



MONASH University

THE EPIGENETIC CONTROL OF MITOCHONDRIAL DNA COPY NUMBER IN TUMORIGENESIS AND DIFFERENTIATION

XIN SUN

MSc

Doctor of Philosophy

A thesis submitted for the degree of *Doctor of Philosophy* at

Monash University in 2018

Department of Molecular and Translational Sciences

Faculty of Medicine, Nursing and Health Sciences

Copyright notice

© Xin Sun (2018).

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.

Acknowledgements

Foremost, I would like to express my sincere gratitude to my supervisor and mentor, Professor Justin St. John for giving me the opportunity to work on this truly inspiring project in the Mitochondrial Genetics Group, for his motivation and patience in guiding me throughout my candidature, as well as for his supervision in training me into a professional and independent researcher.

Special thanks to my co-supervisor Dr. Matthew McKenzie, who always has an open door to give me advice and support. I would also like to recognise my friends and fellow lab mates, who understand the hardness of research and are always there willing to give me a hand. Tesha, for your empathy and advise for my problems not only in work but also for general life. Nok, for your hilarious and positive personality that cheered me up when I was faced with tough situations. Lexie, for being my cold noodle buddy and inspiring me to be a strong female researcher. Rukmali, for being a good friend, who I met from work and came for any random help. Jacqui, for keeping our laboratory an organized and lovely second home. Vijesh, William, Suzan, Chern, Jessica and Kirsten, for their words of wisdom as senior PhD students and postdocs. There were other nice people I have met along the 4 years of my PhD. I appreciate all the kind help.

I would also like to acknowledge the financial and technical support I have received during my PhD. This research was supported by an Australian Government Research Training Program (RTP) Scholarship. I would like to thank Monash University for funding me to present my work at conferences nationally and internationally. I acknowledge the facilities of Hudson Institute of Medical Research

and Monash University, particular thanks to MHTP Medical Genomics Facility and Monash Bioinformatics Platform for the professional assistance.

Last, but by no means least, I would like to thank my husband, Scott, for your love and support to calm down my stress and (very rare) unstable emotions and keep me carrying on. I am so indebted to my little ones, Vincent and Chelsea. Maybe you have heard people said how amazing to do a PhD with two young children, but it is you that always encourage Mum never giving up, and truly you have made my PhD way happier. Heartfelt thanks to my family and friends, particularly my parents Yumin and Jikai, and my parents-in-law Chuanyu and Changfu, for the selfless and endless support in coming over all the way to Australia and giving the little ones the best care. Thank you all for always standing by me through thick and thin. I am forever grateful.

Abstract

The mitochondrion with its own genome, mitochondrial DNA (mtDNA), is critical for generating sufficient cellular energy for development, regulating apoptosis and regulating DNA methylation profiles. During embryonic development, mtDNA replication is strictly down-regulated to establish the mtDNA set point, and then up-regulated in a tissue-specific manner to meet the mtDNA copy number required by each cell type. Tumour cells maintain low mtDNA copy number to support tumorigenesis and, consequently, fail to maintain the mtDNA set point that is required to promote differentiation. The regulation of mtDNA copy number during tumorigenesis and differentiation is driven by a group of nuclear-encoded mtDNA-specific replication factors, such as the mtDNA-specific polymerase gamma with its catalytic subunit A (POLG) and supporting subunits B (POLG2), the mtDNA-specific topoisomerase TOP1MT, the single-strand stabilizing enzyme SSBP1, and the helicase TWNK. These replication factors are tightly associated with mitochondrial function and their aberrant gene expression profiles have been reported in different types of cancers. Previous work has shown that almost complete depletion of mtDNA in tumour-initiating cells caused significantly late onset of tumorigenesis and tumour formation rate. This reflects significant changes to overall gene transcriptional profiles, especially to the developmental networks, and DNA methylation profile of the key regulatory region of *POLG*, namely its second exon. Moreover, recent findings suggest that mtDNA methylation takes place but with undetermined functions, which may be important for regulating mtDNA replication and transcription.

My PhD aimed to further investigate the epigenetic control on mtDNA copy number in tumorigenesis and differentiation. Firstly, by using advanced whole-genome next-

generation sequencing, my work investigated changes to the methylome and transcriptome of the HSR-GBM1 glioblastoma cell line induced by two DNA demethylation treatments, namely 5-Azacytidine and Vitamin C. This enabled identification of the key mtDNA replication factors that are DNA methylated and influence mtDNA copy number through global changes to DNA demethylation patterns. These findings confirmed the role of DNA methylation in regulating mtDNA copy number in glioblastoma cells.

Using an *in vivo* tumour model of glioblastoma derived from cells with various levels of mtDNA content, I identified aberrant DNA methylation patterns of the mtDNA replication factors and tumorigenic regulators that result from the restoration of mtDNA copy number and tumorigenic capability. These findings indicate that mtDNA communicates with the nuclear genome by promoting the modulation of DNA methylation in tumour initiating cells.

Furthermore, I have characterized the degree of mtDNA methylation resulting from 1) the presence of mtDNA from different cell types under the same chromosomal background; 2) different stages of tumour progression; 3) different nuclear genome backgrounds using glioblastoma and osteosarcoma models; and 4) different levels of mtDNA depletion. The outcomes highlight the bi-directional control between the nuclear and mitochondrial genomes through DNA methylation.

My PhD thesis combines multi-disciplinary data from the mitochondrial genome, epigenetics and tumorigenesis, which could provide novel insights into the epigenetic control of mtDNA copy number for further research, especially for cancers and developmental disorders.

Publications during enrolment

1. **Sun X**, Johnson J, St. John JC. Global DNA methylation synergistically regulates the nuclear and mitochondrial genomes. *Nucleic Acids Research*. 2018; 46:5977-5995. DOI: 10.1093/nar/gky339.
2. St. John JC, Srirattana K, Tsai T, **Sun X**. The mitochondrial genome: how it drives fertility. *Reproduction, Fertility and Development*. 2018; 30:118–139. DOI:10.1071/RD17408.
3. Lee W, **Sun X**, Tsai T, Johnson J, Gould AJ, Garama DJ, Gough DJ, McKenzie M, Trounce I, St. John JC. Mitochondrial DNA haplotypes induce differential patterns of DNA methylation that result in differential chromosomal gene expression patterns. *Cell Death Discovery*. 2017 11; 3:17062. DOI:10.1038/cddiscovery.2017.62. **Co-first author**.
4. **Sun X**, St. John JC. The role of the mtDNA set point in differentiation, development and tumorigenesis. *Biochemical Journal*. 2016; 473(19):2955-71. DOI: 10.1042/BCJ20160008.
5. **Sun X**, Lee W, Vaghjiani V, St. John JC. Analysis of mitochondrial DNA copy number and its regulation through DNA methylation of POLGA. *Methods in Molecular Biology*. 2016;1351:131-41. DOI: 10.1007/978-1-4939-2848-4_18.
6. Cagnone G, Vaghjiani V, Lee W, **Sun X**, Johnson J, Yeung KY, St John JC. Analysis of the Mitochondrial DNA and Its Replicative Capacity in Induced Pluripotent Stem Cells. *Methods in Molecular Biology*. 2016;1357:231-67. DOI: 10.1007/7651_2014_156

Thesis including published works declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes three original papers published in peer reviewed journals and two submitted publications. The core theme of the thesis is the epigenetic control of mitochondrial DNA copy number in tumorigenesis and differentiation. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the student, working within the Hudson Institute of Medical Research (School of Clinical Sciences) under the supervision of *Professor Justin St. John*.

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

In the case of the five chapters, my contribution to the work involved the following:

Thesis Chapter	Publication Title	Status	Nature and % of student contribution	Co-author name(s) Nature and % of Co-author's contribution*	Co-author(s), Monash student Y/N*
1	The role of the mtDNA set point in differentiation, development and tumorigenesis.	Published	80%, wrote first draft	20%, Justin St. John input into final manuscript	No
2	Analysis of mitochondrial DNA copy number and its regulation through DNA methylation of POLGA.	Published	60%, drafted the manuscript	5%, William Lee input into final manuscript; 5%, Vijesh Vaghjiani input into final manuscript; 30%, Justin St. John conceived the work, co-ordinated the work, input into final manuscript.	No
3	Global DNA methylation synergistically regulates the nuclear and mitochondrial genomes.	Published	60%, participated in the study design, performed the molecular experiments, performed the computational analysis and wrote	5%, Jacqueline Johnson performed tissue culture; 35%, Justin St. John conceived the work, designed and	No

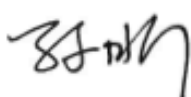
			the first draft of the manuscript.	coordinated the experiments, input into final manuscript.	
4	Modulation of mitochondrial DNA copy number during tumorigenesis induces changes in DNA methylation of the nuclear genome.	Returned for revision	60%, participated in the study design, performed the molecular experiments, performed the computational analysis and wrote the first draft of the manuscript.	40%, Justin St. John conceived the work, designed and coordinated the experiments, input into final manuscript.	No
5	The degree of mitochondrial DNA methylation in tumorigenesis	Submitted	55%, participated in the study design, performed the molecular experiments, performed the computational analysis and wrote the first draft of the manuscript.	10%, Vijesh Vaghjiani, W Samantha N Jayasekara and Jason E. Cain performed the animal works; 35%, Justin St. John conceived the work, designed and coordinated the experiments, input into final manuscript	No

**If no co-authors, leave fields blank*

I have not renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Student signature:

Date: 19th July 2018



The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor signature:

Date: 19th July 2018



Table of Contents

CHAPTER 1	Introduction: The role of the mtDNA set point in differentiation, development and tumorigenesis.	P1
	Overall Hypothesis and Aims	P20
CHAPTER 2	General Methods: Analysis of mitochondrial DNA copy number and its regulation through DNA methylation of POLGA.	P23
CHAPTER 3	Global DNA methylation synergistically regulates the nuclear and mitochondrial genomes in glioblastoma cells.	P37
CHAPTER 4	Rationale leading to Chapter 4	P79
	Modulation of mitochondrial DNA copy number during tumorigenesis induces changes in DNA methylation of the nuclear genome.	P81
CHAPTER 5	Rationale leading to Chapter 5	P135
	The degree of mitochondrial DNA methylation in tumorigenesis.	P137
CHAPTER 6	General Discussion	P205
APPENDIX A	Mitochondrial DNA haplotypes induce differential patterns of DNA methylation that result in differential chromosomal gene expression patterns.	P219
APPENDIX B	The mitochondrial genome: how it drives fertility.	P231
APPENDIX C	Analysis of the Mitochondrial DNA and Its Replicative Capacity in Induced Pluripotent Stem Cells.	P255

**Page numbers of this thesis are indicated at the center footer of each page.*

CHAPTER 1

Introduction

Review article attached.

Review Article

The role of the mtDNA set point in differentiation, development and tumorigenesis

Xin Sun^{1,2} and Justin C. St. John^{1,2}

¹Centre for Genetic Diseases, Hudson Institute of Medical Research, 27–31 Wright Street, Clayton, VIC 3168, Australia and ²Department of Molecular and Translational Sciences, Monash University, 27–31 Wright Street, Clayton, VIC 3168, Australia

Correspondence: Justin St John (justin.stjohn@hudson.org.au)

Mitochondrial DNA replication is critical for maintaining mtDNA copy number to generate sufficient cellular energy that is required for development and for functional cells. In early development, mtDNA copy number is strictly regulated at different stages, and, as a result, the establishment of the mtDNA set point is required for sequential cell lineage commitment. The failure to establish the mtDNA set point results in incomplete differentiation or embryonic arrest. The regulation of mtDNA copy number during differentiation is closely associated with cellular gene expression, especially with the pluripotency network, and DNA methylation profiles. The findings from cancer research highlight the relationship between mitochondrial function, mtDNA copy number and DNA methylation in regulating differentiation. DNA methylation at exon 2 of DNA polymerase gamma subunit A (POLGA) has been shown to be a key factor, which can be modulated to change the mtDNA copy number and cell fate of differentiating and tumour cells. The present review combines multi-disciplinary data from mitochondria, development, epigenetics and tumorigenesis, which could provide novel insights for further research, especially for developmental disorders and cancers.

Introduction

Mitochondria are double-membraned organelles primarily responsible for the generation of ATP as well as several other vital cellular functions. Each mitochondrion hosts multiple copies of the mitochondrial genome (mtDNA), which is a circular double-stranded genome encoding key subunits of the electron transfer chain (ETC). There are programmed changes to mtDNA copy number during cellular development and tumorigenesis, which are directly regulated by a group of nuclear-encoded regulators. It has also been demonstrated that changes to mtDNA copy number are associated with the expression of key regulators of cell fate and the DNA methylation status at exon 2 of the catalytic subunit of DNA polymerase gamma (*POLGA*). These processes are strictly regulated to establish the mtDNA set point, which provides differentiating cells with the potential to increase mtDNA copy number in a cell-specific manner.

Mitochondria and OXPHOS

The mitochondrion is a double-membraned organelle present in nearly all mammalian cells. Mitochondria are associated with a variety of cellular and metabolic activities such as cell signalling, ion homeostasis, cell differentiation and apoptosis [1,2]. They are primarily described as the powerhouses of cellular metabolism because they are the major generators of cellular energy (ATP) through the aerobic process of oxidative phosphorylation (OXPHOS) [3,4] (Figure 1).

The ETC is located on the inner membrane of the mitochondrion. It comprises five enzyme complexes (Figure 1): Complex I (NADH dehydrogenase; ND), Complex II (succinate dehydrogenase; SD), Complex III (cytochrome *c* reductase; CYTB), Complex IV (cytochrome *c* oxidase; COX) and Complex V (ATP synthase). Additionally, coenzyme Q (CoQ) and cytochrome *c* (CytC) are two

Received: 22 January 2016
Revised: 17 May 2016
Accepted: 6 June 2016

Version of Record published:
27 September 2016

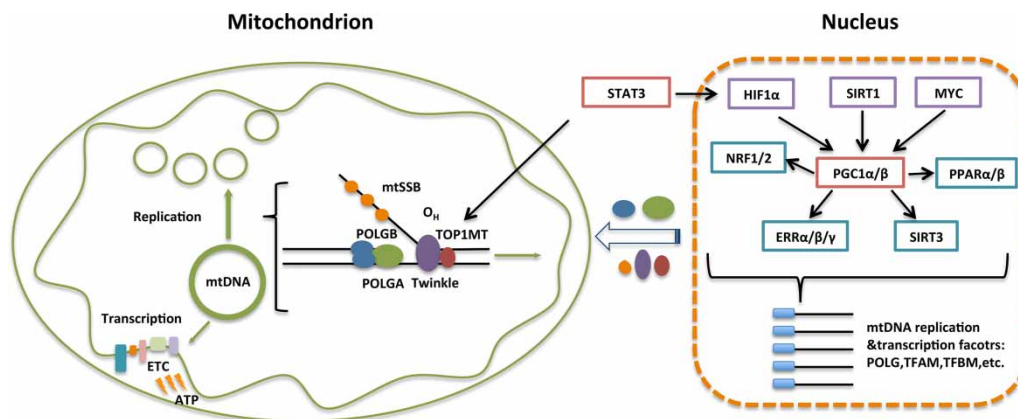


Figure 1. mtDNA interactions.

mtDNA replication and transcription take place inside the mitochondria. mtDNA replication is dependent on nDNA-encoded replication factors including the mitochondrial single-stranded-binding protein (mtSSB), Twinkle, POLGA, POLG accessory subunit (POLGB) and DNA topoisomerase I mitochondrial (TOP1MT). They translocate into the mitochondria. A group of upstream regulators including STAT3, HIF1 α , SIRT1, MYC, PGC1 α/β , NRF1/2, ERR $\alpha/\beta/\gamma$, SIRT3 and PPAR α/β , participate in regulating gene expression of the direct regulators of mtDNA transcription and replication. The ETC consists of five complexes, which contribute to the generation of ATP through OXPHOS.

electron carriers that are extensively involved [3]. Briefly, Complex I and Complex II accept electrons from the electron donors and pass them on to CoQ. The electrons are then transferred through Complex III, CytC and Complex IV and finally accepted by O₂. During this process, a pH gradient across the inner membrane forms due to the protons being pumped out by the complexes. Potential energy stored in the pH gradient is then transformed into biochemical energy, in the form of ATP, through the process of transferring the protons back into the mitochondrial matrix through Complex V [3]. It is noted that the process described above constitutes aerobic respiration. This is a far more efficient process than anaerobic respiration, as it produces 32–36 molecules of ATP per molecule of glucose compared with two molecules of ATP from glycolysis [4].

The mitochondrial genome

The mitochondrion hosts a genome, mtDNA, whose replication and expression apparatuses are distinct from those of the nuclear genome. mtDNA is a circular double-stranded DNA molecule that is 16.6 kb in size in humans (Figure 2). It consists of a heavy (H) and a light (L) strand. The mitochondrial genome encodes 37 genes which include 22 tRNAs, two rRNAs (12S and 16S rRNA) and 13 polypeptides of the ETC complex enzymes [5] (Figure 2). These include genes encoding seven subunits of Complex I (ND1–ND6 and ND4L), one Complex III subunit (CYTB), three Complex IV subunits (COX I, II and III) and two Complex V subunits (ATPase6 and ATPase8) (Figure 2). As for the other subunits of ETC complexes, they are nuclear DNA (nDNA)-encoded proteins that are imported into the mitochondrial matrix. Therefore, efficient and functional OXPHOS requires a well-established synergy between the nucleus and the mitochondria to meet the demands of a specialized cell for the production of energy.

The mitochondrial genome is highly compact with only two non-coding regions and no introns. The main non-coding region contains the displacement loop region (D-loop). It also consists of the origin of replication for the H strand (O_H), and the promoters of transcription for both the H and L strands (HSP1, HSP2 and LSP) (Figure 2). The origin of replication for the L strand is located in the other non-coding region that is located two-thirds of the genome downstream of O_H, between the genes tRNA-Cys and tRNA-Asn [5]. Genes encoding tRNAs are interspersed between genes encoding rRNAs and the ETC subunits, especially on the L strand [6]. Also, the termination codons are not encoded in some of the mitochondrial genes, but are created post-transcriptionally by the poly(A) tail mechanism [6], which means that the initial transcript normally ends with either a U or a UA, and polyadenylation can add adenines after these transcripts to form a UAA termination codon.

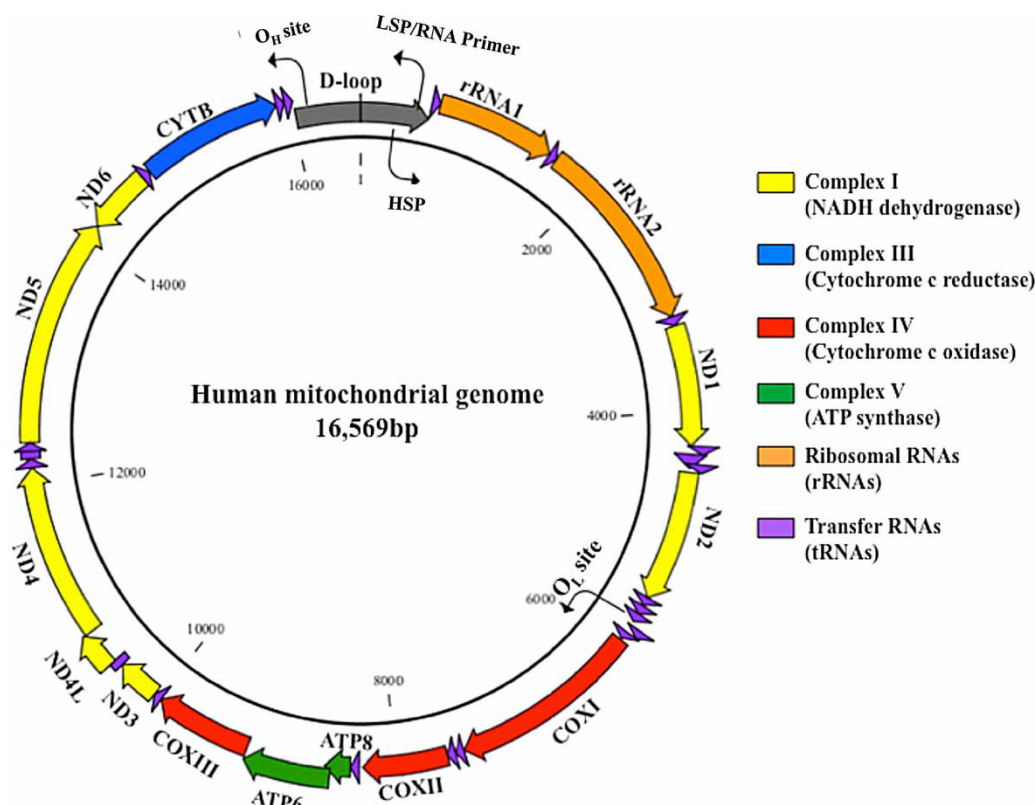


Figure 2. The human mitochondrial genome.

The human mitochondrial genome is a double-stranded circular genome, which is ~16.6 kb in size. It encodes 22 tRNAs, two rRNAs and 13 subunits of the ETC complexes. The 13 subunits are seven subunits of Complex I (NADH dehydrogenase, ND1–ND6 and ND4L), one Complex III subunit (cytochrome c reductase, CYTB), three Complex IV subunits (cytochrome c oxidase, COX I, II and III) and two Complex V subunits (ATPase6 and ATPase8). The main non-coding region contains the D-loop region, the origin replication for the H strand (O_H) and the promoters of transcription for the H and L strands (HSP and LSP). The origin of replication for the L strand (O_L) is located two-thirds of the genome downstream from O_H, between the tRNA-Cys and tRNA-Asn genes.

Another key feature that distinguishes the mitochondrial genome from the nuclear genome is that there are multiple copies of the mitochondrial genome per mitochondrion, and also per cell, as each cell also possesses multiple mitochondria [7]. Indeed, different cell types have different levels of mtDNA copy number depending on their demands for cellular energy [8]. mtDNA copy number could be as high as several thousand in cells such as muscle and neurons, whereas it is much lower in fibroblasts, spleen cells and stem cells [8]. Moreover, the mitochondrial genome can be present in a heteroplasmic manner, which means that both wild-type and mutated forms can coexist in the same cell [9]. It is thought that mtDNA mutations are a constant phenomenon mostly due to the primary targeting of reactive oxygen species (ROS). This is despite the mtDNA-specific replicase DNA polymerase gamma (POLG) having one of the lowest error rates of the eukaryotic polymerases for its exonucleolytic proofreading functions [10]. There is also the presence of diverse repair systems in the mitochondrion including recombinational repair [10]. Nevertheless, the mutations remain mostly silent or have no effect on phenotype, as it has been demonstrated that a mutation level of over 70% is normally required to lead to the onset of one of the many mitochondrial diseases [11].

In mitochondria, mtDNA molecules are organized into DNA–protein complexes similar to nucleoid structures in prokaryotes which supports the theory that the mitochondrion evolved from a proteobacterial partner that invaded eukaryotes [12]. The mitochondrial nucleoid consists of more than 20 binding proteins, which comprise many of the factors responsible for the transcription and replication of mtDNA. Owing to the limited coding capacity of mtDNA, these factors are nuclear-encoded and imported into the mitochondrial matrix.

mtDNA copy number is strictly controlled by its transcription and replication, which take place in mitochondria. mtDNA transcription and replication are driven by a group of nDNA-encoded mitochondrial specific factors (Figure 1), which includes a group of factors that bind directly with mtDNA, such as POLG and mitochondrial transcription factor A (TFAM), and a group of indirect factors, which regulate the gene expression of those direct factors. The best characterized of the indirect factors are the nuclear respiratory factors (NRF1 and NRF2) and the peroxisome-proliferator-activated receptor γ (PPAR γ) co-activator 1 α (PGC1 α) family (Figure 1). It is noted that the regulation of mtDNA transcription and replication requires the coordinated control of all the factors and is at a complex level. As the related data have been well summarized and reviewed by others [13–18], we are not covering this topic here. However, it is important to point out that the regulation of mtDNA copy number involves numerous factors, which have cell- and developmental-stage-specific expression profiles and are associated with other cellular functions.

The mtDNA set point

In order to meet the very different energy demands of a variety of cell types, mtDNA copy number is strictly regulated during development. This is, perhaps, best exemplified in embryonic stem cells (ESCs), which acquire mtDNA copy number in a cell-specific manner during differentiation through the strict regulation of mtDNA replication. This equips mature cells with the appropriate number of mtDNA copies in order that they are able to perform their specialized functions [19–21] (Figure 3).

Very early on in development, the primordial germ cells, which are the first indicative germ cells, are laid down and they possess approximately 200 copies of mtDNA per cell [22,23]. As these cells mature during the extensive process of oogenesis to form fertilizable oocytes, namely metaphase II oocytes, mtDNA copy number increases significantly to more than 180,000 copies per cell. Indeed, there is a ‘critical threshold’ of mtDNA copy number to enable the metaphase II oocyte to proceed through fertilization [24–27] (Figure 3). Failure to surpass this threshold results in either unsuccessful fertilization or embryos arresting during pre-implantation development. However, those oocytes that mature with insufficient mtDNA copy number, namely mtDNA-deficient oocytes, can be rescued through supplementation with exogenous populations of mtDNA [28]. This improves fertilization rates and successful development to the blastocyst stage [29], the final stage of pre-implantation development when the embryo first differentiates into two cellular compartments [24,28]. These cellular compartments comprise the inner cell mass (ICM), which gives rise to the embryo proper and then the fetus, and the trophectoderm, which forms the placenta. Indeed, mitochondrial supplementation is being introduced clinically to promote embryo development, especially in older patients when their embryos either fail to fertilize or they arrest very early during development [30].

Once fertilization has taken place, mtDNA copy number peaks between fertilization and activation of the embryonic genome [25,31,32], a transcriptional event that enables the embryo to initiate transcription of its own gene products, which means that it is no longer dependent on products carried over from the oocyte. The timing of embryonic genome activation is species-specific, but usually occurs between the two- and eight-cell stages [31,33]. Nevertheless, this mtDNA replication event is usually completed by the two-cell embryo stage. Subsequent mtDNA replication is restricted until the blastocyst stage [25,28,34]. As cell division progresses in an exponential manner in the early embryo, mtDNA copy number is diluted out significantly up to the blastocyst stage. Indeed, this appears to be a key developmental process as it is now evident from studies in human embryos cultured *in vitro* that they actively secrete, or shed, mtDNA into the surrounding culture medium. To this extent, it has been hypothesized that this is an indicator of better quality embryo development, as it enhances development to the blastocyst stage and promotes implantation [35]. This outcome is supported by the knowledge that too many copies of mtDNA at the blastocyst stage can lead to aneuploidy and failure of the hatched blastocyst to implant [36].

Pluripotency defines a naïve state that provides the cell with the potential to differentiate into all cell types of the body [37]. Pluripotency is regulated by gene markers that include OCT4, SOX2 and NANOG, which collectively form the pluripotent network, and are expressed in the ICM and undifferentiated embryonic stem cells [38,39]. Indeed, failure to express the markers of the pluripotent network results in loss of pluripotent potential and leads to embryonic death or differentiation [39–41]. Consequently, the trophectodermal cells are the first cells to have undergone differentiation, and they lose their pluripotent potential and express lineage specific markers, such as CDX2 and GATA3 [42–44]. In line with the onset of differentiation, the trophectodermal cells initiate mtDNA replication [25,28,34] (Figure 3). On the other hand, the ICM cells retain low

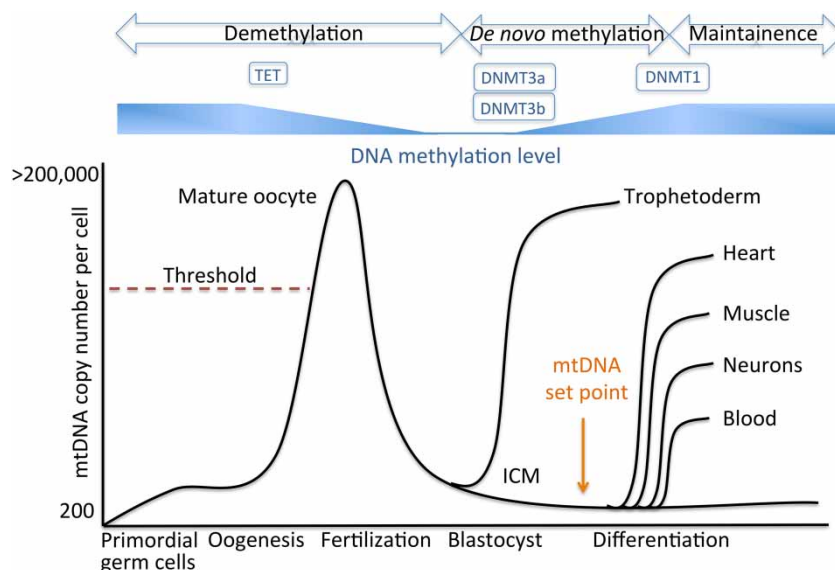


Figure 3. Synchronous changes to mtDNA copy number and DNA methylation during early development.

mtDNA copy number increases during oogenesis, peaks at fertilization, and decreases through to the blastocyst stage. Surpassing the threshold of mtDNA copy number (red broken line) during oogenesis is critical for fertilization. mtDNA replication is initiated in the trophoblast, whereas the ICM continues to reduce mtDNA copy number and establishes the mtDNA set point before differentiation. Once cells commit to a specific lineage, they replicate mtDNA in a cell-specific manner so that they have sufficient mtDNA to transcribe essential subunits of the ETC to produce ATP and perform their specialized functions. Synchronous changes to DNA methylation are also observed to regulate gene expression required for development. TET enzymes reduce parental DNA methylation until the blastocyst stage. *De novo* methylation by DNMT3a and DNMT3b is initiated in the blastocyst. DNMT1 then works to maintain cell-specific DNA methylation profiles.

mtDNA copy number and restrict mtDNA replication so that copy number is further reduced with each cell division [19,44,45] (Figure 3).

As the ICM consists of embryonic cells that are pluripotent, the continual dilution of mtDNA copy number contributes to the establishment of the mtDNA set point, which is required to establish lineage commitment [46]. The mtDNA set point is characterized by low mtDNA copy number. This can range from as few as 20 to 200 copies, which is a similar state as that of the primordial germ cells [22,23]. This provides the template for subsequent differentiation events in order that specialized cells acquire the specific numbers of mtDNA copies required to undertake their specialized functions, as is the case for primordial germ cells that differentiated into mature fertilizable oocytes [20,47–49] (Figure 3). Indeed, it is of great importance for a pluripotent cell to establish the mtDNA set point to maintain pluripotency. Partial knockdown of *PolgA*, the major mtDNA replication factor, in undifferentiated ESCs leads to reduced mtDNA copy number, down-regulation of *Oct4* expression and the onset of early differentiation [19].

There is a close association between pluripotency, mtDNA and mitochondrial function during embryo development, and the consequences of failing to establish the mtDNA set point can be understood from the perspective of induced pluripotency [49]. Induced pluripotent stem cells (iPSCs) are derived from differentiated cells, such as fibroblasts, which undergo reprogramming, or dedifferentiation, to re-establish pluripotency. However, the degree of reprogramming that takes place is a key issue. iPSCs may express pluripotent markers; however, they can still have incomplete pluripotent potential due to the carry-over of epigenetic memory from the cells of origin, which leads them to differentiate into certain cell lineages instead of all cell types [50,51]. Furthermore, in some cases, they spontaneously start to express lineage-specific markers, but mtDNA replication is restricted when compared with differentiated ESCs, which eventually leads to failed differentiation [49]. Indeed, it has been shown that iPSCs, which failed to complete the process of reprogramming, attempt to increase mtDNA copy number during differentiation but this then stalls, a situation that we refer to as ‘pseudo-differentiation’ [8]. Thus complete pluripotency requires not only a genetically and epigenetically

reprogrammed nucleus, but also the establishment of the mtDNA set point to synchronize the amount of mtDNA available to enable cells to perform their specialized functions when fully differentiated.

Cancer cells help to explain the importance of mtDNA to the process of differentiation. Cancer cells are highly proliferative cells, some of which can be induced to differentiate [8]. Indeed, the differentiated status of a cancer cell is widely used to grade cancers [52,53]. A more differentiated cancer cell is more like a normal somatic cell that is non-cancerous in nature, and it will grow at a slower rate than an undifferentiated cancer cell, as is the case for mature differentiated cells [52,53]. A variety of mitochondrial dysfunctions have been discovered in cancer cells, which are tightly related to mtDNA and highly likely to contribute to their failure to undergo differentiation and their ability to adopt a pseudo-differentiated state [8,54]. As an example, glioblastoma multiforme cells can be induced to initiate the process of differentiation. However, the process stalls and is marked by asynchronous increases in mtDNA copy number and perturbed cell-specific gene expression patterns [8]. Undifferentiated cancer cells share some characteristics associated with pluripotency, including cellular energy mechanisms and the regulation of mtDNA replication when they fail to differentiate [8,21]. We now review some of the related mitochondrial dysfunctions identified in tumorigenesis to bring new insights to the understanding and the importance of mtDNA in the overall picture of development.

Mitochondrial dysfunction in tumorigenesis associated with the mtDNA set point

The Warburg effect

The Warburg hypothesis states that cancer cells primarily rely on aerobic glycolysis instead of OXPHOS for energy production even in the presence of oxygen (Figure 4). In 1924, Warburg initially asserted the hypothesis that the primary cause of cancer was the damage to mitochondria and, subsequently, insufficient metabolic respiration [55]. With the finding of oncogenic mutations as the root cause of cancer [56], the Warburg effect currently describes the metabolic features of cancer cells. Indeed, not only do cancer cells utilize aerobic glycolysis for energy production but embryonic cells do so in a similar manner [57]. Low mtDNA copy number in cancer and pluripotent cells limits energy production through OXPHOS, and this is thought to aid cellular proliferation [58]. Even though aerobic glycolysis is not as efficient as OXPHOS in generating large amounts of ATP, one explanation is that it incorporates nutrients and metabolizes them into other organic elements (e.g. nucleotides, amino acids and lipids) to support the requirements to build and generate new cells when in a proliferative mode [58]. Although impaired mitochondrial function is not found in most cancers to be the direct cause mediated by the Warburg effect [58], recent studies have extended the hypothesis to a variety of mitochondrial dysfunctions in cancer cells, which might contribute to or result from the failure to establish the mtDNA set point [59].

Pathological DNA mutations

Sporadic nDNA mutations or abnormal expression levels of the nDNA-encoded factors of mtDNA transcription and replication lead to dysfunctional mtDNA transcription and replication, which could result in reduced oxidative capacity and energy production. Once the reduced level affects cellular metabolism, clinical symptoms are often observed. It is known that mutations to *POLG* could lead to depletion of mtDNA and promote breast cancer [60]. In addition, frequent truncating mutations of *Tfam* have been demonstrated to induce mtDNA depletion and apoptotic resistance in a mouse model of colorectal cancer [61]. Thus these mutations could perturb the establishment of the mtDNA set point, which then limits the differentiation potential of tumour cells.

The impact of ROS production and mitochondrial metabolites on gene expression

ROS are a group of by-products from cellular metabolism that can readily oxidize other molecules. ROS are frequently generated in the mitochondria [62]. Increased ROS results from cellular stress and can translate the stress into cellular signalling pathways for certain responses [63]. It has been demonstrated that increased ROS could trigger mitophagy and apoptosis to either reduce the number of mitochondria, therefore reducing mtDNA copy number, or trigger cell death dependent on the extent of the increase [64,65]. Low levels of ROS are beneficial for healthy metabolism through post-translational modification of phosphatases and kinases [66],

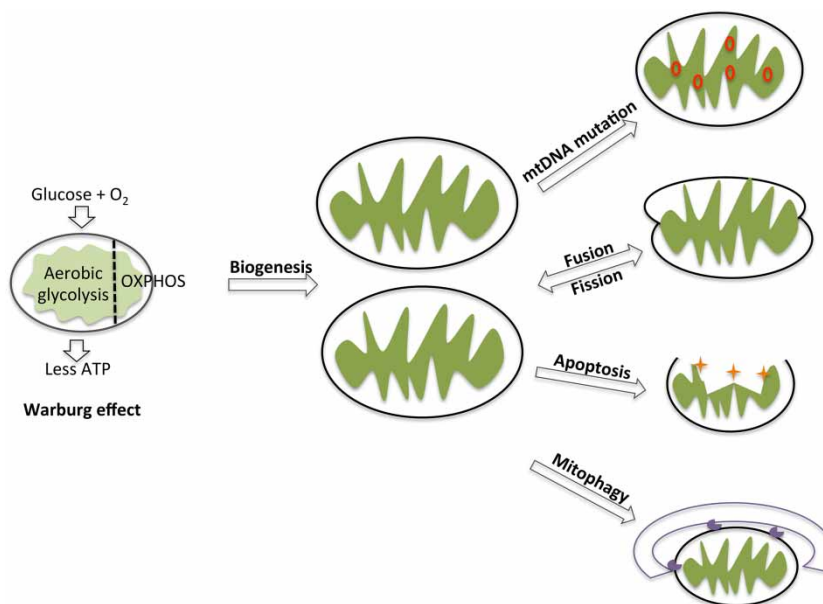


Figure 4. Mitochondrial function and dysfunction in tumour cells and during development.

Tumour cells mainly rely on aerobic glycolysis even in the presence of oxygen, a process known as the Warburg effect. Pluripotent cells also utilize glycolysis to produce ATP probably due to limited numbers of mtDNA copy. Mitochondrial biogenesis involves formation of cristae and a membrane potential. Mature mitochondria in cells exhibit various activities including fusion and fission, mtDNA mutations, cellular apoptosis and mitophagy, which could contribute to developmental disorders and cancers.

whereas excessively high levels of ROS generate lethal levels of oxidative stress and lead to cell death by causing damage to DNA [67], triggering senescence [68] and apoptosis [69].

Elevated ROS has been identified in cancer cells that are induced by oncogenes to promote tumorigenesis. However, to maintain the elevated level at a beneficial level rather than at the excessively high level, it is tightly controlled by aberrant gene expression and metabolism in cancer, as cancer cells can produce higher levels of antioxidants to balance the redox level to a level beneficial for tumorigenesis [64]. It has been reported that PGC1 α can up-regulate the expression of antioxidant enzymes to maintain redox balance [70]. The tumour-suppressor gene p53 is also known to interact with NRF2 to up-regulate the expression of antioxidants [71]. The oncogene MYC regulates antioxidant levels by promoting glutamine uptake and glutaminolysis [72]. Furthermore, down-regulation of TFAM has recently been found to inhibit the tumorigenesis of non-small-cell lung cancer by activating ROS-mediated c-Jun N-terminal kinase (JNK)/p38 mitogen-activated protein kinase (MAPK) signalling and reduce cellular energy [73]. These findings related to the mtDNA replication and transcription factors provide functional explanations as to how the failure to establish the mtDNA set point blocks the differentiation of cancer cells and promotes tumorigenesis, but this has yet to be tested experimentally.

Mutations to mtDNA

The mitochondrial genome encodes several subunits of the ETC complexes and other functional RNAs. Therefore, reduced function in the ETC complexes resulting from homoplasmic mtDNA mutations can lead to OXPHOS inefficiency. Mutations in the mtDNA-encoded *ND* genes and *COX1* have been identified in primary tumours and it is suggested that they promote tumour progression [74–78]. This will contribute to the inefficiency within the ETC, as more electrons are passed through the complexes resulting in the production of more ROS that causes further mtDNA damage or mutation [74] (Figure 4). Nevertheless, mechanistic evidence is needed to determine whether mtDNA mutations are a part of tumorigenesis or just a part of mitochondrial dysfunction in cancer. Even though the levels of mtDNA mutations are not directly related to mtDNA copy number, they affect the cell's respiratory capacity. If the levels of mtDNA mutations affect the function of the mtDNA-encoded ETC subunits, it therefore suggests that the establishment of the mtDNA set point is

disturbed functionally. Consequently, mtDNA copy number should be assessed in the context of mitochondrial function and whether a cell's set point is modulated to compensate for this.

The control of the number of mitochondria

Mitochondrial biogenesis is activated through environmental stimuli or when cells are under redox stress [79–82]. Mitochondrial biogenesis is a complicated procedure requiring mtDNA replication and synthesis of ETC subunits [81,82]. In early differentiation, mitochondrial maturation also involves the formation of cristae and a higher membrane potential for normal mitochondrial functions [83,84] (Figure 4). The Sirtuin family, especially Sirtuin 1 and 3 (encoded by *SIRT1/SIRT3*) have been shown to be involved in mitochondrial biogenesis, in mitochondrial transcription as the target of PGC1 α , and have a role in suppressing ROS production [85,86]. Moreover, the ETC subunits, as well as other mitochondrial structural proteins, are encoded by both the chromosomal and mitochondrial genomes. Thus the process of mitochondrial biogenesis is highly dependent on the ordered and effective co-operation between the hierarchical network of the factors from both the nucleus and the mitochondria. In cancer, some oncogenes and tumour suppressors have been found to affect this network to alter the rate of biogenesis. The tumour suppressor p53 has been demonstrated to repress the transcription of the main initiator of mitochondrial biogenesis, PGC1 α , when telomere function is compromised [87], and therefore causes deregulation of the downstream pathway through cofactors of PGC1 α , namely PPAR γ , NRF1/2 and oestrogen-related receptor (ERR) α/β and, extensively, to key mitochondrial regulators, such as TFAM and POLGA [81]. In addition, the oncogene *MYC* could promote mitochondrial biogenesis by up-regulating expression of PGC1 β [88–90] (Figure 1). The cell proliferation factor signal transducer and activator of transcription 3 (STAT3) also directly binds to the D-loop region of mtDNA to regulate its transcription. More importantly, STAT3 promotes mitochondrial transcription to optimize energy production which is used to promote the proliferation of ESCs [91]. STAT3 also targets hypoxia-induced factor 1 α (HIF1 α), the master transcription regulator of the genes that respond to hypoxia, which also interacts with PGC1 α to promote mitochondrial biogenesis [92,93]. Increased mitochondrial biogenesis in cancer cells is probably a response to offset lower energy production resulting from the Warburg effect and to balance out the elevated levels of ROS.

On the other hand, the number of mitochondria is also controlled by mitochondria-specific autophagy, known as mitophagy. Mitophagy describes the process of degradation of mitochondria by the autophagosome (Figure 4). Mitophagy is triggered when the mitochondrial membrane is depolarized, which, in most cases, results from mitochondrial fission (discussed below) [64,94]. Mitophagy has tumour-suppressor function as the loss of the key regulators of mitophagy, parkin and BNIP3/NIX, promotes tumorigenesis dependent on the type and the stage of the cancers [95–97]. BNIP3/NIX can be induced by hypoxia as the tumour cell microenvironment involves the regulation of p53 [96]. In ESCs, cellular function and pluripotency are also sensitive to the microenvironment, especially in tissue culture. It has been shown that the lack of a growth factor, Gfer, can trigger mitophagy and loss of pluripotency, which could result in reduced mtDNA copy number and failure to establish the mtDNA set point [98].

The dynamics of mitochondrial fission and fusion happen at a balanced rate in normal cells, which closely responds to hypoxia, energy metabolism, apoptosis and other cellular stress [99] (Figure 4). Mitochondrial fission is driven by the GTPase dynamin-related protein 1 (DRP1) with co-factors including MFF1 and FIS1. Mitochondrial fusion requires the GTPases mitofusin 1/2 located on the outer membrane and OPA1 on the inner membrane of mitochondria [99]. Mitochondrial fission is important for mitosis to distribute the organelles evenly into daughter cells, whereas fusion has been found to promote OXPHOS and cell growth probably by increasing the density of the ETC complexes and improving the formation of the ETC supercomplexes [100]. Evidence of increased fission rate has been reported in cancer cells and has been related to the increases in ROS and mitophagy, which again could lead to the reduction in mtDNA copy number [64]. Moreover, it has been noted that mitochondrial dynamics might play vital roles in apoptosis, as observed in interactions between the key regulators of fusion/fission and apoptosis including oncogenic BCL-2 family members [101,102]. As apoptotic evasion is one of the hallmarks of tumorigenesis, these findings have shown the potential contribution of mitochondrial dynamics to tumorigenesis. However, direct experimental results need to be analysed to determine the cellular mechanisms in which the abnormal mitochondrial dynamics identified in tumorigenesis contribute to or resulted from the abnormal regulation of mtDNA copy number and stalled differentiation.

In summary, due to the fundamental role of mitochondria in cellular metabolism and bioenergetics, it is not surprising to see such a variety of mitochondrial dysfunctions associated with tumorigenesis (Figure 4).

However, most of the explanations are not clear enough to clarify how mitochondrial dysfunction mechanistically relates to tumorigenesis and failure to complete differentiation. As it has been shown that, by resetting mtDNA copy number and thus the mtDNA set point, cellular differentiation could be promoted [8], these dysfunctions probably result from or contribute to the failure to establish the mtDNA set point. As reviewed above, the interactions between oncogenes/tumour suppressors and mitochondrial replication and transcription factors are present at multiple levels, which result in a variety of metabolic changes during tumorigenesis. Cancers derived from different cell types exhibit different mitochondrial dysfunctions with different metabolic pathways being perturbed. Moreover, the mtDNA set point is essential to determining cell fate. These findings highlight that the regulation of mtDNA copy number co-operates tightly with other metabolic pathways to regulate various cellular functions, which, if perturbed, could lead to tumorigenesis.

The effect of DNA methylation on mtDNA copy number during differentiation

What is DNA methylation?

DNA methylation is the only known covalent modification of DNA [103]. Typically, DNA methylation occurs at 5' of the pyrimidine ring of cytosine (5 mC) followed by guanine (CpG dinucleotide) in mammalian DNA (Figure 5). It is known that over 28 million CpGs exist in the human genome. In plants, it can also occur at non-CpG cytosines [104]. There is also emerging evidence showing that non-CpG DNA methylation can occur at CpA and CpT dinucleotides in mammalian cells, such as embryonic and neuronal cells [105,106].

Methyl groups are added to cytosines by DNA methyltransferase (DNMT) enzymes [107] (Figure 5). There are two types of DNMT enzymes either working to maintain DNA methylation during DNA replication or initiating DNA methylation at a novel DNA site. Maintenance of DNA methylation is carried out by DNMT1 along with DNA replication, whereas DNMT3A and DNMT3B are believed to function as the *de novo* methylation DNMT enzymes adding methyl groups at new cytosine sites [108] (Figure 5). 5 mC can be demethylated to 5-hydroxymethylcytosine (5 hmC) by ten-eleven translocation methylcytosine dioxygenase (TET; Figure 5) [109]. The TET family, consisting of TET1, TET2 and TET3, is known to oxidize 5 mC to 5 hmC [109,110].

CpG dinucleotides usually appear in repeat clusters termed CpG islands (CGIs). The widely accepted definition of CGI was given by Gardiner-Garden and Frommer in 1987 as a 200-bp sequence of DNA with a CG content of more than 50% and the observed/expected CpG ratio (O/E ratio) is more than 0.6 [111]. Where the O/E ratio is calculated by the formula [111]:

$$\text{Observed/expected CpG ratio} = \frac{\text{Number of CpG} \times \text{length of sequence}}{\text{Number of C} \times \text{number of G}}$$

CGIs contain high CpG components, which could be 5-fold higher than the other regions. The estimated number of CGIs is 30 000 in the human genome. On average, CGIs are ~1 kb in length with 7% being CpG dinucleotides [111,112].

Chemical agents can also inhibit DNA methylation or induce DNA demethylation. They have been widely used in DNA methylation research. 5-Azacytidine is a chemical analogue of cytosine that inhibits DNA methylation at low concentrations [113]. It blocks DNA methyltransferase, DNMT1, which causes global DNA demethylation [114] (Figure 5). Recent research has revealed that vitamin C can also induce DNA demethylation by enhancing the activity of the TET1 enzymes and thus improving the conversion of 5 mC into the DNA demethylated state of 5 hmC [115] (Figure 5).

In mammalian cells, the methylome is mainly composed of genomic regions of low CpG densities, whereas regions of higher CpG densities, such as CGIs, are mainly unmethylated. DNA methylation profiles vary significantly between different tissue types reflecting different gene expression profiles and biological functions [116]. DNA regions with a different methylation status can result in different biological consequences.

Biological function of DNA methylation in different gene regions

DNA methylation plays vital roles in embryonic development, maintenance of pluripotency, genomic imprinting and X chromosome inactivation by regulating transcription and stabilizing chromosomes and chromatin structure [117–119]. It is a dispensable epigenetic system in the regulation of gene expression. Most often,

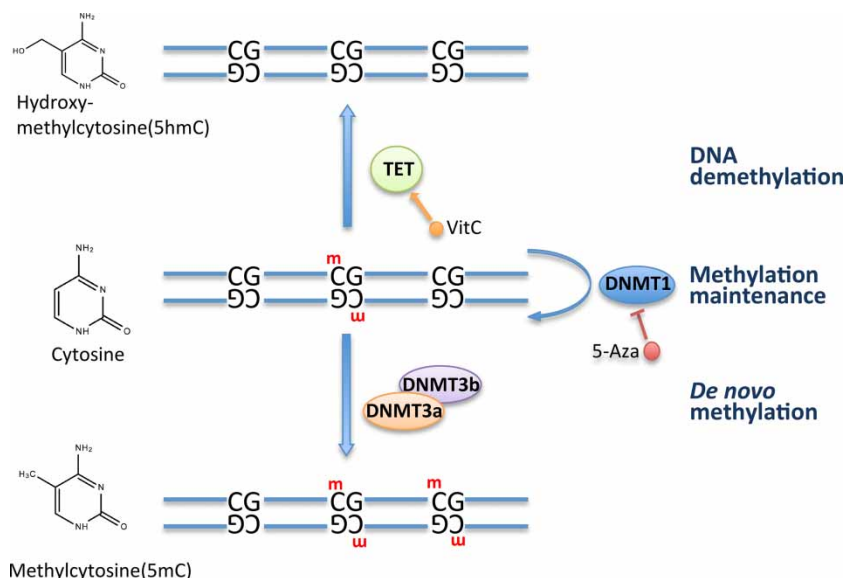


Figure 5. DNA methylation and demethylation mechanisms.

Cytosine in CpG dinucleotides can be methylated to 5 mC by DNMT enzymes: DNMT1 maintains DNA methylation patterns, DNMT3a/b create novel DNA methylation sites. 5 mC can be demethylated by TET enzymes to 5 hmC. DNA methylation inhibitors, such as 5-Azacytidine (5-Aza), inhibit DNMT1, whereas vitamin C (VitC) enhances TET activity.

DNA methylation is best characterized by silencing transcription which occurs at CpG islands in gene promoter regions. Through access to high-throughput methylome datasets, it has been suggested that DNA methylation seems to occur at intragenic (within gene-coding) regions, and at regulatory intergenic (between gene-coding) regions, which appear to regulate gene expression patterns differently.

More than 50% of human genes possess CGIs in their promoter regions [120]. As mentioned above, most of them remain unmethylated in somatic cells. Some suppressed genes have been shown with methylated CGIs in their promoter regions [116]. Hypermethylated promoters have been well characterized to block binding of transcription factors and to silence gene expression. Aberrant methylation of promoter-located CGIs, especially hypermethylated promoters, plays important roles in repressing tumour-suppressor genes [121,122]. Interestingly, nearly all of the known mitochondrial replication and transcription factors have a CGI in their promoter regions, which indicates that their gene expression is probably affected by the DNA methylation status across their CGI regions. DNA hypermethylation across the promoters of *PPARGC1A* (encoding PGC1 α) has been shown to be mediated by DNMT3B and leads to the reduction of mitochondrial content in muscle cells in diabetic patients [123]. Methylation in the promoter region of TFAM also has an impact on mitochondrial function in insulin resistance in peripheral leucocytes [124].

Methylome datasets of single-base resolution have shown that fewer CGIs exist but more DNA methylation occurs in intragenic regions [125,126]. Intragenic methylation also presents in a strong tissue-specific manner. However, the function of intragenic methylation shows diversity and has also been referred to as the DNA methylation paradox [127]. It is believed that high DNA methylation is important for blocking transcription of the repetitive DNA elements and suppressing initiation from spurious promoters within gene regions [122,128]. However, a non-monotonic relationship is observed between intragenic methylation and gene expression [129]. Both positive and negative correlations exist between DNA methylation at exons and gene expression in different cell types [125]. Evidence has also shown that DNA methylation across exons is higher than in introns [130]. The transition of DNA methylation at exon/intron boundaries leads to a proposed function for maintaining transcriptional efficiency by regulating splicing and elongation [130]. Moreover, the dynamics of DNA methylation across gene regions might also be related to the transcriptional unit including the accessibility of chromatin and the relative density of the nucleosomes, DNMT1 and human RNA polymerase II (POL II) [129]. In an inactive gene region, chromatin is closed and dense nucleosomes block DNMT1 from mediating DNA methylation, whereas when chromatin partially opens, DNMT1 methylates gene regions

along with POL II to execute transcription; a completely opened chromatin is fully occupied with POL II for transcription, which interferes with the access of DNMT1.

Overall, there is still a question mark over the relationship between methylome patterns and transcriptional profiles. When talking about DNA methylation patterns, inter- (between gene-coding regions) and intra- (within gene-coding regions) genetically, it is necessary to take gene expression and other regulatory mechanisms into consideration and investigate them all thoroughly based on the particular cell types and their corresponding cell functions.

The effect of DNA methylation on mtDNA copy number during differentiation

During cell differentiation and embryonic development, the DNA methylation profile is re-established. After fertilization, TET enzymes become activate and erase the parental DNA methylation levels significantly [131] (Figure 3). By the blastocyst stage, the extensively demethylated genome of the ICM starts to be re-methylated by DNMT3a and DNMT3b up to their specific levels and maintained by DNMT1 afterwards during DNA replication [132] (Figure 3). The changes that occur during differentiation are believed to be predominantly responsible for transcriptional changes specified for different cell fates. It has been reported that *in vivo* liver differentiation in human ESCs is characterized by DNA demethylation over time [133]. In a DNA methylation study targeting genome-wide promoter regions of mouse ESCs, it was reported that the DNA–protein interactions between the promoter regions and OCT4, NANOG, Polycomb proteins and histone modifications (H3K27 trimethylation) are correlated with the DNA methylation status of the cell [134]. Increasing studies on DNA methylation in embryonic cells have provided insights into how developmental processes are modulated by DNA methylation and more and more cell lineages are characterized by their DNA methylation signatures.

As mentioned before, mtDNA copy number is also reset to establish the mtDNA set point starting at the blastocyst stage leading towards specified mtDNA copy number for different cell lineages (Figure 3). It has been questioned whether there is a relationship between the coinciding changes in DNA methylation and mtDNA copy number. A recent study revealed that the levels of DNA methylation increased at exon 2 of *PolgA* in mouse ESCs as they underwent differentiation into neurons [20]. Furthermore, the level of DNA methylation varies at exon 2 of *PolgA* in a tissue-specific manner. Statistical analysis suggested that *PolgA* expression was negatively correlated with its intragenic DNA methylation at exon 2 and the corresponding mtDNA copy number, which indicates that changes to DNA methylation could modulate mtDNA copy number by regulating the gene expression of the primary mtDNA replication factor POLGA. Similar findings are also observed in the differentiation process from human ESCs to neural stem cells and finally into astrocytes [21]. High DNA methylation was observed in pluripotent and multipotent cells that correlated with low *POLGA* expression and low mtDNA copy number. In terminally differentiated astrocytes, however, DNA methylation at exon 2 of *POLGA* decreased with higher mtDNA copy number. Interestingly, cancer cell lines, such as the glioblastoma multiforme HSR-GBM1 cells, which failed to differentiate, are unable to demethylate exon 2 of *POLGA* [21]. However, in the presence of DNA demethylation agents such as 5-Aza, the high DNA methylation status at exon 2 of *POLGA* was decreased promoting mtDNA replication and cell differentiation. These findings highlight the important role of DNA methylation in modulating mtDNA copy number, both of which act as vital determinants in cell differentiation and development (Figure 6). Likewise, partial depletion of mtDNA in cancer cells changes the methylation status at exon 2 of *POLGA* and affects the tumorigenic capacity of these cells. Nevertheless, when tumours form, mtDNA copy number is reinstated to original levels, suggesting that cells do, indeed, have an ‘in built’ set point [21,135].

Interestingly, the mtDNA set point can be modulated by mtDNA genotype. In cells that have had their mtDNA replaced with mtDNA from different genotypes, either in tumour cell [135] or ESC [136] models, not only are chromosomal gene expression patterns changed, but mtDNA copy number is modulated in response along with DNA methylation patterns at exon 2 of *POLGA* [21]. This indicates that the establishment of the mtDNA set point is a two way process where interactions between the nuclear and mitochondrial genomes reflect adaptation to the environment and cellular phenotype.

Conclusion

mtDNA is a critical part of the cellular energy-production system. Nevertheless, its impact is no longer limited to powering cellular function. mtDNA copy number changes during cellular development which affects fertilization rates and embryo viability. The establishment of the mtDNA set point is essential for the potential of a

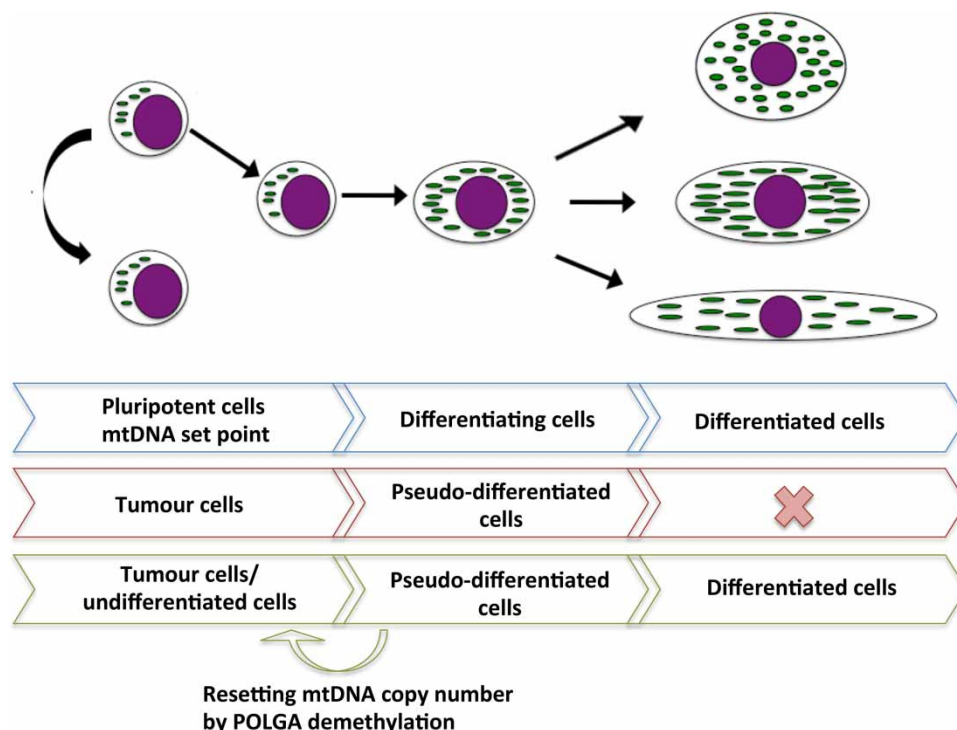


Figure 6. The mtDNA set point is critical for differentiation of pluripotent cells and tumour cells.

During differentiation, mtDNA copy number is strictly regulated in a cell-specific manner accompanied by changes in the DNA methylation status at exon 2 of *POLGA*. Cancer cells fail to complete differentiation and adopt a pseudo-differentiated state. However, they can complete differentiation by resetting mtDNA copy number by DNA demethylation.

cell to complete differentiation. During differentiation, mtDNA copy number increases in a lineage-specific manner as dictated by a cell's specialized cellular functions. The changes to mtDNA copy number are determined not only by the expression of mtDNA transcription and replication factors but also by stage-specific markers, as is the case during pluripotency, which it is very tightly regulated. The synchronous changes to DNA methylation during development, especially at exon 2 of *POLGA*, provide insights into our understanding of the cross-talk between the nucleus and the mitochondria; in other words, the changes to gene expression profiles and mtDNA copy number. These findings can be combined further to understand tumorigenesis and developmental disorders. Aberrant DNA methylation profiles found in early development and tumorigenesis interfere with the expression of *POLGA*, which destabilizes the mtDNA set point and the expansion of mtDNA copy number during differentiation. As a great variety of genes and regulators are involved in regulating mtDNA as well as general mitochondrial function, it is highly likely that similar mechanisms exist that have an impact on development and tumorigenesis.

Abbreviations

5-Aza, 5-azacitidine; CGI, CpG island; CoQ, coenzyme Q; COX, cytochrome c oxidase; CytB, cytochrome c reductase; CytC, cytochrome c; D-loop, displacement loop; DNMT, DNA methyltransferase; ERR, oestrogen-related receptor; ESC, embryonic stem cell; ETC, electron transfer chain; HIF1 α , hypoxia-induced factor 1 α ; 5 hmC, 5-hydroxymethylcytosine; 5 mC, 5-methylcytosine; ICM, inner cell mass; iPSC, induced pluripotent stem cell; ND, NADH dehydrogenase; nDNA, nuclear DNA; NRF, nuclear respiratory factor; O/E ratio, observed/expected CpG ratio; OXPHOS, oxidative phosphorylation; PGC1 α , peroxisome-proliferator-activated receptor γ co-activator 1 α ; POLG, DNA polymerase gamma; POLGA, POLG subunit A; POL II, RNA polymerase II; PPAR, peroxisome-proliferator-activated receptor; ROS, reactive oxygen species; SIRT, sirtuin; STAT3, signal transducer and activator of transcription 3; TET, ten-eleven translocation methylcytosine dioxygenase; TFAM, mitochondrial transcription factor A.

Funding

This work was funded by the Australian National Health and Medical Research Council [grant number GNT1041471] and the Victorian Government's Operational Infrastructure Support Program.

Competing Interests

The Authors declare that there are no competing interests associated with the manuscript.

References

- Duchen, M.R. (1999) Contributions of mitochondria to animal physiology: from homeostatic sensor to calcium signalling and cell death. *J. Physiol.* **516**, 1–17 doi:10.1111/j.1469-7793.1999.001aa.x
- Rossi, C.S., Bielawski, J. and Lehninger, A.L. (1966) Separation of H^+ and OH^- in the extramitochondrial and mitochondrial phases during Ca^{++} -activated electron transport. *J. Biol. Chem.* **241**, 1919–1921 PMID: 5945863
- Mitchell, P. (1961) Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* **191**, 144–148 doi:10.1038/191144a0
- Brown, G.C. (1992) Control of respiration and ATP synthesis in mammalian mitochondria and cells. *Biochem. J.* **284**, 1–13 doi:10.1042/bj2840001
- Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, J. et al. (1981) Sequence and organization of the human mitochondrial genome. *Nature* **290**, 457–465 doi:10.1038/290457a0
- Ojala, D., Montoya, J. and Attardi, G. (1981) tRNA punctuation model of RNA processing in human mitochondria. *Nature* **290**, 470–474 doi:10.1038/290470a0
- Satoh, M. and Kuroiwa, T. (1991) Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell. *Exp. Cell Res.* **196**, 137–140 doi:10.1016/0014-4827(91)90467-9
- Dickinson, A., Yeung, K.Y., Donoghue, J., Baker, M.J., Kelly, R.D.W., McKenzie, M. et al. (2013) The regulation of mitochondrial DNA copy number in glioblastoma cells. *Cell Death Differ.* **20**, 1644–1653 doi:10.1038/cdd.2013.115
- Shoubbridge, E.A. and Wai, T. (2007) Mitochondrial DNA and the mammalian oocyte. *Curr. Top. Dev. Biol.* **77**, 87–111 doi:10.1016/S0070-2153(06)77004-1
- Alexeyev, M., Shokolenko, I., Wilson, G. and LeDoux, S. (2013) The maintenance of mitochondrial DNA integrity: critical analysis and update. *Cold Spring Harb. Perspect. Biol.* **5**, a012641 doi:10.1101/cshperspect.a012641
- Schon, E.A., DiMauro, S. and Hirano, M. (2012) Human mitochondrial DNA: roles of inherited and somatic mutations. *Nat. Rev. Genet.* **13**, 878–890 doi:10.1038/nrg3275
- Gray, M.W., Burger, G. and Lang, B.F. (1999) Mitochondrial evolution. *Science* **283**, 1476–1481 doi:10.1126/science.283.5407.1476
- Falkenberg, M. and Larsson, N.G. (2007) DNA replication and transcription in mammalian mitochondria. *Annu. Rev. Biochem.* **76**, 679–699 doi:10.1146/annurev-biochem.76.060305.152028
- Leigh-Brown, S., Enriquez, J.A. and Odom, D.T. (2010) Nuclear transcription factors in mammalian mitochondria. *Genome Biol.* **11**, 215 doi:10.1186/gb-2010-11-7-215
- Holt, I.J. and Reyes, A. (2012) Human mitochondrial DNA replication. *Cold Spring Harb. Perspect. Biol.* **4**, a012971 doi:10.1101/cshperspect.a012971
- Gustafsson, C.M., Falkenberg, M. and Larsson, N.G. (2016) Maintenance and expression of mammalian mitochondrial DNA. *Annu. Rev. Biochem.* **85**, 133–160 doi:10.1146/annurev-biochem-060815-014402
- Young, M.J. and Copeland, W.C. (2016) Human mitochondrial DNA replication machinery and disease. *Curr. Opin. Genet. Dev.* **38**, 52–62 doi:10.1016/j.gde.2016.03.005
- Kaguni, L.S. and Oliveira, M.T. (2016) Structure, function and evolution of the animal mitochondrial replicative DNA helicase. *Crit. Rev. Biochem. Mol. Biol.* **51**, 53–64 doi:10.3109/10409238.2015.1117056
- Facucho-Oliveira, J.M., Alderson, J., Spikings, E.C., Egginton, S. and St. John, J.C. (2007) Mitochondrial DNA replication during differentiation of murine embryonic stem cells. *J. Cell Sci.* **120**, 4025–4034 doi:10.1242/jcs.016972
- Kelly, R.D.W., Mahmud, A., McKenzie, M., Trounce, I.A. and St John, J.C. (2012) Mitochondrial DNA copy number is regulated in a tissue specific manner by DNA methylation of the nuclear-encoded DNA polymerase gamma A. *Nucl. Acids Res.* **40**, 10124–10138 doi:10.1093/nar/gks770
- Lee, W., Johnson, J., Gough, D.J., Donoghue, J., Cagnone, G.L.M., Vaghjiani, V. et al. (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 doi:10.1038/cddis.2015.34
- Cree, L.M., Samuels, D.C., de Sousa Lopes, S.C., Rajasimha, H.K., Wonnapijit, P., Mann, J.R. et al. (2008) A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes. *Nat. Genet.* **40**, 249–254 doi:10.1038/ng.2007.63
- Wai, T., Teoli, D. and Shoubbridge, E.A. (2008) The mitochondrial DNA genetic bottleneck results from replication of a subpopulation of genomes. *Nat. Genet.* **40**, 1484–1488 doi:10.1038/ng.258
- El Shourbagy, S.H., Spikings, E.C., Freitas, M. and St John, J.C. (2006) Mitochondria directly influence fertilisation outcome in the pig. *Reproduction* **131**, 233–245 doi:10.1530/rep.1.00551
- Spikings, E.C., Alderson, J. and St John, J.C. (2007) Regulated mitochondrial DNA replication during oocyte maturation is essential for successful porcine embryonic development. *Biol. Reprod.* **76**, 327–335 doi:10.1095/biolreprod.106.054536
- Reynier, P., May-Panloup, P., Chretien, M.-F., Morgan, C.J., Jean, M., Savagner, F. et al. (2001) Mitochondrial DNA content affects the fertilizability of human oocytes. *Mol. Hum. Reprod.* **7**, 425–429 doi:10.1093/molehr/7.5.425
- Santos, T.A., El Shourbagy, S. and St. John, J.C. (2006) Mitochondrial content reflects oocyte variability and fertilization outcome. *Fertil. Steril.* **85**, 584–591 doi:10.1016/j.fertnstert.2005.09.017
- May-Panloup, P., Vignon, X., Chretien, M.-F., Heyman, Y., Tamassia, M., Malhière, Y. and Reynier, P. (2005) Increase of mitochondrial DNA content and transcripts in early bovine embryogenesis associated with upregulation of mtTFA and NRF1 transcription factors. *Reprod. Biol. Endocrinol.* **3**, 65 doi:10.1186/1477-7827-3-65

- 29 Cagnone, G.L.M., Tsai, T.S., Makanji, Y., Matthews, P., Gould, J., Bonkowski, M.S. et al. (2016) Restoration of normal embryogenesis by mitochondrial supplementation in pig oocytes exhibiting mitochondrial DNA deficiency. *Sci. Rep.* **6**, 23229 doi:10.1038/srep23229
- 30 Chappel, S. (2013) The role of mitochondria from mature oocyte to viable blastocyst. *Obstet. Gynecol. Int.* **2013**, 1–10 183024 doi:10.1155/2013/183024
- 31 Aoki, F., Worrall, D.M. and Schultz, R.M. (1997) Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. *Dev. Biol.* **181**, 296–307 doi:10.1006/dbio.1996.8466
- 32 St John, J.C., Amaral, A., Bowles, E., Oliveira, J.F., Lloyd, R., Freitas, M. et al. (2006) The analysis of mitochondria and mitochondrial DNA in human embryonic stem cells. *Methods Mol. Biol.* **331**, 347–374 doi:10.1345/1-59785-046-4:347
- 33 Telford, N.A., Watson, A.J. and Schultz, G.A. (1990) Transition from maternal to embryonic control in early mammalian development: a comparison of several species. *Mol. Reprod. Dev.* **26**, 90–100 doi:10.1002/mrd.1080260113
- 34 McConnell, J.M.L. and Petrie, L. (2004) Mitochondrial DNA turnover occurs during preimplantation development and can be modulated by environmental factors. *Reprod. Biomed. Online* **9**, 418–424 doi:10.1016/S1472-6483(10)61277-1
- 35 Stigliani, S., Persico, L., Lagazio, C., Anserini, P., Venturini, P.L. and Scaruffi, P. (2014) Mitochondrial DNA in Day 3 embryo culture medium is a novel, non-invasive biomarker of blastocyst potential and implantation outcome. *Mol. Hum. Reprod.* **20**, 1238–1246 doi:10.1093/molehr/gau086
- 36 Fragouli, E. and Wells, D. (2015) Mitochondrial DNA assessment to determine oocyte and embryo viability. *Semin. Reprod. Med.* **33**, 401–409 doi:10.1055/s-0035-1567821
- 37 Bradley, A., Evans, M., Kaufman, M.H. and Robertson, E. (1984) Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* **309**, 255–256 doi:10.1038/309255a0
- 38 Loh, Y.-H., Wu, Q., Chew, J.-L., Vega, V.B., Zhang, W., Chen, X. et al. (2006) The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* **38**, 431–440 doi:10.1038/ng1760
- 39 Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N. and Lovell-Badge, R. (2003) Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* **17**, 126–140 doi:10.1101/gad.224503
- 40 Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K. et al. (2003) The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631–642 doi:10.1016/S0092-8674(03)00393-3
- 41 Nichols, J., Zevnik, B., Anastasiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I. et al. (1998) Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379–391 doi:10.1016/S0092-8674(00)81769-9
- 42 Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R. and Rossant, J. (2005) Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* **123**, 917–929 doi:10.1016/j.cell.2005.08.040
- 43 Home, P., Ray, S., Dutta, D., Bronshteyn, I., Larson, M. and Paul, S. (2009) GATA3 is selectively expressed in the trophectoderm of peri-implantation embryo and directly regulates Cdx2 gene expression. *J. Biol. Chem.* **284**, 28729–28737 doi:10.1074/jbc.M109.016840
- 44 Wray, J., Kalkan, T. and Smith, A.G. (2010) The ground state of pluripotency. *Biochem. Soc. Trans.* **38**, 1027–1032 doi:10.1042/BST0381027
- 45 Evans, M.J. and Kaufman, M.H. (1981) Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154–156 doi:10.1038/292154a0
- 46 St. John, J.C., Facucho-Oliveira, J., Jiang, Y., Kelly, R. and Salah, R. (2010) Mitochondrial DNA transmission, replication and inheritance: a journey from the gamete through the embryo and into offspring and embryonic stem cells. *Hum. Reprod. Update* **16**, 488–509 doi:10.1093/humupd/dmq002
- 47 St. John, J. (2014) The control of mtDNA replication during differentiation and development. *Biochim. Biophys. Acta* **1840**, 1345–1354 doi:10.1016/j.bbagen.2013.10.036
- 48 St John, J.C. and Campbell, K.H.S. (2010) The battle to prevent the transmission of mitochondrial DNA disease: is karyoplast transfer the answer? *Gene Ther.* **17**, 147–149 doi:10.1038/gt.2009.164
- 49 Kelly, R.D.W., Sumer, H., McKenzie, M., Facucho-Oliveira, J., Trounce, I.A., Verma, P.J. and St. John, J.C. (2013) The effects of nuclear reprogramming on mitochondrial DNA replication. *Stem Cell Rev. Rep.* **9**, 1–15 doi:10.1007/s12015-011-9318-7
- 50 Polo, J.M., Liu, S., Figueroa, M.E., Kulalert, W., Eminli, S., Tan, K.Y. et al. (2010) Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat. Biotechnol.* **28**, 848–855 doi:10.1038/nbt.1667
- 51 Polo, J.M., Anderssen, E., Walsh, R.M., Schwarz, B.A., Nefzger, C.M., Lim, S.M. et al. (2012) A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* **151**, 1617–1632 doi:10.1016/j.cell.2012.11.039
- 52 Umbas, R., Schalken, J.A., Aalders, T.W., Carter, B.S., Karthaus, H.F., Schaafsma, H.E. et al. (1992) Expression of the cellular adhesion molecule E-cadherin is reduced or absent in high-grade prostate cancer. *Cancer Res.* **52**, 5104–5109 PMID: 1516067
- 53 Elston, C.W. and Ellis, I.O. (1991) Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* **19**, 403–410 doi:10.1111/j.1365-2559.1991.tb00229.x
- 54 St. John, J.C. (2016) Mitochondrial DNA copy number and replication in reprogramming and differentiation. *Semin. Cell Dev. Biol.* **52**, 93–101 doi:10.1016/j.semcdb.2016.01.028
- 55 Warburg, O. (1956) On the origin of cancer cells. *Science* **123**, 309–314 doi:10.1126/science.123.3191.309
- 56 Shih, C. and Weinberg, R.A. (1982) Isolation of a transforming sequence from a human bladder carcinoma cell line. *Cell* **29**, 161–169 doi:10.1016/0092-8674(82)90100-3
- 57 Krisher, R.L. and Prather, R.S. (2012) A role for the Warburg effect in preimplantation embryo development: metabolic modification to support rapid cell proliferation. *Mol. Reprod. Dev.* **79**, 311–320 doi:10.1002/mrd.22037
- 58 Vander Heiden, M.G., Cantley, L.C. and Thompson, C.B. (2009) Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**, 1029–1033 doi:10.1126/science.1160809
- 59 Ward, P.S. and Thompson, C.B. (2012) Metabolic reprogramming: a cancer hallmark even Warburg did not anticipate. *Cancer Cell* **21**, 297–308 doi:10.1016/j.ccr.2012.02.014
- 60 Singh, K.K., Ayyasamy, V., Owens, K.M., Koul, M.S. and Vujcic, M. (2009) Mutations in mitochondrial DNA polymerase- γ promote breast tumorigenesis. *J. Hum. Genet.* **54**, 516–524 doi:10.1038/jhg.2009.71
- 61 Guo, J., Zheng, L., Liu, W., Wang, X., Wang, Z., Wang, Z. et al. (2011) Frequent truncating mutation of TFAM induces mitochondrial DNA depletion and apoptotic resistance in microsatellite-unstable colorectal cancer. *Cancer Res.* **71**, 2978–2987 doi:10.1158/0008-5472.CAN-10-3482
- 62 Murphy, M.P. (2009) How mitochondria produce reactive oxygen species. *Biochem. J.* **417**, 1–13 doi:10.1042/BJ20081386

- 63 D'Autreaux, B. and Toledano, M.B. (2007) ROS as signalling molecules: mechanisms that generate specificity in ROS homeostasis. *Nat. Rev. Mol. Cell Biol.* **8**, 813–824 doi:10.1038/nrm2256
- 64 Frank, M., Duvezin-Caubet, S., Koob, S., Occhipinti, A., Jagasia, R., Petcherski, A. et al. (2012) Mitophagy is triggered by mild oxidative stress in a mitochondrial fission dependent manner. *Biochim. Biophys. Acta* **1823**, 2297–2310 doi:10.1016/j.bbamcr.2012.08.007
- 65 Simon, H.-U., Haj-Yehia, A. and Levi-Schaffer, F. (2000) Role of reactive oxygen species (ROS) in apoptosis induction. *Apoptosis* **5**, 415–418 doi:10.1023/A:1009616228304
- 66 Giannoni, E., Buricchi, F., Raugei, G., Ramponi, G. and Chiarugi, P. (2005) Intracellular reactive oxygen species activate Src tyrosine kinase during cell adhesion and anchorage-dependent cell growth. *Mol. Cell. Biol.* **25**, 6391–6403 doi:10.1128/MCB.25.15.6391-6403.2005
- 67 Cooke, M.S., Evans, M.D., Dizdaroglu, M. and Lunec, J. (2003) Oxidative DNA damage: mechanisms, mutation, and disease. *FASEB J.* **17**, 1195–1214 doi:10.1096/fj.02-0752rev
- 68 Takahashi, A., Ohtani, N., Yamakoshi, K., Iida, S.-i., Tahara, H., Nakayama, K. et al. (2006) Mitogenic signalling and the p16INK4a-Rb pathway cooperate to enforce irreversible cellular senescence. *Nat. Cell Biol.* **8**, 1291–1297 doi:10.1038/ncb1491
- 69 Garrido, C., Galluzzi, L., Brunet, M., Puig, P.E., Didelot, C. and Kroemer, G. (2006) Mechanisms of cytochrome c release from mitochondria. *Cell Death Differ.* **13**, 1423–1433 doi:10.1038/sj.cdd.4401950
- 70 St-Pierre, J., Drori, S., Uldry, M., Silvaggi, J.M., Rhee, J., Jäger, S. et al. (2006) Suppression of reactive oxygen species and neurodegeneration by the PGC-1 transcriptional coactivators. *Cell* **127**, 397–408 doi:10.1016/j.cell.2006.09.024
- 71 Faraonio, R., Vergara, P., Di Marzo, D., Pierantoni, M.G., Napolitano, M., Russo, T. et al. (2006) p53 suppresses the Nrf2-dependent transcription of antioxidant response genes. *J. Biol. Chem.* **281**, 39776–39784 doi:10.1074/jbc.M605707200
- 72 Wise, D.R., DeBerardinis, R.J., Mancuso, A., Sayed, N., Zhang, X.Y., Pfeiffer, H.K. et al. (2008) Myc regulates a transcriptional program that stimulates mitochondrial glutaminolysis and leads to glutamine addiction. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 18782–18787 doi:10.1073/pnas.0810199105
- 73 Xie, D., Wu, X., Lan, L., Shangguan, F., Lin, X., Chen, F. et al. (2016) Downregulation of TFAM inhibits the tumorigenesis of non-small cell lung cancer by activating ROS-mediated JNK/p38MAPK signaling and reducing cellular bioenergetics. *Oncotarget* **7**, 11609–11624 doi:10.18632/oncotarget.7018
- 74 Polyak, K., Li, Y., Zhu, H., Lengauer, C., Willson, J.K., Markowitz, S.D. et al. (1998) Somatic mutations of the mitochondrial genome in human colorectal tumours. *Nat. Genet.* **20**, 291–293 doi:10.1038/3108
- 75 Petros, J.A., Baumann, A.K., Ruiz-Pesini, E., Amin, M.B., Sun, C.Q., Hall, J. et al. (2005) mtDNA mutations increase tumorigenicity in prostate cancer. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 719–724 doi:10.1073/pnas.0408894102
- 76 Tseng, L.-M., Yin, P.-H., Yang, C.-W., Tsai, Y.-F., Hsu, C.-Y., Chi, C.-W. and Lee, H.-C. (2011) Somatic mutations of the mitochondrial genome in human breast cancers. *Genes Chromosomes Cancer* **50**, 800–811 doi:10.1002/gcc.20901
- 77 Wang, C.-Y., Wang, H.-W., Yao, Y.-G., Kong, Q.-P. and Zhang, Y.-P. (2007) Somatic mutations of mitochondrial genome in early stage breast cancer. *Int. J. Cancer* **121**, 1253–1256 doi:10.1002/ijc.22822
- 78 Yeung, K.Y., Dickinson, A., Donoghue, J.F., Polekhina, G., White, S.J., Grammatopoulos, D.K. et al. (2014) The identification of mitochondrial DNA variants in glioblastoma multiforme. *Acta Neuropathol. Commun.* **2**, 1 doi:10.1186/2051-5960-2-1
- 79 Attardi, G. and Schatz, G. (1988) Biogenesis of mitochondria. *Annu. Rev. Cell Biol.* **4**, 289–333 doi:10.1146/annurev.cb.04.110188.001445
- 80 Fernandez-Marcos, P.J. and Auwerx, J. (2011) Regulation of PGC-1 α , a nodal regulator of mitochondrial biogenesis. *Am. J. Clin. Nutr.* **93**, 884S–889S doi:10.3945/ajcn.110.001917
- 81 Scarpulla, R.C., Vega, R.B. and Kelly, D.P. (2012) Transcriptional integration of mitochondrial biogenesis. *Trends Endocrinol. Metab.* **23**, 459–466 doi:10.1016/j.tem.2012.06.006
- 82 Wenz, T. (2013) Regulation of mitochondrial biogenesis and PGC-1 α under cellular stress. *Mitochondrion* **13**, 134–142 doi:10.1016/j.mito.2013.01.006
- 83 Pikó, L. and Matsumoto, L. (1976) Number of mitochondria and some properties of mitochondrial DNA in the mouse egg. *Dev. Biol.* **49**, 1–10 doi:10.1016/0012-1606(76)90253-0
- 84 Wilding, M., Dale, B., Marino, M., di Matteo, L., Alviggi, C., Pisaturo, M.L. et al. (2001) Mitochondrial aggregation patterns and activity in human oocytes and preimplantation embryos. *Hum. Reprod.* **16**, 909–917 doi:10.1093/humrep/16.5.909
- 85 Aquilano, K., Vigilanza, P., Baldelli, S., Pagliei, B., Rotilio, G. and Ciriolo, M.R. (2010) Peroxisome proliferator-activated receptor γ co-activator 1 α (PGC-1 α) and sirtuin 1 (SIRT1) reside in mitochondria: possible direct function in mitochondrial biogenesis. *J. Biol. Chem.* **285**, 21590–21599 doi:10.1074/jbc.M109.070169
- 86 Kong, X., Wang, R., Xue, Y., Liu, X., Zhang, H., Chen, Y. et al. (2010) Sirtuin 3, a new target of PGC-1 α , plays an important role in the suppression of ROS and mitochondrial biogenesis. *PLoS One* **5**, e11707 doi:10.1371/journal.pone.0011707
- 87 Sahin, E., Colla, S., Liesa, M., Moslehi, J., Müller, F.L., Guo, M. et al. (2011) Telomere dysfunction induces metabolic and mitochondrial compromise. *Nature* **470**, 359–365 doi:10.1038/nature09787
- 88 Kim, J., Lee, J.H. and Iyer, V.R. (2008) Global identification of Myc target genes reveals its direct role in mitochondrial biogenesis and its E-box usage in vivo. *PLoS One* **3**, e1798 doi:10.1371/journal.pone.0001798
- 89 Li, F., Wang, Y., Zeller, K.I., Potter, J.J., Wonsey, D.R., O'Donnell, K.A. et al. (2005) Myc stimulates nuclearly encoded mitochondrial genes and mitochondrial biogenesis. *Mol. Cell. Biol.* **25**, 6225–6234 doi:10.1128/MCB.25.14.6225-6234.2005
- 90 Morrish, F. and Hockenbery, D. (2014) MYC and mitochondrial biogenesis. *Cold Spring Harb. Perspect. Med.* **4**, a014225 doi:10.1101/cshperspect.a014225
- 91 Carbognin, E., Betto, R.M., Soriano, M.E., Smith, A.G. and Martello, G. (2016) Stat3 promotes mitochondrial transcription and oxidative respiration during maintenance and induction of naive pluripotency. *EMBO J.* **35**, 618–634 doi:10.15252/embj.201592629
- 92 O'Hagan, K.A., Cocchiaglia, S., Zhdanov, A.V., Tambuwala, M.M., Cummins, E.P., Monfared, M. et al. (2009) PGC-1 α is coupled to HIF-1 α -dependent gene expression by increasing mitochondrial oxygen consumption in skeletal muscle cells. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 2188–2193 doi:10.1073/pnas.0808801106
- 93 Pawlus, M.R., Wang, L. and Hu, C.J. (2014) STAT3 and HIF1 α cooperatively activate HIF1 target genes in MDA-MB-231 and RCC4 cells. *Oncogene* **33**, 1670–1679 doi:10.1038/nc.2013.115

- 94 Twigg, G., Elorza, A., Molina, A.J., Mohamed, H., Wikstrom, J.D., Walzer, G. et al. (2008) Fission and selective fusion govern mitochondrial segregation and elimination by autophagy. *EMBO J.* **27**, 433–446 doi:10.1038/sj.emboj.7601963
- 95 Youle, R.J. and Narendra, D.P. (2011) Mechanisms of mitophagy. *Nat. Rev. Mol. Cell Biol.* **12**, 9–14 doi:10.1038/nrm3028
- 96 Zhang, J. and Ney, P.A. (2009) Role of BNIP3 and NIX in cell death, autophagy, and mitophagy. *Cell Death Differ.* **16**, 939–946 doi:10.1038/cdd.2009.16
- 97 Cesari, R., Martin, E.S., Calin, G.A., Pentimalli, F., Bichi, R., McAdams, H. et al. (2003) Parkin, a gene implicated in autosomal recessive juvenile parkinsonism, is a candidate tumor suppressor gene on chromosome 6q25-q27. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5956–5961 doi:10.1073/pnas.0931262100
- 98 Todd, L.R., Damin, M.N., Gomathinayagam, R., Horn, S.R., Means, A.R. and Sankar, U. (2010) Growth factor erv1-like modulates Drp1 to preserve mitochondrial dynamics and function in mouse embryonic stem cells. *Mol. Biol. Cell* **21**, 1225–1236 doi:10.1091/mbc.E09-11-0937
- 99 Otera, H. and Mihara, K. (2011) Molecular mechanisms and physiologic functions of mitochondrial dynamics. *J. Biochem.* **149**, 241–251 doi:10.1093/jb/mvr002
- 100 Detmer, S.A. and Chan, D.C. (2007) Functions and dysfunctions of mitochondrial dynamics. *Nat. Rev. Mol. Cell Biol.* **8**, 870–879 doi:10.1038/nrm2275
- 101 Autret, A. and Martin, S.J. (2009) Emerging role for members of the Bcl-2 family in mitochondrial morphogenesis. *Mol. Cell* **36**, 355–363 doi:10.1016/j.molcel.2009.10.011
- 102 Suen, D.F., Norris, K.L. and Youle, R.J. (2008) Mitochondrial dynamics and apoptosis. *Genes Dev.* **22**, 1577–1590 doi:10.1101/gad.1658508
- 103 Hotchkiss, R.D. (1948) The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. *J. Biol. Chem.* **175**, 315–332 PMID: 18873306
- 104 Lister, R., O'Malley, R.C., Tonti-Filippini, J., Gregory, B.D., Berry, C.C., Millar, A.H. and Ecker, J.R. (2008) Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* **133**, 523–536 doi:10.1016/j.cell.2008.03.029
- 105 Lister, R., Mukamel, E.A., Nery, J.R., Urich, M., Puddifoot, C.A., Johnson, N.D. et al. (2013) Global epigenomic reconfiguration during mammalian brain development. *Science* **341**, 1237905 doi:10.1126/science.1237905
- 106 Ramsahoye, B.H., Biniszkievicz, D., Lyko, F., Clark, V., Bird, A.P. and Jaenisch, R. (2000) Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5237–5242 doi:10.1073/pnas.97.10.5237
- 107 Cheng, X. (1995) DNA modification by methyltransferases. *Curr. Opin. Struct. Biol.* **5**, 4–10 doi:10.1016/0959-440X(95)80003-J
- 108 Okano, M., Bell, D.W., Haber, D.A. and Li, E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell* **99**, 247–257 doi:10.1016/S0092-8674(00)81656-6
- 109 Shen, L. and Zhang, Y. (2012) Enzymatic analysis of Tet proteins: key enzymes in the metabolism of DNA methylation. *Methods Enzymol.* **512**, 93–105 doi:10.1016/B978-0-12-391940-3.00005-6
- 110 Kohli, R.M. and Zhang, Y. (2013) TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* **502**, 472–479 doi:10.1038/nature12750
- 111 Gardiner-Garden, M. and Frommer, M. (1987) CpG islands in vertebrate genomes. *J. Mol. Biol.* **196**, 261–282 doi:10.1016/0022-2836(87)90689-9
- 112 Deaton, A.M. and Bird, A. (2011) CpG islands and the regulation of transcription. *Genes Dev.* **21**, 1074–1086 doi:10.1101/gad.2037511
- 113 Jones, P.A. and Taylor, S.M. (1980) Cellular differentiation, cytidine analogs and DNA methylation. *Cell* **20**, 85–93 doi:10.1016/0092-8674(80)90237-8
- 114 Kelly, T.K., De Carvalho, D.D. and Jones, P.A. (2010) Epigenetic modifications as therapeutic targets. *Nat. Biotechnol.* **28**, 1069–1078 doi:10.1038/nbt.1678
- 115 Blaschke, K., Ebata, K.T., Karimi, M.M., Zepeda-Martínez, J.A., Goyal, P., Mahapatra, S. et al. (2013) Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* **500**, 222–226 doi:10.1038/nature12362
- 116 Eckhardt, F., Lewin, J., Cortese, R., Rakyan, V.K., Attwood, J., Burger, M. et al. (2006) DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat. Genet.* **38**, 1378–1385 doi:10.1038/ng1909
- 117 Bird, A. (2002) DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21 doi:10.1101/gad.947102
- 118 Feil, R. and Khosla, S. (1999) Genomic imprinting in mammals: an interplay between chromatin and DNA methylation? *Trends Genet.* **15**, 431–435 doi:10.1016/S0168-9525(99)01822-3
- 119 Riggs, A.D. (1975) X inactivation, differentiation, and DNA methylation. *Cytogenet. Genome Res.* **14**, 9–25 doi:10.1159/000130315
- 120 Fatemi, M., Pao, M.M., Jeong, S., Gal-Yam, E.N., Egger, G., Weisenberger, D.J. et al. (2005) Footprinting of mammalian promoters: use of a CpG DNA methyltransferase revealing nucleosome positions at a single molecule level. *Nucleic Acids Res.* **33**, e176 doi:10.1093/nar/gni180
- 121 Esteller, M. (2002) CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene* **21**, 5427–5440 doi:10.1038/sj.onc.1205600
- 122 Maunakea, A.K., Chepelev, I., Zhao, K., and Bruneau, B. (2010) Epigenome mapping in normal and disease states. *Circ. Res.* **107**, 327–339 doi:10.1161/CIRCRESAHA.110.222463
- 123 Barrès, R., Osler, M.E., Yan, J., Rune, A., Fritz, T., Caidahl, K. et al. (2009) Non-CpG methylation of the PGC-1 α promoter through DNMT3B controls mitochondrial density. *Cell Metab.* **10**, 189–198 doi:10.1016/j.cmet.2009.07.011
- 124 Gemma, C., Sookoian, S., Dieuzeide, G., García, S.I., Gianotti, T.F., González, C.D. and Pirola, C.J. (2010) Methylation of TFAM gene promoter in peripheral White blood cells is associated with insulin resistance in adolescents. *Mol. Genet. Metab.* **100**, 83–87 doi:10.1016/j.ymgme.2010.02.004
- 125 Aran, D., Toperoff, G., Rosenberg, M. and Hellman, A. (2011) Replication timing-related and gene body-specific methylation of active human genes. *Hum. Mol. Genet.* **20**, 670–680 doi:10.1093/hmg/ddq513
- 126 Ball, M.P., Li, J.B., Gao, Y., Lee, J.-H., LeProust, E.M., Park, I.-H. et al. (2009) Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat. Biotechnol.* **27**, 361–368 doi:10.1038/nbt.1533
- 127 Jones, P.A. (1999) The DNA methylation paradox. *Trends Genet.* **15**, 34–37 doi:10.1016/S0168-9525(98)01636-9
- 128 Walsh, C.P., Chaillet, J.R. and Bestor, T.H. (1998) Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat. Genet.* **20**, 116–117 doi:10.1038/2413
- 129 Jjingo, D., Conley, A.B., Yi, S.V., Lunyak, V.V. and Jordan, I.K. (2012) On the presence and role of human gene-body DNA methylation. *Oncotarget* **3**, 462–474 doi:10.18632/oncotarget.497

- 130 Laurent, L., Wong, E., Li, G., Huynh, T., Tsirigos, A., Ong, C.T. et al. (2010) Dynamic changes in the human methylome during differentiation. *Genome Res.* **20**, 320–331 doi:10.1101/gr.101907.109
- 131 Guo, F., Li, X., Liang, D., Li, T., Zhu, P., Guo, H. et al. (2014) Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. *Cell Stem Cell* **15**, 447–458 doi:10.1016/j.stem.2014.08.003
- 132 Santos, F., Hendrich, B., Reik, W. and Dean, W. (2002) Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev. Biol.* **241**, 172–182 doi:10.1006/dbio.2001.0501
- 133 Brunner, A.L., Johnson, D.S., Kim, S.W., Valouev, A., Reddy, T.E., Neff, N.F. et al. (2009) Distinct DNA methylation patterns characterize differentiated human embryonic stem cells and developing human fetal liver. *Genome Res.* **19**, 1044–1056 doi:10.1101/gr.088773.108
- 134 Fouse, S.D., Shen, Y., Pellegrini, M., Cole, S., Meissner, A., Van Neste, L. et al. (2008) Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. *Cell Stem Cell* **2**, 160–169 doi:10.1016/j.stem.2007.12.011
- 135 Lee, W.T.Y., Cain, J.E., Cuddihy, A., Johnson, J., Dickinson, A., Yeung, K.Y. et al. (2016) Mitochondrial DNA plasticity is an essential inducer of tumorigenesis. *Cell Death Discov.* **2**, 16016 doi:10.1038/cddiscovery.2016.16
- 136 Kelly, R.D.W., Rodda, A.E., Dickinson, A., Mahmud, A., Nefzger, C.M., Lee, W. et al. (2013) Mitochondrial DNA haplotypes define gene expression patterns in pluripotent and differentiating embryonic stem cells. *Stem Cells* **31**, 703–716 doi:10.1002/stem.1313

Overall Hypothesis

In summary, mtDNA is critical for mitochondrial function including supplying energy for cellular metabolism. In cancers, changes to mtDNA copy number have been identified and related to the nuclear-encoded mtDNA replication factor, POLG, and its associated aberrant DNA methylation patterns. As a hallmark of cancer, aberrant DNA methylation occurs extensively across the genome. It is of great importance to identify the other mtDNA replication factors that are epigenetically regulated and to investigate their impact on regulating mtDNA replication. Furthermore, mtDNA copy number varies between cell types based on their energy demands. During differentiation, cells replicate their mtDNA differently in a tissue-specific manner. However, differentiation in tumour cells is restricted. Based on these previous findings, it is important to further investigate how cells regulate mtDNA replication under the influence of DNA methylation during cell differentiation and how the mechanism is altered in tumour cells.

The overall hypothesis of my PhD thesis is that the epigenetic control of mtDNA copy number in tumorigenesis and differentiation relies on the bi-directional communication between the nuclear and mitochondrial genomes through DNA methylation.

Aims

1. To identify through global DNA demethylation patterns, the key mtDNA replication factors that are differentially DNA methylated and influence mtDNA copy number.

2. To determine how mtDNA copy number influences the DNA methylation profiles of the master regulators of mtDNA replication and tumorigenesis in end point tumours.
3. To characterize the plasticity of mtDNA methylation resulting from different combinations of the nuclear and mitochondrial genomes and at different stages of tumorigenesis.

CHAPTER 2

General Methods

Publication attached.

Other methods will be included in the chapters they specifically apply to.

Chapter 10

Analysis of Mitochondrial DNA Copy Number and Its Regulation Through DNA Methylation of *POLGA*

Xin Sun, William Lee, Vijesh Vaghjiani, and Justin C. St. John

Abstract

Replication of mitochondrial DNA (mtDNA) is important for ensuring that cells have sufficient mtDNA copy number to meet their specific requirements for the generation of cellular energy through oxidative phosphorylation. A number of transcription and replication factors are required for this process, with a key factor being the nuclear-encoded mtDNA-specific DNA polymerase γ . DNA polymerase γ has a catalytic subunit (POLGA), whose gene has been shown to be DNA methylated at exon 2. This methylation is considered to be one of the key mechanisms that regulate mtDNA copy number. These findings have made it of great importance to establish optimal methods for investigating the effects of DNA methylation on mtDNA replication. Here, we provide methods to determine the extent of DNA methylation at exon 2 of *POLGA* as well as other gene targets of interest. We also show how mtDNA copy number is assessed and, from these two outputs, define the efficiency of mtDNA replication by calculating the mtDNA-replicative efficiency index.

Key words mtDNA copy number, MeDIP, DNA methylation, POLGA

1 Introduction

Replication of mitochondrial DNA (mtDNA) is key to developmental outcome. To this extent, it has been demonstrated that mtDNA copy number is strictly regulated during the process of development, whereby copy number decreases significantly from the numbers present in the fertilized oocyte as it progresses through pre-implantation development [1]. This continues until the blastocyst stage, the final stage of preimplantation development, when the trophectoderm and the inner cell mass are established. These two cell populations then give rise to the placenta and the embryo proper, respectively [2]. The inner cell mass is also the source of embryonic stem cells, which have the potential to differentiate into all cell types of the body [3]. In the blastocyst, mtDNA replication is initiated in the trophectoderm. However, the cells of the inner

cell mass do not replicate mtDNA. This remains the case until later stages of development when naïve embryonic cells commit to specific lineages [4]. Consequently, embryonic stem cells are excellent models to study mtDNA replication.

Regulation of mtDNA copy number is primarily under the control of the nuclear genome [5, 6]. Indeed, the factors that regulate mitochondrial transcription and replication are encoded by the nuclear genome [7]. Recently, it has been shown that one of these key factors, the mtDNA-specific polymerase, DNA polymerase γ (POLG), undergoes DNA methylation within the second exon of the gene encoding its catalytic subunit (*POLGA*) [8, 9]. This finding has been shown both in mouse [8] and human [10] embryonic and differentiating stem cells, and in tissues. Indeed, continuous replication of mtDNA is indicative of the ability of the naïve cell to differentiate and to generate ATP through oxidative phosphorylation [10], which takes place within the electron transfer chain. This is also most likely to be the case across a large number of species. DNA methylation at exon 2 has also been shown to be a regulator of mtDNA copy number in tumor cells [10].

DNA methylation can be determined through either bisulfite sequencing or immunoprecipitation of DNA using an antibody capture approach, known as *m*ethylated DNA Immunoprecipitation (MeDIP) [8, 10]. The MeDIP approach enables *de novo* DNA methylation and demethylation to be assessed on the same sample using two antibodies. The first antibody captures 5-methylcytosine (5mC), which is indicative of *de novo* DNA methylation. The second antibody, 5-hydroxy methylcytosine (5hmC), enables *de novo* demethylated DNA to be captured. This approach has advantages over bisulfite sequencing, which does not distinguish between *de novo* DNA methylation and demethylation [10]. We have used antibodies to 5mC and 5hmC to assess the levels of DNA methylation at exon 2 of *POLGA* and shown that the ratio of the differences between the two populations of captured mtDNA can generate a mitochondrial replication efficiency index, when expressed in combination with mtDNA copy number [10]. Furthermore, it is worth mentioning that there is emerging evidence showing that mtDNA could be DNA methylated [11]. Even though this appears controversial, it is interesting to consider this point when investigating the regulation of the mitochondrial genome.

The assays we show here demonstrate how to capture *de novo* methylated DNA at exon 2 of *POLGA* and also from regions within the mitochondrial genome. We further show how to quantitatively assess the amount of methylated and non-methylated DNA as well as mtDNA copy number. From this, we demonstrate how to determine the mtDNA-replicative efficiency index so that a variety of cells from different stages of development and disease states can be accurately compared to determine whether mtDNA

copy number and the regulation of mtDNA replication through DNA methylation at exon 2 at *POLGA* are attributable to the disease being investigated. As it becomes evident which other mtDNA replication factors undergo DNA methylation, these assays can be easily applied to the affected gene regions.

2 Materials

2.1 Immuno-precipitation of Methylated DNA (MeDIP)

1. DNA sonication.
 - (a) Covaris Adaptive Focused Acoustics (AFA™) S220 system.
 - (b) Covaris milliTUBE.
2. DNA electrophoresis.
 - (a) Agarose in 1× Tris-EDTA (TE) buffer (pH 8.0, 10 mM Tris, 1 mM EDTA).
 - (b) Sybr safe DNA gel stain.
3. 10× IP buffer: 100 mM sodium phosphate pH 7.0, 1.4 M NaCl, 0.5 % Triton X-100.
4. Immunoprecipitation.
 - (a) Anti-5mC antibody.
 - (b) Anti-5hmC antibody.
 - (c) Dynabeads® Protein G.
5. Suspension mixer wheel.
6. PBS/BSA buffer: 0.1 % bovine serum albumin (BSA) in 1× phosphate buffered saline (PBS).
7. Magnetic rack.
8. 10× Proteinase K digestion buffer: 50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1.0 % Sodium dodecyl sulfate (SDS).
9. Proteinase K (20 mg/mL).
10. Thermo-shaker.
11. PCR purification kit.
12. Autoclaved Milli-Q water.
13. 1.7 mL sterile microcentrifuge tubes.

2.2 Quantifying MeDIP Products Using Real-Time PCR

1. SensiMix Hi-Rox real-time PCR mastermix.
2. Gene-specific primers.
3. Real-Time PCR instrument (*see Note 1*).
4. 0.1 mL sterile strip tubes and caps.

2.3 Determining mtDNA Copy Number

1. PCR reagents:
 - (a) 10× NH₄ reaction buffer.
 - (b) 50 mM MgCl₂ solution.
 - (c) 50 mM dNTP mix dilution.
 - (d) *Taq* polymerase.
2. Gene specific primers.
3. 0.2 mL sterile polypropylene PCR tubes.
4. Gel extraction kit.
5. SensiMix Hi-Rox real-time PCR mastermix.
6. Real-Time PCR machine.
7. 0.1 mL sterile strip tubes and caps.

3 Methods

3.1 MeDIP

1. Pre-cool and degas the Covaris Adaptive Focused Acoustics (AFATM) S220 system (*see Note 2*).
2. In a Covaris milliTUBE, aliquot 5–10 µg of genomic DNA and adjust to 100 µL with autoclaved Milli-Q water. Add 100 µL of 2× TE buffer. Sonicate the DNA sample in the Covaris Adaptive Focused Acoustics (AFATM) S220 system for 195 s to generate fragments sized between 200 and 1000 bp. The sonication setting is 5–10 °C for the temperature range, 105.0 for peak power, 5.0 for duty factor and 200 for cycles/burst.
3. To confirm efficient sonication, check the DNA by electrophoresis on a 1.5 % agarose gel. The DNA should appear as a smear between 200 and 1000 bp. If a smear is not present, repeat the sonication step.
4. In a new tube, add 3 µg of sonicated DNA and adjust to a final volume of 450 µL with autoclaved Milli-Q water.
5. Denature DNA samples by incubating at 95 °C for 10 min, then immediately place on ice for 5 min.
6. Add 50 µL of 10× IP buffer and mix well by pipetting.
7. Add 3 µg of anti-5mC or anti-5hmC antibody into each sample (1 µg per 1 µg DNA).
8. Incubate samples at 4 °C for 2 h by rotation.
9. Vortex Dynabeads[®] Protein G bottle thoroughly and immediately aliquot 30 µL per sample (10 µL per 1 µg DNA) into a new tube.
10. Pre-wash the beads with 1 mL PBS/BSA buffer for 5 min at 4 °C by rotation. Place the tube in a magnetic rack and remove the surrounding fluid and repeat the wash two more times. Each time, place the tube in the magnetic rack to remove the fluid from the beads.

11. Resuspend the beads in 30 μ L of 1 \times IP buffer.
12. Add beads into each sample and incubate at 4 $^{\circ}$ C for 16 h by rotation (*see Note 3*).
13. Place samples in a magnetic rack for 1 min. Remove all fluid.
14. Wash the beads with 1 mL of 1 \times IP buffer for 5 min at 4 $^{\circ}$ C by rotation. Repeat the wash two more times. Place the samples in a magnetic rack. Remove all fluid.
15. Resuspend beads in 250 μ L of proteinase K digestion buffer and add 10 μ L of 20 mg/mL proteinase K.
16. Incubate at 50 $^{\circ}$ C for 3 h on a thermo-shaker.
17. On a magnetic rack, transfer the supernatant into a new tube.
18. Extract DNA using a PCR purification kit. Elute purified DNA in 50 μ L of autoclaved Milli-Q water (*see Note 4*).

3.2 Quantifying MeDIP Products by Real-Time PCR

1. Real-time PCR is applied to quantify the levels of enrichment of 5mC and 5hmC in the gene of interest. This is achieved by comparing levels of 5mC and 5hmC in the gene of interest with a reference DNA sample (Input), which is prepared from the same DNA source but it only undergoes the DNA sonication and PCR purification steps in the MeDIP procedure (*see above*).
2. Run Input, 5mC and 5hmC samples for each DNA sample. Prepare sufficient sample to run reactions in triplicate. Prepare two reactions as non-template control (NTC).
3. Prepare reactions for real-time PCR in a total volume of 20 μ L, containing: 10 μ L of 2 \times SensiMix SYBR HI-ROX mastermix, 1 μ L (5 μ M) of each forward and reverse primer for the gene of interest (primer information for several mtDNA genes and exon 2 of *POLGA* shown in Table 1), 6 μ L autoclaved Milli-Q water, 2 μ L of purified MeDIP sample or H₂O as NTC.
4. Run each reaction on a real-time PCR instrument in triplicate. The reaction profile settings are: 95 $^{\circ}$ C for 10 min, followed by 50 cycles of 95 $^{\circ}$ C for 15 s, followed by gene-specific annealing temperature (Table 1) for 15 s with an extension for 15 s at 72 $^{\circ}$ C (cycling A) and for 15 s at the product-specific melting temperature (cycling B; Table 1). Data are acquired from the Cycling B-FAM/Sybr channel (*see Note 5*).
5. Adjust the cycle threshold (CT) to obtain the CT values for each sample. The enrichment levels for 5mC or 5hmC are calculated using the following formula:

$$\text{Enrichment of 5mC} = 2^{(CT_{\text{Input}} - CT_{5\text{mC}})}$$

$$\text{Enrichment of 5hmC} = 2^{(CT_{\text{Input}} - CT_{5\text{hmC}})}$$

Table 1
qPCR primers for target genes

Human gene	Forward primer (5'-3')	Reverse primer (5'-3')	Product size	Annealing Tm	Cycling B. Tm
<i>β-globin</i>	CAACTTCATCCACGTTTCACC	GAAGAGCCCAAGGACAGGTAC	268	57	80
<i>mtDNA</i>	CGAAAGGACAAAGAGAGAAATAAGG	CTGTAAAGTTTAAAGTTTATGCG	152	53	78
<i>POLGA</i> exon 2	CAGACCTCCACGTCGAACAC	GACAACTGGACCAGCACTT	209	59	83

3.3 Determining mtDNA Copy Number by Real-Time PCR

6. To investigate if the mitochondrial genome is DNA methylated, regions of the mitochondrial genome can be quantified using MeDIP products (Table 2).
1. Two standard curves need to be generated: one for nuclear DNA (*β-globin*) and one for mtDNA. All steps below should be applied to both nuclear DNA (*β-globin*) and mtDNA.
 2. Adjust the concentration of total DNA templates to 100 ng/μL. For amplification of *β-globin* and mtDNA templates using conventional PCR, prepare the reaction mixture to a total volume of 50 μL containing: 5 μL of 10× NH₄ reaction buffer, 1.5 μL of 50 mM MgCl₂ solution, 0.5 μL 50 mM dNTP mix, 0.5 μL of *Taq* polymerase, 1 μL (25 μM) of each forward and reverse primer for *β-globin* or mtDNA (primer information shown in Table 1), 2 μL (200 ng) of total DNA and 38.5 μL autoclaved Milli-Q water.
 3. Reaction profile settings are 95 °C for 10 min, followed by 35 cycles of 95 °C for 30 s, primer-specific annealing temperature (Table 1) for 30 s and 72 °C for 30 s. The reaction is completed by an extension phase at 72 °C for 10 min.
 4. To confirm DNA templates have been amplified, run PCR samples on a 2 % agarose gel.
 5. Purify the targeted DNA by cutting the targeted band from the gel and extracting DNA with a gel extraction kit. Determine the amount of purified DNA template using a spectrophotometer.
 6. Adjust the concentration of purified DNA templates to 1 ng/μL.
 7. Prepare tenfold serial dilutions of the purified DNA from 10⁻¹ to 10⁻⁸ ng/μL as the templates for the standards.
 8. Adjust the concentration of the unknown DNA samples to 10 ng/μL.
 9. Prepare quantitative PCR reactions to a total volume of 20 μL containing: 10 μL of 2× SensiMix SYBR HI-ROX mastermix, 1 μL (5 μM) of each forward and reverse primer for *β-globin* or mtDNA (primer information shown in Table 1), 2 μL of each standard template or unknown sample and 6 μL autoclaved Milli-Q water.
 10. Reactions are performed in a real-time PCR machine. Reaction profile settings are: 95 °C for 10 min followed by 50 cycles of 95 °C for 15 s, gene-specific annealing temperature (Table 1) for 15 s and extension for 15 s at 72 °C (cycling A) and for 15 s at the product specific temperature (cycling B; Table 1). Extension and data acquisition are performed on Cycling B-FAM/Sybr channel.

Table 2
Primers for mitochondrial genes

Human gene	Forward primer (5'-3')	Reverse primer (5'-3')	Product size	Annealing Tm
<i>MT-ND1</i>	TCTCACCATCGCTCTTCTAC	GGTTGGTCTCTGCTAGTGTG	350	57
<i>MT-CTB</i>	ATGACCCCCAATACGCAAAACT	GGGAGGACATAGCCTATGAA	401	54
<i>MT-CO2</i>	CTGCTTCCTAGTCCTGTATG	GTCGGTGTA CTCTAGGTTC	235	56
<i>MT-ND6</i>	CCGCACCAATAGGATCCTCCCGA	GCA TGGGGGTCAGGGGTTGAG	187	63
<i>MT-ATP6</i>	CAGTGATTATAGGCTTTCGCTC	GTGTTGTCGTGCAGGTAGAG	343	57

11. Melt data are acquired from 72 to 98 °C with a 60 s interval before the first step followed by 4 s for each subsequent step of 1 °C increase. The melt curve determines the specificity of the primers, where only one curve should be observed.
12. A standard curve is produced and indicates the efficiency of quantitative PCR. The efficiency should range between 0.95 and 1.
13. Concentrations of the mtDNA and nuclear DNA-specific products are calculated based on the standard curve.
14. mtDNA Copy number per cell is calculated using the formula shown below (*see Note 6*):

$$\text{mtDNA copy number per cell} = 2 \times N_{\text{mtDNA}} / N_{\beta\text{-globin}}$$
 As chromosomal DNA is diploid in cells, mtDNA copy number is doubled.
 N_{mtDNA} and $N_{\beta\text{-globin}}$ are determined by the formula:

$$N = ([\text{ng} / \mu\text{L}] \times 6.023 \times 10^{14}) / (\text{product size in bp} \times 660)$$

where N is the number of molecules per reaction, $[\text{ng}/\mu\text{L}]$ is the concentration determined by the standard curve computationally, 10^{14} is the conversion of 1 mol to 1 nmol using Avogadro's constant, which states 1 mol = 6.023×10^{23} , 660 is the mean molecular weight of a nucleic acid base pairing in Daltons (Da).

15. Replication efficiency index = mtDNA copy number / (enrichment level of 5mC / enrichment of 5hmC) (*see Note 7*).

4 Notes

1. Our reaction and analyses are conducted on a 72-well Rotorgene-3000 Real-Time PCR machine with Rotor-Gene 6 software. Different instruments can be used according to the prescribed operational manual.
2. For MeDIP, it normally takes 2 h for the sonication system to cool down and degas.
3. For MeDIP, it is important to plan the timeline well in advance because the procedure includes a series of incubation steps, especially a 16 h incubation step. We normally start this at 5 p.m. and terminate it at 9 a.m. the next day.
4. The MeDIP products are single-stranded DNA. For any downstream experiments requiring double-stranded DNA, second-strand synthesis should be performed.
5. Collection of data at cycling B ensures that background noise, such as primer-dimers, are excluded from the analysis.

Table 3
Example of real-time PCR dataset for mtDNA standards and samples

Name	Type	Average Ct	Given conc. (ng/μL)	Calc conc. (ng/μL)
STD1	Standard	6.19	0.1	
STD2	Standard	8.51	0.01	
STD3	Standard	11.73	0.001	
STD4	Standard	14.96	0.0001	
STD5	Standard	18.81	0.00001	
STD6	Standard	22.51	0.000001	
STD7	Standard	25.47	0.0000001	
H ₂ O	NTC		0	
Sample	Unknown	10.25		0.00742

6. An example of how to calculate mtDNA copy number:

From the real-time PCR, the data collected from the Cycling B-FAM/Sybr channel (Table 3) are used to generate a standard curve, i.e., an output graph (Fig. 1). As shown in the output graph, the efficiency and R values met the requirements to qualify as standards. In the dataset, the NTC sample was not detected. Based on the standard curve, the concentration of the sample was given by computational analysis. As shown in Table 3, the mtDNA concentration of the sample is 0.00742 ng/μL. In the same way, the concentration for *β-globin* of the sample can also be determined.

As an example, if the concentration of the *β-globin* gene is 0.000007 ng/μL, then:

$$N_{\text{mtDNA}} = (0.00742 \times 6.023 \times 10^{14}) / (152 \times 660) = 44560413.8$$

$$N_{\beta\text{-globin}} = (0.000007 \times 6.023 \times 10^{14}) / (268 \times 660) = 23835.9$$

$$\text{mtDNA copies per cell} = 2 \times N_{\text{mtDNA}} / N_{\beta\text{-globin}} = 3738.9$$

7. An example of how to calculate the replication efficiency index:

MeDIP experiments will have determined the CT values for 5mC and 5hmC at, for example, 26.4 cycles and 27.2 cycles, respectively and for the Input at 25.3 cycles. Therefore:

$$\text{Enrichment of 5mC} = 2^{(\text{CT}_{\text{Input}} - \text{CT}_{5\text{mC}})} = 2^{(25.3 - 26.4)} = 0.467$$

$$\text{Enrichment of 5hmC} = 2^{(\text{CT}_{\text{Input}} - \text{CT}_{5\text{hmC}})} = 2^{(25.3 - 27.2)} = 0.268$$

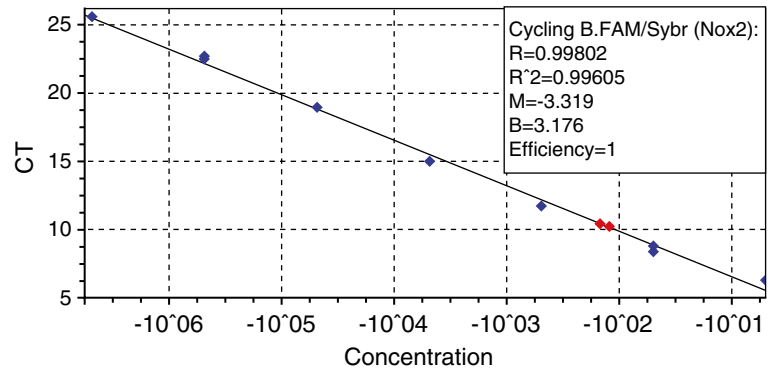


Fig. 1 Example of a standard curve generated for mtDNA copy number from real-time PCR

Also taking the mtDNA copy number determined above, as an example, therefore,

the mtDNA replication efficiency index for a specific cell or sample

$$= \text{mtDNA copy number} / (\text{enrichment level of 5mC} / \text{enrichment of 5hmC})$$

$$= 3738.9 / (0.467 / 0.268) = 2146$$

References

- Spikings EC, Alderson J, St John JC (2007) Regulated mitochondrial DNA replication during oocyte maturation is essential for successful porcine embryonic development. *Biol Reprod* 76:327–335
- Niwa H, Toyooka Y, Shimosato D, Strumpf D, Takahashi K, Yagi R, Rossant J (2005) Interaction between Oct3/4 and Cdx2 determines trophoblast differentiation. *Cell* 123:917–929
- Evans MJ, Kaufman MH (1981) Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292:154–156
- Facucho-Oliveira JM, St John JC (2009) The relationship between pluripotency and mitochondrial DNA proliferation during early embryo development and embryonic stem cell differentiation. *Stem Cell Rev* 5:140–158
- Clayton DA (1992) Transcription and replication of animal mitochondrial DNAs. *Int Rev Cytol* 141:217–232
- Clayton DA (1998) Nuclear-mitochondrial intergenomic communication. *BioFactors* 7:203–205
- Kucej M, Butow RA (2007) Evolutionary tinkering with mitochondrial nucleoids. *Trends Cell Biol* 17:586–592
- Kelly RD, Mahmud A, McKenzie M, Trounce IA, St John JC (2012) Mitochondrial DNA copy number is regulated in a tissue specific manner by DNA methylation of the nuclear-encoded DNA polymerase gamma A. *Nucleic Acids Res* 40:10124–10138
- Oakes CC, La Salle S, Smiraglia DJ, Robaire B, Trasler JM (2007) Developmental acquisition of genome-wide DNA methylation occurs prior to meiosis in male germ cells. *Dev Biol* 307:368–379
- Lee W, Johnson J, Gough DJ, Donoghue J, Cagnone GLM, Vaghjani V, Brown KA, Johns TG, St John JC (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
- Shock LS, Thakkar PV, Peterson EJ, Moran RG, Taylor SM (2011) DNA methyltransferase 1, cytosine methylation, and cytosine hydroxy-methylation in mammalian mitochondria. *Proc Natl Acad Sci U S A* 108:3630–3635

CHAPTER 3

Global DNA methylation synergistically regulates the nuclear and mitochondrial genomes in glioblastoma cells.

Publication attached.

Global DNA methylation synergistically regulates the nuclear and mitochondrial genomes in glioblastoma cells

Xin Sun^{1,2}, Jacqueline Johnson¹ and Justin C. St. John^{1,2,*}

¹Centre for Genetic Diseases, Hudson Institute of Medical Research, 27-31 Wright Street, Clayton, VIC 3168, Australia and ²Department of Molecular and Translational Sciences, Monash University, 27-31 Wright Street, Clayton, VIC 3168, Australia

Received December 15, 2016; Revised March 22, 2018; Editorial Decision April 14, 2018; Accepted April 19, 2018

ABSTRACT

Replication of mitochondrial DNA is strictly regulated during differentiation and development allowing each cell type to acquire its required mtDNA copy number to meet its specific needs for energy. Undifferentiated cells establish the mtDNA set point, which provides low numbers of mtDNA copy but sufficient template for replication once cells commit to specific lineages. However, cancer cells, such as those from the human glioblastoma multiforme cell line, HSR-GBM1, cannot complete differentiation as they fail to enforce the mtDNA set point and are trapped in a ‘pseudo-differentiated’ state. Global DNA methylation is likely to be a major contributing factor, as DNA demethylation treatments promote differentiation of HSR-GBM1 cells. To determine the relationship between DNA methylation and mtDNA copy number in cancer cells, we applied whole genome MeDIP-Seq and RNA-Seq to HSR-GBM1 cells and following their treatment with the DNA demethylation agents 5-azacytidine and vitamin C. We identified key methylated regions modulated by the DNA demethylation agents that also induced synchronous changes to mtDNA copy number and nuclear gene expression. Our findings highlight the control exerted by DNA methylation on the expression of key genes, the regulation of mtDNA copy number and establishment of the mtDNA set point, which collectively contribute to tumorigenesis.

INTRODUCTION

The human mitochondrial genome (mitochondrial DNA, mtDNA) is a circular, double stranded molecule that is ~16.6 kb in size (1). It is essential for the production of energy as it encodes 13 subunits of the electron transfer chain

(ETC), which generates the vast majority of cellular energy through the biochemical process of oxidative phosphorylation (OXPHOS). mtDNA also encodes 22 transfer RNAs (tRNAs) and 2 ribosomal RNAs (rRNAs). The major non-coding region, the D-loop, is the site of interaction for the nuclear-encoded mtDNA transcription and replication factors that bind to the mitochondrial genome to regulate mitochondrial genomic processes (1). For example, this region contains the transcription start sites for the heavy and light strands of the genome, namely the heavy and light strand promoters (HSP1/2 and LSP), and the site for the initiation of heavy strand replication (O_H).

Cells possess multiple copies of mtDNA with mtDNA copy number being cell-type specific. This is achieved during early development when mtDNA replication is strictly regulated in order that a cell acquires the appropriate numbers of mtDNA copy to meet its specific functions when fully differentiated (2,3). They are empowered to do this as they had previously established the mtDNA set point, which is defined as the number of mtDNA copies (~200) that a naïve cell possesses before it initiates the process of differentiation. These copies are then used by undifferentiated cells as the template for mtDNA replication as they differentiate into distinct mature cell types (2–6). Indeed, during differentiation, mtDNA copy number increases in a synchronous manner as cells mature from a naïve state to a fully differentiated state, which is modulated by the expression of genes that are the master regulators of differentiation. They then express genes associated with terminal differentiation and possess their required numbers of mtDNA copy (4–6). However, a number of cancer cell types are unable to expand their mtDNA copy number as they mainly rely on aerobic glycolysis for energy production, which allows for higher rates of cellular proliferation and prevents differentiation from taking place (4). Indeed, cancer cells appear to be trapped in a ‘pseudo-differentiated’ state whereby they are unable to complete differentiation and cannot increase mtDNA copy number (7,8). As a result, they fail to maintain or reinforce the mtDNA set point.

*To whom correspondence should be addressed. Tel: +61 3 8572 2678; Fax: +61 03 9594 7416; Email: Justin.StJohn@hudson.org.au

mtDNA replication is dependent on transcription having first taken place and is driven by a group of nuclear-encoded mitochondrial-specific transcription and replication factors (9,10). There are several key mitochondrial-specific and direct-binding transcription factors including mitochondrial RNA polymerase (POLRMT) (11), mitochondrial transcription factor A (TFAM) (12,13) and mitochondrial transcription factors B1 (TFB1M) and B2 (TFB2M) (14). The initiation of replication requires a short transcript that is specifically transcribed for mtDNA replication (15) and is used by the direct-binding mitochondrial specific DNA Polymerase Gamma (POLG) to drive mtDNA replication (16,17). In the human, this is a heterotrimer enzyme composed of a catalytic subunit (POLGA) and two accessory subunits (POLGB) (18). POLGA (encoded by the gene *POLG*) acts as the DNA polymerase (19) whereas POLGB (encoded by the gene *POLG2*) recognizes the initial RNA primer and enhances the enzymatic activity of POLG (9,18,20). This process is supported by Twinkle (TWNK), the DNA helicase that separates the two strands allowing for replication to take place (21,22), and the mitochondrial single-stranded DNA binding protein (SSBP1), which acts to stabilize the transient status of single-stranded DNA. Furthermore, the double-stranded circular structure of the mitochondrial genome ensures the maintenance of DNA topology during mtDNA replication by allowing the two strands of the circular genome to unwind or rotate around each other. The mitochondrial topoisomerase, which is encoded by the *TOP1MT* gene, can facilitate this by transiently breaking one strand of mtDNA and religating the strand after having passed through the other strand (23).

A group of indirect factors also play vital roles in regulating mtDNA transcription and replication even though they do not directly bind to mtDNA as is the case for the direct-binding factors. The best-characterized regulators are the nuclear respiratory factors (NRF1 and NRF2), the peroxisome proliferator-activated receptor γ coactivator-1 family (PGC1 α), the Sirtuin family, especially Sirtuin 1 and Sirtuin 3, the tumor suppressor p53 and signal transducer and activator of transcription 3 (STAT3). NRFs interact with PGC1 α cofactors to regulate the expression of *TFB1M* and *TFB2M* by binding to their promoter regions (24). In addition, Sirtuin1 and 3, NAD⁺-dependent deacetylases encoded by the *SIRT1/3* genes, are co-factors with PGC1 α that co-localise with TFAM inside mitochondria (25,26). Given that thyroid hormone T3 is essential for mitochondrial biogenesis (27), it has been shown that T3 binding to T3 receptors affects mtDNA expression by either binding to mtDNA directly or interfering with the expression of the transcription factors in the nucleus, such as NRF1 and PGC1 α (28–31).

We have previously shown that mtDNA copy number is associated with DNA methylation at exon 2 of *POLG* (4,32). It appears that tumor cells are extensively DNA methylated, which normally occurs at CpG dinucleotides, and converts the cytosine into 5-methylcytosine (5mC) (33,34). The glioblastoma multiforme HSR-GBM1 cell line is a high-grade malignant glioblastoma cell line (35). It is DNA hypermethylated, which is associated with oncomutations of the isocitrate dehydrogenase (IDH1/2) in the

citric acid cycle (33,36–39). The hypermethylated genome, which leads to and supports its tumorigenic gene profile (37–41), can be reversed through the use of DNA demethylation agents such as 5-azacytidine (5Aza) and vitamin C (VitC) (42,43). 5Aza is a chemical analogue of cytosine that inhibits DNA methylation at low concentrations, and has been applied in cancer therapies (43,44). It blocks the DNA methyltransferase, DNMT1, which causes global DNA demethylation (45). VitC can also induce DNA demethylation by enhancing the activity of the TET1 enzyme and, thus, improves the conversion of 5mC to the DNA demethylated state of 5-hydroxy-methylcytosine (5hmC) (42).

The demethylation of HSR-GBM1 cells results in significant increases in mtDNA copy number, as they are able to modulate mtDNA copy number in synchrony with changes in chromosomal gene expression patterns during differentiation (4,6). Indeed, during development, the DNA methylation status of the genome undergoes significant changes to firstly erase the parental profiles by ten-eleven translocation methylcytosine dioxygenases, the TET family, and reestablish their own profiles through the DNA methyltransferase (DNMT) family of genes during differentiation (46,47). This highlights the cooperation that exists between the two genomes and the importance of the collective regulation of DNA methylation and mtDNA copy number in tumorigenesis and differentiation.

To further investigate the impact of DNA methylation on regulating the mtDNA-specific transcription and replication factors, we have used methylated DNA immunoprecipitation (MeDIP) in combination with whole-genome sequencing (MeDIP-Seq), whole genome RNA-sequencing (RNA-Seq), and a custom designed high-throughput Fluidigm real-time PCR array on the well characterised HSR-GBM1 tumor cell line and compared its DNA methylation profiles to its demethylated state when induced by 5Aza and VitC. Our results show that a number of tumor-specific genes and several key mtDNA transcription and replication factors are modulated by global DNA demethylation.

MATERIALS AND METHODS

Cell culture

HSR-GBM1 cells were cultured as neurospheres on ultra-low attachment plates (Sigma-Aldrich, MO, USA) in Dulbecco's Modified Eagle Medium /Nutrient Mixture (DMEM/F-12) Media (Thermo Fisher Scientific, MA, USA) containing 2% StemPro neural supplement (Thermo Fisher Scientific), 20 ng/ml basic fibroblast growth factor (bFGF; Merck Millipore, MO, USA) and 20 ng/ml epidermal growth factor (EGF; Merck Millipore) at 37°C, 5% CO₂ and 95% humidity.

DNA demethylation assay

To induce DNA demethylation, HSR-GBM1 cells were cultured in the presence of 0.5 μ M of 5Aza (Sigma-Aldrich) for 48 h or 100 μ g/ml of VitC (L-ascorbic acid, Sigma-Aldrich) for 72 h. Culture media was changed every 24 h. Cells were harvested after the treatments for downstream experiments or stored as cell pellets at -80°C .

DNA and RNA extraction

Total genomic DNA and RNA were extracted from cultured cell pellets using the ISOLATE II Genomic DNA Kit (Bioline, London, UK) and RNeasy Mini Kit (Qiagen, CA, USA), respectively, according to the manufacturer's protocols with minor modifications. The DNA samples were treated with 3 μ l of RNase solution (Qiagen) and Proteinase K Solution (20 μ g/ μ l; Qiagen) at 65°C for 10 min and the RNA samples were treated with DNase I (3 kunitz units/ μ l; Qiagen) for 20 min.

Determination of mtDNA copy number per cell

mtDNA copy number per cell was determined using an established protocol, as previously described (4). Quantitative real-time PCR (qPCR) was performed on purified total DNA targeting β -globin and mtDNA using a Rotor-Gene 3000 (Corbett Research, Cambridge, UK). mtDNA copy number was calculated using the formula: mtDNA copy number per cell = $2 \times N_{\text{mtDNA}}/N_{\beta\text{-globin}}$, where N_{mtDNA} and $N_{\beta\text{-globin}}$, determined by the formula: $N = (\text{qPCR product concentration} \times 6.023 \times 10^{14})/(\text{qPCR product size in bp} \times 660)$, as described in (32). The concentration of the qPCR product is determined using standard curves. Primer sequences and reaction conditions are listed in Supplementary Table S1. Biological triplicates ($n = 3$) were analysed.

Immunoprecipitation of methylated DNA (MeDIP)

Five microgram of genomic DNA was sheared into 200–1000 bp fragments in a precooled and degassed Covaris Adaptive Focused Acoustics (AFA™) S220 system. The dsDNA was denatured by incubation at 95°C for 10 min and then underwent MeDIP, as described in (48). 1.5 μ g of anti-5mC antibody (Active Motif) were added to 3 μ g of DNA fragments and 20 μ l per sample of Dynabeads® Protein G (Thermo Fisher Scientific) in 500 μ l of 1 \times IP buffer (100 mM sodium phosphate, pH 7.0; 1.4 M NaCl; 0.5% Triton X-100) at 4°C for 16 h under rotation. The beads were washed with 1 ml of 1 \times IP buffer three times and were resuspended in 250 μ l proteinase K digestion buffer (50 mM Tris-HCl, pH 8.0; 10 M EDTA, pH 8.0; 1.0% SDS) with 10 μ l of proteinase K (20 mg/ml; Bioline) at 50°C for 3 h on a thermo-shaker. The supernatant was then collected into a new tube and placed on a magnetic particle concentrator (Thermo Fisher Scientific). DNA was purified from the elutant using the QIAquick PCR Purification Kit (Qiagen). Purified DNA was eluted in 50 μ l of autoclaved Milli-Q H₂O.

MeDIP-Seq

To adapt MeDIP for next generation sequencing, minor changes were made to the MeDIP procedure described above. Two microgram of each genomic DNA sample were sheared into ~500 bp in a precooled and degassed Covaris Adaptive Focused Acoustics (AFA™) S220 system. These fragments were then end repaired, A-tailed, ligated to Illumina adaptors (including the sample specific indexed) and purified without amplification to ensure conservation of methylated CpGs using the Illumina TruSeq DNA

PCR-Free Kit, as described in the manufacturer's Low Sample (LS) protocol (Illumina Protocol part #15036187 Rev A Jan 2013). The dsDNA libraries then followed the procedure, as described above. After MeDIP, dsDNA libraries were generated and amplified for 10 cycles, as per the TruSeq ChIPSeq protocol with standard Illumina P7 and P5 primers (Illumina protocol part #15023092 Rev A August 2012). The resultant libraries were analysed using Qubit (Agilent) and Bioanalyzer (ThermoFisher Scientific) and quantitated by qPCR. The 12 pM library pool was prepared based upon qPCR and used for clustering by the cBot™ Hybridization system (Illumina, CA, USA). Cluster generation was performed according to the cBot™ user guide (part #15006165 Rev K October 2012). Clusters were in the optimal range (795 k/mm²; optimal 750–900 k/mm²). 50-bp single-read sequencing was then performed on the Illumina HiSeq1500 and HiSeq3000 sequencing platforms using the Illumina protocol (15035788 Rev D, April 2014; 15066493 v01, August 2015). The reads were converted to fastq format and adaptors were removed using the bcl2fastq software (Illumina). The number of reads per sample was evenly distributed. In total, >550 million reads passed the filter, and over 95% of the sequencing reads passed the Illumina sequencing quality score of Q30 that was deemed excellent quality for base-calling. Run quality parameters were deemed excellent with the PhiX spike-in having an error rate of < 0.2% (expected < 0.5%) and phasing/prephasing of < 0.08/0.1 (expected 0.2/< 0.4).

Bioinformatics analysis for MeDIP-Seq

Sequencing files in fastq format firstly underwent and passed the pre-alignment quality and remaining adaptor check using the FastQC software (v 0.11.5; downloaded from Babraham Bioinformatics public projects <https://www.bioinformatics.babraham.ac.uk/projects/>). The raw sequences were then mapped to the human reference genome GRCh38/hg38 (UCSC) using the Burrows-Wheeler Aligner (BWA) software (version 0.7.16a) (49). The commands of 'aln' and 'samse' of the BWA-backtrack algorithm were used with their default settings to undertake the mapping. The output files (*.sam) were then combined, filtered and converted to bam files (*.bam) by Samtools (version 1.4) using the commands of 'view', 'sort' and 'merge'. Further filtering based on the mapping scores (MAPQ \geq 30) was performed to remove ambiguous reads and only uniquely mapped reads were kept for further analysis. Mapping files (*.bam) were exported for downstream analysis. Mapping parameters are listed in Supplementary Table S2.

Analytical analysis was carried out using the MEDIPS package (version 1.24.0) on R programming software (version 3.3.2) (50,51). Command codes were based on the R script document of MEDIPS. MeDIP-Seq specific quality control analysis was performed on the mapping files for each sample, which included saturation analysis, CpG coverage analysis and CpG enrichment analysis. The parameters specified in this analysis were set to: 'extend' of 500, which means all short reads were extended to 500 nucleotides according to the reference genome; 'shift' to 0, which means there was no shift in the genome locations

for the sequences; 'window_size' to 100, which indicates the analysis was performed every 100 bp of the genome; 'uniq' to 1, which indicates that the reads that mapped to exactly the same genomic positions were counted as one read. A coupling factor was set up based on the HSR-GBM1 samples (control group) for normalization. Biological triplicates were grouped into three cohorts: the GBM cohort, the GBM + VitC cohort and the GBM + 5Aza cohort. Differentially methylated windows of 100 bp were identified between groups using the 'edgeR' plug-in function. Among the six samples in each comparison, genomic windows that were covered by a minimal counts threshold of 100 reads were kept for further statistical analysis (minrowsum = 100). Statistically significant results were selected using criteria for the adjusted *P* value ≤ 0.05 . Continuous significant windows were merged as one differentially methylated region (DMR).

Comparisons between the GBM cohort and the GBM + VitC cohort and between the GBM cohort and the GBM + 5Aza cohort were performed. DMRs identified between the cohorts then underwent annotation based on their corresponding regions in the human genome (hg38) using the ChIPSeeker package (version 3.5) (52). The built-in function of region of interest (ROI) analysis in the MEDIPS package was specifically used to investigate the DNA methylation levels at the CpG islands (CGIs) found in gene bodies and promoter regions of the mtDNA replication factors. One-way ANOVA was used to statistically compare the mean relative methylation score of ROIs. Relative methylation score was developed specifically for MeDIP-Seq in order to normalize methylation scores for regions based on the concept of CpG coupling analysis (50,51). Relative methylation score is calculated using the formula of $\log_2(\text{mean MeDIP-Seq signal} \times 10^6 / \text{the corresponding estimated number of reads based on the coupling analysis} \times \text{the total number of short reads})$ (50). It has been shown to correlate well with DNA methylation levels following bisulfite conversion methods, when applied to individual CpG sites within a region (50). In other words, relative methylation score offers a relatively high-throughput and accurate measurement of DNA methylation.

Purification of mtDNA from cells

To further verify whether DNA methylation exists within the mitochondrial genome, mtDNA was purified from cells to eliminate mtDNA pseudo-genes present in the nuclear genome from the analysis. Mitochondrial isolation was performed using a 5 ml Potter-Elvehjem tissue grinder set (Cat. No. 358034; Wheaton, USA). Freshly collected cells (~10 million) were resuspended in 5 ml of solution A (20 mM HEPES-KOH, pH 7.6, 220 mM Mannitol, 70 mM sucrose, 1 mM EDTA and 2 mg/ml BSA freshly added) and placed on ice for 15 min to facilitate swelling. The cells were then homogenized at 4°C in a glass chamber with a drill-fitted pestle for 50 repetitions. The cell homogenate was centrifuged at 800g for 10 min to remove cell debris and nuclei. The supernatant was centrifuged at 10 000 g for 20 min to pellet the mitochondrial fraction. The pellet was then resuspended in 175 µl of solution B (20 mM HEPES-KOH, pH 7.6, 220 mM Mannitol, 70 mM sucrose, 10 mM MgCl₂).

10 µl of DNase I (3 kunitz units/µl; Qiagen) in 20 µl of RDD buffer (Qiagen) were added to the mitochondrial suspension and incubated at 37°C for 30 min to further remove nuclear DNA. 1 ml of solution A without BSA was added to stop DNase activity and the suspension was centrifuged at 10 500 g for 20 min at 4°C to pellet the mitochondria. The pellet of mitochondria was resuspended in 200 µl of lysis buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA and 1% SDS) with 1 µl of proteinase K (20 mg/ml, Bioline) and incubated at 50°C for 60 min. mtDNA was then purified using the phenol:chloroform method by adding 1 volume of phenol: chloroform: isoamyl alcohol (25:24:1, Sigma), vortexing for 1 min and centrifuging at 17 000g for 5 min. The upper layer was collected and underwent another round of chloroform extraction by adding 1 volume of chloroform: isoamyl alcohol (24:1, Sigma). The upper layer was collected and mtDNA was precipitated by adding 2.5 volumes of 100% ethanol, 0.1 volume of 3 M sodium acetate and 1 µl of glycogen (20 µg/µl), and incubated at -80°C for at least 2 h. Samples were centrifuged at 20 000g for 10 min at 4°C and the mtDNA pellets were washed with 500 µl of 70% ethanol. After further centrifugation at 12 000g for 5 min at 4°C, the mtDNA pellets were air-dried for 10 min and dissolved in 50 µl of autoclaved Milli-Q H₂O.

Purified mtDNA samples firstly underwent the MeDIP-long-PCR protocol without sonication to demonstrate the presence of DNA methylation on mtDNA (described below in the section entitled 'Whole-mtDNA MeDIP and amplification by long PCR'). Secondly, purified mtDNA samples and long PCR products underwent MeDIP (with sonication) followed by qPCR to determine the levels of DNA methylation at different regions of the mitochondrial genome (described below in the section entitled 'Quantification of MeDIP products').

Whole-mtDNA MeDIP and amplification by long PCR

Prior to performing MeDIP, a positive and a negative control were generated by long PCR to generate two overlapping long mtDNA fragments (~8500 bp) spanning the whole mitochondrial genome. As PCR products are deemed to be unmethylated, the negative control (Neg-longPCR-5mC) comprised equal concentrations of two long mtDNA fragments generated by long PCR using primer set 1 (Supplementary Table S1). Each long PCR contained 1× High Fidelity PCR buffer, 100 mM MgSO₄, 1 mM dNTPs (Bioline), 1 U of Platinum Taq High Fidelity (Invitrogen), 10 µM each forward and reverse primer (primer set 1; Supplementary Table S1) in a total volume of 50 µl. Reaction cycling profiles were 94°C for 2 min, 35 cycles of 94°C for 15 s, 63°C for 30 s and 68°C for 8 min 45 s. Similarly, the positive control (Pos-longPCR-5mC) comprised equal concentrations of the two long mtDNA fragments that had undergone DNA methylation treatment. DNA methylation treatment was carried out using the CpG methyltransferase *M.SssI* (New England Biolabs; USA), according to the manufacturer's instructions. Briefly, 1 µg of the combined mtDNA fragments was incubated with four units of the enzyme in the presence of 160 µM S-adenosylmethionine (SAM) at 37°C for 4 h followed by 65°C for 20 min.

The purified mtDNA sample (mtDNA-5mC), and the positive (Pos-longPCR-5mC) and negative (Neg-longPCR-5mC) control samples then underwent MeDIP, as described in the section of 'Immunoprecipitation of methylated DNA (MeDIP)', except that the sonication step was not included. This maintains the integrity of the mtDNA, which acts as the template for long PCR following MeDIP. Moreover, the purified mtDNA sample also underwent MeDIP without 5mC antibody (mtDNA-NAC) to account for non-specific pull-down. MeDIP products were purified using the phenol:chloroform method, and assessed by long PCR using primer set 2 (Supplementary Table S1) with the same reaction conditions, as described above. mtDNA sequences targeted by primer set 2 are located just downstream and just upstream of 5' and 3' ends of the sequences used for the negative and positive controls, respectively. A total DNA sample isolated from GBM cells (Total DNA) and a non-template control (H₂O; NTC) were also amplified. PCR products were then run on a 0.8% agarose gel.

Quantification of MeDIP products

After performing MeDIP (with a sonication step), as described in the section 'Immunoprecipitation of methylated DNA (MeDIP)', on the purified mtDNA samples and long PCR products, the levels of DNA methylation for the regions of interest were quantified by qPCR on a Rotor-Gene 3000 machine under primer specific conditions (Supplementary Table S1). The levels of DNA methylation for the mitochondrial genome were determined by firstly normalizing the qPCR signals of the MeDIP products to their corresponding input samples (5mC/Input), and then weighting these values against the positive and negative controls that were used to represent 100% and 0% levels of DNA methylation, respectively. The levels of DNA methylation at exon 2 of *POLG*, and *TOP1MT* exon 8 and intron 10 from MeDIP-Seq were also validated by MeDIP-qPCR using total DNA. The results were determined by normalizing the qPCR signals of the MeDIP products to their corresponding input samples (5mC/Input).

Determination of copy number variation (CNV)

All samples were genotyped for CNV using the Illumina Human Global Screening Array Beadchip (Illumina), which covers ~700K SNPs. Array data were processed using GenomeStudio 2.0 (Illumina), according to the user's manual. Data normalization, clustering and genotype calling were performed using the Genotype Module. The full data report containing log R ratios (LRR) and B allele frequencies (BAF) for each probe was then exported as input data to the Nexus 9.0 software (Biodiscovery Inc., CA, USA) for advanced analysis. The Nexus Copy Number Module was used to carry out the comparisons of DNA copy number between groups. Human genome build hg19 was chosen as the reference genome to assign genome locations. Statistical significance between groups in individual SNPs was determined by *P* value ≤ 0.05 . Genomic regions were aligned to hg38 for overlapping analysis with DMRs.

RNA-Seq analysis

Total RNA was submitted to the Australian Genome Research Facility (AGRF; VIC, Australia) to perform RNA sequencing. An Agilent Bioanalyzer Nano chip (Agilent) was used to determine the integrity of the RNA samples. cDNA libraries were constructed using the Illumina Truseq Standard mRNA Kit (Illumina), according to the manufacturer's instructions. The libraries were sequenced on the HiSeq 2500 platform (Illumina). Image analysis was performed in real time using the HiSeq Control Software (HCS) v2.2.68 and Real Time Analysis (RTA) program v1.18.66.3 running on the instrument's computer. RTA performed real-time base calling on the HiSeq instrument computer. The Illumina bcl2fastq 2.19.0.316 pipeline was then used to generate the sequence data at AGRF. The sequences were mapped to human genome hg38 by STAR aligner (2.5.3a) using default settings (53). The annotation was based on the comprehensive GENCODE gene annotation (v26). Only the uniquely mapped reads were used for downstream analysis. The reads were then summarized to counts for each gene by the 'featureCounts' function in the Rsubread package (1.26.0) using the inbuilt annotation hg38 (54). Identification of differentially expressed genes was performed using the edgeR package (3.16.5) employing the trimmed mean of M-values normalization method (TMM) application (55,56). Gene expression data were uploaded to Ingenuity Pathway Analysis software (IPA; Qiagen). Genes with false discovery rate (FDR) ≤ 0.05 and absolute fold change ≥ 2 were deemed to be significantly differentially expressed genes. These genes were used to perform pathway analysis for determining the most affected canonical pathways, biological functions and networks, and used for undertaking the DMR-overlapping analysis.

Gene expression analysis using the Fluidigm platform

cDNA was synthesized from 1 μ g of total RNA using oligo (dT) primers and the Superscript III First-Strand synthesis system (Thermo Fisher Scientific), according to the manufacturer's instructions. Pre-amplification was performed to increase the number of copies of each gene to detectable levels, as detailed in Gene Expression Preamp with the Fluidigm Preamp MasterMix and Taqman Assays Quick Reference PN 100-5876 B1. All Taqman assays are listed in Supplementary Table S3. Taqman assays were pooled with C1 DNA suspension buffer to produce a final concentration for each assay of 180 nM. 1.25 μ l of each cDNA sample as well as a non-template control underwent pre-amplification for 14 cycles with 3.75 μ l of the pooled assays and Taqman PreAmp Master mix (Thermo Fisher Scientific), according to the manufacturer's instructions.

Following preamplification, reaction products were diluted 5-fold in C1 DNA suspension buffer. Assays and samples were combined in a 96.96 Dynamic array IFC plate, according to the Fluidigm® 96.96 Real-Time PCR Workflow Quick Reference PN 6800088. An Integrated Fluidic Circuit (IFC) controller HX was used to prime and load the plate. 5 μ l of each preamplified sample were loaded in duplicate into each sample inlet and 5 μ l of each Taqman assay (10 \times) were also loaded into each assay inlet of the plate. Gene expression was performed according to the Biomark GE 96.96

Standard v2 protocol. Data were exported as a spreadsheet (.csv) using the Fluidigm Real-Time PCR analysis software (v4.1.1). Differentially expressed genes were analysed using the HTqPCR package (version 1.28) (57). The package uses the 'deltaCt' method of normalization. To improve the analysis, three housekeeping genes were used for normalization. *ActB*, *HPRT1* and *OAZ1* were chosen as the housekeeping genes for the nuclear genes. They are widely used in glioblastoma studies (4,58). *Rn18S*, *HPRT1* and *OAZ1* were chosen as the housekeeping genes for the mitochondrial-related genes. *Rn18S* has higher levels of expression and is, therefore, more suitable for the normalization of the mitochondrial genes, which have much higher levels of expression than nuclear-encoded genes (58,59). The package uses the 'limma' method to perform statistical analysis, as stated in the user guide available from the website. Results were then exported and plotted using GraphPad Prism 7 (GraphPad Software, Inc., CA, USA).

Data deposition

MeDIP-Seq sequences (triplicates) for the HSR-GBM1 cells and HSR-GBM1 cells treated with 5Aza and VitC have been deposited in the Sequence Read Archive (SRA) under accession number SRP080899. CNV data and RNA sequencing data have been deposited in the Gene Expression Omnibus (GEO) under accession number GSE98693.

RESULTS

Whole genome MeDIP-Seq, quality control and CNV analysis

To investigate the impact that DNA methylation has on the whole genome and how it affects mtDNA replication, we performed whole genome MeDIP-Seq on DNA samples from HSR-GBM1 cells (GBM) to identify sites that are DNA methylated, in order to then compare them with the DNA methylome of HSR-GBM1 cells cultured with either 5Aza (GBM + 5Aza) or VitC (GBM + VitC) to determine which sites undergo modulation. The HSR-GBM1 cell line was chosen as it is one of the best-characterized GBM cell lines and is indicative of an undifferentiated cancer cell line, which has the potential to be differentiated into neuronal- and astrocyte-like cells (6,35) once it has undergone DNA demethylation induced by 5Aza or VitC (4).

On average, >40.4 million uniquely mapped sequences per sample were obtained (Table 1). The mapped reads were cleaned to remove duplicates that mapped to exactly the same genomic position and were counted as one read to ensure that the repeat reads resulting from PCR amplification were not biasing the data. To check the reproducibility of the workflow, a Pearson correlation test was performed on the mapped reads within each cohort. On average, high correlation scores (0.94 for the GBM cohort, 0.95 for the GBM + VitC cohort and 0.95 for the GBM + 5Aza cohort) were achieved amongst the triplicates for each cohort (Table 1). The MEDIPS analysis package offers several quality control tests for MeDIP-Seq experiments. Saturation analysis gives the estimated saturation correlation score to assess the coverage of the genome. Overall, scores of 0.95 were achieved from the samples, which indicates sufficient

coverage (Table 1). CpG coverage analysis specifically includes coverage levels of the 28 million CpG dinucleotides in the human genome. On average, >91% of total CpG dinucleotides in the human genome were covered with at least one read which indicates good enrichment of the targets for downstream analysis (Table 1) and the differences in coverage depth are shown in Supplementary Figure S1.

Cancer cells frequently exhibit genomic instability. CNVs between cohorts can have a direct impact on MeDIP-Seq signals that are not reflective of true changes in DNA methylation and, therefore, result in false positive or negative results (60). To preclude the potential impact of CNVs on the MeDIP-Seq data, the three cohorts of cells were genotyped to screen for CNVs over the whole genome. By using an Illumina global screening array, 700 K SNPs were analyzed and statistical comparisons were performed using the Nexus 9.0 Copy Number module. Overall gains in copy number were observed in the GBM genome. In all, 247 regions were identified as copy number gains (Supplementary Table S4). However, no significant CNVs were identified between either the GBM and GBM + VitC or GBM + 5Aza cohorts (Supplementary Figure S2). In other words, the DNA demethylation treatments did not improve or intensify genomic instability among the cohorts of GBM cells, which means that the CNVs did not impact on the analysis by MeDIP-Seq.

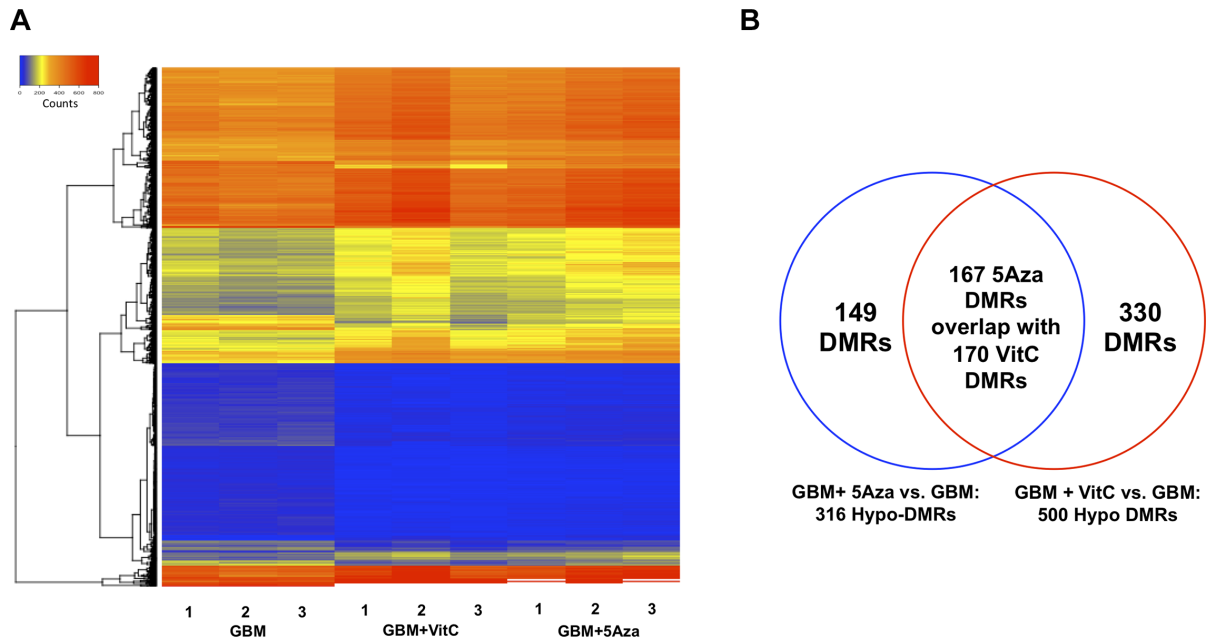
Identification of differentially methylated regions (DMRs)

Using the MEDIPS package, 816 differentially methylated regions (DMRs) were identified between the GBM + 5Aza and the GBM cohorts and 1743 DMRs between the GBM + VitC and GBM cohorts (adjusted P value ≤ 0.05 , minrowsum ≥ 100). The heatmap for the total number of DMRs indicated that the two cohorts that had undergone DNA demethylation treatment exhibited significant changes when compared with the GBM methylome (Figure 1A). Interestingly, both the GBM + 5Aza and the GBM + VitC cohorts exhibited hyper-DMRs and hypo-DMRs throughout the genome (Figure 1A). Indeed, 316 out of 816 DMRs (38.72%) were identified to be demethylated by 5Aza and 500 out of 1743 DMRs (28.68%) were demethylated by VitC. The occurrence of hypermethylation is expected, especially in the GBM + VitC cohort, as the rebound effect can result in the re-establishment of DNA methylation (4). We focused on the demethylated DMRs for further analysis, as they are indicative of the direct effects of the DNA demethylation treatments.

Between the two cohorts of DMRs, 167 DMRs identified from the comparison between the GBM + 5Aza cohort and the GBM cohort were found to overlap with 170 DMRs identified from the comparison between GBM + VitC and GBM cohorts (the unequal distribution of DMRs is due to each of the treatments differentially affecting different regions resulting in some partial overlap) (Figure 1B). In all, 149 DMRs were uniquely identified from the comparison between the GBM + 5Aza and the GBM cohorts, whereas 330 DMRs were uniquely identified from the comparison between the GBM + VitC and the GBM cohorts (Figure 1B). The uniquely identified DMRs probably resulted from their different mechanisms of action for regulating DNA

Table 1. MeDIP-Seq outputs and QC results following analysis using the MEDIPS package. Each number after a sample indicates a biological replicate

Sample	Reads without duplicates (millions)	Correlation between triplicates	Estimated Saturation Correlation	CpGs covered by the reads
GBM 1	41.8	0.94	0.95	91.7%
GBM 2	37.3		0.95	91.51%
GBM 3	38.3		0.95	91.65%
GBM+5Aza 1	37.6	0.95	0.95	91.23%
GBM+5Aza 2	45.6		0.95	91.98%
GBM+5Aza 3	48.6		0.95	92.1%
GBM+VitC 1	37.8	0.95	0.95	90.8%
GBM+VitC 2	44.4		0.95	90.83%
GBM+VitC 3	32.5		0.94	89.56%

**Figure 1.** Overview of the DMRs. (A) Heatmap of the DNA methylation profiles. Total identified DMRs were plotted using values in counts for the three biological replicates from the GBM cohort, the GBM + 5Aza cohort and the GBM + VitC cohort. The colour scheme from blue, yellow to red represents the level of DNA methylation from low to high. (B) Venn diagram of the DMRs identified in the two comparisons. The 316 DMRs identified in the comparison between GBM + 5Aza and GBM are shown in the blue circle; the 500 DMRs identified between GBM + VitC and GBM are shown in the red circle. 167 DMRs identified in the GBM + 5Aza cohort overlap or partially overlap with 170 DMRs identified in the GBM + VitC cohort.

demethylation. As a result, the unique DMRs are likely to be the sites that were only affected by one of the agents. On the other hand, both of the agents could have demethylated the same DNA region but only one of them was able to significantly demethylate that region or maintain DNA demethylation.

As the function of DNA methylation varies dependent on the location of the methylated sites, the identified DMRs were then annotated according to their genome location using the ChIPSeeker package, which identified promoter regions (1000 bp upstream of transcription start site (TSS)), exons, introns, three UTRs and downstream regions (< 3 kb). These regions are collectively referred to as intragenic regions and are normally associated with transcriptional functions. However, the other DMRs (> 50%) are located at intergenic regions, which refer to non-coding genomic regions between genes (Supplementary Figure S3). Regardless of the number of DMRs, it appeared that VitC had a tendency to affect more intragenic regions when com-

pared with 5Aza. This could arise from the enhanced activity of the TET1 enzyme modulated by VitC, as it is known to specifically target more intragenic regions (42), while DNMT1 is more likely to be inhibited by 5Aza (43), thus causing global DNA demethylation.

Furthermore, of the DMRs identified, only seven mapped with the copy number gain regions identified above, which further suggests the regions possessing CNVs were not affected by DNA demethylation and had only a very limited effect on the identification of DMRs (Supplementary Table S4).

Identification of differentially expressed genes and the association with the DMRs

To investigate the changes to the overall gene expression profiles induced by DNA demethylation, total RNA samples from the three cohorts were sequenced and the total differentially expressed gene (FDR \leq 0.05, absolute logFC

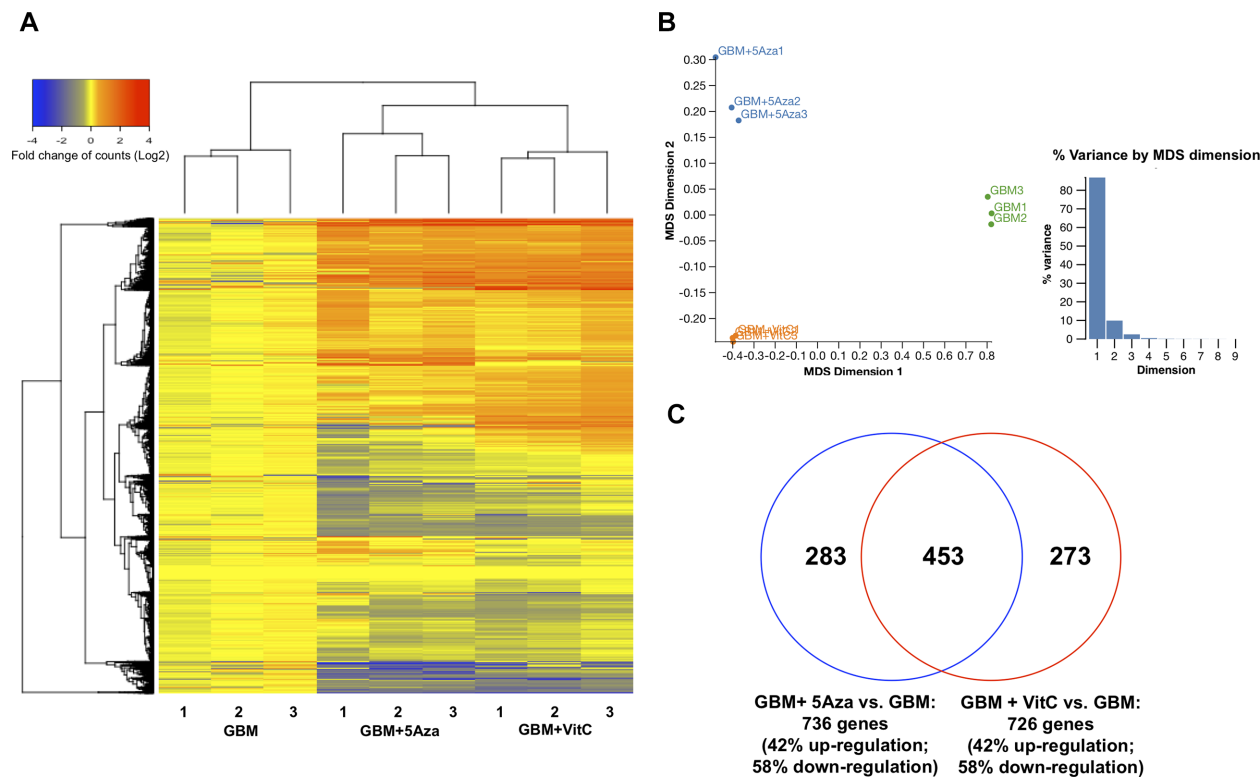


Figure 2. Overview of differentially expressed genes. (A) Heatmap of the RNA-Seq profiles. Total identified differentially expressed genes ($FDR \leq 0.05$, absolute $\log_2 FC \geq 1$) for the GBM + 5Aza cohort and the GBM + VitC cohort were plotted based on the fold changes (\log_2) of read counts for the mean value of the GBM cohort. The color scheme from blue, yellow to red represents the level of expression from low to high. (B) MDS plot of the RNA profiles of the three cohorts. The MDS plot shows the similarity of the cohorts for each group indicated by the distance on the first two dimensions. Dimension 1 accounts for 83% of the total variance, and Dimension 2 accounts for 12% of the total variance. The GBM cohort was plotted in green, the GBM + VitC cohort was plotted in orange and the GBM + 5Aza cohort was plotted in blue. (C) Venn diagram of the differentially expressed genes identified in the two comparisons. 736 genes identified to be differentially expressed in the comparison between GBM + 5Aza and GBM are shown in the blue circle; 726 genes identified to be differentially expressed between GBM + VitC and GBM are shown in the red circle. 453 genes overlapped between these two comparisons. A 42% up-regulation and a 58% down-regulation were identified in both comparisons.

≥ 1) profiles are shown in the heatmap (fold change of counts in \log_2 scale) for each of the triplicates from the treated cohorts (Figure 2A). The multidimensional scaling (MDS) plot analysis shows that the three cohorts are distinct from each other (Figure 2B). The first dimension separated the untreated GBM and the DNA demethylated GBM cohorts, which indicates the primary factor that distinguishes the cohorts is their DNA methylation status. The differences between the GBM + 5Aza cohort and GBM + VitC cohort were also indicated by distance in the second dimension, which suggests that these two treatments caused different responses in gene expression despite the effects of DNA demethylation. Using edgeR for further analysis, 726 differentially expressed genes were identified from the GBM + VitC cohort and 736 differentially expressed genes were identified from the GBM + 5Aza cohort ($FDR \leq 0.05$ and absolute $\log_2 FC \geq 1$; Figure 2C). A total of 453 genes were commonly modulated by both treatments (Figure 2C). However, 273 genes were uniquely identified in the GBM + VitC gene list, whereas 283 genes were uniquely modulated by 5Aza (Figure 2C). Interestingly, both comparisons determined that 42% of genes were up-regulated and 58% were down-regulated (Figure 2C), which highlights the dual effects that DNA methylation has on transcription.

Differentially expressed genes were then used for pathway analysis using IPA software. The top canonical pathways, top diseases and bio-functions affected by each of the treatments are listed in Table 2. In the top canonical pathways, hepatic fibrosis and acute phase response signaling pathways were identified in both comparisons, which suggests the genes in these pathways are greatly sensitive to DNA demethylation. On the other hand, VitC modulated cellular pathways involved in cellular organization and cell cycle, including mitotic roles of the polo-like kinase, and G2/DNA damage checkpoints in cell cycle, whereas 5Aza modulated pathways associated with tumorigenesis and metabolism including colorectal cancer metastasis signaling, Wnt/-catenin signaling and MIF regulation of innate immunity. The most affected disease was cancer resulting from both treatments with 565 molecules being affected by VitC and 572 molecules affected by 5Aza. Interestingly, VitC tended to modulate cell growth by affecting DNA replication, recombination, and repair in the cell cycle and triggering cell death, whereas 5Aza tended to inhibit cellular proliferation. Indeed, the increases in cell number during tissue culture were slightly lower in the treated cohorts than the GBM cohort. From the same initial number of cells, the GBM cohort increased $2.9 (\pm 0.1)$ fold, whereas

Table 2. Pathways affected by differentially expressed genes (FDR \leq 0.05, absolute fold change \geq 2)

Name	P-value	Overlap/no. of molecules
Top canonical pathways affected by VitC		
Hepatic Fibrosis/Hepatic Stellate Cell Activation	5.61E-06	9.8% 18/183
Mitotic Roles of Polo-Like Kinase	1.05E-04	13.6% 9/66
Acute Phase Response Signaling	3.73E-04	8.3% 14/169
Cell Cycle: G2/M DNA Damage Checkpoint Regulation	4.59E-04	14.3% 7/49
Salvage Pathways of Pyrimidine Ribonucleotides	1.60E-03	9.5% 9/95
Diseases and disorders affected by VitC		
Cancer	2.57E-04–3.31E-16	565
Organismal Injury and Abnormalities	2.62E-04–3.31E-16	582
Reproductive System Disease	2.24E-04–3.53E-16	333
Gastrointestinal Disease	2.62E-04–4.33E-13	499
Dermatological Diseases and Conditions	1.01E-04–9.62E-11	380
Molecular and cellular functions affected by VitC		
Cellular Assembly and Organization	2.68E-04–2.88E-14	122
DNA Replication, Recombination, and Repair	2.68E-04–2.88E-14	41
Cell Cycle	2.14E-04–5.22E-13	104
Cell Death and Survival	2.62E-04–2.84E-12	223
Cellular Movement	2.32E-04–3.75E-12	172
Top canonical pathways affected by 5Aza		
Hepatic Fibrosis / Hepatic Stellate Cell Activation	7.52E-15	16.4% 30/183
Colorectal Cancer Metastasis Signaling	8.45E-06	8.5% 21/247
Acute Phase Response Signaling	2.68E-05	9.5% 16/169
Wnt/-catenin Signaling	9.92E-05	8.9% 15/169
MIF Regulation of Innate Immunity	1.39E-04	17.1% 7/41
Diseases and disorders affected by 5Aza		
Cancer	2.38E-05–2.97E-22	572
Organismal Injury and Abnormalities	2.38E-05–2.97E-22	585
Dermatological Diseases and Conditions	4.48E-08–2.42E-19	399
Gastrointestinal Disease	1.18E-05–2.84E-18	510
Metabolic Disease	1.68E-05–1.05E-12	139
Molecular and cellular functions affected by 5Aza		
Cellular Movement	2.39E-05–5.26E-18	173
Cellular Development	2.30E-05–2.14E-10	219
Cellular Growth and Proliferation	2.30E-05–2.14E-10	206
Cell Morphology	1.94E-05–5.13E-10	176
Lipid Metabolism	2.18E-05–6.79E-10	88

GBM+VitC gained 2.5 (\pm 0.14) fold and GBM + 5Aza gained 2.55 (\pm 0.1) fold ($P > 0.05$). The networks analyses described the gene ontology (GO) enrichment for the differentially expressed genes (Supplementary Figures S4–S7). The top network enriched in the VitC-induced differentially expressed genes was cell cycle, cellular assembly and organization, and DNA replication, recombination and repair (Supplementary Figure S4). Interestingly, the glioblastoma disease pathway was also found to be modulated by VitC with a range of associated genes up-regulated (Supplementary Figure S5). On the other hand, 5Aza more effectively influenced the: (i) cell morphology, cellular function and maintenance, and carbohydrate metabolism (Supplementary Figure S6); and (ii) cancer, neurological disease, and organismal injury and abnormalities (Supplementary Figure S7) networks. Even though there was similarity between the 5Aza- and VitC-induced differentially expressed genes and DMRs, the differences in the enrichment of cellular functions and networks indicated the extent of DNA demethylation. This is likely due to differences at different targets as a result of the different mechanisms associated with DNA demethylation resulting from the administration of VitC or 5Aza. This, in turn, leads to differences in transcription and, therefore, distinct responses from cellular pathways.

Validation of transcriptional changes to the DMR-overlapping differentially expressed genes

To specifically investigate the association between the intragenic DMRs and the differentially expressed genes, we mapped the intragenic DMRs with the differentially expressed genes (FDR \leq 0.05, absolute logFC \geq 1). In all, 51 genes overlapped with VitC DMRs; and 27 genes overlapped with 5Aza DMRs. There were 18 genes common to both lists (Table 3). However, due to the limited number of genes, no cellular pathways were significantly enriched ($P < 0.001$).

The 60 DMR-overlapping differentially expressed genes were further analysed using the Fluidigm platform to validate their changes in gene expression. Fold changes (log2) for each gene relative to the GBM cohort were used to generate the heatmap (Figure 3A; detailed results are shown in Supplementary Table S5). Genes with their levels of expression modulated by at least one of the treatments over absolute 2-fold are separately plotted in Figure 3B. The statistical significances are indicated on the sides of the plot. Overall, up-regulation in expression of these genes was observed in both treated cohorts with 5Aza triggering more significant changes than VitC (Figure 3B). *MYT1*, *MAN1C1*, *PLEC*, *TLE3*, *PALM3*, *FAM222A*, *GTF2IRD1*, *PRDM16*, *WNK2*, *KIF21B* and *PLXNA4* were up-regulated by both treatments by over 2-fold, whilst *SEC61G*, *C2CD4C* and

Table 3. Differentially expressed genes that were identified to overlap with DMRs. Differentially expressed genes that were mapped with DMRs in each comparison are listed: common DMR-overlapping differentially expressed genes that were identified in both comparisons and uniquely identified DMR-overlapping differentially expressed genes in each comparison

Comparison		DMR-overlapping differentially expressed genes
GBM+5Aza vs. GBM	Unique	<i>SLCO4A1; PRICKLE2; MEIS3; MAN1C1; HSD17B14; GAS6; FND10; FBXO17; COL22A1;</i>
	Common	<i>TSNARE1; TLE3; SPPL2B; SMARCD3; SEC61G; RAB7B; PRDM16; PLXNA4; PLEC; NACAD; MEST; LINC01224; FZRI; FCGBP; ELN; CLU; CCDC71L; BCL3; ADGRB1; AGAP3; AGPAT4; ASL; ATP6V0E2-AS1; C2CD4C; C7orf57; CLIP2; COL4A2; CYP4F11; FAM20C; FAM222A; FAM49A; GTF2IRD1; HSPG2; KIF17; KIF21B; LRRC4B; LTBP4; LYPLA1; MASP2; MYCNUT; MYT1; NKAIN4; PALM3; PDGFA; PRKAG2; PTPRS; SHC2; SNX10; ST3GAL1; TTYH1; WNK2;</i>
GBM+VitC vs. GBM		
GBM+VitC vs. GBM	Unique	
	Common	

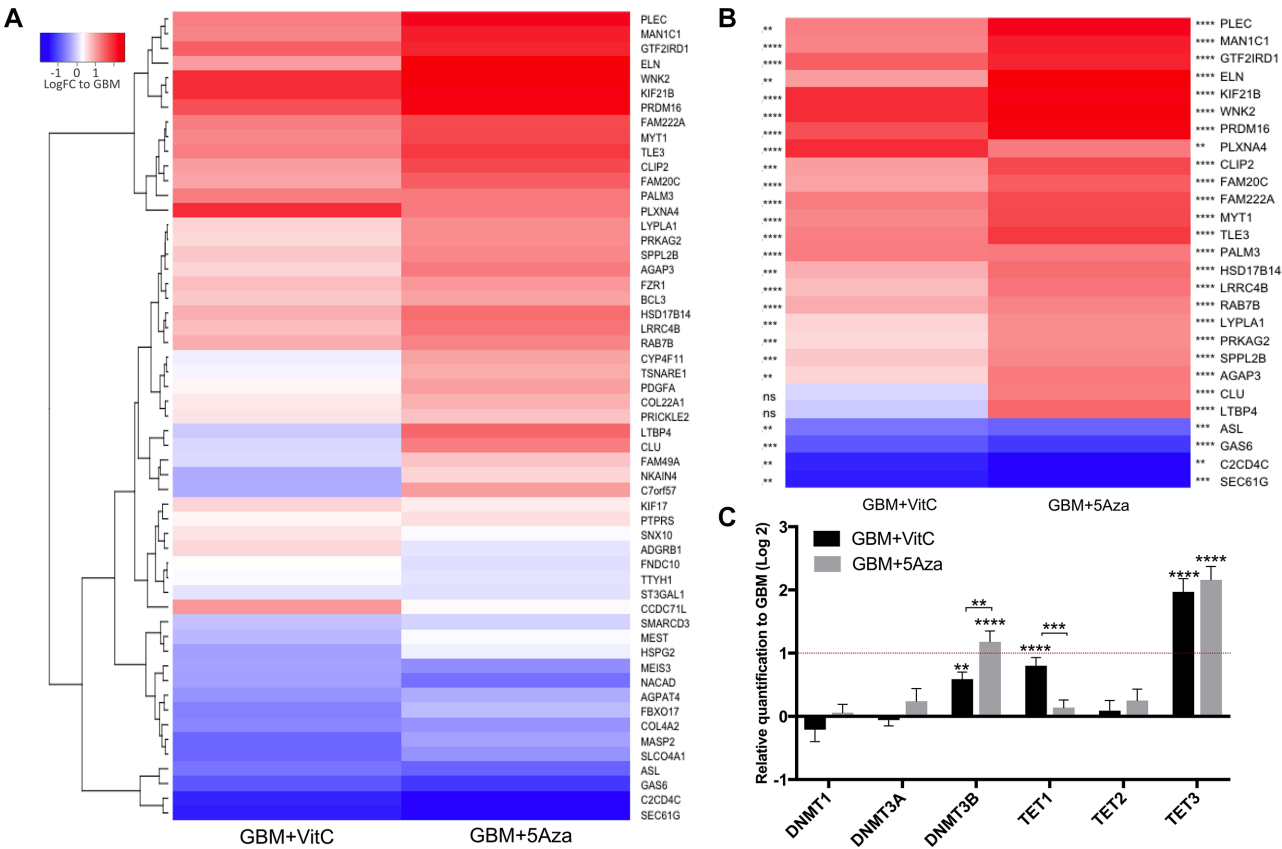


Figure 3. The validation of gene expression using the Fluidigm array. (A) Heatmap of the gene expression levels of the DMR-overlapping differentially expressed genes. For each gene, fold changes to the mean value of the respective gene in the GBM cohort (\log_2) were plotted. (B) Heatmap of the gene expression levels of the DMR-overlapping differentially expressed genes that had more than 2-fold changes after either of the treatments. Statistical significances are indicated at the sides of the plot relevant to each gene: the left side is for the GBM+VitC cohort and the right side is for the GBM + 5Aza cohort. **, ***, **** indicate P values < 0.01, 0.001, 0.0001, respectively; ns = not significant. (C) Differential expression of the regulators of DNA methylation. Bars represent the mean of the relative quantification levels in \log_2 scale normalized to the GBM cohort. Detailed statistical results of the DMR-overlapping differentially expressed genes are shown in Supplementary Table S5. Bars in black represent GBM + VitC; bars in grey represent GBM + 5Aza. Error bars show SEM. The level of significance for each gene is labelled above each bar, **, ***, **** indicate P values < 0.01, 0.001, 0.0001, respectively. The red hash line at 1 indicates 2-fold change of up-regulation.

GAS6 were down-regulated by both treatments by over 2-fold, which indicates their expression levels are sensitively modulated by DNA demethylation. Interestingly, *SEC61G* encodes a subunit of the Sec61 complex as the central component of the protein translocation apparatus of the endoplasmic reticulum membrane. This gene is located upstream of another hallmark of GBM, *EGFR*, and is highly over-expressed in GBM to support tumour cell survival (61). *WNK2* encodes the tumor suppressor WNK lysine defi-

cient protein kinase 2, which is regulated by DNA methylation at its promoter region and loss of expression has been found in GBM as a result of promoter hypermethylation (62). *PRDM16*, PR domain containing 16, was found to be overexpressed when its promoter is hypomethylated and contributes to tumorigenesis by disrupting mitochondrial function in astrocytoma (63). Furthermore, 13 genes (*ASL*, *LTBP4*, *CLU*, *PRKAG2*, *AGAP3*, *LYPLA1*, *SPPL2B*, *LRRC4B*, *HSD17B14*, *RAB7B*, *FAM20C*, *CLIP2*

and *ELN*) were only significantly modulated by 5Aza to over 2-fold. They are likely to be more sensitive to the DNA demethylation mediated by the inhibition of DNMT1 activity resulting from the 5Aza treatment (Figure 3B).

As there were different patterns of DNA methylation and gene expression induced by 5Aza and VitC, we investigated the levels of gene expression for the regulators of DNA methylation including the DNA methylation enzymes, namely methyltransferases DNMT1/3A/3B/3L, and the DNA demethylation enzymes, namely TET1/2/3. It is known that 5Aza inhibits DNA methylation through the DNMT1 enzyme, whereas VitC acts as the co-factor of the TET1 enzyme to enhance DNA demethylation (42,43). TET-mediated DNA demethylation converts 5mC to 5hmC, which is important for the regulation of the tumour- and differentiation-related genes. The DNMT3 family is known to mediate *de novo* DNA methylation and is maintained by DNMT1 during DNA replication. The expression of *DNMT3L*, the co-factor of DNMT3A, was not detected in each of the three cohorts. *TET3* was significantly up-regulated following treatment with VitC by >2-fold compared with the GBM cohort (Figure 3C), whereas 5Aza enhanced the expression of *DNMT3B* and *TET3* by >2-fold (Figure 3C). When we compared the two DNA demethylation agents, treatment with 5Aza resulted in significantly higher levels of gene expression for *DNMT3B* than VitC ($P < 0.01$; Figure 3C). VitC triggered significantly higher levels of expression for *TET1* than 5Aza ($P < 0.001$; Figure 3C). The up-regulation of *DNMT3B* is likely to be a response to global DNA demethylation in order to compensate for the reduced levels of DNA methylation induced by the DNA demethylation agents. Indeed, this is a typical feature, which is similar to the rebound effect when DNA demethylation agents are withdrawn from cancer cells and they return to their hypermethylated state (4,64). It is also interesting to note that the levels of expression for *TET1* were up-regulated by its co-factor VitC, and not by 5Aza. The differences shown in the methylation factors further indicated that these two treatments could mediate different DNA demethylation machineries as different targets in the genome were modulated.

The changes in DNA methylation to the genomic regions of the mtDNA transcription and replication factors

We specifically focused on investigating the changes in DNA methylation to the genomic regions of the mtDNA transcription and replication factors. Indeed, the rationale for undertaking this analysis is supported by the >3-fold increase in mtDNA copy number induced by the DNA demethylation treatments, which strongly indicates that mtDNA copy number is epigenetically regulated (Figure 4A) (4,32). Increased mtDNA copy number has also previously been shown to be tightly associated with the key regulator of mtDNA replication, *POLG* and more specifically associated with the DNA methylation status of a CpG island (CGI) in its second exon (4). Indeed, CGIs are genomic regions with repeated CpGs that usually exhibit low levels of DNA methylation. However, they are more tightly associated with regulatory roles in gene transcription than

CpG dinucleotides, which are present throughout the whole genome (65).

To further investigate changes in DNA methylation to the other mtDNA transcription and replication factors, the relative methylation score of CGIs found in the gene bodies and their promoter regions were analyzed using the MEDIPS package based on the MeDIP-Seq data. A total of 20 mtDNA transcription and replication factors were examined and had a CGI in the promoter regions of at least one transcript variant, which indicates their expression is highly likely to be affected by DNA methylation. Overall, reduced levels of DNA methylation were observed in the treated cohorts (Supplementary Figure S8). VitC significantly reduced the levels of DNA methylation in three promoter CGIs, *TFB1M*, *TWNK* and *SSBP1* (Supplementary Figure S8).

The gene-body CGIs of the mtDNA transcription and replication factors were also analysed in the same way (Figure 4B). Firstly, the significant reduction in DNA methylation at exon 2 of *POLG* induced by both treatments was compatible with our previous findings (4). We also investigated the four CGIs that overlapped within intron 7, exon 8, exon 10 and intron 10 of *TOP1MT*, which would strongly indicate that the expression of *TOP1MT* is probably modulated by DNA methylation, as is the case for *POLG* (4,6). Both treatments caused a significant decrease in DNA methylation at exon 8 of *TOP1MT*, while the CGI at intron 10 was only significantly demethylated by VitC. CGIs at the boundary of exon 1 and intron 1 of *TWNK* and intron 1, exon 2/6 of *ESRRB* did not undergo significant DNA demethylation. Using MeDIP-qPCR, we confirmed that demethylation had taken place at exon 2 of *POLG* and exon 8 of *TOP1MT* through both 5Aza and VitC (Figure 4C). Interestingly, our validation also showed that the CGI at intron 10 of *TOP1MT* was also demethylated by both agents (Figure 4C).

These data show that the DNA methylation status at the promoter-located CGIs and the gene-body CGIs are modulated by the DNA demethylation treatments, which probably affect the gene expression profiles of these transcription and replication factors and their ability to regulate mtDNA copy number. To validate this, we determined the levels of expression of each of the transcription and replication factors. Firstly, we confirmed our previous finding that the expression of *POLG* was significantly up-regulated to ~2-fold in both demethylated cohorts (Figure 4D). Along with *POLG*, the vast majority of the transcription and replication factors showed increased levels of expression after the DNA demethylation treatments. *NRF1*, *NRF2* and *STAT3* more than doubled their levels of expression after VitC treatment (Figure 4D) and even more so after 5Aza treatment. *ESRRB*, *TWNK*, *TOP1MT*, *POLRMT*, *HIF1 α* and *SIRT1* were significantly increased by <2-fold by VitC (Figure 4D). However, 5Aza increased expression for each of these genes to >2-fold (Figure 4D). *SSBP1* was down-regulated in both cohorts, which indicates DNA demethylation treatments are likely to suppress its expression (Figure 4D). *IDH1*, involved in the TCA cycle and associated with TET-mediated DNA demethylation, was down-regulated by VitC by over 2-fold (Figure 4D). Interestingly, *TFB1M*, *SIRT3*, *PGC1 α* and *IDH2* had opposing results

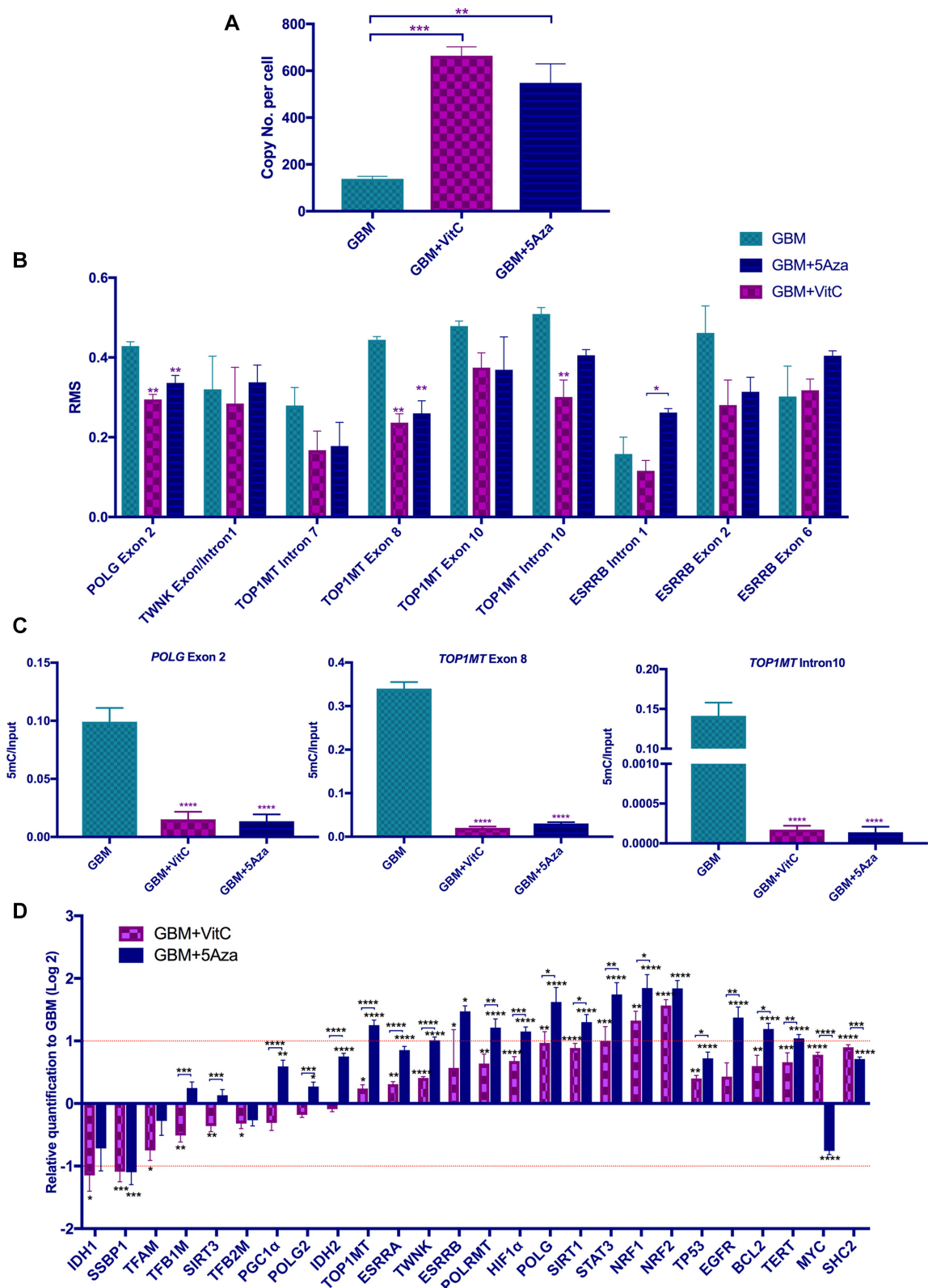


Figure 4. Analysis of mtDNA transcription and replication factor related regions. (A) mtDNA copy number for the three cohorts. mtDNA copy number per cell for the GBM, GBM + VitC and GBM + 5Aza cohorts ($n = 3$). (B) DNA methylation levels expressed as relative methylation score (RMS) over the CGIs located within the gene bodies of the mtDNA transcription and replication factors are shown ($n = 3$). The relative methylation score was determined using the MEDIPS package based on the MeDIP-Seq data. (C) Validation of DNA methylation using MeDIP-qPCR at exon 2 of *POLG*, and exon 8 and intron 10 of *TOP1MT* ($n = 3$). DNA methylation levels were determined by normalizing 5mC to the input (5mC/Input). (D) Validation for gene expression of the mtDNA transcription and replication factors ($n = 6$). Statistical analysis was performed using One-way ANOVA. Bars represent the mean \pm SEM. * to **** indicate P values of < 0.05 , 0.01 , 0.001 , 0.0001 . The GBM cohort is shown by the light blue bars; the GBM + VitC cohort is shown by the purple bars, the GBM + 5Aza cohort is represented by the dark blue bars.

for the two treatments. Their levels of expression were decreased after VitC treatment but slightly up-regulated by 5Aza treatment (Figure 4D). From the comparisons between the two treated cohorts, 5Aza had significantly more effect on increasing the levels of gene expression, especially on the direct-binding transcription and replication factors, which included *TFB1M*, *POLG2*, *TOP1MT*, *TWNK*, *POLRMT* and *POLG*. The indirect transcription and replication factors, including *SIRT3*, *PGC1 α* , *IDH2*, *ESRRA*, *HIF1 α* , *SIRT1*, *STAT3* and *NRF1*, were also more greatly affected by 5Aza than VitC (Figure 4D). Furthermore, we analysed other tumor markers of GBM including *EGFR*, *BCL2*, *TP53*, *TERT*, *SHC2* and *MYC* that interact closely with the mtDNA transcription and replication factors (66–69). Similarly, overall up-regulation was observed after the treatments, but only 5Aza significantly up-regulated *EGFR*, *BCL2* and *TERT* to ~2-fold or above. However, 5Aza was found to repress *MYC* by over 1.5-fold (Figure 4D).

The relationship between mtDNA methylation and the expression of mtDNA-encoded genes

In the human mitochondrial genome (hg38), there are 435 CpG dinucleotides (70). Even though DNA methylation of mtDNA remains controversial, there is an increasing number of findings indicating that mtDNA methylation takes place, especially in the main non-coding region, the D-loop (70,71). Since the nuclear genome harbours mtDNA-pseudo genes, we performed MeDIP on purified populations of mtDNA to investigate the presence of mtDNA methylation.

Firstly, to determine the degree of nuclear DNA elimination from the purified mtDNA samples, we normalized the levels of nDNA (β -globin) to levels of mtDNA. The relative levels of nDNA in the purified mtDNA samples were significantly lower than in the total DNA samples isolated from cells that had not undergone mitochondrial isolation (Figure 5A). Indeed, significant reductions of nDNA (to nearly zero) in all three cohorts were achieved by the purification process. To determine whether mtDNA is DNA methylated, we tested purified mtDNA and long PCR fragments generated from the purified mtDNA, that had been treated with and without a CpG methylation specific enzyme. In this respect, we generated two overlapping fragments of mtDNA (~8500 bp) by long PCR using primer set 1 (Supplementary Table S1) and treated these fragments with the CpG methylation enzyme, *M.SssI* to generate a positive control (Pos-longPCR-5mC) and without the enzyme to generate a negative control (Neg-longPCR-5mC), as PCR products alone do not maintain methylation. As can be seen from Figure 5B, the purified mtDNA sample (mtDNA-5mC) had weaker signals than the positive control (Pos-longPCR-5mC) and stronger signals than the negative control (Neg-longPCR-5mC). Additionally, no bands were observed for the non-antibody control for MeDIP (mtDNA-NAC), and the non-template control (NTC) indicating that the mitochondrial genome can undergo DNA methylation.

Since the purified mtDNA sample had a weaker signal than the positive control, we performed the MeDIP assay (with sonication) on purified mtDNA samples and the

Pos-longPCR-5mC and the Neg-longPCR-5mC controls to quantify the levels of DNA methylation at different regions of the mitochondrial genome. The levels of DNA methylation were determined by firstly normalizing the qPCR signals of the MeDIP products to their corresponding input samples (5mC/Input), and then weighting against the positive and negative controls that were used to represent 100% and 0% levels of DNA methylation, respectively (Figure 5C). We assessed the non-coding regions (O_H , O_L , HSP and LSP), 2 rRNAs and all 13 ETC subunits using region-specific primers (Figure 5C; Supplementary Table S1). In HSR-GBM1 cells, methylated mtDNA molecules represent $\leq 29\%$ of the total population, and levels of DNA methylation in different regions varied. When comparing all regions with *ND5*, which was identified as having the highest level of DNA methylation, all the other regions were significantly lower (Figure 5C). Additionally, HSP and LSP were higher than the replication start sites O_H and O_L . DNA methylation levels at the two rRNAs were relatively low. Amongst the ETC subunits, *ND5*, *ND6*, *ND2*, *ATP6* and *CYTB* had relatively higher levels of DNA methylation than the other ETC-encoding genes. Interestingly, following treatment with the DNA demethylation agents, 5Aza and VitC, there were significantly reduced levels of DNA methylation at most of the regions except for *CO3*, *ND1*, *ND3*, O_L and *RNR2* (Figure 5C). In contrast to the results observed for the nuclear genome, 5Aza induced more extensive DNA demethylation to the mitochondrial genome than VitC with *CO1*, *CO2*, O_H and *RNR1* being only demethylated by 5Aza.

As HSP and LSP are the transcription start sites of the heavy and light strands, we therefore investigated the levels of gene expression for the 13 mtDNA-encoded genes and 2 rRNAs (Figure 5D). In contrast to the overall up-regulation observed for nuclear gene expression, overall down-regulation was observed for the mitochondrial-related genes following induction of DNA demethylation. Through both agents, *RNR2*, *ND3*, *ATP6* and *CO3* were significantly down-regulated by ~2-fold or more. 5Aza also triggered significant down-regulation in *ATP8*. However, there was no overall correlation between gene-specific DNA methylation and gene expression.

We also extended the scope to the chromosomal genes encoding the ETC subunits as they are also indicative of mitochondrial function and cellular metabolism. Through both agents, *CYCI*, encoding a subunit of complex III, had decreased levels of expression induced by both treatments, whereas *SDHC*, encoding a subunit of complex II, had increased levels of expression by >2-fold after 5Aza treatment (Figure 5D). Apart from this, the difference between these two agents reached significant levels in *SDHB/C/D* (encoding subunits of complex II) (Figure 5D). 5Aza promoted the expression of the tumor suppressors *SDHC* more than VitC, whereas VitC repressed the expression of *SDHB/D*.

DISCUSSION

The replication of mtDNA is strictly regulated during differentiation and tumorigenesis to ensure that cellular energy through OXPHOS is generated in a cell type-specific manner (5,6,32,72). Pluripotent and cancer cells have lower

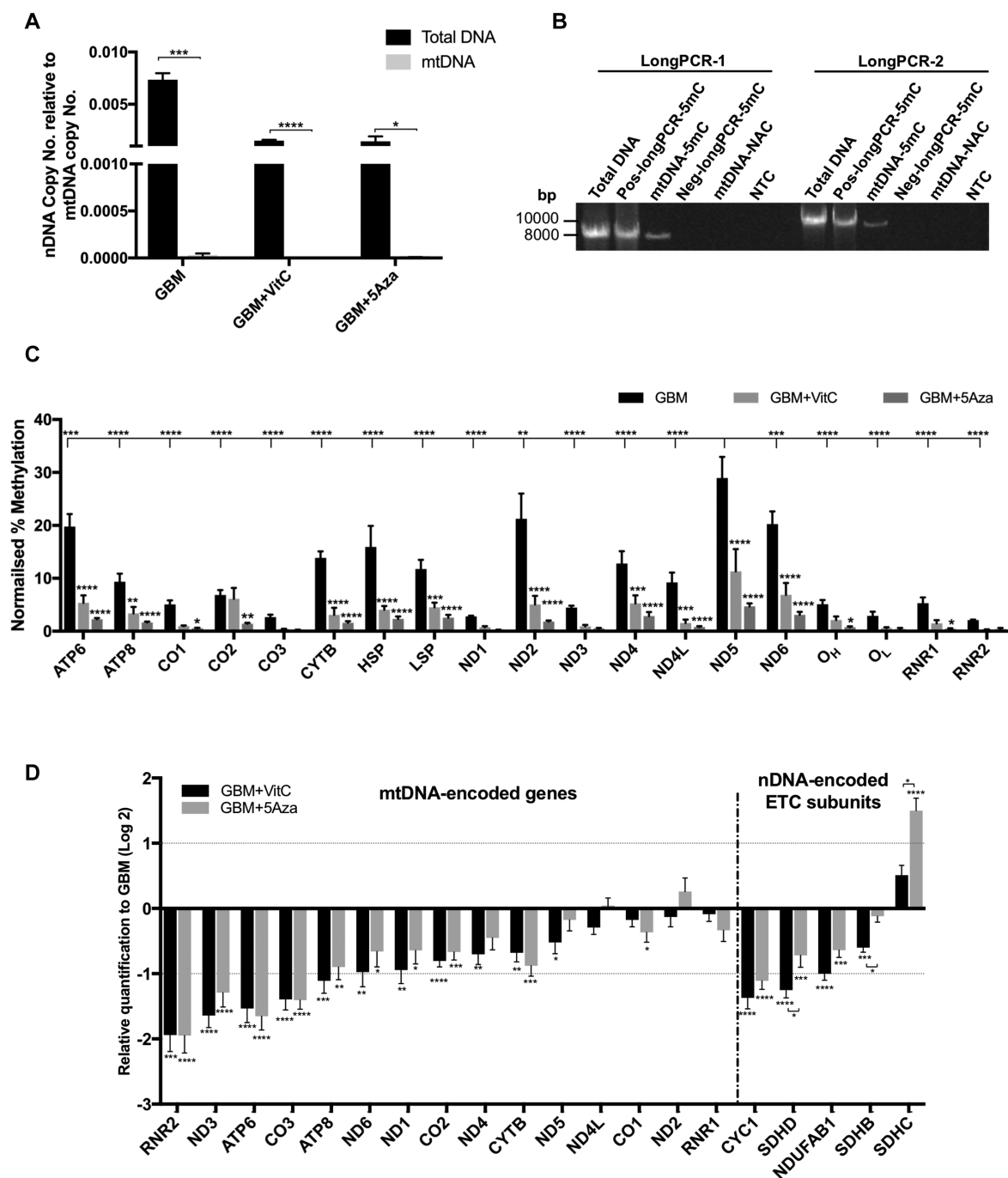


Figure 5. Levels of DNA demethylation in the mitochondrial genome and expression levels of the mtDNA encoded genes. (A) Detection of nDNA contamination after mtDNA purification. Copy number for nDNA (β -globin) relative to the copy number for mtDNA was plotted. Multiple t-tests were performed between the purified mtDNA samples (gray bars) and corresponding total DNA samples (black bars) isolated from GBM cells. Bars represent the mean \pm SEM ($n = 3$). (B) Agarose gel showing long-PCR products from purified mtDNA samples to demonstrate the presence of mtDNA methylation. Long PCR using two pairs of primers (primer set 2) spanning across the mitochondrial genome was performed on: (i) total DNA isolated from non-treated GBM cells (Total DNA); (ii) the positive control - MeDIP performed on the long PCR products previously generated with primer set 1 and treated with the *M. SssI* enzyme (Pos-longPCR-5mC); (iii) purified whole mtDNA having undergone MeDIP - (mtDNA-5mC); (iv) the negative control - MeDIP performed on the long PCR products previously generated with primer set 1 (Neg-longPCR-5mC); (v) the non-antibody control for MeDIP - MeDIP was performed on the purified mtDNA without using the 5mC antibody (mtDNA-NAC); and (vi) a non-template control (H_2O) for PCR (NTC). Fragment sizes are indicated. (C) Normalized levels of DNA methylation within regions of the mitochondrial genome determined on purified mtDNA samples that had undergone MeDIP with sonication. The normalized levels of DNA methylation were determined by normalizing the MeDIP results (5mC/Input) against the positive and negative controls that were used to represent 100% and 0% of DNA methylation, respectively. Statistical significance was determined between the 5Aza and VitC treated cohorts and the GBM cohort and between *ND5* and all the other genes by two-way ANOVA ($n = 3$). (D) Significant differential expression of the mitochondrial and chromosomal genes encoding the ETC subunits identified in the treated and non-treated cohorts using the Fluidigm array. Bars represent the mean of the relative quantification levels normalized to the GBM cohort ($n = 6$). *, **, ***, **** indicate P values of < 0.05 , 0.01 , 0.001 , 0.0001 , respectively.

mtDNA copy number and, thus, primarily rely on glycolysis to promote their high rates of proliferation (4). HSR-GBM1 cells are stem-like cancer cells with restricted potential to differentiate and exhibit a hypermethylated genome (35). We have previously shown that the levels of DNA methylation at exon 2 of *POLG* are negatively correlated with mtDNA copy number in HSR-GBM1 cells to restrict their potential to complete differentiation (4). However, in the presence of DNA demethylation agents, such as 5Aza and VitC, global demethylation, including demethylation at exon 2 of *POLG*, promotes synchrony between mtDNA replication and chromosomal gene expression (4). Consequently, these cells are able to undergo differentiation in a manner similar to adult stem cells. Likewise, the partial and near complete depletion of mtDNA in cancer cells alters the levels of DNA methylation at exon 2 of *POLG* and the tumorigenic capacity of these cells (5).

Here, we have shown that the global DMRs and differentially expressed genes induced by DNA demethylation through 5Aza and VitC overlap greatly with associated cancer pathways, such as cell death, cell proliferation and glioblastoma, by altering the expression levels of tumorigenic-specific genes. Both up- and down-regulation were observed in the RNA-Seq datasets due to DNA demethylation, which probably not only resulted from the direct impact on the targeted genes but also from sequential responses. To this extent, we have shown that tumour markers that overlapped with DMRs included *PRDM16* and *WNK2*, which were up-regulated by 5Aza and VitC, whereas the oncogene of GBM, *SEC61G*, was significantly down-regulated as a result of both treatments. *SEC61G* is located upstream of *EGFR*, which is a biomarker of GBM (61). Indeed, *PRDM16*, *WNK2* and *SEC61G* were identified as DMR-overlapping differentially expressed genes that aligned with the pathway analysis to show the DNA demethylation treatments primarily target cell death, cell growth and cell differentiation in GBM.

In terms of the mtDNA-specific transcription and replication factors, the results firstly confirmed our previous finding that, by using DNA demethylation agents, DNA methylation levels decreased across the key regulatory region, exon 2, of *POLG* and, as a result, the expression of *POLG* was significantly increased. In addition, we have identified differentially methylated CGIs in the gene body of *TOP1MT*, especially at exon 8 and intron 10 by both treatments. The existence of several differentially methylated CGIs in *TOP1MT* strongly indicates the transcriptional control by DNA methylation on this gene. Indeed, *TOP1MT* was found to be significantly up-regulated by 5Aza to over 2-fold. However, TET-mediated DNA demethylation induced by VitC also increased its gene expression but not to significant levels.

We further investigated the impact of DNA demethylation on the other mtDNA transcription and replication factors. Overall, up-regulation of expression profiles was observed, which is likely to be the major contributor to the up-regulation of mtDNA copy number. The transcription and replication factors including *NRF1/2* and *STAT3* were up-regulated by 5Aza and VitC to over 2-fold, whilst 5Aza had greater effect in promoting 10 of the transcription and replication factors to over 2-fold including the

direct-binding mtDNA transcription and replication factors *TWNK*, *TOP1MT*, *POLRMT* and *POLG*. Interestingly, *SSBP1* was down-regulated by both treatments. However, it is the gene that had the highest level of methylation for the promoter CGIs amongst the mtDNA transcription and replication factors. Therefore, there is likely to be other regulatory machinery offsetting the up-regulatory role of demethylation at the promoter region of *SSBP1* and causing the down-regulation of *SSBP1*. Furthermore, *TFB1M*, *SIRT3*, *PGC1 α* and *SIRT3* were down-regulated by VitC but slightly up-regulated by 5Aza. Although the DNA demethylation levels that occurred over their promoter regions were similar following the two treatments, the opposite patterns presented for gene expression likely resulting again from other patterns of regulation, such as to the upstream regulators of *TFB1M*, which include the NRFs and *PGC1 α* (24), rather than from DNA demethylation to *TFB1M* itself.

Additionally, IDH is tightly involved in mitochondrial function through the TCA cycle and considered to contribute to the hypermethylation of many types of cancers including GBM (33,37–39). In the TCA cycle, mutant IDH can alter the metabolic pathway to generate (*R*)-2-hydroxyglutarate, known as an oncometabolite, instead of α -ketoglutarate (33,37,40). This metabolic change disrupts the normal function of the TET enzymes, and results in the hypermethylated genomes of cancers (33,37,40). Our results showed that there was an impact of DNA demethylation on the expression levels of *IDH1* and *IDH2*. *IDH2* could be significantly up-regulated by 5Aza, while *IDH1* was down-regulated only by VitC. *IDH1* encodes the cytosolic IDH isoform, whereas *IDH2* encodes the mitochondrial isoform (39). Moreover, our finding also suggests that up-regulation of the TET family after DNA demethylation treatments compensates for or modulates DNA methylation, which was likely contributed by the up-regulation of *IDH2* in the mitochondria. Thus, our finding suggests that the induced DNA demethylation could interfere with the TCA cycle in the mitochondria, and the differences in changes to gene expression between *IDH1* and *IDH2* might indicate a response to the treatments according to their different sites of action.

Apart from the DMR-overlapping differentially expressed genes, we also analysed other tumor markers including *BCL2*, *TP53*, *EGFR*, *SHC2*, *TERT* and *MYC* that interact tightly with the mtDNA transcription and replication factors in GBM (66–69,73) (Figure 4D). Up-regulation occurred in nearly all of these factors after both treatments, amongst which *EGFR*, *BCL2* and *TERT* were up-regulated by 5Aza by over 2-fold while 5Aza repressed the expression of *MYC* (>1.5 fold). As an oncogene and an important transcription factor for cell proliferation, the different consequences of the up- or down- regulation of *MYC* by VitC and 5Aza could lead to various patterns of downstream metabolic reactions in cancer cells (64). *PGC1 α* , repressed by *MYC* in cancer cells, plays vital roles in regulating mitochondrial function and metabolic features of cancer stem cells (66,74). Our findings show that the levels of gene expression for *PGC1 α* were up-regulated by 5Aza, but not by VitC. Moreover, *MYC* is also reported to have a direct role in regulating mitochondrial biogenesis by binding

to the promoter regions of several key mtDNA transcription and replication factors including *POLG* and *POLG2* (75). *POLG* was significantly up-regulated after both DNA demethylation treatments. These changes could result from DNA demethylation through their direct effect on their own specific gene expression, for example the CGI at exon 2 of *POLG*, or indirectly via their upstream regulators, for example *MYC*.

Even though it is thought that mtDNA is not as extensively methylated as the nuclear genome, there are methylated sites in mtDNA that have previously been reported with unknown functions associated with mtDNA transcription (59,76). Our results confirmed the existence of DNA methylation in mtDNA by using a MeDIP assay. In general, the level of mtDNA methylation in GBM cells was <29% relative to the fully methylated positive control (Figure 5C). Even though mtDNA methylation was not present at a high level as shown by us and others (59,76), DNA demethylation agents were able to take effect in the mitochondrial genome. DNMT and TET enzymes have been found in mitochondria (59,77), which provide explanations as to why VitC, the activator of TET, and 5Aza, the inhibitor of DNMT, were able to significantly reduce the levels of mtDNA methylation. The control region of mtDNA underwent significant demethylation by 5Aza and VitC. The control region contains the start sites of transcription for both strands (HSP and LSP) and the origin of replication for the heavy strand (O_H), where the mtDNA transcription and replication factors bind to initiate transcription and replication of mtDNA (1). Significant DNA demethylation was observed at HSP and LSP induced by both agents. 5Aza also significantly demethylated the O_H site. The methylation status of this region could potentially affect the accessibility of the transcription and replication factors, similar to events observed in the nuclear genome, where, for example, the methylation status of exon 2 of *Polg* is associated with the binding of RNA polymerase II (32). Other than the control region, induced DNA demethylation at most of the mitochondrial genes was observed in both treated cohorts except for *CO3*, *ND1*, *ND3*, O_L and *RNR2*.

In contrast to the negative correlation observed between DNA demethylation and up-regulation in nuclear gene expression, the patterns of mitochondrial gene expression showed overall down-regulation after DNA demethylation. Indeed, the down-regulation of mtDNA expression induced by the DNA demethylation agents is intriguing given that there are increased numbers of mtDNA copy induced by the DNA demethylation agents that would provide more template for mtDNA transcription. Consequently, in the presence of more templates, DNA demethylation repressed the transcription of the primary polycistronic transcripts of mtDNA. *RNR2*, *ND3*, *ATP6*, *CO3* and *ATP8* were significantly down-regulated to around 2-fold or more by both treatments. In synchrony with this, the transcription factors *TFAM* and *TFB1/2M* were also down-regulated or remained at similar levels of expression compared to the up-regulation of the other replication factors. Therefore, one or more factors could mediate the overall down-regulation of mtDNA expression.

Apart from the overall down-regulation, the change in expression for each gene varies, which led us to determine

whether there were gene-specific correlations between the levels of DNA demethylation and transcriptional changes. However, we did not observe any gene-specific correlations. Gene-specific changes to mitochondrial genes that are associated with DNA methylation have been reported in p53-knockout mouse embryonic fibroblasts (59) and diabetic retinopathy (78). However, no validated mechanism has been fully described. The gene-specific expression did not result from the direct effect of DNA methylation, as it probably could only affect transcription at the level of the polycistronic RNA, likely from the potential indirect effects of DNA methylation on the post-transcriptional process. The tRNA punctuation model has been widely proposed to be the primary machinery of post-transcriptional processing of the polycistronic transcripts in human mitochondria (79,80). *CO3*, *ATP6* and *ATP8* are located sequentially on the mitochondrial genome without a recurring tRNA in between, which likely explains why the changes to gene expression induced by DNA demethylation remained similar over the three genes after processing. Indeed, whether mtDNA methylation adds an epigenetic layer in regulating mtDNA transcription has not been fully understood.

Whilst it has been reported that induced CpG methylation by the mitochondrial targeted *M.SssI* enzyme decreased mtDNA copy number in human colon carcinoma HCT116 cells (76,81), we have observed increases in mtDNA copy number resulting from DNA demethylation induced by 5Aza and VitC but decreases in mtDNA transcription. We propose that mtDNA demethylation ensures that the mitochondrial genome is accessible for replication through fewer CpG sites being methylated. DNA demethylation induced by these agents also results in DNA demethylation of key mtDNA replication factors, such as *POLG*, and increases in their expression to further promote mtDNA replication. Even though mtDNA transcription was down-regulated, the RNA-DNA hybrid primer would appear to be available to initiate replication, as recently shown (76,81). Whether transcription is by-passed or specifically down regulated remains to be determined. However, DNA demethylation induced the two genomes to promote mtDNA replication in synchrony with altered patterns of gene expression of the nuclear genome, which would be indicative of a cell going through differentiation unlike mature cells that would likely have less replication and prioritize transcription to promote energy production (6,32). Indeed, it is worth pointing out that there are short waves of mtDNA replication which do not reflect the need for ATP-derived by OXPHOS, as, for example, when mouse embryonic stem cells differentiate into mature cell types (2). As with DNA demethylated GBM cells, these events could represent a checking mechanism to ensure synchrony between the nuclear and mitochondrial genomes in order that cell differentiation can proceed. Consequently, during differentiation, it appears that mtDNA replication can take place independently of mtDNA transcription.

The differences observed for DNA demethylation and gene expression induced by 5Aza and VitC provide insights into their different modes of action in inducing DNA demethylation. 5Aza is an analogue of cytosine that competitively binds to DNMT1 so that DNMT1 could not bind to cytosines to modify methylation (43), whereas VitC

acts as the cofactor of TET1 to enhance its activity (42). The unique DMRs and different proportions of overlapping regions of the genome provide evidence of their different modes of actions. Based on the use of the same methylome provided by the HSR-GBM1 cells, VitC significantly demethylated more genomic regions, especially intragenic regions. However, there were DMRs uniquely regulated by both agents. There is limited knowledge as to whether these agents have specific targets or preference over certain genomic regions. We have shown that the induced DNA demethylation caused by these two treatments differentially regulated gene expression. Even though 5Aza induced fewer DMRs compared with VitC, it affected gene expression more effectively. Indeed, VitC has been shown to have a faster rebound effect than 5Aza after its withdrawal, which allows cells to return to their original states (4). It is likely that, as HSR-GBM1 cells have fast rates of cell division (4), the inhibition of DNMT1 by 5Aza prevented the DNA methylation patterns in the newly formed cells from being reestablished, and allowed the transcription factors to bind to the genome to promote gene expression. However, the TET enzymes, enhanced by VitC, were interacting extensively with the genome to carry out DNA demethylation but the occupancy of the TET enzymes on the genome might result in less accessibility of the transcription factors to promote gene expression. Additionally, we also found several interesting differences in the effects induced by 5Aza and VitC: (i) the changes to the expression of *IDH1* and *IDH2* (discussed above); (ii) the changes to the expression of the key regulator of mitochondrial function and tumorigenesis, *MYC*, were contrasting where VitC up-regulated *MYC* and 5Aza down-regulated *MYC*; and (iii) 5Aza induced greater DNA demethylation in mtDNA than VitC, but the changes to mtDNA expression were more significantly modulated by VitC, which was opposite to what was found in the nucleus.

In all, we have shown that global demethylation of the GBM genome induces changes to several key tumor markers, such as *SEC61G*, *PRDM16* and *WNK2*, and that several key pathways including glioblastoma, cell death, cell growth and cell differentiation are also affected as a result. We further show that increases in mtDNA copy number coincided with increases in the expression of many of the mtDNA replication factors, two of which appear to be modulated by DNA demethylation, namely *POLG* and *TOP1MT*, that were identified from our mtDNA replication factor-specific screen. Furthermore, the expression of the tumour markers *EGFR*, *BCL2*, *TERT* and *MYC*, which interact with the mtDNA transcription and replication factors, were also modulated by the DNA demethylation agents suggesting an association between the key regulators of cellular fate and mtDNA replication. These interactions likely empower HSR-GBM1 cells to undergo differentiation as we have previously shown after culture with both 5Aza and VitC (4) and to break free from their pseudo-differentiated state where they are trapped between two states, an undifferentiated and a fully differentiated state (7,8). It would, thus, appear that synchrony between nuclear gene expression and the regulation of mtDNA copy number, and thus resetting of the mtDNA set point, are impor-

tant to cellular differentiation and that hypermethylation in HSR-GBM1 cells impinges on this process.

DATA AVAILABILITY

MeDIP-Seq sequences in triplicates for HSR-GBM1 cells and the cells treated with 5Aza and VitC have been deposited in the Sequence Read Archive (SRA) under accession number SRP080899. CNV data and RNA sequencing data have been deposited in the Gene Expression Omnibus (GEO) under accession number GSE98693.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Dr Trevor Wilson and Dr Selva Kumari Ramasubramanian, The Medical Genomics Facility Monash Health Translation Precinct, for assistance in performing next generation sequencing and the Fluidigm high-throughput real-time PCR array, respectively. We are grateful to BioDiscovery, Inc. for the use of Nexus Copy Number software in our data analysis. We thank Dr Aidan Sudbury from the School of Mathematical Sciences, Monash University for assistance with statistical analyses.

FUNDING

Hudson Institute of Medical Research Discretionary Funds; Victorian Government's Operational Infrastructure Support Program; Australian Postgraduate Award (to X.S.).

Conflict of interest statement. None declared.

REFERENCES

- 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.
- Facucho-Oliveira, J.M., Alderson, J., Splinkings, E.C., Egginton, S. and St. John, J.C. (2007) Mitochondria DNA replication during differentiation of murine embryonic stem cells. *J. Cell Sci.*, **15**, 4025–4034.
- Facucho-Oliveira, J.M. and St John, J.C. (2009) The relationship between pluripotency and mitochondrial DNA proliferation during early embryo development and embryonic stem cell differentiation. *Stem Cell Rev.*, **5**, 140–158.
- 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.
- Lee, W.T., Cain, J.E., Cuddihy, A., Johnson, J., Dickinson, A., Yeung, K.Y., Kumar, B., Johns, T.G., Watkins, D.N., Spencer, A. *et al.* (2016) Mitochondrial DNA plasticity is an essential inducer of tumorigenesis. *Cell Death Discov.*, **2**, 16016.
- Dickinson, A., Yeung, K.Y., Donoghue, J., Baker, M.J., Kelly, R.D., McKenzie, M., Johns, T.G. and St John, J.C. (2013) The regulation of mitochondrial DNA copy number in glioblastoma cells. *Cell Death Differ.*, **20**, 1644–1653.
- Lee, W.T. and St John, J. (2015) The control of mitochondrial DNA replication during development and tumorigenesis. *Ann. N. Y. Acad. Sci.*, **1350**, 95–106.

8. Sun, X. and St John, J.C. (2016) The role of the mtDNA set point in differentiation, development and tumorigenesis. *Biochem. J.*, **473**, 2955–2971.
9. Xu, B. and Clayton, D.A. (1996) RNA-DNA hybrid formation at the human mitochondrial heavy-strand origin ceases at replication start sites: an implication for RNA-DNA hybrids serving as primers. *EMBO J.*, **15**, 3135–3143.
10. Xu, B. and Clayton, D.A. (1995) A persistent RNA-DNA hybrid is formed during transcription at a phylogenetically conserved mitochondrial DNA sequence. *Mol. Cell. Biol.*, **15**, 580–589.
11. Tiranti, V., Savoia, A., Forti, F., D'Apolito, M.F., Centra, M., Rocchi, M. and Zeviani, M. (1997) Identification of the gene encoding the human mitochondrial RNA polymerase (h-mtRPO) by cyberscreening of the Expressed Sequence Tags database. *Hum. Mol. Genet.*, **6**, 615–625.
12. Fisher, R.P. and Clayton, D.A. (1985) A transcription factor required for promoter recognition by human mitochondrial RNA polymerase. Accurate initiation at the heavy- and light-strand promoters dissected and reconstituted in vitro. *J. Biol. Chem.*, **260**, 11330–11338.
13. Fisher, R.P. and Clayton, D.A. (1988) Purification and characterization of human mitochondrial transcription factor 1. *Mol. Cell. Biol.*, **8**, 3496–3509.
14. Falkenberg, M., Gaspari, M., Rantanen, A., Trifunovic, A., Larsson, N.G. and Gustafsson, C.M. (2002) Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA. *Nat. Genet.*, **31**, 289–294.
15. Hillen, H.S., Parshin, A.V., Agaronyan, K., Morozov, Y.I., Graber, J.J., Chernev, A., Schwinghammer, K., Urlaub, H., Anikin, M., Cramer, P. et al. (2017) Mechanism of transcription anti-termination in human mitochondria. *Cell*, **171**, 1082–1093.
16. Kasamatsu, H., Grossman, L.I., Roberson, D.L., Watson, R. and Vinograd, J. (1974) The replication and structure of mitochondrial DNA in animal cells. *Cold Spring Harb. Symp. Quant. Biol.*, **38**, 281–288.
17. Clayton, D.A. (1982) Replication of animal mitochondrial DNA. *Cell*, **28**, 693–705.
18. Carrodeguas, J.A., Theis, K., Bogenhagen, D.F. and Kisker, C. (2001) Crystal structure and deletion analysis show that the accessory subunit of mammalian DNA polymerase gamma, Pol gamma B, functions as a homodimer. *Mol. Cell*, **7**, 43–54.
19. Kaguni, L.S. and Olson, M.W. (1989) Mismatch-specific 3'→5' exonuclease associated with the mitochondrial DNA polymerase from *Drosophila* embryos. *Proc. Natl. Acad. Sci. U.S.A.*, **86**, 6469–6473.
20. Lim, S.E., Longley, M.J. and Copeland, W.C. (1999) The mitochondrial p55 accessory subunit of human DNA polymerase gamma enhances DNA binding, promotes processive DNA synthesis, and confers N-ethylmaleimide resistance. *J. Biol. Chem.*, **274**, 38197–38203.
21. Takamatsu, C., Umeda, S., Ohsato, T., Ohno, T., Abe, Y., Fukuoh, A., Shinagawa, H., Hamasaki, N. and Kang, D. (2002) Regulation of mitochondrial D-loops by transcription factor A and single-stranded DNA-binding protein. *EMBO Rep.*, **3**, 451–456.
22. Korhonen, J.A., Gaspari, M. and Falkenberg, M. (2003) TWINKLE Has 5' → 3' DNA helicase activity and is specifically stimulated by mitochondrial single-stranded DNA-binding protein. *J. Biol. Chem.*, **278**, 48627–48632.
23. Zhang, H., Meng, L.H. and Pommier, Y. (2007) Mitochondrial topoisomerases and alternative splicing of the human TOP1mt gene. *Biochimie*, **89**, 474–481.
24. Gleyzer, N., Vercauteren, K. and Scarpulla, R.C. (2005) Control of mitochondrial transcription specificity factors (TFB1M and TFB2M) by nuclear respiratory factors (NRF-1 and NRF-2) and PGC-1 family coactivators. *Mol. Cell. Biol.*, **25**, 1354–1366.
25. Aquilano, K., Vigilanza, P., Baldelli, S., Paglicci, B., Rotilio, G. and Ciriolo, M.R. (2010) Peroxisome proliferator-activated receptor gamma co-activator 1alpha (PGC-1alpha) and sirtuin 1 (SIRT1) reside in mitochondria: possible direct function in mitochondrial biogenesis. *J. Biol. Chem.*, **285**, 21590–21599.
26. Kong, X., Wang, R., Xue, Y., Liu, X., Zhang, H., Chen, Y., Fang, F. and Chang, Y. (2010) Sirtuin 3, a new target of PGC-1alpha, plays an important role in the suppression of ROS and mitochondrial biogenesis. *PLoS One*, **5**, e11707.
27. Weitzel, J.M., Iwen, K.A. and Seitz, H.J. (2003) Regulation of mitochondrial biogenesis by thyroid hormone. *Exp. Physiol.*, **88**, 121–128.
28. Enriquez, J.A., Fernandez-Silva, P., Garrido-Perez, N., Lopez-Perez, M.J., Perez-Martos, A. and Montoya, J. (1999) Direct regulation of mitochondrial RNA synthesis by thyroid hormone. *Mol. Cell. Biol.*, **19**, 657–670.
29. Sato, I., Miyado, M., Miwa, Y. and Sunohara, M. (2006) Expression of nuclear and mitochondrial thyroid hormone receptors in postnatal rat tongue muscle. *Cells Tissues Organs*, **183**, 195–205.
30. Casas, F., Rochard, P., Rodier, A., Cassar-Malek, I., Marchal-Victorion, S., Wiesner, R.J., Cabello, G. and Wrutniak, C. (1999) A variant form of the nuclear triiodothyronine receptor c-ErbAalpha1 plays a direct role in regulation of mitochondrial RNA synthesis. *Mol. Cell. Biol.*, **19**, 7913–7924.
31. Fernandez-Vizarra, E., Enriquez, J.A., Perez-Martos, A., Montoya, J. and Fernandez-Silva, P. (2008) Mitochondrial gene expression is regulated at multiple levels and differentially in the heart and liver by thyroid hormones. *Curr. Genet.*, **54**, 13–22.
32. Kelly, R.D., Mahmud, A., McKenzie, M., Trounce, I.A. and St John, J.C. (2012) Mitochondrial DNA copy number is regulated in a tissue specific manner by DNA methylation of the nuclear-encoded DNA polymerase gamma A. *Nucleic Acids Res.*, **40**, 10124–10138.
33. Figueroa, M.E., Abdel-Wahab, O., Lu, C., Ward, P.S., Patel, J., Shih, A., Li, Y., Bhagwat, N., Vasanthakumar, A., Fernandez, H.F. et al. (2010) Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell*, **18**, 553–567.
34. Turcan, S., Rohle, D., Goenka, A., Walsh, L.A., Fang, F., Yilmaz, E., Campos, C., Fabius, A.W.M., Lu, C. and Ward, P.S. (2012) IDH1 mutation is sufficient to establish the glioma hypermethylation phenotype. *Nature*, **483**, 479–483.
35. Galli, R., Binda, E., Orfanelli, U., Cipelletti, B., Gritti, A., De Vitis, S., Fiocco, R., Foroni, C., Dimeco, F. and Vescovi, A. (2004) Isolation and characterization of tumorigenic, stem-like neural precursors from human glioblastoma. *Cancer Res.*, **64**, 7011–7021.
36. Hess, K.R., Broglio, K.R. and Bondy, M.L. (2004) Adult glioma incidence trends in the United States, 1977–2000. *Cancer*, **101**, 2293–2299.
37. Turcan, S., Rohle, D., Goenka, A., Walsh, L.A., Fang, F., Yilmaz, E., Campos, C., Fabius, A.W., Lu, C., Ward, P.S. et al. (2012) IDH1 mutation is sufficient to establish the glioma hypermethylation phenotype. *Nature*, **483**, 479–483.
38. Yan, H., Parsons, D.W., Jin, G., McLendon, R., Rasheed, B.A., Yuan, W., Kos, I., Batnig-Haberle, I., Jones, S., Riggins, G.J. et al. (2009) IDH1 and IDH2 mutations in gliomas. *N. Engl. J. Med.*, **360**, 765–773.
39. Cohen, A.L., Holmen, S.L. and Colman, H. (2013) IDH1 and IDH2 mutations in gliomas. *Curr. Neurol. Neurosci. Rep.*, **13**, 345.
40. Haseeb, A., Makki, M.S. and Haqqi, T.M. (2014) Modulation of ten-eleven translocation 1 (TET1), Isocitrate Dehydrogenase (IDH) expression, alpha-Ketoglutarate (alpha-KG), and DNA hydroxymethylation levels by interleukin-1beta in primary human chondrocytes. *J. Biol. Chem.*, **289**, 6877–6885.
41. Marie, S.K. and Shinjo, S.M. (2011) Metabolism and brain cancer. *Clinics (Sao Paulo)*, **66**(Suppl. 1), 33–43.
42. Blaschke, K., Ebata, K.T., Karimi, M.M., Zepeda-Martinez, J.A., Goyal, P., Mahapatra, S., Tam, A., Laird, D.J., Hirst, M., Rao, A. et al. (2013) Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature*, **500**, 222–226.
43. Jones, P.A. and Taylor, S.M. (1980) Cellular differentiation, cytidine analogs and DNA methylation. *Cell*, **20**, 85–93.
44. Kaminskas, E., Farrell, A.T., Wang, Y.C., Sridhara, R. and Pazdur, R. (2005) FDA drug approval summary: azacitidine (5-azacytidine, Vidaza) for injectable suspension. *Oncologist*, **10**, 176–182.
45. Kelly, T.K., De Carvalho, D.D. and Jones, P.A. (2010) Epigenetic modifications as therapeutic targets. *Nat. Biotechnol.*, **28**, 1069–1078.
46. Guo, F., Li, X., Liang, D., Li, T., Zhu, P., Guo, H., Wu, X., Wen, L., Gu, T.P., Hu, B. et al. (2014) Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. *Cell Stem Cell*, **15**, 447–459.
47. Brunner, A.L., Johnson, D.S., Kim, S.W., Valouev, A., Reddy, T.E., Neff, N.F., Anton, E., Medina, C., Nguyen, L., Chiao, E. et al. (2009) Distinct DNA methylation patterns characterize differentiated human embryonic stem cells and developing human fetal liver. *Genome Res.*, **19**, 1044–1056.

48. Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L. and Schubeler, D. (2005) Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.*, **37**, 853–862.
49. Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.
50. 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 Res.*, **20**, 1441–1450.
51. Lienhard, M., Grimm, 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.
52. Yu, G., Wang, L.G. and He, Q.Y. (2015) ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*, **31**, 2382–2383.
53. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
54. Liao, Y., Smyth, G.K. and Shi, W. (2013) The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.*, **41**, e108.
55. Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139–140.
56. McCarthy, D.J., Chen, Y. and Smyth, G.K. (2012) Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.*, **40**, 4288–4297.
57. Dvinge, H. and Bertone, P. (2009) HTqPCR: high-throughput analysis and visualization of quantitative real-time PCR data in R. *Bioinformatics*, **25**, 3325–3326.
58. Valente, V., Teixeira, S.A., Neder, L., Okamoto, O.K., Oba-Shinjo, S.M., Marie, S.K., Scrideli, C.A., Paco-Larson, M.L. and Carloti, C.G. Jr (2014) Selection of suitable housekeeping genes for expression analysis in glioblastoma using quantitative RT-PCR. *Ann. Neurosci.*, **21**, 62–63.
59. 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. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 3630–3635.
60. Robinson, M.D., Stirzaker, C., Statham, A.L., Coolen, M.W., Song, J.Z., Nair, S.S., Strbenac, D., Speed, T.P. and Clark, S.J. (2010) Evaluation of affinity-based genome-wide DNA methylation data: effects of CpG density, amplification bias, and copy number variation. *Genome Res.*, **20**, 1719–1729.
61. Lu, Z., Zhou, L., Killela, P., Rasheed, A.B., Di, C., Poe, W.E., McLendon, R.E., Bigner, D.D., Nicchitta, C. and Yan, H. (2009) Glioblastoma proto-oncogene SEC61gamma is required for tumor cell survival and response to endoplasmic reticulum stress. *Cancer Res.*, **69**, 9105–9111.
62. Moniz, S., Martinho, O., Pinto, F., Sousa, B., Loureiro, C., Oliveira, M.J., Moita, L.F., Honavar, M., Pinheiro, C., Pires, M. et al. (2013) Loss of WNK2 expression by promoter gene methylation occurs in adult gliomas and triggers Rac1-mediated tumour cell invasiveness. *Hum. Mol. Genet.*, **22**, 84–95.
63. Lei, Q., Liu, X., Fu, H., Sun, Y., Wang, L., Xu, G., Wang, W., Yu, Z., Liu, C., Li, P. et al. (2016) miR-101 reverses hypomethylation of the PRDM16 promoter to disrupt mitochondrial function in astrocytoma cells. *Oncotarget*, **7**, 5007–5022.
64. Yang, X., Han, H., De Carvalho, D.D., Lay, F.D., Jones, P.A. and Liang, G. (2014) Gene body methylation can alter gene expression and is a therapeutic target in cancer. *Cancer Cell*, **26**, 577–590.
65. Deaton, A.M. and Bird, A. (2011) CpG islands and the regulation of transcription. *Genes Dev.*, **25**, 1010–1022.
66. Morrish, F. and Hockenbery, D. (2014) MYC and mitochondrial biogenesis. *Cold Spring Harb. Perspect. Med.*, **4**, a014225.
67. Sahin, E., Colla, S., Liesa, M., Moslehi, J., Muller, F.L., Guo, M., Cooper, M., Kotton, D., Fabian, A.J., Walkey, C. et al. (2011) Telomere dysfunction induces metabolic and mitochondrial compromise. *Nature*, **470**, 359–365.
68. Autret, A. and Martin, S.J. (2009) Emerging role for members of the Bcl-2 family in mitochondrial morphogenesis. *Mol. Cell*, **36**, 355–363.
69. Suen, D.F., Norris, K.L. and Youle, R.J. (2008) Mitochondrial dynamics and apoptosis. *Genes Dev.*, **22**, 1577–1590.
70. Liu, B., Du, Q., Chen, L., Fu, G., Li, S., Fu, L., Zhang, X., Ma, C. and Bin, C. (2016) CpG methylation patterns of human mitochondrial DNA. *Sci. Rep.*, **6**, 23421.
71. Mposhi, A., Van der Wijst, M.G., Faber, K.N. and Rots, M.G. (2017) Regulation of mitochondrial gene expression, the epigenetic enigma. *Front. Biosci. (Landmark Ed.)*, **22**, 1099–1113.
72. Kelly, R.D., Rodda, A.E., Dickinson, A., Mahmud, A., Nefzger, C.M., Lee, W., Forsythe, J.S., Polo, J.M., Trounce, I.A., McKenzie, M. et al. (2013) Mitochondrial DNA haplotypes define gene expression patterns in pluripotent and differentiating embryonic stem cells. *Stem Cells*, **31**, 703–716.
73. Taylor, T.E., Furnari, F.B. and Cavenee, W.K. (2012) Targeting EGFR for treatment of glioblastoma: molecular basis to overcome resistance. *Curr. Cancer Drug Targets*, **12**, 197–209.
74. Sancho, P., Burgos-Ramos, E., Tavera, A., Kheir, T., Jagust, P., Schoenhals, M., Barneda, D., Sellers, K., Campos-Olivas, R. and Graña, O. (2015) MYC/PGC-1 balance determines the metabolic phenotype and plasticity of pancreatic cancer stem cells. *Cell Metab.*, **22**, 590–605.
75. Kim, J., Lee, J.H. and Iyer, V.R. (2008) Global identification of Myc target genes reveals its direct role in mitochondrial biogenesis and its E-box usage in vivo. *PLoS One*, **3**, e1798.
76. Bellizzi, D., D'Aquila, P., Scafione, 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.
77. Dzitoyeva, S., Chen, H. and Manev, H. (2012) Effect of aging on 5-hydroxymethylcytosine in brain mitochondria. *Neurobiol. Aging*, **33**, 2881–2891.
78. 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.
79. Ojala, D., Montoya, J. and Attardi, G. (1981) tRNA punctuation model of RNA processing in human mitochondria. *Nature*, **290**, 470–474.
80. Mercer, T.R., Neph, S., Dinger, M.E., Crawford, J., Smith, M.A., Shearwood, A.M., Haugen, E., Bracken, C.P., Rackham, O., Stamatoyannopoulos, J.A. et al. (2011) The human mitochondrial transcriptome. *Cell*, **146**, 645–658.
81. van der Wijst, M.G., van Tilburg, A.Y., Ruiters, M.H. and Rots, M.G. (2017) Experimental mitochondria-targeted DNA methylation identifies GpC methylation, not CpG methylation, as potential regulator of mitochondrial gene expression. *Sci. Rep.*, **7**, 177.

Supplementary Files

Global DNA methylation synergistically regulates the nuclear and mitochondrial genomes in glioblastoma cells.

Xin Sun^{1,2}, Jacqueline Johnson¹, Justin C. St. John^{1,2}

¹ Centre for Genetic Diseases, Hudson Institute of Medical Research, 27-31 Wright Street, Clayton VIC 3168, AUSTRALIA

² Department of Molecular and Translational Sciences, Monash University, 27-31 Wright Street, Clayton VIC 3168, AUSTRALIA

Corresponding author: Justin St. John, Voice: +61 3 8572 2678, Fax: +61 03 9594 7416;
Email: Justin.StJohn@hudson.org.au

Supplementary Table 1: Primer pairs for long and real time PCR

Human Gene	Forward primer (5'-3')	Product size	Annealing Tm (°C)
β-globin-F	CAACTTCATCCACGTTCCACC	268	57
β-globin-R	GAAGAGCCAAGGACAGGTAC		
mtDNA-F	CGAAAGGACAAGAGAAATAAGG	152	53
mtDNA-R	CTGTAAAGTTTTAAGTTTATGCG		
Long1-F (Set 1)	GACGGGCTCACATCACCCATAA	8382	63
Long1-R (Set 1)	GCGTACGGCCAGGGCTATTGGT		
Long2-F (Set 1)	GCCACAATAACCTCCTCGGACTCCT	8703	63
Long2-R (Set 1)	GGTGGCTGGCACGAAATTGACC		
Long1-F (Set 2)	TACACATGCAAGCATCCCCGTTCCA	7871	63
Long1-R (Set 2)	GGGCAATGAATGAAGCGAACAGATTTTCG		
Long2-F (Set 2)	CAGTGATTATAGGCTTTCGCTCTAAG	7985	63
Long2-R (Set 2)	TCTGTGTGGAAAGTGGCTGTGCAGA		
RNR1-F	TACACATGCAAGCATCCCCG	276	57.5
RNR1-R	GGGGAGGGGGTGATCTAAAA		
RNR2-F	GGGCATAACACAGCAAGACG	295	56.5
RNR2-R	AACATCGAGGTCGTAAACCCT		
CO1-F	CTTCTCCTTACACCTAGCAG	143	55.5
CO1-R	GTAGGACTGCTGTGATTAGG		
ND4-F	CTGCCTACGACAAACAGAC	161	55.5
ND4-R	GGCAGAATAGTAATGAGGATG		
ND4L-F	GGATTAGACTGAACCGAATTG	176	54.5
ND4L-R	AGTAGGGAGGATATGAGGTG		
O _L -F	CAGCTAAGCACCCCTAATCAAC	162	55.5
O _L -R	TCTAAAGACAGGGGTTAGGC		
ND5-F	AAACCCCATTAACGCCTGG	254	56.5
ND5-R	TGCGGTGTGTGATGCTAGG		
CO3-F	TCACCATTTCCGACGGCATC	109	57.5
CO3-R	TGGCGGATGAAGCAGATAGTG		
ND3-F	ACCACAACCTAACGGCTACAT	164	56.5
ND3-R	GGGCTCATGGTAGGGGTAAA		
CYTB-F	ATGACCCCAATACGCAAACT	401	54.5
CYTB-R	GGGAGGACATAGCCTATGAA		
CO2-F	CTGCTTCCTAGTCCTGTATG	235	56.6
CO2-R	GTCGGTGTACTCGTAGGTTT		
ND6-F	CCGCACCAATAGGATCCTCCCGA	187	63.5
ND6-R	GCATGGGGGTCAGGGGTTGAG		
ATP6-F	CAGTGATTATAGGCTTTCGCTC	343	57.5
ATP6-R	GTGTTGTCGTGCAGGTAGAG		
ND2-F	ACCGTACAACCCCTAACATAACC	166	57.5
ND2-R	GAGGAGGGTGGATGGAATTAAG		
ATP8-F	CACTATTCCTCATCACCCAACTAA	140	56.5
ATP8-R	GGGCAATGAATGAAGCGAAC		
LSP-F	AGCCACTTTCCACACAGACATCATA	250	59
LSP-R	TAGGATGGGCGGGGTTGTA		
HSP-F	TCAATACAACCCCGCCCAT	257	57.5
HSP-R	TCGTGGTGATTAGAGGGTGAA		
O _H -F	GAGCACCCCTATGTCGCAGTA	170	57
O _H -R	TCTGTGTGGAAAGTGGCTGT		
POLG exon2-F	CAGACCTCCACGTGCAACAC	209	59.5
POLG exon2-R	GACAACCTGGACCAGCACTT		
TOP1MT intron10-F	AAGGACTGTGAGCCGTTGG	136	59.5
TOP1MT intron10-R	GACGACCTCTTCGACAGGC		
TOP1MT exon8-F	CCACCGGCACTCTGTTGTAG	197	59.5
TOP1MT exon8-R	ACCATTCCCCACCAACTAGC		

Supplementary Table 2: Analytical parameters used for each of the bioinformatics analysis

Burrows-Wheeler Aligner (BWA; 0.7.16a)	
index	-t 12 Homo_sapiens.GRCh38.dna.primary_assembly.fa
aln	-t 12
samse	
Samtools (1.4)	
view	
sort	
merge	
MEDIPS (1.24.0)	
extent	500
window size	100
shift	0
uniq	1
MEDIPS.couplingVector-pattern	CG
MEDIPS.couplingVector-RefObj	GBM[[1]]
MEDIPS.meth-diff.method	edgeR
MEDIPS.meth- MEDIPS	TRUE
MEDIPS.meth-minRowSum	100
MEDIPS.meth-p.adj	bonferroni
MEDIPS.selectSig-pvalue	0.05
MEDIPS.selectSig-agj	TRUE
HTqPCR(1.28)	
readCtData-na.value	40
normalizeCtData-norm	deltaCt
normalizeCtData-deltaCt.genes	OAZ1;HPRT1;ActB/18SrRNA
limmaCtData	
STAR(2.5.3a)	
genomeFastaFiles	Homo_sapiens.GRCh38.dna.primary_assembly.fa
sjdbGTFfile	gencode.v26.annotation.gtf
Rsubread-FeatureCounts (1.26.0)	
annot.inbuilt	hg38
edgeR for RNA seq	
calcNormFactors-method	TMM
estimateGLMCommonDisp	
estimateGLMTrendedDisp	
estimateGLMTagwiseDisp	
glmFit	
glmLRT	

Supplementary Table 3. Taqman assays used in the Fluidigm qPCR arrays

Gene name	Taqman assay number
18SrRNA	Hs99999901_s1
Actb	Hs99999903_m1
ADGRB1	Hs00181777_m1
AGAP3	Hs01553103_g1
AGPAT4	Hs01088853_m1
ASL	Hs00902699_m1
BCL2	Hs00608023_m1
BCL3	Hs00180403_m1
C10orf2	Hs00958168_g1
C2CD4C	Hs00985894_s1
C7orf57	Hs01102699_m1
CCDC71L	Hs00703139_s1
CLIP2	Hs00185593_m1
CLU	Hs00156548_m1
COL22A1	Hs01377218_m1
COL4A2	Hs01098873_m1
CYC1	Hs00357718_g1
CYP4F11	Hs01680107_m1
DNMT1	Hs00154749_m1
DNMT3A	Hs01027166_m1
DNMT3B	Hs00171876_m1
DNMT3L	Hs01081364_m1
EGFR	Hs01076090_m1
ELN	Hs00355783_m1
ESRRA	Hs01067166_g1
Esrrb	Hs01584024_m1
FAM20C	Hs00911619_m1
FAM222A	Hs00757936_m1
FAM49A	Hs01003627_m1
FBXO17	Hs00227695_m1
FCGBP	Hs00175398_m1
FNDC10	Hs00972832_s1
FZR1	Hs00393592_m1
GAS6	Hs01090305_m1
GTF2IRD1	Hs00249456_m1
HIF1a	Hs00153153_m1
HPRT1	Hs02800695_m1
HSD17B14	Hs00212233_m1
HSPG2	Hs01078536_m1
IDH1	Hs01855675_s1
IDH2	Hs00158033_m1
KIF17	Hs00325418_m1
KIF21B	Hs01118430_m1
LRRC4B	Hs00297475_s1
LTBP4	Hs00943217_m1
LYPLA1	Hs00911024_g1
MAN1C1	Hs01057229_m1
MASP2	Hs00373722_m1
MEIS3	Hs00908777_m1
MEST	Hs00853380_g1
MT-ATP6	Hs02596862_g1
MT-ATP6	Hs02596862_g1
MT-ATP8	Hs02596863_g1
MT-CO1	Hs02596864_g1
MT-CO2	Hs02596865_g1
MT-CO3	Hs02596866_g1
MT-CYB	Hs02596867_s1
MT-ND1	Hs02596873_s1

Gene name	Taqman assay number
MT-ND2	Hs02596874_g1
MT-ND3	Hs02596875_s1
MT-ND4	Hs02596876_g1
MT-ND4L	Hs02596877_g1
MT-ND5	Hs02596878_g1
MT-ND6	Hs02596879_g1
MTRNR1	Hs02596859_g1
MTRNR2	Hs02596860_s1
MTRNR2	Hs02596860_s1
MYC	Hs00153408_m1
MYT1	Hs01027966_m1
NACAD	Hs00322395_m1
NDUFAB1	Hs00192290_m1
NFE2L2	Hs00975961_g1
NKAIN4	Hs03806373_m1
NRF1	Hs01031046_m1
OAZ1	Hs00427923_m1
PALM3	Hs00908820_g1
PDGFA	Hs00234994_m1
PLEC	Hs00356986_g1
PLXNA4	Hs01039085_m1
POLG	Hs00160298_m1
POLG2	Hs00200546_m1
POLRMT	Hs04187596_g1
PPARGC1A	Hs01016719_m1
PRDM16	Hs00223161_m1
PRICKLE2	Hs00291033_s1
PRKAG2	Hs00211903_m1
PTPRS	Hs01548375_m1
RAB7B	Hs01088520_m1
SDHB	Hs00268117_m1
SDHC	Hs01698067_s1
SDHD	Hs00829723_g1
SEC61G	Hs00414142_m1
SHC2	Hs01044373_m1
SIRT1	Hs01009006_m1
SIRT3	Hs00953477_m1
SLCO4A1	Hs00249583_m1
SMARCD3	Hs00162003_m1
SNX10	Hs00203362_m1
SPPL2B	Hs01086493_m1
SSBP1	Hs00995376_g1
ST3GAL1	Hs00161688_m1
STAT3	Hs01047580_m1
TERT	Hs00972650_m1
TET1	Hs00286756_m1
TET2	Hs00325999_m1
TET3	Hs00379125_m1
TFAM	Hs00273372_s1
TFB1M	Hs00274971_m1
TFB2M	Hs00915025_m1
TLE3	Hs01032572_m1
TOP1MT	Hs01080056_m1
TP53	Hs01034249_m1
TSNARE1	Hs00417029_g1
TTYH1	Hs01012139_m1
WNK2	Hs00396601_m1

Supplementary Table 4: DMRs overlapping with copy number gain regions

CN gain-chr	CN gain-start	CN gain-end	Overlapping with VitC-DMR (Start-End)	Overlapping with 5Aza-DMR (Start-End)	Gene annotation
chr1	211088244	211213171			
chr1	219926421	220192189			
chr1	223872383	224349137	224012401-224012800; 224013301-224014100; 224014301-224014400	224012401-224013000; 224013301-224013900; 224014401-224014600	RefSeq: NR_136593.1 Status: Validated
chr1	225662969	226344033			
chr1	235871996	236467557	236097001-236097300		RefSeq: NR_136593.1 Status: Validated
chr1	240509953	240721167			
chr2	23830787	24168230			
chr2	24279904	24347049			
chr2	25745122	25923850			
chr2	31822420	32698792			
chr2	36938466	37061471			
chr2	38546335	38790197			
chr2	44102040	44280450			
chr2	47359069	47475052			
chr2	47475248	47478329			
chr2	47806619	48081697			
chr2	55217933	55756611			
chr2	61356413	61981762			
chr2	88374324	88473434			
chr2	128092312	128149383			
chr2	169775289	169886104			
chr2	171836010	172029793			
chr2	200954453	201120309			
chr2	202146246	203571593			
chr2	207636569	207680279			
chr2	227736331	227801294			
chr2	231671719	231767569			
chr3	27369211	27607977			
chr3	32627446	32758839			
chr3	47510474	48192618			
chr3	49195300	49415508			
chr3	49972414	50088975			
chr3	57659275	57987445			
chr3	93794694	93992813			
chr3	109246829	109451710			
chr3	121461715	121575559			
chr3	125446706	125586874			
chr3	136327166	136748776			
chr3	141928486	142383422			
chr3	153007487	153145543			
chr3	155707030	155906268			
chr3	177222323	177276348			
chr3	183596159	183831367			
chr3	185514495	185631566			
chr3	185793497	185916445			

chr3	195311174	195412448				
chr3	196397702	196534331				
chr3	196653381	196850362				
chr4	25732795	25758882				
chr4	39352717	40173338				
chr4	40471689	40801700				
chr4	48810828	48869071				
chr4	56732553	57039205				
chr4	70769273	71116641				
chr4	88138224	88262858				
chr4	108749303	108839687				
chr4	112521720	112697473				
chr4	127921622	128185413				
chr4	128712913	128778268				
chr4	139151702	139343727				
chr4	164745950	164910752				
chr4	188083155	188339073				
chr5	32297962	32409457				
chr5	45888519	46398991				
chr5	50144698	50405648				
chr5	56027372	56160869				
chr5	69077621	70328678				
chr5	80265268	80383186				
chr5	126561616	126796341				
chr5	134653706	134818494				
chr5	157005426	157095755				
chr5	160096285	160228116				
chr5	180401363	180451188				
chr6	10602105	10622676				
chr6	10762201	10794068	10786801-10787000;10787101-10787200	10786801-10787000;10787101-10787200	MAK	
chr6	17299435	17688474				
chr6	18122311	18215201				
chr6	21572400	21701354				
chr6	30512251	30542286				
chr6	31070275	31074635				
chr6	31218147	31222799				
chr6	31586382	31590036				
chr6	32255921	32270672				
chr6	34590409	34823647				
chr6	42489126	42841567				
chr6	73203906	73640630				
chr6	87541110	87552228				
chr6	87584076	87800145				
chr6	88937618	89113196				
chr6	106646935	106691569				
chr6	106861004	107302026				
chr6	110697636	111057269				

chr6	134094530	134290122			
chr6	150898965	150927721			
chr6	154670203	154791629			
chr7	152474058	152553603			
chr7	152687630	152719713			
chr8	39968734	40003202			
chr8	42589865	42731214			
chr8	43548966	43976244			
chr8	45944781	46360719			
chr8	56000933	56084629			
chr8	66695906	67035970			
chr8	69867144	70008543			
chr8	73711496	74029019			
chr8	80150668	80454889			
chr8	94735791	94832007			
chr9	751662	902311			
chr9	6610145	6775089			
chr9	15360401	15417865			
chr9	15664504	15635380			
chr9	18988309	19331183			
chr9	26783678	27038539			
chr9	33807622	34365032			
chr9	36172516	36641994			
chr9	71803749	72152054			
chr9	76340497	76510161			
chr9	77167020	77231597			
chr9	83734515	83820030			
chr9	84345142	84421031			
chr9	86172815	86241971			
chr9	96263562	96844094			
chr9	104981387	105085205			
chr9	107432378	107544092			
chr9	111893087	112513884			
chr9	124849983	125364302			
chr9	129908240	130117264			
chr9	130538508	130623435			
chr10	11935705	12147092			
chr10	12212516	12320488			
chr10	13272317	13296478			
chr10	21373419	21636759			
chr10	30497226	30573492			
chr10	32152920	32195631			
chr10	34750153	34989940			
chr10	35216346	35261106			
chr10	36526302	36853269			
chr10	63089820	63416834			
chr10	67794325	68116947			

chr10		68157998	68197454			
chr10		68421227	69008370			
chr10		72374026	73528699			
chr10		74692645	74752987			
chr10		92112717	92793516			
chr11		3696967	3761630			
chr11		9224608	9585408			
chr11		18297908	18687715			
chr11		32497949	32570975			
chr11		33081598	33365070			
chr11		47335062	47335998			
chr11		47349863	47351432			
chr11		47614047	47962063			
chr11		55035600	55153102			
chr11		107756868	108142529			
chr12		676755	835156			
chr12		928247	1099344			
chr12		1200073	1282631			
chr12		7607454	7964524			
chr12		19229988	19571611			
chr12		31528597	31922673			
chr12		32011128	32082369			
chr12		37497617	38015281			
chr12		50163762	51302894			
chr12		56454563	56593167			
chr12		64455920	64626580			
chr12		95208795	95374695			
chr12		95536069	95694649			
chr12		96409093	96534052			
chr12		98437154	98612485			
chr12		101468859	101560180			
chr12		110115916	110645170			
chr12		111476554	111730027			
chr12		120834751	120980418			
chr12		121167560	121198618			
chr12		122249730	122773223			
chr13		19667630	19959928			
chr13		20860547	21006396			
chr13		28067988	28173463			
chr13		44641318	44695574			
chr13		44969959	45283991			
chr13		48013841	48057543			
chr13		48306330	48309352			
chr13		49448076	50008076			
chr13		72952170	73012708			
chr13		95287449	95387714			
chr14		18226020	18597531			

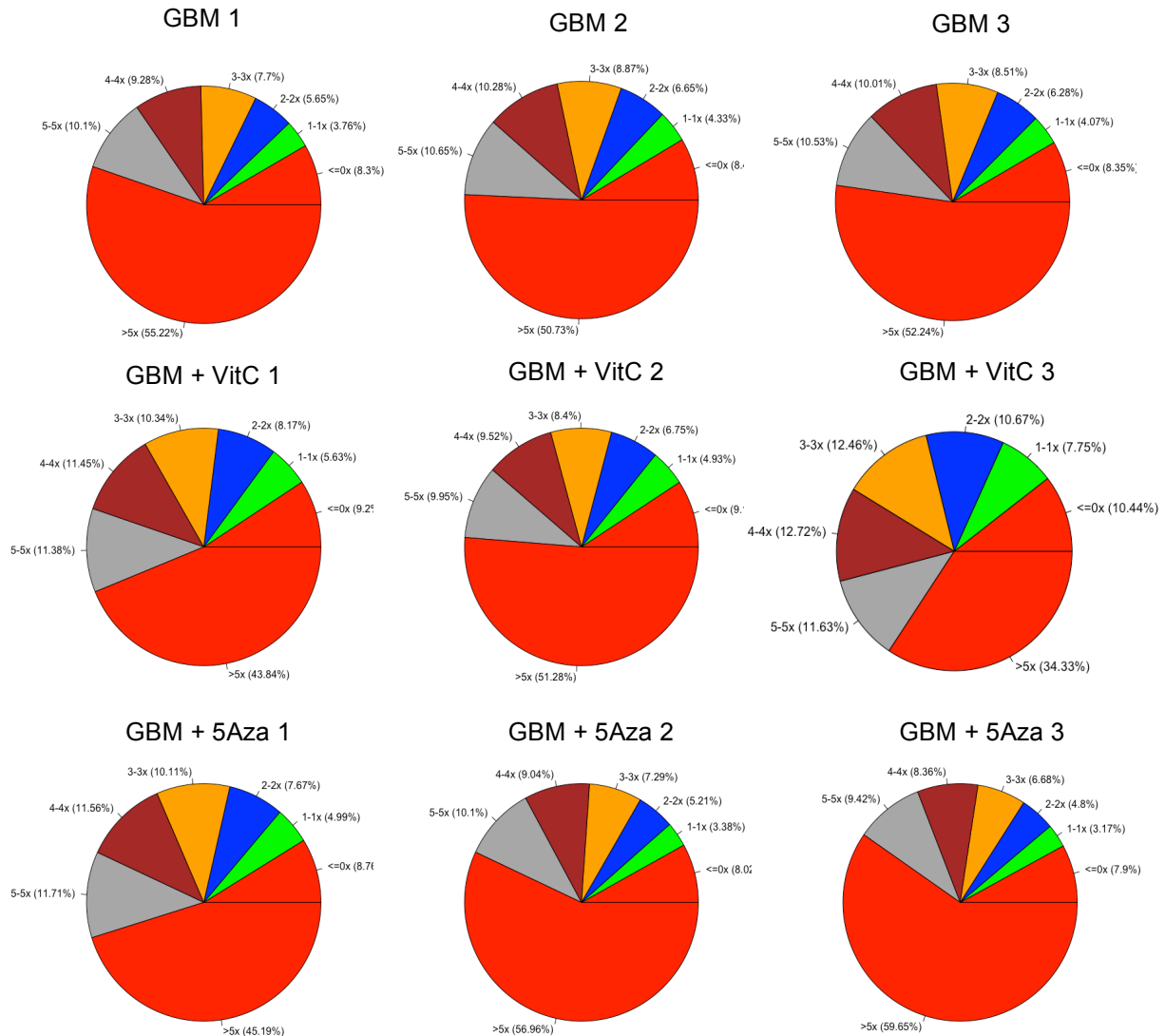
chr14	34414032	35392612		
chr14	49634985	49938786		
chr14	50126803	50298560		
chr14	63649056	63927604		
chr14	102536381	102607800		
chr15	22804033	23039644		
chr15	41030730	41413896		
chr15	44179266	44643219		
chr15	50258031	50737806		
chr15	64205413	64656404		
chr15	98562803	98588562		
chr16	9015048	9122794		
chr16	10264025	10492179		
chr16	18916138	19133410		
chr16	19438839	19616090		
chr16	35760404	36022890		
chr16	58675243	58832542		
chr16	68838934	69808567		
chr17	839486	934864		
chr17	1273696	1397225		
chr17	2488112	2670197		
chr17	2692871	2774484		
chr17	19837231	20050132		
chr17	26941116	27164482		
chr17	29299154	29492282		
chr17	37268235	37477828		
chr17	43063370	43092939		
chr17	43094516	43126420		
chr17	58631423	59375677		
chr17	59714615	59927560		
chr17	61598756	61693477		
chr17	61961298	62394188		
chr17	64544881	64696197		
chr18	20932565	21774317		
chr18	32028017	32147838		
chr18	57649601	57691915		
chr19	6633640	6643289		
chr19	6951095	7103959		
chr19	8134474	8208953		
chr19	11560254	11706482		
chr19	12597220	12649891		
chr19	14571552	14701255	14649801-14649900	ADGRE3
chr19	16880586	16817245	16754501-16754600	NWD1
chr19	20448323	21427482		
chr19	21689836	21838370		
chr19	23590034	23812918		
chr19	24049459	24435205	24179701-24179800	

chr19	27241251	27575884			
chr19	36432067	36632746			
chr19	53539770	53785562	53648201-53648600		
chr20	61563044	61634567			
chr20	63332435	63369826			
chr21	36112594	36192427			
chr22	17184683	17271460			
chr22	28696413	28786494			
chr22	31668448	31772794			
chr22	41003010	41174388			
chrX	32342021	32349467			
chrY	12733073	12846634			
chrY	13075160	13262055			
chrY	16701071	16989164			

Supplementary Table 5: Outputs from the Fluidigm qPCR array, as determined by the HTqPCR package

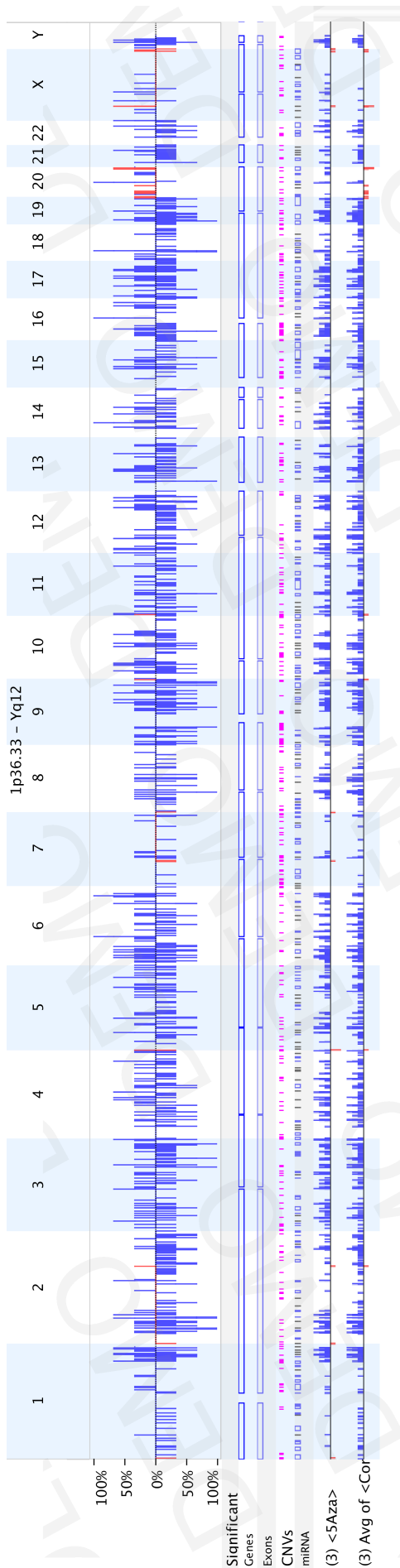
Gene	GBM+VitC vs. GBM-p.value	GBM+VitC vs. GBM-Mean of logFC	GBM+5Aza vs. GBM-p.value	GBM+5Aza vs. GBM-Mean of logFC	GBM+5Aza vs. GBM-SEM of logFC	GBM+5Aza vs. GBM-SEM of logFC
ABHD17A	0.8490108	0.531815748	0.471267664	0.093669061	2.035206082	0.888290972
ABR	0.000779018	1.637500694	0.219994705	4.85E-05	2.177948773	0.371490866
ADGRB1	0.901807243	0.4540474	0.688531043	0.22862308	1.34541738	0.755178138
ANGPT2	0.6272323	0.567438859	0.275089159	0.009109337	-0.087223653	0.700368489
ARHGEF10	0.794460223	0.532310264	0.639460279	0.006046615	2.405845201	0.931744972
ARHGEF10L	0.174136941	1.160629147	0.794525951	0.000274234	3.450230016	0.925723315
ARHGEF7	0.444790778	0.116748582	0.685152105	0.067912114	1.624173397	0.791331168
ASIC3	0.872555138	0.484506505	0.73726712	0.005279999	2.430836802	0.769438041
B4GALNT3	0.188721827	1.06277943	0.717833631	0.000671238	2.967967263	0.910452133
BAHCC1	0.899497648	0.512426905	1.152050092	0.00179199	4.028273369	1.040520502
BMP7	0.172011155	0.810481803	0.450915782	0.001697444	1.978753202	0.720904415
CCDC88C	0.909230707	0.340595972	0.861606777	0.000592546	3.489416587	0.930328459
COL22A1	0.716888933	0.25197351	0.687462134	0.038996333	1.853451457	0.86072692
COL26A1	0.660491347	0.64674025	0.615210997	0.195662046	1.609787815	0.992982915
COL4A2	0.082828394	-0.266978376	0.337412828	0.692421062	0.742221011	0.732515011
COL8A2	0.153290071	-3.877288984	3.861033534	0.342694564	3.67197231	0.912532026
CROCCP2	0.560795708	0.628885328	0.674540634	0.010594962	1.985174326	0.769986149
CUX1	0.418386418	0.762785899	0.640724588	0.00265748	2.970052799	0.849109794
CYP4F11	0.088127715	2.854609297	0.101145164	0.036593216	3.584800989	0.089738439
DCC	0.121052958	-0.379223094	0.352091878	0.70821687	0.705555742	0.907759156
DGCR8	0.524954124	0.735930006	0.860358038	0.000671606	3.09782705	0.829034681
DMRTA2	0.671078557	0.734988459	0.770399736	0.02259973	2.826544555	1.280642807
DNAAF5	0.438932546	0.112988297	0.419509272	0.036320556	1.748867489	0.794028996
EGFR	0.00571896	1.466617051	0.339930905	5.09E-05	2.517367905	0.426870256
ELFN1	0.058460855	4.692104609	0.536038994	0.030896726	5.82679696	0.705224219
ELN	0.596834795	0.061281599	0.934233274	0.000457229	3.843044038	0.998020589
ENO2	0.039918341	1.16520356	0.325032835	0.526237815	1.104155215	0.755054982
ERICH3	0.783017539	0.263015918	0.834178296	0.00115282	3.018152102	0.608623099
EXD3	0.950460622	0.43934338	0.92006191	0.008645986	2.383839331	0.697640694
FAM20C	0.236367676	0.954935219	0.727710299	0.000452798	2.922464468	0.770603566
FAM234B	0.851388205	0.478811176	0.595640685	0.015070112	1.91844585	0.82672592
FAM49A	0.155342164	0.638597776	0.272015875	0.004948555	1.394296545	0.413237723
FZR1	0.429534978	0.749092528	0.477328788	0.005174275	2.285187686	0.903674987
GET4	0.516247405	0.629857459	0.551541197	0.030264064	1.692469056	0.729928408
GPR35	0.184704867	1.007091361	0.819666101	0.05966931	1.769290586	0.782553808
HIVEP3	0.022023753	1.639298763	0.718840156	0.000197309	3.298153169	0.892548001
HSD11B1L	0.372407775	0.126234137	0.461965533	0.074957102	1.479384972	0.683149226
IQCE	0.749193741	0.307979753	0.596453248	0.005612192	1.904330012	0.699370676
JRK	0.183310936	1.079186094	0.749824521	6.94E-05	3.598726556	0.928025804
KIF17	0.042888221	1.054665064	0.384972797	0.010331208	1.749389442	0.731742367
KREMEN1	0.150784988	1.142841279	0.802936154	9.02E-05	3.548327874	0.85345768
KRT7	0.153486987	-2.970079528	3.194383723	0.947768315	1.036736719	0.909057389
LINC01006	0.304569383	0.836487065	0.692540835	0.169730658	1.470145974	0.751196993
MEIS3	0.563825434	0.290231711	0.570791593	0.956106046	0.876685084	0.527024162

MTCL1	0.884138659	0.469029593	0.73036444	0.000140759	3.028597104	0.794233994
MUC12	0.122763335	5.824927242	3.580190106	0.006655766	11.43736147	2.881270576
MUC3A	0.655353067	-1.008341301	0.351205852	0.150456344	3.348173427	2.720679524
MYH7B	0.445250458	0.097644419	0.570342434	0.001873163	2.414622386	0.833969134
MYO7A	0.250117449	-0.040480267	0.541774236	0.001857696	2.332150526	0.816459006
MYRF	0.346945364	1.045647979	0.980357665	0.002034504	3.384967777	1.001549113
MYREL	0.033506797	-5.397535906	3.393563686	0.573033184	2.303053497	0.770183617
NCAPG2	0.981581605	0.399490624	0.452877244	0.056127015	1.67059078	0.754820824
NDUFAF5	0.462794724	0.836094619	0.53627121	0.001808196	2.057395804	0.690379435
NFKBIB	0.810981403	0.33666836	0.445797632	0.073383711	1.461466479	0.692779129
NGEF	0.801052354	0.330759224	0.415671602	0.055146639	0.253319909	0.656874922
NKAIN4	0.778547483	-0.04463983	0.086030808	2.37E-06	1.11238846	0.133641299
NSUN6	0.403709282	0.174119076	0.396671019	0.023048073	1.58413776	0.632358839
NWD1	0.857355446	-0.372601495	3.623711857	0.248410402	6.030528405	3.126932966
OPRL1	0.084834048	1.225585995	0.733649909	0.000383764	2.953950872	0.809003863
P2RX5	0.183474667	0.089802362	0.371379764	0.623622491	1.001818772	0.599141145
PIP5K1C	0.629017123	0.642564811	0.804016979	0.00142785	2.783555429	0.82189893
PLEC	0.872214045	0.507208796	0.838106666	0.00092469	3.434203589	1.046629467
PLXNA4	0.000167548	2.292469529	0.317067072	0.002014605	1.731189525	0.406445124
POMT1	0.168035517	-0.028167793	0.460779463	0.765424874	0.9788782	0.685583749
PPP1R12C	0.000118611	1.386320935	0.156081694	1.43E-05	1.684477014	0.265178343
PRDM16	0.004437841	1.885771033	0.684583544	5.06E-06	4.051596416	0.816521399
PRKAR1B	0.07196897	0.876015598	0.486207345	4.51E-05	2.307608771	0.695967087
PRKCA	0.126022946	1.033887303	0.608802998	0.000344191	2.714153292	0.882939404
PRRT4	0.139283293	-0.47880361	0.835385891	0.420096376	1.357123279	0.989539503
PTPRN2	0.127327969	0.687940617	0.09231575	0.001251632	1.671138999	0.352194864
PTPRS	0.868916686	0.474686332	0.703179598	0.017645025	1.952494037	0.625356348
RFNG	0.599953728	0.584219477	0.565493026	0.001345222	2.208579384	0.7811225
SCNN1D	0.34004755	0.116942652	0.671369645	0.019633044	1.667745566	0.609978177
SDK1	0.347243942	0.840158146	0.658676427	0.000141135	3.220575872	0.912296081
SERINC2	0.015274435	-0.12683558	0.502223245	0.001161811	0.096546829	0.611766953
SH3RF3	0.13987869	0.848202747	0.494633341	1.24E-05	2.773435997	0.735752533
SHC2	2.02E-06	1.477460502	0.094607532	2.79E-06	1.439701646	0.167908811
SHC3	0.161511803	0.773608768	0.505351455	0.000978033	1.923219116	0.681860452
SLC04A1	0.035094065	-0.257425132	0.358249729	0.643159226	0.753365369	0.71470039
SPPL2B	0.000430854	1.335409552	0.224466492	1.19E-05	1.884552985	0.242299864
ST3GAL3	0.463185087	0.669989037	0.536358768	0.001622079	2.252057487	0.847530649
TNS1	0.351653674	0.843156807	0.624777998	0.015812416	2.132530706	0.941698187
TOP1MT	0.195548665	0.815352568	0.517910783	5.76E-05	2.61525311	0.756431561
TSC2	0.278842802	0.807584718	0.679278961	0.000189351	2.691165774	0.759809656
TSNARE1	0.621189961	0.544640625	0.522681019	0.000783903	2.053011817	0.743674529
VSTM2A	0.021010852	0.643538211	0.194012205	0.672437995	0.108329681	0.132271922
WNK2	0.019155665	1.970983345	0.940689734	4.80E-05	4.34589222	0.846637115
ZNF358	0.539483405	0.255344064	0.51564775	0.11489287	1.297478229	0.610122424

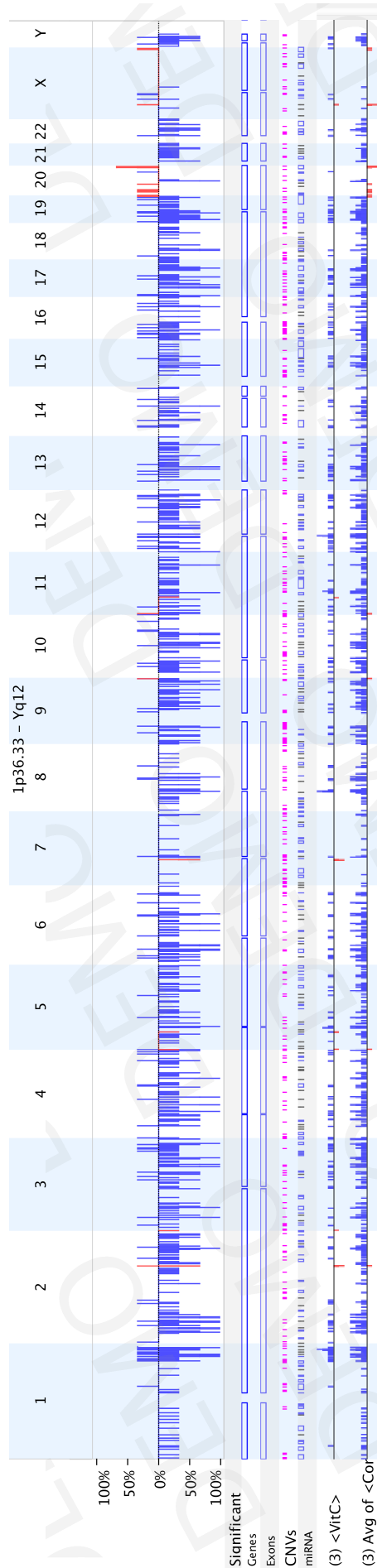


Supplementary Figure 1. Coverage analysis of CpGs from the MeDIP-Seq data using the MEDIPS package. The total number of CpG sites in the human genome was fractionised into pie charts according to the level of coverage. 7 fractions were used. <=0x in the upper red fraction: CpG sites that were not covered; 1-1x in the green fraction: CpG sites that were covered once; 2-2x in the blue fraction: CpG sites that were covered twice; 3-3x in the orange fraction: CpG sites covered three times; 4-4x in the brown fraction: CpG sites covered 4 times; 5-5x in the grey fraction: CpG sites covered 5 times; >5x in the lower red fraction: CpG sites covered more than 5 times. The first row is for the GBM cohort; the second row is for the GBM+VitC cohort; and the third row is for the GBM+5Aza cohort.

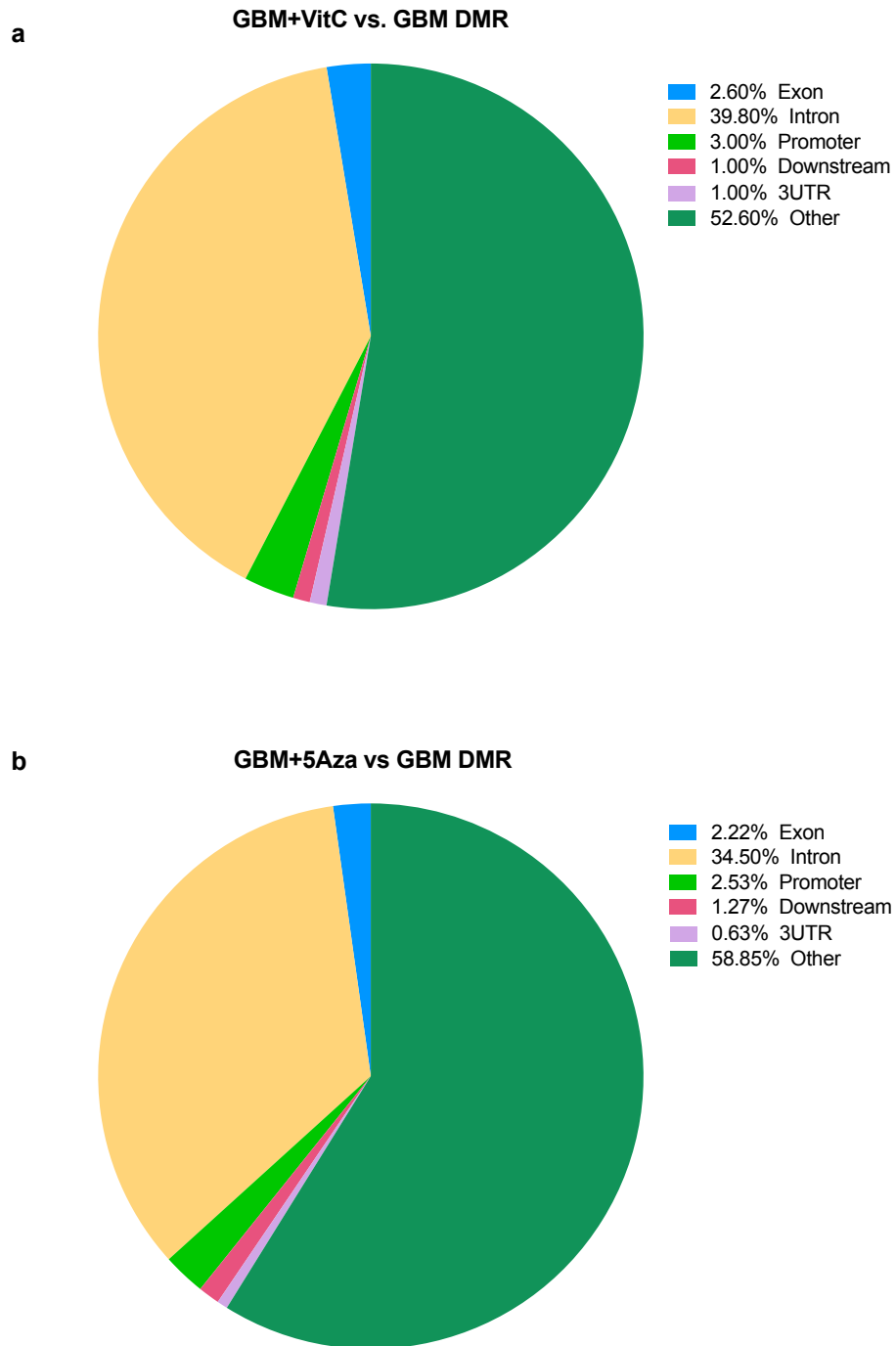
a)



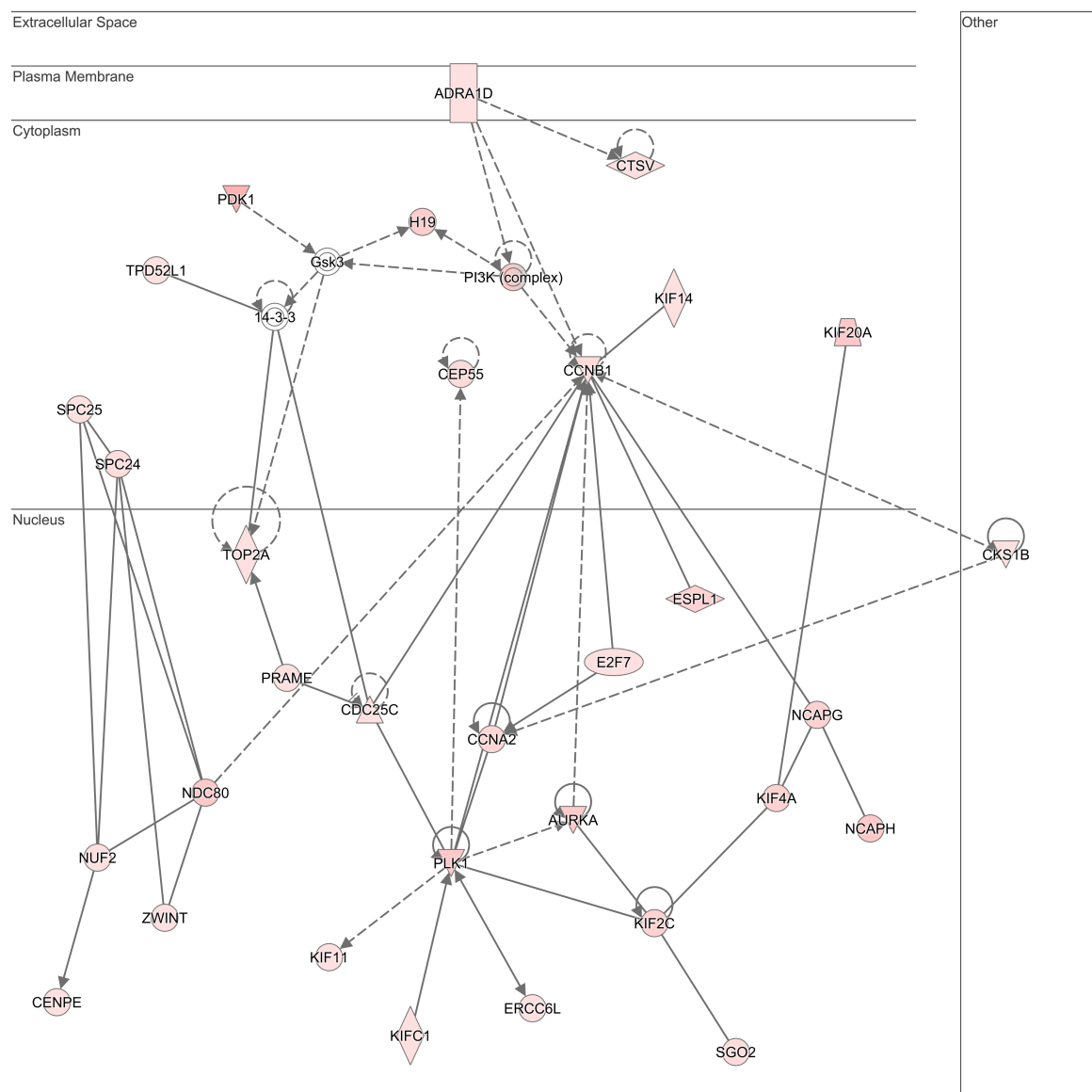
b)



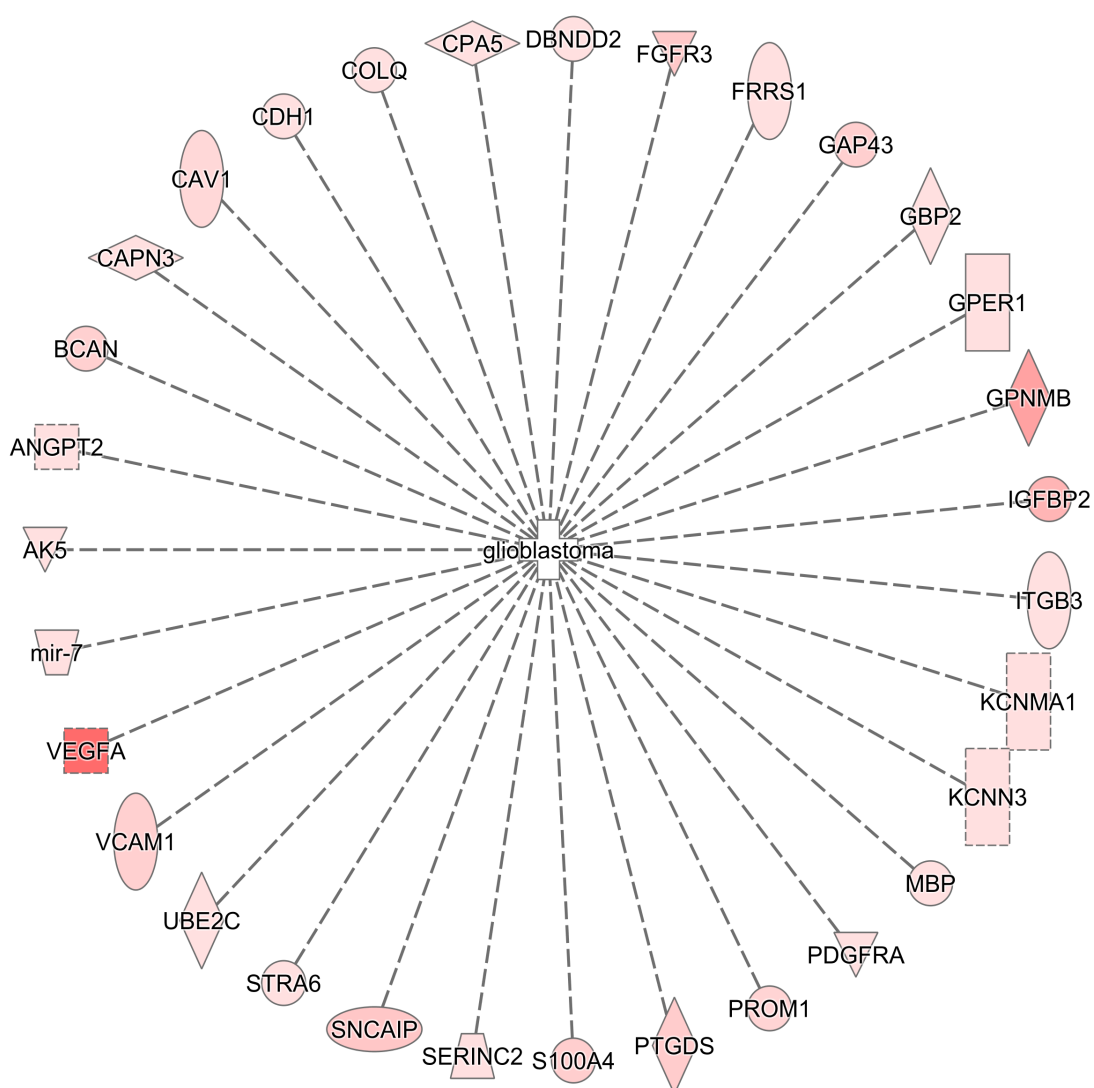
Supplementary Figure 2. Whole genome CNV comparisons between the GBM, the GBM+VitC and the GBM+5Aza cohorts, as determined by Nexus Copy Number Module (9.0). In comparisons between the: **a)** GBM+VitC and GBM cohorts; and **b)** GBM+5Aza and GBM cohorts, the upper panel shows copy number enrichment for each chromosome. Blue bars represent copy number gains, whereas red bars indicate copy number loss. Bars above 0% indicate CNVs identified in the treated cohorts, bars below 0% indicate CNVs in the GBM cohort. The panel 'Significant' indicates the significant CNVs between the two cohorts. The panels of 'Genes', 'Exons', 'CNVs', and 'miRNA' are genome annotations. The two panels at the bottom show the mean copy number value of the triplicates in each cohort.



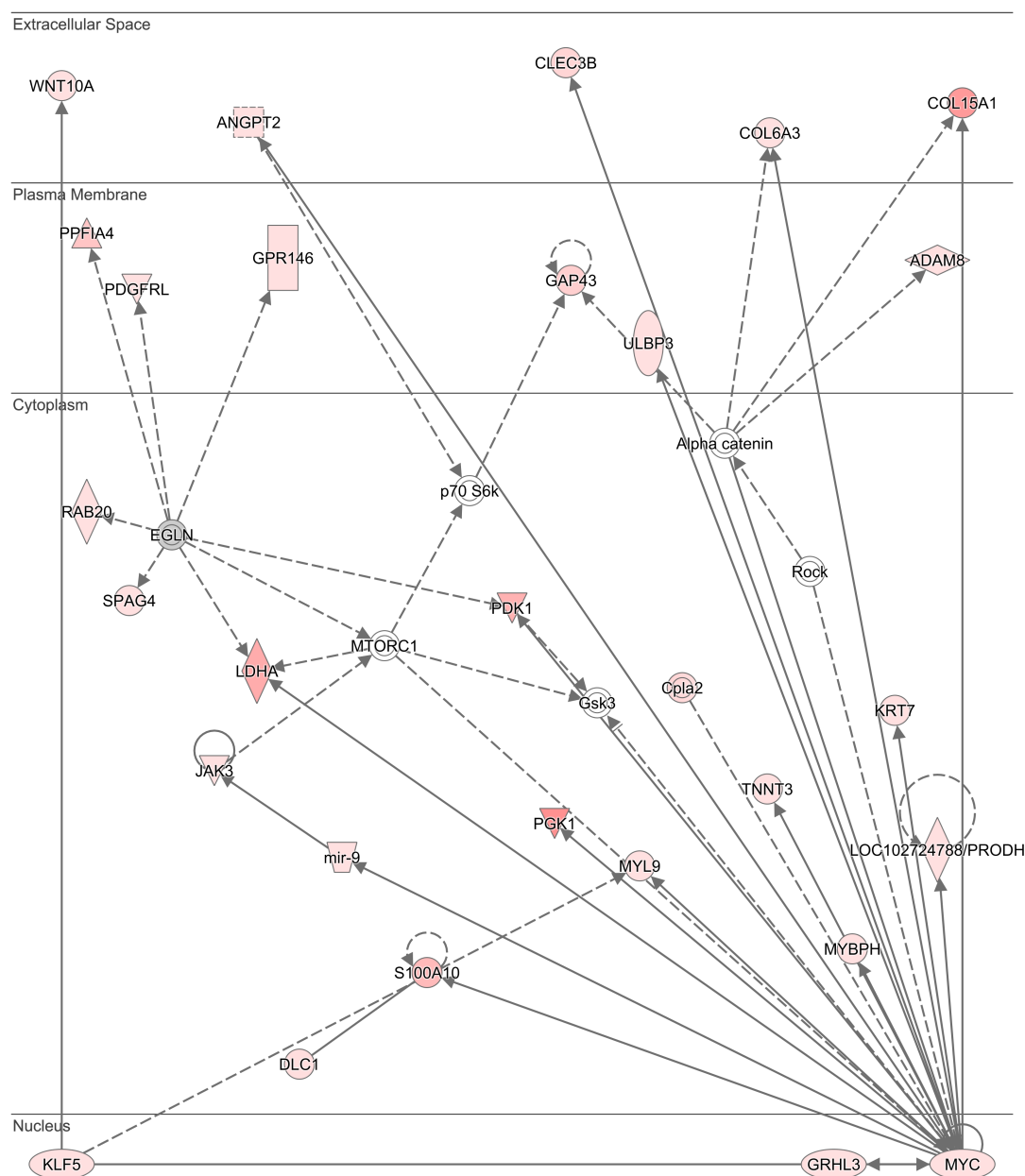
Supplementary Figure 3. Annotations of the DMRs. DMRs were sub-grouped into exon (blue), intron (yellow), promoter (green), downstream regions (≤ 3 kb; pink), 3UTR (purple) and intragenic regions (dark green) based on their genomic locations. The percentages of each part were calculated based on the number of corresponding DMRs out of the total number of DMRs. The basal distribution for genomic elements, based on size, in the human genome is 1.5% for exons, 24% for introns and 75% for other intergenic regions (1).



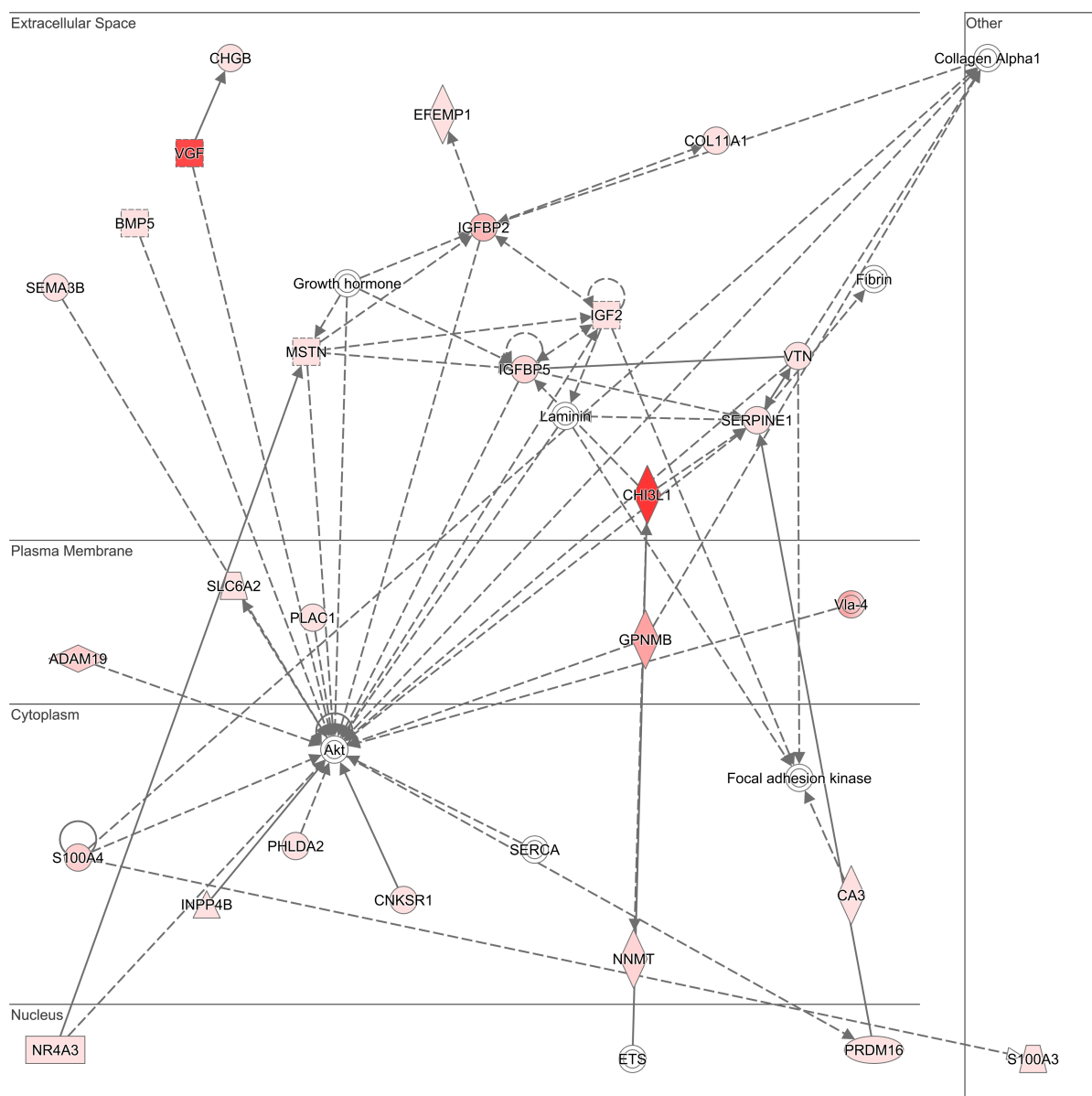
Supplementary Figure 4. The top enriched network of differentially expressed genes induced by VitC. A total of 32 molecules were identified to be associated with the functions of the cell cycle, cellular assembly and organization, and DNA replication, recombination, and repair. The network is presented in relation to its sub-cellular locations. The colour strength (from white to red) correlates with the fold-change in up-regulation of gene expression. Therefore, pink is indicative of up-regulation induced by VitC.



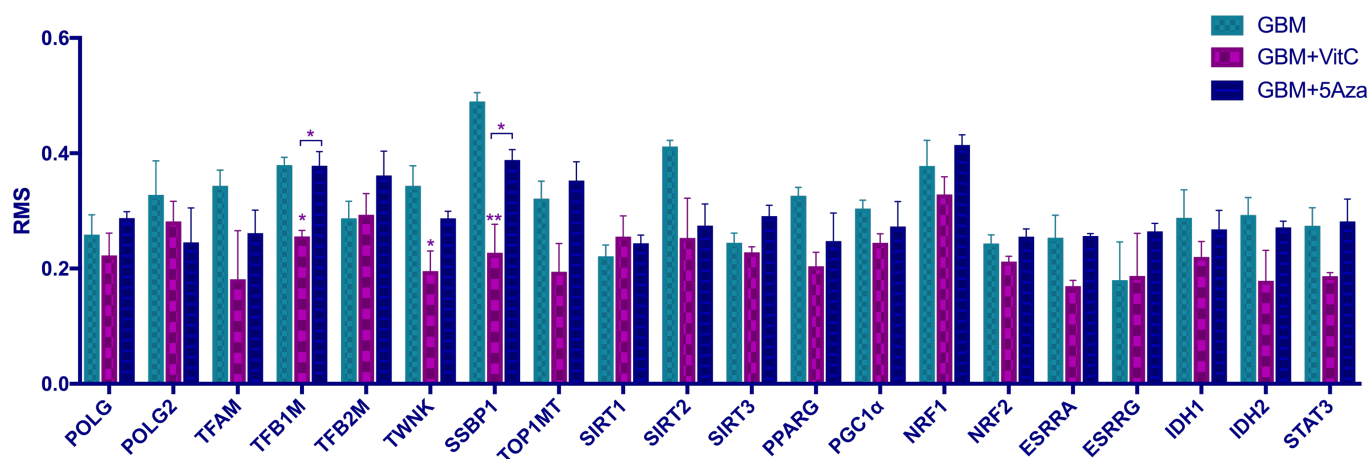
Supplementary Figure 5. The glioblastoma disease pathway affected by VitC. Genes associated with the disease pathway are presented as a circular display. The pathway figure was exported from IPA software. The colour strength (from white to red) correlates with the fold-change in up-regulation. Molecules in red indicate higher levels of increased gene expression whilst pink is more indicative of moderate up-regulation induced by VitC.



Supplementary Figure 6. The top enriched network of differentially expressed genes induced by 5Aza. The 28 molecules were identified to be associated with cell morphology, cellular function and maintenance, and carbohydrate metabolism. The network is presented in relation to its sub-cellular locations. The colour strength (from white to red) correlates with the fold-change in up-regulation of gene expression induced by 5Aza.



Supplementary Figure 7. The second most enriched network of differentially expressed genes induced by 5Aza. 26 molecules were identified to be associated with cancer, neurological disease, and organismal injury and abnormalities. The pathway figure was exported from IPA software. The network is presented in relation to its sub-cellular locations. Molecules in red indicate increased gene expression. The colour strength (from white to red) correlates with the fold-change in up-regulation of gene expression induced by 5Aza.



Supplementary Figure 8. DNA methylation levels over the CGIs located within the promoter regions of the mtDNA transcription and replication factors. DNA methylation levels expressed as RMS over the CGIs located in the promoter regions of the mtDNA transcription and replication factors are shown (n=3). Statistical analysis was performed using one-way ANOVA between the three cohorts. Bars represent the mean \pm SEM. * and ** indicate *p* values of < 0.05 and 0.01, respectively. The GBM cohort is shown by the light blue bars; GBM+VitC cohort is shown by the purple bars, the GBM+5Aza cohort is represented by the dark blue bars.

REFERENCE

1. Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A. *et al.* (2001) The sequence of the human genome. *Science*, **291**, 1304-1351.

CHAPTER 4

Modulation of mitochondrial DNA copy number during tumorigenesis induces changes in DNA methylation of the nuclear genome.

Manuscript attached.

Rationale leading to Chapter 4

In Chapter 3, the DNA methylation profile of GBM cells was characterized, and the sites that could be modulated by DNA demethylation treatments, namely 5-azacytidine and Vitamin C, were identified. I found that the global transcriptional changes after DNA demethylation treatments were tightly associated with networks in cellular differentiation and cancer. I also found that key mtDNA replication factors were extensively DNA demethylated after vitamin C and 5-azacytidine treatments, which up-regulated their expression leading to an increase in mtDNA copy number. Interestingly, mtDNA methylation was also detectable and could be decreased through DNA demethylation agents. These results highlighted the role of both nuclear and mtDNA demethylation in the up-regulation of mtDNA copy number.

Therefore, my second aim was to investigate whether mtDNA copy number has an effect on DNA methylation profiles during *in vivo* tumorigenesis. This aim was achieved by applying the same MeDIP-Seq workflow, as described in Chapter 3, to a GBM tumour model where cohorts of tumour-initiating cells had different levels of mtDNA copy number. Moreover, previous studies indicate that tumours derived from cells with 3% and 0.2% of their original mtDNA copy number had delayed onset and lower rates of tumour formation. Therefore, the findings from Chapter 4 could provide further insights into the relationship between the regulation of mtDNA copy number and DNA methylation in tumorigenesis.

Modulation of mitochondrial DNA copy number during tumorigenesis induces changes in DNA methylation of the nuclear genome

Xin Sun^{1,2}, Justin C St John^{1,2}

¹ Mitochondrial Genetics Group, Hudson Institute of Medical Research, 27-31 Wright Street, Clayton VIC 3168, AUSTRALIA

² Department of Molecular and Translational Sciences, Monash University, 27-31 Wright Street, Clayton VIC 3168, AUSTRALIA

Corresponding author: Justin St. John, Voice: +61 3 8572 2678, Fax: +61 03 9594 7416;

Email: Justin.StJohn@hudson.org.au

ABSTRACT

Background: There are multiple copies of mitochondrial DNA (mtDNA) present in each cell type and they are strictly regulated in a cell-specific manner by a group of nuclear-encoded replication factors. This strict regulation of mtDNA copy number is mediated by cell specific DNA methylation of these replication factors. Glioblastoma multiforme, HSR-GBM1, cells are hypermethylated and maintain low mtDNA copy number to support their tumorigenic status. When HSR-GBM1 cells with 50% of their original mtDNA content were inoculated into mice, tumours grew more aggressively than non-depleted cells. However, when the cells possessed only 3% and 0.2% of their original mtDNA content, tumour formation was less frequent and the onset was significantly delayed. Importantly, the process of tumorigenesis was dependent on mtDNA copy number being restored to pre-depletion levels.

Results: By performing whole genome MeDIP-Seq and RNA-Seq on tumours generated from cells possessing 100%, 50%, 0.3% and 0.2% of their original mtDNA content, we determined that restoration of mtDNA copy number caused significant changes to both the nuclear methylome and its transcriptome. The affected genes were specifically associated with gene networks and pathways involving behaviour, nervous system development, cell differentiation, and regulation of transcription and cellular processes. The mtDNA-specific replication factors were also modulated.

Conclusions: Our results highlight the bi-directional control of the nuclear and mitochondrial genomes through modulation of DNA methylation to control mtDNA copy number, which, in turn, modulates nuclear gene expression during tumorigenesis.

Keywords

Mitochondrial DNA; DNA methylation; gene expression; tumorigenesis; POLG

BACKGROUND

In mammalian cells, there are multiple copies of the mitochondrial genome (mtDNA) within each mitochondrion. The human mitochondrial genome is a circular, double-stranded genome that is 16.6 kb in size. It is essential for the production of cellular ATP, as it encodes 13 subunits of the electron transfer chain that conducts oxidative phosphorylation (OXPHOS) and is the cell's major generator of ATP. It also encodes 22 transfer RNAs and 2 ribosomal RNAs [1].

mtDNA replication is driven by a group of nuclear-encoded mtDNA-specific replication factors. The primary factor is the mtDNA-specific polymerase, Polymerase Gamma [2, 3], which is a heterotrimer enzyme composed of a catalytic subunit, subunit A (POLG), and two supporting subunits, subunits B (POLG2) [4]. The process of mtDNA replication is also supported by the mitochondrial helicase Twinkle (TWNK), the single-stranded DNA binding protein (SSBP1) and the mitochondrial topoisomerase (TOP1MT) [5-7]. There are several other key factors including mitochondrial transcription factors A (TFAM), B1 (TFB1M) and B2 (TFB2M), which generate the precursor transcript used to initiate mtDNA replication [8, 9]. Furthermore, some key upstream regulators of these mtDNA replication factors are highly involved in mitochondrial biogenesis and function, including the nuclear respiratory factors (NRF1/2), the peroxisome proliferator-activated receptor γ (PPARG) and its co-activator PGC1 α (PPARGC1A), the Sirtuin family of genes (SIRT1-3), and the estrogen related receptors (ESRRA/B/G) [10-14].

mtDNA replication is strictly regulated during development and differentiation, which enables mature cells to acquire cell-specific numbers of mtDNA copy to support their specific functions [15-18]. In the first instance, this is achieved by pluripotent (naïve) cells having established the mtDNA set point (reviewed in [19]). These naïve cells each possess low

numbers of mtDNA copy, around 200 copies, which promotes glycolysis, as the favoured form of energy production, and cellular proliferation [20]. These copies of mtDNA then serve as the initial template for mtDNA replication, allowing cells to acquire sufficient copies of mtDNA as they mature into differentiated states [16-18, 20, 21].

The inability to regulate mtDNA copy number affects cellular function and impedes developmental potential. For example, oocytes at the metaphase II stage that possess too few copies of mtDNA frequently fail to fertilise [22-25]. Likewise, certain types of cancer cells exhibit low mtDNA copy number and are unable to successfully complete differentiation. Indeed, cancer cells mainly rely on aerobic glycolysis for energy production, which allows for higher rates of cellular proliferation and prevents differentiation from taking place [16]. However, when they are induced to differentiate, they appear to be trapped in a 'pseudo-differentiated' state and unable to maintain the mtDNA set point and, as a result, the nuclear and mitochondrial genomes do not act in synchrony [19, 26]. Interestingly, when cancer cells are partially depleted of their mtDNA content, they undergo dedifferentiation and are then able to replicate mtDNA as they undergo differentiation, suggesting that they have re-established the mtDNA set point and synchrony between the nuclear and mitochondrial genomes [18].

During the early stages of development, the nuclear genome undergoes DNA demethylation mediated by the ten-eleven translocation methylcytosine dioxygenases (TET enzymes) to erase parental DNA methylation profiles. *de novo* DNA methylation then takes place to establish a new DNA methylation profile directed by the DNA methyltransferase family (DNMTs), which primes naïve cells for cellular differentiation [27, 28]. To this extent, cell-specific gene expression profiles are established during differentiation in synchrony with changes in DNA methylation patterns. Likewise, the DNA methylation status at exon 2 of

POLG is a determinant of when it is expressed and, in turn, regulates mtDNA copy number in a cell-specific manner [15, 16]. This is supported by experiments using DNA demethylation agents, such as 5-azacytidine [29] and vitamin C [30], where modulation of DNA methylation at exon 2 of *POLG* increased mtDNA copy number in HSR-GBM1 cells derived from Glioblastoma multiforme (GBM) tumours [16, 18, 31].

The HSR-GBM1 cell line is a high-grade malignant GBM cell line that is extensively DNA methylated, which contributes to its tumorigenic gene profile [32-39]. Its hypermethylated profile is established by over-expression of the isocitrate dehydrogenase (IDH) that harbours onco-mutations. Under normal circumstances, IDH enzymes act on the citric acid cycle to generate α -ketoglutarate, which is a co-factor of the TET enzymes that modulate DNA demethylation patterns [32-34, 37, 40]. However, overexpression of and mutations to IDH in GBM result in a metabolic switch that produces 2-hydroxyglutarate and restricts DNA demethylation induced by the TET enzymes [32-34, 37, 40].

Interestingly, mtDNA depletion of HSR-GBM1 cells to varying amounts of mtDNA copy number affected tumour progression and frequency when these cells were inoculated into mice [18]. Progression and frequency were greatest in cells depleted to 50% of their original content but tumour formation was less frequent and the onset was significantly delayed in cells possessing only 3% and 0.2% of their original mtDNA content [18]. Notably, mtDNA copy was restored to similar levels during *in vivo* tumorigenesis through modulation of *POLG* at exon 2 [17].

In order to determine if global DNA methylation profiles were modulated following the restoration and maintenance of mtDNA copy number during tumorigenesis, we investigated the DNA methylation profiles of GBM tumours derived from HSR-GBM1 cells possessing

varying amounts of mtDNA copy number using whole genome methylated-DNA-immunoprecipitation (MeDIP)-Seq. We further matched the modulated regions with their transcriptional profiles to focus on their effects on gene expression. We also investigated the mtDNA replication factors to determine how they responded to the newly established interactions between the nucleus and the mitochondrial genome. Our results highlight the bi-directional control of the nuclear and mitochondrial genomes through modulation of DNA methylation to control mtDNA copy number and gene expression during tumour progression using GBM as a model.

RESULTS

Replenishment of mtDNA copy number

Tumours were generated from HSR-GBM1 cells possessing different levels of mtDNA copy number, namely GBM¹⁰⁰ (possessing 100% of their mtDNA content), GBM⁵⁰ (50% mtDNA content), GBM³ (3% mtDNA content) and GBM^{0.2} (0.2% mtDNA content) cells [18]. Tumours arising from GBM³ and GBM^{0.2} cells had significantly delayed onset and lower rates of tumour formation. On the other hand, GBM⁵⁰ tumours exhibited accelerated tumour formation. Furthermore, mtDNA copy number was replenished in each of the tumour types to similar levels (Fig. 1), which highlights the need for sufficient mtDNA copy number to promote tumorigenesis, and likely explains why delayed tumour formation was observed in cells with lower levels of mtDNA copy number.

Modifications to DNA methylation resulting from re-establishment of mtDNA copy number during tumorigenesis

In order to determine whether DNA methylation of the nuclear genome was modulated following re-establishment of mtDNA copy number during tumour progression, we assessed the global DNA methylation profiles of each of the tumour types. Whole genome MeDIP-Seq was performed on DNA extracted from individual tumours from each cohort to identify key regions that might be modulated. Using the MEDIPS package, the quality of the sequencing reads was determined using the saturation and enrichment analyses and each achieved high standards, indicating that there was sufficient coverage and enrichment of the methylated DNA (Additional file 1). Regions with total coverage of $100 \times$ (minrowsum = 100) or greater were used for statistical analysis, which is deemed to be a more stringent filter than the default threshold of 10-fold coverage [31].

Differentially methylated regions (DMRs) were identified between the control group, GBM¹⁰⁰ tumours, and tumours formed from cells possessing varying levels of mtDNA. Table 1 shows the number of DMRs identified in the GBM⁵⁰, GBM³ and GBM^{0.2} tumours when compared with the GBM¹⁰⁰ tumours using different filters for statistical significance. With the increase of statistical power (p value) from ≤ 0.05 , 0.01 to 0.001, fewer DMRs were identified in each comparison. Overall, there were fewer DMRs identified in the GBM³ tumours followed by the GBM⁵⁰ tumours whilst there were far more DMRs identified in the GBM^{0.2} tumours. Using a cut-off p value of 0.001, for instance, 124 DMRs were identified between the GBM⁵⁰ tumours and the GBM¹⁰⁰ tumours, 36 DMRs for the GBM³ tumours and 16565 DMRs for the GBM^{0.2} tumours. The greater number of DMRs identified for the GBM^{0.2} tumours indicated that the DNA methylation profile for this cohort of tumours was extensively modified as a result of the increased levels of mtDNA depletion in their initiating cells and subsequent restoration during tumorigenesis. We further applied the Bonferroni correction to adjust p values to control for false positive calls. After correction, there were 143 DMRs identified in the GBM^{0.2} tumours, but there were no significant DMRs identified in the GBM⁵⁰ and GBM³ cohorts of tumours.

Both hyper- and hypo-methylated states were observed for tumours derived from cells with lower levels of mtDNA copy number. However, the majority of the DMRs were hypo-methylated compared to the GBM¹⁰⁰ tumours (Table 1). There were 31 hyper-methylated-DMRs (21.68%) identified in the comparison between GBM^{0.2} and GBM¹⁰⁰ tumours using an adjusted p-value of 0.05. Indeed, the tendency to hypo-methylation was also observed in the other tumours. Taking the results identified using a p-value of, for instance, 0.001, only 7.26% of the DMRs in the GBM⁵⁰ tumours and 44.44% of the DMRs in the GBM³ tumours were hyper-methylated. The different proportions of hypo- and hyper-methylation further suggest that when tumorigenesis was initiated from different levels of mtDNA copy number, different patterns of DNA methylation were induced.

The annotation of the DMRs and gene ontology analysis for the DMR-overlapping genes

Firstly, the annotation of genomic regions was performed on the four groups of DMRs identified using the filters of a p value ≤ 0.001 and an adjusted p value ≤ 0.05 (GBM⁵⁰ vs GBM¹⁰⁰ ($p \leq 0.001$), GBM³ vs GBM¹⁰⁰ ($p \leq 0.001$), GBM^{0.2} vs GBM¹⁰⁰ ($p \leq 0.001$) and GBM^{0.2} vs GBM¹⁰⁰ (adjusted. $p \leq 0.05$)). Overall, the results showed that over 49% of the DMRs were enriched in intragenic regions, namely exons, introns and promoter regions, and they were the most enriched regions (Fig. 2a - d). Each of these four groups of DMRs showed the greatest enrichment at introns, which was 52.51% of the GBM⁵⁰ DMRs (Fig. 2a), 39.22% of the GBM³ DMRs (Fig. 2b), 34.73% of the GBM^{0.2} DMRs (Fig. 2c; $p \leq 0.001$) and 35.94% of the GBM^{0.2} DMRs (Fig. 2c; adjusted $p \leq 0.05$). For the GBM⁵⁰ and GBM³ DMRs, the second most enriched regions were the distal intergenic regions, which were enriched at 23.46% and 37.25%, respectively (Fig. 2a and b). Interestingly, for the GBM^{0.2} DMRs, identified using the filters of $p \leq 0.001$ and adjusted $p \leq 0.05$, similar patterns were observed regardless of the numbers of DMRs (Fig. 2c and d). In contrast to the GBM⁵⁰ and GBM³ DMRs, the second most enriched region for the GBM^{0.2} DMRs was the promoter regions, which were 23.60% and 25.81% of the DMRs after filtering at $p \leq 0.001$ and adjusted $p \leq 0.05$, respectively (Fig. 2c and d). The distal intergenic region was ranked third accounting for 17.86% and 19.82% of the GBM^{0.2} DMRs identified using the filters of $p \leq 0.001$ and adjusted $p \leq 0.05$, respectively (Fig. 2c and d). Untranslated regions (UTRs) and downstream regions of genes accounted for less than 10% of the DMRs in each comparison.

To understand the related functions of the DMRs, functional classification and statistical overrepresentation analyses were performed on the DMR-overlapping genes using the PANTHER Classification System. The functional classification of the DMR-overlapping

genes showed that the GBM⁵⁰ and GBM³ tumours mostly clustered into cellular and metabolic processes (Fig. 3). Furthermore, the pathways of biological regulation, localization, response to stimuli and developmental processes were present in both cohorts (Fig. 3). Due to the large number of DMR-overlapping genes identified in the GBM^{0.2} tumours using the two different filters, statistical over-representation analysis was, therefore, performed. Interestingly, the biological processes enriched in the DMR-overlapping genes showed strong associations with developmental processes, especially in the DMR-overlapping genes identified with an adjusted p value of 0.05 (Table 2). The common developmental pathways identified in both analyses (shown in bold) included developmental process, nervous system development and cell differentiation processes (Table 2). This is in line with a previous finding that, after partial mtDNA depletion with 7 and 14-day ddC treatment, differentiation was able to take place in the GBM cells [18]. Moreover, regulation of transcription from the RNA polymerase II promoter was found to be commonly modulated, which reflected the potential for transcriptional changes that could result from the DMRs.

CNV analysis and the association with the DMRs

A common feature of cancer cells is genomic instability, which causes significant copy number variation (CNV) when compared to healthy controls [41]. CNV can also affect the outcomes from comparative analysis after MeDIP-Seq, which, if ignored, can introduce false positive results due to CNVs inducing higher or lower levels of enrichment rather than being reflective of the real differences in DNA methylation [41]. To investigate the impact of CNV on the results, tumour samples were genotyped using the Illumina Global Screening array which covers over 700 K SNPs in the human genome. CN-gain and loss, and loss of heterozygosity (LOH) regions were identified in the tumour samples (Table 3). The GBM¹⁰⁰ tumours exhibited 13 CN-gain and 14 CN-loss regions. The GBM³ tumours possessed 10 regions of CN-gain and 7 regions of CN-loss. The GBM⁵⁰ and the GBM^{0.2} tumours had 4 CN-

gain regions and 3 CN-loss regions. Amongst the regions with CNVs, two of them were identified to overlap with the DMRs identified in the GBM^{0.2} tumours ($p \leq 0.001$), none of the other CNV regions identified in the tumours overlapped with the DMRs (Additional file 2). Nevertheless, no CNV was identified to be significantly different between the tumour groups, as assessed using the NEXUS Copy Number Module. This negates the potential effect of CNV on the analysis of MeDIP-Seq outcomes.

Furthermore, as the tumours were derived from the same parental cell line, each group of tumours had similar numbers of LOH: 460 in the GBM¹⁰⁰ tumours, 461 in the GBM⁵⁰ tumours, 471 in the GBM³ tumours, and 464 in the GBM^{0.2} tumours. By contrast to the regions with CN gain and loss, more than 99% of the regions with LOH were commonly identified amongst the tumours.

The validation for the gene expression of the DNA methylation regulators

DNA methyltransferases (DNMTs), namely DNMT1/3A/3B, are responsible for catalysing DNA methylation at cytosines by converting them to 5-methylcytosines (5mC). On the other hand, the Ten-eleven Translocation (TET) enzymes, namely TET1/2/3, can demethylate 5mC to 5-hydroxymethylcytosine (5hmC) [27, 28]. The changes to these DNA methylation and demethylation factors can provide insights into how the changes to the DNA methylation profiles occurred. Therefore, the gene expression levels of the DNA methylation regulators were assessed using a custom-made Fluidigm qPCR array (Fig. 4). There was ≥ 1.5 fold-change in expression of *DNMT1* in GBM⁵⁰ and GBM^{0.2} tumours, suggesting that the reduction in mtDNA copy number strongly induced DNA demethylation in the tumours, and is probably the major reason why the majority of regions were hypo-methylated (Table 1). *DNMT3A* was also significantly down-regulated by over 1.5-fold in the GBM⁵⁰ tumours. None of the DNA methylation factors was found to be differentially expressed in the GBM³ tumours compared

with the GBM¹⁰⁰ tumours, which further supported the results that fewer DMRs were identified in this group. The expression of *TET1* showed a trend of up-regulation only in the GBM^{0.2} tumours. As there was a higher percentage of hyper-DMRs in the GBM^{0.2} tumours, the up-regulation of *TET1* might be a response to counter the increased hyper-methylation observed.

Furthermore, it has been reported that the genome of GBM tumours is extensively methylated to support the tumorigenic transcriptome that it exhibits [32-36]. This is further supported by the findings that the overexpression of the mutated isocitrate dehydrogenase (*IDH1/2*) in the citric acid cycle inhibits the activity of TET2 and thus contributes to the hyper-methylated genome [32-34, 37, 40]. We investigated the expression of *IDH1/2* in these tumours and found that they were down-regulated in the GBM⁵⁰ tumours (Fig. 4), which likely reduced their ability to induce hyper-methylation and supports the hypo-methylation observed in these tumours.

DMR-overlapping differentially expressed genes and differentiation markers after restoration of tumorigenic capacity

To investigate which genes changed their patterns of expression in response to the altered levels of DNA methylation, we compared the identified DMRs with the differentially expressed genes previously reported on the same samples [17]. We focused on the DMRs identified using an adjusted p value of 0.05, which were considered to be the genomic regions that had undergone most significant changes in DNA methylation. There were 9 DMR-overlapping differentially expressed genes, namely *BAIAP2*, *L3MBT1L*, *KCNC1*, *GPSM1*, *SLC27A1*, *MAF*, *OGFR*, *MICALL2* and *RHOT2* (Fig. 5a). Interestingly, the function of these genes is tightly associated with tumorigenesis and differentiation networks. *SLC27A1* was found to be significantly down regulated in a *SOX2*-knockdown GBM cell line that was characterized with abolished dedifferentiation and decreased tumorigenesis [42]. *RHOT2*, a member of the Ras

homologue gene family, was identified to be differentially expressed in different types (CD133⁺ and CD133⁻) of cancer stem cells in GBM [43]. *OGFR*, which encodes the opioid growth factor receptor, is a negative regulator of cell proliferation. Furthermore, activation of OGFR has been shown to reduce proliferation of astrocytes in cell culture [44]. The changes to the levels of gene expression were further validated for these genes (Fig. 5a). Interestingly, the GBM³ tumours, with the least number of DMRs identified, showed more variable patterns in expression amongst these DMR-overlapping genes when compared with the GBM¹⁰⁰ tumours. *RHOT2* and *OGFR* were significantly up-regulated in the GBM³ tumours ($p < 0.05$ and $p < 0.01$, respectively). The greater modification that occurred in the DNA methylation profiles of the GBM⁵⁰ and GBM^{0.2} tumours promoted gene expression profiles more similar to the tumorigenic features of the GBM¹⁰⁰ tumours, whereas less modification in the GBM³ tumours failed to mimic the transcriptional changes in these DMR-overlapping genes. This indicates that the restoration of mtDNA copy number and the DMRs in the end-point tumours probably functioned to drive similar tumorigenic profiles to those of the GBM¹⁰⁰ tumours.

As the DMRs were found to be tightly associated with developmental processes (Fig. 3 and Table 2), we also focused on the neuronal markers and genes in neurogenesis that have been reported in cells having undergone mtDNA depletion to the same levels [18]. Indeed, the partial mtDNA-depletion in GBM cells could lead to increased expression of genes associated with early developmental processes [18]. In tumours formed from these cells, an overview of gene expression identified in RNA-Seq data is shown (Fig. 5b). Both up- and down-regulation were observed. To make a comparison between the *in vivo* tumours and *in vitro* cells, the negative regulator of cell proliferation *ALK* was found to be down-regulated in the mtDNA-depleted tumours. However, *ALK* was found to be up-regulated in the mtDNA-depleted cultured cells. Likewise, the regulators of cell differentiation and proliferation, *VEGFA*, *DLG4*, *CDK*, *NRP2* and *ACHE*, were also found to be down-regulated in the mtDNA-

depleted tumours but presented higher levels of expression in the cells. The regulators of transcription and differentiation *ASCL1* and *DLL1* in the Notch signalling pathway were found to be up-regulated in the mtDNA-depleted tumours but down-regulated in the cells. The regulator of synaptogenesis *APOE* was found to be up-regulated in the GBM⁵⁰ and GBM³ tumours, but down-regulated in the GBM^{0.2} tumours and the mtDNA-depleted cells. Moreover, *FGF13*, *SEMA4D*, *PAPD6B* and *NRCAM* were found to be commonly up-regulated in both cells and tumours. As the cells recovered their tumorigenicity after restoration of mtDNA copy number *in vivo*, the shifts in gene expression between the tumours and the cells indicated that the restoration and maintenance of mtDNA copy number in tumorigenesis is tightly associated with the regulation of differentiation.

Modulation of the mtDNA replication factors for the maintenance of mtDNA copy number in tumorigenesis

As mtDNA copy number was replenished to similar levels in each of the tumour types (Fig. 1), we determined how the mtDNA-specific replication factors were affected. In total, 20 known mtDNA replication factors were screened for differentially methylated intragenic CGIs. The fold changes in relative methylation scores for the GBM¹⁰⁰ tumours of 28 promoter and gene-body CGIs within the mtDNA replication factors were plotted (Fig. 6a). An increasing trend in fold changes was observed from the GBM⁵⁰ to the GBM^{0.2} tumours. Particularly, for the CGIs, higher levels of methylation were found in the GBM⁵⁰ tumours compared to the GBM¹⁰⁰ tumours (shown in red blocks). The fold changes generally increased to an even higher level in the GBM^{0.2} tumours. This trend was mostly observed amongst the gene-body CGIs. For the CGIs that were more hypo-methylated in the GBM⁵⁰ tumours than the GBM¹⁰⁰ tumours (shown in blue blocks), the levels of DNA demethylation tended to be lower in the GBM^{0.2} tumours. However, the demethylation process, namely the transition from 5mC to 5hmC, was significantly enhanced in the tumours formed from GBM^{0.2} tumours [16], which

indicates the degree to which DNA methylation is regulated at a complex level to control the amount of mtDNA copy number available during tumorigenesis.

Furthermore, the gene expression levels were also assessed using a Fluidigm qPCR array (Fig. 6b). Overall, the levels of expression of the mtDNA replication factors in the tumours formed from cells possessing varying levels of mtDNA were found to be down-regulated, which likely restricts mtDNA replication during tumorigenesis after the restoration of mtDNA copy number to its original levels. *ESRRA*, *NRF1*, *POLG*, *PPARGCA1* and *SIRT1* were identified to be significantly down-regulated in all of the mtDNA-depleted tumours. The mtDNA replication factors *POLG2*, *TOP1MT* and *TWNK* were also found to be down-regulated in the GBM⁵⁰ and GBM³ tumours. Moreover, the key mtDNA transcription factors *POLRMT*, *TFAM*, *TFB1M* and *TFB2M* were also down-regulated in the GBM⁵⁰ and GBM³ tumours. *NRF2* and *SIRT3* were only down-regulated in the GBM⁵⁰ tumours. The hypoxia regulator *STAT3* was found to be significantly down-regulated in the GBM⁵⁰ and GBM^{0.2} tumours. *ESRRG* was only down-regulated in the GBM³ tumours. This indicates that mtDNA replenishment had been completed and the low levels of expression of the mtDNA replication factors were indicative of mtDNA turnover during cell division rather than the active repopulation of depleted cells with mtDNA.

DISCUSSION

It appears that mtDNA is essential to the initiation and maintenance of tumorigenesis, as shown in a model of GBM. To this extent, cells with low mtDNA content, namely cells with 3% and 0.2% of their original content, exhibited significantly delayed onset of and lower frequency in tumour formation [17, 18]. The extended period prior to the onset of tumorigenesis allowed these cells to restore mtDNA copy number to sufficient levels to promote and maintain tumorigenesis [17, 18]. The restoration of mtDNA can be achieved by either the cell's mtDNA replication machinery or the horizontal transfer of mitochondria from surrounding cells [17, 45, 46]. Through mtDNA next-generation sequencing of GBM tumours that restored their mtDNA, it is apparent that they possessed only human mtDNA [47]. However, in a model of human osteosarcoma, when 143B cells were completely depleted of their mtDNA, they acquired mtDNA from surrounding mouse cells to initiate tumorigenesis [17]. The differences between these models can be explained by xenomitochondrial studies, where cells preferentially replicate a more closely related population and only revert to a more genetically diverse source of mtDNA when the more closely related population is not available [48]. However, this is at the expense of compromised OXPHOS function [18]. In the 143B model, the more genetically diverse mouse mtDNA did not, however, produce fully developed tumours. These findings highlight the degree of compatibility between the nuclear and mitochondrial genomes that is required to support mtDNA replication and tumorigenesis and the importance of the presence of mtDNA.

In previous studies, we have demonstrated that nuclear-mitochondrial cross talk involves DNA methylation. Two key mtDNA replication factors are *POLG* and *TOP1MT*, which have CGIs that are modulated at exon 2 and exon 8, respectively [31]. During cell differentiation, the levels of DNA methylation at exon 2 of *POLG* play an important role in the cell-specific accumulation of mtDNA copy number initiated from the mtDNA set point [15, 16]. Likewise,

the levels of DNA methylation at exon 2 of *POLG* are correlated with mtDNA copy number in HSR-GBM1 cells, which takes place via modulating the expression of *POLG* [16]. Hypermethylation has been well-characterised in the GBM genome, which has also been identified at exon 2 of *POLG* and exon 8 of *TOP1MT* [31]. In the presence of DNA demethylation agents, such as 5-azacytidine and Vitamin C, the high levels of DNA methylation at exon 2 of *POLG* and exon 8 of *TOP1MT* can be reduced, which promotes mtDNA replication of mtDNA and the synchronous increases in gene expression associated with differentiation [16, 17, 31].

Here, we further show the bi-directional regulation between the two genomes whereby DNA methylation and gene expression profiles were extensively modulated upon the restoration of mtDNA during tumorigenesis. Indeed, once the mtDNA-depleted tumour cells had re-established their tumorigenic capability, mtDNA copy number in each tumour group was restored to similar levels (Fig. 1) [18]. Even though the GBM cells with varying levels of mtDNA restored their mtDNA content and initiated tumorigenesis, this process caused different responses in the context of DNA methylation in the resultant tumours (Table 1). The nuclear genomes were extensively modulated as shown from the DMRs identified from the comparisons between the tumours derived from cells with varying levels and the original level of mtDNA. Particularly, as tumours formed from cells with the highest levels of depletion, the GBM^{0.2} tumours had 16565 regions across the whole genome that were differentially methylated compared with the GBM¹⁰⁰ tumours ($p < 0.001$; Table 1). The number of DMRs identified in the GBM^{0.2} tumours was much higher than the ones identified in the GBM⁵⁰ (124) and GBM³ (36) tumours, which reflects the fact that the GBM^{0.2} cells had required longer to restore mtDNA copy number and required greater modification to promote tumorigenesis than the GBM⁵⁰ and GBM³ tumours [17].

Furthermore, there were two key common features observed amongst the three comparisons. Firstly, both hyper- and hypo-methylation were observed in the tumours initiated from cells with lower levels of mtDNA content, but the majority of the DMRs were DNA hypo-methylated in all three comparisons (Table 1). Therefore, the restoration of mtDNA during tumorigenesis tended to demethylate the nuclear genome. Indeed, GBM^{0.2} tumours had the fewest hyper-DMRs, namely 2.14%, GBM⁵⁰ tumours 7.26%, whilst GBM³ tumours had the most, 44% (Table 1). This was further supported by the changes in the levels of expression of the DNA methylation regulators. The DNA methylation factor *DNMT1* was significantly repressed to around 1.5-fold in the GBM⁵⁰ and GBM^{0.2} tumours to maintain low levels of the DNA methylation, but not in the GBM³ tumours (Fig. 4). However, *DNMT3A* was only down-regulated in the GBM⁵⁰ tumours (Fig. 4). Interestingly, in the GBM^{0.2} tumours, *TET1* was the only factor that tended to be up-regulated, which likely mediated the transition from 5mC to 5hmC to DNA demethylate the GBM genome resulting in the highest levels of hypo-methylation being observed in these tumours.

Secondly, regardless of the different numbers of DMRs identified in each comparison ($p < 0.001$ and adj. $p < 0.05$), the annotation for the genomic locations of the DMRs presented high enrichment over intragenic regions, namely promoters, exons and introns (Fig. 2). It is known that the effects of DNA methylation on transcription varies dependent on the location of the methylated sites. Methylated CpG islands within the promoter regions are known to repress transcription [49], whereas methylation within gene bodies has been found to positively and negatively correlate with transcriptional elongation [50]. Promoter-overlapping DMRs represented 8.38% and 5.88% of the DMRs in the GBM⁵⁰ and GBM³ tumours, respectively. In the GBM^{0.2} tumours, however, more promoter regions (23.6% by $p < 0.001$ and 25.8% by adj. $p < 0.05$) were differentially methylated, which indicates more genes were likely affected by the methylation levels in their promoter regions. Moreover, intron-

overlapping DMRs accounted for the largest proportion (> 34%) amongst all groups of DMRs (Fig. 2a-d). As for exon-overlapping DMRs, these accounted for 6.7% of DMRs in the GBM⁵⁰ tumours, 3.92% of DMRs in the GBM³ tumours and more than 10% of DMRs in the GBM^{0.2} tumours.

As mtDNA restoration caused significant changes to the DNA methylation profiles, we further showed that the identified DMR-overlapping genes affected a variety of biological process, especially developmental and cellular processes (Fig. 3 and Table 2). In the GBM⁵⁰ and GBM³ tumours, the resulting DMR-overlapping genes primarily targeted cellular and metabolic processes (Fig. 3). Developmental processes were more overrepresented in the GBM^{0.2} tumours (both $p < 0.001$ and adj. $p < 0.05$; Table 2), including cell differentiation and nervous system development. This was in line with our previous finding that cancer cells with lower mtDNA content are able to synchronously undertake mtDNA replication and differentiation, which suggests that they have re-established the mtDNA set point [18].

Moreover, we also matched the list of DMRs and the differentially expressed genes reported previously from RNA-Seq analysis on the same set of samples (Fig. 5a). There were 9 DMR-overlapping differentially expressed genes, namely *BAIAP2*, *L3MBT1L*, *KCNC1*, *GPSM1*, *SLC27A1*, *MAF*, *OGFR*, *MICALL2* and *RHOT2* (Fig. 5a). Interestingly, the GBM³ tumours, with the least number of DMRs identified, showed more variable patterns in expression amongst these DMR-overlapping genes when compared to the GBM¹⁰⁰ tumours. *RHOT2* ($p < 0.05$) and *OGFR* ($p < 0.01$) were significantly up-regulated in the GBM³ tumours. The function of these genes is tightly associated with differentiation in GBM tumorigenesis. *RHOT2* is differentially expressed in different types (CD133⁺ and CD133⁻) of GBM stem cells in [43]. *OGFR*, a negative regulator of cell proliferation, has been shown to reduce proliferation of astrocytes in cell culture once activated [44]. On the other hand, the greater

levels of modification that occurred in the DNA methylation profiles of the GBM⁵⁰ and GBM^{0.2} tumours promoted gene expression profiles more similar to the tumorigenic features of the GBM¹⁰⁰ tumours, whereas less modification in the GBM³ tumours failed to mimic the transcriptional changes in these DMR-overlapping genes. This indicates that the DNA methylation profiles were modulated in the end-point tumours to promote similar tumorigenic gene expression profiles to those of the GBM¹⁰⁰ tumours as a result of the restoration of mtDNA copy number.

The strong developmental feature observed in the DMRs and differentially expressed genes led us to further investigate the neuronal markers and genes in neurogenesis that have been reported in cells having undergone mtDNA depletion to the same levels [18]. Indeed, the partial depletion of mtDNA in the GBM cell model led to increased expression of genes associated with early developmental processes [18]. The regulators of cell proliferation and differentiation, *ALK*, *VEGFA*, *DLG4*, *CDK*, *NRP2* and *ACHE*, were down-regulated in the mtDNA-depleted tumours, but up-regulated in the mtDNA-depleted cells. However, the regulators of transcription and differentiation, *ASCL1* and *DLL1*, in the Notch signalling pathway were up-regulated in the mtDNA-depleted tumours but down-regulated in the cells. The regulator of synaptogenesis, *APOE*, was up-regulated in the GBM⁵⁰ and GBM³ tumours, but down-regulated in the GBM^{0.2} tumours and the mtDNA-depleted cells. *FGF13*, *SEMA4D*, *PAPD6B* and *NRCAM* were commonly up-regulated in both cells and tumours. As the cells recovered their tumorigenicity after restoration of mtDNA copy number *in vivo*, the shift between the tumours and the cells indicated that the restoration and maintenance of mtDNA copy number in tumorigenesis is tightly associated with the regulation of neuronal differentiation.

Even though mtDNA had been restored to similar levels in the mtDNA-depleted tumours (Fig. 1), we further showed that the mtDNA replication factors were still extensively regulated in the context of DNA methylation as well as their gene expression levels (Fig. 6). Amongst the mtDNA replication factors, most of them were found to have an intragenic CpG island. The majority of gene-body CGIs in the mtDNA-depleted tumours were more hyper-methylated than the GBM¹⁰⁰ tumours, whereas the promoter CGIs tended to be more hypo-methylated than the GBM¹⁰⁰ tumours (Fig. 6a). Moreover, with the increasing depletion of mtDNA, changes to DNA methylation took place more extensively in all the mtDNA replication factors. Particularly, for the hyper-methylated CGIs found in the GBM⁵⁰ tumours compared to the GBM¹⁰⁰ tumours, the fold changes generally increased to an even higher level in the GBM^{0.2} tumours. For the CGIs that were more hypo-methylated in the GBM⁵⁰ tumours than the GBM¹⁰⁰ tumours, the levels of DNA demethylation tended to be even lower in the GBM^{0.2} tumours. Interestingly, the overall transition from 5mC to 5hmC, which is indicative of demethylation, was also significantly enhanced in the tumours formed from GBM^{0.2} tumours [16]. This indicates the degree to which DNA methylation is regulated at a complex level to control the amount of mtDNA copy number that is available during tumorigenesis.

Furthermore, the expression of mtDNA replication factors in the tumours was found to be generally down-regulated, which likely restricts the levels of mtDNA replication during tumorigenesis after the restoration of mtDNA copy number to its original levels (Fig. 6b). The mtDNA replication factors, including *ESRRA*, *NRF1*, *POLG*, *PPARGCA1* and *SIRT1*, were down-regulated in all the tumours. In the GBM⁵⁰ and GBM³ tumours, *POLG2*, *TOP1MT* and *TWINK* were also found down-regulated. Moreover, the key mtDNA transcription factors, *POLRMT*, *TFAM*, *TFB1M* and *TFB2M*, were found down-regulated in the GBM⁵⁰ and GBM³ tumours. *NRF2* and *SIRT3* were only found to be down-regulated in the GBM⁵⁰ tumours. The hypoxia regulator *STAT3* was down-regulated in the GBM⁵⁰ and GBM^{0.2} tumours, whereas

ESRRG was only down-regulated in the GBM³ tumours. GBM cells with various levels of mtDNA require the restoration of mtDNA to initiate the *in vivo* process of tumorigenesis [18]. The *in vitro* differentiation of tumour cells needs the re-establishment the mtDNA set point, which can be achieved by partial depletion of mtDNA and DNA demethylation treatments [16, 18, 31]. Therefore, we have further shown here that the *in vivo* process of tumorigenesis went through global transcriptional changes induced by the changes to DNA methylation upon mtDNA restoration, especially at the key CGIs of the mtDNA replication factors, which are needed to not only balance the restored mtDNA copy number but also restore tumorigenic features that are similar to the original GBM cells namely the GBM¹⁰⁰ tumours that had not undergone mtDNA depletion. After mtDNA copy number had been restored to sufficient levels for tumorigenesis, mtDNA replication is strictly controlled to not only maintain low levels of mtDNA copy to promote tumorigenesis but ensure sufficient copy is available to repopulate cells as they actively divide to promote cell proliferation.

CONCLUSIONS

In all, we have shown that the restoration of mtDNA copy number during tumorigenesis induces major changes to the nuclear genome, including global DMRs and differentially expressed genes. These changes further enriched developmental processes as well as essential metabolic pathways associated with GBM tumorigenesis. Furthermore, the changes to the nuclear-encoded mtDNA replication factors highlight the feedback from the mitochondrial genome, once the tumours derived from mtDNA depleted GBM cells had restored tumorigenic capacity. This was clearly demonstrated in the GBM^{0.2} tumours, formed from the cells having undergone the highest levels of depletion and the greatest delay in onset of tumour formation. They underwent more extensive DNA methylation at key CGIs within the intragenic regions of the mtDNA replication factors to maintain similar transcriptional levels. Our results highlight the bi-directional control of the nuclear and

mitochondrial genomes through modulation of DNA methylation in response to mtDNA copy number and to control gene expression in tumorigenesis.

METHODS

Cell culture and Xenograft models

GBM tumours were previously generated from HSR-GBM1 cells possessing different levels of mtDNA copy number after treatment with 10 μ M 2'-3'-dideoxycytidine (ddC), a mtDNA depletion agent, in the presence of 50 μ g/mL uridine (Sigma-Aldrich, MO, USA). Cells were depleted to 50% (GBM⁵⁰), 3% (GBM³), 0.2% (GBM^{0.2}), and 100% (GBM¹⁰⁰) of their original mtDNA content, as described in [18]. HSR-GBM1 cells were cultured in complete neural stem cell media consisting of Dulbecco's Modified Eagle Medium / Nutrient Mixture (DMEM/F-12) Media (Thermo Fisher Scientific, MA, USA), 2% StemPro neural supplement (Thermo Fisher Scientific), 20 ng/mL basic fibroblast growth factor (bFGF; Merck Millipore MO, USA) and 20 ng/mL epidermal growth factor (EGF; Merck Millipore) at 37 °C, 5% CO₂ and 95% humidity.

The animal work was approved by the Animal Ethics Committee, Monash University, Approval Number: MMCA/2011/76. Briefly, 0.5 million HSR-GBM1 tumour cells in 100 mL of medium were inoculated subcutaneously into both flanks of 5- to 6-week-old, female BALB/c nude mice (Animal Research Centre, Perth, Australia). Tumour growth rates and volumes were reported in [18].

DNA extraction

Total genomic DNA was extracted from tumours using the DNeasy Blood & Tissue Kits (Qiagen, CA, USA), according to manufacturer's protocols with minor modifications. The DNA samples were treated with 3 μ L of RNase solution (Qiagen) at room temperature.

Determination of mtDNA copy number per cell

mtDNA copy number per cell was determined, as previously described [16]. Concentrations of *β -globin* and mtDNA were determined against standard curves generated by quantitative

real-time PCR (qPCR; Rotor-Gene 3000, Corbett Research, Cambridge, UK) on purified total DNA. Primer sequences and reaction conditions are listed in Additional file 3. mtDNA copy number per cell was calculated using the formula of $2 \times N_{mtDNA} / N_{\beta\text{-globin}}$, where, for N_{mtDNA} and $N_{\beta\text{-globin}}$, $N = (\text{qPCR product concentration} \times 6.023 \times 10^{14}) / (\text{qPCR product size in bp} \times 660)$.

Immunoprecipitation of methylated DNA (MeDIP) - Seq and analysis

5 µg of genomic DNA from each of the tumour samples underwent MeDIP, as previously described [51]. Briefly, each DNA sample was sheared into 200-1000 bp using the Covaris Adaptive Focused Acoustics (AFA™) S220 system (Woburn, MA, USA). dsDNA was then denatured to single-stranded DNA, as required for the antibody, by incubation at 95 °C for 10 min. 3 µg of each DNA sample was immunoprecipitated with 2 µg of anti-5mC antibody (Active Motif) with 20 µL per sample of prewashed Dynabeads® Protein G (Thermo Fisher Scientific). The suspension was incubated in 500 µL of IP buffer (100 mM sodium phosphate (pH 7.0); 1.4 M NaCl ; 0.5% Triton X-100) at 4 °C for 16 h under rotation. The beads were then washed three times with 1 mL of IP buffer and resuspended in 250 µL of proteinase K digestion buffer (50 mM Tris-HCl, pH 8.0; 10 M EDTA, pH 8.0; 1.0% SDS) with 10 µL of proteinase K (20 mg/mL; Bioline, London, UK). The suspension was incubated at 50 °C for 3 h on a thermo-shaker. The supernatant was then collected on a magnetic particle concentrator (Thermo Fisher Scientific). DNA was purified from the supernatant using the QIAquick PCR Purification Kit (Qiagen), according to the manufacturer's protocol.

MeDIP products from each of the tumour samples underwent library construction using the DNA SMART ChIP-Seq Kit (Clontech, CA, USA), as per the manufacturer's protocol (No. 021115). Libraries were quality checked using Qubit fluorometric quantitation (Thermo Fisher Scientific), a 2100 Bioanalyzer (Agilent, CA, USA), and qPCR. All libraries were of a similar size (297-335 bp) and quantities (~ 40 ng/µl). A single equimolar pool was made following

denaturation of the libraries at 94 °C for 2 min. 12 pM of each library was used for on-board cluster generation and sequencing in 4 lanes of two consecutive Illumina HiSeq 1500 Rapid runs. Cluster densities were within the optimal range of 881 - 959 k/mm² (optimal 750 - 1000 k/mm²). This resulted in around 170 million readable clusters per lane. 95% of the clusters passed filter, and 96% of the reads passed the Illumina sequencing quality score of Q30, which was deemed excellent quality for base-calling. Run quality parameters with the PhiX spike-in had an error rate of 0.12% (expected < 0.5%) and phasing / prephasing of < 0.13/ < 0.10 (expected < 0.4/ < 0.2).

Further analysis was performed on the sequences using the customised analytical pipeline, as described in [31]. Briefly, the sequences were checked for quality and adaptors using the FastQC software (v 0.11.5). The raw sequences were aligned to the human reference genome GRCh38/hg38 (UCSC) using the Burrows-Wheeler Aligner (BWA) software (version 0.7.16a) in the single-end mode of the BWA-backtrack algorithm with default settings [52]. The output files (*.sam) were then combined and converted to bam files (*.bam) by Samtools (version 1.4). Uniquely mapped reads with mapping scores (MAPQ) ≥ 30 were kept for analytical analysis using the MEDIPS package (version 1.24.0), according to the command codes of the MEDIPS package. Briefly, the reads were cleaned for PCR duplicates that mapped to exactly the same genomic positions by setting the parameter “uniq” to 1. The reads were then extended to 500 nucleotides according to the reference genome without any shift in the genomic locations (extend = 500; shift = 0). The coverage of every 100-bp window of the genome was summarized (window_size = 100). MeDIP-Seq specific quality control analysis was performed on each sample, which includes saturation analysis and CpG enrichment analysis. A coupling factor was set up based on the GBM¹⁰⁰ sample for normalization. Differentially methylated windows (DMWs) with the total count of reads ≥ 100 (minrowsum = 100) were identified between groups using the ‘edgeR’ method. Statistically

significant DMWs were selected for downstream analysis. Continuous significant DMWs were merged as one differentially methylated region (DMR). DMRs then underwent annotation based on their corresponding genomic regions in the human genome using the ChIPSeeker package (version 3.5) [53]. DNA methylation levels at particular regions of the genome, for example CpG islands (CGIs) associated with mtDNA replication factors, were assessed by the built-in ROI (region of interest) analysis in the MEDIPS package. DNA methylation levels of ROIs were expressed as relative methylation scores, as determined by the MEDIPS package. The relative methylation score was developed specifically for MeDIP-Seq in order to normalise methylation scores for regions based on the concept of CpG coupling analysis [54, 55].

Gene ontology analysis

Gene lists were analysed for gene ontology using the Panther Classification System (Version 12.0) [56]. The functional classification analysis was used to annotate the functions of genes. The statistical overrepresentation test (Fisher's Exact Test with a false discovery rate (FDR) correction) was used to identify the significance of associated biological processes overrepresented amongst the genes.

Determination of Copy Number Variation (CNV)

Total DNA from the tumour samples was submitted to the Australian Genome Research Facility (AGRF; VIC, AUS) to perform genotyping. The Illumina Human Global Screening Array Beadchip (Illumina) was used, which covers approximately 700K SNPs throughout the genome. Array data were normalised, clustered and underwent genotype calling using the Genotype Module of GenomeStudio 2.0 (Illumina), according to the user's manual. The full data report containing log R ratios and B allele frequencies for each probe was then exported and analysed using Nexus 9.0 software (Biodiscovery Inc., CA, USA). The Nexus Copy

Number Module was used to identify regions of CN gain and loss, loss of heterozygosity (LOH) for each group, and statistically significant CNVs between groups ($p \text{ value} \leq 0.05$). Human genome hg19 was chosen as the reference genome to annotate the genomic locations. Genomic regions were then aligned to human genome hg38, allowing overlapping analysis with the DMRs using the 'intersect' function of BEDTools (v2.24.0) [57, 58].

Identification of DMR-overlapping differentially expressed genes

RNA-Seq was previously performed on the GBM tumours [17]. The sequencing files were deposited in the NCBI Sequence Read Archive (SRA) under the accession code PRJNA296542 [17]. The sequences were mapped to the human genome (hg19) using the Tophat aligner (v1.3.1) [17, 59]. Differentially expressed genes were determined using the Cufflinks tool (v2.2.1) [17, 59]. Differentially expressed genes ($FDR \leq 0.05$) underwent Lift-over analysis to match hg38 human genome assembly using the GALAXY platform (usegalaxy.org) and then overlapping analysis was performed with the DMRs using the 'intersect' function of BEDTools (v2.24.0) [57, 58].

Gene expression analysis using real-time quantitative PCR

Total genomic RNA was extracted from the tumour samples using the RNeasy Mini Kit and the QIAshredder (Qiagen, CA, USA), according to manufacturer's protocols with minor modifications. The RNA samples were treated with DNase I (Qiagen) on column for 20 min. cDNA was synthesized from 1 μg of the total RNA using the Superscript III First-Strand synthesis system (Thermo Fisher Scientific), according to the manufacturer's instructions.

cDNA products were assessed using the Rotor-Gene 3000 RT-PCR machine under primer specific conditions (Additional file 3), as described in [16]. Relative gene expression was calculated using the $\Delta\Delta\text{CT}$ method and *OAZ1*, *18S rRNA* and *HPRT1* were used as the

housekeeping genes. Data were represented as the fold change to the GBM¹⁰⁰ group (n=3 biological replicates; mean \pm SEM). One-way ANOVA was used to determine statistical significance between the GBM⁵⁰, GBM³, and GBM^{0.2} tumours and the control group (GBM¹⁰⁰) tumours. Results were plotted using GraphPad Prism 7 (GraphPad Software, Inc., CA, USA).

Gene expression analysis using the Fluidigm platform

Gene expression of targets of interest was assessed using the Fluidigm qPCR array, according to the manufacturer's instructions. Taqman primers are listed in Additional file 4. Taqman primers were pooled and diluted in C1 DNA suspension buffer to a final concentration for each primer of 180 nM. Each cDNA sample and a non-template control underwent pre-amplification for 14 cycles, according to the manufacturer's instructions (Quick Reference PN 100-5876 B1). The pre-amplification reaction consisted of the Taqman PreAmp Master mix (Thermo Fisher Scientific) and the pooled Taqman primers. Products were diluted 5-fold with C1 DNA suspension buffer. The Integrated Fluidic Circuit (IFC) controller HX was then used to prime and load the 96.96 Dynamic array plate. 5 μ L of each sample was loaded in duplicate into each sample inlet and 5 μ L of each Taqman assay (10 x) were loaded into each assay inlet. Real-time qPCR was performed according to the Biomark GE 96.96 Standard v2 protocol. Data were exported using the Fluidigm Real-Time PCR analysis software (v4.1.1). Relative gene expression was calculated using the same method described above.

LIST OF ABBREVIATIONS

Mitochondrial DNA (mtDNA)

Oxidative phosphorylation (OXPHOS)

Methylated-DNA-immunoprecipitation (MeDIP)

Tumours formed from cells possessing 100% of their mtDNA content (GBM¹⁰⁰)

Tumours formed from cells possessing 50% of their mtDNA content (GBM⁵⁰)

Tumours formed from cells possessing 3% of mtDNA content (GBM³)

Tumours formed from cells possessing 0.2% of mtDNA content (GBM^{0.2})

Differentially methylated regions (DMR)

Copy number variation (CNV)

Loss of heterozygosity (LOH)

5-methylcytosine (5mC)

5-hydroxymethylcytosine (5hmC)

ADDITIONAL FILES

Additional file 1: MeDIP-Seq specific QC results determined by the MEDIPS package.

Additional file 2: CNV regions identified in each cohort of tumours determined using the Nexus 9.0 software.

Additional file 3: Primer pairs for real time PCR.

Additional file 4: Taqman assays used in the Fluidigm qPCR arrays.

DECLARATIONS

Ethics approval and consent to participate

- The animal work was approved by the Animal Ethics Committee, Monash University, Approval Number: MMCA/2011/76.

-

Consent for publication

- Not applicable.

-

Availability of data and material

- The MeDIP-Seq datasets supporting the conclusions of this article are available in Sequence Read Archive (SRA) and are accessible through accession number PRJNA397951. The CNV datasets supporting the conclusions of this article are available in the Gene Expression Omnibus (GEO) and are available under accession number GSE98366. The RNA-Seq datasets supporting the conclusions of this article are available in SRA and are obtainable under accession number PRJNA296542, as detailed in [17].

-

Competing interests

- The authors declare that they have no competing interests.

Funding

- This work was supported by Hudson Institute of Medical Research Discretionary Funds, and the Victorian Government's Operational Infrastructure Support Program. XS was supported by an Australian Postgraduate Award.

-

Authors' contributions

XS designed and performed the experiments, analysed the data and drafted the manuscript.

JCSJ conceived the work, designed and coordinated the experiments, analysed the data, drafted the manuscript and obtained funding for the work. All authors edited, read and approved the final manuscript.

Acknowledgements

- We are grateful to Dr. Trevor Wilson and Dr. Selva Kumari Ramasubramanian, The Medical Genomics Facility, Monash Health Translation Precinct, for assistance in performing MeDIP-Seq and the Fluidigm real-time qPCR array, respectively. We are also grateful to BioDiscovery, Inc for the use of Nexus Copy Number software in our data analysis.

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**:457-465.
2. Kasamatsu H, Grossman LI, Robberson DL, Watson R, Vinograd J: **The replication and structure of mitochondrial DNA in animal cells.** *Cold Spring Harb Symp Quant Biol* 1974, **38**:281-288.
3. Clayton DA: **Replication of animal mitochondrial DNA.** *Cell* 1982, **28**:693-705.
4. Carrodegua JA, Theis K, Bogenhagen DF, Kisker C: **Crystal structure and deletion analysis show that the accessory subunit of mammalian DNA polymerase gamma, Pol gamma B, functions as a homodimer.** *Mol Cell* 2001, **7**:43-54.
5. Takamatsu C, Umeda S, Ohsato T, Ohno T, Abe Y, Fukuoh A, Shinagawa H, Hamasaki N, Kang D: **Regulation of mitochondrial D-loops by transcription factor A and single-stranded DNA-binding protein.** *EMBO Rep* 2002, **3**:451-456.
6. Korhonen JA, Gaspari M, Falkenberg M: **TWINKLE Has 5' -> 3' DNA helicase activity and is specifically stimulated by mitochondrial single-stranded DNA-binding protein.** *J Biol Chem* 2003, **278**:48627-48632.
7. Zhang H, Meng LH, Pommier Y: **Mitochondrial topoisomerases and alternative splicing of the human TOP1mt gene.** *Biochimie* 2007, **89**:474-481.
8. Fisher RP, Clayton DA: **Purification and characterization of human mitochondrial transcription factor 1.** *Mol Cell Biol* 1988, **8**:3496-3509.
9. Falkenberg M, Gaspari M, Rantanen A, Trifunovic A, Larsson NG, Gustafsson CM: **Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA.** *Nat Genet* 2002, **31**:289-294.
10. Gleyzer N, Vercauteren K, Scarpulla RC: **Control of mitochondrial transcription specificity factors (TFB1M and TFB2M) by nuclear respiratory factors (NRF-1 and NRF-2) and PGC-1 family coactivators.** *Mol Cell Biol* 2005, **25**:1354-1366.
11. Aquilano K, Vigilanza P, Baldelli S, Pagliei B, Rotilio G, Ciriolo MR: **Peroxisome proliferator-activated receptor gamma co-activator 1alpha (PGC-1alpha) and sirtuin 1 (SIRT1) reside in mitochondria: possible direct function in mitochondrial biogenesis.** *J Biol Chem* 2010, **285**:21590-21599.
12. Chen JQ, Delannoy M, Cooke C, Yager JD: **Mitochondrial localization of ERalpha and ERbeta in human MCF7 cells.** *Am J Physiol Endocrinol Metab* 2004, **286**:E1011-1022.
13. Schreiber SN, Emter R, Hock MB, Knutti D, Cardenas J, Podvinec M, Oakeley EJ, Kralli A: **The estrogen-related receptor alpha (ERRalpha) functions in PPARGgamma coactivator 1alpha (PGC-1alpha)-induced mitochondrial biogenesis.** *Proc Natl Acad Sci U S A* 2004, **101**:6472-6477.
14. Kong X, Wang R, Xue Y, Liu X, Zhang H, Chen Y, Fang F, Chang Y: **Sirtuin 3, a new target of PGC-1alpha, plays an important role in the suppression of ROS and mitochondrial biogenesis.** *PLoS One* 2010, **5**:e11707.
15. Kelly RD, Mahmud A, McKenzie M, Trounce IA, St John JC: **Mitochondrial DNA copy number is regulated in a tissue specific manner by DNA methylation of the nuclear-encoded DNA polymerase gamma A.** *Nucleic Acids Res* 2012, **40**:10124-10138.
16. Lee W, Johnson J, Gough DJ, Donoghue J, Cagnone GL, Vaghjiani V, Brown KA, Johns TG, St John JC: **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**:e1664.

17. Lee WT, Cain JE, Cuddihy A, Johnson J, Dickinson A, Yeung KY, Kumar B, Johns TG, Watkins DN, Spencer A, St John JC: **Mitochondrial DNA plasticity is an essential inducer of tumorigenesis.** *Cell Death Discov* 2016, **2**:16016.
18. Dickinson A, Yeung KY, Donoghue J, Baker MJ, Kelly RD, McKenzie M, Johns TG, St John JC: **The regulation of mitochondrial DNA copy number in glioblastoma cells.** *Cell Death Differ* 2013, **20**:1644-1653.
19. Sun X, St John JC: **The role of the mtDNA set point in differentiation, development and tumorigenesis.** *Biochem J* 2016, **473**:2955-2971.
20. Facucho-Oliveira JM, Alderson, J., Splikings, E.C., Egginton, S., and St. John, J.C.: **Mitochondria DNA replication during differentiation of murine embryonic stem cells.** *J Cell Sci* 2007, **15**:4025-4034.
21. Facucho-Oliveira JM, St John JC: **The relationship between pluripotency and mitochondrial DNA proliferation during early embryo development and embryonic stem cell differentiation.** *Stem Cell Rev* 2009, **5**:140-158.
22. El Shourbagy SH, Spikings EC, Freitas M, St John JC: **Mitochondria directly influence fertilisation outcome in the pig.** *Reproduction* 2006, **131**:233-245.
23. Spikings EC, Alderson J, St John JC: **Regulated mitochondrial DNA replication during oocyte maturation is essential for successful porcine embryonic development.** *Biol Reprod* 2007, **76**:327-335.
24. Reynier P, May-Panloup P, Chretien MF, Morgan CJ, Jean M, Savagner F, Barriere P, Malthiery Y: **Mitochondrial DNA content affects the fertilizability of human oocytes.** *Mol Hum Reprod* 2001, **7**:425-429.
25. Santos TA, El Shourbagy S, St John JC: **Mitochondrial content reflects oocyte variability and fertilization outcome.** *Fertil Steril* 2006, **85**:584-591.
26. Lee WT, St John J: **The control of mitochondrial DNA replication during development and tumorigenesis.** *Ann N Y Acad Sci* 2015, **1350**:95-106.
27. Guo F, Li X, Liang D, Li T, Zhu P, Guo H, Wu X, Wen L, Gu TP, Hu B, et al: **Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote.** *Cell Stem Cell* 2014, **15**:447-459.
28. Brunner AL, Johnson DS, Kim SW, Valouev A, Reddy TE, Neff NF, Anton E, Medina C, Nguyen L, Chiao E, et al: **Distinct DNA methylation patterns characterize differentiated human embryonic stem cells and developing human fetal liver.** *Genome Res* 2009, **19**:1044-1056.
29. Jones PA, Taylor SM: **Cellular differentiation, cytidine analogs and DNA methylation.** *Cell* 1980, **20**:85-93.
30. Blaschke K, Ebata KT, Karimi MM, Zepeda-Martinez JA, Goyal P, Mahapatra S, Tam A, Laird DJ, Hirst M, Rao A, et al: **Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells.** *Nature* 2013, **500**:222-226.
31. Sun X, Johnson J, St John JC: **Global DNA methylation synergistically regulates the nuclear and mitochondrial genomes in glioblastoma cells.** *Nucleic Acids Res* 2018; doi: 10.1093/nar/gky339.
32. Yan H, Parsons DW, Jin G, McLendon R, Rasheed BA, Yuan W, Kos I, Batinic-Haberle I, Jones S, Riggins GJ, et al: **IDH1 and IDH2 mutations in gliomas.** *N Engl J Med* 2009, **360**:765-773.
33. Turcan S, Rohle D, Goenka A, Walsh LA, Fang F, Yilmaz E, Campos C, Fabius AW, Lu C, Ward PS, et al: **IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype.** *Nature* 2012, **483**:479-483.
34. Cohen AL, Holmen SL, Colman H: **IDH1 and IDH2 mutations in gliomas.** *Curr Neurol Neurosci Rep* 2013, **13**:345.
35. Haseeb A, Makki MS, Haqqi TM: **Modulation of ten-eleven translocation 1 (TET1), Isocitrate Dehydrogenase (IDH) expression, alpha-Ketoglutarate**

- (alpha-KG), and DNA hydroxymethylation levels by interleukin-1beta in primary human chondrocytes. *J Biol Chem* 2014, **289**:6877-6885.
36. Marie SK, Shinjo SM: **Metabolism and brain cancer.** *Clinics (Sao Paulo)* 2011, **66 Suppl 1**:33-43.
 37. Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, Li Y, Bhagwat N, Vasanthakumar A, Fernandez HF, et al: **Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation.** *Cancer Cell* 2010, **18**:553-567.
 38. Turcan S, Rohle D, Goenka A, Walsh LA, Fang F, Yilmaz E, Campos C, Fabius AWM, Lu C, Ward PS: **IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype.** *Nature* 2012, **483**:479-483.
 39. Galli R, Binda E, Orfanelli U, Cipelletti B, Gritti A, De Vitis S, Fiocco R, Foroni C, Dimeco F, Vescovi A: **Isolation and characterization of tumorigenic, stem-like neural precursors from human glioblastoma.** *Cancer Res* 2004, **64**:7011-7021.
 40. Hess KR, Broglio KR, Bondy ML: **Adult glioma incidence trends in the United States, 1977-2000.** *Cancer* 2004, **101**:2293-2299.
 41. Robinson MD, Stirzaker C, Statham AL, Coolen MW, Song JZ, Nair SS, Strbenac D, Speed TP, Clark SJ: **Evaluation of affinity-based genome-wide DNA methylation data: effects of CpG density, amplification bias, and copy number variation.** *Genome Res* 2010, **20**:1719-1729.
 42. Berezovsky AD, Poisson LM, Cherba D, Webb CP, Transou AD, Lemke NW, Hong X, Hasselbach LA, Irtenkauf SM, Mikkelsen T, deCarvalho AC: **Sox2 promotes malignancy in glioblastoma by regulating plasticity and astrocytic differentiation.** *Neoplasia* 2014, **16**:193-206, 206 e119-125.
 43. Lottaz C, Beier D, Meyer K, Kumar P, Hermann A, Schwarz J, Junker M, Oefner PJ, Bogdahn U, Wischhusen J, et al: **Transcriptional profiles of CD133+ and CD133- glioblastoma-derived cancer stem cell lines suggest different cells of origin.** *Cancer Res* 2010, **70**:2030-2040.
 44. Campbell AM, Zagon IS, McLaughlin PJ: **Astrocyte proliferation is regulated by the OGF-OGFr axis in vitro and in experimental autoimmune encephalomyelitis.** *Brain Res Bull* 2013, **90**:43-51.
 45. Berridge MV, Schneider RT, McConnell MJ: **Mitochondrial Transfer from Astrocytes to Neurons following Ischemic Insult: Guilt by Association?** *Cell Metab* 2016, **24**:376-378.
 46. Dong LF, Kovarova J, Bajzikova M, Bezawork-Geleta A, Svec D, Endaya B, Sachaphibulkij K, Coelho AR, Sebkova N, Ruzickova A, et al: **Horizontal transfer of whole mitochondria restores tumorigenic potential in mitochondrial DNA-deficient cancer cells.** *Elife* 2017, **6**.
 47. Yeung KY, Dickinson A, Donoghue JF, Polekhina G, White SJ, Grammatopoulos DK, McKenzie M, Johns TG, St John JC: **The identification of mitochondrial DNA variants in glioblastoma multiforme.** *Acta Neuropathol Commun* 2014, **2**:1.
 48. Kenyon L, Moraes CT: **Expanding the functional human mitochondrial DNA database by the establishment of primate xenomitochondrial cybrids.** *Proc Natl Acad Sci U S A* 1997, **94**:9131-9135.
 49. Deaton AM, Bird A: **CpG islands and the regulation of transcription.** *Genes Dev* 2011, **25**:1010-1022.
 50. Yang X, Han H, De Carvalho DD, Lay FD, Jones PA, Liang G: **Gene body methylation can alter gene expression and is a therapeutic target in cancer.** *Cancer Cell* 2014, **26**:577-590.
 51. Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, Schubeler D: **Chromosome-wide and promoter-specific analyses identify sites of**

- differential DNA methylation in normal and transformed human cells.** *Nat Genet* 2005, **37**:853-862.
52. Li H, Durbin R: **Fast and accurate short read alignment with Burrows-Wheeler transform.** *Bioinformatics* 2009, **25**:1754-1760.
 53. Yu G, Wang LG, He QY: **ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization.** *Bioinformatics* 2015, **31**:2382-2383.
 54. Chavez L, Jozefczuk J, Grimm C, Dietrich J, Timmermann B, Lehrach H, Herwig R, Adjaye J: **Computational analysis of genome-wide DNA methylation during the differentiation of human embryonic stem cells along the endodermal lineage.** *Genome Res* 2010, **20**:1441-1450.
 55. Lienhard M, Grimm C, Morkel M, Herwig R, Chavez L: **MEDIPS: genome-wide differential coverage analysis of sequencing data derived from DNA enrichment experiments.** *Bioinformatics* 2014, **30**:284-286.
 56. Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, Thomas PD: **PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements.** *Nucleic Acids Res* 2017, **45**:D183-D189.
 57. Quinlan AR, Hall IM: **BEDTools: a flexible suite of utilities for comparing genomic features.** *Bioinformatics* 2010, **26**:841-842.
 58. Quinlan AR: **BEDTools: The Swiss-Army Tool for Genome Feature Analysis.** *Curr Protoc Bioinformatics* 2014, **47**:11 12 11-34.
 59. Ghosh S, Chan CK: **Analysis of RNA-Seq Data Using TopHat and Cufflinks.** *Methods Mol Biol* 2016, **1374**:339-361.

Table 1. The number of DMRs identified in tumours formed from depleted cells compared with non-depleted cells.

Comparisons	P ≤ 0.05	P ≤ 0.01	P ≤ 0.001	Adj.p ≤ 0.05
GBM⁵⁰ vs GBM¹⁰⁰	8980	1135	124	
	(16.85% :	(10.57% :	(7.26% :	-
	83.15%)	89.43%)	92.74%)	
GBM³ vs GBM¹⁰⁰	4940	597	36	
	(34.76% :	(41.88% :	(44.44% :	-
	65.24%)	58.12%)	55.56%)	
GBM^{0.2} vs GBM¹⁰⁰	51366	34060	16565	143
	(11.25% :	(3.66% :	(2.14% :	(21.68% :
	88.75%)	96.34%)	97.86%)	78.32%)

The ratio of hyper- and hypo-methylated-DMRs (expressed as percentages) are shown in brackets.

Table 2. Pathways affected by the DMR-overlapping genes in the GBM^{0.2} tumours using the PANTHER Classification system.

PANTHER Pathways	Input / Background	Fold Enrichment	P- value	FDR	PANTHER Pathways	Input / Background	Fold Enrichment	P- value	FDR
DMR-overlapping genes in the GBM^{0.2} tumours (p ≤ 0.001)					DMR-overlapping genes in the GBM^{0.2} tumours (adj. p ≤ 0.05)				
Cell differentiation (GO:0030154)	191/548	1.47	6.12E-06	9.95E-05	Cell differentiation (GO:0030154)	9/459	3.93	5.42E-04	4.41E-02
Nervous system development (GO:0007399)	113/314	1.52	1.88E-04	1.76E-03	Nervous system Development (GO:0007399)	7/238	5.89	2.23E-04	5.45E-02
Regulation of transcription from RNA polymerase II promoter (GO:0006357)	196/598	1.39	9.47E-05	9.63E-04	Regulation of transcription from RNA polymerase II promoter (GO:0006357)	10/548	3.66	4.61E-04	5.63E-02
					Cellular process (GO:0009987)	57/7905	1.45	5.54E-04	2.70E-02
					Behaviour (GO:0007610)	4/76	10.55	6.79E-04	2.76E-02
Developmental process (GO:0032502)	537/1501	1.51	1.01E-16	2.46E-14					
Sensory perception of smell (GO:0007608)	6/240	0.11	1.41E-15	1.72E-13					
Defence response to bacterium (GO:0042742)	0/112	< 0.01	9.10E-11	4.44E-09					
Response to biotic stimulus (GO:0009607)	3/144	0.09	2.98E-10	1.21E-08					

B cell mediated immunity (GO:0019724)	0/94	<0.01	3.08E-09	9.40E-08	
Complement activation (GO:0006956)	1/94	0.04	5.91E-08	1.44E-06	
Ectoderm development (GO:0007398)	96/212	1.91	3.82E-07	8.46E-06	
Cell recognition (GO:0008037)	4/105	0.16	3.64E-06	6.84E-05	
Mesoderm development (GO:0007498)	109/269	1.71	4.74E-06	8.27E-05	
Phagocytosis (GO:0006909)	15/182	0.35	9.91E-06	1.42E-04	
Embryo development (GO:0009790)	49/106	1.95	1.96E-04	1.77E-03	
Cellular component morphogenesis (GO:0032989)	144/423	1.44	2.32E-04	2.02E-03	
Biological adhesion (GO:0022610)	124/356	1.47	2.72E-04	2.21E-03	
Cell adhesion (GO:0007155)	124/356	1.47	2.72E-04	2.29E-03	
Segment specification (GO:0007379)	37/78	2.01	7.86E-04	5.99E-03	
Transmembrane receptor protein tyrosine kinase signalling pathway (GO:0007169)	60/151	1.68	1.07E-03	7.90E-03	
Negative regulation of apoptotic process (GO:0043066)	43/99	1.84	1.25E-03	8.73E-03	

Regulation of phosphate metabolic process (GO:0019220)	169/537	1.33	1.34E-03	8.86E-03
Intracellular signal transduction (GO:0035556)	310/1071	1.22	1.55E-03	9.92E-03
Anterior/posterior axis specification (GO:0009948)	12/16	3.17	3.37E-03	1.96E-02
Cytoskeleton organization (GO:0007010)	129/404	1.35	3.74E-03	2.12E-02
Synaptic transmission (GO:0007268)	121/382	1.34	5.92E-03	3.28E-02

Common processes are highlighted in bold.

Table 3. CNV analysis determined using the Nexus Copy Number Module.

		CN gain	CN loss	LOH
GBM¹⁰⁰	Total	13	14	460
	Unique	9	11	1
GBM⁵⁰	Total	4	3	461
	Unique	1	0	2
GBM³	Total	10	7	471
	Unique	8	5	5
GBM^{0.2}	Total	4	3	464
	Unique	2	2	3

Figure Legends

Figure 1. Replenishment of mtDNA copy number in the GBM⁵⁰, GBM³ and GBM^{0.2} tumours. Fold change in mtDNA copy number for GBM⁵⁰, GBM³ and GBM^{0.2} tumours relative to the GBM¹⁰⁰ tumours. Statistical significance was determined using One-way ANOVA. No significant differences were identified between the groups.

Figure 2. Annotations of the DMRs. Based on their genomic locations, DMRs identified in the comparison a) GBM⁵⁰ vs GBM¹⁰⁰ ($p \leq 0.001$); b) GBM³ vs GBM¹⁰⁰ ($p \leq 0.001$); c) GBM^{0.2} vs GBM¹⁰⁰ ($p \leq 0.001$); and d) GBM^{0.2} vs GBM¹⁰⁰ (adjusted. $p \leq 0.05$) were classified into promoters, exons, introns, UTRs, downstream regions and distal intergenic regions using the ChIPSeeker package. DMR-overlapping regions in brown signify promoter regions, violet-red indicates exons, green indicates introns, blue indicates promoters, teal represents UTRs, blue represents downstream regions and orange indicates intergenic regions. The percentage of each group is indicated in the key to each figure.

Figure 3. Functional classification analysis of the DMR-overlapping genes in the GBM⁵⁰ and GBM³ tumours. DMR-overlapping genes in a) GBM⁵⁰ tumours; and b) GBM³ tumours were classified into sub-groups based on their associated biological processes using the PANTHER system.

Figure 4. Differential expression of the regulators of DNA methylation. Bars represent the mean of the relative quantification levels relative to the GBM¹⁰⁰ tumours (relative expression = 1). Error bars show SEM. Statistical significance was determined by One-way ANOVA. *, ** and *** indicate p values < 0.05, 0.01 and 0.001, respectively.

Figure 5. Transcriptional profiles of DMR-overlapping differentially expressed genes and markers in neurogenesis and neural stem cells. a) Differential expression of the DMR-overlapping differentially expressed genes. Bars represent the mean of the relative

quantification levels relative to the GBM¹⁰⁰ tumours (relative expression = 1). Statistical significance was determined by One-way ANOVA. Error bars show SEM. *, ** indicate p values < 0.05 and 0.01, respectively. **b) Heatmap of markers in neurogenesis and neural stem cells from the RNA-Seq profiles.** For the GBM⁵⁰, GBM³ and GBM^{0.2} cohorts, the fold changes (log2) of read counts to the mean value of the GBM¹⁰⁰ cohort were plotted. The colour scheme from blue, white to red represents the level of expression from 2-fold down-regulation to 2-fold up-regulation.

Figure 6. Overview of the changes to DNA methylation and transcription of the mtDNA replication factors. **a) DNA methylation levels of the intragenic CGIs associated with the mtDNA replication factors.** Fold-changes in relative methylation scores to the GBM¹⁰⁰ tumours were plotted for the three biological replicates from the GBM⁵⁰, GBM³ and GBM^{0.2} cohorts. The colour scheme from blue, white to red represents the level of DNA methylation from low to high. “P”, “E” and “I” following each gene name indicates CGIs that are located at the promoter regions, exons and introns, respectively. **b) Differential expression of the mtDNA replication factors.** Bars represent the mean of the relative quantification levels relative to the GBM¹⁰⁰ tumours (relative expression = 1). Error bars show SEM. Statistical significance was determined by One-way ANOVA. *, ** and *** indicate p values < 0.05, 0.01 and 0.001, respectively.

Figure 1

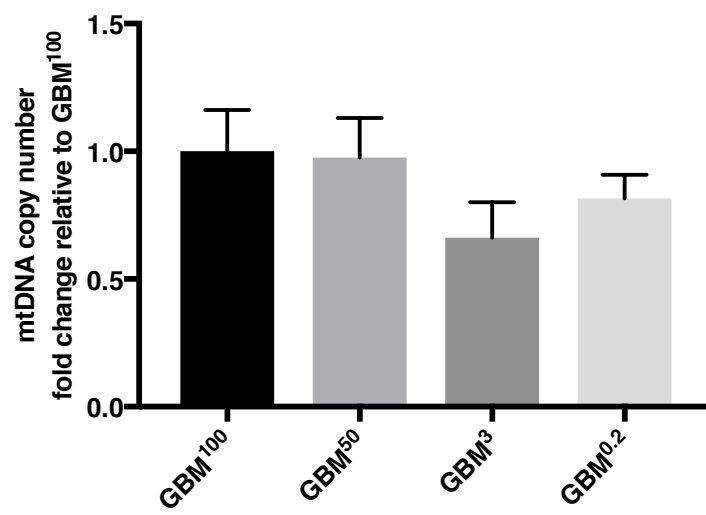


Figure 2

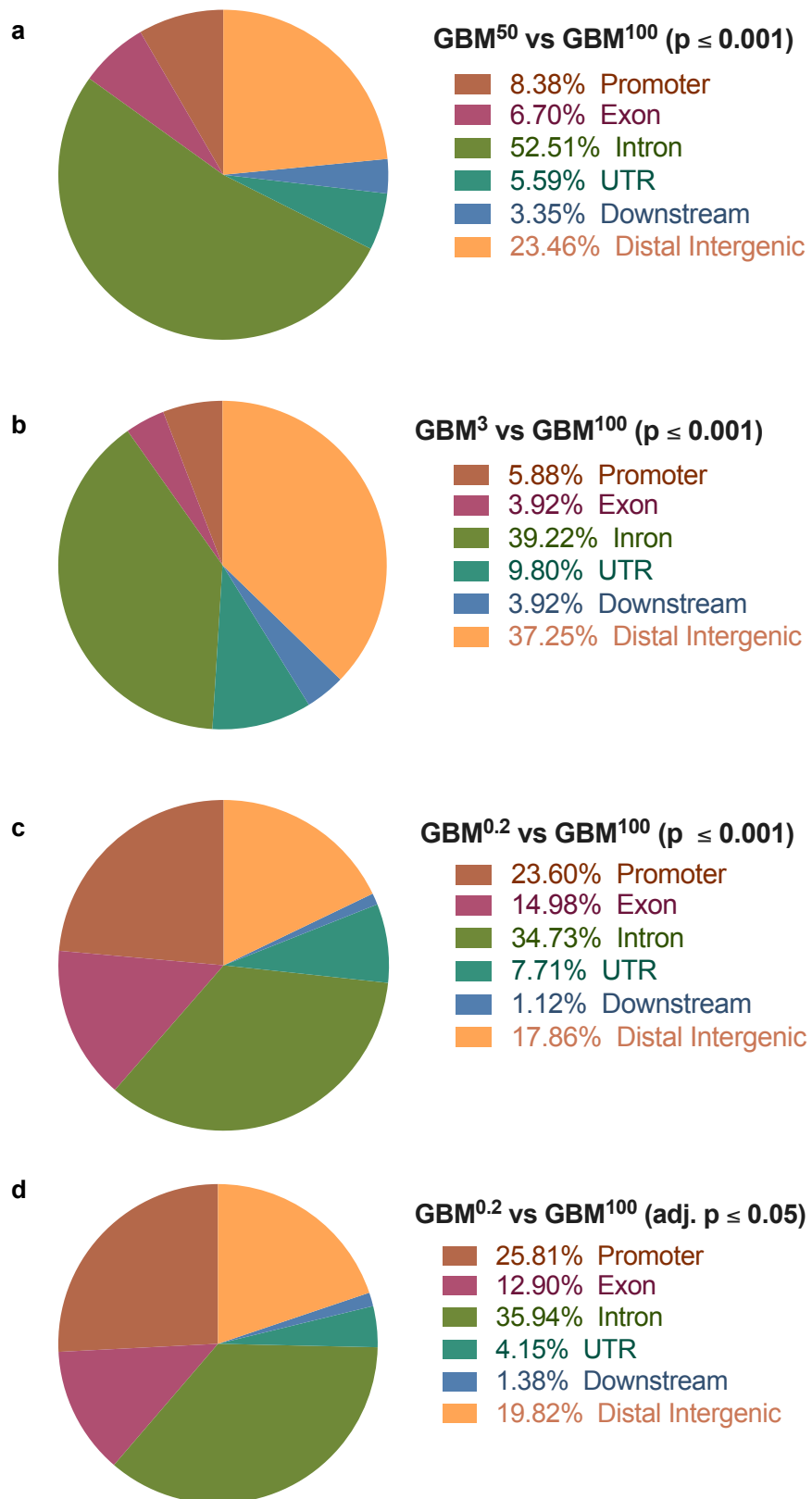


Figure 3

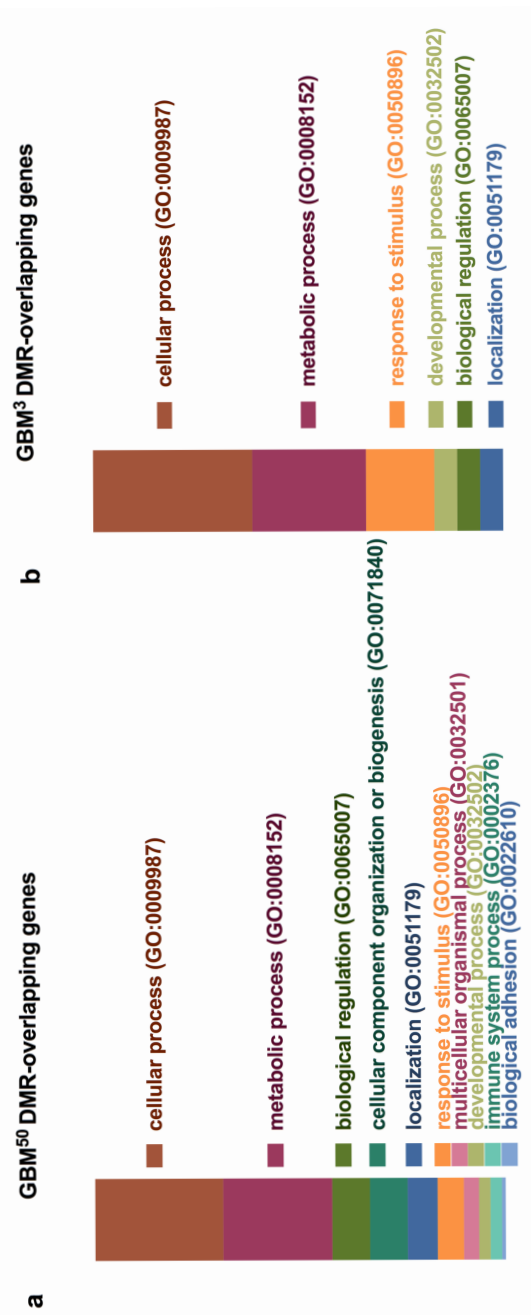


Figure 4

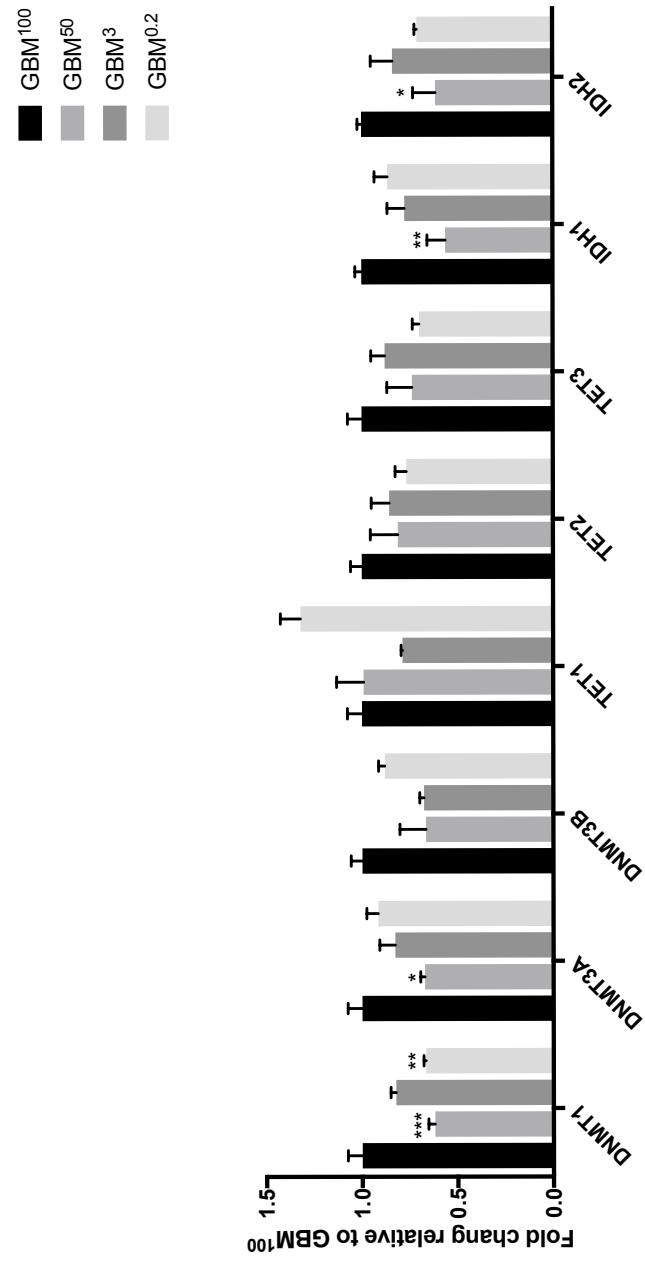


Figure 5

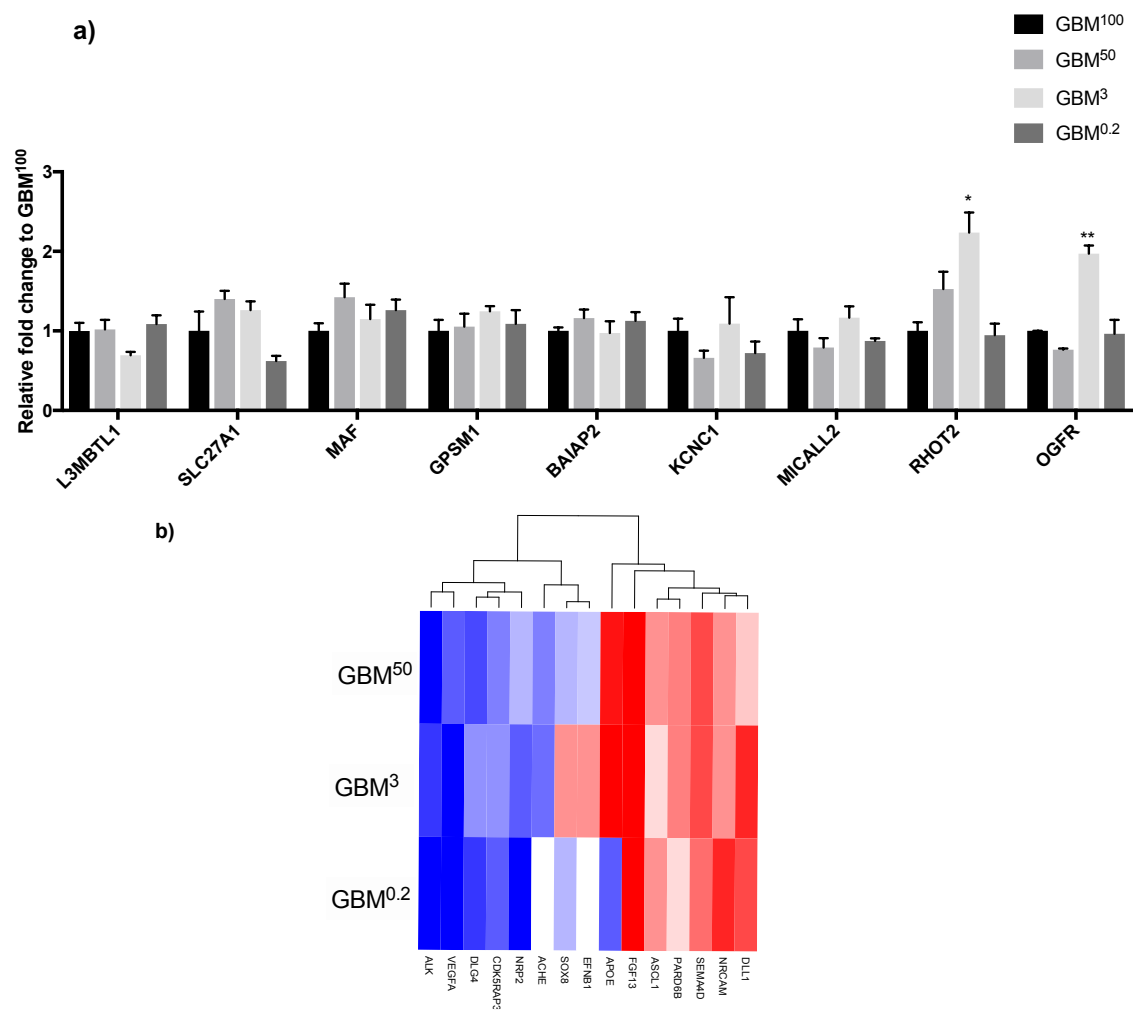
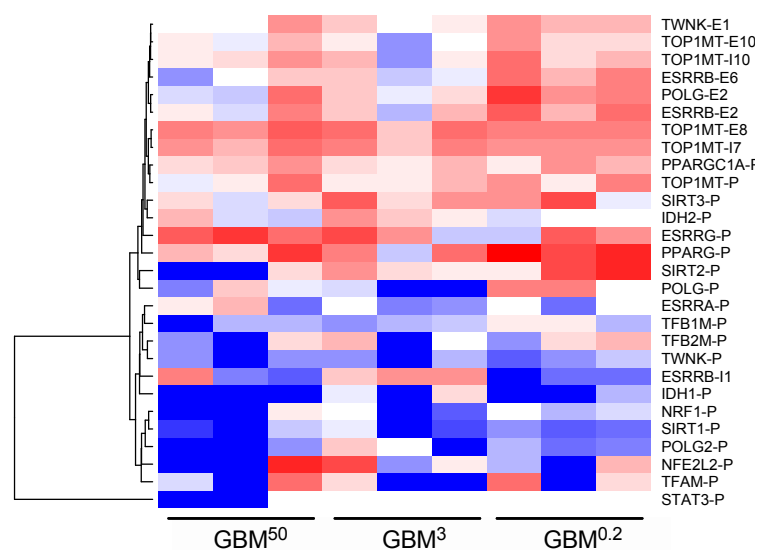
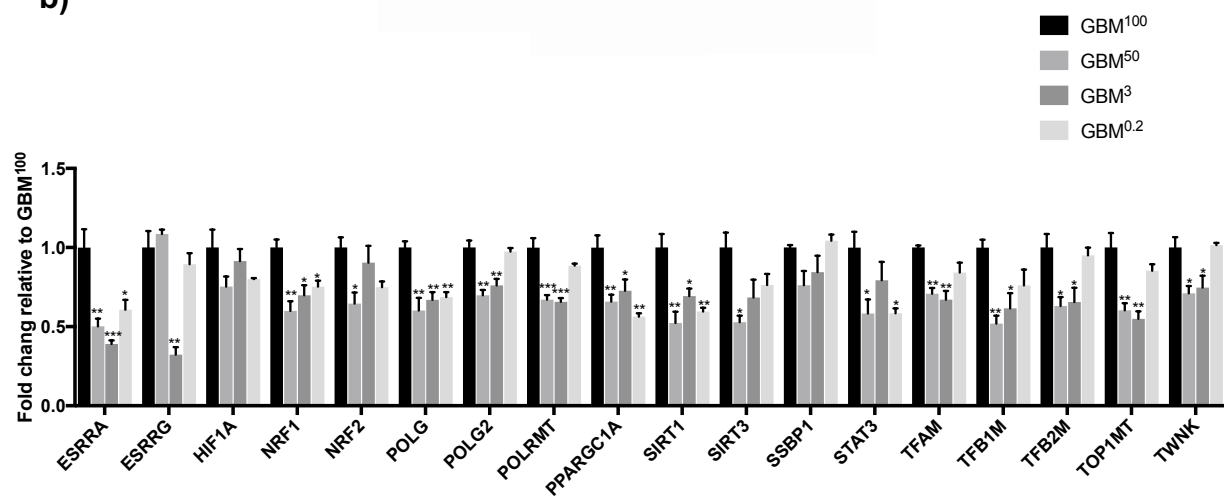


Figure 6

a)



b)



Additional files for

Modulation of mitochondrial DNA copy number during tumorigenesis induces changes in DNA methylation of the nuclear genome

Xin Sun^{1, 2}, Justin C St John^{1, 2}

¹ Mitochondrial Genetics Group, Hudson Institute of Medical Research, 27-31 Wright Street, Clayton VIC 3168, AUSTRALIA

² Department of Molecular and Translational Sciences, Monash University, 27-31 Wright Street, Clayton VIC 3168, AUSTRALIA

Corresponding author: Justin St. John, Voice: +61 3 8572 2678, Fax: +61 03 9594 7416;

Email: Justin.StJohn@hudson.org.au

Additional file 1. MeDIP-Seq specific QC results determined by the MEDIPS package.

	Saturation Score	Enrichment Score
GBM¹⁰⁰-1	0.97	2.07
GBM¹⁰⁰-2	0.98	2.55
GBM¹⁰⁰-3	0.95	1.89
GBM⁵⁰-1	0.98	2.33
GBM⁵⁰-2	0.97	2.17
GBM⁵⁰-3	0.95	1.97
GBM³-1	0.96	1.88
GBM³-2	0.97	2.28
GBM³-3	0.98	2.23
GBM^{0.2}-1	0.98	2.67
GBM^{0.2}-2	0.97	2.39
GBM^{0.2}-3	0.96	2.21

Additional file 2. CNV regions identified in each cohort of tumours determined using the Nexus 9.0 software.

	GBM¹⁰⁰	GBM⁵⁰	GBM³	GBM^{0.2}
CN GAIN	chr1:0-832,780 chr1:150,990,558-151,341,916 chr1:3,205,527-3,282,778 chr2:15,821,548-15,950,064 chr2:166,859,135-166,868,616 chr2:166,915,083-166,915,196 chr4:3,266,069-3,487,234 chr7:54,723,018-55,297,327 chr8:146,179,817-146,364,022 chr15:63,352,476-63,354,438 chrX:155,119,599-155,270,560 chrX:18,528,151-18,593,560 chrX:32,360,246-32,473,852	chr2:15,863,254-15,950,064 chr2:16,146,366-16,212,772 chr2:16,857,519-16,931,905 chrX:32,360,138-32,367,584	chr1:797,934-843,780 chr11:2,389,848-2,448,378 chr11:2,604,703-2,606,505 chr12:195,412-252,925 chr13:115,090,132-115,090,198 chr14:107,188,511-107,349,540 chr18:77,949,664-78,077,248 chr21:48,052,620-48,129,895 chr3:0-185,862 chr7:159,084,752-159,138,663	chr19:59,075,899-59,128,983 chr21:48,006,207-48,129,895 chr5:0-60,908 chr7:159,049,106-159,138,663
CN LOSS	chr1:224,267,066-224,397,010 chr1:235,149,351-235,458,521 chr1:236,217,325-236,261,585 chr11:108,098,592-108,115,578 chr12:34,700,607-35,800,000 chr12:38,223,599-38,312,274 chr13:20,347,985-20,493,343 chr15:44,489,615-44,856,850 chr17:41,256,189-41,278,437 chr2:47,635,575-47,638,531 chr2:47,639,550-47,651,407 chr4:40,034,626-40,103,792 chr6:74,067,561-74,350,353 chr7:61,814,839-62,249,302	chr1:235,861,367-235,914,143 chr12:34,700,607-35,800,000 chr5:45,947,493-46,399,093	chr14:34,950,211-35,075,115 chr22:29,092,959-29,121,341 chr4:57,682,271-57,697,521 chr5:46,117,343-46,399,093 chr6:107,469,470-107,540,766 chr7:61,859,728-62,000,116 chr8:43,457,427-43,831,387	chr10:42,502,256-42,632,033 chr11:50,392,461-51,328,556 chr12:34,507,145-35,800,000

Additional file 3. Primer pairs for real time PCR.

Human Gene	Forward primer (5'-3')	Product size	Annealing T _m (°C)
β-globin-F	CAACTTCATCCACGTTTCACC	268	57
β-globin-R	GAAGAGCCAAGGACAGGTAC		
mtDNA-F	CGAAAGGACAAGAGAAATAAGG	152	53
mtDNA-R	CTGTAAAGTTTTAAGTTTTATGCG		
SLC27A1-F	TCTCTCTGCTTCCCCAGGAT	206	58
SLC27A1-R	CGGATCAGCACAGAGAGACC		
RHOT2-F	CAGAGCGAAAGGCTTGAGGA	136	58
RHOT2-R	ACAGGATCAGCGACGTCTTC		
OGFR-F	CTCCGCATCACACGCATC	207	58
OGFR-R	AAGTGCTCCAGGCGAA		
MICALL2-F	ACATGATTGAGAAGCTGGGCCT	108	58
MICALL2-R	GCCCTACTGGCTACTACTGGG		
MAF-F	AGCAAGTCGACCACCTCAAG	168	58
MAF-R	CTGGAATCGCGTGTGAGACT		
L3MBTL1-F	GGTTTGGCTGGTGTAGCTTG	179	58
L3MBTL1-R	ACTCGCTTATCCGGAAGTGG		
KCNC1-F	CCCGTCATCGTGAACAATTTC	150	58
KCNC1-R	ACTGTGGTGTGGAGAGTTTAC		
GPSM1-F	CCTTCTTTGAGGCTGCTGT	135	58
GPSM1-R	GGAGGTCATGCTTGTGGTATT		
BAIAP2-F	AGAGCACACCCATCATGAAC	145	58
BAIAP2-R	GGGAGTGTGTTGGAGTAGGA		

Additional file 4. Taqman assays used in the Fluidigm qPCR arrays.

Gene name	Taqman assay number
Actb	Hs99999903_m1
18SrRNA	Hs99999901_s1
TWINK	Hs00958168_g1
DNMT1	Hs00154749_m1
DNMT3A	Hs01027166_m1
DNMT3B	Hs00171876_m1
EGFR	Hs01076090_m1
ESRRA	Hs01067166_g1
ESRRB	Hs01584024_m1
ESRRG	Hs00155006_m1
HIF1a	Hs00153153_m1
HPRT1	Hs02800695_m1
IDH1	Hs01855675_s1
IDH2	Hs00158033_m1
OAZ1	Hs00427923_m1
POLG	Hs00160298_m1
POLG2	Hs00200546_m1
POLRMT	Hs04187596_g1
PPARGC1A	Hs01016719_m1
SIRT1	Hs01009006_m1
SIRT2	Hs00247263_m1
STAT3	Hs01047580_m1
TERT	Hs00972650_m1
TET1	Hs00286756_m1
TET2	Hs00325999_m1
TET3	Hs00379125_m1
TFAM	Hs00273372_s1
TFB2M	Hs00915025_m1
TOP1MT	Hs01080056_m1
TP53	Hs01034249_m1

CHAPTER 5

The degree of mitochondrial DNA methylation in tumorigenesis.

Manuscript attached.

Rationale leading to Chapter 5

In the first two aims of my thesis, I have investigated the role of DNA demethylation in promoting mtDNA copy number and the effect of mtDNA modulation on the DNA methylation profiles in endpoint tumours. From the results in Chapter 3, I have also reported the presence of DNA methylation in the mitochondrial genome that can also be induced to demethylate through treatment with DNA demethylation agents. This is considered to be another factor in promoting mtDNA replication. Moreover, recent findings have reported the presence of DNA methylation in the mitochondrial genome from a variety of tissues and diseases, but without its function being determined. To further investigate the hypothesis that levels of mtDNA methylation regulate mtDNA replication, I have characterized the patterns of mtDNA methylation using glioblastoma and osteosarcoma tumour models under different scenarios: 1) different mtDNA genotypes under the same chromosomal background; 2) different tumour nuclear genome backgrounds; 3) different stages of tumour progression; and 4) different levels of mtDNA depletion. To ensure the reliability of my findings, I have taken advantages of three types of technologies including whole-mitochondrial genome bisulfite-sequencing, MeDIP, and pyro-sequencing. The findings from Chapter 5 could provide novel insights into mtDNA methylation in the epigenetic regulation of mtDNA copy number in tumorigenesis.

The degree of mitochondrial DNA methylation in tumorigenesis

Xin Sun^{1,2}, Vijesh Vaghjiani^{2,3}, W Samantha N Jayasekara^{2,3}, Jason E. Cain^{2,3}, Justin C St John^{1,2}

¹ Mitochondrial Genetics Group, Hudson Institute of Medical Research, 27-31 Wright Street, Clayton, VIC 3168, AUSTRALIA

² Department of Molecular and Translational Sciences, Faculty of Medicine, Nursing and Health Sciences, Monash University, 27-31 Wright Street, Clayton, VIC 3168, AUSTRALIA

³ Centre for Cancer Research, Hudson Institute of Medical Research, 27-31 Wright Street, Clayton, VIC 3168, AUSTRALIA

Corresponding author: Justin St. John, Voice: +61 3 8572 2678, Fax: +61 03 9594 7416; Email: Justin.StJohn@hudson.org.au

Abstract

Different cell types possess different copies of mtDNA to support their specific requirements for cellular metabolism. Cell-specific mtDNA copy numbers are established through cell-specific mtDNA replication during cell differentiation. However, cancer cells are trapped in a 'pseudo-differentiated' state as they fail to expand mtDNA copy number. Global DNA methylation can regulate this process, as induced DNA demethylation promotes differentiation of cancer cells and expansion of mtDNA copy number. To determine the role that mtDNA methylation plays in regulating mtDNA replication during tumorigenesis, we have characterized the patterns of mtDNA methylation using glioblastoma and osteosarcoma tumour models that have different combinations of mtDNA genotypes and copy number against common nuclear genome backgrounds at different stages of tumour progression. To ensure the reliability of the findings, we have applied a robust experimental pipeline including three approaches, namely whole-mtDNA bisulfite-sequencing with mtDNA-genotype-specific analysis, pyro-sequencing, and methylated immunoprecipitation against 5mC and 5hmC. We have determined genotype-specific methylation profiles, which were modulated through tumour progression. Moreover, a strong influence from the nuclear genome was also observed on mtDNA methylation patterns using the same mtDNA genotype under different nuclear genomes. Our findings highlight the influences that the nuclear and mitochondrial genomes have in setting mtDNA methylation patterns to regulate mtDNA copy number in tumorigenesis.

Author summary

In mammalian cells, not only the nucleus contains DNA, but also the mitochondria. Mitochondrial DNA (mtDNA) is a small, circular genome that encodes 13 subunits of the electron transfer chain which is located in the inner membrane of mitochondria to produce the majority of cellular energy (ATP), through the biochemical process of oxidative phosphorylation. mtDNA is different to nuclear DNA as there are multiple copies per cell, which need to be present at sufficient levels for each cell type to generate enough energy to perform its specialised functions. However, it is similar to nuclear DNA in that its genetic activities, such as gene transcription and replication, can be potentially regulated by DNA methylation, a modification which occurs to DNA. Cancer cells have specific metabolic profiles and maintain low mtDNA copy number. Our data show the presence of mtDNA methylation and its potential role in manipulating mtDNA copy number in cancer.

Introduction

In human cells, mitochondria possess their own genome, mitochondrial DNA (mtDNA). mtDNA is organized as a circular, double-stranded structure [1], which is maternally-only inherited. The human mitochondrial genome is 16569 bp in size and is present in multiple copies per cell [1]. It is essential for cellular function as it encodes 13 subunits of the electron transfer chain (ETC) complexes, where oxidative phosphorylation (OXPHOS) takes place, the major pathway for generating cellular energy. mtDNA also encodes 22 transfer RNAs (tRNAs) and 2 ribosomal RNAs (rRNAs). In contrast to the nuclear genome, the mitochondrial genome has only two non-coding regions. The major non-coding region, the D-loop, contains the transcription promoter sites for the heavy and light strands (HSP1/2 and LSP), and the initiation site for heavy strand replication (O_H). As a result, this region is the site of interaction for the nuclear-encoded transcription and replication factors that translocate to the mitochondrion to mediate mitochondrial genomic activities [1]. The other non-coding region is located two-thirds downstream, and contains the initiation site for light strand replication (O_L) [1].

mtDNA copy number varies according to cell type. This is established during early development when mtDNA replication is strictly regulated in order that a differentiated cell is able to acquire sufficient copies of mtDNA to meet its specific energy requirements [2, 3]. Indeed, successful differentiation is subject to a naïve cell establishing the mtDNA set point, which is defined as the lowest number of mtDNA copies required to be present (~ 200 copies) to act as initiating templates for replication as the process of differentiation ensues [2-6]. However, mtDNA replication appears to be blocked as cancer cells are induced to differentiate [6]. A

number of cancer cell types maintain low numbers of mtDNA copy as they primarily rely on aerobic glycolysis for energy production, which promotes cellular proliferation rates at the expense of cell differentiation [4]. Indeed, cancer cells are trapped in a 'pseudo-differentiated' state whereby they are unable to maintain or reinforce the mtDNA set point and, therefore, failed to expand mtDNA copy number [7, 8].

Recently, it has become apparent that DNA methylation could take place in the mitochondrial genome, especially in the main non-coding region [9-11]. This is further supported by the finding that DNA methyltransferases, namely the DNMT enzymes, and DNA demethylation enzymes, the ten-eleven translocation methylcytosine dioxygenases (TET), have been found in mitochondria [12, 13]. The DNMT enzymes methylate cytosine into 5-methylcytosine (5mC), whereas the TET enzymes promote DNA demethylation by oxidizing 5mC to 5-hydroxymethylcytosine (5hmC) [12, 13]. This is also consistent with our previous finding that the DNA demethylation agents vitamin C (VitC), the activator of TET, and 5-Azacytidine (5Aza), the inhibitor of DNMT, were able to significantly reduce levels of mtDNA methylation [11]. Consequently, the use of DNA demethylation agents could promote tumour cells to complete differentiation and expand mtDNA copy number, which is mediated by the levels of DNA methylation at exon 2 of *POLG*, the gene encoding the mtDNA-specific polymerase gamma, and exon 8 of *TOP1MT*, the gene encoding the mtDNA-specific topoisomerase [4, 11, 14]. Furthermore, as mtDNA appears to synergistically undergo DNA demethylation [11], it is likely that DNA demethylation to mtDNA would provide extra template to promote mtDNA replication. This indicates an extra epigenetic layer in the regulation of mtDNA replication.

Importantly, another feature of mtDNA is its susceptibility to variants [15]. mtDNA haplotypes, defined by the phylogenetic origins of their maternally inherited lineages [16], have been associated with a range of human diseases such as cancer [17], diabetes [18], Alzheimer's [19] and Parkinson's [20], and infertility [21, 22]. Different mtDNA genotypes are able to induce differential gene expression patterns of the same nuclear genome in stem cell [23] and tumour [5] models. Therefore, it is of great importance to take mtDNA genotypes into consideration when investigating mtDNA methylation. For instance, to determine mtDNA methylation levels using whole-genome bisulfite-sequencing technology, non-specific mapping to the published human reference genome, such as the hg38 Human Genome Assembly, could potentially produce over- or under-estimation of mtDNA methylation levels.

In order to determine the degree that mtDNA methylation regulates mtDNA replication during tumorigenesis, we have characterized the patterns of mtDNA methylation using tumour models of glioblastoma and osteosarcoma that have different combinations of mtDNA genotypes and copy number against the same nuclear genome background at different stages of tumour progression. We have applied three types of approaches including whole-mitochondrial genome bisulfite-sequencing with mtDNA-genotype-specific analysis, methylated immunoprecipitation (MeDIP), and pyro-sequencing. The findings provide novel insights into mtDNA methylation and its role in the epigenetic regulation of mtDNA copy number in tumorigenesis.

Results

Characterization of genotype-specific mtDNA methylation through bisulfite-sequencing.

By isolating mitochondria and purifying mtDNA from early and late stage tumours that possessed the same chromosomal backgrounds but different mtDNA genotypes from osteosarcoma (143B), glioblastoma multiforme (GBM) and human neuron stem cells (NSC), namely, 143B^{143B}, 143B^{GBM} and 143B^{NSC} tumours, and performing whole-genome-bisulfite sequencing, we were able to achieve a coverage of more than 500-fold on average across the mitochondrial genome amongst the tumours. This is deemed to be excellent in terms of mtDNA bisulfite sequencing analysis (S1 Table) and provides sufficient depth to identify the overall levels of mtDNA methylation at individual sites [24].

In order to determine the specific mtDNA methylation profiles for each genotype, we first generated long PCR products as linearized control samples for each mtDNA genotype to ensure the correct context for each comparison. Each long PCR control consisted of equal concentrations of two overlapping sequences (~ 8500 kb) that span the whole mitochondrial genome. The positive control samples were treated with *M.SssI* CpG methyltransferase to identify the methylated bases for each genotype and the levels of methylation at each site. The negative controls were not treated with the enzyme and would act as the baseline, indicative of non-mtDNA methylation. In the published human mitochondrial genome (hg38), there are a total of 870 CpG predicted sites on both strands of the mitochondrial genome [9], which is same for the 143B^{NSC} genotype (S2 Table). The differences in CpG profiles amongst the mitochondrial genomes of the 143B^{143B}, 143B^{GBM} and 143B^{NSC} tumours are evident at

nt 9053 - 9054 and nt11911 - 11912, which are only present in the 143B^{143B} and 143B^{GBM} genomes (S2 Table). CpG sites that were validated by the long PCR samples were kept for further analysis. These had a minimum of 10 reads at each site across all control samples. In total, there were 810 CpG sites from both strands validated for further comparisons (S3 Table). The full potential of DNA methylation at each CpG site determined by the positive and negative linearized control samples demonstrated the feasibility of mtDNA to undergo methylation and provided the baseline for genotypic data normalization in subsequent analyses (S3 Table).

We then determined the overall mtDNA methylation profiles for the 143B^{143B}, 143B^{GBM} and 143B^{NSC} tumours and GBM cells (Fig.1). On average, the normalised methylation levels at CpG sites over the genome were 10% (Fig.1), which is consistent with our previous finding using a MeDIP-qPCR assay and findings reported by others [12, 25].

The changes in mtDNA methylation resulting from different mtDNA genotypes

Assessment of the 143B^{143B}, 143B^{GBM} and 143B^{NSC} tumours at early and late stages showed that several CpG sites exhibited significant differential levels of DNA methylation ($p < 0.05$ and difference in DNA methylation $> 5\%$ against the same on different mtDNA genotypes) as a result of different mtDNA genotypes under the same nuclear genome (Table 1). Between early tumours, there was only one site which is located within the *ND5* region (13209 on light strand) showing 7% lower DNA methylation in the 143B^{GBM} early tumours when compared with the 143B^{143B} early tumours. mtDNA derived from cancer cells, namely 143B^{143B} and 143B^{GBM}, showed more variable patterns compared with the 143B^{NSC} early tumours. 143B^{NSC}

early tumours had higher levels of methylation than the 143B^{143B} early tumours at two CpG sites within the *CYTB* region: 15.75% higher for site 14578 and 12.04% higher for site 14802. Likewise, 143B^{NSC} early tumours showed higher levels of DNA methylation than the 143B^{GBM} early tumours at 6 sites, namely the same two sites identified in the 143B^{143B} early tumours: 12.46% for site 14578, 19.64% for site 14802; and at site 498 within the major non-coding region, site 9143 within the *ATP6* region, site 13209 within the *ND5* region, and site 14829 within the *CYTB* region. These sites were uniquely hyper-methylated in the 143B^{NSC} early tumours when compared with the 143B^{GBM} early tumours.

Table 1. Changes in mtDNA methylation resulting from different mtDNA genotypes.

	Location of C	Strand	Annotation	143B ^{GBM} vs. 143B ^{143B}		143B ^{NSC} vs. 143B ^{143B}		143B ^{NSC} vs. 143B ^{GBM}	
Early Tumours	498	+	Non-coding					*	19.83%
	9143	+	ATP6					*	5.85%
	13209	-	ND5	**	-7.00%			***	7.54%
	14758	+	CYTB			**	15.75%	**	12.46%
	14802	+	CYTB			**	12.04%	**	19.64%
	14829	+	CYTB					**	7.20%
Late Tumours	3695	+	ND1	**	-1.69%	**	10.58%		
	3705	+	ND1	***	-7.57%			**	5.11%
	13712	+	ND5					*	7.31%

*, **, *** indicate *p* values of < 0.05, 0.01, 0.001, respectively.

Amongst the late tumours, similar patterns were observed, but fewer sites showed significant differences as a result of the different mtDNA genotypes. mtDNA derived from GBM cells had lower levels of DNA methylation than mtDNA derived for 143B at sites 3695 and 3705 within the *ND1* region. mtDNA derived from NSC had higher levels of DNA methylation at site 3695 (10.58%) than 143B^{143B} late tumours, and at sites 3705 (5.11%) and 13712 (7.31%) than the 143B^{GBM} late tumours. Therefore, in the presence of the 143B nuclear genome, mtDNA derived from a tumour background tended to have lower levels of mtDNA methylation than mtDNA derived from a non-

tumorigenic cell, such as NSC, suggesting that mtDNA methylation is necessary to enforce the tumour phenotype.

The changes in mtDNA methylation during tumorigenesis and assessment of regulatory regions using pyro-sequencing

We also determined whether the patterns of mtDNA methylation changed during the process of tumorigenesis. Indeed, two sites were significantly modified and showed a decrease in levels of mtDNA methylation from the early stage to the late stage, namely site 3950 within the *ND1* region in the 143B^{143B} tumours and site 10399 within the *ND3* region in the 143B^{NSC} tumours (Table 3).

Table 2. Changes in mtDNA methylation resulting from tumorigenic progression.

	Location of C	Strand	Annotation	P value	Difference
143B^{143B} Late vs. Early	3950	+	ND1	*	-3.97%
143B^{NSC} Late vs. Early	10399	+	ND3	**	-6.67%

*, ** indicate *p* values of < 0.05, 0.01, respectively.

As mtDNA copy number was significantly increased in all three tumour types during the process of tumorigenesis [5], the patterns across the D-loop, the site of interaction for the nuclear-encoded mtDNA replication factors, were further assessed using pyro-sequencing on the 143B^{143B} cells, and 143B^{143B} early and 143B^{143B} late tumours. Pyro-sequencing is a sensitive and accurate technique for determining DNA methylation levels, especially for the regions such as the D-loop that showed relatively lower levels of methylation (as seen in Fig. 1). Regions targeting HSP (Fig 2a), LSP (Fig 2b) and *ND6* (Fig 2c; the only gene on the light strand) were successfully assessed. A decrease in levels of mtDNA methylation were observed at sites 498 and 545 for the HSP region, site 499 for the LSP region and sites 14225 and 14248 for the *ND6* region from the early to the late stages. Furthermore, site 421

within the LSP region had significantly lower levels of DNA methylation in the late tumours compared with the cells.

The impact of different nuclear genomes on mtDNA methylation

To understand the bi-directional control of the mitochondrial and nuclear genomes, it is necessary to investigate the impact of different nuclear genomes on mtDNA methylation. Thus, the mtDNA methylation profile of GBM cells was compared with 143B^{GBM} tumours as they possess the same mitochondrial genome (GBM), but are combined with different nuclear genomes (GBM and 143B).

GBM mtDNA methylation appeared to be extensively modulated under the different nuclear genomes (Table 3). In total, 25 sites were significantly different in the 143B^{GBM} early tumours compared with GBM cells, whereas 19 sites were significantly different in the 143B^{GBM} late tumours compared with GBM cells. Amongst these sites, 16 of them were commonly identified with similar differences in DNA methylation, including both higher and lower levels of DNA methylation in the tumours. Two sites within each of the *RNR2*, *ND1*, *ND2*, and *COI* regions, and one site in each of the *ATP6/8*, *TR* and *ND4L* regions showed higher levels of DNA methylation when GBM mtDNA is in combination with the 143B nuclear genome. Fewer sites showed significantly lower levels of DNA methylation in the 143B^{GBM} tumours, including one site in each of *COI*, *ND4*, *ND5*, *ND6* and *CYTB*. There were also uniquely differentially methylated sites identified in the 143B^{GBM} early and late tumours when compared with GBM cells. For the 143B^{GBM} early tumours, 7 sites including one site within the D-loop region, three sites within *RNR2*, and one site in each of *ND1*, *COI*, and *ND5* were uniquely changed. For the 143B^{GBM} late tumours, 3

sites including one site within *RNR1*, and one site in each of *ATP6* and *TG* were uniquely changed. These findings indicate that the nuclear genome strongly exerts an influence on the patterns of mtDNA methylation.

Table 3. Changes in mtDNA methylation of GBM genotype resulting from different nuclear genomes.

Location of C	Strand	Annotation	143B ^{GBM} Early vs. GBM Cell		143B ^{GBM} Late vs. GBM Cell	
			P	Difference	P	Difference
170	+	Non-coding	*	7.49%		
1597	+	MT-RNR1			*	4.96%
2719	-	MT-RNR2	*	34.26%	**	27.12%
2810	-	MT-RNR2	*	-30.90%		
3022	-	MT-RNR2	*	-7.12%		
3061	-	MT-RNR2	*	8.50%		
3094	-	MT-RNR2	**	5.58%	**	4.74%
3945	-	ND1	*	6.56%		
3950	+	ND1	*	6.81%	*	9.48%
3965	+	ND1	*	4.92%	*	3.03%
4664	-	ND2	*	6.06%	*	4.63%
5110	+	ND2	*	5.61%	*	6.75%
6241	+	CO1	*	9.05%		
6850	+	CO1	*	-12.72%	*	-8.53%
7218	-	CO1	*	5.96%	*	4.56%
7336	-	CO1	*	5.02%	*	3.97%
8544	-	ATP8; ATP6	**	8.51%	*	7.57%
9008	+	ATP6			*	6.14%
10012	+	TG			*	8.09%
10169	-	ND3	*	12.33%		
10174	+	ND3	*	-12.72%		
10426	-	TR	**	7.18%	*	8.24%
10584	+	ND4L	**	4.66%	*	5.54%
11765	-	ND4	*	-5.74%	**	-6.44%
13120	-	ND5	***	-8.22%	***	-7.38%
13938	+	ND5	*	-34.96%		
14248	-	ND6	*	-11.83%	*	-12.43%
15043	+	CYTB	**	-31.92%	*	-31.28%

*, **, *** indicate *p* values of < 0.05, 0.01, 0.001, respectively.

The conversion of 5mC to 5hmC in mtDNA identified by MeDIP

Following the findings identified using whole mtDNA bisulfite-sequencing, we applied MeDIP using antibodies to distinguish between 5mC, the methylated state, and the demethylated status of 5hmC. We focused on the key regulatory regions in mtDNA, namely O_H, O_L, LSP and HSP, to further determine the levels of plasticity of mtDNA methylation. Generally, the levels of mtDNA methylation (5mC/5hmC) at O_H, HSP and LSP in the D-loop were highest in the 143B^{143B} tumours, followed by the 143B^{GBM} and 143B^{NSC} tumours (Fig. 3). O_L had lower levels of DNA methylation than the sites in the major non-coding region. Therefore, for each mtDNA genotype, DNA methylation presented in similar patterns across the genome with the major non-coding region having higher levels of methylation than the minor non-coding region.

We further investigated whether the conversion from 5mC to 5hmC of mtDNA derived from GBM cells and NSCs varied when they were under different nuclear genomes. Amongst the O_H, O_L, HSP and LSP sites (Fig. 4), mtDNA from GBM cells presented similar patterns of DNA methylation when they were under the GBM nuclear genome and the 143B nuclear genome. Interestingly, mtDNA from hNSCs had much higher levels of DNA methylation when they are under the tumorigenic 143B nuclear genome than the healthy NSC nuclear genome. Therefore, mtDNA derived from the GBM tumour model could maintain its specific mtDNA methylation profiles under different tumorigenic nuclear genomes (143B or GBM), but mtDNA methylation derived from healthy cells tended to be reset by the tumorigenic nuclear genome.

Different mtDNA methylation patterns resulted from the restoration of mtDNA copy number initiated from cells with varying levels of mtDNA

Using GBM tumours formed from cells with varying levels of mtDNA, namely GBM¹⁰⁰ (100% of original content), GBM⁵⁰ (50% of original content), GBM³ (3% of original content) and GBM^{0.2} (0.2% of original content) tumours, we investigated the changes to mtDNA methylation (5mC/5hmC) after the restoration of mtDNA during tumorigenesis [6]. The GBM³ and GBM^{0.2} tumours exhibited significantly late onset and lower rates of tumour formation, as mtDNA copy number had to be restored to sufficient levels for the establishment of tumorigenesis [5, 6]. Once mtDNA had been sufficiently replicated to recover tumorigenic capability, mtDNA replication was restricted as tumour cells maintained low mtDNA copy number to support tumorigenesis. Here, we focused on the sites in the non-coding regions, namely O_H, O_L, HSP and LSP (Fig. 5). For the origins of replication, only the O_H site presented differential DNA methylation patterns, as there were no significant differences at O_L. GBM^{0.2} tumours had significantly higher levels of methylation than the other three cohorts, which likely restricts further replication after mtDNA copy number had been restored. GBM³ had significantly higher levels of methylation than the GBM⁵⁰ tumours, which might contribute to the relatively faster onset in the formation of the GBM⁵⁰ tumours [6]. GBM^{0.2} tumours also exhibited higher levels of methylation at HSP and LSP than the GBM¹⁰⁰ tumours.

Discussion

Our findings demonstrate that mtDNA methylation adds another layer in the epigenetic control to mtDNA copy number in tumour cells. It has been reported previously that the DNA methylation status at key regions of nuclear-DNA encoded mtDNA replication factors, namely exon 2 of *POLG* and exon 8 of *TOP1MT* are related to mtDNA copy number [4, 7, 8, 11, 14]. Since the findings that mtDNA demethylation can also be induced by DNA demethylation agents that likely contributes to the up-regulated mtDNA copy number [11], it is important to determine the potential role that mtDNA methylation plays in regulating mtDNA replication. However, the presence of mtDNA methylation has been hotly debated over several decades with its role in mitochondrial genetics remaining largely undetermined [12, 24, 26].

In this study, we employed a technical pipeline to remove ambiguity associated with the results. Firstly, we used purified populations of mtDNA to avoid the noise arising from the mtDNA pseudogenes present in the nucleus [11]. Secondly, it is the first time that genotype-specific analysis has been applied to assess mtDNA methylation, which is a necessary step given that mtDNA is susceptible to variants in, for example, cancers [15] and different mtDNA haplotypes are differently predisposed to various forms of cancer and other diseases [15, 27-29]. Indeed, we identified two CpG sites at nt 9053 - 9054 and nt 11911 - 11912 that were gained by the 143B and GBM mtDNA genotypes compared with the NSC genotype. Thirdly, the strategy of using linearized positive and negative controls for each genotype generated by long PCR and treated in the presence and absence of a DNA methylating enzyme, respectively, provided a baseline for data normalization. This enabled the full potential of each CpG site

amongst the population of mtDNA in the cells and tumours to be determined. Lastly, whole-genome bisulfite sequencing achieved significantly high coverage (> 500 fold coverage on average) to avoid possible over-estimation of mtDNA methylation levels resulting from low coverage of sequencing reads (< 250 fold) as argued in [24]. In addition to the issue of sequencing coverage, the circular and supercoiling structure of mtDNA has been considered as another concern in blocking bisulfite treatment and causing over-estimation in mtDNA methylation [24]. To this extent, the purified mtDNA samples and the linearized long PCR control samples were firstly fragmented before bisulfite conversion to overcome the potential effects of supercoiling, an approach we previously employed when assessing mtDNA methylation by MeDIP-Seq [11]. As a result, our findings were consistent with published results [12, 25]. Moreover, the use of 5mC- and 5hmC-specific antibodies in MeDIP enabled further determination of mtDNA methylation to understand the active demethylation process in the samples, namely the transition from 5mC to 5hmC, which is not distinguishable following bisulfite conversion. Nevertheless, pyro-sequencing, a highly sensitive method for determining the levels of DNA methylation at specific sites, further identified the subtle changes in the key regulatory regions of the early tumours.

In our studies, we have used a model of osteosarcoma, namely 143B cells, and early and late stage tumours derived from these cells that possessed the same chromosomal backgrounds but different mtDNA genotypes, i.e., 143B^{143B}, 143B^{GBM} and 143B^{NSC} cells and tumours. This enabled us to investigate the patterns of mtDNA methylation between different genotypes and under different stages of tumour progression. Indeed, these tumours have been shown to have different patterns of mtDNA replication during tumorigenesis [5]. Even though an increase in mtDNA copy number was

observed during tumour progression across all three types of tumours, the 143B^{NSC} and 143B^{GBM} tumours gained significantly greater mtDNA copy number at the later stage compared with the 143B^{143B} tumours [5]. Therefore, 143B tumours served as excellent models to investigate whether mtDNA methylation regulates mtDNA copy number.

We firstly identified genotype-specific mtDNA methylation profiles. In the presence of the 143B nuclear genome, mtDNA derived from cancer cells, i.e. 143B and GBM mtDNA, tended to have lower levels of mtDNA methylation than NSC mtDNA. All of the differentially methylated CpG sites within non-coding region, *ATP6*, *CYTB*, *ND1* and *ND5* identified in the early and late tumours had higher levels of DNA methylation in the 143B^{NSC} tumours than the 143B^{GBM} and 143B^{143B} tumours. Moreover, there were three sites that were significantly more hypo-methylated in the 143B^{GBM} tumours than the 143B^{143B} tumours, which indicates that different mtDNA genotypes from different cancer backgrounds could also lead to different patterns of mtDNA methylation and indicates that the mitochondrial genome is somehow responsible for establishing its methylation status independent of the nucleus. Indeed, different cancers have been associated with different combinations of mtDNA mutations [29, 30], which could contribute to the tumour-specific profiles and further emphasize the importance of the genotype-specific context of mtDNA in investigating mtDNA methylation. The findings identified by MeDIP further indicate that amongst these methylated CpG sites, there were different rates of 5mC to 5hmC conversion. Even though 143B^{NSC} tumours had higher levels of mtDNA methylation at several CpG sites, they exhibited greater levels of demethylation, especially within the D-loop region. 143B^{143B} tumours maintained the highest levels of 5mC over 5hmC

amongst the three types of tumours. This could be further associated with the changes in mtDNA copy number and that the higher levels of 5mC could contribute to the relatively restricted replication of mtDNA in the 143B^{143B} tumours. However, it appears that not just mtDNA methylation is affected by mtDNA genotype. In the 143B model, mtDNA genotype influenced the expression of chromosomal genes, which also gave clear indications of the genes specific to osteogenic tumorigenic initiation and maintenance [5].

During the process of tumorigenesis, decreases in the levels of mtDNA methylation were observed at two gene regions, namely site 3950 within the *ND1* region in 143B^{143B} tumours and site 10399 within the *ND3* region in the 143B^{NSC} tumours. The genes *ND1* and *ND3* are associated with tumorigenic mutations in complex I of ETC in cancers, changes in mtDNA replication mediated by surrounding mtDNA methylation sites could potentially affect the frequency of these mutated copies [29, 30]. Indeed, there were significant increases in mtDNA copy number amongst these tumours from the early to the late stages [5]. Likewise, a decreasing trend in DNA methylation was observed amongst the CpG sites within the *ND6*, HSP and LSP regions from the early stage (Fig. 2), which could potentially contribute to the active binding of the mtDNA replication factors to mtDNA, as is the case for *POLG* where it has been shown that the methylation status of exon 2 is associated with the binding of RNA polymerase II [14].

In order to investigate whether the levels of mtDNA methylation change under different nuclear genomes, we also assessed GBM cells through the same experimental pipeline. GBM cells had a distinct mtDNA methylation landscape from

the various combinations of 143B tumours. Indeed, GBM mtDNA had 25 CpG sites and 19 CpG sites were differentially methylated when compared with the methylation profiles of 143B^{GBM} early and late tumours, respectively. 16 of these CpGs were differentially methylated in the 143^{GBM} tumours. These sites are located across the whole genome, including *RNR2*, *ND1*, *ND2*, *CO1*, *ATP6/8*, *TR*, *ND4L*, *ND4*, *ND5*, *ND6* and *CYTB*. A number of findings that have focused on either specific regions or the whole mitochondrial methylome have identified regions but with different patterns of DNA methylation, as shown in mouse and human tissues (brain, liver, breast and ESCs) and associated with disease (aging, colon cancer, 143B and GBM) [11-13, 31-34]. These findings not only suggest that these regions can be differentially methylated in different scenarios, but also indicate a strong relationship between the mitochondrial methylome and tissue-specific mtDNA copy number in synchrony with the tissue-specific nuclear genome background. Both hyper- and hypo-methylation were observed, but over 60% of all the identified sites were hyper-methylated when the GBM mtDNA was in the presence of the 143B nuclear genome. Moreover, the 5mC/5hmC rates remained at similar levels across the key regulatory sites in the non-coding regions. Interestingly, NSC mtDNA had a greater conversion rate from 5mC to 5hmC over the D-loop region (O_H, LSP and HSP) when in the presence of the 143B nuclear genome (as indicated by lower 5mC/5hmC values), which indicates that mtDNA methylation derived from healthy cells tended to be reset by the tumorigenic nuclear genome. Indeed, TET enzymes are regulated by the isocitrate dehydrogenase (IDH) in the citric acid cycle (TCA) in mitochondria [35, 36]. In the TCA cycle, IDH generates α -ketoglutarate, which facilitates the activity of the TET enzymes. However, it has been reported that downregulation of hypoxia regulator HIF1 α promotes IDH in osteosarcoma [37], which provide an explanation for the greater

conversion from 5mC to 5hmC mediated by TET enzymes. Collectively, these findings indicate that the nuclear genome exerts a strong influence on the mitochondrial genome and its methylated status.

Another interesting perspective in our understanding of how mtDNA methylation regulates copy number is the adaption of mtDNA methylation during the formation of tumours from mtDNA-depleted cells. mtDNA copy number was restored to similar levels in end-point GBM tumours [6], which explains why the GBM³ and GBM^{0.2} cells took longer to initiate tumour formation [5, 6]. Moreover, various mtDNA variants have been identified within GBM and the accumulation of these variants is proportional to the amount of mtDNA being restored [29], which makes these tumours interesting models to understand GBM-specific replication of mtDNA. GBM^{0.2} tumours had significantly higher levels of DNA methylation within the D-loop region than the other three cohorts, which likely further restricts replication of mtDNA after they had re-established their original copy number. Indeed, once mtDNA had been sufficiently replicated to recover tumorigenic capability, mtDNA replication was restricted as tumour cells maintained low mtDNA copy number to support tumorigenesis. Indeed, instead of OXPHOS, tumours cells mainly rely on aerobic glycolysis for energy production, which switches the metabolic profiles to prioritize cellular proliferation and prevents differentiation from taking place [4]. Likewise, GBM³ tumours had significantly higher levels of methylation than the GBM⁵⁰ tumours, which likely contributes to the relatively faster onset in the formation of the GBM⁵⁰ tumours [6].

In all, by applying robust experimental pipelines to assess the levels of DNA methylation in 143B and GBM cells and tumour models, we not only confirmed the presence of mtDNA methylation, but also determined the levels of mtDNA methylation in different scenarios. Firstly, different mtDNA genotypes under the same nuclear genomic background modulated levels of mtDNA at several sites across the mitochondrial genome, whereby higher levels of 5mC are related to the relatively restricted expansion of mtDNA copy number in the 143B^{143B} tumours. Secondly, the same mtDNA genotype behaved differently under the influence of different nuclear genomes to regulate mtDNA methylation. Thirdly, levels of mtDNA methylation tended to decrease during tumour progression, which could potentially contribute to the increases in mtDNA copy number observed in these tumours. Lastly, after tumours had restored sufficient mtDNA to initiate tumorigenesis, higher levels of 5mC over the D-loop were acquired to potentially restrict further replication of mtDNA. Collectively, mtDNA methylation adds another layer in the epigenetic control of mtDNA replication.

Materials and methods

Cell culture

HSR-GBM1 cells were cultured in complete neural stem cell media consisting of Dulbecco's Modified Eagle Medium/Nutrient Mixture Media (DMEM/F-12; Thermo Fisher Scientific, MA, USA), 2% StemPro neural supplement (Thermo Fisher Scientific), basic fibroblast growth factor (bFGF; 20 ng/mL; Merck Millipore, MO, USA) and epidermal growth factor (EGF; 20 ng/mL; Merck Millipore) at 37 °C in 5% CO₂ and 95% humidity. The mtDNA content of HSR-GBM1 cells was depleted to varying levels, namely 50%, 3% and 0.2% of the original mtDNA content, by additionally culturing with 2'-3'-dideoxycytidine (ddC; 10 µM Sigma-Aldrich, MO, USA) and uridine (50 µg/mL; Sigma-Aldrich) [6].

143B cell lines possessing different mtDNA genotypes derived from HSR-GBM1 cells and human NSCs were previously generated, namely 143B^{GBM} and 143B^{NSC}, as described in [5]. 143B^{143B}, 143B^{GBM} and 143B^{NSC} cell lines were cultured in standard DMEM (SD-DMEM) consisting of DMEM, 10% (v/v) FBS, sodium pyruvate (10 mM), GlutaMax (2 mM) and 1% (v/v) penicillin/streptomycin (Thermo Fisher Scientific) at 37 °C in 5% CO₂ and 95% humidity [5].

Xenograft models

GBM tumours were previously generated from HSR-GBM1 cells possessing 100%, 50%, 3% and 0.2% of their original mtDNA content, namely GBM¹⁰⁰ (non-depleted), GBM⁵⁰, GBM³ and GBM^{0.2} tumours, respectively, as described in [6]. 143B tumours were previously generated from 143B^{143B}, 143B^{GBM} and 143B^{NSC} cell lines, as described in [6]. Briefly, 0.5 million of the depleted and non-depleted HSR-GBM1

cells, or 1 million of each type of the three 143B cell lines in 100 μ L of Matrigel (Corning, New York, USA) were inoculated subcutaneously into the right flank of 8-week-old, female BALB/c nude mice. Tumour growth rates and volumes were reported in [5, 6]. The animal work was approved by Animal Ethics Committee A, Monash University, Approval Numbers: MMCA/2011/76, MMCA 2012/24 and MMCA 2013/05.

DNA and RNA extraction

Total genomic DNA and RNA were extracted from cultured cell pellets or tumour tissues using the DNeasy Blood & Tissue Kit and RNeasy Mini Kit (Qiagen, CA, USA), respectively, according to the manufacturer's protocols with minor modifications. The DNA samples were treated with 3 μ L of RNase A (Qiagen) at room temperature and Proteinase K (20 μ g/ μ L; Bioline, London, UK) at 65 °C for 10 min. The RNA samples were treated with DNase I (3 kunitz units/ μ L; Qiagen) in the presence of RDD buffer at room temperature for 20 min.

Purification of mtDNA from tumour tissues and cells

To eliminate mtDNA pseudo-genes present in the nuclear genome, mtDNA was purified from tumour tissues and cells, as described in [11]. Cells (~ 10 million) and tumour tissues (~ 20 mg) were resuspended in 5 mL of solution A (20 mM HEPES-KOH, pH 7.6, 220 mM Mannitol, 70 mM sucrose, 1 mM EDTA) and 2 mg/mL BSA was added. The suspension was incubated on ice for 15 min to facilitate swelling and then homogenized at 4 °C using a 5 mL Potter-Elvehjem tissue grinder set (Wheaton, USA) for 50 repetitions. The homogenate was centrifuged at 800 g for 10 min to remove cell debris and nuclei. The supernatant was then centrifuged at 10,000 g for

20 min at 4 °C to pellet the mitochondrial fraction. To further remove nuclear DNA, the pellet was resuspended in 175 µL of solution B (solution A without EDTA) with DNase I (30 kunitz units; Qiagen) and incubated at 37 °C for 30 min. 1 mL of solution A was then added to stop DNase activity and the suspension was centrifuged at 10,500 g for 20 min at 4 °C. The supernatant was discarded and the mitochondrial pellet was resuspended in 200 µL of lysis buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA and 1% SDS) with 1 µL of proteinase K (20 mg/mL, Bioline). The suspension was then incubated at 50 °C for 60 min. mtDNA was purified using the DNeasy Blood & Tissue Kit (Qiagen), according to manufacturer's protocol.

Generation of negative and positive controls for mtDNA bisulfite-sequencing by long PCR

To generate experimental controls for whole-mitochondrial-genome bisulfite-sequencing, long PCR products from mtDNA were generated as negative controls. PCR products do not maintain methylation modifications and, therefore, are deemed to be unmethylated. Two overlapping mtDNA sequences spanning the whole mitochondrial genome were generated by long PCR and combined in equivalent concentrations as the negative control. Each long PCR reaction contained 1 unit of Platinum Taq High Fidelity (Thermo Fisher Scientific) in High Fidelity PCR buffer, 100 mM MgSO₄, 1 mM dNTPs (Bioline) and 10 µM of each forward and reverse primer (S4 Table) in a total volume of 50 µL. PCR cycling profiles were initiated at 94°C for 2 min, followed by 35 cycles of 94 °C for 15 sec, 63 °C for 30 sec and 68 °C for 8 min, 45 sec. Long PCR products were then purified using a PCR QIAquick PCR Purification Kit (Qiagen), according to manufacturer's instructions. Importantly, in order to generate mtDNA-genotype specific long PCR controls, DNA samples

purified from 143B^{143B}, 143B^{GBM} and 143B^{NSC} tumours were used as the respective templates.

Similarly, the positive controls comprised the two long PCR products that had been combined and undergone DNA methylation treatment using an *in vitro* CpG methyltransferase, *M.SssI* (New England Biolabs, MA, USA), according to manufacturer's instructions. Briefly, 1 µg of combined long PCR products was treated with 4 units of the *M.SssI* enzyme in the presence of 160 µM S-adenosylmethionine (SAM) at 37 °C for 4 h. The reaction was then terminated by incubating the products at 65 °C for 20 min. Long PCR products were then purified from the mixture using the PCR QIAquick PCR Purification Kit (Qiagen), according to manufacturer's instructions.

Bisulfite sequencing

Purified mtDNA samples and control samples generated using long PCR were submitted to the MHTP Medical Genomics Facility (Clayton, Australia) to perform bisulfite sequencing. Libraries were prepared using the Zymo Pico Methyl-Seq™ Library Prep Kit following protocol 1.0.0 (Zymo Research, CA, USA), which is compatible with Illumina's TruSeq chemistries. Briefly, all samples were processed with 50 ng of DNA based on measurements generated by Qubit Fluorometric Quantitation (Thermo Fisher Scientific). Samples underwent 16 cycles of amplification, as recommended by the manufacturer's protocol. Samples then underwent random fragmentation to linearize mtDNA molecules to overcome the overestimation of DNA methylation caused by the circular and supercoiling structure of mtDNA molecules, as raised in [24]. Samples then underwent bisulfite conversion,

random priming and addition of sequencing indexed adaptors, according to the protocol. Libraries were pooled in equivalent molar ratios. 14 pM of a single pool was assessed by qPCR and then used for generating the sequencing clusters. Size and concentration of the libraries were checked using a 2100 Bioanalyzer with the High Sensitive DNA Kit (Agilent). 150-bp paired-end sequencing was then performed on the Illumina MiSeq v2 platform.

mtDNA- genotype- specific analysis for bisulfite sequencing

Firstly, adaptors and poor-quality reads were cleaned from raw sequences using the TrimGalore program (v0.4.5; http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) in the paired-end mode with 10 bp trimmed at both ends. Trimmed sequences were checked by the FastQC program to ensure the quality of reads and for complete removal of any remaining adaptors.

Since mtDNA is known to be susceptible to variants in both non-coding and coding regions in cancer [15], it is important to select matching mitochondrial genotypes as the reference genomes for the mapping of mtDNA sequences [22]. We had previously sequenced the 143B^{143B}, 143B^{GBM} and 143B^{NSC} mitochondrial genomes and these are available in GenBank under accession numbers KT946592, KT946593 and KT946594, respectively [5]. The sequences were mapped to their corresponding mitochondrial genomes using the Bismark Package (v0.19.0) with the paired-end mode set to the parameter of ‘-bowtie2 -non_directional -bam’ [38]. PCR duplicates were then removed using the function of ‘deduplicate_bismark’ with the parameters “-p -bam”. The coverage of the methylated and unmethylated CpG and non-CpG sites

in the mitochondrial genome were constructed using the function of 'bismark_methylation_extractor' with the parameters '-p -comprehensive -merge_non_CpG -bedgraph -CX -counts -cytosine_report -CX'.

CpG sites that were covered by a minimum of 10 reads in all of the negative and positive long PCR samples generated for each genotype were kept for further analysis. The differences in the percentage methylation between the negative and the positive long PCR samples were determined to be the full potential (100%) for each of the CpG sites in each mtDNA genotype [39]. These full potentials were then applied to normalise the levels of DNA methylation of the respective tumour samples at their corresponding CpG sites. Normalised levels of DNA methylation then underwent comparison using the One-Way ANOVA method in Prism 7.0 software (GraphPad Software, CA, USA).

Immunoprecipitation of methylated DNA (MeDIP)

Purified mtDNA underwent MeDIP, as described in [40]. Briefly, 5 µg of purified mtDNA was sheared into 200 - 1000 bp fragments by the Covaris Adaptive Focused Acoustics (AFATM) S220 system. The DNA was denatured by heating at 95 °C for 10 min and then cooled on ice for 5 min, which also avoided the potential issue raised by the circular, supercoiled structure of mtDNA [24]. 1.5 µg of anti-5mC or anti-5hmC antibody (Active Motif, CA, USA) were added to 3 µg of DNA fragments in the presence of 20 µL of Dynabeads[®] Protein G (Thermo Fisher Scientific) in 500 µL of IP buffer (100 mM sodium phosphate, pH 7.0; 1.4 M NaCl ; 0.5% Triton X-100). The suspension was incubated at 4 °C for 16 h under rotation. The beads were collected on a magnetic particle concentrator (Thermo Fisher Scientific) and washed with 1 mL of

IP buffer three times. The beads were then resuspended in 250 μ L proteinase K digestion buffer (50 mM Tris-HCl, pH 8.0; 10 M EDTA, pH 8.0; 1.0% SDS) with 10 μ L of proteinase K (20 mg/mL; Bioline) and incubated on a thermo-shaker at 50 °C for 3 h. The supernatant was then collected into a new tube. DNA was purified from the elutant using the QIAquick PCR Purification Kit (Qiagen) and collected in 50 μ L of autoclaved Milli-Q H₂O. Levels of DNA methylation for the regions of interest were quantified by qPCR on the Rotor-Gene 3000 machine under primer-specific conditions (S4 Table). The ratio of 5mC/5hmC was determined.

Pyrosequencing

DNA samples were submitted to the Australian Genome Research Facility (Perth, Australia) to perform pyrosequencing. DNA samples were firstly converted using the Epitect Bisulphite Kit (Qiagen), according to the manufacturer's instructions. Bisulfite-converted primers were designed using the PyroMark Assay Design software for the regions of interest (S4 Table). The regions of interest were amplified using the PyroMark PCR Kit (Qiagen) with biotin labelled primers that were purified using HPLC (S1 Table). PCR products were then immobilized to the Streptavidin Sepharose High Performance beads (GE Healthcare Life Sciences, MA, USA). PCR products with the beads underwent denaturation to anneal with the pyrosequencing primers (S4 Table). Pyrosequencing was then performed on a PyroMark 24 PyroSequencing system (Qiagen), according to manufacturer's instructions. DNA methylation levels for each CpG site were exported using PyroMark Q24 software. Data that were identified as having failed bisulfite conversion and QC were ignored for further analysis. Statistical differences were determined amongst groups using One-way ANOVA (comparison between three groups) or Student T-test (comparisons

between two groups) ($n = 3$; mean \pm SEM).

Data availability

The bisulfite sequencing datasets supporting the conclusions of this article are available in Sequence Read Archive (SRA) through accession number PRJNA478095.

Acknowledgements

We are grateful to Dr. Trevor Wilson and Ms. Jodee Gould, MHTP Medical Genomics Facility, Monash Health Translation Precinct, for assistance in performing bisulfite

sequencing. We are also grateful to Dr. David Chandler, Australian Genome Research Facility (Perth), for assistance in performing pyro-sequencing.

References

1. Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, et al. Sequence and organization of the human mitochondrial genome. *Nature*. 1981;290(5806):457-65.
2. Facucho-Oliveira JM, Alderson J, Splikings E.C., Egginton S., and St. John J.C. Mitochondria DNA replication during differentiation of murine embryonic stem cells. *J Cell Sci*. 2007;15(120):4025-34.
3. Facucho-Oliveira JM, St John JC. The relationship between pluripotency and mitochondrial DNA proliferation during early embryo development and embryonic stem cell differentiation. *Stem Cell Rev*. 2009;5(2):140-58.
4. Lee W, Johnson J, Gough DJ, Donoghue J, Cagnone GL, Vaghjiani V, 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(2):e1664.
5. Lee WT, Cain JE, Cuddihy A, Johnson J, Dickinson A, Yeung KY, et al. Mitochondrial DNA plasticity is an essential inducer of tumorigenesis. *Cell Death Discov*. 2016;2:16016.
6. Dickinson A, Yeung KY, Donoghue J, Baker MJ, Kelly RD, McKenzie M, et al. The regulation of mitochondrial DNA copy number in glioblastoma cells. *Cell Death Differ*. 2013;20(12):1644-53.
7. Lee WT, St John J. The control of mitochondrial DNA replication during development and tumorigenesis. *Ann N Y Acad Sci*. 2015;1350:95-106.
8. Sun X, St John JC. The role of the mtDNA set point in differentiation, development and tumorigenesis. *Biochem J*. 2016;473(19):2955-71.
9. Liu B, Du Q, Chen L, Fu G, Li S, Fu L, et al. CpG methylation patterns of human mitochondrial DNA. *Sci Rep*. 2016;6:23421.
10. Mposhi A, Van der Wijst MG, Faber KN, Rots MG. Regulation of mitochondrial gene expression, the epigenetic enigma. *Front Biosci (Landmark Ed)*. 2017;22:1099-113.
11. Sun X, Johnson J, St John JC. Global DNA methylation synergistically regulates the nuclear and mitochondrial genomes in glioblastoma cells. *Nucleic Acids Res*. 2018;46(12):5977-95.
12. 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(9):3630-5.
13. Dzitoyeva S, Chen H, Manev H. Effect of aging on 5-hydroxymethylcytosine in brain mitochondria. *Neurobiol Aging*. 2012;33(12):2881-91.
14. Kelly RD, Mahmud A, McKenzie M, Trounce IA, St John JC. Mitochondrial DNA copy number is regulated in a tissue specific manner by DNA methylation of the nuclear-encoded DNA polymerase gamma A. *Nucleic Acids Res*. 2012;40(20):10124-38.
15. He Y, Wu J, Dressman DC, Iacobuzio-Donahue C, Markowitz SD, Velculescu VE, et al. Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. *Nature*. 2010;464(7288):610-4.
16. Wallace DC. Bioenergetics in human evolution and disease: implications for the origins of biological complexity and the missing genetic variation of common diseases. *Philos Trans R Soc Lond B Biol Sci*. 2013;368(1622):20120267.
17. Shen L, Wei J, Chen T, He J, Qu J, He X, et al. Evaluating mitochondrial DNA in patients with breast cancer and benign breast disease. *J Cancer Res Clin Oncol*. 2011;137(4):669-75.
18. Liou CW, Chen JB, Tiao MM, Weng SW, Huang TL, Chuang JH, et al. Mitochondrial DNA coding and control region variants as genetic risk factors for type 2 diabetes. *Diabetes*. 2012;61(10):2642-51.
19. Ridge PG, Maxwell TJ, Corcoran CD, Norton MC, Tschanz JT, O'Brien E, et al. Mitochondrial genomic analysis of late onset Alzheimer's disease reveals protective haplogroups H6A1A/H6A1B: the Cache County Study on Memory in Aging. *PLoS One*. 2012;7(9):e45134.
20. Ghezzi D, Marelli C, Achilli A, Goldwurm S, Pezzoli G, Barone P, et al. Mitochondrial DNA haplogroup K is associated with a lower risk of Parkinson's disease in Italians. *Eur J Hum Genet*. 2005;13(6):748-52.
21. Ruiz-Pesini E, Lapena AC, Diez-Sanchez C, Perez-Martos A, Montoya J, Alvarez E, et al. Human mtDNA haplogroups associated with high or reduced spermatozoa motility. *Am J Hum Genet*. 2000;67(3):682-96.

22. Tsai TS, Rajasekar S, St John JC. The relationship between mitochondrial DNA haplotype and the reproductive capacity of domestic pigs (*Sus scrofa domestica*). *BMC Genet.* 2016;17(1):67.
23. Kelly RD, Rodda AE, Dickinson A, Mahmud A, Nefzger CM, Lee W, et al. Mitochondrial DNA haplotypes define gene expression patterns in pluripotent and differentiating embryonic stem cells. *Stem Cells.* 2013;31(4):703-16.
24. Mehta M, Ingerslev LR, Fabre O, Picard M, Barres R. Evidence Suggesting Absence of Mitochondrial DNA Methylation. *Front Genet.* 2017;8:166.
25. van der Wijst MG, van Tilburg AY, Ruiters MH, Rots MG. Experimental mitochondria-targeted DNA methylation identifies GpC methylation, not CpG methylation, as potential regulator of mitochondrial gene expression. *Sci Rep.* 2017;7(1):177.
26. Nass MM. 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):155-75.
27. Gomez-Zaera M, Abril J, Gonzalez L, Aguilo F, Condom E, Nadal M, et al. Identification of somatic and germline mitochondrial DNA sequence variants in prostate cancer patients. *Mutat Res.* 2006;595(1-2):42-51.
28. Webb E, Broderick P, Chandler I, Lubbe S, Penegar S, Tomlinson IP, et al. Comprehensive analysis of common mitochondrial DNA variants and colorectal cancer risk. *Br J Cancer.* 2008;99(12):2088-93.
29. Yeung KY, Dickinson A, Donoghue JF, Polekhina G, White SJ, Grammatopoulos DK, et al. The identification of mitochondrial DNA variants in glioblastoma multiforme. *Acta Neuropathol Commun.* 2014;2:1.
30. Chatterjee A, Mambo E, Sidransky D. Mitochondrial DNA mutations in human cancer. *Oncogene.* 2006;25(34):4663-74.
31. Bianchessi V, Vinci MC, Nigro P, Rizzi V, Farina F, Capogrossi MC, et al. Methylation profiling by bisulfite sequencing analysis of the mtDNA Non-Coding Region in replicative and senescent Endothelial Cells. *Mitochondrion.* 2016;27:40-7.
32. Pirola CJ, Gianotti TF, Burgueno AL, Rey-Funes M, Loidl CF, Mallardi P, et al. Epigenetic modification of liver mitochondrial DNA is associated with histological severity of nonalcoholic fatty liver disease. *Gut.* 2013;62(9):1356-63.
33. 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.
34. Bellizzi D, D'Aquila P, Scafone T, Giordano M, Riso V, Riccio A, et al. The control region of mitochondrial DNA shows an unusual CpG and non-CpG methylation pattern. *DNA Res.* 2013;20(6):537-47.
35. Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell.* 2010;18(6):553-67.
36. Turcan S, Rohle D, Goenka A, Walsh LA, Fang F, Yilmaz E, et al. IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature.* 2012;483(7390):479-83.
37. Liu DC, Zheng X, Zho Y, Yi WR, Li ZH, Hu X, et al. HIF-1 α inhibits IDH-1 expression in osteosarcoma. *Oncol Rep.* 2017;38(1):336-42.
38. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics.* 2011;27(11):1571-2.
39. Skvortsova K, Zotenko E, Luu PL, Gould CM, Nair SS, Clark SJ, et al. Comprehensive evaluation of genome-wide 5-hydroxymethylcytosine profiling approaches in human DNA. *Epigenetics Chromatin.* 2017;10:16.
40. Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, et al. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat Genet.* 2005;37(8):853-62.

Supporting information

S1 Table. Read coverage from bisulfite sequencing.

S2 Table. The differences in mtDNA methylation amongst the 143B^{143B}, 143B^{GBM} and 143B^{NSC} mitochondrial genomes.

S3 Table. The full potential for DNA methylation at each CpG site in the mitochondrial genomes for each 143B tumour type determined by inducing DNA methylation on their respective long PCR samples.

S4 Table. Primer pairs for long and real time PCR and pyro-sequencing.

Fig 1. mtDNA methylation profiles determined by whole-mitochondrial genome bisulfite-sequencing. Whole-mitochondrial genome bisulfite-sequencing was applied to the purified mtDNA samples from a) 143B^{143B} early tumours, b) 143B^{143B} late tumours, c) 143B^{NSC} early tumours, d) 143B^{NSC} late tumours, e) 143B^{GBM} early tumours, f) 143B^{GBM} late tumours, and g) GBM cells. Levels of methylation (% Methylation) at CpG sites across the whole genome are shown.

Fig 2. mtDNA methylation identified by pyro-sequencing. Levels of DNA methylation at CpG sites of a) HSP, b) LSP and c) *ND6* regions of 143B cells, 143B^{143B} early and 143B^{143B} late tumours were determined by pyro-sequencing. Statistical significance was determined by One-way ANOVA. Bars represent the mean of the percentage of DNA methylation (mean \pm SEM; n=3). * and ** indicate *p* values of < 0.05 and 0.01, respectively.

Fig 3. Levels of DNA methylation at different regions of the mitochondrial genome in 143B^{143B}, 143B^{GBM} and 143B^{NSC} tumours. DNA methylation levels (5mC/5hmC) within regions of the mitochondrial genome determined on purified mtDNA samples using MeDIP-qPCR. Statistical significance was determined between the 143B^{143B}, 143B^{GBM} and 143B^{NSC} tumours by One-way ANOVA. Bars represent the mean of the relative quantification levels (mean \pm SEM; n=3). *, ** indicate *p* values of < 0.05, 0.01, respectively.

Fig 4. Levels of DNA methylation of the same mtDNA genotype under a different nuclear genome. a) Levels of DNA methylation of GBM mtDNA in association with the GBM (GBM¹⁰⁰) and 143B (143B^{143B}) nuclear genomes; b) Levels of DNA

methylation of NSC mtDNA under the NSC (hNSC) and 143B (143B^{NSC}) nuclear genomes. DNA methylation levels (5mC/5hmC) determined on purified mtDNA samples using MeDIP. Statistical significance was determined between early and late stage tumours by multiple t-tests. Bars represent the mean of the relative quantification levels (mean \pm SEM; n=3). *, ** indicate *p* values of < 0.05, 0.01, respectively.

Fig 5. Levels of DNA methylation at different regions of the mitochondrial genome in GBM¹⁰⁰, GBM⁵⁰, GBM³ and GBM^{0.2} tumours. DNA methylation levels (5mC/5hmC) within regions of the mitochondrial genome determined on purified mtDNA samples using MeDIP. Statistical significance was determined between the GBM¹⁰⁰, GBM⁵⁰, GBM³ and GBM^{0.2} tumours by One-way ANOVA. Bars represent the mean of the relative quantification levels (mean \pm SEM; n=3). *, **, *** indicate *p* values of < 0.05, 0.01, 0.001, respectively.

Figure 1

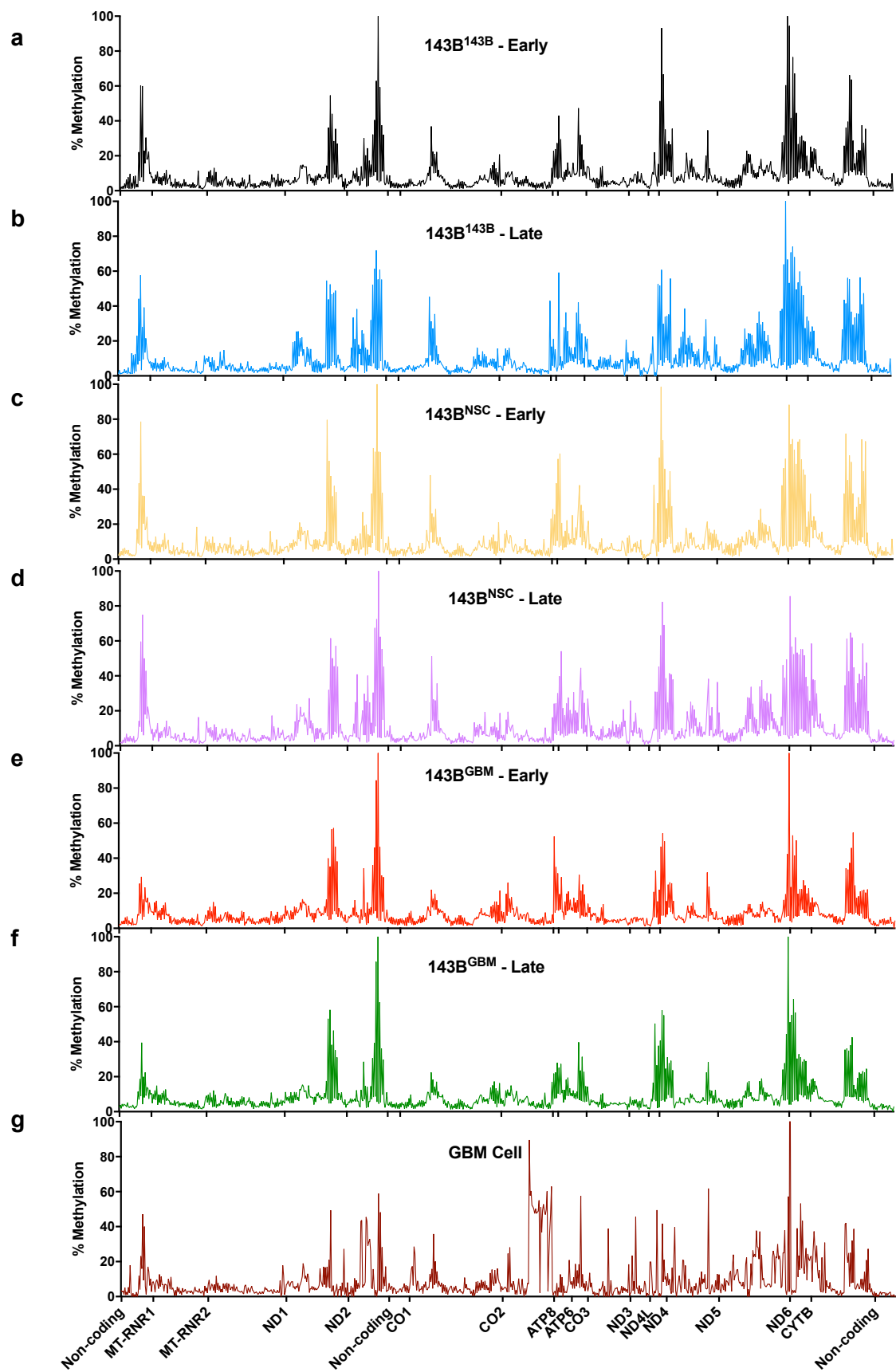


Figure 2

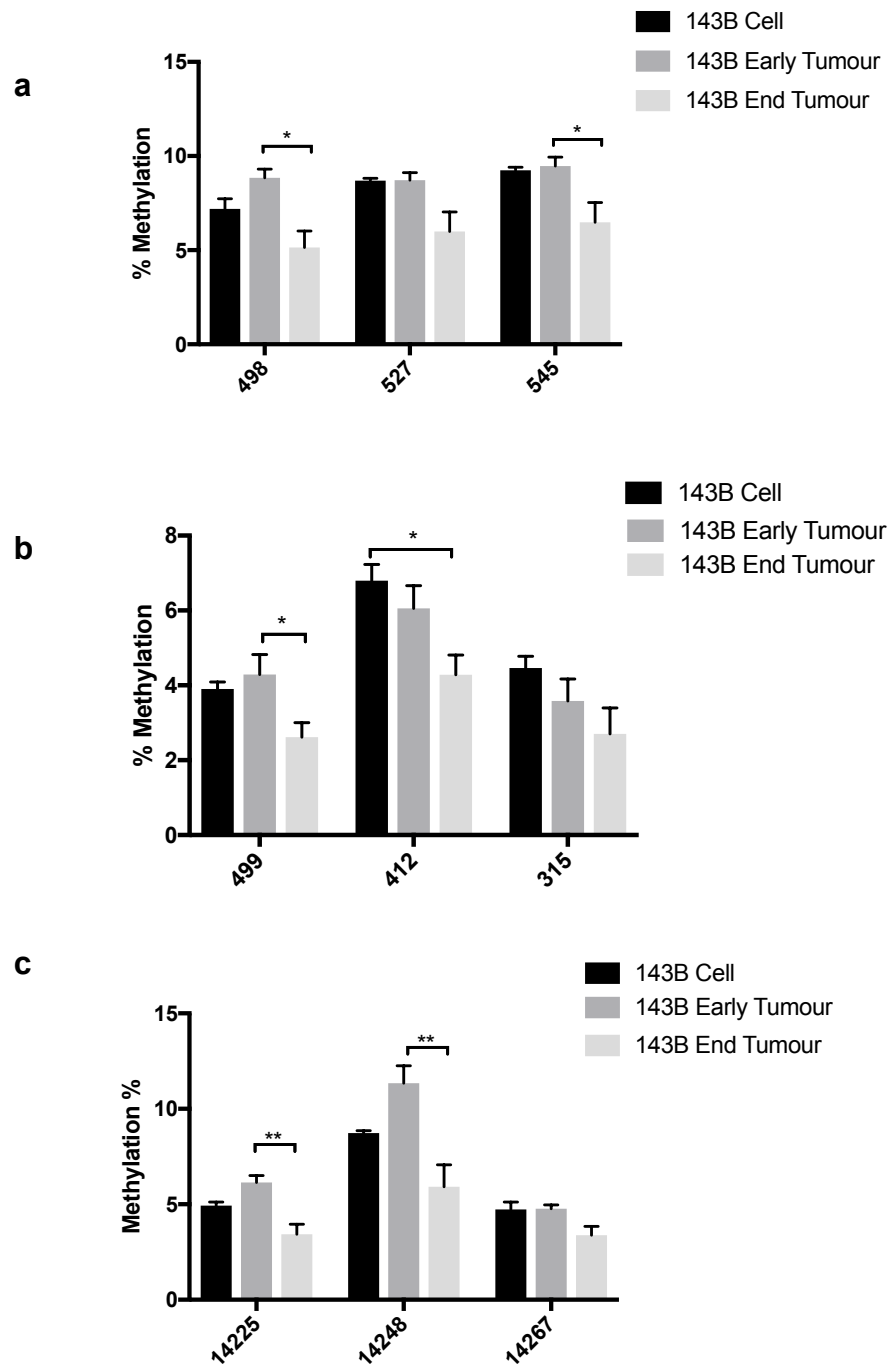


Figure 3

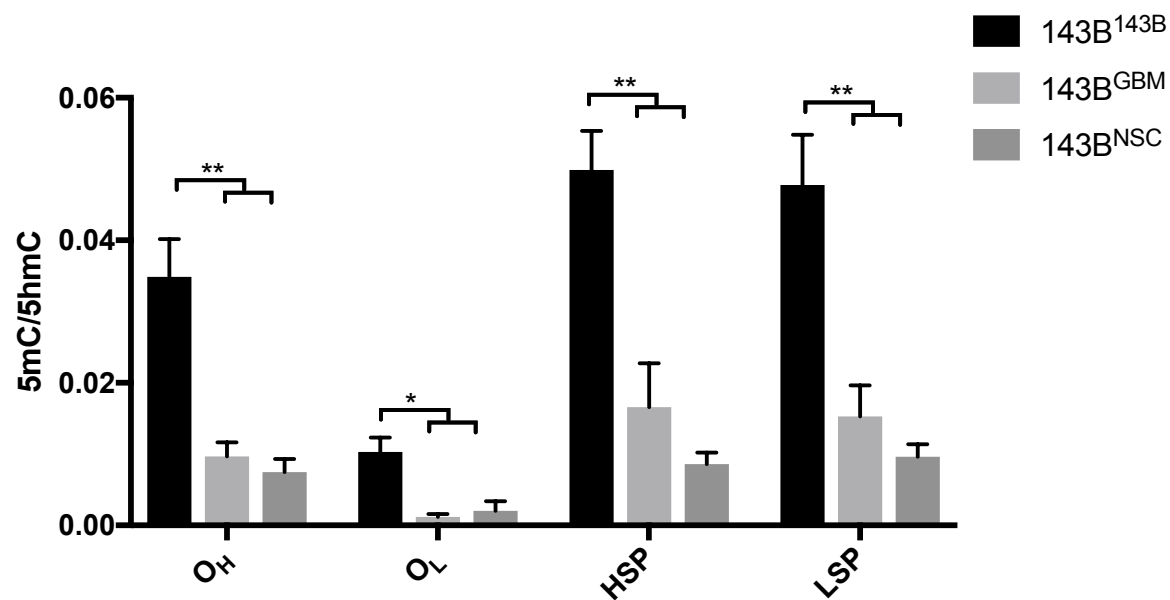


Figure 4

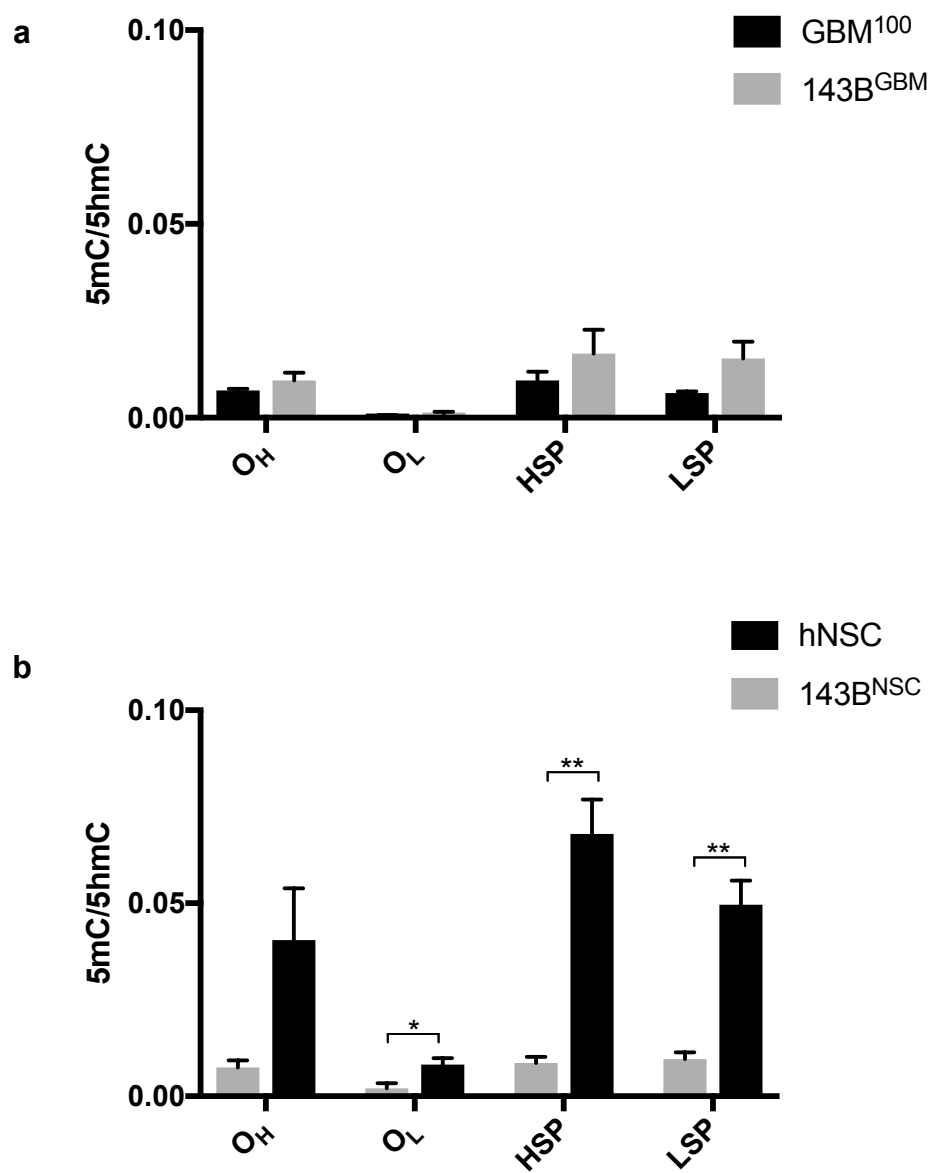
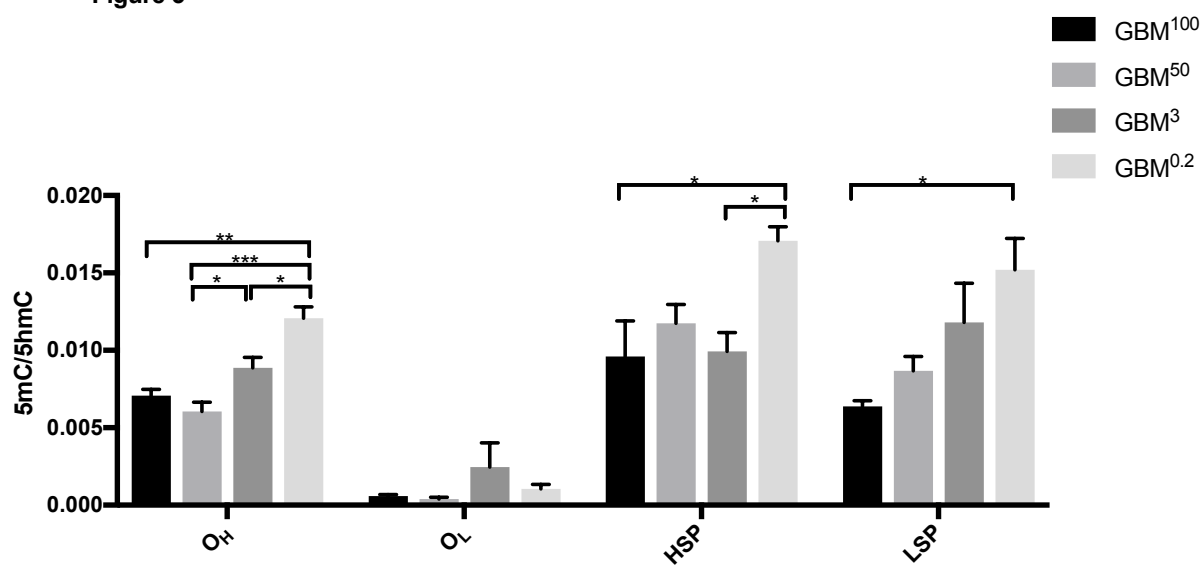


Figure 5



S1 Table. Read coverage from bisulfite sequencing.		
Sample	Uniquelymapped reads	Coverage (Fold)
143B-143B-Early1	165366	649
143B-143B-Early2	170498	669
143B-143B-Early3	176841	694
143B-GBM-Early1	192239	754
143B-GBM-Early2	158296	621
143B-GBM-Early3	171375	672
143B-NSC-Early1	253344	994
143B-NSC-Early2	150401	590
143B-NSC-Early3	141035	553
143B-143B1	95365	374
143B-143B2	196598	771
143B-143B3	127338	500
143B-GBM1	241114	946
143B-GBM2	195700	768
143B-GBM3	120463	473
143B-NSC1	268890	1055
143B-NSC2	279372	1096
143B-NSC3	213592	838

S2 Table. The differences in mtDNA methylation amongst the 143B143B, 143BGBM and 143BNSC mitochondrial genomes.

143B143B	location of C	Strand	Context	143BGBM	location of C	Strand	Context	143BNSC	location of C	Strand	Context
KT946592.1	33	+	CG	KT946593.1	33	+	CG	KT946594.1	33	+	CG
KT946592.1	34	-	CG	KT946593.1	34	-	CG	KT946594.1	34	-	CG
KT946592.1	61	+	CG	KT946593.1	61	+	CG	KT946594.1	61	+	CG
KT946592.1	62	-	CG	KT946593.1	62	-	CG	KT946594.1	62	-	CG
KT946592.1	78	+	CG	KT946593.1	78	+	CG	KT946594.1	78	+	CG
KT946592.1	79	-	CG	KT946593.1	79	-	CG	KT946594.1	79	-	CG
KT946592.1	80	+	CG	KT946593.1	80	+	CG	KT946594.1	80	+	CG
KT946592.1	81	-	CG	KT946593.1	81	-	CG	KT946594.1	81	-	CG
KT946592.1	91	+	CG	KT946593.1	91	+	CG	KT946594.1	91	+	CG
KT946592.1	92	-	CG	KT946593.1	92	-	CG	KT946594.1	92	-	CG
KT946592.1	96	+	CG	KT946593.1	96	+	CG	KT946594.1	96	+	CG
KT946592.1	97	-	CG	KT946593.1	97	-	CG	KT946594.1	97	-	CG
KT946592.1	105	+	CG	KT946593.1	105	+	CG	KT946594.1	105	+	CG
KT946592.1	106	-	CG	KT946593.1	106	-	CG	KT946594.1	106	-	CG
KT946592.1	120	+	CG	KT946593.1	120	+	CG	KT946594.1	120	+	CG
KT946592.1	121	-	CG	KT946593.1	121	-	CG	KT946594.1	121	-	CG
KT946592.1	162	+	CG	KT946593.1	162	+	CG	KT946594.1	162	+	CG
KT946592.1	163	-	CG	KT946593.1	163	-	CG	KT946594.1	163	-	CG
KT946592.1	170	+	CG	KT946593.1	170	+	CG	KT946594.1	170	+	CG
KT946592.1	171	-	CG	KT946593.1	171	-	CG	KT946594.1	171	-	CG
KT946592.1	186	+	CG	KT946593.1	186	+	CG	KT946594.1	186	+	CG
KT946592.1	187	-	CG	KT946593.1	187	-	CG	KT946594.1	187	-	CG
KT946592.1	262	+	CG	KT946593.1	262	+	CG	KT946594.1	262	+	CG
KT946592.1	263	-	CG	KT946593.1	263	-	CG	KT946594.1	263	-	CG
KT946592.1	315	+	CG	KT946593.1	315	+	CG	KT946594.1	315	+	CG
KT946592.1	316	-	CG	KT946593.1	316	-	CG	KT946594.1	316	-	CG
KT946592.1	411	+	CG	KT946593.1	411	+	CG	KT946594.1	411	+	CG
KT946592.1	412	-	CG	KT946593.1	412	-	CG	KT946594.1	412	-	CG
KT946592.1	498	+	CG	KT946593.1	498	+	CG	KT946594.1	498	+	CG
KT946592.1	499	-	CG	KT946593.1	499	-	CG	KT946594.1	499	-	CG
KT946592.1	525	+	CG	KT946593.1	525	+	CG	KT946594.1	525	+	CG
KT946592.1	526	-	CG	KT946593.1	526	-	CG	KT946594.1	526	-	CG
KT946592.1	544	+	CG	KT946593.1	544	+	CG	KT946594.1	544	+	CG
KT946592.1	545	-	CG	KT946593.1	545	-	CG	KT946594.1	545	-	CG
KT946592.1	624	+	CG	KT946593.1	624	+	CG	KT946594.1	624	+	CG
KT946592.1	625	-	CG	KT946593.1	625	-	CG	KT946594.1	625	-	CG
KT946592.1	708	+	CG	KT946593.1	708	+	CG	KT946594.1	708	+	CG
KT946592.1	709	-	CG	KT946593.1	709	-	CG	KT946594.1	709	-	CG
KT946592.1	739	+	CG	KT946593.1	739	+	CG	KT946594.1	739	+	CG
KT946592.1	740	-	CG	KT946593.1	740	-	CG	KT946594.1	740	-	CG
KT946592.1	765	+	CG	KT946593.1	765	+	CG	KT946594.1	765	+	CG
KT946592.1	766	-	CG	KT946593.1	766	-	CG	KT946594.1	766	-	CG
KT946592.1	785	+	CG	KT946593.1	785	+	CG	KT946594.1	785	+	CG
KT946592.1	786	-	CG	KT946593.1	786	-	CG	KT946594.1	786	-	CG
KT946592.1	808	+	CG	KT946593.1	808	+	CG	KT946594.1	808	+	CG
KT946592.1	809	-	CG	KT946593.1	809	-	CG	KT946594.1	809	-	CG
KT946592.1	842	+	CG	KT946593.1	842	+	CG	KT946594.1	842	+	CG
KT946592.1	843	-	CG	KT946593.1	843	-	CG	KT946594.1	843	-	CG
KT946592.1	886	+	CG	KT946593.1	886	+	CG	KT946594.1	886	+	CG
KT946592.1	887	-	CG	KT946593.1	887	-	CG	KT946594.1	887	-	CG
KT946592.1	898	+	CG	KT946593.1	898	+	CG	KT946594.1	898	+	CG
KT946592.1	899	-	CG	KT946593.1	899	-	CG	KT946594.1	899	-	CG
KT946592.1	900	+	CG	KT946593.1	900	+	CG	KT946594.1	900	+	CG
KT946592.1	901	-	CG	KT946593.1	901	-	CG	KT946594.1	901	-	CG
KT946592.1	908	+	CG	KT946593.1	908	+	CG	KT946594.1	908	+	CG
KT946592.1	909	-	CG	KT946593.1	909	-	CG	KT946594.1	909	-	CG
KT946592.1	932	+	CG	KT946593.1	932	+	CG	KT946594.1	932	+	CG
KT946592.1	933	-	CG	KT946593.1	933	-	CG	KT946594.1	933	-	CG
KT946592.1	935	+	CG	KT946593.1	935	+	CG	KT946594.1	935	+	CG
KT946592.1	936	-	CG	KT946593.1	936	-	CG	KT946594.1	936	-	CG
KT946592.1	1023	+	CG	KT946593.1	1023	+	CG	KT946594.1	1023	+	CG
KT946592.1	1024	-	CG	KT946593.1	1024	-	CG	KT946594.1	1024	-	CG
KT946592.1	1133	+	CG	KT946593.1	1133	+	CG	KT946594.1	1133	+	CG
KT946592.1	1134	-	CG	KT946593.1	1134	-	CG	KT946594.1	1134	-	CG
KT946592.1	1146	+	CG	KT946593.1	1146	+	CG	KT946594.1	1146	+	CG
KT946592.1	1147	-	CG	KT946593.1	1147	-	CG	KT946594.1	1147	-	CG
KT946592.1	1177	+	CG	KT946593.1	1177	+	CG	KT946594.1	1177	+	CG
KT946592.1	1178	-	CG	KT946593.1	1178	-	CG	KT946594.1	1178	-	CG
KT946592.1	1216	+	CG	KT946593.1	1216	+	CG	KT946594.1	1216	+	CG
KT946592.1	1217	-	CG	KT946593.1	1217	-	CG	KT946594.1	1217	-	CG
KT946592.1	1226	+	CG	KT946593.1	1226	+	CG	KT946594.1	1226	+	CG
KT946592.1	1227	-	CG	KT946593.1	1227	-	CG	KT946594.1	1227	-	CG
KT946592.1	1262	+	CG	KT946593.1	1262	+	CG	KT946594.1	1262	+	CG
KT946592.1	1263	-	CG	KT946593.1	1263	-	CG	KT946594.1	1263	-	CG
KT946592.1	1302	+	CG	KT946593.1	1302	+	CG	KT946594.1	1302	+	CG
KT946592.1	1303	-	CG	KT946593.1	1303	-	CG	KT946594.1	1303	-	CG
KT946592.1	1314	+	CG	KT946593.1	1314	+	CG	KT946594.1	1314	+	CG
KT946592.1	1315	-	CG	KT946593.1	1315	-	CG	KT946594.1	1315	-	CG
KT946592.1	1322	+	CG	KT946593.1	1322	+	CG	KT946594.1	1322	+	CG

S3 Table. The full potential for DNA methylation at each CpG site in the mitochondrial genomes for each 143B tumour type determined by inducing DNA methylation on their respective long PCR samples.

143B143B location of C	Full potential	LongPCR Counts >10	143BGBM location of C	Scale	ControlCounts >10	143BNSC location of C	Full potential	LongPCR Counts >10
33	93.25	Y	33	87.67	Y	33	91.82	Y
34	92.28	Y	34	70.63	Y	34	89.63	Y
61	80.57	Y	61	87.21	Y	61	74.79	Y
62	82.40	Y	62	62.92	Y	62	75.08	Y
78	94.86	Y	78	92.54	Y	78	92.23	Y
79	90.43	Y	79	66.78	Y	79	87.53	Y
80	92.71	Y	80	90.46	Y	80	90.16	Y
81	87.92	Y	81	64.24	Y	81	85.23	Y
91	74.29	Y	91	90.74	Y	91	68.85	Y
92	76.87	Y	92	62.33	Y	92	67.99	Y
96	82.05	Y	96	90.23	Y	96	72.53	Y
97	80.89	Y	97	64.75	Y	97	71.17	Y
105	80.84	Y	105	90.31	Y	105	75.63	Y
106	85.01	Y	106	66.10	Y	106	78.63	Y
120	88.77	Y	120	89.85	Y	120	81.04	Y
121	92.32	Y	121	74.24	Y	121	86.79	Y
162	93.94	Y	162	92.23	Y	162	89.65	Y
163	88.24	Y	163	72.32	Y	163	82.22	Y
170	89.26	Y	170	94.52	Y	170	84.99	Y
171	78.58	Y	171	71.91	Y	171	68.04	Y
186	83.80	Y	186	93.69	Y	186	81.77	Y
187	76.38	Y	187	70.23	Y	187	71.37	Y
262	68.89	Y	262	84.21	Y	262	58.40	Y
263	75.36	Y	263	67.28	Y	263	64.87	Y
315	71.66	Y	315	70.77	Y	315	63.71	Y
316	66.45	Y	316	71.88	Y	316	61.42	Y
411	72.99	N	411	90.28	N	411	55.70	N
412	76.39	Y	412	72.77	Y	412	68.38	Y
498	79.66	Y	498	89.03	Y	498	79.34	Y
499	44.65	Y	499	63.43	Y	499	31.65	Y
525	75.24	Y	525	89.68	Y	525	70.91	Y
526	64.92	Y	526	68.42	Y	526	58.34	Y
544	91.53	Y	544	88.05	Y	544	87.35	Y
545	56.62	Y	545	66.37	Y	545	51.05	Y
624	73.45	Y	624	74.91	Y	624	69.28	Y
625	90.44	Y	625	71.04	Y	625	86.93	Y
708	84.71	Y	708	49.29	Y	708	81.42	Y
709	79.59	Y	709	64.87	Y	709	69.19	Y
739	84.85	Y	739	56.19	Y	739	81.00	Y
740	92.48	Y	740	68.37	Y	740	89.91	Y
765	88.70	Y	765	62.95	Y	765	84.86	Y
766	93.70	Y	766	67.26	Y	766	91.83	Y
785	66.67	Y	785	55.63	Y	785	51.62	Y
786	75.70	Y	786	61.47	Y	786	65.91	Y
808	86.78	Y	808	60.87	Y	808	80.79	Y
809	93.52	Y	809	71.21	Y	809	91.93	Y
842	80.49	Y	842	63.58	Y	842	71.58	Y
843	82.13	Y	843	58.61	Y	843	73.25	Y
886	81.88	Y	886	69.35	Y	886	75.87	Y
887	91.91	Y	887	64.23	Y	887	89.48	Y
898	68.47	Y	898	54.08	Y	898	51.18	Y
899	81.59	Y	899	67.31	Y	899	75.44	Y
900	85.06	Y	900	62.59	Y	900	76.56	Y
901	77.44	Y	901	66.06	Y	901	64.09	Y
908	89.56	Y	908	65.76	Y	908	84.29	Y

909	94.17	Y	909	77.36	Y	909	93.63	Y
932	61.44	Y	932	57.92	Y	932	49.92	Y
933	59.01	Y	933	64.25	Y	933	49.03	Y
935	81.52	Y	935	62.50	Y	935	72.52	Y
936	93.68	Y	936	74.26	Y	936	88.84	Y
1023	87.27	Y	1023	68.36	Y	1023	84.49	Y
1024	90.28	Y	1024	78.75	Y	1024	85.91	Y
1133	88.19	Y	1133	65.09	Y	1133	82.39	Y
1134	74.45	Y	1134	73.94	Y	1134	64.12	Y
1146	85.90	Y	1146	63.31	Y	1146	78.14	Y
1147	93.96	Y	1147	81.00	Y	1147	90.95	Y
1177	54.01	Y	1177	52.80	Y	1177	39.91	Y
1178	78.34	Y	1178	72.99	Y	1178	67.18	Y
1216	82.71	Y	1216	64.46	Y	1216	75.39	Y
1217	93.23	Y	1217	79.83	Y	1217	90.79	Y
1226	87.03	Y	1226	67.76	Y	1226	82.38	Y
1227	92.23	Y	1227	75.78	Y	1227	85.68	Y
1262	78.17	Y	1262	60.59	Y	1262	68.35	Y
1263	63.13	Y	1263	66.09	Y	1263	48.05	Y
1302	84.81	Y	1302	60.67	Y	1302	77.54	Y
1303	92.61	Y	1303	76.86	Y	1303	88.70	Y
1314	88.87	Y	1314	66.58	Y	1314	86.92	Y
1315	95.91	Y	1315	77.72	Y	1315	93.18	Y
1322	70.71	Y	1322	58.83	Y	1322	59.65	Y
1323	75.77	Y	1323	70.89	Y	1323	63.58	Y
1388	87.40	Y	1388	60.76	Y	1388	82.92	Y
1389	93.72	Y	1389	76.75	Y	1389	90.59	Y
1414	72.41	Y	1414	54.58	Y	1414	61.64	Y
1415	75.38	Y	1415	65.82	Y	1415	64.51	Y
1473	88.08	Y	1473	61.01	Y	1473	82.61	Y
1474	94.64	Y	1474	75.90	Y	1474	94.24	Y
1475	91.99	Y	1475	70.23	Y	1475	90.70	Y
1476	94.14	Y	1476	72.78	Y	1476	91.26	Y
1484	45.95	Y	1484	49.85	Y	1484	34.74	Y
1485	31.60	Y	1485	53.88	Y	1485	22.69	Y
1488	89.36	Y	1488	65.10	Y	1488	85.55	Y
1489	89.00	Y	1489	68.58	Y	1489	81.22	Y
1537	91.08	Y	1537	68.32	Y	1537	89.01	Y
1538	95.69	Y	1538	79.50	Y	1538	95.15	Y
1561	87.71	Y	1561	66.23	Y	1561	81.18	Y
1562	93.47	Y	1562	74.74	Y	1562	90.00	Y
1597	85.47	Y	1597	72.49	Y	1597	75.76	Y
1598	85.20	Y	1598	70.11	Y	1598	77.41	Y
1663	70.27	Y	1663	61.08	Y	1663	57.24	Y
1664	68.69	Y	1664	66.09	Y	1664	54.36	Y
1749	85.01	Y	1749	68.21	Y	1749	74.70	Y
1750	94.52	Y	1750	82.12	Y	1750	92.77	Y
1769	87.35	Y	1769	69.60	Y	1769	80.57	Y
1770	94.60	Y	1770	80.46	Y	1770	93.25	Y
1786	89.24	Y	1786	72.55	Y	1786	84.18	Y
1787	91.73	Y	1787	78.74	Y	1787	86.66	Y
1905	89.12	N	1905	74.07	Y	1905	86.85	Y
1906	84.61	N	1906	72.60	Y	1906	76.46	Y
1915	76.50	Y	1915	63.78	Y	1915	64.68	Y
1916	90.83	Y	1916	74.64	Y	1916	86.70	Y
1948	88.95	N	1948	77.56	Y	1948	87.53	Y
1949	95.48	N	1949	81.85	Y	1949	94.32	Y

1989	85.43		Y		1989	68.67		Y		1989	76.52		Y
1990	92.61		Y		1990	79.64		Y		1990	88.83		Y
2001	74.66		Y		2001	66.46		Y		2001	61.43		Y
2002	87.87		Y		2002	76.07		Y		2002	80.41		Y
2202	88.56		Y		2202	67.66		N		2202	79.44		Y
2203	84.54		Y		2203	73.55		Y		2203	72.15		Y
2332	90.92		N		2332	67.34		Y		2332	85.90		Y
2333	96.28		N		2333	80.75		Y		2333	95.28		Y
2344	90.90		Y		2344	66.04		Y		2344	86.04		Y
2345	84.74		N		2345	70.53		Y		2345	74.63		Y
2476	84.70		N		2476	64.61		Y		2476	76.64		Y
2477	88.75		Y		2477	70.03		Y		2477	78.87		Y
2491	83.75		Y		2491	63.57		Y		2491	75.91		Y
2492	76.85		Y		2492	64.38		Y		2492	63.14		Y
2541	70.20		Y		2541	63.41		Y		2541	56.52		Y
2542	57.94		Y		2542	60.51		Y		2542	44.33		Y
2566	81.42		Y		2566	70.37		Y		2566	70.76		Y
2567	69.39		Y		2567	65.60		Y		2567	57.82		Y
2570	48.87		Y		2570	56.88		Y		2570	33.66		Y
2571	89.04		Y		2571	71.46		Y		2571	82.53		Y
2572	89.96		Y		2572	72.50		Y		2572	82.80		Y
2573	86.76		Y		2573	69.09		Y		2573	77.34		Y
2584	85.63		Y		2584	69.22		Y		2584	76.52		Y
2585	94.38		Y		2585	76.57		Y		2585	92.42		Y
2642	88.79		Y		2642	66.32		Y		2642	80.30		Y
2643	94.77		Y		2643	77.01		Y		2643	93.05		Y
2689	88.18		Y		2689	63.77		Y		2689	86.58		Y
2690	94.18		Y		2690	80.65		Y		2690	94.62		Y
2699	51.89		Y		2699	39.40		Y		2699	38.55		Y
2700	90.51		Y		2700	68.49		Y		2700	84.80		Y
2718	76.69		Y		2718	48.36		Y		2718	67.19		Y
2719	89.90		Y		2719	72.07		Y		2719	86.69		Y
2809	75.01		Y		2809	59.10		Y		2809	67.35		Y
2810	85.73		Y		2810	73.13		Y		2810	76.57		Y
2818	36.95		Y		2818	40.36		Y		2818	30.08		Y
2819	53.54		Y		2819	64.29		Y		2819	46.19		Y
2824	90.73		Y		2824	68.32		Y		2824	87.91		Y
2825	90.74		Y		2825	73.60		Y		2825	83.22		Y
2843	85.50		Y		2843	64.85		Y		2843	79.35		Y
2844	90.45		Y		2844	76.83		Y		2844	83.74		Y
2877	87.53		Y		2877	69.26		Y		2877	80.12		Y
2878	85.84		Y		2878	76.87		Y		2878	77.41		Y
2915	88.98		Y		2915	67.40		Y		2915	82.32		Y
2916	92.76		Y		2916	79.62		Y		2916	90.25		Y
2942	89.78		Y		2942	68.42		Y		2942	82.51		Y
2943	94.13		Y		2943	79.10		Y		2943	92.59		Y
2982	91.51		Y		2982	73.24		Y		2982	87.91		Y
2983	88.63		Y		2983	72.87		Y		2983	80.87		Y
2988	82.88		Y		2988	69.48		Y		2988	73.29		Y
2989	78.40		Y		2989	68.90		Y		2989	66.61		Y
3009	86.12		Y		3009	72.00		Y		3009	77.05		Y
3010	86.07		Y		3010	71.12		Y		3010	76.88		Y
3021	86.23		Y		3021	72.84		Y		3021	78.42		Y
3022	88.14		Y		3022	71.41		Y		3022	77.00		Y
3035	82.37		Y		3035	73.10		Y		3035	74.74		Y
3036	86.60		Y		3036	72.28		Y		3036	76.77		Y
3046	90.10		Y		3046	74.68		Y		3046	84.81		Y

3047	93.37	Y	3047	78.66	Y	3047	91.35	Y
3060	93.58	Y	3060	77.97	Y	3060	91.27	Y
3061	95.55	Y	3061	78.72	Y	3061	93.21	Y
3078	61.09	Y	3078	64.21	Y	3078	47.88	Y
3079	88.58	Y	3079	73.71	Y	3079	80.99	Y
3093	39.62	Y	3093	61.43	Y	3093	30.29	Y
3094	66.69	Y	3094	68.33	Y	3094	55.15	Y
3125	93.45	Y	3125	82.30	Y	3125	91.69	Y
3126	88.75	Y	3126	80.39	Y	3126	85.92	Y
3161	79.37	Y	3161	75.27	Y	3161	67.74	Y
3162	73.15	Y	3162	72.20	Y	3162	59.58	Y
3171	93.66	Y	3171	82.45	Y	3171	92.94	Y
3172	95.76	Y	3172	81.26	Y	3172	93.50	Y
3246	88.48	Y	3246	74.73	Y	3246	85.39	Y
3247	84.72	Y	3247	74.40	Y	3247	74.02	Y
3253	93.60	Y	3253	72.17	Y	3253	87.97	Y
3254	94.06	Y	3254	79.94	Y	3254	91.16	Y
3350	84.37	Y	3350	50.94	Y	3350	77.32	Y
3351	92.58	Y	3351	76.12	Y	3351	89.05	Y
3374	84.30	Y	3374	42.69	Y	3374	82.77	Y
3375	91.19	Y	3375	77.98	Y	3375	85.54	Y
3378	83.83	Y	3378	45.24	Y	3378	72.32	Y
3379	85.29	Y	3379	76.95	Y	3379	79.70	Y
3405	91.18	Y	3405	43.39	Y	3405	87.66	Y
3406	95.99	Y	3406	78.82	Y	3406	95.31	Y
3419	68.64	Y	3419	43.76	N	3419	58.62	Y
3420	73.04	Y	3420	64.04	Y	3420	59.35	Y
3434	90.34	Y	3434	59.57	N	3434	87.21	Y
3435	95.89	Y	3435	76.48	Y	3435	94.66	Y
3452	86.09	Y	3452	70.28	Y	3452	83.36	Y
3453	89.65	Y	3453	64.53	Y	3453	82.05	Y
3458	88.39	Y	3458	69.85	Y	3458	77.90	Y
3459	83.70	Y	3459	67.36	Y	3459	74.21	Y
3494	92.34	Y	3494	78.96	Y	3494	93.07	Y
3495	33.92	Y	3495	35.24	Y	3495	23.28	Y
3524	79.78	Y	3524	70.30	Y	3524	70.55	Y
3525	52.99	Y	3525	53.42	Y	3525	40.83	Y
3529	87.14	Y	3529	67.77	Y	3529	81.09	Y
3530	72.80	Y	3530	59.96	Y	3530	58.22	Y
3548	88.51	Y	3548	71.31	Y	3548	79.80	Y
3549	85.51	Y	3549	71.60	Y	3549	77.34	Y
3641	67.39	Y	3641	60.56	Y	3641	54.72	Y
3642	82.43	Y	3642	79.54	Y	3642	70.60	Y
3686	86.44	Y	3686	45.90	Y	3686	79.33	Y
3687	62.06	Y	3687	69.64	Y	3687	48.79	Y
3695	63.80	Y	3695	47.99	Y	3695	57.79	Y
3696	75.86	Y	3696	74.02	Y	3696	66.07	Y
3698	84.54	Y	3698	48.34	Y	3698	71.53	Y
3699	94.03	Y	3699	79.83	Y	3699	90.27	Y
3705	91.65	Y	3705	56.35	Y	3705	82.20	Y
3706	92.37	Y	3706	77.95	Y	3706	86.93	Y
3888	76.20	Y	3888	43.27	Y	3888	58.81	Y
3889	75.43	Y	3889	72.97	Y	3889	61.80	Y
3899	75.74	Y	3899	54.59	Y	3899	62.72	Y
3900	77.96	Y	3900	72.65	Y	3900	66.03	Y
3908	90.78	Y	3908	55.35	Y	3908	83.52	Y
3909	90.17	Y	3909	80.98	Y	3909	85.75	Y

3920	90.24	Y	3920	63.15	Y	3920	87.11	Y
3921	79.81	Y	3921	72.77	Y	3921	68.34	Y
3944	90.61	Y	3944	60.99	Y	3944	87.65	Y
3945	90.39	Y	3945	79.74	Y	3945	85.64	Y
3950	92.26	Y	3950	69.84	Y	3950	92.67	Y
3951	94.95	Y	3951	80.94	Y	3951	92.89	Y
3953	71.28	Y	3953	49.21	Y	3953	59.03	Y
3954	88.31	Y	3954	75.76	Y	3954	79.65	Y
3965	43.00	Y	3965	39.56	Y	3965	35.97	Y
3966	72.97	Y	3966	66.78	Y	3966	59.43	Y
3983	78.24	N	3983	67.46	N	3983	69.97	N
3984	91.13	Y	3984	76.54	Y	3984	86.28	Y
4049	91.11	Y	4049	80.24	N	4049	75.96	Y
4050	94.74	Y	4050	79.53	Y	4050	92.30	Y
4125	90.01	Y	4125	67.80	N	4125	85.98	Y
4126	91.64	Y	4126	79.14	Y	4126	86.57	Y
4140	93.56	Y	4140	65.57	N	4140	86.35	Y
4141	82.05	Y	4141	71.25	Y	4141	70.83	Y
4146	87.93	Y	4146	75.84	N	4146	89.90	Y
4147	79.16	Y	4147	70.52	Y	4147	68.17	Y
4151	89.79	Y	4151	71.09	Y	4151	91.46	Y
4152	94.81	Y	4152	78.74	Y	4152	92.67	Y
4344	89.78	Y	4344	75.87	Y	4344	83.31	Y
4345	79.54	Y	4345	70.88	Y	4345	66.82	Y
4374	93.71	Y	4374	80.73	Y	4374	93.35	Y
4375	95.33	Y	4375	77.77	Y	4375	94.17	Y
4425	92.80	Y	4425	69.54	Y	4425	89.66	Y
4426	94.12	Y	4426	75.90	Y	4426	91.11	Y
4438	92.41	Y	4438	76.20	Y	4438	90.57	Y
4439	80.18	Y	4439	70.63	Y	4439	70.16	Y
4462	93.61	Y	4462	76.01	Y	4462	92.42	Y
4463	95.83	Y	4463	78.38	Y	4463	94.73	Y
4489	90.75	Y	4489	73.75	Y	4489	90.22	Y
4490	91.92	Y	4490	71.57	Y	4490	84.30	Y
4530	94.77	N	4530	70.90	Y	4530	90.49	Y
4531	93.97	N	4531	75.36	Y	4531	90.09	Y
4539	88.66	Y	4539	69.26	Y	4539	85.30	Y
4540	91.26	Y	4540	73.00	Y	4540	84.68	Y
4619	93.35	Y	4619	86.78	Y	4619	89.73	Y
4620	71.33	Y	4620	65.42	Y	4620	57.21	Y
4653	97.63	Y	4653	87.72	Y	4653	92.27	Y
4654	95.17	Y	4654	79.37	Y	4654	93.97	Y
4663	89.45	Y	4663	80.47	Y	4663	86.53	Y
4664	92.27	Y	4664	76.29	Y	4664	87.72	Y
4711	96.62	Y	4711	77.40	Y	4711	92.73	Y
4712	94.80	Y	4712	81.02	Y	4712	93.66	Y
4846	91.72	Y	4846	75.03	Y	4846	88.18	Y
4847	62.81	Y	4847	65.69	Y	4847	49.15	Y
4918	92.43	Y	4918	76.81	Y	4918	84.80	Y
4919	91.65	Y	4919	77.05	Y	4919	85.01	Y
4994	91.49	Y	4994	73.32	Y	4994	93.54	Y
4995	96.09	Y	4995	81.85	Y	4995	94.79	Y
5052	92.01	Y	5052	85.84	Y	5052	92.37	Y
5053	95.47	Y	5053	81.01	Y	5053	93.21	Y
5110	85.13	Y	5110	76.04	Y	5110	87.34	Y
5111	94.74	Y	5111	79.99	Y	5111	92.29	Y
5145	88.81	Y	5145	79.62	Y	5145	84.29	Y

5146	86.10	Y	5146	74.20	Y	5146	77.56	Y
5162	89.18	Y	5162	78.51	Y	5162	86.05	Y
5163	91.88	Y	5163	77.36	Y	5163	87.08	Y
5235	80.53	Y	5235	77.19	Y	5235	83.45	Y
5236	88.63	Y	5236	75.43	Y	5236	81.79	Y
5242	72.65	Y	5242	79.13	Y	5242	54.31	Y
5243	70.90	Y	5243	72.44	Y	5243	58.25	Y
5269	89.49	N	5269	58.33	N	5269	90.94	N
5270	91.48	Y	5270	79.23	Y	5270	88.93	Y
5351	42.18	Y	5351	25.73	Y	5351	49.31	Y
5352	89.12	Y	5352	75.65	Y	5352	80.02	Y
5398	78.48	Y	5398	74.55	N	5398	86.57	Y
5399	95.20	Y	5399	81.95	Y	5399	92.93	Y
5458	89.89	Y	5458	81.06	Y	5458	79.02	Y
5459	55.48	Y	5459	61.71	Y	5459	39.00	Y
5469	93.84	Y	5469	78.84	Y	5469	90.81	Y
5470	93.52	Y	5470	74.24	Y	5470	86.69	Y
5621	84.16	Y	5621	66.54	Y	5621	71.92	Y
5622	90.89	Y	5622	78.79	Y	5622	88.99	Y
5736	88.03	N	5736	64.02	Y	5736	85.18	Y
5737	77.38	N	5737	74.25	Y	5737	64.11	Y
5739	63.14	N	5739	58.51	Y	5739	51.76	Y
5740	42.42	N	5740	65.49	Y	5740	30.25	Y
5742	87.29	N	5742	64.58	Y	5742	76.09	Y
5743	94.45	N	5743	80.91	Y	5743	89.13	Y
5754	85.62	N	5754	63.81	Y	5754	74.17	Y
5755	92.94	N	5755	77.03	Y	5755	89.15	Y
5766	92.55	Y	5766	67.31	Y	5766	88.27	Y
5767	90.66	N	5767	74.58	Y	5767	82.75	Y
5789	88.11	N	5789	70.03	Y	5789	85.57	Y
5790	91.64	N	5790	79.13	Y	5790	86.66	Y
5819	87.37	Y	5819	63.43	Y	5819	79.53	Y
5820	89.67	N	5820	75.24	Y	5820	81.51	Y
5908	78.13	Y	5908	61.27	Y	5908	66.69	Y
5909	81.46	Y	5909	72.50	Y	5909	71.34	Y
5911	75.36	Y	5911	62.15	Y	5911	65.04	Y
5912	76.44	Y	5912	71.92	Y	5912	65.70	Y
5915	52.02	Y	5915	55.89	Y	5915	42.12	Y
5916	64.79	Y	5916	68.98	Y	5916	52.86	Y
5968	87.31	Y	5968	73.82	Y	5968	81.24	Y
5969	90.88	Y	5969	73.49	Y	5969	85.26	Y
5971	94.53	Y	5971	75.66	Y	5971	91.99	Y
5972	94.58	Y	5972	77.06	Y	5972	93.82	Y
6014	86.01	Y	6014	74.88	Y	6014	78.65	Y
6015	92.76	Y	6015	75.56	Y	6015	90.47	Y
6019	75.56	Y	6019	71.23	Y	6019	63.99	Y
6020	83.09	Y	6020	69.65	Y	6020	74.17	Y
6052	82.13	Y	6052	76.27	Y	6052	72.78	Y
6053	75.16	Y	6053	68.44	Y	6053	60.62	Y
6067	89.81	Y	6067	75.88	Y	6067	80.01	Y
6068	91.13	Y	6068	73.28	Y	6068	85.29	Y
6073	95.67	Y	6073	83.91	Y	6073	94.29	Y
6074	92.18	Y	6074	71.09	Y	6074	88.04	Y
6127	92.09	Y	6127	79.73	Y	6127	87.43	Y
6128	91.61	Y	6128	73.50	Y	6128	87.88	Y
6163	83.81	Y	6163	75.00	Y	6163	74.01	Y
6164	77.35	Y	6164	64.22	Y	6164	64.81	Y

6172	88.81	Y	6172	77.67	Y	6172	87.70	Y
6173	91.31	Y	6173	71.18	Y	6173	89.25	Y
6180	63.57	Y	6180	70.85	Y	6180	51.99	Y
6181	75.87	Y	6181	66.31	Y	6181	62.58	Y
6188	92.70	Y	6188	83.90	Y	6188	92.49	Y
6189	93.88	Y	6189	74.18	Y	6189	91.96	Y
6241	91.49	Y	6241	78.76	Y	6241	90.38	Y
6242	92.67	Y	6242	74.01	Y	6242	89.30	Y
6262	69.44	Y	6262	70.90	Y	6262	61.78	Y
6263	81.57	Y	6263	68.90	Y	6263	73.01	Y
6328	95.32	Y	6328	84.94	Y	6328	94.36	Y
6329	95.08	Y	6329	78.72	Y	6329	92.95	Y
6444	61.57	Y	6444	70.50	Y	6444	48.48	Y
6445	60.64	Y	6445	67.12	Y	6445	49.41	Y
6454	86.56	Y	6454	80.05	Y	6454	83.79	Y
6455	91.22	Y	6455	76.38	Y	6455	86.54	Y
6463	91.62	Y	6463	79.44	Y	6463	89.06	Y
6464	88.99	Y	6464	75.98	Y	6464	81.59	Y
6539	61.42	Y	6539	67.59	Y	6539	47.37	Y
6540	82.27	Y	6540	73.34	Y	6540	72.80	Y
6562	84.22	Y	6562	72.03	Y	6562	74.80	Y
6563	87.51	Y	6563	73.30	Y	6563	78.84	Y
6568	92.31	Y	6568	74.17	Y	6568	87.44	Y
6569	66.26	Y	6569	66.20	Y	6569	52.22	Y
6571	90.54	Y	6571	73.35	Y	6571	84.96	Y
6572	92.53	Y	6572	78.37	Y	6572	90.17	Y
6616	86.84	Y	6616	76.04	Y	6616	76.36	Y
6617	84.73	Y	6617	70.17	Y	6617	76.75	Y
6655	87.87	Y	6655	71.87	Y	6655	80.92	Y
6656	89.02	Y	6656	72.20	Y	6656	83.47	Y
6688	92.72	Y	6688	78.48	Y	6688	92.27	Y
6689	92.59	Y	6689	79.24	Y	6689	91.20	Y
6760	93.44	Y	6760	77.24	Y	6760	93.64	Y
6761	93.23	Y	6761	76.01	Y	6761	92.66	Y
6796	93.30	Y	6796	69.98	Y	6796	86.52	Y
6797	92.59	Y	6797	69.70	Y	6797	88.38	Y
6806	94.71	Y	6806	74.72	Y	6806	93.60	Y
6807	93.24	Y	6807	75.49	Y	6807	92.90	Y
6823	90.65	Y	6823	70.18	Y	6823	87.94	Y
6824	87.40	Y	6824	62.38	Y	6824	78.49	Y
6838	92.65	Y	6838	70.97	Y	6838	90.32	Y
6839	87.94	Y	6839	63.77	Y	6839	78.01	Y
6850	67.68	Y	6850	63.42	Y	6850	53.67	Y
6851	70.31	Y	6851	62.52	Y	6851	57.57	Y
6853	83.62	Y	6853	66.87	Y	6853	71.98	Y
6854	86.44	Y	6854	66.84	Y	6854	75.83	Y
6874	91.95	Y	6874	73.13	Y	6874	90.59	Y
6875	78.19	Y	6875	60.02	Y	6875	67.56	Y
6886	95.56	Y	6886	76.72	Y	6886	95.13	Y
6887	90.23	Y	6887	72.20	Y	6887	86.96	Y
6949	87.67	Y	6949	74.25	Y	6949	79.81	Y
6950	86.53	Y	6950	69.87	Y	6950	77.83	Y
6997	93.08	Y	6997	67.76	Y	6997	90.30	Y
6998	90.65	Y	6998	59.99	Y	6998	83.62	Y
7006	95.82	Y	7006	81.04	Y	7006	95.81	Y
7007	94.12	Y	7007	75.98	Y	7007	92.53	Y
7011	96.43	Y	7011	79.96	Y	7011	96.00	Y

7012	94.76	Y	7012	76.18	Y	7012	92.55	Y
7018	95.38	Y	7018	77.24	Y	7018	93.29	Y
7019	89.26	Y	7019	68.03	Y	7019	82.36	Y
7129	96.11	Y	7129	69.89	Y	7129	92.44	Y
7130	90.99	Y	7130	77.60	Y	7130	82.91	Y
7159	91.53	Y	7159	70.57	Y	7159	87.05	Y
7160	88.46	Y	7160	74.94	Y	7160	78.62	Y
7162	83.46	Y	7162	64.47	Y	7162	74.55	Y
7163	93.55	Y	7163	80.59	Y	7163	90.85	Y
7195	85.27	Y	7195	65.99	Y	7195	76.56	Y
7196	72.67	Y	7196	69.61	Y	7196	59.41	Y
7204	93.81	Y	7204	66.03	Y	7204	88.14	Y
7205	92.66	Y	7205	81.48	Y	7205	90.65	Y
7214	94.50	Y	7214	70.07	Y	7214	88.64	Y
7215	89.63	Y	7215	74.69	Y	7215	81.01	Y
7217	60.31	Y	7217	58.45	Y	7217	41.04	Y
7218	86.59	Y	7218	73.17	Y	7218	76.51	Y
7224	93.22	Y	7224	71.76	Y	7224	87.69	Y
7225	89.25	Y	7225	70.93	Y	7225	81.62	Y
7234	92.58	Y	7234	67.77	Y	7234	87.17	Y
7235	90.64	Y	7235	73.24	Y	7235	84.32	Y
7330	85.25	Y	7330	64.56	Y	7330	76.88	Y
7331	86.93	Y	7331	70.69	Y	7331	76.82	Y
7335	88.94	Y	7335	70.62	Y	7335	84.73	Y
7336	89.38	Y	7336	76.28	Y	7336	83.60	Y
7340	87.50	Y	7340	64.09	Y	7340	80.18	Y
7341	82.98	Y	7341	73.86	Y	7341	73.84	Y
7417	87.67	Y	7417	72.17	Y	7417	86.98	Y
7418	87.08	Y	7418	77.11	Y	7418	85.30	Y
7426	93.85	Y	7426	76.50	Y	7426	95.03	Y
7427	93.45	Y	7427	75.50	Y	7427	92.17	Y
7461	81.27	Y	7461	56.82	Y	7461	68.50	Y
7462	64.16	Y	7462	59.78	Y	7462	50.88	Y
7598	87.83	Y	7598	59.12	Y	7598	77.70	Y
7599	91.93	Y	7599	73.69	Y	7599	87.29	Y
7617	86.41	Y	7617	62.79	Y	7617	77.31	Y
7618	86.67	Y	7618	70.91	Y	7618	77.33	Y
7662	91.37	Y	7662	62.04	Y	7662	83.23	Y
7663	79.96	Y	7663	68.51	Y	7663	67.46	Y
7755	93.34	Y	7755	47.31	Y	7755	79.65	Y
7756	91.94	Y	7756	77.36	Y	7756	85.17	Y
7773	81.76	Y	7773	33.55	Y	7773	62.27	Y
7774	68.57	Y	7774	69.80	Y	7774	53.67	Y
7791	92.76	Y	7791	44.81	Y	7791	91.25	Y
7792	81.27	Y	7792	68.52	Y	7792	68.72	Y
7812	73.19	N	7812	37.10	Y	7812	58.21	Y
7813	65.63	N	7813	63.64	Y	7813	51.60	Y
7828	92.39	N	7828	49.73	Y	7828	91.63	Y
7829	94.00	N	7829	77.80	Y	7829	92.15	Y
7848	80.98	N	7848	39.00	Y	7848	74.93	Y
7849	90.58	N	7849	74.61	Y	7849	86.99	Y
7857	66.52	Y	7857	28.81	Y	7857	50.92	Y
7858	86.43	Y	7858	71.05	Y	7858	78.45	Y
7908	92.44	Y	7908	40.58	Y	7908	86.51	Y
7909	92.73	Y	7909	78.04	Y	7909	90.04	Y
7917	76.49	Y	7917	34.17	Y	7917	64.95	Y
7918	84.14	Y	7918	71.99	Y	7918	73.96	Y

7923	84.83		Y	7923	33.82		Y	7923	84.48		Y
7924	91.37		Y	7924	75.08		Y	7924	85.01		Y
7926	63.93		Y	7926	27.34		Y	7926	53.18		Y
7927	92.17		Y	7927	76.37		Y	7927	88.61		Y
7977	55.85		Y	7977	33.92		Y	7977	40.90		Y
7978	58.57		Y	7978	62.21		Y	7978	45.91		Y
7984	84.41		Y	7984	38.66		Y	7984	76.93		Y
7985	79.54		Y	7985	65.53		Y	7985	66.20		Y
7995	89.87		Y	7995	56.82		Y	7995	76.30		Y
7996	86.87		Y	7996	70.49		Y	7996	78.07		Y
8005	83.78		Y	8005	66.24		Y	8005	71.52		Y
8006	86.05		Y	8006	69.85		Y	8006	78.31		Y
8018	92.97		Y	8018	72.65		Y	8018	88.07		Y
8019	90.59		Y	8019	70.94		Y	8019	85.14		Y
8035	90.58		N	8035	74.92		Y	8035	87.69		Y
8036	93.21		Y	8036	75.31		Y	8036	91.26		Y
8058	77.70		Y	8058	69.86		Y	8058	64.96		Y
8059	73.72		Y	8059	63.87		Y	8059	60.29		Y
8112	91.94		Y	8112	74.57		Y	8112	90.03		Y
8113	88.98		Y	8113	69.51		Y	8113	82.69		Y
8116	66.58		Y	8116	64.50		Y	8116	51.88		Y
8117	77.55		Y	8117	70.65		Y	8117	68.61		Y
8139	92.74		Y	8139	71.88		Y	8139	85.39		Y
8140	89.28		Y	8140	71.29		Y	8140	81.40		Y
8146	92.28		Y	8146	71.14		Y	8146	86.37		Y
8147	86.98		Y	8147	70.33		Y	8147	77.15		Y
8150	73.35		Y	8150	65.12		Y	8150	65.43		Y
8151	92.47		Y	8151	76.86		Y	8151	89.34		Y
8163	88.19		Y	8163	74.34		Y	8163	79.67		Y
8164	80.54		Y	8164	62.80		Y	8164	68.53		Y
8211	88.02		Y	8211	74.94		Y	8211	83.29		Y
8212	65.05		Y	8212	54.53		Y	8212	51.18		Y
8253	94.66		Y	8253	81.81		Y	8253	94.10		Y
8254	91.56		Y	8254	67.46		Y	8254	87.06		Y
8385	90.43		N	8385	65.24		N	8385	88.52		N
8386	93.64		Y	8386	73.68		Y	8386	90.69		Y
8531	83.35		Y	8531	45.31		N	8531	85.80		Y
8532	84.79		Y	8532	66.41		Y	8532	79.24		Y
8543	81.94		Y	8543	55.75		N	8543	77.00		Y
8544	87.71		Y	8544	61.88		Y	8544	83.61		Y
8579	94.75		Y	8579	72.79		N	8579	92.82		Y
8580	54.23		Y	8580	39.49		Y	8580	39.22		Y
8582	77.44		Y	8582	79.86		N	8582	70.50		Y
8583	84.91		Y	8583	55.90		Y	8583	78.20		Y
8646	96.59		N	8646	78.65		N	8646	82.48		N
8647	87.38		Y	8647	65.23		Y	8647	79.38		Y
8721	94.03		Y	8721	75.42		Y	8721	85.84		Y
8722	72.12		Y	8722	59.25		Y	8722	60.10		Y
8780	79.48		Y	8780	48.04		Y	8780	74.11		Y
8781	87.69		Y	8781	70.07		Y	8781	80.72		Y
8854	80.50		Y	8854	60.82		Y	8854	76.43		Y
8855	91.22		Y	8855	73.29		Y	8855	88.52		Y
8858	58.39		Y	8858	52.30		Y	8858	55.55		Y
8859	85.30		Y	8859	70.19		Y	8859	79.81		Y
8877	77.10		Y	8877	54.28		Y	8877	68.09		Y
8878	84.90		Y	8878	70.91		Y	8878	77.62		Y
8957	89.02		Y	8957	61.94		Y	8957	87.40		Y

8958	8232	Y	8958	7403	Y	8958	7728	Y
8996	5569	Y	8996	5431	Y	8996	4984	Y
8997	8460	Y	8997	8858	Y	8997	7826	Y
9000	8729	Y	9000	6182	Y	9000	8554	Y
9001	9060	Y	9001	8994	Y	9001	8648	Y
9008	7781	Y	9008	6100	Y	9008	7408	Y
9009	7530	Y	9009	8956	Y	9009	6797	Y
9053	6402	Y	9053	6011	Y			
9054	7558	Y	9054	8922	Y			
9137	7356	Y	9137	7476	Y	9137	6102	Y
9138	6296	Y	9138	8725	Y	9138	5563	Y
9143	7678	Y	9143	6294	Y	9143	7203	Y
9144	8688	Y	9144	9171	Y	9144	8040	Y
9161	8989	Y	9161	7578	Y	9161	7842	Y
9162	7334	Y	9162	8941	Y	9162	6577	Y
9194	9493	Y	9194	7450	Y	9194	8990	Y
9195	9532	Y	9195	9343	Y	9195	9469	Y
9292	8211	Y	9292	7013	Y	9292	7491	Y
9293	5520	Y	9293	8151	Y	9293	4679	Y
9327	7240	Y	9327	7031	Y	9327	6437	Y
9328	8638	Y	9328	9160	Y	9328	8189	Y
9380	6602	Y	9380	6449	Y	9380	5539	Y
9381	9343	Y	9381	9126	Y	9381	9235	Y
9382	8628	Y	9382	6684	Y	9382	7833	Y
9383	8589	Y	9383	9008	Y	9383	8034	Y
9392	9249	Y	9392	6725	Y	9392	8729	Y
9393	9386	Y	9393	9304	Y	9393	9238	Y
9443	8371	Y	9443	7075	Y	9443	7550	Y
9444	8907	Y	9444	9077	Y	9444	8485	Y
9448	9112	Y	9448	7212	Y	9448	8792	Y
9449	9564	Y	9449	9271	Y	9449	9473	Y
9487	8302	Y	9487	7454	Y	9487	7480	Y
9488	9357	Y	9488	9285	Y	9488	9130	Y
9573	8279	Y	9573	6412	Y	9573	7640	Y
9574	8717	Y	9574	9075	Y	9574	7801	Y
9610	9181	Y	9610	6296	Y	9610	8497	Y
9611	9522	Y	9611	9388	Y	9611	9443	Y
9619	8744	Y	9619	6536	Y	9619	8551	Y
9620	9135	Y	9620	9102	Y	9620	8789	Y
9671	8534	Y	9671	6667	Y	9671	8135	Y
9672	7263	Y	9672	8948	Y	9672	6651	Y
9751	6801	Y	9751	6582	Y	9751	5546	Y
9752	9189	Y	9752	9242	Y	9752	8801	Y
9772	8449	Y	9772	7174	Y	9772	7779	Y
9773	8737	Y	9773	9198	Y	9773	8288	Y
9775	6029	Y	9775	6857	Y	9775	5197	Y
9776	8568	Y	9776	9178	Y	9776	7849	Y
9784	8766	Y	9784	7270	Y	9784	8251	Y
9785	9118	Y	9785	9155	Y	9785	8568	Y
9817	9154	Y	9817	7538	Y	9817	8552	Y
9818	9280	Y	9818	9186	Y	9818	8890	Y
9826	9342	Y	9826	7570	Y	9826	8773	Y
9827	9100	Y	9827	9213	Y	9827	8531	Y
9866	9255	Y	9866	7695	Y	9866	8709	Y
9867	8746	Y	9867	9139	Y	9867	8206	Y
9910	8291	Y	9910	7865	Y	9910	7393	Y
9911	8491	Y	9911	9041	Y	9911	7869	Y

9916	54.50	Y	9916	81.57	Y	9916	44.79	Y
9917	44.10	Y	9917	81.59	Y	9917	38.89	Y
9919	66.73	Y	9919	71.09	Y	9919	58.66	Y
9920	63.10	Y	9920	85.69	Y	9920	57.10	Y
10012	74.96	Y	10012	72.17	Y	10012	76.24	Y
10013	77.41	Y	10013	89.02	Y	10013	71.98	Y
10066	76.34	Y	10066	74.09	Y	10066	76.88	Y
10067	77.73	Y	10067	89.21	Y	10067	69.98	Y
10141	79.74	Y	10141	71.37	Y	10141	69.73	Y
10142	70.01	Y	10142	87.77	Y	10142	61.66	Y
10168	87.65	Y	10168	68.28	Y	10168	81.06	Y
10169	94.82	Y	10169	93.93	Y	10169	93.63	Y
10174	84.88	Y	10174	81.79	Y	10174	80.75	Y
10175	80.28	Y	10175	90.48	Y	10175	75.41	Y
10180	61.58	Y	10180	70.83	Y	10180	48.55	Y
10181	72.31	Y	10181	88.17	Y	10181	66.41	Y
10195	88.39	Y	10195	76.96	Y	10195	85.90	Y
10196	51.64	Y	10196	84.56	Y	10196	47.87	Y
10199	85.58	Y	10199	75.33	Y	10199	78.45	Y
10200	92.65	Y	10200	92.14	Y	10200	90.65	Y
10201	80.97	Y	10201	75.28	Y	10201	80.07	Y
10202	55.17	Y	10202	83.60	Y	10202	48.46	Y
10399	84.73	Y	10399	75.16	Y	10399	80.23	Y
10400	70.87	Y	10400	91.25	Y	10400	73.05	Y
10425	81.17	Y	10425	83.84	Y	10425	84.22	Y
10426	90.63	Y	10426	94.16	Y	10426	84.38	Y
10435	77.51	Y	10435	81.50	Y	10435	79.09	Y
10436	92.46	Y	10436	94.29	Y	10436	89.90	Y
10535	91.37	Y	10535	84.06	Y	10535	85.28	Y
10536	85.89	Y	10536	90.49	Y	10536	77.68	Y
10584	92.27	Y	10584	83.68	Y	10584	88.54	Y
10585	90.94	Y	10585	91.32	Y	10585	85.94	Y
10669	93.07	Y	10669	78.33	Y	10669	84.19	Y
10670	70.02	Y	10670	88.66	Y	10670	62.18	Y
10675	94.32	Y	10675	80.11	Y	10675	86.48	Y
10676	75.13	Y	10676	88.89	Y	10676	67.81	Y
10683	70.36	Y	10683	67.40	Y	10683	62.48	Y
10684	60.16	Y	10684	85.35	Y	10684	53.70	Y
10735	90.08	N	10735	84.26	N	10735	85.50	N
10736	94.24	Y	10736	92.01	Y	10736	91.05	Y
10773	98.15	N	10773	88.85	N	10773	69.39	N
10774	58.20	Y	10774	84.24	Y	10774	50.82	Y
10932	97.78	N	10932	87.84	N	10932	84.56	N
10933	80.27	Y	10933	89.79	Y	10933	70.96	Y
11002	68.82	Y	11002	87.30	N	11002	59.71	Y
11003	68.02	Y	11003	88.14	Y	11003	60.79	Y
11029	90.48	Y	11029	87.93	Y	11029	88.06	Y
11030	90.35	Y	11030	93.41	Y	11030	88.28	Y
11124	87.36	Y	11124	87.07	Y	11124	81.86	Y
11125	88.95	Y	11125	93.00	Y	11125	83.03	Y
11161	91.06	Y	11161	83.40	Y	11161	87.05	Y
11162	93.13	Y	11162	92.63	Y	11162	88.89	Y
11182	69.01	Y	11182	75.16	Y	11182	61.09	Y
11183	74.83	Y	11183	90.13	Y	11183	68.60	Y
11190	60.09	Y	11190	72.39	Y	11190	49.12	Y
11191	91.35	Y	11191	91.72	Y	11191	87.23	Y
11244	84.25	Y	11244	87.42	N	11244	89.68	Y

11245	92.02	Y	11245	91.75	Y	11245	87.84	Y
11388	88.46	Y	11388	83.20	Y	11388	82.56	Y
11389	94.43	Y	11389	92.32	Y	11389	91.71	Y
11421	82.87	Y	11421	78.08	Y	11421	77.00	Y
11422	86.18	Y	11422	90.95	Y	11422	81.25	Y
11433	82.37	Y	11433	83.14	Y	11433	73.88	Y
11434	72.40	Y	11434	86.70	Y	11434	63.98	Y
11454	83.98	Y	11454	84.02	Y	11454	79.38	Y
11455	90.83	Y	11455	91.62	Y	11455	86.26	Y
11475	65.74	Y	11475	72.60	Y	11475	55.29	Y
11476	73.79	Y	11476	88.61	Y	11476	73.62	Y
11491	94.52	Y	11491	84.49	Y	11491	92.20	Y
11492	91.97	Y	11492	91.42	Y	11492	86.97	Y
11590	95.16	Y	11590	87.31	Y	11590	90.55	Y
11591	95.06	Y	11591	93.88	Y	11591	92.86	Y
11609	58.48	Y	11609	84.33	Y	11609	55.12	Y
11610	57.97	Y	11610	87.69	Y	11610	53.16	Y
11646	91.82	Y	11646	89.64	Y	11646	90.20	Y
11647	93.06	Y	11647	92.22	Y	11647	89.45	Y
11688	77.55	Y	11688	82.35	Y	11688	71.18	Y
11689	70.87	Y	11689	88.64	Y	11689	63.27	Y
11691	75.37	Y	11691	78.76	Y	11691	65.18	Y
11692	94.11	Y	11692	93.17	Y	11692	92.03	Y
11709	82.18	Y	11709	85.05	Y	11709	78.09	Y
11710	54.93	Y	11710	84.95	Y	11710	47.53	Y
11715	92.60	Y	11715	85.08	Y	11715	84.12	Y
11716	92.46	Y	11716	93.10	Y	11716	87.89	Y
11760	93.66	Y	11760	87.44	Y	11760	89.07	Y
11761	87.09	Y	11761	90.56	Y	11761	81.08	Y
11764	83.97	Y	11764	84.05	Y	11764	74.77	Y
11765	92.77	Y	11765	91.96	Y	11765	88.49	Y
11776	91.59	Y	11776	84.73	Y	11776	88.10	Y
11777	94.92	Y	11777	92.05	Y	11777	92.42	Y
11850	88.90	Y	11850	85.46	Y	11850	83.06	Y
11851	72.79	Y	11851	88.13	Y	11851	64.89	Y
11859	74.42	Y	11859	84.13	Y	11859	68.96	Y
11860	55.06	Y	11860	84.98	Y	11860	45.64	Y
11912	83.33	Y	11912	83.03	N			
11913	72.48	Y	11913	88.08	Y			
12052	88.15	Y	12052	88.61	Y	12052	92.04	Y
12053	94.43	Y	12053	93.64	Y	12053	93.59	Y
12111	94.62	Y	12111	87.45	Y	12111	93.32	Y
12112	86.73	Y	12112	90.17	Y	12112	81.77	Y
12123	82.38	Y	12123	85.92	Y	12123	73.90	Y
12124	94.58	Y	12124	93.06	Y	12124	92.47	Y
12190	75.58	Y	12190	84.17	Y	12190	65.43	Y
12191	84.50	Y	12191	90.25	Y	12191	79.33	Y
12205	87.89	Y	12205	90.64	Y	12205	80.20	Y
12206	88.88	Y	12206	92.62	Y	12206	84.21	Y
12404	91.30	Y	12404	87.97	Y	12404	86.91	Y
12405	80.51	Y	12405	89.39	Y	12405	71.62	Y
12455	92.38	Y	12455	88.54	Y	12455	90.15	Y
12456	95.26	Y	12456	93.54	Y	12456	94.17	Y
12526	88.44	Y	12526	86.66	Y	12526	86.59	Y
12527	81.71	N	12527	89.48	Y	12527	71.87	Y
12620	83.99	Y	12620	89.07	Y	12620	78.69	Y
12621	89.58	Y	12621	90.48	Y	12621	84.18	Y

12734	90.55	Y	12734	89.22	Y	12734	87.16	Y
12735	80.76	Y	12735	88.70	Y	12735	75.59	Y
12761	92.13	Y	12761	87.70	Y	12761	87.00	Y
12762	85.52	Y	12762	89.27	Y	12762	79.36	Y
12773	68.71	Y	12773	83.28	Y	12773	60.38	Y
12774	79.12	Y	12774	87.99	Y	12774	73.72	Y
12812	92.28	Y	12812	87.46	Y	12812	87.83	Y
12813	85.91	Y	12813	88.62	Y	12813	80.15	Y
12816	88.04	Y	12816	82.33	Y	12816	83.58	Y
12817	89.72	Y	12817	89.34	Y	12817	84.79	Y
12861	90.81	Y	12861	85.52	Y	12861	83.82	Y
12862	92.11	Y	12862	90.80	Y	12862	88.21	Y
12866	85.31	Y	12866	83.76	Y	12866	79.66	Y
12867	90.52	Y	12867	89.66	Y	12867	86.93	Y
12869	70.40	Y	12869	82.79	Y	12869	65.15	Y
12870	86.06	Y	12870	88.69	Y	12870	81.64	Y
12875	69.19	Y	12875	81.04	Y	12875	62.89	Y
12876	72.58	Y	12876	86.89	Y	12876	64.22	Y
12887	87.24	Y	12887	85.76	Y	12887	80.84	Y
12888	69.89	Y	12888	86.30	Y	12888	62.01	Y
12950	71.73	Y	12950	86.02	Y	12950	58.63	Y
12951	88.27	Y	12951	91.04	Y	12951	70.03	Y
13119	91.42	Y	13119	88.70	Y	13119	85.56	Y
13120	88.46	Y	13120	91.37	Y	13120	83.98	Y
13178	86.68	Y	13178	90.93	Y	13178	82.83	Y
13179	91.46	Y	13179	91.54	Y	13179	87.84	Y
13196	84.49	Y	13196	88.64	Y	13196	77.50	Y
13197	93.18	Y	13197	91.81	Y	13197	90.46	Y
13208	87.09	Y	13208	88.28	Y	13208	86.83	Y
13209	52.20	Y	13209	82.85	Y	13209	44.44	Y
13238	71.20	Y	13238	88.35	Y	13238	67.17	Y
13239	77.14	Y	13239	88.88	Y	13239	70.29	Y
13286	82.12	Y	13286	88.13	Y	13286	74.31	Y
13287	85.13	Y	13287	89.58	Y	13287	78.52	Y
13331	82.35	Y	13331	88.33	Y	13331	78.71	Y
13332	80.00	Y	13332	89.03	Y	13332	71.62	Y
13364	90.13	Y	13364	89.30	Y	13364	87.32	Y
13365	95.19	Y	13365	92.91	Y	13365	93.95	Y
13404	89.86	Y	13404	89.68	Y	13404	85.84	Y
13405	88.08	Y	13405	92.11	Y	13405	82.08	Y
13523	90.18	Y	13523	93.05	Y	13523	88.36	Y
13524	81.19	Y	13524	90.69	Y	13524	74.70	Y
13529	79.97	Y	13529	92.68	Y	13529	74.61	Y
13530	83.99	Y	13530	91.46	Y	13530	78.83	Y
13550	83.32	Y	13550	92.67	Y	13550	78.26	Y
13551	78.04	Y	13551	89.19	Y	13551	71.74	Y
13577	83.63	Y	13577	89.58	Y	13577	77.28	Y
13578	80.11	Y	13578	88.21	Y	13578	72.11	Y
13595	85.92	Y	13595	90.52	Y	13595	83.21	Y
13596	70.78	Y	13596	86.79	Y	13596	63.04	Y
13608	91.52	Y	13608	91.43	Y	13608	90.00	Y
13609	88.60	Y	13609	90.18	Y	13609	82.09	Y
13641	76.87	Y	13641	88.18	Y	13641	71.65	Y
13642	65.43	Y	13642	85.07	Y	13642	56.40	Y
13667	92.82	Y	13667	93.87	Y	13667	90.86	Y
13668	87.88	Y	13668	90.41	Y	13668	81.93	Y
13701	70.87	Y	13701	87.60	Y	13701	58.97	Y

13702	67.40	Y	13702	85.94	Y	13702	57.91	Y
13712	77.58	Y	13712	77.72	Y	13712	72.28	Y
13713	88.44	Y	13713	89.20	Y	13713	82.59	Y
13724	82.43	Y	13724	85.79	Y	13724	70.63	Y
13725	94.88	Y	13725	91.77	Y	13725	91.93	Y
13757	93.73	Y	13757	88.88	Y	13757	90.59	Y
13758	93.39	Y	13758	89.98	Y	13758	90.49	Y
13808	88.94	Y	13808	83.03	Y	13808	83.65	Y
13809	75.07	Y	13809	86.69	Y	13809	67.00	Y
13913	93.04	Y	13913	92.40	Y	13913	82.24	Y
13914	87.71	Y	13914	90.38	Y	13914	83.02	Y
13938	82.13	Y	13938	89.08	Y	13938	74.09	Y
13939	88.06	Y	13939	89.06	Y	13939	83.39	Y
13966	88.29	Y	13966	89.10	Y	13966	88.64	Y
13967	88.05	Y	13967	89.09	Y	13967	90.03	Y
14158	8.33	N	14158	18.43	N	14158	0.00	N
14159	92.33	Y	14159	89.67	Y	14159	91.26	Y
14224	66.67	N	14224	94.44	N	14224	68.25	N
14225	56.86	Y	14225	82.90	Y	14225	49.03	Y
14247	83.33	N	14247	87.96	N	14247	100.00	N
14248	93.49	Y	14248	89.43	Y	14248	89.58	Y
14266	88.89	N	14266	90.48	N	14266	94.44	N
14267	88.33	Y	14267	88.88	Y	14267	82.29	Y
14382	85.23	N	14382	88.68	N	14382	88.60	Y
14383	81.11	Y	14383	86.27	Y	14383	76.69	Y
14457	89.86	Y	14457	94.71	Y	14457	88.68	Y
14458	82.25	Y	14458	86.55	Y	14458	72.43	Y
14558	92.06	Y	14558	94.00	Y	14558	93.86	Y
14559	82.67	Y	14559	85.79	Y	14559	75.86	Y
14567	91.58	Y	14567	93.17	Y	14567	82.14	Y
14568	84.30	Y	14568	87.48	Y	14568	78.05	Y
14679	88.26	Y	14679	93.46	N	14679	89.11	Y
14680	90.93	Y	14680	90.80	Y	14680	86.32	Y
14683	98.04	Y	14683	93.57	N	14683	91.63	Y
14684	92.85	Y	14684	91.41	Y	14684	91.26	Y
14696	93.53	Y	14696	85.56	N	14696	82.62	Y
14697	90.81	Y	14697	88.34	Y	14697	85.40	Y
14719	95.61	Y	14719	93.34	N	14719	94.71	Y
14720	75.04	Y	14720	82.97	Y	14720	66.85	Y
14758	95.37	Y	14758	95.97	Y	14758	94.37	Y
14759	95.10	Y	14759	90.62	Y	14759	94.60	Y
14802	79.07	Y	14802	89.91	Y	14802	67.67	Y
14803	61.50	Y	14803	79.07	Y	14803	45.67	Y
14829	92.86	Y	14829	93.96	Y	14829	89.23	Y
14830	94.92	Y	14830	89.85	Y	14830	93.59	Y
14844	78.54	Y	14844	91.59	Y	14844	68.39	Y
14845	80.13	Y	14845	85.22	Y	14845	73.14	Y
14859	70.56	Y	14859	89.70	Y	14859	64.90	Y
14860	86.53	Y	14860	87.65	Y	14860	80.30	Y
14919	81.53	Y	14919	90.30	Y	14919	76.34	Y
14920	66.93	Y	14920	86.71	Y	14920	60.52	Y
14928	67.25	Y	14928	86.89	Y	14928	58.49	Y
14929	67.10	Y	14929	86.39	Y	14929	57.77	Y
14943	87.36	Y	14943	90.02	Y	14943	82.74	Y
14944	53.11	Y	14944	79.13	Y	14944	44.59	Y
14956	87.45	Y	14956	89.56	Y	14956	78.07	Y
14957	81.21	Y	14957	84.25	Y	14957	72.89	Y

14961	84.51	Y	14961	90.96	Y	14961	78.66	Y
14962	93.78	Y	14962	88.20	Y	14962	90.60	Y
14983	90.04	Y	14983	90.60	Y	14983	88.76	Y
14984	89.22	Y	14984	86.71	Y	14984	84.01	Y
14994	93.21	Y	14994	92.12	Y	14994	90.17	Y
14995	90.71	Y	14995	89.25	Y	14995	86.31	Y
15003	83.54	Y	15003	90.60	Y	15003	77.35	Y
15004	84.91	Y	15004	85.20	Y	15004	77.93	Y
15039	86.79	Y	15039	90.24	Y	15039	82.26	Y
15040	91.44	Y	15040	85.39	Y	15040	86.14	Y
15043	35.73	Y	15043	76.41	Y	15043	32.62	Y
15044	76.47	Y	15044	81.81	Y	15044	71.85	Y
15057	95.00	Y	15057	93.30	Y	15057	93.38	Y
15058	94.53	Y	15058	86.68	Y	15058	92.47	Y
15090	79.96	Y	15090	88.29	Y	15090	73.50	Y
15091	79.35	Y	15091	83.06	Y	15091	72.24	Y
15146	93.50	Y	15146	90.42	Y	15146	91.43	Y
15147	93.27	Y	15147	86.71	Y	15147	89.85	Y
15198	86.16	Y	15198	90.27	Y	15198	80.22	Y
15199	78.58	Y	15199	84.00	Y	15199	70.72	Y
15274	95.12	Y	15274	93.93	Y	15274	90.56	Y
15275	93.62	Y	15275	86.88	Y	15275	90.55	Y
15348	93.38	Y	15348	92.14	Y	15348	90.41	Y
15349	89.22	Y	15349	89.60	Y	15349	83.84	Y
15353	67.50	Y	15353	94.37	Y	15353	54.31	Y
15354	88.25	Y	15354	86.61	Y	15354	81.33	Y
15390	89.09	Y	15390	88.59	Y	15390	87.91	Y
15391	94.21	Y	15391	88.87	Y	15391	92.78	Y
15429	74.50	Y	15429	90.97	Y	15429	66.18	Y
15430	64.25	Y	15430	80.44	Y	15430	57.05	Y
15435	76.10	Y	15435	82.04	Y	15435	77.22	Y
15436	78.02	Y	15436	82.13	Y	15436	69.32	Y
15498	66.86	Y	15498	85.93	Y	15498	59.25	Y
15499	72.81	Y	15499	82.23	Y	15499	65.19	Y
15555	95.51	Y	15555	90.55	Y	15555	91.55	Y
15556	90.00	Y	15556	87.59	Y	15556	85.89	Y
15573	87.02	Y	15573	91.72	Y	15573	83.46	Y
15574	90.30	Y	15574	88.77	Y	15574	85.36	Y
15589	90.02	Y	15589	92.10	Y	15589	81.78	Y
15590	78.37	Y	15590	83.91	Y	15590	71.42	Y
15594	89.34	Y	15594	92.74	Y	15594	84.95	Y
15595	72.53	Y	15595	83.24	Y	15595	64.78	Y
15615	50.24	Y	15615	83.72	Y	15615	46.79	Y
15616	50.69	Y	15616	76.06	Y	15616	41.72	Y
15697	69.25	Y	15697	88.43	Y	15697	65.95	Y
15698	66.06	Y	15698	81.28	Y	15698	58.05	Y
15732	60.06	Y	15732	84.25	Y	15732	53.49	Y
15733	89.06	Y	15733	85.98	Y	15733	84.22	Y
15759	87.28	Y	15759	88.23	Y	15759	78.32	Y
15760	89.77	Y	15760	83.31	Y	15760	83.92	Y
15810	93.47	Y	15810	92.24	Y	15810	90.62	Y
15811	94.15	Y	15811	85.39	Y	15811	91.53	Y
15925	74.38	Y	15925	90.30	Y	15925	80.39	Y
15926	85.58	Y	15926	83.58	Y	15926	85.72	Y
16082	89.55	Y	16082	88.96	Y	16082	84.39	Y
16083	71.25	Y	16083	79.75	Y	16083	63.09	Y
16094	87.67	Y	16094	90.93	Y	16094	82.68	Y

16095	92.88	Y	16095	79.89	Y	16095	89.32	Y
16127	94.91	Y	16127	92.97	Y	16127	90.20	Y
16128	92.82	Y	16128	82.03	Y	16128	88.80	Y
16327	94.38	Y	16327	92.25	Y	16327	91.03	Y
16328	88.72	Y	16328	81.94	Y	16328	82.95	Y
16359	82.42	Y	16359	89.64	Y	16359	77.95	Y
16360	69.94	Y	16360	78.74	Y	16360	61.12	Y
16410	93.18	Y	16410	91.11	Y	16410	90.90	Y
16411	94.03	Y	16411	84.81	Y	16411	90.32	Y
16426	94.35	Y	16426	91.71	Y	16426	92.76	Y
16427	87.26	Y	16427	79.96	Y	16427	80.26	Y
16448	86.10	Y	16448	87.60	Y	16448	81.66	Y
16449	69.36	Y	16449	77.62	Y	16449	59.41	Y
16453	85.87	Y	16453	87.99	Y	16453	81.06	Y
16454	91.62	Y	16454	80.24	Y	16454	86.94	Y
16494	92.70	Y	16494	90.11	Y	16494	89.51	Y
16495	92.21	Y	16495	80.51	Y	16495	88.86	Y
16541	89.04	Y	16541	78.04	Y	16541	84.56	Y
16542	88.42	Y	16542	76.91	Y	16542	83.60	Y
16564	83.75	N	16564	76.67	N	16564	71.65	N
16565	88.64	Y	16565	81.55	Y	16565	87.16	Y

S4 Table. Primer pairs for long and real time PCR and pyro-sequencing.

Human Gene	Forward primer (5'-3')	Product size	Annealing Tm (°C)
Long PCR			
Long1-F	GACGGGCTCACATCACCCCATAA	8382	63
Long1-R	GCGTACGGCCAGGGCTATTGGT		
Long2-F	GCCACAACCTAACCTCCTCGGACTCCT	8703	63
Long2-R	GGTGGCTGGCACGAAATTGACC		
qPCR			
O _L -F	CAGCTAAGCACCCCTAATCAAC	162	55.5
O _L -R	TCTAAAGACAGGGGTTAGGC		
LSP-F	AGCCACTTTCCACACAGACATCATA	250	59
LSP-R	TAGGATGGGCGGGGTTGTA		
HSP-F	TCAATACAACCCCGCCCAT	257	57.5
HSP-R	TCGTGGTGATTTAGAGGGTGAA		
O _H -F	GAGCACCCCTATGTCGCAGTA	170	57
O _H -R	TCTGTGTGGAAAGTGGCTGT		
POLG exon2-F	CAGACCTCCACGTCTGAACAC	209	59.5
POLG exon2-R	GACAACCTGGACCAGCACTT		
TOP1MT intron10-F	AAGGACATGTCAGCCGTTGG	136	59.5
TOP1MT intron10-R	GACGACCTCTTCGACAGGC		
TOP1MT exon8-F	CCACCGGCACTCTGTTGTAG	197	59.5
TOP1MT exon8-R	ACCATTCCCCACCAACTAGC		
Pyro-sequencing			
h-MT-ND6_F	ATGGAGGTAGGATTGGTGTT		
h-MT-ND6_RB	Biotin-CACCAACAAACAATATTCAACCAATAACTA		
h-MT-ND6_S1	AGGAGAGGGGTTAGG		
hDloop-OH_FB	Biotin-TGTGGAAAGTGGTTGTGTAGATA		
hDloop-OH_R	CACAAATCTATCACCCCTATTAACCACTCA		
hDloop-OH_S1	AAAACCTCTCCATACATTT		
hDloop-OH_S2	AAACACCCTATATC		
hDloop-OH_S3	CCTACCTCATCCTATTAT		
hmtDNA-LSP_F	GTGTGTGTGTTGGGTAGG		
hmtDNA-LSP_RB	Biotin-CCACTTTCCACACAAACATCATAAC		
hmtDNA-LSP_S1	GTGTGTTGGGTAGGA		
hmtDNA-LSP_S2	GGGTGATTGTTAAAAGTGTA		
hmtDNA-LSP_S3	TTTAAGTGTTGTGGTTAGA		
hmtDNA-HSP_F	TGGTGATTTAGAGGGTGAATT		
hmtDNA-HSP_RB	Biotin-CCTCCCCTCCCATACTACT		
hmtDNA-HSP_S1	ATTTAGAGGGTGAATTTATTG		
hmtDNA-HSP_S2	ATGGGGTGATGTGAG		
hmtDNA-HSP_S3	TTTTTGGGGTTTGGT		

CHAPTER 6

General Discussion

General Discussion

The epigenetic control of mitochondrial DNA (mtDNA) copy number in tumorigenesis and differentiation relies on the bi-directional communication between the nuclear and mitochondrial genomes through, for example, DNA methylation. As introduced in Chapter 1, mtDNA, located inside the mitochondrial matrix, encodes 13 subunits of the electron transfer chain (ETC) complexes and, thus, is essential for efficient oxidative phosphorylation (OXPHOS) and all the other adenosine triphosphate (ATP)-dependent cellular activities [1]. In accordance with its functions, mtDNA copy number of each cell is present in a tissue-specific manner to meet the energy requirements of each cell type, which is established during cell differentiation in synchrony with changes in DNA methylation profiles [2, 3]. In contrast, cancer cells fail to expand mtDNA copy number and are trapped in a ‘pseudo-differentiated’ status, which is associated with dysregulated global DNA methylation patterns [4, 5]. Both tumorigenesis and differentiation require crosstalk between the nuclear and the mitochondrial genomes mediated through DNA methylation, whereby both the nuclear-DNA encoding mtDNA replication and differentiation factors and mtDNA itself can be regulated to achieve the desired mtDNA copy number. This will likely regulate the flow of mitochondrial metabolites and mitochondrial functions that will, in turn, provide feedback to the nucleus to maintain the desired levels of DNA methylation and, thus, gene expression activities.

The first aim of my thesis focused on the key mtDNA replication factors that are DNA methylated and influence mtDNA copy number (Fig. 1) (Chapter 3). It has been previously shown that inducing DNA demethylation at exon 2 of *POLG* can promote replication of mtDNA and cell differentiation in cancer cells [2, 3], therefore, it is

important to further discover other contributing factors. By applying whole-genome MeDIP-Seq to GBM cells and GBM cells treated with 5-azacytidine (5Aza), a DNMT inhibitor [6], and vitamin C (VitC), a TET activator [7], we have determined the sites in the GBM methylome that can be modulated by DNA demethylation treatments. Globally, these sites were able to alter gene expression profiles, which overlap greatly with tumorigenic genes and pathways in cancer, cell death and growth, and cellular development, such as *EGFR* and *SHC2*. Furthermore, DNA demethylation occurred extensively at CpG islands (CGIs) across the promoter regions and the gene bodies of the mtDNA-specific replication factors, especially at exon 2 of *POLG* and exon 8 of *TOP1MT* (Fig. 1). As a result, an overall transcriptional up-regulation of the replication factors was triggered by these treatments, which is considered to be the major cause for the increase in mtDNA copy number. Moreover, we have found that DNA methylation also takes place in the mitochondrial genome and can be modulated by the DNA demethylation treatments to further facilitate mtDNA replication being driven by the nuclear-encoded mtDNA replication factors (Fig. 1). With increasing attention being given to mtDNA methylation, mtDNA hypomethylation has been identified within the D-loop and *ND6* regions mediated by, for example, arsenic exposure resulting in increased mtDNA copy number [8]. These findings highlight the role that DNA demethylation to both genomes plays in boosting mtDNA replication in cancer.

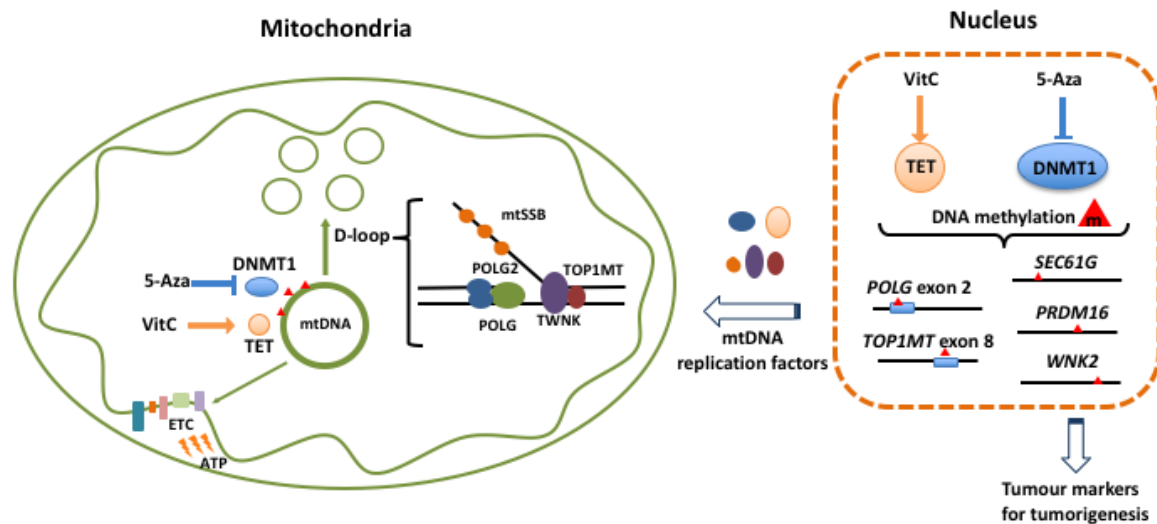


Figure 1. mtDNA genetics and epigenetics. mtDNA transcription of the genes encoding the ETC subunits takes place inside the mitochondrion. mtDNA replication is driven by nDNA-encoded replication factors that bind to sites in the D-loop region, including the single-stranded-binding protein (mtSSB), the helicase TWNK, the catalytic subunit of the mtDNA-specific polymerase, POLG, and its accessory subunits, POLG2, and the mtDNA-specific topoisomerase, TOP1MT. Global DNA demethylation promotes mtDNA replication and increases mtDNA copy number by acting on the nucleus at exon 2 of *POLG*, exon 8 of *TOP1MT* and tumour markers, such as *SEC61G*, *PRDM16* and *WNK2*, and the mitochondrial genome.

Interestingly, even though 5Aza and VitC triggered comparable changes in DNA methylation and gene expression profiles, the uniquely identified differentially methylated regions and expressed genes provided evidence of their different modes of action and their different targets in the genome. In clinics, VitC has been supplemented in the standard epigenetic therapy of 5Aza in cancer patients to enhance epigenetic alterations [9, 10]. Our findings here also demonstrated how the different effects induced by 5Aza and VitC could compensate for each other in modifying DNA methylation and gene expression profiles in cancers. Moreover, several GBM tumour markers including the oncogenes *SEC61G* and *PRDM16* and the tumor suppressor gene *WNK2* have been identified with changes to both DNA methylation and gene expression after both treatments, indicating global DNA demethylation regulated both genomes during tumorigenesis [11-13]. Other key GBM tumour markers, including *EGFR*, *BCL2*, *MYC* and *TERT* that interact closely

with the mtDNA transcription and replication factors, were significantly modulated in terms of their gene expression by 5Aza [14-17], which further indicates the tight relationship between mtDNA replication, DNA methylation and nuclear gene expression in GBM cells.

The second aim of my thesis addressed the question as to how mtDNA copy number during tumorigenesis influenced DNA methylation patterns of the nuclear genome, especially the master regulators of mtDNA replication and differentiation. I used the glioblastoma (GBM) tumour models formed from cells depleted of their mtDNA content to varying amounts, which have been reported to result in significant differences in tumour progression and frequency [18]. Progression and frequency were greatest in the GBM⁵⁰ tumours (50% of the original mtDNA copy number), but less frequent and significantly delayed in the GBM³ (3% of the original mtDNA copy number) and GBM^{0.2} (0.2% of the original mtDNA copy number) tumours [18]. Moreover, mtDNA copy was restored to similar levels through modulation of *POLG* at exon 2 [19]. Previous findings also showed that GBM cells depleted for 7 or 14 days were able to overcome the block in mtDNA replication and complete differentiation, suggesting that they had re-established their mtDNA set points [18] (Fig. 2). My findings have further shown how DNA methylation is involved in mtDNA replication efficiency and the degree of differentiation in determining the fate of cancer cells (Fig. 2). Even though GBM⁵⁰, GBM³ and GBM^{0.2} cells were tumorigenic *in vivo*, there were aberrant patterns of DNA methylation of the mtDNA replication factors and regulators of differentiation compared with the GBM¹⁰⁰ tumours that enabled mtDNA copy number and tumorigenic capacity to be restored (Fig. 2).

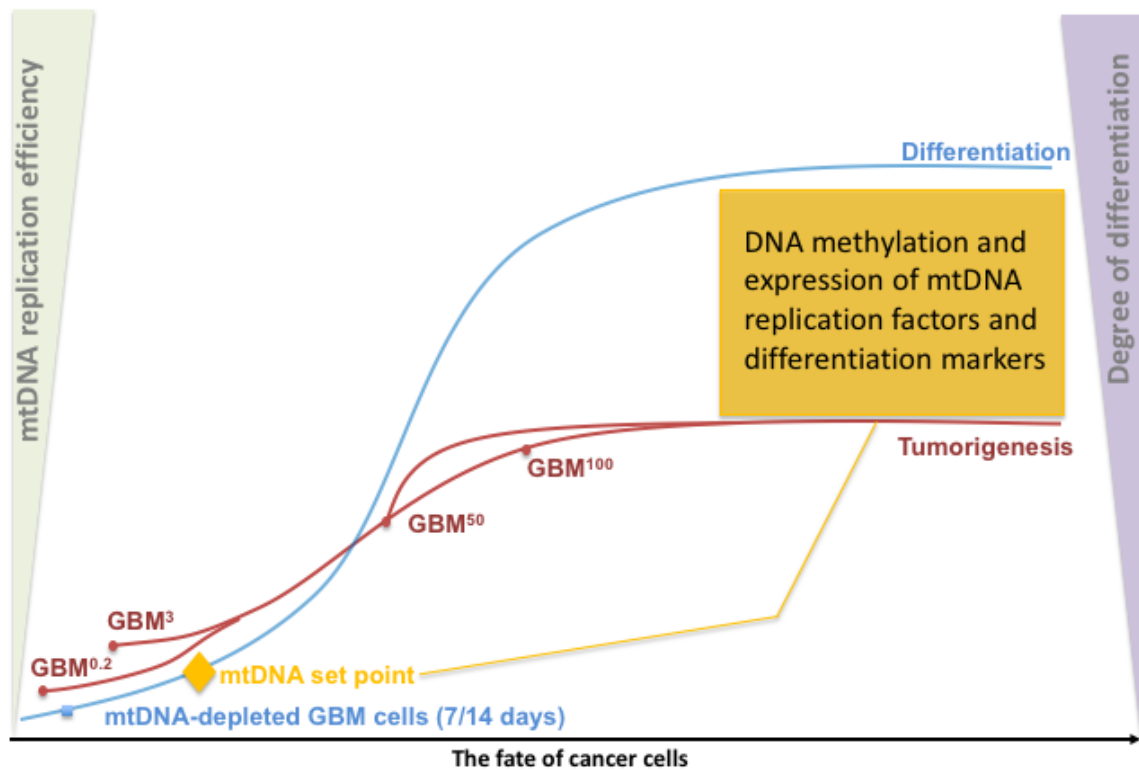


Figure 2. The role of mtDNA copy number in determining the fate of cancer cells. GBM¹⁰⁰, GBM⁵⁰, GBM³ and GBM^{0.2} tumours (red dots) follow the fate of tumorigenesis (red lines), but their mtDNA replication efficiency and degree of differentiation (y axis) are altered. GBM cells undergoing mtDNA depletion for 7 or 14 days (blue dot) re-established their mtDNA set points (orange block) allowing for the resultant changes in DNA methylation and expression of the mtDNA replication factors and differentiation markers required for differentiation (blue line).

As discussed in Chapter 1, the mtDNA set point is achieved by pluripotent (naïve) cells that possess low numbers of mtDNA copy as the initiating templates for mtDNA replication in each cell lineage [3, 18-21]. GBM cells having undergone mtDNA depletion for 7 or 14 days (indicated by the blue dot in Fig. 2) were able to re-establish the mtDNA set point (indicated by the orange block in Fig. 2) and, thus, were able to remodel the DNA methylation profiles and expression of the mtDNA replication factors and differentiation markers allowing for successful differentiation (indicated as the blue line). However, *in vivo* tumorigenesis of GBM⁵⁰ tumours only allowed cells that underwent a change in mtDNA replication efficiency (as indicated by the red dot in Fig. 2), that was not sufficient enough to change the tumorigenic fate to one of differentiation, and, as a consequence, most likely resulted in the faster

progression of tumorigenesis. Likewise, GBM^{0.2} and GBM³ cells failed to switch to the fate of differentiation as they were unable to re-establish their mtDNA set points, which is critical to promoting mtDNA replication efficiency and the capacity for differentiation required for cell lineage-specific differentiation (orange box in Fig. 2). As a result, all of the GBM tumours had to maintain lower numbers of mtDNA of copy to support tumorigenesis, which frequently favours glycolysis to promote cell proliferation [3]. Moreover, mitochondrial functions are tightly associated with DNA methylation as they generate S-adenosylmethionine (SAM), the universal methyl group donor, and α -ketoglutarate in the tricarboxylic acid (TCA) cycle that regulates the activity of the TET family of enzymes [22-26]. However, SAM has been reported to have abnormal levels in one-carbon metabolism in mitochondria that could contribute to tumorigenesis [22], whereas oncomutations in IDH result in the production of the oncometabolite 2-hydroxyglutarate instead of α -ketoglutarate in the TCA cycle [23-26]. To this extent, my findings in Chapter 4 demonstrated that GBM^{0.2} tumours had attempted to modulate global DNA methylation patterns and transcription of mtDNA replication factors and differentiation markers at the expense of tumorigenesis, as shown by the red line of GBM^{0.2} that is the closest to the blue line of differentiation (Fig. 2). This proposed model indicates how modulation of mtDNA load in tumour initiating cells induces changes in DNA methylation of the nuclear genome and, therefore, can affect the fate of cancer cells.

In the last part of my thesis, I have confirmed that DNA methylation takes place in the mitochondrial genome and demonstrated its potential role in regulating mtDNA replication. Following on from the findings in Chapter 3 that induced DNA demethylation takes place in the mitochondrial genome with a potential role in

accelerating mtDNA replication (Chapter 3), I have determined genotype-specific patterns of mtDNA methylation using osteosarcoma cells, 143B cells, and their respective tumour models that have different mtDNA genotypes, namely 143B^{143B}, 143B^{GBM} and 143B^{NSC} tumours. These can change in a synchronized manner with the genotypic-specific expansion of mtDNA copy number from the early to late stages of tumorigenesis. Our group has previously shown mtDNA haplotypes can induce alterations in DNA methylation with consequent transcriptional changes [27] (Appendix A). My findings add another perspective in the understanding as to how levels of mtDNA methylation and mtDNA copy number could communicate with the nuclear genome. Additionally, mtDNA derived from GBM or NSC cells exhibited distinct methylation patterns when it was in combination with a different nuclear genome. This further supports the hypothesis that the epigenetic control of mtDNA copy number is regulated by the nucleus, as demonstrated in Chapter 3.

Importantly, due to several technical issues raised by several studies on mtDNA methylation [28, 29], i.e., sequencing coverage, nuclear mtDNA pseudogenes and supercoiling and the circular structure of mtDNA, we have optimised the experimental pipeline used in Chapter 5. mtDNA samples were purified to avoid the noise arising from the mtDNA pseudogenes present in the nuclear genome [30]. I have also ensured high coverage (> 500 fold coverage on average) in bisulfite-sequencing to avoid bias in determining mtDNA methylation levels resulting from insufficient coverage of < 250 fold, as argued in [28]. Thirdly, the purified mtDNA samples and the linearized long PCR controls were randomly fragmented during bisulfite conversion to eliminate the effects of supercoiling and the circular structure of the mitochondrial genome. The strategy of linearized positive and negative long

PCR controls for each genotype has been carried over from MeDIP-Seq performed in Chapter 3 to provide the baseline for data normalization [30]. Furthermore, since CpG sites at nt 9053 - 9054 and nt 11911 - 11912 were only present in the 143B^{143B} and 143B^{GBM} genomes but not in the 143B^{NSC} mtDNA genomes [19], it is necessary, and it is the first time that genotype-specific analysis has been applied to assess mtDNA methylation. Indeed, mtDNA is known to be susceptible to variants, which can induce different gene expression and DNA methylation profiles and account for different mtDNA genotypes (haplotypes) being predisposed to cancers and other diseases [27, 31-34]. Moreover, the use of 5-methylcytosine (5mC)- and 5-hydroxymethylcytosine (5hmC)-specific antibodies in MeDIP further identified the distinct patterns in transition from 5mC (the methylated state) to 5hmC (the demethylated state) in the tumours investigated, whereas pyro-sequencing served as a sensitive method for determining the subtle changes to mtDNA methylation in the key regulatory regions. This pipeline is a comprehensive method that takes advantage of different techniques and enables careful consideration of mtDNA-specific issues, which can serve as a standard methodology to unlock future mtDNA methylation studies.

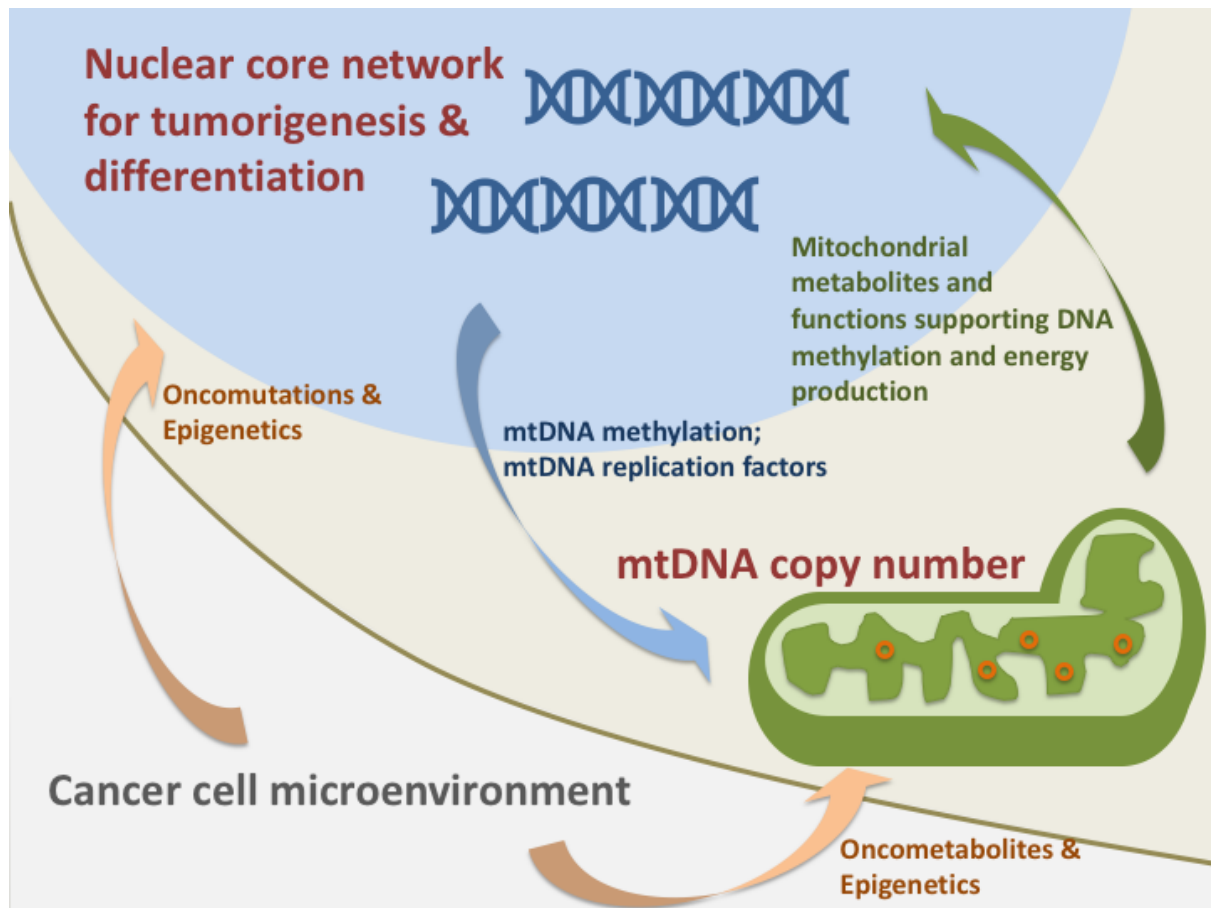


Figure 3. The crosstalk between the nuclear and mitochondrial genomes associated with DNA methylation in cancer cells. In the tumorigenic microenvironment, cancer cells possess oncomutations and aberrant epigenetic profiles in the nucleus to build up the core nuclear network for tumorigenesis and differentiation, and also oncometabolites and aberrant epigenetic profiles in mitochondria to support tumorigenic levels of mtDNA copy number. The nuclear core network exerts epigenetic control on mtDNA copy number, mtDNA replication factors, and mtDNA methylation. The level of mtDNA copy per cell also influences the nuclear core network by maintaining mitochondrial metabolites and functions for DNA methylation and energy production.

Collectively, my findings provide novel insights into our understanding of the epigenetic control of mtDNA copy number in tumorigenesis and differentiation. The crosstalk between the nuclear and the mitochondrial genomes relies on the language encoded by DNA methylation. The proposed model shown in Fig. 3 suggests that cancer cells gained oncomutations and aberrant epigenetic profiles from the tumorigenic microenvironment to build up the core nuclear network for tumorigenesis and differentiation in the nucleus. On the other hand, oncometabolites and the

aberrant epigenetic profiles derived from the mitochondria reflect the presence of the tumorigenic requirements for the levels of mtDNA copy number. The nuclear core network exerts epigenetic control on mtDNA copy number by regulating the DNA methylation status of the mtDNA replication factors and mtDNA methylation. The level of mtDNA copy per cell also influences the nuclear core network by maintaining mitochondrial metabolites and functions for DNA methylation and energy production, which takes place in a mtDNA genotype-specific manner by inducing their own nuclear DNA methylation, and transcriptional and mtDNA methylation profiles. For future directions, it remains to be explored how mtDNA methylation regulates mtDNA replication, e.g. affecting the access and binding of the replication factors to the genome, and how changes in mtDNA copy number or haplotypes alter cellular metabolism such as SAM and α -ketoglutarate to regulate global DNA methylation.

Reference

1. Anderson, S., et al., *Sequence and organization of the human mitochondrial genome*. Nature, 1981. **290**(5806): p. 457-65.
2. Kelly, R.D., et al., *Mitochondrial DNA copy number is regulated in a tissue specific manner by DNA methylation of the nuclear-encoded DNA polymerase gamma A*. Nucleic Acids Res, 2012. **40**(20): p. 10124-38.
3. 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.
4. Lee, W.T. and J. St John, *The control of mitochondrial DNA replication during development and tumorigenesis*. Ann N Y Acad Sci, 2015. **1350**: p. 95-106.
5. Sun, X. and J.C. St John, *The role of the mtDNA set point in differentiation, development and tumorigenesis*. Biochem J, 2016. **473**(19): p. 2955-71.
6. Jones, P.A. and S.M. Taylor, *Cellular differentiation, cytidine analogs and DNA methylation*. Cell, 1980. **20**(1): p. 85-93.
7. Blaschke, K., et al., *Vitamin[thinsp]C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells*. Nature, 2013. **500**(7461): p. 222-226.
8. Sanyal, T., et al., *Hypomethylation of mitochondrial D-loop and ND6 with increased mitochondrial DNA copy number in the arsenic-exposed population*. Toxicology, 2018. **408**: p. 54-61.
9. Sajadian, S.O., et al., *Vitamin C enhances epigenetic modifications induced by 5-azacytidine and cell cycle arrest in the hepatocellular carcinoma cell lines HLE and Huh7*. Clin Epigenetics, 2016. **8**: p. 46.
10. Liu, M., et al., *Vitamin C increases viral mimicry induced by 5-aza-2'-deoxycytidine*. Proc Natl Acad Sci U S A, 2016. **113**(37): p. 10238-44.
11. Lu, Z., et al., *Glioblastoma proto-oncogene SEC61gamma is required for tumor cell survival and response to endoplasmic reticulum stress*. Cancer Res, 2009. **69**(23): p. 9105-11.
12. Moniz, S., et al., *Loss of WNK2 expression by promoter gene methylation occurs in adult gliomas and triggers Rac1-mediated tumour cell invasiveness*. Hum Mol Genet, 2013. **22**(1): p. 84-95.
13. Lei, Q., et al., *miR-101 reverses hypomethylation of the PRDM16 promoter to disrupt mitochondrial function in astrocytoma cells*. Oncotarget, 2016. **7**(4): p. 5007-22.
14. Morrish, F. and D. Hockenbery, *MYC and mitochondrial biogenesis*. Cold Spring Harb Perspect Med, 2014. **4**(5): p. a014225.
15. Sahin, E., et al., *Telomere dysfunction induces metabolic and mitochondrial compromise*. Nature, 2011. **470**(7334): p. 359-65.
16. Autret, A. and S.J. Martin, *Emerging role for members of the Bcl-2 family in mitochondrial morphogenesis*. Mol Cell, 2009. **36**(3): p. 355-63.
17. Suen, D.F., K.L. Norris, and R.J. Youle, *Mitochondrial dynamics and apoptosis*. Genes Dev, 2008. **22**(12): p. 1577-90.
18. Dickinson, A., et al., *The regulation of mitochondrial DNA copy number in glioblastoma cells*. Cell Death Differ, 2013. **20**(12): p. 1644-53.
19. Lee, W.T., et al., *Mitochondrial DNA plasticity is an essential inducer of tumorigenesis*. Cell Death Discov, 2016. **2**: p. 16016.
20. Facucho-Oliveira, J.M., Alderson, J., Splikings, E.C., Egginton, S., and St. John, J.C., *Mitochondria DNA replication during differentiation of murine embryonic stem cells*. J Cell Sci, 2007. **15**(120): p. 4025-34.
21. Facucho-Oliveira, J.M. and J.C. St John, *The relationship between pluripotency and mitochondrial DNA proliferation during early embryo development and embryonic stem cell differentiation*. Stem Cell Rev, 2009. **5**(2): p. 140-58.

22. Cuyas, E., et al., *Metformin regulates global DNA methylation via mitochondrial one-carbon metabolism*. *Oncogene*, 2018. **37**(7): p. 963-970.
23. Figueroa, M.E., et al., *Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation*. *Cancer Cell*, 2010. **18**(6): p. 553-67.
24. Turcan, S., et al., *IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype*. *Nature*, 2012. **483**(7390): p. 479-83.
25. Haseeb, A., M.S. Makki, and T.M. Haqqi, *Modulation of ten-eleven translocation 1 (TET1), Isocitrate Dehydrogenase (IDH) expression, alpha-Ketoglutarate (alpha-KG), and DNA hydroxymethylation levels by interleukin-1beta in primary human chondrocytes*. *J Biol Chem*, 2014. **289**(10): p. 6877-85.
26. Liu, D.C., et al., *HIF-1alpha inhibits IDH-1 expression in osteosarcoma*. *Oncol Rep*, 2017. **38**(1): p. 336-342.
27. Lee, W.T., et al., *Mitochondrial DNA haplotypes induce differential patterns of DNA methylation that result in differential chromosomal gene expression patterns*. *Cell Death Discov*, 2017. **3**: p. 17062.
28. Mehta, M., et al., *Evidence Suggesting Absence of Mitochondrial DNA Methylation*. *Front Genet*, 2017. **8**: p. 166.
29. Liu, B., et al., *CpG methylation patterns of human mitochondrial DNA*. *Sci Rep*, 2016. **6**: p. 23421.
30. Sun, X., J. Johnson, and J.C. St John, *Global DNA methylation synergistically regulates the nuclear and mitochondrial genomes in glioblastoma cells*. *Nucleic Acids Res*, 2018. **46**(12): p. 5977-5995.
31. Gomez-Zaera, M., et al., *Identification of somatic and germline mitochondrial DNA sequence variants in prostate cancer patients*. *Mutat Res*, 2006. **595**(1-2): p. 42-51.
32. Webb, E., et al., *Comprehensive analysis of common mitochondrial DNA variants and colorectal cancer risk*. *Br J Cancer*, 2008. **99**(12): p. 2088-93.
33. He, Y., et al., *Heteroplasmic mitochondrial DNA mutations in normal and tumour cells*. *Nature*, 2010. **464**(7288): p. 610-4.
34. Yeung, K.Y., et al., *The identification of mitochondrial DNA variants in glioblastoma multiforme*. *Acta Neuropathol Commun*, 2014. **2**: p. 1.

Appendix A

ARTICLE

Mitochondrial DNA haplotypes induce differential patterns of DNA methylation that result in differential chromosomal gene expression patterns

William T Lee^{1,2,6}, Xin Sun^{1,2,6}, Te-Sha Tsai^{1,2,6}, Jacqueline L Johnson^{1,2,6}, Jodee A Gould³, Daniel J Garama^{2,4}, Daniel J Gough^{2,4}, Matthew McKenzie^{1,2}, Ian A Trounce⁵ and Justin C St. John^{1,2}

Mitochondrial DNA copy number is strictly regulated during development as naive cells differentiate into mature cells to ensure that specific cell types have sufficient copies of mitochondrial DNA to perform their specialised functions. Mitochondrial DNA haplotypes are defined as specific regions of mitochondrial DNA that cluster with other mitochondrial sequences to show the phylogenetic origins of maternal lineages. Mitochondrial DNA haplotypes are associated with a range of phenotypes and disease. To understand how mitochondrial DNA haplotypes induce these characteristics, we used four embryonic stem cell lines that have the same set of chromosomes but possess different mitochondrial DNA haplotypes. We show that mitochondrial DNA haplotypes influence changes in chromosomal gene expression and affinity for nuclear-encoded mitochondrial DNA replication factors to modulate mitochondrial DNA copy number, two events that act synchronously during differentiation. Global DNA methylation analysis showed that each haplotype induces distinct DNA methylation patterns, which, when modulated by DNA demethylation agents, resulted in skewed gene expression patterns that highlight the effectiveness of the new DNA methylation patterns established by each haplotype. The haplotypes differentially regulate α -ketoglutarate, a metabolite from the TCA cycle that modulates the TET family of proteins, which catalyse the transition from 5-methylcytosine, indicative of DNA methylation, to 5-hydroxymethylcytosine, indicative of DNA demethylation. Our outcomes show that mitochondrial DNA haplotypes differentially modulate chromosomal gene expression patterns of naive and differentiating cells by establishing mitochondrial DNA haplotype-specific DNA methylation patterns.

Cell Death Discovery (2017) 3, 17062; doi:10.1038/cddiscovery.2017.62; published online 11 September 2017

INTRODUCTION

The murine mitochondrial genome (mtDNA) is a double-stranded, ~16.3 kb, circular genome.¹ It encodes 13 proteins of the electron transfer chain, which generates the vast majority of cellular energy through oxidative phosphorylation (OXPHOS). Whilst the majority of the subunits of the electron transfer chain are encoded by the nuclear genome, each of the complexes, except for complex II, has one or more of its proteins encoded by mtDNA.² mtDNA also encodes two rRNAs and 22 tRNAs and has one major non-coding region, the D-loop. The D-loop possesses two hypervariable regions that identify maternal relatives,³ and is the site of interaction for the nuclear-encoded transcription and replication factors.⁴

mtDNA copy number is strictly regulated during development and differentiation.⁵ The primordial germ cells possess ~200 copies of mtDNA,⁶ which exponentially increase during oogenesis until the mature, fertilisable oocyte has >150 000 copies.^{7,8} Following fertilisation, there is active reduction of mtDNA copy number until the blastocyst stage, the final stage of preimplantation development.^{6,8} Whilst the blastocyst's outer ring of cells, the trophectodermal cells, replicate mtDNA as they differentiate into the trophectoderm,⁸ the inner cell mass cells, which form the

embryo proper and are the source of embryonic stem (ES) cells, further reduce mtDNA copy number to establish the mtDNA set point.^{8,9}

The mtDNA set point ensures that all naive (undifferentiated, pluripotent) cells maintain low mtDNA copy number, and, thus, use glycolysis to generate ATP.¹⁰ This promotes cellular proliferation to enable the embryo to generate a critical mass of cells for post-gastrulation development. Once differentiation is initiated, cells replicate their mtDNA in a cell-specific manner,¹¹ which is mediated by the cell-specific DNA methylation of a CpG island in exon 2 of the catalytic subunit of the mtDNA-specific replication factor, DNA polymerase gamma (*PolgA*).¹² Therefore, cells with a high requirement for ATP through OXPHOS, such as heart, muscle and neuronal cells, acquire high numbers of mtDNA copy, whilst cells with a lower requirement for ATP possess fewer copies of mtDNA and use glycolysis.¹³

mtDNA copy number is important to cellular fate. Altering mtDNA copy number in tumour cells can modulate chromosomal gene expression patterns, and promote differentiation.^{11,14} Likewise, mtDNA haplotypes can influence chromosomal gene expression patterns in ES cells¹⁵ and tumours.¹⁴ mtDNA haplotypes are defined by specific regions of mtDNA that identify the

¹Centre for Genetic Diseases, Hudson Institute of Medical Research, 27-31 Wright Street, Clayton, Victoria 3168, Australia; ²Department of Molecular and Translational Science, Monash University, 27-31 Wright Street, Clayton, Victoria 3168, Australia; ³Medical Genomic Facility, Monash Health Translational Precinct, 27-31 Wright Street, Clayton, Victoria 3168, Australia; ⁴Centre for Cancer Research, Hudson Institute of Medical Research, 27-31 Wright Street, Clayton, Victoria 3168, Australia and ⁵Centre for Eye Research Australia, Ophthalmology, University of Melbourne Department of Surgery, 32 Gisborne Street, East Melbourne, Victoria 3002, Australia.

Correspondence: JC St. John (justin.stjohn@hudson.org.au)

⁶These authors all contributed equally to this work.

Received 24 July 2017; accepted 2 August 2017; Edited by A Rufini

phylogenetic origins of maternal lineages.¹⁶ In a range of species, mtDNA haplotypes are associated with adaptation to warm and cold environments,¹⁷ predisposition to diseases of aging such as cancer,¹⁸ diabetes,¹⁹ Alzheimer's²⁰ and Parkinson's,²¹ and fertility.^{22,23}

We have investigated whether chromosomal gene expression patterns can be altered in a haplotype-specific manner due to modulation of global DNA methylation patterns. We assessed global patterns of hypo- and hypermethylation in four ES cell lines each possessing the same chromosomal genotype but different mtDNA genotypes, namely mtDNA divergent ES cell lines. We assessed their mtDNA replicative efficiency during differentiation and, using DNA demethylation agents, determined whether their DNA methylation patterns could be altered to modulate chromosomal gene expression patterns.

RESULTS

Next-generation sequencing of mitochondrial genomes

We sequenced the mitochondrial genomes of four mtDNA divergent mouse ES cell lines (CC9^{mus}, CC9^{spretus}, CC9^{dunni} and CC9^{pahari}) generated from the fusion of enucleated cytoplasts of *Mus musculus*, *Mus spretus*, *Mus dunni* and *Mus pahari* cells to mitochondrial depleted *M. musculus* (CC9.3.1) ES cells; and the parental CC9.3.1 ES cell line to determine their genetic diversity. Figure 1a shows the phylogenetic representation of the four lines relative to the parental line. As the CC9^{mus} line possessed the same mtDNA genotype as the parental line, it was used to exclude bias resulting from generation of the cells. The degree of divergence between the CC9^{mus} and the CC9^{spretus} lines is 1.82 Mya, CC9^{dunni} is 3.88 Mya and CC9^{pahari} is 6.44 Mya (Figure 1b). Supplementary Table S1 shows the single-nucleotide polymorphisms amongst the haplotypes and Supplementary Table S2 the respective changes in amino-acid codons.

Gene expression analysis

To determine whether mitochondrial haplotypes influence the differentiation potential of the four mtDNA divergent ES cell lines, each line was induced to undergo neural differentiation. We analysed cells by real-time PCR for expression of master regulators and endpoint markers of differentiation at 3, 12 and 21 days of differentiation. The nine genes included *Musashi1*, a neural precursor marker; *Nestin*, a primitive neuroepithelial marker; *Ncam1*, an immature neuronal committed progenitor marker; *Sox1*, a neuroectodermal marker; *Pax6*, an advanced neuronal precursor cell marker; *Tubb3*, indicative of newly differentiated neurons; *Map2a*, indicative of mature neurons; *Gfap*, indicative of mature astrocytes; and *Syp*, indicative of mature neurons with synaptic vesicles.

Apart from *Musashi1*, on day 3, the lines showed discordant patterns of expression of the master regulators of neural differentiation (Figure 2a). Whilst there was the anticipated upregulation of *Pax6* with CC9^{spretus} and CC9^{pahari} cells being significantly different, *Ncam1* and *Tubb3* were significantly higher in each of the lines apart from CC9^{mus} cells. Both *Nestin* and *Sox1* also showed upregulation in CC9^{pahari} cells. For the endpoint markers, there was precocious expression of *Syp* in CC9^{spretus} cells and less so in CC9^{dunni} and CC9^{pahari} cells with a similar pattern for *Gfap* in CC9^{spretus} and CC9^{pahari} cells whilst only CC9^{spretus} was upregulated for *Map2a*. On day 12, CC9^{pahari} cells had very high levels of expression for each of the genes, including precocious expression of the endpoint markers *Map2a*, *Gfap* and *Syp* (Figure 2b). On day 21, CC9^{mus} and CC9^{spretus} cells regulated expression at similar levels but there were significant increases in CC9^{dunni} and CC9^{pahari} cells for *Sox1*, and nonsignificant increases for *Map2a*, *Nestin*, *Ncam1*, *Gfap* (CC9^{pahari} only), *Tubb3* (CC9^{dunni} only) and *Sox1*. Consequently, there were discordant patterns of neural gene expression during differentiation and at endpoint

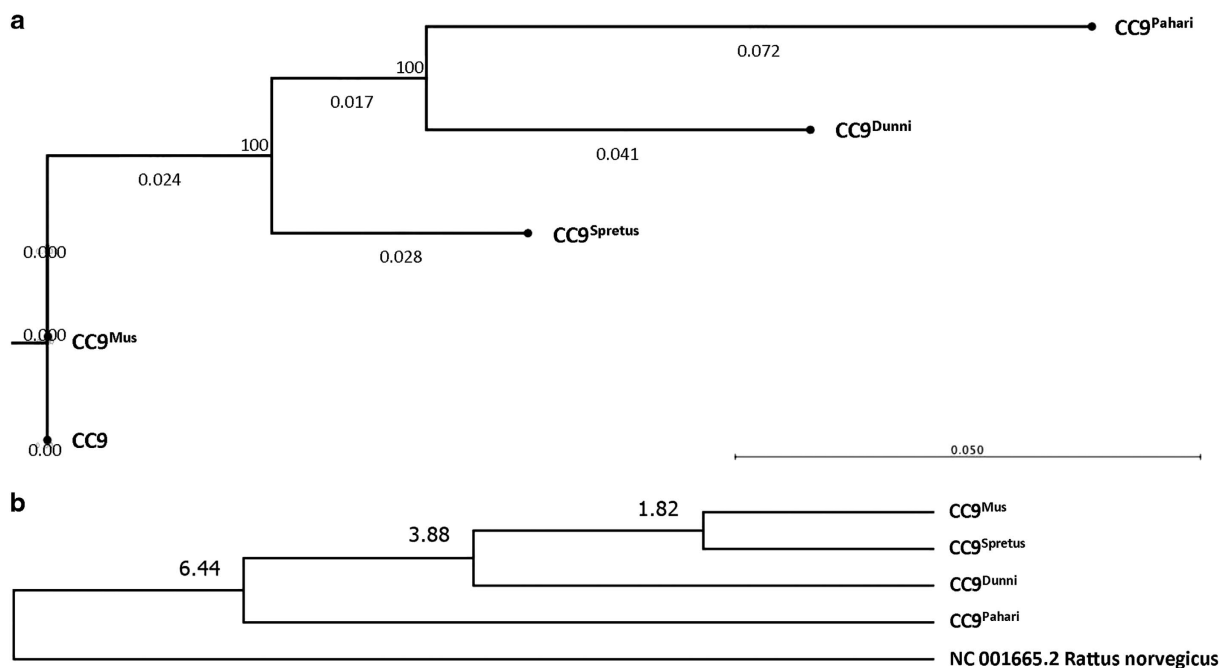


Figure 1. The phylogenetic relationship between the divergent mtDNA ES cell lines. (a) Phylogenetic clustering of mtDNA haplotypes from CC9^{mus}, CC9^{spretus}, CC9^{pahari} and CC9^{dunni} whole mitochondrial genome sequences. A Maximum Likelihood phylogenetic tree was constructed with the GTR model and Neighbor-Joining method with 1000 bootstrap replicates. Bootstrap values are expressed as a percentage; (b) Molecular Phylogenetic analysis by the Maximum Likelihood method. Time of divergence was estimated using the RelTime method. The estimated divergence time for *M. musculus* and *M. spretus* was 1.82 Mya; *M. musculus* and *M. dunni* was 3.88 Mya; and *M. musculus* and *M. pahari* was 6.44 Mya.

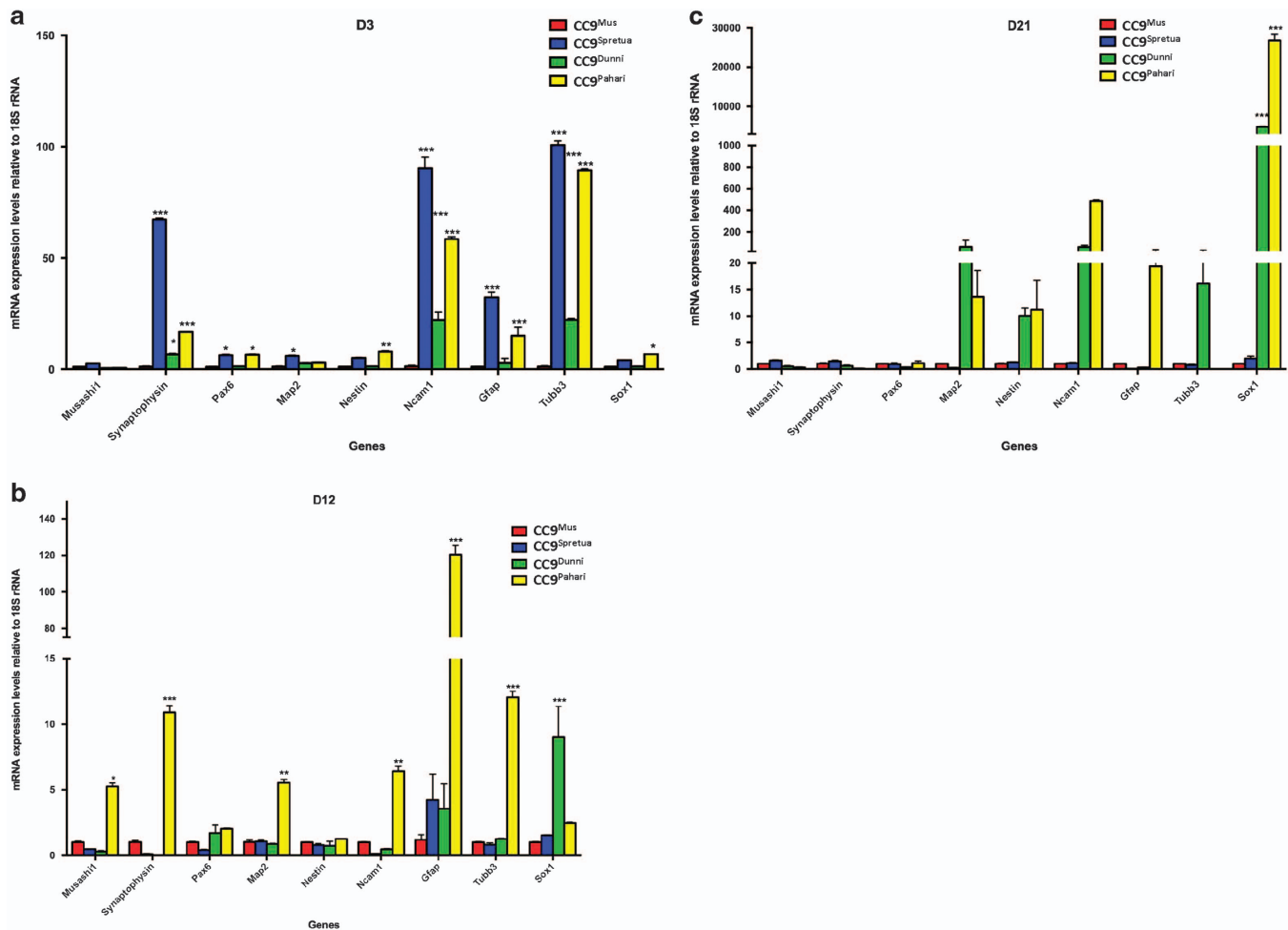


Figure 2. Gene expression during neural differentiation of divergent mtDNA ES lines. CC9^{mus}, CC9^{spretus}, CC9^{dunni} and CC9^{pahari} ES cells were induced to differentiate into neural lineages over 21 days and were assessed at days 3 (a), 12 (b) and 21 (c) for expression of *Musashi1*, *Synaptophysin*, *Pax6*, *Map2*, *Nestin*, *NCAM1*, *Gfap*, *Tubb3* and *Sox1* by real-time PCR. **P* < 0.05; ***P* < 0.01; ****P* < 0.001.

based on a cell's mtDNA haplotype, which were exaggerated as divergence increased (Figure 2c).

mtDNA replicative efficiency

As the divergent ES cell lines showed discordant patterns of expression during neural differentiation, we determined their capacity to regulate mtDNA copy number. We assessed mtDNA copy number per cell for each line and expressed this value as a function of the ratio of 5-methylcytosine (5mC—DNA methylation) to 5-hydroxymethylcytosine (5hmC—DNA demethylation) within exon 2 of *PolgA*, that is, their mtDNA replicative efficiencies. Undifferentiated CC9^{mus} cells exhibited low mtDNA replicative efficiency (Figure 3a), as expected for undifferentiated murine ES cells.^{9,12} However, efficiencies were significantly higher for the other lines. On day 3 of differentiation (Figure 3b), there was a slight increase in mtDNA replicative efficiency for CC9^{mus} cells, indicative of increased mtDNA copy number and the onset of differentiation,⁹ whilst CC9^{spretus} cells returned to very low levels. However, CC9^{dunni} and CC9^{pahari} cells maintained significantly high efficiencies.

On day 12 (Figure 3c), CC9^{mus} cells further increased their replicative efficiency in synchrony with a more differentiated state. However, CC9^{spretus} and CC9^{dunni} cells had significantly lower and CC9^{pahari} cells significantly higher efficiencies. On day 21 (Figure 3d), CC9^{mus} cells increased their mtDNA replicative efficiency, whilst the other lines had significantly lower

efficiencies. Consequently, only CC9^{mus} cells exhibited the potential to replicate mtDNA copy synchronously during differentiation.

Levels of enrichment for POLGA, ESRRB and TFAM

To determine the degree of POLGA affinity for each of the mtDNA haplotypes, using a chromatin immunoprecipitation (ChIP) assay, we assessed its levels of enrichment at its primary binding site in the origin of replication of the heavy strand (O_H) in the D-loop where mtDNA replication is initiated. In undifferentiated cells, the levels of POLGA enrichment were greater in the CC9^{mus} cells (Figure 4a). There was a similar outcome for the enrichment of EsRRB (Figure 4b), a key regulator of pluripotency that binds just upstream of the DNA methylated CpG island in exon 2 of *PolgA*. This suggests that the interaction of a key regulator of pluripotency and the mtDNA-specific replication factor are more tightly regulated in undifferentiated CC9^{mus} ES cells, and more efficient at maintaining the mtDNA set point and restricting precocious differentiation than for the other haplotypes. When we assessed the levels of DNA methylation in the CpG island at exon 2 of *PolgA* by pyrosequencing, undifferentiated CC9^{mus} cells exhibited higher levels of DNA methylation at each of the 11 sites compared with the other divergent lines (Figure 4c). Again, using a ChIP assay, there was discordant binding affinity for TFAM, the mitochondrial transcription factor that initiates mtDNA replication (Figure 4d). Likewise, there was discordant binding

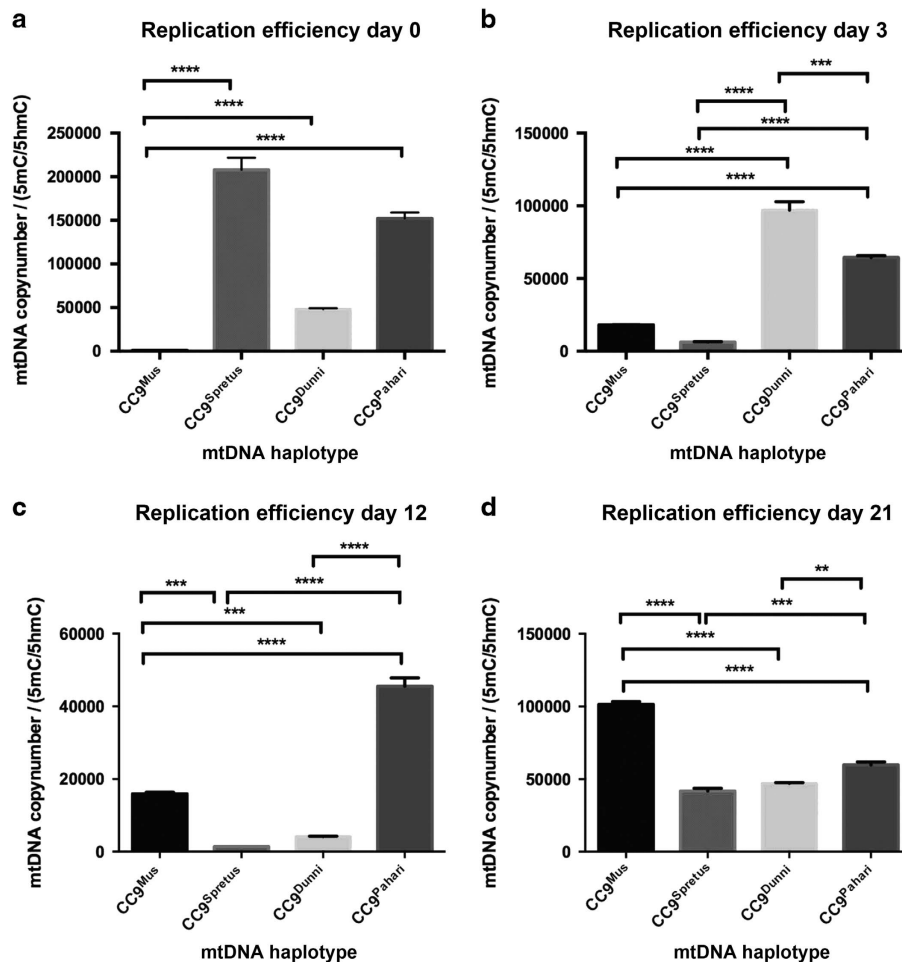


Figure 3. Replicative efficiency of divergent mtDNA ES lines. CC9^{Mus}, CC9^{Spretus}, CC9^{Dunni} and CC9^{Pahari} ES cells were induced to differentiate into neural lineages over 21 days and were assessed at days 0 (a), 3 (b), 12 (c) and 21 (d) of differentiation for mtDNA copy number and ratios of 5mC and 5hmC. MtDNA copy number was assessed by real-time PCR. Levels of enrichment for 5mC and 5hmC were assessed by MeDIP using antibodies against 5mC and 5hmC, and real-time PCR across exon 2 of PolgA. The data are expressed as a ratio of mtDNA copy against 5mC/5hmC, where 5mC and 5hmC are indicative of DNA methylation and DNA demethylation, respectively. ***P* < 0.01; ****P* < 0.001; *****P* < 0.0001.

affinity in the coding genes, namely the *ATPase6* (Supplementary Figure S1A), cytochrome B (Supplementary Figure S1B), *Cox1* (Supplementary Figure S1C) and *Nd1* (Supplementary Figure S1D) genes, where TFAM likely acts as a packaging protein.²⁴

DNA methylation induced by the divergent mtDNA haplotypes in the CC9 chromosomal genome

As there were different patterns of neural gene expression and discordant patterns of binding affinity for POLGA and TFAM for each mtDNA haplotype, we assessed DNA methylation patterns amongst the ES cell lines using the 2 × 105K CpG microarray. We identified 8351 probes, which were assigned to 4243 loci of which 3552 were known genes. The CC9^{Mus} ES cells were the most distinctive with 513 hypermethylated (Figure 5a) and 590 hypomethylated (Figure 5b) genes specific to this line. The CC9^{Spretus} cells had 11 hypermethylated (Figure 5a) and 92 hypomethylated (Figure 5b) genes, the CC9^{Dunni} line 24 hypermethylated (Figure 5a) and 79 hypomethylated (Figure 5b) genes, and the CC9^{Pahari} line 14 hypermethylated (Figure 5a) and 7 hypomethylated (Figure 5b) genes.

Following assignment to DAVID for functional annotation clustering,²⁵ 5 of the 513 hypermethylated CC9^{Mus} genes (Supplementary Table S3) were associated with the mitochondrion, 122 with transcriptional regulation and 40 with neuronal

differentiation and development. Of the hypomethylated CC9^{Mus} genes, 53 genes were associated with the mitochondrion (Supplementary Table S4), 91 with transcriptional regulation and 12 with neuronal differentiation. These DNA methylation patterns are likely to influence mitochondrial respiration, cellular function and differentiation. The 11 hypermethylated CC9^{Spretus} genes mostly affected zinc finger proteins, which regulate DNA- and protein-binding functions whilst the 92 hypomethylated genes affected nuclear function and DNA binding. Most of the hypermethylated CC9^{Dunni} genes affected the regulation of transcription and RNA metabolic processing. Amongst the 79 hypomethylated CC9^{Dunni} genes, 19 were associated with microtubules and cytoskeleton, 8 with cell cycle and 7 with RNA processing.

Modulation of the regulators of DNA methylation in divergent ES cell lines

To determine whether the extensive hypermethylation patterns observed in the divergent ES cells could be modulated by DNA demethylation agents, we cultured CC9^{Mus}, CC9^{Dunni} and CC9^{Pahari} ES cells in the presence of 5-Azacytidine (5-Aza) and vitamin C (VitC) for 48 and 72 h, respectively. 5-Aza modulates DNA methyltransferase 1 (DNMT1) and, therefore, inhibits DNA methylation during cell division,²⁶ whilst VitC acts on TET1 to promote the conversion of 5mC to 5hmC to demethylate DNA.²⁷

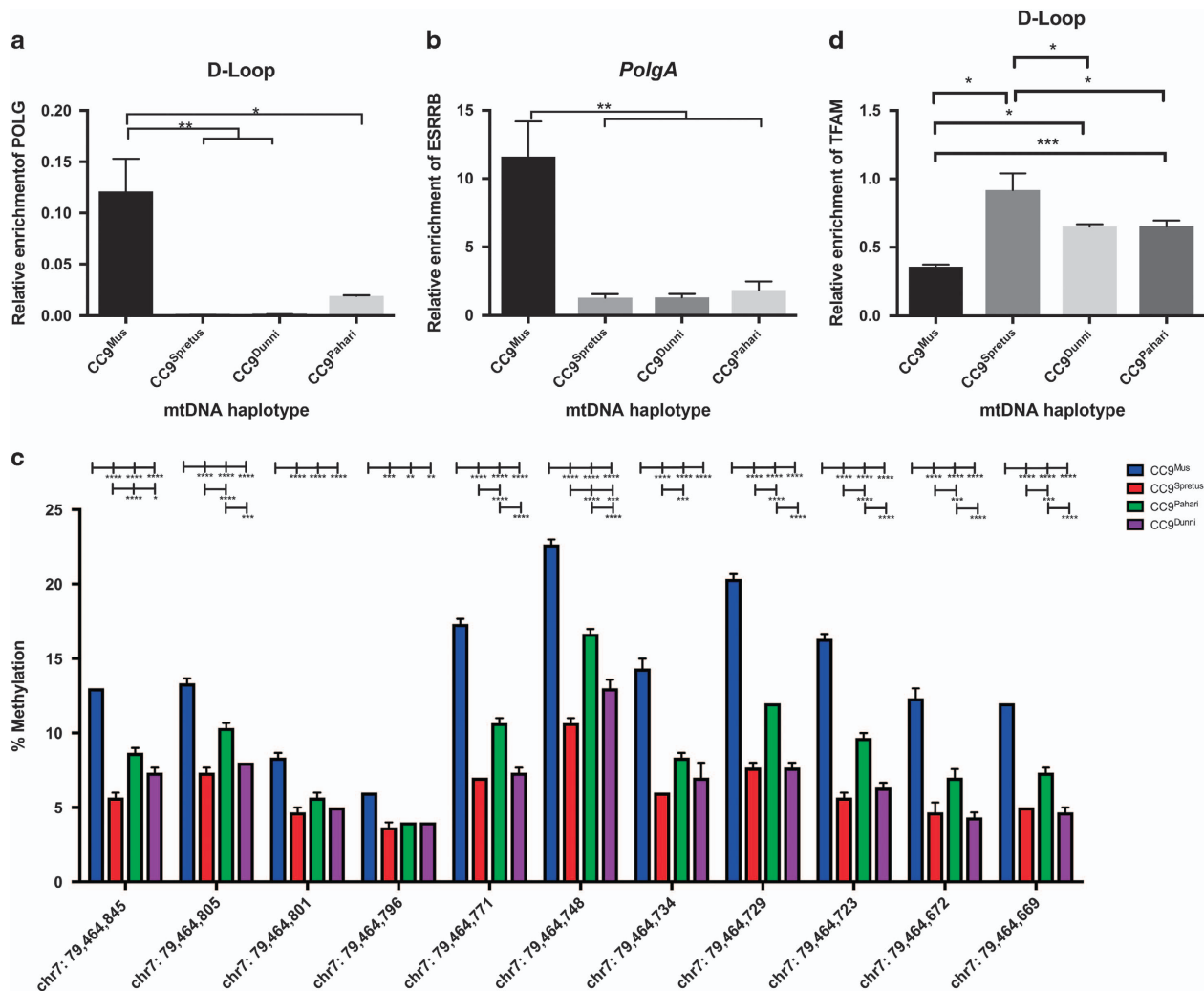


Figure 4. The levels of enrichment for POLGA, ESRRB and TFAM and DNA methylation at exon 2 of *PolgA*. **(a)** Levels of enrichment for *PolgA* in the O_H in the D-loop region of the mitochondrial genome following ChIP using antibodies specific to POLGA and real-time PCR across the O_H region; **(b)** levels of enrichment for ESRRB within the CpG island of *PolgA* as determined by ChIP using an anti-ESRRB antibody and real-time PCR across the region of interest in *PolgA*. **(c)** % Methylation of *PolgA* at exon 2 for CC9^{Mus}, CC9^{Spretus}, CC9^{Pahari} and CC9^{Dunni} cells. Pyrosequencing was performed for 11 CpGs found on mouse *PolgA* exon 2 (chr7: 79 464 669–79 464 845). Primers were designed using the mouse reference sequence from UCSC Genome Browser Dec. 2011 (GRCh38/mm10) Assembly. **(d)** Levels of enrichment for TFAM in the D-loop region of the mitochondrial genome following ChIP using an anti-TFAM antibody and real-time PCR. Data are expressed as mean \pm S.E. M. Statistical analysis was performed using two-way ANOVA followed by Tukey's multiple comparisons test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.

The addition of VitC greatly increased the expression of TET1 in each of the lines (range ~ 20 - to > 60 -fold) whilst 5-Aza marginally affected TET1 activity (Figures 6a and b). Addition of 5-Aza upregulated DNMT1 expression slightly in the three lines though there were also increases with VitC (Figures 6a and c). This is likely to be a response to the inhibition of DNMT1 from interacting within promoter regions by 5-Aza and VitC being unable to inhibit DNMT1 activity. α -Ketoglutarate (α -KG) is a product of the tricarboxylic acid (TCA) cycle and is a cofactor in the conversion of 5mC to 5hmC to demethylate DNA.²⁸ VitC increased levels of α -KG (Figure 6d) whilst 5-Aza reduced levels in CC9^{Dunni} and CC9^{Pahari} cells (Figure 6e). We also examined mitochondrial malate dehydrogenase 2 (MDH2) activity, as MDH2 occurs before α -KG in the TCA cycle, and observed a corresponding fold change decrease in CC9^{Dunni} and CC9^{Pahari} cells with VitC treatment (Figure 6f) but there was no change for 5-Aza (Figure 6g). These results suggest that mtDNA haplotypes modulate the TCA cycle to regulate global DNA methylation patterns.

Modulation of chromosomal gene expression by addition of VitC and 5-Aza to divergent ES cells

To determine whether the changes to the modulators of DNA methylation induced by 5-Aza and VitC affected chromosomal gene expression patterns, we induced cells cultured with 5-Aza and VitC to undergo neural differentiation. Using a Fluidigm array, we analysed undifferentiated, and days 3 and 21 differentiated cells for neural-specific markers associated with neurogenesis, neuronal differentiation, endpoint neural differentiation, neuronal ion channels and neuronal signal transduction that had exhibited either hypermethylation or hypomethylation. We also analysed regulators of DNA methylation, and mtDNA transcription and replication (Supplementary Tables S5; Figure 7).

For the regulators of DNA methylation following VitC treatment, there were overall decreases in gene expression on day 0 except for the CC9^{Dunni} population, increases on day 3 and decreases on day 21 (Supplementary Table S5; Figure 7). Similar patterns were observed for 5-Aza (Supplementary Table S6), although day 21

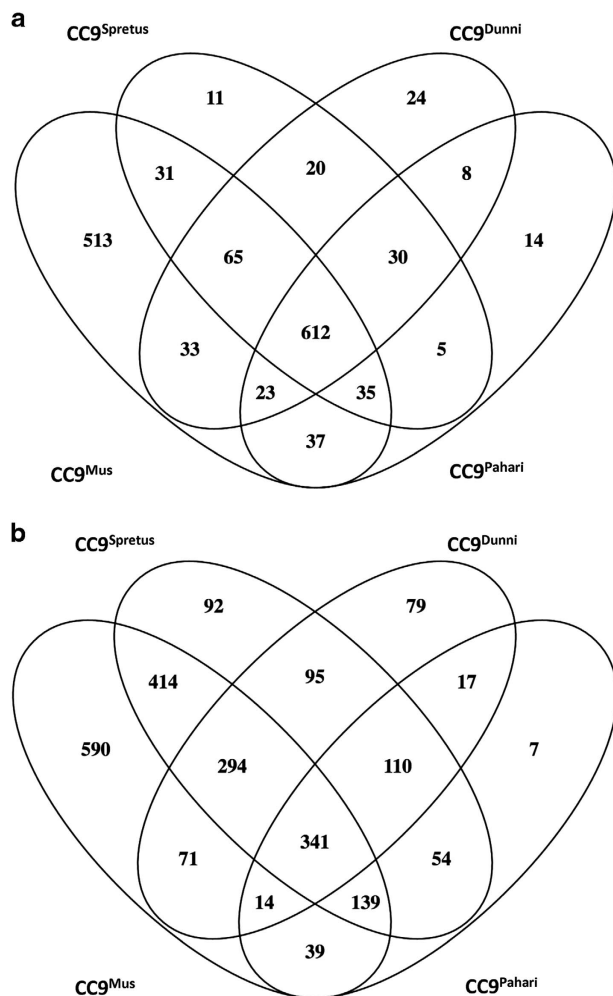


Figure 5. Analysis of hyper- and hypomethylated genes for each of the divergent mtDNA ES cell lines. The methylation status for each of the lines was determined by MeDIP array using an antibody to 5mC, and data were collected using the Agilent CytoGenomics Analytic software (v2.9). (a) Hypermethylated genes; (b) hypomethylated genes.

CC9^{mus} cells did not survive. These outcomes reflect the translation of the transcripts into protein (cf Figure 6; Supplementary Tables S5 and S6). For the mtDNA transcription and replication factors, each of the lines behaved differently with no clear patterns. For the markers of neurogenesis, overall, VitC induced decreases in expression levels in undifferentiated cells and increases on day 3 whilst, on day 21, levels were downregulated in CC9^{mus} and CC9^{pahari} cells but were equally up- and downregulated in CC9^{dunni} cells. 5-Aza treatment only induced an overall change (downregulation) in CC9^{mus} and CC9^{pahari} cells for the neurogenesis markers on day 0, whilst, on day 3, all three lines upregulated expression. On day 21, there was overall downregulation. Consequently, there appears to be effective regulation of the neurogenesis markers with anticipated upregulation on day 3 and downregulation on day 21. For neuronal differentiation, both treatments induced upregulation of gene expression on day 3 but variable outcomes for day 21 with CC9^{mus} and CC9^{pahari} cells downregulating expression after VitC with similar outcomes for CC9^{pahari} and CC9^{dunni} cells after 5-Aza treatment, which is anticipated for genes associated with differentiation. However, neither treatment induced overall increased levels of expression for genes of endpoint markers, neuronal ion channels or neuronal signal transduction. Consequently, resetting DNA methylation

patterns in undifferentiated mtDNA divergent ES cells did not enhance differentiation potential.

DISCUSSION

Whilst mtDNA haplotypes are associated with predisposition to disease,^{18,19,20,21} adaptation to environments^{16,17} and fertility,^{23,29} it has not been apparent how these outcomes are induced. Previously, we had shown that mtDNA haplotypes can modulate neural gene expression patterns in ES cells and their propensity to form beating cardiomyocytes.¹⁵ Here, we show a mechanistic approach where mtDNA haplotypes modulate key regulators of DNA methylation, DNMT1 and TET1. We further show the release of α -KG from the TCA cycle, which is a cofactor in the conversion of 5mC to 5hmC by TET1.²⁸ Nuclear-mitochondrial compatibility is important for establishing functional electron transfer chains.³⁰ However, alternate metabolic pathways, such as the TCA cycle, appear to be more affected when challenged by VitC-induced DNA demethylation. To this extent, whilst each of the lines produced greater levels of α -KG and TET1, the more divergent combination of CC9^{pahari} cells produced twofold more α -KG.

The mtDNA set point is important for establishing and maintaining pluripotency in undifferentiated cells and for regulating mtDNA copy number during differentiation.^{9,12,14,31} The interactions between the nucleus and the mitochondrial genome likely modulate the mtDNA set point to accommodate the requirements of both genetic compartments of the cell,¹⁴ which suggests that the potential to differentiate into certain lineages is regulated by this interaction. Indeed, CC9^{mus} cells were the most efficient at regulating pluripotency through the increased affinity of ESRRB for *PolgA* in undifferentiated cells and synchronising increases in mtDNA copy number with stage-specific changes in gene expression during differentiation. Nevertheless, the use of DNA demethylation agents prevented the cells from completing differentiation, suggesting that having already established the set point had enabled the two genomes to establish their mechanisms of interaction.

During postimplantation development when cells retain their undifferentiated status, mtDNA replication is restricted to replenishing copy number, as is the case in ES cells.⁹ However, mtDNA replication is tested between E7.5 and E10.5 when *PolgA*³² and *Tfam*³³ homozygous knockout mice die *in utero*, which is equivalent to the key mtDNA turnover events identified in ES cells.⁹ Not only is this likely to be important for testing whether the mtDNA set point has been adequately established, but it enables the mitochondrial and chromosomal genomes to determine how they will collectively function. This involves the chromosomal genome-regulating levels of mtDNA copy number to maintain the mtDNA set point to mediate pluripotency and to respond to cues to differentiate, as shown by the differing affinities for the enrichment of *M. musculus*-encoded TFAM, POLGA and ESRRB. In turn, the mitochondrial genome influences the chromosomal genome by regulating OXPHOS activity through the degree of compatibility of the proteins that is encoded for the electron transfer chain. If cells build less functional electron transfer chains, negative feedback could result in reduced TCA cycle activity, as indicated by decreased MDH2 activity following VitC, that would result in increased levels of α -KG and promote the conversion of 5mC to 5hmC,²⁸ as modulated by VitC.²⁷ Similarly, mtDNA-depleted cells have been shown to modulate histone acetylation marks through the TCA cycle as they restore mtDNA copy number.³⁴ Consequently, the initial stages of development involve a genomic 'tug-of-war' that establishes the most advantageous genomic conditions for cells to function effectively, which will affect their ability to complete differentiation as evidenced by the varying degrees of success for the mtDNA divergent ES cells.

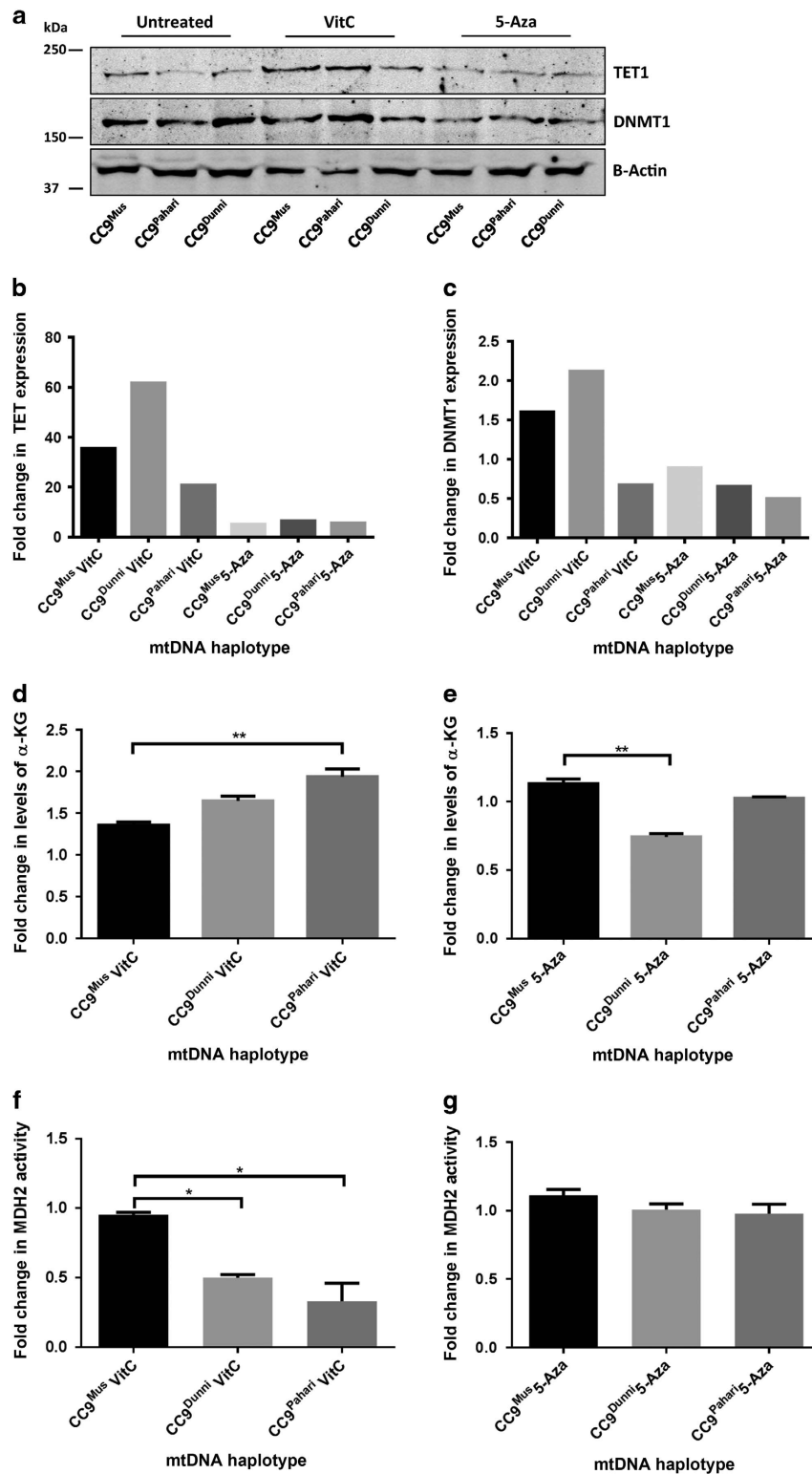
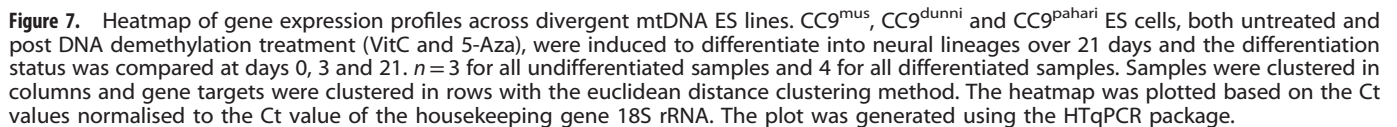


Figure 6. VitC increases TET1 expression and modulates α -KG levels and MDH2 activity. (a) CC9^{Mus}, CC9^{Dunni} and CC9^{Pahari} ES cells were treated with VitC and 5-Aza, and TET1 and DNMT1 protein levels were measured by western blot. Actin protein levels were used to confirm equivalent loading and the precision plus all blue protein marker (Biorad, Gladesville, NSW, Australia) was used to determine protein size. TET1 (b) and DNMT1 (c) expression was normalised to actin and expressed as the fold change in expression compared to vehicle-treated cells. α -KG levels increased as a result of VitC treatment (d) but not 5-Aza (e) whilst MDH2 activity decreased as a result of VitC treatment (f) but 5-Aza had no effect (g) when compared to non-treated cells from the same line with values represented as fold change to non-treated cells. * $P < 0.05$; ** $P < 0.01$.



Adaptation to an oocyte's mtDNA background has important implications for assisted reproductive technologies such as

Cell Death Discovery (2017) 17062

with increased fertility.⁴⁴ Similar problems arise when metaphase II spindle transfer or pronuclear transfer are used to prevent the transmission of mutant mtDNA. Appropriate mtDNA matching would ensure that the resultant cells, tissues and organs had compatible mitochondrial and chromosomal genomes, especially as key imprinting events take place very early during development.⁴⁵

In conclusion, using an ES cell model that tests one set of chromosomes against several divergent mtDNA haplotypes, we have shown that mtDNA haplotypes influence chromosomal gene expression by modulating DNA methylation. Each mtDNA divergent ES cell line established its own DNA methylation profile that could be altered by DNA demethylation agents, which resulted in fold changes in levels of α -KG and perturbed chromosomal gene expression profiles. Furthermore, each of the lines differentially replicated its mtDNA in a specific manner, which was associated with precocious gene expression profiles in the more divergent haplotypes during differentiation. These outcomes could have serious implications for those using nuclear transfer to prevent the transmission of mtDNA disease and account for the disorders associated with somatic cell nuclear transfer.

MATERIALS AND METHODS

Additional materials and methods appear in Supplementary Information.

Mouse ES cell culture and differentiation

M. musculus CC9.3.1 ES cells that were previously reconstructed to harbour *M. musculus* mtDNA (CC9^{mus}) and mtDNA from more divergent subspecies *M. spretus* (CC9^{spretus}), *M. terricolor* (CC9^{dunni}) and *M. pahari* (CC9^{pahari}) were cultured and differentiated, with minor modifications, as previously described¹⁵ (Supplementary Information). For DNA demethylation experiments, undifferentiated ES cells cultured in feeder-free conditions were treated with VitC (Sigma-Aldrich, Castle Hill, NSW, Australia) at a final concentration of 100 μ g/ml or 5-Aza (Sigma-Aldrich) at a final concentration of 0.5 μ M for 72 and 48 h, respectively.

DNA and RNA extraction and cDNA synthesis

Total DNA and RNA were extracted using the DNeasy Blood and Tissue Kit and RNeasy Mini Kit (both Qiagen, Valencia, CA, USA), respectively, according to the manufacturer's protocol. DNA samples were treated with RNase solution (Qiagen) and Proteinase K solution (Qiagen) at 65 °C for 10 min while RNA samples were treated with DNase I (Qiagen) for 20 min. cDNA was synthesised from 1 μ g of total RNA using oligo(dT) primers and the Superscript III First-Strand synthesis system (Thermo Fisher, Scoresby, VIC, Australia), according to the manufacturer's instructions.

Next-generation sequencing of mitochondrial genomes

Next-generation sequencing of complete mitochondrial genomes was performed on amplified long PCR products. Long PCR reactions were prepared, as described in ref. 46 (Supplementary Information) and PCR products purified using the QIAquick PCR Purification Kit (Qiagen), according to the manufacturer's protocol. Purified amplicon pairs were combined at equal concentrations, and amplicon libraries were generated using the recommended workflow procedures from the Ion Fragment Library Kit and Ion Xpress Template kit using 318 chips and run on an Ion Torrent PGM (all Thermo Fisher).

DNA fragments were mapped to a mouse mtDNA reference genome (accession: AP013031), using the CLC Genomics Workbench v7.5.1 (Qiagen), to assemble each mtDNA sequence. The voting strategy was used for base-pair calling. The accession numbers for the mtDNA sequences are KY018919 (*M. musculus*), KY018920 (*M. dunni*), KY018921 (*M. spretus*) and KY038052 (*M. pahari*).

Phylogenetic analysis

Model testing was performed using CLC Genomics Workbench, as described in Tsai *et al.*²³ Using the GTR model,^{47,48} a Maximum Likelihood tree was created with 1000 bootstrap replicates to show the relationship

between the different mtDNA haplotypes. Further details are available in the Supplementary Information.

Evolutionary analyses

Evolutionary analyses were conducted in MEGA6.⁴⁹ The complete mtDNA sequences for each *Mus* species and *Rattus norvegicus* (NC_001662.2) were aligned using ClustalW followed by model testing. The General Time Reversible model⁴⁷ had the lowest Bayesian Information Criterion scores, and was, therefore, selected.⁴⁸ A Maximum Likelihood phylogenetic tree was constructed by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using the Maximum Composite Likelihood approach. A discrete Gamma distribution was used to model evolutionary rate differences amongst sites (five categories (+G, parameter=0.3734)). The tree was drawn to scale, with branch lengths measured by the number of substitutions per site. The tree was supported by 1000 bootstrap replicates. Estimation of divergence time was performed using the RelTime method.⁵⁰ Calibration constraints were based on the *R. norvegicus* and *M. musculus* split of 8–12 Mya.⁵¹

Pyrosequencing of exon 2 of *PolgA*

Pyrosequencing assays were designed using the PyroMark Assay Design Software (Version 2.0.1, Qiagen). A unit of 500 ng DNA samples were converted using the Epitect Bisulphite Conversion Kit (Qiagen), as per the manufacturer's protocol. The region of interest was amplified by PCR using PyroMark PCR Kit (Qiagen) and prepared for pyrosequencing, as described in Supplementary Table S7. Pyrosequencing was performed on a PyroMark 24 Pyrosequencing System (Qiagen), as per the manufacturer's instructions. Data were analysed on the PyroMark Q24 software to determine the % methylation values for each CpG site in the sample.

Immunoprecipitation of methylated DNA

Immunoprecipitation of methylated DNA (MeDIP) was performed, as previously described.^{14,31} Briefly, 3 μ g of the sonicated DNA was immunoprecipitated with 2 μ g of either 5mC (Active Motif, Carlsbad, CA, USA) or 5hmC (Active Motif) at 4 °C overnight, and the immunoprecipitated DNA was purified using the Qiagen PCR Purification Kit (Qiagen). Further details are available in the Supplementary Information.

Chromatin immunoprecipitation

ChIP was performed as previously described.¹⁵ Cells were crosslinked then sonicated to fragment chromatin to an average size of 200–800 bp. Chromatin from 1×10^6 cells was immunoprecipitated with Protein G Dynabeads and an anti-POLGA antibody (G-6, Santa Cruz Biotechnology, Inc., Dallas, TX, USA), or anti-TFAM antibody (Santa Cruz Biotechnology, Inc.), or anti-ESRRB antibody (H6705, R&D Systems, Minneapolis, MN, USA). Crosslinks in immunoprecipitated samples were reversed and pulled-down samples purified using the QIAquick PCR Purification Kit (Qiagen). Further details are available in the Supplementary Information.

Real-time PCR to assess mtDNA copy number, mRNA expression, ChIP and MeDIP

All real-time PCR (quantitative PCR, qPCR) reactions were performed on a RotorGene 3000 real-time PCR machine (Corbett Research, Mortlake, NSW, Australia). The number of mtDNA copies/cell were quantified against external standards for β -actin and mtDNA, as previously described in Kelly *et al.*¹⁵ All primers used are listed in Supplementary Table S7. mRNA expression levels were determined by the $\Delta\Delta$ Ct method, as described in Kelly *et al.*¹⁵ all primers used are listed in Supplementary Table S7. Real-time PCR was performed on MeDIP and ChIP samples using primers amplifying gene regions of interest (Supplementary Table S7) to determine enrichment against input samples, as described in Kelly *et al.*¹⁵

CpG array (MeDIP array)

A unit of 400 ng of input and 150 ng 5mC-containing DNA samples purified from MeDIP, as described above, were used for each CpG microarray (Agilent, Mulgrave, VIC, Australia). Input DNA was Cy3-labelled and the methylated DNA fraction Cy5-labelled using the Agilent SureTag DNA labelling kit for 4 h at 37 °C. Samples were column-purified and combined with Cot-1 DNA, Deionised Formamide, CGH blocking agent and HI-RPM hybridisation buffer (Agilent). Samples were hybridised on mouse

(105K) CpG island microarrays (015279- Agilent) for 40 h at 67 °C. Arrays were washed according to the Agilent CGH protocol, immediately scanned on an Agilent microarray scanner and processed using Agilent Feature extraction software version 11.0.1.1.

The data were processed using the Agilent CytoGenomics Analytic software (v2.9). Once the data were obtained, data filtering was done on SPSS v24.0 (IBM CORP, St. Leonards, NSW, Australia). To identify the differentially methylated probes and genes, we analysed the data by using a cutoff of fourfold differential methylation ratios (\log_2 ratio $> +2$ or < -2) between the samples. We then calculated the mean differential DNA methylation ratios for each of the groups and combined the data with another cutoff obtained from the Z-scores for each of the probes identified. As the Z-scores reflect the relative distance of the log ratios of a probe to the Gaussian distributions of other probes with similar melting temperatures on the array, we further filtered the data using Z-scores $> +5$ or < -5 as cutoffs.

α -KG quantification

α -KG was quantified using the α -Ketoglutarate Assay Kit (Sigma-Aldrich), according to the manufacturer's instructions. In all, 2×10^6 cells were homogenised in α -KG buffer, the samples were mixed with a coupled enzyme, and the resulting fluorometric product measured using a FLUOstar Optima plate reader (BMG Labtech, Mornington, VIC, Australia) at Ex/Em = 544/590 nm. The amount of α -KG per sample in nmol (Ay) was determined from the α -KG standard curve using the equation: $Ay = (\text{corrected absorbance} - (y - \text{intercept})/\text{slope})$. All experimental samples were run as replicates and all samples and standards measured in duplicate. Results are expressed as fold change in α -KG levels compared to untreated cells of each line.

MDH2 assay

MDH2 activity was determined following the manufacturer's protocol (Abcam, 119693). Cells were lysed in extraction buffer, 50 μ g of protein were bound to antibody capture plates, enzyme activity buffer containing a reagent dye was added and absorbance at 450 nm was recorded every 30 s for 30 min on a FLUOstar Optima plate reader (BMG Labtech). Enzyme activity was calculated by: $U = (r_A \times V_{\text{cuvette}})/(l \times \xi \times V_{\text{sample}} \times \rho)$, where r_A = rate of absorbance change; V_{cuvette} = volume of the solution; l = optical path length; ξ = extinction coefficient; V_{sample} = volume sample; ρ = mass concentration of material. The extinction coefficient of the reagent dye was 37/mM/cm. Results are expressed as fold change in MDH2 activity compared to untreated cells of each line.

Fluidigm array and analysis

Pre-amplification was performed on cDNA samples, as described in the Gene Expression Preamp with Fluidigm Preamp MasterMix (Fluidigm, San Francisco, CA, USA) and Taqman Assays Quick Reference PN 68000133 RevC protocol. In all, 96 Taqman assays were selected, as listed in Supplementary Table S8, and pooled with C1 DNA suspension buffer to produce a final concentration for each assay of 180 nM. A volume of 1.25 μ l of each cDNA sample and a non-template control underwent pre-amplification for 14 cycles with 3.75 μ l of pooled assays and Taqman PreAmp Master Mix (Life Technologies), according to the manufacturer's instructions.

Assays and samples were combined in a 96.96 Dynamic array Integrated Fluidic Circuit (IFC) plate, according to the Fluidigm 96.96 Real-Time PCR Workflow Quick Reference PN 6800088 protocol. Using the IFC controller HX, 5 μ l of each pre-amplified sample was loaded as duplicates into each sample inlet and 5 μ l of each Taqman assay ($10 \times$) was loaded into the assay inlet of the plate. Gene expression was performed according to the Biomark GE 96.96 Standard v2 Protocol. Data were exported using the Fluidigm Real-Time PCR analysis software (v4.1.1). Differentially expressed genes were analysed using the HTqPCR package (version 1.26). The normalisation of 'deltaCt' and the Limma method were used.

Genomic data sets

The accession numbers for the MedIP array data sets reported in this paper are deposited as NCBI GEO: GSE94918 (<http://www.ncbi.nlm.nih.gov/geo/>). The mtDNA next-generation sequencing data are deposited at GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>). The respective accession numbers are KY018919 (*M. musculus*), KY018920 (*M. dunni*), KY018921 (*M. spretus*) and KY038052 (*M. pahari*).

ACKNOWLEDGEMENTS

This work was supported by the Victorian Government's Operational Infrastructure Support Program.

AUTHOR CONTRIBUTIONS

WTL, XS, T-ST and JCSJ designed the study. WTL, XS, T-ST, JJJ, JAG, DJG, DJG and MM performed the experiments. WTL, XS, T-ST, JAG, DJG, DJG, MM, IAT and JCSJ analysed the data. WTL, XS, T-ST, JAG, DJG, DJG, MM, IAT and JCSJ wrote and edited the manuscript.

COMPETING INTERESTS

The authors declare no conflict of interest.

PUBLISHER'S NOTE

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Bibb MJ, Van Etten RA, Wright CT, Walberg MW, Clayton DA. Sequence and gene organization of mouse mitochondrial DNA. *Cell* 1981; **26**(2 Pt 2): 167–180.
- Pfeiffer T, Schuster S, Bonhoeffer S. Cooperation and competition in the evolution of ATP-producing pathways. *Science* 2001; **292**: 504–507.
- Ivanov PL, Wadhams MJ, Roby RK, Holland MM, Weedn VW, Parsons TJ. Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. *Nat Genet* 1996; **12**: 417–420.
- Kucej M, Butow RA. Evolutionary tinkering with mitochondrial nucleoids. *Trends Cell Biol* 2007; **17**: 586–592.
- St John JC, Facucho-Oliveira J, Jiang Y, Kelly R, Salah R. Mitochondrial DNA transmission, replication and inheritance: a journey from the gamete through the embryo and into offspring and embryonic stem cells. *Hum Reprod Update* 2010; **16**: 488–509.
- Cree LM, Samuels DC, de Sousa Lopes SC, Rajasimha HK, Wonnapijit P, Mann JR et al. A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes. *Nat Genet* 2008; **40**: 249–254.
- Santos TA, El Shourbagy S, St John JC. Mitochondrial content reflects oocyte variability and fertilization outcome. *Fertil Steril* 2006; **85**: 584–591.
- Spikings EC, Alderson J, St John JC. Regulated mitochondrial DNA replication during oocyte maturation is essential for successful porcine embryonic development. *Biol Reprod* 2007; **76**: 327–335.
- Fachucho-Oliveira JM, Alderson J, Spikings EC, Egginton S, St John JC. Mitochondrial DNA replication during differentiation of murine embryonic stem cells. *J Cell Sci* 2007; **120**(Pt 22): 4025–4034.
- Fachucho-Oliveira JM, St John JC. The relationship between pluripotency and mitochondrial DNA proliferation during early embryo development and embryonic stem cell differentiation. *Stem Cell Rev* 2009; **5**: 140–158.
- Dickinson A, Yeung KY, Donoghue J, Baker MJ, Kelly RD, McKenzie M et al. The regulation of mitochondrial DNA copy number in glioblastoma cells. *Cell Death Differ* 2013; **20**: 1644–1653.
- Kelly RD, Mahmud A, McKenzie M, Trounce IA, St John JC. Mitochondrial DNA copy number is regulated in a tissue specific manner by DNA methylation of the nuclear-encoded DNA polymerase gamma A. *Nucleic Acids Res* 2012; **40**: 10124–10138.
- Moyes CD, Battersby BJ, Leary SC. Regulation of muscle mitochondrial design. *J Exp Biol* 1998; **201**(Pt 3): 299–307.
- Lee WTY, Cain JE, Cuddihy A, Johnson J, Dickinson A, Yeung KY et al. Mitochondrial DNA plasticity is an essential inducer of tumorigenesis. *Cell Death Discovery* 2016; **2**: 16016.
- Kelly RD, Rodda AE, Dickinson A, Mahmud A, Nefzger CM, Lee W et al. Mitochondrial DNA haplotypes define gene expression patterns in pluripotent and differentiating embryonic stem cells. *Stem Cells* 2013; **31**: 703–716.
- Wallace DC. Bioenergetics in human evolution and disease: implications for the origins of biological complexity and the missing genetic variation of common diseases. *Philos Trans R Soc B Biol Sci* 2013; **368**: 20120267.
- Ruiz-Pesini E, Mishmar D, Brandon M, Procaccio V, Wallace DC. Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science* 2004; **303**: 223–226.

- 18 Shen L, Wei J, Chen T, He J, Qu J, He X *et al*. Evaluating mitochondrial DNA in patients with breast cancer and benign breast disease. *J Cancer Res Clin Oncol* 2011; **137**: 669–675.
- 19 Liou CW, Chen JB, Tiao MM, Weng SW, Huang TL, Chuang JH *et al*. Mitochondrial DNA coding and control region variants as genetic risk factors for type 2 diabetes. *Diabetes* 2012; **61**: 2642–2651.
- 20 Ridge PG, Maxwell TJ, Corcoran CD, Norton MC, Tschanz JT, O'Brien E *et al*. Mitochondrial genomic analysis of late onset Alzheimer's disease reveals protective haplogroups H6A1A/H6A1B: the Cache County Study on Memory in Aging. *PLoS One* 2012; **7**: e45134.
- 21 Ghezzi D, Marelli C, Achilli A, Goldwurm S, Pezzoli G, Barone P *et al*. Mitochondrial DNA haplogroup K is associated with a lower risk of Parkinson's disease in Italians. *Eur J Hum Genet* 2005; **13**: 748–752.
- 22 Ruiz-Pesini E, Lapena AC, Diez-Sanchez C, Perez-Martos A, Montoya J, Alvarez E *et al*. Human mtDNA haplogroups associated with high or reduced spermatozoa motility. *Am J Hum Genet* 2000; **67**: 682–696.
- 23 Tsai TS, Rajasekar S, St John JC. The relationship between mitochondrial DNA haplotype and the reproductive capacity of domestic pigs (*Sus scrofa domestica*). *BMC Genetics* 2016; **17**: 67.
- 24 Kaufman BA, Durisic N, Mativetsky JM, Costantino S, Hancock MA, Grutter P *et al*. The mitochondrial transcription factor TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures. *Mol Biol Cell* 2007; **18**: 3225–3236.
- 25 Huang, da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 2009; **37**: 1–13.
- 26 Kelly TK, De Carvalho DD, Jones PA. Epigenetic modifications as therapeutic targets. *Nat Biotechnol* 2010; **28**: 1069–1078.
- 27 Blaschke K, Ebata KT, Karimi MM, Zepeda-Martinez JA, Goyal P, Mahapatra S *et al*. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* 2013; **500**: 222–226.
- 28 Haseeb A, Makki MS, Haqqi TM. Modulation of ten-eleven translocation 1 (TET1), Isocitrate Dehydrogenase (IDH) expression, alpha-Ketoglutarate (alpha-KG), and DNA hydroxymethylation levels by interleukin-1beta in primary human chondrocytes. *J Biol Chem* 2014; **289**: 6877–6885.
- 29 Sutarno, Cummins JM, Greeff J, Lymbery AJ. Mitochondrial DNA polymorphisms and fertility in beef cattle. *Theriogenology* 2002; **57**: 1603–1610.
- 30 McKenzie M, Trounce I. Expression of *Rattus norvegicus* mtDNA in *Mus musculus* cells results in multiple respiratory chain defects. *J Biol Chem* 2000; **275**: 31514–31519.
- 31 Lee W, Johnson J, Gough DJ, Donoghue J, Cagnone GLM, Vaghjiani V *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**: e1664.
- 32 Hance N, Ekstrand MI, Trifunovic A. Mitochondrial DNA polymerase gamma is essential for mammalian embryogenesis. *Hum Mol Genet* 2005; **14**: 1775–1783.
- 33 Larsson NG, Wang J, Wilhelmsson H, Oldfors A, Rustin P, Lewandoski M *et al*. Mitochondrial transcription factor A is necessary for mtDNA maintenance and embryogenesis in mice. *Nat Genet* 1998; **18**: 231–236.
- 34 Martinez-Reyes I, Diebold LP, Kong H, Schieber M, Huang H, Hensley CT *et al*. TCA cycle and mitochondrial membrane potential are necessary for diverse biological functions. *Mol Cell* 2016; **61**: 199–209.
- 35 Stephens PR, Wiens JJ. Testing for evolutionary trade-offs in a phylogenetic context: ecological diversification and evolution of locomotor performance in emydid turtles. *J Evol Biol* 2008; **21**: 77–87.
- 36 Pick JL, Hutter P, Ebner C, Ziegler AK, Giordano M, Tschirren B. Artificial selection reveals the energetic expense of producing larger eggs. *Front Zool* 2016; **13**: 38.
- 37 Paschal JC, Sanders JO, Kerr JL. Calving and weaning characteristics of Angus-, Gray Brahman-, Gir-, Indu-Brazil-, Nellore-, and Red Brahman-sired F1 calves. *J Anim Sci* 1991; **69**: 2395–2402.
- 38 Martin LC, Brinks JS, Bourdon RM, Cundiff LV. Genetic effects on beef heifer puberty and subsequent reproduction. *J Anim Sci* 1992; **70**: 4006–4017.
- 39 Eler JP, Silva JA, Ferraz JB, Dias F, Oliveira HN, Evans JL *et al*. Genetic evaluation of the probability of pregnancy at 14 months for Nellore heifers. *J Anim Sci* 2002; **80**: 951–954.
- 40 Wilmut I, Schnieke AE, McWhir J, Kind AJ, Campbell KH. Viable offspring derived from fetal and adult mammalian cells. *Nature* 1997; **385**: 810–813.
- 41 Meirelles FV, Smith LC. Mitochondrial genotype segregation in a mouse heteroplasmic lineage produced by embryonic karyoplast transplantation. *Genetics* 1997; **145**: 445–451.
- 42 Tachibana M, Sparman M, Sritanandomchai H, Ma H, Clepper L, Woodward J *et al*. Mitochondrial gene replacement in primate offspring and embryonic stem cells. *Nature* 2009; **461**: 367–372.
- 43 Jeanisch R, Eggan K, Humpherys D, Rideout W, Hochedlinger K. Nuclear cloning, stem cells, and genomic reprogramming. *Cloning Stem Cells* 2002; **4**: 389–396.
- 44 Bowles EJ, Campbell KH, St John JC. Nuclear transfer: preservation of a nuclear genome at the expense of its associated mtDNA genome(s). *Curr Top Dev Biol* 2007; **77**: 251–290.
- 45 Doherty AS, Mann MR, Tremblay KD, Bartolomei MS, Schultz RM. Differential effects of culture on imprinted H19 expression in the preimplantation mouse embryo. *Biol Reprod* 2000; **62**: 1526–1535.
- 46 Johnson J, Lee W, Frazier AE, Vaghjiani V, Laskowski A, Rodriguez AL *et al*. Deletion of the complex I subunit NDUFS4 adversely modulates cellular differentiation. *Stem Cells Dev* 2016; **25**: 239–250.
- 47 Yang Z. Estimating the pattern of nucleotide substitution. *J Mol Evol* 1994; **39**: 105–111.
- 48 Nei M, Kumar S. *Molecular Evolution and Phylogenetics*. Oxford University Press: New York, NY, USA 2000.
- 49 Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 2013; **30**: 2725–2729.
- 50 Tamura K, Battistuzzi FU, Billings-Ross P, Murillo O, Filipski A, Kumar S. Estimating divergence times in large molecular phylogenies. *Proc Natl Acad Sci USA* 2012; **109**: 19333–19338.
- 51 Catzeffis FM, Aguilar JP, Jaeger JJ. Muroid rodents: phylogeny and evolution. *Trends Ecol Evol* 1992; **7**: 122–126.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017

Supplementary Information accompanies the paper on the *Cell Death Discovery* website (<http://www.nature.com/cddiscovery>)

Appendix B

The mitochondrial genome: how it drives fertility

Justin C. St. John^{A,B}, Kanokwan Srirattana^{A,B}, Te-Sha Tsai^{A,B} and Xin Sun^{A,B}

^ACentre for Genetic Diseases, Hudson Institute of Medical Research, Clayton, Vic. 3168, Australia.

^BDepartment of Molecular and Translational Sciences, Monash University, Clayton, Vic. 3168, Australia.

^CCorresponding author. Email: justin.stjohn@hudson.org.au

Abstract. In mammalian species, the mitochondrial genome is between 16.2 and 16.7 kb in size and encodes key proteins associated with the cell's major energy-generating apparatus, the electron transfer chain. The maternally inherited mitochondrial genome has, until recently, been thought to be only involved in the production of energy. In this review, we analyse how the mitochondrial genome influences the developing embryo and cellular differentiation, as well as fetal and offspring health and wellbeing. We make specific reference to two assisted reproductive technologies, namely mitochondrial supplementation and somatic cell nuclear transfer, and how modulating the mitochondrial content in the oocyte influences embryo viability and the potential to generate enhanced offspring for livestock production purposes. We also explain why it is important to ensure that the transmission of only one population of mitochondrial (mt) DNA is maintained through to the offspring and why two populations of genetically distinct mitochondrial genomes could be deleterious. Finally, we explain how mtDNA influences chromosomal gene expression patterns in developing embryos and cells primarily by modulating DNA methylation patterns through factors associated with the citric acid cycle. These factors can then modulate the ten–eleven translocation (TET) pathway, which, in turn, determines whether a cell is in a more or less DNA methylated state.

Additional keywords: DNA methylation, embryo, livestock production, mitochondrial DNA, mitochondrial haplotypes, mitochondrial supplementation, oocyte, somatic cell nuclear transfer.

Introduction

Whilst much has been made about the role of mitochondria during the fertilisation process, development and, in somatic cells, the mitochondrial genome has been largely ignored in terms of fertilisation outcome and animal production. However, more recent investigations suggest that the mitochondrial genome does have an important role to play and data from somatic cell biology indicate that different mitochondrial (mt) DNA haplotypes can influence chromosomal gene expression patterns in these cells. This review aims to discuss the roles that the mitochondrial genome has to play in developmental outcomes and how different mtDNA haplotypes influence the phenotype of livestock and can predispose to or protect against specific traits that were previously thought to be the sole domain of the chromosomal genome.

The use of technologies such as mitochondrial supplementation and somatic cell nuclear transfer (SCNT) can propagate specific traits that the livestock and breeding industries would likely require in the future in order to generate animals that are fit to survive. The changes are required due to the world's climate and other factors that would affect production efficiencies, such as infection and disease. However, as little concern

has been given to the mitochondrial genome, much emphasis must now be related to its selection when assisted reproductive technologies (ARTs) are performed because it can significantly influence outcomes in a positive and negative manner. This review will cover mitochondrial supplementation and SCNT, and suggest why regulation of the mitochondrial genome in terms of its haplotype and replication during development must be controlled in order that we produce healthy live offspring that are robust and can carry out the tasks for which their phenotypes have been selected.

Mitochondria

Cells have multiple pathways for the generation of energy (see Fig. 1). In the cytoplasm of the cell, glycolysis mediates the breakdown of glucose that produces pyruvate, NADH and ATP (Pfeiffer *et al.* 2001). Whilst the process of glycolysis is rapid, the amount of ATP generated is limited to four molecules from every two molecules of glucose invested (Pfeiffer *et al.* 2001). However, when oxygen is in supply, pyruvate and NADH enter the mitochondrion. Pyruvate is oxidised in the citric acid cycle, resulting in the production of more NADH and FADH₂. NADH

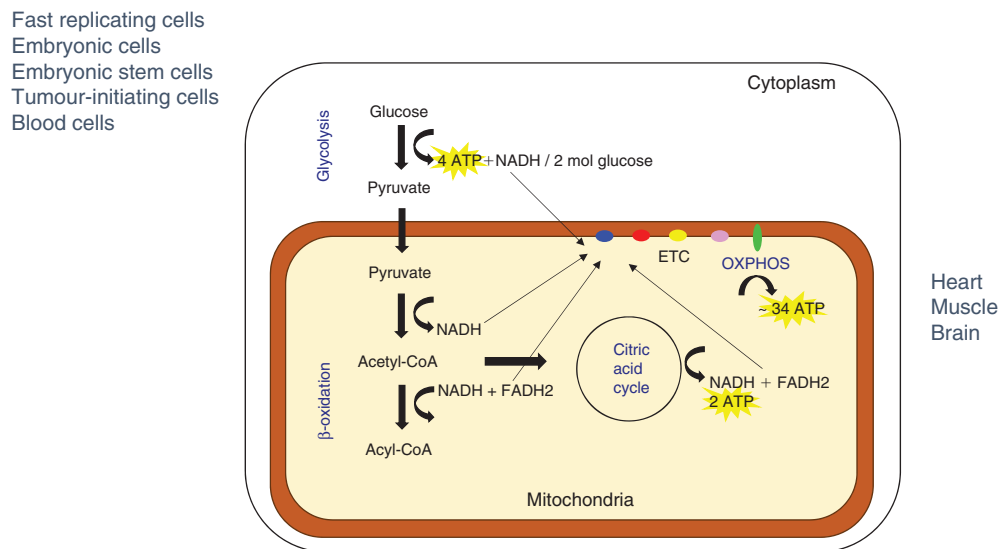


Fig. 1. Cellular pathways for the generation of energy. Glycolysis metabolises 2 mol glucose to generate pyruvate, NADH and four molecules of ATP. Under aerobic conditions, pyruvate enters the mitochondrion and is enzymatically broken down through β -oxidation and the citric acid cycle to produce further molecules of NADH and FADH₂. The electrons of NADH and FADH₂ are transferred to the electron transfer chain (ETC) to synthesise another 34–38 molecules of ATP through oxidative phosphorylation (OXPHOS). Glycolysis is favoured by fast-replicating cells, such as embryonic cells, embryonic stem cells, tumour-initiating cells and blood cells, which have a low requirement for ATP derived through OXPHOS. However, heart, muscle and brain cells have a greater reliance on OXPHOS. CoA, coenzyme A.

and FADH₂ act as electron donors that promote further synthesis of ATP through oxidative phosphorylation, which takes place in the electron transfer chain (Pfeiffer *et al.* 2001). Although oxidative phosphorylation is a longer and more complex process than glycolysis for the production of ATP, it produces 34–38 molecules of ATP for every two molecules of glucose that enter glycolysis.

In maturing oocytes and embryos, mitochondria tend to be naïve, oval-type structures that are limited in their ability to produce energy (Sathananthan *et al.* 2002; Van Blerkom 2004). Mature mitochondria, typically present in differentiated cells, are elongated structures and form networks (Schultz *et al.* 2016). They comprise an outer membrane that regulates the movement of molecules into and out of this organelle (Duchen 1999; Van Blerkom 2004). The inner membrane acts as a backbone for internal structures such as the electron transfer chain. The inner and outer membranes collaborate to form the mitochondrial membrane potential that results from the transit of protons out of and into the intermembrane space, arising from electrons passing through each of the complexes of the electron transfer chain. The inner membrane can also fold to produce cristae, with the resulting inner space, namely the matrix, being far denser in cells that have a high dependency for energy production through oxidative phosphorylation (Quintana-Cabrera *et al.* 2017). In addition, the matrix houses the mitochondrial genome and ribosomes (Shutt and Shadel 2010).

Mature mitochondria that form highly structured networks have other important roles to play in cellular function. For example, they act as Ca²⁺ stores and regulators (Duchen 1999); they initiate the process of steroidogenesis by breaking

down cholesterol (Martinez *et al.* 2015); they provide balanced free radical activity by catalysing excess free radicals that are not used to mediate intercellular activities and can be deleterious to cellular function (Sheshadri and Kumar 2016); and they are direct and indirect mediators of apoptosis and necrosis (Wang *et al.* 2017). In addition, they are important regulators of innate immune responses to pathogens and cell stress through, for example, modulation of cellular metabolic reprogramming events and cytosolic immune signalling (West *et al.* 2011; Cloonan and Choi 2012; Monlun *et al.* 2017). More recently, mitochondria have been described as regulators and activators of the epigenome where by-products of the citric acid cycle can modulate the conversion of 5-methylcytosine to 5-hydroxymethylcytosine, which, in turn, can catalyse the conversion of methylated DNA to demethylated DNA (Haseeb *et al.* 2014). In addition, mature mitochondria can modulate the activity of histone modifiers (Martínez-Reyes *et al.* 2016). Finally, mitochondria act as the vehicles responsible for protecting and transmitting mtDNA within the developing organism and into subsequent generations (Birky 1995).

Determining the origins of the mitochondrion has generated rigorous debate among evolutionary biologists and has been the subject of renewed vigour following the publication of Margulis's classical endosymbiotic theory (Sagan 1967). Endosymbiotic theory describes how mitochondria originated from bacteria and the coming together of two different organisms to live symbiotically (i.e. together; Archibald 2015). To this extent, the multiple functions that the mitochondrion plays in cell activity demonstrate the interdependence of the two entities to undertake complex functions. These would include,

polymorphic among individuals due to the high rate of mtDNA evolution (Sbisà *et al.* 1997). Indeed, variation in sequences, especially within the hypervariable regions of the non-recombining mtDNA genome, can be used to understand recent evolutionary history, phylogeography and maternal genealogies within and between species (Lirón *et al.* 2006). In addition, the control region of the D-loop, which houses the central conserved domain and conserved sequence boxes, is the site of interaction for the nuclear-encoded transcription and replication factors that migrate through the cell's cytoplasm to the mitochondrion to initiate transcription followed by replication of the mitochondrial genome (Clayton 1992, 1998; Kucej and Butow 2007).

The only exceptions to the maternal inheritance of mtDNA arise following interspecific breeding, where the spermatozoa and eggs arise from two different species. The earlier studies demonstrating the transmission of sperm mtDNA used the mouse as a model and showed that, for example, when *Mus musculus* and *Mus spretus* mice were mated, sperm mtDNA was transmitted to the offspring at low levels (<1%; Gyllenstein *et al.* 1991). Paternal transmission of mtDNA has also been described when different species of monkey are mated (St John and Schatten 2004). Nevertheless, the leakage of sperm mtDNA in vertebrates tends to be the exception to the rule rather than a frequent occurrence and has not been demonstrated in livestock species, apart from during early cattle embryo development following, for example, the crossing of domestic cow and the Asian wild gaur (Sutovsky *et al.* 1999). However, a full investigation is warranted to determine the extent of sperm mtDNA leakage in crossbreeds of different livestock species.

Interestingly, several mechanisms have been suggested to account for the elimination of sperm mtDNA (Sato and Sato 2013). Indeed, the contrasting views indicate that elimination can take place in the egg (Sutovsky *et al.* 1999; Nishimura *et al.* 2006; Sato and Sato 2011; Song *et al.* 2016) or before fertilisation (Yu *et al.* 2017). Mussels provide interesting comparisons whereby offspring can inherit male and female mitochondrial genomes (Kenchington *et al.* 2009). Indeed, it was previously argued that this was a sex-specific event, where blue mussels transmit both maternal- and paternal-type genomes to male offspring and only maternal genomes to female offspring (Zouros *et al.* 1992); however, there now appears to be less emphasis on sex-specific outcomes (Kenchington *et al.* 2009).

Assuming a hypothetical situation where biparental inheritance exists, it has been argued that the transmission of two distinct genomes could result in competition between these two genomes where mtDNA variants that had the potential to increase their rate of transmission through higher rates of replication would be favoured (Beekman *et al.* 2014). Indeed, it is further argued that this outcome could be at the expense of the organism's fitness. Therefore, the transmission of one genome ensures that the consequences of different genomes coexisting or one outcompeting another are avoided, and thus negates the effects of the selfish genome (Cosmides and Tooby 1981; Hurst and Hamilton 1992). Indeed, these effects are further reduced through oogenic mtDNA bottlenecks that lead to a filtering of mtDNA variants and mutations that would be present in the primordial germ cells of the next generation

(Stewart *et al.* 2008; Beekman *et al.* 2014). However, others have argued that the uniparental inheritance of mtDNA through the mother can have detrimental effects for the health and wellbeing of male offspring, otherwise described as the 'mother's curse' (Gemmell *et al.* 2004) or the 'sex-specific selective sieve' (Frank and Hurst 1996; Innocenti *et al.* 2011). In this instance, mutations that are advantageous to female offspring could be disadvantageous to male offspring, but, due to uniparental inheritance, will not be transmitted to subsequent generations by male offspring (Beekman *et al.* 2014).

mtDNA transcription and replication

mtDNA replication is preceded by mtDNA transcription, as it requires the generation of a DNA–RNA hybrid primer to facilitate the copying of mtDNA (Clayton 1992). As with mtDNA transcription, there are several nuclear-encoded mtDNA-specific replication factors that translocate to the mitochondrion (Kucej and Butow 2007). These factors are encoded by the mitochondrial transcription factor A (*TFAM*) gene (Larsson *et al.* 1998), the mitochondrial polymerase (DNA directed) gamma (*POLG*) gene (Hance *et al.* 2005), the Twinkle mtDNA helicase (*TWNK*) gene (Spelbrink *et al.* 2001; Korhonen *et al.* 2003), the mitochondrial single-stranded DNA-binding protein (*mtSSB*) gene (Farr *et al.* 2004; Ciesielski *et al.* 2015) and the DNA topoisomerase 1 mitochondrial (*TOP1MT*) gene (Zhang *et al.* 2007). Indeed, these factors are ever-present in the mitochondrion, where they interact with the mitochondrial genome to form a nucleoid structure, a combination of transcriptional and replication factors binding to one or more mitochondrial genomes in each mitochondrion (Kucej and Butow 2007; Bogenhagen 2012).

Interestingly, the mtDNA-specific replicase *POLG*, is differentially methylated at a cytosine–phosphorous–guanine (CpG) island within its second exon (Oakes *et al.* 2007). These changes to DNA methylation appear to coincide with the changes to mtDNA copy number during differentiation, not only in terms of the overall trend, but also in a tissue-specific manner (Facucho-Oliveira *et al.* 2007; Lee and St John 2015; Sun and St John 2016). For example, it has been shown that the levels of DNA methylation at exon 2 of *Polg* increased during neuronal differentiation of mouse embryonic stem cells (ESCs; Kelly *et al.* 2012). Importantly, the levels of DNA methylation at exon 2 of *Polg* presented in a tissue-specific manner and were negatively correlated with the corresponding transcriptional levels of *Polg* and mtDNA copy number (Kelly *et al.* 2012). During the differentiation of human ESCs into terminally differentiated astrocytes, negative correlations were also observed between the decreasing levels of DNA methylation at exon 2 of *POLG* and increasing levels of *POLG* expression and mtDNA copy number (Lee *et al.* 2015).

Nevertheless, cancer stem cell lines, such as HSR-GBM1 cells, which are indicative of glioblastoma multiforme, a Type VI brain tumour, failed to differentiate as the DNA methylation levels remained high at exon 2 of *POLG* (Lee *et al.* 2015). However, by using DNA demethylation agents, such as 5-azacytidine and vitamin C, the DNA methylation status at exon 2 of *POLG* was induced to decrease, which facilitates

mtDNA replication and promotes cell differentiation. In this respect, 5-azacytidine is a chemical analogue of cytosine that inhibits DNA methylation at low concentrations by blocking DNA methyltransferase (DNMT) 1, resulting in global DNA demethylation (Kelly *et al.* 2010), whilst vitamin C acts by enhancing the activity of the ten-eleven translocation (TET) 1 enzyme, and thus improves the conversion of 5-methylcytosine into the DNA-demethylated state of 5-hydroxymethylcytosine (Blaschke *et al.* 2013). Additionally, cancer cells with different loads of mtDNA vary in their levels of DNA methylation at exon 2 of *POLG*, which, in turn, results in significant differences in their tumorigenic capacity (Lee *et al.* 2015, 2016). Indeed, it has been shown that cancer cells with low mtDNA content have delayed tumour formation (Lee *et al.* 2015, 2016).

Furthermore, mtDNA copy number can be modulated by different mtDNA genotypes. In ESCs with the same nuclear genome background but with mtDNA from different genotypes, changes in mtDNA copy number were observed in combination with changes to chromosomal gene expression profiles and DNA methylation patterns at exon 2 of *POLG* (Kelly *et al.* 2013a). This has also been observed in tumour cells (Lee *et al.* 2015, 2016). These findings collectively highlight the impact of DNA methylation on modulating mtDNA copy number, but also the mitochondrial genome in modulating the DNA methylation patterns of the nuclear DNA. It is likely that both of these activities contribute significantly to successful cell differentiation and development.

Regulation of mtDNA copy number during development

It is becoming increasingly evident that mtDNA copy number is strictly regulated during development (see Fig. 3). During mouse oogenesis, mtDNA copy number is clonally expanded from approximately 200 copies present in the primordial germ

cells, the very first germ cells that are laid down, to greater than 100 000 copies in MII oocytes (Cao *et al.* 2007; Cree *et al.* 2008; Wai *et al.* 2008). These oocytes will have between one and two copies of mtDNA per mitochondrion (Pikó and Taylor 1987; Wai *et al.* 2008). However, in larger mammals, mtDNA content in developmentally competent MII oocytes comprises more than 150 000 copies, whereas oocytes that are less developmentally competent have fewer than 100 000 copies (Reynier *et al.* 2001; May-Panloup *et al.* 2005a; El Shourbagy *et al.* 2006; Santos *et al.* 2006; Takeo *et al.* 2013). Indeed, mtDNA-deficient, less developmentally competent oocytes are less likely to fertilise or, when they do, are more likely to arrest during preimplantation development (El Shourbagy *et al.* 2006; Spikings *et al.* 2007; Cagnone *et al.* 2016).

Since the expression of the nuclear-encoded mtDNA replication factors, encoded by *TFAM* and *POLG*, does not take place during preimplantation embryo development, as demonstrated in mouse (Thundathil *et al.* 2005), cattle (May-Panloup *et al.* 2005b) and pig (Spikings *et al.* 2007; Cagnone *et al.* 2016) embryos, apart from a minor replication event that can occur just before embryonic genome activation (McConnell and Petrie 2004; Spikings *et al.* 2007; Cagnone *et al.* 2016), the population of mtDNA that is present at fertilisation is an investment in subsequent developmental outcomes (May-Panloup *et al.* 2005a; El Shourbagy *et al.* 2006; Santos *et al.* 2006; Spikings *et al.* 2007; Cree *et al.* 2008; Takeo *et al.* 2013; Cagnone *et al.* 2016). In many cases, oocytes that have fewer than 150 000 copies of mtDNA often fail to fertilise or, when they fertilise, the resultant embryos arrest during preimplantation development (Spikings *et al.* 2007; Cagnone *et al.* 2016). This effect has also been demonstrated experimentally, where developmentally competent oocytes have been treated during IVM with an mtDNA-depletion agent to reduce mtDNA copy number with the resulting MII oocytes either failing to fertilise or arresting

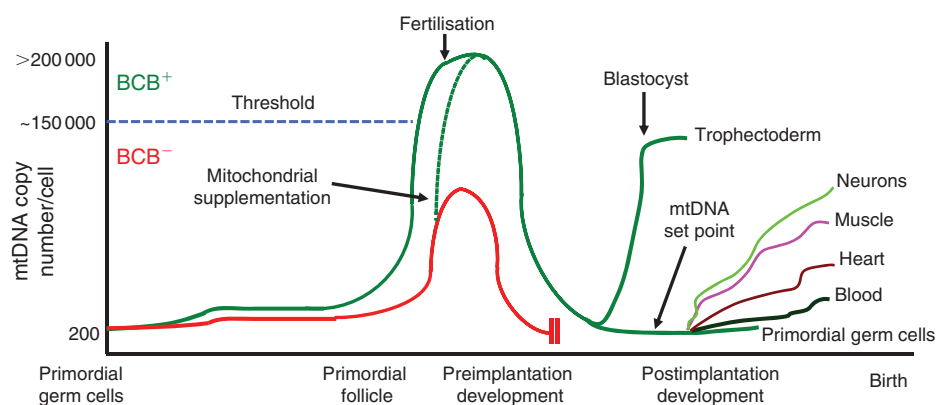


Fig. 3. Regulation of mitochondrial (mt) DNA copy numbers during development. During oogenesis, mtDNA copy numbers peak at fertilisation (green line), decreasing during preimplantation development until the blastocyst stage when replication is initiated in the trophoblast only. The inner cell mass cells continue to reduce mtDNA copy number and establish the 'mtDNA set-point', which is essential for mature cells to acquire the required number of mtDNA copies. Failure of the maturing oocyte to increase mtDNA copy numbers above the threshold (dotted blue line; BCB⁺ (Brilliant Cresyl Blue) oocytes) can result in fertilisation failure or early embryo arrest (red line). Oocytes with low mtDNA copy numbers (BCB⁻ oocytes) can be rescued by supplementing these oocytes with autologous populations of mitochondria as intracytoplasmic sperm injection is performed to fertilise the mature oocyte (superimposed dashed green line).

during preimplantation development (Spikings *et al.* 2007). This investment is especially important, as mtDNA copy is reduced after each stage of embryo cleavage as new blastomeres form (El Shourbagy *et al.* 2006). Moreover, the overall mtDNA content is actively reduced during preimplantation development in larger mammals (May-Panloup *et al.* 2005b; Spikings *et al.* 2007), which has been proposed to arise from the active extrusion of mtDNA from the embryo, as demonstrated in human embryos cultured *in vitro* (Stigliani *et al.* 2014).

During early development, mtDNA replication is first initiated at the late morula/blastocyst stage (May-Panloup *et al.* 2005b; Spikings *et al.* 2007; Cagnone *et al.* 2016; see Fig. 3), whereby mtDNA copy number increases in the trophectodermal cells in response to a greater requirement for energy as the naïve trophectodermal cells differentiate into the more mature type cells that give rise to the placenta (Houghton 2006). However, the inner cell mass cells, which give rise to the embryo proper, continue to dilute out their mtDNA content as replication remains restricted to establish the mtDNA set-point (Spikings *et al.* 2007; Kelly *et al.* 2012; St John 2012; see Fig. 3). The mtDNA set-point establishes a founder population of mtDNA in naïve cells (Lee and St John 2015; Lee *et al.* 2015; Sun and St John 2016), which is similar to the mitochondrial genetic bottleneck that has been proposed to act as a purifying filter to select which mtDNA molecules are segregated to the primordial germ cells (Jansen and de Boer 1998; Cree *et al.* 2008; Wai *et al.* 2008).

The mtDNA set-point is characterised by the persistence of low levels of mtDNA copy number that remain low as cells persist in an undifferentiated state and before undergoing clonal amplification during cellular differentiation (Facucho-Oliveira *et al.* 2007; St John 2012; Sun and St John 2016). As a result, highly proliferative cells, including ESCs and cancerous cells (Vander Heiden *et al.* 2009), primarily utilise glucose metabolism, and this is promoted by having a low mtDNA copy number, which ensures that they cannot effectively use oxidative phosphorylation (Facucho-Oliveira *et al.* 2007; Kelly *et al.* 2012; Krisher and Prather 2012; Lee *et al.* 2015). Indeed, parallels can be drawn between preimplantation embryos and tumour-initiating cells, as both populations of cells have the propensity to use aerobic or anaerobic glycolysis, otherwise described as the Warburg effect (Warburg 1956a, 1956b), to produce various metabolites that are necessary to support significant increases in cell division that are essential at their respective developmental stages (Vander Heiden *et al.* 2009; Krisher and Prather 2012; St John 2014).

As pluripotent ESCs expressing the key markers of the pluripotent network, such as octamer-binding transcription factor 4 (*OCT4*), SRY (sex determining region Y)-box 2 (*SOX2*) and Nanog Homeobox (*NANOG*) (Wray *et al.* 2010), differentiate into mature cell types of varying kinds during organogenesis, mtDNA is replicated to meet their respective requirements for the generation of ATP through oxidative phosphorylation (Facucho-Oliveira *et al.* 2007; see Fig. 3). Indeed, disrupting the steady state relationship between *POLGA* and pluripotency results in the onset of cellular differentiation (Facucho-Oliveira *et al.* 2007) and, before this in the embryo, the ability to regulate mtDNA copy number (Bowles *et al.* 2007b).

For example, cardiac and skeletal muscle cells have high demands for oxidative phosphorylation-derived ATP, reflected in their high levels of mitochondrial mass and respiratory capacity (D'Erchia *et al.* 2015) and high numbers of mtDNA copies per cell (in the range of thousands of copies; Pohjoismäki *et al.* 2010), with multiple copies of mtDNA per mitochondrion that are linked through nucleoid structures (Legros *et al.* 2004), to support their specialised functions. On the other hand, spermatozoa (Amaral *et al.* 2007) and spleen cells (Kelly *et al.* 2012) possess low mtDNA copy numbers (in the range of tens of copies) and thus have fewer copies of the mitochondrial genome per mitochondrion as they primarily rely on glycolysis to generate ATP (Leary *et al.* 1998; Moyes *et al.* 1998; St John 2012). Indeed, it appears that spermatozoa have only one copy of the mitochondrial genome per mitochondrion (Hecht *et al.* 1984). Tissue-specific regulation of mtDNA copy number is modulated by DNA methylation and expression of *POLG* (Kelly *et al.* 2012).

Meeting the energy demands of each specialised cell type is crucial, as failure to do so can lead to many complex diseases, such as metabolic and age-related degenerative diseases (Wallace 2010). To this extent, induced pluripotent stem cells and ESCs generated through SCNT that have not established the mtDNA set-point are unable to regulate mtDNA copy number effectively and frequently fail to complete differentiation (Kelly *et al.* 2013b), as is the case for tumour-initiating cells (Dickinson *et al.* 2013). This arises as both sets of cells are hypermethylated and the only manner in which they can effectively regulate their mtDNA copy number and differentiation is through induced DNA demethylation, which can be mediated by agents such as 5-azacytidine and vitamin C (Kelly *et al.* 2012, 2013b; Lee *et al.* 2015).

Why is mtDNA important to fertilisation outcome?

The investment hypothesis

Oocytes with low levels of mtDNA copy number frequently fail to fertilise or they arrest during early embryonic development. In a pig model, it has been demonstrated that mtDNA copy number below a threshold of 150 000 copies of mtDNA, namely mtDNA deficiency, results in fertilisation failure or arrest during preimplantation development (May-Panloup *et al.* 2005a; El Shourbagy *et al.* 2006; Santos *et al.* 2006; Spikings *et al.* 2007; Cagnone *et al.* 2016). However, low mtDNA copy number appears to be more detrimental to embryo development than fertilisation potential in small animals such as mice (Wai *et al.* 2010). In addition, mtDNA deficiency has been associated with ageing where, for example, younger women have oocytes with larger cytoplasmic volume (<35 vs ≥35 years) and mtDNA copy number (<40 vs ≥40 years) compared with older women (Murakoshi *et al.* 2013). mtDNA deficiency has also been associated with several reproductive disorders in humans, including mild endometriosis (Xu *et al.* 2015) and diminishing ovarian reserve (Boucret *et al.* 2015).

Nevertheless, those oocytes that have greater than 150 000 copies of mtDNA fertilise and frequently give rise to good-quality embryos that have the potential to implant and give rise to live offspring. Given that the primary role of mitochondria is

to produce energy in the form of ATP, it might make sense to assume that the relationship between mtDNA copy number and developmental capacity in the oocyte is one of energetics. Higher ATP production would mean more energy available for the major cellular reorganisation events that occur during early pre- and postimplantation embryo development. However, it is unlikely that mtDNA copy number is indicative of the metabolic capacity of the oocyte, particularly as the mitochondria present within the mature oocyte are immature and metabolically quiescent (Baca and Zamboni 1967; Leese *et al.* 2007). Indeed, mtDNA-deficient oocytes appear to have similar numbers of mitochondria but significantly fewer copies of mtDNA (Cagnone *et al.* 2016), suggesting that this is an mtDNA issue and not a mitochondrial issue. In fact, the energetic needs during preimplantation development shift to the glycolytic pathways (Flood and Wiebold 1988; Krisher and Prather 2012). Indeed, the inhibition of oxidative phosphorylation-derived ATP through the use of the classical inhibitors leads to increased cell division, and thus increased cell numbers, in porcine embryos (Macháty *et al.* 2001). Instead, it is likely that sufficient mtDNA copy numbers are essential for the oocyte to by-pass developmental checkpoints and undergo subsequent cell divisions. Indeed, during ESC differentiation, there is a significant replication event that takes place at Day 6 of differentiation, which equates to Embryonic Day (E) 7.5 in mouse embryonic development (Leahy *et al.* 1999), when *Polg* homozygous knockout mice die *in utero* (Hance *et al.* 2005).

Ooplasmic (cytoplasmic) transfer: the importance of maintaining homoplasmic transmission

Since the maintenance of sufficient mtDNA copy numbers is necessary for fertility, it is unsurprising that a great deal of attention has been paid to altering the mtDNA population *in vitro* to enhance embryo development. Fortunately, in many species, including livestock species, the oocyte is a single cell accessible to manipulation via microinjection, a technique routinely performed in fertility clinics with, for example, intracytoplasmic sperm injection (ICSI). In humans with a history of IVF failure or male partners with poor sperm motility and/or concentration, a single spermatozoon is microinjected directly into the oocyte. Likewise, this technique can be employed in the livestock industry, where it seeks to propagate key genetic traits in animals exhibiting fertility problems where standard breeding approaches through AI fail (Seidel 2015; Tang *et al.* 2015). However, although current ARTs, such as IVF and ICSI, are effective methods for overcoming poor sperm quality, they are not effective for overcoming poor oocyte quality (Tannus *et al.* 2017).

If having a sufficient mtDNA copy number is necessary for successful fertilisation and development outcome, it would be appropriate to supplement mtDNA-deficient oocytes with mitochondria while performing ICSI, as was the perceived case with ooplasmic transfer (Cohen *et al.* 1997). In this instance, Cohen *et al.* (1997) removed a portion of the cytoplasm from an oocyte from a young female donor's oocyte and transferred it into the patient's oocyte during ICSI. Although the cytoplasm also contained mRNAs, proteins and other cytosolic factors, it

was strongly suspected that mitochondria were the main candidates mediating the enhanced outcome (Jansen 2000). Out of the 30 procedures performed, 13 pregnancies were recorded with 16 live births (Barritt *et al.* 2001a). Several other groups also began performing cytoplasmic transfer, albeit with small numbers, but the results appeared promising (Huang *et al.* 1999; Lanzendorf *et al.* 1999).

Unfortunately, as a consequence of cytoplasmic transfer, several of the offspring displayed varying levels of mitochondrial heteroplasmy, namely the coexistence of donor and recipient oocyte mtDNA in the offspring's cells and tissues (Brenner *et al.* 2000). Indeed, this form of heteroplasmy is thought to lead to potentially deleterious health effects (Acton *et al.* 2007; Sharpley *et al.* 2012), although a limited recent follow-up study suggests that the children born from cytoplasmic transfer are normal (Chen *et al.* 2016). The outcomes from the study by Chen *et al.* (2016) must be carefully interpreted given the limited number of criteria assessed. Furthermore, this procedure was the first known instance of germline therapy, as mitochondrial carry over during cytoplasmic transfer has the potential to be transmitted through the germline and be passed on to future generations. Safety concerns were also raised as two of the embryos exhibited 45,XO karyotypes (Barritt *et al.* 2001a). Indeed, these outcomes contributed to the banning of cytoplasmic or ooplasmic transfer in many countries, including the USA, UK and Australia.

An alternative is to transfer mitochondria using an autologous approach (El Shourbagy *et al.* 2006). In this instance, supplementation was performed using mitochondria isolated from sister oocytes obtained from the same ovary using a porcine model, which ensures that mitochondrial genetic integrity can be maintained. Autologous mitochondrial supplementation appears to be effective at rescuing developmentally incompetent porcine oocytes with low mtDNA copy numbers (El Shourbagy *et al.* 2006; Cagnone *et al.* 2016; see Figs 3, 4), as

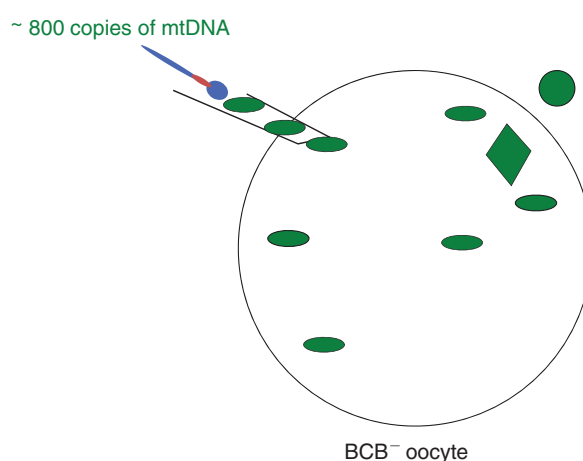


Fig. 4. Autologous mitochondrial supplementation into developmentally incompetent oocytes. Mitochondrial (mt) DNA-deficient oocytes (BCB⁻ (Brilliant Cresyl Blue)) are supplemented with approximately 800 copies of mtDNA as intracytoplasmic sperm injection is performed to produce embryos, fetuses and offspring.

assessed using brilliant cresyl blue, a dye that selects between developmentally competent and developmentally incompetent oocytes by assaying glucose-6-phosphate dehydrogenase (G6PDH) activity at the start of *in vitro* culture (Roca *et al.* 1998), as also demonstrated in mouse (Wu *et al.* 2007), cattle (Pujol *et al.* 2004; Bhojwani *et al.* 2007) and goat (Rodríguez-González *et al.* 2002) oocytes. Not only did this process improve the blastocyst development rates derived from these rescued oocytes, the resultant gene expression profiles more closely resembled those present in blastocysts derived from normal oocytes (Cagnone *et al.* 2016). Others have performed autologous supplementation with mitochondria derived from cumulus cells (Tzeng *et al.* 2004) and oogonal stem cells (Fakih *et al.* 2015) to produce embryos and children; induced pluripotent stem cells (Li *et al.* 2017) to produce mouse embryos; and granulosa cells to produce bovine embryos (Hua *et al.* 2007). However, the suitability of these sources of mitochondria for embryo production still needs to be determined. The population of mtDNA present in oocytes will have been through the mtDNA genetic bottleneck and will therefore possess a population of mtDNA that has been refined, where certain variants will have been selected for or against (Jenuth *et al.* 1997). A full investigation of how other cell types regulate the selection of and against specific variants needs to be conducted before they could be used for supplementation purposes.

Interestingly, the supplementation of developmentally incompetent porcine oocytes with mitochondria results in a significant increase in mtDNA copy number by the 2-cell stage embryo, an increase of more than fourfold over the population present in the developmentally incompetent oocyte (Cagnone *et al.* 2016), based on an initial investment of just 800 copies of mtDNA introduced into the egg at the time of fertilisation (Cagnone *et al.* 2016). Although the exact mechanism by which this occurs remains to be elucidated, a mitochondrial replication event occurs between fertilisation and the first cell division, as previously observed in mice (McConnell and Petrie 2004), cattle (May-Panloup *et al.* 2005b) and pigs (Spikings *et al.* 2007; Cagnone *et al.* 2016). It may be that the supplementation of mitochondria in the oocyte signals the nuclear genome to initiate embryonic genome activation, thereby rescuing the embryo from early developmental arrest due to low mtDNA copy number. In this respect, gene expression analysis of blastocysts derived using this technology demonstrated that these blastocysts appear to be far healthier in terms of the genes they express (Cagnone *et al.* 2016). To this extent, they express key genes associated with cell morphology networks, cell development and amino acid metabolism, which are essential for good embryo development.

There are numerous benefits to autologous mitochondrial supplementation over its heterologous counterpart. Namely, using autologous supplementation from germline tissue will not result in the offspring inheriting two populations of mtDNA. Any potential mismatch between the mitochondrial and nuclear genome is also avoided. Furthermore, many of the public concerns regarding so-called 'three-parent' babies are avoided as the mtDNA is exclusively derived from the mother. Finally, although it cannot be ruled out that autologous mitochondrial supplementation may carry transgenerational epigenetic effects,

there is significantly less risk that this procedure may have negative health implications for the germline and future generations.

The fact that successful mtDNA supplementation is highly dependent upon an mtDNA replication event that takes place between fertilisation and the 2-cell stage most likely explains why the cytoplasm from a young donor's oocyte could contribute to the frequent occurrence of heteroplasmy and to a heteroplasmic load that is as high as 40% of the offspring's mtDNA content (Brenner *et al.* 2000). However, when more divergent populations of mtDNA coexist, such as with intergeneric and interspecies mitochondrial injection and parthenogenetic activation, there is no selection for or against the transferred population of mtDNA at the blastocyst stage (Takeda *et al.* 2012). This demonstrates the potential power of the replication event taking place just after fertilisation. Therefore, for the maintenance of the genetic integrity of livestock and to ensure that the resultant offspring do not suffer from the effects of harbouring two divergent populations of mtDNA coexisting within tissues and cells, autologous mitochondrial transfer is a far more desirable approach (Spikings *et al.* 2006; St John 2014). Importantly, mtDNA contributes to the genetic integrity of an individual and maintenance of uniparental-only inheritance is an important evolutionary mammalian trait (Birky 1995). It is unlike chromosomal inheritance, which utilises cross-over events post-fertilisation to determine the inheritance of each allele, thus making the mitochondrial genome the most stability inherited genome in mammalian cells.

Other ARTs with the propensity to aberrantly regulate mtDNA

It appears that the more sophisticated an ART, the greater the potential for irregular patterns of mtDNA transmission (St John *et al.* 2010). A classic example is SCNT, otherwise known as cloning, which has been effectively used in several species, including sheep (Wilmut *et al.* 1997), cattle (Cibelli *et al.* 1998), mice (Wakayama *et al.* 1998), pigs (Polejaeva *et al.* 2000), goats (Keefer *et al.* 2001), cats (Shin *et al.* 2002), rabbits (Challah-Jacques *et al.* 2003) and horses (Galli *et al.* 2003). As documented by many investigators, SCNT involves the transfer of a somatic cell into an enucleated recipient oocyte, which, following activation, undergoes embryonic division and can then be transferred to a surrogate in order to establish pregnancy and live offspring (Wilmut *et al.* 1997; see Fig. 5). This technology has significant power as it enables animals to be produced, and even reproduced, with desired genetic traits. For breeders in the livestock industry, this can empower them to maintain lines or breeds of animals with enhanced traits associated with, for example, fertility, meat and milk quality (Bowles *et al.* 2007a; Watanabe and Nagai 2008).

However, as with ooplasmic transfer, where donor cytoplasm is transferred, the resultant offspring can harbour two genetically divergent populations of mtDNA, as the donor cells bring accompanying mtDNA into the recipient oocyte that can then be transmitted through the embryo into the fetus and onto the resultant offspring (Steinborn *et al.* 2002; Hiendleder *et al.* 2003; Takeda *et al.* 2003; Fig. 5). Although only 800–2000

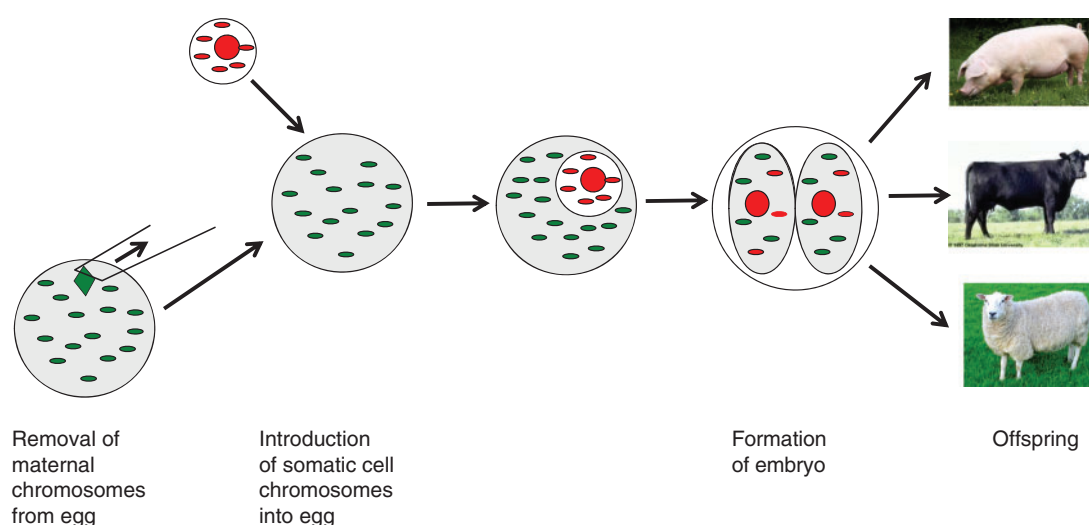


Fig. 5. Somatic cell nuclear transfer. A mature cell, such as a skin cell, is introduced into an oocyte that retains its mitochondrial (mt) DNA but has had its chromosomes removed (enucleation). Once the reconstructed oocyte is activated, an embryo can form, which can then be transferred to a surrogate for implantation and fetal development. Whilst the technology has produced a range of livestock species, many offspring harbour two genetically distinct populations of mtDNA.

copies of mtDNA are introduced into the recipient oocyte with the donor cell, which is not dissimilar to the amount introduced with ooplasmic or mitochondrial supplementation (Cagnone *et al.* 2016), it is likely that the same mtDNA replication event is responsible for propagating the donor cell's mtDNA. In addition, it appears that the mtDNA replication factor machinery is not switched off during preimplantation development following SCNT (Lloyd *et al.* 2006; Bowles *et al.* 2007b). In cattle, this can result in the offspring possessing 0–59% of its total mtDNA originating from the donor cell population (Steinborn *et al.* 2002; Hiendleder *et al.* 2003; Takeda *et al.* 2003; Chiaratti *et al.* 2010; Hammond *et al.* 2016), and up to 44% in pigs (Takeda *et al.* 2006) and 46.5% in sheep (Burgstaller *et al.* 2007).

The transmission of donor cell mtDNA is further promoted by the persistence of the donor cell's mtDNA replication machinery, including POLGA and TFAM, which is still active during preimplantation development following SCNT (Lloyd *et al.* 2006; Bowles *et al.* 2007b). Prior to SCNT, the donor cells actively divide and replicate their mtDNA. Even when serum-starved cells are used, which is frequently the case before SCNT, these cells maintain the potential to replicate their mtDNA. This occurs as mtDNA replication is not cell cycle dependent and is semi-autonomous to nuclear events. Previously, there had been much debate as to whether mtDNA heteroplasmy was simply an unimportant by-product of SCNT or whether it would impact on embryo and fetal development and offspring health and wellbeing. However, documented evidence clearly suggests that the presence of two populations of mtDNA can have serious phenotypic implications for the offspring's health and wellbeing (Barritt *et al.* 2001b; Sharpley *et al.* 2012). Indeed, reports on the first wave of cloned offspring generated highlighted their predisposition to pathologies associated with severe organ and tissue dysfunction (Cibelli *et al.* 2002). Close examination of

these disorders shows remarkable similarity to the mitochondrial diseases associated with pathogenic mtDNA mutations and deletions, suggesting that the potential heteroplasmic nature of the mitochondrial genotypes of these offspring was a contributing factor.

To prevent the transmission of donor cell mtDNA to offspring and their descendants inheriting two populations of mtDNA, donor cells can be depleted of their mtDNA before performing SCNT (see Fig. 6). In other words, an mtDNA-depleted (mtDNA⁰) cell can be directly transferred to the enucleated recipient oocyte and fused, as is normal practise for SCNT (Lloyd *et al.* 2006; Bowles *et al.* 2007a; Srirattana and St John 2017). Cells can be cultured using an mtDNA-depletion agent, commonly ethidium bromide (Desjardins *et al.* 1985; Hayakawa *et al.* 1998; Lloyd *et al.* 2006; Lee *et al.* 2010) or 2,3-dideoxycytidine (Brinkman and Kakuda 2000; Srirattana and St John 2017). Whilst ethidium bromide is potentially toxic (Waring 1965), the actions of 2,3-dideoxycytidine are specific to POLG (Brinkman *et al.* 1998). There are three distinct advantages to using mtDNA⁰ cells. First, the mtDNA content in donor cells is completely depleted and, after 30 days of culture, the donor cells do not have the propensity to replicate mtDNA. Second, as a consequence of depletion during culture, the cells reduce their division rates and are therefore not actively dividing when they are fused to the recipient egg. Third, and just as importantly, the depletion process does not affect the integrity of the chromosomal genome (Srirattana and St John 2017).

Embryos derived from depleted cells can progress to the blastocyst stage and possess recipient oocyte mtDNA only (Lloyd *et al.* 2006; Lee *et al.* 2010; Srirattana and St John 2017). Furthermore, since the resultant offspring possess oocyte-only mtDNA, they thus inherit their mtDNA in a manner similar to offspring derived from natural fertilisation, as demonstrated in a sheep model (Lee *et al.* 2010). Indeed, these sheep

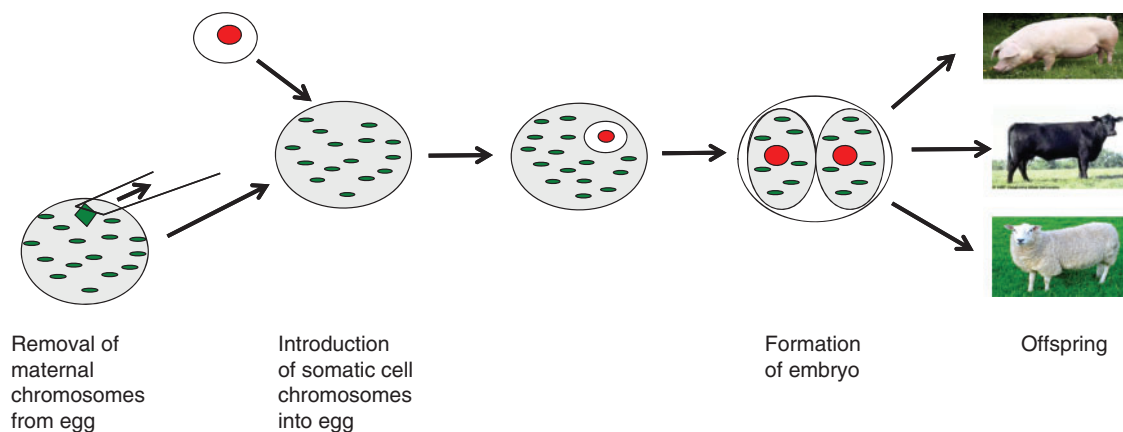


Fig. 6. Somatic cell nuclear transfer using a mitochondrial (mt) DNA-depleted cell (mtDNA⁰-SCNT). Donor cell mtDNA is eliminated using an mtDNA-depletion agent, such as 2,3-dideoxycytidine. The donor cell is transferred into an enucleated recipient oocyte and the reconstructed oocyte is activated. The resultant embryo can be transferred to a surrogate to generate live offspring. This technology can be used to generate a range of livestock species that have mtDNA transmitted from the recipient oocyte only.

are the only genetically identical animals that have been specifically generated to be truly genetically identical. Furthermore, embryos derived using mtDNA⁰-SCNT frequently have increased cell numbers at the blastocyst stage and their levels of mtDNA copy numbers are representative of those observed in IVF-derived embryos (Srirattana and St John 2017). This form of regulation appears to be an important developmental milestone as mtDNA copy number can exceed a threshold in the trophoblastic cells of the blastocyst, which has been associated with chromosomal aneuploidy and failed implantation rates in the human (Fragouli *et al.* 2015).

In addition, the gene expression profiles of blastocysts derived from depleted cells were different compared with those of non-depleted cells (Srirattana and St John 2017). These affected genes are primarily involved in embryonic development. Furthermore, the addition of trichostatin A (TSA), a histone deacetylase inhibitor, to the culture media increased the blastocyst formation rate of embryos derived from depleted cells. When blastocysts generated from depleted cells and non-depleted cells, both cultured in the presence of TSA, were compared, there was several differentially expressed genes. In this comparison, the only difference between these two sets of blastocysts was the presence of donor cell mtDNA when SCNT was performed, resulting in a significantly higher mtDNA copy number at the blastocyst stage for the non-depleted cells. Furthermore, the use of TSA on embryos derived from depleted cells positively modulated expression levels of genes involved in embryonic development, the inner cell mass and DNA methylation (Srirattana and St. John 2017). Consequently, it appears that the use of depleted cells in combination with a reprogramming agent aids developmental outcome.

Effects of maternally inherited mtDNA on male fertility

Since the mitochondrial genome is inherited through the female germline, there should be no apparent direct implications for the offspring following fertilisation in terms of the inheritance of

paternal mtDNA, apart from those hypothetical situations discussed above. However, this does not exclude the effects of male offspring inheriting deleterious mtDNA from their mothers that can affect their health and wellbeing, as indicated by the 'mother's curse' phenomenon (Gemmell *et al.* 2004; Innocenti *et al.* 2011). Nevertheless, as mtDNA is susceptible to mutation and deletion, there are consequences for male fertility and specifically sperm quality (St John *et al.* 2000).

Several studies in human patients attending infertility clinics have shown a correlation between increased levels of single (Kao *et al.* 1995) and multiple (Lestienne *et al.* 1997; St John *et al.* 2001; Hosseinzadeh Colagar and Karimi 2014) large-scale mtDNA deletions and poor sperm quality. Equally so, there are associations with point mutations and sperm motility (Spiropoulos *et al.* 2002). Although controversial, as it appears that early studies indicated that oxidative phosphorylation was not implicated in bovine (Storey 1980) and rabbit (Storey and Kayne 1975) sperm motility, biochemical analysis of human spermatozoa suggests a direct effect of oxidative phosphorylation on sperm motility (Folgerø *et al.* 1993; St John *et al.* 2005a), especially in the context of substrate availability in the female reproductive tract at the time of ovulation (St John *et al.* 2005a).

Poor sperm quality is not limited to mutations and deletions. Whilst mtDNA copy number increases during oogenesis so that the fertilisable oocyte possesses sufficient copies of mtDNA as an investment in post-fertilisation events, the opposite is very much observed during the later stages of spermatogenesis, with a reduction in mtDNA copy as round spermatids mature into elongated spermatids and spermatozoa (Hecht *et al.* 1984), which is partly mediated by the expression of a testis-specific isoform of the gene that encodes *TFAM* that does not possess the mitochondrial targeting sequence (Larsson *et al.* 1996). Indeed, it appears that good-quality sperm possess fewer copies of the mitochondrial genome than poor-quality spermatozoa (Diez-Sánchez *et al.* 2003; May-Panloup *et al.* 2003; Amaral *et al.* 2007). The reduced mtDNA content has two implications for successful fertilisation outcome. First, a few good molecules are

preserved for energy production; second, the risk of leakage of sperm mtDNA is significantly reduced as there would be fewer molecules to target post-fertilisation in the embryo. Consequently, as predicted in many model systems and mammals, the transmission of a mother's mtDNA to her son has implications for her son's fertility (St John *et al.* 1997; Beekman *et al.* 2014).

mtDNA haplotypes

As mtDNA haplotypes have evolved over time, they have become defined by a common set of mtDNA sequences that cluster following alignment with bioinformatics tools (Hasegawa *et al.* 1985). Indeed, the mitochondrial genome appears to be more susceptible to mutation than the nuclear genome, and evolves at a faster rate than the single-copy fraction of the nuclear genome by a factor of around 10 (Brown *et al.* 1979). As mtDNA is maternally inherited, these clusterings of mtDNA sequences can define common maternal origins (Brown *et al.* 1979; Coskun *et al.* 2003). Within a species, it is thought that all individuals originated from one or several 'mtDNA Eves'. As an individual's ancestors migrated from the location of their common ancestor, the changing environments could have resulted in selection, whether natural or artificial, to prefer different sequence variants (Ruiz-Pesini *et al.* 2004). These sequence variants could have direct functional effects on phenotypes as sequence variants can be selected for at a population level to promote functional effects, such as the ability to tolerate heat. As a result, an individual's mtDNA haplotype is indicative of the environmental responses that enable groupings with similar mtDNA sequences to survive and adapt to different influences (Gerber *et al.* 2001; Innocenti *et al.* 2011). Consequently, mtDNA haplotypes have the potential to predispose and protect individuals against several diseases (for a review, see Tsai and St John 2016). In humans, these can include aging disorders, such as cancer (Kaiparettu *et al.* 2010) and diabetes (Hwang *et al.* 2011), and adaptation to high altitude (Ji *et al.* 2012). They can also influence growth and physical performance (Nagao *et al.* 1998), longevity (Tanaka *et al.* 1998), survival in various climatic conditions (Ruiz-Pesini *et al.* 2004) and lipid and glucose metabolism (Latorre-Pellicer *et al.* 2016) in a variety of species. In cattle, mtDNA haplotypes have been associated with milk quality (Schutz *et al.* 1994), meat quality (Mannen *et al.* 2003) and fertility in cows (Sutarno *et al.* 2002; Tamassia *et al.* 2004; Srirattana *et al.* 2017); in pigs, there are associations between mtDNA haplotypes and litter size and reproductive capacity (Tsai *et al.* 2016). Furthermore, mtDNA haplotypes have been shown to influence developmental outcomes associated with certain ARTs, such as SCNT and IVF (Brüggerhoff *et al.* 2002; Bowles *et al.* 2008; Srirattana *et al.* 2017). Below, we focus on how mtDNA haplotypes influence reproductive function and capacity using examples from Australian commercial livestock breeding programs.

Implications of mtDNA haplotypes for pig breeding

Within Australian commercial breeding pigs, there are five different mtDNA haplotypes (Tsai *et al.* 2016), of which three are of Asian origin and two are from Europe. None of the five

mtDNA haplotypes is indicative of a specific breed, with each of the mtDNA haplotypes comprising pigs with phenotypes typical of Large White, Landrace and Duroc pigs. During early pig domestication, although gene flow and admixture took place between early domestic animals and from wild local populations before their recognisable breeds were established (Frantz *et al.* 2013, 2015), the selection of desirable chromosomal-driven phenotypes could have rapidly restricted the mtDNA haplotypes in the population. Similarly, the distribution of haplotypes across various breeds demonstrates how modern breeding strategies have been driven by the distribution of semen throughout Australia, with enhancement of phenotypic traits being favoured through the male lineage, thereby potentially inadvertently eliminating certain mtDNA haplotypes in future generations. Nevertheless, breeding sows have continued to determine inheritance of specific mtDNA haplotypes, which will persist on an ongoing basis through future generations irrespective of the breeding strategy employed. Consequently, the only constant genetic constituent in many breeding lines is the mitochondrial genome, as it is continually passed down the female germline unless technologies such as cytoplasmic transfer or SCNT are used, resulting in the loss of homoplasmic transmission (St John *et al.* 2005b; Takeda *et al.* 2006).

In Australian commercial pigs, the five different mtDNA haplotypes have very different reproductive capacities (Tsai *et al.* 2016). For example, there are significant differences in litter size among the haplotypes across multiple generations. Furthermore, females from three of the haplotypes are significantly more likely to have litters ≥ 15 across three parities than the other two haplotypes. In addition, each of the haplotypes is either more or less efficient at converting developmentally competent oocytes, MII oocytes, fertilised oocytes and blastocysts to live offspring, with differences between each haplotype at each developmental stage. In this respect, two of the haplotypes were determined to be more efficient than the others at each of these developmental stages. Furthermore, each of the haplotypes had significantly different levels of mtDNA copy number in their oocytes, which, again, is indicative of the level of developmental competence, fertilisation outcome and development to the blastocyst stage (El Shourbagy *et al.* 2006; Spikings *et al.* 2007; Cagnone *et al.* 2016).

Each of the mtDNA haplotypes also exhibited key DNA signatures that impacted on their transcriptional and replicative efficiency, and thus the amount of mtDNA copy numbers present in their oocytes at the time of fertilisation. This was especially the case for a variant at 16383delC that was positively correlated with mtDNA copy number and is present in the highly conserved sequence box II. Termination of transcription at conserved sequence box II results in the generation of the RNA–DNA hybrid primer that is required for mtDNA replication, as this transcript forms a G-quadruplex structure to interact with the mitochondrial-specific transcription elongation factor and thus determines whether transcription continues or replication is initiated (Wanrooij *et al.* 2010; Agaronyan *et al.* 2015). Consequently, this direct link between mtDNA haplotype, mtDNA copy number and reproductive strategy could also impact on other common maternally associated reproductive traits, such as time to pregnancy, weaning efficiency, lactation

efficiency and gestational length, although these traits in association with mtDNA haplotypes remain to be investigated. However, pigs with haplotypes that have lower levels of fertility are likely to exhibit other traits, otherwise they probably would not survive as they would be deemed commercially unimportant. For example, such pigs could be indicative of better meat quality or other characteristics associated with high economic breeding values and, as producers infrequently consider mtDNA haplotypes when assessing economic breeding values, they would have been maintained through other criteria.

Implications of mtDNA haplotypes for cattle breeding

Typically, domestic cattle fall into two subspecies, namely *Bos taurus* and *Bos indicus*, which are phenotypically distinct. It is generally understood that these two subspecies originated from *Bos primigenius*, an extinct wild ox that evolved approximately 500 000 years before present, as discussed in Magee *et al.* (2014). Indicine cattle are readily identifiable through a muscular fatty hump located in either the shoulder or neck regions, a more prominent dewlap, droopy ears and a naval flap. In addition, they have a predisposition to tolerate hot and dry environments and exhibit a lower basal metabolic rate. On the other hand, native African taurine cattle have the ability to withstand a wasting disease, namely trypanosomiasis, which is transmitted through the tsetse fly and is a major pest for many cattle (Magee *et al.* 2014). Admixed cattle are a combination of both taurine and indicine cattle, and exhibit a combination of taurine and indicine traits, which results from the interbreeding of two or more populations that were previously isolated and, as a consequence, introduces new genetic lineages into a population (Magee *et al.* 2014). Admixture is likely preceded by introgression, where the transfer of genetic information from one species to another results from hybridisation and repeated back-crossing. These breeding patterns demonstrate the complexity that exists in understanding the origins of domesticated cattle, as exhibited in north-east Asian cattle (Bradley *et al.* 1998) and demonstrated in a study of Japanese cattle, which, through the analysis of their mtDNA sequences, suggested that the populations present arose through three different strains of aurochs (Mannen *et al.* 1998).

In general terms, taurine cattle are centred around northern and western Africa and Eurasia, which ranges from north-western Europe to Japan. On the other hand, indicine cattle originate from India, the Near and Middle East and eastern and southern Africa. Their origins will be reflected in their genetic profiles, which, in turn, reflect their migration patterns. Whilst both *B. taurus* and *B. indicus* cattle have the same chromosomal number, they exhibit distinguishing Y chromosome morphologies, as discussed in Magee *et al.* (2014). Their chromosomal number enables recombination of their nuclear DNA and thus allows effective cross-breeding. From an analysis of *B. indicus* Y chromosome introgression and admixture in *B. taurus* breeds from the Near East region, it is evident there was a significant degree of introgression from *B. indicus* cattle, especially from populations inhabiting Iraq, indicative of east regions, with a decline reported for populations located in more westerly regions close to Anatolia, which indicates the introduction of *B. indicus* cattle from Iran and northern Pakistan (Edwards *et al.* 2007).

The use of mtDNA sequencing indicates that taurine and indicine mtDNA sequence groups diverged approximately 200 000 years ago from a minimum of two geographically separate and genetically divergent ancestors (Magee *et al.* 2014). In all, there appear to be five modern *B. taurus* mtDNA haplogroups (T, T1, T2, T3 and T4), whilst a sixth haplotype (T5) has been observed in a few modern Italian cattle. On the other hand, there are two modern *B. indicus* haplogroups (I1 and I2).

In terms of Australian cattle breeding, it is well understood that *B. taurus* cattle are bred in the more temperate climate of southern Australia, whilst *B. indicus* cattle are more suited to the less temperate, more humid and arid climate of northern Australia (Bolton *et al.* 1987; Chenoweth 1994). Indeed, *B. taurus* cattle experience reduction in milk yield, meat quality, conception and growth rates when exposed to heat stress (Hansen 2004; García-Ispuerto *et al.* 2007), which means that they are best suited to Victoria, the southern part of Australia where the climate is cooler. However, approximately 40% of *B. taurus* cattle in southern Australia harbour mtDNA haplotypes indicative of *B. indicus* cattle, and 55% of *B. indicus* cattle in northern Australia possess *B. taurus* mtDNA genomes (Srirattana *et al.* 2017). This would suggest that the domestic species introduced into Australia were the products of either admixture and/or introgression given that it takes approximately 200 years or 30 generations for the distinctive morphological changes to appear in cattle (Magee *et al.* 2014).

It has been strongly argued that certain mtDNA haplotypes are indicative of tolerance to heat in humans (Ruiz-Pesini *et al.* 2004), and this is directly and indirectly supported by some studies in cattle (Hansen 2004; Hernández-Cerón *et al.* 2004; Gaughan *et al.* 2010; Basiricò *et al.* 2011; Liu *et al.* 2011). That being the case, there is a window of opportunity to maximise the current genetics in Australia whereby *B. taurus* cattle with *B. indicus* mtDNA genomes could be relocated to northern Australia, where these cattle exhibiting good-quality meat could be pastured in the warmer, more arid conditions and survive. This offers the potential for their high meat quality to be retained due to the protection that would be afforded to them by their *B. indicus* mtDNA genomes. Furthermore, the maintenance of *B. taurus* cattle with *B. taurus* mtDNA genomes and the introduction of *B. indicus* cattle with *B. taurus* mitochondrial genomes into southern Australia would be appropriate as their *B. taurus* mtDNA genomes would be more suited to the cooler environment. However, the poor-quality meat and milk yields of *B. indicus* cattle with *B. taurus* mitochondrial genomes would offer very low economic breeding values unless the new environment promoted an as yet undefined trait for such cattle. Consequently, as shown in the Australian model, sought-after cattle traits with appropriate chromosomal and mtDNA traits may, indeed, be present. However, until all genetic characteristics are assessed, many breeders will not fully understand the full potential of their livestock and the appropriate way to rear them.

As with pig mtDNA haplotypes (Tsai *et al.* 2016), cattle mtDNA haplotypes distinguish between good- and poor-quality oocytes. *B. indicus* cattle with *B. indicus* mtDNA haplotypes exhibit very low levels of oocyte mtDNA copy numbers and

their oocytes frequently fail to fertilise and result in reduced blastocyst numbers (Srirattana *et al.* 2017). However, it is evident that *B. taurus* cattle harbouring *B. indicus* mtDNA located in southern Australia possess fewer copies of oocyte mtDNA than *B. taurus* cattle with *B. taurus* mtDNA, but oocytes from both backgrounds exhibit similar levels of fertilisation potential and produce a similar number of blastocysts (Srirattana *et al.* 2017). This suggests that some adaptation has occurred in order that the haplotype is fit for its environment. Consequently, as with Australian commercial pigs (Tsai *et al.* 2016), there appear to be specific reproductive efficiencies associated with each mtDNA haplotype.

Evolutionary compensation

Whilst specific mtDNA haplotypes highlight important economic breeding value traits, this does not preclude them from exhibiting a deficiency for another trait, otherwise described as an 'evolutionary trade-off'. Evolutionary trade-off has been described in birds to explain the relationship between, for example, reproductive capacity and energetic expense (Pick *et al.* 2016). As a consequence, females investing in the production of larger oocytes require larger reproductive organs and increased body mass with higher resting metabolic rates, indicating a metabolic deficiency associated with the maternal regulation of reproductive capacity (Pick *et al.* 2016). Other examples include cattle that exhibit the best estimated breeding values and are trait leaders for growth and carcass traits but, frequently, present with poor fertility (Paschal *et al.* 1991; Martin *et al.* 1992; Eler *et al.* 2002). Furthermore, it has been demonstrated that cattle mtDNA can influence the phenotype of bison. In this instance, bison harbouring cattle mtDNA are physically smaller, and this is irrespective of their nutritional environments (Derr *et al.* 2012). In the context of heat tolerance and mtDNA haplotypes, it remains to be determined whether the influence of the mtDNA haplotype persists in cattle exhibiting *B. taurus* phenotypes and possessing *B. indicus* mtDNA haplotypes or whether the effect is lost through interactions of the specific mtDNA haplotype with various nucleus combinations arising from directed breeding programs. In other words, is the balance in communication between the nuclear and mitochondrial genomes lost due to the nuclear genome promoting other key developmental requisites, such as hybrid vigour, at the expense of mtDNA traits?

In predicting these outcomes, theory centred on mitonuclear evolutionary interactions may provide some indications as to potential outcomes. Whilst it is evident that some of the mitochondrial genes transferred to the nuclear genome result in nuclear-mtDNA (N-mtDNA) genes, the location of these genes and the interplay between the nuclear and mitochondrial genomes will have significant outcomes on phenotype and function by modulating patterns of gene flow and introgression, amongst other factors (Wolff *et al.* 2014; Hill 2015; Sloan *et al.* 2017). However, a common theme appears to relate to the efficiency of the organism to use oxidative phosphorylation and the dependence on the mtDNA-encoded genes that are regulators of cellular, tissue and an individual's functional effectiveness, with mutation and deletion leading to their

dysfunction. Indeed, in larger mammalian systems, such as the human, there appears to be significant mitonuclear linkage disequilibrium throughout the human genome, but the associations tended to be weak (Sloan *et al.* 2015).

How do mtDNA haplotypes influence phenotype?

mtDNA haplotypes most likely influence cellular function by regulating chromosomal gene expression patterns and the cell's metabolic profiles, as highlighted in studies using cellular models. For example, pig cells possessing identical chromosomal genomes but different mtDNA haplotypes have been shown to exhibit altered respiration and ATP production rates (Yu *et al.* 2015), which, in turn, modulated the cell's bioenergetics profile. Similar results have been observed in mouse model systems using similar constructs (McKenzie and Trounce 2000; McKenzie *et al.* 2003), namely the fusion of an mtDNA-depleted cell to an enucleated cell (Fig. 7). Additionally, the binding affinity of TFAM, which, in turn, affects the transcriptional efficiency and replication of mtDNA, has been shown to vary amongst haplotypes, which could affect energy production efficiency (Suissa *et al.* 2009). Likewise, in mouse models, certain mtDNA haplotypes have been shown to induce premature aging in animals (Latorre-Pellicer *et al.* 2016). In addition, in reconstructed mouse ESC lines harbouring identical chromosomal genomes but different mtDNA haplotypes (Fig. 7), undifferentiated and differentiated cells produced very different chromosomal gene expression patterns (Kelly *et al.* 2013a), which, for example, influenced neural gene expression profiles during differentiation and *in vitro* beating cardiomyocyte frequency. Indeed, in certain human haplotypes, specific polymorphisms are associated with distinct gene expression patterns (Cohen *et al.* 2016). These outcomes were suggested to be induced by subtle modifications to DNA methylation patterns that are regulated in an mtDNA haplotype-specific manner (Kelly *et al.* 2013a; Lee *et al.* 2017). There is additional evidence that the metabolites released from the citric acid cycle can regulate histone acetylation (Martínez-Reyes *et al.* 2016) and DNA methylation patterns (Lee *et al.* 2017), which act to modulate chromosomal gene expression patterns. Therefore, the degree of nucleomitochondrial compatibility that arises following fertilisation is critical as it will influence the levels of ATP produced by the electron transfer chain (McKenzie and Trounce 2000; McKenzie *et al.* 2003; Yu *et al.* 2015), given that the electron transfer chain is encoded by both genomes. Consequently, the greater the degree of incompatibility, the greater the likelihood that more metabolites will be processed and released from the citric acid cycle, the feeder metabolic pathway, which normally passes electrons to the electron transfer chain, and can modulate histone acetylation (Martínez-Reyes *et al.* 2016) and DNA methylation (Kelly *et al.* 2013a; Lee *et al.* 2017) patterns. In terms of DNA methylation, α -ketoglutarate operates through the TET pathway that modulates the conversion of 5-methylcytosine to 5-hydroxymethylcytosine to demethylate DNA (Blaschke *et al.* 2013; Haseeb *et al.* 2014). This very early interaction has the potential to lay the very first epigenetic signature specific to the developing embryo, fetus and resultant offspring.

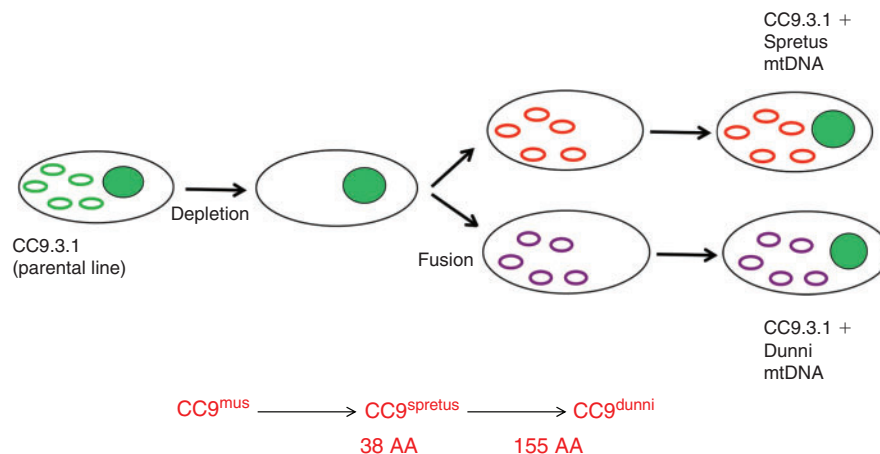


Fig. 7. Generation of mitochondrial (mt) DNA-divergent embryonic stem cells (ESCs) and somatic cell lines. The parental cells, either somatic cells or ESCs, are derived from one mouse strain (*Mus musculus*) and depleted of their mtDNA. In separate reactions, the chromosomes are removed from skin cells (enucleation) from other strains of mice (*Mus spretus* and *Mus dunni*), which have very different mitochondrial haplotypes. These cells are fused with depleted ESCs to produce mtDNA-divergent ESC or somatic cell lines. AA, amino acids. Numbers represent differences in amino acids encoded by each mtDNA haplotype in relation to *M. musculus*.

In the context of livestock breeding, it remains to be determined whether, for example, cattle that have a *B. taurus* phenotype, derived from hybrid back-crossing that results in the retention of a small amount of the *B. indicus* nuclear genotype, with a *B. indicus* mtDNA genotype retain the ability to protect against heat. In other words, the studies that need to be performed in this context are to relocate the relevant animals to hotter, more arid conditions and test their performance levels. It is important to determine whether multiple rounds of breeding through different crossings could have led to or did not lead to the loss of the protective effect brought by the relevant mtDNA haplotype. These types of experiments will determine the transgenerational effects that mtDNA haplotypes can impose on livestock value and viability, or whether other compensatory mechanisms are at play.

mtDNA haplotypes, SCNT and elite breeding programs

By combining knowledge related to mtDNA haplotypes and SCNT, there is a significant potential to generate offspring with the desired traits from the chromosomal and mitochondrial genomes. For example, the ability to produce dairy cattle exhibiting the chromosomal traits for excellent milk quality with mitochondrial genomes associated with heat tolerance would allow cattle with high economic breeding values to be pastured in the more arid and hotter climates of the world. This would offer significant economic benefit to producers and allow otherwise unusable pasture land to be utilised. A similar scenario could be envisaged for beef cattle. Likewise, in the pig, an association between chromosomal traits for lean meat quality with a mitochondrial haplotype for litter size would be beneficial.

Nevertheless, such approaches would require careful assessment using modern molecular genetic tools, and perhaps even

the derivation of cell lines with the desired chromosomal genomes and mtDNA haplotypes (McKenzie *et al.* 2003), to identify suitable genomic partners. To this extent, it has been shown in sheep cloning experiments that successful embryonic development is highly dependent on a slightly more divergent mtDNA haplotype than that of the somatic cell (Bowles *et al.* 2007b). Furthermore, the fusing of a cow somatic cell to several enucleated cow oocytes, each possessing different mtDNA haplotypes, resulted in the somatic cell favouring the propagation of an mtDNA haplotype that was more genetically divergent to its own (Bowles *et al.* 2008). That having been said, these outcomes have significant implications for animal breeding programs. For example, when the somatic cell from a breed of cattle with a high economic breeding value (e.g. an Angus cow with excellent marbling content) favours a more genetically divergent mtDNA haplotype that has greater predisposition for tolerance to increased temperature during earlier development (Hernández-Cerón *et al.* 2004) and when pasturing (Beatty *et al.* 2006; Gaughan *et al.* 2010), then cattle can be generated that offer the potential to expand the current genetics available and promote enhanced livestock production (Bowles *et al.* 2008). This represents a distinct developmental advantage that would confer a specific genetic advantage to the offspring, since they could be propagated and survive in environments that would not normally be to their benefit.

However, the issue of the degree of tolerable genetic divergence needs to be resolved. The degree of sequence divergence between *B. taurus* and *B. indicus* is dependent on the breeds assessed. Nevertheless, within the *Bos* genus, the degree of divergence appears tolerable for the generation of live offspring, as successfully demonstrated using embryonic cells, isolated from preimplantation embryos (Meirelles *et al.* 2001), and somatic cells (Steinborn *et al.* 2002; Bowles *et al.* 2008). In sheep, the predicted tolerable mtDNA sequence divergence

to the blastocyst stage of development lies between 0.0022% and 0.0391%, but when this is extended to cross the species barrier (i.e. when goat cells are introduced into enucleated sheep oocytes), a sequence divergence of 0.4114%, blastocyst formation did not result (Bowles *et al.* 2007a). In mice, increased genetic divergence has resulted in increased embryonic loss and stillborn rates (Ma *et al.* 2016). Furthermore, interspecies nuclear transfer has been attempted in crosses as diverse as cat donor cells and cattle oocytes with poor preimplantation developmental outcome (Thongphakdee *et al.* 2008). Although reprogramming agents, such as scriptaid, 5-azacytidine (Gómez *et al.* 2012) and TSA (Wittayarat *et al.* 2017), can enhance embryo development in interspecies cat nuclear transfer embryos, developmental outcome is still limited. Indeed, interspecies cloning using mouse somatic cells and porcine oocytes appears to only progress to the blastocyst stage when the porcine ooplasm has been partially depleted of mtDNA and murine ESC extract introduced to provide a murine mtDNA template for replication and cytoplasmic factors to promote reprogramming (Jiang *et al.* 2011).

Conclusions

In conclusion, the mitochondrial genome influences the phenotype of the offspring in multiple ways. Through the different mitochondrial haplotypes in a given species, the mitochondrial genome can affect fertility and several other traits that livestock require in order to meet quality expectations. Furthermore, by manipulating the mitochondrial genome when performing either mitochondrial supplementation or SCNT, live offspring can be produced that have the potential to adopt an mtDNA trait promoting enhanced fertility and/or providing protection against diseases and the effects of climate change. These outcomes could promote better husbandry and maintenance of these animals. Understanding the control of the mitochondrial genome on an animal's phenotype is important because it not only accounts for specific traits, but the mitochondrial genome can also regulate the DNA methylation and histone patterns associated with the chromosomal traits and therefore influence some of the phenotypic traits that we previously thought were only directly associated with the chromosomal genome. In all, the mitochondrial genome has an exciting role to play in livestock production for food and other purposes, and should be considered in assessment of an animal's economic breeding value.

Conflicts of interest

JSJ was funded by OvaScience Inc., USA, a company involved in human infertility, to undertake some of the work described in this article. The other authors declare no conflicts of interests.

Acknowledgements

The authors' work reported herein was supported by the Victorian Government's Operational Infrastructure Support Program.

References

- Acton, B. M., Lai, I., Shang, X., Jurisicova, A., and Casper, R. F. (2007). Neutral mitochondrial heteroplasmy alters physiological function in mice. *Biol. Reprod.* **77**, 569–576. doi:10.1095/BIOLREPROD.107.060806
- Agaronyan, K., Morozov, Y. I., Anikin, M., and Temiakov, D. (2015). Mitochondrial biology. Replication–transcription switch in human mitochondria. *Science* **347**, 548–551. doi:10.1126/SCIENCE.AAA0986
- Amaral, A., Ramalho-Santos, J., and St John, J. C. (2007). The expression of polymerase gamma and mitochondrial transcription factor A and the regulation of mitochondrial DNA content in mature human sperm. *Hum. Reprod.* **22**, 1585–1596. doi:10.1093/HUMREP/DEM030
- Anderson, S., de Bruijn, M. H., Coulson, A. R., Eperon, I. C., Sanger, F., and Young, I. G. (1982). Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome. *J. Mol. Biol.* **156**, 683–717. doi:10.1016/0022-2836(82)90137-1
- Archibald, J. M. (2015). Endosymbiosis and eukaryotic cell evolution. *Curr. Biol.* **25**, R911–R921. doi:10.1016/J.CUB.2015.07.055
- Baca, M., and Zamboni, L. (1967). The fine structure of human follicular oocytes. *J. Ultrastruct. Res.* **19**, 354–381. doi:10.1016/S0022-5320(67)80225-9
- Barritt, J., Willadsen, S., Brenner, C., and Cohen, J. (2001a). Cytoplasmic transfer in assisted reproduction. *Hum. Reprod. Update* **7**, 428–435. doi:10.1093/HUMUPD/7.4.428
- Barritt, J. A., Brenner, C. A., Malter, H. E., and Cohen, J. (2001b). Rebuttal: interooplasmic transfers in humans. *Reprod. Biomed. Online* **3**, 47–48. doi:10.1016/S1472-6483(10)61966-9
- Basiricò, L., Morera, P., Primi, V., Lacetera, N., Nardone, A., and Bernabucci, U. (2011). Cellular thermotolerance is associated with heat shock protein 70.1 genetic polymorphisms in Holstein lactating cows. *Cell Stress Chaperones* **16**, 441–448. doi:10.1007/S12192-011-0257-7
- Beatty, D. T., Barnes, A., Taylor, E., Pethick, D., McCarthy, M., and Maloney, S. K. (2006). Physiological responses of *Bos taurus* and *Bos indicus* cattle to prolonged, continuous heat and humidity. *J. Anim. Sci.* **84**, 972–985. doi:10.2527/2006.844972X
- Beekman, M., Dowling, D. K., and Aanen, D. K. (2014). The costs of being male: are there sex-specific effects of uniparental mitochondrial inheritance? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369**, 20130440. doi:10.1098/RSTB.2013.0440
- Bhojwani, S., Alm, H., Torner, H., Kanitz, W., and Poehland, R. (2007). Selection of developmentally competent oocytes through brilliant cresyl blue stain enhances blastocyst development rate after bovine nuclear transfer. *Theriogenology* **67**, 341–345. doi:10.1016/J.THERIOGENOL.2006.08.006
- Birky, C. W., Jr (1995). Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proc. Natl Acad. Sci. USA* **92**, 11331–11338. doi:10.1073/PNAS.92.25.11331
- Blaschke, K., Ebata, K. T., Karimi, M. M., Zepeda-Martinez, J. A., Goyal, P., Mahapatra, S., Tam, A., Laird, D. J., Hirst, M., Rao, A., Lorincz, M. C., and Ramalho-Santos, M. (2013). Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* **500**, 222–226. doi:10.1038/NATURE12362
- Bogenhagen, D. F. (2012). Mitochondrial DNA nucleoid structure. *Biochim. Biophys. Acta* **1819**, 914–920. doi:10.1016/J.BBAGRM.2011.11.005
- Bolton, R. C., Frahm, R. R., Castree, J. W., and Coleman, S. W. (1987). Genotype × environment interactions involving proportion of Brahman breeding and season of birth. I. Calf growth to weaning. *J. Anim. Sci.* **65**, 42–47. doi:10.2527/JAS1987.65142X
- Boucret, L., Chao de la Barca, J. M., Moriniere, C., Desquiret, V., Ferre-L'Hotellier, V., Descamps, P., Marcaillou, C., Reynier, P., Procaccio, V., and May-Panloup, P. (2015). Relationship between diminished ovarian reserve and mitochondrial biogenesis in cumulus cells. *Hum. Reprod.* **30**, 1653–1664. doi:10.1093/HUMREP/DEV114
- Bowles, E. J., Campbell, K. H., and St John, J. C. (2007a). Nuclear transfer: preservation of a nuclear genome at the expense of its associated mtDNA genome(s). *Curr. Top. Dev. Biol.* **77**, 251–290. doi:10.1016/S0070-2153(06)77010-7

- Bowles, E. J., Lee, J. H., Alberio, R., Lloyd, R. E., Stekel, D., Campbell, K. H., and St John, J. C. (2007b). Contrasting effects of *in vitro* fertilization and nuclear transfer on the expression of mtDNA replication factors. *Genetics* **176**, 1511–1526. doi:10.1534/GENETICS.106.070177
- Bowles, E. J., Tecirlioglu, R. T., French, A. J., Holland, M. K., and St John, J. C. (2008). Mitochondrial DNA transmission and transcription after somatic cell fusion to one or more cytoplasts. *Stem Cells* **26**, 775–782. doi:10.1634/STEMCELLS.2007-0747
- Bradley, D. G., Loftus, R. T., Cunningham, P., and MacHugh, D. E. (1998). Genetics and domestic cattle origins. *Evol. Anthropol.* **6**, 79–86. doi:10.1002/(SICI)1520-6505(1998)6:3<79::AID-EVAN2>3.0.CO;2-R
- Brenner, C. A., Barritt, J. A., Willadsen, S., and Cohen, J. (2000). Mitochondrial DNA heteroplasmy after human ooplasmic transplantation. *Fertil. Steril.* **74**, 573–578. doi:10.1016/S0015-0282(00)00681-6
- Brinkman, K., and Kakuda, T. N. (2000). Mitochondrial toxicity of nucleoside analogue reverse transcriptase inhibitors: a looming obstacle for long-term antiretroviral therapy? *Curr. Opin. Infect. Dis.* **13**, 5–11. doi:10.1097/00001432-200002000-00002
- Brinkman, K., ter Hofstede, H. J., Burger, D. M., Smeitink, J. A., and Koopmans, P. P. (1998). Adverse effects of reverse transcriptase inhibitors: mitochondrial toxicity as common pathway. *AIDS* **12**, 1735–1744. doi:10.1097/00002030-199814000-00004
- Brown, W. M., George, M., Jr, and Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. *Proc. Natl Acad. Sci. USA* **76**, 1967–1971. doi:10.1073/PNAS.76.4.1967
- Brüggerhoff, K., Zakhartchenko, V., Wenigerkind, H., Reichenbach, H. D., Prella, K., Scherthaner, W., Alberio, R., Kuchenhoff, H., Stojkovic, M., Brem, G., Hiendler, S., and Wolf, E. (2002). Bovine somatic cell nuclear transfer using recipient oocytes recovered by ovum pick-up: effect of maternal lineage of oocyte donors. *Biol. Reprod.* **66**, 367–373. doi:10.1095/BIOLREPROD66.2.367
- Burgstaller, J. P., Schinogl, P., Dinnyes, A., Muller, M., and Steinborn, R. (2007). Mitochondrial DNA heteroplasmy in ovine fetuses and sheep cloned by somatic cell nuclear transfer. *BMC Dev. Biol.* **7**, 141. doi:10.1186/1471-213X-7-141
- Cagnone, G. L. M., Tsai, T. S., Makanji, Y., Matthews, P., Gould, J., Bonkowski, M. S., Elgass, K. D., Wong, A. S. A., Wu, L. E., McKenzie, M., Sinclair, D. A. S., and St. John, J. C. (2016). Restoration of normal embryogenesis by mitochondrial supplementation in pig oocytes exhibiting mitochondrial DNA deficiency. *Sci. Rep.* **6**, 23229. doi:10.1038/SREP23229
- Cao, L., Shitara, H., Horii, T., Nagao, Y., Imai, H., Abe, K., Hara, T., Hayashi, J., and Yonekawa, H. (2007). The mitochondrial bottleneck occurs without reduction of mtDNA content in female mouse germ cells. *Nat. Genet.* **39**, 386–390. doi:10.1038/NG1970
- Challah-Jacques, M., Chesne, P., and Renard, J. P. (2003). Production of cloned rabbits by somatic nuclear transfer. *Cloning Stem Cells* **5**, 295–299. doi:10.1089/153623003772032808
- Chen, S. H., Pascale, C., Jackson, M., Szvetecz, M. A., and Cohen, J. (2016). A limited survey-based uncontrolled follow-up study of children born after ooplasmic transplantation in a single centre. *Reprod. Biomed. Online* **33**, 737–744. doi:10.1016/J.RBMO.2016.10.003
- Chenoweth, P. J. (1994). Aspects of reproduction in female *Bos indicus* cattle: a review. *Aust. Vet. J.* **71**, 422–426. doi:10.1111/J.1751-0813.1994.TB00961.X
- Chiaratti, M. R., Ferreira, C. R., Meirelles, F. V., Meo, S. C., Perecin, F., Smith, L. C., Ferraz, M. L., de Sa Filho, M. F., Gimenes, L. U., Baruselli, P. S., Gasparini, B., and Garcia, J. M. (2010). Xenoplasmic transfer between buffalo and bovine enables development of homoplasmic offspring. *Cell. Reprogram.* **12**, 231–236. doi:10.1089/CELL.2009.0076
- Cibelli, J. B., Stice, S. L., Golueke, P. J., Kane, J. J., Jerry, J., Blackwell, C., Ponce de Leon, F. A., and Robl, J. M. (1998). Cloned transgenic calves produced from nonquiescent fetal fibroblasts. *Science* **280**, 1256–1258. doi:10.1126/SCIENCE.280.5367.1256
- Cibelli, J. B., Campbell, K. H., Seidel, G. E., West, M. D., and Lanza, R. P. (2002). The health profile of cloned animals. *Nat. Biotechnol.* **20**, 13–14. doi:10.1038/NBT0102-13
- Ciesielski, G. L., Bermek, O., Rosado-Ruiz, F. A., Hovde, S. L., Neitzke, O. J., Griffith, J. D., and Kaguni, L. S. (2015). Mitochondrial single-stranded DNA-binding proteins stimulate the activity of DNA polymerase gamma by organization of the template DNA. *J. Biol. Chem.* **290**, 28697–28707. doi:10.1074/JBC.M115.673707
- Clayton, D. A. (1992). Transcription and replication of animal mitochondrial DNAs. *Int. Rev. Cytol.* **141**, 217–232. doi:10.1016/S0074-7696(08)62067-7
- Clayton, D. A. (1998). Nuclear–mitochondrial intergenomic communication. *Biofactors* **7**, 203–205. doi:10.1002/BIOF.5520070307
- Cloonan, S. M., and Choi, A. M. (2012). Mitochondria: commanders of innate immunity and disease? *Curr. Opin. Immunol.* **24**, 32–40. doi:10.1016/J.COI.2011.11.001
- Cohen, J., Scott, R., Schimmel, T., Levron, J., and Willadsen, S. (1997). Birth of infant after transfer of anucleate donor oocyte cytoplasm into recipient eggs. *Lancet* **350**, 186–187. doi:10.1016/S0140-6736(05)62353-7
- Cohen, T., Levin, L., and Mishmar, D. (2016). Ancient out-of-Africa mitochondrial DNA variants associate with distinct mitochondrial gene expression patterns. *PLoS Genet.* **12**, e1006407. doi:10.1371/JOURN.PGEN.1006407
- Coskun, P. E., Ruiz-Pesini, E., and Wallace, D. C. (2003). Control region mtDNA variants: longevity, climatic adaptation, and a forensic conundrum. *Proc. Natl Acad. Sci. USA* **100**, 2174–2176. doi:10.1073/PNAS.0630589100
- Cosmides, L. M., and Tooby, J. (1981). Cytoplasmic inheritance and intragenomic conflict. *J. Theor. Biol.* **89**, 83–129. doi:10.1016/0022-5193(81)90181-8
- Cree, L. M., Samuels, D. C., de Sousa Lopes, S. C., Rajasimha, H. K., Wonnapijit, P., Mann, J. R., Dahl, H. H., and Chinnery, P. F. (2008). A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes. *Nat. Genet.* **40**, 249–254. doi:10.1038/NG.2007.63
- D'Erchia, A. M., Atlante, A., Gadaleta, G., Pavesi, G., Chiara, M., De Virgilio, C., Manzari, C., Mastropasqua, F., Prazzoli, G. M., Picardi, E., Gissi, C., Horner, D., Reyes, A., Sbisa, E., Tullo, A., and Pesole, G. (2015). Tissue-specific mtDNA abundance from exome data and its correlation with mitochondrial transcription, mass and respiratory activity. *Mitochondrion* **20**, 13–21. doi:10.1016/J.MITO.2014.10.005
- Derr, J. N., Hedrick, P. W., Halbert, N. D., Plough, L., Dobson, L. K., King, J., Duncan, C., Hunter, D. L., Cohen, N. D., and Hedgecock, D. (2012). Phenotypic effects of cattle mitochondrial DNA in American bison. *Conserv. Biol.* **26**, 1130–1136. doi:10.1111/J.1523-1739.2012.01905.X
- Desjardins, P., Frost, E., and Morais, R. (1985). Ethidium bromide-induced loss of mitochondrial DNA from primary chicken embryo fibroblasts. *Mol. Cell. Biol.* **5**, 1163–1169. doi:10.1128/MCB.5.5.1163
- Dickinson, A., Yeung, K. Y., Donoghue, J., Baker, M. J., Kelly, R. D., McKenzie, M., Johns, T. G., and St John, J. C. (2013). The regulation of mitochondrial DNA copy number in glioblastoma cells. *Cell Death Differ.* **20**, 1644–1653. doi:10.1038/CDD.2013.115
- Diez-Sánchez, C., Ruiz-Pesini, E., Lapeña, A. C., Montoya, J., Pérez-Martos, A., Enríquez, J. A., and López-Pérez, M. J. (2003). Mitochondrial DNA content of human spermatozoa. *Biol. Reprod.* **68**, 180–185. doi:10.1095/BIOLREPROD.102.005140
- Doro, M. G., Piras, D., Leoni, G. G., Casu, G., Vaccargiu, S., Parracciani, D., Naitana, S., Pirastu, M., and Novelletto, A. (2014). Phylogeny and patterns of diversity of goat mtDNA haplogroup A revealed by resequencing complete mitogenomes. *PLoS One* **9**, e95969. doi:10.1371/JOURNAL.PONE.0095969
- Duchen, M. R. (1999). Contributions of mitochondria to animal physiology: from homeostatic sensor to calcium signalling and cell death. *J. Physiol.* **516**, 1–17. doi:10.1111/J.1469-7793.1999.001AA.X

- Edwards, C. J., Baird, J. F., and MacHugh, D. E. (2007). Taurine and zebu admixture in Near Eastern cattle: a comparison of mitochondrial, autosomal and Y-chromosomal data. *Anim. Genet.* **38**, 520–524. doi:10.1111/J.1365-2052.2007.01638.X
- El Shourbagy, S. H., Spikings, E. C., Freitas, M., and St John, J. C. (2006). Mitochondria directly influence fertilisation outcome in the pig. *Reproduction* **131**, 233–245. doi:10.1530/REP.1.00551
- Eler, J. P., Silva, J. A., Ferraz, J. B., Dias, F., Oliveira, H. N., Evans, J. L., and Golden, B. L. (2002). Genetic evaluation of the probability of pregnancy at 14 months for Nellore heifers. *J. Anim. Sci.* **80**, 951–954. doi:10.2527/2002.804951X
- Facucho-Oliveira, J. M., Alderson, J., Spikings, E. C., Egginton, S., and St John, J. C. (2007). Mitochondrial DNA replication during differentiation of murine embryonic stem cells. *J. Cell Sci.* **120**, 4025–4034. doi:10.1242/JCS.016972
- Fakih, M., El Shmoumy, M., Szeptycki, J., dela Cruz, D. B., Lux, C., Verjee, S., Burgess, C. M., Cohn, G. M., and Casper, R. F. (2015). The AUGMENTSM treatment: physician reported outcomes of the initial global patient experience. *JFIV Reprod. Med. Genet.* **3**, e116. doi:10.4172/2375-4508.1000154
- Farr, C. L., Matsushima, Y., Lagina, A. T., 3rd, Luo, N., and Kaguni, L. S. (2004). Physiological and biochemical defects in functional interactions of mitochondrial DNA polymerase and DNA-binding mutants of single-stranded DNA-binding protein. *J. Biol. Chem.* **279**, 17047–17053. doi:10.1074/JBC.M400283200
- Flood, M. R., and Wiebold, J. L. (1988). Glucose metabolism by preimplantation pig embryos. *J. Reprod. Fertil.* **84**, 7–12. doi:10.1530/JRF.0.0840007
- Folgerø, T., Bertheussen, K., Lindal, S., Torbergsen, T., and Øian, P. (1993). Mitochondrial disease and reduced sperm motility. *Hum. Reprod.* **8**, 1863–1868. doi:10.1093/OXFORDJOURNALS.HUMREP.A137950
- Fragouli, E., Spath, K., Alfaraawati, S., Kaper, F., Craig, A., Michel, C. E., Kokocinski, F., Cohen, J., Munne, S., and Wells, D. (2015). Altered levels of mitochondrial DNA are associated with female age, aneuploidy, and provide an independent measure of embryonic implantation potential. *PLoS Genet.* **11**, e1005241. doi:10.1371/JOURNAL.PGEN.1005241
- Frank, S. A., and Hurst, L. D. (1996). Mitochondria and male disease. *Nature* **383**, 224. doi:10.1038/383224A0
- Frantz, L. A., Schraiber, J. G., Madsen, O., Megens, H. J., Bosse, M., Paudel, Y., Semiadi, G., Meijaard, E., Li, N., Crooijmans, R. P., Archibald, A. L., Slatkin, M., Schook, L. B., Larson, G., and Groenen, M. A. (2013). Genome sequencing reveals fine scale diversification and reticulation history during speciation in *Sus*. *Genome Biol.* **14**, R107. doi:10.1186/GB-2013-14-9-R107
- Frantz, L. A., Schraiber, J. G., Madsen, O., Megens, H. J., Cagan, A., Bosse, M., Paudel, Y., Crooijmans, R. P., Larson, G., and Groenen, M. A. (2015). Evidence of long-term gene flow and selection during domestication from analyses of Eurasian wild and domestic pig genomes. *Nat. Genet.* **47**, 1141–1148. doi:10.1038/NG.3394
- Galli, C., Lagutina, I., Crotti, G., Colleoni, S., Turini, P., Ponderato, N., Duchi, R., and Lazzari, G. (2003). Pregnancy: a cloned horse born to its dam twin. *Nature* **424**, 635. doi:10.1038/424635A
- García-Ispuerto, I., López-Gatius, F., Bech-Sabat, G., Santolaria, P., Yáñez, J. L., Nogareda, C., De Rensis, F., and López-Béjar, M. (2007). Climate factors affecting conception rate of high producing dairy cows in northeastern Spain. *Theriogenology* **67**, 1379–1385. doi:10.1016/J.THERIOGENOLOGY.2007.02.009
- Gaughan, J. B., Mader, T. L., Holt, S. M., Sullivan, M. L., and Hahn, G. L. (2010). Assessing the heat tolerance of 17 beef cattle genotypes. *Int. J. Biometeorol.* **54**, 617–627. doi:10.1007/S00484-009-0233-4
- Gemmell, N. J., Metcalf, V. J., and Allendorf, F. W. (2004). Mother's curse: the effect of mtDNA on individual fitness and population viability. *Trends Ecol. Evol.* **19**, 238–244. doi:10.1016/J.TREE.2004.02.002
- Gerber, A. S., Loggins, R., Kumar, S., and Dowling, T. E. (2001). Does nonneutral evolution shape observed patterns of DNA variation in animal mitochondrial genomes? *Annu. Rev. Genet.* **35**, 539–566. doi:10.1146/ANNUREV.GENET.35.102401.091106
- Gómez, M. C., Biancardi, M. N., Jenkins, J. A., Dumas, C., Galiguis, J., Wang, G., and Earle Pope, C. (2012). Scriptaid and 5-aza-2'-deoxycytidine enhanced expression of pluripotent genes and *in vitro* developmental competence in interspecies black-footed cat cloned embryos. *Reprod. Domest. Anim.* **47**(Suppl. 6), 130–135. doi:10.1111/RDA.12027
- Gray, M. W. (1989). Origin and evolution of mitochondrial DNA. *Annu. Rev. Cell Biol.* **5**, 25–50. doi:10.1146/ANNUREV.CB.05.110189.000325
- Gray, M. W. (2012). Mitochondrial evolution. *Cold Spring Harb. Perspect. Biol.* **4**, a011403. doi:10.1101/CSHPERSPECT.A011403
- Gyllenstein, U., Wharton, D., Josefsson, A., and Wilson, A. C. (1991). Paternal inheritance of mitochondrial DNA in mice. *Nature* **352**, 255–257. doi:10.1038/352255A0
- Hammond, E. R., Green, M. P., Shelling, A. N., Berg, M. C., Peek, J. C., and Cree, L. M. (2016). Oocyte mitochondrial deletions and heteroplasmy in a bovine model of ageing and ovarian stimulation. *Mol. Hum. Reprod.* **22**, 261–271. doi:10.1093/MOLEHR/GAW003
- Hance, N., Ekstrand, M. I., and Trifunovic, A. (2005). Mitochondrial DNA polymerase gamma is essential for mammalian embryogenesis. *Hum. Mol. Genet.* **14**, 1775–1783. doi:10.1093/HMG/DDI184
- Hansen, P. J. (2004). Physiological and cellular adaptations of zebu cattle to thermal stress. *Anim. Reprod. Sci.* **82–83**, 349–360. doi:10.1016/J.ANIREPROSCI.2004.04.011
- Haseeb, A., Makki, M. S., and Haqqi, T. M. (2014). Modulation of ten-eleven translocation 1 (TET1), isocitrate dehydrogenase (IDH) expression, alpha-ketoglutarate (alpha-KG), and DNA hydroxymethylation levels by interleukin-1beta in primary human chondrocytes. *J. Biol. Chem.* **289**, 6877–6885. doi:10.1074/JBC.M113.512269
- Hasegawa, M., Kishino, H., and Yano, T. (1985). Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**, 160–174. doi:10.1007/BF02101694
- Hayakawa, T., Noda, M., Yasuda, K., Yorifuji, H., Taniguchi, S., Miwa, I., Sakura, H., Terauchi, Y., Hayashi, J., Sharp, G. W., Kanazawa, Y., Akanuma, Y., Yazaki, Y., and Kadowaki, T. (1998). Ethidium bromide-induced inhibition of mitochondrial gene transcription suppresses glucose-stimulated insulin release in the mouse pancreatic beta-cell line betaHC9. *J. Biol. Chem.* **273**, 20300–20307. doi:10.1074/JBC.273.32.20300
- Hecht, N. B., Liem, H., Kleene, K. C., Distel, R. J., and Ho, S. M. (1984). Maternal inheritance of the mouse mitochondrial genome is not mediated by a loss or gross alteration of the paternal mitochondrial DNA or by methylation of the oocyte mitochondrial DNA. *Dev. Biol.* **102**, 452–461. doi:10.1016/0012-1606(84)90210-0
- Hernández-Cerón, J., Chase, C. C., Jr, and Hansen, P. J. (2004). Differences in heat tolerance between preimplantation embryos from Brahman, Romosinuano, and Angus breeds. *J. Dairy Sci.* **87**, 53–58. doi:10.3168/JDS.S0022-0302(04)73141-0
- Hiendleder, S., Lewalski, H., Wassmuth, R., and Janke, A. (1998). The complete mitochondrial DNA sequence of the domestic sheep (*Ovis aries*) and comparison with the other major ovine haplotype. *J. Mol. Evol.* **47**, 441–448. doi:10.1007/PL00006401
- Hiendleder, S., Zakhartchenko, V., Wenigerkind, H., Reichenbach, H. D., Brüggerhoff, K., Prella, K., Brem, G., Stojkovic, M., and Wolf, E. (2003). Heteroplasmy in bovine fetuses produced by intra- and interspecific somatic cell nuclear transfer: neutral segregation of nuclear donor mitochondrial DNA in various tissues and evidence for recipient cow mitochondria in fetal blood. *Biol. Reprod.* **68**, 159–166. doi:10.1095/BIOLREPROD.102.008201
- Hill, G. E. (2015). Mitonuclear ecology. *Mol. Biol. Evol.* **32**, 1917–1927. doi:10.1093/MOLBEV/MSV104

- Hosseinzadeh Colagar, A., and Karimi, F. (2014). Large scale deletions of the mitochondrial DNA in astheno, asthenotero and oligoasthenotero-spermic men. *Mitochondrial DNA* **25**, 321–328. doi:10.3109/19401736.2013.796512
- Houghton, F. D. (2006). Energy metabolism of the inner cell mass and trophectoderm of the mouse blastocyst. *Differentiation* **74**, 11–18. doi:10.1111/J.1432-0436.2006.00052.X
- Hua, S., Zhang, Y., Li, X. C., Ma, L. B., Cao, J. W., Dai, J. P., and Li, R. (2007). Effects of granulosa cell mitochondria transfer on the early development of bovine embryos *in vitro*. *Cloning Stem Cells* **9**, 237–246. doi:10.1089/CLO.2006.0020
- Huang, C. C., Cheng, T. C., Chang, H. H., Chang, C. C., Chen, C. I., Liu, J., and Lee, M. S. (1999). Birth after the injection of sperm and the cytoplasm of tripronucleate zygotes into metaphase II oocytes in patients with repeated implantation failure after assisted fertilization procedures. *Fertil. Steril.* **72**, 702–706. doi:10.1016/S0015-0282(99)00309-X
- Hurst, L. D., and Hamilton, W. D. (1992). Cytoplasmic fusion and the nature of sexes. *Proc. Biol. Sci.* **247**, 189–194. doi:10.1098/RSPB.1992.0027
- Hwang, S., Kwak, S. H., Bhak, J., Kang, H. S., Lee, Y. R., Koo, B. K., Park, K. S., Lee, H. K., and Cho, Y. M. (2011). Gene expression pattern in transmitochondrial cytoplasmic hybrid cells harboring type 2 diabetes-associated mitochondrial DNA haplogroups. *PLoS One* **6**, e22116. doi:10.1371/JOURNAL.PONE.0022116
- Innocenti, P., Morrow, E. H., and Dowling, D. K. (2011). Experimental evidence supports a sex-specific selective sieve in mitochondrial genome evolution. *Science* **332**, 845–848. doi:10.1126/SCIENCE.1201157
- Jansen, R. P. S. (2000). Germline passage of mitochondria: quantitative considerations and possible embryological sequelae. *Hum. Reprod.* **15** (Suppl. 2), 112–128. doi:10.1093/HUMREP/15.SUPPL_2.112
- Jansen, R. P., and de Boer, K. (1998). The bottleneck: mitochondrial imperatives in oogenesis and ovarian follicular fate. *Mol. Cell. Endocrinol.* **145**, 81–88. doi:10.1016/S0303-7207(98)00173-7
- Jenuth, J. P., Peterson, A. C., and Shoubridge, E. A. (1997). Tissue-specific selection for different mtDNA genotypes in heteroplasmic mice. *Nat. Genet.* **16**, 93–95. doi:10.1038/NG0597-93
- Ji, F., Sharples, M. S., Derbeneva, O., Alves, L. S., Qian, P., Wang, Y., Chalkia, D., Lvova, M., Xu, J., Yao, W., Simon, M., Platt, J., Xu, S., Angelin, A., Davila, A., Huang, T., Wang, P. H., Chuang, L. M., Moore, L. G., Qian, G., and Wallace, D. C. (2012). Mitochondrial DNA variant associated with Leber hereditary optic neuropathy and high-altitude Tibetans. *Proc. Natl Acad. Sci. USA* **109**, 7391–7396. doi:10.1073/PNAS.1202484109
- Jiang, Y., Kelly, R., Peters, A., Fulka, H., Dickinson, A., Mitchell, D. A., and St John, J. C. (2011). Interspecies somatic cell nuclear transfer is dependent on compatible mitochondrial DNA and reprogramming factors. *PLoS One* **6**, e14805. doi:10.1371/JOURNAL.PONE.0014805
- Kaipappattu, B. A., Ma, Y., and Wong, L. J. (2010). Functional effects of cancer mitochondria on energy metabolism and tumorigenesis: utility of transmitochondrial cybrids. *Ann. N. Y. Acad. Sci.* **1201**, 137–146. doi:10.1111/J.1749-6632.2010.05621.X
- Kao, S., Chao, H. T., and Wei, Y. H. (1995). Mitochondrial deoxyribonucleic acid 4977-bp deletion is associated with diminished fertility and motility of human sperm. *Biol. Reprod.* **52**, 729–736. doi:10.1095/BIOLREPROD52.4.729
- Keefer, C. L., Baldassarre, H., Keyston, R., Wang, B., Bhatia, B., Bilodeau, A. S., Zhou, J. F., Leduc, M., Downey, B. R., Lazaris, A., and Karatzas, C. N. (2001). Generation of dwarf goat (*Capra hircus*) clones following nuclear transfer with transfected and nontransfected fetal fibroblasts and *in vitro*-matured oocytes. *Biol. Reprod.* **64**, 849–856. doi:10.1095/BIOLREPROD64.3.849
- Kelly, T. K., De Carvalho, D. D., and Jones, P. A. (2010). Epigenetic modifications as therapeutic targets. *Nat. Biotechnol.* **28**, 1069–1078. doi:10.1038/NBT.1678
- Kelly, R. D., Mahmud, A., McKenzie, M., Trounce, I. A., and St John, J. C. (2012). Mitochondrial DNA copy number is regulated in a tissue specific manner by DNA methylation of the nuclear-encoded DNA polymerase gamma A. *Nucleic Acids Res.* **40**, 10124–10138. doi:10.1093/NAR/GKS770
- Kelly, R. D., Rodda, A. E., Dickinson, A., Mahmud, A., Nefzger, C. M., Lee, W., Forsythe, J. S., Polo, J. M., Trounce, I. A., McKenzie, M., Nisbet, D. R., and St John, J. C. (2013a). Mitochondrial DNA haplotypes define gene expression patterns in pluripotent and differentiating embryonic stem cells. *Stem Cells* **31**, 703–716. doi:10.1002/STEM.1313
- Kelly, R. D., Sumer, H., McKenzie, M., Facucho-Oliveira, J., Trounce, I. A., Verma, P. J., and St John, J. C. (2013b). The effects of nuclear reprogramming on mitochondrial DNA replication. *Stem Cell Rev.* **9**, 1–15. doi:10.1007/S12015-011-9318-7
- Kenchington, E. L., Hamilton, L., Cogswell, A., and Zouros, E. (2009). Paternal mtDNA and maleness are co-inherited but not causally linked in mytilid mussels. *PLoS One* **4**, e6976. doi:10.1371/JOURNAL.PONE.0006976
- Korhonen, J. A., Gaspari, M., and Falkenberg, M. (2003). TWINKLE has 5' → 3' DNA helicase activity and is specifically stimulated by mitochondrial single-stranded DNA-binding protein. *J. Biol. Chem.* **278**, 48627–48632. doi:10.1074/JBC.M306981200
- Krisher, R. L., and Prather, R. S. (2012). A role for the Warburg effect in preimplantation embryo development: metabolic modification to support rapid cell proliferation. *Mol. Reprod. Dev.* **79**, 311–320. doi:10.1002/MRD.22037
- Kucej, M., and Butow, R. A. (2007). Evolutionary tinkering with mitochondrial nucleoids. *Trends Cell Biol.* **17**, 586–592. doi:10.1016/J.TCB.2007.08.007
- Lanzendorf, S. E., Mayer, J. F., Toner, J., Oehninger, S., Saffan, D. S., and Muasher, S. (1999). Pregnancy following transfer of ooplasm from cryopreserved-thawed donor oocytes into recipient oocytes. *Fertil. Steril.* **71**, 575–577. doi:10.1016/S0015-0282(98)00504-4
- Larsson, N. G., Garman, J. D., Oldfors, A., Barsh, G. S., and Clayton, D. A. (1996). A single mouse gene encodes the mitochondrial transcription factor A and a testis-specific nuclear HMG-box protein. *Nat. Genet.* **13**, 296–302. doi:10.1038/NG0796-296
- Larsson, N. G., Wang, J., Wilhelmsson, H., Oldfors, A., Rustin, P., Lewandoski, M., Barsh, G. S., and Clayton, D. A. (1998). Mitochondrial transcription factor A is necessary for mtDNA maintenance and embryogenesis in mice. *Nat. Genet.* **18**, 231–236. doi:10.1038/NG0398-231
- Latorre-Pellicer, A., Moreno-Loshuertos, R., Lechuga-Vieco, A. V., Sanchez-Cabo, F., Torroja, C., Acin-Perez, R., Calvo, E., Aix, E., Gonzalez-Guerra, A., Logan, A., Bernad-Miana, M. L., Romanos, E., Cruz, R., Cogliati, S., Sobrino, B., Carracedo, A., Perez-Martos, A., Fernandez-Silva, P., Ruiz-Cabello, J., Murphy, M. P., Flores, I., Vazquez, J., and Enriquez, J. A. (2016). Mitochondrial and nuclear DNA matching shapes metabolism and healthy ageing. *Nature* **535**, 561–565. doi:10.1038/NATURE18618
- Leahy, A., Xiong, J. W., Kuhnert, F., and Stuhlmann, H. (1999). Use of developmental marker genes to define temporal and spatial patterns of differentiation during embryoid body formation. *J. Exp. Zool.* **284**, 67–81. doi:10.1002/(SICI)1097-010X(19990615)284:1<67::AID-JEZ10>3.0.CO;2-O
- Leary, S. C., Battersby, B. J., Hansford, R. G., and Moyes, C. D. (1998). Interactions between bioenergetics and mitochondrial biogenesis. *Biochim. Biophys. Acta* **1365**, 522–530. doi:10.1016/S0005-2728(98)00105-4
- Lee, W. T., and St John, J. (2015). The control of mitochondrial DNA replication during development and tumorigenesis. *Ann. N. Y. Acad. Sci.* **1350**, 95–106. doi:10.1111/NYAS.12873

- Lee, J. H., Peters, A., Fisher, P., Bowles, E. J., St John, J. C., and Campbell, K. H. (2010). Generation of mtDNA homoplasmic cloned lambs. *Cell. Reprogram.* **12**, 347–355. doi:10.1089/CELL.2009.0096
- Lee, W., Johnson, J., Gough, D. J., Donoghue, J., Cagnone, G. L. M., 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. doi:10.1038/CDDIS.2015.34
- Lee, W. T. Y., Cain, J. E., Cuddihy, A., Johnson, J., Dickinson, A., Yeung, K. Y., Kumar, B., Johns, T. G., Watkins, D. N., Spencer, A., and St John, J. C. (2016). Mitochondrial DNA plasticity is an essential inducer of tumorigenesis. *Cell Death Dis.* **2**, 16016. doi:10.1038/CDDISCOVERY.2016.16
- Lee, W. T., Sun, X., Tsai, T.-S., Johnson, J. L., Gould, J. A., Garama, D. J., Gough, D. J., McKenzie, M., Trounce, I. A., and St John, J. C. (2017). Mitochondrial DNA haplotypes induce differential patterns of DNA methylation that result in differential chromosomal gene expression patterns. *Cell Death Discov.* **3**, 17062. doi:10.1038/CDDISCOVERY.2017.62
- Leese, H. J., Sturmey, R. G., Baumann, C. G., and McEvoy, T. G. (2007). Embryo viability and metabolism: obeying the quiet rules. *Hum. Reprod.* **22**, 3047–3050. doi:10.1093/HUMREP/DEM253
- Legros, F., Malka, F., Frachon, P., Lombes, A., and Rojo, M. (2004). Organization and dynamics of human mitochondrial DNA. *J. Cell Sci.* **117**, 2653–2662. doi:10.1242/JCS.01134
- Lestienne, P., Reynier, P., Chretien, M. F., Penisson-Besnier, I., Malthiery, Y., and Rohmer, V. (1997). Oligoasthenospermia associated with multiple mitochondrial DNA rearrangements. *Mol. Hum. Reprod.* **3**, 811–814. doi:10.1093/MOLEHR/3.9.811
- Li, R., Wen, B., Zhao, H., Ouyang, N., Ou, S., Wang, W., Han, J., and Yang, D. (2017). Embryo development after mitochondrial supplementation from induced pluripotent stem cells. *J. Assist. Reprod. Genet.* **34**, 1027–1033. doi:10.1007/S10815-017-0948-9
- Lirón, J. P., Bravi, C. M., Mirol, P. M., Peral-García, P., and Giovambattista, G. (2006). African matrilineages in American Creole cattle: evidence of two independent continental sources. *Anim. Genet.* **37**, 379–382. doi:10.1111/J.1365-2052.2006.01452.X
- Liu, Y., Li, D., Li, H., Zhou, X., and Wang, G. (2011). A novel SNP of the *ATPIA1* gene is associated with heat tolerance traits in dairy cows. *Mol. Biol. Rep.* **38**, 83–88. doi:10.1007/S11033-010-0080-8
- Lloyd, R. E., Lee, J. H., Alberio, R., Bowles, E. J., Ramalho-Santos, J., Campbell, K. H., and St John, J. C. (2006). Aberrant nucleo-cytoplasmic cross-talk results in donor cell mtDNA persistence in cloned embryos. *Genetics* **172**, 2515–2527. doi:10.1534/GENETICS.105.055145
- Ma, H., Marti Gutierrez, N., Morey, R., Van Dyken, C., Kang, E., Hayama, T., Lee, Y., Li, Y., Tippner-Hedges, R., Wolf, D. P., Laurent, L. C., and Mitalipov, S. (2016). Incompatibility between nuclear and mitochondrial genomes contributes to an interspecies reproductive barrier. *Cell Metab.* **24**, 283–294. doi:10.1016/J.CMET.2016.06.012
- Macháty, Z., Thompson, J. G., Abeydeera, L. R., Day, B. N., and Prather, R. S. (2001). Inhibitors of mitochondrial ATP production at the time of compaction improve development of *in vitro* produced porcine embryos. *Mol. Reprod. Dev.* **58**, 39–44. doi:10.1002/1098-2795(200101)58:1<39::AID-MRD6>3.0.CO;2-B
- Agee, D. A., MacHugh, D. E., and Edwards, C. J. (2014). Interrogation of modern and ancient genomes reveals the complex domestic history of cattle. *Anim. Front.* **4**, 7–22. doi:10.2527/AF.2014-0017
- Mannen, H., Tsuji, S., Loftus, R. T., and Bradley, D. G. (1998). Mitochondrial DNA variation and evolution of Japanese black cattle (*Bos taurus*). *Genetics* **150**, 1169–1175.
- Mannen, H., Morimoto, M., Oyama, K., Mukai, F., and Tsuji, S. (2003). Identification of mitochondrial DNA substitutions related to meat quality in Japanese Black cattle. *J. Anim. Sci.* **81**, 68–73. doi:10.2527/2003.81168X
- Martin, L. C., Brinks, J. S., Bourdon, R. M., and Cundiff, L. V. (1992). Genetic effects on beef heifer puberty and subsequent reproduction. *J. Anim. Sci.* **70**, 4006–4017. doi:10.2527/1992.70124006X
- Martin, W. F., Garg, S., and Zimorski, V. (2015). Endosymbiotic theories for eukaryote origin. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **370**, 20140330. doi:10.1098/RSTB.2014.0330
- Martinez, F., Olvera-Sanchez, S., Esparza-Perusquia, M., Gomez-Chang, E., and Flores-Herrera, O. (2015). Multiple functions of syncytiotrophoblast mitochondria. *Steroids* **103**, 11–22. doi:10.1016/J.STEROIDS.2015.09.006
- Martinez-Reyes, I., Diebold, L. P., Kong, H., Schieber, M., Huang, H., Hensley, C. T., Mehta, M. M., Wang, T., Santos, J. H., Woychik, R., Dufour, E., Spelbrink, J. N., Weinberg, S. E., Zhao, Y., DeBerardinis, R. J., and Chandel, N. S. (2016). TCA cycle and mitochondrial membrane potential are necessary for diverse biological functions. *Mol. Cell* **61**, 199–209. doi:10.1016/J.MOLCEL.2015.12.002
- May-Panloup, P., Chretien, M. F., Savagner, F., Vasseur, C., Jean, M., Malthiery, Y., and Reynier, P. (2003). Increased sperm mitochondrial DNA content in male infertility. *Hum. Reprod.* **18**, 550–556. doi:10.1093/HUMREP/DEG096
- May-Panloup, P., Chretien, M. F., Jacques, C., Vasseur, C., Malthiery, Y., and Reynier, P. (2005a). Low oocyte mitochondrial DNA content in ovarian insufficiency. *Hum. Reprod.* **20**, 593–597. doi:10.1093/HUMREP/DEH667
- May-Panloup, P., Vignon, X., Chretien, M. F., Heyman, Y., Tamassia, M., Malthiery, Y., and Reynier, P. (2005b). Increase of mitochondrial DNA content and transcripts in early bovine embryogenesis associated with upregulation of mtTFA and NRF1 transcription factors. *Reprod. Biol. Endocrinol.* **3**, 65. doi:10.1186/1477-7827-3-65
- McConnell, J. M., and Petrie, L. (2004). Mitochondrial DNA turnover occurs during preimplantation development and can be modulated by environmental factors. *Reprod. Biomed. Online* **9**, 418–424. doi:10.1016/S1472-6483(10)61277-1
- McFarland, R., Taylor, R. W., and Turnbull, D. M. (2007). Mitochondrial disease – its impact, etiology, and pathology. *Curr. Top. Dev. Biol.* **77**, 113–155. doi:10.1016/S0070-2153(06)77005-3
- McKenzie, M., and Trounce, I. (2000). Expression of *Rattus norvegicus* mtDNA in *Mus musculus* cells results in multiple respiratory chain defects. *J. Biol. Chem.* **275**, 31514–31519. doi:10.1074/JBC.M004070200
- McKenzie, M., Chiotis, M., Pinkert, C. A., and Trounce, I. A. (2003). Functional respiratory chain analyses in murine xenomitochondrial cybrids expose coevolutionary constraints of cytochrome *b* and nuclear subunits of complex III. *Mol. Biol. Evol.* **20**, 1117–1124. doi:10.1093/MOLBEV/MSG132
- Meirelles, F. V., Bordignon, V., Watanabe, Y., Watanabe, M., Dayan, A., Lobo, R. B., Garcia, J. M., and Smith, L. C. (2001). Complete replacement of the mitochondrial genotype in a *Bos indicus* calf reconstructed by nuclear transfer to a *Bos taurus* oocyte. *Genetics* **158**, 351–356.
- Monlun, M., Hyernard, C., Blanco, P., Lartigue, L., and Faustin, B. (2017). Mitochondria as molecular platforms integrating multiple innate immune signalings. *J. Mol. Biol.* **429**, 1–13. doi:10.1016/J.JMB.2016.10.028
- Moyes, C. D., Battersby, B. J., and Leary, S. C. (1998). Regulation of muscle mitochondrial design. *J. Exp. Biol.* **201**, 299–307.
- Murakoshi, Y., Sueoka, K., Takahashi, K., Sato, S., Sakurai, T., Tajima, H., and Yoshimura, Y. (2013). Embryo developmental capability and pregnancy outcome are related to the mitochondrial DNA copy number and ooplasmic volume. *J. Assist. Reprod. Genet.* **30**, 1367–1375. doi:10.1007/S10815-013-0062-6
- Nagao, Y., Totsuka, Y., Atomi, Y., Kaneda, H., Lindahl, K. F., Imai, H., and Yonekawa, H. (1998). Decreased physical performance of congenic

- mice with mismatch between the nuclear and the mitochondrial genome. *Genes Genet. Syst.* **73**, 21–27. doi:10.1266/GGS.73.21
- Nishimura, Y., Yoshinari, T., Naruse, K., Yamada, T., Sumi, K., Mitani, H., Higashiyama, T., and Kuroiwa, T. (2006). Active digestion of sperm mitochondrial DNA in single living sperm revealed by optical tweezers. *Proc. Natl Acad. Sci. USA* **103**, 1382–1387. doi:10.1073/PNAS.0506911103
- Oakes, C. C., La Salle, S., Smiraglia, D. J., Robaire, B., and Trasler, J. M. (2007). Developmental acquisition of genome-wide DNA methylation occurs prior to meiosis in male germ cells. *Dev. Biol.* **307**, 368–379. doi:10.1016/J.YDBIO.2007.05.002
- Paschal, J. C., Sanders, J. O., and Kerr, J. L. (1991). Calving and weaning characteristics of Angus-, Gray Brahman-, Gir-, Indu-Brazil-, Nellore-, and Red Brahman-sired F₁ calves. *J. Anim. Sci.* **69**, 2395–2402. doi:10.2527/1991.6962395X
- Pfeiffer, T., Schuster, S., and Bonhoeffer, S. (2001). Cooperation and competition in the evolution of ATP-producing pathways. *Science* **292**, 504–507. doi:10.1126/SCIENCE.1058079
- Pick, J. L., Hutter, P., Ebner, C., Ziegler, A. K., Giordano, M., and Tschirren, B. (2016). Artificial selection reveals the energetic expense of producing larger eggs. *Front. Zool.* **13**, 38. doi:10.1186/S12983-016-0172-Y
- Pikó, L., and Taylor, K. D. (1987). Amounts of mitochondrial DNA and abundance of some mitochondrial gene transcripts in early mouse embryos. *Dev. Biol.* **123**, 364–374. doi:10.1016/0012-1606(87)90395-2
- Pittis, A. A., and Gabaldon, T. (2016). Late acquisition of mitochondria by a host with chimaeric prokaryotic ancestry. *Nature* **531**, 101–104.
- Pohjoismäki, J. L., Goffart, S., Taylor, R. W., Turnbull, D. M., Suomalainen, A., Jacobs, H. T., and Karhunen, P. J. (2010). Developmental and pathological changes in the human cardiac muscle mitochondrial DNA organization, replication and copy number. *PLoS One* **5**, e10426. doi:10.1371/JOURNAL.PONE.0010426
- Polejaeva, I. A., Chen, S. H., Vaught, T. D., Page, R. L., Mullins, J., Ball, S., Dai, Y., Boone, J., Walker, S., Ayares, D. L., Colman, A., and Campbell, K. H. (2000). Cloned pigs produced by nuclear transfer from adult somatic cells. *Nature* **407**, 86–90. doi:10.1038/35024082
- Poole, A. M., and Gribaldo, S. (2014). Eukaryotic origins: how and when was the mitochondrion acquired? *Cold Spring Harb. Perspect. Biol.* **6**, a015990. doi:10.1101/CSHPERSPECT.A015990
- Pujol, M., Lopez-Bejar, M., and Paramio, M. T. (2004). Developmental competence of heifer oocytes selected using the brilliant cresyl blue (BCB) test. *Theriogenology* **61**, 735–744. doi:10.1016/S0093-691X(03)00250-4
- Quintana-Cabrera, R., Mehrotra, A., Rigoni, G., and Soriano, M. E. (2017). Who and how in the regulation of mitochondrial cristae shape and function. *Biochem. Biophys. Res. Commun.* doi:10.1016/J.BBRC.2017.04.088
- Reynier, P., May-Panloup, P., Chretien, M. F., Morgan, C. J., Jean, M., Savagner, F., Barriere, P., and Malthiery, Y. (2001). Mitochondrial DNA content affects the fertilizability of human oocytes. *Mol. Hum. Reprod.* **7**, 425–429. doi:10.1093/MOLEHR/7.5.425
- Roca, J., Martinez, E., Vazquez, J. M., and Lucas, X. (1998). Selection of immature pig oocytes for homologous *in vitro* penetration assays with the brilliant cresyl blue test. *Reprod. Fertil. Dev.* **10**, 479–485. doi:10.1071/RD98060
- Rodríguez-González, E., López-Béjar, M., Velilla, E., and Paramio, M. T. (2002). Selection of prepubertal goat oocytes using the brilliant cresyl blue test. *Theriogenology* **57**, 1397–1409. doi:10.1016/S0093-691X(02)00645-3
- Ruiz-Pesini, E., Mishmar, D., Brandon, M., Procaccio, V., and Wallace, D. C. (2004). Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science* **303**, 223–226. doi:10.1126/SCIENCE.1088434
- Sagan, L. (1967). On the origin of mitosing cells. *J. Theor. Biol.* **14**, 225–274. doi:10.1016/0022-5193(67)90079-3
- Santos, T. A., El Shourbagy, S., and St John, J. C. (2006). Mitochondrial content reflects oocyte variability and fertilization outcome. *Fertil. Steril.* **85**, 584–591. doi:10.1016/J.FERTNSTERT.2005.09.017
- Sathananthan, H., Pera, M., and Trounson, A. (2002). The fine structure of human embryonic stem cells. *Reprod. Biomed. Online* **4**, 56–61. doi:10.1016/S1472-6483(10)61916-5
- Sato, M., and Sato, K. (2011). Degradation of paternal mitochondria by fertilization-triggered autophagy in *C. elegans* embryos. *Science* **334**, 1141–1144. doi:10.1126/SCIENCE.1210333
- Sato, M., and Sato, K. (2013). Maternal inheritance of mitochondrial DNA by diverse mechanisms to eliminate paternal mitochondrial DNA. *Biochim. Biophys. Acta* **1833**, 1979–1984. doi:10.1016/J.BBAMCR.2013.03.010
- Sbisà, E., Tanzariello, F., Reyes, A., Pesole, G., and Saccone, C. (1997). Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. *Gene* **205**, 125–140. doi:10.1016/S0378-1119(97)00404-6
- 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. *Ann. Neurol.* **63**, 35–39. doi:10.1002/ANA.21217
- Schultz, J., Waterstradt, R., Kantowski, T., Rickmann, A., Reinhardt, F., Sharoyko, V., Mulder, H., Tiedge, M., and Baltrusch, S. (2016). Precise expression of Fis1 is important for glucose responsiveness of beta cells. *J. Endocrinol.* **230**, 81–91. doi:10.1530/JOE-16-0111
- Schutz, M. M., Freeman, A. E., Lindberg, G. L., Koehler, C. M., and Beitz, D. C. (1994). The effect of mitochondrial DNA on milk production and health of dairy cattle. *Livest. Prod. Sci.* **37**, 283–295. doi:10.1016/0301-6226(94)90123-6
- Seidel, G. E., Jr. (2015). Lessons from reproductive technology research. *Annu. Rev. Anim. Biosci.* **3**, 467–487. doi:10.1146/ANNUREV-ANI-MAL-031412-103709
- Sharpley, M. S., Marciniak, C., Eckel-Mahan, K., McManus, M., Crimi, M., Waymire, K., Lin, C. S., Masubuchi, S., Friend, N., Koike, M., Chalkia, D., MacGregor, G., Sassone-Corsi, P., and Wallace, D. C. (2012). Heteroplasmy of mouse mtDNA is genetically unstable and results in altered behavior and cognition. *Cell* **151**, 333–343. doi:10.1016/J.CELL.2012.09.004
- Sheshadri, P., and Kumar, A. (2016). Managing odds in stem cells: insights into the role of mitochondrial antioxidant enzyme MnSOD. *Free Radic. Res.* **50**, 570–584. doi:10.3109/10715762.2016.1155708
- Shin, T., Kraemer, D., Pryor, J., Liu, L., Rugila, J., Howe, L., Buck, S., Murphy, K., Lyons, L., and Westhusin, M. (2002). A cat cloned by nuclear transplantation. *Nature* **415**, 859. doi:10.1038/NATURE723
- Shutt, T. E., and Shadel, G. S. (2010). A compendium of human mitochondrial gene expression machinery with links to disease. *Environ. Mol. Mutagen.* **51**, 360–379.
- Sloan, D. B., Fields, P. D., and Havird, J. C. (2015). Mitonuclear linkage disequilibrium in human populations. *Proc. Biol. Sci.* **282**, 20151704. doi:10.1098/RSPB.2015.1704
- Sloan, D. B., Havird, J. C., and Sharbrough, J. (2017). The on-again, off-again relationship between mitochondrial genomes and species boundaries. *Mol. Ecol.* **26**, 2212–2236. doi:10.1111/MEC.13959
- Song, W. H., Yi, Y. J., Sutovsky, M., Meyers, S., and Sutovsky, P. (2016). Autophagy and ubiquitin-proteasome system contribute to sperm mitochondria after mammalian fertilization. *Proc. Natl Acad. Sci. USA* **113**, E5261–E5270. doi:10.1073/PNAS.1605844113
- Spelbrink, J. N., Li, F. Y., Tiranti, V., Nikali, K., Yuan, Q. P., Tariq, M., Wanrooij, S., Garrido, N., Comi, G., Morandi, L., Santoro, L., Toscano, A., Fabrizi, G. M., Somer, H., Croxen, R., Beeson, D., Poulton, J., Suomalainen, A., Jacobs, H. T., Zeviani, M., and Larsson, C. (2001).

- Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria. *Nat. Genet.* **28**, 223–231. doi:10.1038/90058
- Spikings, E. C., Alderson, J., and St John, J. C. (2006). Transmission of mitochondrial DNA following assisted reproduction and nuclear transfer. *Hum. Reprod. Update* **12**, 401–415. doi:10.1093/HUMUPD/DML011
- Spikings, E. C., Alderson, J., and St John, J. C. (2007). Regulated mitochondrial DNA replication during oocyte maturation is essential for successful porcine embryonic development. *Biol. Reprod.* **76**, 327–335. doi:10.1095/BIOLREPROD.106.054536
- Spiropoulos, J., Turnbull, D. M., and Chinnery, P. F. (2002). Can mitochondrial DNA mutations cause sperm dysfunction? *Mol. Hum. Reprod.* **8**, 719–721. doi:10.1093/MOLEHR/8.8.719
- Srirattana, K., and St John, J. C. (2017). Manipulating the mitochondrial genome to enhance cattle embryo development. *G3 (Bethesda)* **7**, 2065–2080. doi:10.1534/G3.117.042655
- Srirattana, K., McCosker, K., Schatz, T., and St John, J. C. (2017). Cattle phenotypes can disguise their maternal ancestry. *BMC Genet.* **18**, 59. doi:10.1186/S12863-017-0523-5
- St John, J. C. (2012). Transmission, inheritance and replication of mitochondrial DNA in mammals: implications for reproductive processes and infertility. *Cell Tissue Res.* **349**, 795–808. doi:10.1007/S00441-012-1444-2
- St John, J. (2014). The control of mtDNA replication during differentiation and development. *Biochim. Biophys. Acta* **1840**, 1345–1354. doi:10.1016/J.BBAGEN.2013.10.036
- St John, J. C., and Schatten, G. (2004). Paternal mitochondrial DNA transmission during nonhuman primate nuclear transfer. *Genetics* **167**, 897–905. doi:10.1534/GENETICS.103.025049
- St John, J. C., Cooke, I. D., and Barratt, C. L. (1997). Mitochondrial mutations and male infertility. *Nat. Med.* **3**, 124–125. doi:10.1038/NM0297-124C
- St John, J. C., Sakkas, D., and Barratt, C. L. (2000). A role for mitochondrial DNA and sperm survival. *J. Androl.* **21**, 189–199.
- St John, J. C., Jokhi, R. P., and Barratt, C. L. (2001). Men with oligoasthenoteratozoospermia harbour higher numbers of multiple mitochondrial DNA deletions in their spermatozoa, but individual deletions are not indicative of overall aetiology. *Mol. Hum. Reprod.* **7**, 103–111. doi:10.1093/MOLEHR/7.1.103
- St John, J. C., Jokhi, R. P., and Barratt, C. L. (2005a). The impact of mitochondrial genetics on male infertility. *Int. J. Androl.* **28**, 65–73. doi:10.1111/J.1365-2605.2005.00515.X
- St John, J. C., Moffatt, O., and D'Souza, N. (2005b). Aberrant heteroplasmic transmission of mtDNA in cloned pigs arising from double nuclear transfer. *Mol. Reprod. Dev.* **72**, 450–460. doi:10.1002/MRD.20370
- St John, J. C., Facucho-Oliveira, J., Jiang, Y., Kelly, R., and Salah, R. (2010). Mitochondrial DNA transmission, replication and inheritance: a journey from the gamete through the embryo and into offspring and embryonic stem cells. *Hum. Reprod. Update* **16**, 488–509. doi:10.1093/HUMUPD/DMQ002
- Steinborn, R., Schinogl, P., Wells, D. N., Berghaler, A., Muller, M., and Brem, G. (2002). Coexistence of *Bos taurus* and *B. indicus* mitochondrial DNAs in nuclear transfer-derived somatic cattle clones. *Genetics* **162**, 823–829.
- Stewart, J. B., Freyer, C., Elson, J. L., and Larsson, N. G. (2008). Purifying selection of mtDNA and its implications for understanding evolution and mitochondrial disease. *Nat. Rev. Genet.* **9**, 657–662. doi:10.1038/NRG2396
- Stigliani, S., Persico, L., Lagazio, C., Anserini, P., Venturini, P. L., and Scaruffi, P. (2014). Mitochondrial DNA in Day 3 embryo culture medium is a novel, non-invasive biomarker of blastocyst potential and implantation outcome. *Mol. Hum. Reprod.* **20**, 1238–1246. doi:10.1093/MOLEHR/GAU086
- Storey, B. T. (1980). Strategy of oxidative metabolism in bull spermatozoa. *J. Exp. Zool.* **212**, 61–67. doi:10.1002/JEZ.1402120109
- Storey, B. T., and Kayne, F. J. (1975). Energy metabolism of spermatozoa. V. The Embden–Myerhof pathway of glycolysis: activities of pathway enzymes in hypotonically treated rabbit epididymal spermatozoa. *Fertil. Steril.* **26**, 1257–1265. doi:10.1016/S0015-0282(16)41543-8
- Suissa, S., Wang, Z., Poole, J., Wittkopp, S., Feder, J., Shutt, T. E., Wallace, D. C., Shadel, G. S., and Mishmar, D. (2009). Ancient mtDNA genetic variants modulate mtDNA transcription and replication. *PLoS Genet.* **5**, e1000474. doi:10.1371/JOURNAL.PGEN.1000474
- Sun, X., and St John, J. C. (2016). The role of the mtDNA set point in differentiation, development and tumorigenesis. *Biochem. J.* **473**, 2955–2971. doi:10.1042/BCJ20160008
- Sutarno, J., Cummins, J. M., Greeff, J., and Lymbery, A. J. (2002). Mitochondrial DNA polymorphisms and fertility in beef cattle. *Theriogenology* **57**, 1603–1610. doi:10.1016/S0093-691X(02)00664-7
- Sutovsky, P., Moreno, R. D., Ramalho-Santos, J., Dominko, T., Simerly, C., and Schatten, G. (1999). Ubiquitin tag for sperm mitochondria. *Nature* **402**, 371–372. doi:10.1038/46466
- Takeda, K., Akagi, S., Kaneyama, K., Kojima, T., Takahashi, S., Imai, H., Yamanaka, M., Onishi, A., and Hanada, H. (2003). Proliferation of donor mitochondrial DNA in nuclear transfer calves (*Bos taurus*) derived from cumulus cells. *Mol. Reprod. Dev.* **64**, 429–437. doi:10.1002/MRD.10279
- Takeda, K., Tasai, M., Iwamoto, M., Akita, T., Tagami, T., Nirasawa, K., Hanada, H., and Onishi, A. (2006). Transmission of mitochondrial DNA in pigs and progeny derived from nuclear transfer of Meishan pig fibroblast cells. *Mol. Reprod. Dev.* **73**, 306–312. doi:10.1002/MRD.20403
- Takeda, K., Srirattana, K., Matsukawa, K., Akagi, S., Kaneda, M., Tasai, M., Nirasawa, K., Pinkert, C. A., Parnpai, R., and Nagai, T. (2012). Influence of intergeneric/interspecies mitochondrial injection; parthenogenetic development of bovine oocytes after injection of mitochondria derived from somatic cells. *J. Reprod. Dev.* **58**, 323–329. doi:10.1262/JRD.2011-013
- Takeo, S., Goto, H., Kuwayama, T., Monji, Y., and Iwata, H. (2013). Effect of maternal age on the ratio of cleavage and mitochondrial DNA copy number in early developmental stage bovine embryos. *J. Reprod. Dev.* **59**, 174–179. doi:10.1262/JRD.2012-148
- Tamassia, M., Nuttinck, F., May-Panloup, P., Reynier, P., Heyman, Y., Charpigny, G., Stojkovic, M., Hiendleder, S., Renard, J. P., and Chastant-Maillard, S. (2004). *In vitro* embryo production efficiency in cattle and its association with oocyte adenosine triphosphate content, quantity of mitochondrial DNA, and mitochondrial DNA haplogroup. *Biol. Reprod.* **71**, 697–704. doi:10.1095/BIOLREPROD.103.026104
- Tanaka, M., Gong, J. S., Zhang, J., Yoneda, M., and Yagi, K. (1998). Mitochondrial genotype associated with longevity. *Lancet* **351**, 185–186. doi:10.1016/S0140-6736(05)78211-8
- Tang, L., Gonzalez, R., and Dobrinski, I. (2015). Germline modification of domestic animals. *Anim. Reprod.* **12**, 93–104.
- Tannus, S., Son, W. Y., Gilman, A., Younes, G., Shavit, T., and Dahan, M. H. (2017). The role of intracytoplasmic sperm injection in non-male factor infertility in advanced maternal age. *Hum. Reprod.* **32**, 119–124.
- Thongphakdee, A., Kobayashi, S., Imai, K., Inaba, Y., Tasai, M., Tagami, T., Nirasawa, K., Nagai, T., Saito, N., Techakumphu, M., and Takeda, K. (2008). Interspecies nuclear transfer embryos reconstructed from cat somatic cells and bovine ooplasm. *J. Reprod. Dev.* **54**, 142–147. doi:10.1262/JRD.19159
- Thundathil, J., Fillion, F., and Smith, L. C. (2005). Molecular control of mitochondrial function in preimplantation mouse embryos. *Mol. Reprod. Dev.* **71**, 405–413. doi:10.1002/MRD.20260
- Tsai, T., and St John, J. C. (2016). The role of mitochondrial DNA copy number, variants, and haplotypes in farm animal developmental outcome.

- Domest. Anim. Endocrinol.* **56**(Suppl.), S133–S146. doi:10.1016/J.DOMANIEND.2016.03.005
- Tsai, T. S., Rajasekar, S., and St John, J. C. (2016). The relationship between mitochondrial DNA haplotype and the reproductive capacity of domestic pigs (*Sus scrofa domestica*). *BMC Genet.* **17**, 67. doi:10.1186/S12863-016-0375-4
- Tzeng, C. R., Hsieh, R. H., Au, H. K., Yen, Y. H., Chang, S. J., and Cheng, Y. F. (2004). Mitochondria transfer (MIT) into oocyte from autologous cumulus granulosa cells (cGCs). *Fertil. Steril.* **82**, S53. doi:10.1016/J.FERTNSTERT.2004.07.136
- Ursing, B. M., and Arnason, U. (1998). The complete mitochondrial DNA sequence of the pig (*Sus scrofa*). *J. Mol. Evol.* **47**, 302–306. doi:10.1007/PL00006388
- Van Blerkom, J. (2004). Mitochondria in human oogenesis and preimplantation embryogenesis: engines of metabolism, ionic regulation and developmental competence. *Reproduction* **128**, 269–280. doi:10.1530/REP.1.00240
- Vander Heiden, M. G., Cantley, L. C., and Thompson, C. B. (2009). Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**, 1029–1033. doi:10.1126/SCIENCE.1160809
- Wai, T., Teoli, D., and Shoubridge, E. A. (2008). The mitochondrial DNA genetic bottleneck results from replication of a subpopulation of genomes. *Nat. Genet.* **40**, 1484–1488. doi:10.1038/NG.258
- Wai, T., Ao, A., Zhang, X., Cyr, D., Dufort, D., and Shoubridge, E. A. (2010). The role of mitochondrial DNA copy number in mammalian fertility. *Biol. Reprod.* **83**, 52–62. doi:10.1095/BIOLREPROD.109.080887
- Wakayama, T., Perry, A. C., Zuccotti, M., Johnson, K. R., and Yanagimachi, R. (1998). Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. *Nature* **394**, 369–374. doi:10.1038/28615
- Wallace, D. C. (2010). Colloquium paper: bioenergetics, the origins of complexity, and the ascent of man. *Proc. Natl Acad. Sci. USA* **107**(Suppl. 2), 8947–8953. doi:10.1073/PNAS.0914635107
- Wang, Z., Figueiredo-Pereira, C., Oudot, C., Vieira, H. L., and Brenner, C. (2017). Mitochondrion: a common organelle for distinct cell deaths? *Int. Rev. Cell Mol. Biol.* **331**, 245–287. doi:10.1016/BS.IRCMB.2016.09.010
- Wanrooij, P. H., Uhler, J. P., Simonsson, T., Falkenberg, M., and Gustafsson, C. M. (2010). G-Quadruplex structures in RNA stimulate mitochondrial transcription termination and primer formation. *Proc. Natl Acad. Sci. USA* **107**, 16072–16077. doi:10.1073/PNAS.1006026107
- Warburg, O. (1956a). On respiratory impairment in cancer cells. *Science* **124**, 269–270.
- Warburg, O. (1956b). On the origin of cancer cells. *Science* **123**, 309–314. doi:10.1126/SCIENCE.123.3191.309
- Waring, M. J. (1965). Complex formation between ethidium bromide and nucleic acids. *J. Mol. Biol.* **13**, 269–282. doi:10.1016/S0022-2836(65)80096-1
- Watanabe, S., and Nagai, T. (2008). Health status and productive performance of somatic cell cloned cattle and their offspring produced in Japan. *J. Reprod. Dev.* **54**, 6–17. doi:10.1262/JRD.19090
- West, A. P., Shadel, G. S., and Ghosh, S. (2011). Mitochondria in innate immune responses. *Nat. Rev. Immunol.* **11**, 389–402. doi:10.1038/NRI2975
- Wilmot, I., Schnieke, A. E., McWhir, J., Kind, A. J., and Campbell, K. H. (1997). Viable offspring derived from fetal and adult mammalian cells. *Nature* **385**, 810–813. doi:10.1038/385810A0
- Wittayarat, M., Sato, Y., Do, L. T., Chatdarong, K., Tharasanit, T., Techakumphu, M., Taniguchi, M., and Otoi, T. (2017). Epigenetic modulation on cat–cow interspecies somatic cell nuclear transfer embryos by treatment with trichostatin A. *Anim. Sci. J.* **88**, 593–601. doi:10.1111/ASJ.12676
- Wolff, J. N., Ladoukakis, E. D., Enriquez, J. A., and Dowling, D. K. (2014). Mitonuclear interactions: evolutionary consequences over multiple biological scales. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369**, 20130443. doi:10.1098/RSTB.2013.0443
- Wray, J., Kalkan, T., and Smith, A. G. (2010). The ground state of pluripotency. *Biochem. Soc. Trans.* **38**, 1027–1032. doi:10.1042/BST0381027
- Wu, Y. G., Liu, Y., Zhou, P., Lan, G. C., Han, D., Miao, D. Q., and Tan, J. H. (2007). Selection of oocytes for *in vitro* maturation by brilliant cresyl blue staining: a study using the mouse model. *Cell Res.* **17**, 722–731. doi:10.1038/CR.2007.66
- Xu, B., Guo, N., Zhang, X. M., Shi, W., Tong, X. H., Iqbal, F., and Liu, Y. S. (2015). Oocyte quality is decreased in women with minimal or mild endometriosis. *Sci. Rep.* **5**, 10779. doi:10.1038/SREP10779
- Yu, G., Xiang, H., Tian, J., Yin, J., Pinkert, C. A., Li, Q., and Zhao, X. (2015). Mitochondrial haplotypes influence metabolic traits in porcine transmissible cybrids. *Sci. Rep.* **5**, 13118. doi:10.1038/SREP13118
- Yu, Z., O'Farrell, P. H., Yakubovich, N., and DeLuca, S. Z. (2017). The mitochondrial DNA polymerase promotes elimination of paternal mitochondrial genomes. *Curr. Biol.* **27**, 1033–1039. doi:10.1016/J.CUB.2017.02.014
- Zhang, H., Meng, L. H., and Pommier, Y. (2007). Mitochondrial topoisomerases and alternative splicing of the human TOP1mt gene. *Biochimie* **89**, 474–481. doi:10.1016/J.BIOCHI.2006.11.002
- Zouros, E., Freeman, K. R., Ball, A. O., and Pogson, G. H. (1992). Direct evidence for extensive paternal mitochondrial DNA inheritance in the marine mussel *Mytilus*. *Nature* **359**, 412–414. doi:10.1038/359412A0

Appendix C

Analysis of the Mitochondrial DNA and Its Replicative Capacity in Induced Pluripotent Stem Cells

Gael Cagnone, Vijesh Vaghjiani, William Lee, Claire Sun, Jacqueline Johnson, Ka-Yu Yeung, and Justin C. St. John

Abstract

The mitochondrial genome resides in the mitochondrion of nearly all mammalian cells. It is important for energy production as it encodes 13 of the key subunits of the electron transfer chain, which generates the vast majority of cellular ATP through the process of oxidative phosphorylation. As cells establish pluripotency, they regulate their mtDNA copy number so that they possess few copies but sufficient that they can be replicated to match the differentiated cell-specific requirements for ATP derived through oxidative phosphorylation. However, the failure to strictly regulate this process prevents pluripotent cells from differentiating. We describe a series of protocols that analyze mtDNA copy number, DNA methylation within the nuclear-encoded mtDNA-specific polymerase, and gene expression of the other factors that drive replication of the mitochondrial genome. We demonstrate how to measure ATP-generating capacity through oxygen respiratory capacity and total cellular ATP and lactate levels. Finally, we also describe how to detect mtDNA variants in pluripotent and differentiating cells using next-generation sequencing protocols and how the variants can be confirmed by high-resolution melt analysis.

Keywords: Mitochondrial DNA, Variants, Next-generation sequencing, Pluripotent cells, Differentiation, Transcription and replication, Oxygen consumption, ATP

1 Introduction

The mitochondrial genome is 16.6 kb in size and encodes 13 of the subunits of the electron transfer chain, which is the intramitochondrial apparatus that generates the majority of cellular ATP through the biochemical process of oxidative phosphorylation (OXPHOS) (1). Mitochondrial DNA (mtDNA) also encodes 22 tRNAs and 2 rRNAs and has one noncoding region, known as the D-loop. This is the mitochondrial genome's control region, where the nuclear-encoded transcription and replication factors that translocate to the mitochondrion functionally interact with mtDNA to initiate the first transcription and then replication of this bacterial-originating genome (2). The D-loop also has two hypervariable regions, which are characterized by sequences specific to maternal relatives and demonstrate the maternal inheritance of mtDNA (3).

Embryonic and induced pluripotent stem cells are excellent models to study and understand mitochondrial transcription and replication events during early development (4). mtDNA is present in large numbers in the egg at the time of fertilization (5). Once the egg has been fertilized, the number of mtDNA copies per cell is significantly diluted out as embryonic cells progress through pre-implantation development to the blastocyst stage (5). At the blastocyst stage, mtDNA replication is restricted to the trophoblast cells. These are the first cells of the embryo to undergo differentiation. The inner cell mass cells, the cells that give rise to the fetus and embryonic stem cells, continue to dilute out their mtDNA copy number until levels are between 200 and 400 copies per cell (6). This is a significant developmental step, which establishes the *mtDNA set point* (6, 7).

The mtDNA set point in embryonic stem cells is defined as the state where all undifferentiated cells are under the control of the OCT4-SOX2-NANOG network and possess few copies of mtDNA. This promotes their ability to proliferate and remain pluripotent but, at the same time, maintain the potential to undergo differentiation and expand their mtDNA copy number. Consequently, the mtDNA set point ensures that all cells can expand their mtDNA copy number to meet the metabolic demands of the specialized cell that they fully differentiate into. This allows heart and muscle precursor cells to propagate high numbers of mtDNA copy to meet their high demands for OXPHOS-derived ATP, while blood and other cells from the spleen have fewer copies as their metabolic requirements are primarily fueled through glycolysis (6). The increases in mtDNA copy number take place in a synchronous manner with the expression of key genes during development. For example, during astrocyte differentiation, mtDNA copy number will remain low as cells continue to express NCAM, NESTIN, MUSASHI1, and CD133. However, as cells start to express endpoint markers, such as GFAP, they increase mtDNA copy number significantly (7).

Assessment of mtDNA copy number is also essential to determine whether a somatic cell that has been reprogrammed to a pluripotent state, i.e., induced pluripotent stem cells, has regulated their mtDNA copy number to low levels as part of the induction of dedifferentiation (8). It is equally important to determine whether induced pluripotent stem cells have the potential to increase their mtDNA copy number in a cell-specific manner during differentiation as a means of assessing their capability to fully differentiate.

mtDNA copy number is regulated by nuclear-encoded transcription and replication factors that translocate to the mitochondrion to drive firstly transcription and then replication (2). The key players in mitochondrial replication are mitochondrial transcription factor A (TFAM); the mitochondrial DNA-specific helicase, TWINKLE; the mitochondrial single-stranded binding protein

(MTSSB); and the mtDNA-specific polymerase, DNA polymerase gamma (POLG). POLG consists of two subunits, one is the accessory subunit known as POLGB, and the other is the catalytic subunit POLGA. One key developmental process is the regulation of POLGA expression during differentiation and development, which is mediated by DNA methylation, again in a cell-specific manner (6). The levels of DNA methylation during differentiation can be assessed by measuring the levels of 5-methylcytosine (5-mC) and 5-hydroxymethylcytosine (5-hmC), which are indicative of active and transient DNA methylation (9).

It is becoming increasingly important to measure the levels of mutations, which are present in stem cells that have been derived from patients carrying a mitochondrial disease, and define the selection of background mtDNA variants that are preferentially amplified during the induction of pluripotency. This could be performed using next-generation sequencing protocols, such as those used in conjunction with the Ion Torrent PGM or Illumina platforms.

Here, we describe a series of essential assays that define mtDNA replication, copy number, and ATP-generating capacity, which can be determined through measurement of oxygen consumption rates with an OROBOROS Oxygraph-2k, and total ATP and lactate assays. mtDNA copy number can be assessed through real-time PCR, while the most appropriate mechanism for measuring 5-mC and 5-hmC activities is through the use of immunoprecipitation of methylated DNA (MeDIP). When assessing the interactions of the nuclear transcription and replication factors that drive transcription and replication of the mitochondrial genome, it is essential to assess these genes in terms of their levels of expression. This can be effectively performed by using a real-time RT-PCR assay and comparing the levels of expression to known housekeeping genes. We favor the use of the Ion Torrent PGM with the CLC Genomics Workbench bioinformatics package for detecting mtDNA variants. Our protocols are sensitive to very low levels (approx. 3 %), which will provide an overall determination of the level of variants present. The variants can then be verified using high-resolution melt (HRM) analysis.

2 Materials

2.1 Culture of Human-Induced Pluripotent Stem Cells

1. Dulbecco's Modified Eagle Medium/Nutrient Mixture F12 with GlutaMAX (DMEM/F12) (Life Technologies; cat. no. 10565-018).
2. KnockOut serum replacement (Life Technologies; cat. no. 10828-028).

3. MEM nonessential amino acids (100×) (Life Technologies; cat. no. 11140-050).
4. 55 mM β -mercaptoethanol (Life Technologies; cat. no. 21985-023).
5. Dulbecco's phosphate-buffered saline (PBS) (Life Technologies; cat. no. 14190-144).
6. 10 μ g/mL basic fibroblast growth factor (bFGF) solution (bFGF powder; Millipore; cat. no. GF003, dissolved in PBS).
7. Serum replacement (SR) medium: For 250 mL, combine 200 mL of DMEM/F12 with GlutaMAX, 50 mL KnockOut serum replacement, 2.5 mL MEM nonessential amino acids, and 450 μ L β -mercaptoethanol and filter sterilize to 0.22 μ m. The medium is supplemented with bFGF to a final concentration of 10 ng/mL just prior to use.
8. 4 mg/mL collagenase I solution (collagenase I powder; Worthington; cat. no. 4196, dissolved in DMEM/F12 medium).
9. 10 mg/mL dispase solution (dispase powder; Life Technologies; cat. no. 17105-041, dissolved in SR medium).
10. 0.1 % gelatin in sterile water (Sigma-Aldrich; cat. no. G1890).
11. 60-mm culture dishes (Falcon; cat. no. 353037).
12. 100-mm culture dishes (Falcon; cat. no. 353003).
13. 25-cm² culture flasks (Falcon; cat. no. 353108).
14. 15-mL sterile conical tube (Corning; cat. no. 430791).
15. Sterile Eppendorf tube (Axygen; cat. no. MCT175C).
16. 6-well ultralow attachment plates (Corning; cat. no. 3471).

2.2 Oxygen Consumption Measurement

1. Polarographic oxygen sensors (the Oxygraph-2K, OROBOROS INSTRUMENTS Corp., Innsbruck, Austria) with 10- and 50- μ L syringes.
2. Hank's balanced buffer solution (Life Technologies, #14025076) and sodium hydrosulfite (Sigma-Aldrich; cat. no. 157953).
3. Reagents: oligomycin (Omy, 5 mg/mL, #75351), carbonyl cyanide *p*-(trifluoro-methoxy) phenyl-hydrazine (FCCP, 10 μ M, #C2920), antimycin A (Ama, 5 mM, #A0274), and rotenone (Rot, 1 mM, #R8875). Reagents can be purchased from Sigma-Aldrich and are dissolved in 100 % EtOH. Store reagents at -20°C .

2.3 ATP Assay

1. Luminescence plate reader. FLUOstar OPTIMA (BMG LAB-TECH, Offenburg, Germany) may be used to record luminescence.

2. 96-well white plate (Bio-Rad, #HSP-9601).
3. ATPlite luminescence assay system (PerkinElmer, #6016941). This kit contains cell lysis solution, substrate buffer solution, ATP standard, substrate solution (luciferase/luciferin). Store kit at 4 °C.

2.4 Lactate Assay

1. Absorbance plate reader. FLUOstar OPTIMA (BMG LAB-TECH, Offenburg, Germany) may be used to record absorbance.
2. 96-well transparent plate (U-shaped well) (Thermo Scientific, #262162).
3. Lactate colorimetric assay kit. Lactate Colorimetric Assay Kit II (BioVision, #K627-100) may be used. This kit contains L(+)-lactate standards, lactate enzyme mix, lactate substrate mix, and lactate assay buffer. Store kit at -20 °C.

2.5 mtDNA Copy Number

1. BIOTAQ™ DNA polymerase kit (Bioline; cat. no. BIO-21040).
2. dNTP (Bioline; cat. no. BIO-39028).
3. Primers (Sigma-Aldrich) (*see* primer, Table 1).
4. [QIAquick Gel Extraction Kit](#) (Qiagen; cat. no. 28704).

2.6 Immunoprecipitation of Methylated DNA (MeDIP)

1. Covaris Adaptive Focused Acoustics (AFA™) S220 system including Covaris SonoLab 7 software (Applied Biosystems; cat. no. 4465653) with Covaris milliTUBE (Covaris, cat. no. 520056).
2. Agarose gel is made with agarose (Bioline; cat. no. BIO-41025) with 1× TE buffer.
3. 10× IP buffer consists of 100 mM sodium phosphate (pH 7.0; Sigma-Aldrich; cat. no. 342483), 1.4 M NaCl (Sigma-Aldrich; cat. no. S5886), and 0.5 % Triton X-100 (Sigma-Aldrich; cat. no. T8532).
4. Anti-5mC antibody (Active Motif; cat. no. 39649) and anti-5hmC antibody (Active Motif; cat. no. 39999).
5. Dynabeads® Protein G (Life Technologies; cat. no. 10009D).
6. PBS/BSA buffer is made with 0.1 % BSA (Sigma-Aldrich; cat. no. A8806) in 1× PBS (Life Technologies; cat. no. 14190-144).
7. Magnetic rack (Invitrogen; cat. no. 123-21D).
8. 10× proteinase K digestion buffer consists of 50 mM Tris-HCl (pH 8.0; Merck; cat. no. 1082191000), 10 mM EDTA (pH 8.0; Sigma-Aldrich; cat. no. ED), and 1.0 % SDS (Sigma-Aldrich; cat. no. L6026).
9. Proteinase K (Bioline; cat. no. BIO-37037).
10. QIAquick PCR purification kit (Qiagen; cat. no. 28104).

Table 1
Primer sequences for gene expression and copy number and their melting temperature

Gene	Forward primers	Reverse primers	Product size (bp)	Annealing temp (°C)
<i>β-ACTIN</i>	CAAAACCTAACTTGCGCAGA	TTTTAGGATGGCAAGGGACT	261	53
<i>MTSSB</i>	GAAGCCATGTTTCGAAGACCTG	CTGATATTCTGTGCCATGTTGTC	271	57
<i>POLGA</i>	CACACCTAAACTCATGGCAC	GTCCACGTCGTTGTAAGGTC	436	56
<i>POLGB</i>	GTTTGCCATGAGTCCCATCTAAC	CTCTGTCAGCTGGAAAAGAATC	280	56
<i>TEAM</i>	ATTGGGGTCGGGTCACTGCCCTCA	TACCTGCCACTCCGCCCTATAAGC	361	63
<i>TWINKLE</i>	GCACAAGTCCATCGTATCTTTC	CATACTCACTGATGAATGTCGTC	197	56
<i>β-GLOBIN</i>	CAACTTCATCCACGTTCAAC	GAAAGAGCCAAGGACAGGTAC	268	57
<i>mtDNA</i>	CGAAAGGACAAGAGAAATAAGG	CTGTAAAGTTTTTAAGTTTTTAATGCG	152	53

2.7 Analysis of Gene Expression of mtDNA Replication Factors and mtDNA-Encoded Genes

2.7.1 cDNA Synthesis

1. QIAshredder columns (Qiagen; cat. no. 79654).
2. RNeasy Plus Mini Kit (Qiagen; cat. no. 74104).
3. DNase I kit (Qiagen; cat. no. 79254).
4. Molecular grade ethanol (Sigma-Aldrich; cat. no. E7023).
5. SuperScript III cDNA synthesis kit (Life Technologies; cat. no. 18080-051).

2.7.2 Real-Time Quantitative PCR

1. SensiMix Hi-ROX (Bioline; cat. no. QT605-20).
2. Gene-specific primers (Sigma-Aldrich) (*see* primer, Table 1).

2.8 Analysis of mtDNA Variants by Next-Generation Sequencing

2.8.1 Long PCR

1. Platinum *Taq* high-fidelity DNA polymerase kit (Invitrogen/Life Technologies; cat. no. 11304-029). The kit is composed of:
 - (a) Platinum *Taq* DNA polymerase high fidelity (5 U/ μ L)
 - (b) 10 \times high-fidelity PCR buffer
 - (c) 50 mM MgSO₄
2. 10 mM each dNTP mix (Bioline).
3. Forward primer 1: 5'-GACGGGCTCACATCACCCCATAA-3' (10 μ M, Sigma-Aldrich).
4. Reverse primer 1: 5'-GCGTACGGCCAGGGCTATTGGT-3' (10 μ M, Sigma-Aldrich).
5. Forward primer 2: 5'-GCCACAACCTCCTCG-GACTCCT-3' (10 μ M, Sigma-Aldrich).
6. Reverse primer 2: 5'-GGTGGCTGGCACGAAATTGACC-3' (10 μ M, Sigma-Aldrich).
7. Autoclaved sterilized ddH₂O.
8. 0.2-mL sterile polypropylene PCR tubes (Axygen/Corning; cat. no. 6571).

2.8.2 Purification of PCR Products

1. QIAquick PCR purification kit (Qiagen; cat. no. 28104). The kit is composed of:
 - (a) QIAquick spin columns
 - (b) Buffer PB
 - (c) Buffer PE
 - (d) Buffer EB
 - (e) pH Indicator I
 - (f) Collection tubes (2 mL)
 - (g) Loading dye
2. 3 M sodium acetate, pH 5.0 (Sigma-Aldrich; cat. no. S2889).
3. Autoclaved sterilized ddH₂O.
4. 1.7-mL sterile polypropylene Snaplock microcentrifuge tubes (Axygen/Corning; cat. no. MCT-175-C).

2.8.3 Quantification of Double-Stranded DNA

1. Qubit 2.0 Fluorometer (Invitrogen/Life Technologies; cat. no. Q32866).
2. Qubit dsDNA HS Assay Kit (Invitrogen/Life Technologies; cat. no. Q32851). The kit is composed of:
 - (a) Qubit dsDNA HS reagent (200× concentrate in DMSO)
 - (b) Qubit dsDNA HS buffer
 - (c) Qubit dsDNA HS Standard #1 (0 ng/μL in TE buffer)
 - (d) Qubit dsDNA HS Standard #2 (10 ng/μL in TE buffer)
3. Qubit assay tubes (Invitrogen/Life Technologies; cat. no. Q32856)

2.8.4 Shearing of Long PCR Amplicons

1. Covaris Adaptive Focused Acoustics (AFA™) S220 system including Covaris SonoLab 7 software (Applied Biosystems; cat. no. 4465653).
2. Covaris microTUBE (Covaris; cat. no. 520096).

2.8.5 Sample Purification

1. Agencourt AMPure XP kit (Beckman Coulter: cat. no. A63881), which is composed of Agencourt AMPure XP reagent beads.
2. DynaMag-2 magnetic rack (Invitrogen; cat. no. 123-21D).
3. Low TE buffer (Ion Fragment Library Kit).
4. 1.5-mL LoBind tubes (Eppendorf; cat. no. 022431021).

2.8.6 Size Selection of the DNA Library

1. E-Gel iBase and E-Gel Safe Imager Combo Kit (Invitrogen; cat. no. G6465).
2. E-Gel SizeSelect 2 % agarose gel (Invitrogen; cat. no. G6610-02).
3. Low TE buffer (Ion Fragment Library Kit).
4. 100-bp DNA ladder (Invitrogen; cat. no. 15628-019).

2.8.7 Quantification of the DNA Libraries

1. Agilent 2100 Bioanalyzer, which includes the Agilent 2100 Expert Software (Agilent Technologies; cat. no. G2947CA). This comprises:
 - (a) Chip-priming station
 - (b) IKA vortex mixer
2. Agilent High Sensitivity DNA Kit (Agilent Technologies; cat. no. 5067-4626), which is composed of:
 - (a) High sensitivity DNA chips
 - (b) Electrode cleaner
 - (c) Syringe
 - (d) Spin filters
 - (e) High sensitivity DNA ladder
 - (f) High sensitivity DNA markers 35/10380 bp
 - (g) High sensitivity DNA gel matrix

2.8.8 DNA Template Amplification Through Emulsion PCR

This includes materials for the preparation of the IKA DT-20 solution and the Ion Sphere Particles, the emulsion, emulsion PCR, and recovery of the Ion Sphere Particles.

1. Ion Xpress Template Kit (Invitrogen/Life Technologies; cat. no. 4466457). This is composed of:
 - (a) Emulsion oil
 - (b) Ion Sphere wash solution
2. Ion Template Reagents Kit (Invitrogen/Life Technologies; cat. no. 4466462). This is composed of:
 - (a) Polymerase
 - (b) Amplification buffer
 - (c) MgCl_2 solution
 - (d) dNTPs
 - (e) Thermostable pyrophosphatase (TAP)
 - (f) Clonal amplification primer mix
 - (g) Enrichment primer
 - (h) Ion Sphere Particles
3. Ion Template Solutions Kit (Invitrogen/Life Technologies; cat. no. 4466463). This comprises:
 - (a) Recovery solution
 - (b) Wash solution
4. Eppendorf Combitips Plus, 10 mL, sterile (Eppendorf; cat. no. 022496123).
5. IKA ULTRA-TURRAX Tube Drive for the PGM System (Ion Torrent; cat. no. 4464747).
6. Repeater Plus pipette (Eppendorf; cat. no. 022260201).
7. MicroAmp Optical 96-Well Reaction Plates (Applied Biosystems; cat. no. N8010560).
8. MicroAmp Optical 8-Cap Strips (Applied Biosystems; cat. no. N8010535).
9. GeneAmp 96-well PCR System 9700 (Applied Biosystems; cat. no. N8050200).
10. 1-Butanol (molecular biology grade, $\geq 99\%$, Sigma-Aldrich; cat. no. B7906).
11. PVC basin (VWR; cat. no. 4415129).
12. 1.5-mL LoBind tubes (Eppendorf; cat. no. 022431021).

2.8.9 Enrichment of Template-Positive Ion Sphere Particles

1. Tween-20 (Sigma-Aldrich; cat. no. P9416).
2. 1 M NaOH (Sigma-Aldrich; cat. no. S5881).
3. DynaMag-2 Magnet (Invitrogen; cat. no. 123-21D).

4. Dynabeads MyOne Streptavidin C1 magnetic beads (Invitrogen; cat. no. 650.01).
5. 1.5-mL LoBind tubes (Eppendorf; cat. no. 022431021).
6. Rotating wheel.
7. Wash solution (from the Ion Template Solutions Kit).

2.8.10 Assessment of the Quality of the Enriched Ion Sphere Particles

1. Qubit 2.0 Fluorometer (Invitrogen; cat. no. Q32866).
2. Qubit assay tubes (Invitrogen; cat. no. Q32856).
3. 0.2-mL RNase-free PCR tubes (Ambion; cat. no. AM12225).
4. 1.5-mL DNA LoBind tubes (Eppendorf; cat. no. 022431021).
5. Qubit 2.0 Easy Calculator Microsoft Excel Spreadsheet (download online).
6. Buffers:
 - (a) 1× TEX (10 mM Tris, 1 mM EDTA, 0.01 % Triton X-100, pH 8)
 - (b) 1× SSPE (diluted from a 20× stock, Ambion; cat. no. AM9767)
7. Oligonucleotides, HPLC purified, 250-nM synthesis scale (IDT):
 - (a) B'-FAM 5'-CTGAGACTGCCAAGGCACACAGGGGA TAGG-3'
 - (b) A-CY5 5'-CCATCTCATCCCTGCGTGTCTCCGACT CAG-3'

2.8.11 Generation of the DNA Libraries for Ion Torrent PGM Sequencing

This comprises materials for end repair of the sheared DNA, ligation of adaptors, and amplification of the 200-bp DNA libraries:

1. The Ion Fragment Library Kit (cat. no. 4466464). This comprises the:
 - (a) 5× end repair buffer
 - (b) End repair enzyme
 - (c) 10× ligase buffer
 - (d) DNA ligase
 - (e) Adapters
 - (f) Platinum PCR SuperMix high fidelity
 - (g) Library Amplification Primer Mix
 - (h) Low TE buffer

2.8.12 Sequencing of Libraries with the Ion Torrent PGM

This includes the materials required for initialization of the Ion Torrent PGM sequencer and loading Ion Sphere Particles onto the Ion Chip:

1. The Ion Sequencing Kit (Invitrogen/Life Technologies; cat. no. 4466456). This is composed of:
Ion PGM Supplies Kit (cat. no. 4466458), composed of:
 - (a) Wash bottle sippers
 - (b) Conical tube sippers
 - (c) Conical tubes and labels
 - (d) Wash 1 bottle (250 mL) and label
 - (e) Wash 2 bottle (2 L) and label
 - (f) Wash 3 bottle (250 mL) and label
 - (g) Ion Sequencing Reagents Kit (cat. no. 4466459), composed of:
 - (h) dGTP
 - (i) dCTP
 - (j) dATP
 - (k) dTTP
 - (l) Sequencing polymerase
 - (m) Sequencing primer
 - (n) Control Ion Sphere Particles
 - (o) Ion PGM reagents kit (cat. no. 4466460). This is composed of:
 - W2 solution
 - Annealing buffer
 - PGM cleaning tablet
 - W3 solution
2. Industrial Grade Liquid Argon, 160-L cylinder (230 PSI) (Air-gas; cat. no. AR160LT230).
3. Multistage (dual-stage) gas regulator (0–50 PSI, 2–3 bar output) (VWR International; cat. no. 55850-422).
4. Magnetic stirrer base MS-3000 (Fisher Biotec; cat. no. BS-010301-ABF).
5. 4-cm magnetic stirrer bar (ProSciTech; cat. no. L35-4008).
6. Glycerol, molecular biology grade (Sigma-Aldrich; cat. no. G5516).
7. Sonication water bath and floating PCR tube holder (Branson; cat no: 952-118).

2.9 High-Resolution Melting Analysis

1. PCR thermocycler. MultiGene™ OptiMax Thermal Cycler (Labnet) may be used.
2. Light scanner with high-resolution melting analysis software. LightScanner System and Call-IT 2.0 software (Idaho Technologies, Salt Lake City, Utah) may be used.
3. 96-well white and black plate. Hard-Shell® Low-Profile Thin-Wall 96-Well Skirted PCR plates (Bio-Rad, Hercules; CA #HSP-9665) may be used.
4. Mineral oil (Sigma-Aldrich).
5. High-resolution melting enzyme mix. HRM Mix (TrendBio) may be used.
6. PCR primers (10 µM).
7. Optical sealing tape. Optical sealing tape (Bio-Rad, Hercules, CA #223-9444) may be used.

3 Methods

3.1 Manual Passage of Human-Induced Pluripotent Stem (iPS) Cells (See Note 1)

1. Under the dissecting microscope, assess colonies for differentiation and determine the most undifferentiated colonies for transfer (*see Note 2*).
2. Change medium on all preprepared MEF feeder plates to SR medium (*see Notes 3 and 4*).
3. From colony plates, choose areas of undifferentiated cells only for transfer by selectively cutting the colony into small pieces to generate small clumps of undifferentiated cells. This can be done with a pulled glass capillary or a 30-G needle.
4. Detach the pieces from the plate by gently teasing them up with the end of a capillary, needle, or pipette.
5. Using a P20 pipette, transfer 6–9 colony pieces onto each prepared feeder plate, and place them carefully on plate to avoid damage to the feeder layer.
6. Carefully transfer plate to 37 °C incubator so colony pieces do not move around and leave overnight to allow cells to attach.
7. Medium change plates daily until ready for next passage, usually 6–7 days.

3.2 Transition from Manual to Enzymatic Passage

1. To set up an enzymatic culture from manual passage dishes, cut colonies as per manual passage method, choosing only undifferentiated areas of colonies (*see Notes 5 and 6*).
2. Transfer colony pieces to a sterile Eppendorf containing 400 µL of serum replacement (SR) medium.
3. Use a 200 µL pipette to gently pipette up and down until colony pieces have broken up to a relatively homogeneous

suspension. Transfer suspension to a 25-cm² flask (previously seeded with MEFs) containing SR medium + 10 ng/mL bFGF (*see Note 7*).

4. Medium change flask daily and passage every 4–7 days as necessary.

3.3 Enzymatic Passage of iPS Cells

1. Remove medium from culture flask.
2. Add collagenase I solution to flask (e.g., 2 mL for 25-cm² flask).
3. Incubate flask at 37 °C for 10–15 min, undisturbed, if using freshly made collagenase or 20 min if using thawed collagenase.
4. During incubation, prepare new flasks by adding SR + 10 ng/mL bFGF medium to flasks previously seeded with MEFs.
5. Tap flask to dislodge cells and add sufficient SR + bFGF medium to dilute collagenase solution (*see Note 8*). Wash medium over surface of flask until majority of cells are removed, some differentiated cells and feeder cells may remain attached to the surface of the flask.
6. Transfer cell suspension to a 15-mL conical tube and centrifuge for 3 min at 150 × *g*.
7. Aspirate medium and gently resuspend pellet in 1 mL of SR + bFGF medium using a 1,000-μL tip; add further SR + bFGF medium and gently resuspend until a homogeneous suspension is achieved.
8. For continuous cultures, inoculate new flasks with desired dilution of cell suspension (*see Note 9*).
9. Observe cells and medium change daily. Passage when confluent, usually every 4–7 days.

3.4 Spontaneous Differentiation of iPS

1. Wash iPS plates with PBS and then incubate with dispase solution for 10 min at 37 °C to allow intact colonies to become detached.
2. Add SR medium, without bFGF, and transfer colonies to a 15-mL tube; centrifuge for 3 min at 150 × *g*.
3. Remove supernatant and then gently resuspend in SR medium, without bFGF, until a homogeneous suspension is achieved.
4. Transfer suspension to ultralow attachment plates to induce the formation of embryoid bodies (EBs).
5. Culture EBs in suspension for 7 days then transfer to 0.1 % gelatin-coated tissue culture plates to let attachment and spontaneous differentiation occur.
6. Change SR medium every 3–4 days until differentiation is complete.

3.5 To Measure Oxygen Consumption in iPS Cells

Measuring change in O₂ concentration (oxygen flux) during incubation with intact living cells can inform on the oxidative phosphorylation (OXPHOS) state in the mitochondria. Generally, mitochondria from embryonic as well as pluripotent stem cells show low oxygen consumption rate (OCR), and the utilization of OXPHOS inhibitors as well as uncouplers of the proton gradient is required to assess the respiration capacity of the cells:

1. Power on the Oxygraph-2K (O2K) and initiate the OROBOROS DataLab software (*see Note 10*).
2. Wash chambers with water, EtOH 80 %, and EtOH 100 % and then water before starting calibration of the oxygen sensors.
3. Fill the O2K chambers with 5 % sodium hydrosulfite and wait for the oxygen concentration to be minimal (~0 %). In the calibration parameters, assign 0 % oxygen to each oxygen sensor.
4. Wash chambers with water, EtOH 80 %, and EtOH 100 % and then water before filling up with HBSS. Wait for oxygen concentration to stabilize, and then assign 100 % oxygen to each sensor.
5. Close completely each O2K chamber (no air bubble) and inject up to 50 µL of cell suspension per chamber. Before injection, cells are washed, counted, and resuspended in HBSS. For effective oxygen consumption measurement, inject at least half a million cells per chamber (*see Note 11*). As an experimental parameter, normalize oxygen flux to the cell number corresponding to cell count (*see Note 12*).
6. Five minutes after injection, record the OCR in each chamber. This measure indicates the basal rate (“routine”) of OXPHOS.
7. Omy is an inhibitor of ATP synthase complex V. Inject 2 µL Omy to each chamber and, after stabilization, record the OCR (*see Note 13*). This indicates the non-phosphorylating rate of respiration. By inhibiting complex V, the proton gradient becomes maximal and oxygen flux reflects proton leak (mainly) as well as other oxidative reactions in the cells.
8. FCCP is an uncoupler of the proton gradient. By adding successive volumes of 2 µL FCCP, titrate to obtain a maximal OCR (maximal stimulated respiratory capacity). Uncoupler titration must be performed carefully as over-titration will inhibit OCR (*see Note 14*). Maximal OCR under uncoupling conditions reflects the maximal activity of the electron transport chain (ETC).
9. Rot and Ama inhibit OXPHOS through complexes I and III. Add 1 µL Ama and 1 µL Rot and record OCR. This measure corresponds to the OXPHOS state under maximal inhibition and reflects residual oxygen consumption (ROX) in the cells. The ROX level is used as a background value of non-OXPHOS activity that must be subtracted to previously recorded OCRs.

10. After definitive OXPHOS inhibition, the experiment is finished and each chamber requires thorough washing (*see Note 15*). Rinse the chamber with water, EtOH 80 %, and EtOH 100 %. Adding spare cells to each chamber is effective to eliminate excess of inhibitors.

Interpretations of cellular respiration are based on corrected OCR measurements for basal (6), non-phosphorylating (7), and uncoupled (8) OXPHOS states. Uncoupling control ratio (uncoupled/basal) is an index of the ETC reserve capacity. It decreases under high ATP demand or because of OXPHOS dysfunction.

3.6 To Measure ATP Content in iPS Cells

ATP is a marker of cell viability, and ATP monitoring can be used to assess the level of metabolic quiescence in iPS cells and the proliferative impact of a wide range of drugs or biological compounds. ATP content can be monitored on the basis of the production of light caused by the reaction of ATP with added luciferase and D-luciferin (firefly enzymes). ATPlite luminescence assay is sensitive down to 5 cells in a 100- μ L medium (*see Note 16*), but a higher cell number is recommended to obtain consistent luminescence. Moreover, luminescence can be acquired without the need of injectors as the signal has a half-life of greater than 5 h:

1. Harvest, count, and resuspend cell in 100 μ L of culture medium, and then pipette the 100 μ L cell suspension into the wells of a 96-well plate.
2. In parallel, add 100 μ L of culture medium without cells in separate wells of the same plate that is used for the experimental samples. These wells will contain a serial dilution of ATP standards (from 10^{-4} M down to blank) (*see Note 17*).
3. Add 50 μ L of lysis solution to every well and shake the plate for 5 min in an orbital shaker at 700 rpm. This lyses the cells and stabilizes the ATP.
4. In wells containing culture medium only, add the serial dilution of 10 μ L ATP standards and shake the plate for 5 min in an orbital shaker at 700 rpm.
5. Add 50 μ L substrate solution to each well and shake the plate for 5 min in an orbital shaker at 700 rpm.
6. Cover the plate in foil for 10 min to prevent access to light and measure the luminescence using a plate reader set for luminescence acquisition (*see Note 18*).
7. From the ATP standard luminescence data, calculate the standard curve and extrapolate the ATP concentration to each experimental sample. Normalize the extrapolated ATP concentration by the number of cells in the initial 100 μ L of medium.

3.7 To Measure Lactate Release by iPS Cells

Lactate production is closely related to glycolytic metabolism in the cell, through lactate dehydrogenase (LDH) enzymes that catalyze the interconversion of pyruvate and lactate. In anaerobic glycolysis, production of lactate by the cell can be monitored in the culture environment, as the cell will release lactate in the medium (*see Note 19*). L(+)-Lactate is the major lactate stereoisomer formed in human intermediary metabolism, while D(−)-lactate is also present but only at about 1–5 % of the concentration of L(+)-lactate (blood). The Lactate Assay Kit II (BioVision) is based on oxidation of L(+)-lactate by LDH, generating a product that interacts with a colored probe ($\lambda_{\text{max}} = 450 \text{ nm}$). The kit detects L(+)-lactate in biological samples such as in serum, cells, culture, and fermentation media, at concentrations from 0.02 to 10 mM lactate (*see Note 20*). Here, the measure of lactate concentration in culture medium is described:

1. Culture cells in medium devoid of lactate and/or pyruvate and then collect spent culture medium (*see Note 21*). The medium containing FBS should be deproteinized to remove LDH that can degrade lactate. Record cell number and total medium volume at time of medium collection. These will be used to normalize the amount of lactate produced by the cell.
2. Prepare a standard curve by adding 0, 2, 4, 6, 8, and 10 μL of L(+)-lactate standard (1 nmol/ μL) into separate wells, and then adjust volume to 50 μL /well with lactate assay buffer to generate 0, 2, 4, 6, 8, and 10 nmol/well of the L(+)-lactate standard.
3. Prepare test sample in 50 μL /well with lactate assay buffer. It is suggested to use two different volumes of test sample, for example, 10 and 40 μL , to ensure the reading is within the standard curve range (*see Note 22*).
4. For each well, prepare a total of 50 μL master mix containing 46 μL lactate assay buffer, 2- μL probe, and 2 μL enzyme mix. Add 50 μL of master mix to each well containing the lactate standard or test samples. Mix well by pipetting up and down.
5. Incubate the reaction for 30 min at room temperature, protect from light.
6. Measure OD 570 nm for colorimetric assay in a microplate reader.
7. Correct background by subtracting the value derived from the 0 lactate control from all sample readings. Plot standard curve nmol/well ~OD570nm readings. Apply the sample readings to the standard curve. Calculate the lactate concentrations of the test samples: $C = \text{La}/\text{Sv}$ (nmol/ μL or mM) where La is the lactate acid amount (nmol) of your sample from standard curve and Sv is the sample volume (μL) added into the well. Normalize the lactate concentration to the total volume of medium and the cell number.

3.8 Preparation for Standard Curve Using Quantitative PCR

1. Adjust the concentration of total DNA templates to 100 ng/ μ L.
2. For amplification using conventional PCR, prepare a reaction mixture with a total volume of 50 μ L containing:
 - (a) 5 μ L 10 \times NH₄ reaction buffer
 - (b) 1.5 μ L 50 mM MgCl₂ solution
 - (c) 0.5 μ L 50 mM dNTP mix dilution
 - (d) 0.5 μ L BIOTAQ™ DNA polymerase
 - (e) 1 μ L forward primer* (50 pmol/ μ L)
 - (f) 1 μ L reverse primer* (50 pmol/ μ L)
 - (g) 2 μ L (200 ng) DNA template
 - (h) 38.5 μ L autoclaved Milli-Q H₂O
3. After amplification of nDNA (β -globin) and mtDNA, run PCR samples on a 2 % agarose gel to confirm the amplified DNA templates.
4. Cut out the targeted band from the gel and purify DNA using Qiagen gel purification kit. Determine the amount of DNA templates purified for nDNA and mtDNA, respectively.
5. Adjust the concentration of purified DNA to 10 ng/ μ L.
6. Prepare tenfold serial dilutions of the purified DNA from 10⁻¹ to 10⁻⁸ ng/ μ L for the standard templates.
7. Prepare for quantitative PCR in 20 μ L reactions containing:
 - (a) 10 μ L 2 \times SensiMix SYBR Hi-ROX master mix
 - (b) 1 μ L forward primer* (5 pmol/ μ L)
 - (c) 1 μ L reverse primer* (5 pmol/ μ L)
 - (d) 2 μ L of each standard template (*see Note 23*)
8. Reactions are performed in the 72-well Rotor Gene 3000 real-time PCR machine. Reaction conditions are:
 - (a) 95 °C for 15 min, followed by:
 - 45 cycles of 95 °C for 15 s
 - Primer-specific annealing temperature for 15 s
 - 72 °C for 15 s
 - (b) Extension and data acquisition are performed on FAM/SYBR channel.
 - (c) Melt curve analysis from 72 to 99 °C with a 30-s wait for the first step followed by 5 s for each subsequent step of 1 °C increase. Acquire melt data from the FAM/SYBR channel.
9. A standard curve is produced with the efficiency of quantitative PCR determined (*see Note 24*). The melt curve determines the specificity of the primers, where only one curve should be observed (*see Note 25*).

3.9 Determining mtDNA Copy Number

1. Adjust the concentration of the unknown targeted DNA samples to 10 ng/μL.
2. Prepare reaction mixtures for quantitative PCR, as previously described, with a total volume of 20 μL containing 18 μL of master mix with primers and 2 μL of the unknown samples (20 ng of total DNA). qPCR conditions, as previously described (step 8 of Section 3.8).
3. Concentrations of the mitochondrial and nuclear-specific products will be calculated based on the standard curve computationally.
4. Copy number is then calculated by

$$N = ([\text{ng}/\mu\text{L}] \times 6.023 \times 10^{14}) / (N_{\text{bp}} \times 660)$$
 where:
 - (a) N is the number of molecules per reaction.
 - (b) $\times 10^{14}$ is the conversion of 1 mol to 1 nmol using Avogadro's constant, which states 1 mol contains 6.023×10^{23} .
 - (c) N_{bp} is the product size of the amplicons of interest.
 - (d) 660 is the mean molecular weight of a nucleic acid base pairing in Daltons (Da).
5. mtDNA copy number per cell is then calculated by dividing the mtDNA copy number by half of the chromosomal DNA copy number as chromosomal DNA is present in diploidy in cells.

3.10 MeDIP

1. Dilute 10 μg genomic DNA in 100 μL autoclaved Milli-Q H₂O and add 100 μL 2× TE buffer in a Covaris milliTUBE. Use Covaris Adaptive Focused Acoustics (AFA™) S220 system to shear the DNA for 195 s to fragments sized between 200 and 1,000 bp. The setting is 5–10 °C temperature limit range, 105.0 peak power, 5.0 duty factor, and 200 cycles/burst (*see Note 26*).
2. Run a 1–1.5 % agarose gel to confirm efficient sonication (*see Note 27*).
3. Take at least 3 μg of sonicated DNA and dilute with autoclaved Milli-Q H₂O to a final volume of 450 μL (*see Note 28*).
4. Denature DNA by boiling at 95 °C for 10 min and immediately place on ice for 5 min.
5. Add 50 μL of 10× IP buffer and 3 μL (3 μg; 1 μg per 1 μg DNA) of anti-5mC or anti-5hmC antibody.
6. Incubate at 4 °C for 2 h with rotation.
7. Prewash 30 μL of Dynabeads® Protein G (10 μL per 1 μg DNA) two times with 1 mL PBS/BSA buffer for 5 min per wash at 4 °C with rotation.
8. Remove fluid from beads on a magnetic rack and resuspend beads in 30 μL (original volume) of 1× IP buffer.

9. Add beads into each sample and incubate at 4 °C for 16 h with rotation.
10. Remove all fluid on a magnetic rack.
11. Wash the beads with 1 mL of 1× IP buffer for 5 min per wash at 4 °C with rotation. Remove all fluid on a magnetic rack.
12. Resuspend beads in 250 µL proteinase K digestion buffer and add 7 µL of 20 mg/mL proteinase K.
13. Incubate at 50 °C for 3 h on a shaker.
14. Extract DNA with QIAquick PCR purification kit. Elute purified DNA in 50 µL of autoclaved Milli-Q H₂O.

3.11 Analysis of Gene Expression of mtDNA Replication Factors and mtDNA-Encoded Genes

1. cDNA synthesis is performed using SuperScript III kit from Life Technologies.
2. A total of 1 µg of RNA is converted to cDNA (*see Note 29*). Take 1 µg of RNA and adjust the volume to 8 µL with RNA-free water and add 1 µL of 50 µM oligo(dT) and 1 µL of 10 mM dNTPs.
3. Incubate the reaction at 65 °C for 5 min and chill on ice for at least 1 min.
4. Add 2 µL of 10× RT buffer, 4 µL of 25 mM magnesium chloride, 2 µL of 0.1 M DTT, 1 µL of RNaseOUT (40 U), and 1 µL of SuperScript III enzyme (200 U) to the reaction (*see Note 30*).
5. Incubate the reaction at 50 °C for 50 min followed by 85 °C for 5 min. Chill the reaction on ice.
6. Add 1 µL of RNase H to the reaction and incubate at 37 °C for 20 min. Chill the reaction on ice.
7. Dilute the cDNA with 42 µL of sterile Milli-Q water (1:3 dilutions). cDNA is ready for gene expression or it can be stored at −20 °C.
8. Prepare master mix containing 10 µL of SensiMix, 1 µL of 5 µM forward primer, 1 µL of 5 µM reverse primer, and 6 µL of sterile Milli-Q water.
9. Add 2 µL of cDNA to sterile strip tubes and mix with 18 µL of master mix. Run each reaction in triplicate.
10. A normalization gene (housekeeping gene) should also be run for quantification of gene expression using $\Delta\Delta CT$ method.
11. Run real-time PCR on Rotor Gene 3000 real-time PCR machine (Corbett Research, Cambridge, UK).
12. The reaction conditions for real-time PCR are 95 °C for 10 min followed by 45 cycles of 95 °C for 15 s, followed by gene-specific annealing temperature with an extension for 15 s at 72 °C.

13. The data is acquired in FAM/SYBR channel during extension phase. The second acquisition phase (cycle B) of 78 °C was used to permit quantification of a single specific product only.
14. Gene expression is normalized to a housekeeping gene and data are expressed as relative fold change calculated using $\Delta\Delta CT$ method.

3.12 Next-Generation Sequencing

Long PCR generates the templates for next-generation sequencing.

3.12.1 Long PCR

1. Two long PCR reactions are performed, each in a 0.2-mL PCR tube containing of 1× high-fidelity PCR buffer, 2 mM MgSO₄, 10 mM dNTPs, 10 μM of the forward primer, 10 μM of the reverse primer, and 1 U of Platinum *Taq* DNA polymerase high fidelity (*see* **Note 31**).
2. 50 ng of genomic DNA is used as the template (*see* **Note 32**).
3. A negative control contains the components listed in step 1, except for the genomic DNA.
4. Reactions are run as either 20 or 50 μL volumes and are adjusted accordingly with sterile ddH₂O.
5. Reaction conditions are:
 - (a) 94 °C for 2 min
 - (b) Followed by 35 cycles of:
 - 94 °C for 15 s
 - 63 °C for 30 s
 - 68 °C for 9 min
 - (c) The reaction is completed by 1 cycle of:
 - (d) 68 °C for 10 min
 - (e) 4 °C hold until the reactions

3.12.2 Purification of the PCR Products

1. Add pH indicator I to buffer PB at a ratio of 1:250 to produce “buffer PBI” at a pH of ≤ 7.5 .
2. Transfer the PCR product to a 1.7-mL sterile microcentrifuge tube.
3. Add buffer PBI to the PCR product at a volume ratio of 5:1.
4. If buffer PBI changes color, add 10 μL of 3 M sodium acetate (pH 5.0) to correct the pH. If the color is unaffected, proceed to step 5.
5. Mix by pipetting and then transfer up to 650 μL of the solution to a QIAquick spin column inserted into a 2-mL collection tube.
6. Centrifuge for 1 min at 10,000 × *g*.
7. Discard the flow-through.

8. If there is excess PCR product from step 5, load the remaining solution into the QIAquick spin column and repeat steps 6 and 7. If not, then proceed to step 9.
9. Add 100 % ethanol to buffer PE before use.
10. Add 750 μL of buffer PE to the QIAquick spin column.
11. Repeat steps 6 and 7 twice.
12. Transfer the QIAquick spin column to a 1.7-mL microcentrifuge tube and add 30 μL of sterile ddH₂O.
13. Repeat step 6 and store the purified template for next-generation sequencing at -20°C until use.

3.12.3 Quantification of Double-Stranded DNA

1. Dilute the Qubit dsDNA HS reagent in Qubit dsDNA HS buffer at a ratio of 1:200 to produce the “Qubit working solution” for the number of samples and standards to be tested (*see Note 33*).
2. Standards #1 and #2 are prepared by adding 10 μL of each Qubit standard to 190 μL of Qubit working solution to individual Qubit assay tubes.
3. DNA samples are prepared by adding 1 μL of the DNA to 199 μL of Qubit working solution in individual Qubit assay tubes.
4. Standards and samples are vortexed for 2–3 s and incubated at room temperature for 2 min.
5. Set up the Qubit 2.0 Fluorometer by choosing the assay type “dsDNA High Sensitivity.”
6. Choose: run a “new calibration.”
7. Choose: insert standard #1 and press “Read.”
8. Remove standard #1, then insert standard #2, and choose “Read.” Remove standard #2.
9. Insert the DNA sample and choose “read next sample.”
10. Repeat step 8 for the quantification of each DNA sample.
11. Calculate the concentration of the sample by using:

$$\text{Concentration}(\text{ng/mL}) = \text{QF value} \times (200/x)$$

QF value = the value obtained from the Qubit 2.0 Fluorometer; x = the volume (μL) of sample used for quantification.

3.12.4 Shearing of Long PCR Amplicons (See **Note 34**)

1. Pool the two long PCR mtDNA amplicons at equimolar concentrations.
2. Transfer the pooled DNA to a Covaris microTUBE and keep on ice.
3. Fill the Covaris S220 system water tank to level 12 with 4°C deionized water.

4. Degas the water by choosing the “De-gas” option on the Covaris Sonolab 7 software and run for 30 min.
5. Set the operating conditions to sheer DNA to 200-bp fragments:
 - (a) Duty factor = 10 %
 - (b) Peak incident power (W) = 175
 - (c) Cycles per burst = 200
 - (d) Time = 180 s
 - (e) Number of cycles = 6
6. Load the Covaris sample tube into the tube holder of the Covaris S220 and apply the holder arm.
7. Lower the arm of the Covaris S220 to partially submerge the Covaris tube into the degassed 4 °C deionized water tank.
8. Select “Start” to initiate DNA fragmentation.
9. When the shearing is complete, remove the Covaris sample tube and place on ice.
10. Repeat steps 6–9 for any other samples.

3.12.5 End Repair of the Sheared DNA

1. Transfer 79 µL of the sheared DNA to a 1.5-mL Eppendorf LoBind tube.
2. Add 20 µL of 1× end repair buffer.
3. Add 1 µL end repair enzyme.
4. Mix by pipetting and incubate for 20 min at room temperature.

3.12.6 Sample Purification

1. Add 180 µL of Agencourt AMPure XP reagent beads to the end-repaired DNA sample.
2. Mix by pipetting five times and centrifuge briefly.
3. Incubate at room temperature for 5 min.
4. Load the sample onto a DynaMag-2 magnetic rack.
5. Leave for 3 min.
6. Discard the supernatant by pipetting.
7. Leaving the tube in the magnetic rack, resuspend the Agencourt AMPure XP reagent beads in 500 µL of 70 % ethanol.
8. Incubate for 30 s.
9. Rotate the tube twice within the DynaMag-2 magnetic rack.
10. Discard the supernatant by pipetting.
11. Repeat steps 7–10.
12. Remove the tube from the magnetic rack and centrifuge briefly and reload the tube into the magnetic rack.
13. Discard residual supernatant by pipetting.

14. Air-dry at room temperature for 5 min.
15. Remove the tube from the magnetic rack and add 25 μL of low TE buffer.
16. Pipette up and down five times to mix.
17. Vortex the tube for 10 s.
18. Centrifuge briefly and place the Eppendorf LoBind tube into the magnetic rack.
19. Incubate at room temperature for 1 min.
20. Transfer the product to a sterile 0.2-mL PCR tube.

3.12.7 Adaptor Ligation

1. Add the following reagents to the 0.2-mL PCR tube containing 25 μL of purified DNA: 1 \times ligase buffer, 2 μL of Ion P1 adapter, 2 μL of Ion Xpress Barcode 1, 2 μL of dNTP mix, 2 μL of DNA ligase, and 8 μL of nick repair polymerase.
2. Repeat step 1 but replace 2 μL of Ion Xpress Barcode 1 for 2 μL of Ion Xpress Barcode 2 (up to Barcode 16).
3. Add sterile ddH₂O to a total reaction volume of 100 μL .
4. Mix by pipetting.
5. Place the 0.2-mL PCR tubes into a thermal cycler and run the reactions at:
 - (a) 25 °C for 15 min
 - (b) 72 °C for 5 min
 - (c) 4 °C hold
6. Transfer the adaptor-ligated products to a sterile 1.5-mL Eppendorf LoBind tube.
7. Repeat Section 3.12.6 by adding 140 μL of Agencourt AMPure XP reagent beads (step 1) and eluting the purified DNA in 20 μL of low TE buffer (step 16).

3.12.8 Size Selection

1. Load 20 μL of the target DNA into wells of a 2 % E-Gel SizeSelect agarose gel.
2. Dilute the 50-bp reference DNA ladder 1 in 40 using low TE buffer.
3. Load 10 μL of the diluted DNA ladder onto the agarose gel.
4. Load 25 μL of sterile ddH₂O into the top wells of the gel.
5. Load 25 μL of sterile ddH₂O into the bottom wells of the gel, except the collection well used for electrophoresis of the DNA ladder. This is loaded with 10 μL of sterile ddH₂O.
6. Set the “Run Time To Reference Line” to 15 min.
7. End the electrophoresis when the 200-bp reference band reaches the reference line.

8. Refill the collection wells with 25 μL of sterile ddH_2O .
9. Set the “Run Time from Reference Line to Collection Well” to 2 min.
10. Collect the DNA when the 200-bp band has migrated into the collection well and transfer to a 0.2-mL PCR tube.

3.12.9 Amplification of the 200-bp DNA Libraries

1. Add 25 μL of the size-selected DNA to 100 μL of Platinum PCR SuperMix high-fidelity master mix and 5 μL of Library Amplification Primer Mix.
2. Transfer the sample tubes to a thermal cycler and amplify at:
 - (a) 95 °C for 5 min
 - (b) Followed by 6 cycles of:
 - 95 °C for 15 s
 - 58 °C for 15 s
 - 70 °C for 1 min
 - (c) Followed by 4 °C hold until removed from the thermal cycler
3. Repeat Section 3.12.6 by adding 1.5 \times the sample volume of Agencourt AMPure XP reagent beads, as in step 1, and eluting the purified DNA in 20 μL of low TE buffer, as in step 16.

3.12.10 Quantification of the DNA Libraries

1. Equilibrate the high sensitivity DNA dye concentrate and the high sensitivity DNA gel matrix vials to room temperature for 30 min before use.
2. Mix the high sensitivity DNA dye concentrate by vortexing for 10 s.
3. Add 15 μL of the dye concentrate to a vial of the high sensitivity DNA gel matrix.
4. Vortex the gel-dye mixture for 10 s.
5. Apply the gel-dye mixture to the filter of a spin column and centrifuge at $4,000 \times g$ at room temperature for 10 min.
6. Discard the spin column.
7. Place the Agilent high sensitivity DNA chip onto the chip-priming station.
8. Load 9 μL of the gel-dye mix onto the Agilent high sensitivity DNA chip in the well labeled with a circularized “G.”
9. Set the syringe plunger at the 1-mL position and close the chip-priming station (*see Note 35*).
10. Push the syringe plunger until held in position by the clip.
11. After 60 s, release the syringe plunger so that the plunger reaches the 0.3-mL mark.
12. After 5 s, set the syringe plunger to the 1-mL position.

13. Open the chip-priming station.
14. Load 9 μL of the gel-dye mix onto three wells labeled as “G” on the Agilent high sensitivity DNA chip.
15. Load 5 μL of the high sensitivity DNA marker in the 11 sample wells and in the well labeled with a ladder.
16. Load 1 μL of DNA sample into the 11 sample wells.
17. Place the Agilent high sensitivity DNA chip onto the IKA vortex mixer.
18. Vortex the chip at 2,400 rpm for 1 min.
19. Open the lid of the Agilent 2100 Bioanalyzer and clean the electrodes.
20. Place the Agilent high sensitivity DNA chip into the Agilent 2100 Bioanalyzer.
21. Open the Agilent 2100 Expert Software and select “dsDNA.”
22. Select “start.”
23. Dispose of the DNA chip at the end of the run.

3.12.11 Preparation of the IKA DT-20 Solution and Ion Sphere Particles

1. Transfer 9 mL of emulsion oil to an IKA DT-20 tube using a 10-mL Eppendorf Combitips Plus pipette tip.
2. Store the IKA DT-20 tube on ice (*see* **Note 36**).
3. Vortex the Ion Sphere Particles for 1 min.
4. Transfer 140 μL of the Ion Sphere Particles to 1.3 mL of wash solution in a 1.5-mL Eppendorf LoBind tube.
5. Pipette five times to mix.
6. Vortex briefly the Ion Sphere Particles.
7. Centrifuge the Ion Sphere Particles at $10,000 \times g$ at room temperature for 5 min.
8. Resuspend the Ion Sphere Particles in 100 μL of supernatant.
9. Adjust the total volume of the Ion Sphere Particles solution to 140 μL with Ion Sphere wash solution.

3.12.12 Preparation of the Aqueous Master Mix for the Emulsion PCR

1. Dilute the DNA library quantified in Section 3.12.10 to 9 pM using sterile ddH₂O.
2. Adjust the total volume of the 9 pM DNA library to 18 μL with sterile ddH₂O.
3. Prepare the aqueous master mix in a 1.5-mL Eppendorf LoBind tube by adding:
 - (a) 336 μL of nuclease-free water
 - (b) 105 μL 10 \times amplification buffer
 - (c) 105 μL of MgCl₂ solution
 - (d) 105 μL of dNTPs

- (e) 105 μL of clonal amplification primer mix
- (f) 5 μL of enrichment primer
- (g) 5 μL of thermostable pyrophosphatase (TAP)
- (h) 126 μL of polymerase
- 4. Mix by pipetting five times.
- 5. Transfer the aqueous master mix to the 140 μL of Ion Sphere Particles prepared in Section 3.12.11.
- 6. Vortex the Ion Sphere Particles/amplification master mix solution at maximum speed for 5 s.
- 7. Add the 18 μL of DNA library diluted to 9 pM.

3.12.13 Generation of the Emulsion

- 1. Place the ice-cold IKA DT-20 tube prepared in Section 3.12.11 on the IKA ULTRA-TURRAX Tube Drive.
- 2. Remove the adhesive seal from the lid of the IKA DT-20 tube to create an opening in the tube lid.
- 3. Vortex the Ion Sphere Particles/amplification master mix/DNA solution then centrifuge briefly for 3 s.
- 4. “Start” the IKA ULTRA-TURRAX Tube Drive.
- 5. Transfer the master mix solution into the IKA DT-20 tube through the opening while the tube is in motion.
- 6. Replace the adhesive seal over the opening on the lid of the IKA DT-20 tube.
- 7. Allow 5 min for mixing and place the IKA DT-20 tube on ice for 5 min.

3.12.14 Emulsion PCR

- 1. Transfer 100 μL of the emulsion using an Eppendorf Repeater pipettor into each well of an Applied Biosystems MicroAmp 96-Well Reaction Plate.
- 2. Cover the 96-well PCR plate with Applied Biosystems MicroAmp Optical Cap Strips.
- 3. Place the PCR plate into a thermal cycler and run at:
 - (a) 94 °C for 6 min.
 Followed by 40 cycles of:
 - (b) 94 °C for 30 s
 - (c) 58 °C for 30 s
 - (d) 72 °C for 90 s
 5 cycles of:
 - (e) 94 °C for 30 s
 - (f) 68 °C for 6 min
 End the reaction with 1 cycle of:
 - (g) 10 °C hold, until collected from the machine.

3.12.15 *Recovery of the
Ion Sphere Particles*

1. Mix the emulsion.
2. Transfer ~1 mL of the emulsion from the PVC basin into six 1.5-mL Eppendorf LoBind tubes.
3. Centrifuge at $15,500 \times g$ for 2 min.
4. Prepare the “breaking solution” by adding 2.5 mL of recovery solution to 7.5 mL of 1-butanol in a 15-mL tube.
5. Vortex for 1 min.
6. After completing step 3, discard the oil layer without disturbing the pellet.
7. Rinse the Applied Biosystems MicroAmp 96-Well Reaction Plate by transferring 100 μ L breaking solution to each well in the first row of the 96-well plate.
8. Mix by pipetting and transfer the breaking solution from step 7 to the next row of the 96-well plate.
9. Repeat rinsing for other rows on the 96-well plate until the bottom row is reached.
10. Transfer the breaking solution to a PVC basin.
11. Repeat steps 7–10.
12. Transfer ~1 mL of breaking solution, collected in the PVC basin, to the six Eppendorf tubes from step 6.
13. Vortex Eppendorf tubes for 1 min.
14. Centrifuge Eppendorf tubes at $15,500 \times g$ for 2 min.
15. Discard the top layer without disturbing the pellet.
16. Add 1 mL of recovery solution to the Eppendorf tubes containing the pellets.
17. Vortex the Eppendorf tubes for 1 min.
18. Centrifuge the Eppendorf tubes at $15,500 \times g$ for 3 min.
19. Discard the supernatant from each Eppendorf tube but leave ~50 μ L of solution behind in each of the six Eppendorf tubes.
20. Transfer the material from the six Eppendorf tubes into one 1.5-mL sterile Eppendorf LoBind tube.
21. Rinse three of the tubes by adding 200 μ L of recovery solution to the first tube, mix by pipetting five times, and transfer the recovery solution to the second tube.
22. Pipette the recovery solution five times and transfer to the third tube.
23. Transfer the solution from the third tube to the tube generated in step 20.
24. Repeat steps 21–23 if necessary.
25. Add the recovery solution at a ratio of 1:1 with the combined tube.

26. Vortex for 30 s.
27. Centrifuge at $15,500 \times g$ for 3 min.
28. Discard the supernatant but leave $\sim 100 \mu\text{L}$ of solution in the tube.
29. Pipette up and down five times.
30. Transfer the contents to a sterile 1.5-mL Eppendorf LoBind tube.
31. Rinse the original tube with $100 \mu\text{L}$ of wash solution and transfer to another Eppendorf tube.
32. Add 1 mL of wash solution to the Eppendorf tube.
33. Repeat steps 26–28.
34. Repeat steps 32 and 33.
35. Place the sample on ice.

*3.12.16 Enrichment of
the Template-Positive Ion
Sphere Particles*

1. Prepare melt-off solution (125 mM NaOH and 0.1 % TWEEN 20) by adding $200 \mu\text{L}$ of 1 M NaOH to $16 \mu\text{L}$ of 10 % Tween-20 (diluted in sterile ddH₂O) and 1.38 mL of sterile ddH₂O.
2. Vortex Dynabeads MyOne Streptavidin C1 beads.
3. Add $10 \mu\text{L}$ of MyOne Streptavidin beads to a 1.5-mL Eppendorf LoBind tube.
4. Add $70 \mu\text{L}$ of wash solution to the MyOne Streptavidin beads.
5. Vortex to mix.
6. Place the Eppendorf tube into the DynaMag-2 magnetic rack for 2 min.
7. Remove the supernatant.
8. Resuspend the MyOne Streptavidin beads by adding $10 \mu\text{L}$ of wash solution.
9. Transfer the $10 \mu\text{L}$ of MyOne Streptavidin beads to the Eppendorf tube containing the Ion Sphere Particles (step 35 of Section 3.12.15).
10. Pipette five times to mix.
11. Vortex to mix again.
12. Add $100 \mu\text{L}$ of annealing buffer to the Ion Sphere Particles.
13. Repeat steps 10 and 11.
14. Place the tube containing the Ion Sphere Particles on a rotating wheel for 10 min at room temperature.
15. Centrifuge the sample for 3 s.
16. Repeat step 6.
17. Transfer the supernatant to a new 1.5-mL Eppendorf LoBind tube marked “unbound.”

18. Add 200 μL of wash solution to the Ion Sphere Particles.
19. Pipette 15 times to mix.
20. Repeat steps 6 and 17.
21. Repeat steps 18–20.
22. Add 400 μL of melt-off solution prepared in step 1 to the sample tube.
23. Vortex.
24. Place the tube on a rotating wheel to spin at room temperature for 7 min.
25. Add 400 μL of wash solution to another 1.5-mL Eppendorf LoBind tube labeled “Enriched-1.”
26. When step 24 is completed, vortex the tube and centrifuge briefly.
27. Repeat step 6.
28. Transfer the supernatant to the “Enriched-1” tube without disturbing the MyOne Streptavidin beads.
29. Vortex the “Enriched-1” tube.
30. Centrifuge at $15,500 \times g$ for 4 min.
31. Remove the supernatant but retain 100 μL containing the enriched Ion Sphere Particles.
32. Add 1 mL of wash solution to the “Enriched-1” tube.
33. Repeat steps 29–31.
34. Pipette to mix.
35. Repeat step 29 and centrifuge for 3 s.
36. Place the “Enriched-1” tube into the DynaMag-2 magnetic rack for 2 min.
37. Transfer the supernatant from “Enriched-1” to a new 1.5-mL Eppendorf LoBind tube labeled “Enriched-2.”

*3.12.17 Assessment of
the Quality of the Enriched
Ion Sphere Particles*

1. Dilute the oligonucleotides to 100 μM in $1 \times$ TE buffer, pH 8.0.
2. Combine the following components to create the hybridization master mix:
 - (a) 100 μL of $1 \times$ SSPE
 - (b) 2 μL of 100 μM B'FAM oligonucleotide
 - (c) 2 μL of 100 μL A-CY5 oligonucleotide
3. Add 2 μL of the enriched Ion Sphere Particles to a 0.2 mL PCR tube.
4. For the negative control, add 2 μL of ion wash buffer to a 0.2-mL PCR tube.

5. Add 52 μL of the hybridization master mix from step 2 to both PCR tubes from steps 3 and 4.
6. Pipette five times to mix.
7. Load PCR tubes into a thermal cycler and run at:
 - (a) 95 °C for 2 min
 - (b) 37 °C for 2 min
8. Transfer the products from the PCR tubes to 1.5-mL Eppendorf LoBind tubes.
9. Add 1 mL 1 \times TEX buffer to each tube.
10. Vortex.
11. Centrifuge both tubes at 15,600 $\times g$ for 3 min.
12. Remove 1 mL of the supernatant leaving behind 20 μL containing the Ion Sphere Particles.
13. Repeat steps 9–12 twice.
14. Add 1 \times TEX to the tubes to a total volume of 200 μL .
15. Transfer samples to separate Qubit assay tubes.
16. Load samples into the Qubit 2.0 Fluorometer.
17. Choose “Ion,” followed by “IonFAM,” and then “Read.”
18. Record the value.
19. Choose “Home,” followed by “IonCY5,” and then “Read.”
20. Repeat step 18.
21. Repeat steps 16–20 with the negative control.
22. Open the Qubit 2.0 Easy Calculator Microsoft Excel Spreadsheet.
23. Enter the values.
24. The templated signal ratio is automatically calculated to determine the quality of the enriched Ion Sphere Particles according to the following guidelines:
 - (a) <0.1—insufficient enriched Ion Sphere Particles resulting in suboptimal loading density on the ion chip.
 - (b) 0.1–0.3—the optimal to provide the best loading density on the ion chip.
 - (c) >0.3—will produce mixed reads due to non-clonal amplification of the template on the Ion Sphere Particles during the emulsion PCR.

*3.12.18 Initialization of
the Ion Torrent PGM
Sequencer*

1. Rinse Wash Bottles 1 and 3 twice using ultrapure (18 Ω) water.
2. Add 350 μL of 100 mM NaOH to Wash Bottle 1.
3. Add 50 mL of W3 solution to Wash Bottle 3.
4. Add 1,978 mL ultrapure (18 Ω) water to Wash Bottle 2 and mark the volume level.

5. Empty Wash Bottle 2.
6. Place a magnetic stirrer into the empty Wash Bottle 2.
7. Fill Wash Bottle 2 with argon gas for 5 min using a flowmeter set at 0.5 L per minute.
8. Fill Wash Bottle 2 with ultrapure (18 Ω) water to the 1,978-mL volume mark from step 4. Ensure there are no gaps that would otherwise allow for gas exchange between the bottle and the environment.
9. Place Wash Bottle 2 on a magnetic stirrer, sealing the bottle opening with the argon gas tube while ensuring that the gas nozzle is not in direct contact with the water.
10. Flow of argon gas was set to 0.5 L per minute, and the water was stirred at a speed to create a “whirlpool” effect within Wash Bottle 2.
11. Add 22 mL of W2 solution to Wash Bottle 2.
12. Mix the contents of Wash Bottle 2 for 1 min.
13. Choose “Init PGM” to initialize.
14. Attach a sterile sipper to the bottle cap of Wash Bottle 2 and seal the opening by attaching the bottle to the Ion Torrent PGM to prevent gas exchange.
15. Choose “Next.”
16. Repeat step 13 with Wash Bottles 1 and 3.
17. Repeat step 15.
18. The Ion Torrent PGM automatically determines the pH of Wash Bottle 2.
19. Add 20 μ L of each dNTP solutions to four separate 50-mL conical tubes.
20. After initialization of the wash solutions is completed, attach sippers to each of the 50-mL conical tubes.
21. Attach all dNTP conical tubes to the Ion Torrent PGM, following the order G, C, A, and T.
22. The Ion PGM automatically fills the conical tubes each with 40 mL of solution.
23. At the end of the initialization, choose “Next” to complete the process and initialize the run.

*3.12.19 Preparation of
the Positively Enriched Ion
Sphere Particles for
Sequencing*

1. Transfer 50 % of the positively enriched Ion Sphere Particles to a sterile 0.2-mL PCR tube.
2. Vortex the vial containing the control Ion Sphere Particles and centrifuge briefly.
3. Add 5 μ L of the control Ion Sphere Particles to the PCR tube from step 1.

4. Add 90 μL of wash solution.
5. Pipette to mix.
6. Centrifuge the PCR tube containing the enriched Ion Sphere Particles at $15,500 \times g$ for 1.5 min.
7. Discard the supernatant, leaving between 20 and 30 μL of supernatant behind.
8. Add 150 μL of annealing buffer to the enriched Ion Sphere Particles.
9. Pipette to mix.
10. Repeat steps 6 and 7, leaving behind only 10 μL of supernatant in the PCR tube.
11. Add 5 μL of sequencing primer and pipette five times to resuspend the Ion Sphere Particles.
12. Load the PCR tube into a thermal cycler and run at:
 - (a) 95 °C for 2 min
 - (b) 37 °C for 2 min
13. Repeat steps 8 and 9.
14. Repeat step 6 and 7, leaving behind approximately 20 μL of supernatant containing the Ion Sphere Particles.
15. Repeat steps 13 and 14, leaving behind approximately 5–6 μL of supernatant containing the Ion Sphere Particles.
16. Add 1 μL of sequencing polymerase.
17. Pipette to mix.
18. Incubate at room temperature for 5 min. Meanwhile the ion chip check can be performed.

3.12.20 Testing of a New Ion Chip

1. Choose “Experiment” on the main menu of the Ion Torrent PGM screen.
2. Clamp the ion chip on the Ion Torrent PGM.
3. Scan the chip’s barcode.
4. Choose “Chip Check.”
5. Choose “Next” to continue to the calibration stage.

3.12.21 Loading the Ion Sphere Particles on the Ion Chip (See **Note 37**)

1. Add 100 μL of 100 % glycerol to 100 μL of annealing buffer to a 1.5-mL Eppendorf LoBind tube.
2. Vortex.
3. Remove the ion chip from the Ion Torrent PGM and place on the grounding plate.
4. Add 50 μL of annealing buffer into the large opening on the chip.
5. Remove residual buffer.

6. Sonicate the Ion Sphere Particles that have been incubated for 5 min from step 18 of Section 3.12.19, for 10 s in a sonication water bath at 40 kHz.
7. Add 2 μ L of the glycerol/annealing buffer mix from step 1 to the sonicated Ion Sphere Particles.
8. Pipette to mix.
9. Add the total mix from step 7 to the large opening of the chip.
10. Repeat step 5.
11. Centrifuge the chip for 10 min.
12. Press “#” to check that the number of cycles is set for 65 on the Ion Torrent PGM.
13. Choose “Autoanalysis” and “Pre-analysis.”
14. Choose “Next” to confirm the settings on the Ion Torrent PGM, and then choose “OK.”
15. Replace the loaded ion chip.
16. Allow the Ion Torrent PGM to calibrate the chip and to initiate the sequencing run.

3.12.22 *Analysis of Sequence Outputs (See Note 38)*

CLC Genomics Workbench 7 (www.clcbio.com)

1. Select “Trim” to trim reads of adaptors.
2. Select “Map Reads” and filter reads to exclude those of a nucleotide length of <15 bp.
3. Select “Variant Mapping.”
4. Accept a Phred quality score of ≥ 15 .
5. Set a mismatch cost of 2.
6. Set an insertion/deletion cost of 3.
7. For a single nucleotide polymorphism (SNP) analysis, a minimum mutation threshold of 3 % was applied.

3.13 To Determine Variant Load by High-Resolution Melting Analysis

As the mtDNA is multimeric, it is prone to contain heteroplasmic population of sequence variants. Targeted PCR followed by high-resolution melting of the PCR product can discriminate the presence of a variant nucleotide in the mtDNA. From next-generation sequencing data, HRM analysis can be used to validate the identified variants by designing specific primers around the mtDNA sequences of interest. For quantification of variant frequency, a separate standard curve needs to be run in parallel to the sample. A standard curve is required to isolate reference and variant sequences and to mix them at different ratios (reference/variant), i.e., 100 %:0 %, 90 %:10 % ... 0 %:100 %:

1. Prepare DNA template at 10 ng/ μ L.
2. Prepare HRM master mix with n = sample number:

- (a) $(n + 1) \times 3 \mu\text{L H}_2\text{O}$
- (b) $(n + 1) \times 1 \mu\text{L forward primer}$
- (c) $(n + 1) \times 1 \mu\text{L reverse primer}$
- (d) $(n + 1) \times 4 \mu\text{L HRM enzyme mix}$
- 3. Add 20 μL mineral oil to each well of the plate.
- 4. Add 9 μL HRM master mix to each well of the plate.
- 5. Add 1 μL DNA template/standard in each well of the plate.
- 6. Seal the plate with transparent tape then centrifuge plate.
- 7. Run the plate on PCR thermocycler according to the following cycles:
 - (a) 2 min at 95 °C
 - (b) 45 cycles of 94 °C for 30 s and annealing primer temperature for 30 s
 - (c) 94 °C for 30 s and cooling to 25 °C for heteroduplex formation
- 8. Insert plate into fluorescence scanner and read fluorescence of melting PCR product at temperature ramping from 70 to 96 °C. Data acquisition increases incrementally by 0.1 °C until the reaction is terminated.
- 9. In HRM analysis software, normalize fluorescence shift around the temperature of melting (*see Note 39*).
- 10. Along the standard curve, observe the shift of the melting curve corresponding to the sample and determine the level of variant according to the ratios of reference/variant from the standards (*see Note 40*).
- 11. Collect data outputs from data file.

4 Notes

- 1. These methods are equally applicable to the culture of human embryonic stem cells.
- 2. Manual passage of iPS colonies allows for the transfer of undifferentiated areas of colonies only.
- 3. Undifferentiated iPS cells are cultured on mitomycin C-inactivated MEFs.
- 4. MEFs need to be seeded at 6×10^4 per cm^2 on manual passage plates.
- 5. MEFs need to be seeded at 2×10^4 per cm^2 for enzymatic passage.
- 6. Starting material should contain minimal differentiated cells to ensure maintenance of enzymatic cultures in an undifferentiated state.

7. Approximately 25–30 iPS colonies are required for each 25-cm² culture flask to be set up.
8. Minimum medium volume added should be equal to original volume of collagenase solution.
9. Recommended dilution range is 1:3–1:6.
10. Set temperature in accordance and magnetic stirrer bar to rotate up to 700 rpm.
11. Minimize the time between cell harvest and injection into the chamber. Maintain resuspended cells at 37 °C in the meantime.
12. If OCR does not increase sufficiently after adding cells to the chamber, more cells can be added (avoid air bubble).
13. The interval of OCR recording should be 2–5 min.
14. A higher concentration of FCCP can be used if OCR does not increase substantially.
15. Traces of rotenone can persist. Efficient cleaning must be achieved.
16. Work in a clean environment and wear gloves to avoid any contamination, as ATP is present in many sources of life.
17. ATP standards should be ready and kept on ice before starting the experiment.
18. Luminescence should be measured 30 min after reagent addition, as light production decreases with time, regardless of cell number.
19. Culture medium can contain phenol red, so a low amount of medium should be used for each sample. Normalize OD with background data from blank wells.
20. Cell concentration during culture will impact on lactate production and lactate concentration in the medium. In the case of glucose starvation (prolonged culture or cell confluence), cells will stop producing lactate from glucose but start consuming lactate.
21. Fresh sample is preferred, but frozen sample (–80 °C) can be used.
22. A high concentration of lactate has an inhibitory effect that leads to lower readings. Lactate overdose will turn from pink to brown.
23. The accuracy of pipetting and preparation of the master mix in advance is critical in order to establish the correct quantitation of nuclear and mitochondrial DNA copy number.
24. The efficiency of amplification of the standards should range between 0.95 and 1, and the R^2 value should be as close to 0.99 as possible to establish accurate quantitation of copy number.

25. The temperature of the extension and data acquisition phases should be set to the temperature where the reduction of fluorescent is first detected.
26. It normally takes 2 h for the sonication system to cool down and degas. Moreover, it is important to plan the timeline well in advance because there are several series of incubation steps in the procedure that includes a 16-h incubation (normally from 5 pm to 9 am next day).
27. The expected result is a smear between 200 and 1,000 bp.
28. Turn on the heat block to preheat to 95 °C for the next step.
29. We typically use a NanoDrop to measure RNA concentration and purity. The absorbance ratio of 260:280 should be ~2.0 indicative of pure RNA, whereas a lower ratio indicates impurities such as phenol or proteins. The absorbance ratio of 260:230 should be in the range of 2.0–2.2 for RNA free of impurities.
30. cDNA synthesis should always be run with an additional reaction with no SuperScript III enzyme (often referred to as no RT control), which is used to determine the presence of contaminating DNA. The presence of contaminating DNA in RNA will affect evaluation of the results.
31. It may be necessary to titrate the MgSO_4 concentration (range = 1.5–3 mM).
32. The use of too much template saturates this sensitive reaction.
33. It is essential to avoid quantification by NanoDrop, as it is too inaccurate for this protocol.
34. Another option is to use enzymatic shearing and kits provide the appropriate enzymes to perform this. However, we have achieved better outcomes through shearing.
35. It is essential to ensure that the syringe plunger is held in position.
36. Storage on ice is essential.
37. It is essential to run a previously analyzed sample to ensure reproducibility, as a quality control measure.
38. We extract raw data from the Ion Torrent and solely use the CLC Genomics Workbench for mapping and variant calling.
39. Normalization of the melting curve must be done carefully in order to obtain flat lines on each side of the difference curves.
40. It is suggested to run an internal control on every HRM plate in order to standardize the difference curves from the serial dilutions of reference-variant mtDNA.

Acknowledgments

This work was supported by the Victorian Government's Operational Infrastructure Support Program and NMHRC Project Grants GNT1022222 and GNT 1041471 to J.C.S.J.

References

1. Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. *Nature* 290(5806):457–465
2. Kuciej M, Butow RA (2007) Evolutionary tinkering with mitochondrial nucleoids. *Trends Cell Biol* 17(12):586–592. doi:[10.1016/j.tcb.2007.08.007](https://doi.org/10.1016/j.tcb.2007.08.007), S0962-8924(07)00242-5 [pii]
3. Gill P, Ivanov P, Kimpton C et al (1994) Identification of the remains of the Romanov family by DNA analysis. *Nat Genet* 6:130–135
4. St John JC, Facucho-Oliveira J, Jiang Y, Kelly R, Salah R (2010) Mitochondrial DNA transmission, replication and inheritance: a journey from the gamete through the embryo and into offspring and embryonic stem cells. *Hum Reprod Update* 16(5):488–509. doi:[10.1093/humupd/dmq002](https://doi.org/10.1093/humupd/dmq002), dmq002 [pii]
5. Spikings EC, Alderson J, St John JC (2007) Regulated mitochondrial DNA replication during oocyte maturation is essential for successful porcine embryonic development. *Biol Reprod* 76(2):327–335. doi:[10.1095/biolreprod.106.054536](https://doi.org/10.1095/biolreprod.106.054536), biolreprod.106.054536 [pii]
6. Kelly RD, Mahmud A, McKenzie M, Trounce IA, St John JC (2012) Mitochondrial DNA copy number is regulated in a tissue specific manner by DNA methylation of the nuclear-encoded DNA polymerase gamma A. *Nucleic Acids Res* 40(20):10124–10138. doi:[10.1093/nar/ghs770](https://doi.org/10.1093/nar/ghs770)
7. Dickinson A, Yeung KY, Donoghue J, Baker MJ, Kelly RD, McKenzie M, Johns TG, St John JC (2013) The regulation of mitochondrial DNA copy number in glioblastoma cells. *Cell Death Differ* 20(12):1644–1653. doi:[10.1038/cdd.2013.115](https://doi.org/10.1038/cdd.2013.115)
8. Kelly RD, Sumer H, McKenzie M, Facucho-Oliveira J, Trounce IA, Verma PJ, St John JC (2013) The effects of nuclear reprogramming on mitochondrial DNA replication. *Stem Cell Rev* 9(1):1–15. doi:[10.1007/s12015-011-9318-7](https://doi.org/10.1007/s12015-011-9318-7)
9. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324(5929):930–935. doi:[10.1126/science.1170116](https://doi.org/10.1126/science.1170116)

