2016

Functional Significance of mtDNA Cytosine Modification Tested by Genome Editing

Jason M. Robinson
Virginia Commonwealth University

Functional Significance of mtDNA Cytosine Modification Tested by Genome Editing

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

By

Jason M. Robinson

Director: Dr. Shirley M. Taylor, Ph.D.

Associate Professor of Microbiology and Immunology

Virginia Commonwealth University

Richmond, Virginia

September 2016

Acknowledgment

Throughout my time at MCV there have been many people that have supported me in the completion of my work. As a father of two children, with a wife who works and goes to school fulltime it has been quite the struggle. I would just like to take the time to acknowledge the people that drove me to be my best; that kept me going when I could have turned and cut loose for easier but less fulfilling options. I had no idea until halfway through my work just how much respect I had gained for science and the people in the scientific community. Being around people who have a respect and passion for science is contagious in the best way. These people, each in their own little or big way helped me realize just how far I could push myself as a student, friend and father.

I first would like to thank the current and past members of the Taylor lab. Dr. Shirley Taylor was an incredibly brave woman when she took me on and graciously invited me to work in her lab. In the past two years we have been through a lot together; I believe her mentorship extended beyond the scope of the lab and into a supportive role that a graduate student could wish for. I am fortunate to have had someone so passionate, knowledgeable and understanding as an advisor. Dr. Tim Lochmann and Elliot Burton were always there to give me a little extra advice or encouragement to keep me going. I appreciate greatly the extra time Dr. Lochmann took to prepare me for my presentations. A colossal thanks has to go to the person who had to endure the most, whom is Dr. Lisa Shock. I do not have enough room in my entire thesis to describe how patient and helpful she has been with me. She has fielded so many questions from me; whether it has actually been important or most likely asinine, she has always answered with thought and care. I owe her thanks for teaching me most of the techniques I have learned in the lab and for continuing to listen to all my little rants and especially the ones of the epic variety.

I would also like to thank my committee members, Dr. Rita Shiang, Dr. Mike Grotewiel and Dr. Mike McVoy. I am grateful for all of the insight and advice they have given me throughout. Their direction during committee meetings or just unplanned office visits was invaluable and their support helped me through more than they may know.

I must thank the neighboring labs in the Massey Cancer Center. The members of the Moran, Hackett, Landry, Lloyd and Ginder labs all have contributed to my work in some sort of way. Whether it was to borrow the use of equipment or borrow scientific insight, all the friendship and cooperation was a blessing to me. I hope in any of my next adventures that the coworkers are even half as supportive as my surrounding labs.

Lastly, I must give the biggest thanks to my family. My parents and Uncle have supported me when others told me I was on my own; their contribution will never be forgotten. It may be odd but I would like to thank my children Avery and Crew; countless times I had to get up from what I was doing to get you to clean up or otherwise entertain you. They were my biggest distraction, however, they were the ones that keep me focused on my goals and my ultimate reason for succeeding. To the woman who gave me those beautiful children and works just as hard if not harder than me, Christie Robinson, I would just like to thank you for putting up with me and loving me; it is truly the hardest job of all.

Table of Contents

List of Tables

List of Figures

ABSTRACT

FUNCTIONAL SIGNIFICANCE OF MTDNA CYTOSINE MODIFICATION TESTED BY GENOME EDITING

By Jason M. Robinson

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2016

Major Director: Dr. Shirley M. Taylor, Ph.D.

Associate Professor of Microbiology and Immunology

The field of epigenetics is gaining popularity and speed, due in part to its capability to answer lingering questions about the root cause of certain diseases. Epigenetics plays a crucial role in regulation of the cell and cell survival, particularly by cytosine methylation. It remains controversial if DNMT's which facilitate methylation are present in mammalian mitochondria and what the functional significance they may have on modification of mitochondrial DNA. CRISPR-Cas9 technology enabled genome editing to remove the MTS (mitochondrial targeting sequence) from DNMT1 of HCT116 cells, purposefully minimizing effects on nuclear cytosine methylation, while exclusively impacting mitochondrial modification. Removal of the DNMT1 MTS did not completely prevent the localization of this enzyme to the mitochondria according to immunoblot analysis. As well, deletion of the MTS in DNMT1 revealed only a small decline in transcription; not until removal of DNMT3B did we see a two-fold decrease in transcription from mitochondrial protein coding genes. No significant decline in transcription occurred when a

DNMT3B knockout also lost the MTS of DNMT1; this study is evidencing that DNMT3B is possibly the more significant methyltransferase in the mitochondria. Our aim from this study and future research is to clearly characterize which enzymes in the mitochondria are controlling cytosine modifications and to understand the mechanistic complexities that accompany cause and consequence of epigenetic modifications.

Chapter 1: Introduction

Mitochondrial Overview

If you have ever taken a high school biology class then you probably have heard that the mitochondrion is the powerhouse of the cell; broadly your teacher was trying to explain that mitochondria are responsible for generating most of your cells' energy. In more scientific terms, mitochondria are able to convert ADP into ATP through the process of oxidative phosphorylation, thereby providing the cell with a source of usable chemical energy [1].

The mitochondrial organelle is comprised of several components that have specialized function. The outer membrane encloses the entire organelle and contains integral membrane proteins called porins which are responsible for either actively or passively transporting proteins, depending on size, into the intermembrane space [1]. The inner mitochondrial membrane divides the mitochondrial matrix from the intermembrane space and contains import machinery with specific transport proteins that control metabolite pathways in and out of the matrix as well as ATP synthase and proteins involved in redox reactions [2]. The inner mitochondrial membrane is expanded by the cristae, invaginations that lead to additional surface area thus increasing the capacity to create ATP [2]. At the core of the mitochondrion is the matrix; this compartment is enclosed by the inner membrane and comprises the bulk of the mitochondrial mass. The matrix houses the mitochondrial DNA along with all the replication, transcription and translational machinery; it also contains free ribosomes and enzymes that facilitate the reactions that produce ATP during the Citric Acid Cycle and oxidative phosphorylation [2].

Beyond energy production, the mitochondria also play pivotal roles in cell cycle regulation, cell growth, cell death, cellular signaling and differentiation. Mitochondria have been implicated in

variety of diseases from cardiac dysfunction to neuromuscular myopathies, and, as a result, are under intense scrutiny [3]. Due to recent findings, mitochondrial defects are finally being considered for their possible correlation in children with severe autism [4]. As mitochondria join in fundamental processes of cellular metabolism, research is starting to address the role of mitochondrial DNA in apoptosis, aging, and complex diseases such as ALS and cancer [3][5].

Mitochondrial DNA

MtDNA is a separate, organelle-specific genome; in humans, this genome contains 16,569 base pairs that encode 37 genes and lack noncoding regions [3]. These genes code for 2 rRNAs, 22 tRNAs and 13 proteins that are integral components in complexes of the electron transport chain. MtDNA is double-stranded and circular, comprised of a heavy outer strand and inner light strand [3][6] (Figure 1-1). The outer strand is said to be heavy because it contains more purines than the inner lighter strand which contains more pyrimidines, which lack the extra ring structure that purines have [6].

Although it is small in comparison to the nuclear genome (3 billion base pairs) the mitochondrial genome can contribute about 1% of the cell's total DNA [7]; each mitochondrion contains 1 to 15 copies of mtDNA and human cells can house thousands of mitochondria [8]. However, despite having global influence from its energy production, and other critical roles along with having its own genome, the mitochondrion is nowhere near autonomous. Most mitochondrial proteins, including those that are involved in organellar transcription, translation, genome regulation and replication are transported to the mitochondria from beginnings in the nucleus [9].

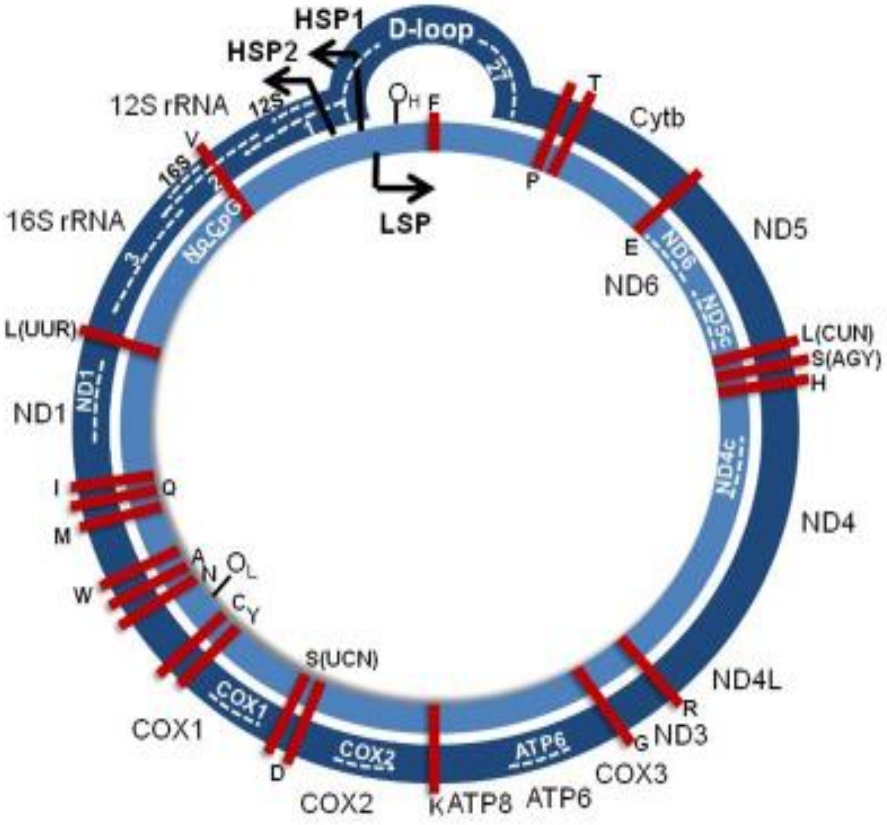


Figure 1-1. **Diagram of the circular mitochondrial genome**. Heavy strand with displacement loop is illustrated in dark blue and light strand in light blue. Genes that are labeled on the outside of the circle are associated with the heavy strand while genes on the inside are associated with the light strand and tRNA genes are depicted in red. Figure made by Lisa Shock

The Mitochondrial Displacement Loop (D loop)

The mitochondrial D loop carries both heavy (HSP) and light strand (LSP) promoters in close proximity to each other. Immediately downstream of the LSP is the origin of replication for the H strand (O_H) (Figure 1-1). The D loop is believed to exist as a triple helical structure in complex with a short LSP-driven transcript [9], which acts as primer in replication of the heavy strand. Super-high resolution microscopy studies [33,34] indicate that relatively few genomes within a given mitochondrion, possible only one, is engaged in replication at any time. Transcription from both HSP and LSP is believed to operate in a similar fashion: TFAM (Transcription Factor A Mitochondria) binds to each promoter, bending the DNA 180^o through interaction of each globular domain with specific binding sites on each promoter. The linker region of TFAM only assumes α-helical structure through interaction with DNA (Taylor and Hackett, unpublished observations). The DNA bend induced by TFAM brings its C-terminal domain in close proximity to the transcription start site, allowing it to recruit TFB2M (Transcription Factor B2 Mitochondria), which partially melts the promoter region, and mitochondrial RNA polymerase (PolRMT), and allowing initiation of transcription. Each functional sequence domain required for replication and strand specific transcription are in very close proximity, leading to potential structural conflict in a genome attempting to carry out replication as well as transcription.

Implications of multiple genomes within the mitochondrion

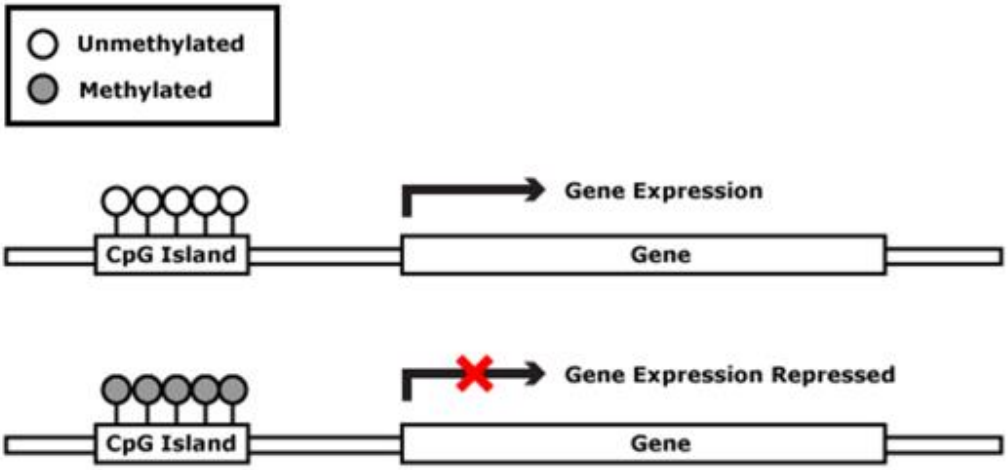
The mitochondrion, as mentioned, contains 1 to 15 copies (genomes) of mtDNA that code for 37 genes. Current research is leading to the hypothesis that the individual copies of mtDNA in the mitochondrion are not working on multiple functions at the same time, and that

each genome has specific instructions to carry out independent functional roles. It is thought that the genomes are partitioned into either replicating, transcribing or inert, packaged within TFAM-coated nucleoids; the underlying mechanism controlling such partitioning is at present unknown. However, this hypothesis allows for the possibility that the cell can modulate its response to energy demands or stress by altering the balance of individual genomes engaged in each activity. The basis of this thesis is that modulation of functional roles among the mitochondrial genomes is controlled, at least in part, by epigenetic modification.

Epigenetics

The field of research that is starting to answer lingering questions about the root cause of certain diseases is the field of epigenetics. Epigenetics is the study of heritable alterations in gene expression that are not caused by a change in DNA sequence [10]. Epigenetics plays a crucial role in regulation of the cell and cell survival, particularly by cytosine methylation [10]. It is an important balance of methylation that must be maintained in order for the cell to display a continual normal transcription. While not involved in any sequence change, epigenetic modifications such as methylation are directly related to how the transcriptional machinery can gain access to a DNA sequence and as a result controls which genes are transcribed or silenced [11].

Epigenetic regulation in the nucleus operates on several levels; DNA methylation/demethylation, histone modification, and assimilation of histone variants [12]. Our studies focus on DNA methylation since mtDNA lack histones [7]. Besides regulation of gene expression DNA methylation in the nucleus is involved in several significant processes, such as chromosomal X inactivation, imprinting, cellular reprograming, immune system development



http://missinglink.ucsf.edu/lm/genes_and_genomes/methylation.html

Figure 1-2. Diagram of CpG Islands. An abundance of genes in the human nuclear genome possess upstream CG-rich areas, known as CpG islands. CpG islands, if methylated can represses gene expression; whether it is helpful or harmful is dependent upon the functionality of the gene.

and conservation of the genome integrity through guarding against endogenous retroviruses and transposons [13].

In eukaryotic nuclei, DNA methylation takes place on cytosine residues at the carbon 5 position, principally in CG dinucleotides or CpG islands [13]. The main consequence of this methylation is an inhibitory tag to repress transcription [14]. The human genome was found to have roughly 56 million CpG sites, of which 60-80% are methylated [13]. Aberrant DNA methylation arrangements have been linked to some human malignancies; in comparison to normal tissue, tumor cells frequently have hypermethylation at promotor CpG islands resulting in gene silencing, while the majority of the genome is hypomethylated [13] (Figure 1-2). In contrast, normal genes predominantly display hypomethylation in the promotor region and elevated intensities of methylation in the rest of the gene. The future of the epigenetic field is promising as modifications have shown to be reversible and are established as important therapeutic targets [15].

DNA Methyltransferases

Methylation is catalyzed by a family of enzymes termed DNA methyltransferases (DNMTs) [16]. S-adenosyl-methionine (SAM) is a cofactor that contributes a methyl group to the methyltransferase substrate, consequently methylating the C5 position of cytosine residues to produce 5-methylcytosine (5mC) [7] (Figure 1-3). It is a family of five genes that comprises the functional methyltransferase enzymes: DNMT1, DNMT2, DNMT3A/3B and DNMT3L [13].

In the nucleus the initial and preserved pattern of methylation during embryogenesis is established by DNMT3A and 3B, otherwise called the de novo methyltransferases which show no preference for unmethylated and hemimethylated DNA [17]. DNMT3A and 3B are the only

active members of the DNMT3 group as DNMT3L is catalytically inactive [17], but assists DNMT3a and 3B in their substrate recognition. DNMT2 is inactive in methylating DNA but instead acts as an RNA methyltransferase [17]. DNMT3A and 3B are both essential for development; DNMT3A knockout in mice causes neonatal lethality, while DNMT3B knockouts die *in utero* [14]. Mutations in either DNMT3 are associated with a variety of disorders such as Tatton-Brown syndrome, a DNMT3A *de novo* mutation that is characterized by overgrowth and intellectual disabilities [18].

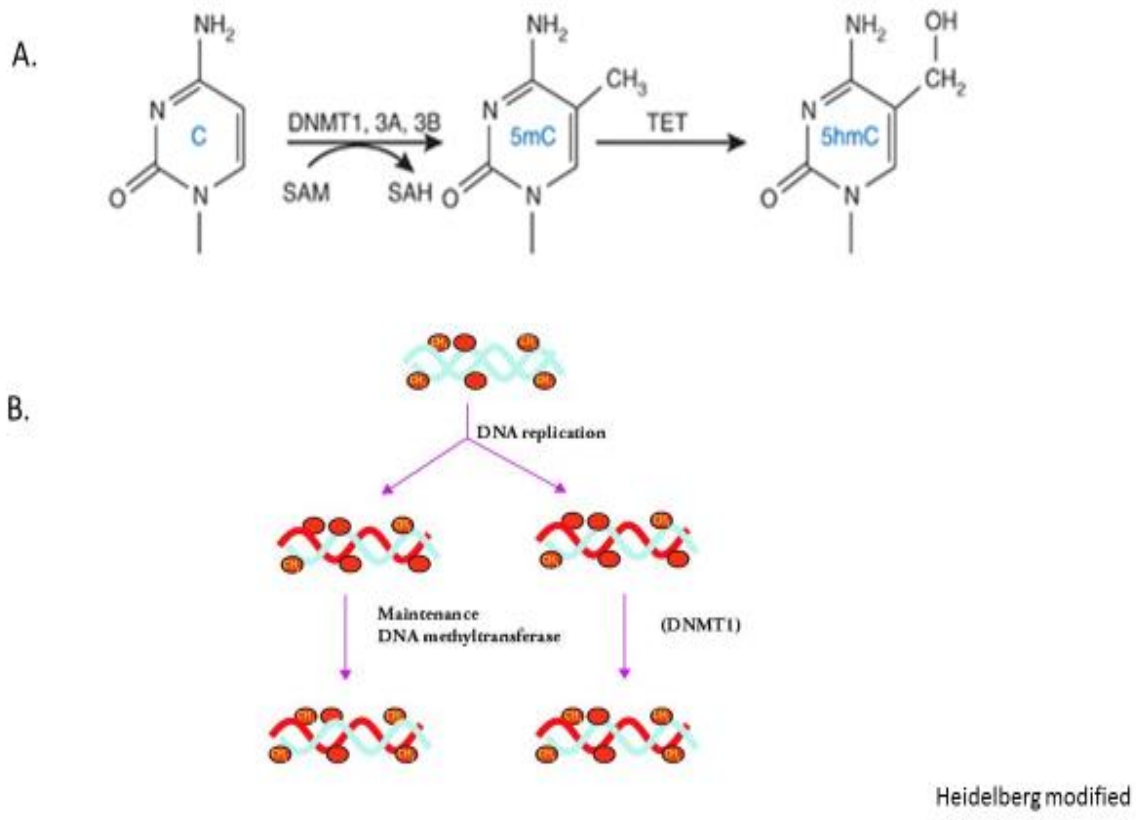


Figure1-3. DNA Methyltransferases. A) Methylation is catalyzed by DNA methyltransferases (DNMTs) on cytosines at the 5th position of the pyrimidine ring in CpG dinucleotides; S-adenosylmethionine (SAM) is the methyl donor that converts cytosine to 5-methylcytosine. In a subsequent reaction, 5mC can be hydroxylated by the ten-eleven translocase (TET) family of methylcytosine dioxygenases. B) The parental strand (red)of DNA has its methylation pattern copied to the newly created strand (blue) during replication. This ensures that the gene expression of the parental strand is passed on and is the same as the daughter strand.

A problem arises during replication where one strand will initially lose its pattern of methylation during each cycle, and this is where DNMT1 comes into play (Figure 1-3). DNMT1 maintains the existing pattern of methylation throughout replication and cell division [13]. DNMT1, unlike DNMT3A/3B, has a preference for hemimethylated DNA, and in addition is highly processive, being able methylate extensive runs without releasing its substrate DNA [13]. Like DNMT3B, DNMT1 is essential for development; DNMT1 knockouts in mice are embryonic lethal soon after gastrulation [14]. These DNMT knockout studies establish the vital nature of DNA methyltransferases and methylation in development and survival of mammalian cells.

Mitochondrial DNA Methylation

It has been long established what the role of DNA methyltransferases are in nuclear DNA methylation and nuclear DNA methylation affects biological function; however the role and presence of DNMTs in mitochondria is a particularly controversial issue that has been at debate for almost half a century. There are disbelievers in the science community about mtDNA being methylated and will exclaim that previous research was excessively sensitive in nature and lacking reproducibility in their experimental methods.

In 1971, DNA methyltransferase activity was detected in the mitochondria of a eukaryotic organism, suggesting the possibility that 5mC was present in mtDNA [19]. However, in those same years contrary evidence was presented stating the absence of 5mC in a mammalian model, consequently spawning the debate as to whether methylation exists in mitochondrial DNA [17].

Years later in 2004, researchers once again failed to observe mtDNA methylation in both gastric and colorectal cancers [20]. Although, they did divulge that frequencies of methylation less than 5% would not have been detected using their site-specific methylation restriction enzyme approach. This offers additional confidence to prior studies in 1984 that studied human fibroblasts and mouse fibroblastoids that identified methylation was present in mitochondrial DNA at 1.5-5% and solely in CpG dinucleotides [21] compared to mammals at 60-90%.

The pro-methylation side of the argument did not come back again until 2011 when researchers started to compile evidence for the cause. Novel methods were used by Shock et al. to establish the existence of 5mC in mtDNA; not only did these researchers display that 5mC is present in mtDNA but also demonstrated the presence of 5hmC, believed to be an intermediate in demethylation, and demonstrated that DNMT1 was the enzyme responsible for this modification [22].

Possibly the most startling observation was that Shock et al. discovered a mitochondrial isoform of DNMT1 (mtDNMT1). The isoform has a conserved in-frame mitochondrial targeting sequence (MTS) with an alternative start site just upstream of the recognized start site for nuclear DNMT1 [22]. A mitochondrial targeting software program (MitoProt II) predicted with >90% probability that the targeting sequence would localize to the mitochondria. They demonstrated, using confocal microscopy, that the mtDNMT1 MTS is able to deliver a heterologous protein (GFP) to the mitochondria [22].

Oxidative stress can cause the two principal regulators of mitochondrial biogenesis, proliferator-activated receptor c coactivator 1 (PGC1α) and nuclear respiratory factor 1 (NRF1) to activate and interact; this activation up-regulates several nuclear-encoded mitochondrial genes

[23]. Shock et al. was able to transiently transfect PGC1α, NRF1, or both together into HCT116 cells to analyze levels of mtDNMT1 by means of immunoblot. Co-transfection of both regulators exhibited a fivefold upturn in total mitochondrial DNMT1, thus showing not only the presence of a DNA methyltransferase in the mitochondria but additionally demonstrating biological association to oxidative stress [22].

Coincidently, the NRF1 binding site overlaps a p53 consensus binding site [24]. This tumor suppressor gene is known to control mitochondrial transcription, which lead Shock et al. to explore if mtDNMT1 mRNA expression is affected when p53 is deleted. They reported a 6-fold increase in mtDNMT1 transcription in p53 null MEFs and HCT116 cells, which reinforces previous reports that p53 has a suppressive influence on expression of DNMT1 through specific DNA binding [23].

Bisulfite modification transforms cytosine residues to uracil but leaves 5mC residues unaffected, and is an effective tool for studying DNA methylation in the genome [7]. Using this technique of bisulfite sequencing, a group reported that 5mC was restricted to the light strand of the D-loop region [16]. The same group was able to show expression of both DNMT1 and DNMT3B in mitochondria from HeLa cells, although, no other group has been able to show DNMT3B in the mitochondria [16].

Besides DNMTs that catalyze the conversion of cytosine to 5mC, there is also a family of enzymes that convert 5mC to 5 hydroxymethylcytosine (5hmC): ten-eleven-translocation (TET) enzymes [7]. The TET enzymes have been functionally identified in the nucleus but translocation of these proteins to the mitochondria remains controversial. The Taylor lab used immunoreactivity and confocal microscopy studies to show that Tet2 but not Tet1 localized to

the mitochondria; thus Tet2 is likely the enzyme responsible for the 5mC to 5hmC conversion [22]. This proof complements the belief that mtDNA is not only methylated but mitochondria contain the enzymes that allow for dynamic methylation and demethylation.

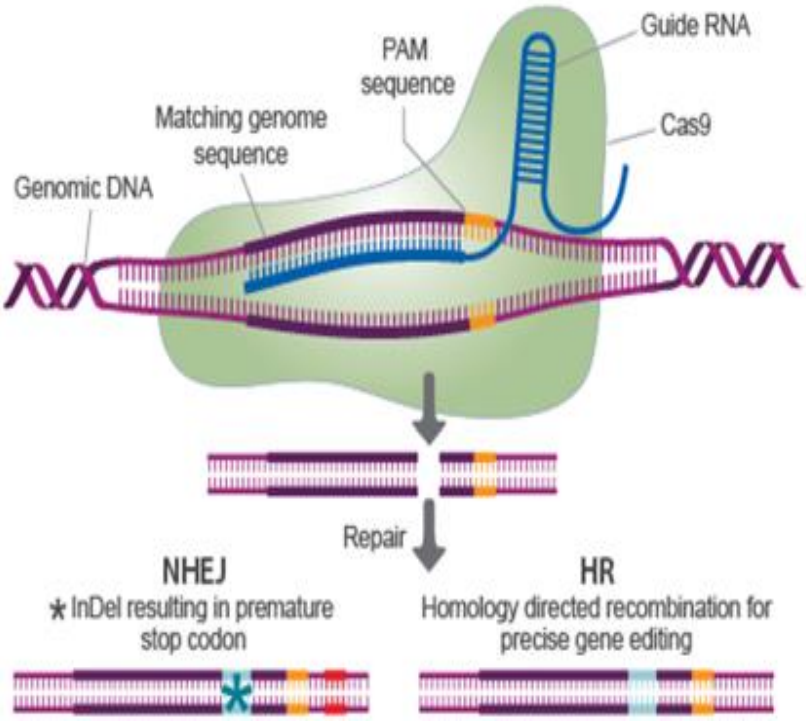
Epigenetics as an essential control mechanism within mitochondria

Previous work in the Taylor Lab has shown that increased methylation of mitochondrial genomes increases the affinity of TFAM for promoter sequences and increases transcription firing from both HSP and LSP *in vitro* and *in vivo* [25]. Epigenetics may therefore provide a mechanism whereby the individual genomes of the mitochondrion are partitioned into functional subgroups; investigating the patterns of methylation in individual genomes could characterize whether the mtDNA is actively replicating, being transcribed or at rest. Generation of cell lines with specific and purposefully altered patterns of methylation will be the starting point of studying functional differences in genomes of the mitochondrion.

CRISPR/Cas9 Genome Editing

Clustered Regularly Spaced Palindromic Repeats (CRISPR)/Cas9 is a bacterial immune system that has been adapted for use in multiple species and is rapidly becoming one of the most prevalent methods for genome editing. The system comprises a guide RNA (gRNA) which delivers an endonuclease, Cas9, to a specific site within the genome [26]. The gRNA contains a scaffold sequence which specifically binds the Cas9 endonuclease and delivers it to a 20bp target sequence [26] (Figure 1-4).

CRISPR/Cas9 can be used to generate site specific knockouts with a unique targeting sequence placed just upstream of a Protospacer Adjacent Motif (PAM); upon expression of the Cas9 endonuclease, it will bind the gRNA to form a riboprotein compound [27]. Upon binding, Cas9 will then undergo a conformational alteration which takes the molecule from an inactive state to an active one [27]. The complex is then able to cleave the target DNA resulting in a double-stranded break. The double-stranded break is then repaired by one of two pathways; non-homologous end joining (NHEJ) or homology directed repair [26]. NHEJ is error-prone and sometimes results in small insertions or deletions (INDELS) in the target DNA; this can result in frameshift mutations leading to premature stop codons [27]. If possible, the ideal end result is a loss-of-function mutation within the desired locus. Homologous recombination will occur at lower frequency if a repair oligo nucleotide, with homology to the sequences surrounding the break site are provided; this mechanism of repair is not error prone, and can be used to generate site specific mutations, insertions or deletions.



http://www.transomic.com/Products/CRISPR-Genome-Editing.aspx

Figure 1-4. Diagram of CRISPR/Cas9 genome editing. The Cas9 endonuclease will create a double stranded break. There are two mechanisms of repair for the break: non-homologous end joining (NHEJ), which generates chance insertions or deletions at the cleavage site, or homologous recombination, which generates targeted mutations centered on DNA template.

Possible Genomic Editing of mtDNMTs

As the interest in mitochondrial epigenetics grows and as the stack of evidence compiles that indicates mtDNA is methylated and that mtDNMTs are of biological significance, it is of interest to understand what the functional consequences are of mtDNMT loss. Mitochondria as mentioned before play pivotal roles in cell signaling and the cell cycle in addition to cellular respiration. Therefore, if DNMTs act in any way in the mitochondria as they do in the nucleus, then assessing their functional significance can lead to understanding of what happens when any of those cellular roles are defective.

Previous methyltransferase assays completed in the Taylor lab by Dr. Lisa Shock provided evidence that DNMT1 and DNMT3B are active in the mitochondria (Figure 1-5). These studies used mouse embryonic stem cells, previously generated by Li et al. [22], in which the catalytic domains of DNMT1, DNMT3A and DNMT3B, and combinations, were knocked out through gene targeting via homologous recombination. Mitochondrial lysates from these cells were used in *in vitro* methyltransferase assays [22]. Graphical analysis of this assay shows that DNMT1 and DNMT3B, but not significantly DNMT3A, are active in these cell lines.

Using methyltransferase assays and manipulating the mitochondrial isoform of DNMT1 that the Taylor lab has developed an approach to test the functional significance of mtDNA cytosine modification. We used CRISPR/Cas9 to perform genome editing to construct innovative cell lines lacking the DNMT1 MTS as well as in combination with DNMT3B KO cell lines.

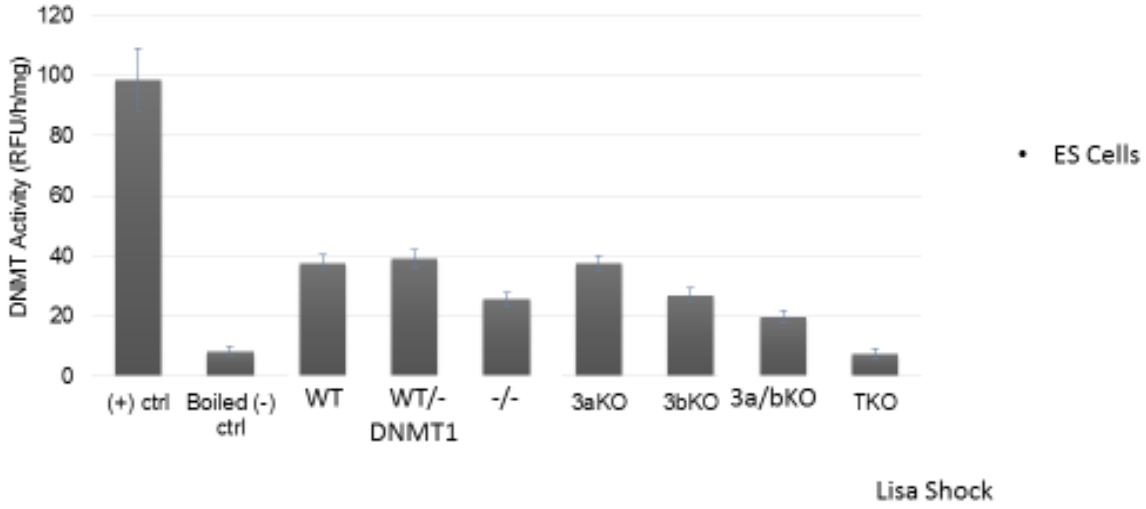


Figure 1-5. Methyltransferase assay. Previous work done by Dr. Lisa Shock in mouse embryonic stem cells. A Fluorometric EpiQuik DNMT Activity Assay Ultra (Epigentek) was used to measure activity in 10ug of protein from isolated mitochondria. Deletion of DNMT's catalytic domain was accomplished through gene targeting and homologous recombination. Heterozygous and homozygous knockouts of DNMT1 were investigated along with DNMT3A and DNMT3B knockouts with TKO being a deletion of all three methyltransferases. This data displays that DNMT3B is active in the mitochondria.

We reasoned that upon removal of the MTS of DNMT1, methyltransferase activity in the mitochondria will decrease. In addition, we proposed that the removal of the MTS of DNMT1 in a cell lacking DNMT3B activity would reduce mitochondrial methyltransferase activity even further. We hypothesized that these changes in mitochondrial methyltransferase activity would lead to changes in levels of DNA modification and functional output of these organelles. These reagents would be critical in determining whether mtDNA epigenetic modification is a mechanism whereby the functional partitioning of genomes within the organelle is dynamically controlled in response to the energetic needs of the cell.

This hypothesis was tested by

1) Methylation analysis using methylated DNA Immunoprecipitation (MeDIP), specifically within the mitochondria.

2) Mitochondrial Purification and Immunoblot Analysis of DNMT levels.

3) Mitochondrial Transcription analysis using RT-PCR.

Chapter 2: Methods

Cell Culture

HCT116 (CRL 1165) neomycin resistant wild-type, HCT 3BKO (gift from Dr. Bert Vogelstein, Johns Hopkins University) and genome edited derivative cell lines were grown at 37ºC in 5% CO_2 in RPMI 1640 media (Gibco/Life Technologies) containing 10% fetal bovine serum. For all experiments performed, cells were fed 48 hours before harvest, trypsinized and replated 24 hours before harvest, identically to ensure consistency. DNMT1 is cell-cycle regulated so it was critical to harvest using this method to ensure that the bulk of the cells were in S-phase, where DNMT1 levels are at their peak.

Cellular Fractionation

To obtain purified mitochondria, crude whole cell lysates and cytosolic fractions, cells were harvested from 3x150mm dishes. The cells were washed two times with chilled 1x PBS, pH 7.4 while on ice. Cells were scraped from the dish and transferred to a pre-weighed 15ml conical tube (BD Falcon). An aliquot of 10% percent from each cell line was separately reserved, pelleted and stored at -80ºC for generation of whole cell lysate. The remaining cells were then pelleted and a whole cell pellet weight was acquired. Three mL of fresh mitochondrial homogenization buffer (10mM Tris-HCl, pH7, 0.25M sucrose, and 1mM EDTA, pH 6.8) containing 1 tablet complete protease inhibitors (Complete EDTA-free Protease Inhibitor Cocktail tablets, Roche) per 25ml buffer were added to resuspend cell pellets. The suspended cells were conveyed into a chilled 7ml glass dounce homogenizer and kept on ice for 5 minutes to permit cells to swell in the hypotonic homogenization buffer. Three series of douncing (15 down and up strokes per series) were performed, each followed by a 2,200rpm centrifugation for

5 minutes at 4°C. The post-nuclear supernatant that contained the cytosolic and mitochondrial fractions were saved after centrifugation. The remaining unbroken nuclear pellet was resuspended and subjected to douncing until three series were complete. The pooled post-nuclear supernatant was again centrifuged to pellet any contaminating nuclear fraction and collected in a 14ml round-bottom tube (BD Falcon). The supernatant was then centrifuged at 9,000 rpm for 15 minutes at 4°C to obtain a crude mitochondrial pellet with the supernatant retained to prepare the cytosolic fraction.

The mitochondrial fraction was further purified by taking the crude mitochondrial pellet and adding 1mL of homogenization buffer with protease inhibitors and centrifuging for 9,000 rpm for 15 minutes at 4°C two times to wash the pellet. The pellet was then washed again as before but without protease inhibitors and resuspended in a volume 10 times the pellet weight in Tyrpsin Digestion Buffer (250mM Sucrose, 2mM EDTA, .5mM EGTA, 10mM HEPES-KOH and 1mM DTT). Typsin-EDTA (Gibco) 100µg/mL was added to reach a final concentration of 10µg/mL. Samples were held at 21°C for 20 minutes with intermittent gentle inversion to mix contents. After incubation, bovine trypsin inhibitor (Sigma) 1mg/mL was added to a final concentration of 10µg/mL; tubes were inverted to mix and placed on ice for 10 minutes. The contents were then centrifuged at 9,000 rpm for 10 minutes at 4°C to pellet intact mitochondria. The pellet was then washed twice with 1mL of homogenization buffer with protease inhibitors and 10µg/mL bovine trypsin inhibitor and centrifuged as before. The final mitochondrial pellet was resuspended and vortexed in SDS Lysis Buffer at 10 times the pellet weight and stored at -20°C.

The cytosolic fraction was further purified by acetone precipitation. Acetone was added to cytosolic fraction at 4x volume and held at -20°C overnight. Tube was then centrifuged at max

speed for 15 minutes at 4°C. Cytosolic pellet was then separated from the supernatant and air-dried for 10 minutes. The cytosolic pellet was then resuspended in 100µl SDS Lysis Buffer and stored at -20°C.

The whole cell lysate was prepared by centrifuging 2500rpm at 4°C for 5 minutes and PBS removed. The pellet was then resuspended in SDS Lysis Buffer at 7.5x volume and passed through a 21G needle to obtain a homogenous lysate, and then stored at -20°C.

Immunoblotting

Protein concentration for lysates was determined using a Bradford assay. A standard curve was made with known concentrations of BSA and was used to determine the protein concentration of the protein lysates. The assay was quantified using a Genesys 10uv spectrophotometer (Thermo). Samples were loaded onto 4-20% SDS-PAGE precast gels. 30µg of protein were loaded for whole cell lysate comparisons. 75µg of whole cell lysate, 25µg of cytosolic lysate and 18µg of mitochondrial lysate were loaded for subcellular fractionation comparisons, to approximate loading of equal cell equivalents. The required volume of lysate was then mixed with an equivalent volume of Laemmli buffer+ 5% BME. Samples were then boiled for 10 minutes without placing on ice prior to being loaded into a gel. Samples were loaded alongside 6 µL of Precision Plus Dual Color ladder (BioRad). The gels were electrophoresed at 120V for ~1 hour in 1x running buffer (25mM Tris base,0.1% SDS, 250mM glycine in 1L total volume). Resolved proteins were then transferred to an Immobilon-FL PVDF membrane in transfer buffer, consisting of 200mL methanol, 700mL ddH_2O and 100 ml 10X transfer buffer (16.879g TrisHCl, 144.134g Glycine, 10g SDS, 17.299g Tris Base). The gel was rinsed with ddH2O and transferred to the 1x transfer buffer for 15 minutes at 4°C for

equilibration. The membrane was activated in methanol for ~1 minute and washed with ddH2O for ~1 minute prior to being transferred into 1x transfer buffer for 15 minutes at 4°C for equilibration. The transfer apparatus (gel, membrane, transfer plate, sponges, Whatman paper) was carefully assembled and placed in the transfer box with 1x transfer buffer at 4^{o}C and transfer was carried out for 1 hour at 100V. After the transfer was complete the apparatus was disassembled and the membrane was washed for 1 minute with ddH_2O and then activated in methanol for 30 seconds. The activated membrane was then ready to be checked for a positive transfer with Ponceau stain (Bio-Rad); 2 minute soak then visualize when slowly rinsing off with ddH_2O. The washed membrane was then placed in StartingBlock T20 (ThermoFisher) overnight at 4^{o}C with shaking. The appropriate primary antibody diluted in StartingBlock T20 was then added to the membrane with light shaking at room temperature for 1 hour. After primary antibody the membrane was washed three times with 1x TBS-T (0.5M Tris-HCl, 2.7mM KCl, 0.14M NaCl and 0.1% Tween 20) with heavy shaking for 15 minutes. The membrane was then lightly shaken in the appropriate secondary antibody diluted in StartingBlock T20 for 1 hour at room temperature. After the secondary antibody incubation the membrane was washed again in 1x TBS-T as before. The membrane was then blotted dry to remove excess buffer and SuperSignal West Dura (Pierce) was added so the blots could be developed on a Li-Cor Odyssey system.

Table 1: List of antibodies and optimal conditions.

Antibody	Manufacturer	Species	Blocking Buffer	Primary Dilution	Secondary Dilution	Secondary Antibody	Protein Size(kD)
Tubulin	Millipore	Mouse	Starting Block	1:25,000	1:10,000	Goat Anti-mouse	50
H3K4me3	Millipore	Rabbit	Starting Block	1:500	1:10,000	Goat Anti-rabbit	32
VDAC	Sigma	Rabbit	Starting Block	1:1,000	1:10,000	Goat Anti-rabbit	17
DNMT1	LS Bio	Mouse	Starting Block	1:1,000	1:10,000	Goat Anti-mouse	185
DNMT3b F2	Santa Cruz Biotechnologies	Mouse	Starting Block	1:200	1:10,000	Goat Anti-mouse	Variable
DNMT3b D1	Santa Cruz Biotechnologies	Mouse	Starting Block	1:200	1:10,000	Goat Anti-mouse	Variable
DNMT3b G9	Santa Cruz Biotechnologies	Mouse	Starting Block	1:200	1:10,000	Goat Anti-mouse	Variable

CRISPR/Cas9 Genome Editing

In order to create a frame-shift mutation within the MTS a single guide RNA was created, which, along with Cas9 endonuclease should cause small insertions or deletions (indels) through the error-prone non-homologous end joining (NHEJ) repair mechanism. Several potential gRNAs were screened for potential off-target cleavage events by blast search and using the MIT gRNA and off-target prediction tools found at *www.**crispr**.mit.edu/*. The selected gRNA is highlighted in yellow in Figure 2-1.

gRNAs were synthesized by extension of overlapping oligonucleotides [33] (guided literature) and inserted into the gRNA expression vector by Gibson Assembly according to the manufacturers protocol (New England Biolabs). Ligation reactions were transfected into competent DH5a and clones were screened by DNA sequence analysis.

HCT116 and its derivatives (3BKO and DKO) were plated at $1x10^6$ cells per well in a 6-well plate. Transfections were carried out using polyjet lipid based transfection reagent (SignaGen Scientific) using 2ug DNA and 3uL polyjet per well. Transfection efficiency was determined to be >80% based on transfection of control wells with pMaxGFP. To minimize prolonged expression of Cas9 endonuclease, cells were co-transfected with a puromycin expression vector (pLVTHMpuro) and transfected cells selected by resistance to 1ug/ml puromycin. Cells were replated into 150mm dishes 48 hours post-transfection, and selection was applied 24 hours later. Additional controls included untransfected cells, and cells transfected with pLVTHM alone.

Approximately 10 days after transfection, individual colonies were picked using polyester-tipped sterile swabs and plated into 48well plates. To generate replica plates, clones

were individually released with trypsin-EDTA (Gibco) and diluted into 2 additional plates. After growth for 5-7 days, one plate was cryopreserved at -80^{0}C in 90% fetal bovine serum 10% DMSO, one plate was harvested for DNA preparation and PCR-based screening, and one was maintained with bi-weekly medium changes.

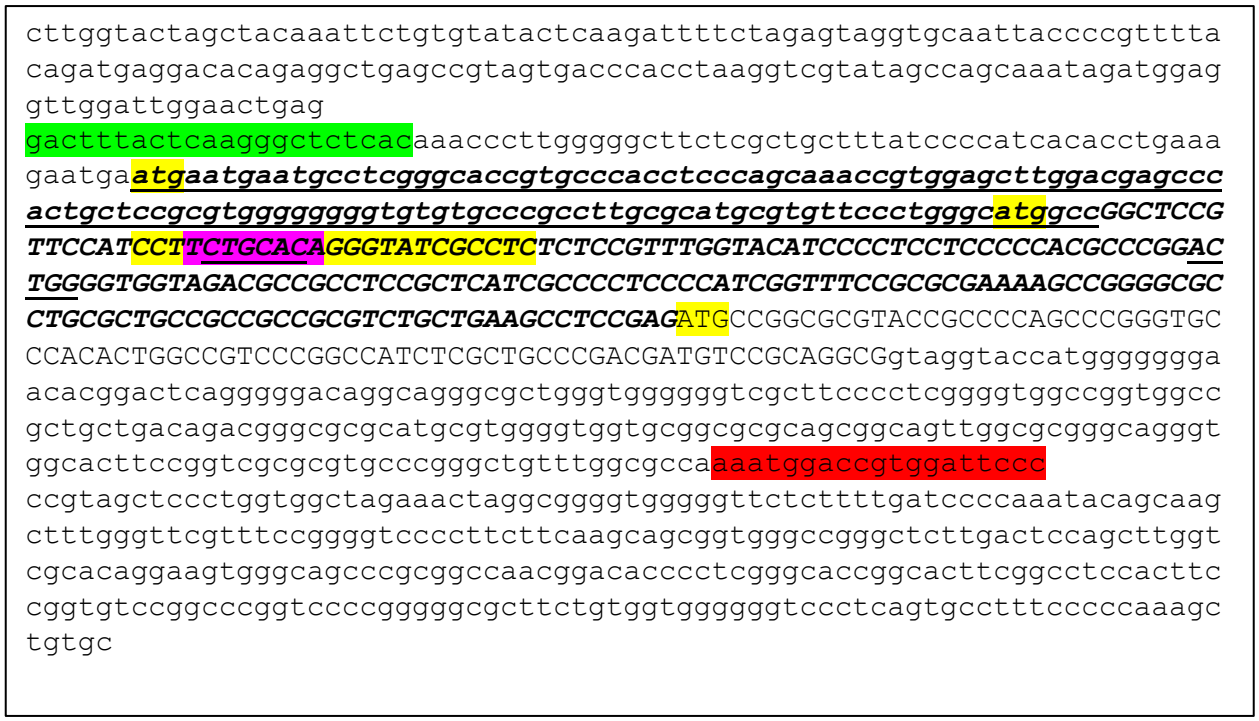


Figure 2-1. Genomic sequence surrounding exon 1 of the human DNMT1 gene.

Selected gRNA is highlighted in yellow. The screening amplicon lies between sense PCR primer (green) and antisense primer (red). Start ATG codons are indicated in yellow and the mitochondrial targeting sequence is in bold italics. The published transcription start site for the nuclear isoform begins at capitalized nucleotides

DNA Extraction and PCR Screening of genome edited cell clones

To select DNA clones for downstream application from CRISPR/Cas9 genome editing, DNA was extracted using MyTaq Extract-PCR Kit (Bioline). An individual well from a 48-well plate was harvested for each clone, using 50ul of MyTaq Extract buffer. Lysates were transferred to a 1.5ml microfuge tube and vortexed. The lysates were incubated for 5 minutes at 75^{o}C with intermittent vortexing then deactivated by heating to 95^{o}C for 10 minutes. Remaining cellular debris was pelleted and the supernatant transferred to a new 1.5ml microfuge tube. The supernatant was then diluted 10-fold in HPLC H_2O for PCR analysis.

PCR reactions consisted of 1µl of diluted template, .5µl of each forward and reverse primers (15µM), 12.5µl of My Taq HS Red Mix and 5.5µl of HPLC H_2O to total 25µl. The reactions were amplified in a Mastercycler (Eppendorf) and cycling conditions were as follows: 95^{o}C initial denaturation phase for 3 minutes, 95^{o}C denaturation phase for 15 seconds, 56^{o}C (primer dependent) annealing phase for 15 seconds and a 72^{o}C extension phase for 20 seconds and held at 4^{o}C as needed.

To screen for positive clones in genome edited samples a T7 endonuclease assay was performed. T7 endonuclease will recognize and cleave mismatches in a imperfectly matched DNA duplex. The reactions were all performed in a Mastercycler (Eppendorf) and set up as follows: 13µl of HPLC H2O, 2µ NEB Buffer 2 and 3µl of wild-type HCT116 amplicon. Reactions were performed in duplicate so that a side by side comparison could be done between wild-type and genome edited samples; 2µl of HPLC H_20 were added to wild-type and 2µl of genome edited sample were added to one of the samples for the duplicate reaction. The samples

were placed in the Mastercycler for a round of hybridization with cycling conditions as follows: initial denaturation at 95oC for five minutes, gradual annealing of -2 ℃ /s to 85 ℃ then -0.1 ℃/s to 25 ℃ and held at 4ºC. T7 endonuclease (NEB) (0.1units/sample) was added to each reaction carrying a mixture of wt and genome edited amplicon, and the reactions were incubated at 37ºC for 15 minutes in the Mastercycler. Products (20ul) were mixed with 1 µL 10x DNA loading dye and resolved on a 1% agarose gel in 1x TAE with EtBr at 80V for 30 minutes. The gels were viewed with a FluorChem Sp (Alpha Innotech). A cleaved band on the gel would indicate a mismatch in DNA indicative of positive genomic editing. Samples with a cleaved band were sent out to be sequenced (Eurofins) to determine specific mutations generated.

Each clone identified to carry mutations likely to disrupt the MTS was analyzed using Mitoprot II to increase confidence that the enzyme would not translocate to the mitochondria. In some instances, the indel generated a new ATG codon with a measurable probability of translocation to the mitochondria. These clones were not analyzed further.

Sub-cloning to validate purity of individual genome edited clones.

Clones carrying indels that produced the desired interruption of the mitochondrial targeting sequence were expanded, cryopreserved, and sub-cloned to ensure purity of the cell line. Cells were plated at a density of 100 cells/100mm dish, and individual colonies were picked 10 days after plating as described above. Each clone was cryopreserved and analyzed for the same indel as the parent clone. Ten sub-clones were analyzed for each parental clone and only those for which all ten showed identical sequences were maintained for further analysis.

RNA Isolation

RNA was isolated by harvesting HCT116 cells from one ~80% confluent 60mm dish. The cells were then lysed in 1mL TRI Reagent (Applied Biosystems) and incubated at room temperature for 5 min for nucleoprotein complexes to dissociate. 200µL of chloroform was added, samples were vortexed and incubated at room temperature for 10 minutes. The samples were then centrifuged at 12,000 x g for 10 minutes at 4^{o}C and the top aqueous phase containing RNA was transferred to a new tube. 500µL of isopropanol was added, samples were vortexed and incubated at room temperature for 5 minutes. After incubation the samples were centrifuged at 12,000 x g for 8 minutes at 4^{o}C and the supernatant was discarded leaving a precipitated RNA pellet. To wash the pellets 1mL of 75% ethanol was added; pellets were centrifuged for 5 minutes at 7,500 x g and the ethanol removed. The remaining RNA pellet was then air dried for 5 minutes and resuspended in 100µL RNase-free H_2O.

RNA was further purified by removal of genomic DNA by DNase I (RNase free) (Life Technologies). 10X DNase I Buffer was added to the RNA sample to make a 1X concentration followed by adding 1µL DNase I per 1µg DNA with a 30 minute incubation at 37^{o}C. After the incubation period the RNA sample was heated to 75^{o}C for 10 minutes to inactivate the DNase I.

The RNA was analyzed for contamination by agarose gel electrophoresis. ~1µg of RNA solution was heated to 70^{o}C for 1 minute and placed on ice before loading. The samples and a 1kb + Ladder (Invitrogen) were resolved on a 1% agarose gel in 1x TAE with EtBr for 50 minutes at 60V. The products were then analyzed for degradation (smearing) and band intensities on a FluorChem Sp (Alpha Innotech).

First Strand Synthesis

First strand cDNA was created with the SuperScript ™ III First-Strand Synthesis System for RT-PCR (Life Technologies). All components used were briefly mixed and centrifuged before being combined in a .5mL tube. Total RNA (5µg) was mixed separately with either 1µL 2µM ATP6 forward, ATP6 reverse primers or 1µL random hexamers, 1µL (10 mM) dNTP mix, and HPLC H_2O to a total volume of 10µL. Negative no template (2µL HPLC H_2O) control and positive Hela (5µg) controls were run as well. The reactions were heated to 65^oC for 10 minutes and then placed on ice for 1 minute. The cDNA synthesis mix was prepared by adding 80µL of 10X RT buffer, 160µL of 25mM MgCl2, 80µL of 0.1M DTT, 40µL of RNaseOUT™ (40 U/µl) and 40µL of SuperScript® III RT (200U/µL). 10µL of cDNA Synthesis Mix was added to each RNA/primer reaction and gently mixed and centrifuged. The reactions were incubated for 50 minutes at 50^oC followed by termination of the reaction at 85^oC for 5 minutes and then chilled on ice. The reactions were collected by centrifugation, 1µL of RNase H was added and samples were incubated at 37^oC for 20 minutes. For long term storage, reactions were stored at -20^oC.

Table 2: Human Mitochondrial Primer Sequences.

Primer Set	Forward	Reverse
Amplification of CRISPR/Cas9 Genome Edited DNMT1		
LSP	TCTGGCCACAGCACTTAAAC	TGTGTGCTGGGTAGGATGG
Endpoint PCR		
HSP	TCCCACTCCCATATACTAATCTCATC	AGGACCAAACCTATTTGTTTATGG
RT-PCR		
16s rRNA	ACCTTACTACCAGACAACCTTAGCC	TAGCTGTTCTTAGGTAGCTCGTCTGG
ND1	TGCGAGCAGTAGCAAACAAT	TGATGGCAGGAGTAATCAGAGG
Cox1	TGGTAACTGCCCATGCTTT	GTGTAAGCATCTGGGTAGTCTG
ND6	AAACACTCACCAAGACCTCAACCC	ATTGATTGTTAGCGGTGTGGTCGG
GAPDH	CCAGGTGGTCTCCTCTGACTTC	TTGGAGGCCATGTGGGCCATGA
Actin	TCATACTCCTGCTTGCTGATCC	CCACGAAACTACCTTCAACTCC
qPCR		
HSP	TCCCACTCCCATATACTAATCTCATC	AGGACCAAACCTATTTGTTTATGG
LSP	TCTGGCCACAGCACTTAAAC	TGTGTGCTGGGTAGGATGG
12s	AGTTCACCCTCTAAATCACCACG	TGACTTGGGTTAATCGTGTGACC
16s rRNA	ACCTTACTACCAGACAACCTTAGCC	TAGCTGTTCTTAGGTAGCTCGTCTGG
no CpG	CTGGTGATAGCTGGTTGTCCAAGA	CCTAGTGTCCAAAGAGCTGTTCCT

RT-PCR

cDNA synthesis was carried out with three different priming methods. ATP6 forward and ATP6 reverse cDNA primers allowed measurement of independent, strand-specific transcription from each mitochondrial promoter (HSP and LSP). For heavy strand transcription, PCR primers located within the 16s, ND1 and Cox1 open reading frames (ORFs) were used in separate qPCR reactions. For light strand transcription, ND6 primers were used. Nuclear genes encoding GAPDH and Actin were used to normalize RT-qPCR data, using cDNA synthesized with random hexamers. The reactions consisted of 10µL SsoFast Universal SYBR Green SuperMix (BioRad), 0.75µL each of forward and reverse primers (10mM), 7.5µL HPLC H2O and 1µL of cDNA template for a total volume of 20µL. The reactions were assembled in a 96-well plate and processed on a BioRad DNA Engine Peltier thermal cycler with a built-in Chromo4 Real-Time Fluorescence Detector. Cycles consisted of a 30 second initial denaturation phase at 98°C, a 10 second denaturation phase at 98^{o}C and a 20 second annealing phase based on the T_m of the primers (56°C- 64°C) with the last three steps repeated for a total of 40 cycles. A melting curve was made after cycling by changing the temperature from 65°C to 90°C and read and recorded every 1°C. Final data were analyzed using the OpticonMonitor3 program.

Methylated DNA Immunoprecipitation

The effect of genome editing on mitochondrial cytosine modification was measured by Methylated DNA Immunoprecipitation (MeDIP), using antibodies specific for 5mC and 5OHmC. To obtain methylated and hydroxymethylated DNA samples HCT116 cells were

harvested from 2-80% confluent dishes. DNA extraction of the cells was performed using a DNeasy Blood and Tissue Kit (Qiagen)

The DNA samples were then purified by phenol chloroform extraction and ethanol precipitation. An equal volume of Phenol:Chloroform:IAA (25:24:1) was added to each sample, vortexed and centrifuged at 16,000rpm for 5 minutes at 4^{o}C. The aqueous phase was transferred to a fresh tube and an equal volume of chloroform was added to each sample, vortexed and centrifuged as before. Again, the aqueous phase was transferred to a fresh tube and 1/10 volume of 3M sodium acetate along with 10µL of glycogen was added and mixed. Ethanol was added at 3X volume, samples were mixed and incubated at -80^{o}C for 20 minutes. After incubation samples were centrifuged at 14,000rpm at 4^{o}C for 30 minutes. The supernatant was removed and 500µL of 70% EtOH was added to the pellet and vortexed. Again, the samples were centrifuged at 14,000rpm at 4^{o}C for 30 minutes and the supernatant was removed. The pellets were allowed to air dry for 10 minutes and resuspended in 200µL HPLC H_2O.

The purified DNA was then sonicated using a Diagenode Biorupter water bath sonicator for 2-15 minutes cycles of 30 seconds on and 30 seconds off at the highest setting. 10µL from each round of sonication plus a non-sonicated aliquot were retained from each sample and resolved on a 1% agarose gel; this gel is a check for a 200-700bp smear indicative of proper sonication.

To set up the IP, 4µg of purified and sonicated DNA per antibody was diluted in TE buffer (10mM Tris, 1mM EDTA) to a total volume of 450µL. The samples were then denatured in boiling water for 10 minutes, directly cooled on ice for 10 minutes and 0.5uL of 10X IP buffer (1.4M NaCl, 0.5% Triton X-100, 100mM $NaPO_4$ pH 7.0) was added in addition to 2µg of 5mC

rabbit monoclonal (Active Motif #61255), 5OHmC rabbit monoclonal (Active Motif #39769), or IgG control rabbit monoclonal antibody (Millipore #12-370). The tubes were then incubated overnight at 4°C while rotating. To pre-block the Protein G Sepharose 4 Fast Flow (GE Healthcare), 30µL per IP sample were centrifuged at 4,500 rpm for 5 minutes at 4°C and the ethanol supernatant was removed. 400µL 0.1% PBS-BSA (10mg BSA, 10mL PBS pH 7.4) was added to wash the beads and this wash was repeated three times. After the third wash the pelleted beads were resuspended in an equal volume 1X IP buffer to make a 50% slurry. To block the beads 5µg of BSA and 5µg sonicated Lambda DNA (NEB) were add per 30µL of slurry and rotated at 4°C overnight. At the end of the overnight incubation the beads were washed three times as before and resuspended in 1X IP buffer. To the DNA samples with antibody approximately 30µL of beads were added to the tubes and rotated for 2 hours at 4°C. At the end of the incubation the beads were centrifuged at 4,500rpm for 5 minutes. The supernatant from this last step was saved in fresh tubes as inputs for qPCR; the beads were further washed three times with 1X IP buffer and rotated at 4°C for 10 minutes. After the third wash the pelleted beads were resuspended in 250µL Proteinase K digestion buffer (10mM EDTA, 0.5% SDS 50mM, Tris pH 8.0) and 3.5µL Proteinase K (20mg/mL stock). The tubes were then incubated in a shaking water bath to prevent sedimentation at 50°C for 3 hours. After the incubation the beads were centrifuged at 4°C for 5 minutes and the supernatant that now contains the DNA retained. The DNA samples were then subject to another phase of purification by phenol/chloroform extraction and ethanol precipitation. The exact steps were followed as before except the last step in which the DNA was resuspended in 75µL TE buffer not HPLC H_2O.

End point PCR

End point PCR was executed using HotStarTaq Master Mix (Qiagen). Each reaction consisted of 12.5 µL of HotStarTaq, 1 µL of forward and reverse primers (10 µM) and 9.5 µL of HPLC H_2O as a master mix; 19µL of master mix were dispensed into individual wells per reaction and 1 µL of cDNA was added to each appropriate well. The reactions were run in a Mastercycler (Eppendorf). Reactions were held at 95°C for 5 minutes prior to cycling. Cycles entailed a 30 second melting phase at 95°C, a 30 second annealing phase based on the T_m of primers (56°C-64°C), and a 1 minute extension phase at 72°C for a total of 25 cycles to assess the quality of the IP prior to qPCR. When cycling was complete the reactions were run at 72°C for 5 minutes and held at 4°C as long as needed. Products were resolved on a 1% agarose gel in 1x TAE with EtBr. 10 µL of the product were mixed with 1 µL 10x DNA loading dye and ran at 80V for 30 minutes. The gels were viewed with a FluorChem Sp (Alpha Innotech) imager.

qPCR

qPCR was set up using QuantiNova SYBR Green PCR mix (Qiagen). A master mix was constructed containing 12.5 µL of QuantiNova SYBR Green PCR mix, 1 µL of 10 µM forward and reverse primers each and 9.5 µL of HPLC H_2O per reaction. 19 µL of master mix were pipetted into each appropriate well of a 96 well plate with 1 µL of cDNA added to each appropriate reaction. Reactions were each performed in triplicate. The plate was processed on a BioRad DNA Engine Peltier thermal cycler with a built-in Chromo4 Real-Time Fluorescence Detector. The samples were held at 95°C for 5 minutes prior to cycling. Cycles consisted of a 30 second melting phase at 95°C, 30 second annealing phase based on the T_m of the primers (56°C- 64°C) and a 1 minute extension phase at 72°C for a total of 40 cycles. A melting curve was produced after cycling by changing the temperature from 65°C to 90°C and read and recorded every 1°C. Final data was examined using an OpticonMonitor3 program.

Chapter 3: Results

Targeted Mutation of The Mitochondrial Targeting Sequence of DNMT1 using CRISPR/Cas9 Genome Editing.

Guide RNAs were developed to target close to the published mitochondrial transcription start site of DNMT1 (Fig. 3-1), with the goal of introducing small insertions or deletions that would disrupt the reading frame of the targeting sequence without disrupting nuclear DNMT1 expression. DNA was extracted from individual clones that were genome edited by CRISPR/Cas9 using MyTaq Extract and PCR amplified using DNMT1 primers to initiate the screening procedure for successful editing (Fig. 3-2); expected PCR products of wild type and genome edited HCT116 cells were 630 bp using these DNMT1 primers. All clones illustrated (Figure 3-2) were positive for DNMT1 PCR amplification, however in the 48-well plate that was screened this was not universally the circumstance. Clones that failed to produce a visible band, usually because of poor growth, were discarded. PCR amplicons where then tested for CRISPR/Cas9 induced mutation using the T7 endonuclease assay to detect mismatches between wild-type and putative mutant DNA. Equal amounts of wt and genome edited clone DNA were mixed, denatured and allowed to reanneal. Annealed PCR products incubated in the absence and presence of the T7 endonuclease, and the T7 assay products were resolved on an agarose gel (Fig. 3-3); Predicted 630bp bands were apparent for wild-type and wt/genome edited clones without endonuclease. Samples where apparent digestion by the T7 endonuclease took place produced products at 400 and ~150-200 base pairs. Samples that had these lower bands indicated a mismatch between wt and genome edited sample, these were sent out for sequencing.

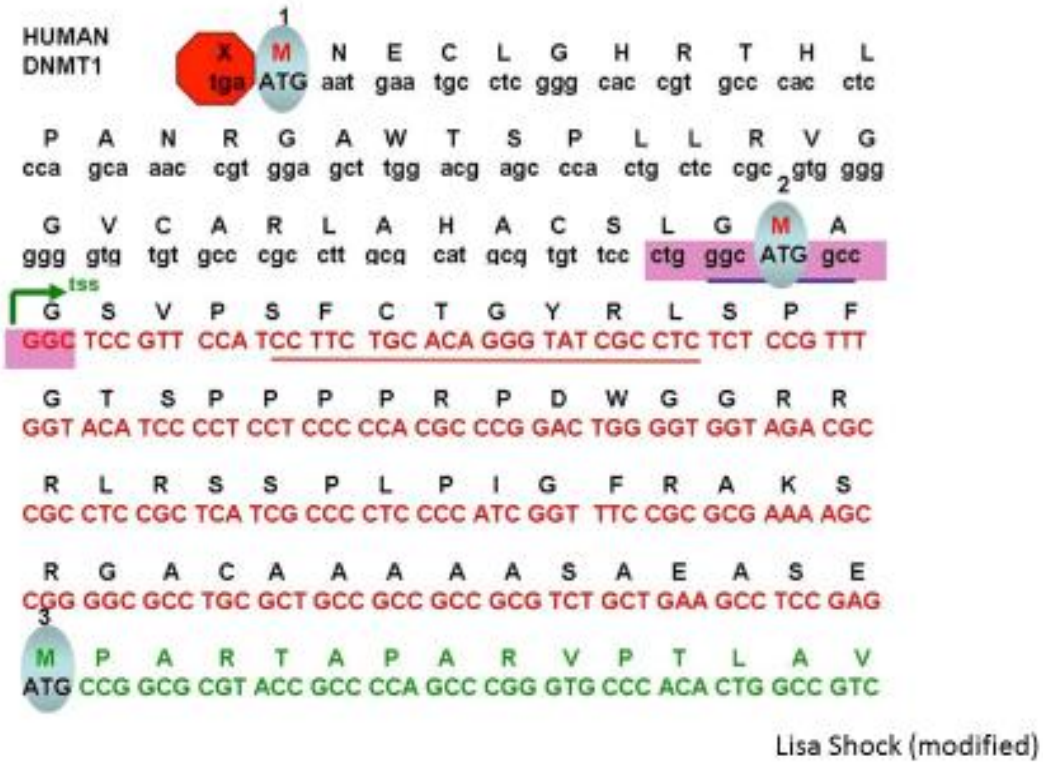


Figure 3-1. Partial human DNMT1 sequence upstream of Exon1 encoding mitochondrial targeting sequence. Upstream of the published nuclear DNMT1 start site (ATG3) there are two additional potential start sites (ATG1, ATG2). Guide RNA (red underline) was designed upstream of ATG3 within the additional 101 in-frame amino acids that encode the MTS (red lettering) to disrupt mtDNMT1 without affecting the nuclear DNMT1 protein (green lettering).

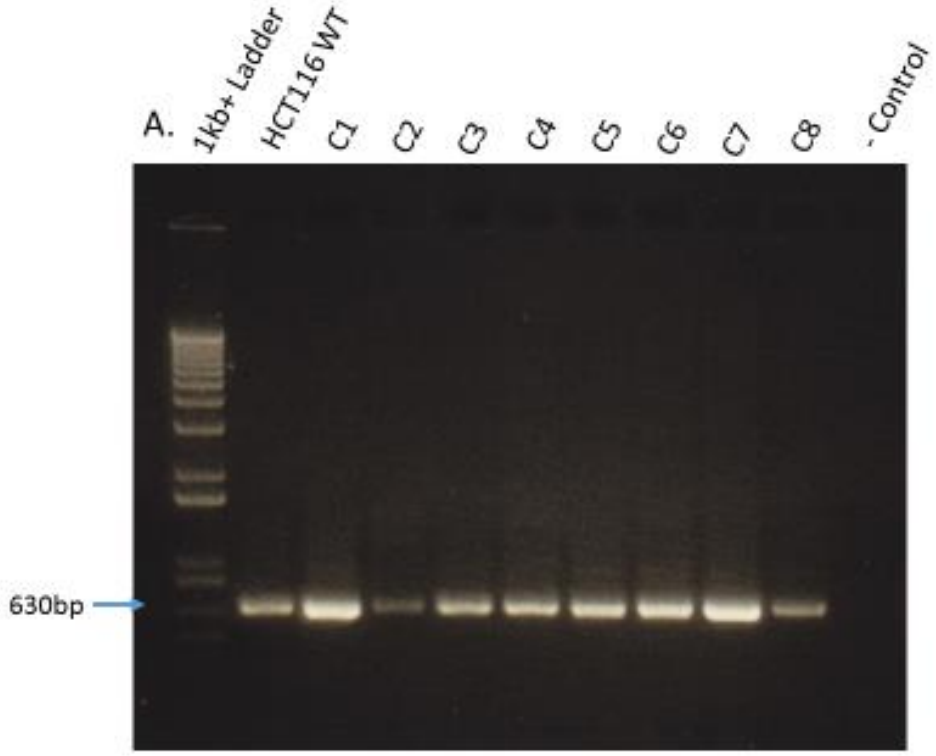


Figure 3-2. Verification of mtDNMT1 DNA amplicon. Agarose gel electrophoresis of PCR assay performed to verify predicted 630bp fragment using DNMT1 primers. Lane 1, 1kb+ Ladder (Invitrogen); 2, positive control amplicon using wt DNA; 3-10, genome edited amplicons; 11, negative (no template) control.

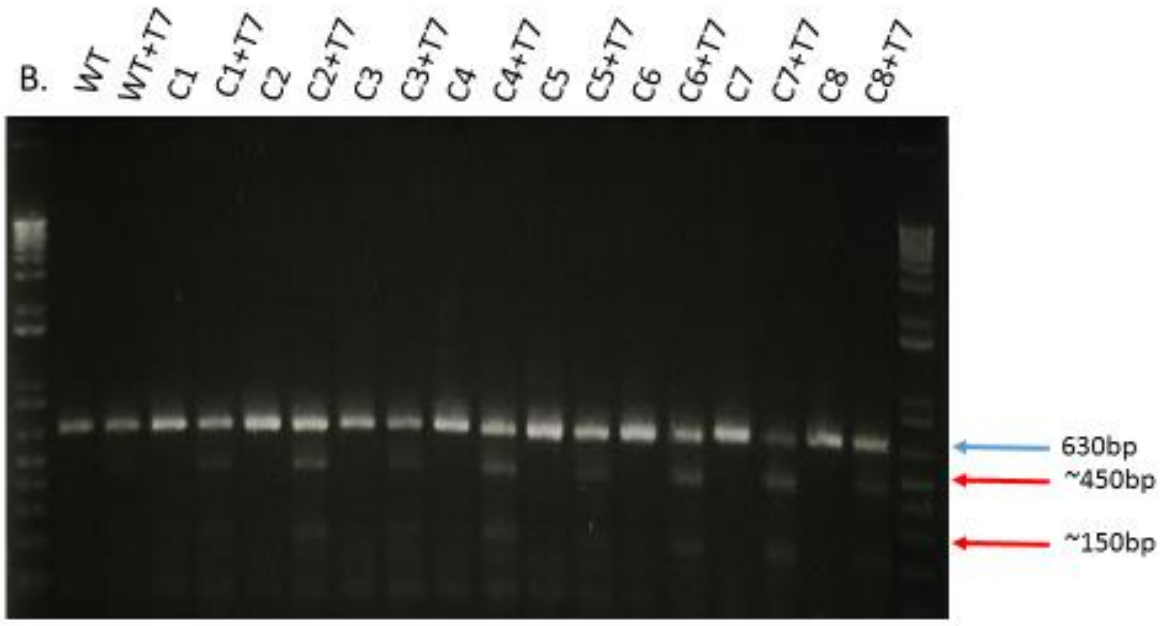


Figure 3-3. Verification of mtDNMT1 mutant construct. Agarose gel electrophoresis of T7 endonuclease assay illustrating DNA amplicons run in parallel with DNA amplicons with T7 enzyme. In samples containing mismatches between wt and mutated DNA, the 630bp predicted DNMT1 fragment (blue arrow) was digested to ~450 and ~150bp (red arrows) with T7 endonuclease fragments. Clones of prospective genome editing were selected by band intensity difference between amplicon and amplicon with T7 enzyme along with digested smaller fragments. Samples C3-C8 were selected for sequencing due to this criteria.

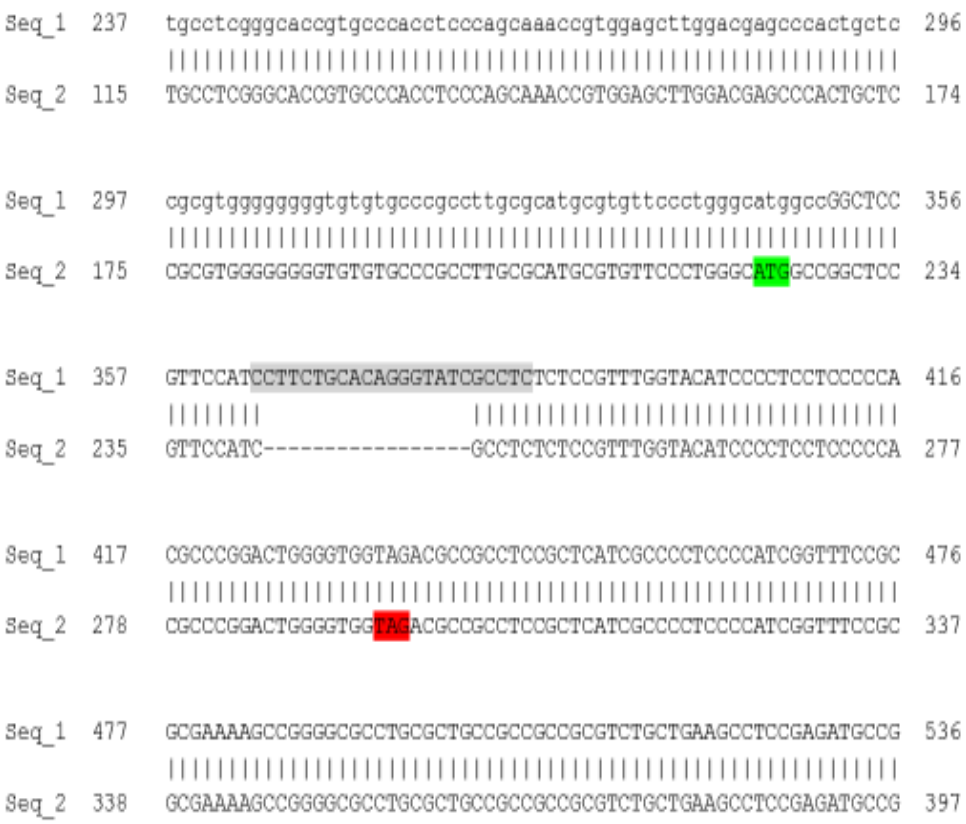


Figure 3-4. Initial sequencing displayed novel stop codon. Sequence results from Eurofins International of DNA amplicons from genome edited HCT116 cell line. Seq_1 is of HCT116 wild-type to contrast Seq_2 which is genome edited. The mutated DNA carries a 17 bp deletion across the Guide RNA (grey highlight), generating a new stop codon (red). ATG 2 (green highlight) to novel stop codon (red) alters and potentially creates an inoperable MTS.

Initial sequence results revealed premature stop codon.

The guide RNAs' specific targeting allowed for the Cas9 endonuclease to create a cleavage site shortly after the transcriptional start site of mtDNMT1, thus leaving the nuclear start site for DNMT1 unaffected. Cellular error prone repair mechanisms (NHEJ) frequently results in small insertions or deletions around the cleavage site. In the example shown in Fig. 3-4, a 17 bp deletion generates a premature stop codon upstream of the nuclear start ATG, resulting in a truncated protein from ATG1 or 2, and presumably a nonfunctioning MTS.

Immunoblot analysis of DNMT1 revealed residual mitochondrial presence of genome edited DNMT1.

Immunoblotting was used to analyze levels of DNMT1 in whole-cell, cytosolic and mitochondrial fractions from HCT116 wildtype and two different HCT genome edited samples. Stringent controls were in place to ensure that cell fractionation was complete, with absence of contamination from other fractions. An antibody for voltage-dependent anion channel (VDAC) located on the outer membrane of the mitochondria was used to certify equal loading. Anti-Histone H3 (tri methyl K4) (H3K4me3) antibody should only detect this nuclear protein in the whole-cell lysate fraction; presence in any other fraction is indicative of nuclear contamination. Tubulin is a globular protein that polymerizes into microtubules; it is absent from the mitochondria and its expression should only be found in whole-cell lysate and cytosolic fractions with its antibody. To note: the tubulin antibody even at exceptionally low concentrations is still an extremely potent antibody and trivial amounts of detection in the mitochondrial fraction should be carefully considered in combination with other controls as a suitably uncontaminated fraction.

Over the course of more than a few cell fraction experiments images on immunoblots would fluctuate due to several possible causes. Each cell fractionation experiment improved upon the last by adjusting technique to better control contamination. Along with technique adjustment there would also be an adjustment in antibody concentration during blotting if needed. Although controls started to improve throughout the runs, the common trend that could be seen on the blot was that genome edited DNMT1 was consistently present in all lanes including the mitochondrial fraction (Fig. 3-5) where the absence of DNMT1 was critical for validation. The near absence of H3K4me3 in the mitochondrial fraction with the expression of DNMT1 in all three lanes initiated a reinvestigation and revalidation of sequences for the genome edited cell lines.

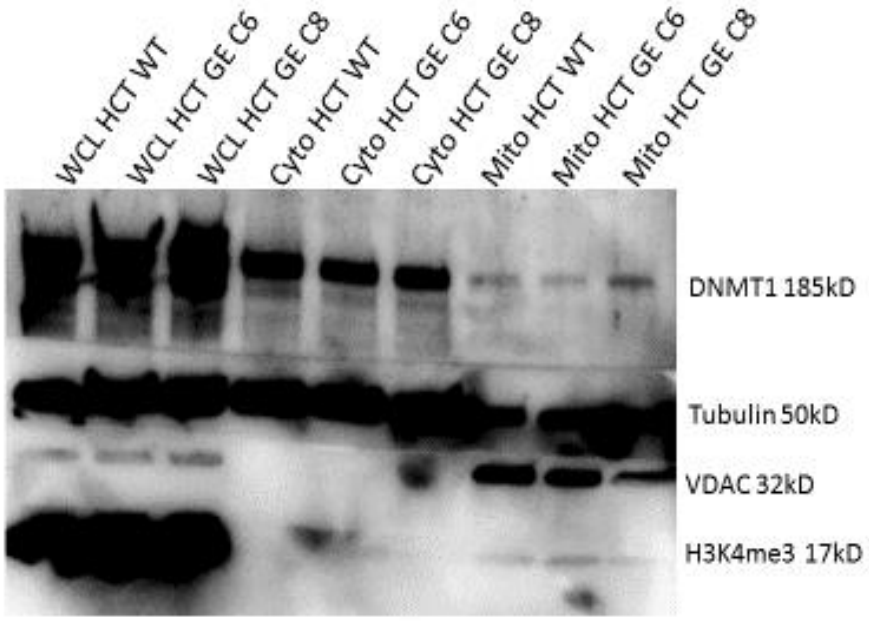


Figure 3-5. Detection of DNMT1 in all sample lanes. Immunoblot investigating expression of DNMT1 found at 185kD. Cellular fractionation provided whole-cell lysate, cytosolic and mitochondrial fractions to be probed. VDAC is a loading control found only in whole-cell and mitochondrial fractions. Tubulin is a nuclear and cytosolic contamination control that should not be present in the mitochondria. H3K4me3 is a nuclear contamination control that should only be found in whole-cell lysate. The presence of DNMT1 in mitochondrial lysates from the two genome edited cell lines suggests that residual targeting activity remains despite successful genomic editing.

Reinvestigation of DNMT1 genome edited sequences uncovers a novel frameshift.

Primary investigation of possible CRISPR/Cas9 indels for HCT116 was performed by selecting sequences that had insertions or deletions in multiples of one or two base pairs, in order to disrupt the reading frame between ATG2 (mitochondrial start) and ATG3(nuclear start). Immunoblot analysis provided evidence that DNMT1 is still able to translocate into mitochondria, suggesting that either DNMT1 is able to translocate without a targeting sequence, as is apparently true for many nuclear encoded mitochondrial proteins, or that residual targeting sequence remains in the engineered locus. Translating the sequencing results shown in Fig. 3-4 does indeed show that from ATG2 and after the 17bp deletion, a novel stop codon does arise; however, translation in the reverse direction from ATG3 reveals inframe coding sequence past ATG2 to a new ATG that is in frame with ATG3. This sequence, while different from that of the original MTS of DNMT1, is evidently able to produce a partially functional MTS. Mitoprot II Software (calculates mitochondrial targeting probability) calculated that the two genome edited samples (C6 and C8: deletions in multiples of two) selected for experimentation had a 27% chance for one and a 70% chance for the other to still be targeted to the mitochondria.

Rescreening and selection of genome edited cell line produces authentic novel stop codon.

Genome edited cell lines were rescreened by previously mentioned methods and sequenced. Sequencing (Fig. 3-6) (19bp del) revealed several clones that had a deletion of n amino acids plus one base pair for both HCT genome edited and 3BKO genome edited cell lines. These deletions produced novel stop codons in the MTS region. Two clones from each cell line were selected; sequences were translated upstream from ATG3 and analyzed as above for the presence of an engineered new upstream ATG. These sequences were analyzed using Mitoprot II; each cell line resulted in a less than 2-3% chance of possessing a functional MTS and no new ATG with their edited sequences (Table 3-1).

Figure 3-6. Sequence illustration of wild-type and genome edited cell lines. Sequences are represented as: intron 1 (blue lower case) with ATG1 and ATG 2(blue oval), new ATG (yellow), mitochondrial targeting sequence (upper case red), ATG3 (green oval), Exon 1 (upper case green), gRNA (underlined red), novel stop (red oval). WT (wild-type) illustrating two ATGs inframe with nuclear ATG3 with location of gRNA. 17bp del illustrates altered sequence with ATG1,2 inframe with novel stop; new ATG now inframe with ATG3. 19bp del illustrates altered sequence with ATG1,2 inframe with novel stop; no ATG inframe with ATG3, presumably non-functional MTS.

Table 3-1: Specification of clone mutation, outcome and probability of processing a functional mitochondrial targeting sequence.

Clone	Specific Mutation	Mutation Outcome	POMTS
HCT WT	N/A	N/A	0.991
HCT C6	17bp deletion	ATG1 & 2 out of frame new ATG2	0.2035
HCT C8	2 bp deletion	ATG1 & 2 out of frame new ATG2	0.7089
HCT B4	10bp deletion	ATG1 & 2 out of frame	0.0233
HCT B7	19bp deletion	ATG1 & 2 out of frame	0.0204
3b A7	19bp deletion	ATG1 & 2 out of frame	0.0312
3b B3	31bp deletion	ATG1 & 2 out of frame	0.0279

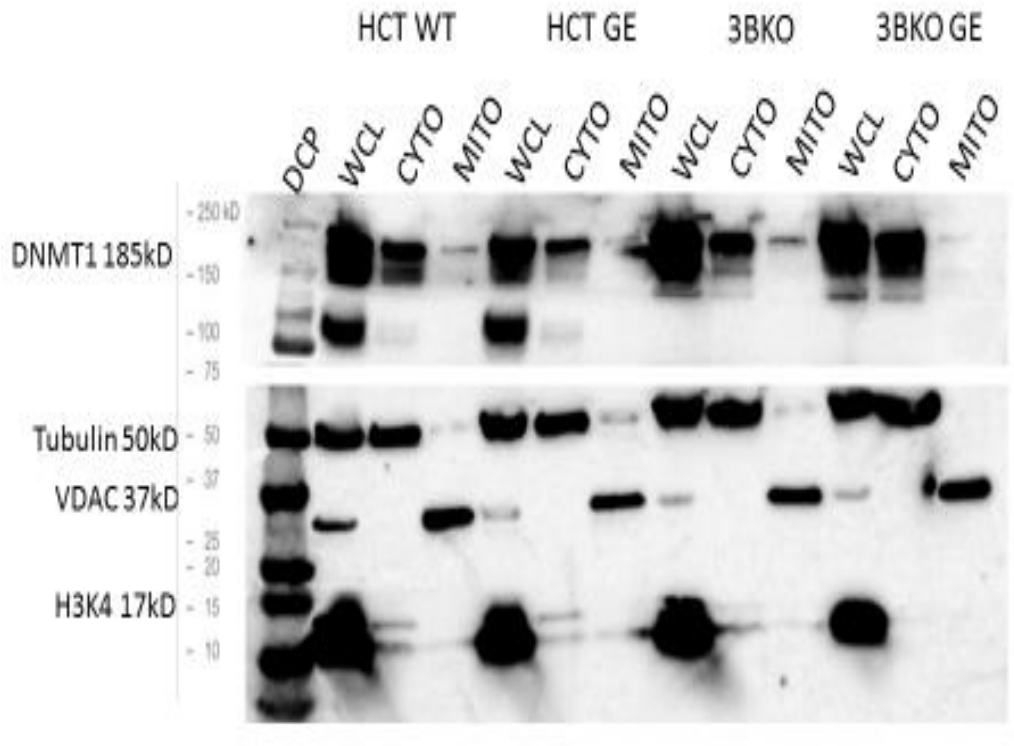


Figure 3-7. Immunoblot analysis for DNMT1 expression of select genome edited cell lines. Lane 1, Dual Color Plus Ladder: lanes 2-4, HCT116 wild-type: lanes 5-7, HCT genome edited clone A7: lanes 8-10, 3BKO: lanes 11-13, 3BKO genome edited clone B3. Expression for DNMT1 at 185kD in all lanes although at a reduced level in fractions that have had the MTS deleted. WCL (whole cell lysate), CYTO (cytosolic fraction) and MITO (mitochondrial fraction).

Immunoblot comparative analysis of DNMT1 revealed a reduction in expression levels.

Immunoblotting was performed on all selectively screened cell lines as well as HCT116 wild-type to determine DNMT1 expression (Fig. 3-7). Previous studies with an analogous DNMT1 antibody have shown DNMT1 expression in HCT116 cells. These current results suggest the same expression of mtDNMT1 in HCT116 wild-type and 3BKO; however, in the genome edited cell lines there is not a complete loss of expression as predicted. In the HCT genome edited and the 3BKO genome edited cell lines the expression of DNMT1 has been significantly reduced.

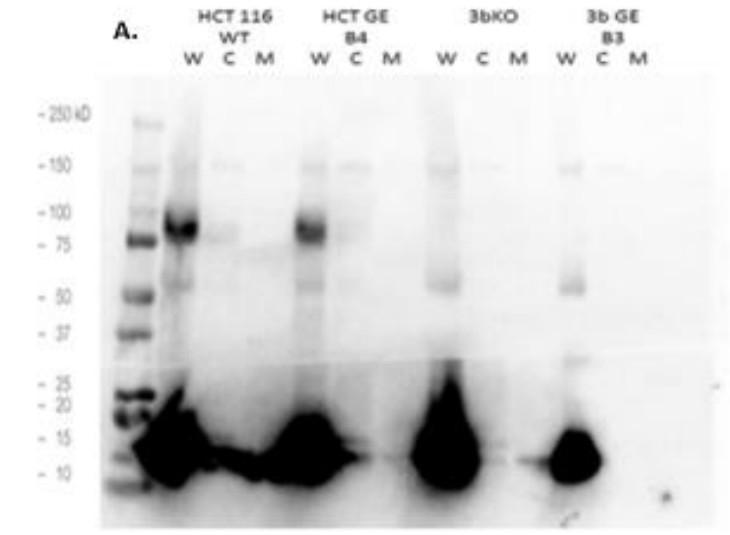
Immunoblot analysis for the expression of DNMT3B may be more complicated than expected.

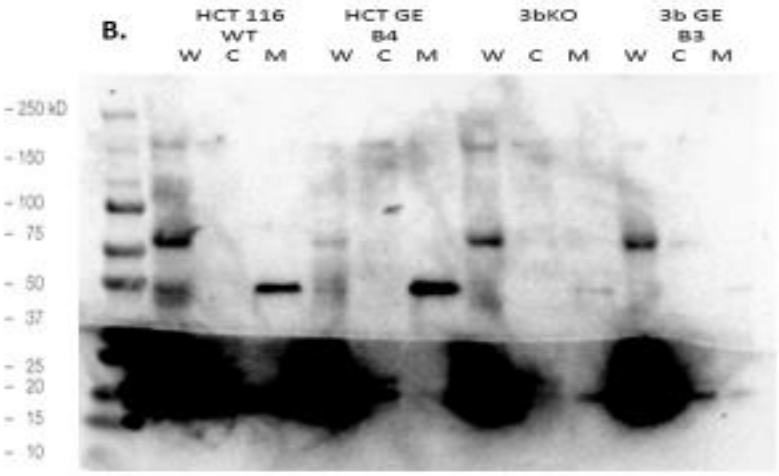
Immunoblot analysis for DNMT3B was performed using three different monoclonal antibodies (Fig. 3-8). The antibodies bound in similar regions so expression was predicted to be comparable. Western blots provided on manufacturer's website (Santa Cruz) place DNMT3B protein anywhere from 80kD-130kD (97.5kD predicted) with identical antibodies, however there is no report of DNMT3b in the mitochondrial fraction for this cell line. Immunoblot analysis of the tested antibodies was variable and inconsistent. Between the three blots, DNMT3B immunoreactivity indicates heterogeneity of size, and further analysis with other methods will be necessary to make a definite account of DNMT3B. For each antibody, we expected to see bands in HCT 116 wt and genome edited cell lines (WCL, cytosol and mitochondrial lysate) that were absent from 3BKO and its genome edited derivatives. For anti-DNMT3B-F2 antibody (Fig 8-3A), WCL bands at >80kDa are absent from 3BKO lysates, but there is no definite band in the mitochondrial fraction for any of these cell line. For antibody D1, 50kDa bands in mitochondrial

extracts of HCT116 wt and derivatives are absent from 3BKO and its derivatives, but no convincing and consistent bands are apparent in WCL (Fig 8-3B). Bands produced by antibody G1 are present in all 4 lines (Fig 8-3C). Interestingly this antibody (Fig. 3-8 C) produced bands in mitochondrial lysates at 37kD and ~60kD, which might represent a cleaved DNMT3B product. Again, these bands are present across all 4 lines. These data raised concerns that the 3BKO cell line obtained from Dr. Vogelstein at Johns Hopkins might have been contaminated with a wt line

Genotyping of 3BKO cell line validates original knock-out.

The genotype of the 3BKO cell line can be identified by wild-type specific (R2) and mutation specific primers (R21) (IDT) (illustration in Fig. 3-9A) for PCR (performed identical as formerly mentioned PCR methods). PCR agarose gel images (Fig. 3-9B) in HCT116 wild-type do not produce a product for the *in silico* knockout sequence but do demonstrate a product for the reverse primers in exon2 indicating an unmodified wildtype cell line with the absence of a knock-out. In the same gel image the 3BKO cell line does not produce a product for reverse primers in exon2 (R2) but do demonstrate a product for the reverse primers in exon21 (R21) that create an in silico sequence indicating that there is an apparent knock-out.


A.
HCT 116
WT
HCT GE
B4
3bKO
3b GE
B3
W C M W C M W C M W C M
- 250 kD
- 150
- 100
- 75
- 50
- 37
- 25
- 20
- 15
- 10


B.
HCT 116
WT
HCT GE
B4
3bKO
3b GE
B3
W C M W C M W C M W C M
- 250 kD
- 150
- 100
- 75
- 50
- 37
- 25
- 20
- 15
- 10

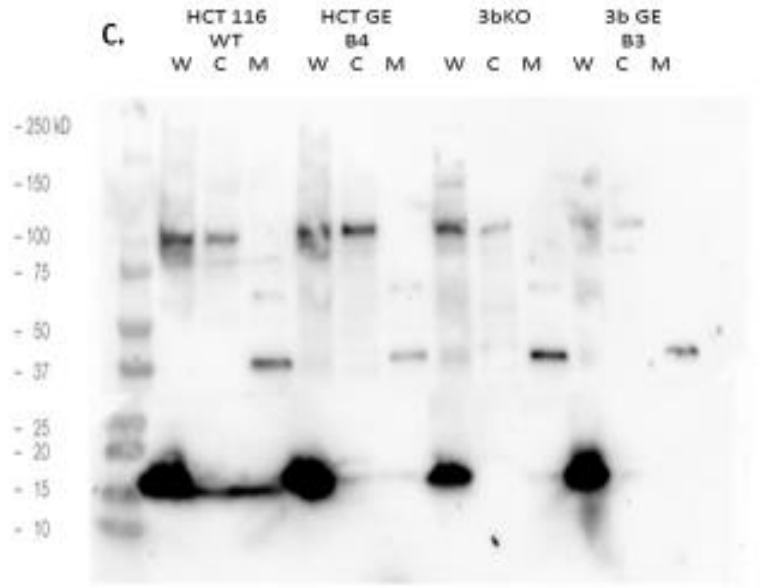


Figure 3-8. Immunoblot analysis of DNMT3B. Lane 1, Dual Color Plus Ladder: lanes 2-4, HCT116 wild-type: lanes 5-7, HCT genome edited clone B4: lanes 8-10, 3BKO: lanes 11-13, 3BKO genome edited clone B3. Protein lysates for all blots were from same cellular fractionation preparation and all protein amounts are consistent. H3K4me3 was run at the bottom of each blot for contamination control. Blots A-C are DNMT3B F2, D1 and G9 antibodies (suppliers?) respectively.

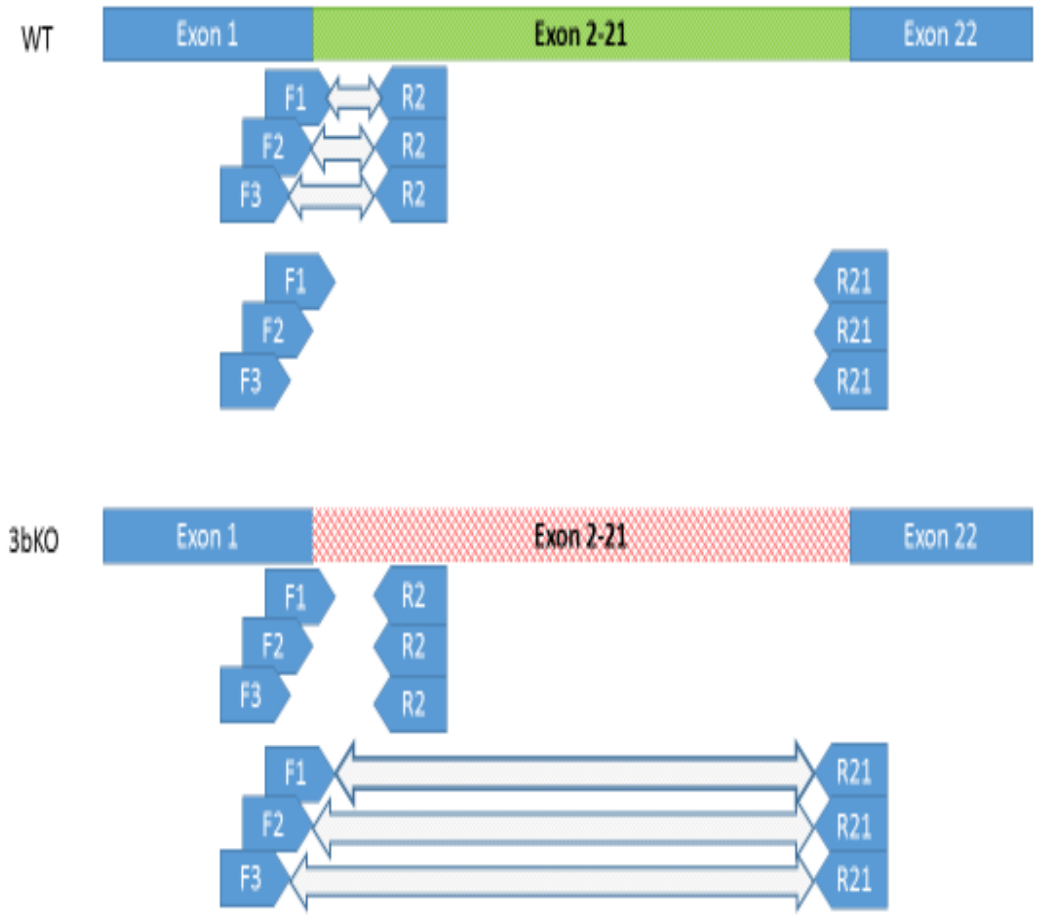


Figure 3-9A: Diagram of expected product of wild-type and 3BKO from PCR. WT illustrates the expected fragments (double-sided arrow) of a wild-type cell line. The reverse primers (R2) that lay in exon 2 will be able to make an estimated ~800-1000bp fragment with forward primers (F1,2,3) in exon 1; since exons 2-21 are present the fragment will be too large to produce for reverse primers (R21) after exon 21 to meet up with forward primers out of exon 1. 3BKO illustrates that when exons 2-21 are knocked out (red hatching) that only reverse 21 primers with in silico sequence will be able to produce a fragment (estimated 2-3kb) with forward primers with in silico sequence.

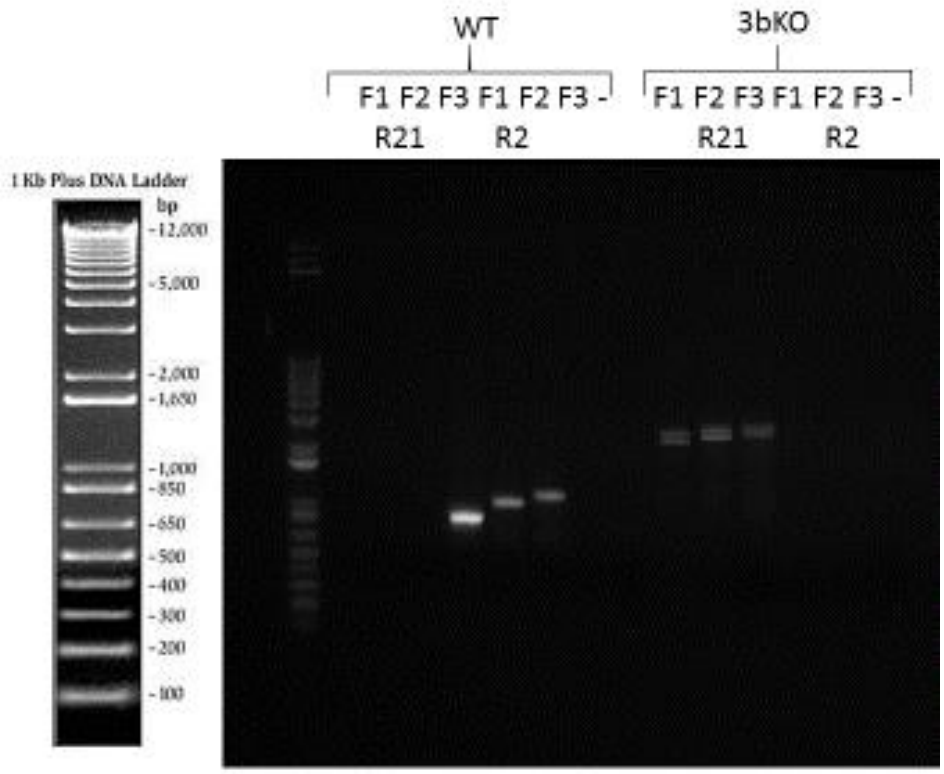


Figure 3-9B. DNMT3B genotyping PCR experiment. Lane 1, 1kb+ ladder: lanes 2-7, HCT wild-type samples with indicated forward and reverse primers: lane 8, negative control: lanes 10-15, 3BKO samples with indicated forward and reverse primers. Visible bands in HCT wild-type and 3BKO are indicative of a positive DNMT3B knock-out.

RTqPCR shows decreased transcription among genome edited cell lines.

RT-qPCR was carried out on strand-specific cDNA of HCT116 wild-type and genome edited cell lines. Primers for 16s RNA, ND6, Cox1 and ND1 were chosen to assess the effect of altering mtDNMT1 levels on transcription in the genome edited cell lines. Using total cDNA synthesized by reverse transcription with random hexamers, the expression levels were normalized to the geometric mean of values for beta-Actin and GAPDH.

Genome editing in the HCT116 wt background indicates modest changes in transcription. We observed a minimal reduction in transcription of 16S rRNA (heavy strand) from wild-type in the genome edited lines (Fig 3-10). ND1 and Cox1, two protein coding heavy strand genes, and ND6, the only protein coding region on the light strand, each show a 20-40% reduction in the HCT genome edited lines. In contrast, deletion of DNMT3B reduces transcription in the protein coding regions analyzed by 2-fold; removal of the DNMT1 MTS did not significantly change this reduced transcription.

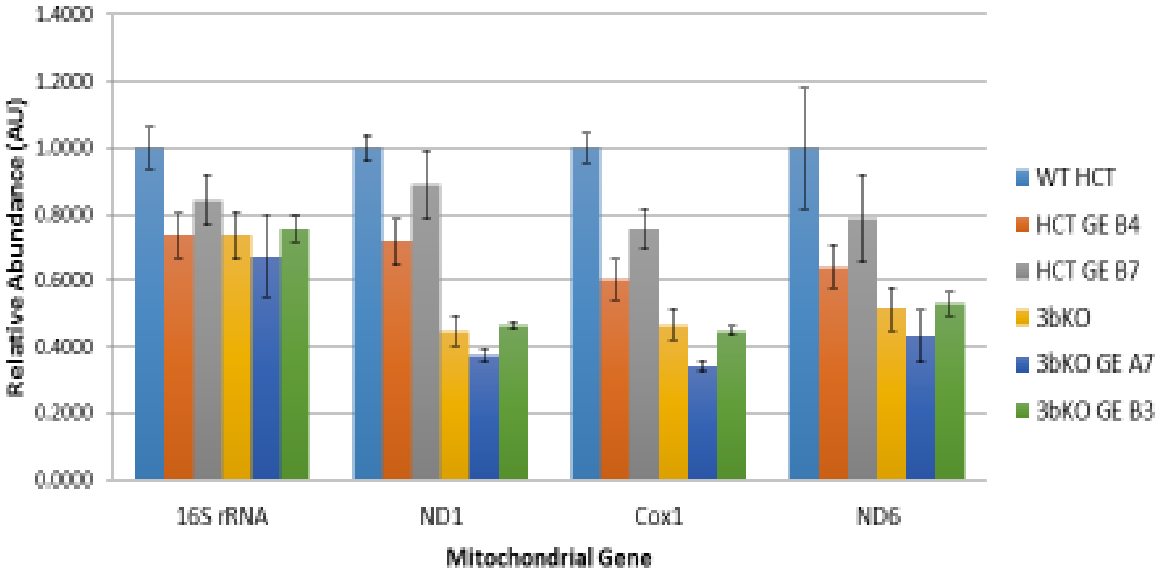


Figure 3-10. Mitochondrial Transcription of genome edited cell lines. Bar graph characterizes real time RT-PCR analysis of mitochondrial genes in HCT116 and 3BKO cell lines. The y-axis signifies normalized mRNA expression to the geometric mean of values for β-Actin and GAPDH.

Quantification of 5-methylcytosine in genome edited cell lines by means of MeDIP.

Since expression of DNMT1 was detected in mitochondrial fractions of the genome edited cell lines at varying levels, and since DNMT3B knockout had a profound effect on mitochondrial transcription, it was important to determine the levels of 5mC in the cell lines to determine whether 5mC levels might be related to the levels of transcription. DNA was purified, sheared to a length of 200-500bp and checked on an agarose gel according to protocol (Fig. 3-11). Challenges occurred when purifying DNA and trying to retain enough to proceed through the IP; after several trouble-shooting experiments, it was concluded that the best way to purify DNA with phenol-chloroform was to use a freshly prepared mixture. The phenol-chloroform must have the right acidity for the DNA to precipitate into the organic phase while the less neutralized RNA remains in the aqueous phase. As well, an RNase treatment step was not original to the protocol; however, after low levels of DNA were observed on analytical gels after sonication, it was thought that RNA might be compromising nanodrop readings. After the RNase treatment step was added the 260/230 ratios significantly improved on the nanodrop, indicating previous unwanted RNA was altering projected DNA quantities.

After MeDIP and a final round of DNA purification, an endpoint PCR was performed to visualize how wild-type and genome edited samples differed by band intensity in regards to specific antibody on an agarose gel. The endpoint PCR was also utilized to see if levels of 5mc could be roughly quantified by comparing band intensity to a non-specific IgG control. Several commercial preparations of IgG yielded a very high background, in many cases higher than the signal obtained with 5mC, and we performed tests on multiple IgG preparations from different suppliers.

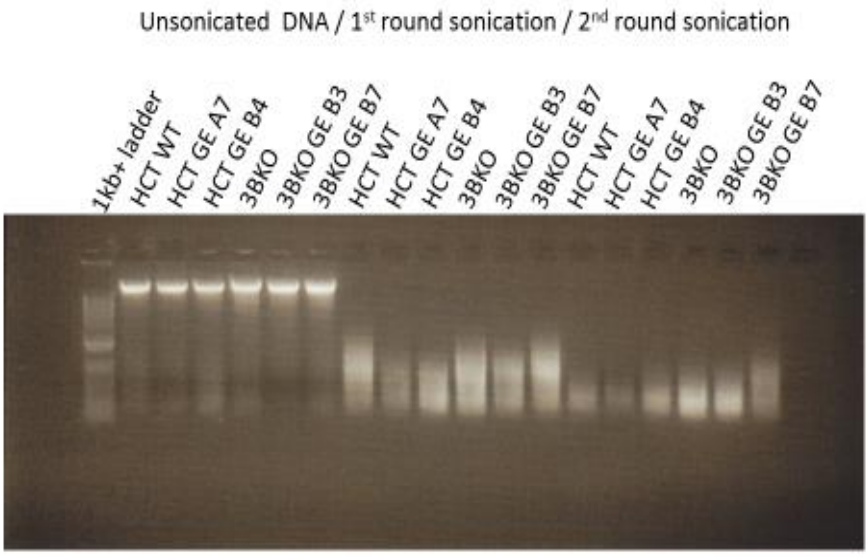


Figure 3-11. Sonication of HCT wild-type and genome edited cell lines. Lane 1: 1Kb Plus DNA Ladder, lanes 2-7: unsonicated samples, lanes 8-13: 15 minutes of 30 seconds on 30 seconds off high power sonication, lanes 14-19: second round of identical sonication. 200-800 base pair smear in last six lanes indicate proper shearing.

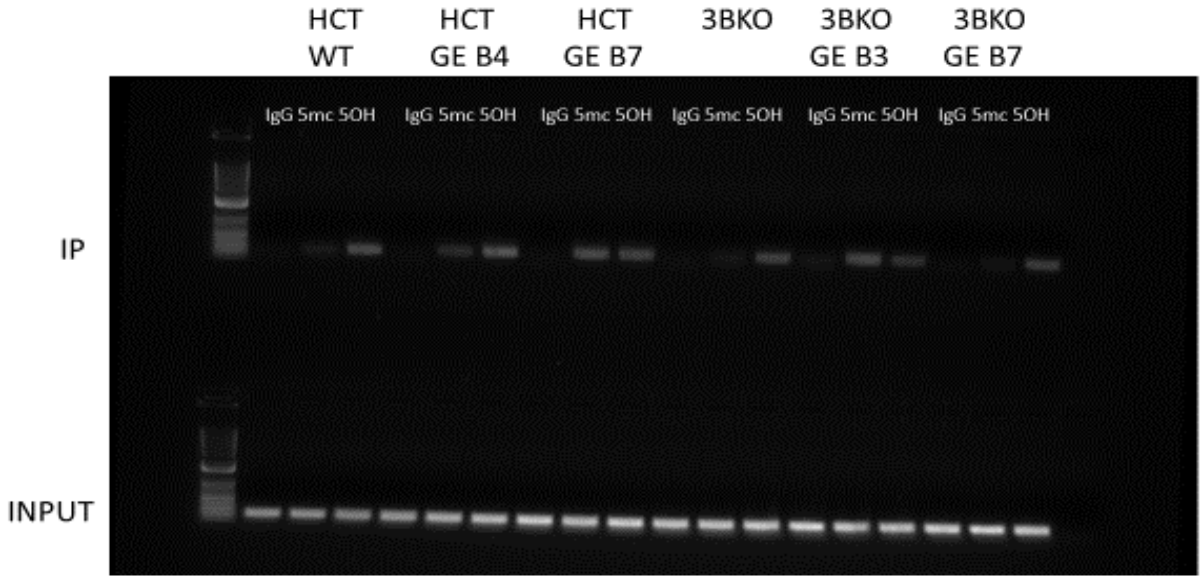


Figure 3-12. Endpoint PCR of HCT116 wild-type and genome edited cell lines. Top segment of the agarose gel is that of the immunoprecipitated samples; each cell line is denoted at the top of figure and lanes of cell line are lined up by IgG, 5mc and 5hmc antibodies respectively. Bottom segment of the gel is the concurrent input samples; even band intensity in this portion indicates and equal amount of DNA went into the IP.

The IgG from Millipore gave a baseline level appropriate to move forward to qPCR seen by a higher band intensity in all IP lanes over IgG (Fig. 3-12).

Quantitative PCR was performed on all DNA samples from the MeDIP using mitochondrial primers for HSP1, LSP and 16s rRNA, and the outcome of the assay was formulated in excel and graphed. The first set of primers evaluated detected HSP1 (Fig 3-13); in comparison to wt HCT116 cells, the levels of 5mC in mtDNA decreased ~25% in HCT-GE cells, 60% in 3BKO cells and 75% in3BKO-GE cells. This data is in accord with the transcription data shown in Fig 3-10, supporting our hypothesis that reduced 5mC results in increased mitochondrial transcription. Subsequent qPCR analysis of LSP and 16s regions, however, showed a different pattern; both LSP and 16s amplicons demonstrated an increase This data contradicts all other evidence from the Taylor lab and from other published works which all point to a role for both DNMT1 and DNMT3B in mtDNA methylation, and a controlling role in mtDNA transcription from both HSP and LSP. Furthermore, the qPCR data did not agree with the endpoint PCR. A closer look at the raw data from the qPCR reveled inconsistencies in the readings for input samples. Since 4ug DNA are used for each IP, and the amount of DNA precipitated is very small, we use the IP supernatants as a measure of input for each sample. We expect each input sample qPCR reading to be very similar, and from prior experimentation find this to be true. However, in the experiment shown in Fig 3-13, input samples were inconsistent from experiment to experiment and varied by as much as 5-fold in some cases. A fault in the protocol used for the process of sonication was brought to our attention after these findings; total purified DNA was sonicated and then the concentration was adjusted, when in fact varying concentrations of DNA as well as total volume going into sonication affects the outcome of shearing.

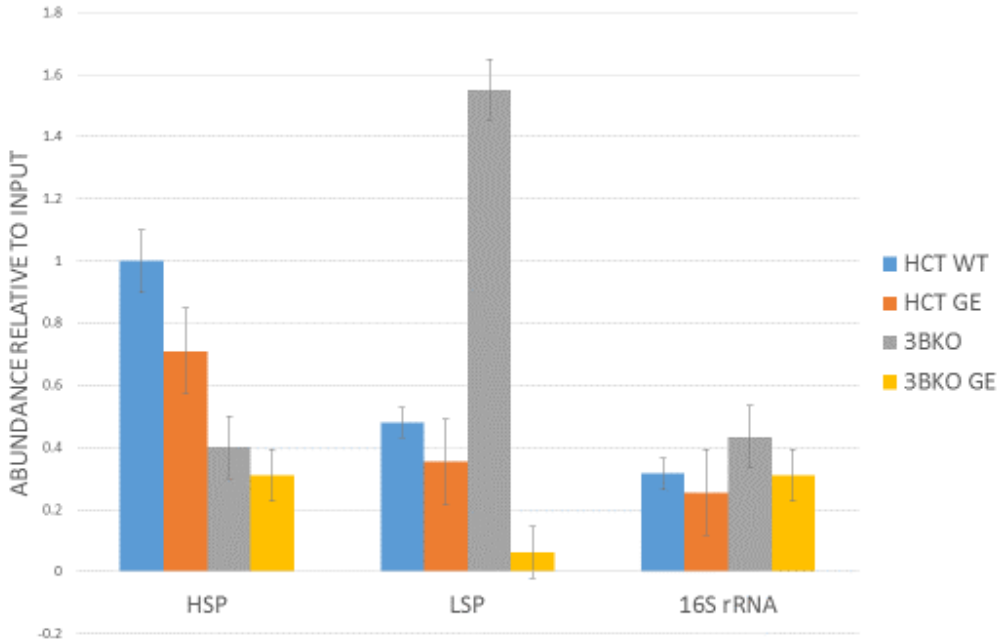


Figure 3-13: Quantification of 5mc from MeDIP. Immunoprecipitated and input DNA was examined by qPCR with mitochondrial primers for HSP, LSP and 16S rRNA. Levels of 5-methyl cytosine are depicted relative to input for each cell line. Data represents average of triplicate readings ± SE. HCT GE clone B4, 3BKO GE clone B3.

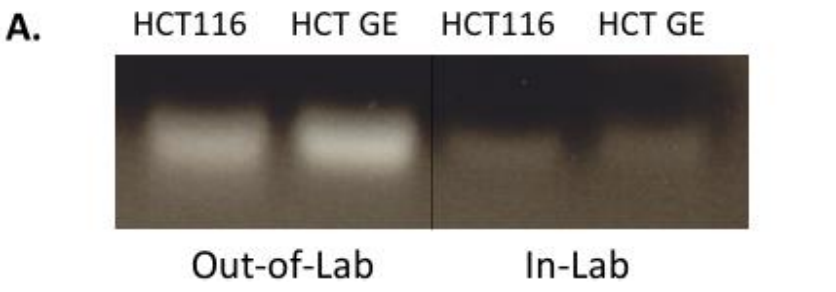
As a result, we began a new IP experiment in which we adjusted the concentration of DNA to equivalent totals prior to sonication. All steps after sonication were taken according to protocol leading into qPCR; again, results for this assay were highly variable and a reason for the variability in outcome was provisionally inaccessible. The importance of this assay in determining the role of methylation in the mitochondria resulted in a large amount of effort in troubleshooting the source of variability in this data; efforts were made in procurement of fresh buffers and reagents, constructing all new primers for qPCR, quantitating and adjusting sonicated DNA prior to IP, performing qPCR assay on different equipment, using several different amplification reagents in qPCR, obtaining new antibodies for IP and attempting to formulate results independently from IgG. These efforts resulted in the same unusable results.

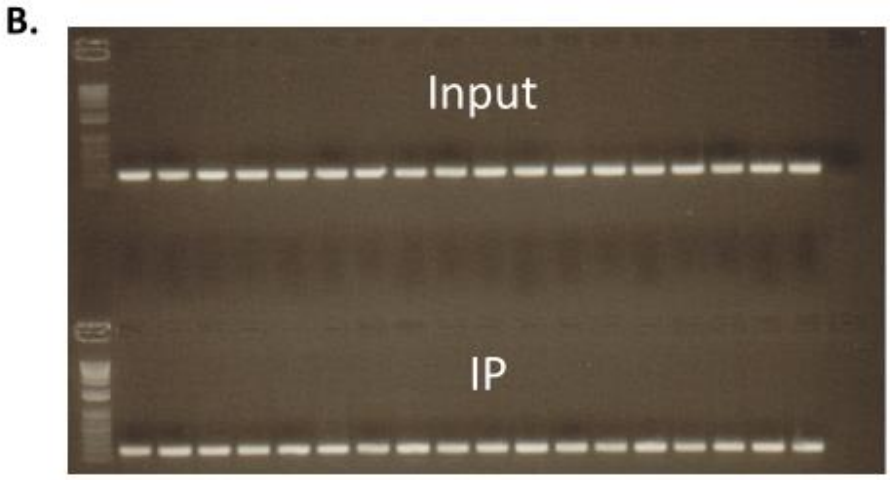
At this point, a possible malfunction in the sonicator was brought to our attention by a colleague. Finally, we determined that the faulty sonicator was shearing the DNA into exceedingly minute fragments and this outcome appeared to vary from sample to sample. These small fragments would be unable to act as template in qPCR reactions. Results of a sonication assay (Fig. 3-14A) where HCT116 wild-type and HCT genome edited samples were of equal concentrations and volume and processed through two separate sonication apparatuses (one out-of-lab new instrument and the one in-lab old instrument) revealed significant differences. Following the original sonication protocol we used two 15 minutes rounds of sonication on high with 30 seconds on and 30 seconds off. Samples were withdrawn for analysis by agarose gel electrophoresis after each round of sonication. For the out-of-lab, new sonicator, both sonications produced bright, broad bands in the gel, ranging in size from 300-600bp with only slight visible smearing, consistent with optimum shearing. However, the older in-lab sonicator

showed a significant decrease in band intensity, even after a single round of sonication, suggestive of excessive DNA degradation. This decrease in band intensity can plausibly be explained by supposing that the in-lab sonicator is processing the samples at such a great force that the fragments it is producing are so minute they are running right through the agarose gel.

Coincidently, results from another MeDIP came back irregular in an endpoint PCR. The usual protocol requires 35 cycles in the Thermocycler, however, bands of equal intensity both in the IP and Inputs were observed in the gel image (Fig 3-14B). The virtually equal intensity in the IP lanes for the gel was something that has not been observed before and needed to be investigated; the assay was run again at 25 cycles, in an attempt to get more semi-quantitative data. Upon reducing the cycle number the gel image for the assay (Fig 3-14C) revealed unequal amounts in the Input lane. These variances in Input are unacceptable to proceed to qPCR and possibly coincides with the sonication malfunction. Reinvestigation of the gel image for the sonication check for the corresponding MeDIP reveals what appears to be over shearing as well (Fig. 3-14D).

As a result of these extensive quality checks, and the clear malfunction of our sonicator, which is used by all investigators in the Cancer Center, a new bath sonicator has been ordered for Massey Cancer Center. Once it is in place, we plan to carry out the necessary replicate assays to allow publication of this data set.


A.
HCT116
HCT GE
HCT116
HCT GE
Out-of-Lab
In-Lab


B.
Input
IP

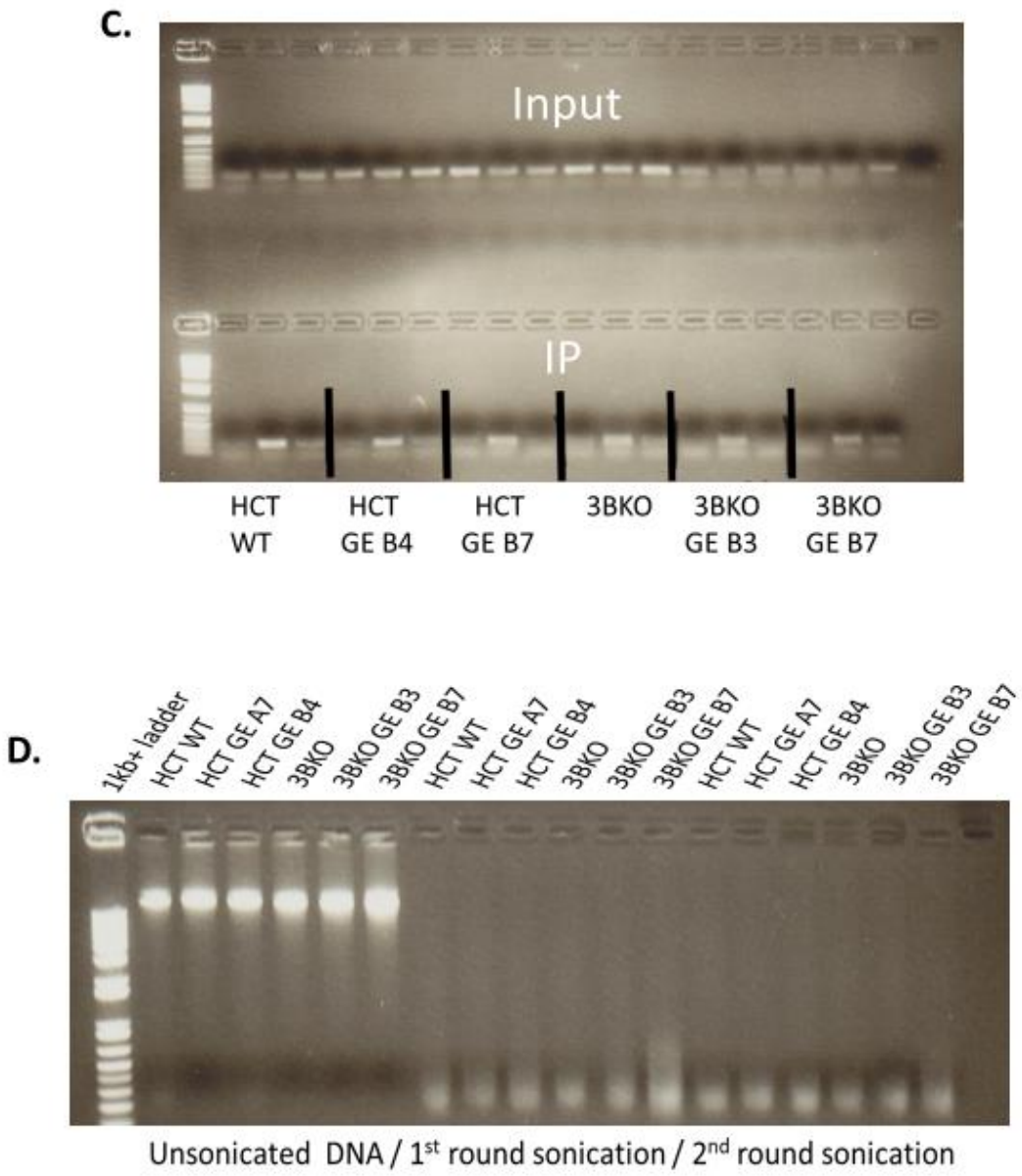


Figure 3-14 A-D: Evidence for inconsistent qPCR results. A: Sonication assay reveals in-lab sonicator has sheared samples to such a small size that a 1% agarose gel cannot contain. First round sonication usually has smear of 500-1000bp, while second round sonication usually has smear of 200-700bp but not below. B: Endpoint PCR (35 cycles) with all input and IP samples from MeDIP on agarose gel. IP samples showing virtually equal band intensities, inconsistent with previous assays requiring further exploration. C: Endpoint PCR (25 cycles) with Input and IP samples; more sensitive assay reveals uneven loading of DNA going into the MeDIP (input). D: Sonication check of concurrent MeDIP reveals overlooked improper shearing of DNA.

Chapter 4: Discussion

The mitochondrial targeting sequence for DNMT1 was discovered by the Taylor lab, and understanding the functional significance of the cytosine modifications that result necessitated disturbing the ability of the proteins responsible from entering the mitochondria. Preventing DNMT1 from being translocated to the mitochondria without affecting the nuclear protein was also a novel and ambitious undertaking. Both HCT116 and 3BKO cell lines were able to have the mitochondrial targeting sequence of DNMT1 altered using CRISPR/Cas9 genome editing. Sequence authentication displays that the deletions created with the endonuclease generate a premature stop codon that disrupts the targeting peptide, further substantiated by mitochondrial targeting software. Efforts to test and create these clones were substantially drawn out because of unconventional and multiple ATP start sites. Within the case of multiple start sites it is important to not only examine for new downstream stop codons after the novel deletion but to also make sure that said deletion does not generate in-frame and functional new upstream start sites that were originally out of frame. These efforts are critical for a complete understanding of the potential outcomes when undertaking genome editing of unique DNA regions.

The immunoblots of this study build upon previous lab work that suggest again that DNMT1 translocates to the mitochondria [22]. The expression levels of DNMT1 in the genome edited cell lines appears to be half of that in the lines that have complete MTSs (Fig. 3-7). This finding suggests that the MTS does carry DNMT1 to the mitochondria because it is in lanes of all samples. However, these data suggest that the MTS it is not the sole contributor to DNMT1 translocation. The expression levels of DNMT1 in the genome edited cell lines suggest that making the MTS nonfunctional will affect the amount of DNMT1 going into the mitochondria

but does not yet answer any questions about the functional significance of the cytosine modifications resulting from mtDNMT1.

Previous studies have shown that the MTS of DNMT1 possesses the capability of taking GFP to the mitochondria [22] and that when the MTS peptide comes in contact with the outer membrane of the mitochondria it is recognized and validated for import; that being said it is of importance to understand that there is always a possibility that DNMT1 uses alternative methods to translocate to the mitochondria. There are several proteins that translocate to the mitochondria without any obvious mitochondrial import signal. These proteins are thought to translocate through interaction with chaperones which maintain an unfolded state [31]. However, although no studies to date have shown interaction of DNMT1 with chaperones, the idea of a chaperone complex shuttling DNMT1 to the mitochondria should not be out of the question. Most peptides imported into mitochondria are recognized by their N-terminal targeting sequences; however, it is also acknowledged that many precursors have internal recognition sequences [9]. The nature of these internal sequences have not been characterized largely because they lack consistent arrangements [9]; the possibility of DNMT1 also containing an internal targeting sequence has not been investigated.

The argument over whether or not DNMT3B is taken to the mitochondria will still be debatable according to formerly presented blots. There has been only one published study where a western blot has shown DNMT3B in the mitochondrial fraction of a cell [16]. Blots from our current study are inconclusive. Each antibody used in this work yielded vastly different results despite the fact that each antibody recognized similar and overlapping regions of DNMT3B. Several publications and manufacturer datasheets using these antibodies show DNMT3B ranging in size from 80-130kD; the predicted molecular mass, based on amino acid sequence is 97.5kD.

Figure 3-8A displays strong bands in the whole-cell lysate and faint bands in cytosolic fractions at ~80kD in the HCT116 wt and the HCT genome edited cell lines; bands are absent in the 3BKO models indicating that there is a true knock-out, however there is no signal in mitochondrial extracts, suggesting that DNMT3B does not get transported to the mitochondria to the extent detectable by immunoblot. Figure 3-8B displays the same size band as the previous figure however it is expressed in every sample and only in the whole-cell lysate fraction. If the 80kD band truly represents DNMT3B, then its presence in the supposed 3BKO line indicates that this line is not a true knock-out. Interestingly there is a band at ~45kD expressed strongly in the mitochondrial fraction of both the HCT lines and faintly in the mitochondrial fraction of the 3BKO lines. These results might indicate cleavage of DNMT3B during translocation into the mitochondria in the HCT lines, and possible DNMT3B contamination from cytosol resulting in the faint bands seen in the 3BKO mitochondrial extracts, perhaps due to improper technique. Figure 3-8C is just as unique as the first two blots; there is a band in the whole-cell lysate and cytosolic fractions of every sample at ~100kD that could be DNMT3B. Again, in the mitochondrial fraction there is a band that could represent cleavage of DNMT3B in the import complexes of the mitochondria at ~40kD; however Figure 3-8C displays these bands in every sample. Several of these figures are suggestive that there is not a true knock-out of DNMT3B; thus our genotyping experiment was crucial to characterization of this line before proceeding on to other characterization experiments. Verification of a knock-out was attained (Fig. 3-9) meaning that any interpretation from the immunoblots would be difficult to substantiate.

DNMT3B does not contain an N-terminal MTS as found for DNMT1. An internal targeting sequence has not been found within DNMT3B, therefore transport to the mitochondria or recognition of the protein at the outer membrane is questionable. N-terminal targeting

sequences are predominantly cleaved on passage through the outer membrane [31], and, for many proteins, multiple cleavage events occur along the length of the peptide upon passage through the inner mitochondrial membrane. The purpose of these proteolytic cleavage events that occur throughout the import complexes has not been identified, although some proteins have been characterized to be cleaved at defined points to specific sizes once inside the mitochondrial matrix. Previous work in the Taylor lab has shown immunoblots of DNMT1 expression in the mitochondrial fraction [22]; however, there has been a consistent lower band that has not been characterized at ~60kD. The DNMT3B immunoblots of this study sometimes display lower bands as well in the mitochondrial fraction. It would be of significance to examine the lower bands in current and past studies with techniques such as mass spectrometry to better identify their relationship to the methyltransferases in the mitochondria.

RTqPCR data was more accommodating to address the functional significance of DNA methyltransferases in the mitochondria. Across the three protein coding mitochondrial genes tested, ND1, Cox1 and ND6, there is an apparent trend towards reduced expression, correlating with genome editing of the MTS in mtDNMT1 alone or in conjunction with knockout of DNMT3B. Our previously published work indicates that increased mtDNMT1 expression results in increased mitochondrial transcription [22], and ongoing *in vitro* experiments in our lab using a reconstituted system with recombinant enzymes and transcription factors, demonstrate a 3-fold increase in transcription from HSP when the 3 CG dinucleotides in and around the promoter are methylated, in comparison to the unmethylated sequence. The 16s ribosomal RNA gene does not showcase this trend as the difference between samples is subtle. The level of rRNA transcription is much higher than that of protein coding mRNAs, and rRNA transcription is believed to operate by a different mechanism that is independent of polycistronic ORF mRNA

production [32]. ND1, Cox1 and ND6 demonstrate a significant pattern of decline amongst the samples. In those gene sets when the MTS of DNMT1 has been genomically edited out of the HCT116 cell line there is an overall decrease in transcription of ~27%. The most significant decline in transcription occurs in the DNMT3B knock-out cell line, which has a decline in transcription of ~55%. Further editing of the DNMT1 MTS in this background did not further decrease the transcription level.

The lack of significance in the difference between the genome edited 3BKO and the 3BKO in itself could possibly mean that DNMT3B and not DNMT1 is the more functionally significant methyltransferase in the mitochondria. Previous studies have shown in the nucleus that in the absence of DNMT1, DNMT3B expression increases in the absence of the DNMT1 [30]. This could possibly be the same case in the mitochondria and explain that in the genome edited cell lines the reduction in transcription is only 27% and not closer to 50%; that DNMT3B is making up for some of the absence of DNMT1. This is possibly evident in (Fig 3-8) where one can see a greater expression level of DNMT3B in the HCT genome edited cell line over the HCT116 wild-type.

Endpoint PCR has been a standard checkpoint throughout multiple investigations in the Taylor Lab and many other labs. The agarose gel image from endpoint PCR gives a basic measurement of success from the MeDIP; however, this assay needs to be closely scrutinized as a lower cycle number reaction can reveal discrepancies obscured with a higher cycle number, as shown in this research. It would be essential to optimize this assay in conjunction with the MeDIP before placing any value on the reliability as a checkpoint.

QPCR was planned to be an essential assay in this research for measuring how genomic modifications to methyltransferases affect levels of methylation within mitochondrial DNA. As results proved to be highly variable, we do not yet have the necessary number of replicates for conclusive support of our hypothesis. The principle effort of having equivalent amounts of DNA going into the MeDIP is critical in order to have reliable qPCR data. However, we were able to pinpoint the source of error, in the faulty sonicator. The troubleshooting to find this equipment failure, although prolonged, did not go to waste; the protocol for the MeDIP was shortened and made to be more efficient and precise, and the technical experience gained was invaluable. Replacement of the faulty equipment should allow future experiments to complete this study for publication.

Chapter 5: Perspectives and Conclusions

There is continued controversy in the literature concerning which methyltransferases reside in the mitochondrial matrix, largely generated because different investigators use different antibodies with largely differing sensitivities. We chose a more specific genetic approach in order to avoid issues of antibody detection.

The goals of this study were therefore to define which methyltransferases were present in mammalian mitochondria and to gain understanding of the role and functional significance of mitochondrial DNA methylation. We approached this study using CRISPR-Cas9, the new and cutting-edge technology for genome editing to remove the MTS from DNMT1, thus minimizing effects on nuclear cytosine methylation, while impacting mitochondrial modification. This same approach is not possible with either DNMT3a or 3b, because they lack N-terminal MTS. Our expectation was that depletion of mtDNMT1 would result in reduced cytosine methylation along with reduced transcription.

To our surprise, removal of the DNMT1 MTS did not completely prevent the localization of this enzyme to the mitochondria, raising the possibility that the enzyme can translocate to this organelle using alternative mechanisms. Significantly reduced protein expression of the cell lines without a DNMT1 MTS conceivably emphasize the cellular importance of necessitating DNMT1 to be in the mitochondria; however, the transcription data presented here supports the notion that other methyltransferases besides DNMT1 localize to mitochondria and may be ever more functionally significant. Deletion of the MTS in DNMT1

shows only a meek reduction in transcription, not until removal of DNMT3B do we see a two-fold decrease in transcription from mitochondrial protein coding genes. As well, no significant reduction in transcription occurred when the DNMT3B knockout also lost the MTS of DNMT1; apparently, contrary to what we expected, DNMT3B is being evidenced to big a bigger player in mitochondrial transcription.

DNMT3B appears to be an attention-grabbing target to further explore. Investigation into possible processing of DNMT3B upon entering the mitochondria is evidenced by this works immunoblots, could lead to the further convincing proof that DNMT3B is translocated to the mitochondria. Not only could research define translocation, it could very well point towards the mechanism of translocation. An understanding if and/or where DNMT3B is being cleaved might expose some internal targeting sequence that could be referenced with other different proteins that have not been able to be exposed in the mitochondria.

As an orthogonal approach to quantifying functional significance of DNA methyltransferases to under-methylated mtDNA, the Taylor Lab has developed a completely methylated representation; accomplished by M.Sss1 which is a CpG methyltransferase that will methylate all CpGs. Since methylation in the mitochondria is known to be exceptionally low compared to the nuclear genome, this over-methylation approach may give a more true quantification. The Taylor Lab has humanized codon usage and placed the TFAM MTS at the N-terminus of M.Sss1. We found that constitutive overexpression is non-permissive since we were not originally able to obtain stable clones. The Taylor Lab has now developed an inducible construct of M.Sss1 and have begun to test its effect on mtDNA methylation and mitochondrial transcription. As well, current investigation suggest that genome edited cell lines have

differences in mtDNA copy number average compared to that of wild-type. Lab efforts have begun to address the consistency of such differences and the most effective way to normalize them.

Further effort will be necessary to conclude what mechanisms in the mitochondria are at work. To really comprehend the impact of mtDNA methylation it will be critical to classify all of the contributors involved and understand the mechanistic complexities that accompany cause and consequence of epigenetic modifications. Genomic editing will significantly support the journey to understanding of epigenetics of the mitochondrion. Mitochondria have been implicated in a variety of neurodegenerative disorders and cancers and the lingering question of what is the plausible biological function of methylation in the mitochondria needs to be answered.

Works Cited

1. Nunnari, J., & Suomalainen, A. (2012) Mitochondria: In sickness and in health. *Cell, 148*(6), 1145-1159.
2. Uzman, A. (2003). Molecular biology of the cell (4th ed.): Alberts, B., Johnson, A., Lewis, J., raff, M., Roberts, K., and Walter, P. *Biochemistry and Molecular Biology Education, 31*(4), 212-214.
3. Birch-Machin, M. A. (2006). The role of mitochondria in ageing and carcinogenesis. *Clinical and Experimental Dermatology, 31*(4), 548-552.
4. Giulivi, C., Zhang, Y., Omanska-Klusek, A., & al, e. (2010). Mitochondrial dysfunction in autism. *Jama, 304*(21), 2389-2396.
5. Qi, Y., Yin, X., Wang, S., Jiang, H., Wang, X., Ren, M., Feng, H. (2015). PGC-1α silencing compounds the perturbation of mitochondrial function caused by mutant SOD1 in skeletal muscle of ALS mouse model. *Frontiers in Aging Neuroscience, 7*, 204.
6. Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., Young, I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature, 290*(5806), 457-465.
7. Iacobazzi, V., Castegna, A., Infantino, V., & Andria, G. (2013). Mitochondrial DNA methylation as a next-generation biomarker and diagnostic tool. *Molecular Genetics and Metabolism, 110*(1–2), 25-34.
8. Satoh, M., & Kuroiwa, T. (1991). Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell. *Experimental Cell Research, 196*(1), 137-140.

9. Neupert, W., & Herrmann, J. M. (2007). Translocation of proteins into mitochondria. *Annual Review of Biochemistry, 76*(1), 723-749.
10. Williams, S. C. P. (2013). Epigenetics. *Proceedings of the National Academy of Sciences of the United States of America, 110*(9), 3209-3209.
11. Eccleston, A., DeWitt, N., Gunter, C., Marte, B., & Nath, D. (2007). Epigenetics. *Nature, 447*(7143), 395-395.
12. Arimondo, P. B., Egger, G., & Tost, J. (2012). Epigenetics. *Biochimie, 94*(11), 2191-2192.
13. Jurkowska, R. Z., Jurkowski, T. P., & Jeltsch, A. (2011). Structure and function of mammalian DNA methyltransferases. *Chembiochem, 12*(2), 206-222.
14. Hermann, A., Gowher, H., & Jeltsch, A. (2004). Biochemistry and biology of mammalian DNA methyltransferases. *Cellular and Molecular Life Sciences CMLS, 61*(19), 2571-2587.
15. Arimondo, P. B., Egger, G., & Tost, J. (2012). Epigenetics. *Biochimie, 94*(11), 2191-2192.
16. Bellizzi, D., D'Aquila, P., Scafone, T., Giordano, M., Riso, V., Riccio, A., & Passarino, G. (2013). The control region of mitochondrial DNA shows an unusual CpG and non-CpG methylation pattern. *DNA Research, 20*(6), 537-547.
17. Maresca, A., Zaffagnini, M., Caporali, L., Carelli, V., & Zanna, C. (2015). DNA methyltransferase 1 mutations and mitochondrial pathology: Is mtDNA methylated? *Frontiers in Genetics, 6*

18. Tatton-Brown, K., Seal, S., Ruark, E., Harmer, J., Ramsay, E., del, V. D., Rahman, N. (2014). Mutations in the DNA methyltransferase gene DNMT3A cause an overgrowth syndrome with intellectual disability. *Nature Genetics, 46*(4), 385-388.
19. Vanyushin, B. F., Kiryanov, G. I., Kudryashova, I. B., & Belozersky, A. N. (1971). DNA-methylase in loach embryos (misgurnus fossilis). *FEBS Letters, 15*(4), 313-316.
20. Maekawa, M., Taniguchi, T., Higashi, H., Sugimura, H., Sugano, K., & Kanno, T. (2004). Methylation of mitochondrial DNA is not a useful marker for cancer detection. *Clinical Chemistry, 50*(8), 1480-1481.
21. Pollack, Y., Kasir, J., Shemer, R., Metzger, S., & Szyf, M. (1984). Methylation pattern of mouse mitochondrial DNA. *Nucleic Acids Research, 12*(12), 4811-4824.
22. Shock, L. S., Thakkar, P. V., Peterson, E. J., Moran, R. G., & Taylor, S. M. (2011). DNA methyltransferase 1, cytosine methylation, and cytosine hydroxymethylation in mammalian mitochondria. *Proceedings of the National Academy of Sciences, 108*(9), 3630-3635.
23. Kohli, R. M., & Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. *Nature, 502*(7472), 472-479.
24. Balinang, Joyce. MS Thesis. *The Regulation of Mitochondrial DNMT1 During Oxidative Stress.* Virginia Commonwealth University, 2012.
25. Burton, Elliot. MS Thesis. *Functional consequences of mtDNA methylation on mitochondrial transcription factor binding and transcription initiation*. Virginia Commonwealth University, 2016

26. Bachu, R., Bergareche, I., & Chasin, L. A. (2015). CRISPR-cas targeted plasmid integration into mammalian cells via non-homologous end joining. *Biotechnology and Bioengineering, 112*(10), 2154-2162.

27. Oude Blenke, E., Evers, M. J. W., Mastrobattista, E., & van der Oost, J.CRISPR-Cas9 gene editing: Delivery aspects and therapeutic potential. *Journal of Controlled Release*

28. Fiskus, W., Herger, B., Rao, R., Atadja, P., & Bhalla, K. (2015). Hydroxamate pan-HDAC inhibitor LBH589 depletes EZH2 and DNMT1, partly through hsp90 inhibition and its chaperone association with DNMT1. *Blood, 108*(11), 2233-2233.

29. Barrès, R., Osler, M. E., Yan, J., Rune, A., Fritz, T., Caidahl, K., Zierath, J. R.Non-CpG methylation of the *PGC-1α* promoter through DNMT3B controls mitochondrial density. *Cell Metabolism, 10*(3), 189-198.

30. Elliott, E. N., Sheaffer, K. L., & Kaestner, K. H. (2016). The 'de novo' DNA methyltransferase Dnmt3B compensates the Dnmt1-deficient intestinal epithelium. *Elife, 5*, e12975.

31. Transport of proteins into mitochondria and chloroplasts. (1979). *The Journal of Cell Biology, 81*(3), 461-483.

32. Martin, M., Cho, J., Cesare, A. J., Griffith, J. D., & Attardi, G.Termination factor-mediated DNA loop between termination and initiation sites drives mitochondrial rRNA synthesis. *Cell, 123*(7), 1227-1240.

33. Lo, C. Y., Chen, S., Creed, S. J., Kang, M., Zhao, N., Tang, B. Z., & Elgass, K. D. (2016). Novel super-resolution capable mitochondrial probe, MitoRed AIE, enables assessment of real-time molecular mitochondrial dynamics. *Scientific Reports, 6*, 30855.

34. Jakobs, S., & Wurm, C. A. (2014). Super-resolution microscopy of mitochondria. *Current Opinion in Chemical Biology, 20*, 9-15.