Blood Mitochondrial DNA Mutations in HIV-Infected Women and Their
Infants Exposed to HAART during Pregnancy

by

Marissa Hélène Jeanne Jitratkosol

B.Sc., University of British Columbia, Vancouver, 2006

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Pathology and Laboratory Medicine)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

June 2010

© Marissa Hélène Jeanne Jitratkosol, 2010

Abstract

Background/Objectives:

Nucleoside reverse transcriptase inhibitors (NRTIs) as part of highly active antiretroviral therapy (HAART) are given to human immunodeficiency virus (HIV)-infected pregnant women to prevent HIV vertical transmission. NRTIs can adversely affect mitochondrial DNA (mtDNA) and may induce mtDNA point mutations. We hypothesised that HAART-exposed/HIV-uninfected infants may show higher blood mtDNA mutation burden than controls born to HIV-uninfected mothers.

Methods:

Blood was collected from infants exposed *in utero* to HIV/HAART and controls (0-6d), as well as from a subset of their mothers (last visit before delivery). MtDNA mutation burden was measured by an assay involving cloning and sequencing mtDNA D-loop PCR amplicons. The presence of transversion mutations A→C/T→G (AC/TG) was analysed by Chi-squared and Wilcoxon signed-rank tests. Relationships with amount of DNA assayed, maternal age, smoking (marijuana/cigarettes) and illicit drug/methadone use in pregnancy were examined. For the HIV/HAART group, relationships with CD4+ count and HIV plasma viral load (pVL) near delivery, as well as length of HAART exposure were also examined.

Results:

The *Taq* error rate from PCR caused a low signal (mutation) to noise (background) ratio. Therefore, only AC/TG mutations, not induced under our assay conditions, were analysed. No significant difference was found between the percentage of HIV/HAART-exposed infants with AC/TG mutations (N=15/57, 26.3%) and controls (N=10/70, 14.3%) before ($p=0.090$) or after ($p=0.058$) controlling for covariates, although a trend was observed. Furthermore, significantly more HIV/HAART-exposed mothers (N=18/42, 42.9%) harboured AC/TG mutations compared to controls (N=7/39, 17.9%) both before ($p=0.015$) and after ($p=0.012$) controlling for covariates. AC/TG mutations were more prevalent in HIV/HAART-exposed mothers than in their infants (N=42, 42.9% vs. 23.8% $p=0.033$), however, this difference disappeared

after controlling for covariates ($p=0.777$). No difference was observed between control mothers and their infants (N=39, both 17.9%). In HIV/HAART-exposed group mothers, only a detectable HIV pVL near delivery predicted AC/TG mutations.

Conclusion:

A subset of mtDNA mutations can be quantified with the developed assay. HIV/HAART exposure in pregnancy may be associated with increased prevalence of maternal mtDNA mutations. Since mtDNA mutations have been linked with aging and age-associated diseases, this raises concerns about the long-term impact of HAART.

Table of Contents

Abstract	ii
Table of Contents	iv
List of Tables.....	viii
List of Figures	ix
List of Abbreviations.....	xi
Acknowledgments.....	xiv
Dedications	xv
1 Introduction	1
1.1 Introduction to the Thesis	1
1.2 HIV/AIDS Epidemic.....	1
1.3 HIV Pathophysiology.....	3
1.3.1 HIV Structure	3
1.3.2 HIV Replication Cycle	3
1.3.3 Natural History of HIV Infection	4
1.4 Highly Active Antiretroviral Therapy (HAART).....	5
1.4.1 PIs.....	7
1.4.2 NRTIs and NtRTIs	7
1.4.3 NNRTIs	7
1.5 HIV and Pregnancy.....	9
1.5.1 Overview.....	9
1.5.2 HAART during Pregnancy.....	9
1.5.3 Prevention of Mother-to-Child Transmission in Low-Income Settings	11
1.5.4 Outcome of Pregnancies with HIV-Infection	12
1.6 Mitochondria	13
1.6.1 Overview.....	13

1.6.2	Mitochondrial DNA (MtDNA) and Heteroplasmy	16
1.6.3	Mitochondrial Disorders	20
1.6.4	Random MtDNA Mutation Accumulation and Oxidative Stress.....	22
1.6.5	Aging	23
1.6.6	NRTIs and Mitochondrial Toxicity.....	27
1.6.7	Total MtDNA Point Mutation Detection.....	31
2	MtDNA Mutation Burden Assay-Background Error Rate Determination, Method Validation and Optimisation	35
2.1	Introduction	35
2.2	Methods and Materials	37
2.2.1	Samples.....	37
2.2.2	DNA Extraction	37
2.2.3	MtDNA Mutation Burden Assay (MMBA).....	37
2.2.4	Determining Background Error Rate of the Assay.....	41
2.3	Results.....	43
2.3.1	Number of Clones to Analyse with Blood Samples	43
2.3.2	Overview of Background Error Rate Determination	44
2.3.3	Time-dependent Effect on Background Error Rate for <i>PfuU</i> at Varying Numbers of PCR Amplification Cycles and C-tract Lengths	45
2.3.4	Background Error Rate for All Mutation Types with Varying C-tract Lengths.....	46
2.3.5	Background Error Rate for Non-C-tract All Mutation Types with Varying Number of PCR Amplification Cycles.....	50
2.3.6	Background Error Rate for Each Mutation Type.....	50
2.3.7	Sequencing Reaction Background Error Rate	53
2.3.8	Background Error Rate Compared to Average Clinical Sample MtDNA Mutation Burden	53
2.3.9	Intra-sample and Inter-sample Variability.....	56
3	Blood Mitochondrial DNA Mutations in HIV-infected Women and their Infants Exposed to HAART during Pregnancy.....	57

3.1	Overview	57
3.2	Study Design	57
3.2.1	Hypothesis	57
3.3	Materials and Methods	58
3.3.1	Funding and Ethical Approval	58
3.3.2	Study Populations	58
3.3.3	Samples Used and Collection Method.....	61
3.3.4	Data Collection	62
3.3.5	Clinical Laboratory Testing	62
3.3.6	DNA Extraction	62
3.3.7	MtDNA Mutation Burden Assay	62
3.3.8	Statistical Analyses	65
3.4	Results.....	66
3.4.1	Study Population	66
3.4.2	Sample Size Calculation	71
3.4.3	MtDNA Mutation Burden	72
3.4.4	Correlation Between Mother/Infant Pairs	78
3.4.5	Correlation Between Maternal Age and MtDNA Mutation Burden in Infants and Mothers .	79
3.4.6	Heteroplasmy Outside the C-tract.....	80
3.4.7	Consensus C-tract Length	80
3.4.8	MtDNA Mutations Within the C-tract	81
3.4.9	Consensus Sequence Between Mother and Infant	84
3.4.10	Sequences with Multiple Mutations.....	84
3.4.11	Location of Point Mutations	85
4	Discussion.....	89
4.1	MtDNA Mutation Burden Assay	89
4.1.1	Limitation.....	93

4.1.2	Conclusion	93
4.2	Clinical Study.....	94
4.2.1	Limitations	102
4.3	Future Directions	104
4.3.1	Different MtDNA Sequencing Method	104
4.3.2	Future Clinical Studies	105
4.3.3	Future Basic Science Studies	106
4.4	Conclusions	107
	References	109
	Appendix: Ethics Review Board Approval Certificate	122

List of Tables

Table 1: Complexes forming the ETC	17
Table 2: Comparison of human nuclear and mitochondrial genome	20
Table 3: Relative frequency of clinical symptoms with various NRTIs.....	28
Table 4: Advantages and disadvantages of different methods/strategies to study mtDNA mutations/heteroplasmy	34
Table 5: List of primers used in the mtDNA mutation burden assay	38
Table 6: Number of times experiments were performed at various PCR amplification conditions.....	45
Table 7: Mutation spectra observed at 25 and 35 PCR amplification cycles with HiFi <i>Taq</i> and <i>PfuU</i> with all C-tract lengths	51
Table 8: Intra-sample total mutation rate variability in 4 individuals	56
Table 9: Intra-sample AC/TG mutation rate variability in 4 individuals.....	56
Table 10: Demographic characteristics, clinical characteristics, and laboratory values for HIV/HAART-exposed mothers and their infants, as well as unexposed control mothers and their infants	70
Table 11: Comparison of the percentage of HIV/HAART-exposed and unexposed control infants and mothers with the AC/TG mutations.....	76
Table 12: Comparisons of the percentage of infants and mothers with the AC/TG mutations with and without HAART-exposure.....	77
Table 13: Consensus C-tract length and heteroplasmy within the C-tract in infants and mothers in the HIV/HAART-exposed and unexposed control groups	81
Table 14: Type and number of mutations from subjects with a single sequence with multiple transition mutations	85
Table 15: Subjects with mutations or heteroplasmy and their consensus base at locations of point mutations referred to in the mtDNA literature.	87
Table 16: Summary of total error rates of different PCR polymerases with 25 and 35 amplification cycles in the literature and with the mtDNA mutation burden assay	91
Table 17: Advantages and disadvantages of each PCR polymerase	94

List of Figures

Figure 1: Natural history of HIV infection	5
Figure 2: Structures of common nucleoside reverse transcriptase inhibitors (NRTIs) and one nucleotide reverse transcriptase inhibitor (NtRTI)	8
Figure 3: An overview of the electron transport chain.....	15
Figure 4: Map of the human mitochondrial genome.....	16
Figure 5: Diagram of heteroplasmy at the mitochondrion and cellular level.....	18
Figure 6: The “vicious cycle” between ROS and mitochondrial dysfunction	24
Figure 7: D-loop region of mtDNA amplified for the assay (nt16559 – 447)	36
Figure 8: Chromatogram of a characteristic “out-of-phase” sequence due to the C-tract	40
Figure 9: Chromatogram of a peak within a peak leading to ambiguous base calling	41
Figure 10: Schematic diagram of the mtDNA mutation burden assay (MMBA) and determination of the background error rate (BER) of the assay	42
Figure 11: Intra-sample CV for blood samples from 2 subjects with HiFi <i>Taq</i> and <i>PfuU</i> at 25 PCR amplification cycles	44
Figure 12: Time dependent effect on background error rates outside the C-tract mutation with <i>PfuU</i> at varying number of PCR amplification cycles and C-tract lengths	46
Figure 13: Background error rates within the C-tract with <i>PfuU</i> at 25, as well as HiFi <i>Taq</i> at 25 and 35 PCR amplification cycles with varying C-tract lengths	47
Figure 14: Background error rates outside the C-tract with <i>PfuU</i> at 25, as well as HiFi <i>Taq</i> at 25 and 35 PCR amplification cycles with varying C-tract lengths	47
Figure 15: Mean percentage of total sequences with a mtDNA mutation by type observed within the C-tract with varying consensus C-tract lengths with HiFi <i>Taq</i> and <i>PfuU</i> at 25 PCR amplification cycles	48
Figure 16: Mean percentage of total sequences with a mtDNA mutation by type observed within the C-tract with varying consensus C-tract lengths with HiFi <i>Taq</i> at 35 PCR amplification cycles.....	49
Figure 17: Background error rates outside the C-tract with HiFi <i>Taq</i> and <i>PfuU</i> at varying numbers of cycles with an 8 C C-tract	50
Figure 18: Background error rates for each type of mutation with <i>PfuU</i> at 25 and HiFi <i>Taq</i> at 25 and 35 PCR amplification cycles with differing C-tract lengths	52

Figure 19: Mean clinical sample mutation rates and background error rate observed for each type of mutation.....	55
Figure 20: Schematic diagram of mtDNA mutation burden assay for clinical samples	64
Figure 21: Flowchart summary of study samples	67
Figure 22: Scatterplot comparing HIV/HAART-exposed and unexposed control infants' and mothers' total mutation rates.	74
Figure 23: Scatterplot comparing infants' to their mothers' total mutation rates in the HIV/HAART-exposed and unexposed control groups.	75
Figure 24: Scatterplots showing lack of significant correlation between infants and their mothers in either HIV/HAART-exposed or unexposed control groups.	78
Figure 25: Scatterplots showing lack of significant correlation between maternal age and infant or maternal total mutation rate of both groups combined.	79
Figure 26: Mean percentage of 80 sequences with a mtDNA mutation by type observed within the C-tract in subjects with varying consensus C-tract lengths	83
Figure 27: Schematic illustration of the formation of AC/TG transversion mutations from 8-oxo-dGTP during DNA replication.....	97

List of Abbreviations

Abbreviation	Full Name
3TC	Lamivudine/2',3'-dideoxy-3'-thiacytidine (chemical name)
A	Adenosine (nucleoside)/Adenine (base)
AD	Alzheimer's disease
ABC	Abacavir
AIDS	Acquired Immune Deficiency Syndrome
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
ART	AntiRetroviral Therapy
ARV	AntiRetroViral
ATP	Adenosine TriPhosphate
ATV	Atazanavir
AZT	Zidovudine(ZDV)/formerly called azidothymidine
BC	British Columbia
BER	Background Error Rate
bp	Base Pair
C	Cytidine (nucleoside)/Cytosine (base)
CD4+	Cluster Designation 4 positive lymphocytes
CDC	Centre for Disease Control and Prevention
chr	Chromosome
CI	Confidence Interval
CPEO	Chronic Progressive External Ophthalmoplegia
C & W	Children's and Women's Health Centre of British Columbia
CV	Coefficient of Variation
d4T	Stavudine/2',3'-didehydro-3'-deoxythymidine (chemical name)
ddC	Zalcitabine/formerly called dideoxycytidine
ddI	Didanosine/formerly called dideoxyinosine
DNA	DeoxyriboNucleic Acid
dNTP	DeoxyriboNucleoside TriPhosphate
EFV	Efavirenz
ETC	Electron Transport Chain
FTC	Emtricitabine (thymidine analogue with a stabilising fluorine atom)

G	Guanosine (nucleoside)/Guanine (base)
gp	GlycoProtein
HAART	Highly Active AntiRetroviral Therapy
HiFi <i>Taq</i>	Expand High Fidelity ^{PLUS} Enzyme Blend (Roche)
HIV	Human Immunodeficiency Virus Type-1
HCV	Hepatitis C virus
IDV	Indinavir
Indel	Insertion and Deletion
IQR	Interquartile Range
kb	kilobases
LPV	Lopinavir
MMBA	MtDNA Mutation Burden Assay
MTCT	Mother-To-Child-Transmission
mtDNA	Mitochondrial DNA
mRNA	messenger RNA
NFV	Nelfinavir
NNRTI	Non Nucleoside Analogue Reverse Transcriptase Inhibitor
NRTI	Nucleoside Analogue Reverse Transcriptase Inhibitor
nt	mtDNA nucleotide position
NtRTI	Nucleotide Reverse Transcriptase Inhibitor
NVP	Nevirapine
OXPHOS	Oxidative Phosphorylation
PBMC	Peripheral Blood Mononuclear Cell
PCR	Polymerase Chain Reaction
<i>PfuU</i>	<i>PfuUltra</i> II Fusion HS DNA polymerase (Stratagene)
PI	Protease Inhibitor
PMTCT	Preventing/Prevent Mother-to-Child Transmission
POLG	DNA Polymerase γ
pVL	Plasma Viral Load
RFLP	Restriction Fragment Length Polymorphism
RNA	RiboNucleic Acid
ROS	Reactive Oxygen Species
rRNA	ribosomal RNA

RT	Reverse Transcriptase
RTV	Ritonavir
S.D.	Standard Deviation
SNP	Single Nucleotide Polymorphism
SQV	Saquinavir
T	Thymidine (nucleoside)/Thymine (base)
TDF	Tenofovir/disoproxil fumarate
tRNA	transfer RNA
UBC	University of British Columbia
UNAIDS	Joint United Nations Programme on HIV/AIDS

Acknowledgments

I would like to thank my supervisor, Dr. Hlne Ct for her guidance and support throughout my research and for having taught me how to become a good scientist and believing in my potential. I would like to thank my advisory committee, Drs. Ed Prydzial, Richard Harrigan and Paula Waters for their valuable time and kind words of encouragement. Several people deserve special mention. Isabelle has always been there to listen, lend a helping hand and advise. Without Beheroz's unrelenting hard work and dedication, this project would not have been completed. Henry has been a great friend and always around to talk about things ranging from nerdy sciency things to politics. Most importantly, with Henry, I was able to share with him Excel excitement and he understood why my eyes twinkled when talking about it. I thank Rachel and Jon for their help and patience in reading my thesis. I thank Evelyn for answering my numerous email questions almost instantaneously. Lastly, it has been a pleasure working with everyone in the Ct lab: Tuhina Imam, Dr. Eszter Papp, Carmen Li and Laura Oliveira.

Dedications

To my parents, partner and friends

1 Introduction

1.1 Introduction to the Thesis

Nucleoside analogue reverse transcriptase inhibitors (NRTIs) are associated with mitochondrial toxicity and this toxicity may be in part due to mitochondrial DNA (mtDNA) mutations accumulating from such treatments. However, this latter link has yet to be confirmed by several sources. This will be reviewed in detail in **section 1**.

The aim of this Master's project was to investigate mtDNA mutation burden in blood of infants and HIV-infected mothers exposed to NRTI containing-highly active antiretroviral therapy (HAART) during pregnancy to prevent mother to child transmission of human immunodeficiency virus (HIV) and to compare these with infants and mothers HIV/HAART-unexposed (**section 3**). The underlying hypothesis was that NRTI containing-HAART exposure is associated with increased mtDNA mutation burden. This is important to study since mtDNA mutations have been linked with aging and age-associated diseases such as Alzheimer's disease and diabetes.

Prior to this, an assay was developed to determine mtDNA mutation burden in a sample involving cloning and sequencing of mtDNA PCR amplicons. The background error rate of this assay was also examined (**section 2**). To complete this project more than 20,000 sequences were analysed.

1.2 HIV/AIDS Epidemic

In 2007, the Joint United Nations Programme on HIV/AIDS (UNAIDS) estimated that there were 33 million people living with the human immunodeficiency virus (HIV) worldwide; 67% of whom were living in sub-Saharan Africa and accounted for 72% of all Acquired Immunodeficiency Syndrome (AIDS) deaths. Globally, 50% of HIV-infected individuals are women of child-bearing age. More importantly, the percentage of women infected with HIV is increasing in many countries. It is estimated that at least 3.28 million pregnancies with HIV infection will occur every year [1]. In 2007 alone, approximately 370,000

children under the age of 15 years became infected with HIV, totalling 2 million children infected worldwide [2]. Ninety percent of these children are thought to have acquired HIV through mother-to-child transmission (MTCT) that is, either *in utero*, during delivery, labour or breastfeeding. Roughly 50% will die before the age of two [3, 4]. Encouragingly, the annual infection and death rates among children worldwide have declined through the expansion of services to prevent mother-to-child transmission (PMTCT) and treatment [2]. Globally, coverage of PMTCT services rose from 10% in 2004 to 45% in 2008 [5]. Moreover, a UNAIDS priority is to eliminate MTCT by 2015, which means providing universal access to PMTCT services to women around the world [6].

In most areas outside of sub-Saharan Africa, the vast majority of those who become infected are injecting drug users, men who have sex with men and sex workers [2]. Canada also follows this trend, of the estimated 65,000 HIV-infected persons in 2008¹, 45% were men who have sex with men and 17% were injection drug users. Aboriginals are overrepresented in the Canadian HIV epidemic, accounting for 8% of HIV-infected persons, despite representing only 4% of the country's population [7].

The percentage of HIV-infected people in Canada who are women is much lower than the global average, at 22%. The main risk factors for HIV acquisition in Canadian women is heterosexual contact and injection drug use [7]. Today, new infections acquired through MTCT are exceedingly low in Canada; this can be attributed to the high rate of prenatal HIV testing and high coverage of free PMTCT services. High rates of testing are due to the implementation of an "opt-out" approach in many provinces, whereby all pregnant women are tested unless they specifically decline. In Alberta, HIV testing occurs in 95% of pregnancies; the highest rates of testing in North America [8]. On the other hand, in BC an "opt-in" approach is used, and as a result, the test rates were only 87% in 2005 [9]. Nationally, the proportion of HIV-infected pregnant women receiving antiretroviral therapy (ART) for PMTCT increased from 60% in 1997 to 89% in 2006. Concurrently, the percentage of children acquiring HIV perinatally decreased

¹ Hereafter HIV refers to HIV type-1 (HIV-1). Two types of HIV have been characterised, HIV-1 and HIV-2. HIV-2 is much less common and is mostly found in West Africa. HIV-1 is predominately responsible for the global AIDS epidemic [8].

significantly, from 22% of at risk pregnancies in 1997 to 3% in 2006 [8]. In BC, there is close to complete antenatal and postpartum HIV care of all HIV-infected pregnant women either directly or indirectly through the Oak Tree Clinic situated at the Children's and Women's Centre of British Columbia (C & W). Between 2003 and 2008, there were approximately 25 pregnant women living with HIV per year, most of whom received ART during their pregnancy for the purpose of PMTCT [9].

1.3 HIV Pathophysiology

1.3.1 HIV Structure

HIV is a member of the retrovirus family; these ribonucleic acid (RNA) viruses characteristically contain the enzyme reverse transcriptase (RT) that produces DNA from their RNA genome. HIV is an enveloped virus composed of host cell plasma membrane imbedded with the glycoprotein (gp) gp41 which is attached to gp120 forming spikes on the envelop surface. Within the virus is a core that resembles a tapered cylinder that consists of capsid protein, which contains viral protease, integrase, RT and other viral proteins, as well as 2 identical copies of the 9.2 kb single-stranded positive RNA genome that encodes 9 viral proteins [10].

1.3.2 HIV Replication Cycle

The HIV replication cycle begins with attachment of the HIV virus to the host cell. This occurs through the interaction between gp120 and, most commonly, the CD4 receptor, a transmembrane glycoprotein involved in T-cell activation that is mainly expressed on T-cells, monocytes/macrophages and dendritic cells. Attachment triggers a conformational change in gp120 that allows binding to the co-receptor CXCR4 or CCR5 (depending on the viral tropism) and fusion to occur. This is followed by the viral core and its contents being released into the cytoplasm [11]. While being transported towards the nucleus, the RNA genome is reverse transcribed to produce full-length double-stranded DNA by viral RT. Subsequently, the double-stranded DNA is integrated into the human genome by the enzyme, integrase with no apparent site preference; the integrated DNA is termed a provirus. Depending on the site of

integration, the provirus can remain dormant for a long period of time [10]. The result of integration is the persistence of virus during the remaining part of the host cell's life [12]. Fabrication of virus occurs when the host cell's RNA polymerase begins to transcribe the proviral DNA, into messenger RNA (mRNA) and new viral HIV genomic RNA for subsequent virions [13]. The mRNA is translated into viral enzymes and viral polypeptides. The final steps of the viral cycle include budding and maturation. During budding, the viral protease is activated within one of the polypeptides produced from viral mRNA translation. Protease cleaves the polypeptide releasing several proteins and ultimately leading to the organisation of the viral contents of the virion [14]. Several steps of HIV's replication are targeted by ART (see **section 1.4**).

1.3.3 Natural History of HIV Infection

HIV is transmitted through contact with infected bodily fluids, including vaginal secretions and semen (during sexual contact), blood (for example during blood transfusions or from sharing contaminated needles during injection drug use), as well as breast milk (during breast feeding).

The most probable route of HIV infection is through dendritic cells, one of the immune cells responsible for antigen presentation to lymphocytes, since they are widespread in mucosal linings (such as the mouth, vaginal walls). Transmission of HIV to a CD4+ T-cell (T-helper cell) occurs once an HIV-infected dendritic cell or a dendritic cell with surface-bound HIV travels to a draining lymph node. CD4+ T-cells are responsible for activating and directing other immune cells and are HIV's primary target since they are most permissive to infection [15]. A productive infection of CD4+ T-cells occurs within days to weeks of infection in the lymph node and is followed by a release of virus throughout the body at which time a high plasma viral load (pVL) is observed. This is termed primary HIV infection [15]. During this period, there is a sharp decline in CD4+ T-cells, an increase in CD8+ T-cells (T-killer cells which are responsible for inducing apoptosis in infected or tumour cells) and viral RNA is high (**Figure 1**). Peak viraemia subsides spontaneously after 2-4 weeks, after which there is a chronic asymptomatic phase

that may last for a decade or more but this varies considerably among individuals. Although this phase is clinically silent, there is continual viral replication and a slow linear decrease in CD4+ T-cells counts which partially explains the decrease in immune function observed in patients with advanced HIV disease [15, 16].

AIDS is diagnosed by the Centre for Disease Control and Prevention (CDC) classification system once CD4+ T-cell counts are below 200 cells/ μl [17] and it is at this stage that most AIDS-defining illnesses occur [18]. Without treatment, in resource-limited settings, patients have a median survival of 6-19 months after AIDS diagnosis [19].

Figure 1: Natural history of HIV infection

During primary infection which lasts 2 – 4 weeks, HIV viral load rapidly increases while CD4+ cell counts decline sharply and CD8+ cell counts increase. After, a partial control over viral replication occurs and a chronic asymptomatic period follows for 2 – 15 years depending on the individual. However, there is continual viral replication which results in a slow decline in CD4+ T-cells and which partially explains the decrease in immune function seen in patients advanced HIV disease [15, 16]. Before the symptomatic phase, the CD8+ T-cell count remains relatively stable but decreases rapidly afterwards, most likely due to the loss of the CD4+ T-cells [16]. When CD4+ cell count is <200 cells/ μl , AIDS is diagnosed and death is imminent without treatment [17].

1.4 Highly Active Antiretroviral Therapy (HAART)

In 1987, zidovudine (AZT), the first HIV antiretroviral (ARV) drug, a nucleoside analogue reverse transcriptase inhibitor (NRTI) became available on the market. In the early 1990's, it was followed by

several other NRTIs. Initially, these drugs decreased viral load and increased patient quality of life; however, quickly, HIV resistance developed leading to treatment failure. The arrival of the first protease inhibitor (PI) in 1995 led to, highly active antiretroviral therapy (HAART) in 1996 and with it, a dramatic decline in HIV mortality rates was observed. In the late 1990's a third class of ARVs was developed called the non-NRTIs (NNRTIs) [20].

HAART is a combination of several HIV ARV drugs that when taken together, lower HIV pVL to undetectable levels, restore immune function by increasing the CD4+ T-cell counts [21] and reduce the risk of developing HIV resistance [20]. Currently, the most common HAART combination is 3 or more drugs from 2 different classes that include: 2 NRTIs and 1 PI or 1 NNRTI. With the advent of HAART, HIV viral replication is virtually inhibited and as a result, the chronic asymptomatic phase can be prolonged significantly. HAART has increased the life expectancy of HIV-infected persons which is still continuing to increase. Individuals who acquire HIV at age 20 and treated with HAART in 2005 could expect to live for another 49 instead of 36 years in 1996 (i.e. an additional 13 years) [22]. Because of HIV viral genome integration within host nuclear DNA (nDNA), the virus persists inside the host cells. Therefore, ART is a treatment but not a cure. As of 2009, the WHO recommends that HAART be initiated if CD4+ T-cell counts are $<350/\mu\text{l}$. The recommendations are different if there are co-infections or if a woman is pregnant [23].

At present, there are more than 20 ARVs available spanning 6 classes: 1. Entry inhibitors, 2. Fusion inhibitors, 3. Integrase inhibitors, 4. PIs, 5. NRTIs and the closely related nucleotide reverse transcriptase inhibitors (NtRTIs) and 6. NNRTIs. Each of these targets a different step of the HIV replication cycle. In this thesis, the focus will be on the latter 3 classes as they are currently the most relevant to HIV-infected pregnant women.

1.4.1 PIs

PIs act by binding to the viral protease and subsequently blocking the cleavage of the viral polyproteins into its constituent proteins and hence maturation of the virions [24]. Some of the frequently used PIs are: indinavir (IDV), saquinavir (SQV), lopinavir (LPV), ritonavir (RTV), atazanavir (ATV) and nelfinavir (NFV).

1.4.2 NRTIs and NtRTIs

Although NRTIs and NtRTIs are slightly structurally different in that NtRTIs are monophosphorylated and NRTIs are unphosphorylated, they both act via the same mechanism i.e. chain termination. Both NRTIs and NtRTIs are modified nucleosides which require intracellular (either cytoplasmic or mitochondrial) phosphorylation by kinases to become triphosphorylated and active. NtRTIs only need to be biphosphorylated, therefore, unlike NRTIs, they bypass the first phosphorylation step which is usually rate-limiting. They directly inhibit reverse transcriptase by causing premature chain termination during viral DNA replication. This is because they lack the 3'-hydroxyl group (3'-OH) on the deoxyribose that is needed to bind to the phosphate group of the next nucleoside triphosphate [25]. Some of the commonly-used NRTIs are: AZT, lamivudine (3TC), stavudine (d4T), didanosine (ddI), abacavir (ABC) and emtricitabine (FTC). The only NtRTI in use today is tenofovir (TDF) (**Figure 2**).

1.4.3 NNRTIs

NNRTIs are non-competitive inhibitors of RT; they bind to a location on the RT that is close to the NRTI binding site. In doing so, they deform the RT polymerase active site and, as a result, affect the polymerisation step of HIV DNA replication. In contrast to NRTIs, NNRTIs do not require cellular activation i.e. intracellular phosphorylation to inhibit RT [25]. Examples of NNRTIs that are frequently used are: efavirenz (EFV) and nevirapine (NVP).

**Thymidine (T)
Analog**

**Cytidine (C)
Analog**

**Guanosine (G)
Analog**

**Adenosine (A)
Analog**

Figure 2: Structures of common nucleoside reverse transcriptase inhibitors (NRTIs) and one nucleotide reverse transcriptase inhibitor (NtRTI). NRTIs and NtRTIs are modified nucleosides and nucleotides respectively. They are DNA chain terminators because they lack a 3'-OH group on the deoxyribose and therefore inhibit HIV replication. They are intracellularly phosphorylated before becoming active metabolites [25].

1.5 HIV and Pregnancy

1.5.1 Overview

MTCT of HIV can occur *in utero* (intrauterine transmission), during labour and delivery (intrapartum transmission) or postnatally through breastfeeding. Without any interventions, the MTCT rate in breastfeeding populations is ~35% and ~20% in those who do not breastfeed [26]. Some of the proposed mechanisms for intrapartum transmission are direct contact of foetal skin or mucous membranes with infectious maternal blood or genital secretions during labour, ingestion of virus from these secretions and ascending infections to the amniotic fluid. Intrauterine transmission, on the other hand, may occur through transfusion of infected cells into the foetoplacental circulation due to placental tears or from progressive HIV infection through the placenta since some placental cells also express CD4 receptors [27]. Although there are some conflicting reports, it is thought that most transmission events occur intrapartum or towards the end of pregnancy. Rouzioux *et al.* estimated that one-third of infants exclusively bottle-fed were infected in the last 2 months of pregnancy and the remaining were infected during the intrapartum period [28]. Other evidence to support this is the lack of manifestations of HIV infection at birth and the fact that only 50% of children who are proven to be infected perinatally, test positive in the first week of life [3]. One of the most important factors that contribute to the risk of transmission is viral load during pregnancy and delivery; as maternal viraemia increases, so does the risk for transmission [3]. This association has led to the usage of ART during pregnancy to decrease viral load and, in turn, MTCT risk.

1.5.2 HAART during Pregnancy

In 1994, the AIDS Clinical Trials Group Protocol 076 (ACTG 076) published their results regarding the use of prophylactic AZT for the first time in asymptomatic non-breastfeeding pregnant women to decrease the risk of MTCT. The ACTG 076 was a multicentre, randomised, double-blind, placebo-controlled study

that showed a MTCT rate of 25.5% in the placebo group compared to a 8.3% in the AZT group, a highly significant decrease ($p = 0.00006$). The regimen for the AZT group consisted of:

1. AZT starting between 14 to 34 weeks of gestation until delivery
2. Intravenous AZT during labour and delivery and
3. Oral liquid AZT for the infants for 6 weeks of life [29].

Since then several studies have been undertaken to examine other regimens that would be less complex and more easily executed in developing countries, however none have shown such extensive efficacy as the use of AZT-containing HAART. Although there have been no randomised controlled clinical trials containing a placebo group with HAART in pregnancy, several retro/prospective cohort studies have shown that MTCT rates are lower with HAART than with than AZT alone [30]. Cooper *et al.* showed decreased rates of MTCT with increasing number of ARVs administered: 20% with no treatment, 10.4% with AZT monotherapy, 3.8% with dual therapy and 1.2% with HAART [31]. This reason for lower MTCT rates with dual therapy and HAART compared to AZT monotherapy may be due to AZT having limited impact on viral load despite its dramatic effect in decreasing MTCT risk [32]. Therefore, combining other ARVs with AZT that have more viral load lowering capabilities has an even greater effect on reducing MTCT risks; since it is well documented that MTCT risk is positively correlated with viral load [33]. With this in mind, HAART administered during pregnancy has been adopted as the standard of care in developed countries including Canada, even though there is limited data on safety, tolerability and efficacy. In Canada, the management guidelines for HIV pregnancies are such that asymptomatic women who are not receiving HAART for their own health should begin HAART for PMTCT in their second or third trimester; this is to reduce the potential teratogenicity of the ARVs. HAART should be discontinued after delivery if, after considering the mother's immune status and HIV pVL, it is thought to be unnecessary. For those who conceive on HAART, the regimen may need revision to exclude ARVs that are known teratogens (e.g. EFV) [20, 34]. In addition, the US Public Health Service recommends that all pregnancy HAART regimens contain AZT. It should also be added on to regimens of pregnant women

who started treatment prior to pregnancy even if their optimal HAART regimen did not include it (unless there is severe AZT-related toxicity or AZT resistance) [35]. In Canada, the recommended PMTCT regimen is 2 NRTIs (e.g. AZT and 3TC) and either a PI (e.g. NFV), or a NNRTI (e.g. NVP) for HIV-infected women who have not previously been exposed to ARVs. In addition to antenatal exposure, it is recommended that HIV-infected pregnant women receive IV AZT during labour and delivery and that infants receive oral AZT for the first 6 weeks of life [34].

1.5.3 Prevention of Mother-to-Child Transmission in Low-Income Settings

At this time, in resource-limited settings, HAART is not readily available due to its high costs. Therefore, simpler ART regimens have been formulated, although the efficacy at PMTCT are lower than that of HAART.

Between 1996 – 2000, Petra, a randomised, double-blind, placebo-controlled trial held in South Africa, Uganda and Tanzania was conducted. It demonstrated that despite breastfeeding, a short-course antepartum (starting at 36 weeks' gestation) with intrapartum, as well as 7 days' maternal and infant postpartum regimen of AZT and 3TC was able to decrease transmission rate from 22% (in the placebo group) to 15% [36].

However, with similar efficacy as with the Petra regimen, the most basic and simple regimen that can be administered to pregnant women with HIV is a single-dose of NVP during labour and another in their infant after delivery. It is thought that this regimen alone has prevented over 30,000 HIV infections in children between 2004 – 2005 [37]. Between 1997 - 1999, the HIVNET 012 trial conducted in Uganda with breastfeeding HIV-infected mothers found that the transmission risk with this drug regimen alone compared to intrapartum AZT and 7 days' infant AZT administration was 13.1% vs. 25.1% [38]. However, the risks of developing resistance to NVP in the mothers are very high (up to 60-89% of women). If feasible, a combination of single-dose NVP and a short-course of AZT or AZT/3TC should be administered as it is more effective [39].

1.5.4 Outcome of Pregnancies with HIV-Infection

As the HIV/AIDS epidemic continues, an increasing number of children will be spared from HIV infection due to the widening availability of HAART. As such, it is even more imperative that we be aware of the potential short- and long-term consequences of *in utero* HAART exposure in HIV-uninfected children.

Although MTCT risk has been dramatically reduced with the introduction of ARVs and even more so with HAART in pregnancy, there are numerous accounts of observed toxicity not only in HAART-treated patients but also in children exposed during pregnancy [40, 41]. Much of this toxicity has been attributed to NRTI usage affecting mitochondrial biology and as such, has been termed mitochondrial toxicity. Mitochondrial toxicity and, more specifically, the mechanisms by which this occurs will be discussed in detail in the section on mitochondria (**section 1.6**).

Antenatal and post-natal exposure in infants has been associated with various laboratory findings such as persistent decreases in lymphocytes, neutrophils and platelets, as well as anaemia. However, most of these findings did not produce any clinical symptoms [42]. Of note, a significant number of ART-exposed infants have transient hyperlactataemia but it is normally asymptomatic. In a few cases, the more serious lactic acidosis occurs [40]. Lactic acidosis can sometimes be fatal albeit rarely (1.3 cases per 1000 patient-years of NRTI-exposure) [43]. Nonetheless, no deaths in infants have been reported from *in utero* ART-exposure. Fatal lactic acidosis, on the other hand, has occurred in pregnant women receiving a combination of ddI and d4T [44]. Both hyperlactataemia and lactic acidosis are manifestations of inadequate aerobic respiration even under normal non-hypoxic conditions. This is due to mitochondrial dysfunction, whereby lactate metabolism is prevented which leads to a build-up of lactic acid [45].

Rarely, other more serious outcomes have resulted such as neuropathy, cardiomyopathy, myopathy, lipoatrophy, pancreatitis and bone marrow suppression. In a large French cohort study of ART-exposed uninfected infants (N = 4392), 2 (0.045%) infants succumbed to encephalopathy [40]. Further analyses found that there was a 0.26% incidence of symptoms similar to those observed in infants with

congenital mitochondrial disease. This incidence was much higher than what is found in the general population of 0.01% [46].

Of note, in a follow-up study that lasted up to 5.6 years of the infants in the ACTG 076 trial, there was a lack of remarkable long-term effects in the AZT-exposed infants (N = 122) compared to the placebo group (N = 112). Several parameters were studied: growth, immunological and cognitive function, neurodevelopment, cancers and deaths. One exposed infant had mild cardiomyopathy at 48 months, however asymptomatic [47].

Several studies have found increased mitochondrial DNA (mtDNA) levels in infants ART-exposed *in utero* [48-50]. In a study by Côté *et al.*, higher levels of mtDNA were found at birth in infants exposed to ART *in utero*, compared to control infants born to HIV-uninfected women; although the trend did not reach statistical significance. There was a further increase during the postpartum prophylaxis period ($P = 0.001$) and this effect persisted up to 8 months when the study ended. With this, there was also a decrease in mtRNA levels at birth in the exposed infants however, this difference later disappeared. The authors postulated that the reduced levels of mtRNA expression may be due to dysregulation of gene expression or a change in mtRNA half-life from ART-exposure. The increase in mtDNA levels was suggested to be a consequence of mitochondrial adaptation in response to mitochondrial dysfunction and/or the low levels of mtRNA [49].

Although there are rare risks of serious adverse effects of *in utero* and postpartum exposure to ARVs, these must be put in context of the clear remarkable benefits of PMTCT.

1.6 Mitochondria

1.6.1 Overview

Mitochondria (plural for mitochondrion), are ancient bacterial symbionts and double membrane-bound organelles found in almost all eukaryotic cells. They have been described as the “power house” of the cell since they are responsible for most of a cell’s adenosine triphosphate (ATP) production through

oxidative phosphorylation (OXPHOS), the main source of energy in a cell. Moreover, mitochondria are also involved in heat production, cell cycle control, apoptosis, intracellular signalling, as well as metabolism of amino acids, lipids, cholesterol, steroids and nucleotides [51]. The density of mitochondria increases with the cell's energy consumption; for example, myocytes (muscle cells) and neurones which have high energy requirements have a higher density of mitochondria than peripheral blood mononuclear cells (PBMCs, i.e. lymphocytes, monocyte/macrophages) [52].

The mitochondria's double-membrane forms 4 distinct compartments: the outer membrane, the intermembrane space, the inner membrane, and the matrix. The inner membrane is the location of the electron transport chain (ETC) where ATP production occurs. This chain comprises 2 mobile electron carriers (ubiquinone and cytochrome c), as well as 5 enzyme complexes (complex I – V) that are individually composed of several subunits. These complexes are proton pumps that act in concert with the mobile electron carriers to shuttle electrons through the inner membrane and simultaneously pump protons (H^+) out of the mitochondrial matrix into the intermembrane space; creating a proton gradient. The mitochondrial membrane potential that results is then exploited by complex V; the energy derived from the return of protons back to the matrix is used by the complex's ATP synthase to synthesise ATP from adenosine diphosphate (ADP) and inorganic phosphate (P_i). OXPHOS refers to the shuttling of electrons, the export of protons, the reduction of O_2 to water at complex IV and the production of ATP; all of which are required in aerobic respiration (**Figure 3**) [51].

Figure 3: An overview of the electron transport chain

The electron transport chain (ETC) is composed of 5 enzyme complexes (complex I – V), each composed of several subunits, as well as 2 mobile electron carriers (ubiquinone and cytochrome c). The complexes and carriers act together to produce a proton gradient by shuttling electrons and simultaneously pumping protons (H^+) from the mitochondrial matrix into the intermembrane space. The protons flow back down their potential gradient through complex V into the matrix, as a result, ATP is generated from ADP + P_i . Reactive oxygen species (ROS) (e.g. $O_2^{\cdot -}$) are toxic by-products of the ETC and are produced at complex I and III [53].

1.6.2 Mitochondrial DNA (MtDNA) and Heteroplasmy

Mitochondria have an independent genome separate from that of the nucleus and it is referred to as mitochondrial DNA (mtDNA). Of the total genomic DNA in a cell, 0.1-2% is mtDNA [54]. This genome is 16,569 bp in size, circular and double-stranded. It encodes 37 genes altogether, including 22 transfer RNAs (tRNAs), 2 ribosomal RNAs (rRNAs, 16S and 12S) and 13 polypeptides (**Figure 4**).

Figure 4: Map of the human mitochondrial genome

The mitochondrial genome is 16,569 bp in size, circular, double-stranded DNA and is composed of 37 genes (13 polypeptides of the ETC, 22 tRNAs, 2 rRNAs). The inner strand is the light (L)-strand rich in cytosine and the outer, the heavy (H)-strand rich in guanine (the heavier bases). The L-strand encodes 9 genes and the H-strand 28. The genes encoding subunits of complex I (ND1 – ND6 and ND4L) are shown in blue; complex IV (COI – COIII) are shown in green; cytochrome *b* of complex III is shown in orange and subunits of complex V (ATPase 6/8) are shown in dark pink. The 2 rRNAs (16s and 12s) are in pale pink. The tRNAs for each amino acid are in gray. The D-loop is the regulatory region containing the replication and transcription initiation sites.

Even though there may be up to 1500 proteins in a mitochondrion, the vast majority of which are encoded by nuclear DNA (nDNA), 4 of the 5 ETC complexes contain subunits that are mtDNA-encoded, without them the ETC would be dysfunctional (**Table 1**) [55].

Table 1: Complexes forming the ETC

Complex		# of subunits		
Name	#	mtDNA encoded	nDNA encoded	Total
NADH Ubiquinone Reductase	I	7	38	45
Succinate Dehydrogenase	II	0	4	4
Ubiquinol Cytochrome c Reductase	III	1	10	11
Cytochrome c oxidase	IV	3	10	13
ATP Synthase	V	2	12	14

The ETC is composed of subunits encoded both by nDNA and mtDNA [55].

There are several marked differences between mtDNA and nDNA (summarised in **Table 2**):

1. *Polyplody*- Each cell can have hundreds and up to thousands of copies of the mitochondrial genome, varying with the energetic requirements of the cell [56] .
2. *Maternal inheritance*-mitochondrial genetics do not follow the standard Mendelian inheritance patterns because during fertilisation, the few mitochondria that originate from the sperm are destroyed in the embryo. As a result, only the mtDNA from the oocyte's mitochondria are transmitted to the child [57]. Although, a woman may transmit her mtDNA to both male and female children, only her daughters will further transmit the mtDNA to their progeny [58].
3. *Heteroplasmy and threshold effect*-heteroplasmy is the co-existence of wild-type and mutated mitochondrial genomes at the level of a mitochondrion, cell, tissue or organism (**Figure 5**). This phenomenon occurs because of the polyplodic nature of mtDNA genomes. Homoplasmy, on the other hand, is when all mitochondrial genomes are identical. Whether the presence of a mtDNA mutation causes mitochondrial dysfunction and produces a clinical phenotype (or disease), depends on the relative proportion of the mutated genomes compared to wild-type and on the energetic demands of the cells, the so-called “threshold effect” [59]. Once the threshold level has been reached, ATP production falls below the requirements of the cell.

Figure 5: Diagram of heteroplasmy at the mitochondrion and cellular level

Heteroplasmy is the co-existence of wild-type and mutated mtDNA genomes at the mitochondrion, cellular, tissue or organism level. On the left, is heteroplasmy at the mitochondrion level since 20% of all the genomes are mutated (depicted by the dot). On the right, is heteroplasmy at the cellular level since the top mitochondrion only harbours mutated genomes while the bottom mitochondrion only has wild-type genomes.

4. *Mitotic segregation*-during cell division, there is a random distribution of mutated and wild-type genomes into daughter cells. Consequently there may be a shift in the proportion of mutated mtDNA in a daughter cell compared to the parental cell and this may also cause a shift in the phenotype. This phenomenon explains the fact that certain patients with a mtDNA mutation may develop a mtDNA-related disorder over time or may improve [58].
5. *Mitochondrial genetic bottleneck*- A mother with heteroplasmy will transmit a varying proportion of mutated mtDNA genomes to her progeny as long as they are present in her germline. During early oogenesis, there is a significant reduction in mtDNA content in the germline (the bottleneck). This is a random process that may cause a “sampling effect” and subsequently modify the levels of heteroplasmy. Therefore, an unaffected heteroplasmic mother with a deleterious mtDNA mutation in her germline can have offspring that are severely, moderately, mildly or unaffected [60].
6. *Organisation of genetic information*-Unlike nDNA, mtDNA does not have introns; there are very few to no intergenic sequences and some genes overlap. Moreover, mitochondria do not have

histones but instead organise mtDNA in nucleoid structures; whereby several mtDNA molecules are intertwined with numerous proteins which localise to the matrix surface of the mitochondrial innermembrane [54]. Replication and transcription is regulated by the D-loop, the non-coding control region (**Figure 4**).

7. *MtDNA replication*-It is solely executed by DNA polymerase γ (POLG). POLG is an RNA-dependent DNA polymerase composed of 2 subunits: 1. POL γ A-a larger catalytic subunit with DNA polymerase, 3' $\rightarrow$ 5' exonuclease (for proofreading ability) and a lyase (required for base excision repair) 2. PolyB-a smaller subunit that enhances DNA binding and processivity [61]. The *in vitro* fidelity of POLG is ~ 1 in 100,000 bp which is considered to be high. The high fidelity can be attributed to high nucleotide selectivity, slow mismatch extension, and efficient proofreading from the exonuclease [62]. Furthermore, mtDNA replication is independent of cell cycle and nDNA replication, this means that even in cells which are terminally differentiated and non-dividing, such as myocytes and nerve cells, mtDNA is still being replicated [63]. There is also evidence of high and random turnover of individual mtDNA molecules that may occur several times per cell cycle [64]. Nonetheless, the copy number of mtDNA within a cell remains relatively stable, a phenomenon described as “relaxed replication” [65]. Although, mtDNA replication occurs independently of nDNA replication, mtDNA replication is controlled by the nucleus since all of the trans-acting factors associated with mtDNA replication are nuclear-encoded [66].
8. *Transcription and translation*-Transcription initiation of mtDNA occurs at the promoters on the light (P_L) and heavy (P_H) strand, both of which are also located in the D-loop close to the origin of replication on the H-strand (O_H) (**Figure 4**). Like in bacteria, transcription generates polycistronic precursor RNA (i.e. one RNA molecule that codes for several gene products), which once processed, produces individual tRNA, rRNAs and mRNA molecules. Translation of mRNAs occurs

at mitochondrial ribosomes with slight variations on codon usage from the universal genetic codons, for example AUA codes for methionine instead of isoleucine [67].

Table 2: Comparison of human nuclear and mitochondrial genome

Characteristic	Mitochondrial Genome	Nuclear Genome
Size	16,569 bp	$\sim 3.3 \times 10^9$ bp
Number of DNA molecules/cell	Polyploidy (varies by energetic requirement)	23 chromosomes in haploid cells, 46 in diploid cells
Number of genes encoded	37 genes (13 polypeptides, 22 tRNAs, 2 rRNAs)	$\sim 20,000 - 30,000$
Gene density	$\sim 1/450$ bp	$\sim 1/40,000$ bp
Introns	Absent	Frequently found in most genes
Percentage of coding DNA	$\sim 93\%$	$\sim 3\%$
Codon usage	AUA codes for methionine; TGA codes tryptophan; AGA and AGG = stop codons	Universal genetic code
Mode of inheritance	Exclusively maternal	Mendelian inheritance for autosomes and the X chromosome; paternal inheritance of the Y chromosome
Polymerases for DNA replication	γ (POLG)	δ and ϵ
DNA repair mechanisms [68]	Less extensive	More sophisticated
Transcription	Polycistronic transcription of all genes	Genes are transcribed individually

Ref [67]

1.6.3 Mitochondrial Disorders

Mitochondrial disorders result from mutations arising in either nuclear- or mitochondrial-encoded genes that cause a mitochondrial respiratory chain dysfunction. Depending on the mitochondrial disorder, a single organ or multiple organs may be affected. These disorders often have neurological or myopathic features such as dementia, ptosis (drooping of the upper eyelid), exercise intolerance and cardiomyopathy, which usually worsen with age. The age of onset of these disorders is variable; however, a nDNA mutation generally results in childhood-onset, while a disorder due to a mtDNA mutation (primary or secondary to a nDNA mutation) generally presents later in life [69]. Many mitochondrial disorders are caused by nDNA mutations since the majority of the ETC subunits are nuclear-encoded. MtDNA replication, maintenance, repair, translation and transcription are all

controlled by the nucleus [52]. Nonetheless, mutations in mtDNA can also have a dramatic effect due to the unusual characteristics of mtDNA with respect to nDNA as described above (**section 1.6.2**). A mtDNA mutation can either be a deletion, point mutation or rearrangement (this latter type of mutation will not be further discussed in the context of this thesis). The severity of a mitochondrial disorder caused by a mtDNA mutation will depend on the extent of its presence in the particular afflicted organ and the threshold effect (see **section 1.6.2**).

1.6.3.1 Disease-causing MtDNA Deletions

Deletions can vary in size from a single base to several kilobases (kb) and can occur throughout the genome. As described by Meissner *et al.*, large deletions probably occur because of polymerase slippage, homologous recombination, double-strand breaks, inefficient repair mechanisms or mutations in the helicase, Twinkle [70]. Deletions are implicated in clinically complex disorders such as Kearns-Sayre syndrome (KSS), Pearson's syndrome, and chronic progressive external ophthalmoplegia (CPEO). The primary clinical feature of both KSS and CPEO is progressive ophthalmoplegia (paralysis or weakness of the extraocular muscles that control eye movement). A deletion that has been observed in KSS, CPEO and Pearson's syndrome is the "common deletion" in mtDNA (4977 bp) [70]. This deletion removes the genes or parts of the genes encoding for subunits of complex I, IV and V, which can lead to dysfunction of mitochondrial OXPHOS [71].

1.6.3.2 Disease-causing MtDNA Point Mutations

As of February 2010, 60 point mutations are confirmed to be pathogenic and over 200 more have been reported in association with mitochondrial disorders in MITOMAP (Human Mitochondrial Genome Database) (<http://mitomap.org>). There are many specific mtDNA point mutations that cause mitochondrial disorders. For example, the A to G mutation at mtDNA nucleotide position (nt) nt3243 (A3243G) of the tRNA^{Leu}, is the cause of 80% of cases of the multisystemic and progressive neurodegenerative disorder, MELAS (mitochondrial encephalomyopathy, lactic acidosis, stroke-like

episodes). This disorder also has other features such as hearing loss, cardiac disease and diabetes mellitus [63].

1.6.4 Random MtDNA Mutation Accumulation and Oxidative Stress

mtDNA is evolving at rates 10 times higher than nDNA [72]; this is because mtDNA is more prone to mutation accumulation than nDNA for several reasons. Firstly, there are no histones to protect mtDNA [73]. Secondly, mitochondria have less extensive DNA repair mechanisms than those present in the nucleus [68].

Lastly and probably most importantly, mtDNA is subject to more oxidative damage than nDNA because it is in close proximity to the OXPHOS process in the cell. As part of normal OXPHOS, approximately 1-2% of the oxygen consumed is incompletely reduced to water; producing $O_2^{\cdot-}$, a reactive oxygen species (ROS) (a free radical that contains an oxygen atom) that can then be converted to H_2O_2 and $OH^{\cdot}$ (another ROS). It is commonly believed that most $O_2^{\cdot-}$ production occurs at complex I and III (**Figure 3**). ROS is considered to be a toxic by-product of aerobic respiration; in states of environmental stresses (e.g. UV radiation, drug exposure), an excessive accumulation of ROS may occur resulting in uncontrolled oxidation of cellular components, a phenomenon referred to as oxidative stress [74]. OXPHOS dysfunction due to oxidative stress leads to energy loss (i.e. reduction in ATP production) [75].

There is evidence of higher oxidative damage in mtDNA compared to nDNA, as illustrated by a greater presence of the biomarkers 8-oxoguanine (8-oxoG) and 8-hydroxy-2'-guanosine (8OHdG) in mtDNA compared to nDNA. These markers are oxidatively modified guanine bases; both of which are repair products of oxidised guanine lesions [76]. These molecules can cause G:C to T:A transversion mutations since they can pair with adenine and cytosine with almost equal affinity [63].

Kang *et al.* and others believe that random mtDNA mutations may accumulate with age and cause mitochondrial dysfunction; eventually leading to cellular dysfunction [63]. Any mutation that arises in mtDNA will more likely have a phenotypic effect than in nDNA since 92.7% of the mtDNA is coding and

the non-coding regions have a regulatory function. This theory is supported by recent findings in a case series studied in the Côté laboratory that points towards the possibility of the existence of acquired mitochondrial disease. It was hypothesised that the development of late-adulthood onset CPEO in an HIV-infected patient was due to additive exposure to ART, as well as HIV that unmasked and amplified a previously subclinical and pre-existing mtDNA deletion [77].

1.6.5 Aging

Aging is the deterioration of an organism's physiological function with the increasing risk of death. It is thought that aging is linked to mitochondria and their loss of function over time due to accumulation of mtDNA mutations or, in other words, acquisition of heteroplasmy from environmental exposures (e.g. UV radiation) or toxins (e.g. ROS). This is the basis of what is called the mitochondrial theory of aging. As ROS induces mtDNA mutations, mitochondrial dysfunction due to the mutations further increases ROS production and, in turn, a "vicious cycle" ensues. With the loss of mitochondrial function, there is a decrease in energy supply and an increased risk of apoptosis (which can be activated by ROS production) **(Figure 6)** [56].

Figure 6: The “vicious cycle” between ROS and mitochondrial dysfunction

ROS (reactive oxygen species) produced from the ETC can oxidatively modify mtDNA bases, that can cause mtDNA mutations which sets-up the “vicious cycle” between ROS and mitochondrial dysfunction. ROS can sometimes lead to apoptosis of the cell [56, 78, 79].

Many studies in brain, liver, as well as skeletal and cardiac muscle have found that oxidative damage, especially 8-oxoG, accumulates in mtDNA in an age-dependent manner. As Druzhyna *et al.* suggest, the persistence of oxidative damage would be expected to cause high mtDNA mutation rates [54]. The common deletion has also been associated with aging. It was detected with high frequencies in adult brains and hearts, in contrast, to young and foetal tissues where none were observed [80]. Michikawa *et al.* found that there were many mutations in fibroblasts of normal old individuals (>65 years) compared to none in the younger healthy individuals at specific positions (e.g. T414G) that are related to replication in the D-loop. Further evidence of age-associated accumulation of mtDNA mutations was demonstrated when samples taken at two different time points were obtained for the same individuals and that an increased number of mutations was found in the later sample [81].

1.6.5.1 Aging in the Mouse Model

Although many human observational studies have shown there is an association between mtDNA mutations or ROS and aging, these do not elucidate whether mtDNA mutations or ROS have a causative

link with aging. To shed light on this matter, several mouse models have been created that either increase or decrease ROS levels or directly increase mtDNA mutation accumulation.

A heterozygous knockout mouse for the gene encoding MnSOD (*Sod2*), the main antioxidant enzyme present in mitochondria, had a 50% decrease in MnSOD activity. Elevated oxidation of mtDNA and nDNA, as well as mitochondrial dysfunction in cardiomyocytes and liver were found in these mice. However, surprisingly, their lifespan seems to be unaffected [82, 83]. It is postulated that perhaps with further decrease in MnSOD activity there would be a more pronounced effect [84].

On the other hand, an increase in lifespan of ~5 months was observed with a mouse model that overexpressed mitochondrial catalase (mCAT mouse), an enzyme responsible for the decomposition of H₂O₂ into innocuous water and oxygen. They developed age-associated pathologies such as cataracts and cardiac pathology later in life than wild-type mice and had less oxidative damage [85].

Two “mutator” homozygous mouse models, *PolgA^{mut}/PolgA^{mut}* and D257 with proofreading deficient POLG were generated independently by Trifunovic *et al.* and Kujoth *et al.*, respectively that increased the rate of mtDNA mutation accumulation (~3-5x higher than wild-type mice) [86, 87]. Both “mutator” mouse models showed several signs of early aging, such as weight loss, kyphosis (abnormal curving of the spine), greying hair, alopecia (hair loss) and decreased lifespan (70% reduction in “mutator” compared to wild-type mice [84]). The *PolgA^{mut}/PolgA^{mut}* mouse also showed a decrease in respiratory chain function. However, what was quite unexpected was the absence of an increase in oxidative stress in D257 mice compared to wild-type mice. This phenomenon can perhaps be explained by the mutation load being so high such that the respiratory transport chain was unable to assemble properly; effectively eliminating ROS production altogether [84].

Even though these mice models have advanced our knowledge in the realm of aging, they have not fully explained or confirmed the link between oxidative stress, mtDNA mutations and aging. The mCAT mice indeed support the mitochondrial theory on aging but the heterozygous knockout MnSOD mice do not.

In addition, a criticism of the D257 and *PolgA^{mut}/PolgA^{mut}* mice is that although their lifespan was reduced, the phenotypes of aging that were observed are typical of humans but not of normal aging inbred mice such as kyphosis [88]. Nonetheless, the data on the “mutator” mice suggests that increased rates of mtDNA mutation accumulation accelerate the aging process.

1.6.5.2 MtDNA Mutations and Age-Associated Diseases

A growing number of studies are reporting links between mtDNA mutations with age-associated diseases such as Alzheimer’s, diabetes and cancers. These late-onset diseases probably develop progressively over time. As mutations arise with age, the mtDNA genomes that harbour these may replicate more often than wild-type genomes. MtDNA genomes harbouring deleterious mutations have been found to be preferentially, clonally amplified in post-mitotic cells [89]. This is because these mutations may confer a “replicative advantage” to mutated mtDNA genomes, which may be especially significant if the mutation occurs in the D-loop where replication is regulated [90, 91]. Another explanation for this “replicative advantage” phenomenon, may be that, increased mtDNA replication occurs as a compensatory mechanism for energy deficiency, arising from dysfunctional mitochondrial proteins as a result of, mtDNA mutations [89]. Eventually the mutated genomes will become predominant and a diseased phenotype may result as the threshold level is reached (**section 1.6.2**).

It is thought that late-onset, sporadic Alzheimer’s disease (AD), a progressive neurodegenerative disorder, is perhaps caused by a defect in OXPHOS and oxidative damage. Studies have shown that patients with AD have structurally abnormal mitochondria and that they have deficient levels of cytochrome c oxidase (complex IV) in the brain and elevated levels of oxidative damage. In approximately 5% of late-onset AD, a T4336C mutation in the tRNA^{Gln} is observed. Recently, Coskun *et al.* found that in brain tissue, all of the AD patients and a subset of those >80 years old analysed had a 63% and 130% rise in somatic mutations in the D-loop respectively, compared to age-matched controls. They also observed that 65% of AD brains had the T414G mutation (previously mentioned in **section 1.6.5**), while this mutation was absent in controls. Interestingly, they found that there was a high

frequency of mutation accumulation in the region that regulates L-strand transcription and H-strand replication. In conjunction with these results, they found that there was mtDNA depletion in AD brains and a decrease in L-strand ND6 mRNA needed for complex I assembly. All together, they believe these mutations may be the cause of sporadic, late-onset AD [92].

The A3243G mutation mentioned previously (**section 1.6.3.2**) with regards to MELAS has also been found in pancreatic β -cells in 1-2% of type 2 diabetes mellitus patients. It is hypothesised that this mutation is the cause of their disease. It is suggested that these patients have a subtype of MELAS that is restricted to pancreatic β -cells [63].

Various mtDNA mutations have been found in several types of cancers such bladder, brain, breast and colon cancer [93]. More specifically, these mutations were base substitutions with homoplasmic characteristics and were most frequently observed in the D-loop [54]. Lu *et al.* suggest that mtDNA mutation accumulation might initiate tumorigenesis. MtDNA genomes harbouring mutations might increase ROS that damages not only the mitochondria themselves, but also nDNA that could in turn activate oncogenic pathways of the cells [94]. This theory is supported by an experiment carried out by Petros *et al.* They generated 2 cytoplasmic hybrids (or cybrids, which are hybrid cells that contain nDNA from one source and mtDNA from another) with a PC3 prostate cancer cell line. The first combination was PC3 with mtDNA harbouring a pathogenic mutation (T8993G) that increases ROS production and the second, was PC3 with wild-type mtDNA (T8993T). They found that this pathogenic mtDNA mutation caused a 7x increase in tumour size compared to cybrids with the wild-type mtDNA genome which showed no tumour growth [95].

1.6.6 NRTIs and Mitochondrial Toxicity

After physicians began to administer NRTIs for HIV treatment, it soon became clear that patients were experiencing adverse effects from these drugs which stemmed from mitochondrial dysfunction.

Mitochondrial dysfunction induced by NRTIs is called mitochondrial toxicity. This was first noted by

Dalakas *et al.* in 1990, whereby HIV-infected patients treated with AZT had myopathy. Upon muscle biopsy, characteristic ‘ragged red fibres’, indicative of abnormal mitochondria, were found in myocytes [96].

1.6.6.1 Signs and Symptoms of Mitochondrial Toxicity

There is a wide variety of signs and symptoms associated with different NRTI treatments (**Table 3**), notably hyperlactataemia and rarely lactic acidosis (mentioned before in **section 1.5.4**), as well as fatigue, cardiomyopathies, myopathies, hepatic steatosis (fatty liver), lipodystrophy (abnormal fat distribution) and neuropathy. Many of these signs and symptoms form clinical syndromes similar to those observed in mitochondrial genetic disorders [97]. Some of these side effects can lead to severe morbidity and can sometimes be fatal [25].

Table 3: Relative frequency of clinical symptoms with various NRTIs

Organ Side Effects	Associated NRTIs
Neuromuscular	
Peripheral polyneuropathy	ddI, d4T
Myopathy	AZT
Cardiomyopathy	ddI, AZT
Gastrointestinal	
Steatosis	All NRTIs (d4T >> other NRTIs)
Pancreatitis	ddI, d4T
Haematological	
Anaemia, neutropaenia	AZT
Nephrological	
Proximal tubule dysfunction, hypophosphataemia	TDF
Metabolic + adiropocytic	
Lipoatrophy	d4T >> ddI > AZT > ABC, TDF, 3TC, FTC
Hyperlactataemia, lactic acidosis	d4T >> ddI > AZT > ABC, TDF, 3TC, FTC
Hyperlipidaemia	d4T

> denotes clinically more frequent and/or more severe [98]

1.6.6.2 Mechanisms of NRTI Mitochondrial Toxicity

Several mechanisms by which NRTIs cause mitochondrial toxicity have been proposed, these include but are not limited to:

1. NRTIs can cause an imbalance in the endogenous nucleotide base pool within mitochondria.

This imbalance can ultimately decrease the fidelity of POLG [99]. The imbalance may occur through many routes, for example:

- a. By simply being present in the mitochondria.
- b. NRTIs have the ability to inhibit the enzymes responsible for nucleoside phosphorylation and, therefore, decrease the endogenous deoxyribonucleoside triphosphate (dNTP) pools.
- c. Transportation of phosphorylated NRTIs may be favoured over endogenous nucleotides [78].

2. NRTIs can directly affect POLG; this can also occur through several routes:

- a. NRTIs are competitive inhibitors of POLG since they compete with endogenous dNTPs at POLG's active site. In this case, the NRTI is not incorporated into the mtDNA elongating strand during mtDNA replication.
- b. Chain termination can occur through incorporation of the triphosphorylated NRTIs (i.e. NRTI-TP) into the mtDNA elongating strand by POLG.

Both a. and b. can lead to mtDNA depletion. In addition, b. can lead to mtDNA deletions.

- c. NRTIs can inhibit the exonuclease activity of POLG which leads to diminished fidelity of POLG.
- d. The efficacy of removal of an NRTI once it has been incorporated is lower than a naturally-occurring dNTP because NRTIs lack the 3'-OH group [100].

Both c. and d. can lead to mtDNA point mutations.

MtDNA depletion, mtDNA deletions and mtDNA point mutations can all result in a decrease in mtRNA and a reduction of mitochondrial protein synthesis or synthesis of defective proteins. Mitochondrial dysfunction or in other words, energy deprivation can ensue. Consequently, there would be an increase in ROS leading to more mtDNA mutations accumulating, adding to the "vicious cycle" of oxidative stress (Figure 6).

POLG has been shown to be particularly sensitive to inhibition by NRTIs in comparison to other polymerases; the relative effectiveness of inhibition by NRTIs is: HIV RT >> POLG > DNA polymerase- α = DNA polymerase- ϵ (both responsible for nDNA replication). The most potent NRTI inhibitors of POLG *in vitro* are the D-drugs (i.e. the dideoxy-NRTIs e.g. zalcitabine (ddC), ddI and d4T) followed by the much less potent 3TC > TDF > AZT > ABC. However, excision of 3'-terminal 3TC residues is 50% as efficient as natural 3'-termini; this is in contrast to AZT for which, although incorporation is unlikely, once incorporated it is inefficiently removed [79]. This may explain the relatively low adverse effects observed with 3TC in the clinical settings. Nevertheless, AZT's ability to inhibit POLG only moderately and its inefficient removal cannot solely explain the potent toxicity that has been reported with this drug (**Table 3**). *In vitro*, it was found that through chemical reduction, AZT and its phosphorylated metabolites could be converted into d4T-TP (one of the more toxic NRTIs). Indeed, recently, d4T-TP was detected in patients despite not being administered d4T but rather AZT [101, 102]. It is thought that AZT is converted to d4T-TP intracellularly [97].

1.6.6.3 NRTIs and MtDNA Mutation Accumulation

It is hypothesised that since NRTIs affect POLG activity, that fidelity of the POLG is diminished, ultimately leading to mtDNA mutations [75]. This hypothesis is the over arching hypothesis to this thesis. A study by Martin *et al.* in 2003, observed an accumulation of mtDNA mutations in individuals treated with NRTIs. They compared blood samples from 16 HIV-infected individuals before and after 6 - 77 months NRTI-treatment. They also compared them to 10 HIV-infected individuals but treatment-naïve controls. They found that, after treatment, mtDNA mutations (heteroplasmy) were observed in 5/16 treated individuals but there were no changes in the controls. Heteroplasmy was found in both coding and non-coding regions [103]. This type of study has not been repeated by others, although many propose that mtDNA mutations would “logically” follow as a result of NRTI exposure [75, 78, 79].

1.6.7 Total MtDNA Point Mutation Detection

Many different types of methods and strategies have been developed to study mtDNA mutations and heteroplasmy. Many of these are designed to screen for specific known point mutations such as polymerase chain reaction (PCR) restriction fragment length polymorphism (RFLP).

PCR RFLP is based on the fact that certain point mutations will result in the loss or gain of a restriction enzyme site. After PCR amplification of the appropriate segment of the mtDNA genome, the sample is digested with the use of a specific restriction enzyme that cleaves around the location of the point mutation of interest. The digested vs. undigested DNA is then separated by gel electrophoresis and the relative intensity of the bands are determined by ethidium bromide staining or radiography and compared to one another. This is the most commonly-used technique in the clinical diagnosis of known pathogenic point mutations such as A3243G [104]. The disadvantage of this technique is it requires the use of dangerous substances (i.e. ethidium bromide or radioactivity) and is not very sensitive; as such, it would be unable to detect very low levels of heteroplasmy/mutations.

For the purpose of exploring total mutation burden of unknown mutations, techniques such as PCR RFLP are too specific. Direct sequencing avoids the problems of specificity however, it has poor sensitivity. Instead methods for screening, such as post-PCR cloning, direct cloning and random mutation capture which have been previously used by others are needed.

1.6.7.1 Post-PCR Cloning

Post-PCR cloning involves PCR amplifying a target mtDNA sequence from a sample. The PCR products are then cloned into a bacterial plasmid vector. The cloned plasmid DNA is re-amplified from selected bacterial colonies and sequenced. The cloning step serves to separate individual mtDNA molecules from one another. Hence each plasmid sequence is considered to be from an individual PCR amplicon derived from one mtDNA molecule. By sequencing large numbers of clones per sample studied, a mtDNA mutation burden per base pair can be determined. A mutation is considered when a base pair change is

observed that is distinct from the consensus sequence of a sample, the consensus sequence being the average sequence all of the clones from one sample [105]. Several groups have used this strategy to study mtDNA: in blood [106], muscle [107], brain [92] and in the “mutator” mice described above (section 1.6.5.1) [86, 87]. This method is considered to be convenient and extremely little starting material is needed because of the PCR step [108]. However, this very step introduces artificial mutations; for example *PfuU* and *Taq* polymerase induce $\sim 10^{-6}$ and $\sim 10^{-4}$ mutations/bp/duplication, respectively [105].

1.6.7.2 Direct Cloning

Direct cloning involves separating mtDNA from nDNA and then performing a restriction enzyme digest to obtain the fragment of interest. These fragments are cloned into a bacterial plasmid vector or into λ phage and individual clones are then sequenced. Each of these clones is considered to be from one mtDNA genome. Proponents of this method believe that levels of artificial mutation induction are much lower than with the post-PCR cloning strategy since the use of polymerases for PCR is avoided. λ phage and bacteria generate as few as $\sim 10^{-7}$ and $\sim 10^{-9}$ errors/bp, respectively [105, 109]. Some of the drawbacks to this strategy are that it requires a considerable amount of starting material and the isolation of mtDNA is laborious [105].

1.6.7.3 Random Mutation Capture Assay

The random mutation capture assay was used to study mtDNA in 2007 by Vermulst *et al.* [110]. It is based on a real-time PCR strategy whereby relative quantification of mutant vs. wild-type mtDNA genomes is performed. MtDNA is first isolated and is digested by *TaqI*, a restriction enzyme that recognises a 4 bp sequence. *TaqI* cleaves wild-type genomes but if any of the 4 bp are mutated, the restriction site is destroyed. After digestion, multiplex real-time PCR is carried out whereby one set of primers binds to sequences nearby the restriction site to quantify all mtDNA genomes while another set binds across the restriction site which only amplifies mutant mtDNA genomes. The ratio of the two is used to determine the mutation rate. The advantage of this method is that it is cost-effective when

compared to sequencing [111]. However, a significant flaw, is that the mutation rate that is observed within the restriction site ($[4 \text{ bp} \times 29 \text{ restriction sites}] / 16568 \text{ bp}$, i.e. 0.7% of the mtDNA genome) may not necessarily reflect the mutation rate elsewhere in the genome.

With both direct and post-PCR cloning strategies, a relatively substantial region of mtDNA genome can be examined simultaneously. Essentially, the maximum length that can be studied is dictated by sequencing abilities. With cycle sequencing, an average 900 bp can be analysed with non difficult sequences [112].

With all 3 methods, the level of sensitivity to mutation load is much more substantial than PCR RFLP. In the case of the random mutation capture assay, the detection limit is said to be 1 mutation in 10^9 wild-type bases [110]. With post-PCR cloning the detection limit is dependent on the PCR polymerase used and its error rate. As for the direct cloning method, the detection limit lies between that of the random mutation capture assay and post-PCR cloning.

Each method has many advantages and disadvantages (**Table 4**) therefore, it is important to consider all aspects of these before implementing the use of a particular method for a study.

Table 4: Advantages and disadvantages of different methods/strategies to study mtDNA mutations/heteroplasmy

Method/Strategy	Advantages	Disadvantages
PCR RLFP	-Established method	-Sometimes uses dangerous substances (e.g. ethidium bromide or radioactivity) -Not very sensitive
Post-PCR cloning	-Little starting material needed -A large area of the mtDNA genome can be analysed at one time -Able to detect unknown mutations	-Use of PCR polymerase that can create artificial mutations -Involves sequencing (which can be costly)
Direct cloning	-A large area of the mtDNA genome can be analysed at one time -Able to detect unknown mutations -Avoids creation of artificial mutations since no PCR polymerase is used	-mtDNA isolation is required (which is labour intensive) -A large amount of starting material needed -Involves sequencing (which can be costly)
Random Capture Mutation Assay	-Avoids creation of artificial mutations since no PCR polymerase is used -Avoids sequencing (therefore more cost-effective) -Detection limit is very high	-Only 0.7% of the genome can be studied

2 MtDNA Mutation Burden Assay-Background Error Rate Determination, Method Validation and Optimisation

2.1 Introduction To quantify mtDNA point mutations in a given sample, an assay was developed based on a post-PCR cloning strategy and hereafter referred to as “mtDNA mutation burden assay” (MMBA). It involves cloning PCR products of a portion of the mtDNA D-loop from a total DNA extract (including both nDNA and mtDNA), sequencing these clones and examining each clone’s sequence for base pair changes compared to the sample’s consensus sequence². Each PCR product/clone can be considered a copy of a single mtDNA molecule [113]. Sequencing of cloned mtDNA allows for the detection of low level mtDNA point mutations that direct sequencing of a PCR product would not reveal.

The region assayed is within the D-loop of the mtDNA genome. The D-loop is the regulatory region of the genome responsible for replication and transcription initiation. It is of interest because it is considered to be evolutionarily the most variable region [114]. This variability may stem from the ageing process; whereby mtDNA point mutations arising over a lifetime in the germline of an individual may eventually lead to polymorphisms observed in a population [115]. Moreover, since the D-loop is a non-coding region, it is under less selective pressure than the coding regions. As a result, the detection of mutations would be more likely in the D-loop region than if the coding region were to be studied. Another reason for choosing this region is since it is not transcribed into RNA, it is less likely to have homologous pseudogenes throughout the nDNA genome, which would hinder the assay.

Here, several experiments were performed to establish PCR conditions that would optimise the fidelity of the assay (i.e. decrease background error/mutation rate). Various numbers of PCR amplification cycles were tested and two polymerase enzyme blends compared. Polymerase error rates vary because of differences in proofreading abilities and tendencies to incorporate incorrect nucleotides [116].

Furthermore, background error rates (BERs) were determined with differing lengths of a cytosine (C)

²A consensus sequence is a representation of a multiple sequence alignment of related sequences, in this case the cloned PCR products derived from a sample containing mtDNA.

mononucleotide repeat (called a C-tract) that lies within the region of the D-loop (at nt303 – 316) PCR amplified (**Figure 7**). This C-tract is composed of 12 – 18 cytosine bases interrupted by one thymine (T) at position nt310 (C₆₋₁₂T₁C₆); this tract is termed D310. There is inter-individual and intra-individual length variation within D310 [117]. There are also 3 other C-tracts within the D-loop but these are outside the region amplified by the assay, at nt16184 – 16193, nt456 – 463 and nt568 – 573, and these also show length variation [118]. Therefore, the presence of such a mononucleotide tract within our amplified fragment could not be avoided. PCR amplification errors, such as insertions and deletions, have previously been shown to occur with T mononucleotide and TA and CA dinucleotide repeats. It has been suggested that these findings can be extrapolated to all types of mononucleotide repeats [119].

The number of clones per subject needed to study mtDNA mutations to obtain a low and stable intrasample variation was also ascertained. Lastly, we determined if the sequencing reaction could also be a source of background error since a polymerase is also involved.

Figure 7: D-loop region of mtDNA amplified for the assay (nt16559 – 447)

HV = hypervariable segment, CSB = conserved sequence block, OH = H-Strand DNA synthesis initiation (₁ = primary site, ₂ = secondary site), TF = mitochondrial transcription binding factor site, 3/4H = mt3/4 H-strand control element, LSP = L-strand promoter, HPR = replication primer, D310 = C-tract (nt303-316)

Based on information gathered from www.mitomap.org, accessed 25 March 2010 and Wang *et al.* [120].

2.2 Methods and Materials

2.2.1 Samples

Control DNA samples were obtained from 2 individuals in the unexposed control group from the Pregnancy Study Cohort (this cohort will be described in detail in **section 3**). These individuals were a mother/infant pair. The mother (PR21C-M) showed a heteroplasmic C-tract whose length varied from 7 – 11Cs ($C_{7-11}TC_6$; this will be denoted as 7 – 11 Cs hereafter), while the infant (PR21C-I) had a fixed 8 C C-tract.

2.2.2 DNA Extraction

Total genomic DNA (i.e. nDNA and mtDNA) was extracted from 0.1 ml of whole blood using QIAamp® DNA Mini Kit (Qiagen) according to manufacturer protocol specifications.

2.2.3 MtDNA Mutation Burden Assay (MMBA)

2.2.3.1 PCR Amplification (1st PCR)

A 509-bp fragment in the D-loop region of the mtDNA genome was amplified with the primers MT16535F (5'-GCC CAC ACG TTC CCC TTA AAT AAG A-3') and MT474R (5'-AGT ATG GGA GTG RGA GGG RAA AA-3') (**Table 5**). The reverse primer contained 2 degenerate bases to accommodate the known single nucleotide polymorphisms (SNPs) found at these positions. Steps were taken to exclude the potential amplification of nuclear mitochondrial pseudogenes that would be a source of PCR contamination [121]. Firstly, this 509-bp fragment using the revised human (Cambridge) mtDNA sequence [122] was megaBLAST searched against all human genome assemblies. The reverse primer sequence was found not to be entirely complementary to positions in the human genome other than the mtDNA site of interest (the forward primer sequence, on the other hand, is complementary to several positions in the human genome). Also, the second best hit after the mitochondrial genome was a fragment on the chromosome (chr) 17, which had 6% gaps and only 82% identity. Furthermore, the reverse primer was only complementary to 7 bases on chr 17. It was, therefore, concluded that there would be either no contaminating amplification occurring or if there was, it would be easily detected

during sequence analysis. Secondly, annealing temperature was set at 60°C, to deter primer mismatch. When using Expand High Fidelity^{PLUS} Enzyme Blend (HiFi *Taq*) (Roche), a blend of *Taq* DNA polymerase and a proofreading protein, the 25 µl PCR mix contained 1 × Reaction buffer with 1.5 mM MgCl₂, 200 µM dNTPs (Invitrogen), 400 pM of each primer (IDTDNA), 2.5 µl of DNA extract, and 0.5 U of HiFi *Taq*. Amplifications conditions were 94°C/30 s, followed by varying numbers of cycles (15, 20, 25, 30, 35 and 45) of 94°C/15 s, 60°C/30 s, 72°C/30 s, with a final extension time of 7 min at 72°C. When using *PfuUltra* II Fusion HS DNA polymerase (*PfuU*) (Stratagene), an extremely high fidelity proofreading thermostable DNA polymerase, the 25 µl PCR mix contained 1× reaction buffer, 250 µM dNTPs, 200 pM of each primer, 2.5 µl of DNA extract and 2.5 U of *PfuU*. Amplifications conditions were 95°C/20 s, followed by the same varying number of cycles as HiFi *Taq*, of 95°C/20 s, 60°C/20 s, 72°C/15 s, with a final extension time of 3 min. The reactions were carried out on the MyCycler Thermal Cycler (Bio-Rad). Slightly differing PCR mixes and amplification conditions were used with each polymerase to adhere as closely as possible to the product instructions provided by their respective manufacturer. PCR products were visualised by agarose gel electrophoresis.

Table 5: List of primers used in the mtDNA mutation burden assay

Application/Primer Name		Primer Sequence
1st PCR		
mt16535F	Forward Primer	5'-GCC CAC ACG TTC CCC TTA AAT AAG A-3'
mt474R	Reverse Primer	5'-AGT ATG GGA GTG RGA GGG RAA AA-3'
2nd PCR		
LongM13R	Forward Primer	5'-CAA TTT CAC ACA GGA AAC AGC TAT GAC-3'
LongT7	Reverse Primer	5'-GAC GGC CAG TGA ATT GTA ATA CG-3'
Sequencing Primers		
T7	Forward Primer	5'-TAA TAC GAC TCA CTA TAG GG-3'
M13R	Reverse Primer	5'-CAG GAA ACA GCT ATG AC-3'

2.2.3.2 Transformation, Selection of Bacterial Cells and PCR Amplification for Sequencing (2nd PCR)

PCR products from HiFi *Taq* amplifications were ligated into pCR®2.1-TOPO® (Invitrogen) and transformed into OneShot® TOP10 Chemically Competent *E. coli* cells (Invitrogen). PCR products from *PfuU* amplifications were ligated into StrataClone Blunt Vector (Stratagene) and transformed into

StrataClone SoloPack Competent Cells (Stratagene) that were part of the StrataClone Blunt PCR Cloning Kit. In both cases, the transformed cells were grown on LB agar plates with 100 ug/ml ampicillin (Roche) and recombinant plasmids were identified by blue-white screening/colour selection using 2% Bluo-Gal (Invitrogen) . After overnight growth, 93 white colonies/clones were picked and each of these was resuspended in 40 µl of UltraPure™ DNase/RNase-Free Distilled Water (Invitrogen) and heated at 95°C/10 min to release DNA from the cells that would subsequently be used as template for PCR. Each resuspended clone was PCR amplified using PCR primers that were complementary to both plasmid vectors: longT7 (5'-GAC GGC CAG TGA ATT GTA ATA CG-3') and longM13R (5'-CAA TTT CAC ACA GGA AAC AGC TAT GAC-3') (**Table 5**). The 25 µl PCR mix contained 1x PCR buffer, 1.5 mM MgCl₂, 200 µM dNTP, 500 pM of each primer, 2 µl of template clone resuspension and 0.05 U *Taq* DNA Polymerase, Native (Invitrogen). Amplification conditions were: 94°C for 30 s followed by 30 cycles of 94°C/15 s, 58°C/30 s and 72°C/30 s with a final extension time of 7 minutes at 72°C. The reactions were carried out on the MyCycler Thermal Cycler (Bio-Rad).

2.2.3.3 Commercial Cycle Sequencing

The PCR products were diluted 1:60 in DNase/RNase-Free Distilled Water (Invitrogen) for commercial PCR product purification and sequencing (Canada's Michael Smith Genome Sciences Centre, Vancouver, B.C.) using the universal sequencing primer M13R (5'-CAG GAA ACA GCT ATG AC-3') (**Table 5**).

Sometimes, a clone's sequence would have an "out of phase" characteristic (mixed sequence problem³) occurring on only one side of the insert i.e. prior to or after the C-tract (**Figure 8**). If this occurred, the clones were resequenced from the opposite direction using the universal sequencing primer, (T7 5'-TAA TAC GAC TCA CTA TAG GG-3') (**Table 5**). The "out of phase" problem occurs due to sequencing polymerase slippage downstream of a long C-tract, resulting in sequencing products of varying sizes.

³ It is when two or more peaks are observed at each base position.

Figure 8: Chromatogram of a characteristic “out-of-phase” sequence due to the C-tract

2.2.3.4 In-House Cycle Sequencing

Diluted PCR products (1:40 in DNase/RNase-free distilled water) were cycle sequenced using BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) with a modified sequencing reaction mix. The 6 µl sequencing reaction mix contained 2.1 µl sequencing dilution buffer (175 mM Tris-HCl pH9.0, 1.25 mM MgCl₂), 0.3 µl of BigDye® Terminator Reaction Mix, 0.33 µM HPLC-purified sequencing primer (IDTDNA) (**Table 5**) and 1 µl of diluted PCR product. The sequencing amplification conditions were: 96°C/10 s, 50°C/5 s, 60°C/55 s for 25 cycles. The products of cycle sequencing were resolved on the ABI 3730xl DNA Analyzer (Applied Biosystems) at the BC Centre for Excellence in HIV/AIDS, Vancouver, BC.

2.2.3.5 Sequence Analysis

Sequences were aligned against the revised Cambridge reference sequence (rCRS) [122] using Sequencher v4.7 sequence analysis software (Gene Codes Corporation). Only 454bp + variable C-tract length between positions nt16558 – 451, were analysed which excluded the degenerate primer sequences. Clones with secondary peaks (a peak within a peak) (**Figure 9**) higher than 50% were discarded. This represented approximately 1% of sequences. Any sequence derived from a clone should, in theory, have one clear peak per base since there is only one DNA species present. Therefore, any base with a secondary peak of less than 50% is considered a negligible artefact. Sequences with secondary peaks likely originate from 2 colonies in too close proximity. All ambiguous positions were verified through manual comparison with the chromatogram. As described by Wilding *et al.*, insertions and deletions (indels) with ≥ 2 mutations were considered a single event [106]. Also, if > 3 clones from one

sample had identical mutations at the same nucleotide position, this was defined as heteroplasmy and these were excluded from the analysis of mutation rate [106]. BER or mutation rate per 10,000 base pair (bp) was calculated using the following equation:

$$= \frac{\text{number of errors/mutations} \times 10,000 \text{ bp}}{\text{number of clones analysed} (454 \text{ bp} + C \cdot \text{tract length})}$$

(454 bp + C-tract length) being the size of the DNA insert without the primer sequences. A minimum of 80 clones was analysed for each sample.

Figure 9: Chromatogram of a peak within a peak leading to ambiguous base calling

2.2.4 Determining Background Error Rate of the Assay

To determine the BER of the assay, DNA from a clone contained in a resuspended heated *E. coli* colony (“cloned control DNA”), was PCR amplified and cloned again using the same protocol as described above. It is assumed that the number of mutations arising from the replication of the D-loop containing-plasmids by *E. coli* polymerase will be negligible. This is because the overall error rate of replication by *E. coli* polymerase is less than 1 in 10^9 nucleotides [109]. Moreover, to determine the fidelity of the assay in relation to the length of the C-tract (between nt303 – 309), clones derived from PR21C-M showing C-tract length heteroplasmy of 7, 8, 9, 10 and 11 Cs were amplified and cloned a second time. See **Figure 10** for schematic diagram of the MMBA.

Figure 10: Schematic diagram of the mtDNA mutation burden assay (MMBA) and determination of the background error rate (BER) of the assay. Total DNA is extracted (nDNA and mtDNA) from a whole blood sample. Subsequently, a 509 bp fragment of the D-loop is PCR amplified (1st PCR) and cloned into a plasmid vector. Ninety-three colonies/clones are picked for a 2nd PCR amplification and sequenced. The sequenced clones are analysed for mutations relative to the sample's consensus sequence (the average of all the sequenced clones). From this, the mutation burden is determined by calculating the number of mutations observed for every 10,000 bp divided by the number of clones multiplied by the length of the sequence analysed i.e. 454 bp (this does not include the length of the primer sequence from the 1st PCR). When determining BER, a clone is picked (represented by the green arrow) and the MMBA assay is performed again starting from the 1st PCR. The solid balls on the DNA represent point mutations.

2.3 Results

2.3.1 Number of Clones to Analyse with Blood Samples

Within the time- and cost-constraints of assaying, the number of clones it would require to achieve low and stable intra-sample variability was determined. Two subjects' blood samples were assayed, with both HiFi *Taq* and *PfuU* and sequences for ~300 clones/sample/PCR polymerase used were obtained. From this, coefficient of variations (CV)⁴ were calculated for varying number of clones analysed/sample i.e. 15 clones were randomly selected 8 times from the pool of ~300 clones and the CV for mutation burden was calculated. This was also repeated for 30, 45, 50, 55, 60, 65, 70, 75, 80, 90 and 100 clones (**Figure 11**). An even larger pool of sequences to randomly select from would have been more ideal however, cost-constraints were limiting.

Given *PfuU*'s lower BER, it had a higher intra-sample CV than HiFi *Taq*. With HiFi *Taq*, the CV stabilised at >55 clones analysed, while with *PfuU*, the CV remained variable at up to 100 clones. Based on these results, *PfuU* would require the analysis of >100 clones; at least in this type of sample (i.e. blood) which shows a low mtDNA mutation burden (vs. for example, muscle tissue) [123]. Consequently, it was decided that when assaying with HiFi *Taq*, a minimum of 80 clones would be analysed; this is 25 more clones than the point at which the CV stabilises, in order to allow for a large margin of error.

⁴ CV = coefficient of variation = ratio of the standard deviation (σ), to the mean (μ), multiplied by 100 % = $\frac{\sigma}{\mu} \times 100\%$. It is a measure of dispersion of data points around the mean.

Figure 11: Intra-sample CV for blood samples from 2 subjects with HiFi *Taq* and *PfuU* at 25 PCR amplification cycles

CVs were calculated for varying number of clones analysed/samples whereby, for each sample, a number of clones (15 - 100) were selected at random, 8 times, from a pool of ~300 clones and a CV for mutation burden was calculated.

2.3.2 Overview of Background Error Rate Determination

For each BER experiment performed, a single bacterial colony or clone (containing a single of species of plasmid with the D-loop insert) was chosen and PCR amplified once more. From the bacterial plate, 93 colonies were resuspended in water, re-amplified and sequenced on one 96-well plate. The remaining wells were set aside for 1 positive (plasmid cloned mtDNA D-loop, which had been previously successfully sequenced) and 2 negative (PCR mix without any DNA added) controls. In some cases, the assay was performed several times for each designated set of PCR conditions (**Table 6**). The majority of the assays were completed at 25 PCR amplification cycles (“PCR amplification cycles” will be shortened to “cycles” from hereafter) with both enzymes. Moreover, extensive assaying was also carried out at 35 cycles with HiFi *Taq* since our clinical samples (**section 3**) were assayed using these conditions. More sampling of clones with 10 or 11C C-tracts was not performed since they are rarely encountered in the general population [124]. With the clinical samples described in section 3, only one subject had a 10 C

consensus C-tract (N = 208); while an 11 C consensus C-tract was never observed. A larger sample size for each condition would have been preferable however; cost- and time-constraints were limiting.

Table 6: Number of times experiments were performed at various PCR amplification conditions

# PCR cycles	Enzyme	C-tract Length				
		7	8	9	10	11
15	HiFi <i>Taq</i>	–	1	–	–	–
	<i>PfuU</i>	–	1	–	–	–
20	HiFi <i>Taq</i>	–	1	–	–	–
	<i>PfuU</i>	–	1	–	–	–
25	HiFi <i>Taq</i>	6	7	3	1	1
	<i>PfuU</i>	6	7	3	1	1
30	HiFi <i>Taq</i>	–	1	–	–	–
	<i>PfuU</i>	–	1	–	–	–
35	HiFi <i>Taq</i>	5	5	4	–	–
	<i>PfuU</i>	–	2	–	–	–
45	HiFi <i>Taq</i>	–	1	–	–	–
	<i>PfuU</i>	–	1	–	–	–

– denotes “Not Assayed”

2.3.3 Time-dependent Effect on Background Error Rate for *PfuU* at Varying Numbers of PCR Amplification Cycles and C-tract Lengths

The mean BER for all mutation types⁵ with *PfuU* at 25 PCR cycles was much higher than that of other numbers of PCR cycles (at 20, 30, 35 and 45 cycles the BERs were <1 vs. ~3 at 25 cycles). Closer inspection of the data points at 25 cycles (**Figure 12**) revealed a time-dependent effect. Namely, samples labelled with solid icons were assayed between April and October 2008, whilst samples labelled with clear icons were assayed in October 2009. The latter, had on average, higher BERs than those assayed between April and October 2008. A possible reason for this discrepancy may have arisen from the *PfuU* enzyme. April 2008 corresponded to the time of purchase of the *PfuU* enzyme; no further purchases were made. During the time lapse between date of purchase in April 2008 and assaying in October 2009, the enzyme may have lost its high-fidelity characteristic. This is despite the fact that the manufacturer did not indicate any expiry date or recommend a time period for usage. Due to this revelation, I will

⁵ “All mutation types” refers to all transversions and transition mutations, as well as insertions and deletions.

hereafter exclude *PfuU* samples assayed in October 2009 from the analyses. Unlike *PfuU*, this time-dependent effect was not observed with samples assayed with HiFi *Taq*.

Figure 12: Time dependent effect on background error rates outside the C-tract mutation with *PfuU* at varying number of PCR amplification cycles and C-tract lengths
Scatterplot of BERs of all mutation types outside the C-tract (454 bp long). Solid icons denote samples assayed between April and October 2008. Clear icons denote samples assayed in October 2009.

2.3.4 Background Error Rate for All Mutation Types with Varying C-tract Lengths

The BERs for all mutation types with varying C-tract lengths (7-11 Cs) were examined using both *PfuU* and HiFi *Taq* with 25 cycles. BERs **within** the C-tract were 10- to 100-fold higher (**Figure 13**) compared to the regions **outside** the C-tract (**Figure 14**). BER within the C-tract itself was almost absent if the C-tract was 7 Cs long, however as the C-tract lengthened the BERs within the C-tract increased dramatically with both enzymes. The most common mutation observed for both enzymes was a one C deletion within the nt303-309 C-tract (**Figure 15**).

Figure 13: Background error rates **within** the C-tract with *PfuU* at 25, as well as HiFi *Taq* at 25 and 35 PCR amplification cycles with varying C-tract lengths

Scatterplot of BER of all mutation types within the C-tract ($C_{7-11}TC_6 = 14-18$ bp long). Green and purple diamonds denote HiFi *Taq* and *PfuU*, respectively.

Figure 14: Background error rates **outside** the C-tract with *PfuU* at 25, as well as HiFi *Taq* at 25 and 35 PCR amplification cycles with varying C-tract lengths

Scatterplot of BERs of all mutation types outside the C-tract (454 bp long). Green and purple diamonds denote HiFi *Taq* and *PfuU*, respectively.

Figure 15: Mean percentage of total sequences with a mtDNA mutation by type observed within the C-tract with varying consensus C-tract lengths with HiFi *Taq* and *PfuU* at 25 PCR amplification cycles. Stacked bar graph of types of mutations as a percentage observed within the C-tract. -#C represents a deletion of cytidine; +#C represents an insertion of cytidine between nt303-309; A, T or G mutation that interrupts the nt303-309 C-tract signifies that a C has been replaced by either an A, T or G base within the nt303-309 portion of the C-tract. A mutation within the nt311-316 C-tract signifies that a C has either been replaced by an A, T or G base or has been deleted. The last category of change incorporates several types of mutations; they have been combined since mutations in the nt311-316 region are far less common than those in the nt303-309 region.

BERs with varying C-tract (7-9 Cs) lengths were also examined at 35 PCR cycles with HiFi *Taq*. Unlike the results at 25 PCR cycles, BERs **within** the C-tract were extremely variable with a large range for the 8C C-tract clones (**Figure 13**). Moreover, the type of mutations that occurred included the insertion of 4 and 5 Cs (**Figure 16**), which had not been observed with other PCR amplification conditions (i.e. HiFi *Taq* and *PfuU*, other numbers of amplification cycles, as well as other clones of varying C-tract lengths). Although the BERs **within** the C-tract for 8C C-tract clones at 35 cycles were highly variable, the BERs **outside** the C-tract were in a similar range to the 7 and 9C C-tract clones (**Figure 14**). These results would suggest that while the HiFi *Taq* polymerase is unable to effectively amplify the C mononucleotide repeats, this inability is confined to this region only.

Figure 16: Mean percentage of total sequences with a mtDNA mutation by type observed within the C-tract with varying consensus C-tract lengths with HiFi *Taq* at 35 PCR amplification cycles. Stacked bar graph of types of mutations as a percentage observed within the C-tract. -#C represents a deletion of cytidine; +#C represents an insertion of cytidine between nt303-309; A, T or G mutation that interrupts the nt303-309 C-tract signifies that a C has been replaced by either an A, T or G base within the nt303-309 portion of the C-tract. No T interruption at nt310 signifies that the T that usually interrupts the nt303-316 C-tract at nt310 is absent. A mutation within the nt311-316 C-tract signifies that a C has either been replaced by an A, T or G base or has been deleted. The last category of change incorporates several types of mutations; they have been combined since mutations in the nt311-316 region are far less common than those in the nt303-309 region.

If the samples of the most commonly found lengths of C-tracts in nature are pooled (i.e. 7-9 C C-tracts), then BER **outside** the C-tract at 25 cycles was ~5.1x higher for HiFi *Taq* than *PfuU*. However, the variation observed was lower amongst the samples amplified with HiFi *Taq* compared to the variation with *PfuU* (mean $\pm$ SD, 3.47 ± 1.14 (N = 16) vs. 0.68 ± 0.36 (N = 4), CV: 33% vs. 53%).

Although the BER **outside** the C-tract is higher at 35 cycles than at 25 with HiFi *Taq*, the difference is not significant (Mann-Whitney $p = 0.20$) and furthermore, the variation observed at 35 cycles was lower compared to that at 25 cycles (mean $\pm$ SD, 3.92 ± 0.87 (N = 14) vs. 3.47 ± 1.14 (N = 16), CV 22% vs. 33%).

After having observed the particularly frequent and wide range of errors induced within the C-tract with either enzyme at both 25 and 35 amplification cycles, it was decided that this region would be excluded in the analyses of mutations. With such high variability, any conclusions made would be questionable or C-tract specific; although it would have been interesting to study.

2.3.5 Background Error Rate for Non-C-tract All Mutation Types with Varying Number of PCR Amplification Cycles

The BERs for HiFi *Taq* and *PfuU* remained relatively stable with increasing numbers of PCR cycles with 8C C-tract clones (**Figure 17**). More concrete conclusions are difficult to make since the number of clones assayed per PCR amplification number was low especially for the 15, 20, 30 and 45 cycles. While, overall the number of PCR cycles did not seem to have an effect on BER, it is recommended that ≥ 20 cycles be used since the small amount of PCR product after 15 PCR cycles would make obtaining 93 clones challenging.

Figure 17: Background error rates outside the C-tract with HiFi *Taq* and *PfuU* at varying numbers of cycles with an 8 C C-tract
Scatterplot of BERs outside the C-tract (454 bp long). Green and purple diamonds denote HiFi *Taq* and *PfuU*, respectively.

2.3.6 Background Error Rate for Each Mutation Type

BERs were also broken down by mutation type, i.e. transition⁶ and transversion⁷ mutations, as well as deletions and insertions. Transition and transversion mutations were further subdivided into point

⁶ Substitution of a purine base with another purine or a pyrimidine base with another pyrimidine.

mutation pairs i.e. a substitution point mutation is paired with its complement, for example A → C (AC) is complemented with T → G (TG).

The mutation spectra of HiFi *Taq* and *PfuU* were different. Whilst at 25 cycles, transition mutations were the most commonly observed mutations of all types of point mutations with HiFi *Taq* (90.2%), *PfuU* was less specific, and induced transition and transversion mutations almost equally often (54.5 % and 45.5%, respectively) (**Table 7**). In addition, the principal specific type of mutations induced by HiFi *Taq* were AG/TG mutations, this was in accordance with several other studies [125]. Also, the most commonly observed transversion mutations with HiFi *Taq* and *PfuU* were the CA/GT mutations. With *PfuU*, and unlike HiFi *Taq*, CG/GC mutations were not induced. Although, the percentage frequency of transversion mutations was very different between the two enzymes, similar absolute mean BERs were observed for these mutations (**Figure 18**). The occurrence of deletions and insertions were low with HiFi *Taq* (1.8%) but these types of mutations were entirely absent in *PfuU* (**Table 7**).

When comparing HiFi *Taq* at 25 and 35 PCR cycles, not only were transversion mutations less frequently observed at 35 PCR cycles than at 25 (5.2% vs. 8.0%, respectively), no AC/TG mutations were induced (**Table 7** and **Figure 18**).

Table 7: Mutation spectra observed at 25 and 35 PCR amplification cycles with HiFi *Taq* and *PfuU* with all C-tract lengths

		Mean Percentage (%)							
		Transversions				Transitions		deletions	insertions
# PCR cycles		A to C, T to G	A to T, T to A	C to G, G to C	C to A, G to T	A to G, T to C	C to T, G to A		
25	HiFi <i>Taq</i> (N = 16)	0.4	0.9	0.5	6.3	59.1	31.1	0.9	0.9
	<i>PfuU</i> (N = 4)	9.1	9.0	0.0	27.5	9.1	45.4	0.0	0.0
35	HiFi <i>Taq</i> (N = 14)	0.0	0.9	0.1	4.1	56.7	36.4	1.3	0.5

Note: At all numbers of PCR cycles, results included 7-9 C C-tract lengths.

⁷ Substitution of a purine base with a pyrimidine base or vice versa.

Figure 18: Background error rates for each type of mutation with *PfuU* at 25 and HiFi *Taq* at 25 and 35 PCR amplification cycles with differing C-tract lengths

Scatterplot of BERs broken down by mutation type outside the C-tract (454 bp long) with differing C-tract lengths (7-11 Cs). Sample sizes: N = 18, for HiFi *Taq* with 25 cycles, N = 6 for *PfuU* with 25 cycles and N = 14 for HiFi *Taq* with 35 cycles.

2.3.7 Sequencing Reaction Background Error Rate

No differences in sequences were detected when 93 clones were taken through the sequencing reaction stage of the MMBA four independent times. It was therefore concluded that any BER observed for the MMBA overall is not due to the sequencing reaction itself.

2.3.8 Background Error Rate Compared to Average Clinical Sample MtDNA Mutation Burden

After having obtained the data from assaying clinical samples, a comparison was made to the BER determined at 35 PCR cycles with HiFi *Taq*. The unexposed control infant clinical samples (**section 3**) (the samples which were hypothesised would have the lowest mutation burden compared to their mothers and HIV/HAART-exposed infants) had similar mean total mutation rate to that determined to represent our BER (mean $\pm$ SD, 3.89 ± 1.19 vs. 3.92 ± 0.87 mutations/10,000 bp). Upon closer inspection of one of the transition mutations (CT/GA), the mean BER was almost twice as high as the mean mutation rate observed for the unexposed control infants (0.73 ± 0.12 vs. 1.42 ± 0.12 mutations/10,000 bp) (**Figure 19**). This last observation seems to go against the assumption that the “cloned control DNA” used to determine BER should either have an equal or lower mutation burden than the clinical samples. The DNA used to determine BERs were derived from a clone and should, in theory, be composed of only one DNA species; consequently, the mutations observed should also only arise because of the infidelity of the PCR polymerase. This is in contrast to clinical samples whose variability would be because of the heteroplasmic mtDNA species due to the biology of mtDNA replication in addition to, the error rate of the PCR polymerase.

Of note, a preliminary analysis had been performed comparing a single HIV/HAART-exposed mother/infant pair to the background error rate, prior to assaying all the clinical samples. Only later was it recognised that these particular samples’ total mutation rates were not only considerably higher than the BER but also happened to be amongst the highest total mutation rates seen in any clinical sample (i.e. mother-6.45 mutations/10,000 bp and infant-6.21 mutations/10,000 bp). Therefore, at the time, it

was thought the margin of error (i.e. the difference between true mutation rates and background error rate) was acceptably large.

The reason for the BER being higher than the clinical samples' mutation rate remains unknown at this time and may be an artefact. It could perhaps be due to the varying amounts of starting DNA material added to the PCR reaction. Because the "cloned control DNA" template originated from a bacterial colony diluted in water, its DNA concentration was unknown. It was later determined through real-time PCR on a short D-loop fragment that the "cloned control DNA" was 10 to 100 fold more concentrated than the average clinical DNA sample. Originally, DNA concentration was not considered as a potential factor that could alter mutation burden. Furthermore, several sample concentrations had been measured and were all very similar. For that reason, DNA concentrations of the blood DNA extracts for the clinical samples and of the "cloned control DNA" samples had not all been quantified prior to assaying. We were unaware of these vastly different concentrations and, as such, the "cloned control DNA" samples were not diluted accordingly before assaying. The different concentrations in starting DNA material observed can most likely be attributed to the source of the DNA for the PCR reaction. "Cloned control DNA" samples were derived from bacteria containing high copy-number plasmids with the mtDNA D-loop insert. In contrast, the clinical samples were derived from a DNA extraction of whole blood consisting of total DNA (i.e. both nDNA and mtDNA), wherein mtDNA only constitutes a small portion of this.

To verify whether PCR template DNA amounts can affect mutation burden as measured in our assay, extensive assaying would need to be carried out at various dilutions. This would also have to be performed with a large sample size at each dilution to account for the wide variation observed between samples.

In addition, unlike the clinical samples which consisted of purified DNA extract, the starting material for the "cloned control DNA" samples was a resuspended bacterial colony that consisted of both DNA and

other cellular debris such as cell membranes. This cellular debris may have interfered with PCR enzyme activity or may have preferentially induced some types of mutations. To test this hypothesis, a plasmid preparation could be performed on bacterial clones to obtain purified plasmid DNA. Currently, this work is ongoing; however, the results will not be obtained in time for inclusion in this thesis.

Figure 19: Mean clinical sample mutation rates and background error rate observed for each type of mutation

Bargraph showing that, in certain instances, the BERs (cloned control DNA) were higher than mean mutation rates in the clinical samples (i.e. CA/GT and CT/GA mutations). The transition mutation rates are ~10x higher than those for transversion mutations in both clinical samples and “cloned control DNA”.

2.3.9 Intra-sample and Inter-sample Variability

To assess intra-sample variability, 2 unexposed control mother/infant pairs from the clinical samples (section 3) were each assayed 8 times. The CV for total mutation rates ranged between 22.0 – 35.3%. This was comparable to the inter-sample CV among all the unexposed infants and mothers (30.4% and 31.0%, respectively) (Table 8).

Table 8: Intra-sample total mutation rate variability in 4 individuals

		Mean $\pm$ S.D. mutations/ 10^4 bp (N = 8)	Coefficient of Variation (%)
Pair 1	Infant	3.79 $\pm$ 1.20	31.8
	Mother	3.79 $\pm$ 0.88	23.2
Pair 2	Infant	4.06 $\pm$ 0.89	22.0
	Mother	3.82 $\pm$ 1.35	35.3

The CVs for AC/TG mutation rates were also assessed (Table 9). The CVs were as expected, much higher than those for total mutation rate because AC/TG mutation rates are much lower than those for total mutations.

Table 9: Intra-sample AC/TG mutation rate variability in 4 individuals

		Mean $\pm$ S.D. mutations/ 10^4 bp (N = 8)	Coefficient of Variation (%)
Pair 1	Infant	0.07 $\pm$ 0.13	185
	Mother	0.10 $\pm$ 0.29	283
Pair 2	Infant	0.03 $\pm$ 0.10	282
	Mother	0.10 $\pm$ 0.14	138

3 Blood Mitochondrial DNA Mutations in HIV-infected Women and their Infants Exposed to HAART during Pregnancy

3.1 Overview

As described in **section 1**, to prevent the transmission of HIV to their infant and/or for their health, HIV-infected women receive NRTI-containing HAART regimens during pregnancy. NRTIs are thought to adversely affect mtDNA and may induce point mutations. MtDNA mutations have been linked to aging and age-associated diseases. Since NRTIs can cross the placenta and accumulate in amniotic fluid [126], they may induce mtDNA mutations in developing foetuses during pregnancy. Throughout embryogenesis and organogenesis, high rates of cell differentiation and proliferation (and therefore active mtDNA replication) occur. This may mean that the unborn child is more susceptible to mtDNA damage when exposed to NRTIs compared to adults who undergo lower rates of mtDNA replication. HAART has dramatically decreased HIV MTCT and the benefits clearly outweigh the risks. However, the full impact of these drugs in uninfected infants is largely unknown, especially in the long-term, and needs to be assessed.

3.2 Study Design

This was an exploratory cohort study and a retrospective/prospective examination of mtDNA mutations present in whole blood samples from HAART-exposed HIV-infected pregnant women and their infants, as well as from HAART-unexposed HIV-uninfected pregnant women and their infants.

3.2.1 Hypothesis

We hypothesised that infants born to HIV-infected mothers treated with HAART during pregnancy will have:

1. a significantly higher mtDNA mutation burden than infants born to HIV-uninfected mothers
2. an increased mtDNA mutation burden relative to their own mother, compared to infants born to HIV-uninfected mothers.

Also,

3. HAART-exposed HIV-infected mothers will have a significantly higher mtDNA mutation burden than HAART unexposed HIV-uninfected mothers.

3.3 Materials and Methods

3.3.1 Funding and Ethical Approval

This study was funded by two grants: “mtDNA damage in infants exposed to HIV drugs *in utero*” (CIHR-MOP-79331, principal investigator: Hlne Ct) and the “CIHR Emerging Team Grant in HIV Therapy and Aging” (CIHR-#HET-85515, principal investigator: Hlne Ct).

The study received ethical approval from the Research Ethics Boards of the University of BC (UBC) and the Hospital Research Review Committee of the Children's and Women's Health Centre of BC (H03-70356 and H04-70540).

3.3.2 Study Populations

3.3.2.1 Study Cohorts

Samples were drawn from two separate prospective observational cohort studies:

1. “Perinatal exposure to antiretroviral therapy in infants born to HIV seropositive mothers: Evaluation of toxicity using mitochondrial DNA and lactate levels” – was a collaboration between Ontario and BC researchers. It was funded by the Ontario HIV Treatment Network (OHTN) between 2003 – 2005 (principal investigators: Susan King and John Forbes), and is hereafter referred to as the “Baby Study”. Blood samples from *in utero* HAART-exposed infants were collected at 0-3 days, 4, 8, 12 and 26 weeks of age, when blood was obtained to perform routine HIV testing. Blood was collected only once from control infants at representative age groups for ethical reasons. An amendment was filed and received approval to use the Baby Study blood samples for mtDNA mutation burden assaying.

2. “Effect of antiretroviral therapy on maternal blood cell mitochondrial DNA levels during pregnancy in HIV-infected women” – was funded by the Canadian Foundation for AIDS Research (CANFAR) between 2005 – 2006 (principal investigator: Deborah Money). It enrolled pregnant women and their infants in BC, and will be hereafter referred to as the “Pregnancy Study”. Blood was collected from pregnant women at <18, 24-28, 32-36 weeks of gestation, delivery, 6 weeks post-partum and from infants at ~2-3 days. Cord blood and placental tissue was collected at time of delivery. For controls, the same samples were collected except no blood was collected at delivery. Once again, an amendment was filed and received approval to use the Pregnancy Study blood samples for mtDNA mutation burden assaying.

For this mtDNA mutation burden study, blood DNA samples from infants in both cohorts were obtained. In addition, blood DNA samples from mothers in the Pregnancy Study were also obtained in the Pregnancy Study. The Pregnancy Study intended not only to capture HAART-exposed infants but to also study their mothers, for comparison.

HAART-exposed

For both cohorts, all HAART-exposed pregnant women and infant subjects were identified as eligible for the study by their physician at the Oak Tree Clinic at the C & W during the mother’s regular prenatal visit. The Oak Tree Clinic specialises in the care of, pregnant women, partners, children and youth who are HIV-infected.

Infant inclusion criteria:

- Infant, born to HIV-infected, HAART-treated mother
- Informed consent was obtained from a parent or guardian

Infant exclusion criterion:

- Subsequent confirmation of HIV infection in the infant

Pregnant woman inclusion criteria (Pregnancy Study):

- HIV-infected and pregnant
- Obtained informed consent

Unexposed Controls

For the Baby Study cohort, healthy newborns were recruited from the C & W in-born units through physician-referrals. A small amount of blood (~0.5 ml) was collected at the same time as routine blood collection for population newborn screening. For older infants, the parents of infants <9 months old having infant blood drawn for any reason were directly approached by research staff at the C & W main accessioning laboratory and a small amount (~0.5 ml) was collected at the same time. For the Pregnancy Study cohort, HIV-uninfected mothers and their infants were recruited by poster advertisement placed throughout the C & W and at one of the co-investigator's private doctor's office at City Square Shopping Centre, Vancouver, BC. Additionally, recruitment by direct approach occurred at the Fir inpatient unit at the C & W, where women struggling with substance use problems in their pregnancy can access prenatal and postpartum care for themselves and their infants. As part of the Pregnancy Study, efforts were made to enrol control mothers/infants of similar ethnicity, socioeconomic status and illicit drug use as those HAART-exposed.

Infant inclusion criteria:

- Infants having blood work done for reasons unrelated to HIV-exposure.
- Informed consent was obtained from a parent or guardian.

Infant exclusion criterion:

- HIV-infection or antiretroviral exposure.

Pregnant woman inclusion criteria (Pregnancy Study):

- HIV-uninfected and pregnant
- Obtained informed consent

3.3.2.2 Inclusion Criteria for the mtDNA Mutation Burden Study

In addition to the criteria above from the 2 study cohorts, the following criteria needed to be met:

Infant inclusion criterion:

- Whole blood collected between 0 – 6 days of life

Samples collected later were not included since infants born to HIV-infected women would be given an oral AZT regimen during the first 6 weeks of life as prophylaxis and may have confounded the analysis of this study.

Mother inclusion criteria (for the Pregnancy Study):

- Having had a blood sample collected from their infant between 0 – 6 days of life
- Whole blood sample collected during their last doctor's visit prior to delivery

3.3.2.3 Special Clarifications on the Groups

Although the HAART-exposed infants have not become HIV-infected, they may be adversely affected by exposure to circulating HIV virus during pregnancy or may have been indirectly affected by the stresses of HIV-infection on the mother such as maternal cytokine imbalance [127]. Therefore, the HAART-exposed infants are also HIV-exposed. Since HAART-exposure cannot be separated from HIV-exposure (as this study does not have a HIV-exposed but HAART-unexposed group), the exposed group will therefore be called "HIV/HAART-exposed". For the unexposed control group, this refers to unexposed to HIV and HAART.

3.3.3 Samples Used and Collection Method

For this mtDNA mutation burden study, the mother's sample from the last perinatal doctor's visit before delivery was used (most often this was the 32-36 weeks of gestation sample). These samples were collected through venipuncture into EDTA or ACD BD Vacutainer® blood collection tubes (BD) and frozen rapidly as whole blood, without processing. For infants, samples obtained between 0-6 days were used

and these were collected by heel prick blood (~0.5 ml) at the same time as sample collection for routine population newborn or HIV screening.

3.3.4 Data Collection

For both cohorts, baseline information collected included maternal demographics, pregnancy history, maternal antiretroviral history and other drug/toxic exposures, as well as delivery and neonatal events, including antenatal antiretroviral use. Data was extracted from clinical records and collected through a detailed and confidential questionnaire. The data collected was stored in a Microsoft Access (Microsoft) database.

3.3.5 Clinical Laboratory Testing

The HIV/HAART-exposed mothers had their CD4+ count and HIV pVL load assessed 1 – 4 weeks prior to delivery (will be referred to as “near delivery”).

3.3.6 DNA Extraction

Total genomic DNA (i.e. nDNA and mtDNA) was extracted from 0.1 ml of whole blood using QIAamp® DNA Mini Kit (Qiagen) according to manufacturer protocol specifications.

3.3.7 MtDNA Mutation Burden Assay

The MMBA performed in this study operates under the same principle as the mtDNA mutation burden assay described in the materials and methods section in 2.2.3 (**Figure 20**). All samples were randomised and blinded prior to performing the MMBA. The number of amplification cycles used for the 1st PCR on the blood DNA extract was 35 cycles using HiFi *Taq* (Roche). To minimise assay condition variability between mother/infant pairs, the pairs were assayed at the same time. With mother/infant pairs, every step following bacterial plating occurred on 96-well plates that were partitioned in half to have both mother and infant clones. When no maternal sample was available, infant sampling took place on one 96-well plate. For each sample, 93 clones were sequenced, the remaining 3 wells were used for 2 negative (PCR mix without any DNA) and 1 positive (plasmid cloned mtDNA D-loop which had been

previously successfully sequenced) controls. The first consecutive 80 readable sequences were analysed. Total mutation rates (# mutations/10,000 bp) and the presence of AC/TG mutations were determined for each sample outside of the C-tract. The type and location of the mutations were also noted.

Figure 20: Schematic diagram of mtDNA mutation burden assay for clinical samples

Total DNA is extracted (nDNA and mtDNA) from a whole blood sample. Subsequently, a 509 bp fragment of the D-loop is PCR amplified (1st PCR) and cloned into a plasmid vector. Ninety-three colonies/clones are picked for a 2nd PCR amplification and sequenced. Eighty clones/sequences are analysed for mutations relative to the sample's consensus sequence (the average of all the sequenced clones). From this, the mutation burden is determined by calculating the number of mutations observed for every 10,000 bp divided by number of clones (80) multiplied by the length of the sequence (454 bp) analysed (this does not include the length of the primer sequence from the 1st PCR). The solid balls on the DNA represent point mutations

3.3.8 Statistical Analyses

When comparing between infants, both the Pregnancy and Baby Study cohorts were pooled. When comparing between mothers or between mother/infant pairs, only the Pregnancy Study cohort was used. Statistical analyses were performed using Microsoft Excel (Microsoft), χ^2 STAT (Addinsoft) and SPSS (SPSS Inc.). A *P* value (2-tailed) of <0.05 was defined as significant.

Total mutation rates were analysed by analysis of variance (ANOVA) between groups (comparison of HIV/HAART-exposed and unexposed controls) and by repeated measures ANOVA within groups (comparison of infants with their mothers). Analysis of covariance (ANCOVA) and repeated measures ANCOVA were used to examine the effects of HIV/HAART exposure during pregnancy on total mutation rate while controlling for potential covariates between and within groups respectively.

In contrast, the AC/TG mutations were analysed in terms of their absence or presence in a sample by Chi-squared test between groups and Wilcoxon signed-rank test within groups. Logistic regression was used to examine the relationship between HIV/HAART exposure during pregnancy on AC/TG mutations while controlling for potential covariates between and within groups respectively.

Hierarchical linear and hierarchical logistic regressions were used to examine predictive models of total mutation rate and presence of AC/TG mutations, respectively, within the HIV/HAART-exposed group.

Comparisons of demographic characteristics, clinical and laboratory values for HIV/HAART-exposed and unexposed control groups were carried out using the two-sample Student's *t*-test (2-tailed) for continuous variables and the chi-squared test for categorical variables.

3.4 Results

3.4.1 Study Population

Subjects were selected from both the Pregnancy and Baby Study cohorts which counted a total of 120 mother/infant pairs and 96 infants enrolled, respectively (**Figure 21**).

Of the initial infants enrolled in the Baby Study cohort, 46 were excluded since no blood sample was collected between 0-6 days of age. Of the remaining subjects, 19 HIV/HAART-exposed and 31 unexposed control subjects were eligible for study.

Of the initial mother/infant pairs enrolled in the Pregnancy Study cohort, 37 pairs were not included in this analysis for various reasons: withdrawal from the study (N = 1), infant blood sample processing error (N = 8), foetal demise (N = 2), misplaced infant samples (N = 2), infant sample collected beyond 6 days of age (N = 11), no *in utero* HAART-exposure (N = 1) and loss to follow-up (N = 12). Of the remaining pairs, 44 HIV/HAART-exposed and 39 unexposed controls pairs were included in this study.

After the samples were assayed, it was realised that 6 of the 83 mothers, had participated in both the Pregnancy and Baby Study cohort. Since samples from 2 infant subjects born to the same mother cannot be independent, only one of the pregnancies can be included in the study. Only in the Pregnancy Study were blood samples collected from both mother and infant and, as a result, the Pregnancy Study cohort mother/infant samples were analysed and the 6 infant samples derived from the Baby Study cohort were ultimately excluded. Two of the eligible mother samples from the Pregnancy Study cohort were eliminated due to labelling errors discovered after assaying and as such, only their corresponding infant samples were included in this study.

In the end, 42 HIV/HAART-exposed and 39 unexposed control mother/infant pairs, as well as 15 HIV/HAART-exposed and 31 unexposed control infants were considered in the final analyses.

Figure 21: Flowchart summary of study samples

3.4.1.1 Demographics and Clinical Characteristics, as well as Laboratory Values

Demographic and clinical characteristics, as well as laboratory values are shown in **Table 10**. Both groups were similar with respect to infant sex, infant Apgar score at 5 minutes and gestational length. The HIV/HAART-exposed infant birth weight was significantly lower by ~0.25 kg compared to the unexposed controls. The unexposed control infants were predominantly delivered vaginally whereas HIV/HAART-exposed infants were almost equally delivered vaginally and by Caesarean-section. HIV/HAART-exposed mothers were younger than the unexposed mothers, they were also more likely to be Aboriginal or Black-African Canadians and the unexposed control mothers were more often Caucasian or Asian. Alcohol consumption, as well as illicit drug and/or methadone use ever during pregnancy were similar between groups except for smoking (marijuana or cigarettes), whereby HIV/HAART-exposed mothers smoked significantly more than unexposed controls. In both groups, all mothers who reported smoking marijuana during their pregnancy also smoked cigarettes, therefore both variables were considered together. Eighty percent (N = 8/10 in both groups) of all mothers receiving methadone during their pregnancy also used illicit drugs. Active HCV-coinfection based on RNA PCR testing was more common in HIV/HAART-exposed mothers than the unexposed controls; however that data was not available for the majority of mothers in the control group.

HIV/HAART-exposed group

None of the infants acquired HIV. Of the 57 HIV/HAART-exposed pregnant women, 12 (21%) conceived on HAART, 5 (9%) started HAART in the 1st trimester, 31 (54%) in the 2nd trimester and 9 (16%) in the 3rd trimester. Once HAART was initiated, all women continued on HAART throughout the remaining pregnancy. The median [interquartile range, IQR] (range) duration of *in utero* HAART-exposure for the group was 19.3 [13.6 – 28.6] (0.6 – 41.1) weeks. The majority of the women were on a combination of AZT + 3TC + NVP or LOP (N = 48, 84%). Those who conceived on HAART were more likely to be on other regimens (8/12, 67%). Thirty-three women (58%) had received HAART prior to pregnancy with a median

duration of 41.1 [19.7 – 96.1] (0.1 – 603.9) weeks. The median CD4+ count for the women was 450 [300 – 630] (90 – 1200) cells/ul and only 8 (14%) of the women had a detectable HIV pVL near delivery, with a median HIV pVL of 432 [264 – 648] (53 – 1280) RNA copies/ml. Based on clinical diagnosis, the median duration of HIV infection at the time of delivery was 3.8 [0.9 – 6.0] (0.1 – 15.3) years.

Table 10: Demographic characteristics, clinical characteristics, and laboratory values for HIV/HAART-exposed mothers and their infants, as well as unexposed control mothers and their infants

Characteristic or Value	HIV/HAART-exposed (N = 57)	Unexposed Controls (N = 70)	P Value
Infant Characteristics			
Male Sex	31 (54)	38 (54)	0.991
Birth Weight, kg	3.05 [2.73 – 3.41] (1.62 – 4.05)	3.30 [2.81 – 3.64] (1.40 – 4.57)	0.026
Gestational length, weeks	38.3 [37.7 – 39.9] (31.3 – 41.6)	39.1 [38.3 – 40.1] (28.9 – 41.9)	0.070
Delivery method, vaginal birth ^a	31 (54)	51 (73)	0.030
Apgar score at 5 minutes	9 [9 – 9] (7 – 10)	9 [9 – 9] (5 – 10)	0.593
Maternal Characteristics			
Maternal age, years	30.6 [25.1 – 35.6] (17.4 – 41.5)	32.3 [29.4 – 35.6] (22.2 – 43.0)	0.049
Active HCV co-infection ^b	8 (19) (N = 42)	1 (6.7) (N = 15)	0.420 ^c
Maternal ethnicity			<0.001
Aboriginal, First Nations, Métis or Inuit	17 (30)	6 (8.6)	
Caucasian	22 (39)	38 (54)	
Black-African Canadian	12 (21)	0 (0)	
Hispanic	0 (0)	1 (1.4)	
Asian and Other	6 (10)	18 (26)	
Missing information	0 (0)	7 (10)	
Maternal use or consumption ever during pregnancy			
Alcohol	19 (35) (N = 54)	16 (23)	0.131
Smoking ^d	30 (55) (N = 55)	20 (29)	0.003
Illicit drugs ^e and/or methadone	20 (36) (N = 55)	16 (23)	0.097
HAART Exposure			
Duration of maternal HAART during pregnancy, weeks	19.3 [13.6 – 28.6] (0.6 – 41.1)	N/A	
Total duration of pre-pregnancy HAART, weeks	41.1 [19.7 – 96.1] (0.1 – 603.9) (N = 33 ^f)	N/A	
HIV Clinical Data			
Duration of HIV infection at delivery, years ^g	3.8 [0.9 – 6.0] (0.1 – 15.3)	N/A	
Undetectable HIV pVL ^h at 1 – 4 weeks prior to delivery ⁱ	49 (86)	N/A	
Detectable HIV pVL, HIV RNA copies/ml (N = 8)	432 [264 – 648] (53 – 1280)	N/A	
CD4+ count, cells/ul at 1 – 4 weeks prior to delivery ^j	450 [300 – 630] (90 – 1200)	N/A	

Note. Data are number (%) of subjects or median [interquartile range] (range) and N, total subjects for which data is available (if not available for entire group). N/A, not applicable. Significant *p* values (<0.05) are highlighted in red

^a Delivery Method-vaginal or Caesarean-Section

^b Data based on HCV RNA PCR testing

^c Smoking includes cigarettes and marijuana

^d Fisher's exact test was used when there were <5 subjects in a given category

^e Illicit drugs included but were not limited to: heroin, cocaine, crack, crystal meth, ecstasy, benzodiazepine, opioids.

^f Other subjects had no prior exposure to HAART

^g Data based on HIV clinical diagnosis date

^h Undetectable HIV viral load signifies <50 copies/ml

ⁱ 1 – 4 weeks prior to delivery will be referred to as "near delivery" hereafter

^j Normal range: 600 – 1,500 cells/ul

3.4.2 Sample Size Calculation

Prior to starting this study, no substantial work had been carried out by other groups in the mtDNA point mutation field, therefore this mtDNA mutation burden study was deemed exploratory. As such, determining appropriate sample size for this study was difficult. One third of the way into the study, a sample size calculation was performed on temporarily unblinded unexposed control infants. To calculate this, the van Belle's equation for sample size as a function of coefficient of variation and ratio of means was used with $\alpha = 0.05$, $\beta = 0.20$ or power = 0.80 and a two-sided test [128]:

$$N \approx \left(\frac{4CV(\text{in percent})}{\ln(f)} \right)^2$$

Where N = number of people to study, f = ratio of the means μ^1/μ^2 , CV = coefficient of variation (variability between subjects).

The inter-sample CV of the unexposed control infants for total mutation rate was 26.1% (N = 23) and rounded up to 30% to be conservative. Additionally, at most what we could probably have expected was a quarter-fold difference in means (i.e. 1.25 : 1 or f = 1.25) between the HIV/HAART-exposed and unexposed control group. Then, using van Belle's equation:

$$N \approx \frac{16 \times (0.3)^2}{(\ln 1.25)^2} \approx 29 \text{ persons}$$

The total number of subjects per category in this study was above this 29 person threshold i.e.

HIV/HAART-exposed infants (N = 57) and mothers (N = 42), as well as unexposed control infants (N = 70) and mothers (N = 39). Of course, for the rare event AC/TG mutations which we did not know would specifically need to be analysed, the sample size calculation would be different.

3.4.3 MtDNA Mutation Burden

MtDNA mutation burden was determined for each sample by amplifying a portion of the mtDNA D-loop (nt16559-447), followed by cloning (80 clones/sample) and sequencing of the PCR products.

3.4.3.1 Covariates

For both the HIV/HAART-exposed and unexposed control groups (mothers and infants), covariates considered were: maternal age at delivery, smoking (cigarettes or marijuana) ever in pregnancy and illicit drug and/or methadone use ever in pregnancy. Maternal age was hypothesised to be a covariate because mtDNA mutations have been observed to accumulate with age [81], and perhaps, mtDNA mutations that accumulate in the oocytes of mothers are transmitted to their progeny [129]. Smoking (cigarettes or marijuana), opioids (such as heroin and morphine), cocaine and amphetamines (such as MDMA/ecstasy and crystal meth) were also hypothesised to be covariates because they have all been associated with oxidative stress [130-134], which may in turn induce mtDNA mutations (**Figure 6**).

In addition, amount of DNA used for 1st PCR was later added as a possible confounder for both groups. We did not deliver equivalent quantities of DNA when assaying the samples, because, unbeknownst to us prior to assaying, there was a wide range of DNA concentrations from the DNA extracts (infants: 23.5 – 724 ng/μl, mothers: 21.5 – 114 ng/μl). Presently, it is unknown whether differing amounts of DNA from a given subject could alter the mutation rate that is detected (but such experiments are currently planned to be carried out by someone else in the laboratory).

None of the covariates mentioned were significantly related to the outcomes (i.e. total mutation rate or the presence of AC/TG mutations) when controlling for the other predictors. However, smoking was highly correlated with illicit drug and/or methadone use ($N = 124$, $r = 0.78$, $p < 0.001$) and mothers who smoked tended to be younger ($N = 125$, $r = -.421$, $p < 0.001$). Gestational length was negatively correlated with HCV co-infection ($N = 57$, $r = -.34$, $p = 0.011$), smoking ($N = 125$, $r = -.33$, $p < 0.001$),

alcohol consumption ($N = 124$, $r = -.24$, $p = 0.007$), as well as illicit drug and/or methadone use ($N = 125$, $r = -.31$, $p < 0.001$).

For the HIV/HAART-exposed infants, possible predictors of mtDNA mutation accumulation also included: duration of mother's HAART-exposure prior to pregnancy, duration of infant *in utero* HAART-exposure, and detectable maternal HIV pVL near delivery. For the HIV/HAART-exposed mothers, this list also included: total lifetime exposure to HAART, detectable maternal HIV pVL and CD4+ count near delivery. A detectable HIV pVL and low CD4+ count can be an indicator that patient adherence to HAART is poor or that HIV has developed resistance to the HAART regimen administered.

In the HIV/HAART-exposed group, illicit drug and/or methadone use did not appear to be a significant marker for a detectable HIV pVL ($N = 55$, $r = 0.22$, $p = 0.10$).

3.4.3.2 Total mtDNA Mutations

Total mtDNA mutation rate was determined by analysing all mutations that occurred in a given sample between nt16559-447 in the D-loop region; substitution mutations (transitions and transversions) and indels, with the exception of any mutations arising in the C-tract region (nt303-316). The total mtDNA mutation rate was not corrected for the background error rate (refer back to **section 2.3.8**)

A one-way between groups ANOVA was conducted to compare the relationship between either HIV/HAART-exposure or no HIV/HAART-exposure on total mutation rates in infants and mothers (**Figure 22**). ANCOVA was subsequently used to control for the amount of DNA, maternal age, smoking and illicit drug and/or methadone use. No significant difference between HIV/HAART-exposed and unexposed control infants was found prior to (mean $\pm$ S.D., 3.94 ± 1.16 mutations/ 10^4 bp ($N = 57$) vs. 3.85 ± 1.17 mutations/ 10^4 bp ($N = 70$), respectively, $F [1, 125] = 0.16$, $p = 0.693$) or after controlling for covariates. Similarly, no significant difference between HIV/HAART-exposed and unexposed control mothers was found prior to (4.44 ± 1.03 mutations/ 10^4 bp ($N = 42$) vs. 4.28 ± 1.33 mutations/ 10^4 bp ($N = 39$), respectively, $F [1, 79] = 0.37$, $p = 0.545$) or after controlling for covariates.

Figure 22: Scatterplot comparing HIV/HAART-exposed and unexposed control infants' and mothers' total mutation rates

The *P* value shown is prior to controlling for covariates. These included: amount of DNA, maternal age at delivery, smoking ever in pregnancy, as well as illicit drug and/or methadone use ever in pregnancy. There was no significant difference between HIV/HAART-exposed and unexposed control infants or mothers prior to or after controlling for covariates.

Note: None of the covariates were significantly correlated to the outcome.

A one-way within group repeated-ANOVA was conducted to compare the total mutation rates between infants and mothers with and without HIV/HAART-exposure (**Figure 23**). Repeated measures ANCOVA was used to control for the previously mentioned covariates. Within the HIV/HAART-exposed group, infants' total mutation rate was significantly lower than their mothers' (3.83 ± 1.10 vs. 4.44 ± 1.03 mutations/10⁴ bp (N = 42), respectively, $F [1, 41] = 6.7, p = 0.013$). However, this difference disappeared when controlling for covariates. Whilst, within the unexposed control group, there was no significant difference between infants' and mothers' total mutation rate found prior to (3.98 ± 1.21 vs. 4.28 ± 1.33 mutations/10⁴ bp (N = 39), respectively, $F [1, 38] = 1.4, p = 0.239$) or after controlling for covariates.

Figure 23: Scatterplot comparing infants' to their mothers' total mutation rates in the HIV/HAART-exposed and unexposed control groups

The P value shown is prior to controlling for covariates. These included: amount of DNA, maternal age at delivery, smoking ever in pregnancy, as well as illicit drug and/or methadone use ever in pregnancy. There was a significant difference between HIV/HAART-exposed infants and mothers ($p = 0.013$), however this significance disappears once controlling for covariates. There was no significant difference within the unexposed control group prior to or after controlling for covariates. Significant p values are highlighted in red.

Note: None of the covariates were significantly correlated to the outcome.

3.4.3.3 AC/TG Mutations

When comparing the BER (determined in **section 2**) to that of the total mutation rate in clinical samples, a high noise to signal ratio was observed. Therefore, only AC/TG mutations were further specifically analysed, which were mutations not induced by HiFi *Taq* under our assay conditions. These mutations were, however, rare; i.e. in all the subjects assayed ($N = 208$), 55 had one $A \rightarrow C$ or $T \rightarrow G$ mutation in the 80 sequences analysed per sample and 5 had two of these mutations (3 HIV/HAART-exposed mothers, 1 HIV/HAART-exposed and 1 unexposed control infant). No subjects had three. Because of the rarity of these mutations, they were statistically analysed in terms of their presence or absence in a given sample. Similar to the total mutation rate analysis, the region examined did not include the C-tract.

A chi-squared test was conducted to compare the relationship between either HIV/HAART-exposure or no HIV/HAART-exposure on the percentage of infants and mothers with the AC/TG mutations. Logistic regression was used to control for the previously mentioned covariates (**Table 11**). A higher percentage of HIV/HAART-exposed infants (26.3%) had these mutations compared to the unexposed control infants (14.3%) but this difference did not reach statistical significance ($\chi^2_{1df} = 2.88, p = 0.090$). However, this difference approached significance when controlling for covariates (Odds Ratio (95% confidence interval (CI)): 2.5 (0.97 – 6.46), $p = 0.058$). In contrast, a significantly higher percentage of HIV/HAART-exposed mothers (42.9%) harboured these mutations compared to unexposed control mothers (17.9%, $\chi^2_{1df} = 5.88, p = 0.015$). This effect persisted when controlling for covariates (Odds Ratio (95% CI): 4.7 (1.4 – 15.6), $p = 0.012$).

Table 11: Comparison of the percentage of HIV/HAART-exposed and unexposed control infants and mothers with the AC/TG mutations

	Infants		Mothers	
	HIV/HAART- exposed (N = 57)	Unexposed Controls (N = 70)	HIV/HAART- exposed (N = 42)	Unexposed Controls (N = 39)
N (%) subjects with AC/TG mutations	15 (26.3)	10 (14.3)	18 (42.9)	7 (17.9)
Ratio between groups	1.84		2.39	
χ ²	2.88		5.88	
P value	0.090		0.015	
Controlling for Covariates				
Odds Ratio(95% CI)	2.5 (0.97 – 6.46)		4.7 (1.4 – 15.6)	
P value	0.058		0.012	

Covariates included: amount of DNA for 1st PCR, maternal age at delivery, smoking ever in pregnancy, as well as illicit drug and/or methadone use ever in pregnancy. Significant p values are highlighted in red.

Note: None of the covariates were significantly correlated to the outcome.

The Wilcoxon signed-rank test was conducted to compare the percentage of infants and mothers with the AC/TG mutations with and without HAART-exposure. Logistic regression was used to control for the previously mentioned covariates (**Table 12**). Within the HIV/HAART-exposed group, the percentage of infants with the AC/TG mutations (23.8%) was significantly lower than that of their mothers (42.9%,

Wilcoxon $Z = -2.1$, $p = 0.033$). However, this difference disappeared when controlling for covariates ($p = 0.777$). Whilst within the unexposed control group, there was no difference between percentage of infants and mothers with these mutations prior to (17.9 % vs. 17.9% respectively, Wilcoxon $Z = 0.0$, $p > 0.999$) or after controlling for covariates ($p = 0.283$).

Table 12: Comparisons of the percentage of infants and mothers with the AC/TG mutations with and without HAART-exposure

	HIV/HAART-exposed (N = 42)		Unexposed Controls (N = 39)	
	Infants	Mothers	Infants	Mothers
# (%) of subjects with AC/TG mutations	10 (23.8)	18 (42.9)	7 (17.9)	7 (17.9)
Ratio within groups	1.80		1.00	
Wilcoxon, Z	-2.1		0.0	
P value	0.033		>0.999	
P value controlling for covariates	0.777		0.283	

Covariates included: amount of DNA for 1st PCR, maternal age at delivery, smoking ever in pregnancy, as well as illicit drug and/or methadone use ever in pregnancy. Significant p values are highlighted in red.

Note: None of the covariates were significantly correlated to the outcome.

3.4.3.4 Controlling for Covariates in the HIV/HAART-exposed Group

Using hierarchical linear regression, in addition to the previously described covariates, the following predictors were considered: a detectable HIV pVL near delivery, duration of mother's HAART-exposure prior to pregnancy and duration of infant *in utero* HAART-exposure. None predicted total mutation rate in HIV/HAART-exposed infants (all p values >0.150). In addition, for the mothers, total lifetime exposure to HAART, a detectable HIV pVL and CD4+ count near delivery did not predict total mutation rate in HIV/HAART-exposed mothers (all p values >0.500).

Using hierarchical logistic regression, a detectable HIV pVL near delivery, duration of mother's HAART-exposure prior to pregnancy and duration of infant *in utero* HAART-exposure did not predict the presence of AC/TG mutations in HIV/HAART-exposed infants (all p values >0.500). Moreover, for the mothers, total lifetime exposure to HAART and CD4+ count near delivery did not predict the presence of AC/TG mutations (both p values >0.100). However, having a detectable HIV pVL near delivery was

associated with increased odds of having these mutations (6/7 HIV/HAART-exposed women with a HIV pVL had AC/TG mutations, Odds Ratio: 14.0 (1.4 – 137.3), $p = 0.024$), albeit the confidence interval was quite large.

3.4.4 Correlation Between Mother/Infant Pairs

As mentioned, since mtDNA mutations have been shown to accumulate with age in fibroblasts [81], they may also be accumulating in the oocytes of mothers and in turn may be passed on to their children [129]. A possible association between mtDNA mutations in mothers and their infants was, therefore, explored.

3.4.4.1 Total Mutation Rate

For total mutation rate, infants and their mothers in either HIV/HAART-exposed or unexposed control groups were not significantly correlated ($N = 42$, $r = -0.02$, $p = 0.914$ and $N = 39$, $r = 0.26$, $p = 0.106$, respectively) (Figure 24).

Figure 24: Scatterplots showing lack of significant correlation between infants and their mothers in either HIV/HAART-exposed or unexposed control groups

3.4.4.2 AC/TG Mutations

For the presence or absence of AC/TG mutations, infants and their mothers in the HIV/HAART-exposed group were positively correlated ($N = 42$, $r = 0.31$, $p = 0.048$). While in the unexposed control groups, they were not significantly correlated ($N = 39$, $r = -.045$, $p = 0.787$).

3.4.5 Correlation Between Maternal Age and MtDNA Mutation Burden in Infants and Mothers

Again, as alluded to in **section 3.4.4**, mtDNA mutation has been shown to accumulate with age and these may also be transmitted from the mother to her child. A possible association between maternal age and mtDNA mutations in mothers and their infants was, therefore, explored.

3.4.5.1 Total Mutation Rate

For total mutation rate, maternal age was not significantly correlated with maternal and infant total mutation rate in both groups combined ($N = 81$, $r = 0.07$, $p = 0.567$ and $N = 127$, $r = -0.08$, $p = 0.376$, respectively) (**Figure 25**).

Figure 25: Scatterplots showing lack of significant correlation between maternal age and infant or maternal total mutation rate of both groups combined

Green = HIV/HAART-exposed group Purple = Unexposed control group

3.4.5.2 AC/TG Mutations

For the presence or absence of AC/TG mutations, maternal age was mildly positively correlated with presence of AC/TG mutations in mothers of both groups combined ($N = 81$, $r = 0.20$, $p = 0.071$). In contrast, maternal age was not significantly correlated with presence of AC/TG mutations in infants of both groups combined ($N = 127$, $r = 0.006$, $p = 0.944$). Of note, maternal age was positively correlated with presence of AC/TG mutations in HIV/HAART-exposed mothers ($N = 42$, $r = 0.31$, $p = 0.049$).

3.4.6 Heteroplasmy Outside the C-tract

If >3 sequences from one sample had the same mutation at the same nucleotide position outside the C-tract, this was considered heteroplasmy and not a mutation. Although such heteroplasmy could arise from a mutation early in the infant's development in the womb, it could also originate from a PCR error in an early cycle. Therefore, to err on the conservative side, these were not considered in these analyses. This phenomenon occurred infrequently (number of subjects (%) in each group with heteroplasmy): 3/57 (5.3%) infants and 3/42 (7.1%) mothers in the HIV/HAART-exposed group and 5/70 (7.1 %) infants and 3/39 (7.7%) mothers in the unexposed control group. If heteroplasmy occurred in a subject, it occurred at only one nucleotide position within the D-loop region analysed and in all cases, it was always a transition type of substitution. In 2 mother/infant pairs (1 HIV/HAART-exposed and 1 unexposed control), heteroplasmy occurred in both (and this was also at the same nucleotide position). In 4 cases of infant heteroplasmy (1 HIV/HAART-exposed and 3 unexposed controls), their respective mothers also harboured the same mutation within the 80 sequences analysed. However, unlike the infants, their mothers did not harbour the same mutation above the cut-off criteria for heteroplasmy (i.e. the same mutation in >3 sequences)

3.4.7 Consensus C-tract Length

There was no difference in consensus C-tract length within and between groups (**Table 13**). In all subjects, 7 and 8C consensus C-tracts were almost equally common (43.8% vs. 43.2% respectively across groups), followed by 9C (10.1% across groups). This distribution was in accordance to study by Sanchez-Céspedes *et al.* [124]. Only one subject, an unexposed control infant had a 10C consensus C-tract. Altogether, there were 6 subjects that had no T that usually interrupts the C-tract at nt310 (C₆₋₁₂TC₆). In 3 instances, there was discordance between the consensus C-tract length between mother and infant (1 HIV/HAART-exposed and 2 unexposed control pairs).

Table 13: Consensus C-tract length and heteroplasmy within the C-tract in infants and mothers in the HIV/HAART-exposed and unexposed control groups

Consensus Length of C tract		Infants		Mothers	
		HIV/HAART-exposed (N = 57)	Unexposed Controls (N = 70)	HIV/HAART-exposed (N = 42)	Unexposed Controls (N = 39)
7	N (%)	27 (47.4)	30 (42.9)	20 (47.6)	14 (35.9)
	N with heteroplasmy (%)	16 (59.3)	20 (66.7)	13 (65.0)	11 (78.6)
	Of those with heteroplasmy, mean % of 80 sequences with a change from the consensus	3.4	4.9	6.2	2.4
8	N (%)	24 (42.1)	29 (41.4)	18 (42.9)	18 (46.2)
	N with heteroplasmy (%)	23 (95.8)	29 (100)	18 (100)	18 (100)
	Of those with heteroplasmy, mean % of 80 sequences with a change from the consensus	16.0	10.5	16.4	14.9
9	N (%)	5 (8.8)	8 (11.4)	3 (7.1)	5 (12.8)
	N with heteroplasmy (%)	5 (100)	8 (100)	3 (100)	5 (100)
	Of those with heteroplasmy, mean % of 80 sequences with a change from the consensus	25.5	23.9	18.8	38.0
10	N (%)	0 (0)	1 (1.4)	0 (0)	0 (0)
	N with heteroplasmy (%)	0 (0)	1 (100)	0 (0)	0 (0)
	Of those with heteroplasmy, mean % of 80 sequences with a change from the consensus	N/A	51.3	N/A	N/A
No T in C-tract	N (%)	1 (1.8)	2 (2.9)	1 (2.4)	2 (5.1)
	N with heteroplasmy (%)	N/A	N/A	N/A	N/A
	Of those with heteroplasmy, mean % of 80 sequences with a change from the consensus	N/A	N/A	N/A	N/A

Note: No T in C-tract signifies that the T that usually interrupts the C-tract at nt310 (C₆₋₁₂TC₆) was not present. N/A = not applicable. Heteroplasmy within the C-tract is defined as any change observed within the C-tract in any given sequence that deviates from the consensus sequence.

3.4.8 MtDNA Mutations Within the C-tract

As mentioned before, mutations in the C-tract were not included in the total mutation rate and AC/TG mutations analyses. The BER study in section 2 showed that the PCR polymerase induced an extremely high level of mutations in the C-tract region compared to the regions outside the C-tract. It would

therefore be very difficult to distinguish “real” mutations from the clinical sample and mutations induced by the PCR polymerase. However, for interest, mutations in the C-tract are described below.

Between and within groups, mtDNA mutations are highly prevalent within the C-tract, unlike the regions outside. The types of mtDNA mutations that were observed and categorised were: 1. insertions of 1 – 3 Cs, deletions of 1 – 2 Cs or substitutions of a C to either A, T or G in the nt306 – 309 region 2. the absence of a T within the C-tract (at nt310) or 3. any of these mutations after the T (nt311 – 316).

MtDNA mutations were most common in the nt306 – 309 region followed by nt310 and then nt311 – 316 (**Figure 26**). When a T was absent from within the C-tract, mixed sequence problems occurred in both sequencing directions (using T7 or M13R sequencing primers) thus making it difficult to accurately read the chromatograms and, as a result, C-tract length could not be determined.

Figure 26: Mean percentage of 80 sequences with a mtDNA mutation by type observed within the C-tract in subjects with varying consensus C-tract lengths

Note: This does not include subjects whose consensus sequence contains no T in the C-tract

Stacked bar graph of types of mutations as a percentage observed within the C-tract. -#C represents a deletion of cytidine; +#C represents an insertion of cytidine between nt303-309; A, T or G mutation that interrupts the nt303-309 C-tract signifies that a C has been replaced by either an A, T or G base within the nt303-309 portion of the C-tract. No T interruption at nt310 signifies that the T that usually interrupts the nt303-316 C-tract at nt310 is absent. A mutation within the nt311-316 C-tract signifies that a C has either been replaced by an A, T or G base or has been deleted. The last category of change incorporates several types of mutations; they have been combined since mutations in the nt311-316 region are far less common than those in the nt303-309 region.

On average, between and within groups, the percentage of 80 sequences that had a mtDNA mutation in the C-tract were similar (this does not include those who had a consensus C-tract with no T): mean $\pm$ S.D., $9.84 \pm 13.0\%$ in the infants (N = 56) and $10.5 \pm 11.7\%$ in the mothers (N = 41) of the HIV/HAART-exposed group and $9.50 \pm 12.7\%$ in the infants (N = 68) and $13.1 \pm 14.2\%$ in the mothers (N = 37) of the unexposed control group.

Between and within groups, an insertion of 1 C was the most common mtDNA mutation and it occurred with similar frequencies (this does not include those who had a consensus C-tract with no T): mean $\pm$ S.D., $6.1 \pm 9.1\%$ in the infants (N = 56) and $7.7 \pm 9.9\%$ in the mothers (N = 41) of the HIV/HAART-exposed group, as well as $6.1 \pm 8.9\%$ in the infants (N = 68) and $8.8 \pm 9.5\%$ in the mothers (N = 37) of the

unexposed control group. However, the insertion of 1 C occurred slightly less frequently, in infants compared to their mothers in both groups (HIV/HAART-exposed group (N = 41): infants- $6.1 \pm 9.9\%$ vs. mothers- $8.2 \pm 10.6\%$, unexposed control group (N = 37): infants- $7.5 \pm 9.9\%$ vs. mothers- $8.8 \pm 9.5\%$, $p = 0.41$).

MtDNA mutations within the C-tract occurred with all subjects whose C-tracts were longer than 7Cs except for one infant with an 8C C-tract, where this did not occur (**Table 13**). Also, between and within groups, the mean frequency of subjects with mtDNA mutations and mean frequency of sequences with a mtDNA mutation increased as the consensus C-tract lengthened.

3.4.9 Consensus Sequence Between Mother and Infant

Outside of the C-tract region, 2 mother/infant pairs (both unexposed controls) had discordant consensus sequences at one nucleotide position. In the first pair, at position nt131, the infant had a C consensus and mother a T, neither of them had heteroplasmy at this nucleotide position. In the second pair, at position nt146, the consensus base for the infant was T and the mother a C. However, there was considerable heteroplasmy in both the infant and the mother (16/80 and 43/80 sequences were Cs at this nucleotide position in the infant and mother, respectively).

3.4.10 Sequences with Multiple Mutations

From the analyses of all the sequences from every subject (i.e. 16640 sequences), if a sequence harboured any mutations, qualitatively, it usually only harboured one, occasionally two and rarely three. Only 3 sequences had >3 mutations (i.e. 4, 5 and 12 mutations).

Intriguingly, there were 5 sequences that harboured multiple mutations (i.e. >3) and for each of these, the same type of transition mutation was observed (except for one sequence, which had 4/5 mutations that were the same transition mutation) (**Table 14**). Interestingly, these sequences were only seen in infants (2 HIV/HAART-exposed and 3 unexposed controls).

Table 14: Type and number of mutations from subjects with a single sequence with multiple transition mutations

Subject	Group	Type of Base		Number of Mutations
		Consensus	Mutation	
Infant 1	HIV/HAART-exposed	G	A	8
Infant 2	HIV/HAART-exposed	C	T	10
Infant 3	Unexposed control	G	A	4
		T	C	1
Infant 4	Unexposed control	G	A	19
Infant 5	Unexposed control	A	G	30

These sequences were not taken into consideration when analysing total mutation rates. Prior to unblinding of groups, the sequences were excluded on the basis that perhaps the mutations were likely an artefact and we chose to be conservative. After having completed all the sequencing analyses, the data on sequences with multiple mutations were reviewed in more detail. It is unlikely that the mutations in these sequences were due to PCR artefacts such as the depletion of a particular dNTP during the PCR. When performing PCR on the clones in a 96-well plate, PCR mastermixes (containing all the PCR reagents including dNTPs) are prepared. This provides a uniform PCR reaction mix between wells. If a particular dNTP had been exhausted, one would expect there to be several sequences with multiple mutations. This is because the majority of the sequences are identical and therefore should exhaust a particular dNTP at the same rate. That said, after extensively researching the literature, no publications that have similar findings with respect to observing sequences with multiple mutations and concomitantly mutations of the same type have been discovered.

3.4.11 Location of Point Mutations

No specific location along the D-loop region studied here seemed to be more prone to mutations than others (except for the C-tract region). There also was no evidence of a specific pattern to the location of the AC/TG mutations.

3.4.11.1 Point Mutations Referred to in the Literature

Point mutations in mtDNA are written as consensus base followed by nucleotide position within the mtDNA genome and lastly the mutation base (ex. A156G). In the literature, 7 specific D-loop point mutations have been the subject of interest: C114T, C150T, A189G, T195C, T408A, T414G and T414C. The C114T and T195C mutations have been reported by Rollins *et al.* to be associated with bipolar disorder and were found in brain tissue [135]. The T414G mutation was reported in brain tissue of AD patients [92] and accumulated with age in cultured fibroblasts [81]. The T414C mutation was claimed to be AD-specific by Coskun *et al.* [92]. The A189G and T408A mutations have shown specificity to muscle tissue and accumulation with age [120]. The C150T mutation was also reported to accumulate with age, but unlike other mutations which are usually detrimental, it has been linked with longevity [136]. The frequency of mutations and heteroplasmy at these particular nucleotide positions are summarised in **Table 15**.

Table 15: Subjects with mutations or heteroplasmy and their consensus base at locations of point mutations referred to in the mtDNA literature

Consensus Base/mtDNA nt position/Mutation Base	N (%) of subjects with:	Infants		Mothers	
		HAART-exposed (N = 57)	Unexposed Controls (N = 70)	HAART-exposed (N = 42)	Unexposed Controls (N = 39)
C114T (associated with bipolar disorder, [135])	Mutations ¹	0 (0)	3 (4.4)	0 (0)	2 (5.1)
	Heteroplasmy ²	0 (0)	0 (0)	0 (0)	0 (0)
	C Consensus	57 (100)	68 (97.1) ³	42 (100)	39 (100)
C150T (associated with longevity, [136])	Mutations	0 (0)	2 (3.3)	2 (5.0)	3 (8.6)
	Heteroplasmy	0 (0)	0 (0)	0 (0)	0 (0)
	C Consensus ⁴	55 (96.5)	61 (87.1)	40 (95.2)	35 (89.7)
A189G (accumulates with age, specifically in muscle, [120])	Mutations	2 (4.2)	3 (4.5)	7 (18.9)	6 (16.2)
	Heteroplasmy	1 (1.8)	0 (0)	0 (0)	1 (2.6)
	A Consensus ⁵	48 (84.2)	67 (95.7)	37 (88.1)	37 (94.9)
T195C (associated with bipolar disorder [135])	Mutations	2 (4.4)	4 (7.7)	2 (6.3)	2 (7.1)
	Heteroplasmy	0 (0)	1 (1.4)	0 (0)	0 (0)
	T Consensus ⁶	45 (78.9)	52 (74.3)	32 (76.2)	28 (71.8)
T408G (accumulates with age, specifically in muscle, [120])	Mutations	0 (0)	1 (1.4)	0 (0)	0 (0)
	Heteroplasmy	0 (0)	0 (0)	0 (0)	0 (0)
	T Consensus	57 (100)	70 (100)	42 (100)	39 (100)
T414G (associated with AD in brain tissue [92] and reported to accumulates with age in cultured fibroblasts, [81])	Mutations	0 (0)	0 (0)	0 (0)	0 (0)
	Heteroplasmy	0 (0)	0 (0)	0 (0)	0 (0)
	T Consensus	57 (100)	70 (100)	42 (100)	39 (100)
T414C (AD-specific mutation, [92])	Mutations	4 (6.9)	3 (4.3)	2 (4.8)	1 (2.6)
	Heteroplasmy	0 (0)	0 (0)	0 (0)	0 (0)
	T Consensus	57 (100)	70 (100)	42 (100)	39 (100)

Notes: The mutations shown in this table cannot be differentiated from those that are “real” and those that are induced by the PCR polymerase error.

¹This is the N (%) of subjects with the specific mutation whose consensus base is C.

²>3 sequences had the same mutation at the same nucleotide position in a subject

³ 1 infant with a T consensus and 1 with a deletion.

⁴All others had a T consensus

⁵All others had a G consensus, except for 1 HIV/HAART-exposed mother/infant pair whose consensus was a C.

⁶All others had a C consensus

Of all the main point mtDNA mutations described in the literature which are mentioned here, the A189G mutation was the most prevalent in our data set, even though this mutation had been labelled as highly 'specific' to muscle (**Table 15**) [120]. Also, the A189G mutation does seem to accumulate with age as the percentage of mothers (~16 - 19%) with this mutation was higher than the infants (<5%). However, this may be purely coincidental since mutations at other locations not shown here were sometimes more prevalent in the infants than in the mothers.

The A189G mutation reached heteroplasmy levels in 2 individuals:

1. A HIV/HAART-exposed infant (4 mutations/80 sequences, 5.0%)
2. A unexposed control mother (11 mutations/80 sequences, 13.8%).

It was previously stated that this mutation has only been shown to reach high degrees of heteroplasmy in skeletal muscle [107]. Moreover, of note, 8.2% of all subjects were homoplasmic for the A189G transition, which is in agreement to a study that describes this position as polymorphic [117].

None of the subjects were homoplasmic for T408G, T414C or T414G and this was the same for C114T except for 1 unexposed control infant. In contrast, some subjects were homoplasmic for C150T and T195C, however, T195C homoplasmy was even more prevalent than C150T (8.2% and 24.0% of all subjects, respectively). As claimed by Rollins *et al.*, T195C is associated with bipolar disorder but it would seem unlikely that 24.0% of those studied here who are homoplasmic for T195C would have a predisposition to bipolar disorder [135].

The mutations at the specific locations explored here, occurred at comparable levels to mutations at other locations. It would therefore suggest that what is being observed are mutations induced by the PCR polymerase and are not "real" mutations. Moreover, all the mutations studied here are transition mutations except for the T414G mutation and as shown in **section 2** the induction rate of transition mutations by HiFi *Taq* is high.

4 Discussion

4.1 MtDNA Mutation Burden Assay

To determine the mtDNA mutation burden of a sample, an assay was developed; named the “mtDNA mutation burden assay” (MMBA), based on a post-PCR cloning strategy. The assay involves the amplification of the mtDNA D-loop (nt16559-447), followed by cloning and sequencing of the PCR products. This method is convenient and is able to detect unknown mtDNA mutations over a substantial region of the mtDNA genome’s D-loop. However, one of the disadvantages of this strategy is the artificial mutations induced by the PCR polymerases’ inherent error rates.

Although, the use of a direct cloning strategy (**section 1.6.7.2**) would have introduced fewer artificial mutations, this strategy was not feasible within the context of our clinical samples, given the amount of starting material that would have been required. Many of our clinical samples were obtained from infants from whom only a small sample of blood can be collected.

As a result of the intrinsic PCR polymerase error rate, several experiments were performed to determine the background error rate (BER) (i.e. PCR polymerase-induced mutations) of the MMBA with 2 different enzymes (HiFi *Taq* and *PfuU*), using various numbers of PCR cycles (15, 20, 25, 30, 35 and 45). BER was also determined with differing lengths of C-tract (i.e. 7 – 11 Cs at D310, between nt303 – 309). In each case, BER was assessed by amplifying previously sequenced single clones and subjecting them to the MMBA once more.

Extensive assaying to obtain an average BER was performed at 35 cycles with HiFi *Taq* because many clinical samples had already been assayed with these conditions prior to commencing BER determination. Much assaying was also performed at 25 cycles (for both HiFi *Taq* and *PfuU*) because it generates enough PCR product for cloning while within the exponential phase (i.e. this refers to early PCR cycles during which products are generated at constant doubling rate with each amplification cycle) [137]. Too few amplification cycles can result in little PCR product and can lead to low cloning efficiency

[138]. Too many amplification cycles can lead to unspecific products being generated since certain PCR reagents become exhausted as the reaction progresses. At this stage, the plateau phase is reached, whereby, PCR products are no longer being generated. In addition, the results of overamplification can include: appearance of unspecific bands, generation of small deletion mutants and in some cases disappearance of the desired product [137]. It should however be noted that as part of this study, no apparent increase in error rate was observed with increasing number of amplification cycles with either *PfuU* or HiFi *Taq*.

As expected the BER of the MMBA was lower with *PfuU* than with HiFi *Taq* (**Table 16**) because, unlike HiFi *Taq*, *PfuU* has proofreading 3' → 5' exonuclease activity [139]. These error rates, were however, much higher than the error rates claimed in the literature for these enzymes. The way in which the “cloned control DNA” was prepared to perform the BER determination experiments may have contributed to the difference in error rates between the MMBA and the literature, as well as the similarity in rates between the MMBA BER and the average clinical sample total mutation rate (as alluded to before in **section 2.3.8**). The “cloned control DNA” consisted of a bacterial colony resuspended in water heated to allow the contents of the bacteria to be released into suspension (including the DNA). The cellular debris from such a DNA preparation may have affected the PCR polymerase activity and as such, lowered the fidelity of the polymerases and favoured the induction of certain types of mutations.

The sequencing step of the MMBA was found not to contribute to BER. This conclusion is based on having repeatedly sequenced 93 clones, four independent times and finding no differences between the repeats.

Table 16: Summary of total error rates of different PCR polymerases with 25 and 35 amplification cycles in the literature and with the mtDNA mutation burden assay

	PCR Polymerase	# of PCR Cycles	Error Rate/10,000 bp ¹
Literature	<i>PfuU</i> (Stratagene)	25	0.10 ²
		35	0.14 ²
	HiFi <i>Taq</i> (Roche)	25	0.38 ³
		35	0.53 ³
	<i>Taq</i>	25	2.23 ²
		35	3.19 ²
Observed MMBA4	<i>PfuU</i> (Stratagene)	25 (N = 4)	0.68
		35 (N = 2)	0.62
	HiFi <i>Taq</i> (Roche)	25 (N = 16)	3.47
		35 (N = 14)	3.92

Human POLG *in vitro* error rate ≈ 0.10 errors/10,000 bp [62]. Nuclear pol $\delta \leq 0.13$ and $\epsilon \leq 0.02$ errors/10,000 bp (from *Saccharomyces cerevisiae*) [140].

Note: ¹Error rate for all types of mutations.

²Calculated based on information in [139] using a *lacI*-based bacterial assay. These numbers may vary depending on the method used to determine PCR enzyme fidelity.

³Calculated based on Roche Expand High Fidelity^{PLUS} PCR system pack insert that states it is “6-fold more accurate compared to *Taq* DNA polymerase” [141].

⁴Calculated based on BER outside the D310 C-tract region.

With high BERs, using the MMBA for absolute quantification of total mutations is not suitable. However, it is more feasible to do so with *PfuU* since its BER is substantially lower than that of HiFi *Taq*.

Furthermore, although absolute quantification of mtDNA mutations may not be a possibility with the MMBA, relative quantification of total mutations can still be of some use to compare different groups within a study.

The mutation spectra of HiFi *Taq* and *PfuU* were different. Almost all mutations induced by HiFi *Taq* were transition mutations (~92 % of all mutations). In contrast, *PfuU* was less specific and induced transition and transversion mutations with almost equal frequency (54.5% and 45.5% of all mutations, respectively). Nonetheless, similar absolute mean BERs were observed with these enzymes with respect to transversion mutation induction. One of the limitations of the MMBA is that PCR-induced mutations are indistinguishable from genuine mutations. However, there is an exception to this, there were particular mutations that were not induced by the polymerases, those being AC/TG mutations with HiFi *Taq* at 35 cycles and CG/TG mutations with *PfuU* at 25 or 35 cycles. With this in mind, if absolute

quantification of a particular type of mutation is to be performed, one should pay particular consideration to the type of PCR polymerase used for the MMBA.

An extremely high and variable error rate was observed within the D310 C-tract as opposed to regions outside the C-tract with either enzyme. The most common error that occurred in the C-tract was a deletion of a C between nt303 – 309. The high error rate was expected since it is known that mononucleotide repeats such as the C-tract cause PCR amplification errors probably because of slipped strand mispairing (polymerase “slippage”) [119]. As a result, it was decided that when assaying a clinical sample, mutations arising within the C-tract would be excluded from calculations of total mutation rate (nonetheless, these mutations could be analysed separately for interest).

Although extensive assaying with *PfuU* at 25 cycles was attempted, this was not successful. Two series of experiments were performed, approximately one year apart. When comparing the first series and the second, the latter showed a dramatic increase in BER. Between the two series, the same lot of *PfuU* enzyme was used. A hypothesis for this is perhaps the proofreading 3'→5' exonuclease degraded during this time and as a result, errors that may have arisen during PCR replication were not corrected as they normally would be. No such observation was made with HiFi *Taq* even after much longer storage times than with *PfuU*. Other possible reasons for the higher *PfuU* BER may have been due to inconsistencies in the PCR mastermix such as the use of different dNTP lot numbers.

When taking into consideration, the time- and cost-constraints of assaying, the number of clones it would require to achieve a low and stable CV was determined. Because of HiFi *Taq*'s high BER, a low/stable CV was achieved at >55 clones analysed when assaying blood samples and evaluating total mutation burden, compared to a CV which still remained high/variable at up to 100 clones with *PfuU*. Low/stable CVs would be achieved with both enzymes if the mtDNA mutation load within a clinical sample was high, which is not the case with blood. However, muscle tissue for example, can harbour

many mtDNA mutations because myocytes have a slow turnover rate compared to blood cells and as such, more mtDNA mutations are able to accumulate [123].

4.1.1 Limitation

Sample size was a limitation in this study. In several cases, only one experiment was carried out per sample condition (for example, N = 1 for *PfuU* with 25 cycles with an 11C C-tract, **Table 6**); this was due to high time and financial costs, mostly from sequencing. With a larger sample size, more precise BERs would have been determined. As the cost of sequencing continues to decrease perhaps this will no longer be a limiting factor.

Of note, in comparison to other groups, we have performed many more experiments to determine BERs at various conditions. In most reports in the literature, only a single clone is evaluated once [106] and sometimes not performed at all [92]. In contrast, we evaluated for example, 16 clones at 25 cycles with HiFi *Taq*. Also, the number of clones that we analysed per sample (i.e. >80 clones/sample) was much higher than in most studies, be it when assaying clinical samples or the “cloned control DNA”. For example, Coskun *et al.* only analysed 10 -20 clones for each patient with AD in their study [92, 142].

4.1.2 Conclusion

Before the full implementation of this assay within a laboratory, each of the factors studied here would need to be carefully taken into consideration, along with the advantages and disadvantages of PCR polymerases (**Table 17**). Nonetheless, this assay can be tailored according to study needs. Moreover, this assay should be adaptable to study other regions of the mtDNA genome and to other animal models (such as the mouse) if pseudogenes are accounted for.

Table 17: Advantages and disadvantages of each PCR polymerase

PCR Polymerase	Advantages	Disadvantages
HiFi <i>Taq</i>	<ul style="list-style-type: none"> Fewer clones needed to obtain low/stable CV = lower cost Suitable to study AC/TG mutations (at 35 cycles) No change in error rate after long storage times 	<ul style="list-style-type: none"> High BER
<i>PfuU</i>	<ul style="list-style-type: none"> More suitable for absolute quantification of mutation burden Suitable to study CG/GC mutations 	<ul style="list-style-type: none"> May require a high number of clones if sample has low mutation burden (to decrease variability) = higher cost Possibly degrades after being stored for a certain time→ high error rate

4.2 Clinical Study

For total mutation rate, the background (HiFi *Taq* error) to clinical sample signal ratio was high.

Therefore, only the transversion mutations A→C/T→G, (AC/TG), which were mutations not induced by *Taq* with 35 PCR cycles, will be discussed here. AC/TG mutations observed in a clinical sample would be attributable to the sample itself and not likely due to a PCR-induced mutation.

The principal finding of this study was that more women had the presence of AC/TG mutations in the mtDNA D-loop (nt16559-447 excluding D310 C-tract) if exposed to HIV/HAART during pregnancy than those unexposed. This may have also tended to be the case in the infants had there been a larger sample size since after controlling for covariates the difference approached significance. In addition, significantly more HIV/HAART exposed mothers had the presence of AC/TG mutations compared to their infants; however, this significance disappeared after controlling for covariates. In contrast, no difference was observed between the infants and their mothers in the unexposed control group. Taken together, these results would suggest that HIV/HAART-exposure during pregnancy is associated with presence of AC/TG mutations. Martin *et al.*'s findings suggest that accumulation of mtDNA mutations is solely due to NRTI-exposure and not HIV-infection (**section 1.6.6.3**) [103]. However, this finding cannot be confirmed in the context of this study since there was no HIV-infected but untreated group. The results from this study also suggest that HIV/HAART-exposure is more strongly associated with mtDNA

AC/TG mutation accumulation in mothers than it is in their infants. This may be explained by the fact that if during pregnancy, the maternal HIV pVL is under control, the embryo/foetus will have minimal to no exposure to HIV which is supported by the fact that ART is able to prevent HIV MTCT. This is unlike the mothers who prior to pregnancy may have had an uncontrolled HIV pVL since they may not have needed HAART for their own treatment.

If HAART were the determining factor for mtDNA mutation accumulation as suggested by Martin *et al.*'s findings, more mtDNA mutations in the HIV/HAART-exposed infants compared to the unexposed infants would have been expected. In addition, mtDNA mutations in the infants would have either been equally or more frequently expected compared to their mothers. This is because the infants are being exposed to equivalent or higher levels of NRTIs than their mothers. NRTIs such as AZT, 3TC and d4T readily cross the placenta (mostly via simple diffusion) and the concentrations of AZT, 3TC and d4T have been found to be approximately equal or higher in the cord blood compared to that of the maternal blood [44, 126]. 3TC has also been shown to accumulate in the amniotic fluid at concentrations 5 times higher than that of maternal blood. Another reason to expect equal or higher mtDNA mutation rates in the HIV/HAART-exposed infants compared to their mothers is that during pregnancy, embryogenesis and organogenesis occurs. During such processes, rapid mtDNA proliferation and hence mtDNA replication take place, as such, the probability that the NRTIs can affect mtDNA replication would be expected to be much higher than in an adult where the frequency of mtDNA replication is lower.

One hypothesis for the reason that more exposed mothers had AC/TG mutations than their infants was because these mutations had accumulated with age or HIV infection. A moderate correlation was found between the presence of AC/TG mutations in mothers and maternal age in the HIV/HAART-exposed group (a moderate but not significant correlation was observed between mothers from both groups and maternal age). However, once the multivariate analysis was performed, maternal age was not found to predict the presence of AC/TG mutations.

A detectable HIV pVL near delivery was found to predict maternal presence of the AC/TG mutations. This is perhaps explained by the fact that HIV may itself be a possible source of oxidative stress which may in turn lead to mtDNA mutation induction. In an *in vitro* study, during HIV infection, a significant number of HIV RNA transcripts were found in mitochondria relative to the cytoplasm and nucleus. Concurrently, mitochondrial viability decreased. The authors postulated that the HIV RNA transcripts compromise mitochondrial function [143]. With dysfunction, there may be altered OXPHOS, which may in turn lead to increased ROS and ultimately increased mtDNA mutations (**Figure 6**). Many transgenic studies have shown that antioxidants which control the levels of ROS, are decreased when Tat protein, responsible for increasing HIV mRNA transcription, is expressed. In one study, the Tat protein was found to repress the expression of the antioxidant, MnSOD in cells transfected with the *tat* gene. Concomitantly, oxidative stress was noted [144]. In addition, transgenic mice that express Tat, showed decreased levels of the antioxidant, glutathione [145]. With uncontrolled ROS levels from the decreases in antioxidants, mtDNA mutation accumulation may result.

However, relying too heavily on this data to interpret the effect of circulating HIV virus would be somewhat fallible. This is because only 7/42 (17%) of the HIV/HAART-exposed group for which mother blood samples had been collected had a detectable HIV pVL near delivery and therefore this was a small sample size.

It was surprising to find that length of HAART-exposure did not seem to predict the occurrence of AC/TG mutations. This would again suggest that HAART may not be the determinant factor of mtDNA mutation induction but that HIV infection itself may play a role.

The occurrence of AC/TG mutations may be linked to an increase in the presence of 8-oxo-dGTP, the oxidised version of dGTP. During replication, 8-oxo-dGTP is capable of pairing with dCTP and dATP with almost equal affinity (**Figure 27**). Such increase in 8-oxo-dGTP may result from increased ROS from mitochondrial dysfunction due to NRTI/HIV-exposure (**sections 1.6.5 and 1.6.6.2** and above).

Figure 27: Schematic illustration of the formation of AC/TG transversion mutations from 8-oxo-dGTP during DNA replication

Once 8-oxo-dGTP is incorporated into DNA, it becomes 8-oxo-G. 8-oxo-dGTP has equal affinity to dA and dC. This illustration is based on information from [146].

Presence of mtDNA mutations can have implications on aging and disease. The effects of these mutations may be even more pronounced if they are acquired early on in life. As alluded to in the introduction, the “mutator” mouse model showed signs of accelerated aging with a ~3 -5x rise in somatic mtDNA point mutations in adult “mutator” mice compared to wild-type mice [86].

If HAART exposure did indeed cause mtDNA mutations, in the case of the children exposed during pregnancy, mutations may accumulate and affect not only them but perhaps also their progeny; this would occur if the mtDNA mutations arose in the germline of a female. In females, the germline is formed between the 2nd and 7th month of gestation [147] and in most cases, this is when HAART is initiated. The long-term consequences of this are that disease phenotypes may not be apparent until the next generation when some of the mtDNA mutations accumulated in the germline are transmitted.

Children who are conceived on HAART may be particularly affected, not only because they will have been exposed to HAART the longest but also because of the event that occurs just prior to conception. That is, the maturation of an oocyte from a primary oocyte to prepare for fertilisation. During this stage, massive mtDNA replication takes place resulting in ~100,000 copies of mtDNA from an initial pool of ~200 [148]. It is plausible to think that the NRTIs present during this time may perhaps have an effect on mtDNA due to such high levels of replication.

Prior to this study, a difference in the number of mtDNA mutations was expected between unexposed mothers and infants. This was because studies have shown that mtDNA mutations accumulate with age. However, no such difference was found. An explanation for this is that maybe mtDNA mutations only begin to accumulate after a certain age and perhaps the mothers in this study have yet to reach this. Before this age, mtDNA mutations may be repaired efficiently, however, as cellular dysfunction begins to occur, mechanisms to prevent and/or repair mtDNA mutations or eliminate damaged mitochondria may begin to fail. Of special interest, Michikawa *et al.* observed that in fibroblasts, more than 50% of the mtDNA molecules had a T414G transversion mutation in 57% of the old individuals analysed (8 of 14 individuals >65 year of age) whereas this mutation was absent in the younger subjects [81]. This mutation was not observed in any of the clinical samples assayed here. This may be due to a difference in tissue analysed or may simply be due to the fact that none of the subjects examined in this study were old enough yet.

When analysing heteroplasmy occurring outside the C-tract (i.e. >3 mutations at the same nucleotide position in 80 clones sequenced/subject), it was only ever observed at 1 nucleotide position in a subject and was always a transition type of substitution mutation. Twice heteroplasmy was found in both mother and their child. Moreover, in 4 instances of heteroplasmy in infants (1 HIV/HAART-exposed and 3 unexposed controls) their respective mother also harboured the mutation but the frequency of this mutation did not exceed the cut-off criteria for heteroplasmy. It was interesting to observe that in

several instances mtDNA mutations present in the mother were indeed being transmitted to the child and that this transmission sometimes resulted in heteroplasmy in the child (**section 3.4.6**).

There was a discordance between 2 mother/infant pairs' consensus sequences outside the C-tract at one nucleotide position (**section 3.4.9**) both of which were in the unexposed control group. In the first case, a homoplasmic transition substitution occurred at position nt131, in which the infant had a C and mother a T. In the second case, substantial heteroplasmy resulted in a difference in consensus (16/80 and 43/80 sequences were Cs at nt146 in the infant and mother, respectively, the remaining were Ts). In the first case, it seems as though accelerated evolution is taking place; typically a mtDNA mutation usually takes 6 generations to establish itself and take foothold (i.e. to depart from heteroplasmy to homoplasmy) [149]. Furthermore, a study of 292 transmissions of mtDNA never found homoplasmic nucleotide substitutions [150]. Perhaps the difference in consensus is due to the "genetic bottleneck" phenomena taking place in the ova of these mothers (**section 1.6.2**). It would be fascinating to see whether or not a difference in base at each respective nucleotide position (i.e. nt131 and nt146) is found in other types of tissues in both these mothers and infants.

It was perplexing to find sequences (N = 5) which all harboured transition mutations multiple times (i.e. >3 transition mutations) and even more specifically, within each of these, the same type of transition mutation was harboured. In one case, the sequence harboured 30 A → G mutations. It was interesting that these types of sequences were only found in infants and not in any mothers. This may simply be because the sample size for mothers was smaller than the infants or that maybe these mutations are eliminated with age. Also, since these sequences occurred in infants in both groups, it would suggest that they are not related to HIV/HAART-exposure. These sequences may have possibly been amplified mitochondrial pseudogenes. Although several measures were taken to avoid such amplifications (**section 2.2.3**), perhaps they occurred nonetheless. When performing the BLAST search of the mtDNA region amplified, no exact match that spanned the entire region was found. However, these databases

do not contain all possible sequences and therefore some pseudogenes may be missed. It is conceivable that between individuals different pseudogenes may be present and the quality of the BLAST search is only as good as the quality of the databases that it searches within. [151]. Another possible reason for such mutations is that these occurrences may be due to a system similar to that of the APOBEC proteins at play. These are cytidine deaminases, which deaminate cytidine to produce uridine; resulting in G to A mutations in the complimentary strand during DNA replication. This group of proteins are part of the innate antiviral immune defence mechanisms. The APOBEC3G protein is the most well studied of the APOBEC group of proteins and targets single-stranded proviral DNA produced during reverse transcription of retroviral RNA genomes. It has been found to have anti-HIV activities. APOBEC3G specifically hypermutates the HIV proviral genomic DNA, which consequently generates non-infectious HIV [152]. In the case of mtDNA, other similar mechanisms may be at play for reasons at this time unknown. These mechanisms probably do not normally target self DNA but for various reasons, protection against these may have failed.

In this study, when performing the MMBA, the amount of DNA added to the PCR reaction was not normalised between samples. This may somehow affect the number of mtDNA mutations that are quantified in each sample i.e. does adding 10 ng of DNA vs. 100 ng from subject A to the assay, result in different amounts of mtDNA mutations observed? However, results from correlation calculations between presence of AC/TG mutations and DNA amount, as well as results from the incorporation of this variable into the regression model seemed to suggest that amount of DNA is not related to the number of mtDNA mutations observed. In future studies, not only will amount of DNA be normalised between samples but more precisely the amount of mtDNA will be normalised. Presently, amount of DNA was quantified based on readings from a spectrophotometer which is unable to differentiate nDNA from mtDNA. Later, amount of mtDNA would be quantified using a real-time PCR method by amplifying a small portion of the D-loop. It is worth mentioning though, that preliminary results (not shown here)

demonstrate that spectrophotometer readings can be a surrogate for mtDNA quantity since these two are highly correlated.

Mutations within the D310 C-tract were examined, although this was not the main focus of this study. This was because the frequency of PCR-induced mutations was found to be high and variable within this region when BER was determined (**section 2**).

Moreover, D310 length heteroplasmy has been described as “intensively complicated and may be difficult to interpret” [153]. In agreement with others [118, 150], a dramatic decrease in stability was observed in our study when the consensus genotype is more than 7C long (i.e. between nt303 – 309). Consistent with these results, in an *in vitro* study, POLG error rates increased with increasing length of homopolymeric runs, because of low frameshift fidelity. The reason for this phenomenon is most likely due to polymerase slippage, where there is a transient misalignment of the template and the elongating strand [62]. Additionally, if a mutation occurred in this C-tract, it usually involved an insertion of a C between nt303-309. This was a striking finding since a deletion of a C usually occurred when BER was determined. This would suggest that the C insertion pattern being observed in the clinical samples is most likely a natural phenomenon and not PCR-induced.

Many groups have studied the D310 region, especially those in the field of cancer research. Sanchez-Céspedes *et al.* suggested that D310 is a mutational hotspot for primary tumours [124]. Mutations in D310 have been identified in but are not limited to bladder, rectal, sigmoid, thyroid and cervical cancer [154-156]. However, no clear link has been made between the frequency of mutations and the severity of the cancers, therefore the study of D310 in cancers is mostly of exploratory nature.

It is not well understood if changes in the D310 will alter mtDNA replication. This tract lies within the CSB2 region and, with CS1, as well as CSB3 (which are outside of the region assayed here) contribute to the formation of persistent RNA-DNA hybrid that leads to the initiation of mtDNA replication. It has been postulated that changes within D310 may provide a growth advantage observed in tumour cells

[124]. At this stage, it is difficult to interpret the findings from this study and to apply it to cancer research since no comprehensive medical history on cancer was collected from the subjects.

Lastly, the location of certain mutations and mutations at specific nucleotide positions were analysed. The location of the AC/TG mutations observed were noted however no apparent pattern was found. Specific mutations described in the literature in association with aging (i.e. C114T, C150T, A189G, T195C, T408A, T414G and T414C) were examined. Nucleotide positions 408 and 414 are located within the promoter region of the L-strand (LSP), while 150, 189, and 195 are located close to the secondary (nt147 – 151) or primary origin (nt191) of mtDNA H-strand synthesis (**Figure 7**) [115]. Therefore, mutations occurring at any one of these locations might have an effect on mtDNA transcription and replication. However, none of these mutations seemed to occur at frequencies above those observed at other locations in the region assayed. It is difficult to study these specific mutations since these are all transition substitution mutations (except for the T414G mutation). At present with our current method of assaying, “real” mutations and those induced by the PCR polymerase are indistinguishable from one another.

4.2.1 Limitations

Although statistical significant differences between and within groups were observed in terms of the presence or absence of AC/TG mutations in within a sample, these mutations were very rare. To illustrate this, in 42 HIV/HAART-exposed mothers, these mutations were observed 21 times, therefore an overall mutation rate of 0.14 AC/TG mutations/10,000 bp (21 mutations/(42 subjects *454 bp*80 sequences)/10,000 bp)) or in other words, 0.23 AC/TG mutations/(mtDNA genome). This was in contrast to the unexposed control mothers (N = 39) in which 7 mutations were found and therefore, 0.05 AC/TG mutations/10,000 bp or 0.08 AC/TG mutations/(mtDNA genome). It is consequently difficult to assign a definite biological or clinical significance to such results. Moreover, the assay allowed the detection of mutations in only 2.7% (454/16559 bp) of the genome, which in addition, only consisted of the non-

coding region. Mutation rates elsewhere in the genome and the mutation rate of other mutations not quantifiable by the MMBA (i.e. mutations other than AC/TG mutations) may vary considerably. For example, since the D-loop is non-coding it may be more permissive to mutations than the coding regions and therefore, show higher mutation rates.

While blood was used in this study, it is known that, unlike adipose, brain and muscle tissues, it is not the most sensitive tissue to mitochondrial toxicity due to its rapid turnover. Also, depending on the NRTI, the affected tissue may be different (as is demonstrated in **Table 3** on the frequency of clinical symptoms with various NRTIs). A study comparing matched blood and adipocyte tissue collected from subjects treated with various NRTIs found that ddI and d4T were strongly associated with mtDNA depletion in adipocyte tissue. However, only ddI was found to be associated with mtDNA depletion in blood. The authors suggest that the determinants are tissue specific [157]. Similar findings were reported in another study by Maagard *et al.*; whereby mtDNA quantity between PBMC and muscle matched samples was not significantly correlated within the NRTI-treated group [158]. Furthermore, studies that have examined mtDNA deletions in different tissues from various age groups have reported that the “common” mtDNA 4977 bp deletion was found to accumulate with age but that this deletion failed to be observed in blood of elderly subjects, even in a 98 year old individual [159]. Therefore, although a significant difference was not observed between, for example, the infants in the HIV/HAART-exposed and the unexposed control group, perhaps a difference would be detected if other tissues were examined.

However, obtaining anything else but blood in “healthy” newborns could be unethical and sometimes difficult. Collecting any tissue that requires invasive procedures (such as muscle biopsies) would be hindered by physicians unwilling to perform these or parents unwilling to consent to. Therefore, in the context of collecting samples from newborns, blood is the most suitable choice and readily available tissue. The Côté lab is currently also collecting mouth swabs to determine if it can be used as a

supplementary or alternative sample tissue. This type of tissue sample is easy to collect with little to no risk to the subject.

In our study, we were unable to separate the effects of HIV infection from those of HAART. In order to do so, another study group would have been required; that being infants born to HIV-infected untreated women. However, in Canada, in an era where HAART is widely available and is the best method of preventing HIV MTCT, not delivering PMTCT services to an HIV-infected pregnant woman would be a breach of ethics.

Although there are several limitations to this study, the results demonstrate that HIV/HAART may have an effect on mtDNA AC/TG mutations in mothers and perhaps infants exposed during pregnancy. Consequently, further investigations are warranted which need not be restricted to simply the realm of pregnancies but rather to the broader field of HIV/HAART, especially patients on long-term therapy.

4.3 Future Directions

4.3.1 Different MtDNA Sequencing Method

As has been demonstrated, cloning and sequencing using polymerases carries a high error rate which is extremely problematic when attempting to detect very low levels of mutations and accurately quantifying mtDNA heteroplasmy. This would still be the case with the newer sequencing technologies developed in the last few years such as 454 (Roche) and Illumina (Solexa) sequencing since they are again based on PCR. Nanopore sequencing may resolve these problems. It is yet to be on the market but Oxford Nanopore Technologies® and IBM are currently developing the technology. This high-throughput technology is based on the use of a nanopore (a very small hole) by which a single molecule of DNA is forced through electrophoretically. As this occurs, there is a small change in the current across the pore that is detected. Each nucleotide gives off a characteristic signal that can be distinguished. This technology has the advantage over other sequencing methods commercially available in that it requires no amplification or labelling step. This decreases costs considerably. The BigDye® labelling in the Sanger

cycle sequencing method that was used here is expensive. The detection of individual DNA molecules provided by nanopore sequencing would be equivalent to the amplification and cloning steps (which serves to separate out individual DNA molecules from one another) in the MMBA; which is currently our source of error. Particularly, long reads of ~10,000 – 50,000 bases may be possible with nanopore sequencing, allowing for the simplification of the sequence assembly process [160]. In the case of mtDNA, it would enable a mtDNA molecule to be sequenced in its entirety, presenting a definite advantage. In addition, several times more mtDNA genomes would be sequenced per sample. We are currently limited to assaying 80 sequences/sample due to costs and labour considerations. A further potential for nanopore sequencing is that many samples could be simultaneously run on a single nanopore chip by DNA barcoding (the addition of a short DNA sequence identifier tag) that would enable each sample to be distinguished from one another, much like is currently the case with deep sequencing.

One impediment to the use of nanopore sequencing that can be foreseen in the context of mtDNA sequencing is that the mtDNA would have to be isolated from the nDNA to prevent sequencing of the chromosomal genome. If nDNA were also sequenced, it would decrease the effectiveness and resolution of the method. The problem with mtDNA isolation is that it requires large amounts of initial sample, it is particularly difficult, laborious and time-consuming with homogenisation, many centrifugation steps and digestion of contaminating nDNA [105, 161].

Nevertheless, as new methods and protocols are formulated, mtDNA isolation may become an easier task. With this in mind and nanopore sequencing in development, there will most likely be a bright future for mtDNA mutation research.

4.3.2 Future Clinical Studies

The results from the study on mtDNA mutations in infants HIV/HAART-exposed *in utero* showed that AC/TG mutations occurred infrequently, pushing the detection limit of our assay. To make a stronger

link between mtDNA mutation and HAART-exposure, it would be of interest to study mtDNA mutations in an investigation that incorporated:

1. patients on long-term HAART
2. HIV-infected but treatment naïve patients
3. HIV-uninfected and treatment naïve individuals

If mtDNA mutations did indeed accumulate with HAART exposure, it would be most apparent amongst patients on long-term HAART compared to an embryo/foetus who is only exposed through their mother for a short-period of time. In our study, many women were on HAART to prevent MTCT and not for treatment purposes, therefore, it might be expected that mtDNA mutations may not yet have accumulated with such a short period of exposure. Also, by studying HIV-infected but treatment naïve patients, the effects of HIV alone could be studied. As discussed previously, HIV infection itself may be a source of oxidative stress which may induce mtDNA mutations.

Although this type of study has been conducted once before by Martin *et al.*, repeating this would be of great benefit. The sample size in Martin *et al.*'s study was small (N = 16 in the NRTI-exposed group) and did not have an HIV-uninfected control group [103].

4.3.3 Future Basic Science Studies

The over arching aim of studying mtDNA mutations in the context of *in utero* HAART-exposure was not only to study mtDNA mutations with HAART, but also to specifically investigate the mtDNA mutation induction capacity of different NRTIs.

It is known that within the clinical setting, not all NRTIs are equally toxic in terms of their propensity to cause mitochondrial toxicity in HIV/HAART-exposed adults or children. This may also be the case in terms of their mutagenicity.

Newer NRTIs that are already available on the market are not currently being widely administered to pregnant women infected with HIV because clinical trials with these drugs have not been conducted and

will most likely never be. Researchers, physicians and pharmaceutical companies are reluctant to conduct clinical trials on pregnant women because of the possible risks of teratogenicity. Therefore, despite AZT's known toxicities, it is still the first line NRTI used since there is little incentive to change the current HAART regimen containing NRTIs AZT and 3TC. At this stage, it is felt that the little risk to the infants with this HAART regimen clearly is outweighed by the more serious risks of acquiring HIV. Perhaps though, lowering the risks of side effects should be considered to further reduce morbidity from *in utero* HAART-exposure. It should be noted however, that despite d4T's known short-term (serious risks of fatal lactic acidosis complications) and long-term (lipoatrophy and neuropathy) toxicity, d4T in combination with 3TC are the first line regimen that many women will receive in developing countries (this regimen is rarely used in developed countries for reasons of toxicity). This is due to financial and accessibility reasons [162]. Since clinical studies are not easily feasible in the context of pregnancies, studies with tissue culture systems would be of great value.

Many studies have been carried out in different tissue culture systems with various NRTIs, however, none of them have ever investigated induction of mtDNA mutations [163]. It would be interesting to study the effects of different NRTIs on cell lines susceptible to NRTI-induced mitochondrial toxicity (such as HepG2, a hepatocellular carcinoma cell line). Moreover, these should be cultured longitudinally over several weeks to mimic the long-term treatment that ART entails. With such a study, the effects of NRTIs alone could be investigated without HIV infection as a confounder.

4.4 Conclusions

The risk of transmission of HIV during pregnancy from mother to child has dramatically decreased from ~ 25 to $\leq 1\%$ with HAART. Despite this beneficial effect, NRTIs are thought to adversely affect mtDNA and may induce point mutations. Additionally, mutations in mtDNA have been linked to aging and age-associated diseases. Since UNAIDS policy is to eliminate HIV MTCT by 2015, this means that the number of HIV-infected pregnant women receiving HAART or ART will rise considerably. As a result, a growing

number of infants will be exposed to HAART/ART and therefore, awareness of their possible side effects is important.

The goal of this thesis was to investigate mtDNA mutation burden in blood of infants and HIV-infected mothers exposed to HAART during pregnancy and to compare these with infants and mothers HIV/HAART-unexposed. To examine mtDNA mutation burden, an assay was developed and the background error rate of this assay was determined. The background (HiFi *Taq* error rate) to clinical mutation signal ratio was high. Therefore, only AC/TG mutations which were mutations not induced by HiFi *Taq* under our assay conditions were analysed.

The clinical study found that the presence of AC/TG mutations in mtDNA was associated with HIV/HAART exposure in mothers and perhaps in the infants. In addition, it found that having a detectable HIV pVL near delivery predicted the presence of AC/TG mutations in the mother.

Given that mtDNA mutations have been associated with aging and age-associated diseases, this raises possible concerns about the long-term impact of HAART exposure. Since only a few specific mutations were analysed in a portion of the mtDNA genome, further studies are warranted to examine all types of mutations throughout the mtDNA genome. This would enable a more in-depth understanding of the effects of HIV/HAART exposure on mtDNA.

References

1. Gray, G.E. and J.A. McIntyre, *HIV and pregnancy*. BMJ, 2007. **334**(7600): p. 950-3.
2. UNAIDS, *2008 Report on the global AIDS epidemic-Executive Summary*. 2008: Geneva.
3. McIntyre, J., *HIV in Pregnancy: a Review*. 1998, WHO/UNAIDS: Geneva.
4. Newell, M.L., et al., *Mortality of infected and uninfected infants born to HIV-infected mothers in Africa: a pooled analysis*. Lancet, 2004. **364**(9441): p. 1236-43.
5. UNAIDS, *AIDS Epidemic Update 2009*. 2009: Geneva.
6. WHO/UNAIDS, *PMTCT Strategic Vision 2010-2015: preventing mother-to-child transmission of HIV to reach the UNGASS and Millennium Development Goals*. 2010, WHO/UNAIDS: Geneva.
7. Public Health Agency of Canada, *Summary: Estimates of HIV Prevalence and Incidence in Canada, 2008*. 2009.
8. Public Health Agency of Canada, *HIV/AIDS Epi Update*. 2007.
9. *STI/HIV Annual Report 2008*. 2008, BC Centre for Disease Control Vancouver.
10. Krogstad, P., *Molecular biology of the human immunodeficiency virus: current and future targets for intervention*. Semin Pediatr Infect Dis, 2003. **14**(4): p. 258-68.
11. Clapham, P.R. and A. McKnight, *Cell surface receptors, virus entry and tropism of primate lentiviruses*. J Gen Virol, 2002. **83**(Pt 8): p. 1809-29.
12. Gelderblom, H.C., et al., *Viral complementation allows HIV-1 replication without integration*. Retrovirology, 2008. **5**: p. 60.
13. Pollard, V.W. and M.H. Malim, *The HIV-1 Rev protein*. Annu Rev Microbiol, 1998. **52**: p. 491-532.
14. Sierra, S., B. Kupfer, and R. Kaiser, *Basics of the virology of HIV-1 and its replication*. J Clin Virol, 2005. **34**(4): p. 233-44.

15. Burger, S. and M.A. Poles, *Natural history and pathogenesis of human immunodeficiency virus infection*. Semin Liver Dis, 2003. **23**(2): p. 115-24.
16. Weber, J., *The pathogenesis of HIV-1 infection*. Br Med Bull, 2001. **58**: p. 61-72.
17. Mindel, A. and M. Tenant-Flowers, *ABC of AIDS: Natural history and management of early HIV infection*. BMJ, 2001. **322**(7297): p. 1290-3.
18. Touloumi, G. and A. Hatzakis, *Natural history of HIV-1 infection*. Clin Dermatol, 2000. **18**(4): p. 389-99.
19. Zwahlen, M. and M. Egger, *Progression and Mortality of Untreated HIV-positive Individuals Living in Resource-Limited Settings: Update of Literature and Evidence Synthesis 2006*, Institute of Social and Preventative Medicine: Bern.
20. Piacenti, F.J., *An update and review of antiretroviral therapy*. Pharmacotherapy, 2006. **26**(8): p. 1111-33.
21. Kalkut, G., *Antiretroviral therapy: an update for the non-AIDS specialist*. Curr Opin Oncol, 2005. **17**(5): p. 479-84.
22. Antiretroviral Therapy Cohort Collaboration, *Life expectancy of individuals on combination antiretroviral therapy in high-income countries: a collaborative analysis of 14 cohort studies*. Lancet, 2008. **372**(9635): p. 293-9.
23. WHO, *Rapid Advice: Antiretroviral therapy for HIV infection in adults and adolescents*. 2009.
24. Bailey, A.C. and M. Fisher, *Current use of antiretroviral treatment*. Br Med Bull, 2008. **87**: p. 175-92.
25. El Safadi, Y., V. Vivet-Boudou, and R. Marquet, *HIV-1 reverse transcriptase inhibitors*. Appl Microbiol Biotechnol, 2007. **75**(4): p. 723-37.
26. UNAIDS, *Rates of Mother-to-Child Transmission and the Impact of Different PMTCT Regimens*. 2005: Geneva.
27. Newell, M.L., *Mechanisms and timing of mother-to-child transmission of HIV-1*. AIDS, 1998. **12**(8): p. 831-7.

28. Rouzioux, C., et al., *Estimated timing of mother-to-child human immunodeficiency virus type 1 (HIV-1) transmission by use of a Markov model. The HIV Infection in Newborns French Collaborative Study Group.* Am J Epidemiol, 1995. **142**(12): p. 1330-7.
29. Connor, E.M., et al., *Reduction of maternal-infant transmission of human immunodeficiency virus type 1 with zidovudine treatment. Pediatric AIDS Clinical Trials Group Protocol 076 Study Group.* N Engl J Med, 1994. **331**(18): p. 1173-80.
30. Rigopoulos, D., et al., *AIDS in pregnancy, part II: Treatment in the era of highly active antiretroviral therapy and management of obstetric, anesthetic, and pediatric issues.* Skinmed, 2007. **6**(2): p. 79-84.
31. Cooper, E.R., et al., *Combination antiretroviral strategies for the treatment of pregnant HIV-1-infected women and prevention of perinatal HIV-1 transmission.* J Acquir Immune Defic Syndr, 2002. **29**(5): p. 484-94.
32. Sperling, R.S., et al., *Maternal viral load, zidovudine treatment, and the risk of transmission of human immunodeficiency virus type 1 from mother to infant. Pediatric AIDS Clinical Trials Group Protocol 076 Study Group.* N Engl J Med, 1996. **335**(22): p. 1621-9.
33. Lorenzi, P., et al., *Antiretroviral therapies in pregnancy: maternal, fetal and neonatal effects. Swiss HIV Cohort Study, the Swiss Collaborative HIV and Pregnancy Study, and the Swiss Neonatal HIV Study.* AIDS, 1998. **12**(18): p. F241-7.
34. Burdge, D.R., et al., *Canadian consensus guidelines for the management of pregnancy, labour and delivery and for postpartum care in HIV-positive pregnant women and their offspring (summary of 2002 guidelines).* CMAJ, 2003. **168**(13): p. 1671-4.
35. Group, U.P.H.G.W., *Public Health Service Task Force Recommendations for Use of Antiretroviral Drugs in Pregnant HIV-Infected Women for Maternal Health and Interventions to Reduce Perinatal HIV Transmission in the United States.* . 2009.
36. Petra Study Team, *Efficacy of three short-course regimens of zidovudine and lamivudine in preventing early and late transmission of HIV-1 from mother to child in Tanzania, South Africa, and Uganda (Petra study): a randomised, double-blind, placebo-controlled trial.* Lancet, 2002. **359**(9313): p. 1178-86.
37. McIntyre, J., *Use of antiretrovirals during pregnancy and breastfeeding in low-income and middle-income countries.* Curr Opin HIV AIDS, 2010. **5**(1): p. 48-53.

38. Guay, L.A., et al., *Intrapartum and neonatal single-dose nevirapine compared with zidovudine for prevention of mother-to-child transmission of HIV-1 in Kampala, Uganda: HIVNET 012 randomised trial*. Lancet, 1999. **354**(9181): p. 795-802.
39. Dao, H., et al., *International recommendations on antiretroviral drugs for treatment of HIV-infected women and prevention of mother-to-child HIV transmission in resource-limited settings: 2006 update*. Am J Obstet Gynecol, 2007. **197**(3 Suppl): p. S42-55.
40. Foster, C. and H. Lyall, *HIV and mitochondrial toxicity in children*. J Antimicrob Chemother, 2008. **61**(1): p. 8-12.
41. Kamemoto, L.E., B. Shiramizu, and M. Gerschenson, *HIV-associated mitochondrial toxicity in pregnancy*. Mitochondrion, 2004. **4**(2-3): p. 153-62.
42. Funk, M.J., et al., *Mitochondrial disorders among infants exposed to HIV and antiretroviral therapy*. Drug Saf, 2007. **30**(10): p. 845-59.
43. Rosso, R., et al., *Fatal lactic acidosis and mimicking Guillain-Barre syndrome in an adolescent with human immunodeficiency virus infection*. Pediatr Infect Dis J, 2003. **22**(7): p. 668-70.
44. Mirochnick, M. and E. Capparelli, *Pharmacokinetics of antiretrovirals in pregnant women*. Clin Pharmacokinet, 2004. **43**(15): p. 1071-87.
45. John, M., et al., *Chronic hyperlactatemia in HIV-infected patients taking antiretroviral therapy*. AIDS, 2001. **15**(6): p. 717-23.
46. Barret, B., et al., *Persistent mitochondrial dysfunction in HIV-1-exposed but uninfected infants: clinical screening in a large prospective cohort*. AIDS, 2003. **17**(12): p. 1769-85.
47. Culnane, M., et al., *Lack of long-term effects of in utero exposure to zidovudine among uninfected children born to HIV-infected women*. Pediatric AIDS Clinical Trials Group Protocol 219/076 Teams. JAMA, 1999. **281**(2): p. 151-7.
48. Aldrovandi, G.M., et al., *Antiretroviral exposure and lymphocyte mtDNA content among uninfected infants of HIV-1-infected women*. Pediatrics, 2009. **124**(6): p. e1189-97.
49. Cote, H.C., et al., *Perinatal exposure to antiretroviral therapy is associated with increased blood mitochondrial DNA levels and decreased mitochondrial gene expression in infants*. J Infect Dis, 2008. **198**(6): p. 851-9.

50. McComsey, G.A., et al., *Increased mtDNA levels without change in mitochondrial enzymes in peripheral blood mononuclear cells of infants born to HIV-infected mothers on antiretroviral therapy*. HIV Clin Trials, 2008. **9**(2): p. 126-36.
51. Chinnery, P.F. and E.A. Schon, *Mitochondria*. J Neurol Neurosurg Psychiatry, 2003. **74**(9): p. 1188-99.
52. Schapira, A.H., *Mitochondrial disease*. Lancet, 2006. **368**(9529): p. 70-82.
53. Stringer, H., *Mitochondrial DNA Alterations and Statin-Induced Myopathy*, in *Pathology and Laboratory Medicine*. 2009, University of British Columbia: Vancouver.
54. Druzhyna, N.M., G.L. Wilson, and S.P. LeDoux, *Mitochondrial DNA repair in aging and disease*. Mech Ageing Dev, 2008. **129**(7-8): p. 383-90.
55. Wallace, D.C., *Mitochondria as chi*. Genetics, 2008. **179**(2): p. 727-35.
56. Greaves, L.C. and R.W. Taylor, *Mitochondrial DNA mutations in human disease*. IUBMB Life, 2006. **58**(3): p. 143-51.
57. Sutovsky, P., et al., *Ubiquitin tag for sperm mitochondria*. Nature, 1999. **402**(6760): p. 371-2.
58. DiMauro, S., *Mitochondrial DNA medicine*. Biosci Rep, 2007. **27**(1-3): p. 5-9.
59. Chinnery, P.F., et al., *Molecular pathology of MELAS and MERRF. The relationship between mutation load and clinical phenotypes*. Brain, 1997. **120** (Pt 10): p. 1713-21.
60. McFarland, R., R.W. Taylor, and D.M. Turnbull, *Mitochondrial disease--its impact, etiology, and pathology*. Curr Top Dev Biol, 2007. **77**: p. 113-55.
61. Falkenberg, M., N.G. Larsson, and C.M. Gustafsson, *DNA replication and transcription in mammalian mitochondria*. Annu Rev Biochem, 2007. **76**: p. 679-99.
62. Longley, M.J., et al., *The fidelity of human DNA polymerase gamma with and without exonucleolytic proofreading and the p55 accessory subunit*. J Biol Chem, 2001. **276**(42): p. 38555-62.
63. Kang, D. and N. Hamasaki, *Alterations of mitochondrial DNA in common diseases and disease states: aging, neurodegeneration, heart failure, diabetes, and cancer*. Curr Med Chem, 2005. **12**(4): p. 429-41.

64. Bogenhagen, D. and D.A. Clayton, *Mouse L cell mitochondrial DNA molecules are selected randomly for replication throughout the cell cycle*. Cell, 1977. **11**(4): p. 719-27.
65. Chinnery, P.F. and D.C. Samuels, *Relaxed replication of mtDNA: A model with implications for the expression of disease*. Am J Hum Genet, 1999. **64**(4): p. 1158-65.
66. Clay Montier, L.L., J.J. Deng, and Y. Bai, *Number matters: control of mammalian mitochondrial DNA copy number*. J Genet Genomics, 2009. **36**(3): p. 125-31.
67. Taylor, R.W. and D.M. Turnbull, *Mitochondrial DNA mutations in human disease*. Nat Rev Genet, 2005. **6**(5): p. 389-402.
68. Berneburg, M., et al., *'To repair or not to repair - no longer a question': repair of mitochondrial DNA shielding against age and cancer*. Exp Dermatol, 2006. **15**(12): p. 1005-15.
69. Chinnery, P.F., *Mitochondrial Disorders Overview*, ed. P. RA, et al. 2000, Seattle, Washington, USA: Gene Reviews-NCBI Bookshelf.
70. Meissner, C., et al., *The 4977 bp deletion of mitochondrial DNA in human skeletal muscle, heart and different areas of the brain: a useful biomarker or more?* Exp Gerontol, 2008. **43**(7): p. 645-52.
71. Mohamed, S.A., et al., *Detection of the 4977 bp deletion of mitochondrial DNA in different human blood cells*. Exp Gerontol, 2004. **39**(2): p. 181-8.
72. Brown, W.M., M. George, Jr., and A.C. Wilson, *Rapid evolution of animal mitochondrial DNA*. Proc Natl Acad Sci U S A, 1979. **76**(4): p. 1967-71.
73. Montoya, J., et al., *20 years of human mtDNA pathologic point mutations: carefully reading the pathogenicity criteria*. Biochim Biophys Acta, 2009. **1787**(5): p. 476-83.
74. Poyton, R.O., K.A. Ball, and P.R. Castello, *Mitochondrial generation of free radicals and hypoxic signaling*. Trends Endocrinol Metab, 2009. **20**(7): p. 332-40.
75. Kohler, J.J. and W. Lewis, *A brief overview of mechanisms of mitochondrial toxicity from NRTIs*. Environ Mol Mutagen, 2007. **48**(3-4): p. 166-72.
76. Arnett, S.D., et al., *Determination of 8-oxoguanine and 8-hydroxy-2'-deoxyguanosine in the rat cerebral cortex using microdialysis sampling and capillary electrophoresis with electrochemical detection*. J Chromatogr B Analyt Technol Biomed Life Sci, 2005. **827**(1): p. 16-25.

77. Pfeffer, G., et al., *Ophthalmoplegia and ptosis: mitochondrial toxicity in patients receiving HIV therapy*. *Neurology*, 2009. **73**(1): p. 71-2.
78. Maagaard, A. and D. Kvale, *Mitochondrial toxicity in HIV-infected patients both off and on antiretroviral treatment: a continuum or distinct underlying mechanisms?* *J Antimicrob Chemother*, 2009. **64**(5): p. 901-9.
79. Lewis, W., B.J. Day, and W.C. Copeland, *Mitochondrial toxicity of NRTI antiviral drugs: an integrated cellular perspective*. *Nat Rev Drug Discov*, 2003. **2**(10): p. 812-22.
80. Bogliolo, M., et al., *Detection of the '4977 bp' mitochondrial DNA deletion in human atherosclerotic lesions*. *Mutagenesis*, 1999. **14**(1): p. 77-82.
81. Michikawa, Y., et al., *Aging-dependent large accumulation of point mutations in the human mtDNA control region for replication*. *Science*, 1999. **286**(5440): p. 774-9.
82. Van Remmen, H., et al., *Knockout mice heterozygous for Sod2 show alterations in cardiac mitochondrial function and apoptosis*. *Am J Physiol Heart Circ Physiol*, 2001. **281**(3): p. H1422-32.
83. Van Remmen, H., et al., *Life-long reduction in MnSOD activity results in increased DNA damage and higher incidence of cancer but does not accelerate aging*. *Physiol Genomics*, 2003. **16**(1): p. 29-37.
84. Jang, Y.C. and H.V. Remmen, *The mitochondrial theory of aging: Insight from transgenic and knockout mouse models*. *Exp Gerontol*, 2009.
85. Schriener, S.E., et al., *Extension of murine life span by overexpression of catalase targeted to mitochondria*. *Science*, 2005. **308**(5730): p. 1909-11.
86. Trifunovic, A., et al., *Premature ageing in mice expressing defective mitochondrial DNA polymerase*. *Nature*, 2004. **429**(6990): p. 417-23.
87. Kujoth, G.C., et al., *Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging*. *Science*, 2005. **309**(5733): p. 481-4.
88. Miller, R.A., *Evaluating evidence for aging*. *Science*, 2005. **310**(5747): p. 441-3; author reply 441-3.

89. Wallace, D.C., *A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: a dawn for evolutionary medicine*. Annu Rev Genet, 2005. **39**: p. 359-407.
90. Fliss, M.S., et al., *Facile detection of mitochondrial DNA mutations in tumors and bodily fluids*. Science, 2000. **287**(5460): p. 2017-9.
91. Chan, D.C., *Mitochondria: dynamic organelles in disease, aging, and development*. Cell, 2006. **125**(7): p. 1241-52.
92. Coskun, P.E., M.F. Beal, and D.C. Wallace, *Alzheimer's brains harbor somatic mtDNA control-region mutations that suppress mitochondrial transcription and replication*. Proc Natl Acad Sci U S A, 2004. **101**(29): p. 10726-31.
93. Santos, C., et al., *Mitochondrial DNA mutations in cancer: a review*. Curr Top Med Chem, 2008. **8**(15): p. 1351-66.
94. Lu, J., L.K. Sharma, and Y. Bai, *Implications of mitochondrial DNA mutations and mitochondrial dysfunction in tumorigenesis*. Cell Res, 2009. **19**(7): p. 802-15.
95. Petros, J.A., et al., *mtDNA mutations increase tumorigenicity in prostate cancer*. Proc Natl Acad Sci U S A, 2005. **102**(3): p. 719-24.
96. Dalakas, M.C., et al., *Mitochondrial myopathy caused by long-term zidovudine therapy*. N Engl J Med, 1990. **322**(16): p. 1098-105.
97. Bienstock, R.J. and W.C. Copeland, *Molecular insights into NRTI inhibition and mitochondrial toxicity revealed from a structural model of the human mitochondrial DNA polymerase*. Mitochondrion, 2004. **4**(2-3): p. 203-13.
98. Esser, S., et al., *Side effects of HIV therapy*. J Dtsch Dermatol Ges, 2007. **5**(9): p. 745-54.
99. Song, S., et al., *DNA precursor asymmetries in mammalian tissue mitochondria and possible contribution to mutagenesis through reduced replication fidelity*. Proc Natl Acad Sci U S A, 2005. **102**(14): p. 4990-5.
100. Pinti, M., P. Salomoni, and A. Cossarizza, *Anti-HIV drugs and the mitochondria*. Biochim Biophys Acta, 2006. **1757**(5-6): p. 700-7.
101. Becher, F., et al., *Significant levels of intracellular stavudine triphosphate are found in HIV-infected zidovudine-treated patients*. AIDS, 2003. **17**(4): p. 555-61.

102. Bonora, S., et al., *Detection of stavudine concentrations in plasma of HIV-infected patients taking zidovudine*. AIDS, 2004. **18**(3): p. 577-8.
103. Martin, A.M., et al., *Accumulation of mitochondrial DNA mutations in human immunodeficiency virus-infected patients treated with nucleoside-analogue reverse-transcriptase inhibitors*. Am J Hum Genet, 2003. **72**(3): p. 549-60.
104. Moraes, C.T., et al., *Techniques and pitfalls in the detection of pathogenic mitochondrial DNA mutations*. J Mol Diagn, 2003. **5**(4): p. 197-208.
105. Kraytsberg, Y., A. Nicholas, and K. Khrapko, *Are somatic mitochondrial DNA mutations relevant to our health? A challenge for mutation analysis techniques*. Expert Opinion in Medical Diagnosis, 2007. **1**(1): p. 109-116.
106. Wilding, C.S., et al., *Mitochondrial DNA mutations in individuals occupationally exposed to ionizing radiation*. Radiat Res, 2006. **165**(2): p. 202-7.
107. Del Bo, R., et al., *Evidence and age-related distribution of mtDNA D-loop point mutations in skeletal muscle from healthy subjects and mitochondrial patients*. J Neurol Sci, 2002. **202**(1-2): p. 85-91.
108. Kraytsberg, Y., et al., *Single molecule PCR in mtDNA mutational analysis: Genuine mutations vs. damage bypass-derived artifacts*. Methods, 2008. **46**(4): p. 269-73.
109. Drake, J.W., *A constant rate of spontaneous mutation in DNA-based microbes*. Proc Natl Acad Sci U S A, 1991. **88**(16): p. 7160-4.
110. Vermulst, M., et al., *Mitochondrial point mutations do not limit the natural lifespan of mice*. Nat Genet, 2007. **39**(4): p. 540-3.
111. Vermulst, M., J.H. Bielas, and L.A. Loeb, *Quantification of random mutations in the mitochondrial genome*. Methods, 2008. **46**(4): p. 263-8.
112. Kieleczawa, J., *Effect of primer proximity to a difficult-to-sequence region on read length and sequence quality*. J Biomol Tech, 2008. **19**(5): p. 335-41.
113. Greaves, L.C., et al., *Quantification of mitochondrial DNA mutation load*. Aging Cell, 2009. **8**(5): p. 566-72.

114. Stoneking, M., et al., *Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes*. Am J Hum Genet, 1991. **48**(2): p. 370-82.
115. Chomyn, A. and G. Attardi, *MtDNA mutations in aging and apoptosis*. Biochem Biophys Res Commun, 2003. **304**(3): p. 519-29.
116. Cline, J., J.C. Braman, and H.H. Hogrefe, *PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases*. Nucleic Acids Res, 1996. **24**(18): p. 3546-51.
117. Attardi, G., *Role of mitochondrial DNA in human aging*. Mitochondrion, 2002. **2**(1-2): p. 27-37.
118. Forster, L., et al., *Evaluating length heteroplasmy in the human mitochondrial DNA control region*. Int J Legal Med, 2010. **124**(2): p. 133-42.
119. Clarke, L.A., et al., *PCR amplification introduces errors into mononucleotide and dinucleotide repeat sequences*. Mol Pathol, 2001. **54**(5): p. 351-3.
120. Wang, Y., et al., *Muscle-specific mutations accumulate with aging in critical human mtDNA control sites for replication*. Proc Natl Acad Sci U S A, 2001. **98**(7): p. 4022-7.
121. Mourier, T., et al., *The Human Genome Project reveals a continuous transfer of large mitochondrial fragments to the nucleus*. Mol Biol Evol, 2001. **18**(9): p. 1833-7.
122. Andrews, R.M., et al., *Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA*. Nat Genet, 1999. **23**(2): p. 147.
123. Lim, K.S., et al., *Pitfalls in the denaturing high-performance liquid chromatography analysis of mitochondrial DNA mutation*. J Mol Diagn, 2008. **10**(1): p. 102-8.
124. Sanchez-Cespedes, M., et al., *Identification of a mononucleotide repeat as a major target for mitochondrial DNA alterations in human tumors*. Cancer Res, 2001. **61**(19): p. 7015-9.
125. Andre, P., et al., *Fidelity and mutational spectrum of Pfu DNA polymerase on a human mitochondrial DNA sequence*. Genome Res, 1997. **7**(8): p. 843-52.
126. Chappuy, H., et al., *Maternal-fetal transfer and amniotic fluid accumulation of nucleoside analogue reverse transcriptase inhibitors in human immunodeficiency virus-infected pregnant women*. Antimicrob Agents Chemother, 2004. **48**(11): p. 4332-6.

127. Mussi-Pinhata, M.M., et al., *Lower respiratory tract infections among human immunodeficiency virus-exposed, uninfected infants*. Int J Infect Dis. 1993. **17**(3): p. 165-7.
128. van Belle, G. and D.C. Martin, *Sample Size as a Function of Coefficient of Variation and Ratio of Means*. The American Statistician, 1993. **47**(3): p. 165-7.
129. Chinnery, P.F., et al., *The inheritance of mitochondrial DNA heteroplasmy: random drift, selection or both?* Trends Genet, 2000. **16**(11): p. 500-5.
130. Dilyara, G., et al., *Systemic effects of smoking*. Chest, 2007. **131**(5): p. 1557-66.
131. Mehra, R., et al., *The association between marijuana smoking and lung cancer: a systematic review*. Arch Intern Med, 2006. **166**(13): p. 1359-67.
132. Mannelli, P., et al., *Opioid use affects antioxidant activity and purine metabolism: preliminary results*. Hum Psychopharmacol, 2009. **24**(8): p. 666-75.
133. Kovacic, P., *Role of oxidative metabolites of cocaine in toxicity and addiction: oxidative stress and electron transfer*. Med Hypotheses, 2005. **64**(2): p. 350-6.
134. Yamamoto, B.K., A. Moszczynska, and G.A. Gudelsky, *Amphetamine toxicities: classical and emerging mechanisms*. Ann N Y Acad Sci. **1187**: p. 101-21.
135. Rollins, B., et al., *Mitochondrial variants in schizophrenia, bipolar disorder, and major depressive disorder*. PLoS One, 2009. **4**(3): p. e4913.
136. Salvioli, S., et al., *The impact of mitochondrial DNA on human lifespan: a view from studies on centenarians*. Biotechnol J, 2008. **3**(6): p. 740-9.
137. Cha, R.S. and W.G. Thilly, *Specificity, efficiency, and fidelity of PCR*. PCR Methods Appl, 1993. **3**(3): p. S18-29.
138. Stratagene, *StrataClone™ Blunt PCR Cloning Kit: Instruction Manual*. 2006.
139. Hogrefe, H.H. and M.C. Borns, *High Fidelity PCR Enzymes*. PCR Primer: A Laboratory Manual ed. C.W. Dieffenbach and G.S. Dveksler. 2003, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
140. McCulloch, S.D. and T.A. Kunkel, *The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases*. Cell Res, 2008. **18**(1): p. 148-61.

141. Roche, *Expand High Fidelity^{Plus} PCR System Pack Insert*, in Roche Applied Science. 2007.
142. Lutz-Bonengel, S., et al., *Different methods to determine length heteroplasmy within the mitochondrial control region*. Int J Legal Med, 2004. **118**(5): p. 274-81.
143. Somasundaran, M., et al., *Localization of HIV RNA in mitochondria of infected cells: potential role in cytopathogenicity*. J Cell Biol, 1994. **126**(6): p. 1353-60.
144. Flores, S.C., et al., *Tat protein of human immunodeficiency virus type 1 represses expression of manganese superoxide dismutase in HeLa cells*. Proc Natl Acad Sci U S A, 1993. **90**(16): p. 7632-6.
145. Choi, J., et al., *Molecular mechanism of decreased glutathione content in human immunodeficiency virus type 1 Tat-transgenic mice*. J Biol Chem, 2000. **275**(5): p. 3693-8.
146. Grollman, A.P. and M. Moriya, *Mutagenesis by 8-oxoguanine: an enemy within*. Trends Genet, 1993. **9**(7): p. 246-9.
147. Gilbert, S.F., *Developmental Biology, 6th Edition* 2000: Sinauer Associates.
148. Lightowlers, R.N., et al., *Mammalian mitochondrial genetics: heredity, heteroplasmy and disease*. Trends Genet, 1997. **13**(11): p. 450-5.
149. Sykes, B., *The Seven Daughters of Eve: The Science That Reveals Our Genetic Ancestry*. 2002: W. W. Norton & Company.
150. Cavelier, L., et al., *MtDNA substitution rate and segregation of heteroplasmy in coding and noncoding regions*. Hum Genet, 2000. **107**(1): p. 45-50.
151. Yao, Y.G., et al., *Pseudomitochondrial genome haunts disease studies*. J Med Genet, 2008. **45**(12): p. 769-72.
152. Romani, B., S. Engelbrecht, and R.H. Glashoff, *Antiviral roles of APOBEC proteins against HIV-1 and suppression by Vif*. Arch Virol, 2009. **154**(10): p. 1579-88.
153. Asari, M., et al., *Differences in tissue distribution of HV2 length heteroplasmy in mitochondrial DNA between mothers and children*. Forensic Sci Int, 2008. **175**(2-3): p. 155-9.
154. Verma, M. and D. Kumar, *Application of mitochondrial genome information in cancer epidemiology*. Clin Chim Acta, 2007. **383**(1-2): p. 41-50.

155. Pinheiro, M., et al., *Mitochondrial genome alterations in rectal and sigmoid carcinomas*. Cancer Lett, 2009. **280**(1): p. 38-43.
156. Tong, B.C., et al., *Mitochondrial DNA alterations in thyroid cancer*. J Surg Oncol, 2003. **82**(3): p. 170-3.
157. Cherry, C.L., et al., *Tissue-specific associations between mitochondrial DNA levels and current treatment status in HIV-infected individuals*. J Acquir Immune Defic Syndr, 2006. **42**(4): p. 435-40.
158. Maagaard, A., et al., *Mitochondrial (mt)DNA changes in tissue may not be reflected by depletion of mtDNA in peripheral blood mononuclear cells in HIV-infected patients*. Antivir Ther, 2006. **11**(5): p. 601-8.
159. Lee, H.C., et al., *Differential accumulations of 4,977 bp deletion in mitochondrial DNA of various tissues in human ageing*. Biochim Biophys Acta, 1994. **1226**(1): p. 37-43.
160. Branton, D., et al., *The potential and challenges of nanopore sequencing*. Nat Biotechnol, 2008. **26**(10): p. 1146-53.
161. Michikawa, Y. and G. Attardi, *Screening for aging-dependent point mutations in mtDNA*. Methods Mol Biol, 2002. **197**: p. 75-92.
162. Currier, J.S., *Sex differences in antiretroviral therapy toxicity: lactic acidosis, stavudine, and women*. Clin Infect Dis, 2007. **45**(2): p. 261-2.
163. Hoschele, D., *Cell culture models for the investigation of NRTI-induced mitochondrial toxicity. Relevance for the prediction of clinical toxicity*. Toxicol In Vitro, 2006. **20**(5): p. 535-46.

Appendix: Ethics Review Board Approval Certificate

The University of British Columbia
Office of Research Services
Clinical Research Ethics Board – Room 210, 828 West 10th Avenue, Vancouver, BC V5Z 1L8

ETHICS CERTIFICATE OF EXPEDITED APPROVAL: RENEWAL

PRINCIPAL INVESTIGATOR: John C. Forbes	DEPARTMENT:	UBC CREB NUMBER: H03-70356
INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:		
Institution		Site
Children's and Women's Health Centre of BC (incl. Sunny Hill)		Children's and Women's Health Centre of BC (incl. Sunny Hill)
Other locations where the research will be conducted: N/A		
CO-INVESTIGATOR(S): David R. Burdge Deborah Money Peter Lansdorp Helene Cote Ariane Alimenti		
SPONSORING AGENCIES: Canadian Institutes of Health Research - "Mitochondrial DNA Damage in Infants Exposed to HIV Antiretroviral Drugs In Utero" Unfunded Research - "Perinatal Exposure to Antiretroviral Therapy in Infants Born to HIV Seropositive Mothers: Evaluation of Toxicity using Mitochondrial DNA and Lactate Levels"		
PROJECT TITLE: Perinatal Exposure to Antiretroviral Therapy in Infants Born to HIV Seropositive Mothers: Evaluation of Toxicity using Mitochondrial DNA and Lactate Levels		

EXPIRY DATE OF THIS APPROVAL: November 6, 2007

APPROVAL DATE: November 1, 2006

CERTIFICATION: In respect of clinical trials: <ol style="list-style-type: none"> 1. The membership of this Research Ethics Board complies with the membership requirements for Research Ethics Boards defined in Division 5 of the Food and Drug Regulations. 2. The Research Ethics Board carries out its functions in a manner consistent with Good Clinical Practices. 3. This Research Ethics Board has reviewed and approved the clinical trial protocol and informed consent form for the trial which is to be conducted by the qualified investigator named above at the specified clinical trial site. This approval and the views of this Research Ethics Board have been documented in writing. <p>The Chair of the UBC Clinical Research Ethics Board has reviewed the documentation for the above named project. The research study, as presented in the documentation, was found to be acceptable on ethical grounds for research involving human subjects and was approved for renewal by the UBC Clinical Research Ethics Board.</p> <p style="text-align: center;"><i>Approval of the Clinical Research Ethics Board by one of:</i></p> <p style="text-align: center;">Dr. James McCormack, Associate Chair</p>
--

**CHILDREN'S & WOMEN'S HEALTH
CENTRE OF BRITISH COLUMBIA**
AN AGENCY OF THE PROVINCIAL HEALTH SERVICES AUTHORITY

UBC C&W Research Ethics Board
A2-136, 950 West 28th Avenue
Vancouver, BC V5Z 4H4
Tel: (604) 875-3103 Fax: (604) 875-2496
Email: cwreb@cw.bc.ca
Website: http://www.cfri.ca/research_support > Research Ethics

ETHICS CERTIFICATE OF MINIMAL RISK APPROVAL: AMENDMENT

PRINCIPAL INVESTIGATOR: Deborah Money	DEPARTMENT: UBC	UBC C&W NUMBER: H04-70540
INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:		
Institution Children's and Women's Health Centre of BC (incl. Sunny Hill) Other locations where the research will be conducted: N/A		Site Child & Family Research Institute
CO-INVESTIGATOR(S): David R. Burdge Julie Van Schalkwyk Peter Lansdorp Dominique Khoo Elizabeth J. Phillips John C. Forbes Helene Cote Ariane Alimenti Julio S.G. Montaner		
SPONSORING AGENCIES: - Canadian Foundation for AIDS Research - "Effect of Antiretroviral Therapy on Maternal Blood Cell Mitochondrial DNA Levels During Pregnancy in HIV-Infected Women" - Canadian Institutes of Health Research (CIHR) - "Mitochondrial DNA Damage in Infants Exposed to HIV Antiretroviral In Utero"		
PROJECT TITLE: Effect of Antiretroviral Therapy on Maternal Blood Cell Mitochondrial DNA Levels During Pregnancy in HIV-Infected Women		

REMINDER: The current UBC Children's and Women's approval for this study expires: **October 23, 2010**

AMENDMENT(S):			AMENDMENT APPROVAL DATE: December 2, 2009
Document Name	Version	Date	
Protocol:			
Protocol Amendment 02, 18 May 2006 - Telomere Length Study Protocol, 13 Oct 2004.v2	N/A	May 18, 2006	
Protocol Amendment 03, 08 Nov 2009 - vit D & PTH	N/A	October 13, 2004	
Protocol Amendment 01, 13 Jan 2006 - Testing mtDNA Quality	N/A	November 8, 2009	
Other Documents:			
Response to UBC proviso received 23 Nov 2009	N/A	January 13, 2006	
CERTIFICATION:			
In respect of clinical trials:			
1. The membership of this Research Ethics Board complies with the membership requirements for Research Ethics Boards defined in Division 5 of the Food and Drug Regulations.			
2. The Research Ethics Board carries out its functions in a manner consistent with Good Clinical Practices.			
3. This Research Ethics Board has reviewed and approved the clinical trial protocol and informed consent form for the trial which is to be conducted by the qualified investigator named above at the specified clinical trial site. This approval and the views of this Research Ethics Board have been documented in writing.			
The amendment(s) for the above-named project has been reviewed by the Chair of the UBC Children's and Women's Research Ethics Board and the accompanying documentation was found to be acceptable on ethical grounds for research involving human subjects.			
Approved by one of:			
Dr. Marc Levine, Chair Dr. Caron Strahlendorf, Associate Chair			