Epigenetic Markers in Forensics and Ageing

By

Shakhawan K. Mawlood

Strathclyde Institute of Pharmacy and Biomedical Sciences (SIPBS)

University of Strathclyde/ Glasgow

A thesis presented in fulfilment of the requirements for the degree of **Doctor** of **Philosophy**

2016

Strathclyde Institute of Pharmacy and Biomedical Sciences (SIPBS)

University of Strathclyde/ Glasgow

Epigenetic Markers in Forensics and Ageing

PhD Thesis

Shakhawan Khalid Mawlood

Signed Date:

Abstract

This PhD project describes the development of two novel forensic approaches the use of trace DNA isolation and profiling from fired gun cartridges, and the development and application of epigenetic analyses of the autosomal and mitochondrial genomes in order to estimate age.

In the first part of the work two novel swabs (forensiX) and a standard cotton swab (EUROTUBE/Deltalab) were compared for the collection of forensic samples. Additionally, the effect of long-term storage of these samples was assessed. The forensiX swabs generally produced a higher (two- to four-fold) DNA yield compared to standard swab (around 750 pg and 250 pg, respectively, from 0.5 µL of saliva). These findings demonstrate the importance of 'active drying' performance in the preservation of DNA and swab selection. Using the forensiX swabs, we subsequently produced data that indicated that DNA samples deposited on cartridges during loading can survive firing from eight different types of gun.

It has been shown that the methylation status of certain human DNA loci correlates with ageing. Blood sample DNA from a cohort of 82 women aged 18 to 91 years was obtained. We used Illumina MiSeq next generation sequencing platform to investigate the promoter regions of 23 genes suggested to have age-correlated methylation levels. Methylation levels at three CpGs located in the *ASPA, ITGA2B* and *PDE4C* genes showed an epigenetic signature of ageing with only a 6 year error range from chronological age.

The methylation of mitochondrial DNA (mtDNA) is a new and incompletely described phenomenon with unknown biological control and significance. We describe the bisulphite sequencing of mtDNA from the same cohort of individuals. We detected low and variable levels of mtDNA methylation at 54 of 133 CpG sites interrogated. Regression analysis of methylation levels at two CpG sites (M1215 and M1313) located within the 12S ribosomal RNA gene showed an inverse relationship with age.

We wished to create a fast and simple forensic tool (compared to next generation sequencing) for practical age estimation. The EpiTect Methyl II PCR system (QIAGEN) was used to compare methylation levels of CpG islands of the promoter regions of 4 age related genes (*NPTX2*, *KCNQ1DN*, *GRIA2* and *TRIM58*). The data obtained from DNA methylation quantification showed successful estimation of subject age (11 year accuracy).

This thesis describes practical steps to obtain forensic DNA samples while also applying novel techniques to explore the use of epigenetic profiling as a means to age prediction. The data generated suggest that the analysis of methylation patterns

of both specific autosomal and mitochondrial gene sequences from biological evidence left at crime scene may help build up an accurate picture of an offender's age and may help investigators obtain better descriptive information about a contributor from DNA, regardless of database inclusion.

Dedication

I’d like to dedicate my PhD thesis to my Father`s soul, Khalid and my great mother, Maryam for all their sacrifices, care, concern, moral support and prays. You made me into who I am. I also dedicate this work to my siblings and all friends for their unconditional support. I must dedicate my work and express my gratitude to Aryan, my queen for her patient and for being here with me and far from her family these years. This PhD is yours too. I also thank my wonderful children: Danaz and Lazo, for always making me smile and for understanding on those weekend mornings when I was writing this thesis instead of playing games. I hope that one day they can read my research and understand why I spent so much time in front of my computer. Most importantly, I dedicate this work to all my friends and the people who read this thesis.

Acknowledgment

After the praises and thank to Allah my God, so… where to begin? First and foremost I want to thank my supervisors, Dr. Nigel Watson for following me to carry out my research under his supervision, but unfortunately he retired before my graduation, special thanks to Dr. Lynn Dennany my second supervisor who supported and helped me all the time, massive thanks to Dr. Benjamin Pickard for all of his help advice over the last two years. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly. Thanks also owed to all staffs in the University of Strathclyde and my great colleagues who I met during my study, Majid, Mohammed, Nicola, Penny, Holly, Gina, Linda, Suaad and Husain hope we can meet again.

Special thanks also go to the Kurdistan Region government for the sponsorship and the opportunity to allow me to be one of the Human capacity development Programme candidates and to pursue one of my lifelong dreams. Finally, thanks to Ministry of the interior for giving me the chance to follow my ambitious childhood and study PhD in the UK.

Publications and presentations relating to this research

Journal Articles:

- Mawlood, S. ; Alrowaithi, M. and Watson, N. `Advantage of ForensiX Swabs in Retrieving and Preserving Biological Fluids`. Journal of Forensic Sciences, 2015. 60(3): p. 686-689.
- Mawlood, S. ; Dennany, L. ; Watson, N. and Pickard, B. (2015), 'Analysis of DNA from Fired Cartridge Casings', World Academy of Science, Engineering and Technology, International Science Index 104, International Journal of Biological, Biomolecular, Agricultural, Food and Biotechnological Engineering, (2015)-9-(8), 805 - 812.
- Mawlood, S. ; Dennany, L. ; Watson, N. ; Dempster1, J. and Pickard, B. `Quantification of global mitochondrial DNA methylation levels and inverse correlation with age at two CpG sites`. In press: AGING, February 2016, (8); 2, P336-341.
- Mawlood, S., Dennany, L., Watson, N. and Pickard, B.S., 2016. The EpiTect Methyl qPCR Assay as novel age estimation method in forensic biology. Forensic Science International, 264, pp.132-138.
- Mawlood, S ; Dennany, L. ; Watson, N. ; Dempster1, J. and Pickard, B. `Tight correlation between methylation status and human age at three autosomal loci: a new forensic profiling tool`. Under revision at Legal Medicine Journal. Ref. No.: LEGMED-D-16-00004.

Oral Presentation:

- Mawlood, S. K., Alrowaithi, M., Dennany, L. and Watson, N. Retrieving DNA From fired Cartridges. Presented at the first Scottish Student Forensic Research Symposium. March 2014 (SSFRS 2014), Glasgow, UK.
- Mawlood, S. K., Dennany, L. and Pickard, B. Estimating Human Age from DNA Methylation. Presented at 2nd Kurdistan Society Science Conference. September 2014 (KSC 2014). Nottingham, UK.

- Mawlood, S., Dennany, L. and Pickard, B. `Potential forensic application of EpiTect methyl system`. Presented at the 2nd Scottish Student Forensic Research Symposium at 27th March 2015 (SSFRS 2015), Dundee, UK.
- Mawlood, S., Dennany, L. and Pickard, B. `Analysis of DNA from Fired Cartridge Casings`. Presented at the XIII International Conference on Forensic Sciences. 27-28th September 2015 (ICFS 2015), Paris, France.
- Mawlood, S. K., Dennany, L., and Pickard, B. `The EpiTect Methyl qPCR Assay as novel age prediction method in forensic biology`. Presented at the 7th European Academy of Forensic Conference. 6-11th September 2015 (EAFS 2015), Prague, Czech Republic.
- Mawlood, S., Dennany, L. and Pickard, B. `Mitoepigenetics: 50 different shades of Grey`. Presented at the 3nd Scottish Student Forensic Research Symposium at 31st March 2016 (SSFRS 2016), Glasgow, UK

Poster Presentation:

Mawlood, S., Alrowaithi, M., Dennany, L., and Watson, N. Importance of active drying using forensiX swabs taken from saliva stains. Presented at 24th International Symposium Human Identification (24 ISHI). 5th- 10th October 2013. Atlanta, GA, the USA.

Mawlood S., Alrowaithi, K., Dennany, L., and Watson, N. Advantage of Prionics Collection Swab in retrieve and preserving Biological Fluide. Presented at Forensic Horizons 2013.The Forensic Science Society and California Association of Criminalists joint Autumn Conference. 6th – 8th November 2013. Manchester, UK.

Mawlood S., Dennany, L., and Watson, N. Estimating Human Age from DNA Methylation. Presented at 9th international Y-chromosme and 6th EMPOP meeting. 14-16 May 2014. DNA in forensic Brussels 2014, Belgium.

List of abbreviations

5-hmC:	5-hydroxymethyl-cytosine
5mC:	5methyl-cytosine
AA:	aplastic anaemia
AIMs:	ancestry informative markers
ALT:	alternative lengthening of telomeres
AR-CpG:	age-related CpG site
AR-DNAm:	age-related DNA methylation
Arg:	arginine
ATP:	adenosine triphosphate
BER:	base excision repair
bp:	base pair
C:	cytoplasm
CGI:	CG islands
CH:	chromosome
CMV:	cytomegalovirus
COBRA:	combined bisulphite/restriction analysis
CpG:	cytosine-phosphate-guanine
CS1:	common sequence 1
CS2:	common sequence 2
Ct:	cycle threshold
CV:	C_T value
DEC:	methylation-dependent enzyme control
D-loop	displacement loop

DM:	densely or uniformly methylated
DMD:	duchene muscular dystrophy
DMSO:	dimethylsulfoxide
DNA	deoxyribonucleic acid
DNAm:	DNA methylation
DNMT:	DNA methyltransferase
DVI:	disaster victim identification
EO:	ethylene oxide
ESC:	embryonic stem cell
F:	forward
FBI:	Federal Bureau of Investigation
FDP:	forensic DNA phenotyping
FDR:	false discovery rate
FISH:	fluorescent in situ hybridization
GRCH38:	genome reference consortium human build38
GWAS:	genome wide association studies
h:	hour(s)
HSP1	H-strand promoter one
H-strand:	heavy strand
HVR:	hyper variable regions
ICC:	immunocytochemistry
IM:	intermediately methylated
IMCFSA:	Indianapolis-Marion County Forensic Service Agency

iPSCs:	induced pluripotent stem cells
ISHI:	International Symposium Human Identification
LHON:	leber hereditary optic neuropathy
LSP:	L-strand promoter
L-strand:	light strand
M_o:	mock digested
M:	methylated
MAD:	mean absolute deviation
M_d:	methylation-dependent digest
MDRE:	methylation dependent restriction enzyme
m:	minute
MLR:	multivariate linear regression
mROS:	mitochondrial reactive oxygen species
M_s:	methylation-sensitive digest
MS:	mass spectrometry
M_{sd}:	double digest
MSRE:	methylation sensitive restriction enzyme
mtDNA:	mitochondrial DNA
MT-HSP2:	minor H-strand promoter two
mtROS:	mitochondrial reactive oxygen species
NARP:	neurogenic muscle weakness, ataxia and retinitis pigmentosa
NGS:	next generation sequencing
Nt:	nucleotide
OH:	origin heavy

OL:	origin light
OR-H:	origin replication of heavy strand
OR-L:	origin replication of light strand
Pers.Comm:	personal communication
PGCs:	primordial germ cells
PH:	promoter sites of heavy strand
PL:	promoter sites of light strand
poly-γ:	polymerase γ
QC:	quality control
R:	refractory percentage
R:	reverse
R:	refractory
RBC:	red blood cell
RMSE:	root mean square error
RNA:	ribonucleic acid
ROI:	region of interest
ROS:	reactive oxygen species
rRNA:	ribosomal RNA
RT:	room temperature
S:	second
SAH:	S-adenosylhomocytosine
SAM:	S-adenosylmethionine
SEC:	methylation-sensitive enzyme control
STR:	short tandem repeat

tDMR:	tissue-specific-differentially-methylated region
tDNA:	touch DNA
TET:	ten-eleven translocations
tRNA:	transfer RNA
tRNA:	transfer RNA
TSS:	transcriptional start sites
UNODC:	United Nations Office on Drug and Crime
W:	analytical window

Table of Contents

List of figures

List of tables

CHAPTER ONE

Thesis Overview

1 Chapter 1 Overall Introduction

This thesis is produced for the purpose of doctoral studies with the Centre for Forensic Science, Department of Pure and Applied Chemistry and latterly, at the Strathclyde Institute of Pharmacy and Biomedical Sciences, both at the University of Strathclyde. The report consists of nine chapters. An Introduction will be the first part and will describe the process by which the aims and objectives of the research were developed and refined. This is followed by a review of the literature. The literature review is initially focused on 'touch sample' analyses, taking into account the physiology of the skin and the occurrence of genetic material in the skin and its secretions and the epigenetic study of DNA. This will be then be widened to include related topics such as the mitochondrial DNA (mtDNA) genome and methylation state of different persons with different ages.

In Chapter Two the report describes the materials and methods that have been utilised since the start of this research. The main work of the research will be concerned with methylation pattern analysis as tool for forensic investigation.

Chapter Three is concerned with 'touch samples'. A preliminary laboratory study was made to assess a new type of evidence collection swab tube (forensiX) in comparison with commercially existing swabs to find the most effective method of recovering DNA from touch samples.

Chapter Four examines DNA methylation levels of several age related genes after the DNA was treated with sodium bisulphite. A large group of genes was studied and the Illumina Next Generation Sequencing (NGS) used for methylation analysis.

Chapter Five focuses on methylation in the mitochondrial genome and its relation to age in an entirely novel study depending on Illumina NGS again for analysis.

Chapter Six focuses on the assessment of a new method to estimate human age from DNA methylation depending on the EpiTect Methyl system.

Chapter Seven is a comparison study among eight types of common weapons to investigate the ability of retrieve DNA from fired cartridge and apply the result with EpiTect Methyl System.

The final chapter (Chapter Eight) will describe the important outputs from this research and explore future lines of research arising from the results reported. The work was designed to supplement the existing laboratory experience and skills of the author, especially in areas such as designing and optimising amplification reactions and the use of certain instruments. In addition to the outcome of papers published in international journals, the work had the advantage that it provided experience in the preparation and delivery of three posters and several oral presentations at an international scientific meeting.

The work has carried out in different places, University of Strathclyde (both the Centre for Forensic Science and Strathclyde Institute of Pharmacy and Biomedical Sciences, SIPBS), the Police DNA analysis lab in Erbil and, for the blood sample collection (venipuncture), Erbil Hospital by an appropriately trained health professional. All the work complied with the ethical policies of the University Ethics Committee and of Erbil hospital.

The profiling was necessary to demonstrate that the methylation study is compatible with the DNA profiling, an important consideration in a forensic analysis.

The identities of the donors of samples were only known to the researcher, to allow samples to be withdrawn from the study at any time at the request of the donor. However, ethical issues concerned this to be addressed by labelling the samples with a code that is only known by the researcher. Any dissemination of the results of the testing was recorded if a profile has been successfully produced or not and that the profiles confirm that the results of the methylation testing are indeed from different people and are not from the same person. The identity of the profile and the profile itself was not reported.

The purpose of the methylation testing was to determine if there is a correlation between the methylation at different genetic loci and the age of the donor. The identities of donors and their country, their age, sex and the type of body fluid were

known to the researcher, but the identities of the donors were not and will not be revealed in any dissemination of the research findings. It will not be possible to identify the origin of the sample from the results. Therefore both issues were addressed by ensuring the anonymity of the donors.

1.1 DNA in Forensic Science

Forensic science can be defined as "a specialism that tends to help investigator solve legal cases, not only in criminal law but also in civil issues" [1]. The application of biological fluids in law enforcement is called forensic biology. The cell is the basic structural unit of life. It can be defined also as a functional and biological unit of all organisms. An ordinary human being is made up of nearly 100 trillion cells, all of which originated from single fertilized (egg) cell [2]. All cells within an individual organism possess the same genetic material called chromosomes made of DNA, except mature red blood cells (RBC).

Along with RNA and proteins, DNA is one of the three main macromolecules that are necessary for life forms. One of the biological tools which have revolutionized forensic investigations is DNA analysis.

The DNA sequence is genetically identical in all somatic cells (excepting T and B cells), thus individual cells can be representative of the organism as a whole [3]. Some "traits" show evidence of heritable characteristics, whereas others vary between individuals but are not heritable. Some others have both a heritable component and a non-heritable component (e.g. weight) [4].

Genetic markers such as DNA will not vary throughout an individual's lifetime. These are the preferred markers for forensic biology. Markers that exist in different forms are polymorphic, meaning "many forms" [5].

On the other hand, a marker that does not vary within a population is described as monomorphic [5-7]. Polymorphic genetic markers that exist in a number of discrete forms (alleles) are considered the most useful markers for forensic analysis purpose. Currently, the use of DNA in forensic science is concerned with testing of selected polymorphic regions to determine genotypes for each, and the combination of these

genotypes is referred to as the DNA profile [8]. The technology uses fluorescently labelled primers to co-amplify a number of short tandem repeat (STR) loci on different chromosomes [2]. Following this the amplification products are separated by capillary electrophoresis and the resulting amplicon peaks interpreted by comparison with standards to infer the genotype combinations, or profile, using a combination of retention time and the type of dye label. The frequency of the combination of STR genotypes is calculated from databases of allele frequencies for different populations as appropriate.

While the chance of a coincidental match between unrelated people cannot be discounted, it is very small: less than one chance in a billion [8]. There are now many kits of reagents prepared specially for the forensic market. In the UK the standard 'work horse' kit was the second generation multiplex or SGM+ kit but in recent years there have been significant improvements such as the Minifiler kit and the next generation kits, or NGM. These offer redesigned primer sequences so that smaller amplicons are made and thus improving the performance with degraded DNA samples. The NGM kits also test more STR loci to improve the power of discriminating between people.

Presently much research effort in the discipline of forensic genetics is directed towards inferring phenotypic information that may benefit the investigator - rather than simple human identification. A major consideration is how robust the tests are for conducting in high throughput circumstances and demonstrating the ability to test small samples that have been exposed to the environment and/or hot gases and pressure such as that on a cartridge case after firing.

The use of mtDNA is most commonly seen in human identification purposes, especially when the nDNA is degraded or when the amount is low [8]. The mitochondrial genome is maternally inherited [9] and has many copies per cell in comparison to only two copies of nDNA [10] thus making mitochondrial DNA very useful in highly degraded specimens and helping to identify people by comparison of sequences with maternal relatives [8].

1.2 **Epithelial Tissue in Forensic Science**

Epithelial tissues are cellular sheets, covering all the free, open surfaces of the body including the skin, and mucous membrane that are exposed to the outside of the body [11]. They comprise sheets of cells bound together strongly [12]. This type of tissue appears as a single (simple epithelium) or multiple (stratified epithelia) cell layers, and all epithelial cells together form an epithelium (plu. Epithelia). Basal laminas are the first layer of epithelial cells within epithelium which occur on basement membrane. This layer of cells supports other epithelial cells and separates them from underlying tissues [13, 14] Depending on their location, epithelial cells are associated with protection, absorption and secretion [11]. Epithelial cells are sloughed from epithelia through the active handling of objects and this increases the value of them during the process of forensic investigation [15]. The largest organ in our body is skin, because 15% of the mass of the body is skin and is considered to be potential source of cells with 110 sweat and oil glands in each square centimetre of the skin for DNA profiling [16].

1.3 **Contact DNA**

Contact DNA' refers to the DNA that left behind through contact of the skin when an individual touches an object or surface in crime scene directly [15]. There might even be transfer indirectly from one person to another first and subsequently to the objects [17]. Contact DNA, also known as trace DNA or 'touch' DNA occurs in relatively low amounts comparing to other body fluids, blood for instance, usually such type of evidence yields less DNA than ordinary amount which is necessary to get a full DNA profile [18-20]. In most cases touch DNA is invisible to the naked eye [21], but still there is an ability to get DNA typing depending on the technique which is uses. Over the last decade various techniques have been used for contact DNA sampling in forensic laboratories around the world and still new techniques are being developed as well as, such as automated extraction processes [17].

It is known that humans shed hundreds to thousands (40,000) skins cells each day [15, 21]. In skin epithelial cells are regenerating cells very fast in the body. When skin cells reach the outer layer, they keratinize and are considered as dead skin (non-nucleated) cells then slough off [22]. The skin matrix contains sebaceous oil glands and sweat glands [16]. Secretions and nucleated cells from these glands are brought to the surface of the skin via the ducts and pores and create a potential source of DNA [15, 23]. The fingers or hands, acting as a vector, have the capability of transferring the cells/DNA as well as transmission of nucleated cells from the eyes and mouth, to objects that have been handled, such as knives, cartridge cases, rope and surfaces [15, 24]. DNA is often recovered from cell-free sweat samples [25]. This discovery supports the hypothesis of nucleic acid-free transferring through handling and it is an answer for those who ask if skin cell are non-nucleated and then dead how DNA can be recovered from touched objects. Figure 1-1 shows handling of cartridge during loading bullets in to magazine [26].

Figure 1-1 handling of Cartridge: The figure shows loading bullets into a magazine.

Exploring improved new technologies has provided forensic scientists with the ability to generate a DNA profile of the contributor from mere contact between the individual and objects. Recently research has shown that contact DNA could be present on gun grips and even fired cartridge cases, as a result of handling the exhibit prior to loading and firing the magazine.

It could be argued that DNA which has been deposited on to the cartridge case during loading of a magazine has been destroyed via the hot temperature and pressure inside the chamber of the weapon. But studies have shown that human identification through DNA typing could be generating from contact DNA. Specifically DNA recovered after firing from an exploded pipe bombs, chisels and hammers [18, 27].

In 2004, Esslinger, with her colleagues, reported one full and three partial profiles from twenty-seven exploded bomb devices [27]. They confirmed that genetic profile could be obtained from materials that have been exposed to extraordinary heat [27]. Several studies have reported DNA recovery from various locations on the firearm, as well as from ejected cartridge cases and non-spent cartridge or ammunition in general [24, 28, 29]. Furthermore a study in Netherlands showed that biological material from cartridge, bullet and casings are regularly recovered [30] and have been used against criminals in court. The Indianapolis-Marion County Forensic

Service Agency (IMCFSA) for instance, provides their officers with special kit (called Trigger Pro) to collect the DNA sample from a firearm crime and other incidents [24]. Polley with her group discovered that unfired cases have less DNA compared to the gun grip and magazine of a firearm, and that more DNA could be recovered compared to a fired one [28]. Hence the quality and quantity of DNA are recovered depend on the item size and textured surface area.

In spite of the hard work by the forensic scientists to get genetic profile from trace DNA cases, there are still considerations and challenges when they face this type of evidence. One issue facing forensic scientists is that they may struggle to recover DNA if the contributor washed their hands before touching objects. The movement and pressure of both hands in liquid cause skin scrub through mechanical agitation. The dirt and dead skin cells are separated from the skin by a combination of the movement of the hands as well as the wet chemical action of the soap. When hand washing occurs, an individual actually removes cells that contain DNA from the skin [31]. One would expect that people would shed less DNA during contact with an object after washing their hands; nevertheless production of the outmost layers of cells in the skin is continuous and rapid. Furthermore, sweat and oil glands in skin cells are proliferating and dying constantly. Van Oorschot, who first worked on touch swabs discovered that dry hands and washed hands have lower DNA yield [32].

Studies have proved that some contacts do not leave enough DNA behind to make a full DNA profile illustrated in Table 1-1. Lowe *et al.,* [17] classified contributors as either ‘good’ or ‘poor’ shedders of DNA. On the contrary Phipps and Petricevic believe that people cannot categorised as being either poor or good shedders because they shed different amount DNA at different time [33]. Hence, amount of DNA shed by an individual could vary on a day to day or even an hour to hour basis. In fact, it was found that an individual could either shed less or more over time depending on the individual. Regarding to the transfer of DNA, journal articles has shown that sometimes DNA is not transferred through contact. Experiments by Lowe *et al.,* demonstrated that there is little or no DNA transferred in approximately half of subjects after handling a sterile tube for 10 s. In study by Rutty, about strangulation,

only about half of the samples yielded DNA results [34]. Interestingly some of these DNA samples did not belong to the victim [34]. A research article by Phipps showed that when individuals held a DNA-free substrate for 10 s, more than 50% of them failed to leave their DNA behind [33]. In addition, when Raymond *et al.,* did their experiment on trace DNA casework samples, they noted that 44% of touched surface did not produce a profile [35].

Table 1-1 DNA amount detected by contact: This table provides information from various studies to show the amount of DNA typically recovered from different substrates.

Substrate/ contact type	Amount of DNA	Mean	Reference
Glass held for 1 min	0-5.2 ng	5.2 ng	[36]
Fabric held for 1min	0-14.8 ng	1.23 ng	[36]
Wood held for 1 min	0-169 ng	5.85 ng	[36]
Swab of hands	0.15-6.4 ng	Blank	[37]
Fingers on substrate for 30s.	0-0.4 ng	Blank	[38]
Various crime evidence items	0-50.8 ng	1.7 ng	[35]
Wallet held for 60 s	4.3 ng	Blank	[35]
Wallet held for various times	3.1-33 ng	11.7 ng	[35]
Touch tools (screw drivers and knfe)	2.3ng	Blank	[35]
Touched firearms	0.6 ng	Blank	[35]
Plastic knife held for 15min	Blank	17.8 ng	[32]
Mug held for 15 min	6.8 ng	Blank	[32]
Glass held for 15 min	34 ng	Blank	[32]
Vinyl gloves worn for 20-90min	Blank	51n g	[32]
Cotton rubbed with palm, finger and side of hand for 15s.	Blank	11.68	[39]
Plastic rubbed with palm, finger and side of hand for 15s.	Blank	0.396 ng	[39]

1.4 **Physical Characteristic Prediction**

With DNA profiling methods, a person of interest can only be identified on the basis of prior inclusion in criminal database, so additional information of an offender from DNA would be of great use for investigators. Obtaining a physical profile of contributor would narrow the search process. Usually the witness(es) of a crime is asked to describe the physical characters of the possible suspect. The height, age, body weight, hair and eye colours are the main physical characteristics in the

description process. The information obtained from the physical features of an individual may not bring the person to justice directly, but could be aid the police in undertaking investigations at a slightly more rapid pace. Although some physical features of a person are influenced by the environment, this estimation would not be an exact match. Environmental and nutritional conditions can complicate final phenotype determination. Plus, phenotypes can be masked: someone can have grey hair or shave their head (or be naturally bald) [40]. Therefore, scientists are now investigating genotype information for the sole purpose of generating a description of an alleged criminal. Some of this information can be used as ancestry informative markers (AIMs) to provide information about a person's physical feature.

1.5 **Epigenetic Analysis as a new tool for Finding Unknown Person**

DNA profiling with numerous highly polymorphic autosomal short tandem repeat (STR) markers has been used in different aspects of human identification in forensic investigation over the past 20 years [41]. Although the use of STR profile persists, human identification is not the only type of information potentially available from the genetic materials in a biological trace. There are some other characteristics that can be used for this purpose.

1.5.1 **Epigenetics**

In all cells the DNA sequence is identical in all tissues but epigenetically they are different. The term epigenetics explains a wide range of biological observations on the heritable changes in genome regulation that do not require a change in DNA sequence, and according to [42] a modern definition is "non-sequence dependant inheritance".

Conrad Hal Waddington was the first person to use the word Epigenetics in 1942 [43]. The phrase 'Epi' is Greek word means 'outer, over, above, in addition' cited by De Wilde [42]. The term epigenetics refers to a biological system which enables changes to be made to the expression level of a gene by switching the gene on or off without any change to the underlying DNA sequence [44]. DNA can be chemically "tagged" to either activate or block transcription of a particular gene through

epigenetic marks. DNA methylation (for example promoter, X-chromosome and imprinted genes), is the most common and best understood epigenetic modification. This is where there is an attachment of a methyl group to the cytosine that usually blocks transcription of DNA, by inhibiting the interaction between DNA and the proteins which normally turn the gene on. As a consequence, the gene is turned off [44].

Evidence suggests that most epigenetic modifications are not passed on directly from generation to generation[45]. In primordial germ cells (PGCs), the precursor cell from which spermatozoa and eggs develop, it has been shown that epigenetic marks are erased to permitting the "reprogramming" or "resetting" of all genes in preparation for the next generation. However, this finding has left scientists unable to answer the question as to how epigenetic inheritance occurs [44]. Surprisingly, this process appears to be so efficient that it "resets" DNA back to its natural state. So it is a good indicator for an individual's life and not his/her previous parental generation.

1.5.2 **DNA Methylation**

DNA methylation in eukaryotes refers to a biochemical process that involves addition of a methyl group to the 5' carbon of cytosine base. This reaction is catalysed by DNA methyltransferases (DNMT) [46]. This pyrimidine ring is one of four bases of the DNA molecule. When cytosine occurs next to guanine the context of the CG dinucleotide it can be described as cytosine-phosphate-guanine (5-CpG-3) it has the capacity to be methylated: hence the methylated dinucleotide sequence will be 5 methyl-cytosine (5mC) [47]. In human DNA 3-6% of all cytosines are thought to be methylated [48, 49]. The forward reaction of methylation is mediated by the enzyme of DNA methyltransferase using S-adenosylmethionine as a donor of methyl group. Demethylation is the backward reaction of methylation which is performed by DNA demethylase enzyme [50]. In addition to 5-mc, 5-hydroxymethyl cytosine (5-hmC) is another cytosine-derived base modification detected in DNA. It is produced from (5mC) by ten-eleven translocation (TET) through hydroxyl methylation reaction as shown in Figure 1-2.

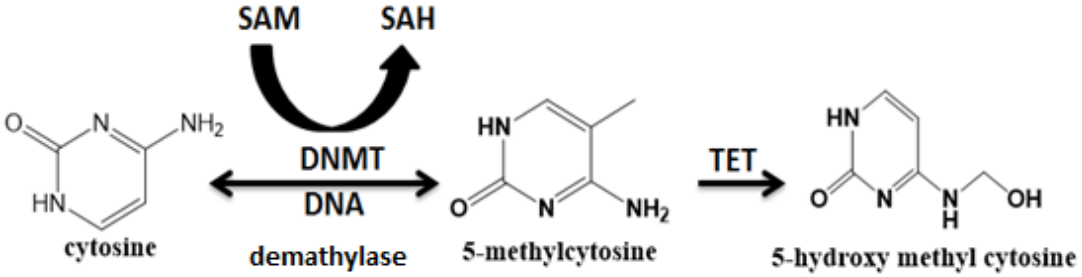

Figure 1-2 DNA methylation reaction: The figure shows methylation of cytosine and hydroxymethylation. Abbreviations: DNMT, DNA methyltransferase; SAM, S-aenosylmethionine; SAH, S-adenosylhomocytosine; TET, ten-eleven translocations. Reproduced from [50]. Drawn using ChemBioDraw [51].

1.6 Epigenetic and Forensic Analysis

One of the important pieces of information for forensic investigators is the age of the donor of biological materials. Analysis of biofluids that are commonly found at the crime scene can potentially be used to determine some physical features like eye and hair colour [52]. So this type of analysis of evidence may aid investigators concerning the gender and genotype of a suspect [53].

Madi *et al.*, showed different methylation patterns across the CpG sites of a set of genetic loci from different tissues (blood, saliva, semen and epithelial tissue) [54]. Hyun *et al.*, also found different methylation levels in some segments of chromosomal DNA that were named tissue-specific-differentially-methylated regions (tDMRs) [55], and demonstrated that the DNA methylation can be used to distinguish different tissue types of the same person. This proves experimentally that DNA methylation is a valuable indicator to distinguish body fluids that could contribute to forensic analysis [56, 57].

1.7 Applications of Forensic Epigenetics

The determination of cell growth, survival and differentiation is known to be controlled by the expression of particular genes, and further influenced by their epigenetic state. The level of DNA methylation of specific genes has relevance within a number of scientific fields [58].

The epigenome changes according to the needs of the cell and also links between the genome and the environment. Hence methylation status can provide information about the activities of genes within a tissue at a certain point of time. In terms of forensic application the analysis of DNA methylation patterns may assist with determining the cause and circumstances of death, the parental origin of alleles, authentication of DNA samples, discrimination of monozygotic twins, identification tissue of type and the age of the individual involved [59-62].

Using DNA methylation technique in crime field is not new. In the early 1990s Naito and his colleague were the first to utilise DNA methylation in forensic biology and established a method for epigenetic female sex typing [63]. Their method was based on the different methylation patterns of the promoter region of *DXZ4* gene in the X chromosome – hypermethylated in the active X chromosome but hypomethylated in the inactive one. The novelty of this work was not just in enabling positive female sex determination, but also in being able to achieve it with as little as 50 pg of DNA. It might also assist in cases of sex-reversed individuals [63].

1.7.1 **Body Fluid Identification**

Besides DNA typing, knowledge of the cellular origin of a recovered biological stain could be beneficial in a criminal investigation. The presence of specific body fluids could be linked to particular types of crime, for example, semen with sexual assault. Most of the current methodologies for identifying body fluids use presumptive and sometimes destructive biochemical tests to identify specific elements in each fluid, whereas the accurate detection of body fluids *in situ* in a non-destructive manner is essential in order to protect the sample and preserve DNA evidence.

The system of body fluid identification is compatible with available DNA typing technologies. However, it would be advantageous to be able to have a panel of specific identification tests in a multiplex reaction which could be performed on a single sample. This would be particularly desirable where the sample size is min.

In 2012, Lindenbergh and his group developed a multiplex assay (19 loci mRNA markers) which was able to differentiate semen, blood, saliva, vaginal mucosa,

menstrual secretion and skin [64]. Although this RNA profile assay shows good sensitivity, RNA is not as stable as DNA and therefore degradation must be considered. This is especially so for older and trace samples [65, 66].

An ideal method would be one which exploits the stability of the DNA molecule that using minimal amounts of the sample. An assay based on DNA methylation should be able to provide direct correlation between DNA type and its source. Furthermore, application of this assay can potentially help cases where only DNA has been recovered; cold cases for instance.

The first study in the forensic field related to DNA methylation-based tissue identification was by Frumkin [61]. Different biological samples were analysed (saliva, blood, semen, urine, skin, vaginal secretion and menstrual blood), screening more than 200 CpG islands. From these, 38 genomic loci showed differential levels of methylation, the authors proposing that these loci could potentially form the basis of a DNA methylation-based assay [61].

However, they further stated that "Loci with higher methylation levels were amplified with higher efficiency compared to loci with lower methylation levels, which yielded a relatively weak signal". As suggested, the peak height of a single locus correlates with its methylation level, although artefacts associated with PCR, such as differences in DNA template concentration or stochastic effects, could also affect peak height. Therefore, only ratios of methylation levels were used for the analysis [61].

Using a panel of several tissue-specific loci in pairs, each tissue showed a unique methylation profile and natural diversity was observed between different volunteers, causing 'noise'. However Gomes *et al*., were unable to validate the results of the previous study and could not differentiate between saliva and skin [67].

Recently, another study of DNA methylation pattern by Madi *et al.*, developed a more sensitive method which was able to identify body fluids in even small amounts of material, which is obviously more helpful in a forensic context [54]. In their method the extracted DNA was treated with sodium bisulphite and analysed by the new techniques of pyro- and next generation sequencing, showing exact amounts of

methylated cytosine at CpG sites. The study tested buccal cells, blood, saliva and semen by testing four loci (*ZC3H12D*, *FGF7*, *C20orf117* and *BCAS4*,). *ZC3H12D* and *FGF7* were thought to be sperm-specific and *C20orf117* to be blood-specific; *BCAS4* appeared to be saliva-specific although this gene was initially reported as semen-specific [54].

The results of this later study are very promising, although the reported differences of methylation levels are not always clear. Further investigation of methylation based tissue identification is therefore needed.

In such cases it is essential to choose markers which show either 'on or off' methylation, where quantification of the methylation status at individual CpG sites is necessary. A good blood epigenetic marker for instance, would be one which fully methylated in WBC (> 90%) but remained completely unmethylated in other biological fluids and vice versa. Although a full validation study and a body-fluid mixture analysis are still required these types of epigenetic markers might also allow analysis of mixed stains. One also has to take into account the inter-individual variation due to pathophysiological conditions or age.

To conclude, the methods described above show potential application of epigenetic markers for body fluid identification, albeit none of them providing accurate quantitative results, with most approaches being qualitative.

1.7.2 **Epigenetics and Age**

Human ageing is very complex, multifactorial process and affects all tissues of our body. Methylation changes at cytosine in the DNA molecule is one of the epigenetic modifications that occur through mammalian life time. Various approaches have been described such as morphological changes of skeletons, teeth and molecular or biochemical changes [68, 69]. The latter includes accumulated changes in mitochondrial DNA such as deletions, the reduction in telomere length associated with successive cell replication, the racemisation of aspartic acid and finally glycation end products. Other indicators are by means of T-cell receptor (TCR) gene deletion[68]. These genes form DNA molecules which form single joint T-cell

receptor excision circles (sjTRECs). It is believed that sjTCRs will decline through ageing. This work has been supported by Ou *et al.,* [69]. However all of these approaches suffer from variation and there are many examples of exceptions such as certain disease states, different between the sexes, and the population group of donors can have effects on these measures as well.

It is possible to examine the DNA methylation status in a sample as it has been shown to be affected by the age of the donor. Generally the methylation component will decrease through a life time; so that a centenarian has a lower DNA methylation component comparing to DNA of new born [70].

Studies of monozygotic twins offer an opportunity to compare individuals with a common starting point of DNA methylation pattern. Earlier studies used a restricted range of ages [71, 72]. Bocklandt *et al.,* more recently studied a wider range; looking at 34 sets of twins of ages 21 to 55 years in age using saliva samples [73]. Previous work had used Illumina human methylation 27 microarrays and demonstrated a trend in the epigenetic pattern with age of donor with an accuracy of 5.2 years [73]. Bocklandt *et al.,* were unable to repeat this trend but did report a set of 88 new loci that did show a trend matching the age of the donors [73]. A regression model was described using three CpG sites, *NPTX2, EDARADD* and *TOM1L1* that determined the age of the donor to a mean accuracy of 5.2 years.

In addition Koch and Wagner used publically available data (from Human Methylation27 BeadChip analysis) of thirteen tissue types. They found 431 hypermethylated and 25 hypomethylated CpG sites that were correlated with age [74]. From these CpG sites of five genes (*TRIM58, KCNQ1DN, NPTX2, BIRC4BP* and *GRIA2*) showed an "Epigenetic-Age" signature. One of these CpG sites (*GRIA2*) was common to both studies [74].

Koch and Wagner study found an average correlation of methylation with age based on selected CpG sites of a precision of 11 years. Estimations of age based on molecular analysis of biological materials are influenced by a complex set of factors but there remains the potential for such measures to make a contribution to ageing studies. Zbieć-Piekarska *et. al*. 2015 claimed that a prediction of biological age

through *ELOVL2* gene marker can be used in forensic science with an average of 7 years error between chronological and predicted age [75]. They used blood sample of 303 individual aged 2-75. Hence, this type of study could be used typically to estimate the biological age for forensic purposes.

1.8 **Ageing Theory**

Human individuals have a long lifespan, as with most other mammalian species, and take a long time to reach sexual maturity [76]. The reproductive stage occupies a significant period of the life span, and is followed by a post-reproductive stage. In most cases, the rate of reproduction is inversely proportional to the productive period and longevity. Ageing is a complicated process and encompasses a variety of mechanisms of varying importance throughout the species' lifespan. Ageing also affects organ function, as this relies on how well its component cells function. When ageing cells die and are not replaced the cell number is reduced and organ function affected. For example, bones become less dense and weaker and thus more prone to break, their calcium content reduced as the amount of absorbed calcium from nutrients decreases [77]. Muscle mass and muscle strength also tend to decrease. This is due in part to muscles being physically used less and thus reducing in size, but also because of decreased levels of growth hormone and testosterone. Ageing skin becomes thinner, less elastic, drier, and finely wrinkled due to decreased production of collagen which provides its strength, and of elastin which provides its flexibility. The fatty layer beneath also becomes thinner, thereby reducing cold tolerance and promoting wrinkling [77]. Furthermore, cognitive change and sensory defeats (sight and hearing) ability will decrease.

Several theories, each with supported evidence have attempted to explain age-dependent DNA changes [78]. Among the most commonly proposed mechanisms are: DNA damage, telomere shortening, and reactive oxygen species (ROS) mediated aberrations to macromolecules [79-81]. These mechanisms are intrinsically degenerative, in that with increasing age, the amount of DNA damage also increases. Such DNA age-dependent changes could have a practical application in forensic science, permitting age determination from samples collected from crime scenes.

Amongst the theories attempting to describe ageing at the genetic level are: somatic mutation theory, error theory, disposable soma model and the dedifferentiation hypothesis [79-81]. According to Cutler, with time, the gradual drifting of cells from their proper differentiation state results in ageing [81]. He argues that controlled differentiation leads to creation of the organism and is maintained for as long as it is necessary to ensure the organism's evolutionary success. Thereafter, the slow random loss of this differentiation begins to set in.

The life span of multicellular organisms includes the stages of development, reproduction and senescence. There is clear evidence to show early development being regulated by successive activation and suppression of specific genes. The gene regulation theory proposes that ageing (senescence) is the result of changes in gene expression due to reproduction and other adult activity. The gene expression needed for both growth and reproductive activity depends on several factors such as optimum levels of nutrition and can be affected by various types of stresses such as temperature, pollution, starvation, and radiation [82].

1.9 **The Importance of Age Estimation**

In recent years, there has been an increase in the number of situations in which a determination of age estimates is required. Therefore, there is a need for new molecular approaches to such age estimation.

1.9.1 **Crime**

In order to use DNA to estimate the age of an individual, the DNA changes associated with the ageing process need to be thoroughly studied and understood.

In many crimes scenes, a biological sample (mostly blood) can be collected, but there is often no direct suspect with which to compare it. If the DNA profile is determined it can be checked for a match in the database, but if no match is identified the profile is simply stored in the database for future comparison[83]. Standard DNA profiles provide no phenotypic information beyond the sex of the individual, but in the last few years, a number of studies have shown the possibility of predicting race and some of physical features of suspects and victims from DNA collected [84, 85].

DNA contains information about many human traits, such as eye colour, hair colour, and height [86]. Forensic DNA phenotyping (FDP) is an investigation method in which the identity from traits of an unknown suspect can be predicted by analysing crime-scene DNA. This prediction could include not only external characteristics, but also behavioural features and geographic origins, and even the surname in some situations [84, 85].

Phenotyping can be done by either of two approaches. The first approach is indirect phenotyping, where the genetic ancestry of the person predicts the external characteristics such as skin colour and hair colour. The other approach is direct phenotyping, where an external characteristic is directly determined from the responsible gene. Indirect phenotyping has been used in a few real criminal cases and it has shown positive results [87]. In contrast, the direct method is feasible, but not routinely conducted [88]. In FDP, many traits need to be predicted to reveal the identity of a suspect, each trait requiring confirmation by different informative markers. In 2002-2003, a panel of 71 SNPs was used to predict the race of a Louisiana serial killer[87]. So some physical feature by DNA methylation marker or a population genetic study including genome wide association studies (GWAS) might be helped better this case earlier [89].

In some crimes of particular interest, mass screening can be carried out, as when forensic DNA profiling was used for the first time in 1986, in England, to identify Colin Pitchfork as responsible for the rape and murder of two young girls [90]. In an attempt to identify a suspect, nearly 4,000 men from three nearby villages, aged between 13 and 34, were tested. The original suspect was excluded however [91], as when the men were asked to volunteer and provide blood or saliva samples, Colin Pitchfork arranged with a friend to give a sample in his name. The truth subsequently emerged when this friend was later overheard talking about the switch, and how he had given his sample under Colin Pitchfork’s name [92]. In similar cases, if the age of the person leaving the biological sample could be determined this would help limit the screening process to a specific age group. It would minimize the number of suspects and the number of samples to be tested, with obvious financial savings and lessening of public anxiety. Moreover, it would also help police to minimize the

search circle for suspects and criminal records to certain age groups or a certain criminal style. For instance, DNA profiling with an ability to provide investigators with clues to the age of the individual (whether perpetrator or victim) could be extremely useful.

In some criminal cases, the suspect has no valid identification documents and age determination may be important to determine if he or she has legal responsibility. Those under 14 are usually below the age of criminal responsibility, while those between 14 and 18 are subject to special criminal standards if accused of an offence. In most countries, anyone over 18 is considered to have full legal responsibility. In the future, in some countries such as Spain, suspects over 18 but under 21 may be subjected to the criminal standards applicable to those under 18 [93]. This shows how it is important to distinguish between biological and chronological ages.

1.9.2 **Immigration**

Immigrants without valid identification documents and who do not know their age may be suspected of making false statements with regard to their age. Age is of legal relevance in criminal, civil or asylum proceedings, with illegal immigrants under 18 potentially being placed under the guardianship of the authorities [94]. Precise age estimation would clearly be very useful in such circumstances.

1.9.3 **Civil Law**

In most develop countries; a birth certificate was not issued until a few decades ago [95]. In Saudi Arabia, for instance, until 1980 there were only a few people with this certificate. The birthday of the majority of Iraqi people who were born before 1970 is listed as either 1st of January or July. As a consequence, there is the possibility that individuals may make false statements regarding their age in order to study or meet job age requirements, when applying for driving licenses, and in other age dependent situations. Job retirement is also age-dependent and people who lie when they applied for a job or were assigned a specific date by the government, may try to misrepresent their age to extend their employment.

1.10 **Age Estimation**

Current age estimation for both living and dead subjects is dependent on several anatomical and morphological characteristics. Samples in which these criteria can be measured are rarely to be found in crime scenes. For instance, bones recovered from human remains after being buried or burned may undergo degradation. Samples recovered from a crime scene are also affected by weather and time, and need to be collected as soon as possible for accurate age estimation and this one of our goal to study.

1.10.1 **Method used for Age Estimation of Living Subjects**

Age estimation in living subjects has been well studied [96-98]. It is based on physical assessments, such as an X-ray of the left hand and the clavicles, and dental inspection, the results being pooled to estimate age. There will also be a physical examination to include measurement of height and weight, and an assessment of sexual maturity [93]. The investigations are carried out by a group of experts including forensic physician, dentist and radiologist, and cooperation between them is clearly required. The criteria obtained from X-ray images of the hand must be evaluated by analysing morphologic maturation status of all epiphyseal cartilages and the size of the sesamoid bone of the metacarpo-phalangeal joint of the thumb. These images must be compared with standard images for the relevant age and sex using, for example, a radiographic atlas [93].

When using teeth for age determination, the main focus will be on the stage of eruption and calcification of the permanent teeth and the size of the dental pulp cavity. It is also important to describe the average number of decayed, missing and filled teeth. In the early childhood teeth undergo significant numbers of morphological/biochemical changes, hence their value in age determination. After reaching maturity however, such alterations in tooth composition are significantly less and so age estimation from adult teeth is more difficult [99]. The forensic physician will summarize the results from these various examinations to give final age estimation.

The procedures are laborious and time consuming and the x-rays pose a radiation hazard, particularly in pregnant women. Magnetic resonance imaging (MRI) can be used to minimize radiation exposure but it is expensive and not often available [93].

1.10.2 **Method used for Age Estimation at Death**

A number of studies have investigated estimation of age at the time of death [100-102]. In 1920, Todd examined 306 males of known age at death and developed a ten-phase system for age determination based on changes in the public face, ranging from 18 to 50 [103]. Epiphyseal closure in the limb bones continues until the age of 25 and is thus useful in younger individuals. Hansen (1954) and Schranz (1959) used both the amount and arrangement of cancellous bone in the proximal end of the humours, the former decreasing with increasing age [104]. The reliability of using this method alone to estimate age is undetermined, and obviously it can only be applied when the proximal end of the humorous is present and intact [104]. Closure of skull sutures is another indicator of age, albeit imprecise, involving reading and scoring ten-suture closer regions [103].

1.11 **Mitochondria**

Mitochondria are double membrane structures and semi-autonomous organelles, typically rod-shaped with diameter of 0.5-1 μm length as shown in Figure 1-3. Each mitochondrion contains its own closed circular DNA chromosome and maternally inherited, because they come from female egg and not male sperm during fertilization [105]. Mitochondria are divided in to an outer and highly folded inner membrane “cristae”. The spaces between folds in the inner membrane called "matrix". The area in between both membranes is termed "intermembrane space” [106, 107]. As they are the place of oxidative phosphorylation in eukaryotic cells, they produce the vast majority of the adenosine tri-phosphate (ATP) energy. In addition to energy production they carry out many essential functions for cell survival like initiate programmed cell death and maintain intracellular homeostasis of inorganic ions [108].

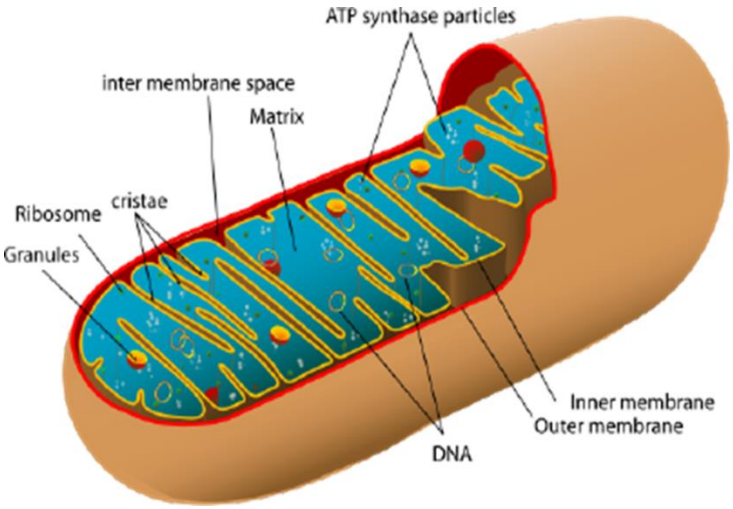

Figure 1-3 Mitochondria Organelle: : The figure shows typical mitochondria organelle. Reprinted from [109].

The mitochondrion particle has been discovered through a gradual process that means there was no single scientist who can be credited with discovering and identifying it. There are several scientists who can be credited with elucidating its structure and functions. Richard Altman in 1894 was the first one to recognize and discovered the existence of this organelle and called them bioblasts, but Carl Banda introduced the name, mitochondrion, in 1898 [110, 111]. The mitochondria, plural for mitochondrion is Greek word, where “mitos” meaning ‘thread’ and “chondros” meaning ‘granule’ [112]. With rare exceptions (erythrocytes), all eukaryotic cells contain mitochondria [113]. Methylation patterns in a mitochondrial study could provide information on many phenotypes such as health status and age. Estimation of the age of contributors is one of the important characteristics that could help the investigator through methylation rate in specific genes. Second, the study provides opportunities to address the age through the more abundant mtDNA.

1.11.1 **The Mitochondrial Genome**

The mitochondrial human genome was first detected by Nass and Nass in an early 1960s [114]. It believed that they existed in physical and genetic isolation for about

two billion years of eukaryotic evolution. The mitochondria contain more than one copy of a DNA genome [115]. Although mtDNA is replicated autonomously within the organelle system, but still there is a highly interdependent with nuclear genomes [113]. The synthesized nuclear coded factors in the cytosolic ribosomes are imported in to the mitochondria as precursor polypeptides through specialized import pores [114]. Excluding mutations, the female passes the mitochondrial genome to her children, thus for offspring the mtDNA type is not solely unique to them proving identical among maternal relatives especially siblings [116]. The mitochondrial genome also encodes some functions which prove essential to the survival of cells, hence there is an 'endosymbiotic' relationship between mitochondria and the cell as that of a host and obligate parasite [113].

Although the DNA sequence of mtDNA represents less than 0.001% of total cellular DNA, the copy number of mitochondria is usually high making up 1/5 of the cell volume, depending on the tissue types ranging from few number to thousands genomes per cell [115]. The two strands of mtDNA have significantly different nucleotide compositions and they can be separated from each other due to their different nucleotide composition: the heavy (H) strand is purine-rich (guanine) and the light (L) strand is pyrimidine-rich (cytosine). Because of this distinction, the DNA strands are traditionally called heavy and light strands. Although, in principle, the mitochondrial DNA is double-stranded, it has been proved that a small portion is known as a triple helix DNA strand structure (7S DNA). Among thousands genes in the human genome the mitochondria contain 37 genes for normal function, from this 28 of the genes use H strand as their sense strand and 9 use L strand. From these, 13 of them encode protein, 22 transfer RNA genes and 2 ribosomal RNA genes [117] as shown in Figure 1-4

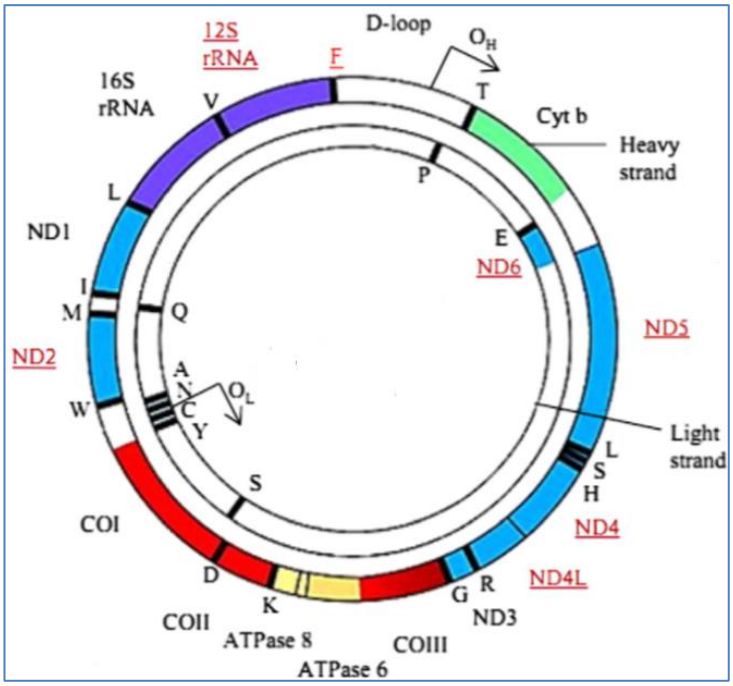

Figure 1-4 The map of Mitochondria: The figure shows all the 37 mitochondrial genes, origin of both heavy (OH) and light strand (OL) replication, respectively. Outer and inner rings correspond to both strands, respectively. Genes that are differently methylated are shown in red and underlined. Taken from [118]**.**

Encoded proteins of mtDNA are vital subunits of the respiratory chain: what is known as energy synthase, specifically seven subunits of complex I (NADH dehydrogenase), complex IV which has three subunits (cytochrome oxidase), another complex which has two subunits known as complex V, and finally one subunit (ubiquinol: cytochrome c oxidoreductase) known as complex III [113]. Anderson *et. al.,* sequenced and organized the human mitochondrial genome and they published their discovery in 1981 [119], followed in 1999 with the re-sequencing of the original mDNA sample and correcting of mistake and infrequent polymorphisms [120]. The original sequence now acts as reference sequence and new sequences are compared to this reference which known as 'Anderson or Cambridge reference sequence' [8].

The mitochondrial DNA sequence is extremely compact compared to the nuclear genome, because nearly 93% of the mitochondrial genome representing coding

sequence [119], in comparison to the human nuclear genome where only ~ 2% represents the coding sequence [121]. In mitochondrial genome the majority of the open reading frames (ORF) for the genes are separated by one or more transfer RNA (tRNA) genes. Hence, unlike nuclear gene DNA sequence, there are no intron with few, if any, extra nucleotides in between [122].

Mitochondria are often referred to as the 'power house' in producing energy for the cell via oxidative phosphorylation process. Through this process the main energy source of the cell adenosine triphosphate (ATP) are produced from oxygen and sugar. Mitochondrial reactive oxygen species (mROS) are produced from oxygen which consumed in mitochondria reviewed in [113]. In 1956, Harman first observed the free radicals' ability to cause DNA damage, and proposed that ageing results from accumulated DNA damage inflicted by free radical atoms: reviewed in [123]. But after twelve years, Harman changed his theory and mentioned that mitochondrial oxidative damage is responsible for ageing [124].

However, until 1989 there was not any clear evidence about the 'mitochondrial theory of ageing'. But, Linnane forwarded this theory and suggested that free radical damage causes somatic mutations in the mitochondrial DNA [125]. Subsequently segregation of cytoplasm during life is "the major contributor to loss of bioenergetics ability within tissues and thus ageing". Furthermore, it has been observed that long-lived species have lower cellular oxidative damage comparing to short-lived species [126]. The absence of mitochondrial DNA repair enzymes makes mtDNA more vulnerable to mutations, and the link to ageing becomes more plausible because of the close neighbourhood of the oxidative electron transport chain to mtDNA [124].

1.11.2 **Mitochondria and Forensic Science**

mtDNA is small comparing to nDNA and much simpler than nDNA, but it is important for forensic investigation for several reasons. One is that the mitochondrial genome mutates approximately ten times the rate of nDNA[127], experiencing a net change of more than 2% per million years. Thus in just 10,000 years, four persisting nucleotide changes might be expected [128]. As a result of this rapid mutation, mtDNA has been used as an indicator to investigate human ethnic groups in

thousands of years [129]. It is believed that exposure of mtDNA to free radicals which generated during respiration causes this elevated mutation rate [130].

Another significant reason for the forensic interest is the maternal inheritance of mitochondria. Even though the presence of the mitochondria in the sperm's tail, but they are destroyed by when they entered the zygote [9, 131]. In contrast, the mammalian egg contains hundreds to thousands copies of mitochondrial DNA [10, 132]. Finally, the large copy numbers of short mtDNA sequence comparing to only two copies of any segment of chromosomal DNA per cell helped mitochondrial genome to isolate and study easier [133]. This feature is especially important in cases in which nDNA is too degraded or there is not enough DNA to get a profile such as in the case of disaster victim identification (DVI) [134]. Hence, the genome of this small organelle can often be typed more successfully than nDNA, especially in low template case, touch samples or hair shafts. Mitochondrial DNA from a maternal relative could even be used to exclude the suspect when the perpetrator is not available for testing [116]. Table 1-2, shows the comparison between nDNA and mtDNA [116]. Numerous studies used mitochondrial genome for forensic [9, 116, 135] medicine purposes [136, 137], archaeological purposes (Identification of the Russian royal Family and Richard III) and mitochondrial mutation as an age indicator [40].

Table 1-2 nDNA vs mtDNA: Comparison of human nuclear and mitochondrial DNA [116].

Feature	nDNA	mtDNA
Cellular location	Nucleus	Mitochondria
Structure	Linear	Circular
Size	3.2 billion bp	16.569 bp
Genetic Unit	2 x 23 chromosomes	mtDNA genome
Mode of Inheritance	Bi-parental	Maternal
Ploidy	Diploid	Haploid
Unique to individual	Yes	No
Copy number	1-2 per cell	>1,000 per cell
Recombinant	Yes	No
Mutation rate	Low	10 times nDNA

1.12 Recent Developments in the Mitochondrial Theory of Ageing

1.12.1 Mitochondrial Theory of Ageing

Mitochondria are central components of cells and work as a site of energy production. As a result of this activity, it generates some chemicals (e.g. ROS) which are unstable; hence they harm other components of the cell and the mitochondrion itself jointly [138]. It is believed that the resulting trauma plays a role in ageing. Proposing that the mitochondrion has crucial roles in ageing, then this trait could very possibly be used on forensic casework samples as an indicator for the age of the donor at the time of formation of a stain or biological trace evidence.

The nDNA and mtDNA are replicated indifferent ways. In mitochondria the genome is replicated continuously and does not depend on the cell cycle, while nDNA it replicates once during cell division [139]. Replication of mtDNA is controlled by several nuclear encoded proteins. DNA polymerase known as polymerase γ (pol-γ) which encoded by the nuclear gene *POLG* is the only enzyme known to be

responsible for mitochondrial DNA replication and maintenance [140]. Two types of mutations are observed; mitochondrial sequence point mutations and large deletions. The exact source of the mutations in the mitochondrial granule is still controversial, but it might be come from the same source of nuclear DNA mutations. Currently research on mitochondrial mutations has been carried out with PolG and a variety of mutation rates observed in different ages of mice [137]. An accumulation of mutation causes disease in mitochondria causing them to lose their function. Hence, they probably have a central role in the ageing. A higher rate of mutation is observed in older mice than in wild-type [141].

In spite of several lines of evidence suggesting of correlation of mutation of mtDNA with age, this may be not the major cause of the ageing process [142]. In addition this data may not be applicable directly to human ageing and the reason for that is some types of mutations like large deletion mutation has been noticed and show an important role in human ageing. However, this type of mutation is not found in mice [137]. Furthermore, unlike mouse mitochondrial genome, lower level mutations in human mitochondrial DNA are associated with deficiency in oxidative phosphorylation (http://ghr.nlm.nih.gov).

Therefore, while for mouse, mutations in mtDNA may not affect the ageing process they may still play a role in human ageing. A previous study has identified mutation accumulation in mitochondrial DNA with age, especially in the hypervariable regions (HV1 and HV2) [136]. Beside the fact that age related deletion and point mutation outside control regions cannot be excluded, but for a forensic application the hypervariable regions in mitochondria do not appear to be a useful indicator to predict the age of the contributor who left body fluid behind at a crime scene [40].

1.12.2 **Mitochondrial Heteroplasmy**

A variety of mutations in mtDNA have been identified in human somatic tissues that fit into one of three categories: deletions, point mutations and small duplications [40]. Such mutations are thought to occur throughout the lifespan, gradually accumulating with age [143]. In the instance that mtDNA has not undergone mutation; all mtDNAs present in a cell would be expected to share 100% identity.

This condition is referred to as homoplasmy [144]. However, following a mutation event cells will be expected to contain a mixture of two types of mtDNA (mutant and wild-type) mtDNA, this condition referred as heteroplasmy [117]. Sometime tissue with three types of heteroplasmy reported Table 1-3.

Table 1-3 Heteroplasmy types and their occurrence: Reproduced from [144]**.**

	Heteroplasmy type	Genetic Characteristics
1	Intercellular	Wild-type and mutant mtDNA are in different cell
2	Intracellular intermitochondria	Wild-type and mutant mtDNA are in different mitochondrial organelle, but in the same cell
3	Intracellular intramitochondria	Wild-type and mutant mtDNA in the same mitochondrial organelle

1.12.3 **Mitochondrial mutations and heteroplasmy that accumulate with age**

One particular region in the mitochondrial genome, the control region, is the most frequent site to undergo mutation; especially point mutations [136]. This makes the control region the most genetically variable section of mtDNA. This is understandable as the control region contains no protein-coding genes, but does instead hold the replication and transcription initiation sites reviewed in [136].

Although most age-accumulating mutations in mtDNA have been localised to the control region, there are a few exceptions. For example, an A→G substitution at position 3243 has been identified by Trifunovic to accumulate with age [136].

1.12.4 **Mitochondrial Deletions Accumulate with Age**

The mtDNA deletions are known to accumulate with age, by far the most recurring is a 5 kb deletion located at position 8470-13,447; consequently known as the "common" deletion [145]. The accumulation of this particular deletion in the

mtDNA genome with age was first identified in brain and heart tissue by Cortopssi in the 1990s [145], and later on wide range of other tissue types have been reported.

Many important protein and tRNA encoding genes are located within the 5 kb mtDNA region affected by this deletion. This includes complex subunits I and IV of the electron transport system and also subunit III of cytochrome oxidase[146].

In normal cells, cytochrome oxidase is considered as main site of oxygen reduction, hence plays an important role in protection of cells against metabolism- derived oxidants. The defective cytochrome oxidase enzyme synthesized as a result of this mtDNA deletion likely leads to increased cellular oxidative damage.

The path of whole genome amplification has been employed by [105, 142, 147-149]. They tried to identify large mtDNA deletions and to look at the correlation between the age of the contributor(s) and the mutations. All researchers were able to confirm the presence of several mtDNA deletions and their accumulation in tissues with increasing age.

1.13 **DNA Methylation and Ageing**

A link between ageing and a reduced DNA methylation level in mammals had been established for more than four decades [150] when it was observed that DNA methylation in a mammalian fibroblast culture was positively correlated with the number of divisions of these cells, while in immortalized cells the methylation level was not correlated [151]. It was proposed that age related dysfunctions in cells and tissues could be caused by impaired DNA methylation [152]. More recent methods of analysis have confirmed that the amount of genomic DNA methylation varies with age. Furthermore these changes can be complex with different effects in different CpG areas, reviewed by Richardson [153]. It is difficult to account for this change by means of purely stochastic processes accompanying DNA methylation and demethylation. This correlation, between methylation and age, agrees well with the concept of programmed ageing that could be caused by changes in DNA methylation [154]. The discovery of biomarkers of ageing epigenetic clocks that enable the prediction of the biological age of most cells and tissues has become the most

important development for gerontology and epigenetics [155]. There are thousands CpG sites with the methylation levels correlated with age that have been precisely mapped. The same report also demonstrated that: "manipulations used to obtain embryonic and induced pluripotent stem cells move the epigenetic clock to the zero mark"[155].

Chronic inflammation is one of the most thoroughly studied factors known to cause changes in the methylation level [156, 157] but the reasons for this remain unknown. According to this phenomenon, referred to as "inflamm-ageing", a prolonged inflammation causes an imbalance of the immune system and facilitates the development of a senile phenotype [158]. Bellizzi and co-worker in 2006 provided a first experimental evidence of mitochondrial role in this process [159] where cellular mtDNA was replaced with another haplotype. They found a presence of a relationship between the mitochondrial and nuclear genomes that can guide the inflammatory processes that accompany ageing. The functioning of the mitochondrial respiratory chain is coupled to the production of toxic mitochondrial reactive oxygen species (mtROS) which causes oxidative stress in mammalian cells. Data suggests that mtROS are involved as messengers in the inflammatory response of endothelium in old mice [160] and they are present in almost any type of oxidative stress caused by either exogenous or endogenous factors [161]. A report has demonstrated that oxidative damage by hydrogen peroxide treatment leads more tightly bound DNMT1 with chromatin and that the DNA methylation level is also changed [162]. It has been found that overall levels of mtDNA mutations are low in tissues. An accumulation of these mutations led to a clonal expansion of mtDNA mutant versions and tissue dysfunction [163].

It is notable that the mitochondria's reparative processes were also underestimated. The same report revealed that mtDNA have multiple repair pathways, including base excision repair (BER) and mismatch repair (MMR) systems [164]. It has been reported that although ageing mice have a 500 fold increase in mutations than normal mice, the mice were able to sustain accumulated heterozygous mutations [165]. These observations do not support the stochastic theory of ageing. To date very few studies of ageing and mtDNA methylation have been reported. The methylated

cytosine level in the 12S mtDNA gene decreased with age [118] and the 5hmC level in cerebral cortex of mice increased with age [166]. It has been found that methylation at sites located before the mitochondrial genes ND6, ATP6, and COX1 decrease during the embryonic development of the human brain [167].

1.14 **Evidence of Epigenetic Mechanisms in Ageing**

The methylation degrees of nuclear genes encoding mitochondrial proteins are correlated with the level of gene expression and mitochondrial activity [168]. J. Hayashi and his group obtained data provide the most striking evidence of the functioning of epigenetic mechanisms in ageing. The respiratory chain does not work so effectively in old people and this leads to a decrease in the oxygen consumption rate and the formation of a “senile” cell phenotype. They also presented data concerning the mechanisms involved. Thus processes of epigenetic cause age related mitochondrial dysfunctions by modifying the expression of genes responsible for the mitochondrial metabolism. Absence of glycine impairs mitochondrial translation and leads to this ageing phenotype [169]. Therefore, reduction in enzyme activity that is involved in synthesis of glycine can significantly affect not only translation, but also methylation, providing epigenetic changes in gene expression leading to metabolic “senile” changes.

Furthermore, some nuclear repair enzymes are targeted to mitochondria under oxidative stress. Interestingly these enzymes apparently perform somewhat different functions in mitochondria than in the nucleus [170]. It is possible that oxidative stress affects mitochondrial replication and transcription. Evidence suggests a close relationship in the regulation of mitochondrial and nuclear gene expression and methylation might be one of these regulation methods that also have a function in ageing. The latest advances in the determination of the CpG sites reflecting biological age [155]. The discovery of key role of the epigenetic modifications during senescence of cell [169], do support the earlier proposed hypothesis, concerning nDNA methylation as a mechanism of programmed ageing [154]. The mtDNA methylation and retrograde signalling might also play an important role in the realization of programmed ageing.

The link between mechanisms at the molecular level and the processes of ageing are emerging as the analytical techniques become more refined. There are also other factors, primarily the telomere length, that are clearly involved. While the purpose of this thesis is to study the potential of the estimation of the age of a donor of biological traces to the investigation of criminal acts, it might be useful to briefly review the molecular ageing processes in order to assess the possible advantages of such an approach.

1.15 **The Problems of Firearm Crimes**

The genetic information of evidence that is left in crime scenes are badly needed for forensic investigation. According to United Nations office on Drug and Crime (UNODC) 2011 annual report thousands of homicides by firearm occur around the world [171]. In 2011 just in the US there were more than 12,000 murders. The majority of those were caused by firearms [172]. In England and Wales homicide rate by firearm is 6.6% of all murders. But the countries mentioned do not have the worst firearm murder rate. The highest rates are Honduras and El Salvador with 83.4% and 76.9% gun-related homicides, respectively [171]. In a country like Iraq the gun number will not be less than the population size, so firearm crime is very common.

In spite of the best efforts of law enforcement agencies and prosecutors, nearly 40% of the US homicides remain unsolved each year cited by [173]. Moreover, fired cartridge cases are considered as a poor source of DNA, thus some analysts report that there was not enough DNA to get a DNA profile in their experience [28]. Others, however, believe that it is probably a waste of time and money to test them [29].

In any crime there may be a multiple suspects, so there is a need to know how to prioritise. There is not a data base for whole population, so just with DNA profile we can only identify people who have previous profile. The information obtained from the physical features of an individual may not bring the person to justice directly, but could be aid the investigator in undertaking investigations at a slightly more rapid pace. Therefore, it is of great importance to develop new methods and techniques to analyse evidence recovered from crime scenes.

1.16 **Aims and Objectives**

The overall aim of the work reported was to study the possible types of information that can be reconstructed from trace DNA recovered from cartridge cases left by automatic weapons used in the commission of a crime. Several different aspects of this aim were considered.

First was an appraisal of commercially available DNA recovery kits, or systems, sold for the purpose of recovering DNA from traces left by touch such as those that might be expected on cartridge cases or magazines.

A second aim was to make use of recent technical advances in the tests of methylation level of certain selected sites within the genome and the mitochondrial DNA. These advances, in the newly emergent field of epigenetics raised the possibility of making an estimate of the age of the person who was the source of the sample.

The relative effectiveness of different types of sample recovery of DNA using three types of commercially available swab was conducted so that the most effective system could be used in subsequent parts of the work.

The potential for making an estimation of donor age from such recovered DNA, was approached by considering the correlation between the nuclear DNA methylation status and the age of the donor at different genetic loci.

The methylation levels found within selected mitochondrial genome sites were studied as a region of human heritable material that has not received wide attention. Further, the study sought to identify the exact CpG sites which have the best epigenetic signature correlated with age which could be used as an epigenetic marker tool to predict donor age, under varying conditions relevant to forensic case work.

Finally, a further study was conducted to assess DNA methylation using an EpiTect qPCR methyl system as a novel methodology in order to investigate the potential forensic contribution this method could offer in estimating the approximate age of the donor of a DNA trace.

CHAPTER TWO

Materials and Methods

2 Chapter 2 Materials and Methods

2.1 Materials

Standard suppliers, such as: Sigma-Aldrich, Invitrogen and Fisher Scientific, provided chemicals materials, molecular biology reagents and plasticware unless otherwise stated DL-Dithiothreitol (DTT) from (Sigma-Aldrich) lot #3H2610.

Agarose (Sigma, Lot# 105H1256) 10X UltraPure™ Tris-Acetate EDTA (TAE Buffer), (Life Technologies, Lot #:025K8400) 6X Gel Loading Dye Buffer, Blue (BioLabs, catalog #: B7021S). 100bp and 1K DNA ladder standard (BioLabs, catalog #: N3232S). SYBR® Safe DNA Gel Stain (Life Technologies, Lot #: S33102). Horizontal Gel Electrophoresis System (Life Technologies, catalog #:00188139. Power supply, Gel casting tray and combs, DNA stain and Staining tray.

2.2 Participant Recruitment

This research project was approved by the University of Strathclyde Ethics Committee and prior to sample donation, participants signed informed consent statement.

2.3 Swabbing

Two buccal samples were taken using ForensiX swabs by the investigator upon instruction. Saliva was collected by rubbing the swab head across the inside of the cheek for approximately half min per swab the buccal cells were collected. Swabs were then placed back into their tube envelope and sealed. To decrease the biological parameter diversity, only females were recruited and their ethnic background and ages as phenotype information were recorded independently by the principal investigator.

2.4 Sample Collection

Saliva and blood collection (via finger) from volunteers was carried out in the Strathclyde University laboratory (the swabbing technique mentioned in 3.2.1). Blood collection via venepuncture was collected in Erbil hospital by a healthcare expert. The analysis steps were done at University of Strathclyde, while the firearm shooting and the DNA extraction from fired cartridge by using robotic method was done at Police Forensic laboratory in Erbil. Finally, the sequencing step was done at Zymo Research Corporation (CA, USA). The purpose of the study was explained to the donors first. Biological samples were collected after ethical approval, and consent forms were signed by the volunteers. Initially, all the volunteers of this study were female age range (18-91 years).

2.5 Collection of Blood

Two methods were used for blood collection.

2.5.1 Blood Collection via Finger Stick

Disposable lancets were used for blood collection, about 50 µL of blood by finger stick was dropped on a clean sterile cotton swabs.

2.5.2 Blood Collection via Venepuncture

Around 5 mL of blood was collected by standard venepuncture from volunteers. Then 100 µL aliquots were immediately placed in Eppendorf tube and extracted and a similar volume was spotted on the clean sterile cotton swabs and stored in -20 ºC.

For EpiTect methylation analysis experiment, a total of 40 unrelated female individuals with age ranging 18-91 years old were sampled.

2.6 QIAamp® DNA Investigator Kit

To isolate total DNA from buccal swab QIAamp® DNA Investigator Kit, lot #56504 (QIAGEN, Crawley, UK), was utilised, following the manufacturer's guidelines with few modifications.

The technique in this kit is well established for extraction of genomic DNA even from small sample sizes [174]. Hence the procedure is suitable for a wide range of forensic and human identity samples. The procedure depends on the combination of selective binding property (silica –based membrane) by using QIAamp MinElute spin columns for purification. The high quality and efficient of the extracted DNA was another reason to utilise this method in which the product will be free of inhibitors (nucleases and proteins). Lastly, this kit is highly recommended for extraction of a number of samples simultaneously and in short time [174].

In first step the head of the swab was separated from the shaft and placed in 2 millilitre (mL) microcentrifuge tubes. To breakdown the cell and release the contents, 20 µL proteinase K and 400 µL of ATL buffer were added, followed by pulse vortex for short time (15 s) and incubated at 56 ºC for one h with shaking. After shaking 400 µL of AL buffer was added. This buffer facilitates breakdown of the cells, plus 1 µL of carrier RNA to remove nuclease. To homogenate, the contents were mixed and incubated again, but this time at 70 ºC for only 10 m. Before transferring the supernatant, 300 µL of absolute ethanol was added. To homogenise the mixture, the content was thoroughly mixed. To harvest more lysate, the remains of the swab were placed in QIA shredder spin column and suspended then collected in MinElute tube. In next step the tissue lysate in QIAamp MinElute spin column was run by centrifugation at 8,000 rounds per min (rpm) for one min. To wash the lysate 500 µL of AW1 was added and centrifuged at 8,000 rpm, followed by similar washes with 700 µL of AW2 buffer. The column was subsequently dried by centrifugation again. To this lysate, 700 µL of absolute ethanol was added and the spin column was suspended again using the same condition. The QIAamp column membrane was dried completely at full speed centrifuge for 4 min. The MiniElute spin column was placed in to new collection tube and left for 10 min at room temperature. Finally DNA was eluted into 40 µL buffer ATE by centrifugation at full speed for 2 min. In this final step the purified DNA was directly used for further analysis or stored at -20°C for later use.

To identify interested gene promoter methylation, EpiTect® Methyl qPCR Array (SA Bioscience, Qiagen, Hilden, Germany) has been used and steps are performed followed instruction of manufacture as mentioned separately in this chapter.

2.7 EpiTect® Methyl qPCR Array

2.7.1 Restriction Digestion

After measurement of DNA concentration by qPCR quantitative method (see Section 3.2.2), the extracted DNA was measured for purity purposes by Nanodrop-1000 UV-vis spectrophotometer. Only those samples which passed both above criteria were exposed to digestion performed by using EpiTect Methyl II DNA Restriction Kit (SA Bioscience, cat#: 335452). The final volume of components was mixed thoroughly and centrifuged briefly in a micro centrifuge tube. Four reaction digestions were carried out: no-enzyme (Mo), Methylation Sensitive enzyme (Ms), Methylation dependent enzyme (Md) and Methylation sensitive and dependent enzymes (Msd) as showed in Table 2-1

Table 2-1 Restriction Digestion Step: Reaction digestions set up for each gene and each sample. Four tubes were labelled as (Mo, Ms, Md and Msd). Sensitive A or dependent B enzymes were added to Ms, Md and Msd tubes respectively, but ddH_2O was added to Mo instead of enzyme.

Component	M_o	M_s	M_d	M_{sd}
Reaction Mix	28 µl	28 µl	28 µl	28 µl
Methylation sensitive Enzyme A	--	1 µl	--	1 µl
Methylation- dependent Enzyme B	--	--	1 µl	1 µl
RNase-/DNase-free water	2 µl	1 µl	1 µl	--
Final volume	30 µl	30 µl	30 µl	30 µl

The components were mixed gently by pipette and briefly spun down. The mixture was incubated at 37 ºC overnight, followed by heating for 20 min at 65 ºC to inactivate the enzymes.

2.7.2 **PCR Set up**

Individual reactions were prepared for each of the four digestions (M_o, M_s, M_d and M_{sd}) in a tube according to Table 2-2.

Table 2-2 PCR Set Up: The table shows the PCR set up preparation by adding right tube contents at right well of the 96 well plate. Taken from [175].

Component	M_o	M_s	M_d	M_{sd}
PCR master mix	12.5 µl	12.5 µl	12.5 µl	12.5 µl
PCR primer mix	1.0 µl	1.0 µl	1.0 µl	1.0 µl
Mo digest	5.0 µl	-	-	-
Ms digest	-	5.0 µl	-	-
M_d digest	-	-	5.0 µl	-
M_{sd} digest	-	-	-	5.0 µl
RNase/ DNase free water	6.5 µl	6.5 µl	6.5 µl	6.5 µl
Final volume	25 µl	25 µl	25 µl	25µl

2.7.3 **QPCR instrument Setup to Quantify DNA Methylation Analysis**

The EpiTect Methyl II PCR Arrays which does not require bisulphite conversion can be run on any real-time PCR instrument with instrument-specific SYBR Green Master Mix acquired from SABiosciences. As a consequence of this, RT2SYBR Green ROX™ qPCR Mastermix (SA, Bioscience, cat# 330520) was selected in this experiment to be used with our Stratagene Mx3005P real time PCR instrument (Agilent Technologies, CA, USA). SYBR® Green (with dissociation curve) protocol was used thermal profile set up with following condition as shown in Table 2-3.

Table 2-3 real time PCR programme: The table shows the PCR condition to quantify genomic DNA for methylation level study.

Temperature	Time	Number of cycles
95°C	10 min	1 cycle
99°C 72°C	30 s 1 min	3 cycles
97°C 72°C	15 s 1 min	40 cycles

For dissociation curve, segment three was added to the programme: 95 ºC for 1 min and one cycle, then annealing at 55 ºC for 30 s and another denaturation at 95 ºC for the same period and one cycle.

2.8 **ZR Genomic DNA ™-Tissue MicroPrep**

Simple and rapid extraction procedure of total DNA from donor samples were performed first using the ZR Genomic DNA ™-Tissue MicroPrep *Kit* (ZYMO RESEARCH, USA), following the manufacturer's guidelines with some changes to optimise performance with the specific swab sample type. In the first step the head of the cotton swab was put in a clean microcentrifuge tube. To this 20 µL of Proteinase K enzyme (Sigma Aldrich Company Ltd., Dorset, UK), with approximately 190 µL of DNA/ RNase-free water (Ambion, Life Technologies, Paisley, UK) and 2 X digestion buffer were mixed and incubated at 56 ºC for 3 h. Then 700 µL of genomic lyses buffer was added (prepared 0.5% (v/v) with β-mercaptoethanol (GE Healthcare, Little Chalfont, UK). To remove insoluble debris of the cells, the tube contents were thoroughly mixed by vortex and then centrifuged for 1 min at 10,000 x *g*. The supernatant from the collection tube was transferred to Zymo-spin™ IC column and spine-downed again at the same speed and for the same time (to facilitate breakdown of the cells). The column content was washed twice with DNA pre-wash and g-DNA wash buffers respectively and after each cleaning step the tube and the spin column were pelleted by centrifugation at 11,200 rpm for 60 s. The

cleaned DNA was eluted in 20 µL elution buffer by centrifugation at top speed. At this stage the DNA was preceded for further analysis or stored in -20 ºC.

2.9 DNA quantity and Quality Assessment Protocol

2.9.1 Nanodrop-1000

DNA concentration was measured with the Nanodrop-1000 UV-vis spectrophotometer (Thermo Scientific, Loughborough, UK). The instrument was first blanked with a sample of DNase-free water or elution buffer (the same solvent as the eluted DNA sample), and 1.5 µL of DNA sample used for calculation of concentration, A260/280 and A260/230.

2.10 DNA Quantification

The extracted DNA was quantified using the Quantifiler® Human DNA Quantification Kit (Life Technologies, CA, USA). A master mix reaction was made up as shown below Table 2-4.

Table 2-4 Master Mix set up: the table shows equal amount of both reaction mix and primer mix per reaction.

Component	Volume per 12.5 µL reaction (µL)	Total volume for reactions (µL)
Reaction Mix FQ	5.75	
Primer Mix IC FQ	5.75	
Total volume (Minus 1 µL DNA sample)	11.5	

A fresh serial dilution of Control DNA Z1 (20 ng/ µL) was prepared according to the manufacturer's protocol as described in Table 2-5. Each diluted control was mixed by pipetting before removing an aliquot for the next dilution. A negative control was also included.

Table 2-5 Serial Dilution: The table shows a serial dilution of the standard DNA which decreasing in concentration, beginning at 20 ng/µL and the last diluted concentration is only 0.00488 ng/µL**.**

Serial dilution of Control DNA Z1	Control DNA Z1	QuantiTect Nucleic Acid Dilution Buffer
20 ng/µl	Undiluted DNA	–
5 ng/µl	10 µl	30 µl
1.25 ng/µl	10 µl	30 µl
0.3125 ng/µl	10 µl	30 µl
0.078125 ng/µl	10 µl	30 µl
0.01953125 ng/µl	10 µl	30 µl
0.0048828125 ng/µl	10 µl	30 µl

The standards and negative control were included in each run in duplicate form. Thermal cycling was conducted on a Stratagene Mx3005P real time PCR instrument (Agilent Technologies, CA, USA) using an absolute quantification (standard curve) assay. The thermocycling was programed as shown in Table 2-6.

Table 2-6 Thermocycler Program: The table shows different steps of qPCR program.

Step	Temp	Time		Additional comments
Initial PCR activation step	95°C	1 min		PCR requires an initial incubation at 95°C for 1 min to activate the DNA polymerase
Two-step cycling:				
Denaturation	95°C	5 s	40 cycles	
Combined annealing/ extension	60°C	32 s		Perform fluorescence data collection

The known concentrations of Z1 DNA in the standard column are used to construct a standard curve which can be used to expect the unknown concentrations of the initial samples.

The analytical software used to interpret the sample concentrations according their cycle threshold (C_T) value. The final DNA quantity for each sample was recorded based on the average result of the two duplicate reactions per sample.

2.11 DNA Amplification and Profiling

Some samples from each set were amplified using a 2720 thermal cycler (Life Technologies, UK) with the Investigator Human Identification PCR Kit (QIAGEN). While fired cartridge samples were processed for STR profiling using the Minifiler kit (Applied Biosystem, part #: 4374618) following the protocols as recommended by manufacture [176]. After running samples on the 3130 Genetic Analyser (Applied Biosystem), the expected DNA profile was produced. There was no evidence of contamination.

2.12 Sanger Sequencing

To sequence the amplicons (PCR products) of all genes of interest, the amplicons were sent with appropriate sequencing primers to Source Bioscience (UK Limited, Nottingham) for sequencing. Then the sequences were visualised and analysed using Finch Tv Version 1.4.0.

2.13 Next Generation Sequencing (NGS)

One of the project's aims was to investigate new technologies. Illumina sequencing was used in this study as it is laborious to use small pool PCR, hence high throughput DNA sequencing was used to analyse methylation pattern of a set of age related genes within target regions (discussed in detail later in chapter four of this report).

2.13.1 NGS Library Preparation

The library preparation and sequencing were carried by Zymo Research. Multiplex amplification of all samples using ROI specific primer pairs and the Fluidigm Access

ArrayTM System was performed according the to the manufacturer's instructions. The resulting amplicons were pooled for harvesting and subsequent barcoding according to the Fluidigm instrument's guidelines. To reduce the time during PCR, sequencer-specific tags and sample-specific barcodes are added to each of the PCR products. The universal forward tag is called common sequence 1 (CS1), and the universal reverse tag is called common sequence 2 (CS2). After barcoding, samples were purified (ZR-96 DNA Clean & Concentrator™ - ZR, Cat#D4023) and then prepared for massively parallel sequencing using a MiSeq V2 300 bp Reagent Kit and paired-end sequencing protocol according to the manufacturer's guidelines. In regards to multiplex sample amplification and Illumina sequencing-based barcoding schemes. The consensus sequences of the Fluidigm adaptors are the following:

CS1= 5'-ACACTGACGACATGGTTCTACA-3'.

CS2 = 5'-TACGGTAGCAGAGACTTGGTCT-3'.

The CS1 and CS2 sequences are included in each of the target or locus-specific primers. The Illumina indexes (or barcodes) are then added in the subsequent barcoding step.

2.14 **Bisulphite Sodium Conversion**

For pre-treatment bisulphite conversion step, EZ DNA Methylation-Direct TM Kit (Cat ≠ D5020) has been chosen, because the manufacture (Zymo Research) claimed that this new product has higher efficiency than other protocol.

2.14.1 **Procedure**

Prior the first use, CT conversion reagent and M-wash buffer were prepared according to the manufacturer's instruction with a few modifications in the PCR set up program (increased denaturation time and decreased input DNA). First step of the protocol started with addition of 130 µL of prepared reagent of CT conversion to 20 µL of purified DNA sample (around 150 ng) in PCR tube. The mixture was pipetted up and down then the liquid centrifuged to the bottom of the tube. The tubes were

placed in thermal cycler and the machine programme set up on alternative parameters to yield improved result as shown in Table 2-7 below.

Table 2-7 DNA Conversion: Treatment of DNA with sodium bisulphite.

Step	Temp	Time
DNA Denaturation	98 ºC	10 min
Nucleotide conversion	64 ºC	3.5 h
Extension of conversion	53 ºC	6 min
Store	4 °C	Up to 24 h

After bisulphite modification, 600 µL of M-binding buffer was added to a zymo-spin™ IC column and the column placed into a provided collection tube. The bisulphite treated sample was loaded into this spin column and mixed gently. The contents were centrifuged at full speed (13,000 rpm) for 30 s then the flow-through was discarded and another spin down similar to previous step was performed. The columns were washed with 100 µL of prepared M-wash buffer. Later on the columns were spun down at full speed again. To remove the un-used bisulphite material from the column, the incubation step was done by addition of 200 µL of M-desulphonation buffer and the content allowed for 15 min at room temperature (20 ºC – 30 ºC). Next to incubation period a similar centrifuge step was repeated again. Finally, the bisulphite treated DNA was eluted by adding 10 µL of M-elution buffer to the column matrix and the eluted volume was collected in the collection tube by full speed centrifuge for 1 min. Upon this stage the DNA was either analysed immediately or stored at -70 ºC for later use.

2.15 **Primer Design Guidelines**

Each pair of the primers were designed according to the region of interest (ROI) sequence and followed as closely as possible, various key, well known aspects and software programme for primer design were used like (http://primer3.ut.ee/) and the parameters that had taken in account were:

2.15.1 **Primers Guidelines for Quantification of DNA (Double Strand)**

1- The percentage of nucleotide base content of guanine (G) and cytosine (C) within the sequence was from 40 to 60%.
2- Wherever possible the sequence of the primer starts with G or C residue (primer 5' end).
3- The melting temperature for designed primers were around 50-65 °C.
4- The length of the primer was within 18-30 bases, however in some exception cases, difficult DNA sequence areas for instance the primer length may be longer.
5- The primers were tested with Auto Dimer programme (http://www.cstl.nist.gov/biotech/strbase) to check for potential primer hairpin and primer-dimer secondary structures [177].

2.15.2 **Primers Guidelines for Bisulphite Treated DNA (Single Strand)**

Different programmes were used for oligonucleotide primer design. Primers for bisulphite treated DNA was designed with Methyl Primer Express Software v1.0 (Applied Biosystem, Foster City, California) or Invitrogen (Life Technologies, CA, USA) at a scale of synthesis of 50nM and were purified via desalting (Invitrogen by Life Technologies, CA, USA).

All primers were assessed for binding specificity during the design process through using the Primer BLAST function provided through National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/tools/primer-blast/).

2.16 **Agarose Gel Electrophoresis**

Agarose gel electrophoresis is easiest technique to separate nucleic acids and proteins [178]. Most biology research laboratories are depending on gel electrophoresis for DNA analysis. Electrophoresis was used to examine the presence of PCR product and determine size depending on the known size ladder.

Electrophoresis is a method of separating substances based on the rate of movement while under the influence of an electric field. Agarose is a polysaccharide purified from seaweed. An agarose gel is created by suspending dry agarose in a buffer solution, boiling until the solution becomes clear, and then pouring it into a casting tray and allowing it to cool. The result is a flexible gelatine-like slab. During electrophoresis, the gel is submersed in a chamber containing a buffer solution and a positive and negative electrode. The DNA to be analysed is forced through the pores of the gel by the electrical current. Under an electrical field, DNA will move to the positive electrode (red) and away from the negative electrode (black). Several factors influence how fast the DNA moves, including; the strength of the electrical field, the concentration of agarose in the gel and most importantly, the size of the DNA molecules. Smaller DNA molecules move through the agarose faster than larger molecules. DNA itself is not visible within an agarose gel. The DNA will be visualized by the use of a dye that binds to DNA.

The purpose of electrophoresis was to determine the presence or absence of PCR products and quantify the size (length of the DNA molecule) of the product.

Work solution of 1X TAE buffer were prepared in 1L, by adding 100 mL of 10X stock TAE buffer in to ~900 mL H_2O. The buffer was poured into 1 L graduated cylinder and a total volume of 1 L completed by adding H_2O. To visualize the DNA in agarose, 5 µL SYBR® Safe DNA Gel Stain Blue was added. According to the target DNA length a 100 bp molecular weight marker was included in order to be able to estimate the molecular weight of the amplicon. The agarose was melted in a microwave until the solution became clear. Then the solution was cooled with swirling the flask occasionally to cool evenly.

To load the gel, 10 µL of the DNA ladder mix was loaded in to the well of the first lane DNA ladder was added into. Next 8 µL of each amplified product was mixed with loading buffer and loaded into each of the wells of the remaining lanes of the gel. After running the gel for around 30 min, DNA fragments were visualised by placing it on UV transilluminator box (Thistle Scientific) and image was captured for the gel by attached camera or other documentation system.

2.17 **Ammunition and Case Loading**

Firstly, a volunteer (known age and sex to the researcher) was asked to wear gloves (non-powdered) for thirty min after hand washing. Then the volunteer was asked to handle ammunition directly and load ten bullets into a labelled and pre-cleaned magazine of one of the weapons. To make the results of the experiment more accurate, the volunteer was asked to load ten bullets repeatedly and with each gun separately. A negative control was prepared by treating the cartridge with UV light using a cross-linker and a mock sample for positive control was prepared by handling a cleaned cartridge with a bare hand after firing of eight tested weapons.

2.18 **Firing and Cartridge Collection**

The surfaces of the firearms were cleaned prior to firing. Clean paper was laid on the floor to reduce contamination by extraneous DNA. A firearms examiner wearing gloves fired the weapons without touching the cartridges. Later on all the guns were prepared for firing and standard firearm safety protocol was followed as below:

The firing process was done at the firearm department in Erbil Forensic Laboratory which is suitable for this type of experiment. The instruction manual of the tested guns was well known and fully understood by the expert shooter before handling the weapons. Each of the weapons was thoroughly checked to make sure it was safe at all time before the test fire was carried out. The air gun was shot into a special standardized tank which is filled with water and never pointed at people. During handling of the air gun, fingers were not placed on the trigger unnecessarily unless when it was loaded and ready for firing. Hands were not permitted over the muzzle of the air gun. The air gun was placed on the proper table to prevent it from falling down and discharging accidentally. The air gun was never cocked until it was ready to be fired.

The ejected cartridge (discharged cartridge) cases were collected and placed into a labelled plastic bag until ready for swabbing. After collection, each shot shell case was swabbed. After double swab technique the swabs were ready for extraction process [179].

2.19 Statistical Analysis

Microsoft Excel and Minitab® 16 software were used to conduct statistical analysis using one-way ANOVA test for significance between the results from different types of tested guns.

2.20 Sequence Analysis of Isolated DNA Fragments

Homology analyses of interest region of the DNA sequences were performed by Basic Local Alignment Search Tool (BLAST) which is available at NCBI Website (http://www.ncbi.nlm.nih.gov/BLAST/). NEWCPGREPORT program was used to report CpG island (CGI) located in sequence of each interested region available at EMBOSS Website (http://emboss.bioinformatics.nl/). PROSCAN-programs were used to predict promoter region in the studied fragment sequence available at (http://www-bimas.cit.nih.gov/molbio/proscan/) website of Advanced Biosciences Computing Centre.

To confirm the location of the target region for both nDNA and mtDNA, UCSC genome browser (http://genome-euro.ucsc.edu/cgi-bin/hgGateway) was utilized depending on the newest version of Genome Reference Consortium Human Reference 38 (GRCh38) assembly [180].

2.21 Bioinformatics

http://www.sabiosciences.com/

https://www.zymoresearch.com/

Blast: http://ncbi,nlm.nih.gov/BLAST/

BLAT: http://genome.ucsc.edu/cgi-bin/hgBlat

http://www.addgene.org/

ClustalX: http://www.clustal.org/.

Ensemble: http://www.ensembl.org/index.html

FinchTV : (http://www.geospiza.com).

Gen Bank:http://www.ncbi.nlm.nih.gov

GraphPad: http://www.graphpad.com/quickcalcs/contingency

Haploview: http://www.broad.mit.edu/mpg/haploview/

Image J: http://rsb.info.nih.gov/ij/

PubMed: http://www.ncbi.nlm.nih.gov/Literature/

Illumina: http://www.illumina.com/).

Plink: http://pnug.mgh.harvard.edu/~purcell/plink

OR calculator: http://www.hutchon.net/ConfidOR.htm.

SwissProt: http://www.expasy.ch/sprot/

UCSC: http://genome.ucsc.edu/

CHAPTER THREE

ForensiX Swab

3 Chapter 3 Advantage of ForensiX Swabs in Retrieving and Preserving Biological Fluids

3.1 Introduction

Human identity by DNA testing is now a highly developed technique. A limiting step in the process is the retrieval of the DNA from biological stains found at crime scenes. There is more than one method for sample collection, but a common approach to recovering dried stains is the double swab technique [179, 181]. In this technique, a wet swab is first used to rehydrate and lift cells while a second dry swab is applied to further recover any remaining cells [182]. For DNA collection purposes various swabs are commercially available with different characteristics and prices. A previous study compared the retrieval ability of the nylon with cotton swabs showed that the retrieval ability of these swabs are different depending on extraction method [183]. Recently a new type of swab packaging that possesses an active drying property has been produced. In these, the active drying is performed by a special internal lining to the swab case that actively absorbs moisture from immediately around the swab when it is returned to the swab case. The manufacturers claim that this feature improves the rate of drying compared to air drying or freezing which is passive drying. The rapid drying limits the degradation of DNA samples prior to their arrival at the laboratory. In addition, the drying occurs within the swab and independently of the outside conditions and this consequently minimises the opportunity for contamination.

This Chapter describes a study that was designed to compare the effectiveness of the ‘actively drying’ swabs with conventional, commercially available, sampling tools. Three types of swabs from two different manufacturers were tested in this experiment (delta lab and Prionics). The cotton swabs used for comparison were the EUROTUBE® Collection Swab from (delta lab, Spain), and the ForensiX Evidence Collection Tube (Prionics AG, Schlieren-Zurich, Switzerland), which contains an active desiccant system and either a cotton swab (forensiX Buccal Swab tube) or a nylon flocked swab (forensiX Nylon® Flocked Swab tube). The forensiX swabs consist of an active drying tube and a cotton swab. All of these were commercially

available at the time of testing. The swabs differ from each other primarily by the swab head, size and material type (Figure 3-1). This study consisted of two phases (120 samples each) with saliva used in both phases. The first phase tested the collection of a different volume of saliva (0.5 to 10 µL) in order to compare the efficiency of each type of swab where measured quantities of saliva, collected from the same person at the same time, were used. The purpose of this phase was to investigate if there was a difference in the quantity of biological material that each type of swab could retrieve.

However, this does not represent the circumstances typically encountered at a crime scene where the amount of saliva contributing to a stain is unknown and where there may not be the opportunity to treat the swab or to let it dry before returning it to its tube. Accordingly, the second phase of the study examined the extraction and subsequent storage of saliva for different periods of time, DNA extractions made after 0 h, 4 h, 1 day and 40 days of preservation. The aim of this, second, phase was to find out if the drying performance had a significant influence on the DNA yield for each of the swab types for different time periods. The saliva was taken from a single donor and at the same time to control the quality and homogeneity of the sample. The three different swab types were employed and the DNA recovery compared.

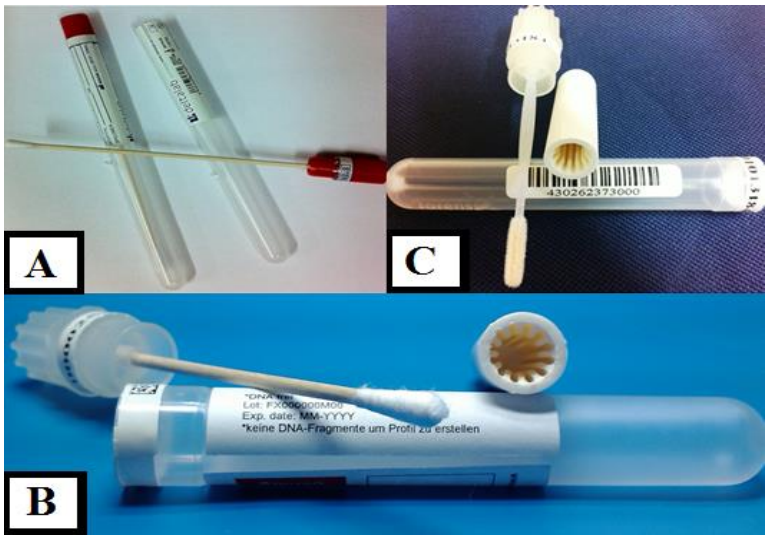

Figure 3-1 The three tested swab types used: A: E UROTUBE® Collection Swab (delta lab, Spain). B: forensiX Buccal Swab tube. C: forensiX Nylon® Flocked Swab tube (Prionics AG, Schlieren-Zurich, Switzerland). Figure of (Swab B) illustrated from Prionics.

3.2 Materials and Methods

3.2.1 Slide Preparation and Swabbing Technique

In the first phase, saliva collected in a clean tube from one volunteer on a single occasion (The participant instructed to spit into the tube (or helped by placing the tip of the tongue behind the front teeth) up to the red fill line marked on the tube). The saliva then mixed thoroughly for 30 s to form a homogenous solution with a brief spin down to remove debris. Different quantities of saliva (0.5, 2.5, 5 and 10 µL) were immediately pipetted onto sterilized, labelled, microscope slide to prepare four sets of glass slides (30 for each set) of the three types of swab to be tested. The slides were left for 30 m to dry at room temperature inside a sterile laminar flow hood. Lastly, the dried saliva on each slide was swabbed using double swab technique mimicking stain collection in real case scenarios at a crime scene, and extracted directly. A separate sterile clean glass slide used with each set as a negative control.

In second phase, saliva was aliquoted onto 120 sterilized, labelled, microscope slides. After 30 min of drying at room temperature, double swab technique was applied on all slides (each 40 slides were swabbed using one of the three tested swab types). The swabs were placed back into their packaging.

The swabs were classified into four groups, marked as 0 h, 4 h, 1 day and 40 days, so that in total each group contains 10 replicates for each swab type. The first group marked as 0 h was proceeded immediately to extraction. The remaining three sets of swabs were stored at room temperature and extracted when they reach the designated times in the second phase (4 h 1 day and 40 days).

Generally, swab heads were removed from the swab shaft while the cotton swabs of EUROTUBE were cut by cleaned scissors. Then swabs were placed into sterile (2 mL) micro centrifuge tubes ready for extraction.

3.2.2 **Extraction and Quantitation**

The total DNA was isolated using the QIAamp DNA Investigator Kit (QIAGEN, Crawley, UK). The elution value was set as 50 µL of ATE buffer (QIAGEN). After extraction, all the sets of swabs from both phases were quantified using the Qiagen Investigator Quantiplex kit with a Stratagene Mx3005P thermalcycler according to manufacturer's instructions. The quantities of DNA were compared to each other and also to the negative control extracted using the same extraction method. If any of the swabs did not work with this protocol, it was noted.

3.2.3 **DNA Amplification and Profiling**

Some samples from each set were amplified using a 2720 thermal cycler (Life Technologies, UK) with the Investigator Human Identification PCR Kit (QIAGEN). After running samples on the 3130 Genetic Analyser (Applied Biosystem, Life Technologies, Paisley, UK), the expected DNA profile was produced. There was no evidence of contamination.

3.3 **Statistical Analysis**

Microsoft Excel and Minitab® 16 software were used to conduct statistical analysis using one-way ANOVA test for significance between the results from the three different types of swabs.

3.4 **Results**

In Phase One, the results of DNA recovery among the three swab types (EUROTUBE collection swab (delta lab), buccal swab tube and nylon flocked swab tube (forensiX)) demonstrated that the nylon flocked swab tube was the most effective in sample recovery among the swabs. In terms of the saliva stains with 2.5, 5 and 10 µL volumes, the three tested swabs showed different DNA recovering abilities. Generally, forensiX Evidence Collection tubes showed higher results compared to the EUROTUBE collection swab, where the average recovered DNA with the nylon flocked swab was two and three fold higher than the buccal swab tube and EUROTUBE collection swab respectively (Figure 3-2).

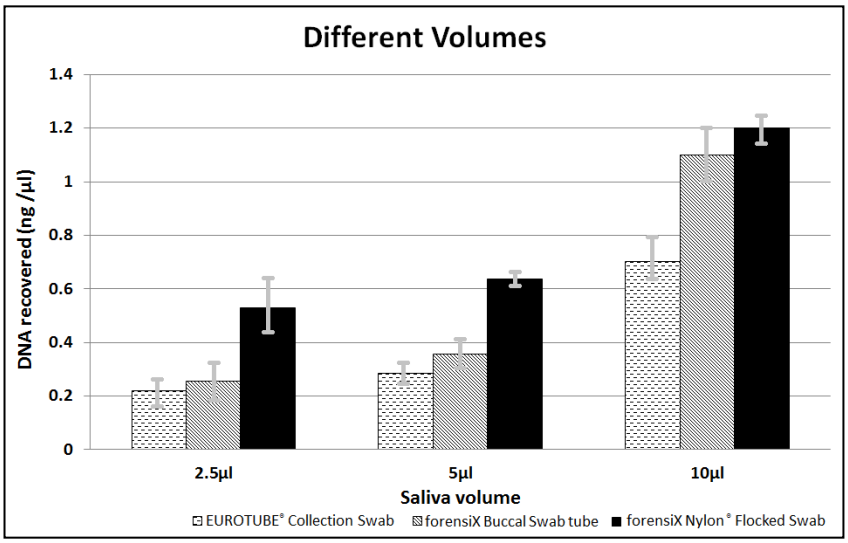

Figure 3-2 The retrieval ability of swab types (EUROTUBE collection swab, forensiX buccal swab tube and forensiX nylon flocked swab: The chart compares the yield of DNA from standard preparations of different volumes, 2.5, 5, and 10 µL, taken from the same person at the same time for the different types of swab. Data based on 10 replicates of each combination of conditions.

One of the interesting results was with smallest volume (0.5 µL) saliva stain, where 750 pg of DNA per microliter was recovered with the nylon flocked swab. While only one third of this result was retrieved with the EUROTUBE swab (Figure 3-3).

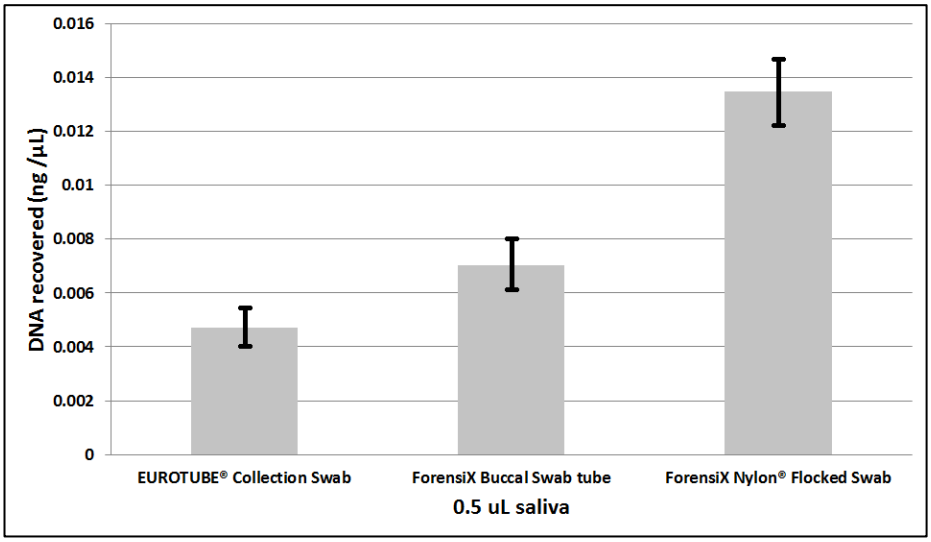

Figure 3-3. The retrieval ability: of 0.5 µL saliva using different swab (10 replicates each) types (EUROTUBE collection swab, forensiX buccal swab tube and forensiX nylon flocked swab).

The results of one-way ANONA showed that the DNA yield of the three types of swabs was significantly different (F (2,117) *p-value* <0.001). In addition, a *post hoc* pairwise comparison was performed in order to identify which DNA yield of the swab type differed from those of the others. The results show that there is a significant difference between the yield of the forensiX nylon flocked tube and those of either forensiX buccal swab tube and EUROTUBE collection swab (F(1,78) *p-value* <0.001 and F (1,78) *p-value* <0.005 respectively). In addition, the yield of forensiX buccal swab tube does not significantly differ from the EUROTUBE collection swab (F (1, 78) *p-value* >0.001).

Figure 3-4 shows the results of Phase Two concerning preservation for a different period of times (0 h, 4 h, 1 day and 40 days) among the swabs. As appear in the figure, the first set of swabs which were extracted immediately (samples 1-30); the amount of recovered saliva was nearly the same among the swab types.

The difference in the DNA yield becomes more prominent among swab types with increase the time of preservation. The amount after 4 h (samples 31-60) and 1 day (samples 61-90) was around 0.30 ng/ µL and 0.35 ng/ µL for the EUROTUBE collection swab and buccal swab tube respectively. For the nylon flocked swab tubes

the average amount of recovered DNA was more than 0.6 ng/ µL for the samples stored for 4 h and one day.

In addition, after 40 days storage in room temperature (samples 91-120), both forensiX swabs (nylon Flocked and buccal swab tube) recovered a higher DNA. This amount was more than four-fold that recovered by the EUROTUBE collection swab. Whilst with preserved saliva stain with EUROTUBE collection swab after the same period (40 days) few picogram of DNA was recovered and in some samples no DNA was recovered.

Comparison of the results under different storage conditions using one-way ANOVA revealed that there is a highly significant difference among the swabs ($P< 0.0001$). The pairwise comparison study shows that the DNA yield of the three types of swabs significantly differ from each other.

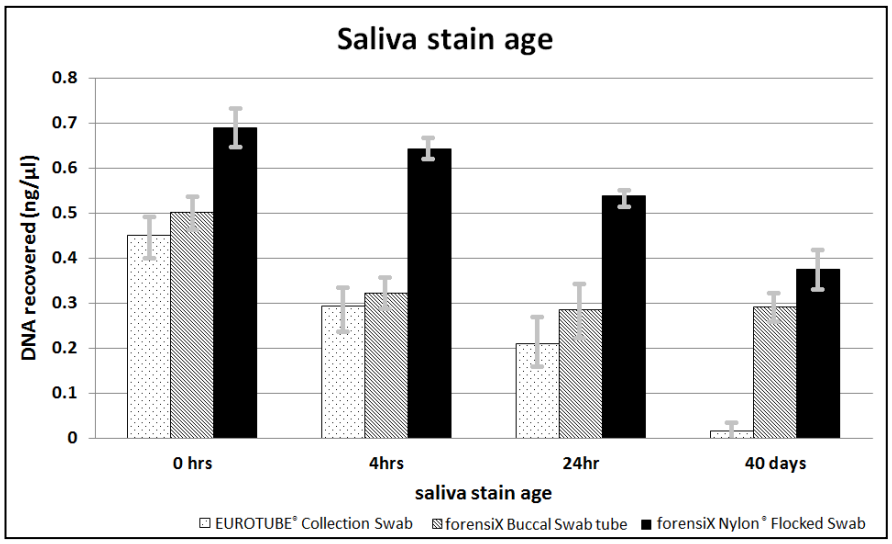

Figure 3-4 DNA preservation of different swabs: Bar chart of retrieval ability of collected saliva stains preserved for different period of times (0 h, 4 h, 1 day and 40 days) with 10 replicates at each period. DNA profiling for a subset of these samples at 40 days indicated that there was limited DNA degradation in the samples.

3.5 **DNA from Cartridge**

To progress the ultimate purpose of the work, another experiment was designed to compare the ability of the evidence collection swabs. After extraction and

quantitation forensiX Nylon flocked swab tube showed again higher result and better retrieval of DNA (Figure 3-5).

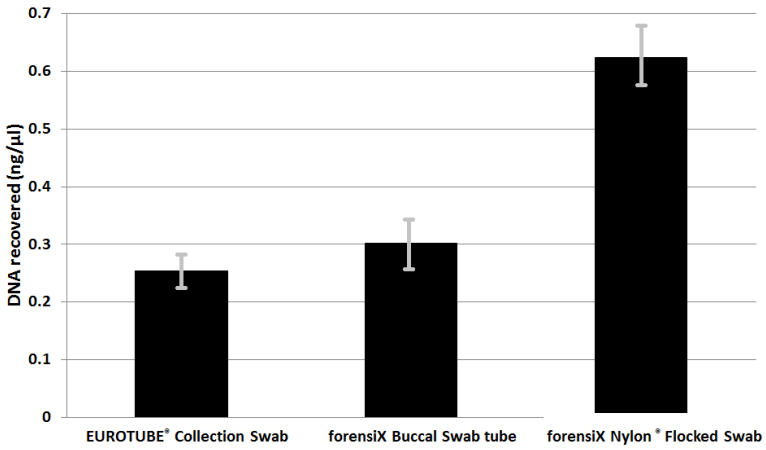

Figure 3-5: Swabs ability to retrieve DNA from Cartridge: Using different swab types (the EUROTUBE collection swab, the forensiX buccal swab tube, and the forensiX nylon-flocked swab).

This result confirms that forensiX Nylon Flock swab tube is one of the best choices to collect biological fluid. Thus we continued to test this swab on fired cartridges in the next step of the experiment.

3.6 **Discussion**

This study was designed to compare the forensiX evidence collection swab tubes with EUROTUBE® Collection swabs simulating the collection of saliva stains encountered at a crime scene. Obtaining the DNA profiles of possible contributors a process involving several steps and the collection of the samples with the swab is the first step. The samples may be stored for days, weeks or even more in the laboratory before the DNA extraction proceeds, and some laboratories do not have facilities for storing large numbers of evidence samples frozen. Hence, biological degradation will occur and this is most problematic for low amount DNA samples trace DNA samples for instance.

In the first phase of the study, using different volumes of saliva, generally, it was found that forensiX nylon flocked swab tubes retrieve a larger DNA quantity than other cotton collection swabs. This was particularly apparent with the smallest tested

volume, 0.5 µL saliva. On the contrary, collection with EUROTUBE swabs recovered only a few picograms of DNA. These were close to results of the negative controls in some samples.

As there are some factors affecting the efficiency of DNA retrieval among the swabs, these factors may result in this difference in DNA yield. One factor that should be taken into consideration is the ability of saliva recovery among three tested swabs. Another significant factor is the efficiently of the swab to release the evidence which has been collected. The swab may be effective at saliva collection, but not efficiently at release during extraction process then the DNA yield is reduced.

The manufacturer claims that the forensiX nylon flocked swabs improve the yield in two ways because they are made of nylon instead of cotton. Firstly, sample absorption into the swab is improved because the sample is more quickly absorbed by capillary action. Secondly, sample retention is reduced as the material is stored in close proximity to the surface of the swab rather than being internalised, thus improving the release of sample material [182, 184]. This hypothesis may explain the differences between forensiX products itself (nylon flocked and buccal swab tubes) in one side and with EUROTUBE collection swab in the other side. In addition, the shaft part of the nylon flocked swab is made of nylon. This could be another reason for the difference in the efficiency ability. The shaft part of the other two swab types is made of wood, which may absorb some moisture and biological material [185].

In the second phase, the results showed the advantage of using forensiX swab tube rather than EUROTUBE collection swabs to preserve samples for long periods of time because the sample is dried faster and, hence, with a reduced opportunity for degradation to occur during the drying process while the sample is still wet. Interestingly, after a significant period of time (up to 40 days) it was still possible to retrieve nearly the same amount of DNA which was retrieved with direct extraction (0 h).

The forensiX collection tool consists of two parts, swab and plastic tube. The inner surface area of the plastic tube contains “SafeDry” desiccant which composed of a molecular sieve drying agent in a polymer matrix. This system allows drying the

sample swab actively by absorbing the moisture and the humidity from the air space around the sample. Therefore, this rapid active drying method improves the retrieval result because it decreases the chance for the bacteria and fungi existing in collected saliva to degrade the DNA. Furthermore, the tube isolates the saliva on the swab head from exposure to the external environment influences. Recently Garvin *et al*., [185] demonstrated that the active drying method is better than other methods which some laboratories follow to store collected biological evidences e.g. passive drying, cooling and freezing [185]. On the other hand, the amount of DNA retrieved with the EUROTUBE collection swab decreased dramatically with increasing the time of storage. This appears clearly in samples stored for 40 days where some samples showed a very low and non-quantifiable recovery of DNA.

The improved drying performance of the forensiX nylon flocked swab tubes may be due not only to the active drying system but also to the nature of the swab head, consisting of a brush like head, from which a solution containing the biological material from a stain might be more readily eluted. It is possible that it is a combination of these factors that contribute to the improved performance of these swabs. Another advantage is the design of the swab shaft of the forensiX where the head was readily cut off for extraction. However, the forensiX swab tubes cost more than commercially collection swabs, but the advantage of performance ability of the forensiX swab tubes will counterbalance the prices by the improved evidence collection capability.

3.7 **Conclusion**

The present study has demonstrated improved yields and the preservation of the DNA by forensiX swab tools compared with EUROTUBE collection swabs. This finding exhibits the importance of using such a tool particularly with small saliva stains that are stored for long periods of time. This is a common circumstance in crime scene collections. Additional investigation should be conducted considering different drying systems, concentrations, longer periods of storage, the use the swabs on other substrates and comparisons with more different types of swab eg. Copan 4N6FLOQSwabs™ (Thermo Fisher Scientific).

CHAPTER FOUR

nDNA Methylation

4 Chapter 4 Tight correlation between methylation status and human age at three autosomal loci: a new forensic profiling tool

4.1 Introduction

DNA profiling with numerous highly polymorphic autosomal short tandem repeat (STR) markers has been used in many different aspects of human identification in forensic investigation over the past 20 years [41]. The DNA that is left behind by a criminal at the scene of a crime is useful only when it matches a DNA database profile. Scientists are now investigating alternative, database-independent, information that can be derived from genotype information in order to generate a description of an alleged criminal. This might predict the gender, physical features of an individual (eye or hair colour) or suggest the ethnic origin (using ancestry informative markers: AIMs) thus accelerating and targeting the investigative process.

The observation that in all cells the DNA sequence is identical but there are quantifiable epigenetic differences between tissues and across life experiences, suggests another direction for forensic profiling. When cytosine occurs next to guanine in the context of the CpG dinucleotide (5-CpG-3) it has the capacity to be methylated on the cytosine to form 5 methyl-cytosine (5mC) [47]. In human DNA 3-6% of all cytosines are thought to be methylated [48, 49]. The forward reaction of methylation is mediated by the DNA methyltransferase enzymes using S-adenosylmethionine as a donor of a methyl group. Demethylation is the backward reaction of methylation which is performed by DNA demethylase enzymes[50]. In addition to 5-mc, 5-hydroxymethyl cytosine (5-hmC) is another cytosine-derived base modification detected in DNA. It is a further derivation from (5mC) catalysed by ten-eleven translocation (TET) enzymes.

Studies have shown that different methylation patterns exist across a set of CpG sites in genomic DNA isolated from different tissues (blood, saliva, semen and epithelial tissue) [54], some of which have been termed tissue-specific-differentially-methylated regions (tDMRs)**.** Furthermore, the DNA methylation can be used to distinguish different tissue types from the same person [56]. This proves

experimentally that DNA methylation is a valuable indicator to distinguish body fluids that could contribute to forensic analysis [56, 57]. In other forensic applications, the analysis of DNA methylation patterns may assist with determining the sex of the sample [63], the cause and circumstances of death, the parental origin of alleles, authentication of DNA samples, discrimination of monozygotic twins, and the age of the individual involved [59-62].

Human ageing is very complex, multifactorial process and affects all tissues of our body. Various age-indicative markers have been described such as morphological changes of skeletons, teeth and molecular or biochemical changes [68, 69]. Increasing age correlates with accumulated genetic changes such as mitochondrial DNA deletions, T-cell receptor deletion and a reduction in telomere length associated with successive cell division as well as biochemical changes such as the racemisation of aspartic acid and glycation end products. However, all of these approaches suffer from high levels of variation and there are many examples of confounding variables, such as certain disease states, sex differences, and the origin of the population under study.

The DNA methylation status in a sample has been shown to be affected by the age of the donor. Generally the methylation component will decrease through a life time; so that a centenarian has a lower overall DNA methylation component compared to DNA of a newborn [70].

Recently Bocklandt and co-workers studied 34 twins aged 21 to 55 years, using DNA from saliva samples [73]. Illumina human methylation27 microarrays were used and, although they were unable to validate the study in different set of samples, a regression model was built using three CpG sites in three genes, *NPTX2*, *EDARADD* and *TOM1L1* to determine the age of the donor to a mean accuracy of 5.2 years [73].

Koch and Wagner used publically available data (from Human Methylation27 Bead Chip analysis) of thirteen tissue types. They found 431 hypermethylated and 25 hypomethylated CpG sites that correlated with age [74]. From these, CpG sites of

five genes (*TRIM58, KCNQ1DN, NPTX2, BIRC4BP* and *GRIA2*) were able to generate an "Epigenetic-Age" signature with a precision of 11 years.

Zbieć-Piekarska *et. al*. 2015 claimed that a prediction of biological age through methylation of *ELOVL2* gene marker can be used in forensic science with only 7 years error between chronological and predicted age. Blood samples of 303 individuals aged 2-75 were used in the study [75].

Here, we describe how DNA methylation profiles from blood sample of 82 women with ages ranging from 18-91 years were generated using the Illumina MiSeq 300 v2 system. This cohort was used to interrogate methylation at multiple loci as an epigenetic marker for the prediction of age for forensic purposes DNA

The main objectives of this study were to:

1. Assess correlations between the chronological age of a person and the methylation state at a number of selected autosomal loci and published studies within the human genome.
2. Determine the optimal experimental techniques for extending the sensitivity of the approach so that samples of limited DNA concentration may be tested.

4.2 **Principle of Sodium Bisulphite Conversion**

Over the years, DNA treatment with bisulphite sodium has become the most widely used technique to study methylation pattern[186]. It is the most appropriate and efficient way analyse bisulphite converted DNA to individual bases. The process of bisulphite conversion is well understood, so DNA treatment with this chemical compound is the first step in numerous downstream analyses and in the different types of applications.

For pre-treatment bisulphite conversion step, EZ DNA Methylation-Direct^{TM} Kit (Cat # 5020 and 5021) has been chosen, because the manufacturer (Zymo Research) claimed that this new product has higher efficiency than other protocols (99.5%), of non- methylated C residues are converted to U, and almost 99.5% to protect methylated cytosine [187]. Genomic DNA was bisulphite converted according to the

manufacturer's protocol[187]. Figure 4-1 shows the principle of bisulphite sodium modification (details in 2.15).

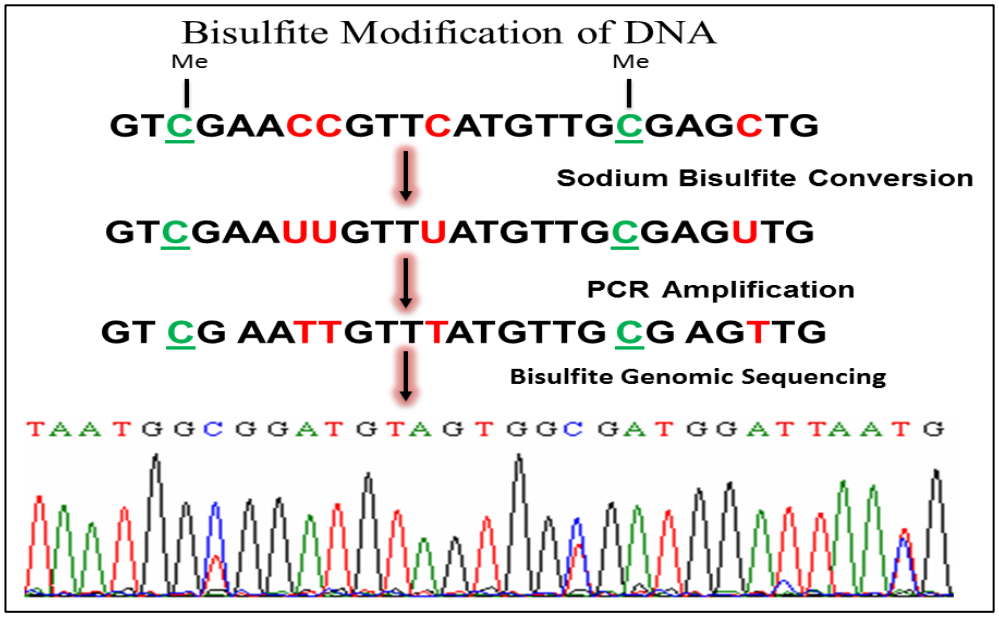

Figure 4-1 DNA sequencing results following bisulphite treatment: Upon bisulphite sodium conversion with EZ DNA Methylation-DirectTM Kit methylated ^{m}CpG (Green and underlined) at nucleotide position#3 and 19 remained intact while non-methylated cytosine (red) at position # 7,8,11 and 22 were converted into uracil (U)and detected as thymine (T) following PCR amplification. Illustration reproduced from www.zymoresearch.com.

4.3 Materials and Methods

This research project was approved by the University of Strathclyde Ethics Committee and prior to sample donation, participants signed informed consent statement. The protocol for swabbing and DNA measurement by qPCR has been previously described [188]. For each reaction conditions a precise DNA concentration was used (120ng) for the bisulphite conversion step, EZ DNA Methylation-DirectTM Kit (Cat ≠ 5020 and 5021, Zymo Research) was applied according to the manufacturers protocol[187].

4.3.1 Regions of Interest (ROI) in Target Genes and their Sequence

The assembly of genome reference consortium human Build38 (GRCH38 or hg38) was used. The GRCh38 assembly offers an "analysis set" that was created to

accommodate next generation sequencing read alignment pipelines. Twenty-three promoters were assessed requiring 50 amplicons which were chosen based on genes identified as showing age-dependent methylation changes in previous publications (Table 4-1).

Table 4-1 Targeted genes: The table shows details and information on the selected genes with their official name, estimated amplicon of each promoter region and the published references linking each gene's methylation level to ageing.

No.	Target Gene (ROI)	NCBI-Gene ID	Official gene name	Location (GRCH38/hg38)Assembly	Estimated amplicons #	Reference
1	*ZC3H12D*	340152	zinc finger CCCH-type containing 12D	chr6:149491968-149492901	5	[54]
2	*SOGA1*	140710	suppressor of glucose, autophagy associated 1	chr20:36863587-36864174	0	[54]
3	*BCAS4*	55653	breast carcinoma amplified sequence 4	chr20:50794865-50795605	1	[189]
4	*BCAS4-pro2*	55653	breast carcinoma amplified sequence 4	chr20:50794865-50795605	2	This study
5	*FGF7*	2252	fibroblast growth factor 7	chr15:49423893-49424213	1	[190]
6	*AHRR*	57491	aryl-hydrocarbon receptor repressor	chr5:373063-373462	1	[191]
7	*NPTX2-pro-1*	4885	neuronal pentraxin II	chr7:98616379-98616828	1	[73]
8	*NPTX2-pro- 2*	4885	neuronal pentraxin II	chr7:98616285-98617284	3	This study
9	*EDARADD-pro1*	128178	EDAR-associated death domain	chr1:236394183-236394603	1	[73]
10	*EDARADD-pro2*	128178	EDAR-associated death domain	chr1:236395160-236396036	2	This study
11	*TOM1L1-pro1*	10040	target of myb1 (chicken)-like 1	chr17:54901012-54901432	1	[192]
12	*TOM1L1-pro2*	10040	target of myb1 (chicken)-like 1	chr17:54900506-54900946	2	This study
13	*GRIA2-pro1*	2891	glutamate receptor, ionotropic, AMPA 2	chr4:157220152-157220947	1	[74, 193, 194]
14	*GRIA2-Pro2*	2891	glutamate receptor, ionotropic, AMPA 2	chr4:157220253-157220684	3	This study

15	*TRIM58-pro1*	25893	tripartite motif containing 58	chr1:247857199-247857648	0	[74, 195]
16	*TRIM58-pro2*	25893	tripartite motif containing 58	chr1:247857029-247857950	3	This study
17	*KCNQ1DN*	55539	KCNQ1 downstream neighbor	chr11:2869033-2870107	4	[74, 196, 197]
18	*ITGA2B-pro1*	3674	integrin, alpha 2b (platelet glycoprotein IIb	chr17:44398822-44399786	4	[193]
19	*ITGA2B-pro2*	3674	integrin, alpha 2b (platelet glycoprotein IIb	chr17:44390160-44390559	1	This study
20	*ASPA-pro1*	443	Aspartoacylase	chr17:3374906-3375337	0	[193, 198, 199]
21	*ASPA-pro2*	443	Aspartoacylase	chr17:3476073-3476472	1	This study
22	*PDE4C-pro1*	5143	phosphodiesterase 4C, cAMP-specific	chr19:18224162-18226665	13	[193, 200]
23	*PDE4C-pro2*	5143	phosphodiesterase 4C, cAMP-specific	chr19:18232892-18233291	1	This study
					Total:50	

4.3.2 Assay Design, Sample Preparation, and Multiplex Targeted Amplification

Assays were designed targeting CpG sites in the specified ROI using primers created with Rosefinch, Zymo Research's proprietary sodium bisulphite converted DNA-specific primer design tool. Within this study, parameters were chosen such that PCR amplicons would be ideally bigger than 100 bp but smaller than 300 bp, cited at (http://www.zymoresearch.com/tools/bisulphite-primer-seeker). In addition, these primers were designed such that they would avoid annealing to CpG sites at the region of interest to the maximum extent possible. In the event that CpG sites within primer-binding sites were absolutely necessary for target amplification, additional primers were synthesized with a pyrimidine (C or T) at the CpG cytosine in the forward primer, or a purine (A or G) in the reverse primer to minimise amplification bias to either a methylated or un-methylated allele. Two pairs of primers were picked to amplify each target region: one pair to sequence DNA before bisulphite sodium treatment Table 4-2 the other to amplify chemically converted DNA. Twenty-eight of the 50 amplicons passed the quality control criteria after DNA conversion.

The sequences of the primers (two pairs) and the estimated amplicon (before and after bisulphite treatment) are shown in (**Appendix 1)**. All primers were re-suspended or ordered in TE solution at 100 μM. Primers were then mixed and diluted to 2 μM. All primers were then tested using Real-Time PCR (Stratagene Mx3005P) with 1 ng of bisulphite- converted control DNA, in duplicate individual reactions. DNA melt analysis was performed to confirm the presence of a specific PCR product. The following guidelines were used to assess performance: sited at (http://www.zymoresearch.com/tools/qmethyl-calculator/single-sample)

- Had average Cytosine-phosphate (Cp) values <40.
- Duplicate Cps do not have a Cp difference >1 (within 5% CV).
- Reached the plateau phase before the run ended at cycle 45.
- Produced melting curves in the expected range for PCR products.
- Duplicate melts had calculated temperatures (Tms) within 10% Ct Value (CV).

Table 4-2 Primers used to amplify target gene amplicons: The table shows forward and reverse primers that amplify the DNA (bisulphite treated) at interest regions of the genes.

Target	Forward primer Sequence	Reverse_Primer_Sequence
ZC3H12D	TTGGGTATYGATTTGTGAATTGAGATTTT	TTCRAATACAAAAATAAAAATATAAAACAAAAA
ZC3H12D_2	TTTTTYGAAAAGTAAAGTTTTGTTTTT	TAACRCAATTAATTAACTACAATAATACC
ZC3H12D_3	GTTTTATYGGTATTATTGTAGTTAATTAATTG	ATACCACTACACTCCAACCTAA
ZC3H12D_4	GTTTAGGTTGGAGTGTAGTGGT	ACRCTTATAATCCCAACACTTTAAAA
ZC3H12D_5	TTTYGAGTAGTTGGGATTATAGG	CACCTATAATCCCAACACTTTAAAA
BCAS4	GGAAATTTTAAAAGAGTTTTAGGTGAT	ACTAACRAAAACTACCTAAAAACC
BCAS4-2	GGATTTYGGGATTTTTGTTGTT	TAAATCTACTTTTCACTTAATACTTTTAAAATAAAA
AHRR	GTGTTGGTAGGATATAGGGGTTGTTTAGG	ACAAAAACCAACCTATCCCCTACCT
NPTX2	TGTTATTTYGTTTTTGAGTTTAGGTTTTGTTGAATAAG	AAAACCCRACCACCTCACTCTC
NPTX2-2_2	GGGGTTGTTTTTGTTGGAGAAAA	AAATCCCRAACTCTACCTCAAAAAAAAAA
EDARADD	AGAGTTTTTATTTAGAAGGTTTGATTTTGGT	TAAACAACCTCTAACTAAAAACTCAACTCTA
EDARADD_promoter-2	AGTTTTGGTGTTTTAGTTGAGG	CTACTTTAACAATTAACAAAACCCC
TOM1L1	GGTTAATTTATTGTAGAATTTTAGTTTTTAAAAA	TCACAATAAAACCACATCACTTAAT
TOM1L1-2	GGYGAAAAGTTTATTTGGGATT	CTTCCCRAACTCTAACAAAAAC
TOM1L1-2_2	TTATGGYGTTTGGTAAGAGTTAT	CACRATCCCCTACCTACCTAAA
GRIA2	TTTTGTGATTTGTTTGTGTGTGTG	AAAAAATCCCTATTTCCCAAATCCTACTAATAAC
GRIA2_Promoter-1	GATAGGGTTTTTGTATTGTTAAATGGGTT	TACTCCAACCTAAACCAAATACAACAA
GRIA2_Promoter-1_2	TTAGTTTTGTTGTATTTGGTTTAGGTTGGA	ACACACACAAACAAATCACAAAAAAAAAC
TRIM58	TGGTTYGTTTGAGTTTTTGAAT	AAAAATACACCTAAAACAAAAACTATAAC
TRIM58_2	AGTTGGYGGGTTTGGTGGAGAG	AAAAACRCCTCACCTAATAACTACC
TRIM58_3	TTTTGTTTTGGTTGTGGTAATTTT	ATTATTACCAAAAATAAACTAAACCACA
KCNQ1DN	GTAGTTTGGGGATTTTAGAGGATT	AAAATAACCAACTAATTAACAACCATATATAAA
KCNQ1DN_2	GGTTTATATATGGTTGTTAATTAGTTGGTTATTTT	ACACACAAAAACCCATTCTTCCTA
KCNQ1DN_3	GGAAGAATGGGTTTTTGTGTGT	TACATCRATCCTACTACCCTCC
KCNQ1DN_4	TTTGTTYGTTTGAAAGGTTTATAAAT	ACRATAATACTAAAACCTTAACCAC
ITGA2B	TTATTAAGATTTGATTTTGGTTGGGGGTT	TCCTTAATATATATTTCCATCCAATCTTTCAACAA
ASPA	TATTTTTGGAGGAATTTATGGGAATGAGTT	CAATACAAAACATAATCTTACCCAAAATTTTCAAA
PDE4C	GYGTTGTAGGAGGTTTTTGTTT	ACCRCCACAACATAACCAAAAC

4.3.3 **Next Generation Sequencing (NGS)**

Illumina MiSEQ 300 V2 sequencing was used in this study to enable a high throughput DNA assessment of multiple loci. Library preparation and sequencing were carried by Zymo Research Corporation. Multiplex amplification of all samples using region-of-interest (ROI) specific primer pairs and the Fluidigm Access ArrayTM System was performed according the to the manufacturer's instructions. The resulting amplicons were pooled and barcoded according to the Fluidigm guidelines. After barcoding, samples were purified (ZR-96 DNA Clean & Concentrator™ ZR, Cat#D4023) and then prepared for massively parallel sequencing using a MiSeq V2 300bp Reagent Kit (cat. # MS-102-2001) and paired-end sequencing protocol according to the manufacturer's guidelines. The consensus sequences of the Fluidigm adaptors were.

CS1= 5'-ACACTGACGACATGGTTCTACA-3'.

CS2 = 5'-TACGGTAGCAGAGACTTGGTCT-3'.

The CS1 and CS2 sequences are included in each of the target or locus-specific primers. The Illumina barcodes were then added in the subsequent barcoding step.

4.3.4 **Sequence Analysis of Isolated DNA Fragments**

Homology analyses of interest region of the DNA sequences were performed by Basic Local Alignment Search Tool (BLAST) which is available at NCBI website (http://www.ncbi.nlm.nih.gov/BLAST/). NEWCPGREPORT program was used to report CpG island (CGI) located in sequence of each interested region available at EMBOSS Website (http://emboss.bioinformatics.nl/). PROSCAN-program was used to predict promoter region in the studied fragment sequence available at (http://www-bimas.cit.nih.gov/molbio/proscan/) website of Advanced Biosciences Computing Centre.

To confirm the location of the target region for nDNA a, UCSC genome browser (http://genome-euro.ucsc.edu/cgi-bin/hgGateway) was utilized depending on the

newest version of Genome Reference Consortium Human Reference 38 (GRCh38) assembly [180].

4.3.5 **Sequence Alignments and Data Analysis**

Sequence reads were identified using standard Illumina base-calling software and then analysed using a Zymo Research proprietary analysis pipeline, which is written in Python. Low quality nucleotides and adapter sequences were trimmed off during analysis of quality control.

Sequence reads were aligned back to the reference genome using Bismark (http://www.bioinformatics.babraham.ac.uk/projects/bismark/), an aligner optimised for bisulphite sequence data and methylation analysis [201]. Paired-end alignment was used as default thus requiring both read were aligned within a certain distance; otherwise both reads were discarded. Index files were constructed using the Bismark genome preparation command and the entire reference genome. The non-directional parameter was applied while running Bismark. All other parameters were set to default. Nucleotides in primers were trimmed off from amplicons during methylation calling. The methylation level of each sampled cytosine was estimated as the number of reads reporting a C, divided by the total number of reads reporting a C or T.

4.4 **Results and Discussion**

Methylation levels were assessed at 50 target regions within 23 ageing-related genes, containing thousands of CpG dinucleotide. Some regions were not amenable to bisulphite primer design due to the fact that they exist in a region of exceptionally high-CpG density, overlap with repetitive elements, or both. In total, 28 amplicon of a total 50 passed the quality control (QC) criteria (**Appendix 2**) of the assay system (Table 4-3).

Table 4-3 The table lists the Regions of Interest (ROI): shows the number of amplicons that passed the quality control (QC) criteria of the Illumina system**.**

Region of Interest (ROI) ID	Number of Amplicons that passed QC
ZC3H12D	5
BCAS4	2
AHRR	1
NPTX2	2
EDARADD	1
EDARADD_promoter-2	1
TOM1L1	1
TOM1L1-2	2
GRIA2	1
GRIA2_Promoter-1	2
TRIM58	3
KCNQ1DN	4
ITGA2B	1
ASPA	1
PDE4C	1
Total	**28**

After bisulphite conversion, Illumina sequencing detected methylation levels at 11 of the 28 amplicons, with 100-300bp length (while the other 17 amplicons either failed to amplify or be sequenced and/or did not show methylation change), with a total read of 397 CpG sites; these data are shown in (**Appendix 3**). An Excel file contains the read data and CpG methylation calling for all sites across all samples for the project. The methylation ratio (meth ratio) is calculated by using methylated CpG count/total CpG count. Furthermore, the coordinates for each of the detected CpG

sites across all of the samples can be seen in the Excel table. If methylation at one CpG site was detected in at least one sample with at least 10 reads, the CpG site has listed in the excel table. For samples where the CpG was not detected (<10 reads), the author left the meth ratio and total CpG count column as blank.

Sequencing data were installed as tracks on the UCSC genome browser for analysis. Evidence for a linear correlation between methylation levels and age was tested by Pearson correlation as detailed in **Table 4-4**.

Table 4-4: The methylation levels: The table shows significant correlation with age at three gene (grey) loci (*ITGA2*, *PDE4C*, and *ASPA*) as assessed by Pearson correlation.

NO.	ROI	Pearson correlation	P-Value
1-	*ITGA2B*	-0.126	0.471
2-	*ITGA2B*	-0.412	**0.012**
3-	*PDE4C*	0.524	**0.001**
4-	*ASPA*	-0.705	**0.00001**
5-	*ASPA-1*	-0.624	0.00002
6-	*ASPA-2*	-0.598	0.00001
7-	*TOM1L1*	-0.123	0.445
8-	*TOM1L1-2*	-0.164	0.304
9-	*GRIA2-1*	0.356	0.148
10-	*TRIM58*	-0.099	0.537
11-	*ZC3H12D*	-0.026	0.870

Methylation profiles of eight of the above 11 age related CpG (AR-CpG) sites showed weak association with ageing, whereas a subset of three CpGs (*ITGA2, ASPA* and *PDE4C*), we have aligned the sequence data using the most up-to-date and advanced human genome assembly - hg38, located at (Ch17: 44,390,357, Ch17: 3,476,272 and Ch19:18,233,090 respectively. These three loci produced better correlations than other sites considered in this study and age estimates were more

accurate when these three sites alone were included, than when all 11 sites were included. Generally, the human genome is hypomethylated during ageing [192], with one study showing newborn DNA having 494,595 more methylated CpG (mCpGs) dinucleotides than centenarian DNA (16,775,090 *vs.* 16,280,495 sites on the Watson and Crick strand, i.e. around 76% *vs.* 75% respectively) [70]. By contrast, promoter CpG island methylation levels tend to increase with ageing. The three regions showing greatest correlation with age displayed both trends: *ASPA* and *ITGA2B* were hypomethylated, while the target region in *PDE4C* was hypermethylated with ageing (figure 4-2) a profile also reported in other studies searching for ageing markers. These results are consistent with those of other recent studies and suggest that age-related CpG sites (AR-CpGs) are either hypomethylated or hypermethylated [196, 198, 199].

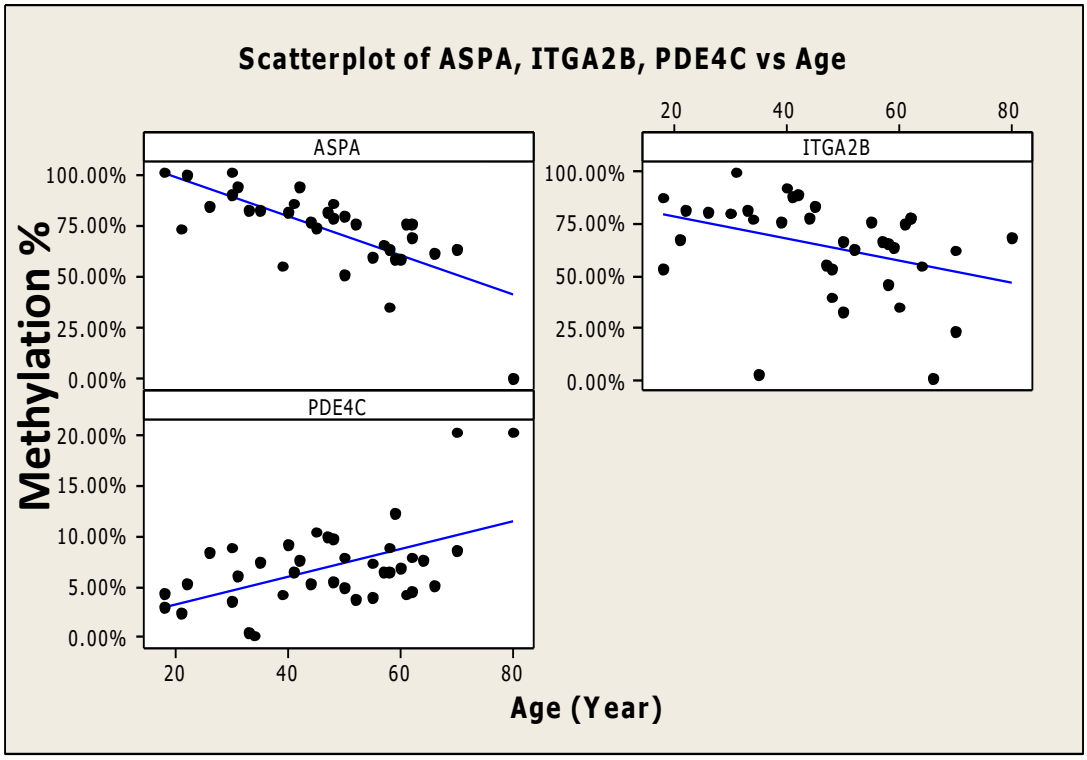

Figure 4-2 A scatterplot of correlation between methylation level at the *ASPA*, *ITGA2B* and *PDE4C*genes and age. Methylation level was decreased (hypomethylated) in both *ASPA* and I*TGA2B* genes, while in *PDEC4*C the level was increased (hypermethylated) with ageing.

Based on the CpG sites of three genes (*ITGA2B, PDE4C* and *ASPA*) a multivariate linear equation was built from a training set (first group) of 41 samples by Pearson

correlation to predict age of the DNA donors. The resulting equation (*Age= 84.93 - 51.92 ASPA+65.2 - 27.4 ITGA2B+34.37 + 189.5 PDE4C*) correlated well with chronological age. The mean absolute deviation (MAD) between real age and estimated age was only 5.3 years. The multivariate linear equation model was validated in a second group of DNA samples (also 41 samples). There was a clear correlation between predicted and real ages (Pearson correlation 0.711 and p-value = 0.0000004); Figure 4-3 , and the MAD was 6.0 years.

Remarkably, from our large set of target genes selected from numerous publications, we have identified the same three informative CpG sites that were also described by Weidner and co-workers who analysed the association by a different methodology - pyrosequencing [193]. A free online calculator has been designed by the Weidner group calculates and predicts the donor ages once you enter the value of methylation percentage of these three epigenetic ageing signatures: http://www.molcell.rwth-aachen.de/epigenetic-ageing-signature. When we implemented the value of methylation levels from our data into this freely online calculator, there were close similarities between the results of this calculator our equation model (*Age= 84.93 - 51.92 ASPA+65.2 - 27.4 ITGA2B+34.37 + 189.5 PDE4C*).

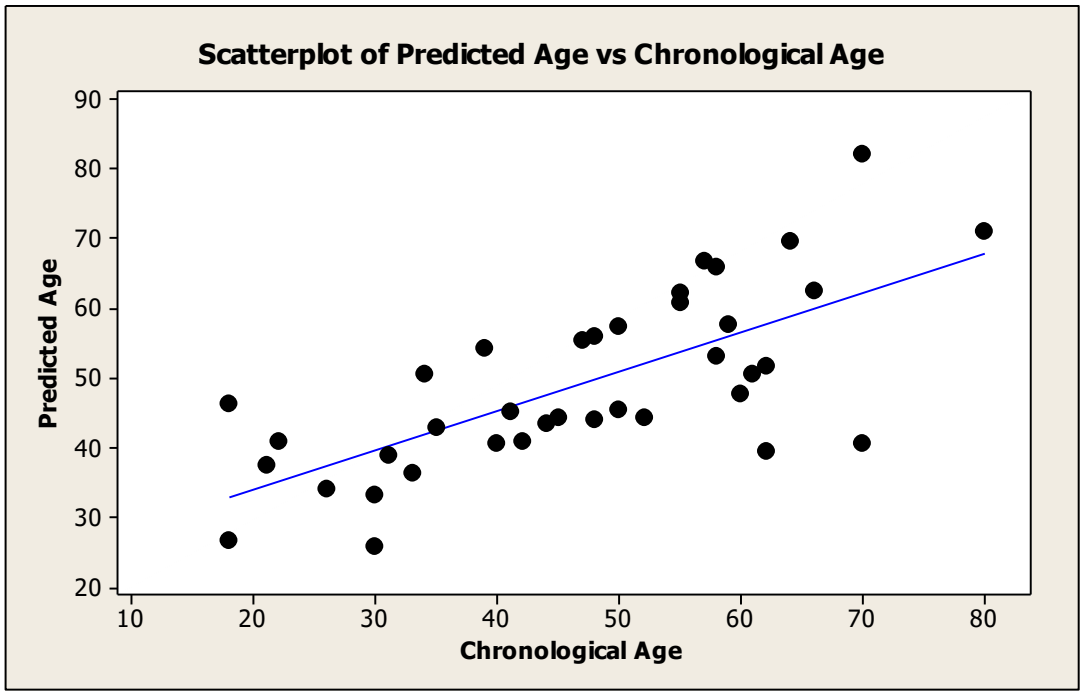

Figure 4-3 A scatter plot of Predicted Ages versus Chronological Age: Clear correlation between predicted age derived from methylation levels at 3 CpG sites and actual age of 41 DNA donors with regression line.

Thus, it is possible to examine the DNA methylation status in a sample as it has been shown to be affected by the age of the donor. Our result is similar but not identical to their study (5 years) and this may be because the two sets of samples were analysed by different assays and obtained from different population. Furthermore, the epigenetic age predictions might also be influenced by differences in the cellular composition in blood that result from ageing [200]. It is still not known if this analysis reflects biological age of the organism or rather of the hematopoietic system (DNA was blood-derived) [193]. Another point that should be taken into account is that epigenetics reflect the interaction of genetic and environmental factors suggesting that clinical or lifestyle parameters in the two cohorts might alter individual and population methylation levels.

Our findings clearly provide further support that the most promising biomarker predictor of age is DNA methylation modification at specific cytosine that occurs through mammalian life time. It is also not yet clear how AR-DNAm changes, which seem to occur in a coordinated and reversible manner, are governed and if they have

biological consequences. The precision of the method used here would enable forensic scientists to estimate the age of perpetrators, especially if the assay could be scaled down to work with small trace evidence of DNA. Furthermore, blood evidence is very common at crime scenes and, as DNA is relatively stable, the age prediction approach might enable scientists to investigate even the oldest forensic samples that had been preserved.

The study reported in this chapter has several limitations that have to be taken into account in data interpretation. The methylation measurement was limited to only females, blood samples (a complex mix of white blood cells) and from the same ethnic background (Kurdish). So, more studies are needed to isolate nDNA and analyse its methylation in specific subtypes of blood cells and other tissues such as saliva, skin and sperm. Further, DNA analyses from other populations and from male samples need to be carried out to assess the generality of our findings. Finally, the impact of epigenetic modifications on determinants of ill health, including smoking, pollution, obesity and chronic life stress should be studied to quantify external influences.

4.5 **Conclusion**

Overall, this experiment has taken a broad look at nDNA methylation patterns in blood samples of different age (18-91 y). Several interesting gene promoters have been studied. The data presented suggests that the use of methylation modification at three genes (*ASPA*, *ITGA2B* and *PDE4C*) can be used as an indicator of age prediction. The outcomes of this work are applicable not only in forensic biology, but also in clinical research where it might aid the study the molecular basis of the phenomenon of ageing in health and disease.

CHAPTER FIVE
mtDNA Methylation

5 Chapter 5 Quantification of Global Mitochondrial DNA Methylation Levels and Inverse Correlation with Age

5.1 Introduction

Quantifying the nuclear DNA (nDNA) methylation of cytosine bases at specific autosomal genetic loci has been used to estimate an individual's age for forensic and medical purposes [74,193,202-204]. The existence of methylated cytosines within mtDNA has been controversial. Originally proposed four decades ago, the presence of mitochondrial methylation has been difficult to confirm [205, 206]. Maekawa *et al.* showed signals for mitochondrial methylation in 2004 [207]. Infantino and co-workers provided the first evidence of methylated bases (5-methyl-2′-deoxycytidine) present in human mtDNA using mass spectrometry [208]. MtDNA methylation has now been reported in the form of methylcytosine (mC) and hydroxymethylcytosine (hmC) modification[209], and a mitochondrially targeted DNA methyltransferase 1 enzyme (mtDNMT1) is suggested to be responsible.[118, 210, 211]-[212-216].

MtDNA methylation has been proposed as a cause of ageing and disease [118,217, 218]. A comprehensive profile of methylation levels across the mitochondrial genome and across ages would be an invaluable starting-point to address the role of epigenetics in these biological phenomena. Here, we describe the next generation sequencing analysis of blood-derived, bisulphite-treated mtDNA taken from a cohort of subjects with ages ranging from 18 to 91, using the Illumina MiSeq sequencing platform.

We selected functional regions of interest (ROI) within the mitochondrial genome and designed primers for methylation analysis for example, the displacement loop (D-loop)/ hypervariable regions that contain the promoters for the 'heavy' and 'light' strand polycistronic mRNAs, specific mitochondrial genes, and regions where published evidence already exists for methylation [211]. In addition, two replicated sites site of mtDNA mutation were chosen because of their potential to create or destroy CpG dinucleotide: an A11778G associated with leber's hereditary optic neuropathy, [219] and a T8993G mutation, which is associated with a neurological phenotype [220, 221].

5.2 **Mitochondrial DNA CpG Sites**

The human mitochondrial genome sequence which is used as a reference by Qiagen (www.Qiagen.com) consists of 16.565 base pairs (bp) with 435 predicted CpG sites (Figure 5-1) and 4746 cytosine residue at non CpG sites (**Appendix 4**).

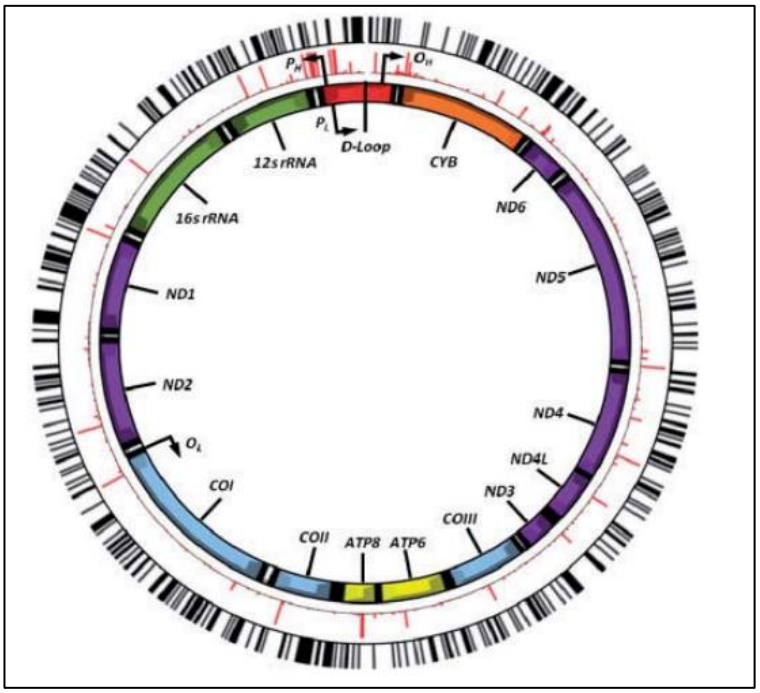

Figure 5-1. The location of CpG sited in the mitochondria genome: The location of CpG sites in the mitochondrial Genome. The figure shows the genome of mitochondria (centre), the size of the 13 genes, and the origin replication of both (heavy and light) strand (OH and OL, respectively) as well as the promoter sites of both strands (PL and PH, respectively). The (red) part in the middle of the image shows relative frequencies and position mtDNA variants. The outer ring (in black) illustrates the possible positions of the 435 CpG sites. Illustrated from [211].

The author hypothesised that methylation of mitochondria could be used as a marker for detection the age of contributor in body fluids for crime investigation purposes. For decades, the methylation of mtDNA in biology, more specifically in forensic science, was somewhat ignored based on an unfounded belief that mitochondrial methylation does not happen because of two reasons: (i) mitochondrial genome lacks histone and is arranged in clusters, and are bound to the mitochondrial membrane. (ii) the methylation enzyme does not access human mitochondria[222]. However,

Maekawa *et.al.,* (2004) study showed that mitochondrial methylation exists, but without any diagnostic or informative value at least for cancer detection[207]. Recently, methylation was detected in mtDNA, but reported as unusual CpG and non-CpG sites, especially in the control region [212]. But mitochondrial methylation remains controversial as still more recent study report that CpG methylation is absent in mtDNA [206].Thus, little attention has been devoted to mtDNA methylation.

Conversely, Chinnery *et. al.,* showed evidence in their experimentation that methylation exists in mitochondrial genome [211]. Furthermore Lacobazzi *et. al.* corroborated this result and showed that in both CpG and non CpG dinucleotide, cytosine can be methylated, but their frequencies would be varied [118]. A segment from mitochondrial genome is known as control region or displacement loop (D-loop) region with position (16,024-0576) [223]. D-loop is the largest non-coding region and has two hypervariable regions 1 and 2 (HVI and HVII). HVI extends from position 16024 to 16365. While HVII extends from approximately position 73 to 340, but HV1 is more variable than HV2[8]. Recently a third hypervariable region (HVIII) was described which extends from position 438-574 nucleotides (nt). The latter region is less polymorphic than HVI and HVII, but can assist resolving ethnic origin [224].

Both the origin of replication and transcriptional promoters are located in D-loop region [225]. Outside of this region is the coding region and the sequence of the mitochondria is nearly universal among all of the population, with few discriminations among individuals. Methylation is predicted to occur in the 13 major gene sequences and individuals with different ages might have varying methylation levels [118]. It can be seen from the mitochondrial reference sequence (**Appendix 5**) that there are various CpG dinucleotides within each of the 13 genes Table 5-1. Since methylation occurs in these genes therefore, they may be used as a marker tool for an age study.

Table 5-1 Mitochondrial gene locations: The table gives information about all the 13 encoding genes of mitochondrial organelle. The location of the genes was illustrated from mitomap website (http://www.mitomap.org) [226].

Gene	Map Locus	Abbreviation	Location
NADH dehydrogenase 1	MTND1	ND1	3307-4262
NADH dehydrogenase 2	MTND2	ND2	4470-5511
NADH dehydrogenase 3	MTND3	ND3	10059-10404
NADH dehydrogenase 4L	MTND4L	ND4L	10470-10766
NADH dehydrogenase 4	MTND4	ND4	10760-12137
NADH dehydrogenase 5	MTND5	ND5	12337-14148
NADH dehydrogenase 6	MTND6	ND6	14149-14673
Cytochrome b	MTCYB	Cytb	14747-15887
Cytochrome c oxidase I	MTCO1	COI	5904-7445
Cytochrome c oxidase II	MTCO2	COII	7586-8269
Cytochrome c oxidase III	MTCO3	COIII	9207-9990
ATP synthase 6	MTATP6	ATP6	8527-9207
ATP synthase 8	MTATP8	ATP8	8366-8572

5.3 **Materials and Methods**

All details on materials and methods can be found in Chapter 2. This section will describe the specific steps relating to the methylation and NGS which are unique to this chapter and others which related to primers design are mentioned in chapter four.

5.3.1 **DNA Samples**

A qualified phlebotomist took 5 mL blood samples from 82 female volunteers (age 18-91 y) at Erbil Hospital, Iraq. All the volunteers were from the Kurdish ethnic group but were recruited randomly without any health history or background information on education, marriage status, BMI, smoking, parity or lifestyle. 100 µL aliquots were immediately placed in Eppendorf tubes and DNA extracted. The extracted DNA was quantified by Stratagene 3005X qPCR instrument (Agilent

Technologies, CA, USA), using the Quantifiler® Human DNA Quantification Kit cat. # 4343895 (Life Technologies, CA, USA).

5.3.2 **Mitochondrial Genome Sequence**

To obtain a DNA sequence which would be in close proximity to the region of interest (i.e. DNA sequence around interested CpG site), the assembly of genome reference consortium human Build38 (GRCH38 or hg38) was used and for mitochondria (ChM) use in the assembly. The GRCh38 assembly offers an "analysis set" that was created to accommodate next generation sequencing read alignment pipelines. Interestingly through mathematic calculation in the D-loop of a mitochondria segment (16321-16570), nine CpG sites have been observed. Another segment from coding region locations (001-190) contains 11 CpG dinucleotides by virtue of the location of replication origin or the promoter of heavy and light strands, the most interesting region for exact positions of the D-loop CpG sites to design primers within the locations below:

1- Origin of replication, locations are at 400-625 or at 700.

2- L-strand promoter (LSP) locations; 392-445.

3- Major H-strand promoter one HSP1 location; 545-567.

4- Minor H-strand promoter two (MT-HSP2) locations; 645-670 [225].

Hence all the CpG dinucleotide which is shown in

Table 5-2, might be used as a tool to investigate methylation status inside D-loop region.

Table 5-2 Region of interest in mtDNA:- Locations and the number of CpG sites of each gene which might be methylated and studied for comparison purpose through age. The information in this table has been compiled by the author from the mtDNA genome sequence**.**

No.	Gene name	mtDNA Location	CG#

1-	Promoter (-strand)	ChrM: 240-600	5
2-	Promoter (+strand)	ChrM: 141-700	9
3-	12Sribosomal RNA	ChrM: 600-1600	31
4-	Origin of Light strand	ChrM: 5640-5940	10
5-	Cytochrome c oxidase subunit	ChrM: 5880-6240	16
6-	Cytochrome c oxidase subunit 2	ChrM: 7561-8340	29
7-	Cytochrome C oxidative subunit3	ChrM: 9141-9940	25
8-	MT-ND6	ChrM: 14100-14640	8
	Total		133

Moreover, this study investigates the consequence of mtDNA mutations and their relationship with the ageing process. We also discovered that some mitochondrial mutations associated with human disease cause CpG sites to be removed or produced. Thus the new CpG form might be tagged with methyl group and in contrast might be come unmethylated when one base pair of CpG form was mutated to another nucleotide. A substitution mutation at G3460A mtDNA which is considered as a RNA mutation type and related to Leigh syndrome[227] was chosen to be investigated. We found that the (G) nucleotide is followed (C) in mtDNA sequence, so when (G) at (3460 bp) was mutated to (A) there is no longer CpG and no longer mCpG proposing that the CpG site was methylated originally. Another mitochondrial DNA point mutation is a G-to-A mutation at nucleotide position (np) 11,778A, which is associated with Leber's Hereditary Optic Neuropathy (LHON) disease [219]. Interestingly, this might be valuable for CpG methylation study, as (C) comes before the mutated (G) and when the latter nucleotide altered to (A) thus CpG will no longer exist at that location as shown in Table 5-3. In contrast some time CG dinucleotide form will appear as a consequence of point mutation as T-to G at nucleotide 8993, which is associated with neurological syndrome [220]. This

heteroplasmic mutation causes change in an amino acid from a highly conserved leucine (Leu) to arginine (Arg) in mitochondrial ATase [221].

Table 5-3 Mitochondrial diseases: This table shows three types of point mutation in mtDNA and might be useful for methylation study. Reproduced from [227].

Type of mutation	Point Mutation	Disease
Proteins	T8993G	Neurogenic muscle weakness, Ataxia and Retinitis Pigmentosa (NARP)
Proteins	G11778A	Leber Hereditary Optic Neuropathy (LHON)
tRNA	G3460A	Leigh Syndrome

5.3.3 **Region of Interest Sequence**

The mitochondrial sequence (GRCH38 or Anderson/Cambridge reference sequence)[8] consists of 16,569 bp in which 435 predicted CpG sites are present (Table 5-4). Regions of interest (ROI) were chosen on the basis of biological function: the promoters for both ‘heavy’ and ‘light’ strand of the D-loop regions, polycistronic mRNAs and most of the functional genes of mitochondria.

Table 5-4:- Target region in Mitochondria: The regions of interest (ROI) with their sequences and primer sequences (forward and reverse).The single strand sequences of the ROI after sodium bisulphite conversion with the primer sequences (forward and reverse) details in **(Appendix 6).**

Target	Forward sequence primer (3'-5')	Reverse sequence primer (3'-5')
Promoter_(-)	TTTATTTGTTTATGGGGTGATGTGAGTT	AACCTAACCAAATTTCAAATTTTATCTTTTAA
Promoter_(+)	ACCCTAACACCAGCCTAACC	TTTGGGGTTTGGTTGGTTCG
12S_Ribosomal_RNA	GAGTTATAGTTTAAAATTTAAAGGATTTGG	AAACTATACCTAAAACTCCAACTCA
Origin_of_Light_Strand	TAGTTAATAGTTAAGTATTTTAATTAATTGGTTTT	AAAAACTAATCAATTACCAAAACCTC
Cytochrome_C_Oxidase subunit-1	TTGTTTTTAGATTTATAGTTTAATGTTTTATTTAG	TACTCCACAAATTTCAAAACATTAAC
Cytochrome_C_Oxidase subunit-2	AAATTAATTGGTTATTAATGGTATTGAATTT	ACTTCCAATTAAATACATAAATAAATAACCTACAA
Cytochrome_C_Oxidase subunit-3	ACCCACCAATCACATGCCTA	GTGTTACATCGCGCCATCAT
T8993G	TTTATGAGYGGGTATAGTGATTATAGGT	ACTTCCAATTAAATACATAAATAAATAACCTACA
G11778A	GGGTTTATATTTTTATTATTATTTTG TTTAGTAAAT	AACRAAACTTACTAAAAATCATCAAAA
G3460A	ACGGGCTACTACAACCCTTC	ATGCTCACCCTGATCAGAGG
MT-ND6	TAAAGTTTACCACAACCACCACC	GTGTGGTCGGGTGTGTTATTATT

Primer sequences used to amplify the ROI (each amplicon 100-300bp in length) and the amplicon sequence (before and after bisulphite conversion) are in (**Appendix 6**). Primers targeting 10 ROI amplicons (Table 5-5) were designed using Methyl Primer Express Software v1.0 (Applied Biosystem, Foster City, California), synthesised (Applied Biosystems, Foster City, California, and Life Technologies, CA, USA), and successfully passed amplification quality control criteria (**Appendix 7**). Some locations within the each ROI were impossible to amplify due to repeat sequences or high CpG content.

Table 5-5 Region of interest in mtDNA: Shows the 10 amplicons which passed Illumina quality control criteria. The location of each ROI in mitochondria genome (ChrM) and estimated amplicon number in the mtDNA are shown as well.

ROI	Location	Estimated number of amplicons
Promoter (+)	ChrM: 141-700	1
Promoter (-)	ChrM: 544,525-624	1
12s Ribosomal RNA	ChrM: 501-1700	1
Origin of light strand	ChrM: 5541-6040	1
Cytochrome C oxidative subunit1	ChrM: 5541-6040	1
Cytochrome C oxidative subunit2	ChrM: 5880-8340340	2
Cytochrome C oxidative subunit3	ChrM: 9141-9940	1
MT-ND6	ChrM: 14001-14740	1
T8993G	ChrM: 8781-9160	1
G11778A	ChrM: 11541-12040	1
		Total:10

5.3.4 **Assay Design for Methylation Study**

Likewise at chromosomal genome section mentioned, design of the assay, sample preparation and multiplex targeted amplification were done depending on the same parameter for target region. After ROI report, amplicon performance testing, sample processing and targeted amplification, the subsequent steps (sequencing and bioinformatics analysis) were carried out at Zymo Research.

5.3.5 **Bisulphite Analysis and NGS Library Construction**

The mtDNA was modified (non-methylated C residues converted to U) via a bisulphite conversion step, EZ DNA Methylation-DirectTM Kit (Cat. 5020 and 5021,

Zymo Research).[187] Multiplex PCR amplification of the ROI (Fluidigm Access ArrayTM System, BioMark, USA) generated pools of amplicons that also employed a universal forward tag common sequence 1 (CS1), and universal reverse tag common sequence 2 (CS2) (CS1= 5'-ACACTGACGACATGGTTCTACA-3', CS2 = 5'-TACGGTAGCAGAGACTTGGTCT-3'). Each individual amplicon pool was subsequently barcoded as described in the previous chapter. Two age-matched libraries, each consisting of 41 individuals, were created and purified (ZR-96 DNA Clean & Concentrator™ - ZR, Cat.# D4023) and then prepared for massively parallel sequence by Illumina next generation sequencing (NGS) using a MiSeq V2 300bp Reagent Kit (cat. # MS-102-2001)**,** (paired-end sequencing protocol) according to the manufacturer's guidelines.

5.4 **Results and discussion**

5.4.1 **Results**

Bisulphite treatment and high-throughput sequencing (Illumina) were used to assess methylation levels within regions of interest chosen because of their functional importance within the mitochondrial genome. In total the study queried methylation levels at 133/435 CpG sites in the sample set.

We identified unambiguous evidence for non-zero mtDNA methylation at 54 (**Appendix 7**) of the 133 CpG sites sequenced with average read depth of 2000. Methylation levels were, on average, very low (often between 2-6%) but showed regional differences across the mitochondrial genome and, more importantly, a great variance between individuals, as the methylation maxima and averages indicate (Figure 5-2).

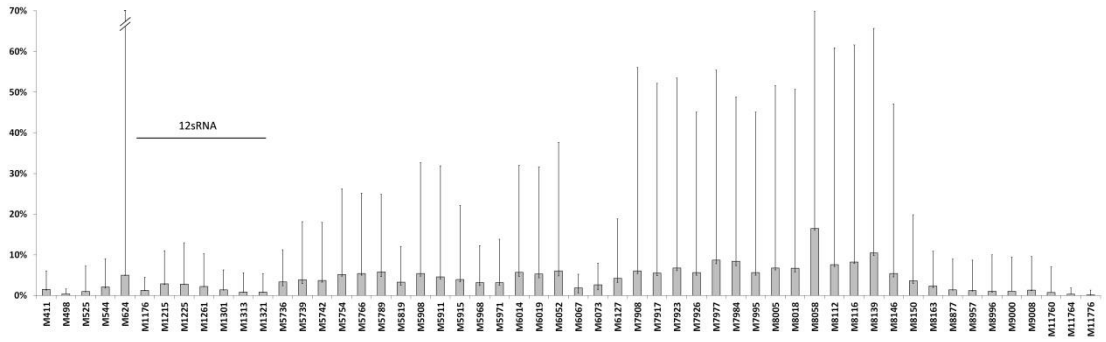

Figure 5-2: Mean methylation levels at 54 CpG sites across the mitochondrial genome (category labels denote reference base position). While methylation levels are typically 2-6%, considerable inter-individual variation was observed, indicated by minimum and maximum range bars. CpG sites within the 12S RNA gene are highlighted (see below).

The correlation between methylation and age was determined in the data set sample using Pearson's correlation coefficient for the 54 sites. Sites MT1215 and MT1313, both in the 12S RNA gene (*MT-RNR1*) showed significant methylation changes (hypomethylation) with increasing age (MT1215, R=-0.322, p=0.043; mt1313 R=-0.383, p=0.015). These findings were consistent with age-related changes reported in the general 12S RNA gene region by others [118, 167].

In the next step, multivariate linear regression was applied to the methylation data from both M1215 and M1313 (Figure 5-3a, b) and a linear regression prediction model was developed relating age to methylation.

The prediction accuracy of the model and contribution of particular predictors were assessed using the adjusted R2 statistic, which measures the proportion of age variation explained by the developed model. The model yielded an R2=0.509 indicating that 50% of the variation in methylation could be explained by age. The mean absolute deviation (MAD) between predicted and chronological age was 9.3 years (Figure 5-3c).

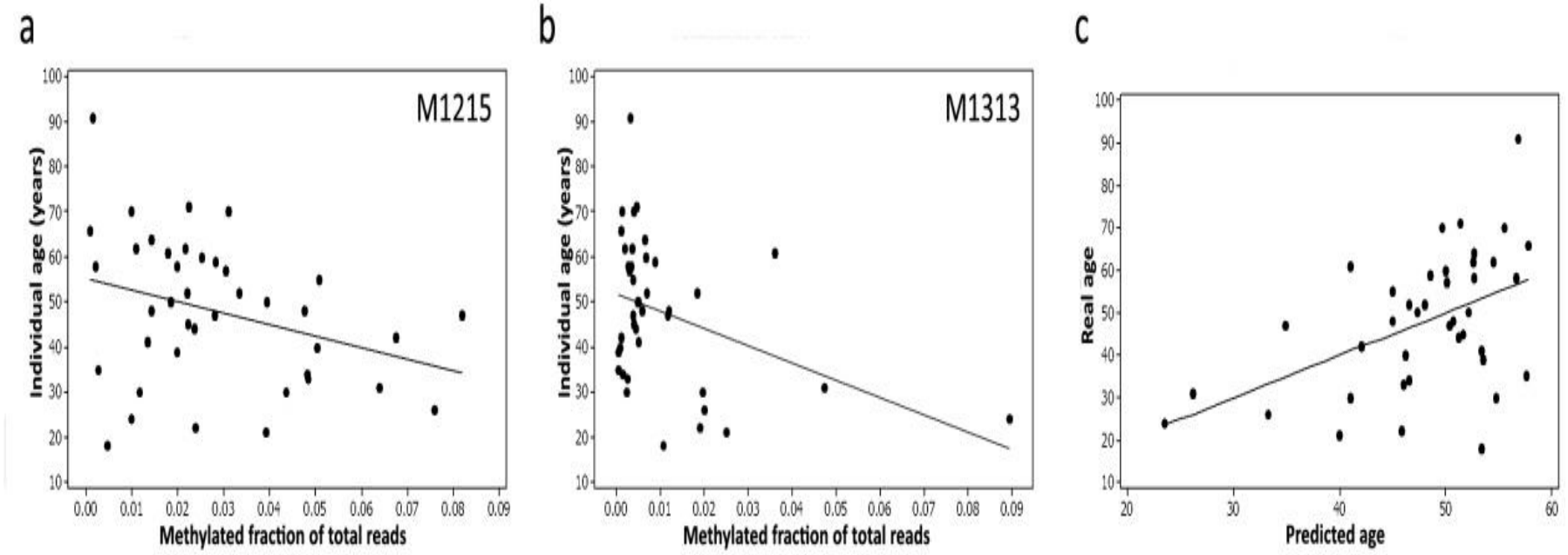

Figure 5-3: Methylation levels (expressed as a fraction) at two CpG sites within the 12S RNA gene: M1215 (a) and M1313 (b), correlate with age and can be used to construct an accurate predictive model (c) using Pearson's correlation**.**

The second sequenced library dataset (n= 41 individuals) was used to evaluate the model. From this set, 31 samples (64.5%) had ages correctly predicted to within a ± MAD. The analysed samples were also divided into three age categories: I: 18–39 years, II: 40–59 years, III: 60 years and older. Prediction results were significantly better for samples from category II (83.3% correct predictions) than for categories I (53.8%) and category III (20% correct predictions), suggesting that innate or environmental methylation-determining factors, in addition to age, are at play in early and late life.

5.4.2 **Discussion**

Although previous reports have shown correlation of mtDNA damage and mutation with ageing, [200] it was only very recently that studies suggested that the mitochondrial methylation could be used as an ageing biomarker [118, 213, 228]. Dzitoyeva and co-workers have reported the existence of mtDNA (hydroxyl) methylation and its increase with ageing in subregions the mouse brain [209]. Our study provides the most comprehensive map of methylation levels at 133 of the available 435 CpG sites in the mitochondrial genome from 82 individuals. Quantified changes in the methylation state (hypomethylation) of two 12S RNA gene CpG sites (M1215 and M1313) correlate with age. This novel discovery opens new avenues to investigate the mtDNA changes associated with health status and in the processes of criminal investigation. The CpG pair at M1215 and M1313 could only have been detected using high throughput sequencing, such as by Illumina MiSeq300 v2.

Therefore, such sensitive and advanced methodologies can solve the issue of failure of others to detect low-level methylation in the mtDNA genome that may explain reported absence of Mitoepigenetics [189, 190], its failure as a health tool marker [207] or its description as having an unusual CpG pattern [212]. Also contributing to this confusion might have been the use of quantification techniques such as ELISA and MeDIP that may struggle with low-level methylation.

The correlation between M1215/M1313 methylation and age was especially accurate for younger and middle-aged individuals in contrast to the poorer predictive power after the age of 60. This might reflect differences in individual lifetime metabolic

health or exposure to environmental influences, thus giving rise to discrepancies between biological and chronological age. Furthermore, we speculate that these life-accumulated environmental influences cause the substantial general mtDNA methylation variance observed between individuals. Alternatively, methylation differences might also reflect varying representation of cell-types in blood samples again, a potential surrogate for cytological investigation that might provide useful insights into immunological status. Our study did not differentiate between 5mC and 5hmC as the chemical conversion by sodium bisulphite (the gold standard detection technique for methylation studies) distinguishes only converted cytosines from non-methylated cytosines. This study has some limitations that have been taken into account in data interpretation. Further studies are needed to isolate mtDNA and analyse its methylation in specific subtypes of blood cells, such as platelet progenitor and lymphocytes. Further, mitochondrial DNA analysis from other populations and from male samples needs to be carried out to assess the generality of our findings. Lastly, the impact on mtDNA methylation of determinants of ill health, including smoking, pollution, obesity and chronic life stress are should be studied to quantify external influences.

To conclude, Illumina sequencing of 82 human blood samples indicated clear methylation patterns in mtDNA at 54 CpG sites in mitochondrial genome and that the level of the methylation was variable among different ages (18-91 years). Two of these CpG sites (M1215 and M1313) showed stronger correlation between predicted and chronological age with only about 9 years mean absolute difference (MAD). This finding implies that mtDNA methylation will be an available biological marker for forensic age-prediction and health status measurement.

CHAPTER SIX
EpiTect Methyl System

6 Chapter 6 The EpiTect Methyl qPCR Assay as a novel Age Estimation Method in Forensic Biology

6.1 Introduction

In previous chapters (Four and Five) we have identified age-related methylation signals but the technology used was expensive and may not be suitable for day-to-day forensic us. Thus we tested a cheaper and a faster method called EpiTect Methyl Assay as a novel age predictor in this chapter. Genome-wide analysis shows that global hypomethylation is typically associated with ageing [70,151,152,229]. Using epigenetic signature differences to classify and distinguish age level is a current topic area of forensic interest. Blood samples have been collected from female volunteers aged 18-91 years, as previously described. The aim of the study was to find a correlation between methylation level and human age by using EpiTect System a rapid and simple methodology originally intended for cancer epigenetics research, but here applied to forensic ageing research where it can operate with low concentrations of template DNA and without the need for bisulphite chemical conversion.

Methylation changes at specific gene regulatory regions can be used as an indicator of age of an individual [193]. Researchers have found that 60–70% of promoter regions of the human genome overlap with CpG islands [230-234]. These islands are regions of DNA (not less than 200 bp) with a high frequency (>50%) of CpG sites [231,235,236]. CpG Islands located in promoter regions of specific genes of interest which are known to show altered methylation levels as a function of age were analysed using the EpiTect Methyl qPCR Assay.

The methylation analysis technique used in our study relies on the detection of the remaining input genome after digestion with methylation-sensitive restriction enzymes. Following cleavage, intact DNA is quantified by real-time PCR. Subsequently, by comparing this with assay results from mock digest (no enzyme added), the relative fractions of methylated or unmethylated digestion can determine by using the ΔCt method, and the result displayed as percentage methylated (M) and unmethylated (uM). The promoter region of genes which are highly expressed are mostly unmethylated, regardless of the high GC content [234]. Existing technologies to quantify DNA methylation pattern include: combined bisulphite/restriction analysis (COBRA) [237],

methylation-specific PCR[238], bisulphite sequencing [239], high resolution melting (HRM) analysis, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) [240], next-generation sequencing [241], Pyrosequencing®[242] and fluorescence-based real-time Methy Light PCR[243]. The workflows commonly start with an initial stage of sodium bisulphite DNA conversion. This step is critical for the analysis because any technical mistakes made here profoundly alter apparent methylation level and cannot be corrected without repeating the treatment: however, additional DNA may not be available for this in the forensic setting. Thus, despite their single-base resolution of the methylome, these methods are not free from drawbacks such as protocol optimisation difficulties, the laborious nature of processing many samples, as well as the requirement for special instruments and equipment.

6.2 **Materials and Methods**

Although the details on materials and methods can be found in Chapter Two, this section will describe the specific steps relating to the EpiTect System which are unique to this chapter. To identify gene promoter methylation, EpiTect® Methyl qPCR Array (SA Bioscience, Qiagen, Hilden, Germany) was used according to the manufacturer's instructions. The EpiTect® Methyl II PCR Array System is accurate and reliable technology for DNA methylation studies of several genes simultaneously[244]. In this fast and an innovative technology, predesigned and ready-to-use primers were provided to detect reliable screening of high frequency of CpG sites termed CpG islands (CGIs) at promoter region which is generally lack DNA methylation in the gene [244, 245].

6.2.1 **Principle of the Commercial EpiTect System**

The EpiTect Methyl II qPCR system (SA Bioscience, Qiagen, Hilden, Germany) is based on the quantitative detection of remaining input DNA within a sample population after treatment with methyl-sensitive and methyl-dependent enzymes [246, 247]. Primers are designed by an optimized computer algorithm to ensure that the amplicon contains cutting sites for both digestion types and are specifically designed for analysing the DNA methylation status of CpG islands. SYBR Green-based real time PCR detection is employed after digestion. To make the system work, each genomic

DNA is subjected to four separate digestion treatments according to the manufacturer`s protocol and is shown in Table 6-1.

Table 6-1 EpiTect Methyl II PCR System Principle: Figure indicates the principle of EpiTect system and shows the four conditions of Mo (mock or without digestion), Ms (digestion with sensitive enzyme to methylation), Md (digestion with dependent restriction enzyme to methylation) and Msd (digestion with double or both enzyme). Adapted from [175].

Enzyme	How it works	Remaining DNA
Mock	No enzyme added	Input fraction
MSRE	Digests unmethylated and partially methylated DNA copies	Methylated fraction
MDRE	Digests methylated and partially unmethylated DNA copies	Unmethylated fraction
Both enzymes	Digests both methylated and unmethylated DNA copies	Fraction resistant to enzyme digestion (analytical window)

The first condition is Mock digest (Mo) which means no enzymes were added and the product of this digestion represents the total amount of input DNA for real time PCR detection. The second condition is a methylation-sensitive digest (M_s) which cleaves unmethylated DNA and real time PCR therefore hypermethylated DNA. The methylation-dependent digest (M_d) is the third condition, which digests methylated DNA and the qPCR detects the remaining unmethylated DNA. The final condition is for the double digest (M_{sd}), both enzymes were added, and all DNA molecules (both methylated and unmethylated) are digested. This reaction measures the background fraction of input DNA refractory (R) to enzyme digestion and double digestion (M_{dd}). The relative amounts of DNA species from the targeted regions are assessed a comparative ΔCT calculation using an automated Excel-based data analysis template provided by the manufacturer. As little as 2 µg genomic DNA can be used (recommended by manufacture) to profile the methylation status of target genes and methylated DNA can be reliably detected in heterogeneous samples [175]. This technology yields data comparable in quality to bisulphite Sanger sequencing and Illumina® assays, without the need for the bisulphite conversion step [244]. The basic principles of the EpiTect Methyl II PCR System along with performance, verification, and application data to demonstrate its robust potential for methylation profiling in various biological systems and for the screening of DNA methylation biomarkers [175, 244].

6.2.2 **Data Analysis**

After the quantification programme has completed, the row C_T values were obtained according to manufacturer`s instructions for the real-time PCR instrument. A standard Microsoft Excel-based template automatically performed all threshold cycle difference (ΔC_T)-based calculations from the raw value of the (C_T) to determine the methylation level of specific gene promoters [248]. The C_T values of both (sensitive and dependent) digests with the mock digestion values were analysed with a free software program which calculates and reports the methylated and unmethylated proportion of the DNA. This free software is available online and can be downloaded at: (http://www.sabiosciences.com/dna_methylation_data_analysis.php).

6.2.3 **Validation of Enzyme Digestion Efficiency**

The success of the EpiTect Methyl II PCR system depends on efficient DNA digestion by a methylation specific restriction enzyme (MSRE) and methylated dependent restriction enzyme (MDRE). Each assay includes specific enzyme control methylation-sensitive enzyme control (SEC) and methylation-dependent enzyme control (DEC) for monitoring the cutting efficiencies of these enzymes and to ensure that the result is reliable and reproducible [244]. After the C_T values are pasted into the Microsoft® Excel® data analysis spread sheet, a “Pass” or “Fail” result is returned for the SEC and DEC controls Figure 6-1. Each of SEC and DEC specific primers are included in each EpiTect assay and are available separately as primer assay (Cat.#.EPHS115450-1A and EPHS115451-1A) respectively.

Two control DNA molecules (one completely methylated and other unmethylated) are supplied spiked in buffer of EpiTect Methyl II DNA restriction kit Cat. #. (335452), each control DNA has specific enzyme-sensitive target regions flanked by unique primer regions. The results were accepted when both SEC and DEC assays passed the quality control. Thus for SEC, the C_T values of Ms and Mo digests were pasted in to Excel data analysis template. If the C_T values difference was equal or greater than four (ΔC_T $[M_s\text{-}M_o] > 4$), the software showed “Pass” in the analysis for. Likewise the difference between C_T values of Md and Mo digests should be (ΔC_T $[M_d\text{-}M_o] > 4$) for DEC assay. These results confirm that more than 93.6% of the control DNA molecules were cleaved, so the enzyme was active and digested input DNA samples efficiently.

	Samples		1		2	
2	Symbol	Catalog Number	UM	M	UM	M
3	NPTX2	EPHS113305-1A	35.22%	64.78%	54.04%	45.96%
4	KCNQ1DN	EPHS102112-1A	44.78%	55.22%	11.12%	88.88%
5	GRIA2	EPHS111325-1A	25.10%	74.90%	42.00%	58.00%
6	TRIM58	EPHS101433-1A	56.91%	43.09%	12.10%	87.90%
7	SEC	EPHS115450-1A	Pass		Pass	
8	DEC	EPHS115451-1A	Pass		Pass	

Instructions | Raw Data | QC Data Report | **Results** | Calculations

Figure 6-1 SEC and DEC control: The figure shows the result of the Microsoft excels data analysis and how both controls (SEC and DEC) passed.

An analytical window (W) in the Microsoft Excel data analysis template represents the C_T value differences between double digest and mock digests. To be a reliable result, (W) should be greater than 3 (ΔC_T $[M_{sd}-M_o] > 3$) which means more than 87.5% of input DNA was digested. In contrast if (W) was less than three ($W < 3$) this means that refractory DNA percentage (R) is greater than 12.5 percent ($R > 12.5\%$). Thus data quality control report (QC Report) worksheet in the Excel software is reported as a "Failure" Figure 6-.

	A	B	C	D	E	F
1	Samples		1		2	
2	Gene Symbol	Catalog Number	W	R	W	R
3	BCAS4-1	EPHS109350-1A	2.74	Failure	-4.45	Failure
4	KCNQ1DN	EPHS102112-1A	-4.81	Failure	-0.81	Failure
5	GRIA2	EPHS111325-1A	-0.93	Failure	2.71	Failure
6	TRIM58	EPHS101433-1A	1.6	Failure	1.4	Failure
7	SEC	EPHS115450-1A	11.78	0.028%	10.62	0.064%
8	DEC	EPHS115451-1A	5.58	2.091%	10.02	0.096%

Instructions | Raw Data | **QC Data Report** | Results | Calculations

Figure 6-2 QC Data Report worksheet: The figure shows the result of the quality control report and the value in analytical window (W) was abnormal (i.e. less than three) for the all age related genes. Similarly the refractory factor (R) of digestion was not complete for the input DNA means that the digestion efficiency was not high for DNA sample through of the Microsoft Excel data analysis.

Although it is recommended by the manufacturer to use (0.5-1 µg) of DNA template, we diluted the input DNA to the minimum amount possible to mimic a real crime case scenario. Hence, the minimum amount which passed the system criteria was 120 ng of

the template. The final amount of the template will become only 20 ng (i.e. 120/6) per gene. This is because the tube content was divided to six tubes (4 target genes and the controls (SEC and DEC)) as it is mentioned in detail in the enzyme digestion step. This amount of DNA sample is still considered high as it cannot be guarantee to retrieve such high exhibit in the scenes of the crime. Adjusting the system to produce results with less input DNA will be necessary of this technique is to be implemented into forensic work in the future. Furthermore, expand the study to other population and/or different gender (male) to see if the assay will predict the same rate as current study. Further study needs to validate the method with training sample and examine the EpiTect system to see if different tissue type can be discriminated, especially in sexual assault case.

6.2.4 **Statistical Analysis for the EpiTect Assay**

Statistical analyses were performed for all data using Excel 2010 (Microsoft Corporation), SPSS and Minitab v16 (Minitab Inc.) the detail is in (**Appendix 8**) First simple linear regression (LR) for each group of females ranging in age from 18 to 91 years was used to analyse the relationship (Correlation R^2) between the methylation percentages of each single gene. The result was considered significant at Statistical Significance (i. e. P-values) below 0.05 (p-value < 0.05). Standardized regression coefficients (β) were used to compare the effect size of particular gene. In the next step, multivariate linear regression (MLR) was applied, allowing simultaneous analysis of all the tested genes. In the second step, a linear regression prediction model between predicted and real age was developed based on the methylation data obtained for the same samples.

6.3 **Results and Discussion**

The EpiTect system was originally used in detection of breast [249] and brain cancer [248]. Therefore, this is the first study to use the method and gene set selection in a manner that could be adapted for forensic purposes. Human DNA was digested and amplified using pre-designed primers in the kit that flank the studied gene loci. The results of the 80 blood samples (each repeated twice) were displayed as percentage of unmethylated and methylated fraction (**Appendix 9**). This shows that DNA methylation successfully discriminated volunteered age.

It is known that DNA methylation has a crucial role in normal cell differentiation and development [250, 251]. Generally the promoter regions are hypomethylated and CGIs lack DNA methylation in the gene [244, 245]. However a small percentage are hypermethylated as well [251]. Based on the assumption that epigenetic changes may drive the heterogeneity of the cell, the promoter methylation levels of 13 candidate genes (*BCAS4, KCNQ1DN, GRIA2, TRIM58, EDARADD, TOM1L1*, *SOGA1, NPTX2, FGF7, ZC3H12D, ASPA, PDE4C* and *ITGA2B*) and details are summarized in Table 6-2). Previous studies have showed that these genes can be used as forensic markers for either age discriminate or tissue identification [74,204,252].

Table 6-2 Screening of 13 candidate genes using EpiTect Methyl qPCR Array: The table shows description, symbol, location and PCR product size of all 13 studied genes which screened by using EpiTect Methyl system.

No.	Gene description	Symbol	UniGen no.	NCBI ID	Catalog #	CpG island location	PCR size
1-	KCNQ1 downstream neighbour	KCNQ1DN	Hs. 127821	55539	EPHS102112-1A	Chr11: 2890388-2891337	157bp
2-	Breast carcinoma amplified seq4	BCAS4	Hs. 708239	55653	EPHS109350-1A	Chr20: 49411401-49412142	254bp
3-	Glutamate receptor	GRIA2	Hs.32763	2891	EPHS111325-1A	Chr4:158141404-158141836	189bp
4-	Tripartite motif containing 58	TRIM58	Hs.269151	25893	EPHS101433-1A	Chr1:248020330-248021552	262bp
5-	EDAR-associated death domain	EDARADD	Hs.352224	128178	EPHS101394-1A	Chr1: 236558459 - 236559336	196bp
6-	Aspartoacylase	ASPA	Hs.171142	443	EPHS105621-1A	Chr17: 3375006 - 3375237	215bp
7-	Target of myb1	TOM1L1	Hs.153504	1040	EPHS106189-1A	Chr17: 52977866 - 52978307	147BP
8-	Phosphodiesterase 4C	PDE4C	Hs.132584	5143	EPHS107250-1A	Chr19: 18335072 - 18337375	217bp
9-	KIAA0889	SOGA1	Hs.460807	140710	EPHS109245-1A	Chr20: 35491001 - 35492805	257BP
10-	Neuronal pentraxin II	NPTX2	Hs.3281	4885	EPHS113305-1A	Chr7: 98245805 - 98247759	174bp
11-	Integrin, alpha 2b	ITGA2B	Hs.411312	3674	No products	Chr17:42476290-42477054	765bp
12-	Fibroblast growth factor 7	FGF7	Hs.567268	2252	No products	Chr:1549549862-49820528	64148bp
13-	Zinc finger CCCH-type	ZC3H12D	Hs.632618	340152	No products	Chr6: 149447630..149485012	37382bp

However, in our study only 4 of 13 genes (**Appendix 10)** showed significantly different methylation levels according to age Table 6-3.

Table 6-3 Location of four genes of interest: The table shows bioinformatics and location of four genes which correlated with age and passed EpiTect system criteria. In addition to methylation analysis of both sensitive enzyme control **(**SEC) and dependent enzyme control **(**DEC).

Gene Symbol	Catalogue Number	CpG island Position	UniGene number	NCBI ID	PCR product size
GRIA2	EPHS111325-1A	Ch4:158141404-158141836	Hs.32763	2891	189
KCNQ1DN	EPHS102112-1A	Chr11:2890388-2891337	Hs.127821	55539	157
NBTX2	EPHS113305-1A	Chr7:98245805-98247759	Hs.3281	4885	174
TRIM58	EPHS101433-1A	Chr1:248020330-248021252	Hs.269151	25893	262
EP-SEC	EPHS115450-1A	Chr0:0-0			159
EP-DEC	EPHS115451-1A	Chr0:0-0			220

The percentage level of methylation was used to test for association between examined genes (*NPTX2, GRIA2, TRIM58* and *KCNQ1DN*) and age. Correlation values (R-square) for methylation at these genes were 0.452, 0.501, 0.808 and 0.550 respectively. The Pearson correlation (p-value) was significant in all cases (p-value < 0.05). Meanwhile the assay results for *EDARADD, FGF7, ZC3H12D, TOM1L1, SOGA1* and *BCAS4* genes did not show any correlation between age and methylation level (data not shown) and no significant differences were found (p-value >0.05). As discussed in Chapter Four the CGI methylation analysis of *ASPA, PDE4C* and *ITGA2B* genes reveals a strong correlation with age.

When information about these three genes was sent to the bioinformatics department of QIAGEN, response was that it would not be possible to create an assay for *ASPA, PDE4C* and *ITGA2B*, "*because there were not enough restriction sites in the provided sequences or within 100 bp on either side of them in the genomic context, so no primer design is feasible*" (Pers. Comm., from QIAGEN [253]. Blood DNA was used in this study, but more experiments need for other tissue types to investigate the discrimination power among different ages. Other biological fluids like saliva might give similar result as it is popularly thought that buccal epithelial

cells are the main source of DNA in saliva, but one study shows that up to 74% of the DNA in saliva comes from white blood cells [254].

Furthermore, all oral samples are not equal, because DNA quality from the same amount of saliva can be very variable and can explain the differing methylation extent of same age donors [254]. Saliva can be enriched in epithelial and blood cells, especially when the donor has gingivitis (i.e. teeth gum diseases) and resulting bleeding. Epigenetics can be thought of as an interaction between genetics and environment, so external factors such lifestyle, diet, pregnancy and even natural disasters, can have an effect on methylation patterns [255]. It could be argued that the EpiTect technique is a favourable route because it assesses multiple CpG sites simultaneously rather than single CpG sites. Most biological changes of DNA methylation are known to occur at multiple CpG sites simultaneously [234,256,257]. This regional analysis may better represent the methylation status of a CpG island than a specific analysis of single sites.

Generally the human genome is hypomethylated through ageing. But CGI methylation levels increase in promoter region in most cases, as these entire four target regions were hypermethylated with ageing as the scatter plots between each gene and age shown in Figure 6-2.

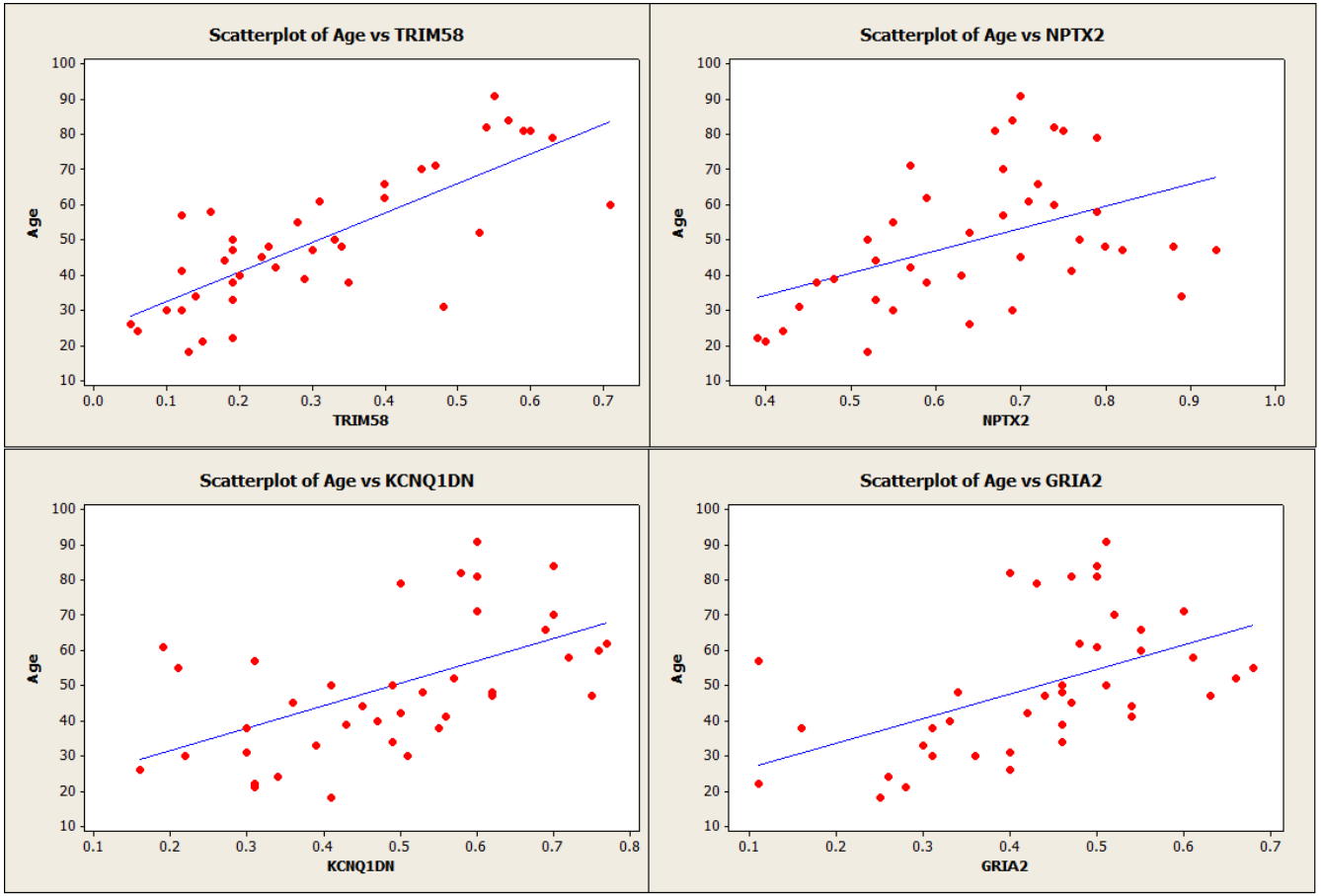

Figure 6-2 Prediction of age by methylation level: The figure shows scatterplot of promoter methylation level against sample age for four genes *NPTX2*, *KCNQ1DN*, *GRIA2* and *TRIM58*).**In each case, a positive correlation was observed.**

Multivariate linear regression used to predict age from methylation data and found the "standard error" value from the analysis report. The result of R value, R squared value, adjusted R squared value and "standard error of estimate" is 0.848, 0.720, 0.687, and 10.80734, respectively the detail is in (**Appendix 8**). The equation model that built in previous step was applied to build a linear regression to predict the real age (Age = 3.39986 + 17.8083 *NPTX2* + 18.5765 *GRIA2*+77.0834 *TRIM58*+7.30819 *KCNQ1DN*) for the female blood samples as shown in Figure 6-. Overall the calculation differences between predicted and real age was about 11 years and the absolute mean deviation (AMD) only in error by 7.2 years.

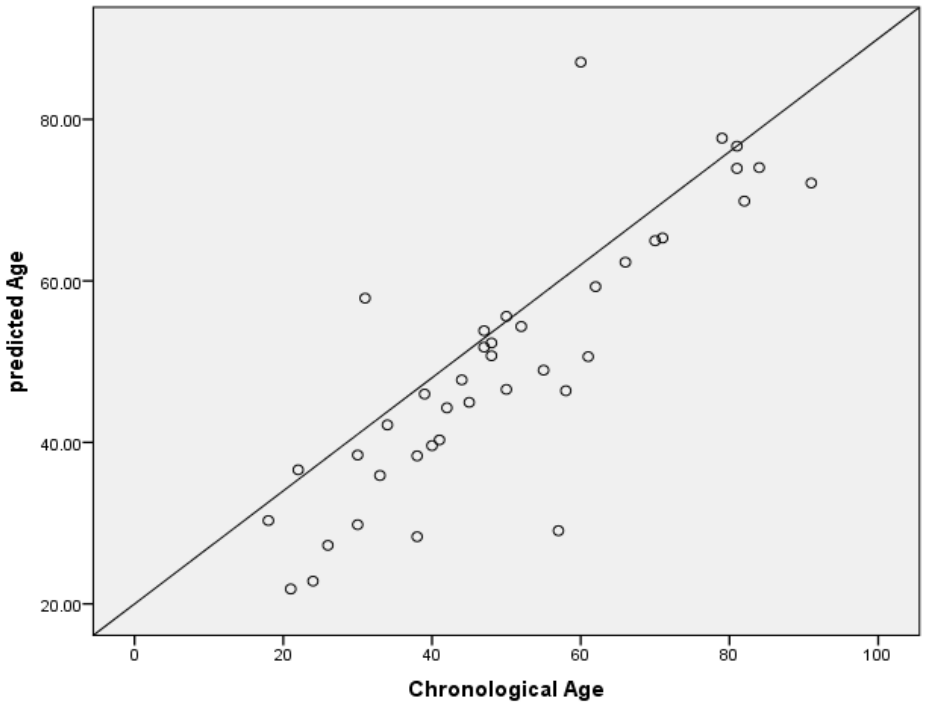

Figure 6-3 Predicted age versus chronological age: Predicted age versus chronological age: The figure shows a scatter plot of predicted age against chronological age of blood samples with regression line with the equation of (Age = 3.39986 + 17.8083 *NPTX2* + 18.5765 *GRIA2*+77.0834 *TRIM58*+7.30819 *KCNQ1DN*).

The results of this study show that the EpiTect system can be used to estimate the age of the donor through DNA methylation levels with an 11 year error range. While the accuracy would be practical use for forensic purpose, some present studies have presented an estimation of age with 5-7 years error [62, 193]. However, our result is

still better than the correlation of age with telomere length, mutation accumulation and T-cell rearrangement. This is a relatively small sample size; caution must be applied, as the findings might not be transferable to other gender (men) or females from different population. The present findings seem to be consistent with Koach *et. al*. (2011) who have shown that the methylation level of CpG sites in *NPTX2*, *KCNQ1DN*, *GRIA2* and *TRIM58* promoters strongly correlates with age with an error range of 11 years [74].

What is important to highlight in our study is that the analysis was done without DNA conversation with sodium bisulphite. Therefore, EpiTect Methyl qPCR system is an interesting alternative, as only a qPCR instrument is needed in any laboratory. It is a complete system, from sample isolation to result.

In comparison to other methods, only 20 ng DNA per gene was required for methylation analysis. In contrast Cygenia (http://www.cygenia.com/) for instance, needs at least 150 ng DNA or 1 mL of blood for methylation diagnostic using pyrosequencing. ZymoResearch (http://www.zymoresearch.com/) is another bioscience company for methylation pattern analysis service, but 0.5-1 µg DNA is required. Furthermore they charge about £ 75 (€ 95) / £ 90 ($150) per sample, compare this to the EpiTect Assay which only costs £10-15 per sample as shown in Table 6-4.

Table 6-4 EpiTect system VS other Techniques: The table compares EpiTect with pyrosequencing and Illumina next generation sequencing.

Technique	**EpiTect Methyl System**	**Pyrosequencing**	**Next Generation Sequencing**
Product	SA bioscience	Cygenia	ZymoResearch
Amount require	20ng/ Gene	150 ng	0.5-1 µg
Quality Absorbance: 260/280 ratio	high > 1.7)	high > 1.7)	high > 1.7)
Instrumentation	Only qPCR	Pyromark	Illumina
Bisulphite Sodium Treatment	No	Yes	Yes
Cost	£ 10-15	£75	£90
Time (hs)	12	35	35

6.3.1 **Validation of the Method**

To assess the statistic result analysis for unknown age samples (training set), the technique of cross validation model or sometime called 'rotation estimation' [258] was used. Thus, 40 samples as training set samples were applied to the equation model and calculate the error rate between real age and predicted age as in Table 6-5.

Table 6-5. Evaluation of the prediction potential of the EpiTect assay for training set of 40 blood samples. The prediction result was considered correct if the predict age matched the chronological age ± 11 years.

Age Prediction	**Age Category (year)**			**Total**
	18-39	40-59	≥ 60	
Incorrect (n)	5 (41.6%)	5 (27.7%)	8 (80.0%)	18 (45.0%)
Correct (n)	7 (58.3%)	13 (72.2%)	2 (20.0%)	22 (55.0%)

The table of the training set showed that the method is quite good for middle age (40-59 years), but interestingly for elderly women (over 60) the equation was able to predict only 20.0% (under the level of chance) of the samples and this result needs further investigation by studying other factors like biomedical and environmental causes.

6.3.2 **Analysis of Noise Factors Affecting Methylation Ratio**

Depending on the fact that methylation modification is occurring in different regions of the chromosome [259]. EpiTect methyl II qPCR assay is based on detection of quantified amount of DNA which survived upon enzyme digestion. The CpG loci of our targets are differentially methylated among individuals and the methylation patterns are generally age-dependent. However, as a result of natural diversity in methylation ratio among people with different ages, even with the same age (but not same environmental exposure and lifestyle impact), there could be potentially some overlap in the observed methylation ratios [254, 260]. Some artefacts such as

differences in DNA template concentration for instance (stochastic effect and pipetting error) are associated during experimental steps [61]. Furthermore, recent studies showed that epigenetic patterns can be affected by a wide variety of external influences including diet, exercise, smoking, maternal environment, and more [261, 262]. It was noted that chronological age is not identical with biological age and it is conceivable that some of the discrepancy between predicted and real age can be attributed to this difference further research might facilitate determination of the biological age for personalized medicine.

6.4 **Conclusion**

This is the first report describing the use of the EpiTect Methyl II PCR system (QIAGEN, SA-Bioscience) to estimate human age through methylation analysis. Although the system does not have the best age estimation accuracy published to date, it is better than telomere-length-based approaches and much faster and cheaper than DNA methylation quantification technologies that use bisulphite conversion chemistry. The system can be used as a simple and useful methodology in forensic work, or non-criminal instances where an objective estimation of age is required (e.g. family law, immigration, and insurance purposes etc.), with only an 11.3 year error between estimated and chronological age.

This study has also some limitations that have been taken into account in data interpretation. The error rate is quite high, so further studies are needed to develop better EpiTect assays that have better accuracy in age prediction. Furthermore, DNA analysis from other populations and from male samples needs to be carried out to assess the generality of our findings. Lastly, the impact on DNA methylation of determinants of ill health, including smoking, pollution, obesity and chronic life stress are should be studied to quantify external influences.

CHAPTER SEVEN
DNA from Cartridge

7 Chapter 7 Analysis of DNA from Fired Cartridge Casings

7.1 Introduction

The primary objective of this Chapter is to conduct an investigation on the survival of DNA in skin cell and sebaceous secretions when subjected to the stresses encountered in materials deposited on firearm cartridges when the cartridge has been fired. This information may assist the interpretation of results from analysing sebaceous secretions, i.e. touch, which are less readily controlled. It is necessary to determine first if the DNA sample that used in the tests will survive or not. Then we will determine if, with this survived evidence, we can build up a prediction of the contributors' physical feature, more precisely the age of contributor. This follows the earlier work in this thesis, and elsewhere, that shows that the chemical tags that modify DNA (methylation) can accurately predict the age of human tissues and cells [62, 203]. Recent work has raised the possibility of estimating some physical feature information such the age of the donor [68, 263]. If possible, this would provide the forensic scientists with more information to aid the investigation. DNA analysis has been widely accepted as providing valuable evidence concerning the identity of the source of biological traces [264].

It may also be possible to infer the ethnic origin of the donor [265]. Such information may be invaluable in certain cases where spent cartridges are left behind at the scene of a crime. Therefore, just a small quantity of DNA can give a rough estimate of the age of an offender. A crime investigator and/or forensic analyst may have numerous questions to answer during an investigation. One of these might be whether DNA and epigenetically analysable material can be obtained from fired cartridges. Professional investigators believe that there are no perfect crimes and according to Locard's exchange principle there will be some exhibits in crime scenes from perpetrator during handling any object present in the place [266]. So depending on both above facts, any possible touch DNA transferred to a cartridge case or bullet either before or during the gun loading process would be obliterated when the gun was discharged or handled.

Collection and extraction of DNA from fired cases will simulate the finding of such items in the areas of a firearm crime scene and subsequently if possible, analysing the DNA methylation state from the DNA. The challenging point in this study will be to assess if the temperature in the firearm chambers, pressure and potentially corrosive gasses generated during firing, will destroy or damage the DNA or can methylated DNA withstand this intense heat and survive. Presently forensic scientists are hardly able to obtain genetic profiles from cells shed onto touched or handled objects by using advanced DNA typing methods as mentioned in the literature review [15]. This type of DNA is called “contact DNA or touch DNA” and currently investigators and attorneys send request to the crime laboratory to try to find any evidence at a crime scene and analyse it [21].

Some common types of evidence at gun crime scenes are cartridge and firearms and these are potential sources of contact DNA and perpetrators may not be wearing gloves all the time. Thus they may handle cartridges directly during loading of a magazine. For this reason, in such type of firearm crime, forensic experts are often asked to swab cartridge cases for DNA evidence.

7.2 **Materials and Methods**

7.2.1 **Workflow Design**

We have designed a study to evaluate the suitability of DNA derived from cartridge cases for use in criminal investigations and also to determine if epigenetic information can be analysed to get an idea about perpetrator′s age. The workflow is showed as a diagram in Figure 7-1.

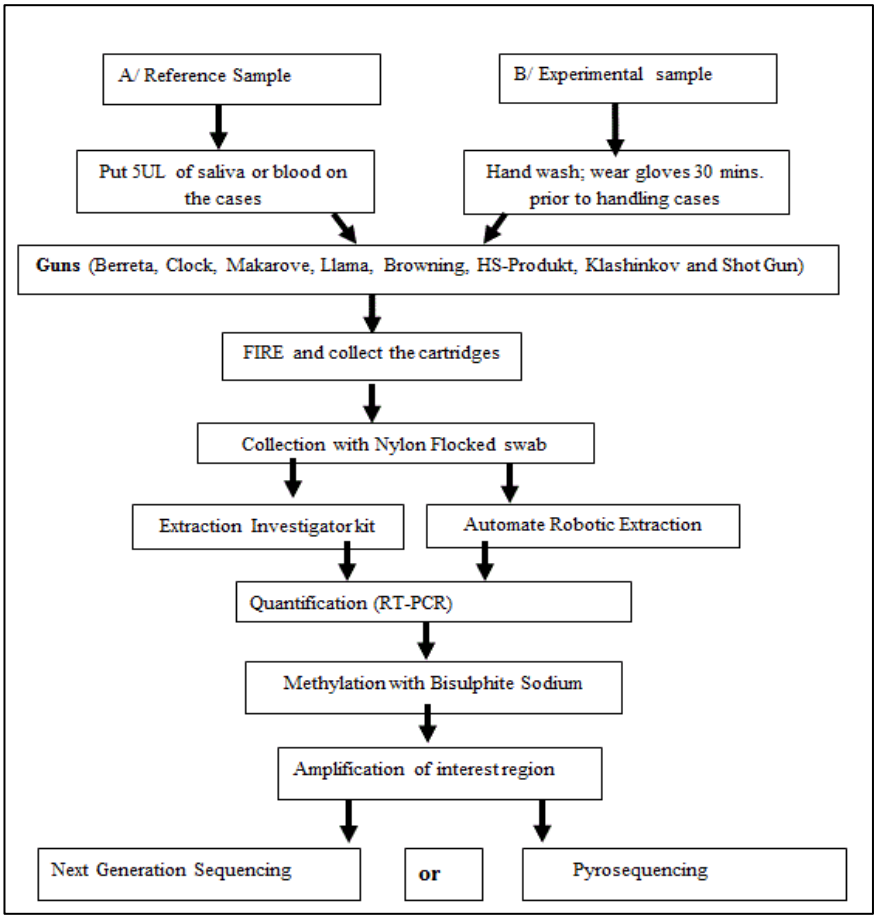

Figure 7-1 Schematic diagram of the steps of the experiment: The diagram shows both reference and experimental sample steps from handling to analysis.

7.2.2 Experimental Design

The experiment consisted of exposing the substrates to UV light before the experiment to destroy pre-existing DNA. After touching and loading all cartridges in to magazines or to the guns directly (Shotgun) by a known shedder (to avoid DNA shedder discrimination), the weapons were discharged under controlled circumstances. In all cases, each weapon was cleaned well (out and inside) and exposed to UV light to avoid source of contamination of ammunition: i.e. contamination could be possible not only from the person who loaded the gun but also those people at manufacture who involved in making, packaging and handling the bullets before the experiment start

Spent cartridges collected to recover DNA and conduct for further analysis. These was done by using advance and automate technique for DNA recovery to find the best result and then applying the method giving the best result to real samples.

The volunteer shedder was asked to load bullets into magazines (both cleaned before) of eight different types of weapons (shotguns have no magazine; hence, the bullet was directly loaded in to the gun). The name and shape of the guns are summarized in Figure 7-2. After the firing process the fired cartridges were swabbed by Nylon Flocked Swab (forensiX) chosen for the isolation step because the swab has higher ability to retrieve and reserve biological fluids, as described in a previous Chapter [267].

Figure 7-2 Eight different types of gun: The figure shows the weapon which were used in the study (from left to right). The top row is Berretta, Browning and Glock), in the middle row Shotgun one and double bore showed, whilst in the lower row, Llama, Kalashnikov and Makarov are shown.

The extraction was done using Prep Filer Express™ Forensic DNA Extraction Kit (Applied Biosystem). The advantages of this kit are its ability to cope with trace DNA and the extracted product will be free of PCR inhibitors.

Automating the DNA isolation process on this system enables appropriate retrieval of DNA in high yields from the cartridge samples.

To study fired cartridge cases, AutoMate Express™ DNA Extraction system was chosen, which is an automated extraction platform that reduces DNA extraction times from many hours to 40 minutes. Extraction with robotic system allows researchers to check the reliability, reproducibility, sensitivity and limitations of the methods. This system is able to remove PCR inhibitors (dyes) from samples, and prevent sample-to-sample contamination [268, 269].

The most critical step of the analysis is the extraction of the DNA from forensic evidence sample, especially poor DNA samples [270]. The cells of interest (nucleated cells) contain other substances besides DNA such as, proteins, carbohydrates, lipids, etc. Therefore, the DNA molecules must be separated from this cellular material before it can be further examined [271]. Although, some other non-automated methods like organic extraction are used for DNA extraction. But multiple steps are required with these manual extractions can lead to contamination and/or loss of sample. It is also time consuming and does not remove all PCR inhibitors [270, 271].

Automated systems for DNA extraction have proven to be extremely useful, especially when processing reference samples (samples of known origin that contain ample amounts of DNA). Automated extraction is much faster approximately thirty min for completion compared to other extraction methods. With automated systems there are less manipulations of the sample and no organic solvents are used. In the context of the laboratory, the absence of organic solvents makes the process safer for the person performing the extraction and the automation allows for faster extraction. Automated extraction methods use solid-phase extraction, compared to liquid-phase extraction, which is used in the organic extraction. Liquid-phase extraction requires multiple manipulations, requires organic solvents and is harder to automate, whereas solid-phase extraction is much easier to automate and does not require organic solvents.

The Applied Biosystem Corporation developed an automated DNA extraction system (Figure 7-3). This automated system can purify high quality DNA from up to 13 samples simultaneously and in approximately thirty min. The system extracts DNA by magnetic bead particle technology.

Figure 7-3. AutoMate Express™ DNA Extraction System: The figure show front side of the Robot Extraction System which used for DNA extraction from fired cases.

7.3 Results and Discussion

7.3.1 Collection, Quantification and DNA Typing

The first feature under investigation was the total yield amount of DNA obtained from the cartridge surface, to establish whether or not the DNA survive after firing through hot temperature and gas pressure. In order to collect the most biological genome on the fired case surfaces and the DNA collected. After DNA measurement with qPCR the samples were processed for STR profiling using the Minifiler kit (Applied Biosystem, part #: 4374618) following the protocols as recommended by manufacture [176]. After running samples on the 3130 Genetic Analyser (Applied Biosystem), the expected DNA profile was produced and the result showed that the

cartridges of some weapons were more likely to produce viable amount of DNA than others as illustrated in Figure 7-4.

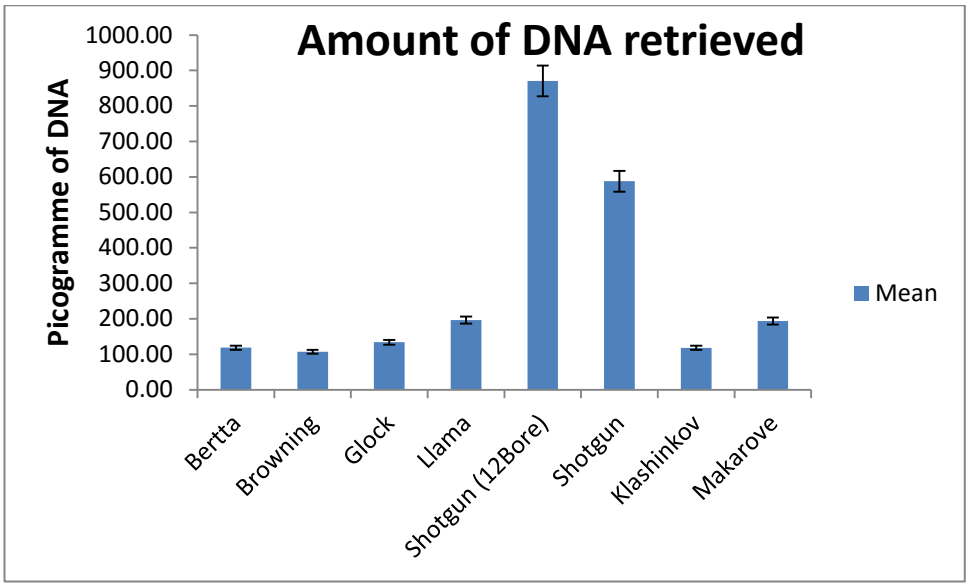

Figure 7-4 Comparison data: The table compares the average (mean ± standard error) amount of collected DNA (pg) from each of the eight tested guns.

The results from the shot shell surface of 12 Bore shotgun was around 0.9 ng (900 pg) of DNA even after shooting. By contrast, cartridges of the Browning pistol represented the smallest amount of evidence items submitted for DNA analysis, around 0.1 ng (107 pg). When PCR amplification was done using Minifiler kit (Life Technologies, Paisley, UK), we were able to produce full DNA profile or at least partial profile for the majority of shot shells from both types of shotgun.

We can explain the reasons that caused higher amounts of retrieved DNA from shot shell than other guns. First of all, the shot gun which we used was a non-automatic weapon that means every single shot shell needed to be manually loaded before shooting and discharged again after firing. Hence, there is no high gas pressure effect the evidence on surface of the ammunition. The size of the bullet is another reason which is bigger and longer (35 mm) than other types of cartridge. Therefore there is more area surface between contacts during handling and this is Locard`s exchange principle "every contact leaves a trace" [266]. There is another possible explanation which is the fact that shotgun has no magazine. Thus the biological sample left

directly to the gun. While in other automatic weapons which tested the, cartridges may lost a portion of biological fluid inside the magazine surface as in direct contact.

7.3.2 **STR Analysis**

The amount of collected DNA from cartridges of other tested guns was too low and not efficient to build a DNA profile. The 3130 Genetic Analyser results in most cases showed only a partial profile. Most precisely for the Beretta, Browning and Kalashinkov weapons there were no clear peaks in most, if not all, the loci using Minifiler Kit as illustrated Figure 7-5. For STR profiling purpose, the Minifiler kit (Applied Biosystem) was used following the protocols as recommended by the manufacturer. This assay is optimized for genotyping degraded and/or inhibited DNA samples. In addition to sex determination (*AMEL*) locus, the kit amplifies eight autosomal STR loci. These loci (*D13S17*, *D7S820*, *D2S1338*, *D21S11*, *D16S539*, *D18S51*, *CSF1PO* and *FGA*) span a range between 70 to 283 nucleotides with the aid of non-nucleotide linkers to achieve appropriate spacing between loci [176].

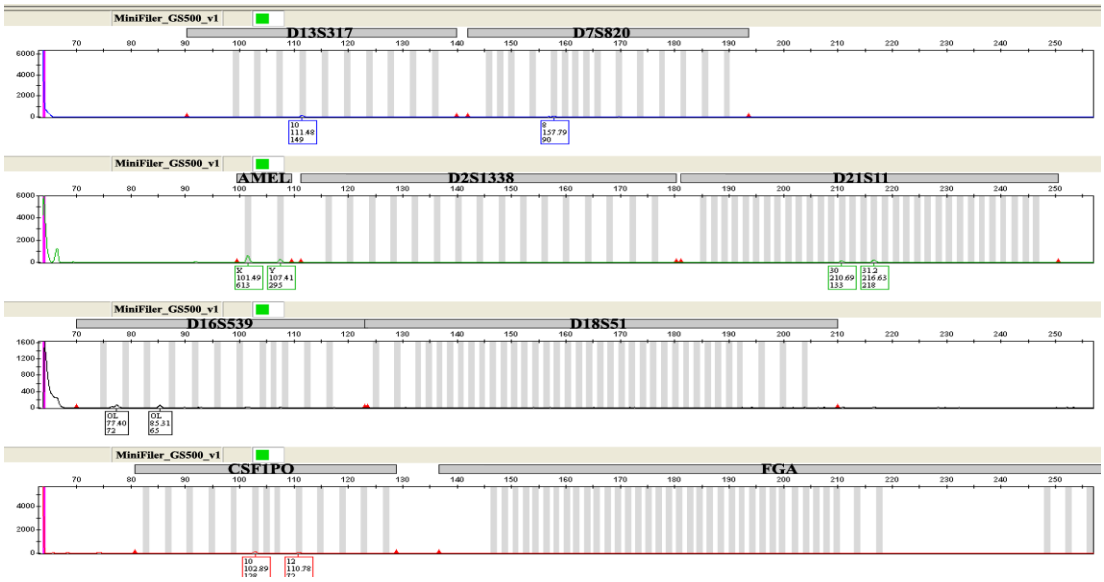

Figure 7-5 MiniFilerTM profile obtained from a fired cartridge case: The figure shows the Genetic Electrophoresis result to build a DNA profile from a spent case fired from the Kalashinkov, but only a few loci (*D13S17*, *AMEL* and *CSF1PO*) showed some peaks – of mostly poor quality

The rifle's shot shells yielded enough DNA quantities that were suitable for further processing. Moreover, with GeneMapper® Software (Applied Biosystem) the STR result was more likely to produce viable DNA fingerprints and in three of ten cases nearly full DNA profile (except *D21S11* locus which the span range is greater than 200 nt) were gained as shown in Figure 7-6.

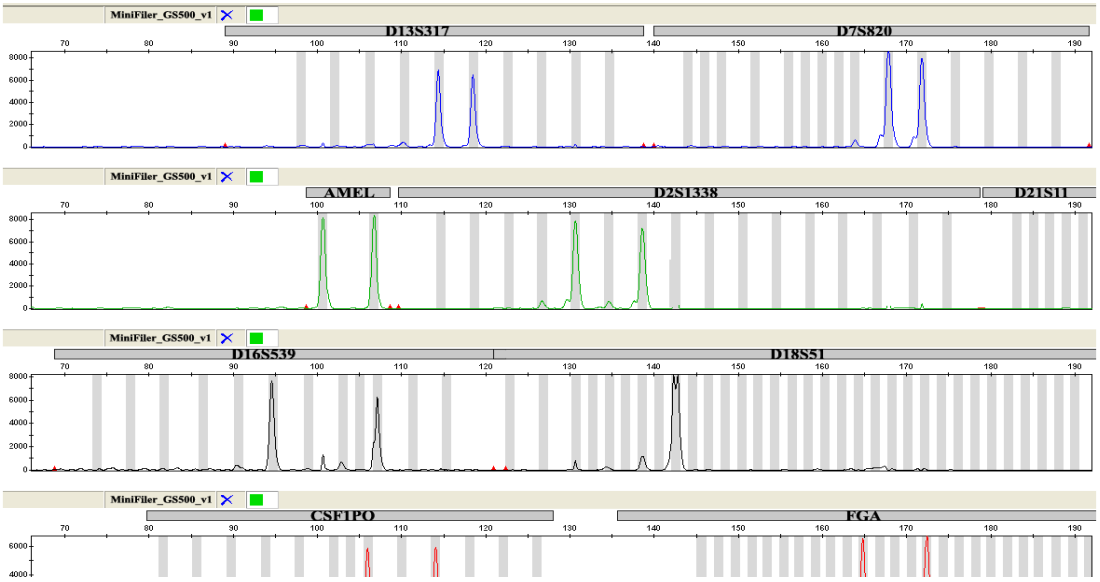

Figure 7-6 MiniFiler™ profile obtained from a fired Shot shell: The figure shows a DNA profile obtained from fired shot shell of Shotgun-Bore. In addition to sex determination locus (Amelogenin), the sample yielded 7 STR (short tendum repeat) loci (except D21S11 locus).

The DNA profile result generally showed a single donor and if there was a complex mixture (e.g. more than two peaks in one locus) the experiment was repeated. Furthermore, even if the shooter (with a known profile) could only be attributed to the mixture as a minor contributor of DNA then still the result was not accepted. Generally the DNA samples of the guns (except shotgun) did not yield sufficient DNA, or the profiles obtained were partial. Among the eight tested weapons, Kalashinkov was found to yield DNA quantities less than 200 pg amplification threshold and the mean was only 58 pg as none of the spent cartridge cases produced sufficient amount of DNA (Table 7-1).

Table 7-1 Statistical data of the fired guns: The table shows all the statistical details of the eight tested gun including average amount of retrieved DNA from spent cases and shot shells.

Guns	N	Mean	SE Mean	StDev	Minimum	Maximum	Median
Berta	10	118.7	42.8	135.2	1.6	346.1	57.2
Browning	10	107.5	42.9	135.6	0.7	405.0	37.0
Glock	10	133.9	25.7	81.1	19.5	297.1	138.1
Llama	10	196.2	53.4	168.7	8.0	495.0	175.0
Gunshot-Bore	10	871	201	637	230	2000	643
Gunshot	10	574	119	375	24	1100	500
Kalashnikov	10	58.0	19.4	61.3	6.5	180	27.8
Makarov	10	193.9	68.2	215.6	13.5	550	86.0

One of the reasons might be that this weapon is considered as a selective firearm which has both semi-automatic and automatic mode. Thus beside the high temperature there is a high pressure gas during shooting and this could affect the DNA sample quality and quantity in negative way. Retrieved DNA samples from the cartridge cases that discharged from the all pistols that were examined in this study (Beretta, Browning, Glock, Llama and Makarov) showed nearly similar result with small portion difference due to the size and material composition of the cases, but the retrieved DNA from any cartridges was not enough to build full eight STR loci. So, the results of this study support previous work by Polley *et al*. when they were not able to only build a partial DNA profile in the firearm crime field [28]. On the other hand, these findings of the current study are consistent with those of Rayan who found that DNA is not completely destroyed by the intense heat created during the firing process, although it has been estimated that internal temperatures of the firearm chambers can reach up to 1800 $^{\circ}$C for between 0.5 and 5 m [272]. It seems possible that the result of our study shows that bullets and cartridge casings (both fired and unfired) can routinely be examined in the DNA laboratory.

As there is a statistical difference among people in terms of DNA shedding (poor or good shedder), so might the volunteer in our experiment was not best shedder [273]. The cartridge case composition is another factor that can be takes in account which may due to variety amount of DNA as they made from different elements. Virtually all cartridge cases are made of brass (70% copper and 30% zinc). Some may also contain aluminium and few have a nickel coating [274]. So the surface or substrate type which contacted, the environmental factors and time contact have effect on the DNA yield. This is mean if the temperature was high like in summer then the loader hand will sweat and more cells are expected to transfer to the bullet surface [25]. Likewise as the time of contact increased there is more chance for cell exchange between two surface [266]. Moreover, bullets within a box or lot unfortunately, do not have uniform composition, but there may be distinct groups of bullets within a box, so the fired cartridges were not the same all the time [275], Therefore different cartridge surface might have an effect on amount of yielded DNA.

7.3.3 **Methylation Analysis of DNA Yielded from Cartridge**

There is no a database for the whole population, so just with DNA profile we can only identify people who have a previously generated profile. Therefore, the ability to drive some additional information about external feature like ages of the criminals becomes pivotal. Although prediction of age will not bring the person to justice directly, it could aid the police in undertaking investigations at a slightly more rapid pace. Using epigenetic signature differences to classify and distinguish age level is a current topic area of forensic interest. DNA methylation on cytosines is the best characterized among the epigenetic modifications which occurs through mammalian life time. The EpiTect Methyl II qPCR system was used. The system is based on the quantitative detection of remaining input DNA within a sample population after treatment with methyl-sensitive and methyl-dependent enzymes (described in previous chapter) [246, 247]. Primers are designed by an optimized computer algorithm to ensure that the amplicon contains cutting sites for both digestion types and are specifically designed for analysing the DNA methylation status of CpG islands. SYBR Green-based real time PCR detection is employed after digestion. To make the system work, each genomic DNA is subjected to four separate digestion treatments according to the manufacturer's protocol.

Methylation analysis of collected DNA on ammunition was performed using EpiTect qPCR assay. Unfortunately, no successful results were obtained from this procedure. A possible explanation may be due to the lack of adequate input DNA as the manual recommend using 0.5-1 µg, and not less than 150 ng of DNA, while in our study the maximum amount of retrieved DNA was from shot shell (1 ng) as was shown in (**Appendix 11**). A free Microsoft worksheet is available online to analyse the results of the EpiTect system. There are some criteria requirements as a control for any result with this free software. This Quality Control (QC) report calculates analytical window (W) and the percentage of DNA refractory (R) for each input DNA sample as illustrated Figure 7-7. If the analytical window is less than three ($W < 3$) meaning that the refractory DNA percentage is greater than 12.5 percent ($R > 12.5\%$), then the digestions are not complete and the analysis is reported as a "Failure" [175].

	A	B	C	D	E	F
1	Samples		SHOTGUN (12bore		SHOTGUN	
2	Gene Symbol	Catalog Number	W	R	W	R
3	BCAS4-1	EPHS109350-1A	2.74	Failure	-4.45	Failure
4	KCNQ1DN	EPHS102112-1A	-4.81	Failure	-0.81	Failure
5	GRIA2	EPHS111325-1A	-0.93	Failure	2.71	Failure
6	TRIM58	EPHS101433-1A	1.6	Failure	1.4	Failure
7	SEC	EPHS115450-1A	11.78	0.028%	10.62	0.064%
8	DEC	EPHS115451-1A	5.58	2.091%	10.02	0.096%

Instructions | Raw Data | QC Data Report | Results | Calculations

Figure 7-7 The QC data report worksheet for EpiTect analysis of trace DNA recovered from fired gun cartridges: The figure shows the result of the quality control report and the value in analytical window (W) was abnormal (i.e. less than three) for all the age related genes. Similarly the refractory factor (R) of digestion was not complete for the input DNA means that the digestion efficiency was not high.

When the sample result passes the QC report, then the template automatically calculates the methylation status of each gene of each sample as the percentage of methylated (M) and unmethylated (UM) DNA in the "Results" worksheet. The “Result” improve enzyme digestion`s efficiency of the test. If the result was “Pass” for the control primer sets (SEC and DEC) which monitor both sensitive and dependent enzymes respectively, the result would be acceptable. That means real time PCR quantification and digestion steps were done successfully. In contrast, the result worksheet for input DNA after digestion reported the result as failure in Figure 7-8. Therefore the result was not acceptable due to the amount and quality of input DNA of the DNA sample.

	A	B	C	D	E	F
1		Samples	SHOTGUN (12bore)		SHOTGUN	
2	Gene Symbol	Catalog Number	UM	M	UM	M
3	BCAS4-1	EPHS109350-1A	Failure	Failure	Failure	Failure
4	KCNQ1DN	EPHS102112-1A	Failure	Failure	Failure	Failure
5	GRIA2	EPHS111325-1A	Failure	Failure	Failure	Failure
6	TRIM58	EPHS101433-1A	Failure	Failure	Failure	Failure
7	SEC	EPHS115450-1A	Pass		Pass	
8	DEC	EPHS115451-1A	Pass		Pass	

Instructions | Raw Data | QC Data Report | Results | Calculations

Figure 7-8: The results worksheet for EpiTect analysis of trace DNA recovered from fired gun cartridges: The figure illustrates the result worksheet of excel template which failed to show methylated (M) and unmethylated (UM) percentage of input DNA.

Logically, skin cells and saliva are the most probable source of DNA in a firearm cases and in a very few crime scenes blood can be find from fired casings, bullets and cartridges. But in our experiment the majority DNA source comes from skin cells and this is another reason why EpiTect qPCR system did not give result as the tested gene assays are optimized for age prediction in other types of tissues like blood and saliva, rather than skin cells. Finally, there are numerous studies about DNA recovery from various locations of firearms, for example gun grip, gun trigger, magazine surface, ejected cartridge cases and non-spent cartridge or ammunition in general [24, 28, 29]. But still few laboratories report successful DNA typing with

spent cartridge. These results of our study differ from some published reports [276], but they are consistent with those of others [28, 29].

The large amount of DNA which was required (120 ng) for an age study can be considered challenge to face and a limitation to applying this approach in real casework. In this study, only one volunteer loaded all the weapons. Hence, we don't know if the loader was a poor or good shedder. In this case, repeating the experiment with different volunteers, and across the day, will be a good idea as there is always inter- and intra-individual variation. This would enable us to be more confident about the success rate of DNA retrieval. Future studies need to develop the EpiTect qPCR assay with more appropriate application method and investigating into ways of decreasing amount of input DNA. Furthermore, a good epigenetic ageing signature at a single gene that provides an accurate biomarker to predict the state of age needs to be found, rather than the current multiple gene age signature process. This would remove the need to divide the amount of recovered DNA into six equal parts (four genes and two controls) as required in the assay.

7.4 **Conclusion**

DNA analysis has been widely accepted as providing valuable evidence concerning the identity of the source of biological traces. Our work has showed that DNA samples can survive on cartridges even after firing. The study also raised the possibility of determining other information such as the age of the donor. Such information may be invaluable in certain cases where spent cartridges from automatic weapons are left behind at the scene of a crime. In spite of the nature of touch evidence and exposure to high chamber temperatures during shooting, but we were still capable to retrieve enough DNA for profile typing in some instances. In terms of using EpiTect system for retrieved DNA, results were not possible due to the low amount of input DNA.

CHAPTER EIGHT
General Discussion

8 Conclusion and Recommendation for Future Work

8.1 General Results and Discussion

Methylation analysis presents a unique opportunity to answer a wide range of as yet unanswered questions in ageing. This thesis generally has focused on two related research themes, age prediction of donor for forensic investigation and control of telomere shortening for biomedical purpose. In CHAPTER THREE we studied and compared two novel swabs (forensiX) with commercial cotton swab and showed their advantage for collection more amount of DNA. The outcome of this study will be useful for low copy number DNA sample (i.e. trace DNA) and also highlighted the importance of drying process for the best preservation.

The next chapter (CHAPTER FOUR) studied tight correlation between methylation status and human age at autosomal loci as a new forensic profiling tool. Thus we used the Illumina MiSeq next generation sequencing platform to investigate promoter regions of 23 genes in total. Methylation levels at three CpGs located in *ASPA, ITGA2B* and *PDE4C* genes showed an epigenetic signature of aging with a mean absolute deviation of only 6 years from chronological age an entirely independent confirmation of an earlier study that identified these loci. The implementation of the methylation pattern of these specific autosomal gene sequences from biological evidence left at crime scene may help build up an accurate picture of an offender's age.

In CHAPTER FIVE, a quantification of global mitochondrial DNA methylation levels and inverse correlation with age was studied. To the author's knowledge, there are no other published works using mitochondrial methylation patterns at exact CpG sites as was achieved here, and using only two CpG (1215 and 1313) dinucleotides to estimate the chronological age of the donor. As such, these outcomes are valuable to researchers not only in forensic science but also those studying mitochondria in their research processing workflow in a clinical setting. The methylation of mitochondrial DNA (mtDNA) is a new and incompletely described phenomenon with unknown biological control and significance.

CHAPTER SIX explored the use of EpiTect Methyl qPCR Assay as novel age estimation tool in forensic biology. The system was used to compare methylation

levels of CpG Island of the promoter regions of 4 age related genes (*NPTX2, KCNQ1DN, GRIA2* and *TRIM58*). The study examined 80 female samples of various ages (18-91 years) obtained from blood, using designed primers that flanked the studied gene loci. The data obtained from DNA quantification showed successful discrimination among volunteered ages for throughput DNA methylation. Overall, the difference between predicted and real age was about 11 years and absolute mean differences (AMD) was only 7.2 years error. The outcome of this chapter showed that EpiTect system can be used as fast and simple innovative predictor tool in future forensic age estimation.

In the final chapter (CHAPTER SEVEN), experiments provided valuable evidence concerning the identity of the source of biological traces. This work has showed that DNA samples can survive on cartridges even after firing. The study also tried to determine other information such as the age of the donor. Such information may be invaluable in certain cases where spent cartridges from automatic weapons are left behind at the scene of a crime. In spite of the nature of touch evidence and exposure to high chamber temperatures during shooting, we were still able to retrieve enough DNA for profile typing. In order to estimate age of contributor, DNA methylation levels were analysed using EpiTect system for retrieved DNA. However, results were negative, due to low amount of input DNA.

8.2 **Conclusions and Recommendations for Future Work**

8.2.1 **Further Studies of Epigenetic Changes with Ageing**

Additional experiments will be necessary to fully elucidate the role of biomarker role of methylation in epigenetics. The first is the study of DNA hydroxymethylcytosine (hmC) as another epigenetic marker, and the second is the new understanding that RNA methylation of base A at the N6 position (m6A) appears to be most prevalent epigenetic mark in eukaryotic mRNA [277], which is a very interesting field and needs to be taken in account for future study for biology and forensic studies. It is necessary to focus and study in future the ratio between DNA mC and hmC, so this gap in our knowledge will be addressed in future work, especially considering whether inhibition of one of the methylation stages produces damage that affects the

whole methylation phenomenon or not. Monitoring the methylation profile longitudinally, from young to aged people is necessary to understand epigenetics phenomenon and how lifestyle can influence change. The study of identical twins will be another future goal to see if their epigenetic ages match the chronological age.

Finally, I have identified a number of specific genes that seem to have methylation profiles that reflect ageing – what are the functions of those genes in our body, and is there a good reason why they should be part of a biological clock?

8.2.2 **Putting Age Prediction into Practice**

Today, the cost and time consumed for forensic investigation is considered very important. Can the very detailed analysis of methylation possible in the academic setting, with next generation sequencing, ever be translated into something with practical forensic use? Our findings suggest that it can be used in future as a forensic day-to-day tool for age prediction. The use of the EpiTect system makes analysis very cheap and faster than other advanced techniques like Pyrosequencing or Illumina NGS. The second major task is to validate the test by using more samples, different type of tissues (or even different type of blood cells) to find universal CpG sites in human body will be good idea to be studied. Further tissue identification is very important for sexual assault crime. The opposite sex (male) and other population around the world must be studied as they have different life style and environment. Although the percentage of crime by elder people is lower than young and middle age, more investigation is required to answer why the selected methylation markers are not accurate for age prediction in this group compared to younger age groups. In terms of forensic investigation our result may not lead to the contributor directly but it aids the investigators to know how old s/he was. Importantly, this ‘epi-clock’ technology also has the potential to exclude the innocent from potential miscarriages of justice.

References

1. Jobling, M.A. and P. Gill, *Encoded evidence: DNA in Forensic Analysis.* Nature Reviews Genetics, 2004. **5**(10): p. 739-751.
2. Butler, J.M. and J.M. Butler, *Fundamentals of Forensic DNA typing*. 2010, Amsterdam; Boston: Academic Press/Elsevier.
3. Jobling, M., M. Hurles, and C. Tyler-Smith, *Human evolutionary genetics: origins, peoples & disease*. 2013: Garland Science.
4. Fiuzat, Y. and R. Atkins, *Genetic and environmental variability in segregating barley populations.* Agronomy Journal, 1953. **45**(9): p. 414-420.
5. Hartl, D.L. and A.G. Clark, *Principles of Population Genetics*. Vol. 116. 1997: Sinauer associates Sunderland.
6. Grant, D.M., et al., *Monomorphic and polymorphic human arylamine N-acetyltransferases: a comparison of liver isozymes and expressed products of two cloned genes.* Molecular pharmacology, 1991. **39**(2): p. 184-191.
7. Hedrick, P.W., *Perspective: highly variable loci and their interpretation in evolution and conservation.* Evolution, 1999: p. 313-318.
8. Butler, J.M. *Forensic DNA typing biology, technology, and genetics of STR markers*. 2005; Available from: http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=187185.
9. Giles, R.E., et al., *Maternal inheritance of human mitochondrial DNA.* Proceedings of the National academy of Sciences, 1980. **77**(11): p. 6715-6719.
10. Robin, E.D. and R. Wong, *Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells.* Journal of Cellular Physiology, 2005. **136**(3): p. 507-513.
11. Zhang, S.-x., *An atlas of histology*. 1999, New York: Springer.
12. Lawrence, E., *Henderson's dictionary of biological terms* P.E. Limited, Editor. 2000: Essex.
13. Burkitt H G, Y.B., Heath JW and Deakin PJ *Wheater s Functional Histology: A Text and Colour Atlas* 3rd Eddition ed. 1993, Churchil Livingston, UK.
14. Wheater, P.R., et al., *Wheater's interactive histology*. 1995, Churchill Livingstone: San Mateo, CA.
15. Wickenheiser, R.A., *Trace DNA: a review, discussion of theory, and application of the transfer of trace quantities of DNA through skin contact.* Journal of Forensic Sciences, 2002. **47**(3): p. 442.
16. Raven, P.H. and G.B. Johnson, *Biology*. 6th ed. 2002, Boston: McGraw-Hill.
17. Lowe, A., et al., *The propensity of individuals to deposit DNA and secondary transfer of low level DNA from individuals to inert surfaces.* Forensic Science International, 2002. **129**(1): p. 25-34.
18. Van Hoofstat, D., et al. *DNA typing of fingerprints and skin debris: sensitivity of capillary electrophoresis in forensic applications using multiplex PCR*. in *Second European Symposium on Human Identification*. 1998.
19. Saito, Y., et al., *DNA fingerprinting from single cells.* Nature, 1997. **389**: p. 555.
20. Gill, P., *Application of low copy number DNA profiling.* Croatian Medical Journal, 2001. **42**(3): p. 229-232.
21. Williamson, A.L., *Touch DNA: Forensic Collection and Application to Investigations.* J Assoc Crime Scene Reconstr, 2012. **18**(1).

22. French, C.E.V., et al., *A novel histological technique for distinguishing between epithelial cells in forensic casework.* Forensic Science International, 2008. **178**(1): p. 1-6.
23. Schiffner, L.A., et al., *Optimization of a simple, automatable extraction method to recover sufficient DNA from low copy number DNA samples for generation of short tandem repeat profiles.* Croatian Medical Journal, 2005. **46**(4): p. 578.
24. Nunn, S., *Touch DNA Collection Versus Firearm Fingerprinting: Comparing Evidence Production and Identification Outcomes.* Journal of forensic sciences, 2013.
25. Quinones, I. and B. Daniel, *Cell free DNA as a component of forensic evidence recovered from touched surfaces.* Forensic Science International: Genetics, 2012. **6**(1): p. 26-30.
26. Company, T.M.N. *http://www.google.co.uk/image/* 2010 [cited 20133 January, 25].
27. Esslinger, K.J., et al., *Using STR analysis to detect human DNA from exploded pipe bomb devices.* Journal of Forensic Sciences, 2004. **49**(3): p. 481-484.
28. Polley, D., et al., *An investigation of DNA recovery from firearms and cartridge cases.* JOURNAL-CANADIAN SOCIETY OF FORENSIC SCIENCE, 2006. **39**(4): p. 217.
29. Horsman-Hall, K.M., et al., *Development of STR profiles from firearms and fired cartridge cases.* Forensic Science International: Genetics, 2009. **3**(4): p. 242-250.
30. Dieltjes, P., et al., *A sensitive method to extract DNA from biological traces present on ammunition for the purpose of genetic profiling.* International Journal of Legal Medicine, 2011. **125**(4): p. 597-602.
31. Brouwer, D.H., M.F. Boeniger, and J. van Hemmen, *Hand wash and manual skin wipes.* Annals of Occupational Hygiene, 2000. **44**(7): p. 501-510.
32. Van Oorschot, R.A. and M.K. Jones, *DNA fingerprints from fingerprints.* NATURE-LONDON-, 1997: p. 767-767.
33. Phipps, M. and S. Petricevic, *The tendency of individuals to transfer DNA to handled items.* Forensic Science International, 2007. **168**(2): p. 162-168.
34. Rutty, G., *An investigation into the transference and survivability of human DNA following simulated manual strangulation with consideration of the problem of third party contamination.* International Journal of Legal Medicine, 2002. **116**(3): p. 170-173.
35. Raymond, J.J., et al., *Trace DNA success rates relating to volume crime offences.* Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 136-137.
36. Daly, D.J., C. Murphy, and S.D. McDermott, *The transfer of touch DNA from hands to glass, fabric and wood.* Forensic Science International: Genetics, 2012. **6**(1): p. 41-46.
37. Bright, J.-A. and S.F. Petricevic, *Recovery of trace DNA and its application to DNA profiling of shoe insoles.* Forensic Science International, 2004. **145**(1): p. 7-12.
38. Alessandrini, F., et al., *Fingerprints as evidence for a genetic profile: morphological study on fingerprints and analysis of exogenous and individual factors affecting DNA typing.* Journal of Forensic Sciences, 2003. **48**(3): p. 586-592.
39. Goray, M., R.J. Mitchell, and R.A.v. Oorschot, *Investigation of secondary DNA transfer of skin cells under controlled test conditions.* Legal Medicine, 2010. **12**(3): p. 117-120.
40. Hewakapuge, S., R. Van Oorschot, and S. Baindur-Hudson, *Evaluation of mtDNA mutations to predict age.* Forensic Science International: Genetics Supplement Series, 2008. **1**(1): p. 561-562.
41. Kayser, M. and P. de Knijff, *Improving human forensics through advances in genetics, genomics and molecular biology.* Nat Rev Genet, 2011. **12**(3): p. 179-192.

42. De Wilde, J., et al., *hTERT promoter activity and CpG methylation in HPV-induced carcinogenesis.* BMC Cancer, 2010. **10**(1): p. 271.
43. Speybroeck, L., *From epigenesis to epigenetics.* Annals of the New York Academy of Sciences, 2002. **981**(1): p. 61-81.
44. Rando, T.A. and H.Y. Chang, *Aging, rejuvenation, and epigenetic reprogramming: resetting the aging clock.* Cell, 2012. **148**(1-2): p. 46-57.
45. Seisenberger, S., et al., *Reprogramming DNA methylation in the mammalian life cycle: building and breaking epigenetic barriers.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2013. **368**(1609).
46. Clark, S.J., J. Harrison, and M. Frommer, *CpNpG methylation in mammalian cells.* Nature Genetics, 1995. **10**(1): p. 20-27.
47. Jeltsch, A., *Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases.* Chembiochem, 2002. **3**(4): p. 274-293.
48. Holliday, R. and G. Grigg, *DNA methylation and mutation.* Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 1993. **285**(1): p. 61-67.
49. Razin, A. and A.D. Riggs, *DNA methylation and gene function.* Science (New York, NY), 1980. **210**(4470): p. 604.
50. Armstrong, L., *Epigenetics*. 2013: Garland Science.
51. Kerwin, S.M., *ChemBioOffice Ultra 2010 suite.* Journal of the American Chemical Society, 2010. **132**(7): p. 2466-2467.
52. Saeed, M., R.M. Berlin, and T.D. Cruz, *Exploring the utility of genetic markers for predicting biological age.* Legal Medicine, 2012. **14**: p. 279-285.
53. Juusola, J. and J. Ballantyne, *Multiplex mRNA profiling for the identification of body fluids.* Forensic Sci Int, 2005. **152**(1): p. 1-12.
54. Madi, T., et al., *The determination of tissue-specific DNA methylation patterns in forensic biofluids using bisulfite modification and pyrosequencing.* Electrophoresis, 2012. **33**(12): p. 1736-1745.
55. Lee, H.Y., et al., *Potential forensic application of DNA methylation profiling to body fluid identification.* International Journal of Legal Medicine, 2012. **126**(1): p. 55-62.
56. Lee, H.Y., et al., *Potential forensic application of DNA methylation profiling to body fluid identification.* 2012. **126**(1).
57. An, J., et al., *DNA methylation-specific multiplex assays for body fluid identification.* 2011. **127**(1).
58. Herman, J.G. and S.B. Baylin, *Gene silencing in cancer in association with promoter hypermethylation.* New England Journal of Medicine, 2003. **349**(21): p. 2042-2054.
59. Das, P.M. and R. Singal, *DNA methylation and cancer.* Journal of Clinical Oncology, 2004. **22**(22): p. 4632-4642.
60. Fragou, D., et al., *Epigenetic mechanisms in metal toxicity.* Toxicology Mechanisms and Methods, 2011. **21**(4): p. 343-352.
61. Frumkin, D., et al., *DNA methylation-based forensic tissue identification.* Forensic Science International: Genetics, 2011. **5**(5): p. 517-524.
62. Bocklandt S, L.W., Sehl ME, Sanchez FJ, Sinsheimer JS, Steve Horvath and Eric Vilain, *Epigenetic Predictor of Age* PLoS ONE, 2011. **6**(6).
63. Naito, E., et al., *Sex determination using the hypomethylation of a human macro-satellite DXZ4 in female cells.* Nucleic Acids Research, 1993. **21**(10): p. 2533-2534.

64. Lindenbergh, A., et al., *A multiplex (m) RNA-profiling system for the forensic identification of body fluids and contact traces.* Forensic Science International: Genetics, 2012. **6**(5): p. 565-577.
65. Wang, Z., et al., *A model for data analysis of microRNA expression in forensic body fluid identification.* Forensic Science International: Genetics, 2012. **6**(3): p. 419-423.
66. Wang, Z., et al., *Screening and confirmation of microRNA markers for forensic body fluid identification.* Forensic Science International: Genetics, 2013. **7**(1): p. 116-123.
67. Gomes, I., F. Kohlmeier, and P. Schneider, *Genetic markers for body fluid and tissue identification in forensics.* Forensic Science International: Genetics Supplement Series, 2011. **3**(1): p. e469-e470.
68. Zubakov, D., et al., *Estimating human age from T-cell DNA rearrangements.* Current Biology, 2010. **20**(22): p. R970-R971.
69. Ou, X., et al., *Detection and quantification of the age-related sjTREC decline in human peripheral blood.* International Journal of Legal Medicine, 2011. **125**(4): p. 603-608.
70. Heyn, H., et al., *Distinct DNA methylomes of newborns and centenarians.* Proceedings of the National Academy of Sciences, 2012. **109**(26): p. 10522-10527.
71. Sahin, K., et al., *DNA methylation analyses of monozygotic twins.* Current Opinion in Biotechnology, 2011. **22**: p. S105.
72. Li, C., et al., *Identical but not the same: The value of DNA methylation profiling in forensic discrimination within monozygotic twins.* Forensic Science International: Genetics Supplement Series, 2011. **3**(1): p. e337-e338.
73. Bocklandt, S., et al., *Epigenetic predictor of age.* PLoS One, 2011. **6**(6): p. e14821.
74. Koch, C.M. and W. Wagner, *Epigenetic-aging-signature to determine age in different tissues.* Aging, 2011. **3**(10): p. 1018-1027.
75. Zbieć-Piekarska, R., et al., *Examination of DNA methylation status of the ELOVL2 marker may be useful for human age prediction in forensic science.* Forensic Science International: Genetics, 2015. **14**(0): p. 161-167.
76. Kanungo, M.S., *Genes and aging*. 2005: Cambridge University Press.
77. Beers, M.H., T.V. Jones, and M. Berkwits, *The Merck manual of health & aging*. 2005: Random House LLC.
78. Brierley, E.J., et al., *Mitochondrial involvement in the ageing process. Facts and controversies.* Molecular and cellular biochemistry, 1997. **174**(1-2): p. 325-328.
79. Olovnikov, A.M., *Telomeres, telomerase, and aging: origin of the theory.* Experimental gerontology, 1996. **31**(4): p. 443-448.
80. Orgel, L.E., *The maintenance of the accuracy of protein synthesis and its relevance to ageing.* Proceedings of the National Academy of Sciences of the United States of America, 1963. **49**(4): p. 517.
81. Cutler, R.G., *Urate and ascorbate: their possible roles as antioxidants in determining longevity of mammalian species.* Archives of gerontology and geriatrics, 1984. **3**(4): p. 321-348.
82. Kanungo, M. and B.S. Gandhi, *Induction of malate dehydrogenase isoenzymes in livers of young and old rats.* Proceedings of the National Academy of Sciences, 1972. **69**(8): p. 2035-2038.
83. Kaye, D.H. and M.E. Smith, *DNA identification databases: legality, legitimacy, and the case for population-wide coverage.* Wis. L. Rev., 2003: p. 413.
84. Koops, B.-J. and M. Schellekens, *Forensic DNA phenotyping: regulatory issues.* Colum. Sci. & Tech. L. Rev., 2008. **9**: p. 158, 160 & n. 4.

85. Kayser, M. and P.M. Schneider, *DNA-based prediction of human externally visible characteristics in forensics: Motivations, scientific challenges, and ethical considerations.* Forensic Science International: Genetics, 2009. **3**(3): p. 154-161.
86. Allen, H.L., et al., *Hundreds of variants clustered in genomic loci and biological pathways affect human height.* Nature, 2010. **467**(7317): p. 832-838.
87. Frudakis, T., *Molecular photofitting: predicting ancestry and phenotype using DNA*. 2010: Elsevier.
88. Graham, E.A., *DNA reviews: predicting phenotype.* Forensic science, medicine, and pathology, 2008. **4**(3): p. 196-199.
89. HE, W., et al., *Genetic structure of Europeans: a view from the North-East.* PLoS One, 2009. **4**: p.
90. Wambaugh, J., *The blooding*. 2011: Open Road Media.
91. Friedman, A.L., *Forensic DNA profiling in the 21st century.* International Journal of Offender Therapy and Comparative Criminology, 1999. **43**(2): p. 168-179.
92. Wambough, J., *The Blooding: True Story of the Narborough Village Murders*. 1989, London: Bantam Press.
93. Schmeling, A., et al., *Forensic age diagnostics of living individuals in criminal proceedings.* HOMO-Journal of Comparative Human Biology, 2003. **54**(2): p. 162-169.
94. Garamendi, P., et al., *Reliability of the methods applied to assess age minority in living subjects around 18 years old: a survey on a Moroccan origin population.* Forensic science international, 2005. **154**(1): p. 3-12.
95. Kumar, R. and M.L. Best, *Impact and sustainability of e-government services in developing countries: Lessons learned from Tamil Nadu, India.* The Information Society, 2006. **22**(1): p. 1-12.
96. Alvarez, M. and J. Ballantyne, *The identification of newborns using messenger RNA profiling analysis.* Analytical biochemistry, 2006. **357**(1): p. 21-34.
97. Tsuji, A., et al., *Estimating age of humans based on telomere shortening.* Forensic Science International, 2002. **126**(3): p. 197-199.
98. von Wurmb-Schwark, N., et al., *Quantification of human mitochondrial DNA in a real time PCR.* Forensic science international, 2002. **126**(1): p. 34-39.
99. Panchbhai, A., *Dental radiographic indicators, a key to age estimation.* 2011.
100. Meissner, C., et al., *Estimation of age at death based on quantitation of the 4977-bp deletion of human mitochondrial DNA in skeletal muscle.* Forensic science international, 1999. **105**(2): p. 115-124.
101. Meissner, C., P. Bruse, and M. Oehmichen, *Tissue-specific deletion patterns of the mitochondrial genome with advancing age.* Experimental gerontology, 2006. **41**(5): p. 518-524.
102. Takasaki, T., et al., *Age estimation in dental pulp DNA based on human telomere shortening.* International journal of legal medicine, 2003. **117**(4): p. 232-234.
103. Meindl, R.S. and C.O. Lovejoy, *Ectocranial suture closure: A revised method for the determination of skeletal age at death based on the lateral-anterior sutures.* American journal of physical anthropology, 1985. **68**(1): p. 57-66.
104. Kerley, E.R., *The microscopic determination of age in human bone.* American Journal of Physical Anthropology, 1965. **23**(2): p. 149-163.
105. Meissner, C., *Mutations of mitochondrial DNA–cause or consequence of the ageing process?* Zeitschrift für Gerontologie und Geriatrie, 2007. **40**(5): p. 325-333.
106. Moraes, C.T., et al., *Mitochondrial DNA structure and function.* International review of neurobiology, 2002. **53**: p. 3-23.

107. Alberts, B., et al., *Molecular biology of the cell 4th edition*. 2002: National Center for Biotechnology InformationÕs Bookshelf.
108. Kang, Y.-H., et al., *Caspase-Independent Cell Death by Arsenic Trioxide in Human Cervical Cancer Cells Reactive Oxygen Species-Mediated Poly (ADP-ribose) Polymerase-1 Activation Signals Apoptosis-Inducing Factor Release from Mitochondria.* Cancer research, 2004. **64**(24): p. 8960-8967.
109. Kenneth Day, L.L.W., Anna Thalacker-Mercer, Andrew West, Marcas M Bamman, and R.M.M.a.D.A. James D Brooks, *Differential DNA methylation with age displays both common and dynamic features across human tissues that are influenced by CpG landscape.* Genome Biology, 2013. **2013**(14:R102).
110. Moran, M.E., *Evolution of Stone Disease*, in *Urolithiasis*. 2014, Springer. p. 85-99.
111. Lea, P.J., et al., *Variations in mitochondrial ultrastructure and dynamics observed by high resolution scanning electron microscopy (HRSEM).* Microscopy research and technique, 1994. **27**(4): p. 269-277.
112. Ernster, L. and G. Schatz, *Mitochondria: a historical review.* The Journal of cell biology, 1981. **91**(3): p. 227s-255s.
113. Singh, K.K., *Mitochondrial DNA mutations in aging, disease, and cancer*. 1998, Berlin; New York: Springer.
114. Moraes, C.T., et al., *Mitochondrial DNA structure and function.* International review of neurobiology, 2002. **53**: p. 3-23.
115. Larsson, N. and D.A. Clayton, *Molecular genetic aspects of human mitochondrial disorders.* Annual review of genetics, 1995. **29**(1): p. 151-178.
116. Butler, J.M. and B.C. Levin, *Forensic applications of mitochondrial DNA.* Trends in Biotechnology, 1998. **16**(4): p. 158-162.
117. Lodish, H., *Molecular cell biology*. 2013, New York, NY: Freeman.
118. Iacobazzi, V., et al., *Mitochondrial DNA methylation as a next-generation biomarker and diagnostic tool.* Molecular Genetics and Metabolism, 2013. **110**(1–2): p. 25-34.
119. Anderson, S., et al., *Sequence and organization of the human mitochondrial genome.* Nature 1981. **290**: p. 457-465
120. Andrews, R.M., et al., *Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA.* Nature genetics, 1999. **23**(2): p. 147-147.
121. Dudek, R.W., *High-yield cell and molecular biology*. 2007: Lippincott Williams & Wilkins.
122. Scheffler, I.E., *Mitochondria* 2007, Hoboken, New jersey: Wiley-Liss.
123. Ozawa, T., *Genetic and functional changes in mitochondria associated with aging.* Physiological reviews, 1997. **77**(2): p. 425-464.
124. de Souza-Pinto, N.C. and V.A. Bohr, *The mitochondrial theory of aging: involvement of mitochondrial DNA damage and repair.* International Review of Neurobiology, 2002. **53**: p. 519-534.
125. Linnane, A., et al., *Mitochondrial DNA mutations as an important contributor to ageing and degenerative diseases.* The Lancet, 1989. **333**(8639): p. 642-645.
126. Kenyon, C., *The plasticity of aging: insights from long-lived mutants.* Cell, 2005. **120**(4): p. 449-460.
127. Lesk, A., *Introduction to genomics*. 2012: Oxford University Press.
128. Wilson, A.C., et al., *Mitochondrial DNA and two perspectives on evolutionary genetics.* Biological Journal of the Linnean Society, 1985. **26**(4): p. 375-400.

129. Kosinski, R.J., D.R. Weinbrenner, and M.G. Cross. *Extraction, Sequencing, and Analysis of Mitochondrial DNA*. in *Tested Studies for Laboratory Teaching*. 2007. University of Kentucky Lexington, Kentucky: Proceeding of the 29th Workshope/Conference of the Association for Biology Laboratory Education (ABLE).
130. BECKMAN, K.B. and B.N. AMES, *The Free Radical Theory of Aging Matures.* PHYSIOLOGICAL REVIEWS, 1998. **78**(2).
131. Birky Jr, C.W., *The inheritance of genes in mitochondria and chloroplasts.* BioScience, 1976: p. 26-33.
132. Ankel-Simons, F. and J.M. Cummins, *Misconceptions about mitochondria and mammalian fertilization: implications for theories on human evolution.* Proceedings of the National Academy of Sciences, 1996. **93**(24): p. 13859-13863.
133. Holland, M.M., et al., *Mitochondrial DNA sequence analysis of human skeletal remains: identification of remains from the Vietnam War.* Journal of forensic sciences, 1993. **38**: p. 542-542.
134. Prinz, M., et al., *DNA Commission of the International Society for Forensic Genetics (ISFG): recommendations regarding the role of forensic genetics for disaster victim identification (DVI).* Forensic Science International: Genetics, 2007. **1**(1): p. 3-12.
135. Foran, D.R., M.E. Gehring, and S.E. Stallworth, *The Recovery and Analysis of Mitochondrial DNA from Exploded Pipe Bombs.* J Forensic Sci, 2008. **54**(1).
136. Trifunovic, A., *Mitochondrial DNA and ageing.* Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2006. **1757**(5–6): p. 611-617.
137. Hudson, G. and P.F. Chinnery, *Mitochondrial DNA polymerase-γ and human disease.* Human molecular genetics, 2006. **15**(suppl 2): p. R244-R252.
138. Villa, S.T., *Environmental Stress Response of Protein Repair L-isoaspartyl Methyltransferase and Its Physiological Role in Repairing Age-damaged Proteins in Higher Plants*. 2007: ProQuest.
139. 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.
140. Kaguni, L.S., *DNA polymerase γ, the mitochondrial replicase 1.* Annual review of biochemistry, 2004. **73**(1): p. 293-320.
141. Krishnan, K.J., et al., *What causes mitochondrial DNA deletions in human cells?* Nat Genet, 2008. **40**(3): p. 275-279.
142. Khrapko, K. and J. Vijg, *Mitochondrial DNA mutations and aging: devils in the details?* Trends in Genetics, 2009. **25**(2): p. 91-98.
143. Nagley, P. and Y.-H. Wei, *Ageing and mammalian mitochondrial genetics.* Trends in Genetics, 1998. **14**(12): p. 513-517.
144. Attardi, G., *Role of mitochondrial DNA in human aging.* MITOCHONDRION - AMSTERDAM-, 2002. **2**(1-2): p. 27-37.
145. Cortopassi, G.A. and N. Arnheim, *Detection of a specific mitochondrial DNA deletion in tissues of older humans.* Nucleic Acids Research, 1990. **18**(23): p. 6927-6933.
146. Pak, J.W., et al., *Mitochondrial DNA mutations as a fundamental mechanism in physiological declines associated with aging.* Aging cell, 2003. **2**(1): p. 1-7.
147. Kovalenko, S.A., et al., *Tissue-specific Distribution of Multiple Mitochondrial DNA Rearrangements during Human Aginga.* Annals of the New York Academy of Sciences, 1998. **854**(1): p. 171-181.
148. Kopsidas, G., et al., *An age-associated correlation between cellular bioenergy decline and mtDNA rearrangements in human skeletal muscle.* Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 1998. **421**(1): p. 27-36.

149. Melov, S., et al., *Marked increase in the number and variety of mitochondrial DNA rearrangements in aging human skeletal muscle.* Nucleic acids research, 1995. **23**(20): p. 4122-4126.
150. Vanyushin, B., et al., *The 5-methylcytosine in DNA of rats.* Gerontology, 1973. **19**(3): p. 138-152.
151. Wilson, V.L. and P.A. Jones, *DNA methylation decreases in aging but not in immortal cells.* Science, 1983. **220**(4601): p. 1055-1057.
152. Holliday, R., *The significance of DNA methylation in cellular aging*, in *Molecular biology of aging*. 1985, Springer. p. 269-283.
153. Richardson, B., *Impact of aging on DNA methylation.* Ageing research reviews, 2003. **2**(3): p. 245-261.
154. Vanyushin, B., *Epigenetics today and tomorrow.* Russian Journal of Genetics: Applied Research, 2014. **4**(3): p. 168-188.
155. Horvath, S., *DNA methylation age of human tissues and cell types.* Genome biology, 2013. **14**(10): p. R115.
156. Issa, J.-P.J., et al., *Accelerated age-related CpG island methylation in ulcerative colitis.* Cancer research, 2001. **61**(9): p. 3573-3577.
157. Kang, G.H., et al., *Aberrant CpG island hypermethylation of chronic gastritis, in relation to aging, gender, intestinal metaplasia, and chronic inflammation.* The American journal of pathology, 2003. **163**(4): p. 1551-1556.
158. Franceschi, C., et al., *Inflamm-aging: an evolutionary perspective on immunosenescence.* Annals of the New York Academy of Sciences, 2000. **908**(1): p. 244-254.
159. Bellizzi, D., et al., *Gene expression of cytokines and cytokine receptors is modulated by the common variability of the mitochondrial DNA in cybrid cell lines.* Genes to Cells, 2006. **11**(8): p. 883-891.
160. Zinovkin, R.A., et al., *Role of mitochondrial reactive oxygen species in age-related inflammatory activation of endothelium.* Aging (Albany NY), 2014. **6**(8): p. 661.
161. Zorov, D.B., et al., *Reactive oxygen species (Ros-Induced) Ros release a new phenomenon accompanying induction of the mitochondrial permeability transition in cardiac myocytes.* The Journal of experimental medicine, 2000. **192**(7): p. 1001-1014.
162. O'Hagan, H.M., et al., *Oxidative damage targets complexes containing DNA methyltransferases, SIRT1, and polycomb members to promoter CpG Islands.* Cancer cell, 2011. **20**(5): p. 606-619.
163. Khrapko, K. and D. Turnbull, *Chapter Two - Mitochondrial DNA Mutations in Aging*, in *Progress in Molecular Biology and Translational Science*, D.O. Heinz, Editor. 2014, Academic Press. p. 29-62.
164. Kazak, L., A. Reyes, and I.J. Holt, *Minimizing the damage: repair pathways keep mitochondrial DNA intact.* Nature reviews Molecular cell biology, 2012. **13**(10): p. 659-671.
165. Vermulst, M., et al., *Mitochondrial point mutations do not limit the natural lifespan of mice.* Nature genetics, 2007. **39**(4): p. 540-543.
166. Dzitoyeva, S., H. Chen, and H. Manev, *Effect of aging on 5-hydroxymethylcytosine in brain mitochondria.* Neurobiology of aging, 2012. **33**(12): p. 2881-2891.
167. Ghosh, S., S. Sengupta, and V. Scaria, *Comparative analysis of human mitochondrial methylomes shows distinct patterns of epigenetic regulation in mitochondria.* Mitochondrion, 2014. **18**: p. 58-62.

168. Takasugi, M., et al., *DNA methylation status of nuclear-encoded mitochondrial genes underlies the tissue-dependent mitochondrial functions.* BMC genomics, 2010. **11**(1): p. 481.
169. Hashizume, O., et al., *Epigenetic regulation of the nuclear-coded GCAT and SHMT2 genes confers human age-associated mitochondrial respiration defects.* Scientific reports, 2015. **5**.
170. Sage, J.M. and K.L. Knight, *Human Rad51 promotes mitochondrial DNA synthesis under conditions of increased replication stress.* Mitochondrion, 2013. **13**(4): p. 350-356.
171. *www.unodc.org/documents/data-and-analysis/statistics/Homicide/Homicides_by_firearms.xls*. 2012 [cited 2013 February, 28].
172. Verdon, T.J., et al., *The Pinpoint DNA Isolation System as a novel DNA sampling method in forensic biology.* Forensic Science International: Genetics Supplement Series, 2011.
173. *Getting Away With Murder* http://www.nytimes.com/2007/10/23/opinion/23tue3.html?_r=0 2007 [cited 2013 February, 28].
174. QIAGEN. *QIAamp® DNA Investigator Kit: Handbook (online)*. : December 2013; Available from: http://www.qiagen.com/gb/products/catalog/sample-technologies/dna-sample-technologies/genomic-dna/qiaamp-dna-investigator-kit.
175. MQIAGEN, *EpiTect Methyl II PCR Assay Handbook*. 2012, www.qiagen.com.
176. Biosystems, A., *AmpFLSTR® MiniFiler™ PCR Amplification Kit User Guide*. July 2014.
177. Vallone, P.M. and J.M. Butler, *AutoDimer: a screening tool for primer-dimer and hairpin structures.* Biotechniques, 2004. **37**(2): p. 226-231.
178. Wieslander, L., *A simple method to recover intact high molecular weight RNA and DNA after electrophoretic separation in low gelling temperature agarose gels.* Analytical biochemistry, 1979. **98**(2): p. 305-309.
179. Sweet, D., et al., *An improved method to recover saliva from human skin: the double swab technique.* Journal of forensic sciences, 1997. **42**: p. 320-322.
180. Kent, W.J., et al., *The human genome browser at UCSC.* Genome research, 2002. **12**(6): p. 996-1006.
181. Sweet, D., et al., *PCR-based DNA typing of saliva stains recovered from human skin.* Journal of forensic sciences, 1997. **42**(3): p. 447.
182. Pang, B. and B. Cheung, *Double swab technique for collecting touched evidence.* Legal Medicine, 2007. **9**(4): p. 181-184.
183. Brownlow, R.J., K.E. Dagnall, and C.E. Ames, *A Comparison of DNA Collection and Retrieval from Two Swab Types (Cotton and Nylon Flocked Swab) when Processed Using Three QIAGEN Extraction Methods.* Journal of Forensic Sciences, 2012. **57**: p. 713-717.
184. Dalmaso, G., et al., *Qualification of high-recovery, flocked swabs as compared to traditional rayon swabs for microbiological environmental monitoring of surfaces.* PDA Journal of Pharmaceutical Science and Technology, 2008. **62**(3): p. 191-199.
185. Garvin, A.M., et al., *The forensiX Evidence Collection Tube and Its Impact on DNA Preservation and Recovery.* BioMed Research International, 2013. **2013**: p. 7.
186. El-Maarri, O., *SIRPH Analysis*, in *Epigenetics Protocols*, T. Tollefsbol, Editor. 2004, Humana Press. p. 195-205.
187. RESEARCH, Z., *EZ DNA Methylation-Direct™ Kit*. 2014, Zymo Research.

188. Mawlood, S.K., M. Alrowaithi, and N. Watson, *Advantage of ForensiX Swabs in Retrieving and Preserving Biological Fluids.* Journal of Forensic Sciences, 2015. **60**(3): p. 686-689.
189. Park, J.-L., et al., *Identification of body fluid-specific DNA methylation markers for use in forensic science.* Forensic Science International: Genetics, 2014. **13**: p. 147-153.
190. Antunes, J.P., et al. *DNA methylation markers as a powerful technique to discriminate body fluids present in crime scenes*. in *Proceedings of the 24th International Symposium on Human Identification*. 2014.
191. Gunter, T.D. and R.A. Philibert, *Smoking, Methylation at AHRR, and Recidivism Risk in a Community Correction Sample of Individuals at High Risk for Recidivism.* Behavioral sciences & the law, 2015. **33**(5): p. 691-700.
192. Lee, H.Y., et al., *Epigenetic age signatures in the forensically relevant body fluid of semen: a preliminary study.* Forensic Science International: Genetics, 2015.
193. Weidner, C.I., et al., *Aging of blood can be tracked by DNA methylation changes at just three CpG sites.* Genome Biology, 2014. **15**(2): p. R24.
194. Koch, C.M., et al., *Specific age-associated DNA methylation changes in human dermal fibroblasts.* PLoS One, 2011. **6**(2): p. e16679.
195. Zbieć-Piekarska, R., et al., *Development of a forensically useful age prediction method based on DNA methylation analysis.* Forensic Science International: Genetics, 2015.
196. Gršković, B., et al., *DNA methylation: the future of crime scene investigation?* Molecular biology reports, 2013. **40**(7): p. 4349-4360.
197. Vidaki, A. and B. Daniel, *Forensic DNA methylation profiling—Potential opportunities and challenges.* Forensic Science International: Genetics, 2013. **7**(5): p. 499-507.
198. Bekaert, B., et al., *Improved age determination of blood and teeth samples using a selected set of DNA methylation markers.* Epigenetics, 2015. **10**(10): p. 922-930.
199. Huang, Y., et al., *Developing a DNA methylation assay for human age prediction in blood and bloodstain.* Forensic Science International: Genetics, 2015. **17**: p. 129-136.
200. Xu, C., et al., *A novel strategy for forensic age prediction by DNA methylation and support vector regression model.* Scientific reports, 2014. **5**: p. 17788-17788.
201. Krueger, F. and S.R. Andrews, *Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications.* Bioinformatics, 2011. **27**(11): p. 1571-1572.
202. Yi, S.H., et al., *Isolation and identification of age-related DNA methylation markers for forensic age-prediction.* Forensic Science International: Genetics, 2014. **11**: p. 117-125.
203. Burgess, D.J., *Human epigenetics: Showing your age.* Nature Reviews Genetics, 2012. **14**(1): p. 6-6.
204. Bocklandt, S., et al., *Epigenetic Predictor of Age.* PLoS ONE, 2011. **6**(6): p. e14821.
205. Manev, H., S. Dzitoyeva, and H. Chen, *Mitochondrial DNA: a blind spot in neuroepigenetics.* Biomolecular concepts, 2012. **3**(2): p. 107-115.
206. Hong, E.E., et al., *Regionally specific and genome-wide analyses conclusively demonstrate the absence of CpG methylation in human mitochondrial DNA.* Molecular and cellular biology, 2013. **33**(14): p. 2683-2690.
207. Maekawa, M., et al., *Methylation of mitochondrial DNA is not a useful marker for cancer detection.* Clinical chemistry, 2004. **50**(8): p. 1480-1481.

208. Stewart, S.A. and R.A. Weinberg, *Senescence: does it all happen at the ends?* Oncogene, 2002. **21**(4): p. 627-630.
209. Stewart, S.A. and R.A. Weinberg, *Telomeres: cancer to human aging.* Annu. Rev. Cell Dev. Biol., 2006. **22**: p. 531-557.
210. Shock, L.S., et al., *DNA methyltransferase 1, cytosine methylation, and cytosine hydroxymethylation in mammalian mitochondria.* Proceedings of the National Academy of Sciences, 2011. **108**(9): p. 3630-3635.
211. Chinnery, P.F., et al., *Epigenetics, epidemiology and mitochondrial DNA diseases.* International Journal of Epidemiology, 2012. **41**(1): p. 177-187.
212. Bellizzi, D., et al., *The Control Region of Mitochondrial DNA Shows an Unusual CpG and Non-CpG Methylation Pattern.* DNA Research, 2013: p. dst029.
213. Byun, H.-M., et al., *Effects of airborne pollutants on mitochondrial DNA methylation.* Part Fibre Toxicol, 2013. **10**(1): p. 18.
214. O'Sullivan, M., et al., *Mitochondrial m. 1584A 12S m62A rRNA methylation in families with m. 1555A> G associated hearing loss.* Human molecular genetics, 2014: p. ddu518.
215. Metodiev, M.D., et al., *Methylation of 12S rRNA is necessary for in vivo stability of the small subunit of the mammalian mitochondrial ribosome.* Cell metabolism, 2009. **9**(4): p. 386-397.
216. Rebelo, A.P., S.L. Williams, and C.T. Moraes, *In vivo methylation of mtDNA reveals the dynamics of protein–mtDNA interactions.* Nucleic acids research, 2009: p. gkp727.
217. Feng, S., et al., *Correlation between increased ND2 expression and demethylated displacement loop of mtDNA in colorectal cancer.* Molecular medicine reports, 2012. **6**(1): p. 125-130.
218. Pirola, C.J., et al., *Epigenetic modification of liver mitochondrial DNA is associated with histological severity of nonalcoholic fatty liver disease.* Gut, 2012: p. gutjnl-2012-302962.
219. Komaki, H., et al., *A novel mtDNA C11777A mutation in Leigh syndrome.* Mitochondrion, 2003. **2**(4): p. 293-304.
220. Harding, A., et al., *Prenatal diagnosis of mitochondrial DNA8993 T----G disease.* American journal of human genetics, 1992. **50**(3): p. 629.
221. Holt, I., et al., *A new mitochondrial disease associated with mitochondrial DNA heteroplasmy.* American journal of human genetics, 1990. **46**(3): p. 428.
222. Klukas, C.K. and I.B. Dawid, *Characterization and mapping of mitochondrial ribosomal RNA and mitochondrial DNA in Drosophila melanogaster.* Cell, 1976. **9**(4): p. 615-625.
223. Kasamatsu, H., D.L. Robberson, and J. Vinograd, *A novel closed-circular mitochondrial DNA with properties of a replicating intermediate.* Proceedings of the National Academy of Sciences, 1971. **68**(9): p. 2252-2257.
224. Yang, I.S., et al., *mtDNAprofiler: A Web Application for the Nomenclature and Comparison of Human Mitochondrial DNA Sequences.* Journal of forensic sciences, 2013.
225. Copeland, W.C., *Defects in mitochondrial DNA replication and human disease.* Crit Rev Biochem Mol Biol, 2012. **47**(1): p. 64-74.
226. Ruiz-Pesini, E., et al., *An enhanced MITOMAP with a global mtDNA mutational phylogeny.* Nucleic acids research, 2007. **35**(suppl 1): p. D823-D828.
227. Zapico, S.C. and D.H. Ubelaker, *mtDNA mutations and their role in aging, diseases and forensic sciences.* Aging and disease, 2013. **4**(6): p. 364.

228. Giordano, M., et al. *Methylation of the human mitochondrial 12S rRNA gene is correlated with aging*. in *12th International FISV Congress, Rome, Italy*. 2012.
229. Liu, L., et al., *Aging, cancer and nutrition: the DNA methylation connection.* Mechanisms of ageing and development, 2003. **124**(10): p. 989-998.
230. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.
231. Bird, A., et al., *A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA.* Cell, 1985. **40**(1): p. 91-99.
232. Larsen, F., et al., *CpG islands as gene markers in the human genome.* Genomics, 1992. **13**(4): p. 1095-1107.
233. Saxonov, S., P. Berg, and D.L. Brutlag, *A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(5): p. 1412-1417.
234. Weber, M., et al., *Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome.* Nature genetics, 2007. **39**(4): p. 457-466.
235. Cooper, D.N., M.H. Taggart, and A.P. Bird, *Unmethlated domains in vertebrate DNA.* Nucleic acids research, 1983. **11**(3): p. 647-658.
236. Gardiner-Garden, M. and M. Frommer, *CpG islands in vertebrate genomes.* Journal of molecular biology, 1987. **196**(2): p. 261-282.
237. Xiong, Z. and P.W. Laird, *COBRA: a sensitive and quantitative DNA methylation assay.* Nucleic acids research, 1997. **25**(12): p. 2532-2534.
238. Herman, J.G., et al., *Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands.* Proceedings of the National Academy of Sciences, 1996. **93**(18): p. 9821-9826.
239. Frommer, M., et al., *A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands.* Proceedings of the National Academy of Sciences, 1992. **89**(5): p. 1827-1831.
240. Ehrich, M., et al., *Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(44): p. 15785-15790.
241. Taylor, K.H., et al., *Ultradeep bisulfite sequencing analysis of DNA methylation patterns in multiple gene promoters by 454 sequencing.* Cancer research, 2007. **67**(18): p. 8511-8518.
242. Shames, D.S., J.D. Minna, and A.F. Gazdar, *Methods for detecting DNA methylation in tumors: from bench to bedside.* Cancer letters, 2007. **251**(2): p. 187-198.
243. Eads, C.A., et al., *MethyLight: a high-throughput assay to measure DNA methylation.* Nucleic acids research, 2000. **28**(8): p. e32-00.
244. Jiang, Q., et al., *EpiTect Methyl II PCR Array System: A simple tool for screening regional DNA methylation of a large number of genes or samples without bisulfite conversion.* Frederick, MD, 2012. **21703**.
245. Illingworth, R.S. and A.P. Bird, *CpG islands–'a rough guide'.* FEBS letters, 2009. **583**(11): p. 1713-1720.
246. Sutherland, E., L. Coe, and E.A. Raleigh, *McrBC: a multisubunit GTP-dependent restriction endonuclease.* Journal of molecular biology, 1992. **225**(2): p. 327-348.
247. Stewart, F. and E. Raleigh, *Dependence of McrBC cleavage on distance between recognition elements.* Biological chemistry, 1997. **379**(4-5): p. 611-616.

248. Ordway, J., et al., *Comprehensive DNA methylation profiling in a human cancer genome identifies novel epigenetic targets.* Carcinogenesis, 2006. **27**(12): p. 2409-2423.
249. Balic, M., et al., *Genetic and epigenetic analysis of putative breast cancer stem cell models.* BMC cancer, 2013. **13**(1): p. 358.
250. Esteller, M., *Epigenetic gene silencing in cancer: the DNA hypermethylome.* Human molecular genetics, 2007. **16**(R1): p. R50-R59.
251. Reik, W., *Stability and flexibility of epigenetic gene regulation in mammalian development.* Nature, 2007. **447**(7143): p. 425-432.
252. Wong, I.H., et al., *Detection of aberrant p16 methylation in the plasma and serum of liver cancer patients.* Cancer research, 1999. **59**(1): p. 71-73.
253. (Hachemi.Zeraia@qiagen.com), D.H. 9th June 2014.
254. Thiede, C., et al., *Buccal swabs but not mouthwash samples can be used to obtain pretransplant DNA fingerprints from recipients of allogeneic bone marrow transplants.* Bone marrow transplantation, 2000. **25**(5): p. 575.
255. Cao-Lei, L., et al., *DNA methylation signatures triggered by prenatal maternal stress exposure to a natural disaster: project ice storm.* PloS one, 2014. **9**(9): p. e107653.
256. Elango, N. and V.Y. Soojin, *DNA methylation and structural and functional bimodality of vertebrate promoters.* Molecular biology and evolution, 2008. **25**(8): p. 1602-1608.
257. Zhang, Y., et al., *DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution.* PLoS genetics, 2009. **5**(3): p. e1000438.
258. Geisser, S., *Predictive inference*. Vol. 55. 1993: CRC Press.
259. Park, J.-G. and V.M. Chapman, *CpG island promoter region methylation patterns of the inactive-X-chromosome hypoxanthine phosphoribosyltransferase (Hprt) gene.* Molecular and cellular biology, 1994. **14**(12): p. 7975-7983.
260. Zampieri, M., et al., *Reconfiguration of DNA methylation in ageing.* Mechanisms of Ageing and Development, 2015.
261. Dunn, J., et al., *Flow-dependent epigenetic DNA methylation regulates endothelial gene expression and atherosclerosis.* The Journal of clinical investigation, 2014. **124**(124 (7)): p. 0-0.
262. Lindholm, M.E., et al., *An integrative analysis reveals coordinated reprogramming of the epigenome and the transcriptome in human skeletal muscle after training.* Epigenetics, 2014. **9**(12): p. 1557-69.
263. Meissner, C. and S. Ritz-Timme, *Molecular pathology and age estimation.* Forensic Science International, 2010. **203**(1–3): p. 34-43.
264. Van Oorschot, R., K.N. Ballantyne, and R.J. Mitchell, *Forensic trace DNA: a review.* Investig genet, 2010. **1**(1): p. 14.
265. Evett, I., R. Pinchin, and C. Buffery, *An investigation of the feasibility of inferring ethnic origin from DNA profiles.* Journal of the Forensic Science Society, 1992. **32**(4): p. 301-306.
266. Locard, E., *The analysis of dust traces. Part I.* The American Journal of Police Science, 1930. **1**(3): p. 276-298.
267. Mawlood, S.K., M. Alrowaithi, and N. Watson, *Advantage of ForensiX Swabs in Retrieving and Preserving Biological Fluids.* Journal of Forensic Sciences, 2015.
268. Naseem, M. *DNA Extraction from Casework Samples using the Applied Biosystems® AutoMate Express™ Forensic DNA Extraction System*. 2011 [cited 214 10,15]; Available from: http://www.appliedbiosystems.com/ForensicNews-Past

269. Liu, J.Y., et al., *AutoMate Express™ Forensic DNA Extraction System for the Extraction of Genomic DNA from Biological Samples*.* Journal of forensic sciences, 2012. **57**(4): p. 1022-1030.
270. Montpetit, S.A., I.T. Fitch, and P.T. O'Donnell, *A simple automated instrument for DNA extraction in forensic casework*, in *J Forensic Sci*. 2005. p. 555-63.
271. Butler, J.M., Y. Shen, and B.R. McCord, *The development of reduced size STR amplicons as tools for analysis of degraded DNA.* Journal of forensic sciences, 2003. **48**(5): p. 1054-1064.
272. Ryan, S.R. *Successful DNA profiling from fired*

casings, bullets and cartridges. JUNE 2012. **VOL.1, ISSUE 4. (http://www.ryanforensicdna.com/yahoo_site_admin/assets/docs/June_Newsletter.163101437.pdf)**.
273. Dominick, A., et al., *Is there a relationship between fingerprint donation and DNA shedding?* Journal of Forensic Identification, 2009. **59**(2): p. 133.
274. Ravreby, M., *Analysis of long-range bullet entrance holes by atomic absorption spectrophotometry and scanning electron microscopy.* Journal of forensic sciences, 1982. **27**(1): p. 92-112.
275. Koons, R.D. and D.M. Grant, *Compositional variation in bullet lead manufacture.* Journal of forensic sciences, 2002. **47**(5): p. 950-958.
276. Branch, L., *Generating forensic DNA profiles from'contact'DNA on cartridge cases and gun grips*. 2011.
277. Capuano, F., et al., *Cytosine DNA methylation is found in Drosophila melanogaster but absent in Saccharomyces cerevisiae, Schizosaccharomyces pombe, and other yeast species.* Analytical chemistry, 2014. **86**(8): p. 3697-3702.

Appendixes

Appendix 1 Primers and sequence of target region of nDNA: This Appendix (DVD attached) contains the sequence of target region before and after bisulphite conversion with two pairs of primers for each sequence of each DNA fragment.

Appendix 2 Quality control Report for nDNA: The content of this appendix is in a DVA (attached).

Appendix 3 Total read of NGS : An Excel file that contains the nDNA read data and CpG methylation calling for all sites (397) across all samples for the project. The methylation ratio (meth-ratio) is calculated by using methylated CpG count total CpG count. Furthermore, you will find in the excel table the coordinates for each of the detected CpG sites across all of the samples. If one CpG site was detected in at least one sample with at least 10 reads, we listed that CpG site in the table. For samples where the CpG was not detected (<10 reads), we left the meth ratio and total CpG count column blank (Attached in the DVD).

Appendix 4: Reference Sequence: The Mitochondrial sequence (hg38:2013) shows 53 red colour with underline and bold CG.

```
   1 GATCACAGGTCTATCACCCTATTAACCACTCACGGGAGCTCTCCATGCATTTGGTATTTT
  61 CGTCTGGGGGGTATGCACGCGATAGCATTGCGAGACGCTGGAGCCGGAGCACCCTATGTC
 121 GCAGTATCTGTCTTTGATTCCTGCCTCATCCTATTATTTATCGCACCTACGTTCAATATT
 181 ACAGGCGAACATACTTACTAAAGTGTGTTAATTAATTAATGCTTGTAGGACATAATAATA
 241 ACAATTGAATGTCTGCACAGCCACTTTCCACACAGACATCATAACAAAAAATTTCCACCA
 301 AACCCCCCCTCCCCCGCTTCTGGCCACAGCACTTAAACACATCTCTGCCAAACCCCAAAA
 361 ACAAAGAACCCTAACACCAGCCTAACCAGATTTCAAATTTTATCTTTTGGCGGTATGCAC
 421 TTTTAACAGTCACCCCCCAACTAACACATTATTTTCCCCTCCCACTCCCATACTACTAAT
 481 CTCATCAATACAACCCCCGCCCATCCTACCCAGCACACACACACCGCTGCTAACCCCATA
 541 CCCCGAACCAACCAAACCCCAAAGACACCCCCCACAGTTTATGTAGCTTACCTCCTCAAA
 601 GCAATACACTGAAAATGTTTAGACGGGCTCACATCACCCCATAAACAAATAGGTTTGGTC
 661 CTAGCCTTTCTATTAGCTCTTAGTAAGATTACACATGCAAGCATCCCCGTTCCAGTGAGT
 721 TCACCCTCTAAATCACCACGATCAAAAGGAACAAGCATCAAGCACGCAGCAATGCAGCTC
 781 AAAACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCAGTGATTAACCTTTAGCAATAA
 841 ACGAAAGTTTAACTAAGCTATACTAACCCCAGGGTTGGTCAATTTCGTGCCAGCCACCGC
 901 GGTCACACGATTAACCCAAGTCAATAGAAGCCGGCGTAAAGAGTGTTTTAGATCACCCCC
 961 TCCCCAATAAAGCTAAAACTCACCTGAGTTGTAAAAAACTCCAGTTGACACAAAATAGAC
1021 TACGAAAGTGGCTTTAACATATCTGAACACACAATAGCTAAGACCCAAACTGGGATTAGA
1081 TACCCCACTATGCTTAGCCCTAAACCTCAACAGTTAAATCAACAAAACTGCTCGCCAGAA
1141 CACTACGAGCACAGCTTAAAACTCAAAGGACCTGGCGGTGCTTCATATCCCTCTAGAGGA
1201 GCCTGTTCTGTAATCGATAAACCCCGATCAACCTCACCACCTCTTGCTCAGCCTATATAC
1261 CGCCATCTTCAGCAAACCCTGATGAAGGCTACAAAGTAAGCGCAAGTACCCACGTAAAGA
1321 CGTTAGGTCAAGGTGTAGCCCATGAGGTGGCAAGAAATGGGCTACATTTTCTACCCCAGA
1381 AAACTACGATAGCCCTTATGAAACTTAAGGGTCGAAGGTGGATTTAGCAGTAAACTAAGA
1441 GTAGAGTGCTTAGTTGAACAGGGCCCTGAAGCGCGTACACACCGCCCGTCACCCTCCTCA
1501 AGTATACTTCAAAGGACATTTAACTAAAACCCCTACGCATTTATATAGAGGAGACAAGTC
1561 GTAACATGGTAAGTGTACTGGAAAGTGCACTTGGACGAACCAGAGTGTAGCTTAACACAA
```

1621 AGCACCCAACTTACACTTAGGAGATTTCAACTTAACTTGAC**CG**CTCTGAGCTAAACCTAG
1681 CCCCAAACCCACTCCACCTTACTACCAGACAACCTTAGCCAAACCATTTACCCAAATAAA
1741 GTATAGG**CG**ATAGAAATTGAAACCTGG**CG**CAATAGATATAGTAC**CG**CAAGGGAAAGATGA
1801 AAAATTATAACCAAGCATAATATAGCAAGGACTAACCCCTATACCTTCTGCATAATGAAT
1861 TAACTAGAAATAACTTTGCAAGGAGAGCCAAAGCTAAGACCCC**CG**AAACCAGA**CG**AGCTA
1921 CCTAAGAACAGCTAAAAGAGCACACC**CG**TCTATGTAGCAAAATAGTGGGAAGATTTATAG
1981 GTAGAGG**CG**ACAAACCTAC**CG**AGCCTGGTGATAGCTGGTTGTCCAAGATAGAATCTTAGT
2041 TCAACTTTAAATTTGCCCACAGAACCCTCTAAATCCCCTTGTAAATTTAACTGTTAGTCC
2101 CAAAGAGGAACAGCTCTTTGGACACTAGGAAAAAACCTTGTAGAGAGAGTAAAAAATTTA
2161 CACCCATAGTAGGCCTAAAAGCAGCCACCAATTAAGAAAG**CG**TTCAAGCTCAACACCCAC
2221 TACCTAAAAAATCCCAAACATATAACTGAACTCCTCACACCCAATTGGACCAATCTATCA
2281 CCCTATAGAAGAACTAATGTTAGTATAAGTAACATGAAAACATTCTCCTC**CG**CATAAGCC
2341 TG**CG**TCAGATTAAAACACTGAACTGACAATTAACAGCCCAATATCTACAATCAACCAACA
2401 AGTCATTATTACCCTCACTGTCAACCCAACACAGGCATGCTCATAAGGAAAGGTTAAAAA
2461 AAGTAAAAGGAACT**CG**GCAAATCTTACCC**CG**CCTGTTTACCAAAAACATCACCTCTAGCA
2521 TCACCAGTATTAGAGGCAC**CG**CCTGCCCAGTGACACATGTTTAA**CG**GC**CGCG**GTACCCTA
2581 AC**CG**TGCAAAGGTAGCATAATCACTTGTTCCTTAAATAGGGACCTGTATGAATGGCTCCA
2641 **CG**AGGGTTCAGCTGTCTCTTACTTTTAACCAGTGAAATTGACCTGCC**CG**TGAAGAGG**CG**G
2701 GCATAACACAGCAAGA**CG**AGAAGACCCTATGGAGCTTTAATTTATTAATGCAAACAGTAC
2761 CTAACAAACCCACAGGTCCTAAACTACCAAACCTGCATTAAAAATTT**CG**GTTGGGG**CG**AC
2821 CT**CG**GAGCAGAACCCAACCTC**CG**AGCAGTACATGCTAAGACTTCACCAGTCAAAG**CG**AAC
2881 TACTATACTCAATTGATCCAATAACTTGACCAA**CG**GAACAAGTTACCCTAGGGATAACAG
2941 **CG**CAATCCTATTCTAGAGTCCATATCAACAATAGGGTTTA**CG**ACCT**CG**ATGTTGGATCAG
3001 GACATCC**CG**ATGGTGCAGC**CG**CTATTAAAGGTT**CG**TTTGTTCAA**CG**ATTAAAGTCCTA**CG**
3061 TGATCTGAGTTCAGAC**CG**GAGTAATCCAGGT**CG**GTTTCTATCTACNTTCAAATTCCTCCC
3121 TGTA**CG**AAAGGACAAGAGAAATAAGGCCTACTTCACAAAG**CG**CCTTCCCC**CG**TAAATGAT
3181 ATCATCTCAACTTAGTATTATACCCACACCCACCCAAGAACAGGGTTTGTTAAGATGGCA
3241 GAGCC**CG**GTAAT**CG**CATAAAACTTAAAACTTTACAGTCAGAGGTTCAATTCCTCTTCTTA
3301 ACAACATACCCATGGCCAACCTCCTACTCCTCATTGTACCCATTCTAAT**CG**CAATGGCAT

3361 TCCTAATGCTTAC**CG**AA**CG**AAAAATTCTAGGCTATATACAACTA**CG**CAAAGGCCCCAA**CG**
3421 TTGTAGGCCCCTA**CG**GGCTACTACAACCCTT**CG**CTGA**CG**CCATAAAACTCTTCACCAAAG
3481 AGCCCCTAAAACC**CG**CCACATCTACCATCACCCTCTACATCAC**CG**CCC**CG**ACCTTAGCTC
3541 TCACCAT**CG**CTCTTCTACTATGAACCCCCCTCCCCATACCCAACCCCCTGGTCAACCTCA
3601 ACCTAGGCCTCCTATTTATTCTAGCCACCTCTAGCCTAGC**CG**TTTACTCAATCCTCTGAT
3661 CAGGGTGAGCATCAAACTCAAACTA**CG**CCCTGAT**CG**G**CG**CACTG**CG**AGCAGTAGCCCAAA
3721 CAATCTCATATGAAGTCACCCTAGCCATCATTCTACTATCAACATTACTAATAAGTGGCT
3781 CCTTTAACCTCTCCACCCTTATCACAACACAAGAACACCTCTGATTACTCCTGCCATCAT
3841 GACCCTTGGCCATAATATGATTTATCTCCACACTAGCAGAGACCAAC**CG**AACCCCCTT**CG**
3901 ACCTTGC**CG**AAGGGGAGTC**CG**AACTAGTCTCAGGCTTCAACAT**CG**AATA**CG**C**CG**CAGGCC
3961 CCTT**CG**CCCTATTCTTCATAGC**CG**AATACACAAACATTATTATAATAAACACCCTCACCA
4021 CTACAATCTTCCTAGGAACAACATATGA**CG**CACTCTCCCCTGAACTCTACACAACATATT
4081 TTGTCACCAAGACCCTACTTCTAACCTCCCTGTTCTTATGAATT**CG**AACAGCATACCCC**C**
4141 **G**ATTC**CG**CTA**CG**ACCAACTCATACACCTCCTATGAAAAAACTTCCTACCACTCACCCTAG
4201 CATTACTTATATGATATGTCTCCATACCCATTACAATCTCCAGCATTCCCCCTCAAACCT
4261 AAGAAATATGTCTGATAAAAGAGTTACTTTGATAGAGTAAATAATAGGAGCTTAAACCCC
4321 CTTATTTCTAGGACTATGAGAAT**CG**AACCCATCCCTGAGAATCCAAAATTCTC**CG**TGCCA
4381 CCTATCACACCCCATCCTAAAGTAAGGTCAGCTAAATAAGCTAT**CG**GGCCCATACCC**CG**A
4441 AAATGTTGGTTATACCCTTCC**CG**TACTAATTAATCCCCTGGCCCAACC**CG**TCATCTACTC
4501 TACCATCTTTGCAGGCACACTCATCACAG**CG**CTAAGCT**CG**CACTGATTTTTTACCTGAGT
4561 AGGCCTAGAAATAAACATGCTAGCTTTTATTCCAGTTCTAACCAAAAAAATAAACCCT**CG**
4621 TTCCACAGAAGCTGCCATCAAGTATTTCCTCA**CG**CAAGCAAC**CG**CATCCATAATCCTTCT
4681 AATAGCTATCCTCTTCAACAATATACTCTC**CG**GACAATGAACCATAACCAATACTACCAA
4741 TCAATACTCATCATTAATAATCATAATAGCTATAGCAATAAAACTAGGAATAGCCCCCTT
4801 TCACTTCTGAGTCCCAGAGGTTACCCAAGGCACCCCTCTGACATC**CG**GCCTGCTTCTTCT
4861 CACATGACAAAAACTAGCCCCCATCTCAATCATATACCAAATCTCTCCCTCACTAAA**CG**T
4921 AAGCCTTCTCCTCACTCTCTCAATCTTATCCATCATAGCAGGCAGTTGAGGTGGATTAAA
4981 CCAAACCCAGCTA**CG**CAAAATCTTAGCATACTCCTCAATTACCCACATAGGATGAATAAT
5041 AGCAGTTCTAC**CG**TACAACCCTAACATAACCATTCTTAATTTAACTATTTATATTATCCT

5101 AACTACTAC**CG**CATTCCTACTACTCAACTTAAACTCCAGCACCA**CG**ACCCTACTACTATC
5161 T**CG**CACCTGAAACAAGCTAACATGACTAACACCCTTAATTCCATCCACCCTCCTCTCCCT
5221 AGGAGGCCTGCCCC**CG**CTAAC**CG**GCTTTTTGCCCAAATGGGCCATTAT**CG**AAGAATTCAC
5281 AAAAAACAATAGCCTCATCATCCCCACCATCATAGCCACCATCACCCTCCTTAACCTCTA
5341 CTTCTACCTA**CG**CCTAATCTACTCCACCTCAATCACACTACTCCCCATATCTAACAA**CG**T
5401 AAAAATAAAATGACAGTTTGAACATACAAAACCCACCCCATTCCTCCCCACACTCAT**CG**C
5461 CCTTACCA**CG**CTACTCCTACCTATCTCCCCTTTTATACTAATAATCTTATAGAAATTTAG
5521 GTTAAATACAGACCAAGAGCCTTCAAAGCCCTCAGTAAGTTGCAATACTTAATTTCTGTA
5581 ACAGCTAAGGACTGCAAAACCCCACTCTGCATCAACTGAA**CG**CAAATCAGCCACTTTAAT
5641 TAAGCTAAGCCCTTACTAGACCAATGGGACTTAAACCCACAAACACTTAGTTAACAGCTA
5701 AGCACCCTAATCAACTGGCTTCAATCTACTTCTCC**CG**C**CG**C**CG**GGAAAAAAGG**CG**GGAGA
5761 AGCCC**CG**GCAGGTTTGAAGCTGCTTCTT**CG**AATTTGCAATTCAATATGAAAATCACCT**CG**
5821 GAGCTGGTAAAAAGAGGCCTAACCCCTGTCTTTAGATTTACAGTCCAATGCTTCACTCAG
5881 CCATTTTACCTCACCCCCACTGATGTT**CG**C**CG**AC**CG**TTGACTATTCTCTACAAACCACAA
5941 AGACATTGGAACACTATACCTATTATT**CG**G**CG**CATGAGCTGGAGTCCTAGGCACAGCTCT
6001 AAGCCTCCTTATT**CG**AGC**CG**AGCTGGGCCAGCCAGGCAACCTTCTAGGTAA**CG**ACCACAT
6061 CTACAA**CG**TTAT**CG**TCACAGCCCATGCATTTGTAATAATCTTCTTCATAGTAATACCCAT
6121 CATAAT**CG**GAGGCTTTGGCAACTGACTAGTTCCCCTAATAAT**CG**GTGCCCC**CG**ATATGG**C**
6181 **G**TTTCCC**CG**CATAAACAACATAAGCTTCTGACTCTTACCTCCCTCTCTCCTACTCCTGCT
6241 **CG**CATCTGCTATAGTGGAGGC**CG**GAGCAGGAACAGGTTGAACAGTCTACCCTCCCTTAGC
6301 AGGGAACTACTCCCACCCTGGAGCCTC**CG**TAGACCTAACCATCTTCTCCTTACACCTAGC
6361 AGGTGTCTCCTCTATCTTAGGGGCCATCAATTTCATCACAACAATTATCAATATAAAACC
6421 CCCTGCCATAACCCAATACCAAA**CG**CCCCTCTT**CG**TCTGATC**CG**TCCTAATCACAGCAGT
6481 CCTACTTCTCCTATCTCTCCCAGTCCTAGCTGCTGGCATCACTATACTACTAACAGAC**CG**
6541 CAACCTCAACACCACCTTCTT**CG**ACCC**CG**C**CG**GAGGAGGAGACCCCATTCTATACCAACA
6601 CCTATTCTGATTTTT**CG**GTCACCCTGAAGTTTATATTCTTATCCTACCAGGCTT**CG**GAAT
6661 AATCTCCCATATTGTAACTTACTACTC**CG**GAAAAAAAGAACCATTTGGATACATAGGTAT
6721 GGTCTGAGCTATGATATCAATTGGCTTCCTAGGGTTTAT**CG**TGTGAGCACACCATATATT
6781 TACAGTAGGAATAGA**CG**TAGACACA**CG**AGCATATTTCACCTC**CG**CTACCATAATCAT**CG**C

6841 TATCCCCACCGGCGTCAAAGTATTTAGCTGACTCGCCACACTCCACGGAAGCAATATGAA
6901 ATGATCTGCTGCAGTGCTCTGAGCCCTAGGATTCATCTTTCTTTTCACCGTAGGTGGCCT
6961 GACTGGCATTGTATTAGCAAACTCATCACTAGACATCGTACTACACGACACGTACTACGT
7021 TGTAGCCCACTTCCACTATGTCCTATCAATAGGAGCTGTATTTGCCATCATAGGAGGCTT
7081 CATTCACTGATTTCCCCTATTCTCAGGCTACACCCTAGACCAAACCTACGCCAAAATCCA
7141 TTTCACTATCATATTCATCGGCGTAAATCTAACTTTCTTCCCACAACACTTTCTCGGCCT
7201 ATCCGGAATGCCCCGACGTTACTCGGACTACCCCGATGCATACACCACATGAAACATCCT
7261 ATCATCTGTAGGCTCATTCATTTCTCTAACAGCAGTAATATTAATAATTTTCATGATTTG
7321 AGAAGCCTTCGCTTCGAAGCGAAAAGTCCTAATAGTAGAAGAACCCTCCATAAACCTGGA
7381 GTGACTATATGGATGCCCCCCACCCTACCACACATTCGAAGAACCCGTATACATAAAATC
7441 TAGACAAAAAAGGAAGGAATCGAACCCCCCAAAGCTGGTTTCAAGCCAACCCCATGGCCT
7501 CCATGACTTTTTCAAAAAGGTATTAGAAAAACCATTTCATAACTTTGTCAAAGTTAAATT
7561 ATAGGCTAAATCCTATATATCTTAATGGCACATGCAGCGCAAGTAGGTCTACAAGACGCT
7621 ACTTCCCCTATCATAGAAGAGCTTATCACCTTTCATGATCACGCCCTCATAATCATTTTC
7681 CTTATCTGCTTCCTAGTCCTGTATGCCCTTTTCCTAACACTCACAACAAAACTAACTAAT
7741 ACTAACATCTCAGACGCTCAGGAAATAGAAACCGTCTGAACTATCCTGCCCGCCATCATC
7801 CTAGTCCTCATCGCCCTCCCATCCCTACGCATCCTTTACATAACAGACGAGGTCAACGAT
7861 CCCTCCCTTACCATCAAATCAATTGGCCACCAATGGTACTGAACCTACGAGTACACCGAC
7921 TACGGCGGACTAATCTTCAACTCCTACATACTTCCCCCATTATTCCTAGAACCAGGCGAC
7981 CTGCGACTCCTTGACGTTGACAATCGAGTAGTACTCCCGATTGAAGCCCCCATTCGTATA
8041 ATAATTACATCACAAGACGTCTTGCACTCATGAGCTGTCCCCACATTAGGCTTAAAAACA
8101 GATGCAATTCCCGGACGTCTAAACCAAACCACTTTCACCGCTACACGACCGGGGTATAC
8161 TACGGTCAATGCTCTGAAATCTGTGGAGCAAACCACAGTTTCATGCCCATCGTCCTAGAA
8221 TTAATTCCCCTAAAAATCTTTGAAATAGGGCCCGTATTTACCCTATAGCACCCCCTCTAC
8281 CCCCTCTAGAGCCCACTGTAAAGCTAACTTAGCATTAACCTTTTAAGTTAAAGATTAAGA
8341 GAACCAACACCTCTTTACAGTGAAATGCCCCAACTAAATACTACCGTATGGCCCACCATA
8401 ATTACCCCCATACTCCTTACACTATTCCTCATCACCCAACTAAAAATATTAAACACAAAC
8461 TACCACCTACCTCCCTCACCAAAGCCCATAAAAATAAAAAATTATAACAAACCCTGAGAA
8521 CCAAAATGAACGAAAATCTGTTCGCTTCATTCATTGCCCCCACAATCCTAGGCCTACCCG

```
 8581 CCGCAGTACTGATCATTCTATTTCCCCCTCTATTGATCCCCACCTCCAAATATCTCATCA
 8641 ACAACCGACTAATCACCACCCAACAATGACTAATCAAACTAACCTCAAAACAAATGATAA
 8701 CCATACACAACACTAAAGGACGAACCTGATCTCTTATACTAGTATCCTTAATCATTTTTA
 8761 TTGCCACAACTAACCTCCTCGGACTCCTGCCTCACTCATTTACACCAACCACCCAACTAT
 8821 CTATAAACCTAGCCATGGCCATCCCCTTATGAGCGGGCACAGTGATTATAGGCTTTCGCT
 8881 CTAAGATTAAAAATGCCCTAGCCCACTTCTTACCACAAGGCACACCTACACCCCTTATCC
 8941 CCATACTAGTTATTATCGAAACCATCAGCCTACTCATTCAACCAATAGCCCTGGCCGTAC
 9001 GCCTAACCGCTAACATTACTGCAGGCCACCTACTCATGCACCTAATTGGAAGCGCCACCC
 9061 TAGCAATATCAACCATTAACCTTCCCTCTACACTTATCATCTTCACAATTCTAATTCTAC
 9121 TGACTATCCTAGAAATCGCTGTCGCCTTAATCCAAGCCTACGTTTTCACACTTCTAGTAA
 9181 AGCCTCTACCTGCACGACAACACATAATGACCCACCAATCACATGCCTATCATATAGTAA
 9241 ACCCAGCCCATGACCCCTAACAGGGGCCCTCTCAGCCCTCCTAATGACCTCCGGCCTAGC
 9301 CATGTGATTTCACTTCCACTCCATAACGCTCCTCATACTAGGCCTACTAACCAACACACT
 9361 AACCATATACCAATGATGGCGCGATGTAACACGAGAAAGCACATACCAAGGCCACCACAC
 9421 ACCACCTGTCCAAAAAGGCCTTCGATACGGGATAATCCTATTTATTACCTCAGAAGTTTT
 9481 TTTCTTCGCAGGATTTTTCTGAGCCTTTTACCACTCCAGCCTAGCCCCTACCCCCCAATT
 9541 AGGAGGGCACTGGCCCCCAACAGGCATCACCCCGCTAAATCCCCTAGAAGTCCCACTCCT
 9601 AAACACATCCGTATTACTCGCATCAGGAGTATCAATCACCTGAGCTCACCATAGTCTAAT
 9661 AGAAAACAACCGAAACCAAATAATTCAAGCACTGCTTATTACAATTTTACTGGGTCTCTA
 9721 TTTTACCCTCCTACAAGCCTCAGAGTACTTCGAGTCTCCCTTCACCATTTCCGACGGCAT
 9781 TCTACGGCTCAACATTTTTTGTAGCCACAGGCTTCCACGGACTTCACGTCATTATTGGCT
 9841 AACTTTCCTCACTATCTGCTTCATCCGCCAACTAATATTTCACTTTACATCCAAACATCA
 9901 CTTTGGCTTCGAAGCCGCCCGCCTGATACTGGCATTTTGTAGATGTGGTTTGACTATTTCT
 9961 GTATGTCTCCATCTATTGATGAGGGTCTTACTCTTTTAGTATAAATAGTACCGTTAACTT
10021 CCAATTAACTAGTTTTGACAACATTCAAAAAAGAGTAATAAACTTCGCCTTAATTTTAAT
10081 AATCAACACCCTCCTAGCCTTACTACTAATAATTATTACATTTTGACTACCACAACTCAA
10141 CGGCTACATAGAAAAATCCACCCCTTACGAGTGCGGCTTCGACCCTATATCCCCCGCCCG
10201 CGTCCCTTTCTCCATAAAATTCTTCTTAGTAGCTATTACCTTCTTATTATTTGATCTAGA
10261 AATTGCCCTCCTTTTACCCCTACCATGAGCCCTACAAACAACTAACCTGCCACTAATAGT
```

```
10321 TATGTCATCCCTCTTATTAATCATCATCCTAGCCCTAAGTCTGGCCTATGAGTGACTACA
10381 AAAAGGATTAGACTGAACCGAATTGGTATATAGTTTAAACAAAACGAATGATTTCGACTC
10441 ATTAAATTATGATAATCATATTTACCAAATGCCCCTCATTTACATAAATATTATACTAGC
10501 ATTTACCATCTCACTTCTAGGAATACTAGTATATCGCTCACACCTCATATCCTCCCTACT
10561 ATGCCTAGAAGGAATAATACTATCGCTGTTCATTATAGCTACTCTCATAACCCTCAACAC
10621 CCACTCCCTCTTAGCCAATATTGTGCCTATTGCCATACTAGTCTTTGCCGCCTGCGAAGC
10681 AGCGGTGGGCCTAGCCCTACTAGTCTCAATCTCCAACACATATGGCCTAGACTACGTACA
10741 TAACCTAAACCTACTCCAATGCTAAAACTAATCGTCCCAACAATTATATTACTACCACTG
10801 ACATGACTTTCCAAAAAACACATAATTTGAATCAACACAACCACCCACAGCCTAATTATT
10861 AGCATCATCCCTCTACTATTTTTTAACCAAATCAACAACAACCTATTTAGCTGTTCCCCA
10921 ACCTTTTCCTCCGACCCCCTAACAACCCCCCTCCTAATACTAACTACCTGACTCCTACCC
10981 CTCACAATCATGGCAAGCCAACGCCACTTATCCAGTGAACCACTATCACGAAAAAAACTC
11041 TACCTCTCTATACTAATCTCCCTACAAATCTCCTTAATTATAACATTCACAGCCACAGAA
11101 CTAATCATATTTTATATCTTCTTCGAAACCACACTTATCCCCACCTTGGCTATCATCACC
11161 CGATGAGGCAACCAGCCAGAACGCCTGAACGCAGGCACATACTTCCTATTCTACACCCTA
11221 GTAGGCTCCCTTCCCCTACTCATCGCACTAATTTACACTCACAACACCCTAGGCTCACTA
11281 AACATTCTACTACTCACTCTCACTGCCCAAGAACTATCAAACTCCTGAGCCAACAACTTA
11341 ATATGACTAGCTTACACAATAGCTTTTATAGTAAAGATACCTCTTTACGGACTCCACTTA
11401 TGACTCCCTAAAGCCCATGTCGAAGCCCCCATCGCTGGGTCAATAGTACTTGCCGCAGTA
11461 CTCTTAAAACTAGGCGGCTATGGTATAATACGCCTCACACTCATTCTCAACCCCCTGACA
11521 AAACACATAGCCTACCCCTTCCTTGTACTATCCCTATGAGGCATAATTATAACAAGCTCC
11581 ATCTGCCTACGACAAACAGACCTAAAATCGCTCATTGCATACTCTTCAATCAGCCACATA
11641 GCCCTCGTAGTAACAGCCATTCTCATCCAAACCCCCTGAAGCTTCACCGGCGCAGTCATT
11701 CTCATAATCGCCCACGGGCTTACATCCTCATTACTATTCTGCCTAGCAAACTCAAACTAC
11761 GAACGCACTCACAGTCGCATCATAATCCTCTCTCAAGGACTTCAAACTCTACTCCCACTA
11821 ATAGCTTTTTGATGACTTCTAGCAAGCCTCGCTAACCTCGCCTTACCCCCCACTATTAAC
11881 CTACTGGGAGAACTCTCTGTGCTAGTAACCACGTTCTCCTGATCAAATATCACTCTCCTA
11941 CTTACAGGACTCAACATACTAGTCACAGCCCTATACTCCCTCTACATATTTACCACAACA
12001 CAATGGGGCTCACTCACCCACCACATTAACAACATAAAACCCTCATTCACACGAGAAAAC
```

12061 ACCCTCATGTTCATACACCTATCCCCCATTCTCCTCCTATCCCTCAACCCCGACATCATT
12121 ACCGGTTTTCCTCTTGTAAATATAGTTTAACCAAAACATCAGATTGTGAATCTGACAAC
12181 AGAGGCTTACGACCCCTTATTTACCGAGAAAGCTCACAAGAACTGCTAACTCATGCCCCC
12241 ATGTCTAACAACATGGCTTTCTCAACTTTTAAAGGATAACAGCTATCCATTGGTCTTAGG
12301 CCCCAAAAATTTTGGTGCAACTCCAAATAAAAGTAATAACCATGCACACTACTATAACCA
12361 CCCTAACCCTGACTTCCCTAATTCCCCCCATCCTTACCACCCTCGTTAACCCTAACAAAA
12421 AAAACTCATACCCCCATTATGTAAAATCCATTGTCGCATCCACCTTTATTATCAGTCTCT
12481 TCCCCACAACAATATTCATGTGCCTAGACCAAGAAGTTATTATCTCGAACTGACACTGAG
12541 CCACAACCCAAACAACCCAGCTCTCCCTAAGCTTCAAACTAGACTACTTCTCCATAATAT
12601 TCATCCCTGTAGCATTGTTCGTTACATGGTCCATCATAGAATTCTCACTGTGATATATAA
12661 ACTCAGACCCAAACATTAATCAGTTCTTCAAATATCTACTCATCTTCCTAATTACCATAC
12721 TAATCTTAGTTACCGCTAACAACCTATTCCAACTGTTCATCGGCTGAGAGGGCGTAGGAA
12781 TTATATCCTTCTTGCTCATCAGTTGATGATACGCCCGAGCAGATGCCAACACAGCAGCCA
12841 TTCAAGCAATCCTATACAACCGTATCGGCGATATCGGTTTCATCCTCGCCTTAGCATGAT
12901 TTTATCCTACACTCCAACTCATGAGACCCACAACAAATAGCCCTTCTAAACGCTAATCCA
12961 GCCTCACCCCACTACTAGGCCTCCTCCTAGCAGCAGCAGGCAAATCAGCCCAATTAGGTC
13021 TCCACCCCTGACTCCCCTCAGCCATAGAAGGCCCCACCCCAGTCTCAGCCCTACTCCACT
13081 CAAGCACTATAGTTGTAGCAGGAATCTTCTTACTCATCCGCTTCCACCCCCTAGCAGAAA
13141 ATAGCCCACTAATCCAAACTCTAACACTATGCTTAGGCGCTATCACCACTCTGTTCGCAG
13201 CAGTCTGCGCCCTTACACAAAATGACATCAAAAAAATCGTAGCCTTCTCCACTTCAAGTC
13261 AACTAGGACTCATAATAGTTACAATCGGCATCAACCAACCACACCTAGCATTCCTGCACA
13321 TCTGTACCCACGCCTTCTTCAAAGCCATACTATTTATGTGCTCCGGGTCCATCATCCACA
13381 ACCTTAACAATGAACAAGATATTCGAAAAATAGGAGGACTACTCAAAACCATACCTCTCA
13441 CTTCAACCTCCCTCACCATTGGCAGCCTAGCATTAGCAGGAATACCTTTCCTCACAGGTT
13501 TCTACTCCAAAGACCACATCATCGAAACCGCAAACATATCATACACAAACGCCTGAGCCC
13561 TATCTATTACTCTCATCGCTACCTCCCTGACAAGCGCCTATAGCACTCGAATAATTCTTC
13621 TCACCCTAACAGGTCAACCTCGCTTCCCCACCCTTACTAACATTAACGAAAATAACCCCA
13681 CCCTACTAAACCCCATTAAACGCCTGGCAGCCGGAAGCCTATTCGCAGGATTTCTCATTA
13741 CTAACAACATTTCCCCCGCATCCCCCTTCCAAACAACAATCCCCCTCTACCTAAAACTCA

13801 CAGCCCTCGCTGTCACTTTCCTAGGACTTCTAACAGCCCTAGACCTCAACTACCTAACCA
13861 ACAAACTTAAAATAAAATCCCCACTATGCACATTTTATTTCTCCAACATACTCGGATTCT
13921 ACCCTAGCATCACACACCGCACAATCCCCTATCTAGGCCTTCTTACGAGCCAAAACCTGC
13981 CCCTACTCCTCCTAGACCTAACCTGACTAGAAAAGCTATTACCTAAAACAATTTCACAGC
14041 ACCAAATCTCCACCTCCATCATCACCTCAACCCAAAAAGGCATAATTAAACTTTACTTCC
14101 TCTCTTTCTTCTTCCCACTCATCCTAACCCTACTCCTAATCACATAACCTATTCCCCCGA
14161 GCAATCTCAATTACAATATATACACCAACAAACAATGTTCAACCAGTAACTACTACTAAT
14221 CAACGCCCATAATCATACAAAGCCCCCGCACCAATAGGATCCTCCCGAATCAACCCTGAC
14281 CCCTCTCCTTCATAAATTATTCAGCTTCCTACACTATTAAAGTTTACCACAACCACCACC
14341 CCATCATACTCTTTCACCCACAGCACCAATCCTACCTCCATCGCTAACCCCACTAAAACA
14401 CTCACCAAGACCTCAACCCCTGACCCCCATGCCTCAGGATACTCCTCAATAGCCATCGCT
14461 GTAGTATATCCAAAGACAACCATCATTCCCCCTAAATAAATTAAAAAAACTATTAAACCC
14521 ATATAACCTCCCCCAAAATTCAGAATAATAACACACCCGACCACACCGCTAACAATCAAT
14581 ACTAAACCCCCATAAATAGGAGAAGGCTTAGAAGAAAACCCCACAAACCCCATTACTAAA
14641 CCCACACTCAACAGAAACAAAGCATACATCATTATTCTCGCACGGACTACAACCACGACC
14701 AATGATATGAAAAACCATCGTTGTATTTCAACTACAAGAACACCAATGACCCCAATACGC
14761 AAAACTAACCCCCTAATAAAATTAATTAACCACTCATTCATCGACCTCCCCACCCCATCC
14821 AACATCTCCGCATGATGAAACTTCGGCTCACTCCTTGGCGCCTGCCTGATCCTCCAAATC
14881 ACCACAGGACTATTCCTAGCCATGCACTACTCACCAGACGCCTCAACCGCCTTTTCATCA
14941 ATCGCCCACATCACTCGAGACGTAAATTATGGCTGAATCATCCGCTACCTTCACGCCAAT
15001 GGCGCCTCAATATTCTTTATCTGCCTCTTCCTACACATCGGGCGAGGCCTATATTACGGA
15061 TCATTTCTCTACTCAGAAACCTGAAACATCGGCATTATCCTCCTGCTTGCAACTATAGCA
15121 ACAGCCTTCATAGGCTATGTCCTCCCGTGAGGCCAAATATCATTCTGAGGGGCCACAGTA
15181 ATTACAAACTTACTATCCGCCATCCCATACATTGGGACAGACCTAGTTCAATGAATCTGA
15241 GGAGGCTACTCAGTAGACAGTCCCACCCTCACACGATTCTTTACCTTTCACTTCATCTTG
15301 CCCTTCATTATTGCAGCCCTAGCAACACTCCACCTCCTATTCTTGCACGAAACGGGATCA
15361 AACAACCCCCTAGGAATCACCTCCCATTCCGATAAAATCACCTTCCACCCTTACTACACA
15421 ATCAAAGACGCCCTCGGCTTACTTCTCTTCCTTCTCTCCTTAATGACATTAACACTATTC
15481 TCACCAGACCTCCTAGGCGACCCAGACAATTATACCCTAGCCAACCCCTTAAACACCCCT

15541 CCCCACATCAAGCC**CG**AATGATATTTCCTATT**CG**CCTACACAATTCTC**CG**ATC**CG**TCCCT
15601 AACAAACTAGGAGG**CG**TCCTTGCCCTATTACTATCCATCCTCATCCTAGCAATAATCCCC
15661 ATCCTCCATATATCCAAACAACAAAGCATAATATTT**CG**CCCACTAAGCCAATCACTTTAT
15721 TGACTCCTAGC**CG**CAGACCTCCTCATTCTAACCTGAAT**CG**GAGGACAACCAGTAAGCTAC
15781 CCTTTTACCATCATTGGACAAGTAGCATC**CG**TACTATACTTCACAACAATCCTAATCCTA
15841 ATACCAACTATCTCCCTAATTGAAAACAAAATACTCAAATGGGCCTGTCCTTGTAGTATA
15901 AACTAATACACCAGTCTTGTAAAC**CG**GAGATGAAAACCTTTTTCCAAGGACAAATCAGAG
15961 AAAAAGTCTTTAACTCCACCATTAGCACCCAAAGCTAAGATTCTAATTTAAACTATTCTC
16021 TGTTCTTTCATGGGGAAGCAGATTTGGGTACCACCCAAGTATTGACTCACCCATCAACAA
16081 C**CG**CTATGTATTT**CG**TACATTACTGCCAGCCACCATGAATATTGTA**CG**GTACCATAAATA
16141 CTTGACCACCTGTAGTACATAAAAACCCAATCCACATCAAAACCCCCTCCCCATGCTTAC
16201 AAGCAAGTACAGCAATCAACCCTCAACTATCACACATCAACTGCAACTCCAAAGCCACCC
16261 CTCACCCACTAGGATACCAACAAACCTACCCACCCTTAACAGTACATAGTACATAAAGCC
16321 ATTTAC**CG**TACATAGCACATTACAGTCAAATCCCTTCT**CG**TCCCCATGGATGACCCCCCT
16381 CAGATAGGGGTCCCTTGACCACCATCCTC**CG**TGAAATCAATATCC**CG**CACAAGAGTGCTA
16441 CTCTCCT**CG**CTC**CG**GGCCCATAACACTTGGGGGTAGCTAAAGTGAACTGTATC**CG**ACATC
16501 TGGTTCCTACTTCAGGGTCATAAAGCCTAAATAGCCCACA**CG**TTCCCCTTAAATAAGACA
16561 TCA**CG**

Appendix 5 Primers and sequence of target region of mtDNA: Please see attached DVD.

Appendix 6 Quality control Report for mtDNA: Please see attached DVD.

Appendix 7 Total read of CpG found in mtDNA :(Attached in DVD).

Appendix 8: The results of Regression Analysis of EpiTect System

[Data set]

Variable Entered/ Removed [a]

Model	Variables Entered	Variables Removed	Method
1	KCNQ1DN, NPTX2, GRIA2, TRIM58[b]	.	Enter

a. All requested variables entered

b. Dependent variables: age

Model Summary [b]

Model	R	R-square	Adjusted R square	Std. Error of the estimate	Statistics variation		
					R-squared variation	F	df1
1	0.848[a]	0.720	0.687,	10.80734	.720	22.448	4

a. Predictor: (Constant), KCNQ1DN, NPTX2, GRIA2, TRIM58

Model Summary [b]

Model	Statistics variation		Durbin-Watson
	df2	Note that the probability F Variation	
1	35	.000	1.534

a. Predicted value (constant), KCNQ1DN, NPTX2, GRIA2, TRIM58

b. Dependent variables: age

ANOVA [a]

Model	Sum of squares	DOF	Mean Squared	F	Sig.
Regression	10487.651	4	2621.913	22.448	.000[b]
Residuals	4087.949	35	116.799		
Total	14575.600	39			

a. Dependent variables: age

b. Predicted value (constant), KCNQ1DN, NPTX2, GRIA2, TRIM58

Coefficients [a]

Model	Unstandardized Coefficients		Standardized Coefficients			Collinearity
	B	Std. Error	Beta	t	Sig.	Tolerance
1 (constant)	3.400	8.690		.391	.698	
NPTX2	17.808	13.133	.138	1.356	.184	.775
GRIA2	18.577	14.725	.133	1.262	.215	.721
TRIM58	77.083	11.815	.705	6.524	.000	.686
KCNQ1DN	7.308	14.008	.063	.522	605	.544

a. Dependent variables: age

Model		Collinearity VIF
1	(constant)	
	NPTX2	1.290
	GRIA2	1.388
	TRIM58	1.457
	KCNQ1DN	1.839

a. Dependent variables: age

b.

Collinearity Diagnostic [a]

Model / Dimension	Eigenvalue	Condition Index	Variance proportions				
			Constant	NPTX2	GRIA2	TRIM58	KCNQ1DN
1	4.699	1.000	.00	.00	.00	.01	.00
1	.177	5.150	.03	.03	.00	.074	.00
2	.054	9.299	.01	.09	.92	.04	.08
3	.048	9.856	.20	.02	.01	.20	.76
4	.021	14.920	.76	.86	.07	.02	.16
5							

a. Dependent variables: age

A residual statistics

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted value	21.8443	87.0786	49.9000	16.39860	40
Residuals	-27.07857	27.93152	.00000	10.23813	40
Std. predicted value	-1.711	2.267	.000	1.000	40
Std. residuals	-2.506	2.584	.000	.947	40

a. Dependent variables: age

Appendix 9: The results of methylation fraction: The data of 80 blood samples (each repeated twice) were displayed as percentage of unmethylated and methylated fraction :(Attached in DVD).

Appendix 10: The EpiTect methylation result of other nine genes (Attached in DVD).

Appendix 11 Amount of DNA retrieved: The table shows the cartridge size and details of DNA amount which retrieved from each fired cartridge of eight tested guns.

Gun Type	Cartridge Size	Amount of DNA (picogram) retrieved from fired cartridge										Mean
Berta	9x19 mm	9.5	327.15	26.07	252.65	62.05	27.05	52.3	346.1	1.6	82	118.73
Browning	9x19 mm	135	14	24	0.7	8.5	3.1	240	195	50	405	107.53
Glock	9x19 mm	105.15	82.6	138.75	159.85	40.68	19.53	142.5	137.5	215.1	297	133.88
Kalashnikov	7.65 mm	17	49.6	6.5	180	93.5	26.5	29	6.5	25	146	118.11
Llama	7.65 mm	8	365	240	130	495	56	37	45.5	365	220	196.15
Makarov	9x18 mm	550	50	460	70	134	46.5	21.45	13.5	491.5	102	193.89
Shotgun	0.7 inch* 35mm	110	445	50	24	830	550	895	1000	395	450	587.66
Shotgun (Bore)	0.7 inch* 35mm	600	685	2045	1100	230	500	450	440	700	1955	870.5