites (approximately 4.68% of total number of sites in the cohort) that had a mtDNA methylation level greater than 5%. To determine whether these increases were associated with age, sex or tissue type, a Fisher's exact test for enrichment was performed for each. However, no significant enrichment for relatively high levels of methylation being associated to any of the co-variates in this study could be identified in this study.

***Figure 5.4: Genome-wide plots of DNA methylation.*** *The average level of methylation at each site was calculated for each sex and tissue. Shown are average DNA methylation levels across ChrM for (**A**) Males in the CER, (**B**) Males in STG, (**C**) Females in CER and (**D**) Females in STG.* A separate plot for each sex and tissue was then created to show the average percentage of DNA methylation.



## 5.5. Discussion

Here, we present the first genome-wide map of mtDNA methylation, at single nucleotide resolution, in mitochondria isolated from post-mortem human brain tissue. Previous studies have investigated changes across the mitochondrial epigenome using publicly available data (63,70,252). However, these studies have been limited by low sequencing depth, the use of non-isolated mtDNA, and in the case of MeDIP-Seq, relatively low resolution. Hong et al., used publicly available BS-Seq data to show that mtDNA methylation does not exist at biologically relevant levels (<1%). However, an ability to determine this conclusively, given the multi-copy nature of the mitochondrial genome, as well as the relatively low sequencing depth derived from non-enriched BS-Seq, may have led to a gross underrepresentation of true mtDNA methylation levels. In our study, though average levels of mtDNA methylation across the 869 sites interrogated were low, (STG mean = 1.239%, CER mean = 1.230%), there was considerable variability in methylation across the genome of each individual. Interestingly, a site of the *D-LOOP*, 545bp, was found to have relatively high levels of mtDNA methylation, with an average of 15.3% and 14.6% methylation in the STG and CER respectively. This adds further evidence to a recent study that identified high levels of mtDNA methylation across the mitochondrial *D-LOOP* (71), potentially highlighting a region of the mitochondrial genome which could have an important impact on the organelle's function.

Importantly, the enrichment process seemed to strongly affect the methylation levels of STG, whereas the methylation differences between enriched and non-enriched CER samples were less obvious. It is important to note that this

analysis was only performed on a single sample for each tissue type and a more extensive analysis should be conducted in the future with more individuals. Despite this, given that a number of probes in both tissues were still affected, our results suggest that mtDNA enrichment prior to NGS library preparation could be very important to the study outcome.

Establishing the extent to which confounders such as age, tissue type and sex play a role in ncDNA methylation studies has led to more robust and reliable analyses in epigenetic studies (191). However, previous studies into mtDNA methylation have not studied the genome-wide effect of these factors. Our preliminary analysis identified many nominally significant DMPs for all of the covariates, with five DMPs also passing a Benjamani-Hochberg correction for multiple testing when examining the effect of sex. Of these five, two also passed a more strict Bonferroni correction. The variation between these two results is likely due a reportedly high correlation that exists between nearby DNA methylation sites across the genome (259), suggesting that sites of DNA methylation are not truly independent, which is assumed by Bonferroni corrections. Furthermore, our post-hoc analysis in the separate tissues also demonstrated the same direction of effect for the sex-associated DMPs we identified within *MT-ATP6* and *D-LOOP*. The small sample size of our study, as well as the lack of functional work carried out on these DMPs makes it difficult to ascertain the physiological relevance of the differences. However, their identification in the *D-LOOP,* a region known to regulate mitochondrial transcription and replication, could possibly effectuate changes in overall mitochondrial output. Additionally, the finding that within the STG, there is an

enrichment for hypermethylation in females and that in both STG and CER, there is an enrichment for hypermethylation with increasing age, highlights the need for carefully controlled epidemiological studies in the future.

Age related changes have also been identified in other mtDNA methylation studies (74,260). Hypomethylation in ncDNA has been frequently attributed to increasing age (261) and age-associated hypermethylation has therefore been suggested to play a role in age–related disorders (262). As such, it is of real interest to investigate why mtDNA methylation tends to increase across the genome with age. If it is to be assumed that mtDNA methylation exists at a relatively low level across the mitochondrial genome, then increases in mtDNA methylation could be a sign of dysregulation within the mitochondria. It is therefore possible that this general increase across the genome could be attributed to the increase in mitochondrial dysfunction seen during ageing (176,263), potentially by altering regulation of mtDNA expression. Despite this, few functional studies have investigated the effect of changes in mtDNA methylation, where mtDNA hypermethylation was shown to be associated with apoptosis in diabetic retinopathy (86). However, to date, no study has looked at the effect of multiple, site-specific mtDNA methylation changes occurring simultaneously.

The ubiquitous nature of mitochondrial distribution throughout the body, except for red blood cells, highlights the importance of maintaining correct mitochondrial function throughout healthy ageing. Despite levels of mtDNA methylation being relatively low in this study in comparison to ncDNA

methylation, a recent study identified that those individuals with increased mtDNA methylation levels at a site within *MT-RNR1* were associated with increased mortality risk when followed up nine years later (74). As such, low levels of mtDNA methylation could be due to the averaging effect of analysing the mitochondrial epigenome across mitochondria in multiple states and as such, underestimate true levels existing within mitochondria.

Further, studies have also shown that mtDNA methylation changes at specific sites in the *D-LOOP* negatively correlate with changes in gene expression. One study demonstrated that increases in *D-LOOP* methylation were associated with reduced *MT-CYB*, *MT-ND6*, and *MT-CO2* transcript levels (86), whilst another showed that reduced *D-LOOP* methylation was associated with increased *MT-ND2* expression in colorectal cancer tumour tissue (75). Taken together, this suggests that mtDNA methylation does have a functional role within the organelle. Further, whilst previous studies have demonstrated that mtDNA methylation is typically very low, with values of approximately 0.5 – 1.0% (33); we have identified several cytosines with DNA methylation levels greater than 10% across the mitochondrial genome.

Our study does however, have several limitations. Most importantly, whilst age related hypermethylation changes were observed in this cohort, given the small sample size, most ages were only represented once. Furthermore, we only investigated a relatively narrow range of ages, with all individuals between the ages of 77 and 87 years. It is therefore possible that the effects of age on DNA methylation merely represent inter-individual differences. However, the enrichment for hypermethylated DMPs associated with increasing age make

this less likely. Our study also does not attempt to functionally characterise nominated DMPs, for example to look at the correlation between mtDNA methylation and gene expression levels. Our primary aim was instead to define a methodology capable of analysing mtDNA methylation at high resolution across the mitochondrial genome and, importantly, to provide strong evidence to support the existence of mtDNA methylation in the brain.

## 5.6. Conclusion

Our study provides the first genome-wide analysis of mtDNA methylation at single base resolution, providing a map of mtDNA methylation across human brain as well as identifying differences in mtDNA methylation between different tissue types, sex and age. This not only highlights a potential biological importance for mtDNA methylation but also highlights the need to control for these factors in future studies. Further, the identification of these traits, particularly the enrichments identified to be associated with age and sex, suggest that mtDNA methylation has potential functional relevance, giving evidence to its existence as an independent mark. Interestingly our data also alluded to a multitude of relatively highly methylated positions across the mitochondrial genome; however, there was no enrichment between these increased levels and each of the covariates analysed. Detecting these variable sites in a larger cohort may however, reveal clues to their biological relevance.

# Chapter 6. Interrogating the Role of MtDNA methylation in AD

## 6.1. Aims

1. To validate sex- and age-specific DMPs previously identified in Chapter 5.
2. To determine whether altered mtDNA methylation patterns are seen in AD, using Braak stage as a measure of neuropathology.
3. To determine whether altered nuclear-encoded mitochondrial gene DNA methylation is associated with AD, using Braak stage as a measure of neuropathology.
4. To correlate mtDNA methylation differences associated with AD with nuclear-encoded mitochondrial gene methylation differences associated with AD in matched samples.

## 6.2. Introduction

The brain is an organ that represents only 2% of the average human's body weight. However despite this, the brain accounts for approximately 20% of the body's total oxygen output (264). This discrepancy between weight and consumption is largely due to the neuronal population and the requirement to maintain an environment critical for the establishment of regular and correct action potential firing (265). Given that mitochondria provide the main source of total cellular ATP, the requirement for the critical refinement and correct regulation of mitochondrial homeostasis is of upmost importance for the physiological health of the brain.

AD is a complex disease characterised by the accumulation of Aβ plaques, tangles of hyperphosphorylated tau, and mitochondrial dysfunction (266). The neuropathological spread of tangles throughout AD brain has been well documented; in the earliest stages of the disease these are largely isolated to the ECX and hippocampus, whilst by the end stages of the disease these have spread throughout the CTX (191). However, at present little is known about the exact molecular mechanisms that may be driving disease progression.

In recent years, it has been hypothesised that AD has an epigenetic component (191). In 2014, the first EWAS in AD brain were published. These studies demonstrated robust and reproducible changes in ncDNA methylation at several loci in AD CTX (95,198), implicating a potential role for epigenetic mechanisms in disease aetiology. However, one caveat of these studies is that the platform used (Illumina Infinium 450K Methylation array) has no coverage of the mitochondrial genome. As such, no genome-wide study to date has investigated whether the mitochondrial dysfunction that is seen in AD may be driven by mitochondrial epigenetic changes.

We have previously hypothesised that, given that mitochondrial dysfunction and epigenetic modifications have both been identified in AD (95,198,267), mtDNA methylation may have an important role in AD pathology (1). Furthermore, a recent study has identified epigenetic changes in two regions of the mitochondrial genome that are significantly associated with Braak stage, a standardised measure of

disease pathology in AD, using a candidate based gene approach (BS-pyrosequencing) (200). Whilst this has added weight to our original hypothesis, a genome-wide, high resolution study of mtDNA methylation in AD has yet to be undertaken. In Chapter 5 we presented a new method for studying mtDNA methylation using targeted BS-Seq, which captures the mitochondrial methylome at single-base resolution. Whilst there was little variation in the methylation levels observed in our pilot study, an enrichment for hypermethylation in females for sex-associated DMPs in the STG was observed, with a similar trend in the CER. Furthermore, we also identified an enrichment for hypermethylated DMPs associated with increasing age in both the STG and CER.

Here, we looked to identify differences in mtDNA methylation in AD brain samples using the targeted BS-Seq method we previously developed in Chapter 5. Furthermore, given that previous studies have highlighted a complex interplay between the nucleus and mitochondria on a genetic and proteomic level (1,267,268), we also hypothesised that this may occur at an epigenetic level. We therefore analysed DNA methylation levels of nuclear-encoded mitochondrial genes, which encode for proteins that translocate into the mitochondria (herein referred to as nuclear-mitochondrial genes), using data generated on the same samples using the Illumina 450K array, to investigate this further.

## 6.3. Methods

### *6.3.1. Sample Demographics for Mitochondrial Isolations*

200mg of ECX brain tissue was obtained from 96 donors archived in the LBB, for mitochondrial isolation. From a subset of these donors, we also had access to ncDNA methylation data generated on the Illumina 450K array (N=49). Samples were selected to have varying degrees of AD pathology (Braak Stage 0-VI) and all AD patients were over the age of 65 at diagnosis, see **Table 6.1** for sample demographics.

***Table 6.1: Sample demographics of samples used for mtDNA analysis in Chapter 6.*** *Shown are Braak stage, age at death, sex, sequencing depth and mapping efficiency (from our study) and finally whether samples had available matched Illumina 450K array data. Data is only shown for samples that passed QC*

| Individual | Braak Stage | Age (at death) | Sex | Sequencing Depth | Mapping Efficiency (%) | Illumina 450K array data available |
|---|---|---|---|---|---|---|
| 1 | VI | 73 | Female | 33084 | 81.10 | Yes |
| 2 | VI | 83 | Male | 1780 | 79.90 | Yes |
| 3 | III | 91 | Male | 1220 | 76.70 | Yes |
| 4 | V | 88 | Male | 1089 | 79.00 | No |
| 5 | 0 | 84 | Male | 1659 | 81.70 | Yes |
| 6 | VI | 77 | Male | 242 | 59.50 | Yes |
| 7 | IV | 81 | Male | 2110 | 84.80 | Yes |
| 8 | II | 76 | Female | 2086 | 77.20 | Yes |
| 9 | V | 85 | Female | 2034 | 79.90 | Yes |
| 10 | V | 65 | Male | 1778 | 76.60 | Yes |
| 11 | VI | 76 | Male | 2211 | 71.40 | Yes |
| 12 | IV | 79 | Male | 2032 | 71.40 | No |
| 13 | V | 85 | Female | 2239 | 84.60 | Yes |
| 14 | V | 89 | Female | 3300 | 83.70 | Yes |
| 15 | V | 88 | Male | 2397 | 83.00 | Yes |
| 16 | 0 | 90 | Male | 701 | 66.70 | Yes |
| 17 | VI | 65 | Female | 2271 | 81.50 | Yes |
| 18 | II | 80 | Male | 2212 | 81.80 | Yes |
| 19 | 0 | 66 | Female | 3097 | 81.40 | Yes |
| 20 | V | 90 | Female | 1577 | 65.20 | Yes |

| Individual | Braak Stage | Age (at death) | Sex | Sequencing Depth | Mapping Efficiency (%) | Illumina 450K array data available |
|---|---|---|---|---|---|---|
| 21 | V/VI | 70 | Female | 1642 | 83.20 | Yes |
| 22 | IV | 79 | Male | 4157 | 76.60 | Yes |
| 23 | 0 | 77 | Female | 17485 | 43.90 | Yes |
| 24 | V | 89 | Female | 1860 | 79.70 | Yes |
| 25 | III | 98 | Female | 1561 | 82.40 | No |
| 26 | VI | 83 | Male | 1490 | 85.90 | Yes |
| 27 | IV | 86 | Male | 3762 | 78.90 | Yes |
| 28 | VI | 84 | Male | 1166 | 85.40 | Yes |
| 29 | II | 90 | Female | 1550 | 79.50 | Yes |
| 30 | II | 67 | Female | 3315 | 76.30 | Yes |
| 31 | 0 | 80 | Female | 969 | 83.50 | No |
| 32 | II | 89 | Female | 1151 | 79.60 | Yes |
| 33 | V | 97 | Female | 2111 | 86.40 | No |
| 34 | V | 97 | Female | 311 | 86.80 | Yes |
| 35 | VI | 78 | Male | 624 | 85.90 | Yes |
| 36 | III | 97 | Male | 3277 | 83.30 | Yes |
| 37 | II | 82 | Female | 2168 | 84.10 | No |
| 38 | VI | 80 | Male | 1892 | 81.60 | Yes |
| 39 | VI | 85 | Male | 3602 | 81.60 | Yes |
| 40 | II | 85 | Female | 4558 | 83.80 | Yes |
| 41 | I | 55 | Female | 3583 | 84.50 | No |
| 42 | VI | 87 | Male | 10291 | 78.30 | Yes |
| 43 | 0 | 63 | Male | 2025 | 86.20 | Yes |
| 44 | I | 67 | Male | 523 | 85.10 | Yes |
| 45 | V | 89 | Male | 18057 | 63.70 | Yes |
| 46 | III | 92 | Female | 3130 | 84.20 | Yes |

| Individual | Braak Stage | Age (at death) | Sex | Sequencing Depth | Mapping Efficiency (%) | Illumina 450K array data available |
|---|---|---|---|---|---|---|
| 47 | VI | 72 | Male | 1438 | 81.20 | No |
| 48 | III | 88 | Female | 832 | 82.40 | Yes |
| 49 | III | 91 | Female | 755 | 84.10 | No |
| 50 | VI | 74 | Male | 3392 | 82.20 | Yes |
| 51 | VI | 81 | Female | 597 | 80.60 | Yes |
| 52 | 0 | 69 | Female | 1073 | 78.00 | Yes |
| 53 | IV | 92 | Male | 2071 | 79.70 | Yes |
| 54 | VI | 78 | Female | 2242 | 75.70 | Yes |
| 55 | VI | 85 | Male | 11332 | 79.10 | Yes |
| 56 | II | 74 | Male | 1761 | 76.20 | No |

### *6.3.2. Mitochondrial Isolation*

Mitochondria were isolated from frozen, post-mortem brain tissue according to the protocol outlined in Section 2.3.4. MtDNA was further enriched through the use of a custom library designed to amplify the mitochondrial genome (Agilent Technologies, California, U.S.A.) as described in Section 2.3.6.

### *6.3.3. Custom Capture of the Mitochondrial Epigenome*

To capture the mitochondrial epigenome, a custom library of RNA baits (Agilent, California, U.S.A.: Design ID 0687721) were designed to provide 100% coverage of the genome at 5X tiling density, as shown in Chapters 2 and 5. Isolated mtDNA extracted from frozen, post-mortem brain tissue was subjected to an adapted version of the Agilent 1µg Methyl-Seq protocol (see Section 2.3.6). After generating indexed libraries for each sample, samples were then pooled in equimolar concentrations and sequenced using the Illumina HiSeq 2500 by Exeter Sequencing Service (Exeter, U.K.), who returned 96 fastq file pairs.

### *6.3.4. Analytical Workflow for MtDNA Methylation*

After sequencing, 100bp paired reads were de-multiplexed and following quality assessment using FastQC, reads were trimmed for adaptor content using TrimGalore (253). Using the same package, we also trimmed the first base pair from the 3' end of both reads. Trimmed files were then aligned to human mitochondrial genome, rCRS, through Bismark, using Bowtie 1 (204). On average,

643,143 reads mapped per sample, providing an average sequencing depth of 3439 times. The average mapping efficiency for each sample to the mitochondrial genome was 78.28% (S.D = 8.81). Mapped reads were then de-duplicated using the deduplicate_bismark function before CpG and non-CpG methylation was called by bismark_methylation_extractor. Given the variation in sequencing depth between samples, 40 samples were removed in sample QC. Reasons for removal of samples included low mapping efficiency (N=2), and in the remaining samples, over 30% of total sites receiving less than 15 calls for methylation/unmethylated cytosine. Further, of the 869 cytosines profiled, 601 were called in the majority (75%) of all remaining samples. Only these 601 cytosines were considered in later statistical analysis. Of note, samples that were primarily affected were of low yield after RNA-DNA hybridisation during the Agilent protocol, suggesting that the quality of DNA in these samples may have been reduced.

All statistical analyses were carried out in R (v3.3.0) (254). Briefly, a linear regression was used to define the effect of each co-variate: age, sex and Braak stage, on mtDNA methylation. For this analysis, Braak stage was grouped into three bins, "low", "medium" and "high", corresponding to Braak stage 0-II, III-IV and V-VI respectively. This method was used to improve power and reduce the differences between group sizes in the analysis. Nominal significance was deemed $P<0.05$. To correct for multiple testing, a Bonferroni correction threshold was used ($P=8.319E-05$).

*6.3.5. DNA Methylation Analysis of Nuclear-Mitochondrial Genes*

NcDNA methylation levels for ~485,000 loci across the human genome were available from ECX samples from two cohorts of individuals (**Table 6.2**) with varying degrees of AD pathology from other experiments performed within our laboratory (N=217 in total) (198,269). From this data, we focussed our analyses on nuclear-mitochondrial probes, which we defined as probes covering genes encoding for proteins that translocate into the mitochondria (N=13,170). These genes were identified as nuclear-mitochondrial genes in MitoCarta, a mitochondrial database (270). Probes corresponding to these gene names were then extracted from each sample file for later use. For each ECX cohort we performed a linear regression analysis on all nuclear-mitochondrial probes. These linear regressions looked to associate ncDNA methylation with Braak stage bins (as previously defined), whilst controlling for age, sex and cellular proportion (as estimated by the R package CETS (271)).

Probes that were associated with Braak stage in both cohorts, with the same direction of effect, were used for a subsequent meta-analysis using metaphor, a package in R (272). From each regression analysis, the respective standard errors and methylation differences of these probes were extracted and used as parameters in the meta-analysis. This gave the overall effect size and significance level of each individual probe across the two cohorts. To correct for multiple testing, a Bonferroni correction for ncDNA was used (P=3.798E-06).

***Table 6.2: Summary sample demographics for nuclear-mitochondrial gene analysis.***

| Braak stage | Cohort 1 | | | Cohort 2 | | |
|---|---|---|---|---|---|---|
| | **0-II** | **III-IV** | **V-VI** | **0-II** | **III-IV** | **V-VI** |
| Total Number | 29 | 18 | 66 | 24 | 18 | 62 |
| Females/ Males | 16/13 | 11/7 | 40/26 | 12/12 | 11/7 | 39/23 |
| Age (Mean +/- SD) | 77.62 +/- 12.57 | 88.50 +/- 5.60 | 85.41 +/- 8.07 | 80.33 +/- 9.75 | 88.50 +/- 5.60 | 85.59 +/- 8.16 |

Of the 56 mtDNA samples we analysed in our study (**Table 6.1**), a subset of 49 had nuclear-mitochondrial probe methylation data available from the 450K array (269) (noted on **Table 6.1**). An analysis was carried out to determine the significance of correlations between ncDNA and mtDNA loci that were identified as being significantly differentially methylated with respect to Braak stage bins by creating a correlation matrix in R (v3.3.0) (254) and using it to plot a heatmap. To determine whether probes (or DMPs) in one genome were significantly correlated with more probes (or DMPs) in the other genome than expected by chance, the frequency and distribution of significant correlations for each was calculated. A Fisher's exact test for enrichment was then performed, analysing the significance of multiple significant correlations within one locus, when compared to the total

number of significant, and non-significant, correlations identified throughout the dataset.

## 6.4. Results

### *6.4.1. MtDNA Methylation Shows Sex-Specific Patterns*

We identified 23 nominally significant DMPs associated with sex (**Table 6.3**). In Chapter 5 we previously identified five sex-specific DMPs which survived Benjamani-Hochberg, of which two also survived Bonferroni correction. In the current study, only three of these five loci passed QC, with the two passing Bonferroni correction in Chapter 5 not passing our selection criteria. Of all sex-specific DMPs identified in both STG and CER, only one, at 15,761bp (*MT-CYB*) was identified in the ECX. However, we also identified that three of the 23 DMPs that were nominally significantly differentially methylated in the ECX were also in the STG, but not CER. Our previous analysis in Chapter 5 also showed that there was a significant enrichment for hypermethylation in females for sex-associated DMPs in the STG. Of the 23 nominally significant sex-associated DMPs we identified in the current study, 22 were hypermethylated in females, with a Fisher's Exact test showing a significant enrichment of hypermethylation in females for sex-associated DMPs (P = 1.891E-03; Odds Ratio = 10.390), again confirming that sex-associated DMPs tend to be hypermethylated in female CTX.

*6.4.2. Conserved Patterns of MtDNA Methylation Exist between Cohorts*

In Chapter 5, we also showed a relatively conserved, low level of mtDNA methylation present across the mitochondrial genome, with methylation levels typically between 1-3%. Interestingly in the current study, which utilises a larger cohort, a similar pattern was also identified, with an average methylation level of 0.95%. However, there were on average 10 sites within each individual that demonstrated relatively high levels (>5%) of mtDNA methylation. To test whether these relatively high levels of methylation were associated with any of the covariates used in our analysis, a Fishers exact test for enrichment was performed for each. However, the frequency of relatively high levels of mtDNA methylation was not enriched for associations with sex, age, or disease.

***Table 6.3: Methylation differences and corresponding P values for sex-associated DMPs in ECX.*** *Also shown are the methylation levels and P values for these loci in the STG and CER from the data generated in Chapter 5. Hypermethylation in males is indicated by a positive methylation value. DMPs that were present in more than one tissue are highlighted in bold. Effect sizes are determined as the percentage methylation differences between the two sexes. Abbreviations: Non-significant (N.S), Complex I (C1), Complex III (C3), Complex IV (C4), Complex V (C5), Non-coding regions (NCR).*

| **Genomic Information** | | | **ECX** | | **STG** | | **CER** | |
|---|---|---|---|---|---|---|---|---|
| **Genomic Position (bp)** | **Gene** | **Complex** | **Effect Size** | **P Value** | **Effect Size** | **P Value** | **Effect Size** | **P Value** |
| 1488 | *MT-RNR1* | rRNA | 1.579 | 0.045 | 0.251 | N.S | 0.034 | N.S |
| 1489 | *MT-RNR1* | rRNA | -0.092 | 0.020 | -0.264 | N.S | -0.095 | N.S |
| 1562 | *MT-RNR1* | rRNA | -0.145 | 0.005 | 0.041 | N.S | -0.139 | N.S |
| 3010 | *MT-RNR2* | rRNA | -0.144 | 0.030 | -0.268 | N.S | -0.217 | N.S |
| 3530 | *MT-ND1* | C1 | -2.159 | 0.001 | 0.432 | N.S | -0.220 | N.S |
| 4996 | *MT-ND2* | C1 | -0.147 | 0.016 | -0.134 | N.S | -0.067 | N.S |
| 6021 | *MT-CO1* | C4 | -0.073 | 0.020 | -0.225 | 0.054 | 0.092 | N.S |
| **6164** | ***MT-CO1*** | **C4** | **-0.907** | **0.024** | **-1.461** | **0.032** | 0.057 | N.S |
| 6618 | *MT-CO1* | C4 | -0.146 | 0.019 | -0.090 | N.S | -0.134 | N.S |
| **7161** | ***MT-CO1*** | **C4** | **-0.112** | **0.005** | **-0.381** | **0.020** | 0.048 | N.S |
| **7332** | ***MT-CO1*** | **C4** | **-0.158** | **0.007** | **-0.213** | **0.048** | 0.177 | N.S |
| 7814 | *MT-CO2* | C4 | -0.100 | 0.006 | -0.088 | N.S | 0.050 | N.S |
| 7850 | *MT-CO2* | C4 | -0.087 | 0.042 | -0.208 | N.S | -0.049 | N.S |
| 8114 | *MT-CO2* | C4 | -0.144 | 0.035 | -0.359 | N.S | 0.072 | N.S |

| Genomic Information | | | ECX | | STG | | CER | |
|---|---|---|---|---|---|---|---|---|
| **Genomic** | **Gene** | **Complex** | **Effect** | **P Value** | **Effect** | **P Value** | **Effect** | **P Value** |
| *8879* | *MT-ATP6* | *C5* | -0.079 | 0.043 | -0.262 | 0.069 | -0.096 | N.S |
| 8959 | *MT-ATP6* | C5 | -0.144 | 0.000 | -0.044 | N.S | 0.012 | N.S |
| 9621 | *MT-CO3* | C4 | -0.152 | 0.013 | -0.152 | N.S | 0.066 | N.S |
| 11456 | *MT-ND4* | C1 | -0.094 | 0.011 | -0.234 | N.S | -0.025 | N.S |
| 13531 | *MT-ND5* | C1 | -0.079 | 0.036 | -0.165 | N.S | 0.009 | N.S |
| 15557 | *MT-CYB* | C3 | -0.142 | 0.002 | -0.191 | N.S | -0.106 | N.S |
| **15761** | ***MT-CYB*** | **C3** | **-0.120** | **0.017** | **-0.257** | **0.028** | **-0.491** | **0.001** |
| 16566 | *D-LOOP* | NCR | -0.840 | 0.048 | 0.263 | N.S | -0.562 | N.S |

*6.4.3. MtDNA Methylation Changes are Associated with Age*

We identified 16 age-associated DMPs (**Table 6.4**). Of the 16 identified, nine were found to be positively correlated with age. As in Chapter 5, we ran a Fisher's exact test, to test for enrichment for hypermethylation within age-associated DMPs. However, despite observing an enrichment for these in both the STG and CER in Chapter 5, we did not observe an enrichment in ECX.

***Table 6.4: Methylation differences and P values for age-associated DMPs in ECX.*** *Positive effect sizes correspond to hypomethylation with increasing age. Effect sizes are determined as differences in percentage methylation per year increase in age. Abbreviations: C1 (Complex I), C3 (Complex III), C4 (Complex IV), NA (Not annotated), NCR (non-coding region), rRNA (ribosomal RNA).*

| Genomic Position (bp) | Gene | Complex | Effect Size | P Value |
|---|---|---|---|---|
| 9673 | *MT-CO3* | C4 | 0.006 | 7.00E-03 |
| 1537 | *MT-RNR1* | rRNA | -0.100 | 9.00E-03 |
| 5740 | *NA* | NA | -0.063 | 9.00E-03 |
| 1217 | *MT-RNR1* | rRNA | -0.008 | 0.01 |
| 808 | *MT-RNR1* | rRNA | -0.048 | 0.012 |
| 15355 | *MT-CYB* | C3 | 0.008 | 0.014 |
| 13759 | *MT-ND5* | C1 | 0.013 | 0.017 |
| 499 | *D-LOOP* | NCR | -0.047 | 0.017 |
| 8007 | *MT-CO2* | C4 | 0.006 | 0.019 |
| 6016 | *MT-CO1* | C4 | 0.004 | 0.020 |
| 12888 | *MT-ND5* | C1 | -0.067 | 0.023 |
| 13579 | *MT-ND5* | C1 | 0.007 | 0.024 |
| 10775 | *MT-ND4* | C1 | 0.005 | 0.027 |
| 9329 | *MT-CO3* | C4 | 0.005 | 0.033 |
| 7342 | *MT-CO1* | C4 | -0.008 | 0.046 |
| 14846 | *MT-CYB* | C3 | 0.006 | 0.049 |

*6.4.4. Site-Specific Changes in MtDNA Methylation are Associated with AD Pathology.*

In this study, we used binned Braak stages as a metric to quantify the different stages of AD neuropathology. We identified 29 nominally significant DMPs that were associated with Braak stage (**Figure 6.1; Table 6.5**). Of the 29 DMPs significantly associated with Braak stage, six corresponded to genes within complex I of the ETC, three within Complex III and ten within complex IV. Of the remainder, two were present in DMPs corresponding to tRNAs, four in rRNA and one within the non-coding region (NCR). Despite this, no enrichment was identified for significant DMPs in any ETC complex. The most significant DMP associated with Braak stage bins, (3531bp; *MT-ND1;* P=3.603E-03) can be seen in **Figure 6.1**. However, this did not reach our bonferroni significant threshold of 8.179E-05.

***Figure 6.1: A Manhattan plot of the mitochondrial genome, showing association between DNA methylation and Braak stage.*** *The blue line denotes nominal significance (P=0.05). The genomic position for each gene across the mitochondrial genome is denoted by a grey rectangular box.*

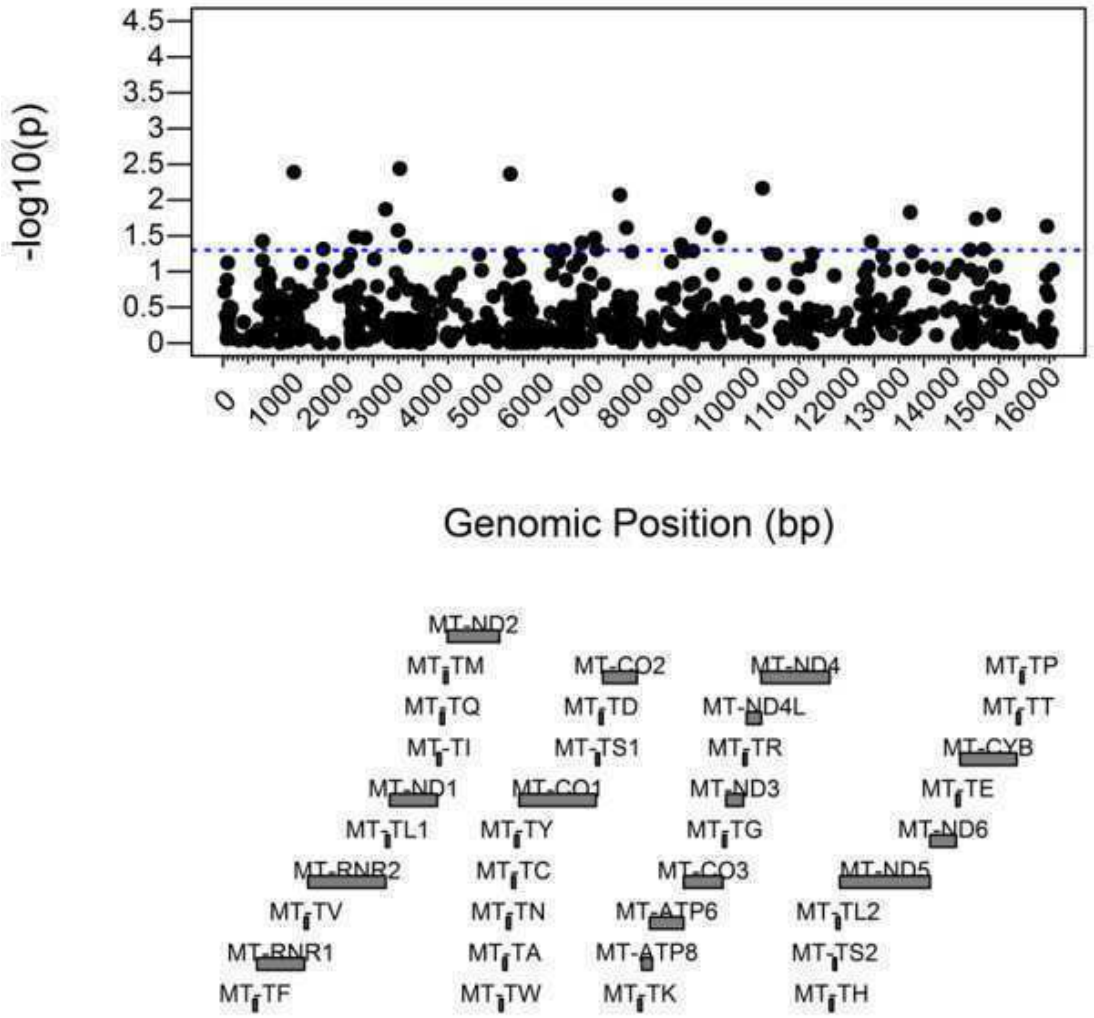


***Figure 6.2: DNA Methylation levels corrected for age and sex of the most significant DMP associated with Braak stage.*** *After adjusting for age and sex, individuals were grouped according to Braak stage bins, low (N=17), Medium (N=11) and High (N=28). A DMP at 3531bp, residing in gene MT-ND1 showed significantly increased DNA methylation with Braak stage (P=3.603E-03).*

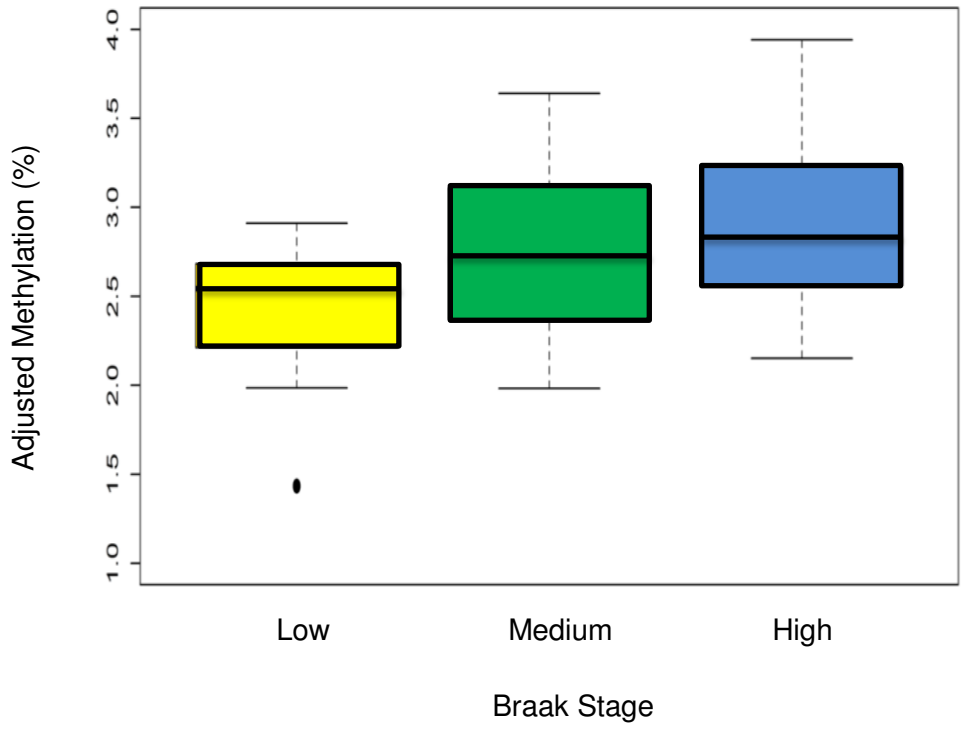


***Table 6.5: DNA Methylation and P values for Braak stage-associated DMPs in ECX.*** *Positive effect sizes correspond to hypomethylation with increasing Braak stage. Effect sizes are determined as differences in methylation percentage per binned Braak stage. Abbreviations: C1 (Complex I), C2 (Complex II), C3 (Complex III), C4 (Complex IV), C5 (Complex V), NA (not annotated), NCR (non-coding region), rRNA (ribosomal RNA), tRNA (transfer RNA).*

| Genomic Position (bp) | Gene | Complex | Effect Size | P Value |
|---|---|---|---|---|
| 5740 | *NA* | NA | 0.787 | 4.00E-03 |
| 1415 | *MT-RNR1* | rRNA | -0.081 | 4.11E-03 |
| 3531 | *MT-ND1* | C1 | 0.227 | 5.08E-03 |
| 10775 | *MT-ND4* | C1 | -0.068 | 6.85E-03 |
| 7928 | *MT-CO2* | C4 | -0.067 | 7.81E-03 |
| 8059 | *MT-CO2* | C4 | 0.456 | 0.013 |
| 15392 | *MT-CYB* | C3 | -0.079 | 0.015 |
| 3248 | *MT-TL1* | tRNA | -0.104 | 0.017 |
| 13726 | *MT-ND5* | C1 | 0.094 | 0.017 |
| 15041 | *MT-CYB* | C3 | -0.057 | 0.017 |
| 9611 | *MT-CO3* | C4 | -1.012 | 0.019 |
| 9574 | *MT-CO3* | C4 | -0.819 | 0.022 |
| 16454 | *D-LOOP* | NCR | -0.605 | 0.025 |
| 2642 | *Mt-RNR2* | rRNA | 0.464 | 0.032 |
| 7419 | *MT-CO1* | C4 | 0.111 | 0.036 |
| 3495 | *MT-ND1* | C1 | -0.687 | 0.038 |
| 7164 | *MT-CO1* | C4 | 0.050 | 0.038 |
| 9145 | *MT-ATP6* | C5 | -0.058 | 0.041 |
| 3643 | *MT-ND1* | C1 | -0.062 | 0.042 |
| 9921 | *MT-TG* | tRNA | 0.043 | 0.043 |
| 12951 | *MT-ND5* | C1 | 0.535 | 0.043 |
| 15200 | *MT-CYB* | C3 | -0.048 | 0.046 |
| 786 | *MT-RNR1* | rRNA | -0.061 | 0.046 |
| 9383 | *MT-CO3* | C4 | 0.963 | 0.047 |
| 14920 | *MT-CYB* | C3 | 0.617 | 0.047 |
| 2844 | *Mt-RNR2* | rRNA | -0.056 | 0.048 |
| 8164 | *MT-CO2* | C4 | 0.413 | 0.048 |
| 6564 | *MT-CO1* | C4 | -0.055 | 0.049 |
| 6808 | *MT-CO1* | C4 | -0.059 | 0.049 |

*6.4.5. Differentially Methylation Nuclear-Mitochondrial Genes are Associated with AD pathology*

We had access to nuclear-mitochondrial DNA methylation data (N=13,170 probes) from the ECX in two separate cohorts of individuals with varying degrees of AD neuropathology. Initially we looked to identify DNA methylation differences associated with Braak stage in each cohort separately, whilst controlling for age, sex and cellular proportions. In total, we identified 714 nominally significant DMPs in cohort 1 and 635 in cohort 2(269)(269). Of these, 44 DMPs overlapped between cohorts with the same direction of effect and were taken forward for a meta-analysis (**Table 6.6**). Of the 44 DMPs identified, one probe, which resided at Chr16: 89,598,950 (cg03169557), located within the Spastic Paraplegia 7 (*SPG7*) gene, was found to pass Bonferroni correction for multiple testing (P=2.430E-06 (**Figure 6.3**)

***Figure 6.3: DNA methylation residuals after correcting for age, sex and cell proportion, of SPG7 the most significant Braak-associated DMP in a nuclear-mitochondrial gene.*** *Shown are data for (**A**) cohort 1 (N=113) and (**B**) cohort 2 (N=104).*

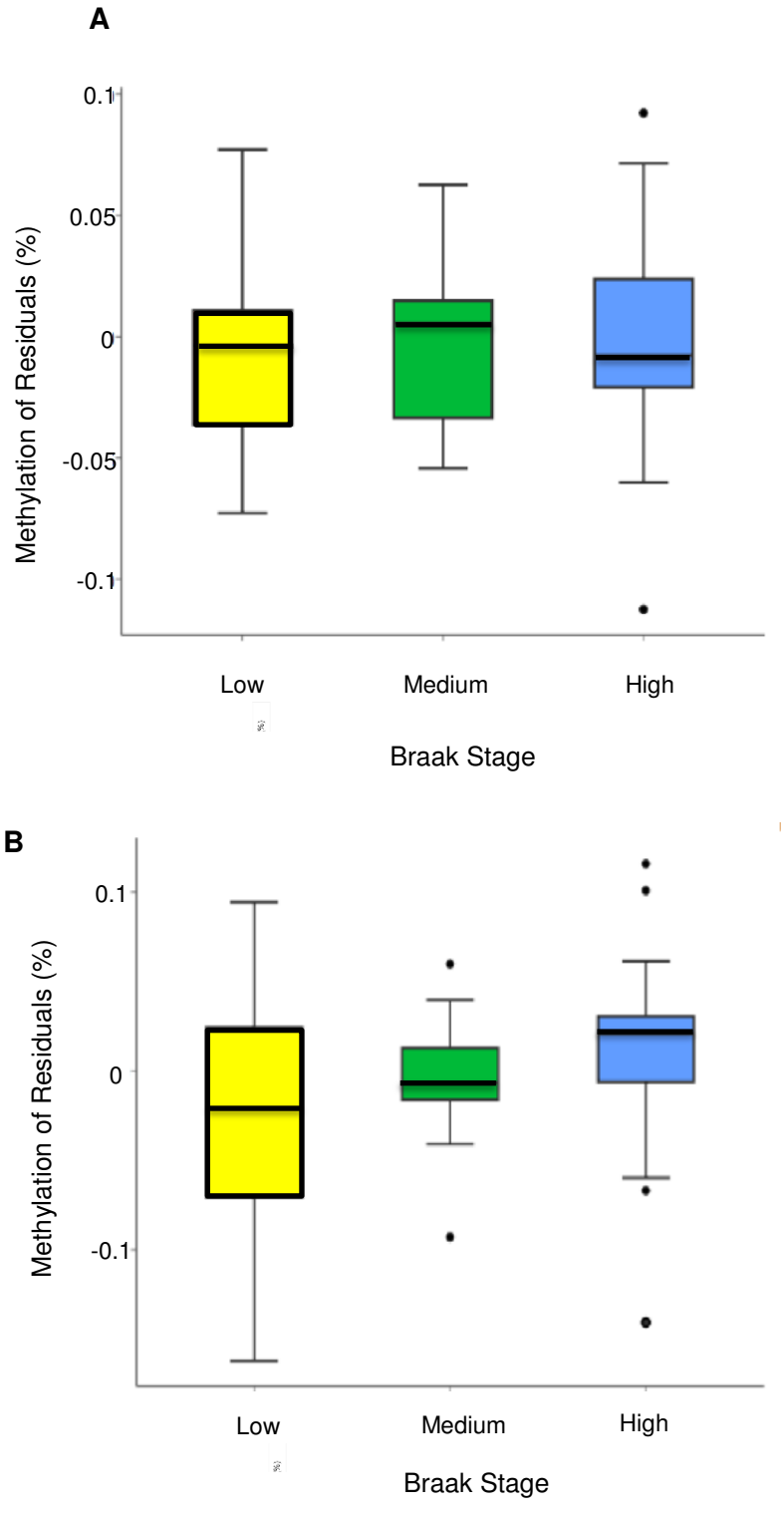


***Table 6.6: Methylation differences and P values for nuclear-mitochondrial DMPs associated with Braak-stage in meta-analysis of ECX.*** *Shown are the effect size and P Values for linear regressions in each ECX cohort separately, in addition to overall P value from meta-analysis. Positive effect sizes denote hypomethylation with increasing Braak stage. Effect sizes are determined as differences in methylation percentage per binned Braak stage Data is ordered by most significant P values in the meta-analysis.*

| Probe | Gene | Meta-analysis | | Cohort 1 | | Cohort 2 | |
|---|---|---|---|---|---|---|---|
| | | Methylation Difference (%) | P Value | Methylation Difference (%) | P Value | Methylation Difference (%) | P Value |
| cg03169557 | *SPG7* | 0.300 | 2.43E-06 | 0.200 | 9.40E-05 | 0.800 | 1.18E-03 |
| cg05242371 | *DMPK* | -0.600 | 1.86E-05 | -0.400 | 0.022 | -0.800 | 1.599E-04 |
| cg14025831 | *PANK2* | 0.500 | 2.58E-05 | 0.400 | 7.15E-04 | 0.600 | 0.016 |
| cg22656126 | *WDR81* | 0.600 | 8.91E-05 | 0.500 | 1.52E-03 | 0.800 | 0.024 |
| cg24547396 | *COMT* | -0.400 | 1.01E-04 | -0.400 | 9.77E-04 | -0.300 | 0.045 |
| cg26182964 | *ECHS1* | -0.400 | 1.12E-04 | -0.400 | 4.27E-03 | -0.500 | 0.012 |
| cg20140950 | *RDH13* | -0.300 | 2.41E-04 | -0.300 | 5.29E-03 | -0.500 | 0.015 |
| cg01200965 | *HK2* | 0.300 | 2.45E-04 | 0.200 | 2.38E-03 | 0.400 | 0.047 |
| cg22011804 | *FASTK* | -0.500 | 3.61E-04 | -0.500 | 6.94E-03 | -0.500 | 0.026 |
| cg12615866 | *MRPL1* | 0.200 | 3.66E-04 | 0.200 | 0.027 | 0.200 | 6.73E-03 |
| cg05191825 | *ISCU* | 0.100 | 3.93E-04 | 0.100 | 0.020 | 0.100 | 9.62E-03 |
| cg07837729 | *MACROD1* | 0.100 | 4.86E-04 | 0.100 | 0.022 | 0.100 | 0.011 |
| cg03611265 | *HEBP1* | 0.700 | 4.94E-04 | 0.600 | 0.034 | 0.800 | 6.46E-03 |
| cg27490193 | *HINT2* | 0.100 | 5.07E-04 | 0.100 | 7.92E-03 | 0.100 | 0.029 |
| cg20938103 | *COX8C* | -0.400 | 5.54E-04 | -0.300 | 5.76E-03 | -0.700 | 0.027 |
| cg02996269 | *PPA2* | -0.200 | 6.19E-04 | -0.300 | 0.015 | -0.200 | 0.018 |
| cg03462975 | *HSD17B8* | -0.300 | 6.96E-04 | -0.300 | 0.022 | -0.200 | 0.015 |

| Probe | Gene | Meta-analysis | | Cohort 1 | | Cohort 2 | |
|---|---|---|---|---|---|---|---|
| | | Methylation Difference (%) | P Value | Methylation Difference (%) | P Value | Methylation Difference (%) | P Value |
| cg08889044 | *SERHL2* | -0.200 | 8.32E-04 | -0.200 | 0.016 | -0.200 | 0.025 |
| cg01192051 | *POLG* | 0.500 | 8.58E-04 | 0.500 | 0.029 | 0.700 | 0.012 |
| cg19191600 | *C6orf136* | -0.300 | 8.76E-04 | -0.300 | 9.69E-03 | -0.400 | 0.043 |
| cg27249304 | *TYSND1* | -0.200 | 9.25E-04 | -0.100 | 0.019 | -0.200 | 0.024 |
| cg15558129 | *LYPLAL1* | -0.200 | 9.51E-04 | -0.200 | 0.025 | -0.200 | 0.019 |
| cg01972394 | *ALDH4A1* | -0.200 | 1.12E-03 | -0.300 | 0.010 | -0.200 | 0.042 |
| cg26215887 | *SLC25A28* | 0.100 | 1.18E-03 | 0.100 | 0.033 | 0.200 | 0.01 |
| cg22920873 | *C7orf55* | 0.200 | 1.28E-03 | 0.200 | 0.041 | 0.200 | 0.016 |
| cg06538141 | *COX4I2* | -0.200 | 1.34E-03 | -0.200 | 0.021 | -0.500 | 7.20E-03 |
| cg15466587 | *MRPS14* | 0.200 | 1.48E-03 | 0.100 | 0.018 | 0.200 | 0.034 |
| cg15035364 | *PMPCB* | -0.100 | 1.57E-03 | -0.100 | 0.022 | -0.100 | 0.035 |
| cg25110782 | *SHMT2* | -0.300 | 1.87E-03 | -0.300 | 0.022 | -0.300 | 0.041 |
| cg07274998 | *OAT* | -0.200 | 1.92E-03 | -0.200 | 0.04 | -0.200 | 0.023 |
| cg25934700 | *SELO* | 0.200 | 2.26E-03 | 0.100 | 0.012 | 0.400 | 0.031 |
| cg19371708 | *GRPEL1* | 0.100 | 2.34E-03 | 0.100 | 0.028 | 0.100 | 0.037 |
| cg23407809 | *WDR81* | 0.300 | 2.55E-03 | 0.200 | 0.032 | 0.400 | 0.027 |
| cg03090204 | *ACSF3* | -0.200 | 2.87E-03 | -0.200 | 0.026 | -0.400 | 0.026 |
| cg26070540 | *COX4I2* | -0.400 | 3.17E-03 | -0.400 | 0.05 | -0.600 | 0.027 |
| cg24266622 | *SLC25A14* | 0.300 | 3.18E-03 | 0.200 | 0.041 | 0.600 | 0.012 |
| cg14192396 | *ALDH18A1* | 0.100 | 3.52E-03 | 0.000 | 0.035 | 0.100 | 0.049 |
| cg09038794 | *MOCS1* | 0.100 | 3.64E-03 | 0.100 | 0.034 | 0.100 | 0.049 |
| cg17518978 | *FARS2* | -0.300 | 4.07E-03 | -0.200 | 0.023 | -0.600 | 0.041 |
| cg01531949 | *IDH3A* | 0.100 | 5.68E-03 | 0.100 | 0.049 | 0.100 | 0.049 |
| cg13716849 | *SARDH* | 0.300 | 6.33E-03 | 0.200 | 0.039 | 0.600 | 0.031 |
| cg06161915 | *SND1* | 0.200 | 7.09E-03 | 0.100 | 0.031 | 0.500 | 0.029 |
| cg17886330 | *ALDH4A1* | -0.200 | 7.10E-03 | -0.200 | 0.044 | -0.500 | 0.031 |

*6.4.6. Nuclear and MtDNA Methylation Patterns are Correlated in Braak-Associated Loci*

Given that the mitochondrion and nucleus have been shown to have a complex interplay at the genomic and proteomic level, with many nuclear genes encoding for proteins responsible for the correct regulation of mitochondrial functioning, we were interested in examining whether there was a correlation between methylation differences within the nuclear and mitochondrial genomes. To this end, we used a correlation matrix to investigate the relationship between the 44 nuclear-mitochondrial Braak-associated probes we identified in our meta-analysis (**Table 6.6**) and the 29 mtDNA Braak-associated mitochondrial DMPs (**Table 6.5**). In total, this identified 55 nominally significant correlations between these mtDNA DMPs and ncDNA loci (**Figure 6.4; Table 6.7**). To determine whether any probes or DMPs were enriched for significant correlations with the other genome, Fisher's exact tests for enrichment were performed (see Section 5.3.5). This identified two mitochondrial DMPs that were enriched for significant correlations with nuclear-mitochondrial DMPs: 3495bp (*MT-ND1*; P=5.825E-05; Odds Ratio = 6.609), and 7164bp (*MT-CO1*; P=0.037; Odds Ratio = 3.027). Further, two nuclear-mitochondrial DMPs were identified as being enriched for significant correlations with mitochondrial DMPs: cg20938103 (*Cytochrome C Oxidase Subunit 8C* (*COX8C*)*;* P=0.041, Odds Ratio = 3.463) and cg25110782 (*Serine Hydroxymethyltransferase 2* (*SHMT2*)*;* P=0.041, Odds Ratio = 3.463).

***Figure 6.4: Correlations were calculated between each mitochondrial DMP and each nuclear-mitochondria DMP associated with Braak stage and plotted as a heatmap.*** *Probes in the nuclear genome (x-axis) are ordered alphabetically whilst DMPs encoded by the mitochondrial genome (y-axis) are ordered by genomic location. To show the strength of correlations a colour scale was used, with dark red corresponding to strongly negative correlations and dark blue being associated with strongly positive correlations.*

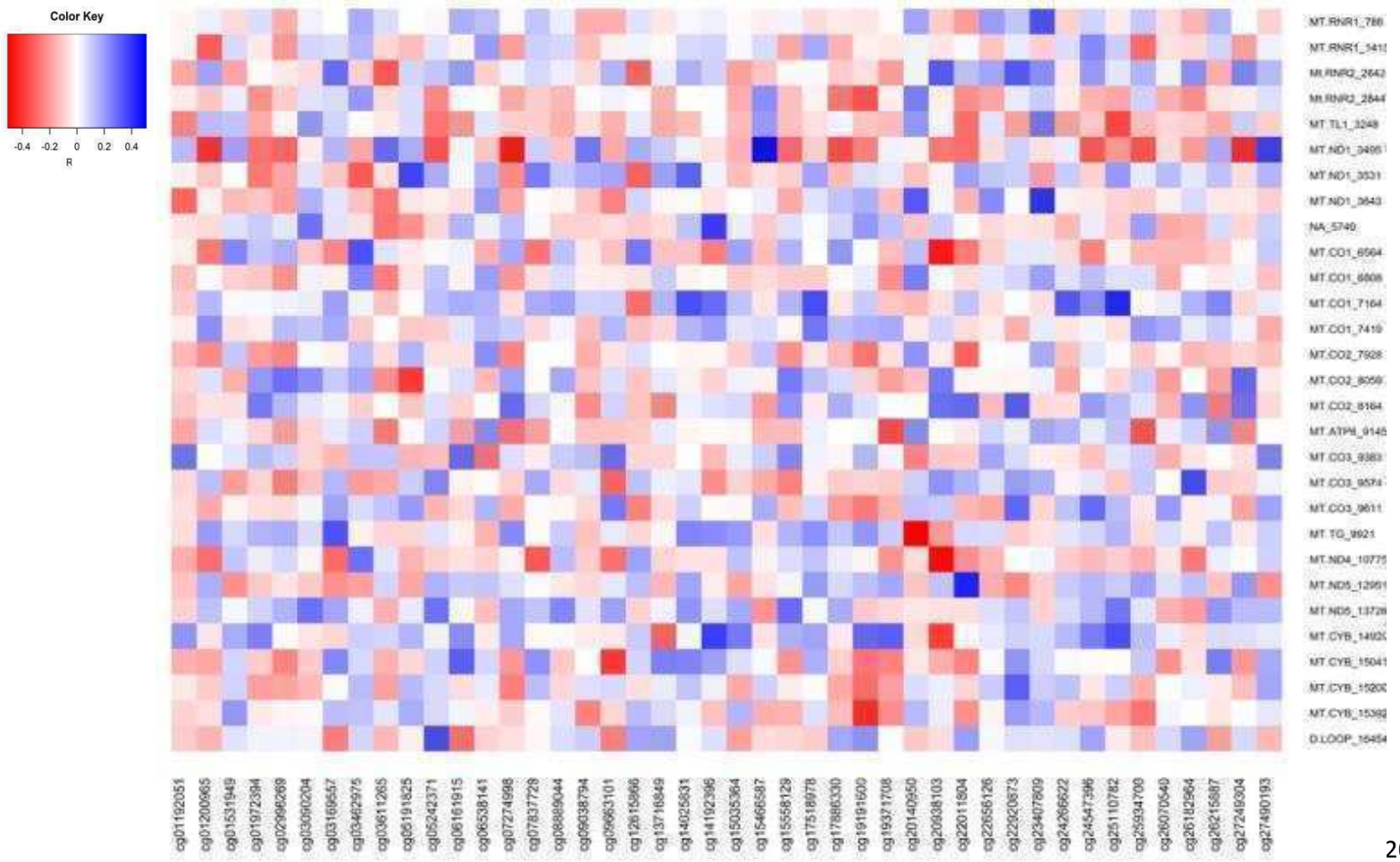


***Table 6.7: List of significant correlations between mitochondrial DMPs and nuclear-mitochondrial DMPs that are significantly associated with Braak stage.*** *Data is ordered by P value.*

| Mitochondrial DMP (bp) | Mitochondrial Gene | Nuclear Probe Correlated | Nuclear Gene | Correlation (r) | P Value |
|---|---|---|---|---|---|
| 9921 | *MT-TG* | cg20140950 | *RDH13* | -0.507 | 3.243E-04 |
| 10775 | *MT-ND4* | cg20938103 | *COX8C* | -0.490 | 5.470E-04 |
| 6564 | *MT-CO1* | cg20938103 | *COX8C* | -0.457 | 1.427E-03 |
| 3495 | *MT-ND1* | cg15466587 | *MRPS14* | 0.480 | 2.664E-03 |
| 7164 | *MT-CO1* | cg25110782 | *SHMT2* | 0.428 | 3.000E-03 |
| 15392 | *MT-CYB* | cg19191600 | *C6orf136* | -0.420 | 3.670E-03 |
| 3495 | *MT-ND1* | cg07274998 | *OAT* | -0.463 | 3.895E-03 |
| 3643 | *MT-ND1* | cg23407809 | *WDR81* | 0.413 | 4.310E-03 |
| 12951 | *MT-ND5* | cg22011804 | *FASTK* | 0.441 | 6.310E-03 |
| 15041 | *MT-CYB* | cg09038794 | *MOCS1* | -0.393 | 6.900E-03 |
| 3495 | *MT-ND1* | cg27249304 | *TYSND1* | -0.426 | 8.510E-03 |
| 3531 | *MT-ND1* | cg05191825 | *ISCU* | 0.376 | 0.010 |
| 5740 | *NA* | cg14912396 | *ALDH18A1* | 0.392 | 0.010 |
| 3248 | *MT-TL1* | cg25110782 | *SHMT2* | -0.370 | 0.011 |
| 3495 | *MT-ND1* | cg01200965 | *HK2* | -0.401 | 0.012 |
| 7164 | *MT-CO1* | cg17518978 | *FARS2* | 0.358 | 0.015 |
| 9145 | *MT-ATP6* | cg19371708 | *GRPEL1* | -0.357 | 0.015 |
| 7164 | *MT-CO1* | cg14025831 | *PANK2* | 0.352 | 0.016 |
| 786 | *MT-RNR1* | cg23407809 | *WDR81* | 0.231 | 0.017 |
| 3495 | *MT-ND1* | cg27490193 | *HINT2* | 0.387 | 0.018 |

| Mitochondrial DMP (bp) | Mitochondrial Gene | Nuclear Probe Correlated | Nuclear Gene | Correlation (r) | P Value |
|---|---|---|---|---|---|
| 6564 | *MT-CO1* | cg03462975 | *HSD17B8* | 0.348 | 0.018 |
| 9921 | *MT-TG* | cg03169557 | *SPG7* | 0.347 | 0.018 |
| 14920 | *MT-CYB* | cg14192396 | *ALDH18A1* | 0.388 | 0.018 |
| 2844 | *MT-RNR2* | cg19191600 | *C6orf136* | -0.341 | 0.020 |
| 9145 | *MT-ATP6* | cg25110782 | *SHMT2* | -0.343 | 0.020 |
| 14920 | *MT-CYB* | cg20938103 | *COX8C* | -0.382 | 0.020 |
| 8059 | *MT-CO2* | cg05191825 | *ISCU* | -0.393 | 0.021 |
| 16454 | *D-LOOP* | cg05242371 | *DMPK* | 0.362 | 0.022 |
| 7164 | *MT-CO1* | cg24266622 | *SLC25A14* | 0.334 | 0.023 |
| 9574 | *MT-CO3* | cg26182964 | *ECHS1* | 0.364 | 0.023 |
| 10775 | *MT-ND4* | cg07837729 | *MACROD1* | -0.330 | 0.025 |
| 3531 | *MT-ND1* | cg03462975 | *HSD17B8* | -0.330 | 0.026 |
| 3643 | *MT-ND1* | cg20140950 | *RDH13* | 0.329 | 0.026 |
| 2642 | *MT-RNR2* | cg22920873 | *C7orf55* | 0.336 | 0.027 |
| 1415 | *MT-RNR1* | cg01200965 | *HK2* | -0.325 | 0.028 |
| 15041 | *MT-CYB* | cg06161915 | *SND1* | 0.324 | 0.028 |
| 3495 | *MT-ND1* | cg17886330 | *ALDHA4* | -0.360 | 0.029 |
| 2642 | *MT-RNR2* | cg03611265 | *HEBP1* | -0.333 | 0.029 |
| 2642 | *MT-RNR2* | cg20938103 | *COX8C* | 0.331 | 0.030 |
| 15200 | *MT-CYB* | cg22920873 | *C7orf55* | 0.321 | 0.030 |
| 14920 | *MT-CYB* | cg25110782 | *SHMT2* | 0.353 | 0.032 |
| 8164 | *MT-CO2* | cg22920873 | *C7orf55* | 0.323 | 0.033 |
| 3531 | *MT-ND1* | cg14025831 | *PANK2* | 0.314 | 0.034 |
| 3531 | *MT-ND1* | cg12615866 | *MRPL1* | -0.312 | 0.035 |
| 3495 | *MT-ND1* | cg25934700 | *SELO* | -0.345 | 0.037 |

| Mitochondrial DMP (bp) | Mitochondrial Gene | Nuclear Probe Correlated | Nuclear Gene | Correlation (r) | P Value |
|---|---|---|---|---|---|
| 3495 | *MT-ND1* | cg05191825 | *ISCU* | 0.166 | 0.038 |
| 7928 | *MT-CO2* | cg22011804 | *FASTK* | -0.307 | 0.038 |
| 3495 | *MT-ND1* | cg24547396 | *COMT* | -0.337 | 0.041 |
| 3643 | *MT-ND1* | cg01192051 | *POLG* | -0.303 | 0.041 |
| 2642 | *MT-RNR2* | cg12615866 | *MRPL1* | -0.312 | 0.042 |
| 10775 | *MT-ND4* | cg03169557 | *SPG7* | -0.297 | 0.045 |
| 15200 | *MT-CYB* | cg19191600 | *C6orf136* | -0.296 | 0.046 |
| 13726 | *MT-ND5* | cg15558129 | *LYPLAL1* | 0.294 | 0.047 |
| 1415 | *MT-RNR1* | cg25934700 | *SELO* | -0.294 | 0.048 |
| 7164 | *MT-CO1* | cg14192396 | *ALDH18A1* | 0.294 | 0.048 |

## 6.5. Discussion

The two major aims of this study were to validate existing findings identified in Chapter 5 and to identify differences in mtDNA methylation associated with AD neuropathology (Braak stage). The strongest associations between mtDNA methylation in Chapter 5 were seen with respect to sex. Unfortunately, due to the strict selection criteria for mitochondrial probes included in the current study, the two probes that passed Bonferroni correction in our initial mixed model, 8722bp and 9055bp (Chapter 5), were removed during QC. Despite this, we again demonstrated a significant enrichment for sex-associated DMPs to be hypermethylated in females compared to males in CTX.

In Chapter 5, we also identified an enrichment for hypermethylated DMPs associated with increasing age. We failed to identify a similar enrichment in the current study. However, it is important to note that in the present study, we had the added covariate of Braak stage in our regression analysis, which given that AD is an age-related disorder, will also be associated with age. Whilst linear regression models can help to better define the role of covariates in determining mtDNA methylation, it remains impossible to account for all variance. As such, the lack of validation for the age-specific DMPs identified in Chapter 5 may be due to the use of a more complex model. To better understand the role of ageing on mtDNA methylation status, it would be important to conduct future studies on healthy non-demented control donors, over a wider range of ages than those investigated in the current study. The ideal experiment would utilise longitudinal samples from the

same individuals to determine mtDNA methylation changes over time. However, this is not possible when using post-mortem human tissue from an inaccessible tissue such as the brain. An alternative experiment may seek to examine longitudinal changes in mtDNA methylation with age, using genetically identical mice.

Of the 29 mitochondrial Braak-associated DMPs identified, six corresponded to genes within Complex I of the ETC. A further ten corresponded to genes within Complex IV of the ETC. Taken together, these findings could indicate dysregulation within the ETC at the methylome level. However despite this, no enrichment for genes within any ETC complex was identified, which could be due to the fact that a large proportion of the assay consists of methylated cytosines within ETC complexes. Interestingly, previous studies have identified that mutations in genes within Complex I of the ETC lead not only to decreases in OXPHOS efficiency (273) but also to an increase in free radical production (274). Both of these changes have been routinely observed in AD (275) and associated with AD pathology (276). A recent study identified that mtDNA methylation differences were also associated with Braak stage in AD patients within both the *D-LOOP* and *MT-ND1* (200). This pyrosequencing study identified an increase in *D-LOOP* methylation and a decrease in methylation in *MT-ND1* at multiple sites across both regions. Whilst Blanch and colleagues did indeed identify three DMPs within *MT-ND1* and one within *D-LOOP*, only 3495bp (*MT-ND1*) and 16,454bp (*D-LOOP*) identified in our study were in the same direction as the general pattern of methylation observed in their study.

Given that robust and reproducible DNA methylation changes in the nuclear genome have previously been identified in several EWAS of AD using the 450K array (198,242,243,277,278), we were interested to determine whether we could identify differential DNA methylation in nuclear-mitochondrial genes. A meta-analysis of two cohorts (94,279) identified 44 DMPs with respect to AD neuropathology. Interestingly one DMP, *SPG7*, reached Bonferroni significance and has been previously shown to be associated with AD from a published EWAS in AD prefrontal CTX (95). *SPG7* is the only gene to have been associated with spastic paraplegia 7, which has been shown to present with a number of neurological factors such as white-matter damage in the frontal lobe (280); ALS, a chronic, neurodegenerative disease (281); and in several models of cerebellar ischemia (282,283). Importantly, downregulation of this gene has been associated with the opening of the mPTP, leading to subsequent $Ca^{2+}$ retention and mitochondrial dysfunction (34), all of which have been identified in AD (284).

Our final analysis aimed to better understand the relationship between methylation patterns in the mitochondrial and nuclear genomes. By investigating the frequency of significant correlations between Braak-associated DMPs we identified in both genomes we identified four DMPs (two mtDNA and two ncDNA), which were significantly enriched for correlations to the other genome. Within the mitochondrial genome, both 3495bp (*MT-ND1*) and 7164bp (*MT-CO1*) were enriched for significant correlations, whilst in the nuclear genome cg20938103 *(COX8C*) and cg25110782 *(SHMT2*) had more significant correlations with mtDNA DMPs than expected by chance. The functional relevance of this cross-talk between the

genomes warrants further investigation. However, it is worth noting that six of nine nuclear-mitochondrial probes correlated to 3495bp (*MT-ND1*) shared an inverse relationship, which could suggest a compensatory mechanism, by which changes in the levels of mtDNA methylation could potentially lead to changes in expression and protein translation could invoke a compensatory reaction in nuclear-mitochondrial genes. Whilst the relationships between these two methylomes at this time can only be hypothesised, the data remains intriguing and studies should seek to investigate this potential complex interplay further.

## 6.6. Conclusions

Despite a strict and conservative selection criterion for loci in the current study, we were first able to validate sex-specific trends of hypermethylation in females observed in Chapter 5. Importantly, our study has identified a number of novel mtDNA loci that were differentially methylated with respect to Braak stage. Functional exploration could provide new insights into this field. Finally, our study was the first to attempt to correlate differences between the nuclear and mitochondrial methylomes, pointing to the possibility of a new layer of complex regulation between the two and potentially, an additional way in which to regulate mitochondrial functionality.

# Chapter 7. General Discussion

## 7.1. Introduction

Throughout this chapter, I aim to summarise the key findings made throughout each chapter of the thesis and the limitations present within these studies. I will then look to discuss and address the limitations when studying the field of mtDNA methylation, with emphasis on how my research has aimed to tackle some of these issues. Finally, I will discuss my opinion on the future perspectives for the field in an attempt to outline what studies could allow for its further development.

## 7.2. Key Findings from the Thesis

### *7.2.1. Chapter 3: Investigating the use of Publicly Available Data for the Determination of MtDNA Methylation Patterns Across Different Tissue Types.*

In our study, we aimed to provide an *in-silico* pipeline to control for the presence of *NUMTs* in publicly available genome-wide methylomic datasets from post-mortem brain tissue, to gain a better understanding of the mitochondrial epigenome in human brain. A number of different approaches have been used to define mitochondrial variants in whole-genome sequencing (285-288) and RNA-Seq data (289) given the high relative-abundance of reads that align to the mitochondrial genome. However, the method we employ in Chapter 3, despite being more computationally intensive and somewhat slower than these previous approaches, is likely to yield the most true representation of mtDNA profiles. Indeed, similar approaches have been suggested for complementary analysis of mtDNA from

total-genome data using other sequencing technologies, as reviewed by Ye et al., 2014 (290); however our methodology is the first to be implemented on MeDIP-Seq data. Although MeDIP-Seq data has been employed to interrogate the mitochondrial epigenome in a variety of cell lines and tissues (63), the pipeline used in this study may not have adequately addressed the homology between the nuclear and mitochondrial genomes. That study aligned reads derived from whole-genome data and aligned directly to the mitochondrial genome (63). Whilst the study does attempt to address this limitation, via simultaneous modelling to the nuclear genome, forcing reads derived from whole-genome data into a targeted sequence could lead to a mixed nuclear-mitochondrial profile of mtDNA methylation being established.

One caveat of the method employed in Chapter 3 is that it removes any regions of the mitochondrial genome that share a high sequence homology to regions of the nuclear genome, meaning we were only able to examine a truncated sequence. However, our study did identify a relatively conserved pattern of mtDNA methylation between samples. This pattern of conservation was also seen in previous studies (63).

Our study also demonstrated that mtDNA methylation patterns were able to separate tissues based on their Euclidean distance, a measure of the variation between all samples. Furthermore, we were able to identify eight DMRs that were significantly differentially methylated after Bonferroni correction between total CTX

and CER. When taken together, our findings show tissue-specific differences in mtDNA methylation, and suggest that these could be influenced by brain-specific cell populations. Overall, our findings highlight that the mitochondrial epigenome warrants further investigation and that, ideally, such studies should be undertaken in isolated mtDNA in order to prevent the need to truncate the consensus sequence.

### *7.2.2. Chapter 4: A Comparison of MtDNA Isolation Methods in Post-Mortem Archived Human Brain Tissue: Applications for Studies of Mitochondrial Genetics in Brain Disorders.*

Our study represents an important step in developing a protocol to better interrogate the mitochondrial epigenome in the absence of *NUMTs*. Currently, most available mtDNA isolation techniques have been optimised on fresh tissue or cell lines (291,292), and have not been tested in frozen post-mortem brain tissue. Further, another common isolation methodology employs the use of long-range PCR to amplify the mitochondrial genome prior to sequence analysis (292). However, this method cannot be used prior to bisulfite treatment, as PCR abolishes the DNA methylation signature. As such, a method to specifically isolate mtDNA from frozen post-mortem tissue, which could then be used for downstream (epi)genomic analyses, was of upmost importance.

In Chapter 4, we compared different methodologies to enrich for mtDNA relative to ncDNA. Interestingly, one method (Method B), which utilised the exonuclease

$RecJ_f$ to preferentially cleave linear DNA (ncDNA), gave a lesser enrichment than the standard phenol: chloroform DNA extraction technique. One possible explanation for this could be that potential nicks introduced into mtDNA via the freeze-thawing process may have linearised the genome. Our study found that an antibody based enrichment technique from Miltenyi Biotec (Method E), significantly enriched for mtDNA compared to all other techniques tested. This method used an anti-TOM-22 antibody conjugated to magnetic microbeads to capture mitochondria, from which we later extracted DNA. Interestingly, this is in contrast to a previous study, which showed similar enrichments between Method E and the use of Percoll (Method A) (293). However, that study was carried out on cultured human cells. Method A relies on the establishment of a density gradient for adequate enrichment of mitochondria. Whilst mitochondria in cell lines and fresh tissue are likely to have similar weights and conformations, the shock of freeze-thawing tissue can lead to degradation of mitochondria and may therefore increase the size of the mitochondrial fraction, resulting in more nuclear contaminants being extracted from a less distinct band. Method E is dependant not on the physical properties of mitochondria, but on the organelle's protein content. TOM-22 is a protein expressed on the OMM, as such our findings could be explained by the specificity of this antibody, which is independent of tissue quality.

In summary, in Chapter 4, we therefore identified a methodology that is both fast and reliable for the enrichment of mtDNA from archived, post-mortem brain tissue, providing a basis for further studies into the field of mitochondrial epigenetics.

*7.2.3. Chapter 5: A Genome-wide Interrogation of MtDNA Methylation via NGS*

In Chapter 3 we identified mtDNA DMRs between CTX and CER. However, one caveat of this approach is that the antibody enrichment and fragment based library preparation makes it impossible to examine mtDNA methylation differences at single nucleotide resolution. Therefore, in Chapter 5 we employed an adapted NGS methodology from Agilent to identify specific patterns of mtDNA methylation across the mitochondrial epigenome in mtDNA isolated using the anti-TOM22 antibody method we selected in Chapter 4.

Our primary analysis identified that mtDNA methylation exists with a distinct prevalence towards non-CpG sites. This corroborates a previous study which shows that mtDNA methylation in the mitochondrial *D-LOOP* also has a striking prevalence towards non-CpG methylation (256). Of interest, this pattern of methylation is of stark contrast to that primarily exhibited in mammalian nuclear genomes (294-296), where only a small percentage of DNA methylation sites are found in this context and as such, the pattern of mtDNA methylation shares more similarity to that of a prokaryote (256), which is corroborated by our findings.

Our analysis focused on identifying DMPs associated with tissue type, sex and age. Sexual dimorphism of mitochondrial output has previously been identified (297) and has also been associated with neuronal cell death (298). Despite these studies being limited to nuclear-mitochondrial genes, they do give evidence to support a sex-specificity for mitochondrial function. Interestingly our study

highlighted five DMPs remaining significantly associated with sex after Benjamani-Hochberg correction. Further, two DMPs, at 9055bp and 8722bp, also passed a more strict Bonferroni correction for multiple testing. Two of these DMPs were identified within *MT-ATP6*, a gene encoding for a mitochondrial protein within Complex V of the ETC. This complex is critical for sustainability of the proton gradient required for ATP generation (299). Whilst the functional consequences of these differences have yet to be determined, it remains plausible that they could contribute towards some of the sex-specific variability observed in mitochondrial outputs.

An interesting observation in ncDNA epigenetics is the association of epigenetic modifications with ageing. In general, global levels of ncDNA methylation show hypomethylation with increasing age (300-303). However, EWAS has demonstrated that these age-associated ncDNA methylation changes are loci-specific, with multiple CpGs showing hypermethylation with increasing age (304,305). Furthermore, in both EWAS of age the ratio of hyper- to hypomethylated probes is relatively close to one. In Chapter 5, we showed an enrichment for age-associated DMPs to be hypermethylated with increasing age in both the STG and CER, suggesting that mtDNA methylation is dissimilar to ncDNA methylation in this respect.

Our findings also disputes previous research that has suggested that mtDNA methylation exists at low, biologically irrelevant levels (70). Whilst in Chapter 5 we

report that in general, mtDNA methylation levels were low (1-3%), we found that an average of 4.68% of all loci in our cohort displayed DNA methylation levels greater than 5%. The average sequencing depth of our study was of an order of magnitude greater than the previous study by Hong and colleagues and, as such, is more likely to represent average methylation levels in a multi-copy, heteroplasmic genome. Importantly, we also identified one locus within the mitochondrial *D-LOOP* with average DNA methylation levels of 15.3% and 14.6% methylation in the STG and CER respectively. Interestingly, this region has been strongly implicated in mitochondrial transcription and replication (306-308). As such, relatively high levels of mtDNA methylation at this site could have strong functional implications for mitochondrial biogenesis.

Our method was able to specifically and significantly enrich for mtDNA in the absence of *NUMTs*. To demonstrate this, we sequenced matched samples that were enriched for mtDNA and their respective ncDNA using our mtDNA isolation method (Method E- Chapter 3), prior to library preparation. We then calculated the average loci-wise mean and standard error of each methylated cytosine to determine the potential variability at each site. By comparing this to methylation levels of matched, ncDNA samples, we were able to infer whether the enrichment affected DNA methylation levels. Loci with DNA methylation levels greater than the standard error for that tissue type were said to be affected by the enrichment. In total, only 167 of 870 probes were unaffected by the enrichment for STG. However, 679 of 870 probes were unaffected by the enrichment process in CER. Whilst the sample size makes drawing clear conclusions from this analysis difficult, it does

suggest that enriching for mtDNA prior to sequencing can affect DNA methylation levels observed.

In summary, this study identified a number of DMPs associated with tissue, sex and age. As such, our study highlights the importance of controlling for these factors in future epidemiological studies and also potentially implicates mtDNA methylation as a functional epigenetic mark. Further, we were able to develop and optimise a custom, targeted BS-Seq method for the mitochondrial epigenome, providing the basis for more detailed investigations into mtDNA methylation in future studies.

### *7.2.4. Chapter 6: Interrogating the Role of MtDNA Methylation in AD*

Given that AD is characterised by mitochondrial dysfunction, we hypothesised that this may be driven by epigenetic processes. In the final data Chapter we aimed to use the methods developed in Chapters 4 and 5 to investigate mtDNA methylation differences associated with AD neuropathology, as well as validating sex and age associated DMPs identified in Chapter 5. In addition, we investigated DNA methylation differences in AD in nuclear-encoded genes encoding proteins that are imported into the mitochondria.

Initially we sought to identify sex-associated DMPs and compare these to those we identified in Chapter 5. Initially, we sought to identify sex-associated DMPs and

compare these to those we identified in Chapter 5. However, only one, at 15,761bp (*MT-CYB*) was replicated in all three regions. Unfortunately, the majority of studies investigating *MT-CYB* have focussed on its role in disease (309-311) and as such, sex-specific variation within this gene has yet to be thoroughly investigated. However, a further three were found to be in common between ECX and STG, as well as a further two which were significant in ECX and trended towards significance in STG. We were unable to find any common sex-associated DMPs between ECX and CER. This suggests that sex-specific DMPs may also be tissue-specific, with the two cortical brain regions (ECX and STG) sharing more sex-associated DMPs than ECX and CER. It is worth noting however, that we did find more sex-associated DMPs in common between the STG and CER than between ECX and STG; however, this is possibly due to the STG and CER being isolated and analysed at the same time (i.e. within Chapter 5), whilst the ECX analysis (Chapter 6) was independent, and thus samples were isolated and analysed at a different time.

Our primary aim in Chapter 6 was to identify mtDNA methylation differences associated with AD neuropathology, as defined by Braak stage. We identified 29 nominally significant, disease-associated DMPs. Interestingly, many of these DMPs occurred in genes encoding for ETC proteins, however no enrichment for DMPs in any ETC complex was identified. The deregulation of OXPHOS and generation of ROS has been tightly linked to mitochondrial dysfunction and AD (267,312). However to date, no study has identified the molecular mechanisms driving this phenomenon. It would be important to investigate the consequence of

disease-associated mtDNA methylation differences we identified, on mitochondrial function.

Our meta-analysis aimed to investigate DNA methylation differences in nuclear-encoded mitochondrial genes and identified 44 nominally significant DMPs that were consistently differentially methylated in both cohorts. Of these, one DMP residing in *SPG7* was Bonferroni significant. This DMP has been previously identified as an EWAS locus in AD in a cohort of 708 prefrontal CTX samples (95). Our findings further highlight the robustness of this DMP in AD, by identifying it in a cohort of 217 ECX samples from independent donors. *SPG7* is the only gene to have been associated with spastic paraplegia 7, which not only presents with limbic deterioration but also a number of neurological factors such as white-matter damage in the frontal lobe (280). Further, genetic and transcriptomic variation within *SPG7* have also been associated with ALS, a chronic, neurodegenerative disease (281), as well as in spastic ataxia (313), and has also been found to play a role in various models of cerebellar ischemia (282,283). *SPG7* downregulation has been shown to lead to the establishment of the mPTP (283). Interestingly, the establishment of the mPTP leads to $Ca^{2+}$ retention and mitochondrial depolarisation and has been associated with AD pathogenesis in previous studies, as reviewed in Qi, et al., 2016 (284).

Finally in Chapter 6 we investigated the relationship between the 44 nuclear-encoded Braak-associated DMPs and the 29 mitochondrial-encoded Braak-

associated DMPs. In total we identified 55 nominally significant correlations between the two genomes. Further, we identified four DMPs (two mitochondrial, two nuclear) that were enriched for significant correlations within the other genome. This final analysis therefore provides a potential new mechanism for the complex interplay between the two genomes and should be further investigated to better understand the molecular mechanisms in place to aid correct mitochondrial functioning. Negative correlations between DMPs within the two genomes could indicate a potential compensatory mechanism, whereby changes in the methylation profile of specific loci in one genome could lead to an alteration of ETC complex function, resulting in changes to the methylation profiles of other loci in the other genome. However, this assumes that mtDNA methylation has an effect on gene expression. Whilst a recent study has shown that changes in mtDNA methylation at the mitochondrial promoter regions are associated with differences in the expression of a number of mitochondrial genes (314), a thorough genome-wide investigation into mtDNA methylation and gene expression has yet to be undertaken. Therefore, understanding the potential complex relationship between the two genomes at the epigenetic level requires careful consideration.

## 7.3. General Discussion and Future Perspectives

Despite identifying nominally significant differences associated with all covariates analysed in Chapters 5 and 6, the effect sizes seen in both studies are relatively small and elucidating their functional consequence could prove very difficult. It is important to note however that this could be due to the high levels of heterogeneity

taken from a tissue sample. The multi-copy nature of mitochondria ensures that, not only does mtDNA show a great level of heteroplasmy across tissues, but also that it may occur within a single cell (315). As such, a detailed investigation into the actual level of mtDNA methylation at the single-cell level would significantly advance the field.

The development of single cell techniques for the investigation of genetic heteroplasmy within mitochondria has aided in elucidating our understanding of such variation within the organelle (316,317). For example Jayaprakash et al., (2015) developed a technique, MSeek, to isolate and sequence mtDNA from cells at different timepoints, allowing for the accurate determination of heteroplasmy levels under different conditions. This has served to improve our understanding of the role of mitochondrial replication in maintaining levels of heteroplasmic variants (317) and could aid in more precisely determining mitochondrial variants within cells causing mitochondrial dysfunction. To determine the functional effect of mtDNA methylation within a cell at a high precision would be of great benefit to the field, however at present, there is no methodology capable of accurately processing such low levels of mtDNA for epigenetic sequencing.

To determine the role mtDNA methylation may play on the governance of changes to mitochondrial gene expression, ROS levels and ETC activity is of upmost importance for future studies and would aid in determining whether mtDNA methylation may contribute to disease. One way in which this could addressed is

through the use of a cybrid cell line (see Section 1.5.4). Briefly, native mtDNA is knocked down within a cell line and patient mtDNA with a (genetic) variant of interest is inserted. This allows for a control cell line with the same ncDNA as the cell line with the inserted mtDNA. As such, variation in mitochondrial output would be driven by the mitochondrial variant (318). To date, this has only been performed to establish the effects of mt-SNPs (319), however it remains possible to utilise this methodology to explore the functional effects of mtDNA methylation. By treating cells with 5-aza, a commonly used DNA methylation inhibitor (320-322) or SAM, a substrate used to increase DNA methylation (323), parent cell mtDNA methylation could be abolished or increased. Whilst delineating the specific effects of mtDNA methylation in the parent cells would be difficult, given that ncDNA methylation would also be affected, transferring mtDNA with altered methylation profiles into a cybrid cell line could help to answer key questions in the field of mitochondrial epigenetics. One caveat of this approach is that it would only investigate global alterations in mtDNA methylation levels, and would not be able to distinguish the functional effects of site-specific mtDNA methylation changes.

Adaptations on the original CRISPR-Cas9 genetic editing technology have allowed for site specific epigenetic modifications (324). Traditionally, the CRISPR-Cas9 construct employs a synthetic guide RNA to recognise and target the sequence of interest. Cas9 is then directed towards this sequence, cutting the DNA and allowing for targeted, site-specific genetic alterations (325). To allow for the addition of specific epigenetic modifications to the target site, the cleavage sites of Cas9 are deactivated, producing a dCas9 variant that binds to the target sequence

without cutting DNA (324). Further, by addition of a mitochondrial specific targeting sequence to the original CRISPR construct, specific mtDNA alterations can be achieved (326). It is therefore possible that a coupling of these two adaptations could create a construct that would specifically target specific loci in the mitochondrial epigenome. By tethering DNMT or TET proteins to this mitochondrial dCas9, mtDNA methylation could therefore be increased or decreased respectively at target sites.

In Chapter 6 we identified a number of significant correlations between disease-associated DMPs encoded by the nuclear and mitochondrial genomes. However, the potential importance of these correlations has yet to be elucidated. For example, inverse correlations could represent a compensatory mechanism, whereas positive correlations could indicate the development of a positive feedback loop. However, it is important to note that these differences were identified in a small cohort and none passed any testing for multiple corrections, nor have they been validated by subsequent experiments. Despite this, with the development of technologies such as CRISPR-Cas9 epigenetic editing, it would remain of significant interest to determine whether ncDNA methylation differences can lead to specific alterations in the mitochondrial methylome and whether these changes may in turn, effect mitochondrial output.

A major limitation of this study is the pre-enrichment methodology we chose to employ. Our adapted mitochondrial isolation method, utilised throughout the course

of this thesis, showed significant enrichments for mtDNA. However, we showed in Chapter 5 that mtDNA methylation varied in matched samples that were and were not enriched for mtDNA prior to capture. This raises two important questions. First, how similar is the methylome of pure mtDNA to the methylome of mitochondria isolated using the anti-TOM22 antibody approach? Second, if mitochondria were isolated from matched samples using different isolation methodologies that achieved a similar purity, would the mtDNA methylome of these matched samples be similar; does the enrichment methodology we employ throughout this thesis have any unexplored biases and would a different methodology yield different methylation profiles? Both questions would be important to address in the future to provide further evidence to support the reproducibility of our experimental design.

Despite these limitations, it is important to note the stage in which the field of mtDNA epigenetics was prior to these studies, being restricted to candidate based gene approaches to explore small sections of the mtDNA epigenetic variation in disease (1,221). The development of EWAS has led to a significant improvement in our understanding of the complex role that ncDNA epigenetics plays in health and disease, as reviewed by Flanagan., 2015 (327). It is therefore possible that the development of this mitochondrial epigenome-wide pipeline could similarly aid the field of mitochondrial epigenetics.

Our studies have now provided a strong basis for the future exploration of mtDNA methylation in any number of clinical phenotypes or experimental settings, with

many questions still remaining to be answered. The most important of which are first, whether variation in mtDNA methylation has an effect on mitochondrial function and second, whether in certain phenotypes, for example in those that are characterised by high levels of mitochondrial dysfunction, the effect sizes seen may be greater than those identified in Chapters 5 and 6. Experimental design remains critical to answering these questions and coupling a discovery based approach such as the one defined throughout the course of this thesis, with a target validation methodology such as the dCas9 approach suggested could aid in the elucidation of a previously ill-defined layer of epigenetic regulation.

## Appendix 1

# The mitochondrial epigenome: a role in Alzheimer's disease?

Matthew Devall, Jonathan Mill and Katie Lunnon

**Keywords:**

Alzheimer's disease, AD, dementia, mitochondria, epigenetics, DNA methylation, mtDNA, heteroplasmy, 5-methylcytosine, 5-hydroxymethylcytosine

## SUMMARY

**Considerable evidence suggests that mitochondrial dysfunction occurs early in Alzheimer's disease, both in affected brain regions and in leukocytes, potentially precipitating neurodegeneration through increased oxidative stress. Epigenetic processes are emerging as a dynamic mechanism through which environmental signals may contribute to cellular changes, leading to neuropathology and disease. Until recently little attention was given to the mitochondrial epigenome itself, as preliminary studies indicated an absence of DNA modifications. However, recent research has demonstrated that epigenetic changes to the mitochondrial genome do occur, potentially playing an important role in several disorders characterized by mitochondrial dysfunction. This review explores the potential role of mitochondrial epigenetic dysfunction in Alzheimer's disease etiology and discusses some technical issues pertinent to the study of these processes.**

Alzheimer's disease (AD) is a chronic, currently incurable, neurodegenerative disorder, accounting for more than 60% of dementia cases, with current estimates predicting more than 135 million dementia cases worldwide by 2050 [1]. The classic neuropathological hallmarks associated with AD include the formation of amyloid beta (Aβ) plaques and neurofibrillary tangles. These are suggested to play a role in the further development of other characteristics of the disease, such as disruption of calcium homeostasis, loss of connectivity, the generation of reactive oxidative species (ROS) and altered plasticity, ultimately leading to neurodegeneration [2-6]. Mitochondrial dysfunction is a consistent feature of AD pathology in both the brain and white blood cells [7-10] although the molecular mechanism(s) mediating this phenomena are yet to be fully elucidated.

**Mitochondrial dysfunction: a prominent feature of AD**

Being the site of ATP generation, mitochondria provide the cell with the energy required to properly function; as such they are often described as 'the powerhouse of the cell'. Mitochondria are cylindrical organelles containing ~16.6kb of DNA (mtDNA) [11], which is separate to the nuclear genome and inherited in a maternal, non-Mendelian fashion. The mitochondrial genome consists of 37 genes, 13 of which encode for polypeptides required for the electron transport chain (ETC) (Figure 1), in addition to two ribosomal RNAs and 22 transfer RNAs. The mitochondria play a vital role in a variety of key biological functions, including apoptosis via caspase dependent and independent mechanisms [12], the

regulation of calcium homeostasis [13, 14] and the production of ROS [15]. For these reasons, mitochondrial dysfunction has been implicated in the pathogenesis associated with AD [16, 17] and forms the basis of the mitochondrial cascade hypothesis [18]. Proposed by Swerdlow et al, this hypothesis states that an individual's genetic code will determine their basal mitochondrial function and that, throughout ageing, this function will decline due to a combination of genetic and environmental factors, determining an individual's time of disease onset [18].

Mitochondrial-encoded ETC gene expression has been shown to be altered in both early and late stages of AD, with decreased expression of complex I and increased expression of complexes III and IV [7]. Increased expression of mitochondrial-encoded ETC complex genes has also been associated with aging, with increased expression of complexes I, III, IV and V in 12- and 18-month wild-type mice compared to 2-month mice, which was accompanied by increased oxidative damage [19]. However decreased expression of these genes was seen in older, 24-month old mice. Further evidence for a role of mitochondria in AD pathogenesis comes from a study demonstrating increased levels of mitochondrial gene expression and oxidative damage in a transgenic Amyloid Precursor Protein (*APP*) mutant mouse model of AD [20]. In addition, various components of the mitochondrial permeability transition pore (mPTP), which acts as a voltage-dependent channel regulating mitochondrial membrane permeability, have been shown to interact with Aβ in various murine models of AD. For example, one recent study found that, in *APP* transgenic mice, Aβ acts to upregulate VDAC1, a component of the mPTP, leading to mPTP blockade. [21]. Interestingly, this study also reports that VDAC1 may interact with hyperphosphorylated tau, suggesting another

mechanism of mitochondrial dysfunction. An earlier study found that Aβ present in mitochondria interacts with CypD, another component of the mPTP, in cortical samples from post-mortem AD patients and *mAPP* transgenic mice [22]. In the mouse model, this was shown to lead to increased ROS production and neuronal cell death. Taken together, this illustrates how mitochondrial-encoded gene expression is altered in AD, a variety of mechanisms by which Aβ interacts with mitochondria in AD, and how mitochondrial dysfunction can lead to changes associated with AD, thus highlighting the need for continued research into the field.

## Epigenetics and AD

Given the high heritability estimates for AD [23], considerable effort has focussed on understanding the role of genetic variation in disease etiology, although more recently it has been hypothesized that epigenetic dysfunction may also be important [24]. A number of studies have shown reduced global levels of the DNA modifications 5-methylcytosine (5-mC) and 5-hydroxymethylcytosine (5-hmC) in AD brain [25-28] with only a handful of studies have looked at changes occurring at specific loci (reviewed in [24]). Recent methodological advances in microarray and genomic sequencing technologies have enabled researchers to undertake epigenome-wide association studies (EWAS) in AD brain, identifying several consistent differentially methylated regions (DMRs) associated with disease [29-31]. Many of these DMRs are tissue-specific, restricted to regions of the brain associated with AD pathology, and correlate strongly with quantitative measures of neuropathology. As such, a strong case is being built for a role of epigenetics in the etiology of AD.

## Epigenetic regulation of the mitochondrial genome

Although hypotheses about the importance of mtDNA modifications are by no means recent, research in this area has been marred by contradictory results since the 1970s [32-35]. The confirmation in 2011 of both 5-mC and 5-hmC occurring in mtDNA prompted a resurgence of interest in mitochondrial epigenomics [36]. The mitochondrial epigenome has some notable differences compared to the nuclear epigenome, and an overview of the mitochondrial genome, including its CpG sites, can be seen in Figure 1. Unlike the nuclear genome, the mitochondrial genome does not contain classical CpG islands [36], and is not associated with chromatin; instead it is structurally organised by nucleoids [37, 38]. As a result, mtDNA is not associated with histone proteins and relies on transcription factors such as mitochondrial transcription factor A (TFAM) to mediate compaction [39]. Histone modifications do not therefore play a direct role in regulating mitochondrial gene expression, highlighting the potential importance of DNA modifications in the regulation of mitochondrial function [40]. Evidence suggests that mtDNA methylation largely influences mtDNA structure and replication and is affected by factors which influence nucleoid compaction and DNA methyltransferase (DNMT) binding [41]. It has been shown that different areas of mtDNA are packaged differently and that a depletion of the nucleoid protein ATAD3 can reduce mtDNA methylation**,** resulting in an open circular state mitochondrial genome, although evidence for an effect of TFAM on mtDNA methylation was inconclusive [41].

DNMTs are a family of enzymes that catalyse the removal of a methyl group from methyl donors such as S-adenosylmethionine (SAM) for addition to the 5-position of cytosine.

Recently, a DNMT isoform, mitochondrial DNMT1 (mtDNMT1), has been found to contain a mitochondrial targeting sequence allowing it to bind to the D-loop of the mitochondrial genome, which contains the promoter sites for both the light and heavy strand of mtDNA and can therefore influence mitochondrial gene expression by altering transcriptional activity [36]. Furthermore, it has been suggested that the presence of these methyltransferases in mitochondria may be tissue-specific. Although Shock *et al*, did not observe mitochondrial localization of DNMT3a in the two cell lines they investigated, a later paper has found that DNMT3a is present, and in higher levels than mtDNMT1, in the mitochondria of motor neurons [42]. This study also demonstrated significantly higher global levels of both mitochondrial DNMT3a and 5-mC in Amyotrophic Lateral Sclerosis (ALS) motor neurons *in vivo*, suggesting a potential role for mtDNA methylation in motor neurons. DNMT1 and DNMT3b have also been observed in the mitochondria, with their inactivation reducing methylation at CpG sites [43].

Recently, it has been debated whether 5-hmC is just an intermediary product of the demethylation process of 5-mC to cytosine or could represent an independent epigenetic mark [44]. Growing evidence now suggests that 5-hmC could be a mark in its own right, produced from the conversion of 5-mC by TET1, TET2 and TET3 [45], with both TET1 and TET2 being present in the mitochondria [43]. Taken together with the presence of 5-hmC in the mitochondrial D-Loop [36], this strengthens the evidence suggesting that demethylation pathways are not only important in nuclear epigenetics, but may also play a role in the mitochondria. Furthermore, recent evidence suggests that 5-mC and 5-hmC exist stably within mtDNA at cytosines not preceding a guanine base, suggesting a role for non-CpG

methylation in mtDNA [46]. Further, CpG and non-CpG methylation has been observed in the mitochondrial D-loop at conserved regions associated with DNA-RNA hybrid formation during transcription, suggesting that DNA methylation in mitochondria shares similarities with plants and fungi and that this methylation may play a role in regulating mtDNA transcription and replication in a cell type-specific fashion [43].

## MtDNA modifications in disease

Despite little being known about the physiological impact of variation in mtDNA methylation, some recent studies have shown that it may be associated with a variety of diseases. The majority of studies have focussed on diseases where mitochondrial dysfunction is known to be prevalent, for example in cancer, which has been previously linked with mitochondrial dysfunction [47] and more recently in Down's Syndrome, where mitochondrial abnormalities have also been reported [48]. Particularly, for the purpose of this review, mitochondrial dysfunction and mtDNA methylation aberrations in Down's syndrome cells (see Table 1) are interesting given that these patients have an increased likelihood of presenting with AD-like phenotypes throughout aging [49, 50] due to possessing an extra copy of *APP*. An overview of studies of mtDNA epigenetics in disease is given in Table 1.

## MtDNA modifications: evidence for a role in AD and aging

Until recently the role of mtDNA modifications in AD has been largely ignored, despite the evidence that mitochondrial dysfunction is involved in AD [18] and that ncDNA methylation

differences are associated with the disease [29-31]. At a global level, an initial dot blot study showed some evidence for increased mitochondrial 5-hmC in AD superior temporal gyrus tissue, although definitive conclusions could not be drawn given the small number of samples used [51]. Mitochondrial DNA modifications in the brain have been shown to be associated with aging, with global mtDNA 5-hmC levels reduced in the frontal cortex of aged mice and specifically decreased 5-hmC levels being found in the regulatory D-Loop, as well as in two genes encoding ETC complex I polypeptides (*MT-ND2* and *MT-ND5*) [52]. Aging was not only found to be associated with overall decreased mtDNA 5-hmc levels but also with increased cortical expression of the mitochondrial ETC genes *MT-ND2, MT-ND4, MT-ND4L*, *MT-ND5*, and *MT-ND6* [52]. A post-mortem study of frontal cortex described differential mtDNA gene expression of these genes, and other mitochondrial-encoded genes, in both early and late-stage AD. [7] Taken together, these findings illustrate that alterations in mitochondrial-encoded genes do occur with aging and in age-related diseases, yet without further studies, the exact role of mtDNA methylation on mitochondrial gene expression in these instances remains uncertain.

## Two genomes are better than one: interactions between the nuclear and mitochondrial genomes

As research into the field of mitochondrial epigenetics gains momentum, studies have focused on a potential *trans*-acting role of mtDNA in the epigenetic regulation of ncDNA, whereby covalent modifications across the mtDNA genome may affect not only the expression of a gene in *cis*, but also have *trans*-acting effects on the transcription of genes in the nuclear genome. Evidence for this is provided by cybrid models, which combine the

nuclear genome of one source with the mitochondrial genome of another in an attempt to determine the functional role of the mtDNA. Using Restriction Landmark Genomic Scanning (RLGS) and $Rho^0$ cells, a form of cybrid cell line designed for investigating mtDNA depletion, one study found that mtDNA depletion significantly altered DNA methylation at CpG islands in nuclear encoded genes [53], indicating that there are functional interactions between the two genomes. Re-introduction of wild-type mtDNA restored DNA methylation levels, at some RLGS spots, suggesting that, at least for some genes, mitochondria may play a role in nuclear DNA methylation. This is corroborated by a recent study demonstrating that mitochondrial haplotype variation can affect ncDNA methylation, with mtDNA haplotype J exhibiting higher global DNA methylation levels, reduced ATP, and overexpression of the nuclear gene methionine adenosyltransferase I, alpha (*MAT1A*), which is required for SAM production thus regulating methylation patterns in the nuclear genome [54]. Therefore genetic variations in mtDNA are capable of influencing epigenetic modifications in both the mitochondrial and nuclear genomes. As such, it is possible that mitochondrial dysfunction in AD could lead to alterations in mtDNA methylation, affecting nuclear gene expression.

The mitochondria comprises approximately 1500 proteins, however of these, only 13 are encoded by the mitochondrial genome; the remainder are encoded by the nuclear genome and imported into the mitochondria. A recent study found that >600 of these genes have tissue-specific differentially methylated regions, ultimately leading to changes in mitochondrial function dependent upon tissue type [55]. This suggests that there is an additional level of complexity to consider in the study of mitochondrial epigenetics,

whereby epigenetic changes in one genome may affect transcriptional control in another in a tissue-specific manner.

## Interrogating the mitochondrial epigenome: technical caveats

Despite the potential importance of mitochondrial DNA modifications in AD, there are a number of technical challenges specific to interrogating the mitochondrial epigenome that have hampered widespread studies to date. These issues can be broadly summarized as encompassing genetic issues and specificity issues, which are outlined briefly with potential solutions in Table 2.

1. Genetic issues

### 1.1. Nuclear pseudogenes

By far the greatest concern when analyzing mtDNA methylation arises from regions of homology between the mitochondrial genome and nuclear mitochondrial pseudogenes (NUMTs). These genes are nuclear paralogs of mtDNA which have been translocated and inserted into the nuclear genome during evolution of both genomes [56]. This phenomena has been shown to be evolutionarily conserved across many species including cats [57], mice, chimpanzees, rhesus macaques [58] and hominins [59]. These insertions were thought to typically occur in non-coding regions; however, more evolutionary recent translocations have actually been shown to prefer integration into coding regions, thus leading to potential alterations in gene function with implications for disease [60]. NUMTs are generally small and typically comprise ~0.1% of the nuclear genome [61]. However in humans, it has been shown that some NUMTs can be as large as 14.7kb, representing a

significant portion of the ~16.6kb human mitochondrial genome [62]. As such, the presence of NUMTs can cause major issues in genomic analyses using pre-sequencing enrichment methods such as custom capture or long-range PCR as the likelihood of NUMT co-amplification, or even preferential amplification, increases due to the strong sequence similarity between the two segments of genome [63]. As such, this sequence similarity can lead to the misclassification of NUMTs as mtDNA during analysis, and has led to a number of publications wrongly describing NUMTS as mtDNA [64, 65]. NUMT misclassification has also been observed in AD genetic studies whereby amplification of the NUMT sequence has led to false heteroplasmies (see below) being reported [66, 67]. One potential solution is to separate mitochondria prior to DNA extraction in an attempt to reduce the risk of contaminating the mitochondrial and nuclear genomes. However, despite extensive research being dedicated to mtDNA analysis, existing methods for mitochondrial isolation and mtDNA extraction via the use of fractional precipitation or gradient ultracentrifugation remain time consuming and labour intensive [68] and often leave residual nuclear DNA contamination following mitochondrial isolation [69].

### 1.2. Variation in mtDNA: haplogroups and genetic and epigenetic heteroplasmy

Each mitochondrion contains between 2-10 copies of mtDNA. However, not all mtDNA in each mitochondrion share the same DNA sequence. Indeed, mutations in some copies of mtDNA mean that the cell itself may be made up of a mixture of different sequences. This phenomenon is known as mitochondrial heteroplasmy and has been linked to various mitochondrial diseases, [70]. It is a potential confounder in studies of mitochondrial diseases, because inter- and intra-individual heteroplasmic variation can confuse the

248 association between a haplogroup with its corresponding phenotype. The importance of
249 this issue, in the context of this review, is highlighted by a recent study demonstrating that
250 mitochondrial heteroplasmy alters DNA methylation across the nuclear-encoded
251 mitochondrial genes TFAM and POLMRT [71]. Finally, if mtDNA methylation is altered across
252 different mtDNA in the same mitochondrion, it could create an epigenetic mosaic within the
253 mitochondrion, the cell and across the tissue, whereby each copy of mtDNA may possess its
254 own methylation profile. If this 'methylomic heteroplasmy' were to occur it could be very
255 difficult to tease apart the effects of such a mosaic in functional studies.

256

257 On a larger scale, mutations in mtDNA can be used to help group cohorts or "haplogroups".
258 Throughout evolution, mutations in mtDNA may be conserved and passed on through
259 maternal inheritance, thus allowing for the tracing of common ancestral lineage by
260 comparing haplogroups. Numerous studies have identified both contributory and protective
261 effects of different haplogroups in AD. For example, haplogroup K reduces the risk of
262 developing sporadic AD in Apolipoprotein ε4 (*APOε4*) carriers in an Italian population [72]
263 but not in the Polish population [73]. This presents an additional potential caveat in
264 mitochondrial epigenetics, as mitochondrial haplogroups have been found to affect global
265 levels of DNA methylation [54]. As such, extra care should be taken to account for
266 haplogroup variability in AD mitochondrial epigenetic studies.

267 2. Specificity and technical issues

268 The brain is a complex, heterogeneous organ with numerous functionally-distinct sub-
269 regions, each with their own different composition of cell types. Unsurprisingly, there are

clear tissue-specific epigenetic differences across brain regions [74, 75]. There is an added level of complexity with respect to the mitochondrial epigenome because each mitochondrion contain between 2-10 copies of mtDNA and each cell contains varying levels of mitochondria; therefore the amount of mtDNA copies in each cell can vary between 100-10,000 dependent upon cell type. In neurodegenerative diseases such as AD, the issue becomes more complicated in that the disease itself is characterized by the loss of neuronal cells and the activation of glia, a process that has been associated with changes in mitochondrial morphology and fission [76, 77]. A recent study using laser capture microdissection demonstrated that alterations in mitochondrial 5-hmC are seen with age in dissected mouse cerebellar purkinje cells, which was not evident in whole cerebellar tissue [52], demonstrating the importance of cell-specific analyses in heterogeneous tissue, particularly when investigating functional impact.

Currently, the most common method of measuring DNA methylation is via the conversion of DNA with sodium bisulfite followed by subsequent sequence analysis. However, these approaches are unable to distinguish between 5-mC and 5-hmC [78], an important limitation given recent studies confirmed the presence of 5-hmC in mitochondria in brain tissue [52]. Studies have found that although both DNA modifications are present in the mitochondria, they occur at much lower levels compared to in ncDNA [36, 51], and thus methods used for quantification may need to be more sensitive. Furthermore, variation in mitochondrial copy number may lead to the dilution of signals and reduce detection if the tissue is largely heterogeneous. Importantly, the mitochondrial genome is not interrogated using tools such as the Illumina Infinium 450K methylation array, the current gold standard

for methylomic analyses in large numbers of samples; thus methods for detecting mtDNA modifications across the entire mitochondrial genome are largely restricted to antibody-based enrichment, such as MeDIP-Seq, which may be less sensitive for detecting low levels of modified cytosine and does not interrogate methylation levels at single base resolution [79].

## Future Perspective: the potential for biomarkers in AD

Two important goals of research into the etiology of AD are a) a fast, non-invasive, inexpensive and reliable biomarker and b) an effective treatment that targets the underlying neuropathology. A potential utility for DNA methylation biomarkers has been proposed for diseases in which traditional biomarkers are either too expensive, invasive, unspecific or insensitive for clinical purposes [80]. Epigenetic modifications have been widely studied in a variety of different cancers and other conditions such as preeclampsia to check for their suitability as prognostic and/or diagnostic biomarkers [81-83]. Differential methylation of mtDNA has yet to be examined with respect to its potential utility as an AD biomarker, but certainly warrants further investigation.

## Conclusion

With mitochondrial epigenetics only recently emerging as a focus for biomedical research, the role of the mitochondrial epigenome in AD has yet to receive much attention. However, it is possible that deregulation of the mitochondrial methylome may lead to aberrant changes in many of the intricately controlled processes that it helps to govern, such as

apoptosis, which may play a key role in pathogenesis. Furthermore, as mitochondrial dysfunction occurs early in AD pathogenesis, it is plausible that alterations in the mitochondrial methylome may play a major role in the onset and development of the disease. Despite the field presenting numerous challenges the links between mitochondrial epigenetics and AD provide good bounds for future research directions.

## EXECUTIVE SUMMARY

### Mitochondrial dysfunction: a prominent feature of AD

- The mitochondrial genome plays a vital role in a variety of key biological functions, including apoptosis via caspase dependant and independent mechanisms, regulating calcium homeostasis and production of ROS.
- Mitochondrial dysfunction is reported to occur in both the brain and blood of AD patients.

### EWAS and AD

- Studies focussing on global levels of 5-mC and 5-hmC have found a reduction in levels of both marks in AD brain.
- Three recent EWAS studies have found differential methylation at specific loci in AD brain.

### Epigenetic regulation of the mitochondria genome

- Despite early controversial results, both 5-mC and 5-hmC have been recently reported in mitochondria.

- MtDNA is not tightly wrapped by histones and is instead condensed by nucleoids, suggesting methylation could play an important role in gene regulation.
- DNMT1 can bind to the D-Loop of the mitochondrial genome and can influence gene expression.
- MtDNA methylation occurs at both CpG sites and non-CpG site in the mitochondrial genome.

**MtDNA methylation: a key player in AD?**

- Very few empirical studies have examined the role of mtDNA methylation in brain.
- Decreased mtDNA 5-hmC levels and increased expression of some mitochondrial-encoded genes has been seen in the pre-frontal cortex of aged mice.

**Technical caveats**

- NUMT misclassification has been observed in AD genetic studies whereby amplification of the NUMT sequence has led to false heteroplasmies being reported.
- MtDNA methylation could be altered in different mitochondria, creating a methylomic heteroplasmy.
- MtDNA methylation patterns could be cell-specific and is an important consideration when investigating heterogeneous tissues such as brain.

**Figure 1: The structure of the mitochondrial genome showing genes encoded by the mitochondria.** 3358 mtDNA genetic variants are shown in red and black lines highlight the predicted CpG sites relative to mutations that define the mitochondrial haplogroup. $P_H$ and $P_L$ represent the heavy and light strand promoter regions and $O_H$ and $O_L$ represent the origins of heavy-strand and light-strand replication respectively. Image taken from [84].

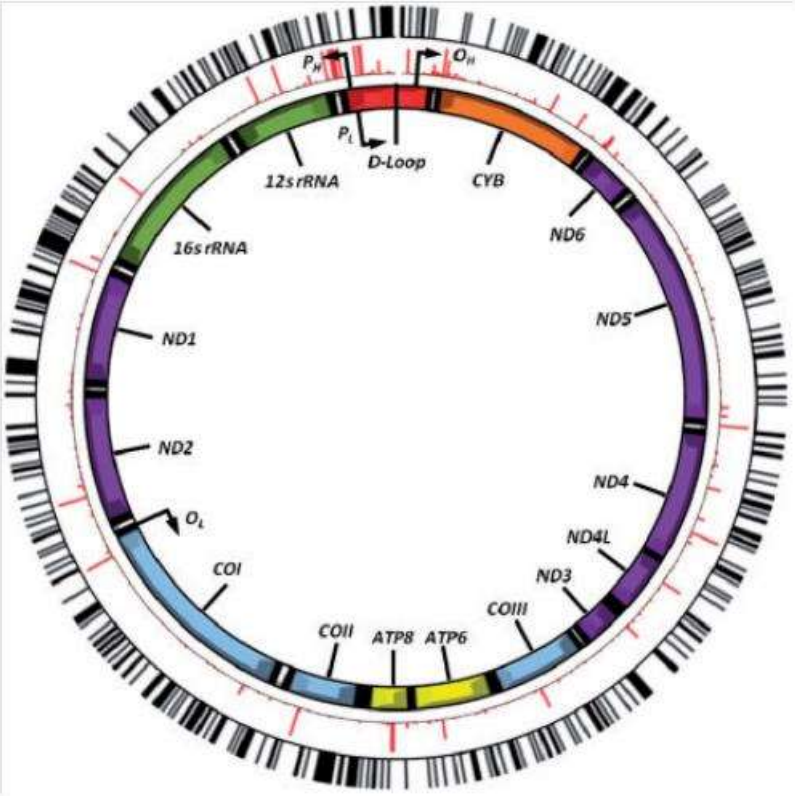


**Table 1: An overview of current studies of mitochondrial epigenetics in disease.** Abbreviations: Quantitative Real-Time PCR (qRT-PCR), mitochondrial DNA (mtDNA), simple steatosis (SS), non-alcoholic steatohepatitis (NASH), S-adenosylmethionine (SAM) , Liquid chromatography-electrospray ionization tandem mass spectrometry (LC-ESI-MS), Liquid chromatography mass spectrometry (LC-MS), Immunofluorescence (IF), Amyotrophic lateral sclerosis (ALS).

| Research Question | Techniques | Main Findings | Reference |
|---|---|---|---|
| The effect of different environmental exposures (metal-rich particulate matter, air benzene levels and traffic derived elemental carbon levels) on mitochondria | Pyrosequencing<br><br>qRT-PCR | Increased exposure to particulate matter increases *MT-RNR1* and *MT-TF* gene methylation<br><br>Increased *MT-RNR1* methylation is associated with a significant increase in mtDNA copy number. | [85] |
| The effect of mtDNA methylation in the mitochondrial D-Loop on gene expression in colorectal cancer cells. | Methylation-specific PCR<br><br>Western blotting | An increased level of demethylated sites in the D-Loop of tumour cells is strongly associated with increased *MT-ND2* expression and mtDNA copy number. | [47] |
| The effect of methylation in the D-Loop, *MT-ND6* and *MT-CO1* on disease progression in simple steatosis (SS) and non-alcoholic steatohepatitis (NASH) | Methylation-specific PCR<br><br>qRT-PCR | Increased *MT-ND6* methylation and decreased *MT-ND6* protein levels in NASH compared to SS. Physical activity reduced *MT-ND6* methylation in NASH. | [86] |
| The effect of decreased S-adenosylmethionine (SAM) on mtDNA methylation in Down's Syndrome lymphoblastoid cells | LC-ESI-MS<br><br>LC-MS/MS | Decreased SAM availability in Down's syndrome lymphoblastoid cells reduces methyl uptake to mitochondria and leads to mtDNA hypomethylation. | [87] |
| The tissue specificity of DNMTs and 5-mC in the mitochondria in relation to ALS models. | IF<br><br>Pyrosequencing | Increased methylation at six cytosine sites in the *16S rRNA* gene in the spinal cord of an ALS mouse cell line.<br><br>Reduced levels of mtDNMT3a protein in skeletal muscle and spinal cord early in disease. | [88] |
| The effect of mtDNA methylation on mtDNA copy number in gastric cancer. | qRT-PCR<br><br>Pyrosequencing | Reduced mtDNA copy number levels in late clinicopathological stages.<br><br>Demethylation of mtDNA increases mtDNA copy number. | [89] |

**Table 2: A summary of the major issues and potential solutions in the field of mitochondrial epigenetics.** Abbreviations: nuclear mitochondrial pseudogenes (NUMTs), fluorescence-activated cell sorting (FACS), laser capture microdissection (LCM).

| Caveat | Potential Issues | Potential Solutions |
|---|---|---|
| Genetic Issues | ~~Wrongful~~ Incorrect determination of pseudogenes as mtDNA affects the validity of results. | 1. Isolate mitochondria before mtDNA extraction to avoid nuclear contamination<br>2. Specific primers designed with the consideration of NUMT amplification [90].<br>3. BLAST search to identify known NUMTs |
| | Genetic mutations in mtDNA may have specific associated methylation signatures. | Haplogroup and heteroplasmy studies should consider mtDNA methylation as a potential variable |
| Cell Specificity and Technical Issues | Different brain regions have differential methylation patterns and different cell population compositions. | 1. Larger samples sizes in specific brain subregions will improve statistical significance<br>2. FACS or LCM to separate cell types such as glia and neurons prior to analysis. |
| | Reduced methylation levels in mitochondria and variation in mtDNA copy number may increase noise and dilute signals. | Comparative analysis of techniques for their suitability to mitochondrial methylation studies should be considered. |
| | Bisulfite based methodologies cannot distinguish between 5-mC and 5-hmC. | Using oxidative bisulfite-sequencing allows for the distinction of 5-mC and 5-hmC at single base resolution[91]. |

**REFERENCE ANNOTATIONS**

- Bellizzi et al., 2013 –reported both 5-mC and 5-hmC in mtDNA at both CpG and non-CpG sites. The study also found that inactivation of DNMT1, DNMT3a and DNMT3b reduced CpG methylation levels markedly, but failed to impact non-CpG methylation to the same extent. As such, this study poses the question as to whether DNMT activity is important for mitochondrial methylation, or whether other factors may also be important.
- Chestnut et al., 2011 – found that DNMT3a was localized in the mitochondria of motor neurons, potentially indicating tissue-specific localization of this methyltransferase. This study also found 5-mC in mitochondria *in vivo*, suggesting that mitochondrial methylation may play a role in motor neurons.
- Dzitoyeva et al., 2012- found that the global levels of 5-hmC in mtDNA show an age-associated decrease in murine frontal cortex and that this was inversely correlated with the expression of some mitochondrial genes, suggesting a potential role of mtDNA methylation in aging.
- Lunnon et al., 2014 – used Illumina Infinium 450K methylation beadarray to demonstrate DNA methylation changes in AD cortex.
- Manczak et al., 2004 –investigated expression levels of mitochondrial-encoded ETC genes in AD, reporting decreased expression of complex I and increased expression of complex III and IV in early and late-stage disease.

- Shock et al., 2011 – reported 5-mC and 5-hmC in mtDNA, leading to a resurgence of interest in the field of mitochondrial epigenetics. This paper also identified an isoform of DNMT1, mtDNMT1, in the mitochondria.

## ACKNOWLEDGMENTS

This work was supported by an Alzheimer's Research UK pilot grant to KL and a NIH grant AG036039 to JM.

## REFERENCES

1. Aten, J.E., T.F. Fuller, A.J. Lusis, and S. Horvath, *Using genetic markers to orient the edges in quantitative trait networks: the NEO software.* BMC Syst Biol, 2008. **2**: p. 34.
2. Hardy, J.A. and G.A. Higgins, *Alzheimer's disease: the amyloid cascade hypothesis.* Science, 1992. **256**(5054): p. 184-5.
3. Hardy, J. and D.J. Selkoe, *The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics.* Science, 2002. **297**(5580): p. 353-6.
4. Mattson, M.P., B. Cheng, D. Davis, K. Bryant, I. Lieberburg, and R.E. Rydel, *beta-Amyloid peptides destabilize calcium homeostasis and render human cortical neurons vulnerable to excitotoxicity.* J Neurosci, 1992. **12**(2): p. 376-89.
5. Lacor, P.N., M.C. Buniel, P.W. Furlow, et al., *Abeta oligomer-induced aberrations in synapse composition, shape, and density provide a molecular basis for loss of connectivity in Alzheimer's disease.* J Neurosci, 2007. **27**(4): p. 796-807.
6. Kadowaki, H., H. Nishitoh, F. Urano, et al., *Amyloid beta induces neuronal cell death through ROS-mediated ASK1 activation.* Cell Death Differ, 2005. **12**(1): p. 19-24.
7. Manczak, M., B.S. Park, Y. Jung, and P.H. Reddy, *Differential expression of oxidative phosphorylation genes in patients with Alzheimer's disease: implications for early mitochondrial dysfunction and oxidative damage.* Neuromolecular Med, 2004. **5**(2): p. 147-62.

8. Ankarcrona, M., F. Mangialasche, and B. Winblad, *Rethinking Alzheimer's disease therapy: are mitochondria the key?* J Alzheimers Dis, 2010. **20 Suppl 2**: p. S579-90.
9. Lunnon, K., Z. Ibrahim, P. Proitsi, et al., *Mitochondrial dysfunction and immune activation are detectable in early Alzheimer's disease blood.* J Alzheimers Dis, 2012. **30**(3): p. 685-710.
10. Lunnon, K., M. Sattlecker, S. Furney, et al., *A blood gene expression marker of early Alzheimer's disease.* J Alzheimers Dis, 2013. **33**(3 ): p. 737-53.
11. Anderson, S., A.T. Bankier, B.G. Barrell, et al., *Sequence and organization of the human mitochondrial genome.* Nature, 1981. **290**(5806): p. 457-65.
12. Pradelli, L.A., M. Beneteau, and J.E. Ricci, *Mitochondrial control of caspase-dependent and -independent cell death.* Cell Mol Life Sci, 2010. **67**(10): p. 1589-97.
13. Chan, S.L., D. Liu, G.A. Kyriazis, P. Bagsiyao, X. Ouyang, and M.P. Mattson, *Mitochondrial uncoupling protein-4 regulates calcium homeostasis and sensitivity to store depletion-induced apoptosis in neural cells.* J Biol Chem, 2006. **281**(49): p. 37391-403.
14. Fu, W., A. Ruangkittisakul, D. MacTavish, G.B. Baker, K. Ballanyi, and J.H. Jhamandas, *Activity and metabolism-related Ca2+ and mitochondrial dynamics in co-cultured human fetal cortical neurons and astrocytes.* Neuroscience, 2013. **250**: p. 520-35.
15. Zhao, Y. and B. Zhao, *Oxidative stress and the pathogenesis of Alzheimer's disease.* Oxid Med Cell Longev, 2013. **2013**: p. 316523.
16. Devi, L. and M. Ohno, *Mitochondrial dysfunction and accumulation of the beta-secretase-cleaved C-terminal fragment of APP in Alzheimer's disease transgenic mice.* Neurobiol Dis, 2012. **45**(1): p. 417-24.
17. Pinto, M., A.M. Pickrell, H. Fukui, and C.T. Moraes, *Mitochondrial DNA damage in a mouse model of Alzheimer's disease decreases amyloid beta plaque formation.* Neurobiol Aging, 2013. **34**(10): p. 2399-407.
18. Swerdlow, R.H., J.M. Burns, and S.M. Khan, *The Alzheimer's disease mitochondrial cascade hypothesis.* J Alzheimers Dis, 2010. **20 Suppl 2**: p. S265-79.
19. Manczak, M., Y. Jung, B.S. Park, D. Partovi, and P.H. Reddy, *Time-course of mitochondrial gene expressions in mice brains: implications for mitochondrial dysfunction, oxidative damage, and cytochrome c in aging.* J Neurochem, 2005. **92**(3): p. 494-504.
20. Reddy, P.H., S. McWeeney, B.S. Park, et al., *Gene expression profiles of transcripts in amyloid precursor protein transgenic mice: up-regulation of mitochondrial metabolism and apoptotic genes is an early cellular change in Alzheimer's disease.* Hum Mol Genet, 2004. **13**(12): p. 1225-40.
21. Manczak, M. and P.H. Reddy, *Abnormal interaction of VDAC1 with amyloid beta and phosphorylated tau causes mitochondrial dysfunction in Alzheimer's disease.* Hum Mol Genet, 2012. **21**(23): p. 5131-46.
22. Du, H., L. Guo, F. Fang, et al., *Cyclophilin D deficiency attenuates mitochondrial and neuronal perturbation and ameliorates learning and memory in Alzheimer's disease.* Nat Med, 2008. **14**(10): p. 1097-105.
23. Gatz, M., C.A. Reynolds, L. Fratiglioni, et al., *Role of genes and environments for explaining Alzheimer disease.* Arch Gen Psychiatry, 2006. **63**(2): p. 168-74.
24. Lunnon, K. and J. Mill, *Epigenetic studies in Alzheimer's disease: current findings, caveats, and considerations for future studies.* Am J Med Genet B Neuropsychiatr Genet, 2013. **162B**(8): p. 789-99.
25. Mastroeni, D., A. McKee, A. Grover, J. Rogers, and P.D. Coleman, *Epigenetic differences in cortical neurons from a pair of monozygotic twins discordant for Alzheimer's disease.* PLoS One, 2009. **4**(8): p. e6617.
26. Mastroeni, D., A. Grover, E. Delvaux, C. Whiteside, P.D. Coleman, and J. Rogers, *Epigenetic changes in Alzheimer's disease: decrements in DNA methylation.* Neurobiol Aging, 2010. **31**(12): p. 2025-37.

27. Chouliaras, L., D. Mastroeni, E. Delvaux, et al., *Consistent decrease in global DNA methylation and hydroxymethylation in the hippocampus of Alzheimer's disease patients.* Neurobiol Aging, 2013. **34**(9): p. 2091-9.
28. Condliffe, D., A. Wong, C. Troakes, et al., *Cross-region reduction in 5-hydroxymethylcytosine in Alzheimer's disease brain.* Neurobiol Aging, 2014. **35**(8): p. 1850-4.
29. Bakulski, K.M., D.C. Dolinoy, M.A. Sartor, et al., *Genome-wide DNA methylation differences between late-onset Alzheimer's disease and cognitively normal controls in human frontal cortex.* J Alzheimers Dis, 2012. **29**(3): p. 571-88.
30. Lunnon, K., R. Smith, E.J. Hannon, et al., *Cross-tissue methylomic profiling in Alzheimer's disease implicates a role for cortex-specific deregulation of ANK1 in neuropathology.* Nat Neurosci, 2014([In Press]).
31. De Jager, P.L., G. Srivastava, K. Lunnon, et al., *Alzheimer's disease pathology is associated with early alterations in brain DNA methylation at ANK1, BIN1 and other loci.* Nat Med, 2014([Submitted]).
32. Nass, M.M., *Differential methylation of mitochondrial and nuclear DNA in cultured mouse, hamster and virus-transformed hamster cells. In vivo and in vitro methylation.* J Mol Biol, 1973. **80**(1): p. 155-75.
33. Dawid, I.B., *5-methylcytidylic acid: absence from mitochondrial DNA of frogs and HeLa cells.* Science, 1974. **184**(4132): p. 80-1.
34. Cummings, D.J., A. Tait, and J.M. Goddard, *Methylated bases in DNA from Paramecium aurelia.* Biochim Biophys Acta, 1974. **374**(1): p. 1-11.
35. Groot, G.S. and A.M. Kroon, *Mitochondrial DNA from various organisms does not contain internally methylated cytosine in -CCGG- sequences.* Biochim Biophys Acta, 1979. **564**(2): p. 355-7.
36. Shock, L.S., P.V. Thakkar, E.J. Peterson, R.G. Moran, and S.M. Taylor, *DNA methyltransferase 1, cytosine methylation, and cytosine hydroxymethylation in mammalian mitochondria.* Proc Natl Acad Sci U S A, 2011. **108**(9): p. 3630-5.
37. Alan, L., J. Zelenka, J. Jezek, A. Dlaskova, and P. Jezek, *Fluorescent in situ hybridization of mitochondrial DNA and RNA.* Acta Biochim Pol, 2010. **57**(4): p. 403-8.
38. Tauber, J., A. Dlaskova, J. Santorova, et al., *Distribution of mitochondrial nucleoids upon mitochondrial network fragmentation and network reintegration in HEPG2 cells.* Int J Biochem Cell Biol, 2013. **45**(3): p. 593-603.
39. Kaufman, B.A., N. Durisic, J.M. Mativetsky, et al., *The mitochondrial transcription factor TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures.* Mol Biol Cell, 2007. **18**(9): p. 3225-36.
40. Manev, H., S. Dzitoyeva, and H. Chen, *Mitochondrial DNA: A Blind Spot in Neuroepigenetics.* Biomol Concepts, 2012. **3**(2): p. 107-115.
41. Rebelo, A.P., S.L. Williams, and C.T. Moraes, *In vivo methylation of mtDNA reveals the dynamics of protein-mtDNA interactions.* Nucleic Acids Res, 2009. **37**(20): p. 6701-15.
42. Chestnut, B.A., Q. Chang, A. Price, C. Lesuisse, M. Wong, and L.J. Martin, *Epigenetic regulation of motor neuron cell death through DNA methylation.* J Neurosci, 2011. **31**(46): p. 16619-36.
43. Bellizzi, D., P. D'Aquila, T. Scafone, et al., *The control region of mitochondrial DNA shows an unusual CpG and non-CpG methylation pattern.* DNA Res, 2013. **20**(6): p. 537-47.
44. Ooi, S.K. and T.H. Bestor, *The colorful history of active DNA demethylation.* Cell, 2008. **133**(7): p. 1145-8.
45. Ito, S., A.C. D'Alessio, O.V. Taranova, K. Hong, L.C. Sowers, and Y. Zhang, *Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification.* Nature, 2010. **466**(7310): p. 1129-33.
46. Sun, Z., J. Terragni, J.G. Borgaro, et al., *High-resolution enzymatic mapping of genomic 5-hydroxymethylcytosine in mouse embryonic stem cells.* Cell Rep, 2013. **3**(2): p. 567-76.

47. Feng, S., L. Xiong, Z. Ji, W. Cheng, and H. Yang, *Correlation between increased ND2 expression and demethylated displacement loop of mtDNA in colorectal cancer.* Mol Med Rep, 2012. **6**(1): p. 125-30.
48. Phillips, A.C., A. Sleigh, C.J. McAllister, et al., *Defective mitochondrial function in vivo in skeletal muscle in adults with Down's syndrome: a 31P-MRS study.* PLoS One, 2013. **8**(12): p. e84031.
49. Coyle, J.T., M.L. Oster-Granite, and J.D. Gearhart, *The neurobiologic consequences of Down syndrome.* Brain Res Bull, 1986. **16**(6): p. 773-87.
50. Wisniewski, K.E., H.M. Wisniewski, and G.Y. Wen, *Occurrence of neuropathological changes and dementia of Alzheimer's disease in Down's syndrome.* Ann Neurol, 1985. **17**(3): p. 278-82.
51. Bradley-Whitman, M.A. and M.A. Lovell, *Epigenetic changes in the progression of Alzheimer's disease.* Mech Ageing Dev, 2013. **134**(10): p. 486-95.
52. Dzitoyeva, S., H. Chen, and H. Manev, *Effect of aging on 5-hydroxymethylcytosine in brain mitochondria.* Neurobiol Aging, 2012. **33**(12): p. 2881-91.
53. Smiraglia, D.J., M. Kulawiec, G.L. Bistulfi, S.G. Gupta, and K.K. Singh, *A novel role for mitochondria in regulating epigenetic modification in the nucleus.* Cancer Biol Ther, 2008. **7**(8): p. 1182-90.
54. Bellizzi, D., P. D'Aquila, M. Giordano, A. Montesanto, and G. Passarino, *Global DNA methylation levels are modulated by mitochondrial DNA variants.* Epigenomics, 2012. **4**(1): p. 17-27.
55. Takasugi, M., S. Yagi, K. Hirabayashi, and K. Shiota, *DNA methylation status of nuclear-encoded mitochondrial genes underlies the tissue-dependent mitochondrial functions.* BMC Genomics, 2010. **11**: p. 481.
56. Lang, B.F., M.W. Gray, and G. Burger, *Mitochondrial genome evolution and the origin of eukaryotes.* Annu Rev Genet, 1999. **33**: p. 351-97.
57. Antunes, A., J. Pontius, M.J. Ramos, S.J. O'Brien, and W.E. Johnson, *Mitochondrial introgressions into the nuclear genome of the domestic cat.* J Hered, 2007. **98**(5): p. 414-20.
58. Calabrese, F.M., D. Simone, and M. Attimonelli, *Primates and mouse NumtS in the UCSC Genome Browser.* BMC Bioinformatics, 2012. **13 Suppl 4**: p. S15.
59. Ovchinnikov, I.V., *Hominin evolution and gene flow in the Pleistocene Africa.* Anthropol Anz, 2013. **70**(2): p. 221-7.
60. Ricchetti, M., F. Tekaia, and B. Dujon, *Continued colonization of the human genome by mitochondrial DNA.* PLoS Biol, 2004. **2**(9): p. E273.
61. Triant, D.A. and J.A. DeWoody, *Molecular analyses of mitochondrial pseudogenes within the nuclear genome of arvicoline rodents.* Genetica, 2008. **132**(1): p. 21-33.
62. Mourier, T., A.J. Hansen, E. Willerslev, and P. Arctander, *The Human Genome Project reveals a continuous transfer of large mitochondrial fragments to the nucleus.* Mol Biol Evol, 2001. **18**(9): p. 1833-7.
63. Ho, S.Y. and M.T. Gilbert, *Ancient mitogenomics.* Mitochondrion, 2010. **10**(1): p. 1-11.
64. Thangaraj, K., M.B. Joshi, A.G. Reddy, A.A. Rasalkar, and L. Singh, *Sperm mitochondrial mutations as a cause of low sperm motility.* J Androl, 2003. **24**(3): p. 388-92.
65. Yao, Y.G., Q.P. Kong, A. Salas, and H.J. Bandelt, *Pseudomitochondrial genome haunts disease studies.* J Med Genet, 2008. **45**(12): p. 769-72.
66. Hirano, M., A. Shtilbans, R. Mayeux, et al., *Apparent mtDNA heteroplasmy in Alzheimer's disease patients and in normals due to PCR amplification of nucleus-embedded mtDNA pseudogenes.* Proc Natl Acad Sci U S A, 1997. **94**(26): p. 14894-9.
67. Davis, R.E., S. Miller, C. Herrnstadt, et al., *Mutations in mitochondrial cytochrome c oxidase genes segregate with late-onset Alzheimer disease.* Proc Natl Acad Sci U S A, 1997. **94**(9): p. 4526-31.

68. Iacobazzi, V., A. Castegna, V. Infantino, and G. Andria, *Mitochondrial DNA methylation as a next-generation biomarker and diagnostic tool.* Mol Genet Metab, 2013. **110**(1-2): p. 25-34.
69. Zhou, J., L. Liu, and J. Chen, *Method to purify mitochondrial DNA directly from yeast total DNA.* Plasmid, 2010. **64**(3): p. 196-9.
70. Wallace, D.C. and D. Chalkia, *Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease.* Cold Spring Harb Perspect Biol, 2013. **5**(11): p. a021220.
71. Hua, S., C. Lu, Y. Song, et al., *High levels of mitochondrial heteroplasmy modify the development of ovine-bovine interspecies nuclear transferred embryos.* Reprod Fertil Dev, 2012. **24**(3): p. 501-9.
72. Carrieri, G., M. Bonafe, M. De Luca, et al., *Mitochondrial DNA haplogroups and APOE4 allele are non-independent variables in sporadic Alzheimer's disease.* Hum Genet, 2001. **108**(3): p. 194-8.
73. Maruszak, A., J.A. Canter, M. Styczynska, C. Zekanowski, and M. Barcikowska, *Mitochondrial haplogroup H and Alzheimer's disease--is there a connection?* Neurobiol Aging, 2009. **30**(11): p. 1749-55.
74. Davies, M.N., M. Volta, R. Pidsley, et al., *Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood.* Genome Biol, 2012. **13**(6): p. R43.
75. Sanchez-Mut, J.V., E. Aso, N. Panayotis, et al., *DNA methylation map of mouse and human brain identifies target genes in Alzheimer's disease.* Brain, 2013. **136**(Pt 10): p. 3018-27.
76. Banati, R.B., R. Egensperger, A. Maassen, G. Hager, G.W. Kreutzberg, and M.B. Graeber, *Mitochondria in activated microglia in vitro.* J Neurocytol, 2004. **33**(5): p. 535-41.
77. Park, J., H. Choi, J.S. Min, et al., *Mitochondrial dynamics modulate the expression of pro-inflammatory mediators in microglial cells.* J Neurochem, 2013. **127**(2): p. 221-32.
78. Nestor, C., A. Ruzov, R. Meehan, and D. Dunican, *Enzymatic approaches and bisulfite sequencing cannot distinguish between 5-methylcytosine and 5-hydroxymethylcytosine in DNA.* Biotechniques, 2010. **48**(4): p. 317-9.
79. Clark, C., P. Palta, C.J. Joyce, et al., *A comparison of the whole genome approach of MeDIP-seq to the targeted approach of the Infinium HumanMethylation450 BeadChip((R)) for methylome profiling.* PLoS One, 2012. **7**(11): p. e50233.
80. How Kit, A., H.M. Nielsen, and J. Tost, *DNA methylation based biomarkers: practical considerations and applications.* Biochimie, 2012. **94**(11): p. 2314-37.
81. Liu, C., L. Liu, X. Chen, et al., *Decrease of 5-hydroxymethylcytosine is associated with progression of hepatocellular carcinoma through downregulation of TET1.* PLoS One, 2013. **8**(5): p. e62828.
82. Sandoval, J., J. Mendez-Gonzalez, E. Nadal, et al., *A prognostic DNA methylation signature for stage I non-small-cell lung cancer.* J Clin Oncol, 2013. **31**(32): p. 4140-7.
83. Anderson, C.M., J.L. Ralph, M.L. Wright, B. Linggi, and J.E. Ohm, *DNA methylation as a biomarker for preeclampsia.* Biol Res Nurs, 2013.
84. Chinnery, P.F., H.R. Elliott, G. Hudson, D.C. Samuels, and C.L. Relton, *Epigenetics, epidemiology and mitochondrial DNA diseases.* Int J Epidemiol, 2012. **41**(1): p. 177-87.
85. Byun, H.M., T. Panni, V. Motta, et al., *Effects of airborne pollutants on mitochondrial DNA methylation.* Part Fibre Toxicol, 2013. **10**: p. 18.
86. Pirola, C.J., T.F. Gianotti, A.L. Burgueno, et al., *Epigenetic modification of liver mitochondrial DNA is associated with histological severity of nonalcoholic fatty liver disease.* Gut, 2013. **62**(9): p. 1356-63.
87. Infantino, V., A. Castegna, F. Iacobazzi, et al., *Impairment of methyl cycle affects mitochondrial methyl availability and glutathione level in Down's syndrome.* Mol Genet Metab, 2011. **102**(3): p. 378-82.

88. Wong, M., B. Gertz, B.A. Chestnut, and L.J. Martin, *Mitochondrial DNMT3A and DNA methylation in skeletal muscle and CNS of transgenic mouse models of ALS.* Front Cell Neurosci, 2013. **7**: p. 279.
89. Wen, S.L., F. Zhang, and S. Feng, *Decreased copy number of mitochondrial DNA: A potential diagnostic criterion for gastric cancer.* Oncol Lett, 2013. **6**(4): p. 1098-1102.
90. Song, H., J.E. Buhay, M.F. Whiting, and K.A. Crandall, *Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified.* Proc Natl Acad Sci U S A, 2008. **105**(36): p. 13486-91.
91. Booth, M.J., T.W. Ost, D. Beraldi, et al., *Oxidative bisulfite sequencing of 5-methylcytosine and 5-hydroxymethylcytosine.* Nat Protoc, 2013. **8**(10): p. 1841-51.

**Appendix 2**

**Epigenetic regulation of mitochondrial function in neurodegenerative disease: new insights from advances in genomic technologies**

Matthew Devall[1], Janou Roubroeks[1,2], Jonathan Mill[1,3] Michael Weedon[1] and Katie Lunnon[1,*]

[1] Institute of Clinical and Biomedical Science, University of Exeter Medical School, University of Exeter, Devon, UK.

[2] Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience (MHENS), Maastricht University, Maastricht, The Netherlands.

[3] Institute of Psychiatry, Psychology & Neuroscience (IoPPN), King's College London, De Crespigny Park, London, UK.

[*] Corresponding author: Dr Katie Lunnon, University of Exeter Medical School, RILD, Barrack Road, University of Exeter, Devon, UK. UK. Tel: + 44 1392 408 298 Email address: k.lunnon@exeter.ac.uk

**Abstract**

The field of mitochondrial epigenetics has received increased attention in recent years and changes in mitochondrial DNA (mtDNA) methylation has been implicated in a number of diseases, including neurodegenerative diseases such as amyotrophic lateral sclerosis. However, current publications have been limited by the use of global or targeted methods of measuring DNA methylation. In this review, we discuss current findings in mitochondrial epigenetics as well as its potential role as a regulator of mitochondria within the brain. Finally, we summarize the current technologies best suited to capturing mtDNA methylation, and how a move towards whole epigenome sequencing of mtDNA may help to advance our current understanding of the field.

**Highlights**

- Variation in nuclear DNA methylation has been associated with neuropsychiatric and neurodegenerative diseases.
- The mitochondrial genome has been shown to be subject to regulation by DNA methylation.
- A potential role for alterations in mitochondrial DNA methylation in brain disorders has not yet been investigated.
- Next generation sequencing allows for an interrogation of the mitochondrial epigenome at single base resolution.

**Keywords**

Mitochondria; DNA methylation; Epigenetics; mtDNA; Alzheimer's disease; AD

**Introduction**

There is a resurgence of interest in the field of mitochondrial epigenetics, as mitochondrial dysfunction has been implicated in a variety of complex diseases including cancer [1, 2], amyotrophic lateral sclerosis [3, 4], and Alzheimer's disease (AD) [5]. Unlike modifications to nuclear DNA (ncDNA), which are now recognized as a mechanism through which the environment can influence biological processes and contribute to the development of a range of different disease phenotypes [6-8], mtDNA methylation has only recently started to be investigated for its role in health and disease.

Despite this recent interest, few studies of mtDNA methylation have extended past global studies of mtDNA methylation using immunocytochemical techniques or the investigation of single candidate genes within the mitochondrial genome [9]. Throughout the course of this review we consider the caveats of studying epigenetic variation in mitochondria, which are particularly pertinent for studies of brain disorders given that changes in cell type composition and mtDNA copy number variation, which occur frequently in neurodegenerative diseases [10, 11], could have profound effects on the interpretation of findings in the field of mitochondrial epigenetics. Finally we consider how recent advances in genomic technologies could be applied to allow accurate and quantifiable measurements of the mitochondrial epigenome in the brain.

## The Role of Mitochondria in the Brain

Mitochondria are the "powerhouse" of the cell and their dynamic nature allows for the organelle to be moved to areas of high ATP demand, for example cortical neurons require approximately 4.7 billion ATP per second to ensure continuous function [12]. Mitochondria are also implicated in many other important neurophysiological functions, for example synaptic mitochondria are believed to be involved in the regulation of neurotransmission by buffering extra intracellular $Ca^{2+}$ [13], making the mitochondria a vital organelle in the establishment of $Ca^{2+}$ homeostasis. Further, mitochondrial morphology and dynamics vary between synaptically immature and mature cortical neurons, with shorter mitochondrial lengths being observed in immature neurons allowing for increased movement and greater ability to meet the high energy demands of immature neurons, providing a key role for this organelle in neuronal development [14]. However, despite being the major site of reactive oxygen species (ROS) production, the mitochondria lack protective histones and defective mtDNA repair has been suggested to play a role in a number of neurodegenerative diseases [15]. As a by-product of ATP generation, ROS accumulation in areas of high ATP demand, such as post-mitotic neuronal cells, has been associated with neuronal loss [16]. Mutations in mitochondrial-encoded genes have been shown to cause a number of maternally heritable, monogenic diseases.

Most notably, mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes (MELAS) syndrome, a disease affecting multiple organs which can lead to the development of a number of syndromes, ranging from muscular weakness, fatigue and stroke-like episodes to dementia, epilepsy and diabetes in later stages [17]. As such, mtDNA variants have been shown to be an important driver of mitochondrial dysfunction and can have significant neuropathological effects [18].

**MtDNA methylation and regulation of the mitochondria**

The diverse range of mitochondrial functions coupled with their importance in the governance of cellular energy demands means that the expression of mitochondrial proteins requires sophisticated levels of fine-tuning. Although ~99% of mitochondrial proteins are encoded by the nuclear genome, the 16.569 Kb mitochondrial genome contains a total of 37 genes: 22 tRNAs, two rRNAs and 13 genes that encode for polypeptides along the electron transport chain (ETC) [19]. Within the nuclear genome, epigenetic processes mediate the reversible regulation of gene expression, occurring independently of DNA sequence, acting principally through chemical modifications to DNA and nucleosomal histone proteins. Epigenetic mechanisms orchestrate a diverse range of important neurobiological and cognitive processes in the brain and epigenetic modifications to the nuclear genome have been widely hypothesized to play a role in many neurological disorders, including Alzheimer's disease (AD) [11, 20, 21]. Given that mitochondrial dysfunction is a prominent feature of AD, we recently hypothesized that epigenetic modifications to the mitochondrial genome could be important in disease progression and pathology [5].

Despite rapid progress in the field of nuclear epigenetics, the field of mitochondrial epigenetics has received little attention since initial, but contradictory studies in the field were published in the 1970's [22-24]. These controversies continue, with one recent report concluding an absence of biologically significant levels of mtDNA methylation in four regions of mtDNA analyzed in human HEK293 cells and publically available data [25]. However, given the low sequencing depth of mtDNA in the publically available data used in this study (94x), and that multiple copies of the mtDNA genome are present in any given cell, it is possible that the true extent of mtDNA methylation was not determined. In contrary to this finding, both 5-methylcytosine (5-mC) and 5-hydroxymethylcystosine (5-hmC) have been identified in mammalian mitochondria from cell lines [26]. Further 5-mC has been found to co-localise in motor neurons with the mitochondrial marker superoxide dismutase 2 (SOD2).[4]. In addition, the identification of DNA methyltransferase 3a,

(DNMT3a), an enzyme that plays an important role in catalyzing the transfer of the methyl moiety from the methyl donor protein S-adenoysl methionine (SAM), in mitochondria isolated from a mouse motor neuron cell line [4], and in mitochondria isolated from mouse brain and spinal cord [26], suggests potential tissue-specific mtDNA methylation. As the mitochondrial genome lacks protective histones, mtDNA instead being packaged via nucleoid proteins [27], post-translational histone modifications will not have a direct role on mitochondrial function. MtDNA methylation, on the other hand, may play an important role in regulating mitochondrial function, either independently, or via a complex mechanism involving nuclear-mitochondrial crosstalk [5]. Recent studies have suggested that mtDNA may have patterns of DNA methylation similar to that of plant and fungi, with methylation of cytosines adjacent to adenosine and thymine, in addition to the traditional CpG dinucleotide [28]. Since we recently suggested that mtDNA methylation could play a role in AD and other disorders characterized by mitochondrial dysfunction [5], further studies have been published to support our hypothesis. For instance, one study assessed mtDNA methylation in platelets in subjects with cardiovascular disease across seven genes using pyrosequencing and identified four genes with significantly increased levels of mtDNA methylation compared to controls [9]. In particular one gene, *MT-C01*, demonstrated >18% difference between groups. Another study showed differences in DNA methylation of the D-Loop and *MT-RNR1* were concordant with methylation differences in the subtelomeric region *D4Z4* in umbilical cord blood as well as the fetal and maternal side of the placenta, suggesting a potentially common epigenetic signature in these tissues [29]. The same group also recently demonstrated a positive correlation between airborne particulate matter exposure and placental mtDNA methylation in the D-Loop and *MT-RNR1* using targeted bisulfite-pyrosequencing [30]. This study also showed that increased mtDNA methylation is accompanied by a reduction in mtDNA content, a potential marker of mitophagy, potentially indicating that alterations in mtDNA methylation of the D-Loop, a region highly involved in mtDNA transcription and replication, may lead to changes in mitochondrial biogenesis.

The role of mtDNA genetic variation on underlying mitochondrial function has become increasingly studied in recent years. MtDNA haplogroups and genetic variation have been found to have pronounced effects on mitochondrial functions, leading to altered ETC functions and ROS levels, which are associated with increased breast cancer risk [31]. However, given that a single mitochondrion can contain up to 10 copies of the mitochondrial genome [32], and as there are multiple mitochondria in any given cell, the extent of heteroplasmic mutations and their effect on mitochondrial function is far from clear [33]. Further, relatively few of the studies investigating mtDNA methylation have related identified changes to alterations in gene expression. One recent study showed significant hypermethylation of the D-Loop is negatively associated with expression of

three mitochondrial-encoded genes (*CYTB, ND6, COXII)* in retinal microvessels derived from diabetic retinopathy donors [34]. However, the group failed to find significant differences in the methylation of *CYTB* and conclude that, given the importance of the D-Loop in replication and transcription of the mitochondrial genome, changes in D-Loop methylation may result in transcriptional changes across the mitochondrial genome and potentially contribute to the pathogenesis of diabetic retinopathy. However, a recent *in vivo* study showed D-loop methylation did not correlate with the expression of mitochondrial-encoded genes during inflammation [35]. Briefly, it was found that lipopolysaccharide (LPS)-mediated inflammation in mice led to significant changes in mtDNA transcript levels which were reversed upon treatment with α-Lipoic acid (LA), an agent with anti-inflammatory properties. However, despite these transcriptional changes, neither treatment altered D-Loop methylation. Taken together, these two studies suggest that mitochondrial transcription may be driven by mtDNA methylation of the D-Loop; however, this phenomenon is not the sole driver of mitochondrial gene expression. Further studies should potentially employ a more holistic approach to investigating the relationship of mtDNA methylation, gene expression and potential nuclear-encoded effectors of mtDNA transcription. Further, the use of whole-methylome sequencing of the mitochondria may aid in the elucidation of base-specific mtDNA methylation changes which may be important in gene expression.

**The Bi-directionality of mitoepigenetics**

Interestingly, a number of nuclear factors are also known to affect mitochondrial function. A recent study found DNA methylation levels in the nuclear-encoded, mitochondrial-specific transcription factor DNA polymerase gamma A (*POLGA*) inversely correlate with mtDNA copy number in pluripotent and multipotent cell types [36]. The removal of mtDNA [37] or changes in mtDNA haplogroups [38] have been found to be associated with differences in ncDNA methylation levels, suggesting an important interplay between the two genomes and a bi-directionality to mitoepigenetics [39].

Cytoplasmic hybrids (cybrids) are an important cell type for studying mitochondrial function that contain identical nuclei to the parent cell, but different mtDNA, allowing for a controlled investigation into the role of mtDNA variants. One study found that cybrids with mtDNA derived from haplogroup J had consistently higher levels of global ncDNA methylation than haplogroup H, suggesting that mtDNA variants may play an important role in influencing ncDNA methylation [40]. This difference in DNA methylation was accompanied by significant decreases in the expression of six of the 11 nuclear-encoded genes assayed that are involved in DNA methylation and acetylation processes, as well as significant increases in the transcription of *MAT2B* and *MBD4*, both of which

are important for DNA methylation. Given the importance of these two processes in epigenomic regulation, this further illustrates how mtDNA genetic variation can impact on the nuclear genome. Another recent study investigating the mitochondrial 3243A>G heteroplasmy, which causes several clinical phenotypes including MELAS, showed that changes in heteroplasmy levels led to large and widespread changes in mitochondrial and transcriptional regulation, including changes in transcripts involved in DNA methylation and histone acetylation. Interestingly, high levels of this heteroplasmic variant (50-90%) resulted in ncDNA gene expression patterns that are similar to those seen in AD, Parkinson's disease and Huntington's disease. With alterations in ncDNA methylation in neurodegenerative disorders now well established [11, 41-44], future studies may aim to study mtDNA and ncDNA epigenomes in parallel to gain an understanding of the complete cellular epigenetic landscape.

## Other mitochondrial DNA modifications

The majority of mtDNA methylation research has focussed on the DNA methylation mark 5-mC. However a number of other epigenetic marks such as 5-hmC, 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC) have also been identified. Despite this, functional roles of 5-caC and 5-fC have yet to be elucidated, however evidence suggests that they are involved in the Tet-mediated oxidation pathway, which converts 5-mC back to unmodified cytosine [45]. In a recent study, 5-fC and 5-caC, but not 5-hmC were found to lead to modest blocks in DNA transcription when mediated by either T7 RNA polymerase (T7 RNAP), or by human RNA polymerase II in an *in vitro* transcription assay, which was also replicated in 2 human cell lines [46]. Interestingly, T7 RNAP shares a high degree of homology with mitochondrial polymerases, and given the identification of 5-hmC and Tet proteins in mitochondria [26, 47], this study hypothesized a mechanism by which mtDNA demethylation products may be important for the regulation of mtDNA transcription. However, at present, no study has investigated 5-fC and 5-caC levels in mtDNA and further investigation into the modulation of mtDNA methylation needs to be undertaken.

## Methodologies appropriate for epigenetic studies of mtDNA

Given the growing interest in studying epigenetic changes in mtDNA in a range of pathologies, there are a number of important considerations pertinent to this field of research. First, tissue-specific effects of DNA methylation mean that profiling a disease-relevant tissue is likely to be critical. Further, different cell types will have varying levels of mtDNA copy number and different DNA

methylation levels dependent upon cellular requirements and localized environments [5]. As such, studies investigating mtDNA methylation should consider mtDNA copy number as a potential confounder to subsequent analysis. This is a pertinent issue in mtDNA methylation analysis given that a number of studies have implicated an association between mtDNA methylation status and mtDNA copy number [48] and that tissue homogenates contain a number of different cell types, many with varying levels of mtDNA copy numbers. One potential way to solve this problem would be to first isolate different cell types from tissue homogenate using Fluorescence-activated cell sorting (FACS). However isolating sufficient levels of mitochondria from individual cell populations to yield sufficient mtDNA for methylation analyses seems unlikely with current technologies requiring high input amounts of DNA.

Another important consideration is the most appropriate method for studying mtDNA methylation. A host of different platforms have been developed that are capable of assessing DNA methylation in the nuclear genome. However, given the relative nascence of the field of mtDNA epigenetics, the appropriateness of the wide range of genomic technologies available has never been reviewed for its utility in this field. Advances in genomic technologies has seen ncDNA methylation studies move from profiling DNA methylation at a handful of candidate CpG sites to epigenome-wide association studies (EWAS). However, the current workhorse for such studies, the Illumina 450K methylation array, and its successor the Illumina EPIC methylation array, provide no coverage of the mitochondrial genome. Given that mtDNA methylation is now becoming increasingly investigated, an understanding of the technologies available to assess genome-wide changes in mitochondrial methylation is of upmost importance. An overview of genomic technologies that could be utilised to measure mtDNA methylation is provided in **Table 1**.

To capture an accurate and precise representation of potential mtDNA methylation changes, a technique should ideally be able to detect DNA modifications with low error rate and at single nucleotide resolution. This is more pertinent to studies of mtDNA epigenetics than in studies of ncDNA methylation given the low levels of mtDNA methylation identified in recent studies [30, 49]. Currently, the only published genome-wide study investigating mtDNA methylation utilized publically-available methylated DNA immunoprecipitation sequencing (MeDIP-Seq) data [50]. Whilst this study was able to identify spatio-temporal patterns of mtDNA methylation in a variety of cells and tissues, the coverage was relatively low (5x). Given sequencing and genomic biases, low levels of coverage may result in some bases not being sequenced adequately for analysis. Further, the resolution of MeDIP-Seq also remains a limitation. Despite high levels of concordance between

MeDIP-Seq data, which averages the methylation values across a region, and other technologies that focus on single CpG sites [51], it is possible that a correlation may not be observed across every region of the mitochondrial genome. For example, averaging methylation in regions as small as 100bp in the mitochondrial genome could include averaging methylation across many gene boundaries of tRNAs in some regions and could greatly mask the true methylation status of these genes. Although this study was able to utilize publically available data that had been generated to inexpensively assess ncDNA methylation, studies to specifically assess mtDNA methylation would ideally investigate DNA methylation at single base resolution. Whole genome-bisulfite sequencing (WGB-Seq) is another source of publically available data that could be exploited to investigate mtDNA methylation, however although this allows the interrogation of DNA methylation at single nucleotide resolution, data with an appropriate depth of sequencing is lacking due to the expense of sequencing the entire ~3,000 Mb human genome. An alternative, and more appropriate next generation sequencing (NGS), approach would specifically capture the entire mitochondrial genome prior to sequencing and due to its relatively small size would allow for inexpensive sequencing at single base resolution analysis. Currently, two custom capture methylation technologies are available from Agilent and Nimblegen, and coverage can be tailored to suit the researcher's needs, allowing for the interrogation of epigenetic variability within the mitochondrial genome. Despite costing considerably less than WGB-Seq, a targeted capture of the mitochondrial genome is still relatively expensive and requires a higher input of DNA.

Aside from cost versus coverage issues, sequencing the mitochondrial genome is associated with a number of unique caveats; most notably, the presence of nuclear mitochondria pseudogenes (*NUMT*s) in the nuclear genome. *NUMT*s are regions of the mitochondrial genome that, over an evolutionary period of time, have translocated to the nuclear genome and therefore share a high sequence homology with their mitochondrial paralogue. As such, failure to account for these regions has led to misinterpretations of genetic sequencing data [52, 53], and is one pitfall of using publically available data, although informatic approaches to account for these regions are being developed [50]. Ideally to avoid misinterpretation of data, mitochondria should be isolated from the tissue of choice prior to sequencing [5], and in epigenetic studies of brain disorders, this therefore requires the isolation of large quantities of mtDNA from small amounts of frozen post-mortem tissue [54].

Finally, the dependence on NGS technologies on PCR amplification prior to sequencing leads to a potential PCR bias and over-representation of some sequences. Given the tissue-specific nature of mitochondrial heteroplasmy, it is possible that these amplifications could influence findings. As such,

a move towards third generation sequencing systems such as the Pac-Bio RS II and Oxford's Nanopore is an appealing prospect. Furthermore, despite both technologies being associated with a high individual base call error rate at present, the Pac-Bio's long read based deep sequencing has been used to accurately map complex regions of the human genome [55] as well as mitochondrial genomes in other species [56-58] [59]. Interestingly, given that both technologies use native DNA, both technologies provide the possibility of simultaneous genetic and epigenetic analysis, including a range of DNA modifications [60-64]. The circular consensus reads of PacBio RS II could provide an interesting platform for the simultaneous identification of the four major epigenetic marks in real time at a high coverage. However, isolation of high quality mtDNA from the range of frozen tissues commonly used in nuclear epigenetics could prove to be difficult and would require further investigation. The continual development of more accurate calling algorithms and likely future cost reductions may therefore lead to these third generation sequencing technologies being the optimal choice for providing specific, deep and targeted analysis of the mitochondrial (epi)genome.

## Conclusions

Given the growing interest in studying mtDNA methylation in a range of brain disorders characterized by mitochondrial dysfunction, the advent of third generation sequencing technologies may allow for the accurate study of these small genomes in the near future. The circular consensus nature of Pac-Bio RS II potentially allows for the entire mitochondrial genome to be investigated from one long read, removing current NGS biases. However, this process would require isolation of intact mtDNA from a variety of potentially frozen and partially degraded tissues, which may prove challenging for researchers. Currently the use of NGS, particularly the more quantitative approaches of bisulfite-sequencing, may allow for the elucidation of mtDNA methylation and future studies should consider the utility of this approach versus candidate-based technologies such as pyrosequencing.

1. Feng, S., et al., *Correlation between increased ND2 expression and demethylated displacement loop of mtDNA in colorectal cancer.* Mol Med Rep, 2012. **6**(1): p. 125-30.
2. Wen, S.L., F. Zhang, and S. Feng, *Decreased copy number of mitochondrial DNA: A potential diagnostic criterion for gastric cancer.* Oncol Lett, 2013. **6**(4): p. 1098-1102.
3. Wong, M., et al., *Mitochondrial DNMT3A and DNA methylation in skeletal muscle and CNS of transgenic mouse models of ALS.* Front Cell Neurosci, 2013. **7**: p. 279.
4. Chestnut, B.A., et al., *Epigenetic regulation of motor neuron cell death through DNA methylation.* J Neurosci, 2011. **31**(46): p. 16619-36.
5. Devall, M., J. Mill, and K. Lunnon, *The mitochondrial epigenome: a role in Alzheimer's disease?* Epigenomics, 2014. **6**(6): p. 665-75.
6. Coppede, F., *The potential of epigenetic therapies in neurodegenerative diseases.* Front Genet, 2014. **5**: p. 220.
7. Rodriguez-Paredes, M. and M. Esteller, *Cancer epigenetics reaches mainstream oncology.* Nat Med, 2011. **17**(3): p. 330-9.
8. Robison, A.J. and E.J. Nestler, *Transcriptional and epigenetic mechanisms of addiction.* Nat Rev Neurosci, 2011. **12**(11): p. 623-37.
9. Baccarelli, A.A. and H.M. Byun, *Platelet mitochondrial DNA methylation: a potential new marker of cardiovascular disease.* Clin Epigenetics, 2015. **7**(1): p. 44.
10. Coskun, P., et al., *A mitochondrial etiology of Alzheimer and Parkinson disease.* Biochim Biophys Acta, 2012. **1820**(5): p. 553-64.
11. Lunnon, K. and J. Mill, *Epigenetic studies in Alzheimer's disease: Current findings, caveats, and considerations for future studies.* Am J Med Genet B Neuropsychiatr Genet, 2013.
12. Zhu, X.H., et al., *Quantitative imaging of energy expenditure in human brain.* Neuroimage, 2012. **60**(4): p. 2107-2117.
13. Tang, Y.G. and R.S. Zucker, *Mitochondrial involvement in post-tetanic potentiation of synaptic transmission.* Neuron, 1997. **18**(3): p. 483-491.
14. Chang, D.T.W. and I.J. Reynolds, *Differences in mitochondrial movement and morphology in young and mature primary cortical neurons in culture.* Neuroscience, 2006. **141**(2): p. 727-736.
15. Yang, J.L., et al., *Mitochondrial DNA damage and repair in neurodegenerative disorders.* DNA Repair, 2008. **7**(7): p. 1110-1120.
16. Uttara, B., et al., *Oxidative Stress and Neurodegenerative Diseases: A Review of Upstream and Downstream Antioxidant Therapeutic Options.* Current Neuropharmacology, 2009. **7**(1): p. 65-74.
17. El-Hattab, A.W., et al., *MELAS syndrome: Clinical manifestations, pathogenesis, and treatment options.* Mol Genet Metab, 2015.
18. Schaefer, A.M., et al., *Prevalence of mitochondrial DNA disease in adults.* Annals of Neurology, 2008. **63**(1): p. 35-39.
19. Boengler, K., G. Heusch, and R. Schulz, *Nuclear-encoded mitochondrial proteins and their role in cardioprotection.* Biochimica Et Biophysica Acta-Molecular Cell Research, 2011. **1813**(7): p. 1286-1294.
20. Chouliaras, L., et al., *Epigenetic regulation in the pathophysiology of Alzheimer's disease.* Prog Neurobiol, 2010. **90**(4): p. 498-510.
21. Lardenoije, R., et al., *The epigenetics of aging and neurodegeneration.* Prog Neurobiol, 2015.
22. Dawid, I.B., *5-methylcytidylic acid: absence from mitochondrial DNA of frogs and HeLa cells.* Science, 1974. **184**(4132): p. 80-1.
23. Nass, M.M., *Differential methylation of mitochondrial and nuclear DNA in cultured mouse, hamster and virus-transformed hamster cells. In vivo and in vitro methylation.* J Mol Biol, 1973. **80**(1): p. 155-75.

24. Groot, G.S. and A.M. Kroon, *Mitochondrial DNA from various organisms does not contain internally methylated cytosine in -CCGG- sequences.* Biochim Biophys Acta, 1979. **564**(2): p. 355-7.
25. Hong, E.E., et al., *Regionally specific and genome-wide analyses conclusively demonstrate the absence of CpG methylation in human mitochondrial DNA.* Mol Cell Biol, 2013. **33**(14): p. 2683-90.
26. Shock, L.S., et al., *DNA methyltransferase 1, cytosine methylation, and cytosine hydroxymethylation in mammalian mitochondria.* Proc Natl Acad Sci U S A, 2011. **108**(9): p. 3630-5.
27. Kaufman, B.A., et al., *The mitochondrial transcription factor TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures.* Mol Biol Cell, 2007. **18**(9): p. 3225-36.
28. Bellizzi, D., et al., *The control region of mitochondrial DNA shows an unusual CpG and non-CpG methylation pattern.* DNA Res, 2013. **20**(6): p. 537-47.
29. Janssen, B.G., et al., *Variation of DNA methylation in candidate age-related targets on the mitochondrial-telomere axis in cord blood and placenta.* Placenta, 2014. **35**(9): p. 665-72.
30. Janssen, B.G., et al., *Placental mitochondrial methylation and exposure to airborne particulate matter in the early life environment: An ENVIRONAGE birth cohort study.* Epigenetics, 2015. **10**(6): p. 536-44.
31. Bai, R.K., et al., *Mitochondrial genetic background modifies breast cancer risk.* Cancer Research, 2007. **67**(10): p. 4687-4694.
32. Wallace, D.C., *Mitochondrial-DNA Sequence Variation in Human-Evolution and Disease.* Proceedings of the National Academy of Sciences of the United States of America, 1994. **91**(19): p. 8739-8746.
33. Wallace, D.C. and D. Chalkia, *Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease.* Cold Spring Harb Perspect Med, 2013. **3**(10): p. a021220.
34. Mishra, M. and R.A. Kowluru, *Epigenetic Modification of Mitochondrial DNA in the Development of Diabetic Retinopathy.* Invest Ophthalmol Vis Sci, 2015. **56**(9): p. 5133-42.
35. Liu, Z., et al., *alpha-Lipoic acid attenuates LPS-induced liver injury by improving mitochondrial function in association with GR mitochondrial DNA occupancy.* Biochimie, 2015.
36. Lee, W., et al., *Mitochondrial DNA copy number is regulated by DNA methylation and demethylation of POLGA in stem and cancer cells and their differentiated progeny.* Cell Death Dis, 2015. **6**: p. e1664.
37. Smiraglia, D.J., et al., *A novel role for mitochondria in regulating epigenetic modification in the nucleus.* Cancer Biol Ther, 2008. **7**(8): p. 1182-90.
38. Bellizzi, D., et al., *Global DNA methylation levels are modulated by mitochondrial DNA variants.* Epigenomics, 2012. **4**(1): p. 17-27.
39. Manev, H., S. Dzitoyeva, and H. Chen, *Mitochondrial DNA: A Blind Spot in Neuroepigenetics.* Biomol Concepts, 2012. **3**(2): p. 107-115.
40. Atilano, S.R., et al., *Mitochondrial DNA variants can mediate methylation status of inflammation, angiogenesis and signaling genes.* Hum Mol Genet, 2015.
41. Lunnon, K., et al., *Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease.* Nat Neurosci, 2014. **17**(9): p. 1164-70.
42. De Jager, P.L., et al., *Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci.* Nat Neurosci, 2014. **17**(9): p. 1156-63.
43. Ng, C.W., et al., *Extensive changes in DNA methylation are associated with expression of mutant huntingtin.* Proc Natl Acad Sci U S A, 2013. **110**(6): p. 2354-9.
44. Moore, K., et al., *Epigenome-wide association study for Parkinson's disease.* Neuromolecular Med, 2014. **16**(4): p. 845-55.
45. Song, C.X. and C. He, *Potential functional roles of DNA demethylation intermediates.* Trends Biochem Sci, 2013. **38**(10): p. 480-4.

46. You, C., et al., *Effects of Tet-mediated oxidation products of 5-methylcytosine on DNA transcription in vitro and in mammalian cells.* Sci Rep, 2014. **4**: p. 7052.
47. Dzitoyeva, S., H. Chen, and H. Manev, *Effect of aging on 5-hydroxymethylcytosine in brain mitochondria.* Neurobiol Aging, 2012. **33**(12): p. 2881-91.
48. Gao, J., et al., *De-methylation of displacement loop of mitochondrial DNA is associated with increased mitochondrial copy number and nicotinamide adenine dinucleotide subunit 2 expression in colorectal cancer.* Mol Med Rep, 2015. **12**(5): p. 7033-8.
49. Byun, H.M., et al., *Effects of airborne pollutants on mitochondrial DNA methylation.* Part Fibre Toxicol, 2013. **10**: p. 18.
50. Ghosh, S., S. Sengupta, and V. Scaria, *Comparative analysis of human mitochondrial methylomes shows distinct patterns of epigenetic regulation in mitochondria.* Mitochondrion, 2014. **18**: p. 58-62.
51. Clark, C., et al., *A comparison of the whole genome approach of MeDIP-seq to the targeted approach of the Infinium HumanMethylation450 BeadChip((R)) for methylome profiling.* PLoS One, 2012. **7**(11): p. e50233.
52. Thangaraj, K., et al., *Sperm mitochondrial mutations as a cause of low sperm motility.* J Androl, 2003. **24**(3): p. 388-92.
53. Yao, Y.G., et al., *Pseudomitochondrial genome haunts disease studies.* J Med Genet, 2008. **45**(12): p. 769-72.
54. Devall, M., et al., *A comparison of mitochondrial DNA isolation methods in frozen post-mortem human brain tissue—applications for studies of mitochondrial genetics in brain disorders.* Biotechniques, 2015. **59**(4): p. 241-246.
55. Larsen, P.A., A.M. Heilman, and A.D. Yoder, *The utility of PacBio circular consensus sequencing for characterizing complex gene families in non-model organisms.* BMC Genomics, 2014. **15**: p. 720.
56. Simpson, L., et al., *Comparison of the Mitochondrial Genomes and Steady State Transcriptomes of Two Strains of the Trypanosomatid Parasite, Leishmania tarentolae.* PLoS Negl Trop Dis, 2015. **9**(7): p. e0003841.
57. Wolters, J.F., K. Chiu, and H.L. Fiumera, *Population structure of mitochondrial genomes in Saccharomyces cerevisiae.* BMC Genomics, 2015. **16**: p. 451.
58. Taniguti, L.M., et al., *Complete Genome Sequence of Sporisorium scitamineum and Biotrophic Interaction Transcriptome with Sugarcane.* PLoS One, 2015. **10**(6): p. e0129318.
59. Peng, Y., et al., *De novo genome assembly of the economically important weed horseweed using integrated data from multiple sequencing platforms.* Plant Physiol, 2014. **166**(3): p. 1241-54.
60. Flusberg, B.A., et al., *Direct detection of DNA methylation during single-molecule, real-time sequencing.* Nat Methods, 2010. **7**(6): p. 461-5.
61. Powers, J.G., et al., *Efficient and accurate whole genome assembly and methylome profiling of E. coli.* BMC Genomics, 2013. **14**: p. 675.
62. Lee, W.C., et al., *The complete methylome of Helicobacter pylori UM032.* BMC Genomics, 2015. **16**: p. 424.
63. Schreiber, J., et al., *Error rates for nanopore discrimination among cytosine, methylcytosine, and hydroxymethylcytosine along individual DNA strands.* Proc Natl Acad Sci U S A, 2013. **110**(47): p. 18910-5.
64. Laszlo, A.H., et al., *Detection and mapping of 5-methylcytosine and 5-hydroxymethylcytosine with nanopore MspA.* Proc Natl Acad Sci U S A, 2013. **110**(47): p. 18904-9.

## Appendix 3

**BENCHMARKS**

# A comparison of mitochondrial DNA isolation methods in frozen post-mortem human brain tissue: applications for studies of mitochondrial genetics in brain disorders.

Matthew Devall[1], Joe Burrage[1], Richard Caswell[1], Matthew Johnson[1], Claire Troakes[2], Safa Al-Sarraj[2], Aaron R Jeffries[1,2], Jonathan Mill[1,2], Katie Lunnon[1,3]

[1] University of Exeter Medical School, University of Exeter, Devon, UK.

[2] Institute of Psychiatry, King's College London, De Crespigny Park, London, UK.

[3] Corresponding author

Corresponding author: Dr Katie Lunnon, University of Exeter Medical School, RILD, Barrack Road, University of Exeter, Devon, UK. UK. Tel: + 44 1392 408 298 Email address: k.lunnon@exeter.ac.uk

**Keywords:** mitochondria, genetics, epigenetics, isolation, DNA, mtDNA, brain, post-mortem

**Word count for abstract:** 124 words

**Word count for method summary:** 36 words

**Word count for body of manuscript:** 1724 words

## Abstract

Given that many brain disorders are characterized by mitochondrial dysfunction, there is a growing interest in investigating genetic and epigenetic variation in mitochondrial DNA (mtDNA). One major caveat for such studies is the presence of nuclear-mitochondrial pseudogenes (*NUMTs*), which are regions of the mitochondrial genome that have been inserted into the nuclear genome over evolution and, if not accounted for, can confound genetic studies of mtDNA. Here we show the first systematic study to compare methods for isolating mtDNA from frozen post-mortem human brain tissue, and show that a commercial method from Miltenyi Biotech, which magnetically isolates mitochondria using antibodies raised against TOM22, gives a significant enrichment of mtDNA, and should be considered the method of choice for mtDNA studies in frozen brain tissue.


## Method Summary

Here we compare five methods of isolating mitochondrial DNA (mtDNA) to standard phenol-chloroform DNA extraction (that isolates nuclear DNA (ncDNA) and mtDNA) to determine the optimal method for enriching mtDNA from frozen post-mortem human brain tissue.

## Main Text

Mitochondria generate ATP, regulate calcium homeostasis [1, 2], mediate apoptosis [3], and produce reactive oxygen species (ROS). Mitochondrial dysfunction has been implicated in a number of diseases, including in the pathogenesis of brain disorders such as Alzheimer's disease [4-6]. Mitochondria are unique mammalian organelles in that they contain their own genome; the mitochondrial genome is ~16.6kb of circular DNA (mtDNA) [7], separate to the nuclear genome (ncDNA) and inherited in a maternal, non-Mendelian fashion. The mitochondrial genome comprises 37 genes; 13 encode for electron transport chain polypeptides, two for rRNAs and 22 for tRNAs. Because of its role in ROS production, mtDNA has a higher mutation rate (10-17 fold) than ncDNA [8]. Mutations in mtDNA are relatively common, with at least one in 200 healthy humans harboring a potentially pathogenic mtDNA mutation [9]. Indeed more than 300 point mutations in mtDNA are associated with disease risk and pathology in MitoMAP [10]. Interestingly, as each mitochondrion contains 2-10 copies of mtDNA and there are multiple mitochondria in any given cell, somatic mutations result in a mosaic of different mtDNA sequences within a given tissue. This phenomenon is known as mitochondrial heteroplasmy and is linked to various mitochondrial diseases [11]. Such heterogeneity is a potential confounder in studies of mitochondrial diseases, because inter- and intra-individual heteroplasmic variation can confuse the association between a haplogroup and its corresponding phenotype. Therefore, unlike studies of ncDNA variation, it is important to use the specific tissue of interest for etiological research. Another interesting feature of the mitochondria is that over evolution sequences of mtDNA have translocated to the nuclear genome. Traditional mitochondrial genetic research, and more recently studies of mitochondrial epigenetics, can be hampered by the presence of these nuclear-mitochondrial pseudogenes (*NUMTs*) as they share a high homology with their mitochondrial paralogs [12, 13]. Given the interest in studying mtDNA genetic and epigenetic changes in the pathology of brain diseases characterized by mitochondrial dysfunction, it is imperative that *NUMT*s are correctly accounted for [14].

The specific isolation of mitochondria prior to downstream processing is vital to fully exclude issues relating to *NUMT* contamination. For this purpose, a number of methods have been developed to

specifically isolate mtDNA, although few of these approaches have been specifically optimized for use on post-mortem tissue, a major resource in many epidemiological studies. In fact most studies investigating mtDNA use fresh animal tissue or cell lines. The insult of freezing tissue prior to isolation will potentially alter the effectiveness of these techniques and increase the risk of *NUMT* inclusion in downstream analysis. In this study we compared the effectiveness of five different mitochondrial isolation methods on post-mortem brain tissue using quantitative real-time PCR (qRT-PCR), to determine the optimal method for the specific enrichment of mtDNA, which was subsequently validated by next generation sequencing (NGS). We tested protocols based on A) Percoll gradients, B) linear DNA digestion, C) differential centrifugation, D) rapid differential centrifugation using a commercial kit and E) magnetic isolation of mitochondria using anti-TOM22 antibodies.

Method A was a modification of the method by Sims and Anderson (3). 150mg of tissue was dissociated using the gentleMACS dissociator (Miltenyi Biotech:130-093-235) and a mitochondrial extraction kit (Miltenyi Biotech:130-097-340). After removal of the nuclear fraction, the supernatant was spun at 13,000xg for 30 minutes at 4°C to form a crude mitochondrial pellet. The pellet was homogenized in a 12% Percoll solution and added above two layers (26% and 40%) of Percoll solution. Samples were spun at 30,700xg for 5 minutes at 4°C with the lower band containing the enriched mitochondrial fraction. Each mitochondrial fraction was diluted in four volumes of isolation buffer and centrifuged at 16,700xg for 10 minutes at 4°C to form a loose mitochondrial fraction. The supernatant was discarded, and mtDNA extracted using a DNA Mini kit (Qiagen:51304). Method B was based on the method by Zhou *et al* [15] that digests linear DNA but leaves circular DNA intact. 20ug genomic DNA (previously extracted using a phenol-chloroform protocol) was treated with 4µl lambda exonuclease (5 U/µl) (New England Biolabs:M0262S) and 12µl $RecJ_f$ (30 U/µl) (New England Biolabs:M0264S) in 400µl 1x lambda exonuclease buffer (New England Biolabs:B0262S) at 37°C for 16 hours. Samples were incubated at 65°C for 10 minutes to inactivate the enzymes and subsequently purified using a DNA Mini Kit (Qiagen:51304). Method C was based on the method by Clayton and Shadel [16]. 100mg brain tissue was homogenized in 1ml chilled homogenization buffer

(0.25M sucrose/10mM EDTA/30mM Tris-HCl, pH 7.5). The homogenate was centrifuged at 1,000xg for 15 minutes at 4$^{o}$C and supernatant removed. The pellet was re-homogenized in 600µl chilled homogenization buffer and spun at 1,000xg for 10 minutes at 4$^{o}$C. The supernatant was combined with the supernatant from the previous step and centrifuged at 12,000xg for 30 minutes at 4$^{o}$C to pellet the mitochondria. MtDNA was extracted using a DNA Mini Kit (Qiagen:51304). Method D was a modification of the method by Clayton and Shadel [16] and purchased as a commercial kit (Promokine:PK-CA577-K280). 100mg of brain tissue was homogenized with the reagents provided, according to the manufacturer's instructions. In Method E 200mg of tissue was dissociated using the gentleMACS dissociator and a mitochondrial extraction kit (Miltenyi Biotech:130-097-340) according to the manufacturer's protocol, with the exception of using an increased quantity of extraction buffer (40µl Solution 1 and 1ml 1x Solution 2). After homogenization the sample was spun at 200xg for 30 seconds and passed through a 70µm pre-separation filter (Miltenyi Biotec:130-095-823) and washed with Solution 3. The homogenate was spun at 500xg for 5 minutes at 4°C and the supernatant removed. The supernatant was magnetically labelled with 100µl anti-human TOM22 antibody-microbeads (Miltenyi Biotech:130-094-532) for 1 hour at 4°C under continuous agitation. The eluate was added to a LS column (Miltenyi Biotech:130-042-401) and placed in a MACS separator (Miltenyi Biotech:130-042-302) and washed. Upon removing the column from the magnetic field the mitochondria were pelleted by centrifugation at 13,000xg for 2 minutes at 4°C, washed in 1ml Storage Buffer and centrifuged at 13,000xg for 2 minutes at 4°C. The supernatant was discarded and mtDNA extracted using a DNA Mini Kit (Qiagen:51304). We compared these approaches to DNA we had previously isolated [17] using a phenol-chloroform protocol, which isolates both ncDNA and mtDNA We assessed the purity of each method using qRT-PCR as previously described (6). Briefly, the number of copies of mtDNA relative to ncDNA was determined by dividing the calculated number of copies of mtDNA (MT-CYB assay) by the calculated number of copies of ncDNA (B2M assay).

Our data showed that Method B (linear DNA digestion) gave the lowest purity (1,242 mtDNA copies/ncDNA copy), and Method E (magnetic-microbeads) the highest purity (14,654 mtDNA

copies/ncDNA copy) (**Figure 1A; Table 1**). Of particular interest to the study was the relative enrichment compared to a standard phenol-chloroform extraction (**Figure 1B**). The only method to show no enrichment was Method B (linear DNA digestion). All other techniques showed a positive enrichment of mtDNA compared to phenol-chloroform. Methods A (Percoll), C (differential centrifugation) and D (rapid differential centrifugation) all gave modest positive enrichments of 2.4-, 1.7- and 2.9- fold respectively. Although giving one of the lowest yields (3.2µg), the optimal method for enrichment relative to ncDNA was Method E (magnetic-microbeads), which gave a 10.7-fold enrichment, and was the only method to show significantly more copies of mtDNA/ncDNA copy compared to phenol-chloroform extraction ($P=1.88x10^{-3}$). Using this method we saw a significant enrichment of mtDNA/ncDNA compared to Methods A ($P=0.019$), B ($P=6.97x10^{-4}$) and C ($P=8.48x10^{-3}$). To validate our enrichment, two of the biological replicates from Method E were compared to a non-enriched standard from phenol-chloroform extraction using NGS. DNA samples were fragmented by sonication using a Bioruptor (Diagenode:UCD-200) to an average size of ~240bp. Sequencing libraries were prepared using the NEXTflex Rapid DNA-Seq kit (Bioo Scientific) and ligated to pre-indexed adapters (NEXTflex-96 DNA Barcodes; Bioo Scientific). Adapter-ligated DNA was amplified for 10 cycles using Herculase II Fusion DNA Polymerase (Agilent Technologies) and NEXTflex PCR primer mix, then pooled for sequencing on an Illumina HiSeq2500 (100bp paired-end, rapid run mode). Raw reads were quality and adaptor trimmed using TrimGalore! (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) before being aligned to GRCH37. Only high quality (Phred <20) reads, uniquely mapping to the genome were considered and total read counts were taken. For the non-enriched standard 1.1% of reads mapped to the mitochondrial genome, compared to an average of 18.7% (16.2% and 21.2% respectively) of reads with Method E, demonstrating an average 16.8 fold enrichment.

Despite Method E providing a greater enrichment than Method C (differential centrifugation) in mouse liver [18] and a similar enrichment to Percoll in an osteosarcoma cell line [19], higher levels of mitochondrial enrichment have previously been reported [20]. However, this method, like Method B

in our study, relies on the circular nature of the intact mitochondrial genome, which, whilst present in cell lines and blood, may be more degraded in frozen, archived brain. In the context of genomic studies of mtDNA, where the exclusion of *NUMTs* is imperative, the relative enrichment of mtDNA is of far greater importance than the yield. Thus, although we saw a lower yield with magnetic-microbeads (Method E) compared to the majority of methods tested, we observed the greatest purity with this method. The reasons for the observed greater enrichment of mtDNA/ncDNA relative to Percoll (Method A) in our study compared to the analysis by Hornig-Do and colleagues potentially may include i) our use of qRT-PCR, which is more sensitive than western blot, ii) our use of frozen samples, rather than fresh samples and iii) the use of brain, rather than a cell line, as brain has high levels of mitochondria. To our knowledge our study represents the first to systematically compare and contrast methods for isolating mtDNA from small quantities of frozen, post-mortem human brain. Our findings suggest that magnetic-microbeads provide a significant enrichment of mtDNA compared to any other method tested. This may be due to a number of reasons, for example the automated homogenization of tissue in this protocol could provide a more consistent and gentle approach than other techniques and the use of magnetically labelled antibodies provides a specific capture of intact mitochondria, which may also contain less degraded mtDNA. We recommend that given the current interest in studying the mitochondrial genome in human brain, that the magnetic-microbead method from Miltenyi Biotech is used prior to DNA extraction to minimize the inclusion of NUMTs in downstream analyses.

## Author Contributions

M.D. J.M. and K.L. conceived the idea for the study. M.D., J.B., R.C., M.J. and A.R.J. performed the experiments. C.T. and S.A-S provided tissue from the London Neurodegenerative Disease Brain Bank. M.D. and K.L. analyzed the data and drafted the manuscript. All authors approved of the final manuscript prior to submission.

## Acknowledgements

This work was funded by an Alzheimer's Research UK grant to K.L. (ARUK-PPG2013A-5). The authors would like to thank the London Neurodegenerative Disease Brain Bank and Brains for Dementia Research (BDR) and their donors for provision of tissue for the study.

## Competing Interests

The authors declare that they have no conflicts of interest in regard to this work.

## References

1. Fu, W., et al., *Activity and metabolism-related Ca2+ and mitochondrial dynamics in co-cultured human fetal cortical neurons and astrocytes.* Neuroscience, 2013. **250**: p. 520-35.
2. Chan, S.L., et al., *Mitochondrial uncoupling protein-4 regulates calcium homeostasis and sensitivity to store depletion-induced apoptosis in neural cells.* J Biol Chem, 2006. **281**(49): p. 37391-403.
3. Pradelli, L.A., M. Beneteau, and J.E. Ricci, *Mitochondrial control of caspase-dependent and -independent cell death.* Cell Mol Life Sci, 2010. **67**(10): p. 1589-97.
4. Devi, L. and M. Ohno, *Mitochondrial dysfunction and accumulation of the beta-secretase-cleaved C-terminal fragment of APP in Alzheimer's disease transgenic mice.* Neurobiol Dis, 2012. **45**(1): p. 417-24.
5. Pinto, M., et al., *Mitochondrial DNA damage in a mouse model of Alzheimer's disease decreases amyloid beta plaque formation.* Neurobiol Aging, 2013. **34**(10): p. 2399-407.
6. Lunnon, K., et al., *Mitochondrial dysfunction and immune activation are detectable in early Alzheimer's disease blood.* J Alzheimers Dis, 2012. **30**(3): p. 685-710.
7. Anderson, S., et al., *Sequence and organization of the human mitochondrial genome.* Nature, 1981. **290**(5806): p. 457-65.

8. Tuppen, H.A., et al., *Mitochondrial DNA mutations and human disease.* Biochim Biophys Acta, 2010. **1797**(2): p. 113-28.
9. Elliott, H.R., et al., *Pathogenic mitochondrial DNA mutations are common in the general population.* Am J Hum Genet, 2008. **83**(2): p. 254-60.
10. Kogelnik, A.M., et al., *MITOMAP: a human mitochondrial genome database.* Nucleic Acids Res, 1996. **24**(1): p. 177-9.
11. Wallace, D.C. and D. Chalkia, *Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease.* Cold Spring Harb Perspect Biol, 2013. **5**(11): p. a021220.
12. Hazkani-Covo, E., R.M. Zeller, and W. Martin, *Molecular Poltergeists: Mitochondrial DNA Copies (numts) in Sequenced Nuclear Genomes.* Plos Genetics, 2010. **6**(2).
13. Bensasson, D., et al., *Mitochondrial pseudogenes: evolution's misplaced witnesses.* Trends in Ecology & Evolution, 2001. **16**(6): p. 314-321.
14. Devall, M., J. Mill, and K. Lunnon, *The mitochondrial epigenome: a role in Alzheimer's disease?* Epigenomics, 2014. **6**(6): p. 665-75.
15. Zhou, J., L. Liu, and J. Chen, *Method to purify mitochondrial DNA directly from yeast total DNA.* Plasmid, 2010. **64**(3): p. 196-9.
16. Clayton, D.A. and G.S. Shadel, *Isolation of mitochondria from tissue culture cells.* Cold Spring Harb Protoc, 2014. **2014**(10): p. pdb prot080002.
17. Lunnon, K., et al., *Cross-tissue methylomic profiling in Alzheimer's disease implicates a role for cortex-specific deregulation of ANK1 in neuropathology.* Nat Neurosci, 2014. **Sept; 17**(9): p. 1164-70.
18. Franko, A., et al., *Efficient isolation of pure and functional mitochondria from mouse tissues using automated tissue disruption and enrichment with anti-TOM22 magnetic beads.* PLoS One, 2013. **8**(12): p. e82392.
19. Hornig-Do, H.T., et al., *Isolation of functional pure mitochondria by superparamagnetic microbeads.* Anal Biochem, 2009. **389**(1): p. 1-5.
20. Jayaprakash, A.D., et al., *Stable heteroplasmy at the single-cell level is facilitated by intercellular exchange of mtDNA.* Nucleic Acids Res, 2015. **43**(4): p. 2177-87.

232

**Table 1. An overview of starting material for each isolation technique and resulting yield.**

| Method | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| Overview | Percoll | DNAse digestion | Differential centrifugation | Rapid differential centrifugation | Magnetic microbeads (anti-TOM22) | Phenol-chloroform |
| Quantity of starting Material | 150mg (tissue) | 20µg (DNA) | 100mg (tissue) | 100mg (tissue) | 200mg (tissue) | 100mg (tissue) |
| Number of samples | 4 | 6 | 4 | 3 | 5 | 5 |
| Average concentration (ng/µl) DNA collected (±SEM) | 162.4 (11.6) | 40.3 (10.1) | 23.7 (8.0) | 257 (138.4) | 32.4 (7.0) | N/A |
| Average yield (µg) DNA collected (±SEM) | 8.1 (0.58) | 2.0 (0.50) | 4.7 (1.60) | 13.7 (6.35) | 3.2 (0.70) | N/A |
| Average copies of mtDNA (±SEM) | 41,530 (15,468) | 12,092,501 (7,742,804) | 28,886,594 (12,405,939) | 403,381 (194,557) | 5,321,960 (1,246,724) | 459,264 (751,075) |
| Average copies of ncDNA (±SEM) | 155 (85) | 8,768 (3,731) | 20,746 (11,256) | 720 (449) | 402 (86) | 4,011 (585) |
| Average ratio mtDNA/ncDNA (±SEM) | 3,337 (1,988) | 1,242 (309) | 2,270 (960) | 3,949 (3,424) | 14,654 (2,922) | 1,367 (35) |
| Fold Enrichment mtDNA/ncDNA relative to phenol-chloroform (P-value) | 2.44 (0.297) | 0.91 (0.725) | 1.66 (0.320) | 2.89 (0.342) | 10.72 ($1.88x10^{-3}$) | N/A |

In brief we compared five methods of isolating mtDNA in post-mortem human brain tissue; (A) discontinuous Percoll gradient, (B) DNase digestion of linear DNA, (C) differential centrifugation, (D) rapid (commercial) mitochondrial isolation via differential centrifugation (E) magnetic labelling and pull-down of mitochondria using an antibody to TOM22. We compared the yield and enrichment to a non-enriched standard (phenol-chloroform) using an unpaired two-tailed t-test.

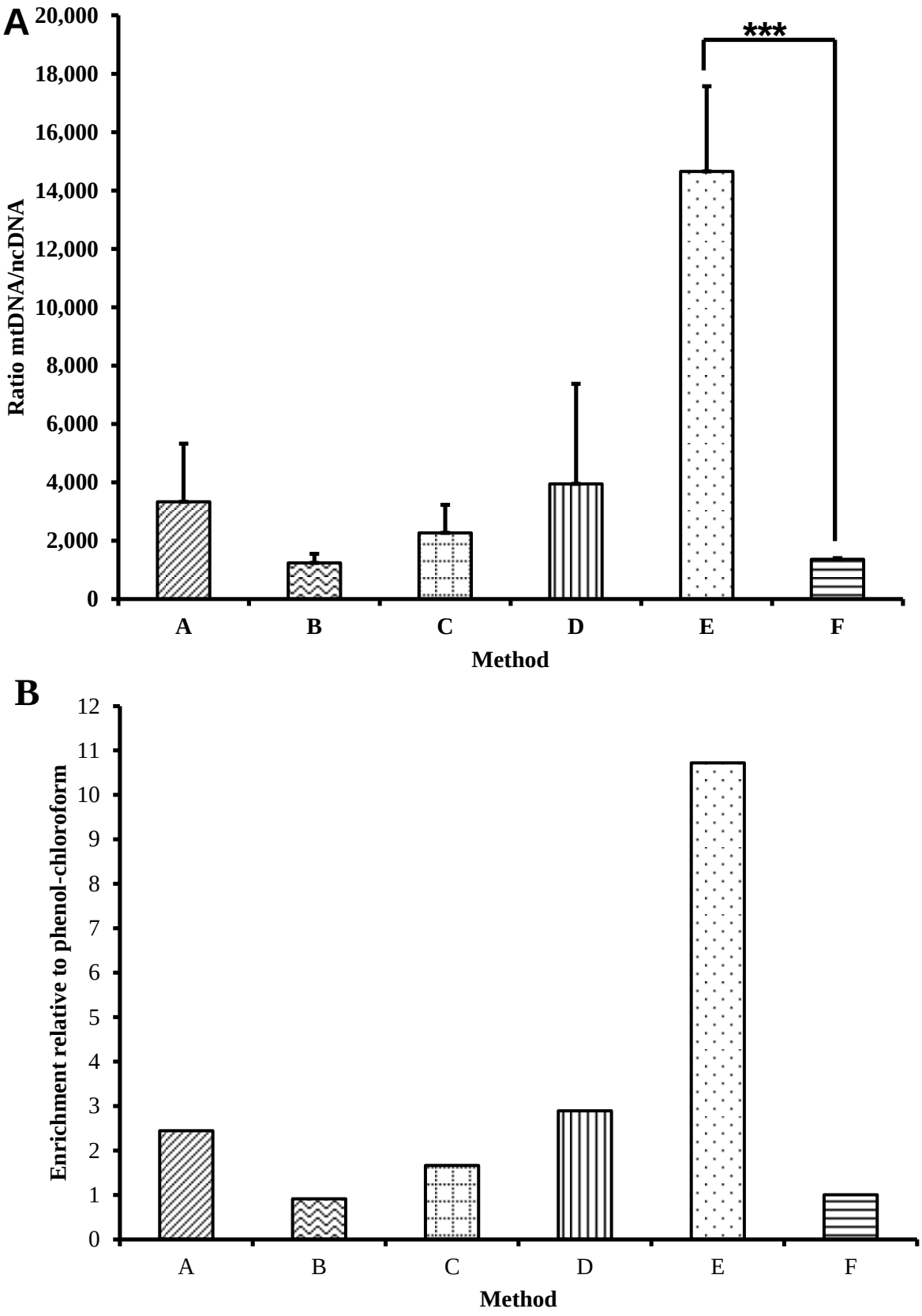



**Figure 1: Enrichment of mtDNA relative to ncDNA.** In total five methods (Percoll (A), DNase digestion (B), differential centrifugation (C), rapid differential centrifugation (D) and magnetic microbeads (E) were compared to a non-enriched standard (phenol-chloroform (F)). Shown is the ratio of mtDNA/ncDNA (±SEM) (**Fig 1A**) and the relative enrichment compared to phenol-chloroform (**Fig 1B**). *** = P<0.005

**Appendix 4**

Devall et al. Clinical Epigenetics (2017) 9:47
DOI 10.1186/s13148-017-0337-3

Clinical Epigenetics

SHORT REPORT Open Access

# Regional differences in mitochondrial DNA methylation in human post-mortem brain tissue

Matthew Devall[1], Rebecca G. Smith[1], Aaron Jeffries[1,2], Eilis Hannon[1], Matthew N. Davies[3], Leonard Schalkwyk[4], Jonathan Mill[1,2], Michael Weedon[1] and Katie Lunnon[1*]

## Abstract

**Background:** DNA methylation is an important epigenetic mechanism involved in gene regulation, with alterations in DNA methylation in the nuclear genome being linked to numerous complex diseases. Mitochondrial DNA methylation is a phenomenon that is receiving ever-increasing interest, particularly in diseases characterized by mitochondrial dysfunction; however, most studies have been limited to the investigation of specific target regions. Analyses spanning the entire mitochondrial genome have been limited, potentially due to the amount of input DNA required. Further, mitochondrial genetic studies have been previously confounded by nuclear-mitochondrial pseudogenes. Methylated DNA Immunoprecipitation Sequencing is a technique widely used to profile DNA methylation across the nuclear genome; however, reads mapped to mitochondrial DNA are often discarded. Here, we have developed an approach to control for nuclear-mitochondrial pseudogenes within Methylated DNA Immunoprecipitation Sequencing data. We highlight the utility of this approach in identifying differences in mitochondrial DNA methylation across regions of the human brain and pre-mortem blood.

**Results:** We were able to correlate mitochondrial DNA methylation patterns between the cortex, cerebellum and blood. We identified 74 nominally significant differentially methylated regions ($p < 0.05$) in the mitochondrial genome, between anatomically separate cortical regions and the cerebellum in matched samples ($N = 3$ matched donors). Further analysis identified eight significant differentially methylated regions between the total cortex and cerebellum after correcting for multiple testing. Using unsupervised hierarchical clustering analysis of the mitochondrial DNA methylome, we were able to identify tissue-specific patterns of mitochondrial DNA methylation between blood, cerebellum and cortex.

**Conclusions:** Our study represents a comprehensive analysis of the mitochondrial methylome using pre-existing Methylated DNA Immunoprecipitation Sequencing data to identify brain region-specific patterns of mitochondrial DNA methylation.

**Keywords:** 5-mC, 5-Methylcytosine, Blood, Brain, DNA methylation, Epigenetics, MeDIP-seq, Mitochondria, *NUMT*s

## Introduction

Mitochondria are unique organelles in that they have their own circular genome, approximately 16.6 kb in size [1]. Mitochondrial DNA (mtDNA) consists of 37 genes, 22 encoding for transfer RNAs (tRNAs), two for ribosomal RNAs (rRNAs) and 13 encoding for proteins important in the electron transport chain. Each of these 13 proteins are directly involved in the regulation of cellular respiration, generating the majority of ATP required for the process. However, mitochondria have an array of other important cellular roles such as calcium homeostasis [2] and neural stem cell differentiation [3]. As such, abnormal mitochondrial function, dynamics and trafficking have been associated with a number of brain disorders including Alzheimer's disease [4, 5], schizophrenia [6], bipolar disorder [7] and major depressive disorder [8].

* Correspondence: k.lunnon@exeter.ac.uk
[1]University of Exeter Medical School, RILD, University of Exeter, Barrack Road, Devon, UK
Full list of author information is available at the end of the article

Epigenetic processes mediate the reversible regulation of gene expression, occurring independently of DNA sequence variation, acting principally through chemical modifications to DNA and nucleosomal histone proteins and orchestrate a diverse range of important physiological functions. DNA methylation is the best characterized and most stable epigenetic modification modulating the transcription of mammalian genomes and, because it can be robustly assessed using existing genomic DNA resources, is the focus of most human epidemiological epigenetic research to date [9]. The most widely used method for epigenome-wide analysis of DNA methylation is the Illumina 450K methylation array, and a number of studies have recently shown differential DNA methylation of the nuclear genome (ncDNA), between different tissue types [10–12] and also in a range of complex diseases, from brain disorders such as Alzheimer's disease [13–15] and schizophrenia [16, 17], to systemic diseases such as type 2 diabetes [18] and Crohn's disease [19]. However, with no representation of the mitochondrial genome on this platform, as well as a lack of analysis on other genome-wide platforms, the role of mtDNA methylation has been largely neglected [20, 21].

Since the identification of 5-methylcytosine (5-mC) in mitochondria, research into mtDNA methylation as an independent and potentially relevant mark has received more regular attention [22, 23]. However, most research is either focussed on low resolution, global DNA methylation, or candidate gene DNA methylation changes using techniques such as bisulfite pyrosequencing [20]. These recent publications have indicated that differences in mtDNA methylation are present in a variety of different phenotypes [24–29] and may have potential utility as a biomarker [30]. In addition, a recent study has explored the use of Methylated DNA Immunoprecipitation Sequencing (MeDIP-seq) to investigate changes in mtDNA methylation across 39 cell lines and tissues from publicly available data [31]. At present, genome-wide sequencing technologies have not yet been used to interrogate alterations in the mtDNA methylome across tissues in the same individuals.

A high proportion of current, publicly available, genome-wide DNA methylation data has been generated through the use of MeDIP-seq, a method designed to interrogate genome-wide changes in methylation at high throughput and low cost [32]. However, given the presence of nuclear-mitochondrial pseudogenes (*NUMT*s), regions of the nuclear genome that share a high sequence homology with their mitochondrial paralogue [33, 34], mitochondrial reads are often discarded from further analysis. The development of bioinformatic pipelines to investigate regions of differential mtDNA methylation from whole genome data would provide a novel way in which to interrogate the mtDNA methylome in publicly available data. Here, we control for the presence of *NUMT*s in a previously published MeDIP-seq dataset, to investigate differential DNA methylation across the mitochondrial genome in human post-mortem brain samples.

## Results

### MtDNA methylation patterns are correlated between the cortex, cerebellum and blood

To date, no study has investigated differences in mtDNA methylation across different matched regions of human brain and blood samples. Our sample (Table 1) consisted of MeDIP-seq data from three individuals, free of any neuropathology and neuropsychiatric disease, for five different regions of the cortex (Brodmann areas (BA) 8, 9 and 10, superior temporal gyrus (STG) and entorhinal cortex (ECX)), the cerebellum (CER) and pre-mortem blood [35]. Given that MeDIP-seq data has been generated from standardly extracted total genomic DNA and thus contains a mixture of ncDNA and mtDNA [36], we initially controlled for regions of high sequence homology between the two genomes within our data by realigning mtDNA reads to a series of custom reference genomes using an in-house pipeline (see the Methods section) to specifically analyze mtDNA methylation (Fig. 1). Briefly, after an initial alignment to the GRCH37 reference genome using BWA, uniquely mapped reads were extracted and aligned to a custom GRCH37 reference genome not containing the mitochondrial sequence. Reads that did not map to this custom genome were found to share less homology with the nuclear genome and were taken forward and realigned once more to the full reference genome. Initially, we were interested to investigate whether changes in mtDNA across the mitochondrial genome were highly correlated between different tissue types. Using principal component analysis (PCA), we found that mtDNA methylation patterns are highly correlated between different cortical regions ($r > 0.99$, $p < 2.2E{-}16$), with a slightly weaker correlation between the cerebellum and cortex ($r > 0.97$, $p < 2.2E{-}16$) (Fig. 2). Due to the small number of blood samples available, deriving a significance level for the correlations between the cerebellum and blood could not be made.

**Table 1** Demographic information

| Individual | Age at death (years) | Age at bloods sampled (years) | Post-mortem delay (hours) | Gender |
|---|---|---|---|---|
| 1 | 82 | 79 | 43 | Female |
| 2 | 92 | N/A | 17 | Female |
| 3 | 78 | 78 | 10 | Male |

MeDIP-seq data was available from post-mortem brain samples obtained from three individuals free of any neuropathology and neuropsychiatric disease. Data was available for five different regions of the cortex (Brodmann areas (BA) 8, 9 and 10, superior temporal gyrus (STG), entorhinal cortex (ECX), the cerebellum (CER) and pre-mortem blood (BLD). MeDIP-seq data was available for all individuals from cortical and cerebellar samples; however, blood MeDIP-seq data was not available for individual 2. Data is freely available to download from http://epigenetics.iop.kcl.ac.uk/brain

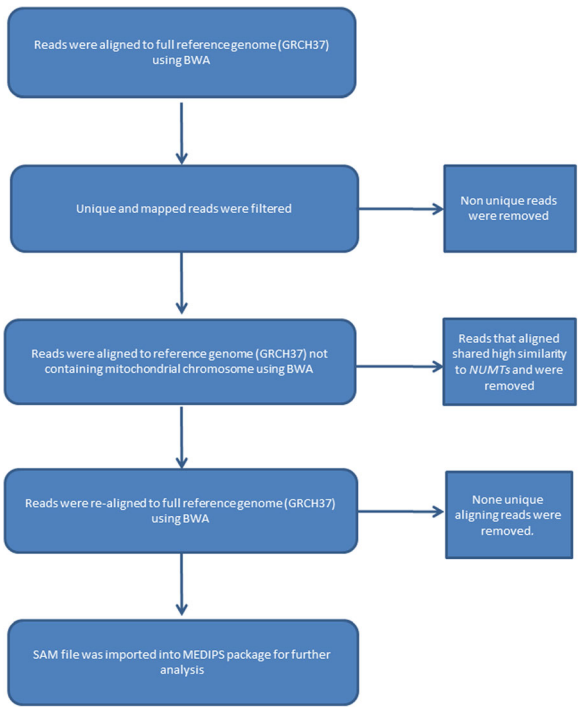



**Fig. 1** Overview of the analysis pipeline

Instead, in an attempt to explore the similarity between matched blood and cerebellum samples, the direction of differential methylation with respect to the cortex was used. Here, we found that 93.1% of the windows analyzed in the cerebellum and blood had the same direction of methylation difference with respect to the cortex, further suggesting a strong correlation between the two tissue types.

### Differentially methylated regions of the mitochondrial genome can be identified between anatomically distinct cortical regions and the cerebellum

Having identified correlated mtDNA methylation patterns across different brain regions, we were interested to investigate whether we could identify differentially methylated regions (DMRs) in the mitochondrial genome between different regions of the cortex and cerebellum. To identify such tissue-specific DMRs within the mitochondrial genome, paired *t* tests were performed across matched cortical and cerebellum samples at 100 bp windows across the mitochondrial genome (see the Methods section). In total, we identified 74 nominally significant DMRs ($p < 0.05$) between the five individual cortical regions and the cerebellum (Table 2; Fig. 3). Of these DMRs, seven (Table 2, bold face) were found to be present across all prefrontal cortex areas (BA8, BA9, BA10). Furthermore, the direction of methylation

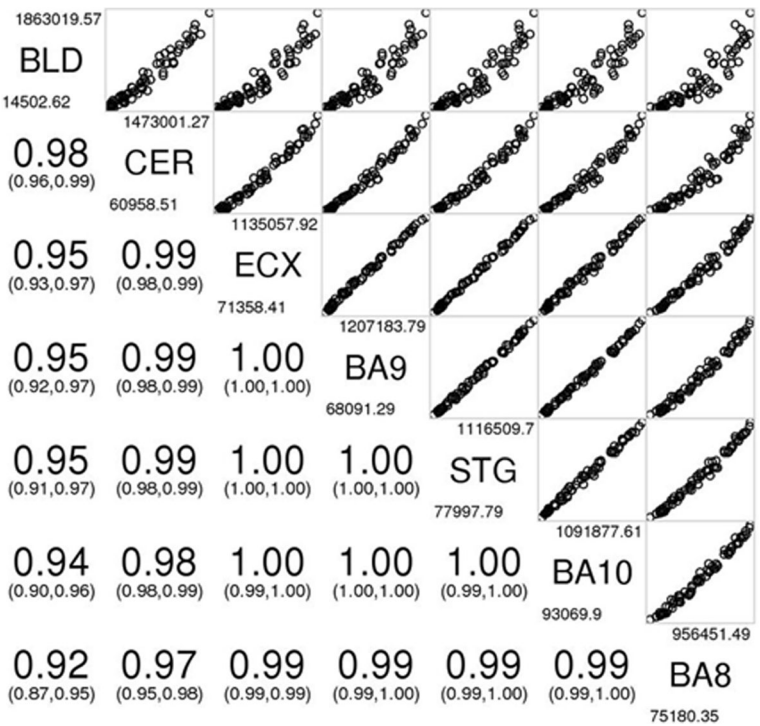



**Fig. 2** MtDNA methylation patterns are correlated between the cortex, cerebellum and blood. Samples were ordered based upon the similarity of their principal components, RPKM values, with *r* calculated for the correlations between each tissue. *BLD* blood, *BA8* Brodmann area 8, *BA9* Brodmann area 9, *BA10* Brodmann area 10, *CER* cerebellum, *CTX* cortex, *ECX* entorhinal cortex, *STG* superior temporal gyrus

**Table 2** List of DMRs identified between five anatomically discreet cortical regions and cerebellum

| Start (bp) | Stop (bp) | Gene(s) | BA8 | | BA9 | | BA10 | | EC | | STG | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | *p* value | Δ RPKM | *p* value | Δ RPKM | *p* value | Δ RPKM | *p* value | Δ RPKM | *p* value | Δ RPKM |
| 1 | 100 | D-Loop | - | - | - | - | - | - | - | - | - | - |
| 101 | 200 | D-Loop | - | - | - | - | - | - | - | - | - | - |
| 201 | 300 | D-Loop | - | - | - | - | - | - | - | - | - | - |
| 301 | 400 | D-Loop | - | - | - | - | - | - | - | - | - | - |
| 401 | 500 | D-Loop | - | - | - | - | - | - | - | - | - | - |
| 501 | 600 | MT-TF | - | - | - | - | - | - | - | - | - | - |
| 601 | 700 | MT-TF/MT-RNR1 | - | - | - | - | - | - | - | - | - | - |
| 701 | 800 | MT-RNR1 | - | - | - | - | - | - | 3.21E−02 | 79156 | - | - |
| 801 | 900 | MT-RNR1 | 4.31E−02 | 94246 | 2.93E−02 | −43592 | - | - | - | - | 1.79E−02 | 95257 |
| 901 | 1000 | MT-RNR1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 1001 | 1100 | MT-RNR1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 1101 | 1200 | MT-RNR1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 1201 | 1300 | MT-RNR1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 1301 | 1400 | MT-RNR1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 1401 | 1500 | MT-RNR1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 1501 | 1600 | MT-RNR1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 1601 | 1700 | MT-RNR1/MT-TV/MT-RNR2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 1701 | 1800 | MT-RNR2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 1801 | 1900 | MT-RNR2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 1901 | 2000 | MT-RNR2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 2001 | 2100 | MT-RNR2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 2101 | 2200 | MT-RNR2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 2201 | 2300 | MT-RNR2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| **2301** | **2400** | **MT-RNR2** | **2.54E−02** | **−97351** | **3.20E−02** | **−29509** | **4.31E−02** | **−82048** | - | - | - | - |
| 2401 | 2500 | MT-RNR2 | 8.90E−03 | 70950 | - | - | - | - | - | - | - | - |
| 2501 | 2600 | MT-RNR2 | 3.30E−03 | 51937 | - | - | - | - | - | - | - | - |
| 2601 | 2700 | MT-RNR2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 2701 | 2800 | MT-RNR2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 2801 | 2900 | MT-RNR2 | - | - | - | - | - | - | - | - | - | - |
| 2901 | 3000 | MT-RNR2 | - | - | - | - | - | - | - | - | - | - |
| 3001 | 3100 | MT-RNR2 | - | - | - | - | - | - | - | - | - | - |
| 3101 | 3200 | MT-RNR2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 3201 | 3300 | MT-RNR2/MT-TL | - | - | - | - | - | - | 2.09E−02 | 242424 | - | - |
| 3301 | 3400 | MT-TL1/MT-ND1 | - | - | - | - | - | - | - | - | - | - |
| 3401 | 3500 | MT-ND1 | - | - | 1.54E−02 | 864685 | - | - | - | - | - | - |
| 3501 | 3600 | MT-ND1 | 4.41E−02 | 1180914 | - | - | 1.72E−02 | 226066 | - | - | - | - |
| **3601** | **3700** | **MT-ND1** | **4.73E−02** | **1250436** | **4.52E−02** | **1228936** | **3.73E−02** | **1183681** | - | - | - | - |
| 3701 | 3800 | MT-ND1 | - | - | - | - | 2.16E−02 | 20818 | - | - | - | - |
| 3801 | 3900 | MT-ND1 | - | - | - | - | - | - | - | - | - | - |
| 3901 | 4000 | MT-ND1 | - | - | - | - | - | - | - | - | - | - |
| 4001 | 4100 | MT-ND1 | 2.41E−02 | −39498 | 1.55E−02 | −38704 | - | - | 4.68E−02 | −49753 | 1.56E−02 | 243767 |
| 4101 | 4200 | MT-ND1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |

**Table 2** List of DMRs identified between five anatomically discreet cortical regions and cerebellum *(Continued)*

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4201 | 4300 | MT-ND1/MT-TI | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 4301 | 4400 | MT-TI/MT-TQ | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 4401 | 4500 | MT-TM/MT-ND2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 4501 | 4600 | MT-ND2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 4601 | 4700 | MT-ND2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 4701 | 4800 | MT-ND2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 4801 | 4900 | MT-ND2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 4901 | 5000 | MT-ND2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 5001 | 5100 | MT-ND2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 5101 | 5200 | MT-ND2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 5201 | 5300 | MT-ND2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 5301 | 5400 | MT-ND2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 5401 | 5500 | MT-ND2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 5501 | 5600 | MT-ND2/MT-TW/MT-TA | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 5601 | 5700 | MT-TA/MT-TN | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 5701 | 5800 | MT-TN/MT-TC | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 5801 | 5900 | MT-TC/MT-TY | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 5901 | 6000 | MT-CO1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 6001 | 6100 | MT-CO1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 6101 | 6200 | MT-CO1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 6201 | 6300 | MT-CO1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 6301 | 6400 | MT-CO1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 6401 | 6500 | MT-CO1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 6501 | 6600 | MT-CO1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 6601 | 6700 | MT-CO1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 6701 | 6800 | MT-CO1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 6801 | 6900 | MT-CO1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 6901 | 7000 | MT-CO1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 7001 | 7100 | MT-CO1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 7101 | 7200 | MT-CO1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 7201 | 7300 | MT-CO1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 7301 | 7400 | MT-CO1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 7401 | 7500 | MT-CO1/MT-TS1 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 7501 | 7600 | MT-TS1/MT-TD/MT-CO2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 7601 | 7700 | MT-CO2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 7701 | 7800 | MT-CO2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 7801 | 7900 | MT-CO2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 7901 | 8000 | MT-CO2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 8001 | 8100 | MT-CO2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 8101 | 8200 | MT-CO2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 8201 | 8300 | MT-CO2/MT-TK | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 8301 | 8400 | MT-TK/MT-ATP8 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 8401 | 8500 | MT-ATP8 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |

**Table 2** List of DMRs identified between five anatomically discreet cortical regions and cerebellum *(Continued)*

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 8501 | 8600 | MT-ATP8/MT-ATP6 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 8601 | 8700 | MT-ATP6 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 8701 | 8800 | MT-ATP6 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 8801 | 8900 | MT-ATP6 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 8901 | 9000 | MT-ATP6 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 9001 | 9100 | MT-ATP6 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 9101 | 9200 | MT-ATP6 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 9201 | 9300 | MT-ATP6/MT-CO3 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 9301 | 9400 | MT-CO3 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 9401 | 9500 | MT-CO3 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 9501 | 9600 | MT-CO3 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 9601 | 9700 | MT-CO3 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 9701 | 9800 | MT-CO3 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 9801 | 9900 | MT-CO3 | - | - | - | - | - | - | - | - | - | - |
| 9901 | 10000 | MT-CO3/MT-TG | - | - | - | - | - | - | - | - | - | - |
| 10001 | 10100 | MT-TG/MT-ND3 | - | - | - | - | - | - | - | - | - | - |
| 10101 | 10200 | MT-ND3 | - | - | - | - | - | - | - | - | - | - |
| 10201 | 10300 | MT-ND3 | 4.68E−02 | −662126 | - | - | - | - | 3.14E−02 | −591067 | 7.30E−03 | −57012 |
| **10301** | **10400** | **MT-ND3/MT-TR/MT-ND4L** | **3.61E−02** | **−132719** | **3.61E−02** | **−135759** | **4.02E−02** | **−115364** | 1.40E−02 | −73063 | 3.36E−02 | −105106 |
| 10401 | 10500 | MT-ND4L | - | - | 4.68E−02 | −71246 | - | - | - | - | - | - |
| 10501 | 10600 | MT-ND4L | - | - | - | - | - | - | - | - | - | - |
| 10601 | 10700 | MT-ND4L | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 10701 | 10800 | MT-ND4L/MT-ND4 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 10801 | 10900 | MT-ND4 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 10901 | 11000 | MT-ND4 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 11001 | 11100 | MT-ND4 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 11101 | 11200 | MT-ND4 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 11201 | 11300 | MT-ND4 | - | - | 1.94E−02 | −68903 | 1.84E−02 | −79625 | 4.71E−02 | 58742 | - | - |
| 11301 | 11400 | MT-ND4 | - | - | 4.78E−02 | −139847 | - | - | - | - | - | - |
| 11401 | 11500 | MT-ND4 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 11501 | 11600 | MT-ND4 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 11601 | 11700 | MT-ND4 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 11701 | 11800 | MT-ND4 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 11801 | 11900 | MT-ND4 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 11901 | 12000 | MT-ND4 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 12001 | 12100 | MT-ND4 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 12101 | 12200 | MT-ND4/MT-TH | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 12201 | 12300 | MT-TS2/MT-TL2 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 12301 | 12400 | MT-TL2/MT-ND5 | 4.71E−02 | 1787 | - | −197592 | - | - | - | - | - | - |
| 12401 | 12500 | MT-ND5 | - | - | - | - | - | - | - | - | - | - |
| 12501 | 12600 | MT-ND5 | - | - | - | - | - | - | - | - | - | - |
| 12601 | 12700 | MT-ND5 | - | - | - | - | - | - | - | - | - | - |
| 12701 | 12800 | MT-ND5 | 2.05E−02 | −5702 | - | - | 4.58E−02 | −388706 | - | - | 4.36E−02 | −394978 |
| 12801 | 12900 | MT-ND5 | 2.26E−02 | −89668 | - | - | - | - | - | - | - | - |

**Table 2** List of DMRs identified between five anatomically discreet cortical regions and cerebellum *(Continued)*

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 12901 | 13000 | MT-ND5 | - | - | 3.10E−02 | −130016 | - | - | - | - | - | - |
| 13001 | 13100 | MT-ND5 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 13101 | 13200 | MT-ND5 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 13201 | 13300 | MT-ND5 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 13301 | 13400 | MT-ND5 | 1.48E−02 | −175917 | - | - | 9.80E−03 | −144949 | 3.40E−02 | −133694 | - | - |
| 13401 | 13500 | MT-ND5 | 3.64E−02 | −104010 | - | - | 3.40E−02 | −81264 | - | - | - | - |
| 13501 | 13600 | MT-ND5 | - | - | - | - | - | - | - | - | - | - |
| 13601 | 13700 | MT-ND5 | - | - | - | - | - | - | - | - | - | - |
| 13701 | 13800 | MT-ND5 | 1.31E−02 | −422610 | - | - | 1.20E−03 | 3714 | - | - | - | - |
| 13801 | 13900 | MT-ND5 | - | - | 3.72E−02 | 708761 | 2.59E−02 | 123249 | - | - | - | - |
| 13901 | 14000 | MT-ND5 | - | - | - | - | 2.17E−02 | −75118 | - | - | - | - |
| 14001 | 14100 | MT-ND5 | - | - | - | - | - | - | - | - | - | - |
| 14101 | 14200 | MT-ND5/MT-ND6 | - | - | - | - | - | - | - | - | 4.18E−02 | −534766 |
| 14201 | 14300 | MT-ND6 | - | - | - | - | - | - | - | - | - | - |
| 14301 | 14400 | MT-ND6 | - | - | - | - | - | - | - | - | - | - |
| 14401 | 14500 | MT-ND6 | - | - | - | - | - | - | - | - | - | - |
| 14501 | 14600 | MT-ND6 | - | - | - | - | - | - | - | - | 4.86E−02 | −82767 |
| 14601 | 14700 | MT-ND6/MT-TE | - | - | - | - | - | - | - | - | - | - |
| 14701 | 14800 | MT-TE/MT-CYB | - | - | - | - | - | - | - | - | - | - |
| 14801 | 14900 | MT-CYB | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 14901 | 15000 | MT-CYB | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 15001 | 15100 | MT-CYB | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| 15101 | 15200 | MT-CYB | - | - | - | - | - | - | - | - | - | - |
| 15201 | 15300 | MT-CYB | - | - | - | - | - | - | 1.20E−02 | 254183 | - | - |
| 15301 | 15400 | MT-CYB | - | - | - | - | - | - | - | - | - | - |
| 15401 | 15500 | MT-CYB | 4.38E−02 | 10628 | - | - | 2.10E−02 | −554527 | - | - | 3.28E−02 | −554225 |
| 15501 | 15600 | MT-CYB | 3.73E−02 | −671707 | - | - | 1.23E−02 | −685630 | 3.82E−02 | 9301 | 3.51E−02 | −695515 |
| 15601 | 15700 | MT-CYB | - | - | - | - | 4.33E−02 | −746270 | - | - | - | - |
| 15701 | 15800 | MT-CYB | - | - | - | - | - | - | - | - | - | - |
| 15801 | 15900 | MT-CYB/MT-TT | - | - | - | - | - | - | - | - | - | - |
| 15901 | 16000 | MT-TT/MT-TP | - | - | - | - | - | - | - | - | - | - |
| 16001 | 16100 | MT-TP | - | - | - | - | - | - | - | - | - | - |
| 16101 | 16200 | D-Loop | - | - | 3.40E−03 | 3014 | - | - | - | - | - | - |
| **16201** | **16300** | **D-Loop** | **4.64E−02** | **380407** | **2.50E−02** | **1117943** | **3.68E−02** | **296890** | - | - | - | - |
| **16301** | **16400** | **D-Loop** | **3.65E−02** | **−941720** | **1.38E−02** | **−112453** | **2.73E−02** | **−177146** | - | - | - | - |
| **16401** | **16500** | **D-Loop** | **4.29E−02** | **444051** | **2.20E−02** | **1151402** | **1.68E−02** | **313156** | - | - | - | - |
| **16501** | **16600** | **D-Loop** | **3.71E−02** | **784572** | **2.10E−02** | **764994** | **2.46E−02** | **836662** | - | - | - | - |

Shown is the location of the DMR within the mitochondrial genome (ChrM) (based on GENCODE), the gene(s) residing within the 100 bp window, and *p* value from paired *t* tests between each of the five cortical regions: Brodmann areas 8, 9 and 10 (BA8, BA9, BA10), entorhinal cortex (ECX) and superior temporal gyrus (STG) compared to the cerebellum (CER). Results are displayed in order of genomic position. RPKM and corresponding *p* values are shown for windows if $p < 0.05$. Key: - denotes data not significant ($p > 0.05$); ND denotes not determined as the window was not included in analysis due to removal in *NUMT* pipeline. Results shown in bold represent those found to be present across all prefrontal cortex areas (BA8, BA9, BA10)

difference was maintained in all Brodmann area regions, with three conserved regions of hypomethylation and four conserved regions of hypermethylation, with respect to the cerebellum. Furthermore, four of the seven conserved regions were adjacent to each other within the mitochondrial displacement loop (D-Loop) (16201–16600 bp), a region associated with gene transcription and DNA replication.

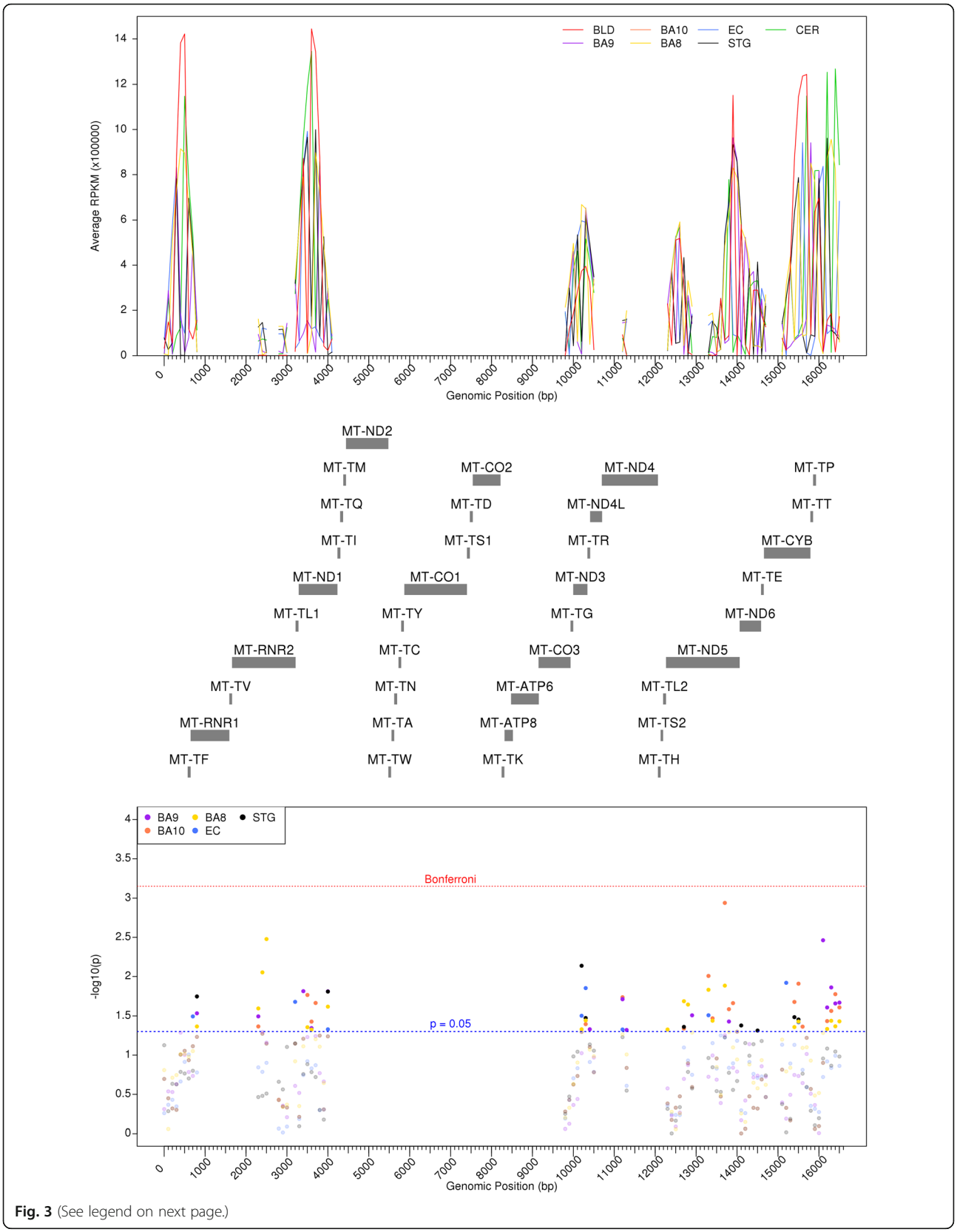



**Fig. 3** (See legend on next page.)

(See figure on previous page.)
**Fig. 3** DNA methylation differences are seen in the mitochondrial genome between brain regions and blood. Average raw RPKM values across the mitochondrial genome for each individual cortical brain region alongside matched blood and cerebellum samples are shown in the *top panel*, with gene positions downloaded from GENCODE shown in the *middle panel*. For each 100 bp window, paired *t* tests were performed to compare each cortical brain region and the cerebellum, with -log10 (*p*) shown in the *bottom panel*. *BLD* blood, *BA8* Brodmann area 8, *BA9* Brodmann area 9, *BA10* Brodmann area 10, *CER* cerebellum, *CTX* cortex, *ECX* entorhinal cortex, *RPKM* reads per kilobase of transcript per million mapped reads, *STG* superior temporal gyrus. *Red dashed line* denotes the Bonferroni significance, whilst *blue dashed line* denotes $p < 0.05$ in the *lower panel*

### A number of differentially methylated regions in mtDNA can be observed between the cortex and cerebellum

We were also interested to see whether total cortical tissue was significantly different to matched cerebellum samples. Given the paired nature of the different anatomical regions of the cortex, we used a multilevel mixed effects model to compare total cortex to cerebellum (see the Methods section). This analysis revealed 48 nominally significant ($p < 0.05$) windows (Table 3; Fig. 4), of which eight passed the Bonferroni correction (Table 3, bold face). Interestingly, three of these eight were adjacent to each other, lying between 10301 and 10600 bp and covering MT-ND3/MT-ND4L and MT-TR. We also saw a Bonferroni significant difference in DNA methylation in the D-Loop, where we earlier noted DNA methylation changes across all three Brodmann area regions.

### MtDNA methylation patterns can distinguish between tissue types

Although we have shown that mtDNA methylation patterns are highly similar between distinct anatomical regions of the human brain and blood, we were also interested to identify whether mtDNA methylation patterns could distinguish between these tissue types. Through unsupervised hierarchical clustering, we showed that average mtDNA methylation patterns can segregate these tissues (Fig. 5a). Importantly, ncDNA methylation profiles in the same samples have also been previously shown to separate the cortex, cerebellum and blood [35]. Interestingly, when we performed unsupervised hierarchical clustering on the individual samples, we found that, in most cases, intra-individual differences across tissue types are greater than inter-individual differences within each tissue type, as the cortex, cerebellum and blood samples clustered with their own tissue type, respectively (Fig. 5b).

## Discussion

The availability of publicly available epigenomic data provides a great resource for mitochondrial epigenetics, a field that is relatively nascent and has yet to be thoroughly investigated in a range of complex diseases. Here, we present evidence that mtDNA methylation patterns across mtDNA are brain region specific. However, data such as that presented here is confounded by a lack of isolation of mtDNA prior to antibody enrichment and sequencing. As such, the potential of including *NUMT*s in datasets derived from data generated using total genomic DNA could lead to misleading results. Here, we controlled for regions of high sequence homology between the nuclear and mitochondrial genomes. However, this approach is likely over-conservative and does lead to the generation of a somewhat truncated consensus sequence. PCA of the mitochondrial epigenome after corrections for nuclear homology was able to separate individuals belonging to the three main tissue types, the blood, cortex and cerebellum based on mtDNA methylation variation among tissue types. This tissue specificity is further highlighted by the identification of eight DMRs that pass the Bonferroni correction for multiple testing between total cortex and cerebellum. MtDNA methylation has been shown to be cell line dependent in the past. [31] Although overall DNA methylation levels were low in all tissues, it is worth noting that the study was performed on non bisulfite-treated DNA. As such, the low percentage of mtDNA methylation is not a pitfall due to a lack of a total bisulfite treatment efficiency. One limitation of the current study is the unavailability of publicly available MeDIP-seq datasets of matched cortical and cerebellum tissue from other cohorts for validation purposes. Future work would aim to replicate our findings in additional study cohorts and also to investigate the relationship between mitochondrial DNA methylation and gene expression.

Despite a number of nominally significant windows being identified between each individual cortical region and the cerebellum, these did not pass the Bonferroni correction, although it is likely this method is too stringent. Nevertheless, the conservation of seven nominally significant windows across each Brodmann area is interesting to note. Four of these windows lie adjacent to each other and correspond to the mitochondrial D-Loop, a region containing the only two mitochondrial promoters which is typically associated with gene transcription and DNA replication. However, one limitation of this study is owed to the use of antibody-based enrichment, resulting in the analysis being limited to a window-based approach. Despite this, studies of the nuclear genome have shown high correlation between window-based approaches and, more sensitive, single-site assays such as the Illumina 450K beadarray [32]. However, given the small size of the mitochondrial genome and that

**Table 3** List of DMRs identified between total cortex and cerebellum

| Start (bp) | Stop (bp) | Gene(s) | *p* value |
|---|---|---|---|
| 1 | 100 | D-Loop | - |
| 101 | 200 | D-Loop | - |
| 201 | 300 | D-Loop | - |
| 301 | 400 | D-Loop | - |
| 401 | 500 | D-Loop | 7.99E−03 |
| 501 | 600 | MT-TF | 4.49E−03 |
| 601 | 700 | MT-TF/MT-RNR1 | 1.16E−02 |
| 701 | 800 | MT-RNR1 | 3.22E−03 |
| **801** | **900** | **MT-RNR1** | **1.91E−04** |
| 901 | 1000 | MT-RNR1 | ND |
| 1001 | 1100 | MT-RNR1 | ND |
| 1101 | 1200 | MT-RNR1 | ND |
| 1201 | 1300 | MT-RNR1 | ND |
| 1301 | 1400 | MT-RNR1 | ND |
| 1401 | 1500 | MT-RNR1 | ND |
| 1501 | 1600 | MT-RNR1 | ND |
| 1601 | 1700 | MT-RNR1/MT-TV/MT-RNR2 | ND |
| 1701 | 1800 | MT-RNR2 | ND |
| 1801 | 1900 | MT-RNR2 | ND |
| 1901 | 2000 | MT-RNR2 | ND |
| 2001 | 2100 | MT-RNR2 | ND |
| 2101 | 2200 | MT-RNR2 | ND |
| 2201 | 2300 | MT-RNR2 | ND |
| 2301 | 2400 | MT-RNR2 | 8.10E−03 |
| 2401 | 2500 | MT-RNR2 | 7.50E−03 |
| 2501 | 2600 | MT-RNR2 | 5.14E−03 |
| 2601 | 2700 | MT-RNR2 | ND |
| 2701 | 2800 | MT-RNR2 | ND |
| 2801 | 2900 | MT-RNR2 | - |
| 2901 | 3000 | MT-RNR2 | - |
| 3001 | 3100 | MT-RNR2 | - |
| 3101 | 3200 | MT-RNR2 | - |
| 3201 | 3300 | MT-RNR2/MT-TL | 1.51E−03 |
| 3301 | 3400 | MT-TL1/MT-ND1 | - |
| 3401 | 3500 | MT-ND1 | 1.37E−02 |
| 3501 | 3600 | MT-ND1 | 4.27E−03 |
| 3601 | 3700 | MT-ND1 | 5.56E−03 |
| 3701 | 3800 | MT-ND1 | 8.21E−03 |
| 3801 | 3900 | MT-ND1 | - |
| 3901 | 4000 | MT-ND1 | - |
| **4001** | **4100** | **MT-ND1** | **3.07E−06** |
| 4101 | 4200 | MT-ND1 | - |
| 4201 | 4300 | MT-ND1/MT-TI | ND |
| 4301 | 4400 | MT-TI/MT-TQ | ND |
| 4401 | 4500 | MT-TM/MT-ND2 | ND |
| 4501 | 4600 | MT-ND2 | ND |
| 4601 | 4700 | MT-ND2 | ND |
| 4701 | 4800 | MT-ND2 | ND |
| 4801 | 4900 | MT-ND2 | ND |
| 4901 | 5000 | MT-ND2 | ND |
| 5001 | 5100 | MT-ND2 | ND |
| 5101 | 5200 | MT-ND2 | ND |
| 5201 | 5300 | MT-ND2 | ND |
| 5301 | 5400 | MT-ND2 | ND |
| 5401 | 5500 | MT-ND2 | ND |
| 5501 | 5600 | MT-ND2/MT-TW/MT-TA | ND |
| 5601 | 5700 | MT-TA/MT-TN | ND |
| 5701 | 5800 | MT-TN/MT-TC | ND |
| 5801 | 5900 | MT-TC/MT-TY | ND |
| 5901 | 6000 | MT-CO1 | ND |
| 6001 | 6100 | MT-CO1 | ND |
| 6101 | 6200 | MT-CO1 | ND |
| 6201 | 6300 | MT-CO1 | ND |
| 6301 | 6400 | MT-CO1 | ND |
| 6401 | 6500 | MT-CO1 | ND |
| 6501 | 6600 | MT-CO1 | ND |
| 6601 | 6700 | MT-CO1 | ND |
| 6701 | 6800 | MT-CO1 | ND |
| 6801 | 6900 | MT-CO1 | ND |
| 6901 | 7000 | MT-CO1 | ND |
| 7001 | 7100 | MT-CO1 | ND |
| 7101 | 7200 | MT-CO1 | ND |
| 7201 | 7300 | MT-CO1 | ND |
| 7301 | 7400 | MT-CO1 | ND |
| 7401 | 7500 | MT-CO1/MT-TS1 | ND |
| 7501 | 7600 | MT-TS1/MT-TD/MT-CO2 | ND |
| 7601 | 7700 | MT-CO2 | ND |
| 7701 | 7800 | MT-CO2 | ND |
| 7801 | 7900 | MT-CO2 | ND |
| 7901 | 8000 | MT-CO2 | ND |
| 8001 | 8100 | MT-CO2 | ND |
| 8101 | 8200 | MT-CO2 | ND |
| 8201 | 8300 | MT-CO2/MT-TK | ND |
| 8301 | 8400 | MT-TK/MT-ATP8 | ND |
| 8401 | 8500 | MT-ATP8 | ND |
| 8501 | 8600 | MT-ATP8/MT-ATP6 | ND |
| 8601 | 8700 | MT-ATP6 | ND |

**Table 3** List of DMRs identified between total cortex and cerebellum *(Continued)*

| | | | |
|---|---|---|---|
| 8701 | 8800 | MT-ATP6 | ND |
| 8801 | 8900 | MT-ATP6 | ND |
| 8901 | 9000 | MT-ATP6 | ND |
| 9001 | 9100 | MT-ATP6 | ND |
| 9101 | 9200 | MT-ATP6 | ND |
| 9201 | 9300 | MT-ATP6/MT-CO3 | ND |
| 9301 | 9400 | MT-CO3 | ND |
| 9401 | 9500 | MT-CO3 | ND |
| 9501 | 9600 | MT-CO3 | ND |
| 9601 | 9700 | MT-CO3 | ND |
| 9701 | 9800 | MT-CO3 | ND |
| 9801 | 9900 | MT-CO3 | - |
| 9901 | 10000 | MT-CO3/MT-TG | - |
| 10001 | 10100 | MT-TG/MT-ND3 | 1.81E−02 |
| 10101 | 10200 | MT-ND3 | 1.39E−02 |
| **10201** | **10300** | **MT-ND3** | **3.53E−04** |
| **10301** | **10400** | **MT-ND3/MT-TR/MT-ND4L** | **1.19E−05** |
| **10401** | **10500** | **MT-ND4L** | **2.61E−04** |
| 10501 | 10600 | MT-ND4L | 9.05E−04 |
| 10601 | 10700 | MT-ND4L | ND |
| 10701 | 10800 | MT-ND4L/MT-ND4 | ND |
| 10801 | 10900 | MT-ND4 | ND |
| 10901 | 11000 | MT-ND4 | ND |
| 11001 | 11100 | MT-ND4 | ND |
| 11101 | 11200 | MT-ND4 | ND |
| **11201** | **11300** | **MT-ND4** | **8.86E−05** |
| 11301 | 11400 | MT-ND4 | 1.65E−03 |
| 11401 | 11500 | MT-ND4 | ND |
| 11501 | 11600 | MT-ND4 | ND |
| 11601 | 11700 | MT-ND4 | ND |
| 11701 | 11800 | MT-ND4 | ND |
| 11801 | 11900 | MT-ND4 | ND |
| 11901 | 12000 | MT-ND4 | ND |
| 12001 | 12100 | MT-ND4 | ND |
| 12101 | 12200 | MT-ND4/MT-TH | ND |
| 12201 | 12300 | MT-TS2/MT-TL2 | ND |
| 12301 | 12400 | MT-TL2/MT-ND5 | - |
| 12401 | 12500 | MT-ND5 | - |
| 12501 | 12600 | MT-ND5 | - |
| 12601 | 12700 | MT-ND5 | - |
| 12701 | 12800 | MT-ND5 | 1.81E−03 |
| 12801 | 12900 | MT-ND5 | 1.05E−02 |
| 12901 | 13000 | MT-ND5 | 8.23E−03 |
| 13001 | 13100 | MT-ND5 | ND |
| 13101 | 13200 | MT-ND5 | ND |
| 13201 | 13300 | MT-ND5 | ND |
| 13301 | 13400 | MT-ND5 | 1.44E−03 |
| 13401 | 13500 | MT-ND5 | 9.13E−04 |
| 13501 | 13600 | MT-ND5 | 1.61E−02 |
| 13601 | 13700 | MT-ND5 | 3.89E−02 |
| **13701** | **13800** | **MT-ND5** | **2.77E−04** |
| 13801 | 13900 | MT-ND5 | 2.80E−03 |
| 13901 | 14000 | MT-ND5 | 1.88E−02 |
| 14001 | 14100 | MT-ND5 | 9.31E−03 |
| 14101 | 14200 | MT-ND5/MT-ND6 | - |
| 14201 | 14300 | MT-ND6 | - |
| 14301 | 14400 | MT-ND6 | 1.04E−02 |
| 14401 | 14500 | MT-ND6 | 1.99E−02 |
| 14501 | 14600 | MT-ND6 | 2.26E−02 |
| 14601 | 14700 | MT-ND6/MT-TE | 3.82E−03 |
| 14701 | 14800 | MT-TE/MT-CYB | 1.92E−02 |
| 14801 | 14900 | MT-CYB | ND |
| 14901 | 15000 | MT-CYB | ND |
| 15001 | 15100 | MT-CYB | ND |
| 15101 | 15200 | MT-CYB | 3.30E−02 |
| 15201 | 15300 | MT-CYB | - |
| 15301 | 15400 | MT-CYB | - |
| 15401 | 15500 | MT-CYB | 8.52E−04 |
| 15501 | 15600 | MT-CYB | 7.43E−04 |
| 15601 | 15700 | MT-CYB | 1.16E−02 |
| 15701 | 15800 | MT-CYB | 2.24E−02 |
| 15801 | 15900 | MT-CYB/MT-TT | - |
| 15901 | 16000 | MT-TT/MT-TP | - |
| 16001 | 16100 | MT-TP | - |
| 16101 | 16200 | D-Loop | 2.23E−03 |
| **16201** | **16300** | **D-Loop** | **5.02E−04** |
| 16301 | 16400 | D-Loop | 1.84E−03 |
| 16401 | 16500 | D-Loop | 1.12E−03 |
| 16501 | 16600 | D-Loop | 2.20E−03 |

Shown is the location of the DMR within ChrM (based on GENCODE), the gene(s) residing within the 100 bp window, and *p* value from a multilevel mixed effects model. Results are displayed in order of genomic position. RPKM and corresponding *p* values are shown for windows if $p < 0.05$. Key: - denotes data not significant ($p > 0.05$); ND denotes not determined as the window was not included in analysis due to removal in *NUMT* pipeline; bold denotes windows that reached our Bonferroni significant threshold of $p < 7.04E{-}04$

23 of the 37 genes present in the genome are below 100 bp in size, this window-based approach may not be the most appropriate for future studies designed to specifically assess mtDNA methylation as it can result

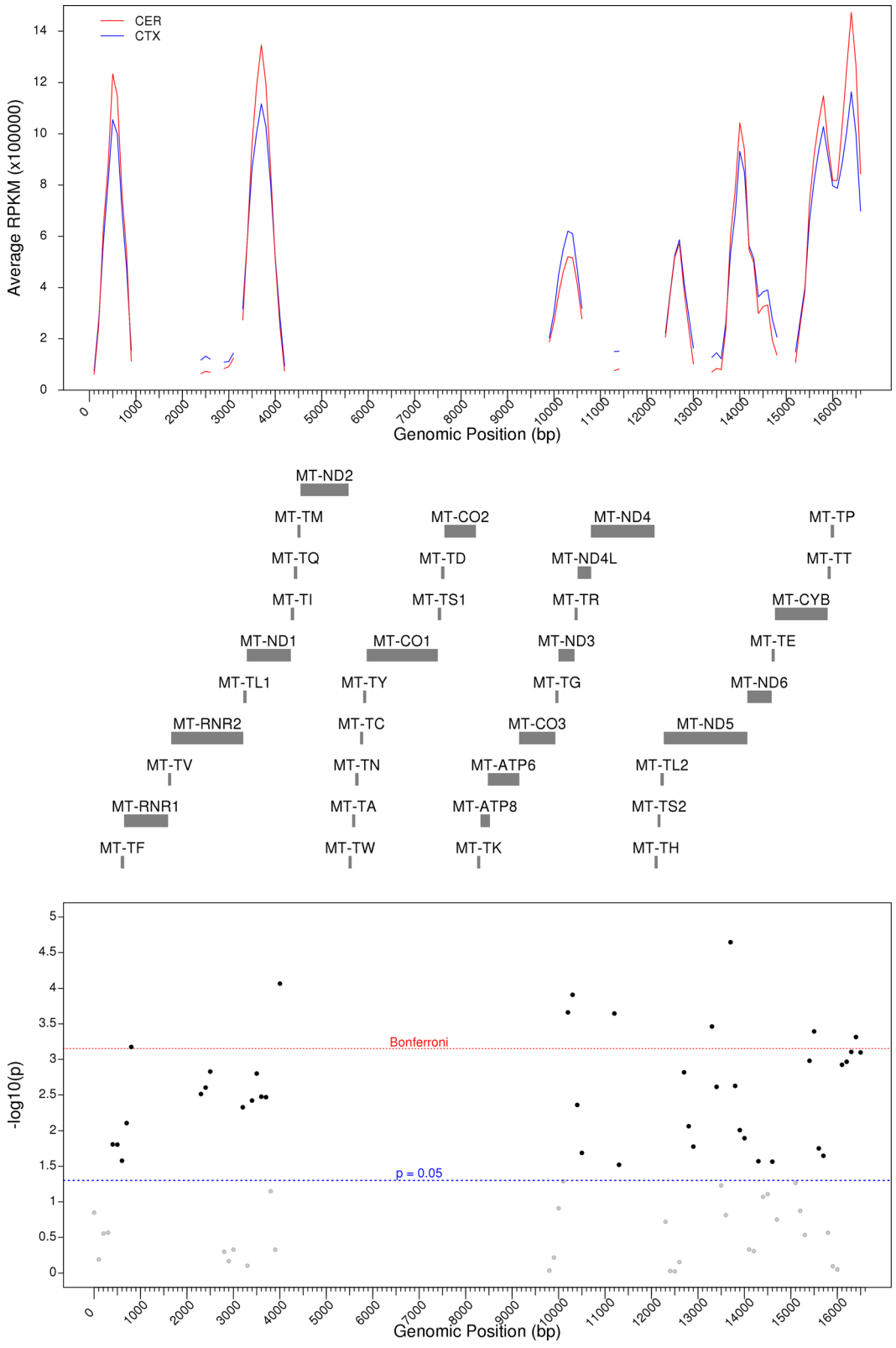



**Fig. 4** (See legend on next page.)

(See figure on previous page.)
**Fig. 4** DNA methylation differences are seen in the mitochondrial genome between the cerebellum and cortex. RPKM values in the total cortex and cerebellum across the mitochondrial genome are shown in the *top panel*, with gene positions downloaded from GENCODE shown in the *middle panel*. For each 100 bp window, paired *t* tests were performed to compare the cortex to the cerebellum, with -log10 (*p*) shown in the *bottom panel*. *BLD* blood, *BA8* Brodmann area 8, *BA9* Brodmann area 9, *BA10* Brodmann area 10, *CER* cerebellum, *CTX* cortex, *ECX* entorhinal cortex, *RPKM* reads per kilobase of transcript per million mapped reads, *STG* superior temporal gyrus. *Red dashed line* denotes the Bonferroni significance, whilst *blue dashed line* denotes $p < 0.05$ in the *lower panel*

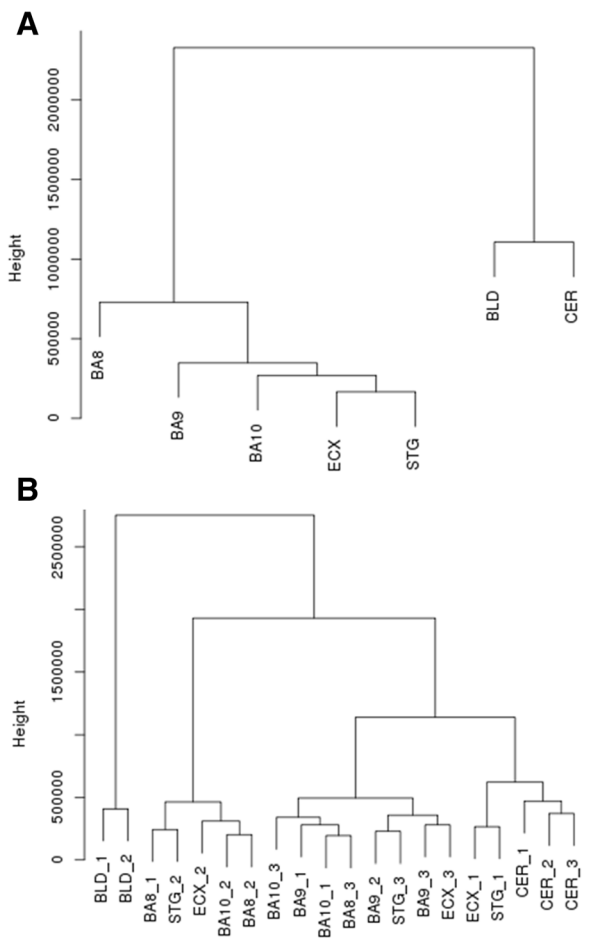



**Fig. 5** MtDNA methylation patterns can distinguish between tissue types. **a** Average RPKM values for each cortical brain region, cerebellum and blood samples were clustered based upon the Euclidean distance, identifying two major clusters; the cortex and blood-cerebellum. **b** When clustering RPKM values in the individual samples from the cortex, cerebellum and blood, we observed that individual cortex samples clustered together, whilst cerebellum and blood samples formed separate clusters. This highlights that tissue-specific differences between the cortex, cerebellum and blood are greater than intra-individual variability within a tissue. *BLD* blood, *BA8* Brodmann area 8, *BA9* Brodmann area 9, *BA10* Brodmann area 10, *CER* cerebellum, *CTX* cortex, *ECX* entorhinal cortex, *RPKM* reads per kilobase of transcript per million mapped reads, *STG* superior temporal gyrus

in a window intersecting two genes in the polycistronic transcript.

## Conclusions

This method provides a conservative approach to determine mtDNA methylation across the genome for data previously generated using next-generation sequencing approaches such as MeDIP-seq. Its conservative nature reduces the risk of the inclusion of *NUMT*s in the final analysis of whole genome data but may also lead to the inclusion of false negatives as well as potential gaps in the reference sequence. As such, it is best suited to analyzing previously generated whole genome data and is not a replacement for the isolation of mitochondrial DNA [36] prior to targeted methylation studies, which would be the optimal approach for investigating mitochondrial epigenomics. However, our method has allowed the identification of novel brain-region-specific DMRs in a previously generated publicly available dataset. Furthermore, the identification of brain region-specific mtDNA methylation patterns across the mitochondrial epigenome suggests the importance of a focussed, tissue-specific study design when investigating mtDNA methylation. As previously discussed, one caveat when utilizing MeDIP-seq data is the segregation of data into neighbouring windows, meaning that determining the exact corresponding gene of a DMR is difficult and, as such, future studies should aim to sequence the mitochondrial DNA methylome at single-base resolution to address this.

## Methods

### Data collection

We utilized publicly available MeDIP-seq data from Davies et al. [35]. In brief, this data was generated using 5 μg fragmented gDNA, which, following end repair <A> base addition and adaptor ligation, was immunoprecipitated using an anti-5-mC antibody (Diagenode, Liège, Belgium). MeDIP DNA was purified and then amplified using adaptor-mediated PCR, with DNA fragments between 220 and 320 bp subjected to highly parallel 50 bp paired-end sequencing on the Illumina HiSeq platform. The paired-end, raw fasta files were provided by the authors and quality checked using FastQC. Sample information is provided in Table 1.

### Quality control and *NUMT* exclusion

Fasta files were subjected to adaptor and Phred score ($q < 20$) trimming. In an attempt to remove any potential contamination of possible *NUMT*s, multiple alignments to the reference genome were undertaken. Paired fasta files were aligned to GRCH37 using BWA. Unique and mapped reads aligning to the mitochondria were then re-mapped to a custom GRCH37 reference without the mitochondrial chromosome. Reads not mapping to the custom reference were then taken forward and realigned to the full GRCH37 reference to eliminate the possibility of homologous regions mapping falsely to the mitochondrial genome (Fig. 1). All alignments were carried out using BWA mem and default settings. Reads per kilobase of transcript per million mapped reads (RPKM) values for each sample were calculated using the MEDIPS package [37]. Methylation was averaged across 100 bp non-overlapping windows (default parameter setting in MEDIPS), and only windows with read counts >10 were considered for analysis. Due to the non-normal distribution of all cohorts, RPKM values were $\log^2$ transformed before statistical analysis.

### Statistical analyses

All analyses were performed in the *R* statistical environment version 3.2.1 [38]. For all analyses, a nominally significant threshold of $p < 0.05$ and a Bonferroni significant threshold of $p < 7.04E{-}04$ were used. Given the matched sample nature of this cohort, two-tailed, paired *t* tests were performed at each window along the mitochondrial genome to identify DMRs between the individual cortical regions and cerebellum. To compare the total cortex to cerebellum, we performed a multilevel mixed effects model in the Lme4 package in *R* [39], using the brain region as the random effect and individual as the fixed effect. To assess the similarity of the brain regions, we used the *R* function "hclust" to cluster average RPKM values for the brain regions using the Euclidean distance. We used the *R* function "corrgram" within the corrgram package [40] to order samples based upon the similarity of their principal components.

#### Abbreviations

5-mC: 5-methylcytosine; BA: Brodmann area; DMR: Differentially methylated region; MeDIP-seq: Methylated DNA Immunoprecipitation Sequencing; mtDNA: Mitochondrial DNA; ncDNA: Nuclear DNA; *NUMT*s: Nuclear-mitochondrial DNA; PCA: Principal component analysis; rRNA: Ribosomal ribonucleic acid; tRNA: Transfer ribonucleic acid

#### Acknowledgements

Not applicable.

#### Funding

This work was funded by an Alzheimer's Society project grant to KL (grant number AS-PG-14-038), an Alzheimer's Research UK pilot grant to KL (grant number ARUK-PPG2013A-5) and an Alzheimer's Association New Investigator Research Grant to KL (grant number NIRG-14-320878).

#### Availability of data and materials

Paired-end, raw fastq files from Davies et al. were provided by the authors [35].

#### Authors' contributions

MD, MW, RS, AJ and EH undertook the data analysis and bioinformatics. MD and MW developed the pipeline for the analysis. MND and JM provided the data for the analysis. KL conceived and supervised the project. MD and KL drafted the manuscript. All authors read and approved the final submission.

#### Competing interests

The authors declare that they have no competing interests.

#### Consent for publication

Not applicable.

#### Ethics approval and consent to participate

Not applicable.

## Publisher's Note



#### Author details

[1]University of Exeter Medical School, RILD, University of Exeter, Barrack Road, Devon, UK. [2]Institute of Psychiatry, Psychology and Neuroscience, King's College London, De Crespigny Park, London, UK. [3]Department of Twin Research & Genetic Epidemiology, King's College London, Lambeth Palace Road, London, UK. [4]School of Biological Sciences, University of Essex, Essex, UK.

Received: 29 October 2016 Accepted: 30 March 2017
Published online: 03 May 2017

#### References

1. Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, et al. Sequence and organization of the human mitochondrial genome. Nature. 1981;290:457–65.
2. Murgia M, Rizzuto R. Molecular diversity and pleiotropic role of the mitochondrial calcium uniporter. Cell Calcium. 2015;58:11–7.
3. Wang W, Esbensen Y, Kunke D, Suganthan R, Rachek L, Bjoras M, Eide L. Mitochondrial DNA damage level determines neural stem cell differentiation fate. J Neurosci. 2011;31:9746–51.
4. Lunnon K, Ibrahim Z, Proitsi P, Lourdusamy A, Newhouse S, Sattlecker M, Furney S, Saleem M, Soininen H, Kloszewska I, et al. Mitochondrial dysfunction and immune activation are detectable in early Alzheimer's disease blood. J Alzheimers Dis. 2012;30:685–710.
5. Lunnon K, Keohane A, Pidsley R, Newhouse S, Riddoch-Contreras J, Thubron EB, Devall M, Soininen H, Kloszewska I, Mecocci P, et al. Mitochondrial genes are altered in blood early in Alzheimer's disease. Neurobiol Aging. 2017;53:36–47.
6. Prabakaran S, Swatton JE, Ryan MM, Huffaker SJ, Huang JT, Griffin JL, Wayland M, Freeman T, Dudbridge F, Lilley KS, et al. Mitochondrial dysfunction in schizophrenia: evidence for compromised brain metabolism and oxidative stress. Mol Psychiatry. 2004;9:684–97. 643.
7. Clay HB, Sillivan S, Konradi C. Mitochondrial dysfunction and pathology in bipolar disorder and schizophrenia. Int J Dev Neurosci. 2011;29:311–24.
8. Chang CC, Jou SH, Lin TT, Lai TJ, Liu CS. Mitochondria DNA change and oxidative damage in clinically stable patients with major depressive disorder. PLoS One. 2015;10:e0125855.
9. Lunnon K, Mill J. Epigenetic studies in Alzheimer's disease: current findings, caveats, and considerations for future studies. Am J Med Genet B Neuropsychiatr Genet. 2013;162B:789–99.
10. Lunnon K, Hannon E, Smith RG, Dempster E, Wong C, Burrage J, Troakes C, Al-Sarraj S, Kepa A, Schalkwyk L, Mill J. Variation in 5-hydroxymethylcytosine across human cortex and cerebellum. Genome Biol. 2016;17:27.
11. Lowe R, Slodkowicz G, Goldman N, Rakyan VK. The human blood DNA methylome displays a highly distinctive profile compared with other somatic tissues. Epigenetics. 2015;10:274–81.
12. Lokk K, Modhukur V, Rajashekar B, Martens K, Magi R, Kolde R, Koltsina M, Nilsson TK, Vilo J, Salumets A, Tonisson N. DNA methylome profiling of

human tissues identifies global and tissue-specific methylation patterns. Genome Biol. 2014;15:r54.
13. Lunnon K, Smith R, Hannon EJ, De Jager PL, Srivastava G, Volta M, Troakes C, Al-Sarraj S, Burrage J, Macdonald R, et al. Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. Nat Neurosci. 2014;17:1164–70.
14. De Jager PL, Srivastava G, Lunnon K, Burgess J, Schalkwyk LC, Yu L, Eaton ML, Keenan BT, Ernst J, McCabe C, et al. Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. Nat Neurosci. 2014;17:1156–63.
15. Smith AR, Smith RG, Condliffe D, Hannon E, Schalkwyk L, Mill J, Lunnon K. Increased DNA methylation near TREM2 is consistently seen in the superior temporal gyrus in Alzheimer's disease brain. Neurobiol Aging. 2016;47:35–40.
16. Pidsley R, Viana J, Hannon E, Spiers HH, Troakes C, Al-Saraj S, Mechawar N, Turecki G, Schalkwyk LC, Bray NJ, Mill J. Methylomic profiling of human brain tissue supports a neurodevelopmental origin for schizophrenia. Genome Biol. 2014;15:483.
17. Hannon E, Spiers H, Viana J, Pidsley R, Burrage J, Murphy TM, Troakes C, Turecki G, O'Donovan MC, Schalkwyk LC, et al. Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. Nat Neurosci. 2016;19:48–+.
18. Florath I, Butterbach K, Heiss J, Bewerunge-Hudler M, Zhang Y, Schottker B, Brenner H. Type 2 diabetes and leucocyte DNA methylation: an epigenome-wide association study in over 1,500 older adults. Diabetologia. 2016;59:130–8.
19. Adams AT, Kennedy NA, Hansen R, Ventham NT, O'Leary KR, Drummond HE, Noble CL, El-Omar E, Russell RK, Wilson DC, et al. Two-stage genome-wide methylation profiling in childhood-onset Crohn's Disease implicates epigenetic alterations at the VMP1/MIR21 and HLA loci. Inflamm Bowel Dis. 2014;20:1784–93.
20. Devall M, Mill J, Lunnon K. The mitochondrial epigenome: a role in Alzheimer's disease? Epigenomics. 2014;6:665–75.
21. Devall M, Roubroeks J, Mill J, Weedon M, Lunnon K. Epigenetic regulation of mitochondrial function in neurodegenerative disease: New insights from advances in genomic technologies. Neurosci Lett. 2016;625:47–55.
22. Shock LS, Thakkar PV, Peterson EJ, Moran RG, Taylor SM. DNA methyltransferase 1, cytosine methylation, and cytosine hydroxymethylation in mammalian mitochondria. Proc Natl Acad Sci U S A. 2011;108:3630–5.
23. Chestnut BA, Chang Q, Price A, Lesuisse C, Wong M, Martin LJ. Epigenetic regulation of motor neuron cell death through DNA methylation. J Neurosci. 2011;31:16619–36.
24. Feng S, Xiong L, Ji Z, Cheng W, Yang H. Correlation between increased ND2 expression and demethylated displacement loop of mtDNA in colorectal cancer. Mol Med Rep. 2012;6:125–30.
25. Pirola CJ, Gianotti TF, Burgueno AL, Rey-Funes M, Loidl CF, Mallardi P, Martino JS, Castano GO, Sookoian S. Epigenetic modification of liver mitochondrial DNA is associated with histological severity of nonalcoholic fatty liver disease. Gut. 2013;62:1356–63.
26. Infantino V, Castegna A, Iacobazzi F, Spera I, Scala I, Andria G, Iacobazzi V. Impairment of methyl cycle affects mitochondrial methyl availability and glutathione level in Down's syndrome. Mol Genet Metab. 2011;102:378–82.
27. Wong M, Gertz B, Chestnut BA, Martin LJ. Mitochondrial DNMT3A and DNA methylation in skeletal muscle and CNS of transgenic mouse models of ALS. Front Cell Neurosci. 2013;7:279.
28. Wen SL, Zhang F, Feng S. Decreased copy number of mitochondrial DNA: a potential diagnostic criterion for gastric cancer. Oncol Lett. 2013;6:1098–102.
29. Blanch M, Mosquera JL, Ansoleaga B, Ferrer I, Barrachina M. Altered mitochondrial DNA methylation pattern in Alzheimer disease-related pathology and in Parkinson disease. Am J Pathol. 2016;186:385–97.
30. Iacobazzi V, Castegna A, Infantino V, Andria G. Mitochondrial DNA methylation as a next-generation biomarker and diagnostic tool. Mol Genet Metab. 2013;110:25–34.
31. Ghosh S, Sengupta S, Scaria V. Comparative analysis of human mitochondrial methylomes shows distinct patterns of epigenetic regulation in mitochondria. Mitochondrion. 2014;18:58–62.
32. Clark C, Palta P, Joyce CJ, Scott C, Grundberg E, Deloukas P, Palotie A, Coffey AJ. A comparison of the whole genome approach of MeDIP-seq to the targeted approach of the Infinium HumanMethylation450 BeadChip® for methylome profiling. PLoS One. 2012;7:e50233.
33. Thangaraj K, Joshi MB, Reddy AG, Rasalkar AA, Singh L. Sperm mitochondrial mutations as a cause of low sperm motility. J Androl. 2003;24:388–92.
34. Yao YG, Kong QP, Salas A, Bandelt HJ. Pseudomitochondrial genome haunts disease studies. J Med Genet. 2008;45:769–72.
35. Davies MN, Volta M, Pidsley R, Lunnon K, Dixit A, Lovestone S, Coarfa C, Harris RA, Milosavljevic A, Troakes C, et al. Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. Genome Biol. 2012;13:R43.
36. Devall M, Burrage J, Caswell R, Johnson M, Troakes C, Al-Sarraj S, Jeffries AR, Mill J, Lunnon K. A comparison of mitochondrial DNA isolation methods in frozen post-mortem human brain tissue-applications for studies of mitochondrial genetics in brain disorders. Biotechniques. 2015;59:241–6.
37. Lienhard MGC, Morkel M, Herwig R, Chavez L. MEDIPS: genome-wide differential coverage analysis of sequencing data derived from DNA enrichment experiments. Bioinformatics. 2014;30:284–6.
38. R Development Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2012. p. 2012.
39. Vazquez AI, Bates DM, Rosa GJ, Gianola D, Weigel KA. Technical note: an R package for fitting generalized linear mixed models in animal breeding. J Anim Sci. 2010;88:497–504.
40. Wright K. Package corrgram. R package version 1.9. 2016. Available from: https://cran.r-project.org/web/packages/corrgram/index.html. Accessed 30 Sept 2016.

**Bibliography**


1. Devall, M., Mill, J. and Lunnon, K. (2014) The mitochondrial epigenome: a role in Alzheimer's disease? *Epigenomics*, **6**, 665-675.
2. Devall, M., Roubroeks, J., Mill, J., Weedon, M. and Lunnon, K. (2016) Epigenetic regulation of mitochondrial function in neurodegenerative disease: New insights from advances in genomic technologies. *Neuroscience letters*.
3. Waddington, C. (1957) *The Strategy of the Genes. A Discussion of Some Aspects of Theoretical Biology*. Allen & Unwin, Ltd.
4. Lunnon, K. and Mill, J. (2013) Epigenetic studies in Alzheimer's disease: Current findings, caveats, and considerations for future studies. *American journal of medical genetics. Part B, Neuropsychiatric genetics: the official publication of the International Society of Psychiatric Genetics*.
5. Tucker, K.L. (2001) Methylated cytosine and the brain: A new base for neuroscience. *Neuron*, **30**, 649-652.
6. Bogdanovic, O. and Veenstra, G.J.C. (2009) DNA methylation and methyl-CpG binding proteins: developmental requirements and function. *Chromosoma*, **118**, 549-565.
7. Zampieri, M., Ciccarone, F., Calabrese, R., Franceschi, C., Burkle, A. and Caiafa, P. (2015) Reconfiguration of DNA methylation in aging. *Mechanisms of ageing and development*, **151**, 60- 70.
8. He, Y. and Ecker, J.R. (2015) Non-CG Methylation in the Human Genome. *Annual review of genomics and human genetics*, **16**, 55-77.
9. Ziller, M.J., Muller, F., Liao, J., Zhang, Y., Gu, H., Bock, C., Boyle, P.,

Epstein, C.B., Bernstein, B.E., Lengauer, T. *et al.* (2011) Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS genetics*, **7**, e1002389.

10. Guo, W., Zhang, M.Q. and Wu, H. (2016) Mammalian non-CG methylations are conserved and cell-type specific and may have been involved in the evolution of transposon elements. *Scientific reports*, **6**, 32207.

11. Meller, S., Zipfel, L., Gevensleben, H., Dietrich, J., Ellinger, J., Majores, M., Stein, J., Sailer, V., Jung, M., Kristiansen, G. *et al.* (2016) CDO1 Promoter Methylation Is Associated with Gene Silencing and Is a Prognostic Biomarker for Biochemical Recurrence-free Survival in Prostate Cancer Patients. *Epigenetics*, **11**, 871-880.

12. Lim, Y.C., Chia, S.Y., Jin, S., Han, W., Ding, C. and Sun, L. (2016) Dynamic DNA methylation landscape defines brown and white cell specificity during adipogenesis. *Molecular metabolism*, **5**, 1033-1041.

13. Jones, P.A. (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond.*Nature reviews. Genetics*, **13**, 484-492.

14. Haupt S., S.V.S.A., Leipe J., Schulze-Koops H., Skapenko A. Methylation of an intragenic alternative promoter regulates transcription of GARP. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, **1859**, 223-234.

15. Schmidl, C., Klug, M., Boeld, T.J., Andreesen, R., Hoffmann, P., Edinger, M. and Rehli, M. (2009) Lineage-specific DNA methylation in T cells correlates with histone methylation and enhancer activity. *Genome research*, **19**, 1165-1174.

16. Yang, X.J., Han, H., De Carvalho, D.D., Lay, F.D., Jones, P.A. and Liang,

G.N. (2014) Gene Body Methylation Can Alter Gene Expression and Is a Therapeutic Target in Cancer. *Cancer Cell*, **26**, 577-590.

17. Goll, M.G. and Bestor, T.H. (2005) Eukaryotic cytosine methyltransferases. *Annu Rev Biochem*, **74**, 481-514.
18. Mohan, K.N. (2016) Stem Cell Models to Investigate the Role of DNA Methylation Machinery in Development of Neuropsychiatric Disorders. *Stem Cells Int*.
19. Wood, K.H. and Zhou, Z. (2016) Emerging Molecular and Biological Functions of MBD2, a Reader of DNA Methylation. *Frontiers in genetics*, **7**, 93.
20. Bardhan, K. and Liu, K. (2013) Epigenetics and colorectal cancer pathogenesis. *Cancers (Basel)*, **5**, 676-713.
21. Ooi, S.K.T. and Bestor, T.H. (2008) The colorful history of active DNA demethylation. *Cell*, **133**, 1145-1148.
22. Branco, M.R., Ficz, G. and Reik, W. (2012) Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nature reviews. Genetics*, **13**, 7-13.
23. Nestor, C.E., Ottaviano, R., Reddington, J., Sproul, D., Reinhardt, D., Dunican, D., Katz, E., Dixon, J.M., Harrison, D.J. and Meehan, R.R. (2012) Tissue type is a major modifier of the 5- hydroxymethylcytosine content of human genes. *Genome research*, **22**, 467-477.
24. Wu, H. and Zhang, Y. (2015) Charting oxidized methylcytosines at base resolution. *Nat Struct Mol Biol*, **22**, 1-6.
25. Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F. *et al.* (1981) Sequence and organization of the human mitochondrial genome. *Nature*,

**290**, 457-465.

26. Wiesner, R.J., Ruegg, J.C. and Morano, I. (1992) Counting Target Molecules by Exponential Polymerase Chain-Reaction - Copy Number of Mitochondrial-DNA in Rat-Tissues. *Biochem Bioph Res Co*, **183**, 553-559.

27. Chinnery, P.F., Elliott, H.R., Hudson, G., Samuels, D.C. and Relton, C.L. (2012) Epigenetics, epidemiology and mitochondrial DNA diseases. *International journal of epidemiology*, **41**, 177- 187.

28. Gellerich, F.N., Trumbeckaite, S., Opalka, J.R., Seppet, E., Rasmussen, H.N., Neuhoff, C. and Zierz, S. (2000) Function of the mitochondrial outer membrane as a diffusion barrier in health and diseases. *Biochemical Society transactions*, **28**, 164-169.

29. Kuhlbrandt, W. (2015) Structure and function of mitochondrial membrane protein complexes. *Bmc Biol*, **13**.

30. Księżakowska-Łakoma K, Ż.M., Wilczyński JR. (2014) Mitochondrial dysfunction in cancer. *Prz Menopauzalny*, **13**, 136-144.

31. Cogliati, S., J.A., E. and Scorrano, L. (2016) Mitochondrial Cristae: Where Beauty Meets Functionality. *Trends in Biochemical Sciences*, **41**, 261-273.

32. Brini, M., Cali, T., Ottolini, D. and Carafoli, E. (2014) Neuronal calcium signaling: function and dysfunction. *Cellular and Molecular Life Sciences*, **71**, 2787-2814.

33. Singh, A.K., Pandey, P., Tewari, M., Pandey, H.P. and Shukla, H.S. (2014) Human mitochondrial genome flaws and risk of cancer. *Mitochondrial DNA*, **25**, 329-334.

34. Yang, Y., Karakhanova, S., Hartwig, W., D'Haese, J.G., Philippov, P.P.,

Werner, J. and Bazhin, A.V. (2016) Mitochondria and Mitochondrial ROS in Cancer: Novel Targets for Anticancer Therapy. *J Cell Physiol*.

35. Yuan, Y., Cruzat, V.F., Newshome, P., Cheng, J., Chen, Y. and Lu, Y. (2016) Regulation of SIRT1 in aging: Roles in mitochondrial function and biogenesis. *Mechanisms of ageing and development*, **155**, 10-21.

36. Bonora, M. and Pinton, P. (2014) The mitochondrial permeability transition pore and cancer: molecular mechanisms involved in cell death. *Front Oncol*, **4**, 302.

37. Biasutto L, Azzolini M, Szabòa I and M, Z. (2016) The mitochondrial permeability transition pore in AD 2016: An update. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*.

38. Azzolin, L., von Stockum, S., Basso, E., Petronilli, V., Forte, M.A. and Bernardi, P. (2010) The mitochondrial permeability transition from yeast to mammals. *FEBS Lett*, **584**, 2504-2509.

39. Bhola, P.D. and Letai, A. (2016) Mitochondria-Judges and Executioners of Cell Death Sentences. *Mol Cell*, **61**, 695-704.

40. Halestrap, A.P. and Richardson, A.P. (2015) The mitochondrial permeability transition: a current perspective on its identity and role in ischaemia/reperfusion injury. *J Mol Cell Cardiol*, **78**, 129- 141.

41. Faizi, M., Salimi, A., Rasoulzadeh, M., Naserzadeh, P. and Pourahmad, J. (2014) Schizophrenia induces oxidative stress and cytochrome C release in isolated rat brain mitochondria: a possible pathway for induction of apoptosis and neurodegeneration. *Iran J Pharm Res*, **13**, 93-100.

42. Yu, T., Chen, C., Sun, Y., Sun, H., Li, T.H., Meng, J. and Shi, X.H. (2015) ABT-737 sensitizes curcumin-induced anti-melanoma cell

activity through facilitating mPTP death pathway. *Biochemical and biophysical research communications*, **464**, 286-291.

43. Sun, N., Youle, R.J. and Finkel, T. (2016) The Mitochondrial Basis of Aging. *Mol Cell*, **61**, 654-666.

44. Lazarou, M., Sliter, D.A., Kane, L.A., Sarraf, S.A., Wang, C., Burman, J.L., Sideris, D.P., Fogel, A.I. and Youle, R.J. (2015) The ubiquitin kinase PINK1 recruits autophagy receptors to induce mitophagy. *Nature*, **524**, 309-314.

45. Kalinderi, K., Bostantjopoulou, S. and Fidani, L. (2016) The genetic background of Parkinson's disease: current progress and future prospects. *Acta Neurol Scand*.

46. Cai, Q. and Tammineni, P. (2016) Alterations in Mitochondria! Quality Control in Alzheimer's Disease. *Frontiers in cellular neuroscience*, **10**.

47. Chen, H.C., Vermulst, M., Wang, Y.E., Chomyn, A., Prolla, T.A., McCaffery, J.M. and Chan, D.C. (2010) Mitochondrial Fusion Is Required for mtDNA Stability in Skeletal Muscle and Tolerance of mtDNA Mutations. *Cell*, **141**, 280-289.

48. Lee, Y.J., Jeong, S.Y., Karbowski, M., Smith, C.L. and Youle, R.J. (2004) Roles of the mammalian mitochondrial fission and fusion mediators Fis1, Drp1, and Opa1 in apoptosis. *Molecular biology of the cell*, **15**, 5001-5011.

49. Senft, D. and Ronai, Z. (2016) Regulators of mitochondrial dynamics in cancer. *Current Opinion in Cell Biology*, **39**, 43-52.

50. Chen, H.C., Detmer, S.A., Ewald, A.J., Griffin, E.E., Fraser, S.E. and Chan, D.C. (2003) Mitofusins Mfn1 and Mfn2 coordinately regulate mitochondrial fusion and are essential for embryonic development. *J Cell Biol*, **160**, 189-200.

51. Kuzmicic, J., del Campo, A., Lopez-Crisosto, C., Morales, P.E., Pennanen, C., Bravo-Sagua, R., Hechenleitner, J., Zepeda, R., Castro, P.F., Verdejo, H.E. *et al.* (2011) Mitochondrial Dynamics: a Potential New Therapeutic Target for Heart Failure. *Rev Esp Cardiol*, **64**, 916-923.

52. Stewart, J.B. and Chinnery, P.F. (2015) The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease. *Nature Reviews Genetics*, **16**, 530-542.

53. Wallace, D.C. and Chalkia, D. (2013) Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease. *Cold Spring Harbor perspectives in medicine*, **3**, a021220.

54. Chinnery, P.F., Mowbray, C., Patel, S.K., Elson, J.L., Sampson, M., Hitman, G.A., McCarthy, M.I., Hattersley, A.T. and Walker, M. (2007) Mitochondrial DNA haplogroups and type 2 diabetes: a study of 897 cases and 1010 controls. *Journal of medical genetics*, **44**.

55. Cano, D., Gomez, C.F., Ospina, N., Cajigas, J.A., Groot, H., Andrade, R.E. and Torres, M.M. (2014) Mitochondrial DNA haplogroups and susceptibility to prostate cancer in a colombian population. *ISRN Oncol*, **2014**, 530675.

56. Kofler, B., Mueller, E.E., Eder, W., Stanger, O., Maier, R., Weger, M., Haas, A., Winker, R., Schmut, O., Paulweber, B. *et al.* (2009) Mitochondrial DNA haplogroup T is associated with coronary artery disease and diabetic retinopathy: a case control study. *Bmc Med Genet*, **10**.

57. Anderson, C.D., Biffi, A., Rahman, R., Ross, O.A., Jagiella, J.M., Kissela, B., Cole, J.W., Cortellini, L., Rost, N.S., Cheng, Y.C. *et al.* (2011) Common mitochondrial sequence variants in ischemic stroke. *Annals of neurology*, **69**,

471-480.

58. Flaquer, A., Baumbach, C., Kriebel, J., Meitinger, T., Peters, A., Waldenberger, M., Grallert, H. and Strauch, K. (2014) Mitochondrial Genetic Variants Identified to Be Associated with BMI in Adults. *PloS one*, **9**.

59. Rossignol, R., Faustin, B., Rocher, C., Malgat, M., Mazat, J.P. and Letellier, T. (2003) Mitochondrial threshold effects. *Biochem J*, **370**, 751-762.

60. Zhu, D.P., Economou, E.P., Antonarakis, S.E. and Maumenee, I.H. (1992) Mitochondrial-DNA Mutation and Heteroplasmy in Type-I Leber Hereditary Optic Neuropathy. *Am J Med Genet*, **42**, 173-179.

61. Thangaraj, K., Joshi, M.B., Reddy, A.G., Rasalkar, A.A. and Singh, L. (2003) Sperm mitochondrial mutations as a cause of low sperm motility. *Journal of andrology*, **24**, 388-392.

62. Yao, Y.G., Kong, Q.P., Salas, A. and Bandelt, H.J. (2008) Pseudomitochondrial genome haunts disease studies. *Journal of medical genetics*, **45**, 769-772.

63. Ghosh, S., Sengupta, S. and Scaria, V. (2014) Comparative analysis of human mitochondrial methylomes shows distinct patterns of epigenetic regulation in mitochondria. *Mitochondrion*, **18**, 58-62.

64. Devall, M., Smith, R.G., Jeffries, A.R., Hannon, E., Davies, M.N., Schalkwyk, L., Mill, J., Weedon, M. and Lunnon, K. (2016) Regional differences in mitochondrial DNA methylation in human post- mortem brain tissue. *Clinical epigenetics*, **In Press**.

65. Ho, S.Y. and Gilbert, M.T. (2010) Ancient mitogenomics. *Mitochondrion*, **10**, 1-11.

66. Tsuji, J., Frith, M.C., Tomii, K. and Horton, P. (2012) Mammalian NUMT insertion

is non-random. *Nucleic acids research*, **40**, 9073-9088.

67. Dawid, I.B. (1974) 5-methylcytidylic acid: absence from mitochondrial DNA of frogs and HeLa cells. *Science*, **184**, 80-81.

68. Nass, M.M. (1973) Differential methylation of mitochondrial and nuclear DNA in cultured mouse, hamster and virus-transformed hamster cells. In vivo and in vitro methylation. *Journal of molecular biology*, **80**, 155-175.

69. Groot, G.S. and Kroon, A.M. (1979) Mitochondrial DNA from various organisms does not contain internally methylated cytosine in -CCGG- sequences. *Biochimica et biophysica acta*, **564**, 355- 357.

70. Hong, E.E., Okitsu, C.Y., Smith, A.D. and Hsieh, C.L. (2013) Regionally specific and genome-wide analyses conclusively demonstrate the absence of CpG methylation in human mitochondrial DNA. *Mol Cell Biol*, **33**, 2683-2690.

71. Bellizzi, D., D'Aquila, P., Scafone, T., Giordano, M., Riso, V., Riccio, A. and Passarino, G. (2013) The control region of mitochondrial DNA shows an unusual CpG and non-CpG methylation pattern. *DNA research : an international journal for rapid publication of reports on genes and genomes*, **20**, 537-547.

72. Shock, L.S., Thakkar, P.V., Peterson, E.J., Moran, R.G. and Taylor, S.M. (2011) DNA methyltransferase 1, cytosine methylation, and cytosine hydroxymethylation in mammalian mitochondria. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 3630-3635.

73. Chestnut, B.A., Chang, Q., Price, A., Lesuisse, C., Wong, M. and Martin, L.J. (2011) Epigenetic regulation of motor neuron cell death through DNA methylation. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **31**, 16619-16636.

74. D'Aquila, P., Giordano, M., Montesanto, A., De Rango, F., Passarino, G. and Bellizzi, D. (2015) Age-and gender-related pattern of methylation in the MT-RNR1 gene. *Epigenomics*, **7**, 707-716.

75. Feng, S., Xiong, L., Ji, Z., Cheng, W. and Yang, H. (2012) Correlation between increased ND2 expression and demethylated displacement loop of mtDNA in colorectal cancer. *Mol Med Rep*, **6**, 125-130.

76. Phillips, A.C., Sleigh, A., McAllister, C.J., Brage, S., Carpenter, T.A., Kemp, G.J. and Holland, A.J. (2013) Defective mitochondrial function in vivo in skeletal muscle in adults with Down's syndrome: a 31P-MRS study. *PloS one*, **8**, e84031.

77. Coyle, J.T., Oster-Granite, M.L. and Gearhart, J.D. (1986) The neurobiologic consequences of Down syndrome. *Brain research bulletin*, **16**, 773-787.

78. Wisniewski, K.E., Wisniewski, H.M. and Wen, G.Y. (1985) Occurrence of neuropathological changes and dementia of Alzheimer's disease in Down's syndrome. *Annals of neurology*, **17**, 278- 282.

79. Byun, H.M., Panni, T., Motta, V., Hou, L., Nordio, F., Apostoli, P., Bertazzi, P.A. and Baccarelli, A.A. (2013) Effects of airborne pollutants on mitochondrial DNA methylation. *Part Fibre Toxicol*, **10**, 18.

80. Pirola, C.J., Gianotti, T.F., Burgueno, A.L., Rey-Funes, M., Loidl, C.F., Mallardi, P., Martino, J.S., Castano, G.O. and Sookoian, S. (2013) Epigenetic modification of liver mitochondrial DNA is associated with histological severity of nonalcoholic fatty liver disease. *Gut*, **62**, 1356-1363.

81. Infantino, V., Castegna, A., Iacobazzi, F., Spera, I., Scala, I., Andria, G. and Iacobazzi, V. (2011) Impairment of methyl cycle affects mitochondrial

methyl availability and glutathione level in Down's syndrome. *Mol Genet Metab*, **102**, 378-382.

82. Wong, M., Gertz, B., Chestnut, B.A. and Martin, L.J. (2013) Mitochondrial DNMT3A and DNA methylation in skeletal muscle and CNS of transgenic mouse models of ALS. *Front Cell Neurosci*, **7**, 279.

83. Wen, S.L., Zhang, F. and Feng, S. (2013) Decreased copy number of mitochondrial DNA: A potential diagnostic criterion for gastric cancer. *Oncol Lett*, **6**, 1098-1102.

84. Bai, R.K., Leal, S.M., Covarrubias, D., Liu, A. and Wong, L.J.C. (2007) Mitochondrial genetic background modifies breast cancer risk. *Cancer Res*, **67**, 4687-4694.

85. Wallace, D.C. (1994) Mitochondrial-DNA Sequence Variation in Human-Evolution and Disease. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 8739- 8746.

86. Mishra, M. and Kowluru, R.A. (2015) Epigenetic Modification of Mitochondrial DNA in the Development of Diabetic Retinopathy. *Invest Ophthalmol Vis Sci*, **56**, 5133-5142.

87. Liu, Z., Guo, J., Sun, H., Huang, Y., Zhao, R. and Yang, X. (2015) alpha-Lipoic acid attenuates LPS- induced liver injury by improving mitochondrial function in association with GR mitochondrial DNA occupancy. *Biochimie*.

88. Lee, W., Johnson, J., Gough, D.J., Donoghue, J., Cagnone, G.L., Vaghjiani, V., Brown, K.A., Johns, T.G. and St John, J.C. (2015) Mitochondrial DNA copy number is regulated by DNA methylation and demethylation of POLGA in stem and cancer cells and their differentiated progeny. *Cell Death Dis*, **6**, e1664.

89. Smiraglia, D.J., Kulawiec, M., Bistulfi, G.L., Gupta, S.G. and Singh, K.K.

(2008) A novel role for mitochondria in regulating epigenetic modification in the nucleus. *Cancer biology & therapy*, **7**, 1182-1190.

90. Bellizzi, D., D'Aquila, P., Giordano, M., Montesanto, A. and Passarino, G. (2012) Global DNA methylation levels are modulated by mitochondrial DNA variants. *Epigenomics*, **4**, 17-27.

91. Manev, H., Dzitoyeva, S. and Chen, H. (2012) Mitochondrial DNA: A Blind Spot in Neuroepigenetics. *Biomolecular concepts*, **3**, 107-115.

92. Atilano, S.R., Malik, D., Chwa, M., Caceres-Del-Carpio, J., Nesburn, A.B., Boyer, D.S., Kuppermann, B.D., Jazwinski, S.M., Miceli, M.V., Wallace, D.C. *et al.* (2015) Mitochondrial DNA variants can mediate methylation status of inflammation, angiogenesis and signaling genes. *Human molecular genetics*.

93. Takasugi, M., Yagi, S., Hirabayashi, K. and Shiota, K. (2010) DNA methylation status of nuclear- encoded mitochondrial genes underlies the tissue-dependent mitochondrial functions. *BMC genomics*, **11**, 481.

94. Lunnon, K., Smith, R., Hannon, E., De Jager, P.L., Srivastava, G., Volta, M., Troakes, C., Al-Sarraj, S., Burrage, J., Macdonald, R. *et al.* (2014) Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. *Nat Neurosci*, **17**, 1164-1170.

95. De Jager, P.L., Srivastava, G., Lunnon, K., Burgess, J., Schalkwyk, L.C., Yu, L., Eaton, M.L., Keenan, B.T., Ernst, J., McCabe, C. *et al.* (2014) Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. *Nat Neurosci*, **17**, 1156-1163.

96. Ng, C.W., Yildirim, F., Yap, Y.S., Dalin, S., Matthews, B.J., Velez, P.J.,

Labadorf, A., Housman, D.E. and Fraenkel, E. (2013) Extensive changes in DNA methylation are associated with expression of mutant huntingtin. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 2354-2359.

97. Moore, K., McKnight, A.J., Craig, D. and O'Neill, F. (2014) Epigenome-wide association study for Parkinson's disease. *Neuromolecular medicine*, **16**, 845-855.

98. Janssen, B.G., Byun, H.M., Gyselaers, W., Lefebvre, W., Baccarelli, A.A. and Nawrot, T.S. (2015) Placental mitochondrial methylation and exposure to airborne particulate matter in the early life environment: An ENVIRONAGE birth cohort study. *Epigenetics*, **10**, 536-544.

99. Ghosh, S., Sengupta, S. and Scaria, V. (2016) Hydroxymethyl cytosine marks in the human mitochondrial genome are dynamic in nature. *Mitochondrion*, **27**, 25-31.

100. Clark, C., Palta, P., Joyce, C.J., Scott, C., Grundberg, E., Deloukas, P., Palotie, A. and Coffey, A.J. (2012) A comparison of the whole genome approach of MeDIP-seq to the targeted approach of the Infinium HumanMethylation450 BeadChip((R)) for methylome profiling. *PloS one*, **7**, e50233.

101. Hirst, M. and Marra, M.A. (2010) Next generation sequencing based approaches to epigenomics. *Briefings in functional genomics*, **9**, 455-465.

102. Stroud, H., Feng, S., Morey Kinney, S., Pradhan, S. and Jacobsen, S.E. (2011) 5- Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome biology*, **12**, R54.

103. Condliffe, D., Wong, A., Troakes, C., Proitsi, P., Patel, Y., Chouliaras, L.,

Fernandes, C., Cooper, J., Lovestone, S., Schalkwyk, L. *et al.* (2014) Cross-region reduction in 5-hydroxymethylcytosine in Alzheimer's disease brain. *Neurobiology of aging*, **35**, 1850-1854.

104. Dzitoyeva, S., Chen, H. and Manev, H. (2012) Effect of aging on 5-hydroxymethylcytosine in brain mitochondria. *Neurobiology of aging*, **33**, 2881-2891.

105. Fouse, S.D., Nagarajan, R.O. and Costello, J.F. (2010) Genome-scale DNA methylation analysis. *Epigenomics*, **2**, 105-117.

106. Tan, L., Xiong, L., Xu, W., Wu, F., Huang, N., Xu, Y., Kong, L., Zheng, L., Schwartz, L., Shi, Y. *et al.* (2013) Genome-wide comparison of DNA hydroxymethylation in mouse embryonic stem cells and neural progenitor cells by a new comparative hMeDIP-seq method. *Nucleic acids research*, **41**, e84.

107. Huang, J., Renault, V., Sengenes, J., Touleimat, N., Michel, S., Lathrop, M. and Tost, J. (2012) MeQA: a pipeline for MeDIP-seq data quality assessment and analysis. *Bioinformatics*, **28**, 587- 588.

108. Laird, P.W. (2010) Principles and challenges of genomewide DNA methylation analysis. *Nature reviews. Genetics*, **11**, 191-203.

109. Chavez, L., Jozefczuk, J., Grimm, C., Dietrich, J., Timmermann, B., Lehrach, H., Herwig, R. and Adjaye, J. (2010) Computational analysis of genome-wide DNA methylation during the differentiation of human embryonic stem cells along the endodermal lineage. *Genome research*, **20**, 1441-1450.

110. Lee, E.J., Luo, J., Wilson, J.M. and Shi, H. (2013) Analyzing the cancer methylome through targeted bisulfite sequencing. *Cancer*

*letters*, **340**, 171-178.

111. Gu, H., Smith, Z.D., Bock, C., Boyle, P., Gnirke, A. and Meissner, A. (2011) Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nature protocols*, **6**, 468-481.

112. Stirzaker, C., Taberlay, P.C., Statham, A.L. and Clark, S.J. (2014) Mining cancer methylomes: prospects and challenges. *Trends in genetics : TIG*, **30**, 75-84.

113. Garrett-Bakelman, F.E., Sheridan, C.K., Kacmarczyk, T.J., Ishii, J., Betel, D., Alonso, A., Mason, C.E., Figueroa, M.E. and Melnick, A.M. (2015) Enhanced reduced representation bisulfite sequencing for assessment of DNA methylation at base pair resolution. *Journal of visualized experiments : JoVE*.

114. Guo, H., Zhu, P., Wu, X., Li, X., Wen, L. and Tang, F. (2013) Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome research*, **23**, 2126-2135.

115. Guo, H., Zhu, P., Guo, F., Li, X., Wu, X., Fan, X., Wen, L. and Tang, F. (2015) Profiling DNA methylome landscapes of mammalian cells with single-cell reduced-representation bisulfite sequencing. *Nature protocols*, **10**, 645-659.

116. Krueger, F., Kreck, B., Franke, A. and Andrews, S.R. (2012) DNA methylome analysis using short bisulfite sequencing data. *Nat Meth*, **9**, 145-151.

117. Stewart, S.K., Morris, T.J., Guilhamon, P., Bulstrode, H., Bachman, M., Balasubramanian, S. and Beck, S. (2015) oxBS-450K: A method for

analysing hydroxymethylation using 450K BeadChips. *Methods*, **72**, 9-15.

118. Booth, M.J., Branco, M.R., Ficz, G., Oxley, D., Krueger, F., Reik, W. and Balasubramanian, S. (2012) Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science*, **336**, 934-937.

119. Booth, M.J., Ost, T.W., Beraldi, D., Bell, N.M., Branco, M.R., Reik, W. and Balasubramanian, S. (2013) Oxidative bisulfite sequencing of 5-methylcytosine and 5-hydroxymethylcytosine. *Nature protocols*, **8**, 1841-1851.

120. Booth, M.J., Marsico, G., Bachman, M., Beraldi, D. and Balasubramanian, S. (2014) Quantitative sequencing of 5-formylcytosine in DNA at single-base resolution. *Nature chemistry*, **6**, 435-440.

121. Xi, Y. and Li, W. (2009) BSMAP: whole genome bisulfite sequence MAPping program. *BMC bioinformatics*, **10**, 232.

122. Schadt, E.E., Turner, S. and Kasarskis, A. (2010) A window into third-generation sequencing. *Human molecular genetics*, **19**, R227-R240.

123. Davis, B.M., Chao, M.C. and Waldor, M.K. (2013) Entering the era of bacterial epigenomics with single molecule real time DNA sequencing. *Current opinion in microbiology*, **16**, 192-198.

124. Flusberg, B.A., Webster, D.R., Lee, J.H., Travers, K.J., Olivares, E.C., Clark, T.A., Korlach, J. and Turner, S.W. (2010) Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nature methods*, **7**, 461-465.

125. Sims, D., Sudbery, I., Ilott, N.E., Heger, A. and Ponting, C.P. (2014) Sequencing depth and coverage: key considerations in genomic

analyses. *Nature reviews. Genetics*, **15**, 121-132.

126. Steinbock, L. and Radenovic, A. (2015) The emergence of nanopores in next-generation sequencing. *Nanotechnology*, **26**, 074003.

127. Song, C.X., Clark, T.A., Lu, X.Y., Kislyuk, A., Dai, Q., Turner, S.W., He, C. and Korlach, J. (2012) Sensitive and specific single-molecule sequencing of 5-hydroxymethylcytosine. *Nature methods*, **9**, 75-77.

128. Ono, Y., Asai, K. and Hamada, M. (2013) PBSIM: PacBio reads simulator--toward accurate genome assembly. *Bioinformatics*, **29**, 119-121.

129. Roberts, R.J., Carneiro, M.O. and Schatz, M.C. (2013) The advantages of SMRT sequencing.*Genome Biol*, **14**, 405.

130. McGinn, S. and Gut, I.G. (2013) DNA sequencing–spanning the generations. *New biotechnology*, **30**, 366-372.

131. Kilianski, A., Haas, J.L., Corriveau, E.J., Liem, A.T., Willis, K.L., Kadavy, D.R., Rosenzweig, C.N. and Minot, S.S. (2015) Bacterial and viral identification and differentiation by amplicon sequencing on the MinION nanopore sequencer. *Gigascience*, **4**, 12.

132. Branton, D., Deamer, D.W., Marziali, A., Bayley, H., Benner, S.A., Butler, T., Di Ventra, M., Garaj, S., Hibbs, A. and Huang, X. (2008) The potential and challenges of nanopore sequencing. *Nature biotechnology*, **26**, 1146-1153.

133. Hyun, B.R., McElwee, J.L. and Soloway, P.D. (2015) Single molecule and single cell epigenomics. *Methods*, **72**, 41-50.

134. Devall, M., Burrage, J., Caswell, R., Johnson, M., Troakes, C., Al-Saraj, S., Jeffries, A., Mill, J. and Lunnon, K. (2015) A comparison of mitochondrial

DNA isolation methods in frozen post-mortem human brain tissue—applications for studies of mitochondrial genetics in brain disorders. *BioTechniques*, **59**, 241-246.

135. Larsen, P.A., Heilman, A.M. and Yoder, A.D. (2014) The utility of PacBio circular consensus sequencing for characterizing complex gene families in non-model organisms. *BMC genomics*, **15**, 720.

136. Simpson, L., Douglass, S.M., Lake, J.A., Pellegrini, M. and Li, F. (2015) Comparison of the Mitochondrial Genomes and Steady State Transcriptomes of Two Strains of the Trypanosomatid Parasite, Leishmania tarentolae. *PLoS Negl Trop Dis*, **9**, e0003841.

137. Wolters, J.F., Chiu, K. and Fiumera, H.L. (2015) Population structure of mitochondrial genomes in Saccharomyces cerevisiae. *BMC genomics*, **16**, 451.

138. Taniguti, L.M., Schaker, P.D., Benevenuto, J., Peters, L.P., Carvalho, G., Palhares, A., Quecine, M.C., Nunes, F.R., Kmit, M.C., Wai, A. *et al.* (2015) Complete Genome Sequence of Sporisorium scitamineum and Biotrophic Interaction Transcriptome with Sugarcane. *PloS one*, **10**, e0129318.

139. Peng, Y., Lai, Z., Lane, T., Nageswara-Rao, M., Okada, M., Jasieniuk, M., O'Geen, H., Kim, R.W., Sammons, R.D., Rieseberg, L.H. *et al.* (2014) De novo genome assembly of the economically important weed horseweed using integrated data from multiple sequencing platforms. *Plant Physiol*, **166**, 1241-1254.

140. Powers, J.G., Weigman, V.J., Shu, J., Pufky, J.M., Cox, D. and Hurban, P. (2013) Efficient and accurate whole genome assembly and methylome profiling of E. coli. *BMC genomics*, **14**, 675.

141. Lee, W.C., Anton, B.P., Wang, S., Baybayan, P., Singh, S., Ashby, M., Chua, E.G., Tay, C.Y., Thirriot, F., Loke, M.F. *et al.* (2015) The complete methylome of Helicobacter pylori UM032. *BMC genomics*, **16**, 424.

142. Schreiber, J., Wescoe, Z.L., Abu-Shumays, R., Vivian, J.T., Baatar, B., Karplus, K. and Akeson, M. (2013) Error rates for nanopore discrimination among cytosine, methylcytosine, and hydroxymethylcytosine along individual DNA strands. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 18910-18915.

143. Laszlo, A.H., Derrington, I.M., Brinkerhoff, H., Langford, K.W., Nova, I.C., Samson, J.M., Bartlett, J.J., Pavlenok, M. and Gundlach, J.H. (2013) Detection and mapping of 5-methylcytosine and 5- hydroxymethylcytosine with nanopore MspA. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 18904-18909.

144. Song, H., Buhay, J.E., Whiting, M.F. and Crandall, K.A. (2008) Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 13486-13491.

145. Chen, H., Dzitoyeva, S. and Manev, H. (2012) Effect of valproic acid on mitochondrial epigenetics. *European journal of pharmacology*, **690**, 51-59.

146. Zhou, J., Liu, L. and Chen, J. (2010) Method to purify mitochondrial DNA directly from yeast total DNA. *Plasmid*, **64**, 196-199.

147. Lang, B.F., Gray, M.W. and Burger, G. (1999) Mitochondrial genome

evolution and the origin of eukaryotes. *Annual review of genetics*, **33**, 351-397.

148. Ricchetti, M., Tekaia, F. and Dujon, B. (2004) Continued colonization of the human genome by mitochondrial DNA. *PLoS biology*, **2**, E273.

149. Triant, D.A. and DeWoody, J.A. (2008) Molecular analyses of mitochondrial pseudogenes within the nuclear genome of arvicoline rodents. *Genetica*, **132**, 21-33.

150. Mourier, T., Hansen, A.J., Willerslev, E. and Arctander, P. (2001) The Human Genome Project reveals a continuous transfer of large mitochondrial fragments to the nucleus. *Molecular biology and evolution*, **18**, 1833-1837.

151. Hirano, M., Shtilbans, A., Mayeux, R., Davidson, M.M., DiMauro, S., Knowles, J.A. and Schon, E.A. (1997) Apparent mtDNA heteroplasmy in Alzheimer's disease patients and in normals due to PCR amplification of nucleus-embedded mtDNA pseudogenes. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 14894-14899.

152. Davis, R.E., Miller, S., Herrnstadt, C., Ghosh, S.S., Fahy, E., Shinobu, L.A., Galasko, D., Thal, L.J., Beal, M.F., Howell, N. *et al.* (1997) Mutations in mitochondrial cytochrome c oxidase genes segregate with late-onset Alzheimer disease. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 4526-4531.

153. Iacobazzi, V., Castegna, A., Infantino, V. and Andria, G. (2013) Mitochondrial DNA methylation as a next-generation biomarker and diagnostic tool. *Molecular genetics and metabolism*, **110**, 25- 34.

154. Wallace, D.C. and Chalkia, D. (2013) Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease. *Cold Spring Harbor*

*perspectives in biology*, **5**, a021220.

155. Hua, S., Lu, C., Song, Y., Li, R., Liu, X., Quan, F., Wang, Y., Liu, J., Su, F. and Zhang, Y. (2012) High levels of mitochondrial heteroplasmy modify the development of ovine-bovine interspecies nuclear transferred embryos. *Reproduction, fertility, and development*, **24**, 501-509.

156. Carrieri, G., Bonafe, M., De Luca, M., Rose, G., Varcasia, O., Bruni, A., Maletta, R., Nacmias, B., Sorbi, S., Corsonello, F. *et al.* (2001) Mitochondrial DNA haplogroups and APOE4 allele are non- independent variables in sporadic Alzheimer's disease. *Human genetics*, **108**, 194-198.

157. Maruszak, A., Canter, J.A., Styczynska, M., Zekanowski, C. and Barcikowska, M. (2009) Mitochondrial haplogroup H and Alzheimer's disease--is there a connection? *Neurobiology of aging*, **30**, 1749-1755.

158. Davies, M.N., Volta, M., Pidsley, R., Lunnon, K., Dixit, A., Lovestone, S., Coarfa, C., Harris, R.A., Milosavljevic, A., Troakes, C. *et al.* (2012) Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome biology*, **13**, R43.

159. Sanchez-Mut, J.V., Aso, E., Panayotis, N., Lott, I., Dierssen, M., Rabano, A., Urdinguio, R.G., Fernandez, A.F., Astudillo, A., Martin-Subero, J.I. *et al.* (2013) DNA methylation map of mouse and human brain identifies target genes in Alzheimer's disease. *Brain : a journal of neurology*, **136**, 3018-3027.

160. Nestor, C., Ruzov, A., Meehan, R. and Dunican, D. (2010) Enzymatic approaches and bisulfite sequencing cannot distinguish between 5-methylcytosine and 5-hydroxymethylcytosine in DNA. *BioTechniques*, **48**,

317-319.

161. Bradley-Whitman, M.A. and Lovell, M.A. (2013) Epigenetic changes in the progression of Alzheimer's disease. *Mechanisms of ageing and development*, **134**, 486-495.

162. Sun, Z., Terragni, J., Borgaro, J.G., Liu, Y., Yu, L., Guan, S., Wang, H., Sun, D., Cheng, X., Zhu, Z. *et al.* (2013) High-resolution enzymatic mapping of genomic 5-hydroxymethylcytosine in mouse embryonic stem cells. *Cell reports*, **3**, 567-576.

163. Gao, J., Wen, S., Zhou, H. and Feng, S. (2015) De-methylation of displacement loop of mitochondrial DNA is associated with increased mitochondrial copy number and nicotinamide adenine dinucleotide subunit 2 expression in colorectal cancer. *Molecular medicine reports*, **12**, 7033-7038.

164. You, C., Ji, D., Dai, X. and Wang, Y. (2014) Effects of Tet-mediated oxidation products of 5- methylcytosine on DNA transcription in vitro and in mammalian cells. *Sci Rep*, **4**, 7052.

165. Pellerin, L. and Magistretti, P.J. (2003) How to balance the brain energy budget while spending glucose differently. *J Physiol-London*, **546**, 325-325.

166. Stauch, K.L., Purnell, P.R. and Fox, H.S. (2014) Aging synaptic mitochondria exhibit dynamic proteomic changes while maintaining bioenergetic function. *Aging-Us*, **6**, 320-334.

167. Tang, Y.G. and Zucker, R.S. (1997) Mitochondrial involvement in post-tetanic potentiation of synaptic transmission. *Neuron*, **18**, 483-491.

168. Chang, D.T.W. and Reynolds, I.J. (2006) Differences in mitochondrial movement and morphology in young and mature primary cortical neurons in culture. *Neuroscience*, **141**, 727-736.

169. Yang, J.L., Weissman, L., Bohr, V.A. and Mattson, M.P. (2008) Mitochondrial DNA damage and repair in neurodegenerative disorders. *DNA repair*, **7**, 1110-1120.

170. Uttara, B., Singh, A.V., Zamboni, P. and Mahajan, R.T. (2009) Oxidative Stress and Neurodegenerative Diseases: A Review of Upstream and Downstream Antioxidant Therapeutic Options. *Curr Neuropharmacol*, **7**, 65-74.

171. El-Hattab, A.W., Adesina, A.M., Jones, J. and Scaglia, F. (2015) MELAS syndrome: Clinical manifestations, pathogenesis, and treatment options. *Molecular genetics and metabolism*.

172. Schaefer, A.M., McFarland, R., Blakely, E.L., He, L., Whittaker, R.G., Taylor, R.W., Chinnery, P.F. and Turnbull, D.M. (2008) Prevalence of mitochondrial DNA disease in adults. *Annals of neurology*, **63**, 35-39.

173. Dening, K.H., King, M., Jones, L., Vickestaff, V. and Sampson, E.L. (2016) Advance Care Planning in Dementia: Do Family Carers Know the Treatment Preferences of People with Early Dementia? *PloS one*, **11**.

174. Braak, H. and Braak, E. (1995) Staging of Alzheimer's disease-related neurofibrillary changes. *Neurobiology of aging*, **16**, 271-278; discussion 278-284.

175. Hyman, B.T., Phelps, C.H., Beach, T.G., Bigio, E.H., Cairns, N.J., Carrillo, M.C., Dickson, D.W., Duyckaerts, C., Frosch, M.P., Masliah, E. *et al.* (2012) National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association*, **8**, 1-13.

176. Swerdlow, R.H., Burns, J.M. and Khan, S.M. (2010) The Alzheimer's Disease Mitochondrial Cascade Hypothesis. *Journal of Alzheimers*

*Disease*, **20**, S265-S279.

177. Pradelli, L.A., Beneteau, M. and Ricci, J.E. (2010) Mitochondrial control of caspase-dependent and -independent cell death. *Cellular and molecular life sciences : CMLS*, **67**, 1589-1597.

178. Chan, S.L., Liu, D., Kyriazis, G.A., Bagsiyao, P., Ouyang, X. and Mattson, M.P. (2006) Mitochondrial uncoupling protein-4 regulates calcium homeostasis and sensitivity to store depletion-induced apoptosis in neural cells. *The Journal of biological chemistry*, **281**, 37391- 37403.

179. Fu, W., Ruangkittisakul, A., MacTavish, D., Baker, G.B., Ballanyi, K. and Jhamandas, J.H. (2013) Activity and metabolism-related Ca2+ and mitochondrial dynamics in co-cultured human fetal cortical neurons and astrocytes. *Neuroscience*, **250**, 520-535.

180. Zhao, Y. and Zhao, B. (2013) Oxidative stress and the pathogenesis of Alzheimer's disease.*Oxidative medicine and cellular longevity*, **2013**, 316523.

181. Devi, L. and Ohno, M. (2012) Mitochondrial dysfunction and accumulation of the beta-secretase- cleaved C-terminal fragment of APP in Alzheimer's disease transgenic mice. *Neurobiology of disease*, **45**, 417-424.

182. Pinto, M., Pickrell, A.M., Fukui, H. and Moraes, C.T. (2013) Mitochondrial DNA damage in a mouse model of Alzheimer's disease decreases amyloid beta plaque formation. *Neurobiology of aging*, **34**, 2399-2407.

183. Swerdlow, R.H., Burns, J.M. and Khan, S.M. (2010) The Alzheimer's disease mitochondrial cascade hypothesis. *Journal of Alzheimer's disease : JAD*, **20 Suppl 2**, S265-279.

184. Reddy, P.H., McWeeney, S., Park, B.S., Manczak, M., Gutala, R.V., Partovi, D., Jung, Y., Yau, V., Searles, R., Mori, M. *et al.* (2004) Gene

expression profiles of transcripts in amyloid precursor protein transgenic mice: up-regulation of mitochondrial metabolism and apoptotic genes is an early cellular change in Alzheimer's disease. *Human molecular genetics*, **13**, 1225-1240.

185. Du, H. and ShiDu Yan, S. (2010) Unlocking the Door to Neuronal Woes in Alzheimer's Disease: Abeta and Mitochondrial Permeability Transition Pore. *Pharmaceuticals (Basel)*, **3**, 1936-1948.

186. Manczak, M. and Reddy, P.H. (2012) Abnormal interaction of VDAC1 with amyloid beta and phosphorylated tau causes mitochondrial dysfunction in Alzheimer's disease. *Human molecular genetics*, **21**, 5131-5146.

187. Du, H., Guo, L., Fang, F., Chen, D., Sosunov, A.A., McKhann, G.M., Yan, Y., Wang, C., Zhang, H., Molkentin, J.D. *et al.* (2008) Cyclophilin D deficiency attenuates mitochondrial and neuronal perturbation and ameliorates learning and memory in Alzheimer's disease. *Nature medicine*, **14**, 1097-1105.

188. Chouliaras, L., Rutten, B.P., Kenis, G., Peerbooms, O., Visser, P.J., Verhey, F., van Os, J., Steinbusch, H.W. and van den Hove, D.L. (2010) Epigenetic regulation in the pathophysiology of Alzheimer's disease. *Progress in neurobiology*, **90**, 498-510.

189. Lardenoije, R., Iatrou, A., Kenis, G., Kompotis, K., Steinbusch, H.W., Mastroeni, D., Coleman, P., Lemere, C.A., Hof, P.R., van den Hove, D.L. *et al.* (2015) The epigenetics of aging and neurodegeneration. *Prog Neurobiol*.

190. Gatz, M., Reynolds, C.A., Fratiglioni, L., Johansson, B., Mortimer, J.A., Berg, S., Fiske, A. and Pedersen, N.L. (2006) Role of genes and environments for explaining Alzheimer disease. *Archives of general psychiatry*, **63**, 168-174.

191. Lunnon, K. and Mill, J. (2013) Epigenetic studies in Alzheimer's disease:

current findings, caveats, and considerations for future studies. *Am J Med Genet B Neuropsychiatr Genet*, **162B**, 789-799.

192. Mastroeni, D., McKee, A., Grover, A., Rogers, J. and Coleman, P.D. (2009) Epigenetic differences in cortical neurons from a pair of monozygotic twins discordant for Alzheimer's disease. *PloS one*, **4**, e6617.

193. Mastroeni, D., Grover, A., Delvaux, E., Whiteside, C., Coleman, P.D. and Rogers, J. (2010) Epigenetic changes in Alzheimer's disease: decrements in DNA methylation. *Neurobiology of aging*, **31**, 2025-2037.

194. Chouliaras, L., Mastroeni, D., Delvaux, E., Grover, A., Kenis, G., Hof, P.R., Steinbusch, H.W., Coleman, P.D., Rutten, B.P. and van den Hove, D.L. (2013) Consistent decrease in global DNA methylation and hydroxymethylation in the hippocampus of Alzheimer's disease patients. *Neurobiology of aging*, **34**, 2091-2099.

195. Celarain, N., Sanchez-Ruiz de Gordoa, J., Zelaya, M.V., Roldan, M., Larumbe, R., Pulido, L., Echavarri, C. and Mendioroz, M. (2016) TREM2 upregulation correlates with 5- hydroxymethycytosine enrichment in Alzheimer's disease hippocampus. *Clinical epigenetics*, **8**, 37.

196. Smith, A.R., Smith, R.G., Condliffe, D., Hannon, E., Schalkwyk, L., Mill, J. and Lunnon, K. (2016) Increased DNA methylation near TREM2 is consistently seen in the superior temporal gyrus in Alzheimer's disease brain. *Neurobiol Aging*, **47**, 35-40.

197. Bakulski, K.M., Dolinoy, D.C., Sartor, M.A., Paulson, H.L., Konen, J.R., Lieberman, A.P., Albin, R.L., Hu, H. and Rozek, L.S. (2012) Genome-wide DNA methylation differences between late-onset Alzheimer's disease and cognitively normal controls in human frontal cortex. *Journal of Alzheimer's*

*disease : JAD*, **29**, 571-588.

198. Lunnon, K., Smith, R., Hannon, E.J., De Jager, P.L., Srivastava, G., Volta, M., Troakes, C., Al-Sarraj, S., Burrage, J., Macdonald, R. *et al.* (2014) Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. *Nat Neurosci*, **Sept; 17**, 1164-1170.

199. De Jager, P.L., Srivastava, G., Lunnon, K., Burgess, J., Schalkwyk, L.C., Yu, L., Eaton, M.L., Keenan, B.T., Ernst, J., McCabe, C. *et al.* (2014) Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. *Nature neuroscience*, **Sep;17**, 1156-1163.

200. Blanch, M., Mosquera, J.L., Ansoleaga, B., Ferrer, I. and Barrachina, M. (2016) Altered Mitochondrial DNA Methylation Pattern in Alzheimer Disease-Related Pathology and in Parkinson Disease. *Am J Pathol*, **186**, 385-397.

201. Marzi, S.J., Meaburn, E.L., Dempster, E.L., Lunnon, K., Paya-Cano, J.L., Smith, R.G., Volta, M., Troakes, C., Schalkwyk, L.C. and Mill, J. (2016) Tissue-specific patterns of allelically-skewed DNA methylation. *Epigenetics*, **11**, 24-35.

202. Devall, M., Burrage, J., Caswell, R., Johnson, M., Troakes, C., Al-Sarraj, S., Jeffries, A.R., Mill, J. and Lunnon, K. (2015) A comparison of mitochondrial DNA isolation methods in frozen post- mortem human brain tissue--applications for studies of mitochondrial genetics in brain disorders. *Biotechniques*, **59**, 241-242, 244-246.

203. Malik, A.N., Shahni, R. and Iqbal, M.M. (2009) Increased peripheral blood mitochondrial DNA in type 2 diabetic patients with nephropathy. *Diabetes Res Clin Pract*, **86**, e22-24.

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

205. R Development Core Team. (2012) R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing, Vienna, Austria 2012*.

206. Murgia, M. and Rizzuto, R. (2015) Molecular diversity and pleiotropic role of the mitochondrial calcium uniporter. *Cell Calcium*, **58**, 11-17.

207. Wang, W., Esbensen, Y., Kunke, D., Suganthan, R., Rachek, L., Bjoras, M. and Eide, L. (2011) Mitochondrial DNA damage level determines neural stem cell differentiation fate. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **31**, 9746-9751.

208. Lunnon, K., Ibrahim, Z., Proitsi, P., Lourdusamy, A., Newhouse, S., Sattlecker, M., Furney, S., Saleem, M., Soininen, H., Kloszewska, I. *et al.* (2012) Mitochondrial dysfunction and immune activation are detectable in early Alzheimer's disease blood. *Journal of Alzheimer's disease : JAD*, **30**, 685-710.

209. Lunnon, K., Keohane, A., Pidsley, R., Newhouse, S., Riddoch-Contreras, J., Thubron, E.B., Devall, M., Soininen, H., Kloszewska, I., Mecocci, P. *et al.* (2017) Mitochondrial genes are altered in blood early in Alzheimer's disease. *Neurobiol Aging*, **53**, 36-47.

210. Prabakaran, S., Swatton, J.E., Ryan, M.M., Huffaker, S.J., Huang, J.T., Griffin, J.L., Wayland, M., Freeman, T., Dudbridge, F., Lilley, K.S. *et al.* (2004) Mitochondrial dysfunction in schizophrenia: evidence for compromised brain metabolism and oxidative stress. *Molecular psychiatry*,

**9**, 684- 697, 643.

211. Clay, H.B., Sillivan, S. and Konradi, C. (2011) Mitochondrial dysfunction and pathology in bipolar disorder and schizophrenia. *International journal of developmental neuroscience : the official journal of the International Society for Developmental Neuroscience*, **29**, 311-324.

212. Chang, C.C., Jou, S.H., Lin, T.T., Lai, T.J. and Liu, C.S. (2015) Mitochondria DNA change and oxidative damage in clinically stable patients with major depressive disorder. *PloS one*, **10**, e0125855.

213. Lunnon, K., Hannon, E., Smith, R.G., Dempster, E., Wong, C., Burrage, J., Troakes, C., Al-Sarraj, S., Kepa, A., Schalkwyk, L. *et al.* (2016) Variation in 5-hydroxymethylcytosine across human cortex and cerebellum. *Genome biology*, **17**, 27.

214. Lowe, R., Slodkowicz, G., Goldman, N. and Rakyan, V.K. (2015) The human blood DNA methylome displays a highly distinctive profile compared with other somatic tissues. *Epigenetics: official journal of the DNA Methylation Society*, **10**, 274-281.

215. Lokk, K., Modhukur, V., Rajashekar, B., Martens, K., Magi, R., Kolde, R., Koltsina, M., Nilsson, T.K., Vilo, J., Salumets, A. *et al.* (2014) DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. *Genome biology*, **15**, r54.

216. Smith, A.R., Smith, R.G., Condliffe, D., Hannon, E., Schalkwyk, L., Mill, J. and Lunnon, K. (2016) Increased DNA methylation near TREM2 is consistently seen in the superior temporal gyrus in Alzheimer's disease brain. *Neurobiology of aging*, **47**, 35-40.

217. Pidsley, R., Viana, J., Hannon, E., Spiers, H.H., Troakes, C., Al-Saraj, S.,

Mechawar, N., Turecki, G., Schalkwyk, L.C., Bray, N.J. *et al.* (2014) Methylomic profiling of human brain tissue supports a neurodevelopmental origin for schizophrenia. *Genome biology*, **15**, 483.

218. Hannon, E., Spiers, H., Viana, J., Pidsley, R., Burrage, J., Murphy, T.M., Troakes, C., Turecki, G., O'Donovan, M.C., Schalkwyk, L.C. *et al.* (2016) Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. *Nature neuroscience*, **19**, 48-+.

219. Florath, I., Butterbach, K., Heiss, J., Bewerunge-Hudler, M., Zhang, Y., Schottker, B. and Brenner, H. (2016) Type 2 diabetes and leucocyte DNA methylation: an epigenome-wide association study in over 1,500 older adults. *Diabetologia*, **59**, 130-138.

220. Adams, A.T., Kennedy, N.A., Hansen, R., Ventham, N.T., O'Leary, K.R., Drummond, H.E., Noble, C.L., El-Omar, E., Russell, R.K., Wilson, D.C. *et al.* (2014) Two-stage genome-wide methylation profiling in childhood-onset Crohn's Disease implicates epigenetic alterations at the VMP1/MIR21 and HLA loci. *Inflamm Bowel Dis*, **20**, 1784-1793.

221. Devall, M., Roubroeks, J., Mill, J., Weedon, M. and Lunnon, K. (2016) Epigenetic regulation of mitochondrial function in neurodegenerative disease: New insights from advances in genomic technologies. *Neurosci Lett*, **625**, 47-55.

222. Lienhard M, G.C., Morkel M, Herwig R and Chavez L (2014) MEDIPS: genome-wide differential coverage analysis of sequencing data derived from DNA enrichment experiments. *Bioinformatics*, **30**, 284-286.

223. Vazquez, A.I., Bates, D.M., Rosa, G.J., Gianola, D. and Weigel, K.A. (2010) Technical note: an R package for fitting generalized linear mixed

models in animal breeding. *Journal of animal science*, **88**, 497-504.

224. Wright, K. (2016) Package corrgram. R package version 1.9. Available from: https://cran.r- project.org/web/packages/corrgram/index.html.

225. Devall, M., Burrage, J., Caswell, R., Johnson, M., Troakes, C., Al-Sarraj, S., Jeffries, A.R., Mill, J. and Lunnon, K. (2015) A comparison of mitochondrial DNA isolation methods in frozen post- mortem human brain tissue-applications for studies of mitochondrial genetics in brain disorders. *BioTechniques*, **59**, 241-246.

226. Lunnon, K., Ibrahim, Z., Proitsi, P., Lourdusamy, A., Newhouse, S., Sattlecker, M., Furney, S., Saleem, M., Soininen, H., Kloszewska, I. *et al.* (2012) Mitochondrial dysfunction and immune activation are detectable in early Alzheimer's disease blood. *Journal of Alzheimer's disease : JAD*, **30**, 685-710.

227. Tuppen, H.A., Blakely, E.L., Turnbull, D.M. and Taylor, R.W. (2010) Mitochondrial DNA mutations and human disease. *Biochimica et biophysica acta*, **1797**, 113-128.

228. Elliott, H.R., Samuels, D.C., Eden, J.A., Relton, C.L. and Chinnery, P.F. (2008) Pathogenic mitochondrial DNA mutations are common in the general population. *American journal of human genetics*, **83**, 254-260.

229. Kogelnik, A.M., Lott, M.T., Brown, M.D., Navathe, S.B. and Wallace, D.C. (1996) MITOMAP: a human mitochondrial genome database. *Nucleic acids research*, **24**, 177-179.

230. Hazkani-Covo, E., Zeller, R.M. and Martin, W. (2010) Molecular Poltergeists: Mitochondrial DNA Copies (numts) in Sequenced Nuclear Genomes. *Plos Genet*, **6**.

231. Bensasson, D., Zhang, D.X., Hartl, D.L. and Hewitt, G.M. (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends Ecol Evol*, **16**, 314-321.

232. Lunnon, K., Smith, R., Hannon, E.J., De Jager, P.L., Srivastava, G., Volta, M., Troakes, C., Al-Sarraj, S., Burrage, J., Macdonald, R. *et al.* (2014) Cross-tissue methylomic profiling in Alzheimer's disease implicates a role for cortex-specific deregulation of ANK1 in neuropathology. *Nature neuroscience*, **Sept; 17**, 1164-1170.

233. Sims, N.R. and Anderson, M.F. (2008) Isolation of mitochondria from rat brain using Percoll density gradient centrifugation. *Nature protocols*, **3**, 1228-1239.

234. Clayton, D.A. and Shadel, G.S. (2014) Isolation of mitochondria from tissue culture cells. *Cold Spring Harb Protoc*, **2014**, pdb prot080002.

235. Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* , **17**, 10-12.

236. Hornig-Do, H.T., Gunther, G., Bust, M., Lehnartz, P., Bosio, A. and Wiesner, R.J. (2009) Isolation of functional pure mitochondria by superparamagnetic microbeads. *Anal Biochem*, **389**, 1-5.

237. Scarpulla, R.C. (2006) Nuclear control of respiratory gene expression in mammalian cells. *J Cell Biochem*, **97**, 673-683.

238. Horbay, R. and Bilyy, R. (2016) Mitochondrial dynamics during cell cycling. *Apoptosis*, **21**, 1327- 1335.

239. Prudent, J., Popgeorgiev, N., Gadet, R., Deygas, M., Rimokh, R. and Gillet, G. (2016) Mitochondrial Ca2+ uptake controls actin cytoskeleton dynamics during cell migration. *Sci Rep*, **6**, 36570.

240. Maccari, I., Zhao, R., Peglow, M., Schwarz, K., Hornak, I., Pasche, M., Quintana, A., Hoth, M., Qu, B. and Rieger, H. (2016) Cytoskeleton rotation relocates mitochondria to the immunological synapse and increases calcium signals. *Cell Calcium*, **60**, 309-321.

241. Ogura, M., Inoue, T., Yamaki, J., Homma, M.K., Kurosaki, T. and Homma, Y. (2016) Mitochondrial reactive oxygen species suppress humoral immune response through reduction of CD19 expression in B cells in mice. *Eur J Immunol*.

242. De Jager, P.L., Srivastava, G., Lunnon, K., Burgess, J., Schalkwyk, L.C., Yu, L., Eaton, M.L., Keenan, B.T., Ernst, J., McCabe, C. *et al.* (2014) Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. *Nature Neuroscience*, **17**, 1156-1163.

243. Smith, R.G., Hannon, E., De Jager, P.L., Chibnik, L., Lott, S., Condliffe, D., Smith, A.R., Haroutunian, V., Troakes, C., Al-Saraj, S. *et al.* (2016) Elevated DNA methylation across a 48kb region spanning the HOXA gene cluster on chromosome 7 is associated with Alzheimer's disease neuropathology in the prefrontal cortex and superior temporal gyrus. *Nature communications*.

244. Hannon, E., Dempster, E., Viana, J., Burrage, J., Smith, A.R., Macdonald, R., St Clair, D., Mustard, C., Breen, G., Therman, S. *et al.* (2016) An integrated genetic-epigenetic analysis of schizophrenia: evidence for co-localization of genetic associations and differential DNA methylation. *Genome biology*, **17**, 176.

245. Moran, S., Martinez-Cardus, A., Sayols, S., Musulen, E., Balana, C., Estival-Gonzalez, A., Moutinho, C., Heyn, H., Diaz-Lagares, A., de Moura, M.C. *et al.* (2016) Epigenetic profiling to classify cancer of

unknown primary: a multicentre, retrospective analysis. *Lancet Oncol*.

246. Stefansson, O.A., Moran, S., Gomez, A., Sayols, S., Arribas-Jorba, C., Sandoval, J., Hilmarsdottir, H., Olafsdottir, E., Tryggvadottir, L., Jonasson, J.G. *et al.* (2015) A DNA methylation-based definition of biologically distinct breast cancer subtypes. *Mol Oncol*, **9**, 555-568.

247. Swan, E.J., Maxwell, A.P. and McKnight, A.J. (2015) Distinct methylation patterns in genes that affect mitochondrial function are associated with kidney disease in blood-derived DNA from individuals with Type1 diabetes. *Diabetic Med*, **32**, 1110-1115.

248. Gao, J.H., Wen, S.L., Zhou, H.Y. and Feng, S. (2015) De-methylation of displacement loop of mitochondrial DNA is associated with increased mitochondrial copy number and nicotinamide adenine dinucleotide subunit 2 expression in colorectal cancer. *Mol Med Rep*, **12**, 7033-7038.

249. Wong, M., Gertz, B., Chestnut, B.A. and Martin, L.J. (2013) Mitochondrial DNMT3A and DNA methylation in skeletal muscle and CNS of transgenic mouse models of ALS. *Front Cell Neurosci*, **7**.

250. Blanch, M., Mosquera, J.L., Ansoleaga, B., Ferrer, I. and Barrachina, M. (2016) Altered Mitochondrial DNA Methylation Pattern in Alzheimer Disease-Related Pathology and in Parkinson Disease. *Am J Pathol*, **186**, 385-397.

251. Baccarelli, A.A. and Byun, H.M. (2015) Platelet mitochondrial DNA methylation: a potential new marker of cardiovascular disease. *Clinical epigenetics*, **7**, 44.

252. Ghosh S, S.S., Scaria V. (2016) Hydroxymethyl cytosine marks in the human mitochondrial genome are dynamic in nature. *Mitochondrion*,

**27**, 25-31.

253. Krueger, F. Trim Galore!

254. Team, R.C. (2016) R: A Language and Environment for Statistical Computing.

255. Guo, J.U., Su, Y.J., Shin, J.H., Shin, J.H., Li, H.D., Xie, B., Zhong, C., Hu, S.H., Le, T., Fan, G.P. *et al.* (2014) Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nat Neurosci*, **17**, 215-222.

256. Bellizzi, D., D'Aquila, P., Scafone, T., Giordano, M., Riso, V., Riccio, A. and Passarino, G. (2013) The Control Region of Mitochondrial DNA Shows an Unusual CpG and Non-CpG Methylation Pattern. *DNA Res*, **20**, 537-547.

257. McCarthy, N.S., Melton, P.E., Cadby, G., Yazar, S., Franchina, M., Moses, E.K., Mackey, D.A. and Hewitt, A.W. (2014) Meta-analysis of human methylation data for evidence of sex-specific autosomal patterns. *BMC genomics*, **15**, 981.

258. Amigo, I., da Cunha, F.M., Forni, M.F., Garcia-Neto, W., Kakimoto, P.A., Luevano-Martinez, L.A., Macedo, F., Menezes-Filho, S.L., Peloggia, J. and Kowaltowski, A.J. (2016) Mitochondrial form, function and signalling in aging. *The Biochemical journal*, **473**, 3421-3449.

259. Hsiao, C.L., Hsieh, A.R., Lian Ie, B., Lin, Y.C., Wang, H.M. and Fann, C.S. (2014) A novel method for identification and quantification of consistently differentially methylated regions. *PLoS One*, **9**, e97513.

260. Mawlood, S.K., Dennany, L., Watson, N., Dempster, J. and Pickard, B.S. (2016) Quantification of global mitochondrial DNA methylation levels and inverse correlation with age at two CpG sites. *Aging (Albany NY)*.

261. Jung, M. and Pfeifer, G.P. (2015) Aging and DNA methylation. *Bmc Biol*, **13**.

262. Rakyan, V.K., Down, T.A., Maslau, S., Andrew, T., Yang, T.P., Beyan, H., Whittaker, P., McCann, O.T., Finer, S., Valdes, A.M. *et al.* (2010) Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome research*, **20**, 434-439.

263. Trifunovic, A. and Larsson, N.G. (2008) Mitochondrial dysfunction as a cause of ageing. *J Intern Med*, **263**, 167-178.

264. Moreira, P.I., Carvalho, C., Zhu, X.W., Smith, M.A. and Perry, G. (2010) Mitochondrial dysfunction is a trigger of Alzheimer's disease pathophysiology. *Bba-Mol Basis Dis*, **1802**, 2-10.

265. Harris, J.J., Jolivet, R. and Attwell, D. (2012) Synaptic Energy Use and Supply. *Neuron*, **75**, 762- 777.

266. Rhein, V. and Eckert, A. (2007) Effects of Alzheimer's amyloid-beta and tau protein on mitochondrial function -- role of glucose metabolism and insulin signalling. *Arch Physiol Biochem*, **113**, 131-141.

267. Swerdlow, R.H., Burns, J.M. and Khan, S.M. (2014) The Alzheimer's disease mitochondrial cascade hypothesis: progress and perspectives. *Biochim Biophys Acta*, **1842**, 1219-1231.

268. Koch, L. (2016) Genetic variation: Nuclear and mitochondrial genome interplay. *Nat Rev Genet*, **17**, 502.

269. Smith, A.R., Smith, R., Roubroeks, J., Hannon, E., Troakes, C., Al-Sarraj, S., Schalkwyk, L.C., Mill, J., van den Hove, D. and Lunnon, K. (In Prep) DNA methylomic and hydroxymethylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease brain. **[In Preparation]**.

270. Calvo, S.E., Clauser, K.R. and Mootha, V.K. (2016) MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. *Nucleic Acids Res*, **44**, D1251-1257.

271. Guintivano, J., Aryee, M.J. and Kaminsky, Z.A. (2013) A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression. *Epigenetics*, **8**, 290-302.

272. Viechtbauer, W. (2010) Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, **36**.

273. Ugalde, C., Janssen, R.J., van den Heuvel, L.P., Smeitink, J.A. and Nijtmans, L.G. (2004) Differences in assembly or stability of complex I and other mitochondrial OXPHOS complexes in inherited complex I deficiency. *Hum Mol Genet*, **13**, 659-667.

274. Sharma, L.K., Fang, H., Liu, J., Vartak, R., Deng, J. and Bai, Y. (2011) Mitochondrial respiratory complex I dysfunction promotes tumorigenesis through ROS alteration and AKT activation. *Hum Mol Genet*, **20**, 4605-4616.

275. Wiseman, F.K., Al-Janabi, T., Hardy, J., Karmiloff-Smith, A., Nizetic, D., Tybulewicz, V.L., Fisher, E.M. and Strydom, A. (2015) A genetic cause of Alzheimer disease: mechanistic insights from Down syndrome. *Nat Rev Neurosci*, **16**, 564-574.

276. Spilsbury, A., Miwa, S., Attems, J. and Saretzki, G. (2015) The role of telomerase protein TERT in Alzheimer's disease and in tau-related pathology in vitro. *J Neurosci*, **35**, 1659-1674.

277. Chibnik, L.B., Yu, L., Eaton, M.L., Srivastava, G., Schneider, J.A., Kellis, M., Bennett, D.A. and De Jager, P.L. (2015) Alzheimer's loci: epigenetic

associations and interaction with genetic factors. *Ann Clin Transl Neurol*, **2**, 636-647.

278. Watson, C.T., Roussos, P., Garg, P., Ho, D.J., Azam, N., Katsel, P.L., Haroutunian, V. and Sharp, A.J. (2016) Genome-wide DNA methylation profiling in the superior temporal gyrus reveals epigenetic signatures associated with Alzheimer's disease. *Genome Med*, **8**, 5.

279. Smith, A.R., Smith, R., Roubroeks, J., Hannon, E., Troakes, C., Al-Sarraj, S., Schalkwyk, L.C., Mill, J., van den Hove, D. and Lunnon, K. (In Prep) DNA methylomic and hydroxymethylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease brain. *In Preparation*.

280. Warnecke, T., Duning, T., Schirmacher, A., Mohammadi, S., Schwindt, W., Lohmann, H., Dziewas, R., Deppe, M., Ringelstein, E.B. and Young, P. (2010) A novel splice site mutation in the SPG7 gene causing widespread fiber damage in homozygous and heterozygous subjects. *Mov Disord*, **25**, 413-420.

281. Kruger, S., Battke, F., Sprecher, A., Munz, M., Synofzik, M., Schols, L., Gasser, T., Grehl, T., Prudlo, J. and Biskup, S. (2016) Rare Variants in Neurodegeneration Associated Genes Revealed by Targeted Panel Sequencing in a German ALS Cohort. *Front Mol Neurosci*, **9**, 92.

282. Choi, T.M., Yun, M., Lee, J.K., Park, J.T., Park, M.S. and Kim, H.S. (2016) Proteomic Analysis of a Rat Cerebral Ischemic Injury Model after Human Cerebral Endothelial Cell Transplantation. *J Korean Neurosurg Soc*, **59**, 544-550.

283. Fu, F., Wu, D. and Qian, C. (2016) The MicroRNA-224 Inhibitor Prevents Neuronal Apoptosis via Targeting Spastic Paraplegia 7 After Cerebral

Ischemia. *J Mol Neurosci*, **59**, 421-429.

284. Qi, H. and Shuai, J. (2016) Alzheimer's disease via enhanced calcium signaling caused by the decrease of endoplasmic reticulum-mitochondrial distance. *Med Hypotheses*, **89**, 28-31.

285. Ding, J., Sidore, C., Butler, T.J., Wing, M.K., Qian, Y., Meirelles, O., Busonero, F., Tsoi, L.C., Maschio, A., Angius, A. *et al.* (2015) Assessing Mitochondrial DNA Variation and Copy Number in Lymphocytes of ~2,000 Sardinians Using Tailored Sequencing Analysis Tools. *PLoS Genet*, **11**, e1005306.

286. Larman, T.C., DePalma, S.R., Hadjipanayis, A.G., Cancer Genome Atlas Research, N., Protopopov, A., Zhang, J., Gabriel, S.B., Chin, L., Seidman, C.E., Kucherlapati, R. *et al.* (2012) Spectrum of somatic mitochondrial mutations in five cancers. *Proc Natl Acad Sci U S A*, **109**, 14087-14091.

287. Goto, H., Dickins, B., Afgan, E., Paul, I.M., Taylor, J., Makova, K.D. and Nekrutenko, A. (2011) Dynamics of mitochondrial heteroplasmy in three families investigated via a repeatable re- sequencing study. *Genome Biol*, **12**, R59.

288. Guo, Y., Li, J., Li, C.I., Shyr, Y. and Samuels, D.C. (2013) MitoSeek: extracting mitochondria information and performing high-throughput mitochondria sequencing analysis. *Bioinformatics*, **29**, 1210-1211.

289. Zhang, P., Samuels, D.C., Lehmann, B., Stricker, T., Pietenpol, J., Shyr, Y. and Guo, Y. (2016) Mitochondria sequence mapping strategies and practicability of mitochondria variant detection from exome and RNA sequencing data. *Brief Bioinform*, **17**, 224-232.

290. Ye, F., Samuels, D.C., Clark, T. and Guo, Y. (2014) High-throughput

sequencing in mitochondrial DNA research. *Mitochondrion*, **17**, 157-163.

291. Azimzadeh, P., Asadzadeh Aghdaei, H., Tarban, P., Akhondi, M.M., Shirazi, A. and Khorram Khorshid, H.R. (2016) Comparison of three methods for mitochondria isolation from the human liver cell line (HepG2). *Gastroenterol Hepatol Bed Bench*, **9**, 105-113.

292. Quispe-Tintaya, W., White, R.R., Popov, V.N., Vijg, J. and Maslov, A.Y. (2013) Fast mitochondrial DNA isolation from mammalian cells for next-generation sequencing. *Biotechniques*, **55**, 133- 136.

293. Hornig-Do, H.T., Gunther, G., Bust, M., Lehnartz, P., Bosio, A. and Wiesner, R.J. (2009) Isolation of functional pure mitochondria by superparamagnetic microbeads. *Anal Biochem*, **389**, 1-5.

294. Fuso, A., Nicolia, V., Pasqualato, A., Fiorenza, M.T., Cavallaro, R.A. and Scarpa, S. (2011) Changes in Presenilin 1 gene methylation pattern in diet-induced B vitamin deficiency. *Neurobiol Aging*, **32**, 187-199.

295. Fuso, A., Ferraguti, G., Grandoni, F., Ruggeri, R., Scarpa, S., Strom, R. and Lucarelli, M. (2010) Early demethylation of non-CpG, CpC-rich, elements in the myogenin 5 '-flanking region A priming effect on the spreading of active demethylation? *Cell Cycle*, **9**, 3965-3976.

296. Nishino, K., Hattori, N., Sato, S., Arai, Y., Tanaka, S., Nagy, A. and Shiota, K. (2011) Non-CpG Methylation Occurs in the Regulatory Region of the Sry Gene. *J Reprod Develop*, **57**, 586-593.

297. Vijay, V., Han, T., Moland, C.L., Kwekel, J.C., Fuscoe, J.C. and Desai, V.G. (2015) Sexual dimorphism in the expression of mitochondria-related genes in rat heart at different ages. *PLoS One*, **10**, e0117047.

298. Sharma, J., Johnston, M.V. and Hossain, M.A. (2014) Sex differences in

mitochondrial biogenesis determine neuronal death and survival in response to oxygen glucose deprivation and reoxygenation. *BMC Neurosci*, **15**, 9.

299. Mitchell, P. (2011) Chemiosmotic coupling in oxidative and photosynthetic phosphorylation. 1966. *Biochim Biophys Acta*, **1807**, 1507-1538.

300. Wilson, V.L., Smith, R.A., Ma, S. and Cutler, R.G. (1987) Genomic 5-methyldeoxycytidine decreases with age. *J Biol Chem*, **262**, 9948-9951.

301. Romanov, G.A. and Vanyushin, B.F. (1981) Methylation of reiterated sequences in mammalian DNAs. Effects of the tissue type, age, malignancy and hormonal induction. *Biochim Biophys Acta*, **653**, 204-218.

302. Wilson, V.L. and Jones, P.A. (1983) DNA methylation decreases in aging but not in immortal cells. *Science*, **220**, 1055-1057.

303. Fairweather, D.S., Fox, M. and Margison, G.P. (1987) The in vitro lifespan of MRC-5 cells is shortened by 5-azacytidine-induced demethylation. *Exp Cell Res*, **168**, 153-159.

304. Florath, I., Butterbach, K., Muller, H., Bewerunge-Hudler, M. and Brenner, H. (2014) Cross- sectional and longitudinal changes in DNA methylation with age: an epigenome-wide analysis revealing over 60 novel age-associated CpG sites. *Hum Mol Genet*, **23**, 1186-1201.

305. Horvath, S. (2013) DNA methylation age of human tissues and cell types. *Genome Biol*, **14**, R115.

306. Jemt, E., Persson, O., Shi, Y.H., Mehmedovic, M., Uhler, J.P., Lopez, M.D., Freyer, C., Gustafsson, C.M., Samuelsson, T. and Falkenberg, M. (2015) Regulation of DNA replication at the end of the mitochondrial D-loop involves the helicase TWINKLE and a conserved sequence element. *Nucleic Acids*

*Research*, **43**, 9262-9275.

307. Cantatore, P. and Attardi, G. (1980) Mapping of Nascent Light and Heavy Strand Transcripts on the Physical Map of Hela-Cell Mitochondrial-DNA. *Nucleic Acids Research*, **8**, 2605-2625.

308. Montoya, J., Christianson, T., Levens, D., Rabinowitz, M. and Attardi, G. (1982) Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial DNA. *Proc Natl Acad Sci U S A*, **79**, 7195-7199.

309. Mitrofanov, K.Y., Zhelankin, A.V., Shiganova, G.M., Sazonova, M.A., Bobryshev, Y.V., Postnov, A.Y., Sobenin I capital A, C.I.A. and Orekhov, A.N. (2016) Analysis of mitochondrial DNA heteroplasmic mutations A1555G, C3256T, T3336C, capital ES, Cyrillic5178capital A, Cyrillic, G12315A, G13513A, G14459A, G14846capital A, Cyrillic and G15059A in CHD patients with the history of myocardial infarction. *Exp Mol Pathol*, **100**, 87-91.

310. Synowiec, E., Hoser, G., Bialkowska-Warzecha, J., Pawlowska, E., Skorski, T. and Blasiak, J. (2015) Doxorubicin Differentially Induces Apoptosis, Expression of Mitochondrial Apoptosis-Related genes, and Mitochondrial Potential in BCR-ABL1-Expressing Cells Sensitive and Resistant to Imatinib. *Biomed Res Int*, **2015**, 673512.

311. Hagen, C.M., Aidt, F.H., Havndrup, O., Hedley, P.L., Jensen, M.K., Kanters, J.K., Pham, T.T., Bundgaard, H. and Christiansen, M. (2015) Private mitochondrial DNA variants in danish patients with hypertrophic cardiomyopathy. *PLoS One*, **10**, e0124540.

312. Swerdlow, R.H. (2012) Mitochondria and cell bioenergetics: increasingly

recognized components and a possible etiologic cause of Alzheimer's disease. *Antioxid Redox Signal*, **16**, 1434-1455.

313. Choquet, K., Tetreault, M., Yang, S., La Piana, R., Dicaire, M.J., Vanstone, M.R., Mathieu, J., Bouchard, J.P., Rioux, M.F., Rouleau, G.A. *et al.* (2016) SPG7 mutations explain a significant proportion of French Canadian spastic ataxia cases. *Eur J Hum Genet*, **24**, 1016-1021.

314. Monique G. P. van der Wijst, A.Y.v.T., Marcel H. J. Ruiters & Marianne G. Rots. (2017) Experimental mitochondria-targeted DNA methylation identifies GpC methylation, not CpG methylation, as potential regulator of mitochondrial gene expression. *Scientific Reports*.

315. Carelli, V., Maresca, A., Caporali, L., Trifunov, S., Zanna, C. and Rugolo, M. (2015) Mitochondria: Biogenesis and mitophagy balance in segregation and clonal expansion of mitochondrial DNA mutations. *Int J Biochem Cell Biol*, **63**, 21-24.

316. Payne, B.A., Cree, L. and Chinnery, P.F. (2015) Single-cell analysis of mitochondrial DNA.*Methods Mol Biol*, **1264**, 67-76.

317. Jayaprakash, A.D., Benson, E.K., Gone, S., Liang, R., Shim, J., Lambertini, L., Toloue, M.M., Wigler, M., Aaronson, S.A. and Sachidanandam, R. (2015) Stable heteroplasmy at the single-cell level is facilitated by intercellular exchange of mtDNA. *Nucleic Acids Research*, **43**, 2177-2187.

318. Swerdlow, R.H., Koppel, S., Weidling, I., Hayley, C., Ji, Y. and Wilkins, H.M. (2017) Mitochondria, Cybrids, Aging, and Alzheimer's Disease. *Prog Mol Biol Transl Sci*, **146**, 259-302.

319. Boominathan, A., Vanhoozer, S., Basisty, N., Powers, K., Crampton, A.L., Wang, X., Friedricks, N., Schilling, B., Brand, M.D. and O'Connor, M.S. (2016)

Stable nuclear expression of ATP8 and ATP6 genes rescues a mtDNA Complex V null mutant. *Nucleic Acids Res*, **44**, 9342-9357.

320. Yang, S.M., Li, S.Y., Yu, H.B., Li, J.R. and Sun, L.L. (2017) Repression of DOK7 mediated by DNMT3A promotes the proliferation and invasion of KYSE410 and TE-12 ESCC cells. *Biomed Pharmacother*, **90**, 93-99.

321. Connelly, J.J., Cherepanova, O.A., Doss, J.F., Karaoli, T., Lillard, T.S., Markunas, C.A., Nelson, S., Wang, T., Ellis, P.D., Langford, C.F. *et al.* (2013) Epigenetic regulation of COL15A1 in smooth muscle cell replicative aging and atherosclerosis. *Hum Mol Genet*, **22**, 5107-5120.

322. Teoh-Fitzgerald, M.L., Fitzgerald, M.P., Jensen, T.J., Futscher, B.W. and Domann, F.E. (2012) Genetic and epigenetic inactivation of extracellular superoxide dismutase promotes an invasive phenotype in human lung cancer by disrupting ECM homeostasis. *Mol Cancer Res*, **10**, 40-51.

323. Luo, J., Li, Y.N., Wang, F., Zhang, W.M. and Geng, X. (2010) S-adenosylmethionine inhibits the growth of cancer cells by reversing the hypomethylation status of c-myc and H-ras in human gastric cancer and colon cancer. *Int J Biol Sci*, **6**, 784-795.

324. Thakore, P.I., D'Ippolito, A.M., Song, L.Y., Safi, A., Shivakumar, N.K., Kabadi, A.M., Reddy, T.E., Crawford, G.E. and Gersbach, C.A. (2015) Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat Methods*, **12**, 1143-+.

325. Lu, Q., Livi, G.P., Modha, S., Yusa, K., Macarron, R. and Dow, D.J. (2017) Applications of CRISPR Genome Editing Technology in Drug Target Identification and Validation. *Expert Opin Drug Discov*.

326. Jo, A., Ham, S., Lee, G.H., Lee, Y.I., Kim, S., Lee, Y.S., Shin, J.H. and

Lee, Y. (2015) Efficient Mitochondrial Genome Editing by CRISPR/Cas9. *Biomed Research International*.

327. Flanagan, J.M. (2015) Epigenome-wide association studies (EWAS): past, present, and future.*Methods Mol Biol*, **1238**, 51-63.