Prediction of age from DNA

A thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

By

SUDINNA KULANGANA HEWAKAPUGE
BSc. (Science), MSc. (Biotechnology)

School of Biomedical Sciences
Faculty of Health, Engineering and Science
Victoria University
Melbourne
Australia
October 2009

Declaration

"I, Sudinna Kulangana Hewakapuge declare that the PhD thesis entitled Prediction of age from DNA is no more than 100,000 words in length, exclusive of tables, figures, appendices, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work".

Signature Date

Acknowledgments

My sincere gratitude goes to my principal supervisor, Dr. Swati Baindur-Hudson, for her constant willingness to give her time and mental energy for lengthy discussions during the planning and undertaking of this project and preparation of this thesis.

I am very grateful to Dr. Roland van Oorschot, at the Victoria Police Forensic Science Services Centre, for facilitating my research at these laboratories and constantly giving me guidance and help in collecting samples. My special thanks go to Dr. Irene Koukoulas in the Forensic Biology section in helping me in the DNA quantification.

My thanks go to Dr. Paul Lewandowski, at Deakin University, for his help in methodologies and critical analysis of my study. I would like to express my gratitude for his willingness to discuss and provide advice.

My gratitude goes to Victoria University for providing me with sufficient funds to carry out my PhD studies. My thanks go to Head of School, Associate Professor Grant Stanley, and all the staff in School of Molecular Sciences for their constant support.

My extra special thanks goes to Dr. Sarah Fraser for helping me with the Real Time PCR and Dr. Joshua Johnson for helping in DNA sequencing and analysis. Thanks to my dear friend Mr. Luc te Marvedale for helping me with the statistical analysis.

My thanks to all the past and present research students in School of Molecular Sciences, Victoria University, for their moral-boosting chats during the studies, in particular, Shruti Harnal, Idris Mohammad, Min-Tze Liong, Lydia Ong, Michelle Vaughn, Wun Keat-Ding, Anne Lise Tang, Warren Flentje, Thomas Dantoft, Natalie Nemaza, Lata Ramchandran and Debashree Lala Ray. Thanks to the technical staff at Victoria University, especially to Dale Tomlinson, Ira Prasatya, Stacey Lloyd, Mary Marshall, Melinda Dillon, Min Nguyen and Joseph Pelle for their assistance.

Thanks to staff in postgraduate office especially to Angela Rojter for correcting English in my manuscript and Lesley Birch, the scholarship coordinator for her constant support.

Thanks to my ever loving parents to and to my sister, for their constant support and love.

Abstract

Currently DNA profiling methods only compare a suspect's DNA with DNA left at the crime scene. When there is no suspect, it would be useful for the police to be able to predict what the person of interest looks like by analysing the DNA left behind in a crime scene. Determination of the age of the suspect is an important factor in creating an "identikit" (set of drawings of different features that can be put together to form the face of a person). This study investigated if one could use a correlation between telomere length and age, to predict the age of an individual from their DNA. Telomere length, in buccal cells, of 167 individuals aged between 1 and 96 years old was measured using quantitative real time PCR. Telomere length decreased with age ($r = -0.185$, $P<0.05$) and the age of an individual could be roughly determined by the following formula: (age = relative telomere length – 1.5 / -0.005). The regression (R^2) value between telomere length and age was ~ 0.04, which is too low to be used for forensics. The causes for the presence of large variation in telomere lengths in the population were further investigated. The age prediction accuracies were low even after dividing samples into non-related Europeans, males and females (5%, 9% and 1% respectively). Mean telomere lengths of eight age groups representing each decade of life showed a non-linear decrease in telomere length with age. There were variations in telomere lengths even among similarly aged individuals aged 26 years old (n = 10) and age 54 years old (n = 9). One of the factors that causes large inter individual variation could be the inheritance of telomere length. If there is a strong paternal or maternal influence, this could be incorporated into the age prediction formula. Parents' telomere lengths were compared with children's telomere lengths. No significant correlation was found between parents' telomere length and children's telomere lengths ($P > 0.05$) possibly due to small sample sizes. There was no

statistically significant influence of fathers' and mothers' ages at conception on individuals' telomere length ($P > 0.05$).

In this study we also analysed hypervariable region sequence variations and mitochondrial deletions within twelve separate families of maternal lineage, eleven of three generations and one with four generations (ages 5 to 96). Since mitochondria are maternally inherited, any sequence difference or differences in deletion pattern between the youngest member of the family compared with the others was considered to be age related. No age related mitochondrial mutation/s was observed in the twelve families analysed. Attempts were also made to analyse larger mitochondrial deletions in the sample. However, due to technical problems and lack of time, these studies could not be completed.

Therefore, telomere length measurement by Real-Time Quantitative PCR and direct sequencing of mitochondrial hypervariable regions cannot be used to predict the age of a person, from whom a biological sample was collected at a crime scene.

Publications

Journal Publications

Hewakapuge Sudinna, Van Oorschot Roland, Paul Lewandowski, Baindur-Hudson Swati (2008). Telomere length and its relationship to age. Legal Medicine, Volume 10, Issue 5, September 2008, Pages 236-242. Copy attached at the end.

Conference proceedings

Hewakapuge Sudinna, Van Oorschot Roland, Baindur-Hudson Swati (2008). Evaluation of mtDNA mutations to predict age. Forensic Science International Genetics Supplement Series, Volume 1, Issue 1, August 2008, pages 561-562. Copy attached at the end.

Oral Presentations

Hewakapuge Sudinna, Van Oorschot Roland, Baindur-Hudson Swati (2006) Analysis of telomere length and its relationship to age. 18th International Symposium of the Forensic Sciences Forensic Science: Classroom to Courtroom, 2 - 7 April 2006 Fremantle Western Australia.

Hewakapuge Sudinna, Van Oorschot Roland, Baindur-Hudson Swati (2006) Evaluation of the use of mitochondria DNA mutations to predict age. 18th International Symposium of the Forensic Sciences Forensic Science: Classroom to Courtroom, 2 - 7 April 2006 Fremantle Western Australia.

Poster Presentations

Hewakapuge Sudinna, Van Oorschot Roland, Baindur-Hudson Swati (2007) Evaluation of mtDNA mutations to predict age. 22nd Annual Congress of the International Society for Forensic Genetics, 21 – 25 August 2007 Copenhagen, Denmark.

Hewakapuge Sudinna, Van Oorschot Roland, Baindur-Hudson Swati (2007) Mitochondrial control region polymorphisms remain constant during ageing. International Society for Applied Biological Sciences (ISABS) Conference in Forensic Genetics and Molecular Anthropology, 3 – 7 September 2007 Split, Croatia.

Abbreviations

ANOVA – analysis of variance

ALFIE – allele frequency for the inference of ethnicity

ALT – alternative lengthening of telomeres

ATP – adenosine tri phosphate

Bp – base pairs

DNA – deoxyribonucleic acid

df – degrees of freedom

DGGE – denaturing gradient-gel electrophoresis

DHPLC – denaturing high pressure liquid chromatography

DIG – Digoxigenin

dNTP – deoxyribonucleotide tri phosphate

E. Coli - *Escherichia coli*

EDTA – ethylenediaminotetra-acetic acid

g - grams

HV1 – hypervariable region 1

HV2 – hypervariable region 2

Kb – kilo bases

M – molar

$MgCl_2$ – Magnesium Chloride

MDA - multiple displacement amplification

Mt. DNA – mitochondrial DNA

ml – millilitres

µl – micro litres

µm – micrometers

mRNA – messenger RNA

ng – nanograms

PCR – polymerase chain reaction

Q –FISH – quantitative fluorescence in situ hybridization

RNA – ribonucleic acid

tRNA – transfer RNA

rRNA – ribosomal RNA

RFLP – restriction fragment length polymorphism

ROS – reactive oxygen species

SDS – sodium dodecyl sulphate

Sig. - significance

SNP – single nucleotide tri phosphate

SSC – Sodium chloride- Sodium citrate

SSO - sequence-specific oligonucleotide

STELA – single telomere length analysis

STR – shot tandem repeat

TERC – telomere RNA component

TERT – telomere reverse transcriptase

TRF1 – telomere repeat binding factor 1

TRF 2 – telomere repeat binding factor 2

UV – ultraviolet

V - volts

V/v – volume per volume

W/v – weight per volume

XL PCR – extra long polymerase chain reaction

Table of contents

List of Figures and Tables

Chapter 1

Introduction and aims

1.1 Introduction

It would be useful to investigators if DNA left at the scene of a crime could act as an "eye-witness", i.e. give some clue as to what a person of interest looks like. The ability to predict the outward appearance of an individual such as age, height, weight, ethnicity, hair and skin colour would enhance the efficiency of criminal investigation by narrowing down the range of suspects. The aim of this research is to investigate whether the age of a person could be predicted by analysing DNA of an individual. The prediction of age and other physical characteristics would not be used as evidence in courts but rather as an investigative tool to narrow down the range of suspects which would be very useful in forensics as an investigative tool. In addition, it may be assist in the identification of an unknown victim of crime or disaster.

To predict age from DNA, we need to study the changes DNA undergoes with the ageing process. Ageing is a complex procedure and involves multiple mechanisms at different levels and stages in life. There are numerous theories which attempt to explain the possible causes and mechanisms that DNA under go with ageing (Brierley *et al.* 1997). Two main theories were considered: "telomere hypothesis" on ageing and "mitochondrial theory on ageing".

The "telomere hypothesis" suggests that the continuous shortening of telomeres during cell division is primarily responsible for ageing (Boukamap 2001). Telomeres are specialised DNA-protein structures found at the ends of linear chromosomes (Greider 1999). The main role of telomeres is to protect chromosome ends from

recombination, fusion and from being recognised as damaged DNA (Greider 1998). Telomeric DNA consists of simple tandemly repeated sequences characterized by clusters of G residues in one strand. In humans it is 5′-TTAGGG-3′ (Blackburn 1990). During DNA replication the lagging strand is replicated in so-called Okazaki fragments. The synthesis of these 5′→3′ fragments requires RNA primers, which are finally replaced by DNA with the exception of outermost ones. Therefore the far end of the telomere is not replicated and the telomeres shorten with each round of DNA replication (Boukamap 2001) and reaches a critical length, which triggers cellular senescence (Bekaert *et al.* 2004). Even though there is no direct correlation between cellular senescence and ageing (Bekaert *et al.* 2004), both *in vitro* (Harely *et al.* 1990) and *in vivo* (Lindsey *et al.* 1991) studies show that telomere length decreases with age. Germ line cells and tumour cells of higher eukaryotes maintain their telomere lengths by expression of the enzyme telomerase (Rattner 1995). In some tumour cells the telomere length is kept constant by a mechanism called alternative lengthening of telomeres (Neumann *et al.* 2002).

Telomere length has been shown to be inherited, in which inheritance values varied depending on the sample population analysed (Graakjaer *et al.* 2004; Njajou *et al.* 2007; Rufer *et al.* 1999; Slagboom *et al.* 1994). The mode of inheritance of telomere length is under debate, since some studies hypothesis X-linked inheritance of telomere lengths (Nawrot *et al.* 2004), while others hypothesis paternal inheritance of telomere lengths (Njajou *et al.* 2007; Nordfjäll *et al.* 2005). The traditional method of measuring telomere length is to determine mean Terminal Restriction Fragment (TRF) length using Southern blots. The two previous studies attempted by Tsuji *et al.,* (Tsuji *et al.* 2002) and Takasaki *et al.,* (Takasaki *et al.* 2003) on predicting age

by analysing telomere length have used the TRF method. This method requires large amounts of DNA (0.5-5μg/individual) and time (3-5 days). Often the amount of DNA available for forensic analysis is low. Therefore, in this study the method developed by Cawthon (Cawthon 2002; Njajou *et al.* 2007; Nordfjäll *et al.* 2005) which involves real time PCR was used, which requires lesser amounts of DNA and time for analysis.

According to the "mitochondrial theory of ageing", the accumulation of mitochondrial mutations during life is a major contributor to the ageing process (Linnane, AW *et al.* 1989). Mitochondria are the cellular organelles that produce most of the energy used by a cell. They produce more than 80% of the energy in the form of the adenosine triphosphate (ATP) needs of a cell. The human mitochondria contain multiple copies of a 16,569 bp-closed circular genome that is replicated and expressed within the organelle system (Larsson *et al.* 1995). Anderson *et al,* published the sequence and organization of the human mitochondrial genome in 1981 (Anderson *et al.* 1981). Today the original sequence is considered the reference sequence to which new sequences are compared, and it is commonly known as the Anderson sequence or the Cambridge reference sequence (Butler 2005). Proteins encoded by mitochondrial DNA (mt DNA) are important subunits of the respiratory chain and ATP synthase (Nagley *et al.* 1998).

Mt DNA is a useful target for the analysis of forensic materials because of its high copy number and maternal inheritance (Butler 2005). A large number of mtDNA mutations (point mutations, deletions and duplications) have been reported to occur in human somatic tissues, many of which accumulate during the ageing process. This

is mainly because mitochondria are the major source of reactive oxygen species (ROS) (Kang *et al.* 1998). Mitochondrial DNA is located inside the mitochondrial inner membrane where reactive oxygen species (ROS) always escape from the respiratory chain causing mtDNA mutations. In addition, mtDNA has an less comprehensive repair system compared with nuclear DNA (Kang *et al.* 1998). The mutations found in mitochondria are classified as large deletions, point mutations and small duplications (Nagley *et al.* 1998). The high copy number poses a problem when analysing mitochondrial DNA mutations. Unlike in nuclear DNA, there is more than one mtDNA sequence within an individual cell and both wild-type and mutations could exist together: a condition known as heteroplasmy. The level of heteroplasmy also been reported to accumulate with age (Calloway *et al.* 2000). In previous research, aimed at predicting the age of an individual, Meissner *et al.* (Meissner, C *et al.* 2006) and N. Von Wurmb-Schwark *et al.* (Wurmb-Schwark *et al.* 2002) analysed the 5kb deletion. Both the studies observed that variability of the 5kb mutation between individuals of the same age groups was relatively large and they could only discriminate between "young" and "old" samples. Investigating any one single deletion or point mutation may result in a distorted perspective as to the extent of somatic mtDNA mutation accumulation in the ageing process (Kovalenko *et al.* 1998).

Thus, this study was designed to analyse mutations and levels of heteroplasmy in the hypervariable region 1 and 2 (HV1 and HV2) of the mitochondrial genome, as well as the presence of mitochondrial deletions in the whole mitochondrial genome. The mitochondrial study sample was composed of maternal lineage families. Mitochondrial sequences and deletion patterns were compared within a family, using

the youngest members' sequence or deletion pattern as the reference. Mitochondria are maternally inherited and, therefore, any sequence/deletion pattern difference observed within a family was considered as age related, with the assumption that most maternally related individuals started off with the same sequence.

Buccal swabs were collected for this study, since it is a non-invasive method of collecting DNA from individuals, and because saliva is one of the major types of potential biological material found at crime scenes.

1.2 Aims of the project

Age would be one of the key characteristic in establishing a physical character profile of an individual. It is already one of the criteria listed in the personal physical description form asked to be filled out by eye-witnesses in police investigation. As stated above, no validated molecular method has been developed to predict the age of an individual by analysing DNA. So far, in previous attempts the methods could only provide a 'rough estimation'. The main aim of this study was to develop a reliable method to predict the age of an individual, which would be applicable in forensic investigations.

1.2.1 Specific Aims of the project

- To determine the correlation between age and telomere length from a population consisting of various age groups.

- To investigate the influence of gender and ethnicity on telomere length of individuals

- To examine inheritance patterns of telomere length, for example, if they are linked to the X chromosome or paternally inherited.

- To examine how mitochondrial mutations/deletions accumulate with age in three/four-generation maternal linage families.

The achievement of these aims would provide methodology to predict the age of an individual whose biological sample was inadvertently left behind at the scene of a crime. The following chapters describe the literature review, methods used to address these aims, results obtained and discuss the extent to which the results achieved the aims. The last chapter summarises the results obtained, provides a general discussion and suggestions for future directions.

Chapter 2

Literature Review

2.1 Use of DNA in forensic science

Every human being has genetic information in the form of DNA which is unique to an individual, except in the case of identical twins (Butler 2005). DNA is present in all nucleated cells and is therefore present in biological samples left behind in a crime scene (Butler 2005). The use of DNA in forensics was first described by Dr. Alec Jeffrey in 1985 (Jeffreys *et al.* 1985). Dr. Jeffery discovered variable number tandem repeats (VNTRs) in non-coding regions of DNA and the fact that they are specific to an individual. He developed a technique that could detect the VNTRs in humans, which is called DNA profiling (Butler 2005). The VNTRs were detected by using a technique called Restriction Fragment Length Polymorphism (RFLP). The most common sources of DNA in a crime scene could come from blood, semen, bones, teeth, hair, saliva, urine, or faeces (Solomon *et al.* 2008). The discovery of the Polymerase Chain Reaction (PCR) method by Kary Mullis in 1986 revolutionised DNA fingerprinting, enabling scientists to use tiny amounts of amplifiable DNA (Butler 2005). Thus, DNA obtained from the small number of cells present in cigarette butts, licked envelopes or postage stamps, dandruff, fingerprints, razor blades, chewing gum, wristwatches, earwax, debris from under fingernails, and toothbrushes could be used to develop a DNA profile of an individual (Butler 2005). Currently, short tandem repeat (STR) loci, which have higher discriminating power than VNTRs, and can be amplified by PCR without the problems of differential amplification, are used in forensic DNA profiling (Butler 2005). Human forensic casework is now done using commercially developed kits that include autosomal STR multiplexes (single-tube PCR reactions that amplify multiple loci) (reviewed in)

(Jobling *et al.* 2004). In the United States of America 13 STR loci are used while the United Kingdom and much of Europe use 10 STR loci with XY homologous amelogenin gene, which is as determinant due to size differences in the gene between males and females (Butler 2006). Australian forensic laboratories use 9 STR loci along with the sex determining gene (Database 2008). The current discriminating power of DNA profiling is one in a trillion among unrelated individuals (Butler 2005). DNA profiles of the victim and potential suspects are compared with profiles from DNA left at the scene of a crime (Butler 2005). If no match is found then the resultant DNA profile is compared against the DNA database of the state or of the country (Tully 2007). The UK, in 1994, was the first country to begin their database called National DNA Database (NDNAD) (Home Office 2007) while the USA soon followed with their database called the combined DNA index system (CODIS) (Investigation 2007). Many other countries, have since introduced similar databases and supporting legislation. Here in Australia, the National Criminal Investigation DNA Database (NCIDD) was established in 2001 (Database 2008). The probability of two people having the same DNA profile varied from $<3 \times 10^{-11}$ to 5×10^{-19} (reviewed in (Jobling *et al.* 2004)). If applied properly, DNA fingerprinting has the power to identify the guilty and exonerate the innocent. The main point is that the forensic geneticist is not to make presumptions of guilt or innocence but to provide unbiased information to judge and jury (Jobling *et al.* 2004)**.**

The majority of genetic materials in humans are in the nucleus of each cell, and the energy producing organelle, the mitochondrion also contains its own genome. The nuclear genome accounts for 99.9995% of the total genetic information, and the mitochondrial genome accounts for the remaining 0.0005% (Strachan *et al.* 1999).

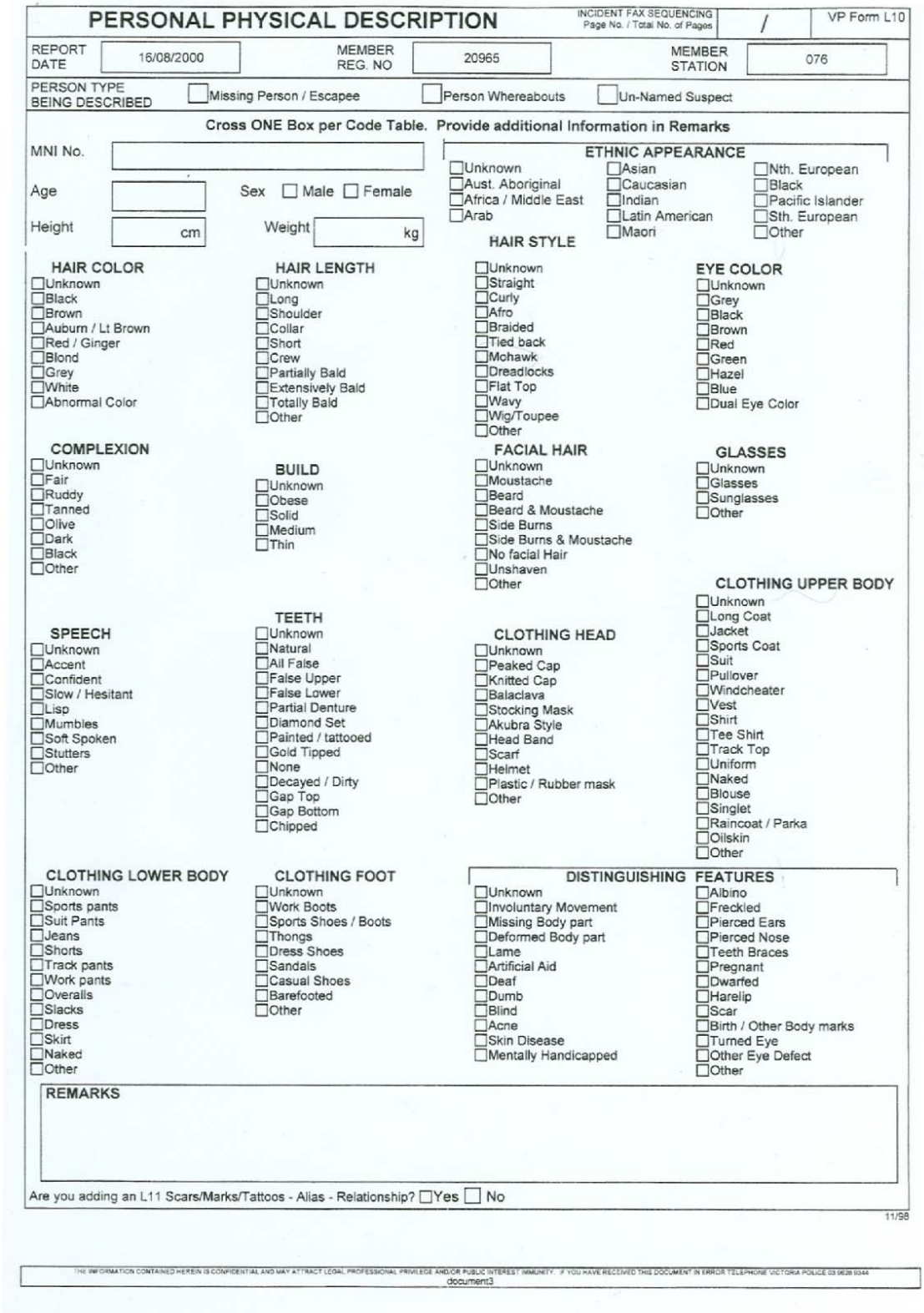
Mitochondrial DNA is used in individual identification purposes, mainly when nuclear DNA is degraded and/or when extracted DNA amount is very low (Butler 2005). Mitochondria are maternally inherited (Giles, RE *et al.* 1980) and have thousands of copies per cell compared to two copies of nuclear DNA (Robin, ED *et al.* 1988); this makes mitochondrial DNA very useful in highly degraded specimens and helps to identify people by comparison of sequences with maternal relatives. The famous example of the use of mitochondrial DNA in forensics is the identification of the remains of the Russian royal family (Butler 2005).

If the investigators find a match between crime scene samples and suspects, and can proceed with their investigations, further details might not be required from the DNA. But if there is no match and no suspects, no eyewitnesses (or unreliable eyewitness reports) to a crime, an investigation can be stalled for years (Genomics 2007). It would be useful to investigators if DNA from biological tissue inadvertently left at the scene of a crime could act as an "eye-witness", that is give some clues as to what a person of interest looks like.

2.2 Prediction of physical characters

Currently used DNA profiling methods can only include or exclude a person of interest in a crime investigation. It would be helpful for the police if further information could be extracted from the DNA. Obtaining a character or physical profile of suspect is often the primary profile step of an investigation. Usually an eye witness of a crime scene is asked to fill in a form called a personal physical description form (Figure 2.1) regarding the possible perpetrator. To predict physical characters from DNA, characters must have a genetic link. The main physical

characters included in the description form that have a clear genetic link include: hair colour, eye colour, and height. Other traits such as weight, body build, some rare forms of teeth presentation and age may also be determined by one's genetic blueprint. So if a test could be developed to predict the above mentioned physical characters of an individual by analysing DNA it would be valuable for the investigators. The information obtained would not be directly used to bring a person to justice. Rather it would help the police in these investigations. Since physical characters of a person are greatly influenced by the environment, predicted characters would not be an exact match. For example a person may have the genes to be tall but, may not attain his/her full height due to malnutrition as a child. A person could shave, dye his/her hair or wear a wig, and wear different coloured contact lenses. Practical values of the physical character prediction should also be taken into account. In ethnicity prediction, if the perpetrator is predicted to be of African ancestry, further tests on hair and eye colour are not much of use. On the other hand, if the perpetrator is predicted to be a European and the majority of suspects are Europeans it is of little help. Therefore just predicting one character has less importance. Obtaining all the information from DNA, for example, analysing DNA at a crime scene and predicting that the perpetrator is a European male, brown haired, has hazel eyes and in his early thirties, is of much greater value.



PERSONAL PHYSICAL DESCRIPTION

INCIDENT FAX SEQUENCING Page No. / Total No. of Pages / VP Form L10

REPORT DATE 16/08/2000 MEMBER REG. NO 20965 MEMBER STATION 076

PERSON TYPE BEING DESCRIBED ☐ Missing Person / Escapee ☐ Person Whereabouts ☐ Un-Named Suspect

Cross ONE Box per Code Table. Provide additional Information in Remarks

MNI No.

Age Sex ☐ Male ☐ Female

Height cm Weight kg

ETHNIC APPEARANCE
☐ Unknown ☐ Asian ☐ Nth. European
☐ Aust. Aboriginal ☐ Caucasian ☐ Black
☐ Africa / Middle East ☐ Indian ☐ Pacific Islander
☐ Arab ☐ Latin American ☐ Sth. European
☐ Maori ☐ Other

HAIR COLOR
☐ Unknown
☐ Black
☐ Brown
☐ Auburn / Lt Brown
☐ Red / Ginger
☐ Blond
☐ Grey
☐ White
☐ Abnormal Color

HAIR LENGTH
☐ Unknown
☐ Long
☐ Shoulder
☐ Collar
☐ Short
☐ Crew
☐ Partially Bald
☐ Extensively Bald
☐ Totally Bald
☐ Other

HAIR STYLE
☐ Unknown
☐ Straight
☐ Curly
☐ Afro
☐ Braided
☐ Tied back
☐ Mohawk
☐ Dreadlocks
☐ Flat Top
☐ Wavy
☐ Wig/Toupee
☐ Other

EYE COLOR
☐ Unknown
☐ Grey
☐ Black
☐ Brown
☐ Red
☐ Green
☐ Hazel
☐ Blue
☐ Dual Eye Color

COMPLEXION
☐ Unknown
☐ Fair
☐ Ruddy
☐ Tanned
☐ Olive
☐ Dark
☐ Black
☐ Other

BUILD
☐ Unknown
☐ Obese
☐ Solid
☐ Medium
☐ Thin

FACIAL HAIR
☐ Unknown
☐ Moustache
☐ Beard
☐ Beard & Moustache
☐ Side Burns
☐ Side Burns & Moustache
☐ No facial Hair
☐ Unshaven
☐ Other

GLASSES
☐ Unknown
☐ Glasses
☐ Sunglasses
☐ Other

CLOTHING UPPER BODY
☐ Unknown
☐ Long Coat
☐ Jacket
☐ Sports Coat
☐ Suit
☐ Pullover
☐ Windcheater
☐ Vest
☐ Shirt
☐ Tee Shirt
☐ Track Top
☐ Uniform
☐ Naked
☐ Blouse
☐ Singlet
☐ Raincoat / Parka
☐ Oilskin
☐ Other

SPEECH
☐ Unknown
☐ Accent
☐ Confident
☐ Slow / Hesitant
☐ Lisp
☐ Mumbles
☐ Soft Spoken
☐ Stutters
☐ Other

TEETH
☐ Unknown
☐ Natural
☐ All False
☐ False Upper
☐ False Lower
☐ Partial Denture
☐ Diamond Set
☐ Painted / tattooed
☐ Gold Tipped
☐ None
☐ Decayed / Dirty
☐ Gap Top
☐ Gap Bottom
☐ Chipped

CLOTHING HEAD
☐ Unknown
☐ Peaked Cap
☐ Knitted Cap
☐ Balaclava
☐ Stocking Mask
☐ Akubra Style
☐ Head Band
☐ Scarf
☐ Helmet
☐ Plastic / Rubber mask
☐ Other

CLOTHING LOWER BODY
☐ Unknown
☐ Sports pants
☐ Suit Pants
☐ Jeans
☐ Shorts
☐ Track pants
☐ Work pants
☐ Overalls
☐ Slacks
☐ Dress
☐ Skirt
☐ Naked
☐ Other

CLOTHING FOOT
☐ Unknown
☐ Work Boots
☐ Sports Shoes / Boots
☐ Thongs
☐ Dress Shoes
☐ Sandals
☐ Casual Shoes
☐ Barefooted
☐ Other

DISTINGUISHING FEATURES
☐ Unknown
☐ Involuntary Movement
☐ Missing Body part
☐ Deformed Body part
☐ Lame
☐ Artificial Aid
☐ Deaf
☐ Dumb
☐ Blind
☐ Acne
☐ Skin Disease
☐ Mentally Handicapped
☐ Albino
☐ Freckled
☐ Pierced Ears
☐ Pierced Nose
☐ Teeth Braces
☐ Pregnant
☐ Dwarfed
☐ Harelip
☐ Scar
☐ Birth / Other Body marks
☐ Turned Eye
☐ Other Eye Defect
☐ Other

REMARKS

Are you adding an L11 Scars/Marks/Tattoos - Alias - Relationship? ☐ Yes ☐ No

11/98

THE INFORMATION CONTAINED HEREIN IS CONFIDENTIAL AND MAY ATTRACT LEGAL PROFESSIONAL PRIVILEGE AND/OR PUBLIC INTEREST IMMUNITY. IF YOU HAVE RECEIVED THIS DOCUMENT IN ERROR TELEPHONE VICTORIA POLICE 03 9628 0344
document3

Figure 2.1 - An example of a personal physical description form given to an eye witness by police.

2.2.1 Physical character prediction and biological ancestry tests already been developed

One of the main physical characters, the gender of a person can be predicted with currently used DNA typing methods. This is done by analysing the size of the amelogenin gene on the X chromosome (Butler 2005) and has been used in forensic labs around the world. Tests for biological ancestry prediction and the prediction of other physical characters are already being used by police in the UK and USA (Genomics 2007; Home Office 2007). For example, the Forensic Science Service department in the UK uses tests that can predict the ethnicity and red hair colour of an individual by analysis of his/her DNA. The ethnicity prediction test compares six STR loci database (Lowe *et al.* 2001), using a software called ALFIE (allele frequency for the inference of ethnicity) and can predict if the DNA came from one of the 5 major ethnic groups in Britain (European, Afro-Caribbean, Indian Subcontinent, South East Asian and Middle Eastern). The red hair colour test was developed by Grimes *et al.,* in 2001 (Grimes *et al.* 2001), and analyses two types of mutations in human melanocortin 1 receptor gene (MC1R). DNAPrint Genomics (Genomics 2007), a company based in the USA, has several kits that claim to be able to predict eye colour and biological ancestry of an individual. One of the kits made by them, DNAWitness™ 2.5, is used by the police in the USA to predict biological ancestry of an individual. DNAWitness™ 2.5 is claimed to be able to predict if an individual's ethnicity is Sub-Saharan African, Native American, East Asian or European by analysing ancestry markers on DNA. Recently the same company developed a test called European DNA 2.0 that is claimed to be able to predict which European sub ancestry an individual belongs to (South Eastern European, Iberian, Basque, Continental European and North Eastern European) by using 1349 ancestry

informative markers. The company also has an eye colour test called RETINOME™, in which an assay uses the polymorphisms in genes determining iris colour in combination with ancestry markers (Frudakis *et al.* 2003).

This study was aimed at predicting age from DNA extracted from biological tissue left behind in a crime scene. Most human physical characters are the result of the expression of a gene or combination of genes. Unlike the above mentioned physical character prediction tests an individual's age cannot be predicted by analysing a gene or combination of genes. However, analysing the natural instability of both nuclear and mitochondrial DNA of an individual (discussed in section 2.4) may be useful in predicting an individual's age.

2.3 Prediction of age

Age estimation of unidentified corpses and skeletons has long been practised in forensic investigations (Schmeling *et al.* 2007). Predicting the age of a person not only has direct relevance, it could also be helpful in interpreting other genetic traits. For example, a 50 year old with genetic determinants that indicate red hair might no longer appear red-headed and an 18 year old male with genetic determinants for a particular type of baldness may not yet appear bald.

Previous studies have focused on predicting age at the time of death (Meissner, C *et al.* 2006; Meissner, Christoph *et al.* 1999; Takasaki *et al.* 2003) as well as on predicting the age of a living person (Alvarez *et al.* 2006; Centre 2008; Tsuji *et al.* 2002; Wurmb-Schwark *et al.* 2002). Studies by Takasaki *et al.,* (Takasaki *et al.* 2003) and Tsuji *et al.,* (Tsuji *et al.* 2002) were carried out to predict age by analysing

telomere length shortening. Alvaraz *et al.,* (Alvarez *et al.* 2006) analysed the pattern of gene expression of isoforms of gamma haemoglobin and studies done by Meissner *et al.,* (Meissner, C *et al.* 2006) and Wurmb-Schwark *et al.,* (Wurmb-Schwark *et al.* 2002) analysed mitochondrial 5kb deletion, to develop a method to predict age of an individual. Both Takasaki *et al.,* (Takasaki *et al.* 2003) ($R^2 = 0.562$) and Tsuji *et al.,* (Tsuji *et al.* 2002) ($R^2 = 0.6922$) devised formulas to roughly predict the age of a person by analysing telomere length. In the Alvaraz *et al.,* (Alvarez *et al.* 2006) study, the two isoforms of gamma haemoglobin showed a restricted pattern of gene expression specific to newborns that could be used to differentiate newborns from older samples (Alvarez *et al.* 2006). Both studies done on 5kb mitochondrial deletion (Meissner, C *et al.* 1997; Wurmb-Schwark *et al.* 2002) developed formulas that could roughly predict the age of the individual and distinguish individuals as young and old (Meissner, C *et al.* 2006; Meissner, Christoph *et al.* 1999; Wurmb-Schwark *et al.* 2002). However, at present, there is no validated molecular assay to predict the age of an living individual (discussed in (Alvarez *et al.* 2006)). A reliable method developed to predict age would be very useful in forensics as an investigative tool.

There has also been interest in prediction of age in animals, especially in wild animal populations, where in some cases; prediction of age is difficult by conventional ecological methods, where the animal has to be killed. For example, researchers at the Whale Research Centre Southern Cross University have started work on analysing telomere lengths in humpback whales using DNA from skin samples (Centre 2008). Development of a method that could predict age in humpback whales could contribute towards ending the current dispute over killing of whales for

research. Age structure data of the wild animal population is important for managing, understanding and conserving them.

2.4 Ageing

To predict the age of an individual using DNA, it is necessary to have understanding of the ageing process at molecular level. Ageing is commonly observed as a process that causes progressive impairment of function, increases the vulnerability to environmental challenges, diseases , decrease in fertility and death (Kirkwood 2005). It is unlikely that some specific genes exist that causes ageing. Ageing occurs largely due to accumulation of somatic damage, due to inefficiencies in maintenance and repair mechanisms. Numerous models have been proposed to explain the mechanisms of ageing at the molecular level. The major theories are:

1. Telomeres (ends of chromosomes) get shortened in each round of cell division causing ageing (Harley 1991).
2. Mutations accumulate in the mitochondrial genome with age which cause ageing (Linnane, AW *et al.* 1989).
3. In the case of age-related increase in somatic mutations and other forms of DNA damage, the capacity of DNA repair is a key factor on determining its life span of an organism (Promislow 1994). The enzyme poly (ADP-ribose) polymerase-1 (PARP-1) is expressed during stress- induced DNA damage. Higher levels of PARP – 1 have been observed in organisms that have a longer lifespan (Kirkwood 2005).
4. The decline in protein metabolism in cells causes accumulation of waste materials, such as various oxidized proteins, non-degradable material and

lipofuscin (age pigment). This decline of protein metabolism functionally affects mitochondria and causes ageing (Terman *et al.* 2004).

The most recent hypothesis on ageing is:

5. Changes or instability to the nuclear architecture (heterochromatin, euchromatin and the nuclear matrix) causes ageing (Oberdoerffer *et al.* 2007).

Although there is evidence to support each of the above theories by itself, with correlations to age-related frailty, disability, and disease, a. recent development is the "network" theories of ageing in which all the mechanisms are considered together (Kirkwood 2005).

The most common biological samples left in a crime scene are blood, semen and saliva, from which DNA can readily be obtained. Thus this study was based on the first two mentioned theories on ageing: the telomere hypothesis of ageing and the mitochondrial theory of ageing. Experiments based on these utilise DNA and thus these were used as the basis for the research in this thesis. Furthermore, as the amount of DNA obtainable from such samples is often low (Butler 2005), techniques based on the Polymerase Chain Reaction (PCR) were selected, as small amounts of DNA are required here.

From the point of view of the aims of this project, only the first two phenomena listed above would be measurable at detectable levels from small amounts of samples left behind at the scene of a crime. For this reason, only the first two theories are discussed below.

2.4.1 Telomere hypothesis of ageing

The telomere hypothesis of ageing was developed to explain the loss of the proliferative ability of normal animal cells first observed by Hayflick *et al.,* in 1961 (Harley 1991). The hypothesis is, "Proliferation-dependent, continuous shortening of the specific structures at the ends of the chromosomes, the telomeres, is the counting mechanism- the internal clock of ageing" (Boukamap 2001).

2.4.2 Mitochondria theory of ageing

In 1989 Linnane *et al*., (Linnane, AW *et al.* 1989) proposed the mitochondrial theory of ageing which is "The somatic accumulation of mitochondrial mutations and the subsequent cytoplasmic segregation of these mutations during life is a major contributor to the gradual loss of cellular bioenergetic capacity within tissues and organs associated with general senescence and diseases of ageing".

2.5 Telomeres

The end of the chromosome is called the telomere and represents the end of the coiled chromatid fibre as well as the end of the single DNA strand, which extends along the length of the chromatid (Rattner 1995). Telomeres are specialized DNA-protein structures (Figure 2.2). The main role of telomeres is to protect chromosome ends from recombination and fusion, and also to prevent them from being recognised as damaged DNA (Greider 1998). Telomeres had been identified by B. McClintock and defined by H.J. Muller, in 1938, as the functional chromosome elements that protect chromosome ends (Greider *et al.* 2004). Muller coined the term "telomere" from the Greek for "end" (telos) and "part" (meros) (Alberts *et al.* 2008). Telomeric DNA consists of simple tandem repeated sequences characterized by clusters of G

residues in one strand. In humans this sequence is 5'-TTAGGG-3', although variant forms such as TTGGGG and TGAGGG exist sub terminally (Lindsey *et al.* 1991). In humans telomere lengths vary from 3 to 20 kb (Harely *et al.* 1990). The number of copies of this basic repeat unit in telomeres varies from species to species, from chromosome to chromosome within a species and even on the same chromosome at different stages of the life cycle (Lewin 1999). Previously, it was thought that no nucleosomes were present within the telomeric region and that hence this region was more accessible to damage (Rattner 1995). Recent studies, however, show that in higher eukaryotes, about 80% of telomeric DNA is organised in tightly packed nucleosomes separated by 10-20bp of linker DNA (Pisano *et al.* 2008).

Human somatic cells have 92 telomeres (Lewis 1998). A more complex set of repeats called subtelomeric/telomere associate repeats are present just internal to the telomeric repeats. Their sequences are not conserved in eukaryotes and their function is unknown (Strachan *et al.* 1999). Telomere sequences do not contain protein-encoding genes but are critical for the preservation of genome integrity (Bekaert *et al.* 2004). The $3^{/}$ end of the telomeric DNA is the G-rich telomeric strand and it forms a $3^{/}$ terminal overhang, 50 to 100 nucleotides in length in humans , protruding from the duplex (Greider 1999). The current model for the telomeric complex is that it forms a T-loop structure. According to the model, the single stranded G-rich overhang at the telomere end invades the double stranded telomere strand, and in combination with telomere binding proteins TRF1 and TRF2, forms the T-loop structure (Figure 2.3) (Griffith *et al.* 1999).

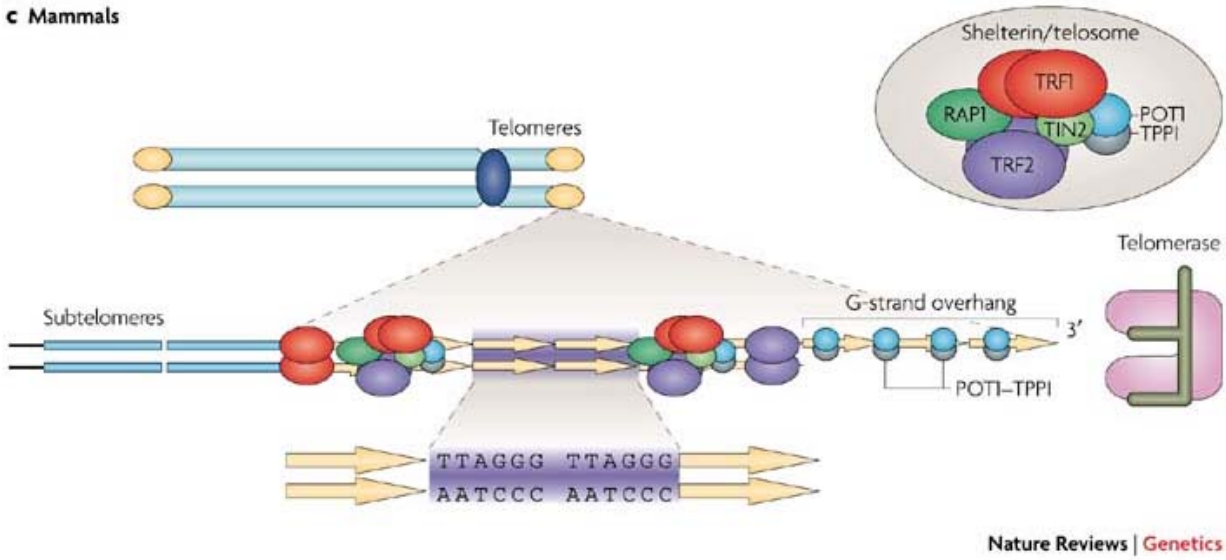


Figure 2.2 – The structure of the mammalian telomere showing telomere repeats, subtelomeres, and G strand overhang. {reproduced from Blasco (Blasco 2007)}.

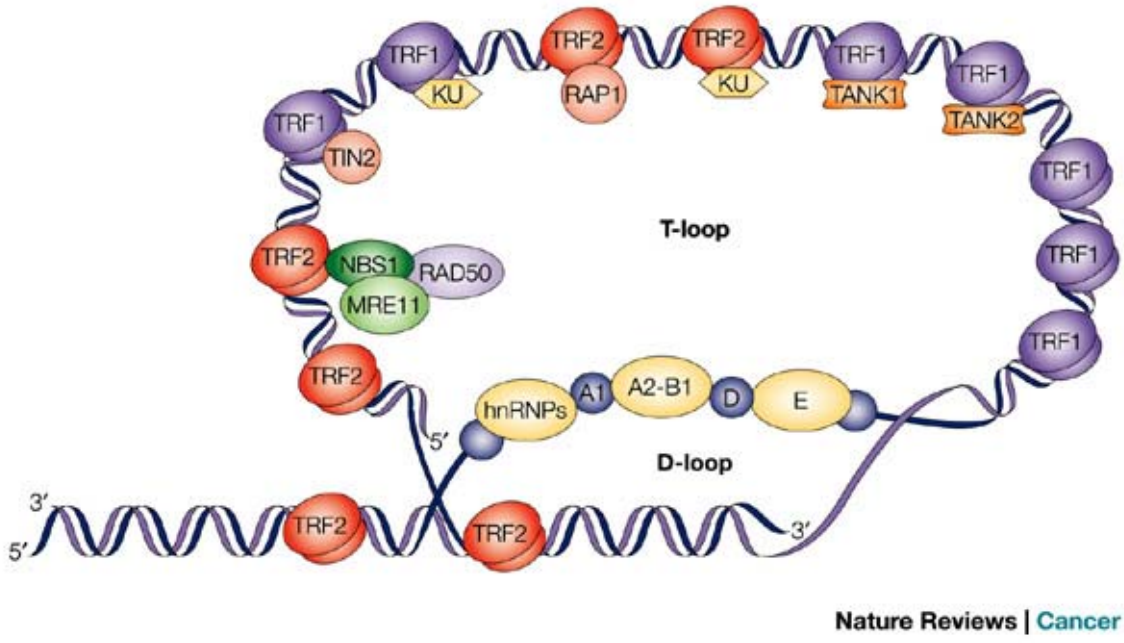


Figure 2.3 - Telomere T loop structure. {reproduced from Neumann et al., (Neumann et al. 2002)}.

2.6 Telomere length and replicative senescence

Cells grown in culture have a limited number of cell divisions and undergo cellular senescence after approximately 50 cell divisions. This was first observed by Hayflick and Moorhead in 1961 in cultured human fibroblasts (Harley 1991). Average normal cells that undergo a finite number of cell divisions and ultimately stop dividing undergo a process called the "replicative senescence" (Kufe *et al.* 2003). There are many theories to explain the "replicative senescence". One of the most favoured is the telomere hypothesis on ageing, which was first proposed as the "marginotomy theory" by Olovkinov in 1971 (Harley 1991).

According to Olvkinov, in each round of DNA replication, ends of chromosomes/telomeres get deleted and after a certain number of cell divisions, a critical deletion occurs and this causes cell death (Harley 1991). Watson and Olovnikov (1971, 1973) independently described the "end replication" problem, which ultimately causes telomere length to be shortened with each round of cell division (Harley 1991).

When a cell divides, the telomere is not fully replicated because of limitations of the DNA polymerases in completing the replication of the ends of the linear chromosomes (Chan, SR *et al.* 2004). During DNA replication, the lagging strand is replicated in discontinuous pieces called Okazaki fragments. A the synthesis of 5' →3' DNA fragment requires RNA primers, which are finally replaced by DNA with the exception of those at the very end (Figure 2.4). Because of this, the far end of the telomere is not replicated and the telomeres shorten with each round of DNA replication (Boukamap 2001).

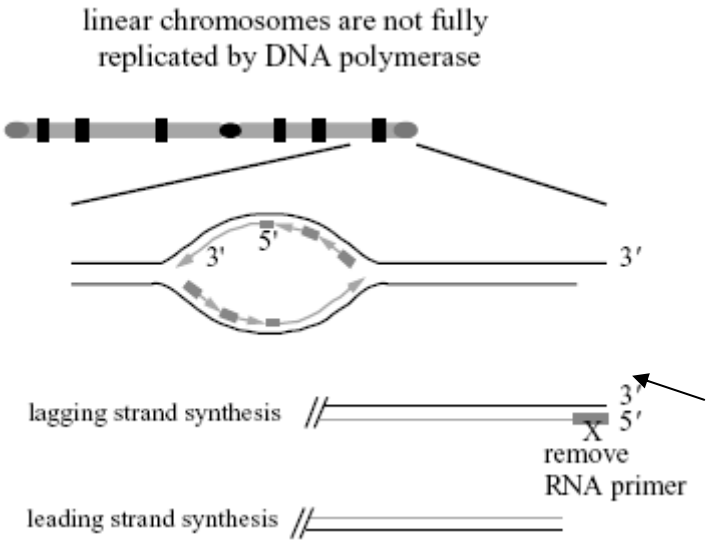


Figure 2.4 – Inability of DNA polymerase to replicate the very end of chromosome (shown by arrow) {reproduced from Chan et al., (Chan, SR et al. 2004)}.

2.6.1 Telomere shortening with age

The first direct evidence for telomere loss during cellular ageing was obtained by analysis of cultured human fibroblasts (Harely *et al.* 1990). *In vivo* loss of telomere length with age was first observed in skin cells by Lindsay *et al.,* (Lindsey *et al.* 1991) in 1991. Thereafter, there were many observations of human telomeres shortening with age in tissues such as fibroblasts, skin, blood, epithelium, liver, thyroid, parathyroid, brain, kidney, vascular tissue, myocardium, spleen, and pancreas (Allsopp, R *et al.* 1995; Allsopp, RC *et al.* 1992; Benetos *et al.* 2001; Butler, MG *et al.* 1998; Chang *et al.* 1995; Counter *et al.* 1992; Harely *et al.* 1990; Ishii *et al.* 2006; Kammori *et al.* 2002; Lindsey *et al.* 1991; Melk *et al.* 2000; Nakamura *et al.* 2002; Rufer *et al.* 1999; Takubo *et al.* 2000; Tsuji *et al.* 2002; Vaziri *et al.* 1994; Vaziri *et al.* 1993) .

Short telomere length has also been demonstrated to increase the risk of mortality (Cawthon *et al.* 2003). One study of 143 individuals over 60 years of age concluded that individuals with shorter telomere length are three times more likely to die of heart disease and eight times more likely to die of infectious disease than those with longer telomere lengths (Cawthon *et al.* 2003). Shorter telomere lengths have also been observed in individuals with premature ageing diseases such as Hutchinson-Gilford progeria (Harely *et al.* 1990) and Down's syndrome (Vaziri *et al.* 1993). In addition, the first cloned mammal from adult somatic cells, Dolly the sheep, was found to have shorter telomere lengths, which was thought to be the cause of her premature ageing (Shiels *et al.* 1999).

In germ line cells, stem cells and cancer cells, telomere length is kept constant by expression of the enzyme telomerase. Human telomerase is a ribonucleoprotein complex composed of two proteins {human telomerase reverse transcriptase (hTERT) and dyskerin) and the human telomerase RNA component (hTR) (Cohen *et al.* 2007). The hTR RNA contains the sequence CCCUAA, the reverse complement of the TTAGGG found in the telomeres (Lewis 1998). The reverse transcriptase (hTERT) uses hTR as a template and reverse-transcribes telomere DNA (Kufe *et al.* 2003). In some cancer cells, in the absence of telomerase, telomere length is elongated by another mechanism called the alternative lengthening of telomeres (ALT) (Kufe *et al.* 2003; Neumann *et al.* 2002) mechanism. A recombination based model was suggested for ALT, in which telomeres on another strand act as a template for telomeric DNA synthesis (Neumann *et al.* 2002).

2.7 Telomere length inheritance

Telomere length and chromosome-specific telomere length patterns have been observed to be inherited (Graakjaer *et al.* 2004; Njajou *et al.* 2007; Rufer *et al.* 1999; Slagboom *et al.* 1994). The reported degree of inheritance values vary depending on the sample population analysed. A strong value of inheritance, up to 78%, was observed in a sample composed of Dutch monozygotic and dizygotic twins (Slagboom *et al.* 1994). Telomere length had a lesser inheritance value of 44% among the sample of a Caucasian founder population of the old order Amish (Njajou *et al.* 2007). Such a wide range of estimates may be a result of differences in study populations, age ranges, environmental conditions, or pedigree structures (Njajou *et al.* 2007).

To evaluate the mode of inheritance of telomere length, the correlation between telomere lengths of father-son, father-daughter, mother-son and mother-daughter combinations have been compared with each other (Table 2.1). There are two hypotheses on the inheritance of telomere length. According to Nowrot *et al.*, telomere length inheritance is linked to the X chromosome (Nawrot *et al.* 2004), whereas, according to Nordfjall *et al.* and Njajou *et al.*, telomere length is paternally inherited (Njajou *et al.* 2007; Nordfjäll *et al.* 2005).

Observation of high correlation between telomere lengths of mothers and children and of fathers and daughters, but notably not with fathers and sons, led Nowrot *et al.* to construct the hypothesis that X-linkage is the mechanism of inheritance for telomere length (Nawrot *et al.* 2004). It was suggested that the DKC1 gene present on the X chromosome plays a role in X linked inheritance (Nawrot *et al.* 2004). The

protein dyskerin is coded by DKCI, which aids in the stable accumulation of the hTERT component of telomerase. A missense mutation in the DKC1 gene causes a premature ageing syndrome called dyskeratosis congenita (Reviewed in (Collins *et al.* 2002)). Polymorphisms present in the DKC1 gene may regulate the telomere length of individuals by regulating the expression of dyskerin and hence telomerase. This theory of X-linked inheritance of telomere lengths is challenged by the observations made by Nordfjall *et al.* and by Njajou *et al.* (Njajou *et al.* 2007; Nordfjäll *et al.* 2005). Both observed high correlation of telomere lengths between fathers and children but not between mothers and children (Table 2.1), and they concluded that telomere length is paternally inherited (Njajou *et al.* 2007; Nordfjäll *et al.* 2005).

Table 2.1 - Comparisons of telomere lengths between parents and children and the hypothesis on mode of inheritance of telomere lengths.

Family pair	Number of pairs	Ethnic background of individuals in sample	Proposed mode of inheritance of telomere length	Reference
Father-Daughter Father-Son Mother-Daughter Mother-Son	47 34 71 51	Northern Belgium	X chromosome- linked telomere length inheritance	Nowrot *et al* (Nawrot *et al.* 2004)
Father-Daughter Father-Son Mother-Daughter Mother-Son	22 20 23 18	Northern Sweden	Paternal inheritance of telomere length	Nordfjall *et al* (Nordfjäll *et al.* 2005)
Father-Daughter Father-Son Mother-Daughter Mother-Son	102 62 105 63	Old order Amish of Pennsylvania, United States of America	Paternal inheritance of telomere length	Njajou *et al* (Njajou *et al.* 2007)

The influence of parental age at conception, and its role in telomere length has also been investigated. Studies carried out by Unryn *et al.* (Unryn *et al.* 2005), Njajou *et al.* (Njajou *et al.* 2007), and Meyer *et al.* (Meyer *et al.* 2007) indicated that the father's age at conception/ birth is an important factor in determining an individual's telomere length. They showed that for each year that the father is older at conception/birth of the child, the child has on average an increase in telomere lengths by 22 bp in Unryn *et al.,* (Unryn *et al.* 2005), 10 bp in Njajou *et al.,* (Njajou *et al.* 2007) and 17 bp in Meyer *et al.,* (Meyer *et al.* 2007). The difference in the amount of telomere length increase in these three studies could be due to the different methods

used. Both Unryn *et al.* (Unryn *et al.* 2005) and Meyer *et al.* (Meyer *et al.* 2007) used the TRF method, while Njajou *et al.* (Njajou *et al.* 2007) used real time PCR. It has been shown that telomere length is increased in sperm cells with the age of the individual and this could be due to the presence of enzyme telomerase in adult testes (Allsopp *et al.* 1992). As for the correlation between the mothers' age at conception and the children's telomere lengths, only a borderline correlation could be demonstrated Njajou *et al.,* (Njajou *et al.* 2007). To date, there has been no study reported on the presence of telomerase in the female germ line (Unryn *et al.* 2005). It is possible that both parents' ages at conception influence children's telomere lengths by varying amounts.

2.8 Telomere length measurement

There are various methods that have been used to measure telomere lengths in humans (discussed in section 2.8.1). In this study, telomere lengths were measured using quantitative real time PCR (Cawthon 2002). The main aim was to investigate the correlation between telomere length and age, and to see if this relationship could be used to predict the age of an individual from their DNA.

2.8.1 Methods in telomere length measurement

2.8.1.1 Telomere restriction fragment (TRF) analysis of telomere lengths

Various methods have been used to measure telomere length in humans. The standard method is mean terminal restriction fragment (TRF) length analysis (Chang *et al.* 1995; Harely *et al.* 1990; Takasaki *et al.* 2003; Tsuji *et al.* 2002). Genomic DNA is digested by frequently cutting restriction enzymes that do not cut telomere segments. The resulting terminal restriction fragments are then separated by gel

electrophoresis and blotted to a membrane; the TRF s are visualised by hybridization with a labelled probe, either radioactive or non-radioactive (Harely *et al.* 1990).

One of the major drawbacks of the TRF method is that it requires a large amount of DNA (0.5-5 µg/individual), and is also time consuming (3-5 days) (Cawthon 2002). Often the amount of DNA available for forensic analysis is very low. The TRF method is also considered to have a low overall sensitivity, and is biased against the detection of short telomeres (Baird *et al.* 2003). TRF length analysis sometimes includes subtelomeric regions (Baird *et al.* 2005; Nordfjäll *et al.* 2005) and this makes it questionable whether the results accurately represent the real telomere lengths. In addition, TRF analysis produces auto-radiographic smears, which makes analysis subjective (Nakagawa *et al.* 2004). Differences in TRF-determined length depend on the restriction enzyme used, and the resultant estimated telomere lengths can differ by up to 5% (Baird *et al.* 2005; Cawthon 2002). This could be due to the existence of polymorphisms in restriction sites near the telomeric and subtelomeric regions on different chromosomes (Cawthon 2002; Nakagawa *et al.* 2004).

2.8.1.2 Quantitative real time PCR analysis of telomere length

A method based on PCR would solve the problems associated with low amounts of DNA and also reduce the time required for analysis. Cawthon (Cawthon 2002) developed a method to measure telomere lengths using quantitative real time PCR. This method requires relatively small quantities of DNA (35ng) and the analysis takes only a few hours. Primers were designed by Cawthon (Cawthon 2002) that hybridize specifically to human telomere regions to avoid the forming of primer

dimmers. Only telomeric (TTAGGG) repeats are amplified; subtelomeric regions are not included (Nordfjäll *et al.* 2005).

In quantitative real time PCR, the amount of DNA in a sample is measured in comparison to the amount of a single copy gene that is assumed to be constant. The single copy gene used in our study was 36B4, which encodes acidic ribosomal phosphoprotein PO located on chromosome 12 (Cawthon 2002).

In this relative quantification approach, concentrations of the telomere hexameric repeat and the 36B4 gene are calculated for the same sample. An external standard curve is used to determine the concentration of these two parameters. Then the absolute concentration of the telomere hexameric repeat (T) is divided by the absolute concentration of the 36B4 gene (S). The resulting value (T/S ratio) is divided by the T/S ratio obtained for the calibrator DNA. The resulting ratio expresses the amount of telomere hexameric repeats, called the relative telomere length (Cawthon 2002). Since the telomere length is quantified as a ratio, the absolute length of the telomeres in base pairs is not obtained. The relative telomere length is believed to reflect the actual differences in telomere length among individuals (Nakagawa *et al.* 2004).

2.8.1.3 Other factors that may influence telomere lengths

Various environmental and life-style factors have been reported to influence telomere lengths, some through the amount of free-radicals present and antioxidants available in the body. These factors include diet, amount of exercise, amount of alcohol consumed regularly, whether a person is a regular smoker, etc. If a person is

afflicted by cancer, this may also affect telomere lengths in certain cells of their body, as telomere lengthening mechanisms are often active in cancer cells, as described in Section 2.6.1.

Initially some attempt was made to obtain self-reported lifestyle of volunteers by asking them to fill out a questionnaire (with informed, written consent). Some of the factors in the questionnaire were the amounts of certain food types consumed, e.g. red meat, fish, poultry, green vegetables, citrus fruit, amount of regular exercise carried out and the amount of alcohol consumed and cigarettes smoked regularly. However, feedback indicated that many of the volunteers were not able to accurately estimate these factors in their lives, especially as they had not been previously asked to specifically observe these. Thus, this self-reported data collected was considered unreliable and hence it was not used in this thesis. Such studies would need to be conducted under controlled conditions, with placebos included where possible, and with the level of factors administered or consumed measured by objective methods. Such a study was beyond the scope of this research but would be interesting to consider in the future.

2.9 Mitochondria

Altman in 1894 first discovered the existence of mitochondria, but Banda provided the name in 1898: mitochondrion from Greek for thread (mitos) and grain (chondros). Mitochondria are present in virtually all-eukaryotic cells, with the exception of erythrocytes (Rasmussen, LJ *et al.* 1998). Mitochondria are semi-autonomous organelles that contain their own genome and own protein synthesis mechanism (reviewed in (Meissner, C 2007)). Mitochondria are usually depicted as stiff, elongated cylinders with a diameter of 0.5 to 1μm, resembling bacteria (Figure

2.5). Time-lapse microcinematography of living cells, however, shows that mitochondria are able to change their shape through fission and fusion, resulting in continuous remodelling of the mitochondrial network (Voccolia *et al.* 2009). Mitochondria are among the larger organelles in the cell, each one being about the size of an *E. coli.* (Alberts *et al.* 1994). Each mitochondrion has a double membrane structure (Figure 2.5). The outer mitochondrial membrane surrounds the inner membrane (Moraes *et al.* 2002). The outer membrane defines the smooth outer perimeter of the mitochondrion. In contrast, the inner membrane has numerous invaginations called cristae. These membranes define two sub-mitochondrial compartments (Alberts *et al.* 1994). These are: the space between the two membranes, i.e. intermembrane space (Moraes *et al.* 2002), and the space formed between the inner membrane with its cristae, the matrix, or central compartment. Mitochondria are the organelles of the cells essential for respiration and oxidative phosphorylation. They produce more than 80% of the energy (ATP) needs of a cell (Alberts *et al.* 1994). The mitochondrial inner membrane, cristae, and matrix are the sites of most reactions involving the oxidation of pyruvate and fatty acids to CO_2 and H_2O and the coupled synthesis of ATP from ADP and Pi.(Alberts *et al.* 1994). The matrix contains a large variety of enzymes belonging to citric acid cycle. Mitochondria, in addition to producing energy, also perform many important cellular functions, all of which are essential for cell survival. Mitochondria carry out respiration, synthesize heme, lipids, amino acids, nucleotides, maintain intracellular homeostasis of inorganic ions and initiate programmed cell death (Kang *et al.* 1998).

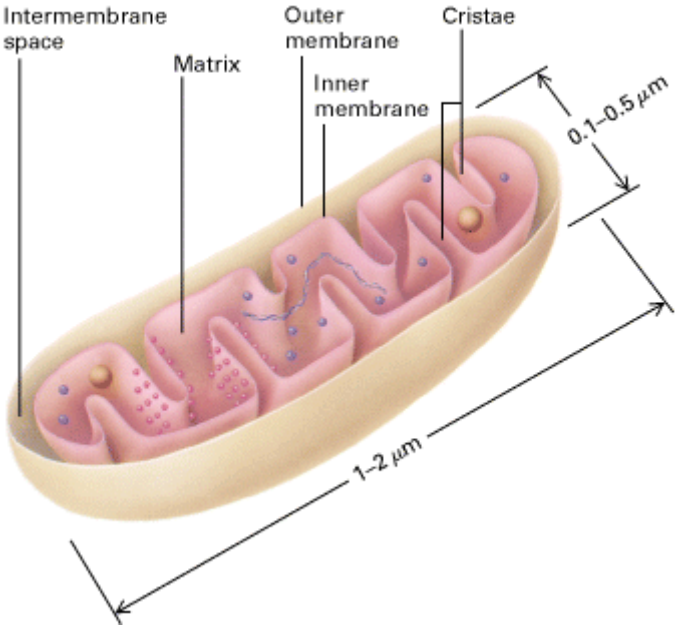


Figure 2.5 – A schematic representation of a mitochondrion. {Reproduced from Lodish et al., (Lodish et al. 2000) }.

2.10 The Mitochondrial Genome

Mitochondria have a separate autonomously replicating DNA genome, and it is maternally inherited by offspring (Giles, RE *et al.* 1980). Barring mutations, the mother passes her mitochondria to her children. Therefore one's mitochondrial DNA type is not unique to them; it is similar between siblings and maternal relatives (Butler, JM *et al.* 1998). The mitochondrial genome was first detected by Nass and Nass in 1963 (Moraes *et al.* 2002). The human mitochondria contain multiple copies of a 16,569bp closed circular DNA genome that is replicated and expressed within the organelle system (Larsson *et al.* 1995). Human mitochondrial DNA (mtDNA) is described as a separate genome because it has existed in physical and genetic isolation from the nuclear genome for 1.5 to 2 billion years of eukaryotic evolution. Nuclear and mitochondrial genomes are, however, highly interdependent, in the sense that mtDNA depends for almost all of its genetic functions on genes resident in

the nucleus (Jacobs *et al.* 1998). The nuclear coded factors that are synthesized in the cytosolic ribosomes as precursor polypeptides are imported in to the mitochondria via specialized import pores (Moraes *et al.* 2002). On the other hand, the mitochondrial genome encodes some functions essential for cellular survival. Hence the relationship resembles that of a host and parasite, or more correctly a symbiotic (Jacobs *et al.* 1998).

Although mtDNA represents less than 1% of total cellular DNA, the copy number of mitochondrial genomes is high, ranging from 10^3 to 10^4 genomes per cell (Larsson *et al.* 1995). The two DNA strands have significantly different base compositions: the heavy (H) strand is rich in guanines; the light (L) strand is rich in cytosines. Although the mitochondrial DNA is principally double-stranded, a small section is defined by a triple DNA strand structure. The human mitochondrial genome contains 37 genes: 13 protein encoding genes, 22 tRNA genes and 2 rRNA genes (Figure 2.6) (Lodish *et al.* 2000). MtDNA encoded proteins are important subunits of the respiratory chain and ATP synthase. Specifically, seven subunits of complex I (NADH dehydrogenase), three subunits of complex IV (cytochrome oxidase), two subunits of complex V (ATP synthase), and one subunit of complex III (ubiquinol: cytochrome c oxidoreductase) (Nagley *et al.* 1998).

Anderson *et al.,* published the sequence and organization of the human mitochondrial genome in 1981 (Anderson *et al.* 1981). Andrews *et al* in 1999 re-sequenced the original mitochondrial DNA sample and corrected for errors and rare polymorphisms (Andrews *et al.* 1999). Today the original sequence (GenBank accession: M63933) is the reference sequence to which new sequences are compared and it is commonly

known as the Anderson sequence or the Cambridge reference sequence (Butler 2005). Unlike its nuclear counterpart, the human mitochondrial genome is extremely compact: approximately 93% of the DNA sequence represents the coding sequence (Anderson *et al.* 1981), whereas only ~ 2% of the human nuclear genome represents coding sequence (Dudek 2006). Most of the open reading frames (not all) for the peptides and ribosomal RNA (rRNA) genes in the mitochondrial genome are separated by one or more transfer RNA (tRNA) genes - with few, if any extra nucleotides in between (Scheffler 2007).

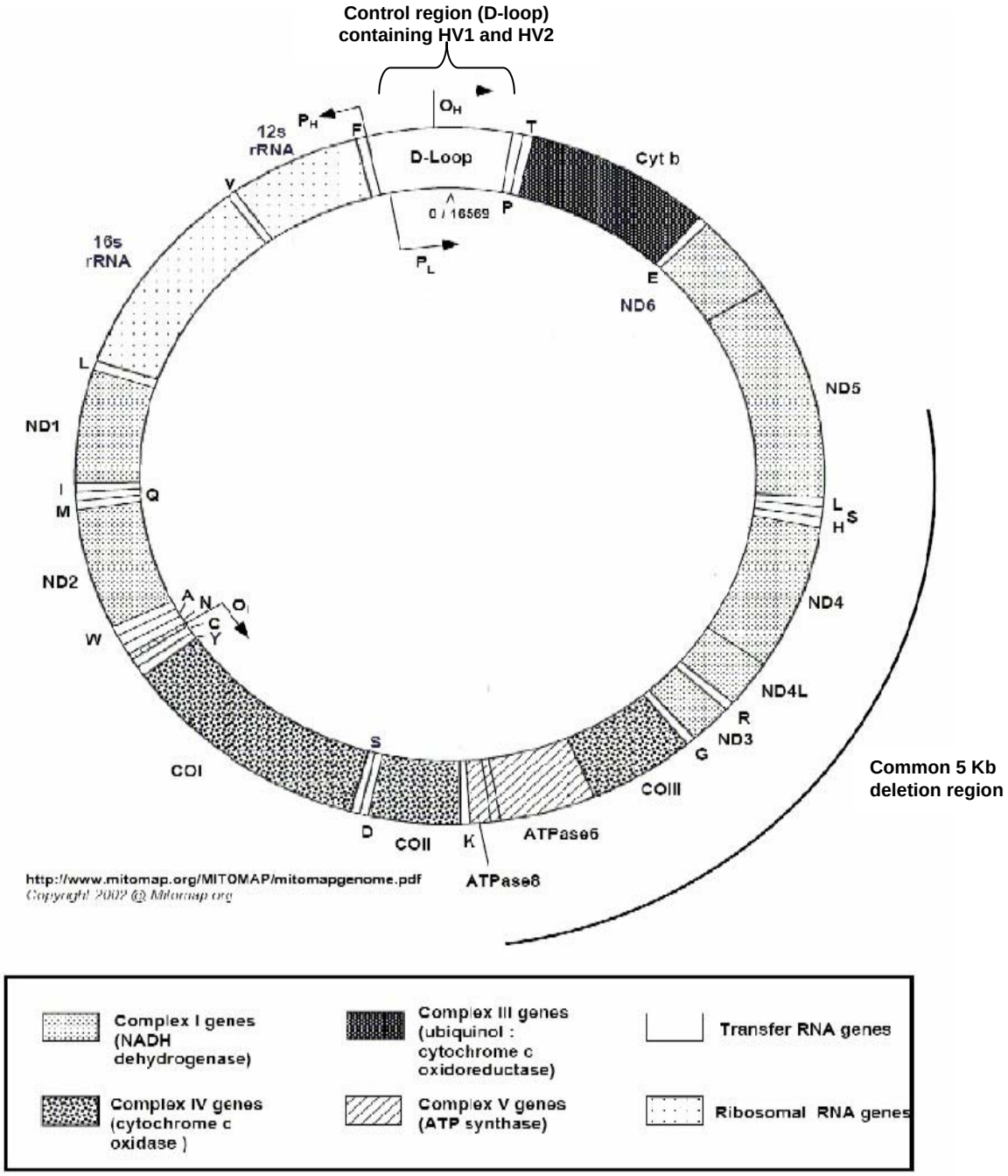


Figure 2.6 – The map of human mitochondrial genome showing the various enzyme-coding genes as well as genes coding for tRNA and rRNA. {reproduced from http://www.mitomap.org/ (MITOMAP 2007)}.

2.11 Oxidative damage to mitochondrial DNA

Mitochondria are the power house of the cell by producing energy by oxidative phosphorylation. In normal physiological conditions around 2% of oxygen consumed in mitochondria are converted in to free radicals/reactive oxygen species (ROS) (reviewed in (Kang *et al.* 1998)). In 1956, Harman first proposed that free radicals cause damage to DNA, lipids and proteins, which in turn causes ageing (reviewed in (Ozawa 1997). In 1972, Harman modified his theory and proposed that oxidative damage to mitochondria is the determining factor behind ageing (Souza-Pinto *et al.* 2002). It was in 1989 that, the 'mitochondrial theory on ageing' was put forward by Linnane *et al* (Linnane, AW *et al.* 1989): that somatic mutations occurring as a result of free radical damage in somatic cell mitochondria and their subsequent cytoplasmic segregation during life is the major contributor to loss of bioenergetic capacity within tissues and thus ageing. Mitochondrial DNA is not associated with nucleosomes and is more accessible to damage by free radicals. The close proximity of mitochondrial DNA to the electron transport chain and the lack of DNA repair enzymes makes it more vulnerable to mutations and thus, the mitochondrial theory on ageing more plausible (reviewed in (Souza-Pinto *et al.* 2002)).

2.12 Recent developments on the mitochondrial theory on ageing

The mitochondrial genome is continuously replicated independent from cell cycle, unlike nuclear DNA which replicates only during cell division. Mitochondrial DNA replication is carried out by several nuclear encoded proteins. From those proteins only one DNA polymerase known as polymerase γ (POLYG) is presently known to be responsible for mitochondrial DNA replication and repair (Hudson *et al.* 2006). Recent studies on mitochondrial mutations were carried out with PolyG mice. PolyG

mice have mutations in the proofreading exonuclease activity of the DNA polymerase γ (POLYG). The $Polyg^{muty/mut}$ (homozygous) mice accumulated mitochondrial mutations with age and aged prematurely. The $Polyg^{mut/+}$ (heterozygous) mice had much higher rates of mutation than the wild-type older mice, but still sustained their health and lifespan (discussed in (Khrapko, Konstantin *et al.* 2007)). This discovery questioned the very basis of the 'mitochondrial theory on ageing' in mice, and it was suggested that whilst mitochondrial mutations /deletions do accumulate with age, they may not contribute the ageing process (Khrapko, Konstantin *et al.* 2007). These data on mice may not be directly applicable to human ageing because of two main reasons. Firstly, some types of mitochondrial DNA mutations such as large deletions have been found to be important in human ageing and age-related diseases, but not so in mice. For example, patients with inherited defects in mitochondrial DNA polymerase γ /POLYG (which leads to increased levels of mtDNA deletions) suffer from psychiatric diseases, Parkinsonian syndrome and male infertility (Hudson *et al.* 2006). This has not been reported in mice. The second reason is that lower levels of mitochondrial DNA mutations are associated with defects in oxidative phosphorylation in humans but not in mice. Therefore, whilst mtDNA mutations may not be essential to ageing in mice, one cannot rule out their effects on the human ageing process.

2.13 Mitochondrial heteroplasmy

A large number of mtDNA mutations have been reported to occur in human somatic tissues which are suggested to accumulate during the ageing process. These mutations are classified as large deletions, point mutations, and small duplications (Nagley *et al.* 1998). If all mtDNA molecules in a cell are identical, this condition is

called homoplasmy. However, when mutations in mtDNA are found, cells contain mixtures of wild type and mutant type mtDNA: a condition known as heteroplasmy (Lodish *et al.* 2000). Three types of heteroplasmy could occur in a tissue: intercellular heteroplasmy, intracellular and intermitochondrial heteroplasmy and intracellular and intramitochondrial heteroplasmy. In intercellular heteroplasmy, wild-type and mutant mtDNA occur in different cells. Intracellular intermitochondrial heteroplasmy is where wild-type and mutant mtDNA occur in different mitochondria in the same cell. The presence of wild and mutant mtDNA in the same mitochondria is called intracellular and intermitochondrial heteroplasmy (Attardi 2002).

2.14 Mitochondrial point mutations and heteroplasmy that accumulate with age

Accumulation of mitochondrial point mutations as well as heteroplasmy has been observed in many tissues. The highest number of point mutations and heteroplasmy has been observed in the mitochondrial control region (Table 2.2). The mitochondrial control region contains initiation sites for transcription and replication of mtDNA (reviewed in (Trifunovic 2006)). This region is the most variable segment of mammalian mtDNA and thus has a higher possibility of harbouring mutations during ageing (reviewed in (Trifunovic 2006)). Hypervariable region 1 (HV1) extends from position ~16024 to ~16365, while hypervariable region 2 (HV2) extends from ~73 to ~340 (the boundaries are not rigidly defined, and vary among particular studies or laboratories) (Holland *et al.* 1999). Liu *et al.*, (Liu *et al.* 1997) found an A→G base substitution at position 3243 (which lies outside the hypervariable region) accumulates with age.

Table 2.2 - Mitochondrial control region mutations that accumulate with age.

Nucleotide position	Tissue	Age range	Method	Presence of mitochondrial mutations	Reference
T414G	Fibroblast (n=18)	20 weeks- 101 years	Denaturant gradient gel electrophoresis (DGGE)	T414G present in 57% of the individuals above 65 years old and was absent in 13 younger individuals.	Michikawa *et al.*,(Michikawa *et al.* 1999)
A73G, T146C, C150T, T152C, A189G, T195C, C198T, A200G, G247A	Heart, brain, muscle and blood (n=43)	11 – 85 years	Immobilized sequence-specific oligonucleotide (SSO) probe system	Heteroplasmy increases with age.	Calloway *et al.,* (Calloway *et al.* 2000)
T414G	Muscle (n=23)	19 - 80 years	Peptide nucleic acid (PNA) directed PCR clamping technique	T414G present in individuals older than 30 years old. Absence of T414G in brain.	Murdock *et al.,* (Murdock *et al.* 2000)
A189G, T408A	Skeletal muscle (n= 40)	20-92 years	Denaturant gradient gel electrophoresis (DGGE)	A189G, T408A present in most of the individuals aged 53 to 92 years old, absent or marginally present in 19 individuals younger than 34 years old. Absence of T414G in skeletal muscle.	Wang *et al.,* (Wang et al. 2001)

G73A, A189G, A214G, G251C, T252C,303-310C insertions, T16028C, T16029C, G16033A, G16034A, G16035A, G16036A, G16049A, 16052-3Cdel, A16054G	Cardiac muscle (n=6) and buccal cells (n=7)	3 months – 109 years old and 2-93 years	Single cell sequence analysis	High proportions of cells in tissues of buccal epithelium and heart muscle contain high proportions of clonal mutant heteroplasmy expanded from single initial mutant mtDNA molecules.	Nekheava *et al.,* (Nekhaeva et al. 2002)
A189G, T408A, T414G	Bicep muscle (n=4)	81, 94, 96, 97 years	DNA sequencing	Presence of age related mutations in all 4 donors.	Del-Bo *et al.,* (Del-Bo *et al.* 2003)
A189G	Buccal cells (n=37) and muscle (n=69)	4-85 years in individuals belonging to 10 maternally related families and 1-97 years	Automated DNA sequencing, Southern blot hybridization, Peptide nucleic acid (PNA) and quantitative real time PCR	Presence of A189G heteroplasmy in older individuals not in younger individuals in same maternal lineage. Presence of A189G heteroplasmy in very high levels in individuals aged 60 years old and older.	Theves *et al.,* (Theves *et al.* 2006)

2.15 Mitochondrial deletions accumulate with age

The most frequently observed age associated mitochondrial deletion is the 5 kb which is often referred to as the 'common deletion'(Cortopassi *et al.* 1990). Cortopassi *et al*., (Cortopassi *et al.* 1990) was the first to report that the 5 kb deletion accumulates with age in brain and heart tissue. The 5 kb deletion occurs between position 8470 bp and 13447 bp in the mitochondrial genome. This results in the loss of subunits of electron transport system complexes I & IV as well as many tRNA (Pak *et al.* 2003). A critical part of the enzyme cytochrome oxidase (subunit III) is included in the common deletion. Cytochrome oxidase is the major site of reduction of molecular oxygen in normal cells. The presence of a deficient cytochrome oxidase complex could lead to increased levels of cellular oxidative damage (Cortopassi *et al.* 1990). The presence of 5 kb deletion with age has since been reported in a wide range of tissues (Table – 2.3).

Other types of deletions observed to accumulate with age in mitochondria are a 8.04 kb deletion (between position of 8030 bp and 16701 bp) in skeletal, cardiac muscles and brain tissue (Baumer *et al.* 1994), a 7.4 kb deletion (between adenosine triphosphate 6 gene and D-loop region) in cardiac muscle (Hattori *et al.* 1991), a 3.6 kb deletion (between position of 1837 bp – 5447 bp, from the 16S rRNA gene to ND2 gene) in the skeletal muscle (Katayama *et al.* 1991), and a 6 kb deletion (between position of 7842 bp-13905 bp) in the liver (Yen *et al.* 1992).

Studies carried out by Melov *et al.,* Kopisdas *et al.*, Kovalenko *et al.*, Khrapko *et al.*, and Meissner *et al.* (Khrapko, K *et al.* 1999; Kopsidas *et al.* 1998; Kovalenko *et al.* 1998; Meissner, C *et al.* 2006; Melov *et al.* 1995) were focused on amplifying the whole mitochondrial genome and identifying large mitochondrial deletions that

accumulate with age. All the studies observed various types of deletions in mitochondria that accumulate with age at the cellular and tissue level. Studies done by Kopisdas *et al.,* (Kopsidas *et al.* 1998) Kovalenko *et al.,* (Kovalenko *et al.* 1998) and Meissner *et al.* (Meissner, C *et al.* 2006) did not classify the deletions they observed. Khrapko *et al.* (Khrapko, K *et al.* 1999) classified the deletions, which varied in size from 3.7 kb to 12 kb. Interestingly in the Khrapko *et al.* (Khrapko, K *et al.* 1999) study, they did not detect the 'common deletion'.

Table 2.3 – Mitochondrial "common deletion" presence in different tissue types that accumulate with age.

Tissue	Age range	Method	Summary of findings	Reference
Heart, liver, kidney, lung, skeletal muscle, spleen, mid brain, etc.	80 minutes - 87 years old	PCR	Present in all the individuals aged 40 - 87 years old.	Linnane *et al.,* (Linnane, A *et al.* 1990)
Brain, heart , spleen (n = 12)	22 weeks - 58 years old	PCR	Present in all individuals aged over 30 years old	Cortopassi *et al.,* (Cortopassi *et al.* 1990)
Liver (n = 55)	2 still born babies, and 27 – 86 years old.	PCR and restriction mapping	Present in all the individuals aged over 50 and 70% of the individuals aged 30-50 years old.	Yen *et a*l.*,* (Yen *et al.* 1991)
Lung (n = 127)	34 weeks-79 years old	PCR	Present in individuals after the age of 30.	Fahn *et al.,* (Fahn *et al.* 1996)
Cortex, putamen, cerebellum (n = 7)	24 - 94 years old	PCR	Present in all the individuals aged 67-94 years old.	Corral-Debrinski *et al.,* (Corral-Debrinski *et al.* 1992)
Ovary (n = 34)	22 weeks-77 years old	PCR	Present in individuals aged over 30 - 77 years old.	Suganuma *et al.,* (Suganuma *et al.* 1993)
Skeletal, heart, kidney (n =37)	1h – 90 years old	PCR	Present in individuals aged 37 – 90 years old.	Liu *et al.,* (Liu *et al.* 1998)
Skeletal muscle (n = 93)	3 months -97 years old	PCR	Present in most of the individuals aged over 20 years old.	Meißner *et al*., (Meißner *et al.* 1997)
Ovary (n = 39)	Fetus – 77 years old	PCR	Present in individuals over 40 -77 years old.	Kitagawa *et al.,* (Kitagawa *et al.* 1993)
Heart (n = 33)	3 month - 93 years old	Real time PCR and DHPLC	Present in individuals above 40 years old.	Mohamed *et al.,* (Mohamed *et al.* 2006)
Unfertilised oocytes (n = 52)	25-42 years old	Real time PCR	Present in all the individuals older than 35 years.	Chan *et al.,* (Chan, CCW *et al.* 2005)

2.16 Purpose and plan of this study in brief

The aim of this study was to investigate the possibility of developing a reliable test to predict age in humans by analysing telomere lengths and mitochondrial mutations/deletions. This would help the forensic community as an investigative tool to predict the age of a criminal using DNA from tissues inadvertently left behind at the scene of a crime. Note that the methods developed are not intended to be used as evidence in courts, rather as a preliminary investigative tool for police to assist in narrowing down the range of suspects. More specifically, the aims were:

- To find or develop a method of reliably and accurately measuring average telomere lengths in humans
- To use this method to measure telomere lengths in a random sample of individuals from various age groups
- To determine if there was sufficient correlation between average telomere lengths and age to be able to predict the age or age group of an individual from an unknown DNA sample
- To determine whether gender, ethnicity or inheritance influenced telomere lengths in individuals
- To examine mitochondrial mutations from a sample consisting of three or four maternal generations to see if mutations did develop with age as some theories in the past have asserted

The next chapter presents the materials and methods used in this study.

Chapter 3

Materials and methods

3.1 Sample Collection

This project was conducted with approval from the Human Research Ethics Committee of Victoria University (HRETH .020/03). Four buccal swabs (COPAN ITALIA S.P.A. Brescia, Italy) were collected from each individual. A person was asked to rub the cotton side of the swab inside the mouth. It was made sure that each individual had not had a main meal for at least two hours prior to buccal swab collection and if they had not, they were asked to rinse the mouth thoroughly with water. The collected buccal swabs were air-dried and stored at –20^{o}C prior to DNA extraction. Upon sample collection, a questionnaire was given to each subject requesting their details: age, gender, ethnicity, longevity of deceased family members, diet and exercise habits (Appendix I and II). All procedures, including filling out of the questionnaire were carried out after informed, written consent from the volunteers.

3.1.1 Sample collection for telomere length analysis

Samples from a total of 167 individuals aged 1 to 96 years old were collected for telomere length analysis. The sample was composed of people from various ethnic backgrounds (Table 3.1). There were 89 individuals, who had at least one relative of his/her family within the sample. The rest of the 78 individuals had no relatives within the sample. The sample was divided into nine age groups and consisted of 71 males and 96 females (Table 3.2). From the sample of 89 related individuals, telomere lengths of father-daughter, father-son, mother-daughter and mother-son

combinations were chosen to study telomere length inheritance in humans (Table – 3.3).

Table 3.1 – Ethnic background of the sample population.

Ethnicity	Number of individuals	Percentage of total sample %
African Negroid	3	1.8
European	117	70.1
Middle Eastern	1	0.6
Native American	1	0.6
Oriental Asian	7	4.2
South Asian	38	22.8
Total	167	100

Table 3.2 – Age and gender distribution of the population.

Age group (Years)	Total number of individuals	Male	Female
0-9	9	3	6
10-19	20	11	9
20-29	39	15	24
30-39	33	15	18
40-49	18	6	12
50-59	26	11	15
60-69	9	6	3
70-79	10	3	7
80-89	2	1	1
90<	1	0	1
Total	167	71	96

Table 3.3 - Parent and children combinations.

Family pair	Number of pairs
Father-Daughter	10
Father-Son	11
Mother-Daughter	23
Mother-Son	16

3.1.2 Sample collection for mitochondrial DNA analysis

Samples from a total of 12 three/ four maternal generation families were collected for the mitochondrial DNA analysis (Table 3.4). Each family included the maternal grandmother, mother and the youngest member could be either a daughter or a son.

Table 3.4 – Maternal lineage families.

Family	Age (Years)			
	Maternal great grandmother	Maternal grandmother	Mother	Daughter /Son
1	-	87	44	13
2	-	67	48	09
3	-	80	59	31
4	-	68	45	25
5	-	76	50	12
6	-	71	49	25
7	-	96	58	28
8	-	76	48	16
9	-	71	43	13
10	-	74	47	14
11	78	54	28	05
12	-	62	34	09

3.2 DNA Extraction

DNA was extracted from two buccal swabs from each individual using the GenomicPrep Cells & Tissue DNA Extraction Kit (GE Healthcare Australia & New Zealand) according to the supplier's protocol. This involved various steps. Firstly, the swabs were dissected and cells on them were lysed by adding chilled cell lysis solution. Proteins and RNA were removed by adding Proteinase K (20mg/ml, QIAGEN Pty. Ltd Australia) and RNaseA (GE Healthcare Australia & New Zealand) respectively. The DNA was then precipitated using 100% isopropanol (MERCK & Co. Inc. NJ, USA). Then the DNA pellet was washed with 70% ethanol (BDH). The DNA pellet was air-dried for 10-15 minutes, and resuspended overnight at room temperature in DNA hydration solution provided in the kit (GE Healthcare Australia & New Zealand). DNA was stored at -20^{o}C until further use.

3.3 Quantification of DNA

DNA concentration was determined using the Quantifiler Human DNA quantification kit (Applied Biosystms Australia, Vic, Australia) at the Victorian Police Forensic Science Services laboratories. This method quantifies the amount of amplifiable human DNA present in a sample. The Quantifiler Human DNA quantification kit contains two independent sets of primers and TaqMan® probes. One primer/ probe set amplifies and detects the presence of the Human Telomerase Reverse Transcriptase gene. (Note that the fact that the telomerase gene is the PCR target in this kit is purely coincidental and has no bearing on experiments in this research thesis conducted on telomere lengths). The other primer/probe set works as an Internal PCR Control (IPC) system. A sample that showed neither amplification of the human telomerase reverse transcriptase gene nor IPC detectors would indicate

the presence of PCR inhibitor in the sample, or a failure of instrument or reagent components; in either case the user is notified that human DNA may be present but not amplifiable (Green *et al.* 2005). The Quantifiler kit assays were performed according to the protocol specified in the user's manual. Commercially available human DNA, K562 High molecular weight DNA (Promega Corporation, Melbourne, Australia) was used to generate the standard curves. Dilutions of the standards were first made by serially diluting the 200 ng/µl stock solution of K562 High molecular weight DNA (Promega) to the following concentrations: 50.0, 16.7, 5.56, 1.85, 0.62, 0.21, 0.068 and 0.023 ng/µl in TE Buffer (that contained 20 µg /mL of molecular biology grade glycogen (Roche). A 'master mix' was prepared by adding 12.5 µl of Quantifiler PCR reaction mix and 10.5 µl of primer mix included in the kit (Applied Biosystems). Then 23µl each was placed in an ABI PRISM^{TM} 96-well Optical reaction plate (Applied Biosystems). 2 µl each of standards and sample DNA were loaded into the plate in duplicates. After all reactions were set up, the Optical Adhesive Cover (Applied Biosystems) was pressed into place with a plastic applicator tool, and covered with a foam compression pad. The plate was then placed in the 96-well sample block of an ABI PRISM 7900HT Sequence Detection System (SDS) instrument. The SDS software controlling the 7900 system was programmed for the following thermal cycling: 95^{o}C for 10 minutes and 40 cycles of [95^{o}C for 15 seconds, 60^{o}C for 60 seconds]. After thermal cycling was completed, data analysis was performed by the SDS software to generate a standard curve for quantification of unknown samples.

3.4 Telomere length analysis

In quantitative real time PCR, the amount of the gene of interest is measured compared to the amount of a single copy gene that is assumed to be constant (the reference gene). The single copy gene used in our study was 36B4, which encodes acidic ribosomal phosphoprotein PO located on chromosome 12 (Cawthon 2002). In this relative quantification approach, concentrations of the telomere hexameric repeats and concentrations of the 36B4 gene are calculated for the same sample and the relative telomere length value calculated is adjusted using values for this reference gene.

The method used was essentially that of Gil and Coetzer in 2004 (Gil *et al.* 2004), using the same primers developed by Cawthon (Cawthon 2002), for use with the LightCycler 2.0 (Roche Diagnostic Australia Pty. Ltd. NSW, Australia) with Some minor modifications were required for use of this method with our samples (see below). The tel 1 primer is designed to hybridise to any available partially complementary 31bp stretch along the strand of telomere DNA oriented 5/ to 3/ toward the centromere. The tel 2 primer is designed to hybridise to any partially complementary 33bp stretch along the strand oriented 5' to 3' toward the end of the chromosome (Figure 3.1). Primers were designed by Cawthon to hybridise specifically to telomere hexamer repeats (TTAGGG and CCCTAA) without producing primer dimers. When the two primers anneal to each other (Figure 3.1) the 3' terminal base of each primer cannot form a stable base pair with the base opposite it, thereby blocking the addition of bases by DNA polymerase (Cawthon 2002).

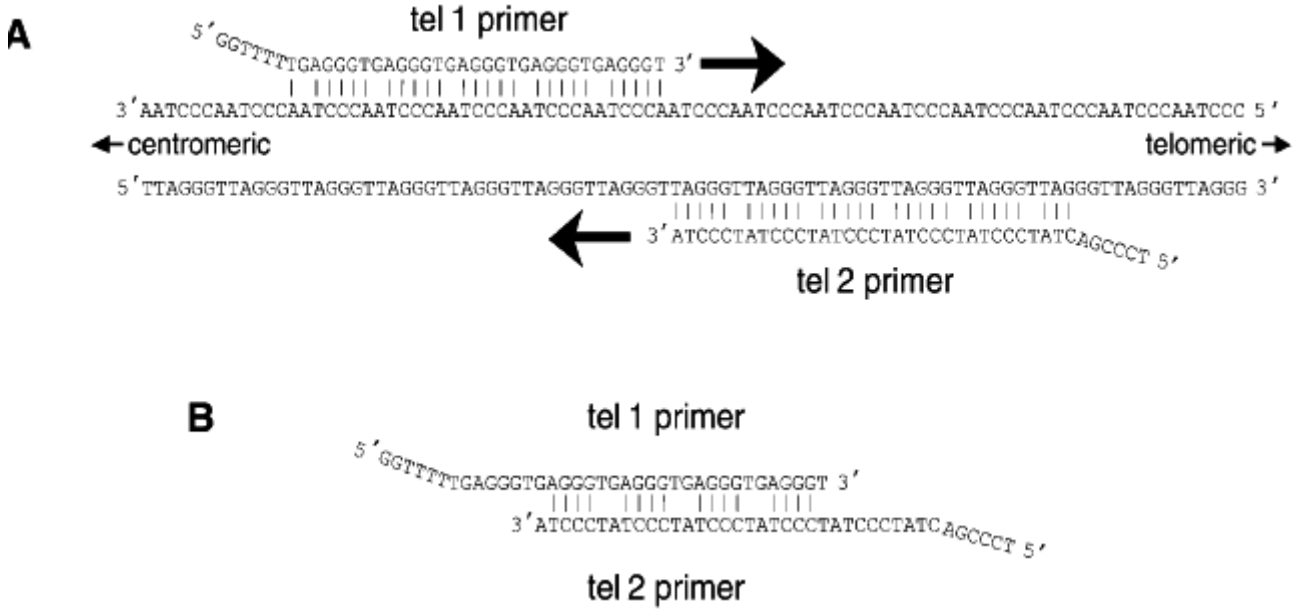


Figure 3.1 –A- Annealing of primers tel 1 and tel 2 to genomic DNA, **B –** Annealing of primers to each other (Cawthon 2002).

3.4.1 Real time quantitative PCR

Both telomere lengths and the 36B4 gene were amplified in the sample of 167 individuals and the family samples. Two master mixes of PCR reagents were prepared: one with the primers for telomere hexameric repeat, and the other with primers for the 36B4 gene. All the conditions were the same for the two master mixes except for the primers. The master-mix consisted of 3mM $MgCl_2$, 0.5μM of forward and reverse primers (Table 3.5), 2 μl of LightCycler Fast start DNA master SYBR Green (Roche Diagnostic Australia Pty. Ltd. NSW, Australia), 10 ng of DNA, and H_2O (PCR grade) up to 20 μl (Gil *et al.* 2004). For each run with the samples

K562, High molecular weight DNA (Promega Corporation, Melbourne, Australia) was amplified as a calibrator.

Table 3.5- The primers used in telomere length analysis.

	Primer pairs and sequences 5/ - 3/	Expected PCR product (bp)
Telomere hexameric repeats	**Tel 1** - 5/ CGG TTT GTT TGG GTT TGG GTT TGG GTT TGG GTT TGG GTT 3/ **Tel 2** - 5/ GGC TGG CCT TAC CCT TAC CCT TAC CCT TAC CCT TAC CCT 3/	~ 500bp Individual specific
Reference gene 36B4	**36B4u** - 5/ CAG CAA GTG GGA AGG TGT AAT CC 3/ **36Bd** - 5/ CCC ATT CTA TCA TCA ACG GGT ACA A 3/	74bp

3.4.1.1 PCR conditions for telomere hexameric repeats

The following conditions were used for telomere amplification: 95^{o}C for 10 minutes, followed by 50 cycles of 95^{o}C for 5 seconds, 56^{o}C for 10 seconds, and 72^{o}C for 60 seconds.

3.4.1.2 PCR conditions for 36B4 gene

The reference gene 36B4 was amplified for all samples. The amplification conditions used were: 95^{o}C for 10 minutes, followed by 50 cycles of 95^{o}C for 5seconds, 58^{o}C for 10 seconds, and 72^{o}C for 40 seconds.

3.4.2 Standard Curves

High molecular weight DNA, K562 (200 ng/µl) (Promega Corporation, NSW, Australia) was serially diluted using double distilled H_2O to produce 8 DNA standards (Table 3.6) (concentration ranging from 0.023ng/µl to 50.0ng/µl). The 8 DNA standards were also amplified at both the telomere and 36B4 gene target regions. This produced two standard curves: one for telomere repeats and other one for 36B4 gene. The LightCycler Relative Quantification software version 1.01 (Roche Diagnostic Australia Pty. Ltd. NSW, Australia) was used to generate the standard curves.

Table 3.6 – Dilution factors and final concentrations of the 8 DNA standards.

	Dilution Factor	Final concentration of DNA standards ng/µl
1	4X	50.0
2	3 X	16.7
3	3 X	5.56
4	3 X	1.85
5	3 X	0.62
6	3 X	0.21
7	3 X	0.068
8	3 X	0.023

3.4.3 Measurement of telomere length

For each sample the relative telomere length was calculated using LightCycler Relative Quantification software version 1.01 (Roche Diagnostic Australia Pty. Ltd. NSW, Australia). The absolute concentration of telomere hexameric repeat (T) and absolute concentration of 36B4 gene (S) were obtained using the standard curves

mentioned in Section 3.4.2 using the software LightCycler Relative Quantification software version 1.01 (Roche Diagnostic Australia Pty. Ltd. NSW, Australia). Then the absolute concentration of the telomere hexameric repeat (T) was divided by the absolute concentration of the 36B4 gene (S). The resulting value (T/S ratio) was divided by the T/S ratio obtained for calibrator DNA. A calibrator DNA K562 (5ng/µl) (Promega Corporation, NSW, Australia) was always included in each run of the real time PCR to minimise the errors between PCR runs. The resulting value represents the quantity of telomere hexameric repeats, called relative telomere length (Cawthon 2002).

3.5 Statistical analysis

All the statistical tests were performed using the software SPSS V15 (SPSS Inc, Chicago, USA). Details follow below.

3.5.1 Reproducibility of telomere length measurement

Telomere lengths of 5 different individuals aged 13, 22, 27, 59, and 68 years old were measured on three consecutive days in duplicates. One way ANOVA (analysis of variance) was carried out to check whether there is any variance above at $P > 0.05$ in measured telomere lengths in each individual telomere lengths on different days. This analysis was carried out to check the reproducibility of the real time quantitative PCR method.

3.5.2 Test for normal distribution of telomere lengths of the sample population

A One-sample Kolmogorov-Smirnov test was carried out on telomere lengths of the sample population (n = 167, aged 1-96 years old). This test analyses whether the telomere lengths of the sample population are normally distributed. Most statistical tests assume a normal distribution of data and thus this was tested to see if the standard tests could be used or non-normal distribution statistical tests were required.

3.5.3 Regression analysis between age and telomere length

Linear regression analysis between telomere length and age was carried out to obtain a formula to the predict age of a person.

$$\text{Age } (x) = \frac{\text{Telomere length } (y) - C \text{ (intercept)}}{m \text{ (rate of telomere length loss)}}$$

The R^2 value from the regression analysis, which normally indicates the goodness to fit to a line of regression here, would indicate the age prediction accuracy of the formula.

3.5.4 Telomere length and age in non-related European origin population

A regression analysis between telomere length and age of unrelated Europeans (n = 60) was performed to exclude or minimise the effect of ethnicity and inheritance on telomere length.

3.5.5 Telomere length and age groups

The samples were divided into nine groups, each a decade long, according to their age (Table 3.2). Mean telomere lengths of each age group were plotted against the

corresponding decade of age. The variation in mean telomere length of each age group was compared with each other using One-way ANOVA.

3.5.6 Telomere length differences in young and old

The population was divided into two groups: "young" aged 1-39 years (n = 101) and "old" aged 40-89 years (n = 65) to see if there were major differences in telomere lengths between these groups. The mean telomere lengths of these two groups were compared using the independent t-test.

3.5.7 Comparison of telomere lengths between similarly aged individuals

Two groups of similarly aged individuals, age 26 (n = 10) and age 54 (n = 9), were chosen for this part of the analysis. These two groups were chosen as they had the most number of samples compared to the other age groups. The minimum, maximum, mean and standard deviations of the relative telomere lengths of each age group were compared with each other.

3.5.8 Comparison of male and female telomere length

Regression analysis of telomere length and age was also carried out separately on males (n = 71) and females (n = 96). The mean telomere length was also calculated for each group and the median age noted.

3.5.9 Relationship between telomere length and age in families

Pearson correlation and regression analysis between telomere length and age were carried out on data from 18 families which had more than 3 members per family.

3.5.10 Telomere length differences between families

The variations between telomere length and age between the families were compared with each other using One-way ANOVA.

3.5.11 Effect of parental telomere length on children's telomere length

Partial correlation between parental telomere length and children's telomere lengths were calculated, with adjustment for age. The age adjustment was included in order to minimise the effect of parents' and children's age on the analysis.

3.5.12 Effect of parents' age at conception with telomere lengths of children

Partial correlation between parents' age at conception and telomere lengths of children was calculated. Pearson correlation between parent's age at conception and children's age was carried out. Further, to remove the effect of the children's ages, a data group of the same age, 26 years (n = 10), was selected for a separate Pearson correlation analysis with their telomere lengths and their parents age at conception.

3.6 Mitochondrial mutation analysis

3.6.1 Amplification of HV1 and HV2 regions of mitochondria

HV1 and HV2 regions of mitochondria were amplified using the following primers (Table 3.7) as previously designed (Parson *et al.* 1998). The HV2 region was amplified twice using two sets of forward and reverse primers (discussed in section 6.4.2). All the primers were commercially synthesized at GeneWorks (GeneWorks Pty Ltd, South Australia, Australia). Stock solutions of primer with an initial concentration of 100μM were made and stored at -20^{o}C.

Table 3.7 – Primer sequences used in the mitochondrial DNA mutation analysis.

	Primer pairs and sequences 5/ - 3/	Expected PCR product (bp)
HV1 region	**F15997** - 5/ CAC CAT TAG CAC CCA AAG CT 3/ **R16401** - 5/ TGA TTT CAC GGA GGA TGG TG 3/	~ 400bp
HV2 region	**F29** - 5/ GGT CTA TCA CCC TAT TAA CCA C 3/ **R408** - 5/ CTG TTA AAA GTG CAT ACC GCC A 3/	~ 379bp
	F 15 - 5/ CAC CCT ATT AAC CAC TCA CG 3/ **R 389** - 5/ CTG GTT AGG CTG GTG TTA GG 3/	~ 374bp

PCR was performed in a total volume of 25 µl consisting of 3 ng of DNA, 10X PCR buffer, 25 mM $MgCl_2$, 2.5 mM dNTP each, 5 µM forward primer, 5 µM reverse primer, and Platinum Taq polymerase (Invitrogen Australia Pty Limited). Amplification was carried out by incubation at 94^{o}C for 11 minutes, followed by 32 cycles of 94^{o}C for 20 seconds, 56^{o}C for 10 seconds and 72^{o}C for 30 seconds in a PTC100, Programmable Thermal controller (MJ Research).

3.6.2 Gel electrophoresis

PCR products were run in 1X TBE buffer [10.8g Tris base (Sigma Aldrich Pty Ltd. NSW, Australia), 5.5g Boric acid (Sigma Aldrich Pty Ltd. NSW, Australia), 0.5M EDTA (BDH), pH 8.00 in 1L], 2% DNA grade agarose (Progen Pharmaceuticals Limited, Queensland, Australia), and 0.5 µg/ml Ethidium bromide (BDH), and a 100 bp molecular weight marker (Roche Diagnostic Australia Pty. Ltd. NSW, Australia) was included in order to be able to estimate the molecular weight of the amplified products.

3.6.3 Purification of PCR products

PCR products were directly purified from the solution using GFX PCR DNA and Gel Band purification kit (GE Healthcare Australia & New Zealand). The manufacturer's protocol was followed. This method involves purifying samples using a column based method. The GFX column was placed in a collection tube. 500 µl of capture buffer was added to the GFX column. The PCR product (~21µl) was transferred into the capture buffer and mixed thoroughly with the buffer by pipetting up and down 4 to 6 times. The sample was centrifuged for 30 seconds, and the flowthrough was discarded. 500 µl of wash buffer was added to the column and centrifuged for 30 seconds. The GFX column was placed in a new eppendorf tube and the collection tube was discarded. 50 µl of autoclaved double distilled water were placed directly on top of the glass matrix in the GFX column. The samples were then incubated for 1 minute at room temperature. All the samples were centrifuged at 1 minute for 14,000g. The purified PCR product was recovered in eppendorf tubes. Samples were stored at 4^{o}C.

3.6.4 Estimation of DNA concentration by agarose gel electrophoresis

The concentration of purified PCR products was estimated by comparing the intensity of the purified DNA with that of the bands of DNA Molecular weight marker XIV (Roche Diagnostic Australia Pty. Ltd. NSW, Australia), run simultaneously on a gel made of DNA-grade agarose (Progen Pharmaceuticals Limited, Queensland, Australia) gel. The concentration of each band in the Molecular weight marker XIV set was provided by the manufacturer (Roche Diagnostic Australia Pty. Ltd. NSW, Australia). These were used to visually estimate the concentration of the PCR products. For example, 4 µl of the 500 bp band DNA

was run on the 2% agarose gel. This was at a concentration of 105ng/4 µl. Thus, there would have been ~ 100 ng of DNA in this band. By visually comparing the intensity of this band with that of the PCR product, one could roughly estimate how much DNA was there in the PCR reaction represented by a particular sample band. The appropriate amount of PCR product (75-100ng) was then used directly in sequencing reactions.

3.6.5 Sequencing of Mitochondrial HV1 and HV2 regions

Both forward and reverse strands of HV1 were sequenced, and the forward strand of HV2, were sequenced twice using two different set of primers. The Bigdye® Terminator V3.1 Cycle Sequencing Kit (Applied Biosystems Australia, Vic, Australia) was used for the sequencing. The reason for sequencing the forward strand of HV2 region twice was that sequencing of HV2 reverse sequence repeatedly gave poor sequence chromatograms (discussed in Section 6.4.2).

3.6.5.1 Amplification of HV1 and HV2 regions for sequencing

DNA sequencing was performed as follows. Master mix was prepared containing 3µl of dilution buffer [250mM Tris HCl (Sigma Aldrich Pty Ltd. NSW, Australia) pH 9.0), 10mM $MgCl_2$ (BDH)], 1 µl of forward or reverse primer (3.2µM), 2µl of Bigdye (Applied Biosystems Australia, Vic, Australia), and double distilled H_2O up to 20 µl. 8µl of template DNA from the experiments described in 3.6.3 (50-100 ng), were added to the PCR tubes, followed by 12 µl of master mix added to each tube. The PCR program used for this sequencing was 44 cycles of 96^oC for 30seconds, 56^oC for 15 seconds, and 60^oC for 4 minutes, then followed by 5 cycles of 96^oC for

30 seconds, and 60^{o}C for 4 minutes, and finally at 25^{o}C for 60 seconds and storage at 4^{o}C until required (Parson *et al.* 1998).

3.6.5.2 Sequencing precipitation of DNA

The sequencing reaction products were precipitated using sodium acetate and ethanol. 80 µl of the precipitation master mix (3µl of Sodium Acetate, pH 5.2, 63µl of 95% ethanol, and 14 µl of double distilled water) was added up to 20 µl of sequencing PCR product. The samples were then vortexed (MT 19 Auto vortex Mixer Chiltern) for 20 seconds and incubated at room temperature for 15 minutes. Then the samples were then centrifuged (Eppendorf Centrifuge 5415C) for 20 minutes. The supernatant was removed using a 1ml syringe for each sample. Precipitated DNA samples were washed with 250 µl of 75% ethanol (BDH). The samples were centrifuged for 5 minutes and supernatant was removed using a 1ml syringe. Samples were air dried and sent to the Micromon DNA sequencing facility, Bdg 75 Monash Strip, Room 220, Monash University, VIC 3800.

3.6.5.3 Analysis of mitochondrial DNA sequences

Sequence editing and alignments were conducted using the BioEdit Sequence Alignment Editor version 7.0.5.2, [1997-2005 Tom Hall (www.mbio.ncsu.edu/BioEdit/bioedit.html)]. The quality of the DNA sequence was assessed by visual analysis of the trace file using Chromas LITE 2.0, 1998-2004 Technelysium Ltd, Australia (http://www.technelsium.com.au/chromas_lite.html).

3.6.5.4 Mitochondrial DNA sequence alignments within a family

The HV1 and HV2 sequences of all the members of each family were compared with each other within the family, using the ClustalW Multiple alignment programme in BioEdit. The youngest member's sequence was considered as the reference sequence and compared to those of the older members of the family. All the alignments were saved using the 'Graphic View' option in BioEdit. .

A similar analysis comparing mitochondrial HV1 and HV2 sequences within a family, using the youngest member's sequence as the reference, was carried out using software called Mutation surveyor version 3.1 (Soft Genetics, PA, USA) (Song *et al.* 2005). Unlike in the ClustalW program, where the sequences are compared, in Mutation Surveyor, sequence chromatograms are compared with each other.

3.6.5.5 Mutation Surveyor

Mutation Surveyor compares sequence chromatograms (not DNA sequences) and identifies differences/mutations present (Song *et al.* 2005). A probable mutation is determined by the mutation score calculated by the software. The mutation score is calculated based on background noise level, the overlapping factor and the dropping factor. The overlapping factor is calculated by comparing the sample and reference bases. The dropping factor is the relative intensity drop of the sample trace compared to the reference trace. The SnRatio determines how large the signal is compared with the background noise level. If the mutation score exceeds 500 on the sequences of both strands or of both forward sequences, it is considered to be a probable mutation (SoftGenetics 2006).

3.7 Analysis of mitochondrial deletions

3.7.1 Amplification of whole mitochondrial genome in humans

3.7.1.1 Amplification of whole mitochondrial genome by Long PCR

The mitochondrial genome of the youngest member of the 37 individuals of the multiple maternal generational families (Table 3.4) was attempted to be amplified by Long PCR as described by Meissner *et al.,* (Meissner, C *et al.* 2006) using triplicate primers as described in Cheng *et al*., (Cheng *et al.* 1994) (Table 3.8). The Long PCR master mix contained 20 ng of DNA, 1µl of 10X Opti buffer (Bioline, Lukenwalde, Germany), 25 mM of $MgCl_2$, 0.5 µl each of dNTPS, 2 µM each of forward and reverse primers, 0.2 µl of Platinum Taq antibody (Invitrogen), 0.25 µl of Bio-X-Long DNA Polymerase (Bioline, Lukenwalde, Germany) and ddH_20 added up to 10 µl. The PCR conditions for the reaction were 94^{o}C for 2 minutes, followed by 20 cycles at 93^{o}C for 10 seconds and 68^{o}C for 13 minutes and 30 seconds. This was followed by an additional 5 cycles under identical conditions, but with a 20 second elongation added to each extension cycle.

Table 3.8 – Primer sequences used in the whole mitochondrial DNA amplification using Long PCR.

Primer pairs and sequences 5/ - 3/	Expected PCR product (bp)
Forward - 5/ TGAGGCCAAATATCATTCTGAGGGGC3/ **Reverse** - 5/ TTTCATCATGCGGAGATG TTG GAT GG3/	~16kb

3.7.1.2 Amplification of whole mitochondrial genome using REPLI–g-Mitochondrial DNA kit

As the Long PCR was not successful after several attempts under various conditions, another method was tried. DNA samples from daughters of family #8 and family # 9 (Table 3.4) were amplified using the REPLI–g-Mitochondrial DNA kit (QIAGEN Pty. Ltd Australia). The manufacturer's instructions were followed. The sample DNA (1.5 ng/µl) was denatured by incubation with REPLI-g mt reaction buffer (QIAGEN Pty. Ltd Australia) for 5 minutes at 75^{o}C. The denaturation was stopped by cooling the solution to room temperature. Then REPLI-g DNA polymerase was added and the isothermal amplification was allowed to proceed overnight at 33^{o}C. The REPLI-g DNA polymerase was inactivated by heating the samples for 3 minutes at 65^{o}C.

3.7.1.3 Gel electrophoresis

Amplified PCR products (4µl each), DNA Molecular Weight Marker λDNA, EcoRI + HindII marker (Fermentas) were run on a 1% agarose gel, the composition of which was the same as that described in Section 3.6.2 above, except for the concentration of agarose.

3.7.2 Preparation of the digoxigenin (DIG) labelled probe

The DIG High Prime DNA Labelling and Detection Starter Kit II (Roche Diagnostic Australia Pty. Ltd. NSW, Australia) were used to label the probe.

3.7.2.1 Purification of the PCR product

The amplified whole mitochondrial genome of a younger individual was purified. At first the PCR product was run on a gel, and the ~16 kb band was excised. Purification of this was attempted using the QIAEX II kit (QIAGEN Pty. Ltd Australia) kit. However, as this was not successful after repeated attempts, the whole PCR product was used in the purification, which was purified directly from the solution, using the same kit. The purification of the DNA is based on selective and quantitative adsorption of nucleic acids to the QIAEX II silica-gel particles in the presence of high salt.

3.7.2.2 DIG-labelling of the probe

The purified PCR product was labelled with DIG. 1µg of purified PCR product was added to 16µl of double distilled water. The DNA was denatured by heating in a boiling water bath for 10 minutes. 4µl of DIG-High prime (Roche Diagnostic Australia Pty. Ltd. NSW, Australia) was added in to the denatured DNA. The vial was incubated at 37^{o}C overnight. The reaction was stopped by heating the vial for 65^{o}C for 10 minutes.

3.7.2.3 Determination of labelling efficiency

The labelling efficiency of the probe was evaluated. A dilution series of labelled probe and DIG-labelled control DNA was prepared as stated in the table below (Table 3.9).

Table 3.9 – Dilution series of labelled probe and the control DNA.

Tube	DNA (µl)	From tube #	DNA dilution Buffer (µl)	Dilution (µl)	Final Concentration
1		Diluted original			1 ng/µl
2	5	1	495	1:100	10 pg/µl
3	15	2	35	1:3.3	3 pg/µl
4	5	2	45	1:10	1 pg/µl
5	5	3	45	1:10	0.3 pg/µl
6	5	4	45	1:10	0.1 pg/µl
7	5	5	45	1:10	0.03 pg/µl
8	5	6	45	1:10	0.01 pg/µl
9	0	-	50	-	0

1 µl each of the dilution series of DIG labelled probe and the DIG-labelled control DNA was applied on to a small strip of positively charged nylon membrane, Hybond N+ (Amersham Biosciences). The DNA was fixed to the membrane by cross-linking with ultraviolet light (UV) on a UV 2011 Macrovue Transilluminator (LKB, Bromma, Sweden) for 5 minutes. The membrane was then transferred into a hybridisation bottle filled with 20 ml of Maleic acid buffer (0.1 M Maleic acid, 0.15 M NaCl pH 7.5) and incubated for 2 minutes at 15-25^{o}C with constant agitation. The Maleic acid buffer was discarded and then the membrane was incubated with 10 ml of 1 X Blocking solution [diluted 10 X Blocking solution (Roche Diagnostic Australia Pty. Ltd. NSW, Australia), in Maleic acid buffer] for 30 minutes. The membrane was next put into 10ml of Antibody solution [Diluted 1:10 000 anti-digoxigenin–AP (Roche Diagnostic Australia Pty. Ltd. NSW, Australia), in Blocking solution] for 30 minutes. After the above incubations, the membrane was washed

twice with 10ml of washing buffer (0.1 M Maleic acid, 0.15 M NaCl pH 7.5, 0.3% 9 (v/v) Tween 20). Then the membrane was equilibrated for 5 minutes in 10 ml of Detection buffer (0.1 M Tris-HCl, 0.1 M NaCl). The membrane was then put in a stomacher bag facing upwards, and 0.1 ml of CSPD Ready Use (Roche Diagnostic Australia Pty. Ltd. NSW, Australia) was added to the top of the membrane. This was immediately covered with another stomacher bag, and CSPD was spread smoothly over the membrane, taking care to avoid and remove air bubbles. The membrane was incubated for 5 minutes and was exposed to Luminescent Image Analyser LAS – 1000 plus (Fuji Film). The intensity of the DIG labelled probe spots were compared with control DIG labelled DNA (Roche). According to the manufacturer's instructions, if the 1 pg dilution of the DIG labelled probe and of the control DIG labelled DNA (Roche) were visible, the labelled probe had reached the expected labelling efficiency.

3.7.2.4 Termination of these experiments due to technical difficulties and time restraints

At this stage, a major problem remained that Long PCR was still not successful in amplifying whole mitochondrial DNA from the samples. Whilst the product from the REPLI–g-Mitochondrial DNA kit (QIAGEN Pty. Ltd Australia) was suitable for use as a probe, it could not be used for amplifying sample DNA for use in Southern blotting to detect deleted mitochondrial DNA. This is because it uses Multiple Displacement Amplification (MDA), which results in several incomplete mtDNA fragments (as well as whole mtDNA), which would confound results in the Southern blots. This is explained in more detail in the Discussion (Section 7.3) in Chapter 7.

The following chapters present the results and the analysis of data obtained following the methods described in this Chapter 3.

Chapter 4

Telomere length and its and relationship to age

4.1 Introduction

One of the main theories on ageing is the "telomere hypothesis", suggesting that the continuous shortening of telomeres during cell division is primarily responsible for ageing (Boukamap 2001). As discussed in section 2.5, telomeres are specialised DNA and protein structures that are found at the ends of linear chromosomes (Greider 1999). The main role of telomeres is to protect chromosome ends from recombination and fusion, and also to prevent them from being recognised as damaged DNA and thus degraded (Greider 1998). Telomeric DNA consists of simple tandem repeated sequences characterized by clusters of G residues in one strand. In humans this sequence is 5'-TTAGGG-3' (Blackburn 1990). During DNA replication, the lagging strand is replicated in discontinuous pieces called Okazaki fragments. The synthesis of 5'→3' DNA fragments require RNA primers, which are finally replaced by DNA with the exception of those at the very end. Because of this, the far end of the telomere is not replicated and the telomeres shorten with each round of DNA replication (Boukamap 2001). Thus, human somatic cells gradually lose telomeric repeats. In germ cells, stem cells and many cancer cells, telomere length is kept constant by expression of the enzyme telomerase (Kufe *et al.* 2003), which adds back telomeric DNA.

The first direct evidence for telomere loss during cellular ageing was obtained by analysis of cultured human fibroblasts (Harely *et al.* 1990). *In vivo* loss of telomere length with age was first observed in skin cells by Lindsay *et al.,* (Lindsey *et al.* 1991) in 1991. Thereafter, there were many observations of human telomeres

shortening with age in tissues such as fibroblasts, skin, blood, epithelium, liver, thyroid, parathyroid, brain, kidney, vascular tissue, myocardium, spleen, and pancreas (Butler, MG *et al.* 1998) .

Previous attempts have been made to devise a formula to predict the age of an individual by analysing telomere length (Takasaki *et al.* 2003; Tsuji *et al.* 2002) using the mean terminal restriction fragment (TRF) length. Studies by Tsuji *et al.* (Tsuji *et al.* 2002) and Takahashi *et al.*(Takasaki *et al.* 2003), devised formulas to predict the age of an individual within a range of ± 7.5 years, from the person's actual age. One of the major drawbacks of the TRF method is that it requires a large amount of DNA (0.5-5 µg/individual), and is also time consuming (3-5 days) (Cawthon 2002). Often the amount of DNA available for forensic analysis is very low. Another disadvantage is that telomere lengths measured by the TRF method include subtelomeric regions (Nakagawa *et al.* 2004). A method based on PCR developed by Cawthon (Cawthon 2002) solved the problems associated with low amounts of DNA to 17.5ng/ µl and also reduce the time required for analysis to a couple of hours. In addition, primers were designed by Cawthon (Cawthon 2002) that hybridize specifically to human telomere regions and only telomeric (TTAGGG) repeats are amplified; subtelomeric regions are not included. Thus, this would give a more accurate representation of the actual telomere lengths. Telomere length measured using the real time PCR method could increase the prediction accuracy by cutting down the limitations existing in TRF method. The narrower the range of predictable age the more beneficial it would be for forensic investigations.

Here, telomere lengths were measured using real time PCR in individuals from different age groups. Buccal swabs were collected for this study, since it is a non-invasive method of collecting DNA from individuals and because saliva is one of the major types of potential biological material found at crime scenes (Butler 2005). The aim was to investigate the correlation between telomere length and age, and to see if this relationship could be used to develop a formula to predict the age of an individual within a narrow range suitable for forensic application. This study was published in journal 'Legal Medicine' as "Investigation of telomere lengths measurements by quantitative real time PCR to predict age" (article in press, 2008).

4. 2 Results

4.2.1 Standard curves

K562 DNA, High molecular weight DNA (Promega) was serially diluted using ddH_2O to produce 8 standard DNA concentrations ranging from 50.00 ng/µl to 0.023 ng/µl. Standard curves were generated by the Light Cycler software, RelQuant (Roche) using the second derivative maximum method for both telomere and 36B4 reactions. In this the method the crossing point (CP), which is the point at which the fluorescence of a reaction is detectable above the background fluorescence set at a particular threshold, is determined automatically for individual samples. Base line adjustment, noise band or crossing line settings are not done manually. Therefore there is no user influence in this method (Roche light Cycler manual).

The slope of the standard curve represents overall reaction efficiency. To achieve efficiency for the standard curve between 1.5 and 2.2, the slope should be between -5.7 and -2.9 (Roche Light Cycler manual). The slopes of telomeres and the 36B4 gene both were within the acceptable range (-4.308 and -3.566 respectively).

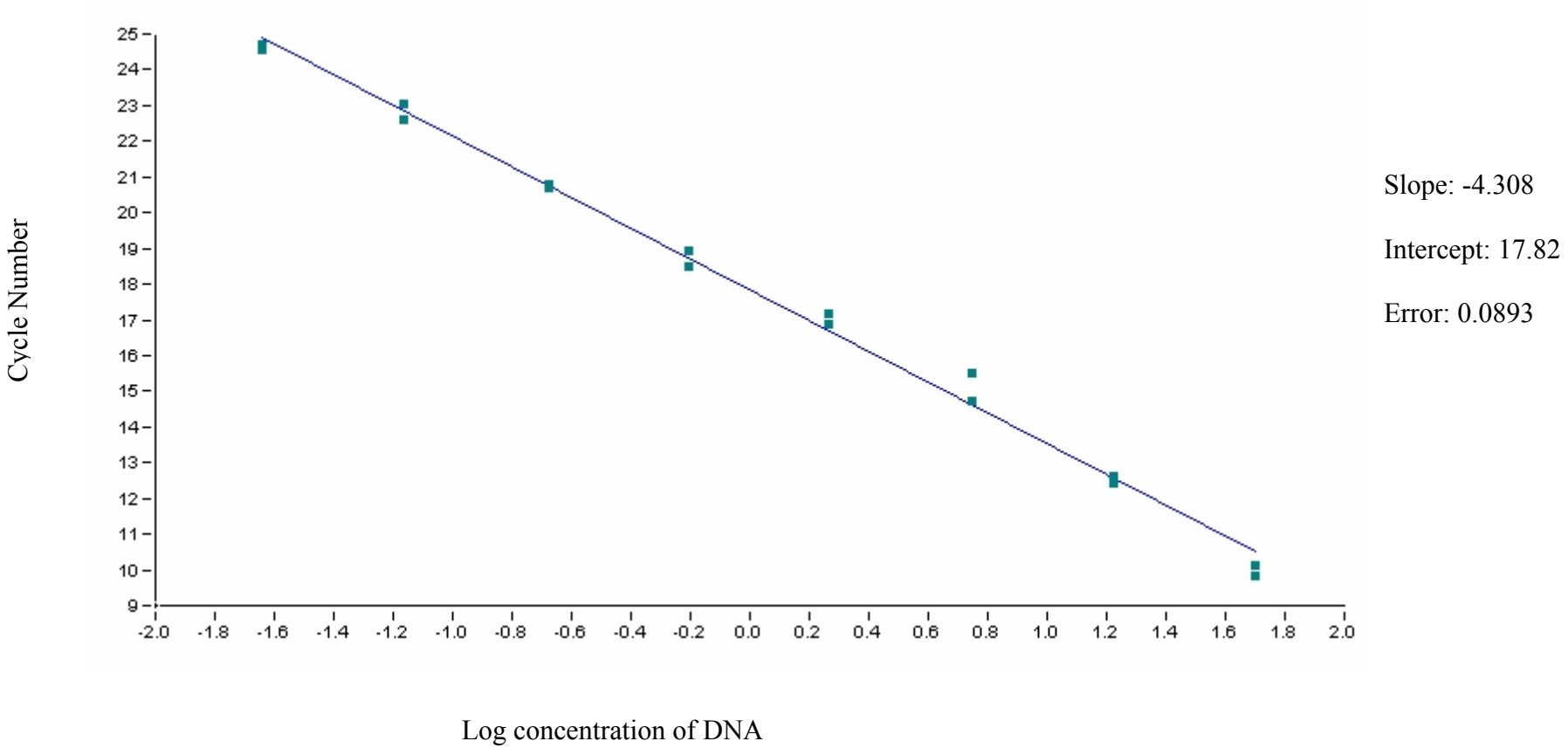


Figure 4.1 –Standard curve of cycle number at which the crossing point was reached versus log concentration of DNA standards. This experiment was conducted using telomere-region-specific primers.

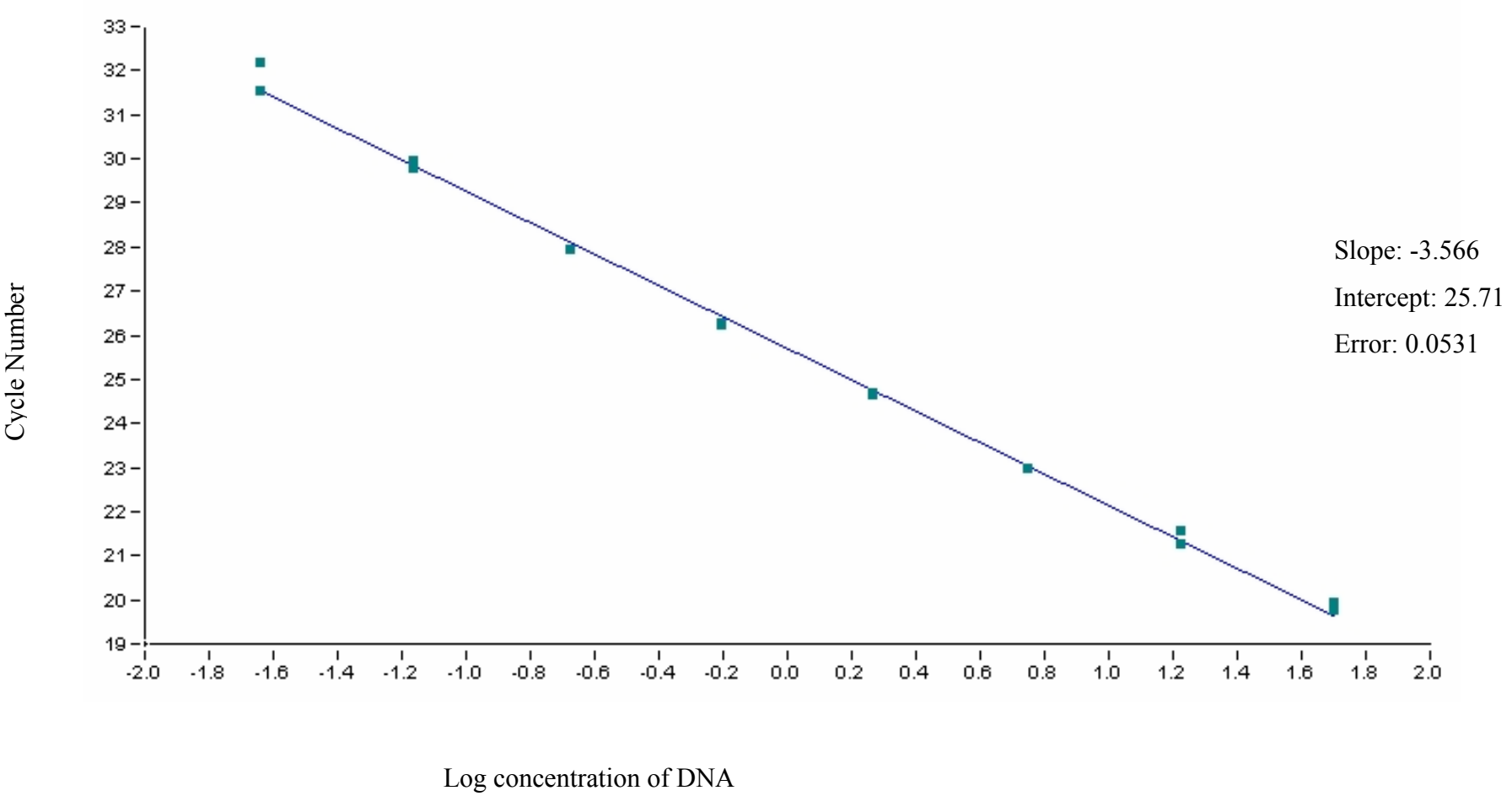


Figure 4.2 –Standard curve of cycle number at which the crossing point was reached versus log concentration of DNA standards. This experiment was conducted using primers targeted at the reference gene (36B4).

4.2.2 Change of DNA concentrations in Real Time PCR

As mentioned above, in real time PCR, the cycle at which fluorescence of a sample is detectable above the background fluorescence, set at a particular threshold, is called the crossing point (CP). The crossing point of a sample depends on the initial concentration of DNA. Samples with initial low DNA concentration need more amplification cycles than those with initial high DNA concentrations (Roche light Cycler manual). In the Cawthon method of quantitative real time PCR analysis of telomere length, the initial DNA concentration used was 17.5 ng/ µl (Cawthon 2002; Gil *et al.* 2004). In this study the concentration of DNA extracted from buccal swabs varied from 0.238 to 85 ng/µl. Since many of the samples we had collected had less than 17.5 ng/µl, the starting DNA concentration was changed to 5 ng/µl. Because of this, the number of cycles in the real time PCR had to be increased to obtain the CP value (crossing point). Therefore, the number of cycles was increased to 50 in both Real time PCR cycles (for telomere length amplification: 20 to 50, for amplification of the reference gene, 36B4: 30 to 50).

4.2.3 Reproducibility of telomere length measurement

Telomere length measurement using Real Time Quantitative PCR was evaluated for reproducibility. Telomere lengths of five samples aged 13, 22, 27, 59 and 68 years old were measured on three different days in duplicate. There was no significant difference ($P<0.05$) between telomere length measurements for any of the samples evaluated on the three different days (Table 4.1). (With the degrees of freedom for these samples, at 1 % significance level the critical region is $F >= 9.55$. As all F values in the table are well below this, the null hypothesis is accepted i.e. there was

no significant level of variation between measurements of each sample in duplicate on the three different days) (Koosis 1997).

Table 4.1 - One way ANOVA table of Telomere length measurements of five samples on three different days.

		Sum of Squares	df	Mean Square	F	Sig.
Sample1	Between Groups	.002	2	.001	1.367	.378
	Within Groups	.002	3	.001		
	Total	.005	5			
Sample2	Between Groups	.000	2	.000	0.475	.662
	Within Groups	.000	3	.000		
	Total	.000	5			
Sample3	Between Groups	.000	2	.000	0.532	.634
	Within Groups	.001	3	.000		
	Total	.002	5			
Sample4	Between Groups	.013	2	.006	5.308	.103
	Within Groups	.004	3	.001		
	Total	.016	5			
Sample5	Between Groups	.002	2	.001	1.296	.393
	Within Groups	.002	3	.001		
	Total	.004	5			

4.2.4 Telomere length measurement in the sample population

The telomere length of 167 individuals aged 1-96 years was measured. Some of the older individuals had a longer telomere length than younger individuals (Figure 4.3). One-sample Kolmogorov-Smirnov test of the results showed that telomere length measurements were normally distributed ($P<0.05$) (data not shown).

4.2.5 Regression analysis between age and telomere length

Relative telomere length decreased with age ($r = -0.185$, $P < 0.05$). The formula for age prediction derived from these results was: Age = (relative telomere length – 1.5) / -0.005. However, the relationship between telomere length and age has an R^2 value of 0.03679. This would give an age prediction accuracy of approximately 4%. This is far too low to be used in forensic investigations, which needs higher accuracy (~70%).

4.2.6 Telomere length and age in a non-related European origin population

Telomere lengths in people of unrelated European origin (n = 60) were analysed in order to exclude or minimise the effects of ethnicity and inheritance effects on telomere length. Relative telomere length decreased with age ($r = -0.233$, $P<0.05$), the R^2 increased slightly to 0.0545 (Figure 4.4), thus the prediction accuracy also increased slightly to ~5.5% However, this is still very far from the accuracy required for forensic investigations.

4.2.7 Telomere length and age groups

Due to the large variation in telomere lengths in the population, the population was divided into 10 year age groups representing each decade of life. Mean telomere lengths of each age group were plotted against age groups (Figure 4.5). The mean telomere length appeared to remain stable in the age groups of 0-9 years, 10-19 years and 20-29 years, then decline from 30-39 years and continued until the age group of 50-59 years. However, one-way ANOVA showed that there was no significant difference in telomere lengths between any of the age groups ($P>0.05$).

4.2.8 Telomere length differences in young and old

Since no significant difference in telomere length was seen between age groups, the sample was divided into two main groups: 'young' individuals aged 1-39 years (n=101) and 'old' individuals aged 40-89 (n= 65) years old. An independent t-test (Table 4.2) showed there is significant difference in telomere lengths between the "young" and the "old" ($P<0.05$). (The null hypothesis here is that there was no significant difference between the two groups. At 164 degrees of freedom, the critical region for rejection of the null hypothesis is at 5% significance t>= 1.65 or t<= -1.65 and for 1% significance, t >= 2.33 or t<= -2.33, the t values in Table 4.2 are 2.971 (equal variances assumed) and 3.117 (equal variances not assumed). Thus both these t values indicate that the null hypothesis is to be rejected, i.e. there was significant difference in telomere length between the two groups at the 5% and 1% confidence levels).

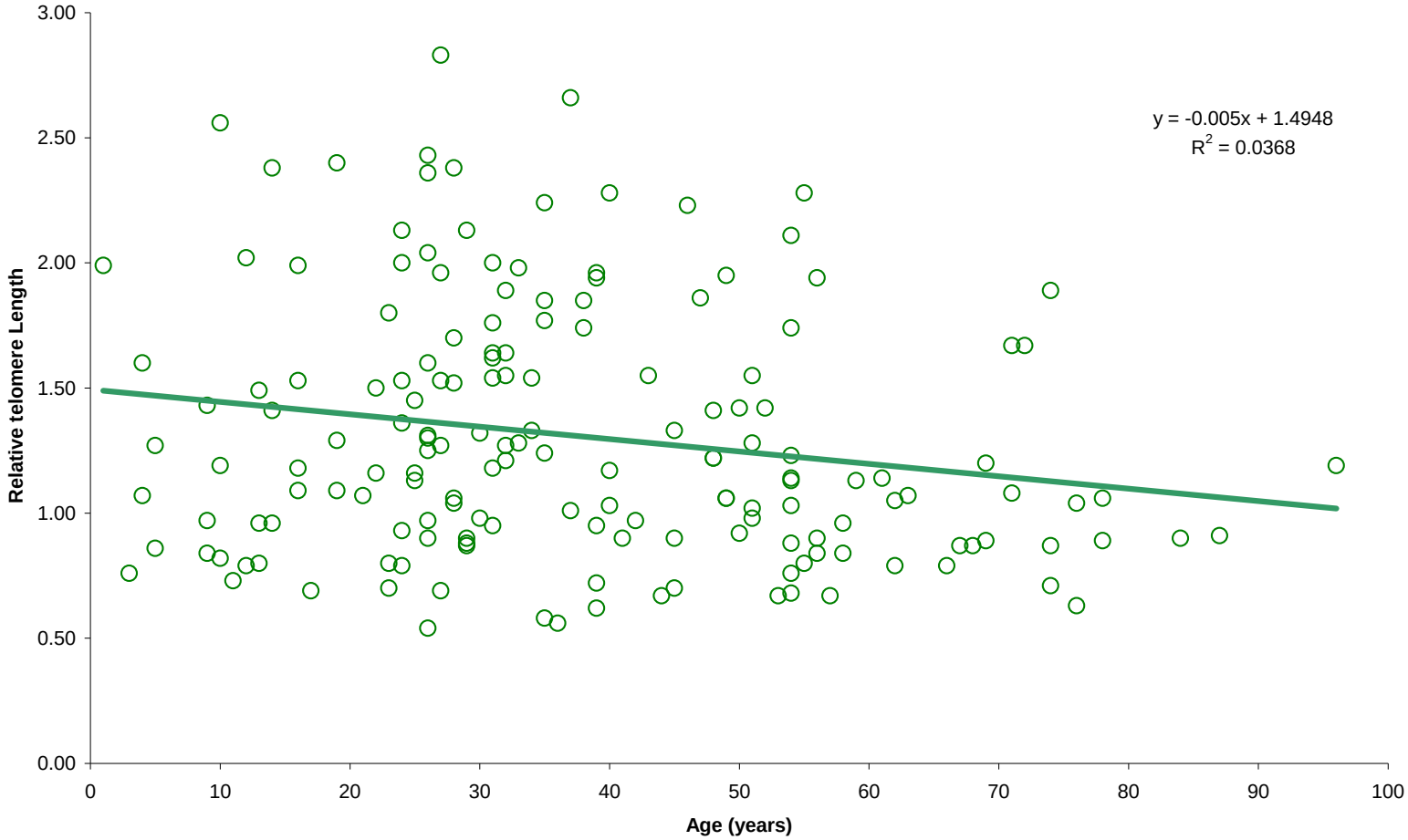


Figure 4.3 – Relative telomere length in relation to age in individuals aged 1 to 96 years old (n = 167).

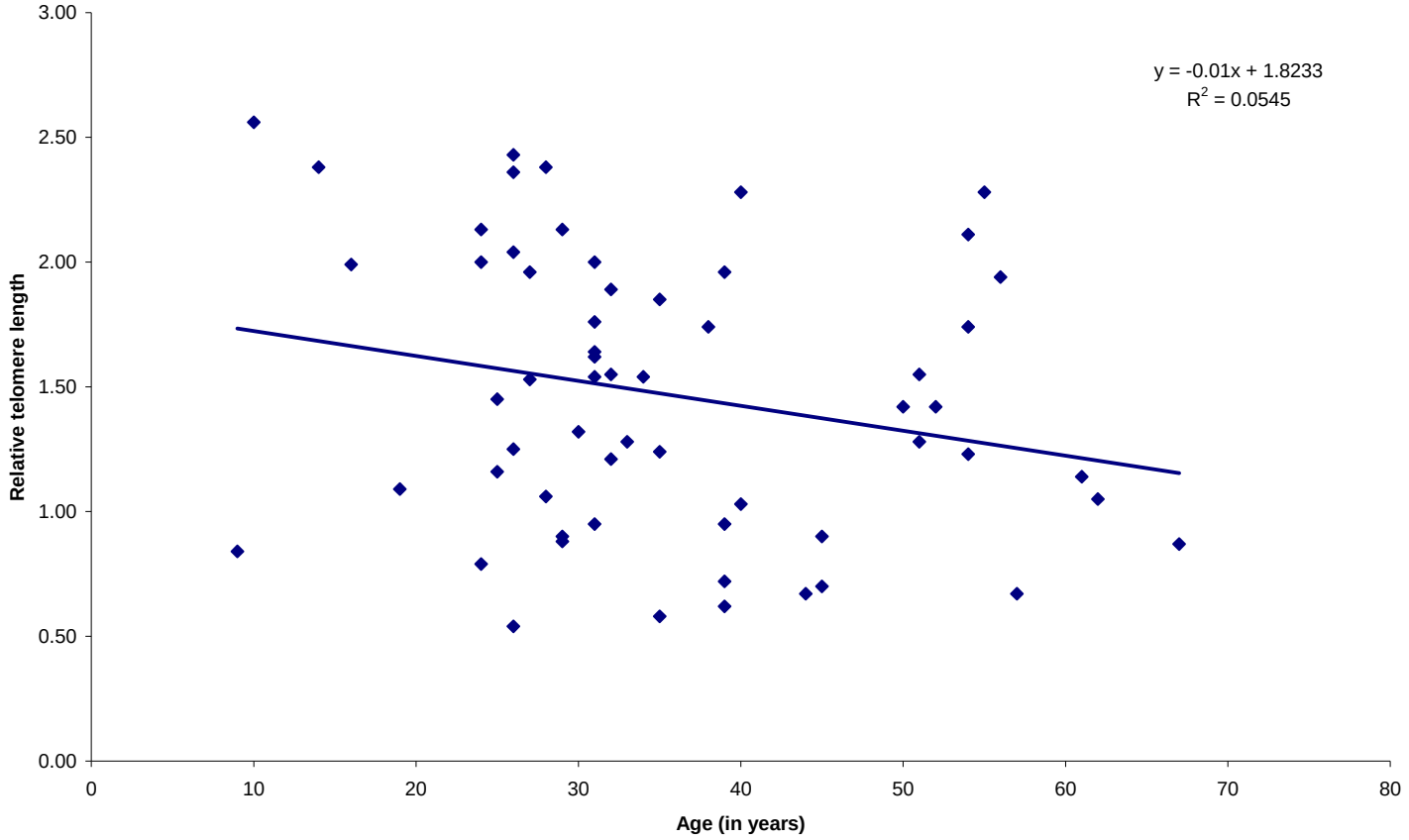


Figure 4.4 - Relative telomere lengths of unrelated European origin individuals (n= 60) decrease with age (r = 0.233, P <0.05). The age prediction accuracy increased to 5%.

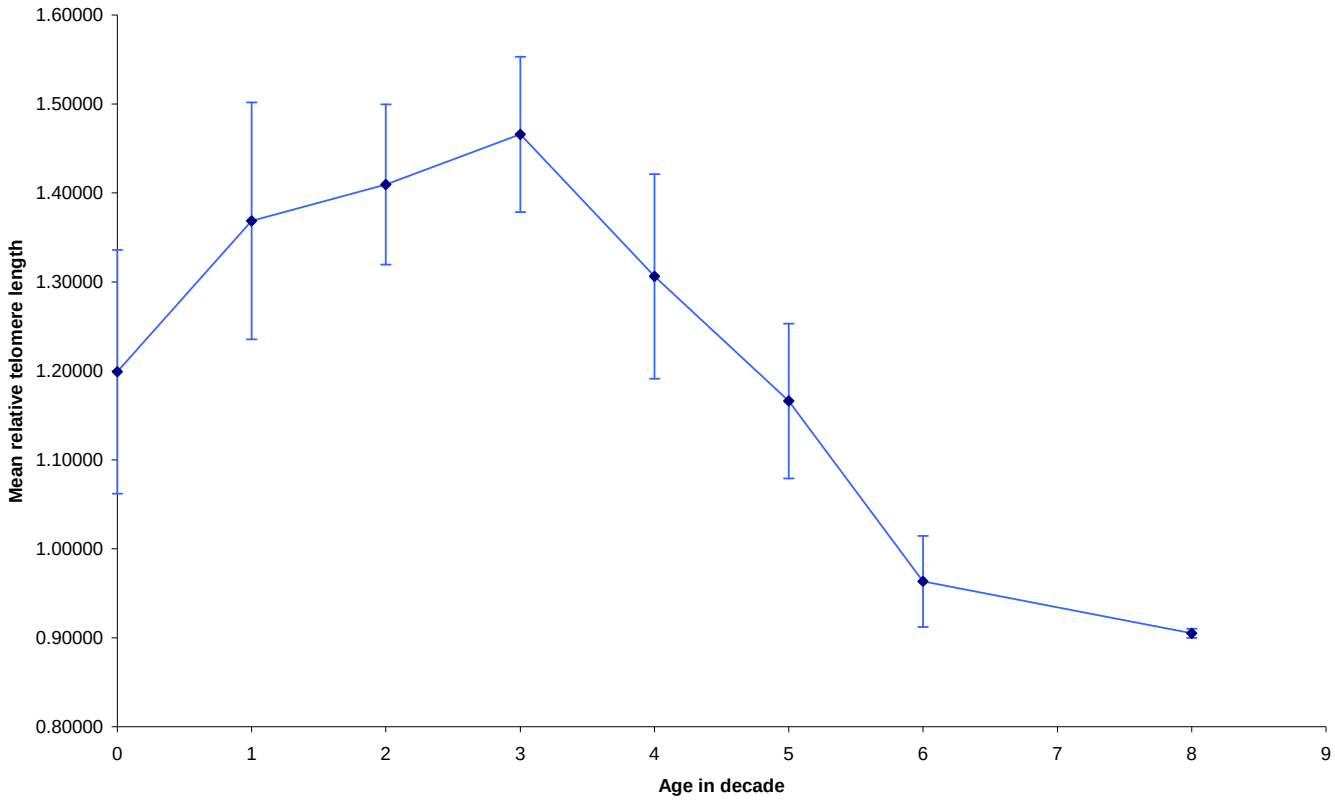


Figure 4.5 – Mean relative telomere length in relation to its age group in decades. Results are expressed as ± standard error.

Table 4.2 – Results of independent t-test for mean telomere lengths of 1-39 year olds ('Young') and 40-89 years olds ('old').

		Levene's test for Equality of Variances		t-test for Equality of Means						
		F	Significance	t	df	Significance (2-tailed)	Mean Difference	Standard error difference	95% Confidence Interval of the Difference	
									Lower	Upper
Telomere length	Equal variances Assumed	6.071	0.015	2.971	164	0.003	0.23453	0.07894	0.07866	0.39040
	Equal variances Not assumed			3.117	156.303	0.002	0.23453	0.07524	0.08590	0.38316

4.2.9 Telomere length differences in same age individuals

Figure 4.3 shows that there is a high amount of variation in telomere lengths between individuals. Individuals of two ages (26 and 54 years old) were selected for analysis of within age-group variation. These groups were chosen as they had the highest number of individuals when compared to other ages. Relative telomere lengths of 10 individuals aged 26 years and 9 individuals aged 54 years were analysed (Table 4.3). From the large values of standard deviations that were found, it can be seen that there is a high variation in telomere lengths in individuals of similar age.

Table 4.3 - Comparisons of relative telomere lengths between two groups of similar aged individuals and the sample population.

Relative telomere length	Age 26 (n=10)	Age 54 (n=9)	Sample (n=167)
Minimum	0.54	0.68	0.54
Maximum	2.43	2.11	2.83
Mean	1.47	1.19	1.31
Standard Deviation	0.63	0.46	0.51

4.2.10 Comparisons of male and female telomere length

Regression analysis was performed on the relationship between age and telomere length in males (n = 71) and females (n = 96) separately. In the sample, females were aged 3-96 years old and males 1-78 years old. The median age of the female sample was 37.72 years and in males it was 36.59 years. The relative telomere length decreased with age in both the sexes. In males there was significant correlation between age and telomere length ($r = -0.300$, $P<0.05$). In females the correlation between age and telomere length was not statistically significant ($r = -0.108$, $P>0.05$). Mean telomere lengths were similar between males and females (1.41 and

1.24 respectively). The steeper slope of the downward line of regression in Figure 4.6 appears to indicate that males lost telomere length more rapidly than females. However, this was not a longitudinal study, but rather the study of different individuals of various ages, and telomere length loss was not actually measured.

Even after separating the sample population into males and females, the age prediction accuracy still did not increase to the point that would be considered of use for forensic purposes ($R^2 = 0.09$ in males and $R^2 = 0.01$ in females) (Figure 4.6).

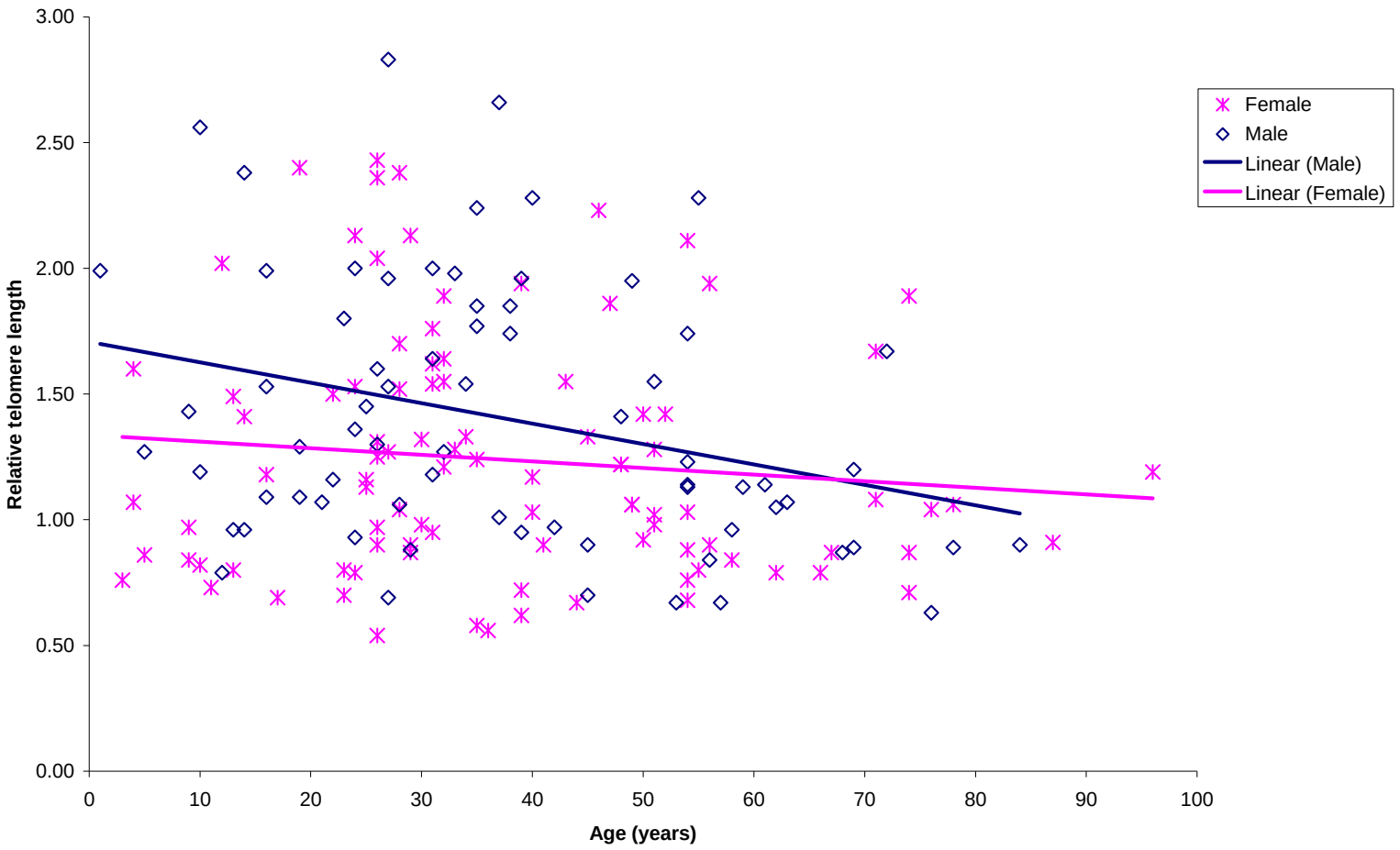


Figure 4.6 – The relationship between relative telomere length in men (n = 71) and women (n= 96). The median age of the female sample was 37.72 years and in males it was 36.59 years (R2 = 0.09 in males and 0.01 in females).

4.3 Discussion

Quantitative Real Time PCR is easy and robust to work with and gives reproducible results. When the entire sample was analysed, there was a negative correlation between telomere length and age ($r = 0.185$, $P<0.05$), indicating that telomere length shortens with age (Figure 4.3). However, the results showed that telomere length measurement using this method cannot be used to accurately predict the age of a person in forensic investigations, since the age prediction value is too low (R^2 approximately 1 to 9% depending on gender). Not much improvement was seen when a sub-sample of unrelated European individuals was analysed separately (Figure 4.4), which contained a mixture of male and female samples, where the predictability was ~ 5.5%. The most likely explanation for all of these observations is that there was a large amount of inter-individual variation in telomere length in the population as a whole.

The real-time PCR approach was chosen instead of the TRF method, since a method based on PCR requires a lesser amount of DNA and can be completed in a smaller amount of time. In addition, it was hoped that by using a method that targets only the telomeric region and does not include the sub-telomeric region as the TRF method does (Cawthon 2002), a better correlation may be seen between telomere length and age. The Pearson correlation coefficient of $r = - 0.185$ in the general population was comparable in value to those reported by Nordfjall *et al.,* (Nordfjäll *et al.* 2005) and Njajou *et al.,* (Njajou *et al.* 2007) ($r = - 0.356$ and $r = - 0.40$), where the studies were also conducted using quantitative real time PCR to measure telomere lengths. On the other hand, in studies conducted by Tsuji *et al.* (Tsuji *et al.* 2002) and Takasaki *et al.* (Takasaki *et al.* 2003), where the TRF method was used to estimate telomere lengths,

the Pearson correlation coefficients were relatively high (r = 0.832 and r = 0.749 respectively) (Takasaki *et al.* 2003; Tsuji *et al.* 2002). Thus, at first it may appear that the real-time quantitative PCR method was less efficient at determining correlation between telomere length and age. However, this is not necessarily the case.

Firstly, it may, in fact, be a more accurate representation of the biological situation as sub-telomeric regions are not included in the analysis. Perhaps differences in subtelomeric region lengths may be contributing to results. Although this is not likely, given that there have not been any reported decreases in subtelomere lengths with age, it cannot be ruled out as a possibility. Recently, it was found that a decrease in methylation in subtelomeric regions is associated with shorter telomere lengths (Maeda *et al.* 2009; Yehezkel *et al.* 2008). As the TRF method relies on the activity of restriction enzymes on subtelomeric regions to extract telomere-containing sections of DNA, the level of methylation in these regions may affect enzymes that are sensitive to methylation. For example, hypomethylated subtelomeric regions may have more restriction sites available close to the telomeric regions, and hence produce smaller fragments. Thus it may not be telomere lengths alone that are reflected in TRF analysis.

Secondly, the studies of Nordfjall *et al.,* (Nordfjäll *et al.* 2005), Njajou *et al.,* (Njajou *et al.* 2007), Tsuji *et al.* (Tsuji *et al.* 2002) and Takasaki *et al.* (Takasaki *et al.* 2003) were conducted independently and were not related. Thus, the differences in the degree of correlation between telomere length and age seen among the studies could also have been due to differences in ethnicity, sample size, tissue type, laboratory conditions and many other possible differences.

One of the factors could be the ethnically mixed sample used in this study, compared with the ethnically harmonious Japanese samples used in Tsuji *et al.,* and Takasaki *et al.,* studies (Takasaki *et al.* 2003; Tsuji *et al.* 2002). The individuals used in this study included individuals from European, South Asian and Oriental Asian, African, Native American, and Middle Eastern backgrounds. Thus there could be expected to be more heterogeneity in telomere lengths in these samples.

In order to minimise the effects of ethnicity and inheritance, a sub-sample of data from unrelated European individuals (n = 60) were selected from this sample for separate analysis. The r value only increased to 0.233 (Figure 4.4), much different to the values reported by Tsuji *et al.,* (Tsuji *et al.* 2002) and Takasaki *et al.,* (Takasaki *et al.* 2003) One possible explanation for this could be that, historically, the Japanese population was much more isolated than populations in Europe (Chesnasis 1992), and this could have resulted in a much more homogeneous population (and hence telomere lengths) than the Europeans. However, all this is purely speculative at this stage and many more studies, using ethnically homogeneous versus ethnically mixed populations would need to be carried out to confirm these hypotheses.

Although there was significant difference in lengths between 'young' (age 1-39 years) and 'old' (40- 89 years) ($p < 0.05$), some of the older individuals had longer telomeres than the younger ones (Figure 4.3). The shortest relative telomere length was observed in a 26 year old female (0.54) and the longest relative telomere length in a 27 year old male (2.83) in this study. The age gap between the two individuals being

just one year, this clearly illustrates the high level of inter-individual variation in telomere lengths.

In previous studies, a high level of telomere length variation was observed in newborns (Okuda *et al.* 2002), foetus (Youngren *et al.* 1998), and zygotes (Baird *et al.* 2003). These telomere length differences between individuals appear to be maintained during life. In fact, telomere length variation among newborns is similar to variation observed later in life (Okuda *et al.* 2002). In our study, relative telomere length varied among similarly aged individuals. Telomere lengths of 10 individuals aged 26 years old showed a relative telomere length range of 0.54 to 2.43 (Table 4.3). Similar observations were made in 9 individuals aged 54, with a telomere range of 0.68 to 2.11 (Table 4.3). The variation in telomere lengths is higher in the younger group (aged 26) compared to the older group (aged 54). In another study, nine individuals aged 35 years had a telomere range of 8.86-17.29kb (Baird *et al.* 2005), indicating that variation of telomere length exists even among similarly aged individuals.

Telomere lengths in the eight age groups (Figure 4.5) analysed showed that mean telomere lengths of age groups 0-9 years, 10-19 years and 20-29 years were similar, indicated by a telomere length plateau in the early stages of life. A decline of mean telomere length was observed from age group of 30-39 years to the age group of 50-59 years followed by a lesser decline thereafter. Telomere length has been reported to be maintained differentially during different stages of life (Nordfjäll *et al.* 2005). A bi-segmental analysis of telomere length versus age was proposed by Rufer *et al.*, (Rufer *et al.* 1999). This increased the statistical significance of the linear regression they performed. In our study, there were two clear phases, perhaps three, of mean

telomere length declining with age (Figure 4.5). However, the age range of individuals in the two phases observed by the reported studies (Frenck *et al.* 1998; Rufer *et al.* 1999) was very different to ours. They observed a rapid decline of telomere length in early childhood from newborns to 4 years old (Frenck *et al.* 1998) and from newborns to 18 months old (Rufer *et al.* 1999). Thus comparison cannot be made between the studies. As our sample had very few specimens from individuals under the age of 4 (n = 2), we cannot comment on the rate of telomere length decline in this very young age group.

Males appear to lose telomere length more rapidly than females (Figure 4.6). Similarly, in another study of peripheral mononuclear cells, males lost telomere length 25 bp per year compared to a 16 bp per year loss in women (Nordfjäll *et al.* 2005). Females usually live longer than men (Perls *et al.* 1998) and this could be due to differences in the rate of their telomere loss and in their telomere length regulation. However, to test this hypothesis accurately, one would have to conduct a longitudinal study, following the same group of males and females over the course of their lifetime and measuring their relative telomere lengths at intervals.

There has also been interest in understanding the genes, and the genetic control involved in telomere length regulation. It has been shown that telomere length and chromosome specific telomere length patterns are inherited (Graakjaer *et al.* 2004; Njajou *et al.* 2007; Rufer *et al.* 1999; Slagboom *et al.* 1994). The inheritance values vary according to the sample population analysed: for instance 78% inheritance was observed in twins (Slagboom *et al.* 1994) and 44% was observed in the Amish population (Njajou *et al.* 2007). Two hypotheses have been put forward on the mode

of inheritance of telomere length. According to Nowrot *et al.*, telomere length inheritance is linked to the X chromosome, (Nawrot *et al.* 2004) whereas, according to Nordfjall *et al.,* and Njajou *et al.*, telomere length is paternally inherited (Njajou *et al.* 2007; Nordfjäll *et al.* 2005) .

The father's age at conception (Unryn *et al.* 2005) and father's life span (Njajou *et al.* 2007) are also reported to have an influence on an individual's telomere length. The next chapter is focussed on analysing the effect of inheritance of individuals' telomere length. In both the TRF method and real time Quantitative PCR, telomere length is measured as a representative value for telomere length of all chromosome ends (humans have 23 pairs of chromosomes). Measurement of the telomere length of a single chromosome may reduce the variation observed. While Quantitative Fluorescence in Situ Hybridization (Q-FISH) method (Slijepcevic 2004) would not suit general forensic applications, as metaphase stage samples are required, a method called single telomere length analysis (STELA), developed and first reported by Baird, *et al.,* (Baird *et al.* 2003) may be worthwhile investigating. If a good correlation of telomere length to age is found by using any of these methods, the next step would be to consider the impact of DNA degradation on crime scene samples and its effect on telomere length measurement.

However, in addition to the factors that were investigated, i.e. age, gender and heritability, there could be other factors that influence telomere lengths. For example, there could have been variation in the activity of telomerase and other telomere proteins in germ line cells of the parents of individuals. In addition, other factors could affect telomere length, such as oxidative damage. Lifestyles e.g. smoking and

drinking could elevate oxidative damage, dietary differences could affect the level of anti-oxidants, and there could be differences in rates of repair of oxidative damage. Therefore direct measurement of telomere length alone may not be sufficient to predict the age of a human. In the future, longitudinal studies of telomere lengths and factors that affect them, conducted by following individuals at different stages of life, may assist our understanding of what regulates telomere length and give a clearer picture of the association between telomere lengths and age. Similarly, multi-generation studies in populations matched for ethnicity, lifestyle, and gender may shed light on the mode of inheritance of telomere lengths. Such studies may be difficult to conduct in human populations, and the use of shorter-lived animal models such as mice may help provide some answers.

4.4 Conclusions

Telomere length measurement by real time quantitative PCR is not a good predictor of an individual's age. The presence of large inter-individual variation in telomere length prevents developing a formula to predict age from DNA with accuracy. Telomere length differences at birth (Ohki *et al.* 2001), telomere attrition differences between males and females (Baird *et al.* 2005), non-linear telomere length loss during different stages of life (Rufer *et al.* 1999), genetic factors and environmental factors give rise to the large amount of variation of telomere lengths present in the population.

Chapter 5

Influence of inheritance on telomere length

5.1 Introduction

As discussed in the preceding Chapter 4, Telomere lengths of 167 individuals aged 1-96 years old were measured using quantitative real time PCR. The aim was to develop a formula to predict the age of a person by analysing telomere length in the general population. Although, relative telomere length was found to decreases with age ($r = -0.185$, $P < 0.05$), the age prediction value was only 4% in the general population (1% in females, 9% in males), and it is too low for use in forensic purposes. The presence of large inter individual variation in telomere length prevents developing a formula to predict age from DNA with accuracy. This may be due to the fact that telomere length is inherited with varying degrees of paternal and maternal influences.

As stated in Section 2.7, telomere length and chromosome specific telomere length patterns have been reported to be inherited (Graakjaer *et al.* 2004; Njajou *et al.* 2007; Rufer *et al.* 1999; Slagboom *et al.* 1994), with the percentage of inheritance varying according to the sample population analysed: for instance 78% inheritance observed in twins (Slagboom *et al.* 1994) and 44% observed in Amish population (Njajou *et al.* 2007). The differences in inheritance values could be due to the different study population, age ranges, and pedigree structures (Njajou *et al.* 2007). The mode of inheritance of telomere length is not yet clear, and two hypotheses have been put forward on this (Nawrot *et al.* 2004; Njajou *et al.* 2007; Nordfjäll *et al.* 2005). According to Nowrot *et al.*, telomere length inheritance is linked to the X chromosome (Nawrot *et al.* 2004), whereas Nordfjall *et al.* and Njajou *et al.* have

suggested that telomere length is paternally inherited (Njajou *et al.* 2007; Nordfjäll *et al.* 2005).

Studies that have been carried out to determine the mode of inheritance examined the correlation between parental telomere length and children's telomere length (Table 5.1) (Nawrot *et al.* 2004; Njajou *et al.* 2007; Nordfjäll *et al.* 2005). In the Nowrot *et al.*(Nawrot *et al.* 2004) Study, the presence of statistically significant high correlation of telomere lengths between mother and children, and father and daughter, but notably not with father and son, helped to construct the hypothesis that the X chromosome plays a major role in telomere length determination. It was suggested that the DKC1 gene present on the X chromosome is the factor behind X linked inheritance. The protein dyskerin coded by the DKC1, aids in stable accumulation of the hTR component of telomerase. Missense mutations in the DKC1 gene cause a premature ageing syndrome called dyskeratosis congenita (Reviewed in (Collins *et al.* 2002)). Polymorphisms present in the DKC1 gene possibly regulate the telomere length of individuals by regulating the expression of telomerase.

Contrary to observations made by Nowrot *et al* (Nawrot *et al.* 2004), both Nordfjall *et al.* and Njajou *et al.* observed high correlation of telomere lengths between father and children but not between mother and children (Table 5.1), concluding that telomere length is paternally inherited (Njajou *et al.* 2007; Nordfjäll *et al.* 2005). They failed to observe X-linked inheritance, since the correlation of telomere lengths between mothers and children was low.

Table 5.1 – The correlation values of parent –children telomere lengths in 3 different studies.

Mode of inheritance	Parent children combinations	Correlation (r)	Significance (p)
X – linked inheritance Nowrot *et al* (Nawrot *et al.* 2004)	Father – daughter	0.62	<0.0001
	Father- son	0.09	0.63
	Mother-daughter	0.59	<0.0001
	Mother-son	0.34	0.013
Paternal inheritance Nordfjal *et al* (Nordfjäll *et al.* 2005)	Father – daughter	0.578	0.005
	Father- son	0.558	0.011
	Mother-daughter	0.161	0.463
	Mother-son	0.141	0.577
Paternal inheritance Njajou *et al* (Njajou *et al.* 2007)	Father – daughter	0.43	<0.001
	Father- son	0.56	<0.001
	Mother-daughter	0.18	0.05
	Mother-son	0.21	0.05

The influence of parental age at conception/birth on the telomere length of an individual has been studied (Meyer *et al.* 2007; Njajou *et al.* 2007; Unryn *et al.* 2005). Unryn *et al.* (Unryn *et al.* 2005) and Meyer *et al.* (Meyer *et al.* 2007) discovered that father's age at conception/birth had a statistically significant correlation with children's telomere length. This could be explained by the fact that telomere length increases with age in sperm cells due to the activity of the enzyme telomerase (Allsopp *et al.* 1992). The positive correlation between father's age at conception with individual's telomere length meant that the older the father's age at conception the longer the telomere lengths of the children (Unryn *et al.* 2005). In a study done on an Amish population, Njajou *et al.* found a positive borderline significance of both the parents age at conception with children's telomere length (Njajou *et al.* 2007).

The present study was also aimed at assessing the influence of inheritance of telomere length and the effect of parents' age at conception on the telomere length of an individual. From the total sample of 167 individuals, telomere lengths of 18 families (n = 79), which had more than 3 members per family, were used for this study. From the sample population, telomere lengths of father-daughter (n = 10), father-son (n = 11), mother – daughter (n = 23) and mother- son (n = 16) pairs were compared to assess the mode of inheritance of telomere length. Telomere lengths of the sample were measured by real-time quantitative PCR (Cawthon 2002), as described in section 2.4.1. If there is strong maternal and/or paternal influence on the length of telomeres of an individual, the effect of these factors could be incorporated into the age prediction formula. The more variables included into the age prediction formula the more its accuracy will be increased.

5.2 Results

5.2.1 Relationship between telomere length and age in families

The correlation and regression analysis of telomere length and age was carried out within each family, in individuals belonging to 18 families (Table 4.2 and figure 4.1). Only family # 6 was found to have a statistically significant correlation between age and telomere length ($r = 0.99$, $p<0.05$), and this was the opposite of what was expected, which would have been a negative r value. The rest of the families' ages had minimal correlation with telomere length. Although the R^2 values were high, the P values were > 0.05, thus the values had low confidence levels.

Table 5.2 – Correlation and regression between telomere length and age in 18 families.

Family	Pearson Correlation (r)	Significance (p)	R Square	Std. Error of the Estimate
1	-0.796	0.413	0.739	0.777
2	-0.507	0.493	0.538	1.17
3	-0.638	0.362	0.238	1.198
4	-0.341	0.508	0.716	1.026
5	0.928	0.072	0.936	0.432
6	0.999	0.023	0.979	0.172
7	-0.169	0.831	0.816	0.524
8	-0.098	0.902	0.875	0.402
9	0.671	0.145	0.894	0.401
10	0.044	0.972	0.915	0.342
11	-0.360	0.640	0.714	0.697
12	-0.789	0.421	0.648	1.01
13	-0.179	0.773	0.660	0.827
14	-0.113	0.928	0.645	0.728
15	0.040	0.960	0.780	0.600
16	-0.434	0.138	0.568	0.707
17	-0.839	0.366	0.534	0.887
18	0.928	0.242	0.947	0.373

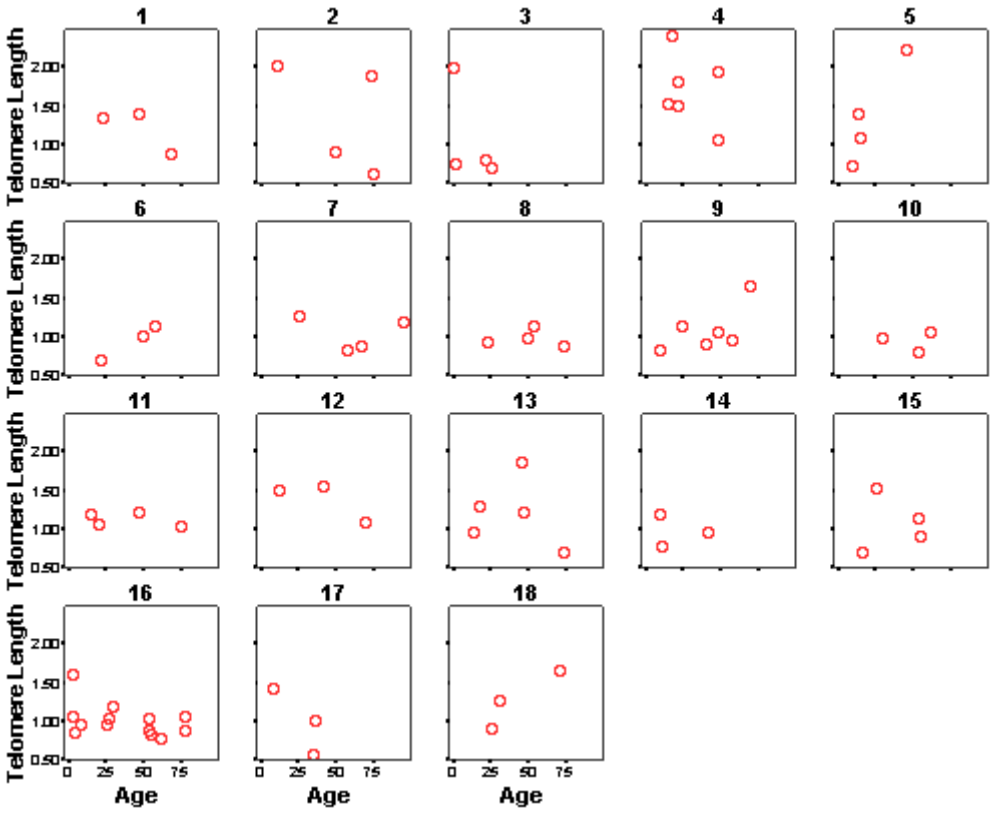


Figure 5.1 – Relationship between telomere length and age in 18 families**.**

5.2.2 Telomere length differences between families

Telomere length between families was compared to check whether any family had relatively 'longer' or 'shorter' telomere lengths. However, analysis of variance (ANOVA) showed that there was no significant difference in telomere lengths between the families ($P > 0.05$) (Table 5.3), confirming that no particular family had rather 'longer' or 'shorter' telomere lengths.

Table 5.3 – ANOVA of telomere lengths between 18 families.

	Sum of Squares	df	Mean Square	F	Significance
Between Groups	3.34	22	0.196	1.28066	0.2182
Within Groups	8.38	66	0.137		
Total	11.719	88			

5.2.3 Influence of parental telomere length on children's telomere length

The influence of parents' telomere lengths on children's telomere lengths was evaluated. Age adjusted partial correlation between parents' telomere length and children's telomere length was carried out to minimise the effect of age on telomere lengths in both parents and children. The regression analysis was also carried out between the parents' and children's telomere lengths.

5.2.3.1 Influence of fathers' telomere length on children's telomere length

Telomere lengths of father-daughter (n = 10) and father-son (n = 11) were compared with each other. The age adjusted partial correlation was not statistically significant ($p>0.05$) in either case; the correlation values been father-daughter ($r = 0.346$, $p = 0.402$) and father-son ($r = 0.525$, $p = 0.147$) respectively. The regression values between father-daughter ($R^2 = 0.08$) and father-son ($R^2 = 0.07$) had similar values (Figures 5.2 and 5.3).

5.2.3.2 Influence of mothers' telomere length on children's telomere length

Telomere lengths of mother-daughter (n = 23), and mother-son (n = 16) pairs were analysed. There was also no significant age adjusted partial correlation between either mother-daughter ($r = 0.149$, $p = 0.520$) or mother-son telomere lengths ($r = -0.001$, $p = 0.996$). The regression values obtained indicate that there was minimal effect of the mothers' telomere length on daughters' telomere length ($R^2 = 0.005$) and sons' telomere length ($R^2 = 0.0046$) (Figures 5.4 and 5.5).

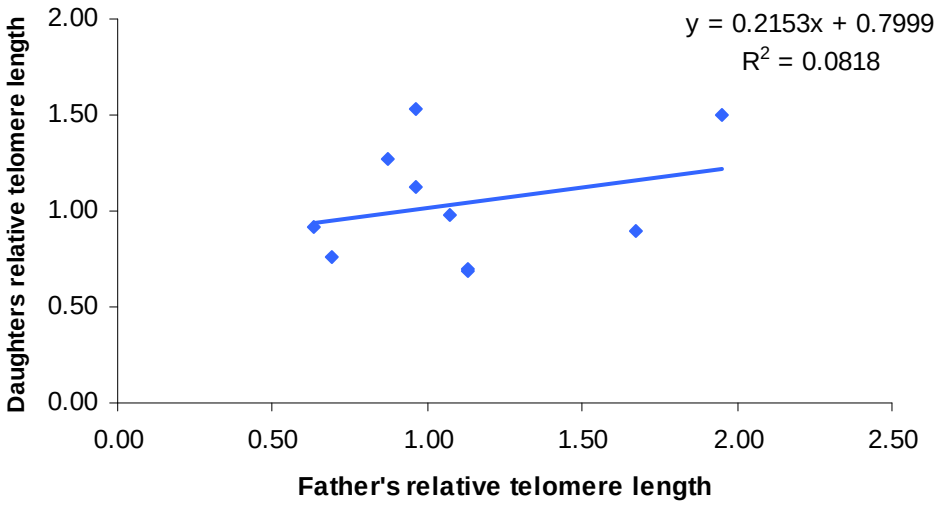


Figure 5.2 – The relationship between fathers' telomere length and daughters' telomere length (r = 0.346, p= 0.402).

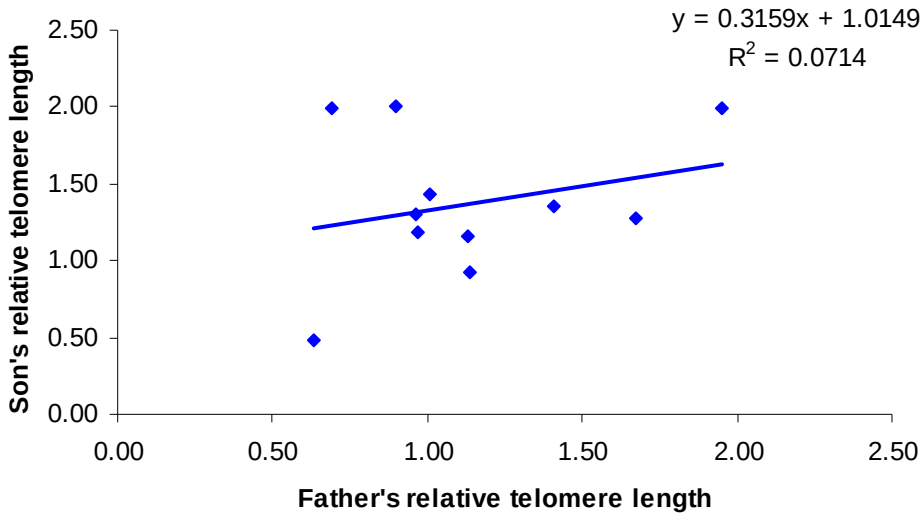


Figure 5.3 –The relationship between fathers' telomere length and sons' telomere length (r = 0.525, p = 0.147).

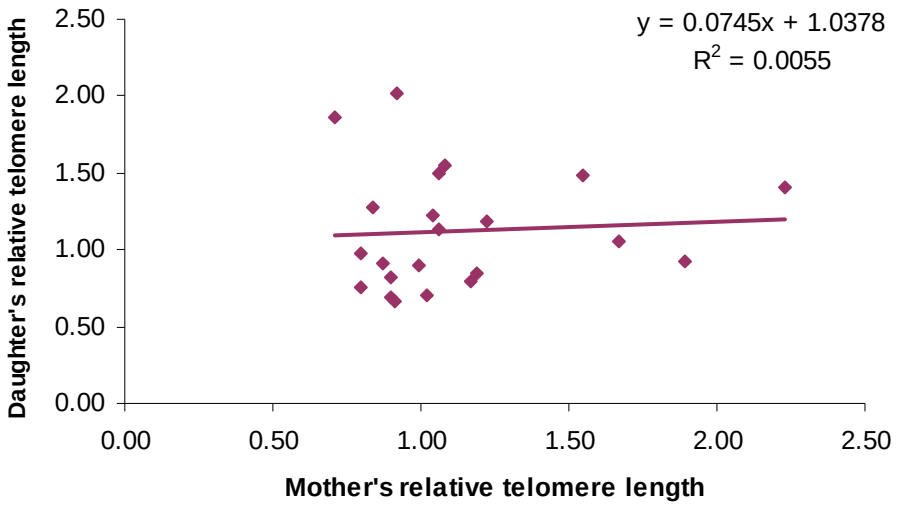


Figure 5.4 – The relationship between mothers' telomere length and daughters' telomere length (r = 0.149, p = 0.520).

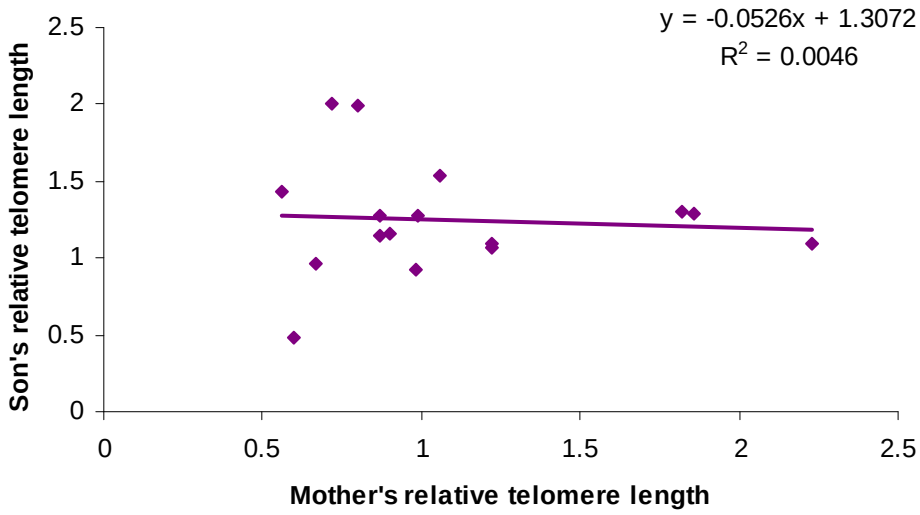


Figure 5.5 – The relationship between mothers' telomere length and sons' telomere length (r = -0.001, p = 0.996).

5.2.4 Influence of parents' age at conception on telomere lengths of children

The fathers' and mother's age at conception (Table 5.3) respectively were compared with their children's telomere length. Neither the fathers' nor mothers' age at conception had a statistically significant correlation with children's telomere lengths ($r = 0.064$, $p = 0.506$), ($r = 0.075$, $p = 0.414$) respectively. However, within the family samples collected for this study, there was statistically significant negative correlation between both fathers' and mother's age at conception with children's age ($r = -.203$, $p = 0.034$), ($r = -0.282$, $p = 0.002$) respectively. This meant that younger individuals had older parents compared to their parents' generation.

Table 5.4 - The age distribution of parents' age at conception and children's age.

	Minimum (years)	Maximum (years)	Mean (years)	Standard deviation
Fathers' age at conception	19	46	29.42	5.43
Mothers' age at conception	18	43	27.59	6.24

5.2.4.1 Influence of parental age at conception on telomere lengths of 26 year olds

Telomere lengths of individuals aged 26 years old ($n = 10$) were compared with fathers' age at conception, to exclude the effect of children's age. There was no significant correlation between fathers' and mothers' age at conception ($r = -0.015$, $p = 0.967$), ($r = 0.187$, $p = 0.607$) respectively on the telomere lengths of individuals aged 26 years (Table 5.4).

Table 5.5 - The age distribution of parents' age at conception of 10 individuals' age 26 years.

	Minimum (years)	Maximum (years)	Mean (years)	Standard deviation
Fathers' age at conception	22	46	30.5	7.32
Mothers' age at conception	20	43	29.60	7.09

5.3 Discussion

In all except one of the 18 families analysed (#6), telomere length had no statistically significant correlation with age. The family # 6 had a positive correlation between telomere length and age. This observation could be due to many factors that regulate telomere length or just simply due to a coincidental anomaly. From the 18 families analysed, no family had either particularly long or short telomere lengths. The variation of telomere length present within a family mirrors the variation present in the population.

Telomere length is regulated by mechanisms that control the telomere replication, elongation and telomerase activity (Saldanha *et al.* 2003). The genes and regulatory pathways that control the telomere length should be investigated. For example, mutations found in genes TERT (encoding hTERT component of telomerase), and TERC (encoding hTR component of telomerase), were responsible for diseases with short telomeres and premature ageing (reviewed in (Garcia *et al.* 2007)). An additional 32 proteins have also been discovered to assist in the function of telomerase (Cohen *et al.* 2007). Polymorphisms that are present in genes that code for telomere regulatory proteins could explain the amount of variances observed within a family.

To evaluate the mode of inheritance of telomere length, the correlation between telomere lengths of father-son, father-daughter, mother-son and mother-daughter combinations were compared with each other. The mothers' telomere length had no statistically significant correlation with sons' or daughters' telomere length, and the regression values were also close to zero (Figures 5.4 and 5.5). The X-linked inheritance of telomere length could not be established. As discussed in chapter 4, measurement of telomere length in an animal model, such as mice, in different stages of life, may be useful in understanding telomere length regulation. Additionally, analysing single nucleotide polymorphisms (SNPs) and mRNA expression levels of major telomere length regulating proteins such as TERT, TERC, TRF1 and TRF2 in a sample of mice, as a longitudinal study, would give an idea how the genes and gene expression levels change with age.

Age adjusted partial correlation of telomere lengths between father-daughter ($r = 0.346$) and father-son ($r = 0.525$) were in agreement with those reported in previous studies of Nordfjall *et al.* (Nordfjäll *et al.* 2005) of father-daughter ($r = 0.578$) and father-son ($r = 0.558$) and of Njajou *et al.* (Njajou *et al.* 2007) of father-daughter ($r = 0.43$) and father-son ($r = 0.56$). However, the correlation values obtained were not statistically significant ($P > 0.05$). Therefore it cannot be concluded that telomere length is paternally inherited in the sample population in this study. The sample size was smaller compared to other research studies, making it difficult to obtain statistically significant values. The aim was to get more pairs, at least 20 each, but time constraints and the difficulty in obtaining samples limited the sample size. Failure to establish the mode of inheritance of telomere length also could be due to the ethnically mixed sample used in this study, compared with the ethnically homogenous

European samples from northern Belgium in Nawrot *et al.* (Nawrot *et al.* 2004), northern Sweden in Nordfjäll *et al.* (Nordfjäll *et al.* 2005) and Old order Amish in Njajou *et al.* (Njajou *et al.* 2007). The sample used in this study included individuals from European, South Asian and Oriental Asian backgrounds.

A possible explanation for the previously reported correlation of telomere lengths of children with either the father or the mother could be due to genomic imprinting (Nordfjäll *et al.* 2005). Genomic imprinting is mono-allelic, from the two alleles of a gene, where one allele is expressed while the other one is not expressed and remains silent. The allele that is imprinted could come from either of the parents (Murphy, SK *et al.* 2003). According to Mendel's law of inheritance, the phenotypic outcome of a gene is not related to the gender of the parent. Interestingly due to genomic imprinting, gene expression becomes dependent on the gender of the parent (Murphy, SK *et al.* 2003). So far, telomere length regulating genes are not yet known to be imprinted (Nordfjäll *et al.* 2005), but involvement of genomic imprinting cannot be ruled out.

The father's age at conception and mother's age at conception had no statistically significant correlation with children's telomere length. Interestingly there was a negative correlation between parents' age at conception with children's age in the sample. In other words, older aged individuals had parents who decided to have children at a younger age. This could be explained by cultural changes, since many women and men currently have children when they are much older, when compared with the generation of their parents. To remove the effect of children's age, the

telomere size of a subset of similarly aged individuals of 26 years old were compared to those of their parents, but no statistically significant correlation was found.

Unryn *et al.* (Unryn *et al.* 2005) and Meyer *et al.* (Meyer *et al.* 2007) reported that the father's age at conception/birth had a positive correlation [($r = 0.46$, $P = 0.01$) ($r = 0.127$, $P < 0.001$) respectively] with children's telomere length (Unryn *et al.* 2005). The positive correlation indicates that the older the father's age at conception the longer the children's telomere length. This observation was suggested to be due to the increase of telomere length during male germ line differentiation into spermatozoa (Achi *et al.* 2000). Interestingly, Njajou *et al.* observed the borderline correlation of both parents' age at conception with their children's telomere length (Njajou *et al.* 2007). Unryn *et al.* (Unryn *et al.* 2005) also observed, even though it is not statistically significant, a trend that children with older mothers had longer telomeres. Unlike in the case of male germ cells, there has been no study done on telomere length in the female germ line, even though the presence of telomerase activity was reported in most female reproductive tissues (Unryn *et al.* 2005). It could be that in both germ line cells telomere length increases with age, and thus influences children's telomere length.

5.4 Conclusions

From this study, the mode of inheritance of telomere lengths could not be established mainly due to lack of samples. To assess the mode of inheritance of telomere length, it would be better to analyse multiple generational families and make comparisons within and between generations. A larger sample size would be required to obtain statistically significant results. As stated before, telomere length is regulated by

different proteins and these need to be identified and analysed. Genomic imprinting and silencing of these genes would also be useful to study. Understanding and identifying factors that determine the telomere length of individuals would assist in developing a formula to predict age. However, it appears that the factors are many and complex; many studies would be required to understand the telomere length regulation.

The subsequent two chapters will discuss mitochondrial mutations and deletions and their relationship with age in the maternal lineage.

Chapter 6

Evaluation of mitochondrial DNA mutations to predict age

6.1 Introduction

Human mitochondria contain multiple copies of a 16,569 bp closed circular deoxyribonucleic acid (DNA) genome that is replicated and expressed within the organelle system (Larsson *et al.* 1995). The expression of the whole genome is essential to maintain cellular bioenergetic capacity for a cell to survive (Ozawa 1997). Approximately 90% of the cellular oxygen is consumed in mitochondria, and 2% of that oxygen is converted into superoxide free radicals (Lee *et al.* 1997). In 1956 Harman proposed the "free radical theory of ageing", that free radicals cause non-specific damage to DNA, lipids, and proteins (reviewed in (Ozawa 1997)). In 1972, Harman postulated that free radical damage to DNA plays a role in the ageing process (reviewed in (Lee *et al.* 1997)). The theory that mitochondrial mutations play a role in ageing was put forward by Linnane *et al*. (Linnane, AW *et al.* 1989). They suggested that somatic accumulation of mitochondrial mutations caused by free radicals and the subsequent cytoplasmic segregation of these mutations during life is a major contributor to the ageing process.

Mitochondria are maternally inherited (Giles, R *et al.* 1980). Only the mother passes mitochondria on to her children. Therefore individual mtDNA type is not unique: maternal relatives and siblings have essentially the same mitochondrial sequence (Butler 2005) and this can be used to determine the maternal ancestry of an individual (Butler, J *et al.* 1998). Due to maternal inheritance and the high copy number, mitochondrial sequence variations have been used in studies on evolution, migration patterns, and individual identification cases (Butler 2005). However, this high copy

number poses a problem when analysing mitochondrial DNA mutations. Unlike in nuclear DNA, there is more than one mtDNA sequence within an individual. This phenomenon is known as heteroplasmy (mixture of wild type and mutant type mtDNA in one individual). If all the mitochondrial sequences are the same it is called homoplasmy (Lodish *et al.* 2000).

6.2 Mitochondrial mutations that accumulate with age

A large number of mtDNA mutations have been reported to occur in human somatic tissues, many of which accumulate over time. Increases of heteroplasmy and the highest number of age-associated point mutations are reported in the mitochondrial control region (Section 2.13). This control region is the largest non-coding region and is the most variable portion of the mitochondrial genome (Calloway *et al.* 2000). The control region of mitochondria has two sub-regions each approximately 440bp long – hypervariable region 1 and 2 (HV1 and HV2) (Figure 6.1).

The age relatedness of point mutations in the mitochondrial control region were first reported by Michikawa *et al.* (Michikawa *et al.* 1999) and Wang *et al.*(Wang et al. 2001) who observed that some point mutations are present in high proportions in older individuals and not in younger individuals. The level of heteroplasmy in the mitochondrial control region also increased with age in different tissue types, with the highest level present in muscle and brain tissues compared with buccal cells, blood and heart (Calloway *et al.* 2000). The level of heteroplasmy differed even among maternal relatives, older members having higher levels of heteroplasmy (Theves *et al.* 2006). These findings indicate both point mutations and heteroplasmy increases are associated with age.

6.3 Prediction of age using mitochondrial DNA

The present study was focused on sequencing HV1 and HV2 regions of mitochondria, in 37 individuals from 12 families of three-or four-generational families (maternal lineage). Note that the initial plan was to study more families, but finding volunteers of three or more generational families of maternal lineage willing to give samples proved to be very difficult. Buccal swabs were collected for this study, since it is a non-invasive method of collecting DNA from individuals. The sequences of the youngest members within each family was considered as the reference, and compared with the rest of the family. Since mitochondrial DNA is maternally inherited (Giles, R *et al.* 1980), the youngest member would be more likely to have the 'original' mitochondrial DNA than any other member, and differences found in the others could be considered to be age related. Thus, the sequences of the youngest members within each family was considered as the reference, and compared with the rest of the family.

It was proposed that investigating one single deletion or point mutation would not reflect the extent of mitochondrial mutations that accumulate with age (Kovalenko *et al.* 1998). Therefore, a region of approximately 440bp from each of HV1 and HV2 was amplified, purified and sequenced to detect any deletions, insertions or substitutions. As any mutation may be present together with the original sequence, the sequence data was also assessed for the presence of heteroplasmy. This was achieved with the aid of software called Mutation Surveyor that allows detection down to 3% heteroplasmy.

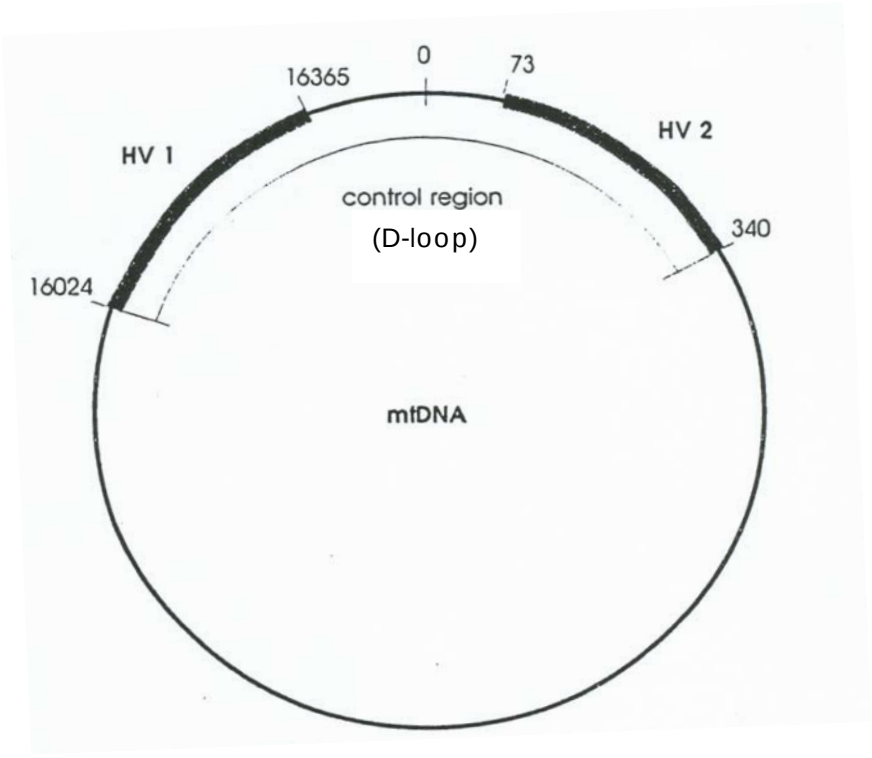


Figure 6.1 – Diagrammatic representations of the mitochondrial control region (D-loop containing the HV1 and HV2 regions. {Reproduced from Gillham (Gillham 1994)}. Note that this region, as shown above is not proportional to the length of whole mitochondrial genome.

6.4 Results

6.4.1 Amplification of HV1 and HV2 regions

PCR amplification was carried out on the HV1 and HV2 regions in all of the 37 samples from 12 families using the same primers and conditions as in Parson *et al.*, (Parson *et al.* 1998). Products of approximately 440bp were amplified (Figures 6.2 and 6.3).

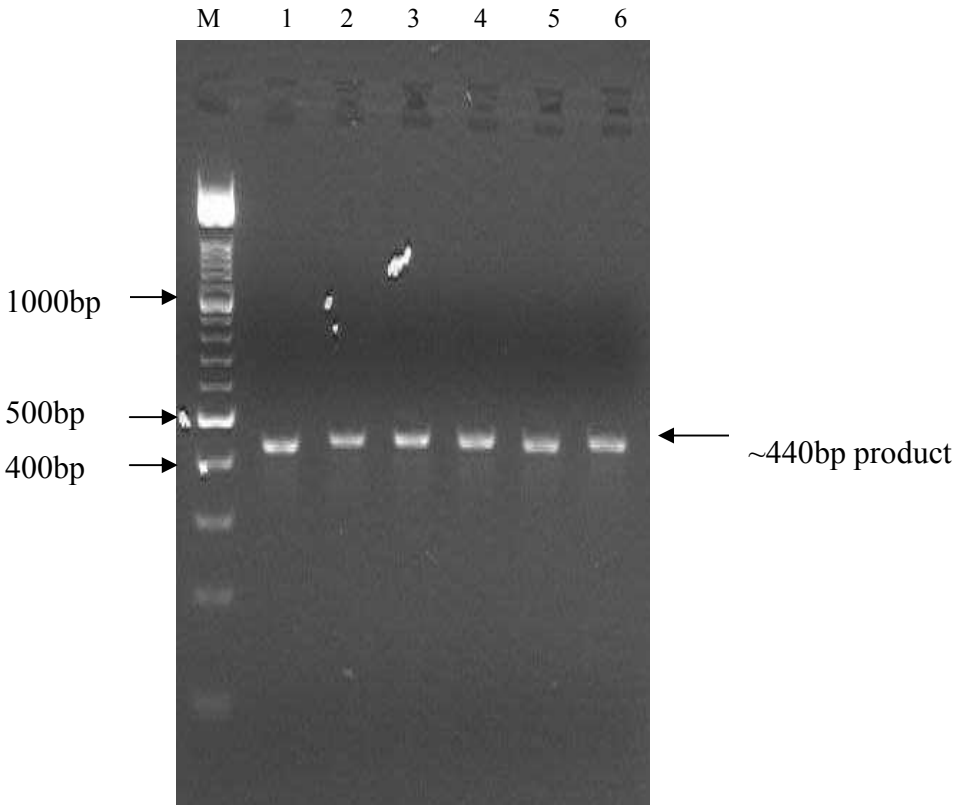


Figure 6.2 - PCR products of the HV1 region in mitochondrial DNA from members of various families. M-100bp molecular weight marker (Roche), Lane 1, 2, 3 – maternal grandmother, mother, mother and daughter of Family 4, Lane 4 – a daughter in Family 3, Lane 5 – a son in Family 1, Lane 6 – maternal grandmother in Family 2.

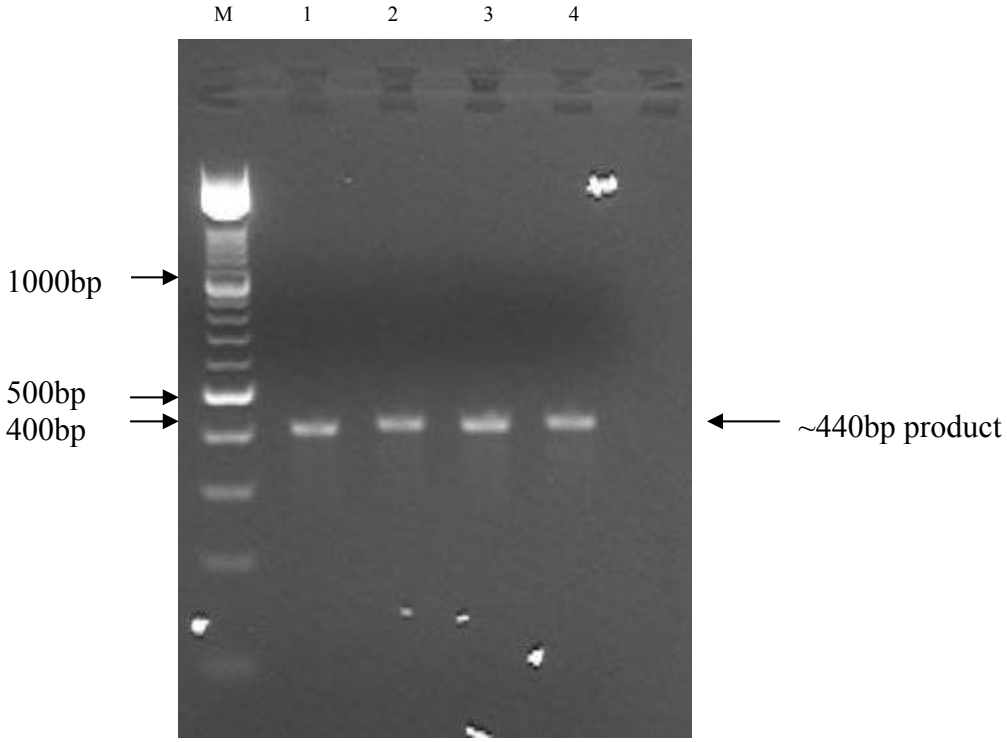


Figure 6.3 - PCR products of the HV2 region in mitochondrial DNA from various families. M-100bp molecular weight marker (Roche), Lane 1 – grandmother of family 4, Lane 2, 3, 4 – maternal grandmother, mother and daughter of family 3.

6.4.2 Sequencing of HV1 and HV2 regions

The amplified PCR products were purified and sequenced in both forward and reverse directions. Both forward and reverse sequences of the HV1 region and the forward sequence in the HV2 region (Figures 6.4a and 6.5a) gave clear sequence chromatograms. However in all samples, the reverse sequence of the HV2 region gave poor sequence chromatograms after the long stretch of G's. Two samples are shown in Figures 6.4b and 6.5b. In the forward sequence this is a homopolymeric tract of cytosines between nucleotides 303 and 315, interrupted by a thymidine at position 310. The presence of the poor sequence, in the reverse direction in HV2 just after a long stretch of Gs, could be due to slippage of DNA polymerase when it reaches a poly base region (Nekhaeva *et al.* 2002).

Usually, both forward and reverse sequences of DNA are sequenced, in order to allow comparison with each other to help eliminate PCR artefacts and sequencing errors. However, as reverse sequencing gave poor results in this case, the HV2 region was sequenced using two forward primers (F29 and F15). The resulting forward sequences could then be confirmed by comparison with each other.

C Stretch from position 303 to 315

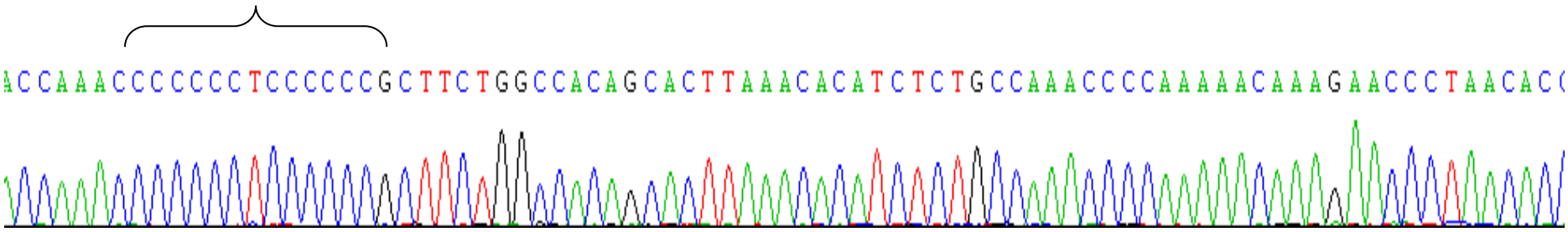


Figure 6.4 a – Part of the HV2 forward sequence chromatogram of daughter of Family # 4.

G stretch (C Stretch forward sequence from position 303 to 315)

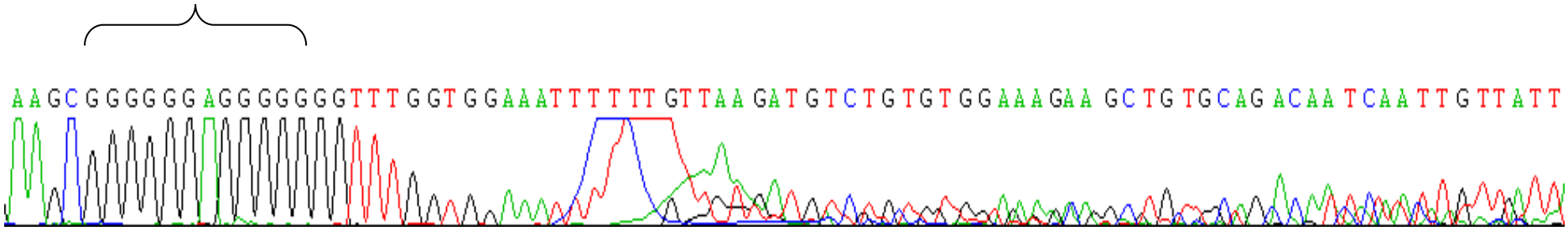


Figure 6.4 b – Part of the HV2 reverse sequence chromatogram of daughter of Family # 4.

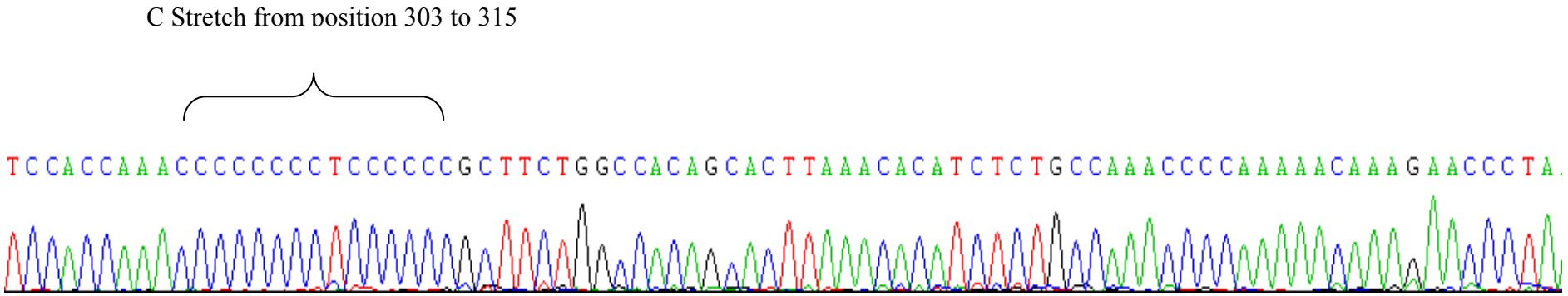


Figure 6.5 a – Part of the HV2 forward sequence chromatogram of daughter of Family # 6.

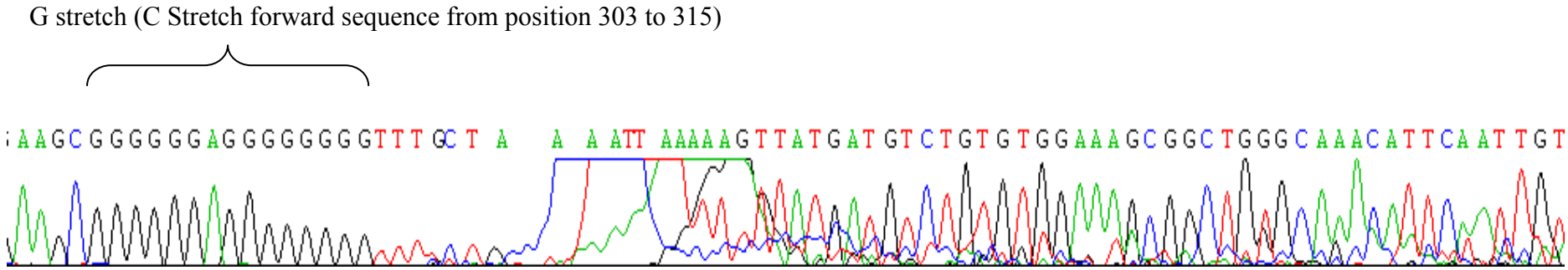


Figure 6.5 b – Part of the HV2 reverse sequence chromatogram of daughter of Family # 6.

6.4.3 Mitochondrial DNA sequence analysis

HV1 and HV2 sequences within a family were compared, using the ClustalW Multiple alignment in BioEdit (Figures 6.6 and 6.7). As mentioned before, the sequence from the youngest family member's sequence was compared with that from the older members of the family. No age related mitochondrial point mutations, deletions or substitutions were detected in any of the 12 families analysed. Examples are shown in Figures 6.6 and 6.7.

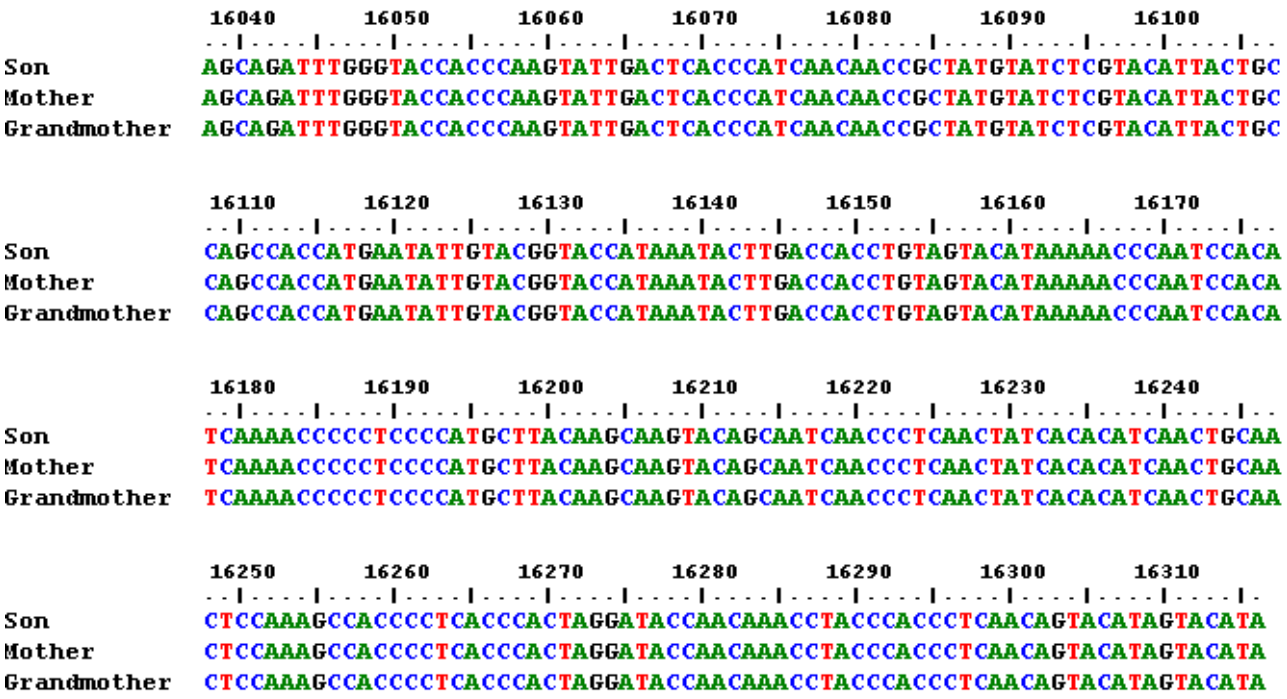


Figure 6.6 - ClustalW analysis of HV1 Region (16024-16365bp) in Family # 1.

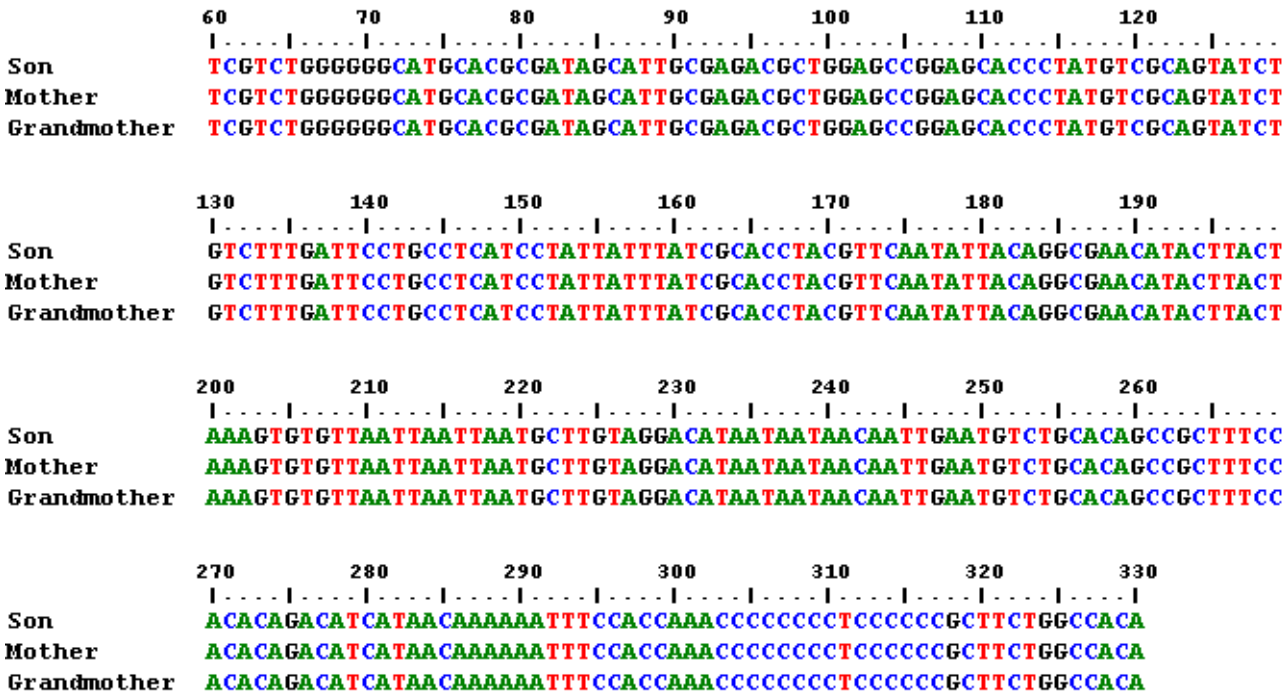


Figure 6.7 - ClustalW analysis of HV2 Region (73-340bp) in Family # 1.

6.4.4 Mitochondrial DNA sequence analysis using the Mutation Surveyor

The sequence chromatograms within a family were analysed using Mutation Surveyor Version 3.1 (Soft Genetics, PA, USA) (Song *et al.* 2005). Mutation surveyor physically compares a sample sequence chromatogram to a reference sequence chromatogram and identifies differences/mutations present. If a possible mutation is detected it is screened for a mutation score. The mutation score is a measure of the probability of error and is based on the ratios of noise level, overlapping factor and dropping factor. The default setting is 500 for the mutation score. If a peak height in a mutation in the mutation electropherogram is greater than 500 and other parameters are satisfied, then it is called a mutation peak (SoftGenetics 2006).

In each of the twelve families we analysed, the sequence chromatogram of the youngest member was used as the reference sequence and was compared with that of the older members of the family. Significant heteroplasmy was not found by this software as seen in the mutation report (Figures 6.8 and 6.9), where possible mutation peaks were not above 500 in both forward and reverse sequences in HV1 and forward sequences in HV2 region. In the mutation report, the left axis indicates mutation peaks. The green centre line in the mutation report indicates there were indels (insertions or deletions) present (Figures 6.8 and 6.9).

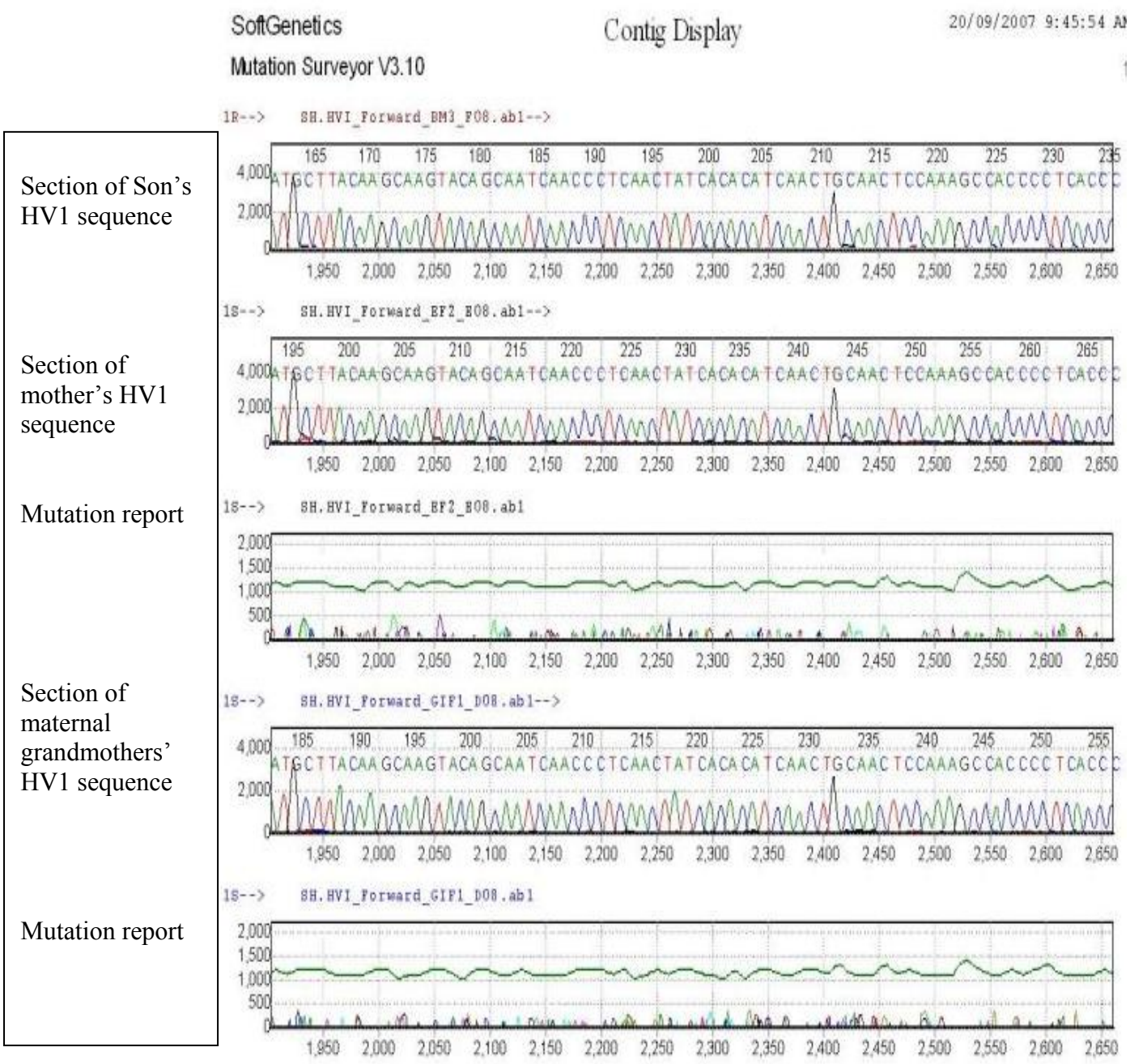


Figure 6.8 – Section of the HV1 region comparisons in Family # 1 by using Mutation surveyor. The HV1 regions of the mother and maternal grandmother were compared to that of the son in the family. The X-axes on the sequence chromatograms represents the frame number and the Y-axes on the sequence represent the relative peak intensity, while numbers in the top part of the chromatogram are the sequencing position numbers. In the Mutation Report, the X-axes represent data points created by the program and Y-axes indicate mutation peak scores. The line in centre of the Y-axes indicates whether indels (insertions or deletions) are present (green = no indel; red = indel).

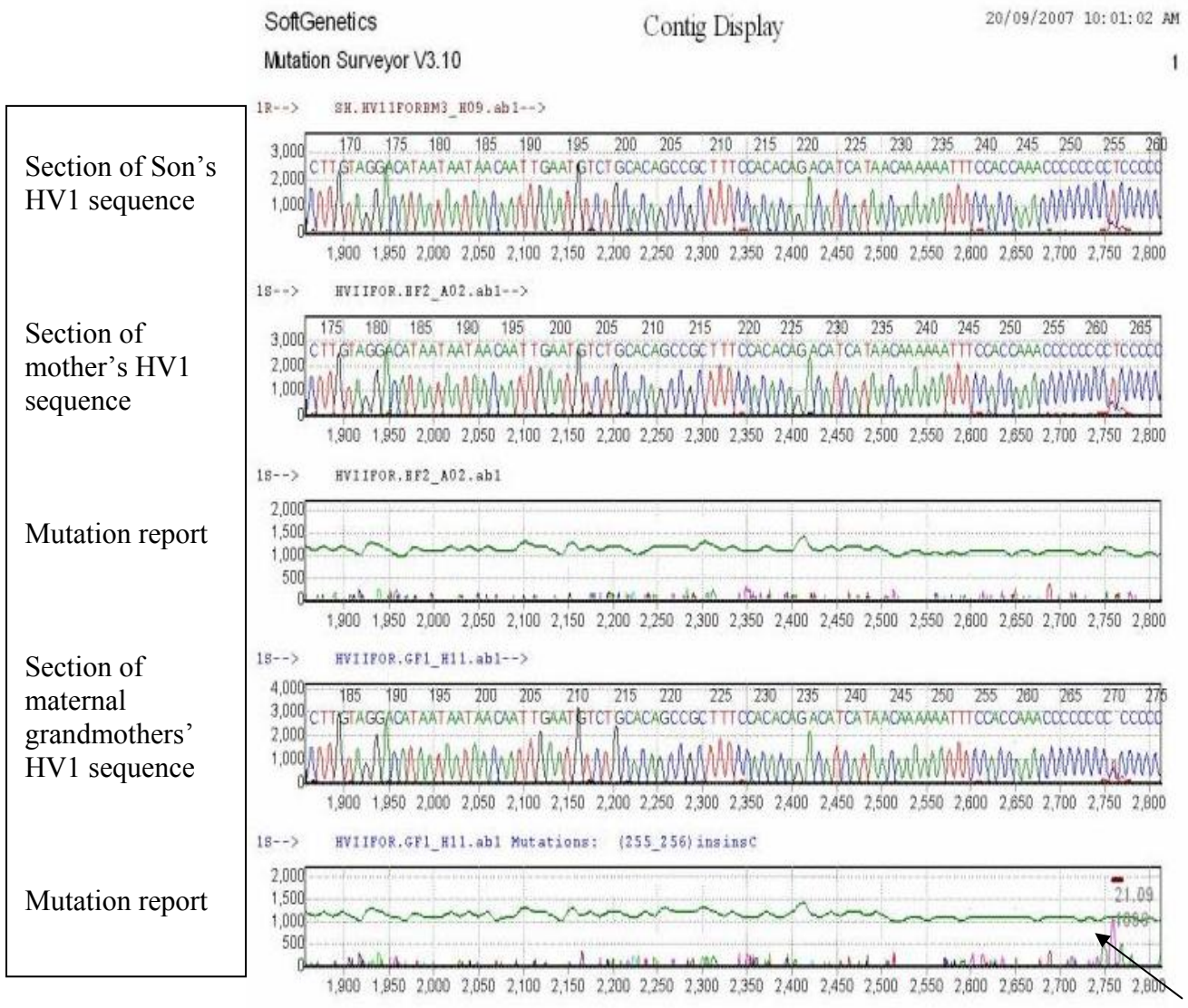


Figure 6.9 - Section of the HV2 region comparisons in Family # 1 by using Mutation surveyor. The HV2 regions of the mother and maternal grandmother were compared to that of the son in the family. The X-axes on the sequence chromatograms represents the frame number and the Y-axes on the sequence represent the relative peak intensity, while numbers in the top part of the chromatogram are the sequencing position numbers. In the Mutation Report, the X-axes represent data points created by the program and Y-axes indicate mutation peak scores. The line in centre of the Y-axes indicates whether indels (insertions or deletions) are present (green = no indel; red = indel). The area indicated by the arrow was not considered as mutation/heteroplasmy since the mutation confidence score (21.09) given by the program was < 30.

6.5 Discussion

In the twelve families analysed, no age-related point mutations were detected in the HV1 and HV2 regions of mitochondrial DNA (Figures 6.6 and 6.7). Previously reported age related mitochondrial mutations were present in some families. A73G was present in four families out of 12 analysed; C150T was observed in Family # 2; T146C was present in Family # 9. However, all of these point mutations were present in every member of the families, including the youngest. Thus, it becomes questionable whether these reported mutations are age-related or rare, in fact, simply found in some families and are maternally inherited.

Even after analysing the sequence chromatograms with Mutation Surveyor (Soft Genetics USA), which is reported to detect heteroplasmy levels down to 3%, no heteroplasmy was detected in any of the families. A similar observation was made by Song *et al.,* (Song *et al.* 2005) in mice. They analysed mtDNA control regions of brain, skeletal muscle, heart, liver, spleen and kidney tissues of 4 older (25-26 months old) mice by analysing the sequences with Mutation Surveyor (Soft Genetics, USA) and they failed to detect age related mutations (Song *et al.* 2005). The Mutation Surveyor (Soft Genetics USA) has been previously used to detect heteroplasmy in mitochondria successfully in breast cancer tissue (Zhu *et al.* 2005) and Leigh's disease patients (Brautbar *et al.* 2008). However, this could have been due to the nature of the tissue they used compared to ours (breast cancer and Leigh's disease muscle tissue vs. our buccal tissue). Elevated level of mtDNA mutations and heteroplasmy have been reported in studies of both these conditions (Tan *et al.* 2005) (Graeber *et al.* 1998).

The study by Theves *et al.,* (Theves *et al.* 2006) compared individuals' mitochondrial sequence with their maternal relative's mitochondrial DNA sequence and was able to detect an age related accumulation of the A189G heteroplasmy in two maternal families out of the

10 examined by the peptide nucleic acid (PNA) directed PCR clamping method. In both families the A189G heteroplasmy was present in higher levels (5-12.6%) in all the older members and absent or lesser (1.5-8.9%) in the youngest individual of the family.

Earlier studies which have reported mitochondrial point mutations that accumulate with age did not compare the DNA sequence of their samples with that of a maternal relative. In these studies the presence of mutations at a higher frequency in the older individuals as compared to the younger individuals led them to the conclusion that the observed mutations were age-related (Del-Bo *et al.* 2003; Michikawa *et al.* 1999; Wang *et al.* 2001).

However, the tissue specificity and age distribution of mitochondrial mutations and heteroplasmy in the mitochondrial control region mutations strongly support the theory that these mutations are somatically accumulated and not inherited mutations (Wang *et al.* 2001). It was observed that age-associated mutations are found frequently in high energy-requiring tissues with slow mitotic activity such as muscle, heart and brain (Theves *et al.* 2006). Buccal cells do not fall under this category.

Thus, detecting point mutations in mitochondrial DNA from buccal tissues may require a method that could detect very low levels of heteroplasmy. Two methods that have been reported to be good at detecting heteroplasmy are the Denaturing Gradient-Gel Electrophoresis (DGGE) and PNA/qPCR method. The DGGE method has been claimed to detect heteroplasmy levels as low as 1%. The study by Theves *et al.,* (Theves et al. 2006) was able to detect the A189G heteroplasmy using the PNA/qPCR combined method. In this method, the PNA molecule is a DNA mimic, in which the negatively charged sugar-phosphate DNA backbone is replaced by an achiral, neutral polymide backbone formed by

repetitive units of *N*-(2-aminoethyl) glycine. The PNA probe hybridises to a target region and prevents PCR primer binding. When the amplification process is monitored by real time PCR it allows a relative quantification of the heteroplasmy. The disadvantage of both these methods is that prior knowledge of the possible mutation/heteroplasmy is required in designing the PCR primers in DGGE and PNA probe in PNA/qPCR. However, much of the information regarding mitochondrial mutations that are related to age is available in the literature, so, in theory, it would have been possible to try these techniques to detect heteroplasmy. However, this is an expensive method as each mutation has to be targeted separately, and resource limitations did not permit this method to be attempted in this research.

The question whether mitochondrial mutations and heteroplasmy accumulate with age remains unanswered. It has previously been proposed that, rather than just one type of mitochondrial mutation, it could be a combination of mitochondrial point mutations, deletions and duplications that contribute to the ageing process (Melov *et al.* 1995). Therefore, analysing only point mutations in the mitochondrial control region in maternally related individuals may be inconclusive. The small sample size in this study (due to time and budget constraints) also makes it difficult to draw definite conclusions. The next chapter focuses on amplifying the whole mitochondrial genome of the same 12 families used in this study and on the detection of possible mitochondrial deletions.

6.6 Conclusions

The failure of detection of mitochondrial mutations or the increase of heteroplasmy with age in HV1 and HV2 regions in buccal samples, indicate that direct sequencing of HV1 and HV2 cannot be used as an age predicting tool for forensic purposes. To get a picture of the

correlation of mitochondrial mutations with age of a person, it may be necessary to analyse all the mutations that occur in the mitochondrial genome, using a larger sample size from various populations, and three to four maternal generation families.

Chapter 7

Mitochondrial deletions and age

7.1 Introduction

As described in Chapter 6, direct sequencing of the HV1 and HV2 regions could not be used to predict the age of an individual. Analysis of mutations/deletions in the HV1 and HV2 area are limited to a small area of approximately 800 bp from the 16 kb of the whole mitochondrial genome. To understand the mitochondrial mutations that accumulate with age, it may be necessary to analyse all the mutations and deletions that occur in the mitochondrial genome.

As discussed in section 2.15, there are several mitochondrial deletions reported to accumulate with age. The most frequently observed large mutation related to age is the 5 kb deletion, normally called the 'common deletion' (Cortopassi *et al.* 1990). Previous attempts have been made by Meissner *et al.* (Meissner, C *et al.* 1997) and Wurmb-Schwark *et al.* (Wurmb-Schwark *et al.* 2002) to predict the age of an individual by analysing the 'common deletion' in skeletal muscle. In studies by both Meissner *et al.*(Meissner, C *et al.* 1997), which analysed 50 individuals aged 24 to 97 years old, and by Wurmb-Schwark *et al.* (Wurmb-Schwark *et al.* 2002), which analysed 42 individuals aged 3 months to 102 years old, the 'common deletion' was present in the samples in varying percentages. The percentage presence of the 'common deletion' increased with age from 0.00049% to 0.14% in the study by Meissner *et al.* (Meissner, C et al. 1997) and from 0.00004 to 0.25% in the study by Wurmb-Schwark *et al.* (Wurmb-Schwark *et al.* 2002). There was large variation among similarly aged individuals, therefore in both the studies distinction with any statistical

significance could only be made between 'young' and 'old' individuals (Meissner, C *et al.* 1997; Wurmb-Schwark *et al.* 2002).

Most of the mitochondrial deletions reported are involved in short direct repeat sequences where the deletion occurs. For example, the 5 kb 'common deletion' is located between two 13 bp direct repeat sequences at positions 8469 and 13447 (Yen *et al.* 1991). Within the 16 kb of mitochondrial genome there are thousands of 4 to 13bp direct repeat sequences, which could be potential mitochondrial deletions sites {discussed in (Kovalenko *et al.* 1998)}. Therefore analysing the entire mitochondrial genome for deletions would enhance the possibility of detecting more than one deletion.

In previous studies, analysis of mitochondrial deletions across the entire mitochondrial genome was done by Southern blot analysis, high resolution restriction analysis of purified mitochondrial genomes, or, more recently, either amplification of the whole mitochondrial genome with a series of overlapping primers {discussed in (Cheng *et al.* 1994)}. Development of a method of amplifying the complete mitochondrial genome using PCR with one set of primers and high fidelity DNA polymerase has made detection of variation across the entire 16kb of mitochondrial genome more straightforward (Cheng *et al.* 1994).

When a part of the mitochondrial genome is deleted, the resulting size is less than the full-length 16 kb. Therefore, when the mitochondrial genome is amplified, if the amplified product is less than 16 kb in length, it is considered to be the result of a deletion. Previous studies on whole mitochondrial amplification (Table 7.1) carried out by Kopisdas *et al.*, Kovalenko *et al.*, Khrapko *et al.* and Meissner *et al.* (Khrapko, K *et al.* 1999; Kopsidas *et al.*

1998; Kovalenko *et al.* 1998; Meissner, C *et al.* 2006) reported several deletions that accumulate with age.

Table 7.1 –Whole mitochondrial genome amplification and detection of deletions that accumulate with age.

Tissue	Age range (years)	Method	Reference
Skeletal muscle (n = 35)	19 to 77	Long extension PCR	Melov *et al.,* (Melov *et al.* 1995)
Skeletal muscle, cardiac muscle and liver (n = 16)	5 to 90	Quantitative XL-PCR	Kovalenko *et al.,* (Kovalenko *et al.* 1998)
Skeletal muscle (n = 2)	5 to 90	XL PCR	Kopsidas *et al.,* (Kopsidas *et al.* 1998)
Cardiac muscle (n = 7)	31 to 109	Long-distance nested PCR	Khrapko *et al.*, (Khrapko, K *et al.* 1999)
Brain, cardiac muscle and skeletal muscle (n = 50)	2 months to 98	Long PCR	Meissner *et al.*, (Meissner, C *et al.* 2006)

Here, the method described by Meissner *et al.* (Meissner, C *et al.* 2006) was attempted. The method involves amplifying the whole mitochondrial genome of an individual in Long PCR (PCR which amplifies more than 5-10kb of PCR product) and probing with a digoxigenin (DIG) labelled mitochondrial genome. If any mitochondrial deletions are present, the amplified PCR product would be less than 16 kb full length mitochondrial DNA.

The first step in this method was to make a working probe. In order to do this, the whole 16.5kb mitochondrial genome has to be amplified and purified and then labelled. In this study the long PCR method described by Meissner *et al.* (Meissner, C *et al.* 2006) was attempted . After several failed attempts at optimisation, we used the REPLI–g-Mitochondrial DNA kit

from QIAGEN (QIAGEN Pty. Ltd Australia) to amplify the whole mitochondrial genome. This method is based upon rolling circle replication (Spits *et al.* 2006) and the results are expected to contain not only whole mitochondrial genome but also incomplete products. The average product length is estimated to be over 10 kb by the manufacturer, which implies that there are products less than and greater than 10kb (QIAGEN 2006). Thus, though it would not be an appropriate method for amplification of test mtDNA it would still be able to provide an adequate probe.

In previous studies (Table 7.1), the whole mitochondrial genome analysis for deletions was carried out on unrelated individuals within a population. In all the individuals the mitochondrial deletions were highly heterogeneous and differed from individual to individual (Meissner, C *et al.* 2006; Melov *et al.* 1995) and sometimes among tissues in the same individual (Meissner, C *et al.* 2006). However, as mitochondria are inherited from the mother, individuals from the same maternal lineage may not have as much of a difference as that observed among unrelated individuals. This study attempted to examine accumulation of mitochondrial deletions with age in the same 12 maternal lineage families used in the point mutation analysis as described in Chapter 6. The patterns of mitochondrial deletion were to be compared within each family. The youngest members' mitochondrial deletion patterns would be considered as the reference, assuming that these would have the least amount of deletions in the family (Giles, R *et al.* 1980). Any differential deletion pattern from the youngest member of the family could be considered to be age related.

7.2 Results

7.2.1 Amplification of mitochondrial genomes in maternal lineage families

7.2.1.1 Amplification of whole mitochondrial genome using Long PCR

We attempted to amplify the mitochondrial genome of the youngest aged individual in the sample using various $MgCl_2$ conditions (Figure 7.1)

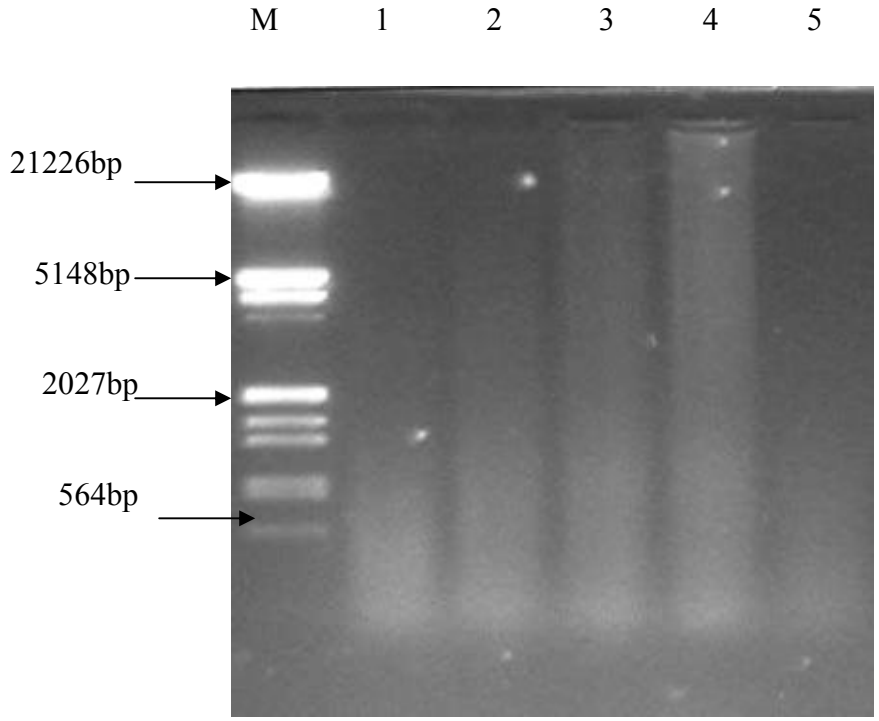


Figure 7.1 – Results of Long PCR amplification attempts using various concentrations of $MgCl_2$. 4 µL of PCR reaction was run on a 1% agarose gel at 75 V for 2 hours. M – λDNA, EcoRI + HindII marker (Fermentas), Lane 1 – 0.5µl of $MgCl_2$, Lane 2 - 1µl of $MgCl_2$, Lane 3 - 1.5µl of $MgCl_2$, Lane 4 - 2µl of $MgCl_2$, Lane 5 – Negative control (0µl of $MgCl_2$).

As can be seen from the gel, not much mitochondrial DNA was amplified to a level where it could be extracted and used as a probe. Only faint, long smears were seen on the gel at any of the concentrations of $MgCl_2$. Several more attempts were made, varying conditions such as

altering template DNA concentrations and annealing temperature, but these were also not successful. Finally another method of amplifying the mitochondrial DNA for use as a probe was sought. The next section shows the results of such an attempt using the REPLI–g-Mitochondrial DNA kit from QIAGEN (Figure 7.2). This kit was selected as the manufacturers claim that it provides highly uniform amplification across the entire human mitochondrial genome and the primers provided in the kit are exonuclease-resistant to achieve high yields (*REPLI-g Mitochondrial DNA Handbook* 2006). Furthermore, when compared with other whole genome amplification methods, the multiple displacement amplification (MDA) used in the REPLI–g-Mitochondrial DNA kit was reported to give the most reliable genotypes, highest call rates, best genomic coverage, and lowest amplification bias (Lovmar *et al.* 2006).

7.2.1.2 - Amplification of whole mitochondrial genome using REPLI–g-Mitochondrial DNA kit from QIAGEN (QIAGEN Pty. Ltd Australia).

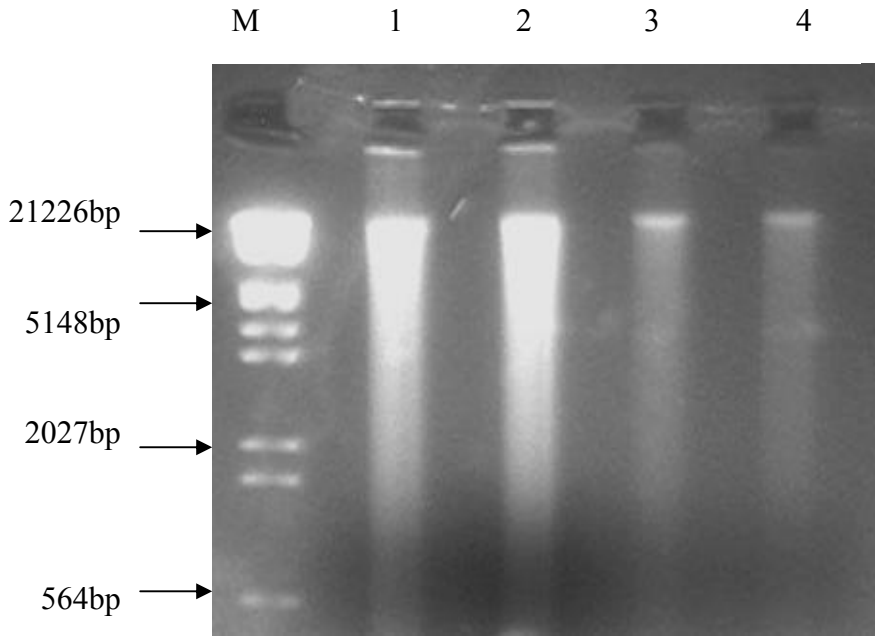


Figure 7.2 - Amplified mitochondrial genomes of two samples and their 1/10 dilutions using the REPLI-g-Mitochondrial DNA (QIAGEN) kit. 4 µL of PCR reaction was run on a 1% agarose gel at 75 V for 2 hours. M – λDNA, EcoRI + HindII marker (Fermentas), Lane 1 – Amplified product of sample from daughter of family # 8, Lane 2 – Amplified product of sample from daughter of Family # 9, Lane 3 – 1 in 10 dilution of amplified product of sample from daughter of Family # 8, Lane 4 – 1 in 10 dilution of amplified product of sample from of daughter of family # 9.

7.2.2 Evaluation of DIG labelling efficiency

Before making and using the final probe, we did a test to evaluate the efficiency of DIG labelling. We attempted to purify the amplified mitochondrial genome of the 'youngest' (5

years of age) of the 37 samples from the excised 16 kb band on the gel. Several attempts at this using the QIAEX II kit (QIAGEN) failed. Finally, the product was purified directly from the solution using the same kit and labelled with DIG (Roche). The labelling efficiency of the DIG- labelled probe was evaluated as discussed in section 3.7.2.2. The 1pg/µl dilution of the DIG labelled probe and of the control DIG labelled DNA (Roche) were clearly visible (Figure 7.3).

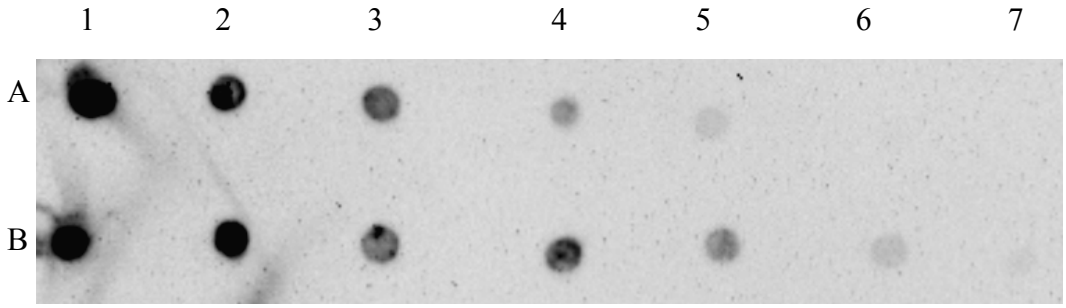


Figure 7.3 – Labelling efficiency of the DIG labelled probe. A – DIG labelled probe, B – Control DNA DIG labelled, Lane 1 – 1 ng/µl, Lane 2- 10 pg/µl, Lane 3 – 3 pg/µl, Lane 4 – 1 pg/µl, Lane 5 – 0.3 pg/µl, Lane 6 – 0.1 pg/µl, Lane 7 – 0.03 pg/ µl.

Thus, the probe was ready to be tested with amplified mitochondrial DNA and then with the samples. However, at this point, a major difficulty was encountered and, unfortunately, this set of experiments had to be stopped. The reasons for this are discussed in the discussion section below.

7.3 Discussion

Initially, attempts were made to amplify the whole mitochondrial genome by methods described by Meissner *et al.* (Meissner, C *et al.* 2006). The same long PCR reagents from the company, Bioline (Bioline, Lukenwalde, Germany), and conditions as per Meissner *et al.* (Meissner, C *et al.* 2006) were used, but all the attempts failed to amplify a clear 16kb band representing the mitochondrial genome (Figure 7.1). Therefore, the REPLI-g Mitochondrial DNA kit by QIAGEN was used as an alternative method to amplify the whole mitochondrial genome for use as a probe. The kit uses Φ29 DNA polymerase, which amplifies DNA by multiple displacement amplification (MDA). The Φ29 DNA polymerase is reported to have a high degree of processivity and fidelity, and the error rate is lower than that of *Taq* DNA polymerase or *Taq* polymerase in combination with *Pwo* polymerase which was used in Long PCR (reviewed in (Lovmar *et al.* 2006)).

A typical DNA yield from this kit is around 150-250 µg per reaction, and average product length is typically greater than 10 kb (*REPLI-g Mitochondrial DNA Handbook* 2006). Samples from the youngest members of Families # 8 and 9 were amplified. The product was found to be very concentrated and thus 1 in 10 dilutions of the PCR products were made to better visualise the 16 Kb band. In the diluted samples, DNA bands of approximately 16 kb were clearly visible (Figure 7.2). As described in Meissner *et al.* (Meissner, C *et al.* 2006), attempts were made to purify this band from the agarose gel but all of them failed. However, as the purified product was only going to be used as a probe, the whole PCR reaction was purified using the kit QIAEX II and labelled with DIG using the 'DIG high Prime DNA Labelling and Detection Starter Kit II' (Roche Diagnostic Australia Pty. Ltd. NSW, Australia).

The labelling efficiency of this kit on this purified product was tested and found to be very good. The 1pg/µl dilution of the DIG labelled probe and of the control DIG labelled DNA (Roche) were clearly visible, thus the labelled probe had reached the expected labelling efficiency, in accordance with the instructions provided in the kit (Figure 7.1).

Whilst it was considered acceptable to label and use the whole product of the REPLI–g-Mitochondrial DNA amplification reaction as a probe, this method of amplification of mitochondrial DNA was not suitable for amplification of mitochondrial DNA from samples to determine the level of deletion mutations. The reason for this is that this method uses a multiple displacement procedure (see Figure 7.4 below).

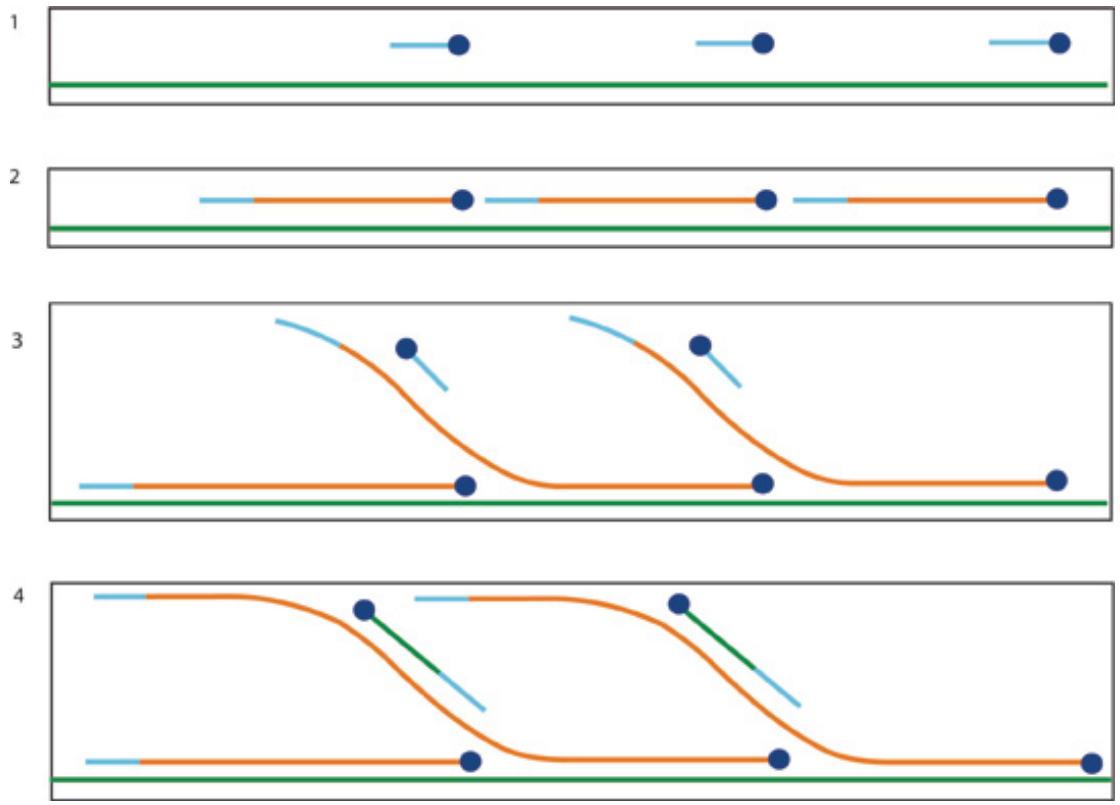


Figure 7.4 - Representation of the multiple displacement amplification method used in the REPLI–g-Mitochondrial DNA kit. (1) The random hexamers (represented by a blue line) bind to the denatured DNA (represented by a green line). (2) The ϕ29 DNA polymerase (represented by a dark blue circle) extends the primers until it reaches newly synthesized double-stranded DNA (represented by an orange line). (3) The enzyme proceeds to displace the strand and continues the polymerisation, while primers bind to the newly synthesized DNA. (4) Polymerisation starts on the new strands, forming a hyperbranched structure. {Reproduced from Spits et al., (Spits et al. 2006)}.

This finally results in amplification of the whole mitochondrial genome, as was evidenced by the ~16 Kb band on the gel in our experiments (Figure 7.2). However, by the very nature of the method, the reaction mixture also contains the smaller intermediate products of the reaction, as shown in Figure 7.4 above, the reason for this being that DNA replication occurring only in the 5' to 3' direction, some of the displaced strands may remain incomplete at their 5' ends. This was evidenced in the smear seen on the gel in Figure 7.2. As the purpose of these experiments was to detect deletion mutations which result in smaller products of whole mitochondrial DNA, these would be lost in the smear produced in this reaction and would either not be detectable or else false positives would be created by this reaction artefact. Thus, it was necessary to amplify the whole mitochondrial DNA from samples by a method such as Long PCR. As this was not successful even in repeated attempts using various conditions, and time constraints did not allow for further experimentation, this study had to be terminated at this point.

7. 4 Conclusions

As deletion mutations in mitochondrial DNA have been reported to be correlated with the age of a person in several studies, it would have been worthwhile investigating this in the context of this research, the goal of which was to predict the age of a person from his/her DNA. The REPLI–g-Mitochondrial DNA kit was successfully used to amplify whole mitochondrial DNA for use as a probe. This was successfully labelled with digoxigenin using the DIG labelling kit from Roche. However, as Long PCR was not successful, the experiments had to be discontinued, as explained above. This could be continued by a researcher in the future.

Chapter 8

Summary, Conclusions and Future Directions

8.1 Summary

The main aim of this study was to develop a method and a formula to predict the age of a person for forensic purposes. Two hypotheses on ageing were considered: the telomere hypothesis (Boukamap 2001), and the mitochondrial hypothesis (Linnane, AW *et al.* 1989). Therefore, telomere length and mitochondrial mutations (point mutations and deletions) were analysed in a sample population composed of different age groups for age dependent changes.

Telomere length decreased with age and the age of an individual could be roughly determined. As discussed in Chapter 4, the presence of large inter-individual variation in telomere lengths prevented the formula from being used in forensic investigations. The causes for the presence of large variation in telomere lengths were further investigated. As described in Chapter 5, when data from families were considered, both mothers' telomere lengths and fathers' telomere lengths had minimal effect on children's telomere lengths. However, the sample size for the families was small, and thus the results are inconclusive. There was no effect of fathers' ages at conception or mothers' age at conception on their children's telomere lengths.

Mitochondrial hypervariable regions 1 and 2 were analysed in twelve families of three to four maternal generations for age dependent mutations or presence of heteroplasmy. There were no age dependent mutations or more than 3% heteroplasmy detected in the two hypervariable regions of mitochondria in the twelve families analysed. The study on mitochondrial

deletions and its relationship with age was not attempted due to technical problems and lack of time to overcome these and get meaningful results.

One of the major obstacles of the study was obtaining a sufficient number of family samples. Even though we used one of the least invasive methods of sampling- buccal swabs- to obtain DNA, we could not obtain many families to volunteer for our study. This was mainly due to the reluctance of parents to give consent for taking samples from their children, although most of the parents were willing to give samples themselves. This limited the total number of family samples for the study as discussed in Chapters 5 and 6.

8.2 Summary of major findings from this study

- Quantitative real time PCR measurement of telomere length was found to be a robust and reproducible method to measure telomere lengths. In the sample of individuals aged 1 -96 years (n = 167), telomere length decreased with age ($r = - 0.185$, $P < 0.05$). The regression analysis between telomere length and age gave a regression coefficient (R^2 value of 0.03679, which indicated the presence of large inter-individual variation.

- Even in the unrelated European population (n = 60), the correlation between telomere length and age was not good. The regression coefficient of the correlation between mean telomere length and age only increased to 0.0545.

- When the population was categorised into 10 year age groups representing each decade of life, the mean telomere lengths of the age groups studied declined in bi-segmental manner against age. From the young (0-9 and 10-19 years old) until early

childhood (20-29 years old) telomere lengths remained similar and a sharp decline was observed from middle age (30-39 years old) until the age of 50-59 years old.

- Even though there was high variation present, there was significant differences in telomere lengths between the "young" [1-39 years (n=101)] and the "old" [40-89 years (n= 65)] ($P<0.05$).

- A high variation in telomere lengths was observed even when similarly aged individuals were analysed [26 year olds (n = 10) and 54 year olds (n = 9)].

- When the data of males and females were analysed separately, telomere length decreased with age in both genders. However, the regression values of the correlation between telomere length and age did not change much, $R^2 = 0.09$ in males and $R^2 = 0.01$ in females.

- There was a more rapid decline of telomere lengths with age in males than in females (Figure 4.4).

- In related families, the mothers' and fathers' telomere lengths were not correlated with the children's telomere length. However as mentioned in Chapter 5, the sample size for the families was small and thus results were inconclusive. Furthermore, there are various complicating factors in studying correlations between parents' and children's telomere lengths using this methodology, some of which are discussed below under "General discussion".

- The mothers' ages at conception and fathers' ages at conception also had no effect on their children's telomere lengths.

- Interestingly, there was a negative correlation between parents' ages at conception and children's ages. Thus, it appears that older individuals had parents who decided to have children at a younger age, while younger individuals tended to have older parents. This could be due to recent life style changes which have resulted in men and women deciding to have children much later in their lives compared to their own parents.

- Mitochondrial control region (HV1 and HV2) mutations did not accumulate with age in the twelve families analysed.

- Mitochondrial heteroplasmy was not present above 3% in the HV1 and HV2 regions of the mitochondria in the 12 families analysed.

8.3 General discussion

In this study, we could not develop a reliable method to predict the age of an individual by analysing either telomere lengths or mitochondrial DNA. Recently, a study carried out by Karlsson *et al.* in 2007 (Karlsson *et al.* 2007) reported at the 22nd Congress of The International Society for Forensic Genetics, Copenhagen, Denmark, made a similar observation after analysing telomere length in leukocytes (n = 96) by real time PCR. In their study, also due to presence of large inter-individual variation, telomere measurement by real-time PCR could not be used to predict the age of an individual.

The observation of high variation in telomere lengths among individuals indicates that there are several factors involved in telomere length regulation. As shown in Figure 8.1, age, inheritance, genetics, gender and environmental factors (obesity, cigarette smoking, stress, etc.) may all interact to regulate the telomere length in an individual.

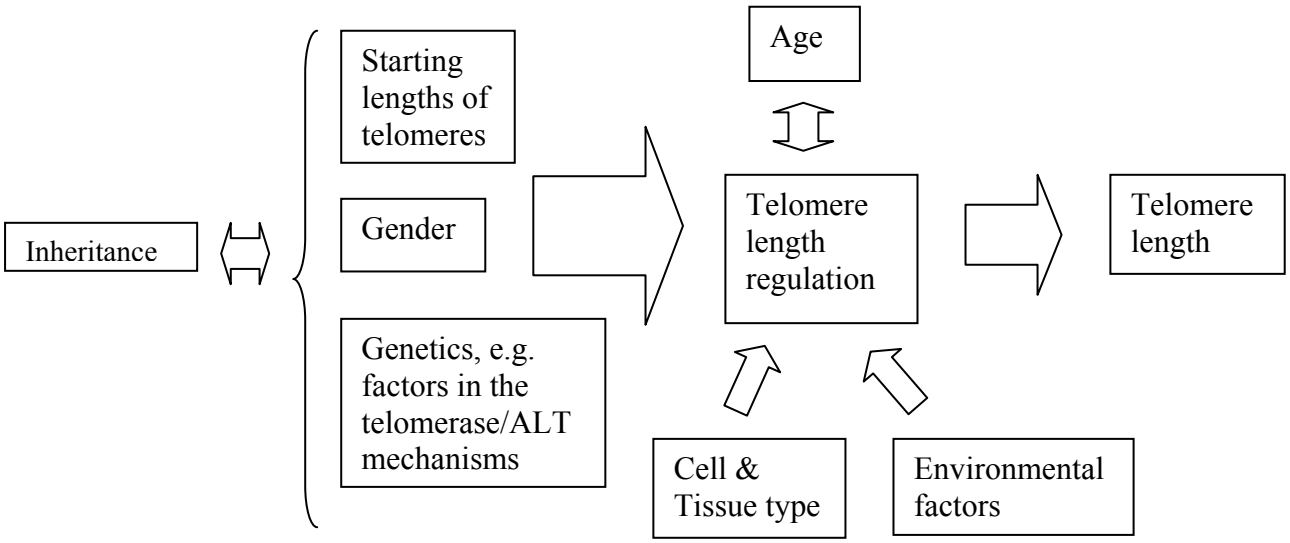


Figure 8.1 – Possible determinants of telomere length.

In hindsight, conducting the telomere length study on a homogeneous population matched for ethnicity, gender, lifestyle, etc. would have probably yielded better information than was obtained in this study. After such an initial study, the effects of these factors could have been studied, one by one, and perhaps use made of various markers such as the amelogenin short tandem repeat (STR), routinely used in forensic analysis to indicate the gender of the person, ethnicity markers, etc. in combination with telomere length results to predict the age of the person. However, it is often difficult to obtain specific samples like that in sufficient numbers in a human population. Furthermore, from results in this study on the limited number of samples matched for gender and ethnicity, it is not likely that a clear relationship would be found between telomere length and the age of a person.

It is recognised that the question of the inheritance mode of telomere lengths is a complex one and could not have been answered through this study alone. However, as there were several family samples in the larger sample, it was considered that comparing telomere sizes within families may be of some interest. Some of the complicating factors are:

a. Most studies so far, including this one, have been based on measurements of the average telomere lengths of all chromosomes. The random distribution and recombination of homologous chromosomes during meiosis in the formation of gametes would complicate the inheritance pattern of the average telomere lengths of all chromosomes together. Studies on telomere lengths of single chromosomes, such as the one conducted by Baird *et al.* (Baird *et al.* 2003) using the STELA (single telomere length analysis) method, may shed some light on the question of inheritance.

b. Even if there may be a link to the DKC 1 gene and its efficiency, the product of which is involved in the regulation of telomerase activity, which could affect telomere lengths in germline cells, there are many other factors involved in the maintenance of telomere lengths in cells (Figure 8.1). A comprehensive study of all these factors would be required before the mode of inheritance can be determined.

Thus, a much more comprehensive study would be required to determine the mode of telomere size inheritance, perhaps using families with markedly long or short telomere lengths, and preferably analysing telomere lengths and factors involved at a very early age in the children, newborn if possible to avoid the effects of telomere maintenance or shortening as these children grow up.

Note that had this study been successful in the development of a method that predicts age with an acceptable degree of accuracy, other factors would have been needed to be considered and further studied, including tissue types of DNA origin, DNA degradation and PCR inhibitors often found in crime scene samples. For example, telomere lengths in various tissue samples from the same set of individuals, such as saliva, blood, hair and semen (in males), which are the most common biological samples left behind in a crime scene (Butler 2005), could have been analysed for comparison. The most common problems faced in DNA analysis in forensics is DNA degradation and presence of PCR inhibitors (Butler 2005). Experiments could have been planned to measure telomere lengths of samples with various degrees of DNA degradation, and other experiments that tested the effects of common PCR inhibitors found in forensic DNA samples (heme, melanin, polysaccharides, bile salts, humic compounds, urea and textile dyes) (Butler 2005). Evaluation of the efficiency of the age prediction formula under the above conditions would be essential in determining how reliable it would be in handling day- to-day forensic samples.

Although the study of point mutations in the HV1 and HV2 regions was unlikely to result in an age prediction tool by itself, it was conducted as there were several reports that linked certain mutations to ageing. However, in this study, no mutations were observed in the older individuals of maternal multigenerational families. Furthermore, there was no sign of even heteroplasmy accumulating with age. This raises the question of whether point mutations in the control region of mitochondria in fact do accumulate with age or not. However, this study was conducted on a limited sample size. Furthermore, the methods used may not have been sensitive enough to detect very low levels of heteroplasmy that may be expected in buccal cells, which are not high-energy requiring. Time and resources did not allow the deletion mutation study to be completed. Hence, the "jury is still out" on this question.

There is an interesting suggestion by Zhang *et al.* (Zhang *et al.* 2003) that the presence of inherited or somatically acquired polymorphisms in mitochondrial DNA could be associated with longevity. In their studies, the mutation C150T in HV2 region was found in 17% of the individuals aged 99-106 years old in the Italian population. In other studies in Irish, Finnish and Japanese populations, C150T was also found to be associated with longevity (Santoro *et al.* 2006). It is suggested that conversion of C to T in position 150 changes the origin of the replication from position 149 to 151, thus providing a survival advantage in humans (Zhang *et al.* 2003). Therefore, it could be that inherited polymorphisms may be associated with ageing and that certain inherited polymorphisms are advantageous.

In any study on mitochondrial DNA mutations and heteroplasmy in particular, other considerations also need to be taken into account. The two main considerations are the number of mitochondrial genomes that exist per mitochondrion as well as the number of mitochondria per cell. A single cell can have 1-1000 mitochondria, and each carries multiple copies of mtDNA (Lodish *et al.* 2000). In humans, there can be 2 to 10 mitochondrial genomes per mitochondrion. The number of mitochondria per cell differs between tissue types: tissues with high energy needs have more mitochondria (for example several hundred in fibroblast cells to over a hundred thousand in oocytes) (Griffiths *et al.* 2005). Thus a single cell could have thousands of copies of the mitochondrial genome (Robin, E *et al.* 1988). This makes picking up heteroplasmy by direct sequencing difficult.

Heteroplasmy occurs when mutant and wild-type mtDNA exist together within a single cell (intracellular heteroplasmy) or when there is a mixture of cells, each of which contains either wild-type or mutant mtDNA (intercellular heteroplasmy) within a tissue. For a mitochondrial heteroplasmy to become physiologically relevant for a given cell, it should reach a certain

threshold level (Khrapko, K *et al.* 1999). To assess the presence of mitochondrial heteroplasmy more accurately, a single cell, or even a single mitochondrion would have to be analysed rather than homogenised tissue. Using the method that has been reported to be good at detecting heteroplasmy up to 1%, denaturing gradient-gel electrophoresis (DGGE) (Michikawa *et al.* 1999), it may be possible to detect actual heteroplasmy levels in a single mitochondrion or in a single cell.

Such a study was recently reported by Lutz-Bonengel *et al.* (Lutz-Bonengel *et al.* 2007) which aimed to detect mitochondrial mutations, level of heteroplasmy in single leukocytes separated by flow cytometry. DNA sequencing, minisequencing and RFLP were used for detecting these. They observed that 4% of the cells analysed were in fact heteroplasmic.

Molecular cloning of mitochondrial DNA genomes either from single cells followed by sequencing could be used to determine the exact amount of heteroplasmy. However, this would be very expensive and time consuming and not practical for use in forensic analysis (Rasmussen, EM *et al.* 2002).

As for the general-interest questions of the extent and rate of telomere shortening and accumulation of mitochondrial mutations, including deletion mutations, with age, the best way to answer these questions would be to conduct a longitudinal study, following the same individuals over their lifetime. Practically, this would be difficult to carry out in humans, and it may be useful to initially conduct such studies on mammalian animal models with shorter life spans such as mice, rats or guinea pigs. If a large enough sample can be studied, it may be possible to study groups matched for gender and the effects of lifestyle also studied, e.g. overfeeding or caloric restriction, amount of exercise and added antioxidants.

8.4 Suggestions for future directions

- Use of techniques such as STELA (Baird *et al.* 2003) and development of other techniques to study telomere lengths of single chromosomes in samples from individuals who are matched for ethnicity, gender, lifestyle, etc.

- Southern blot studies to study the extent of deletion mutations in mitochondrial DNA, both in a cross-section of the population and maternal multigenerational families. The probe could be made from mitochondrial DNA purified from isolated mitochondria, eliminating the need to purify the large, ~ 16 Kb fragment from agarose gels.

- Analysis of the whole mitochondrial genome for mutations in samples from multiple maternal generation families. This could be achieved by amplifying the mitochondrial genome in sections and sequencing of those sections. This would cover point mutations, duplications and well as possible deletions. There is a commercial kit available called, Variant SEQr™ Resequencing System from Applied Biosystem, which includes all the necessary primers and solutions for amplifying the whole mitochondrial genome in sections and sequencing it.

- The stability of microsatellites is reported to decrease during the ageing process (Coolbaugh-Murphy *et al.* 2005). Quantification of microsattellite instability in relation to age by small pool PCR (SP- PCR) may be a way of investigating this phenomenon in “young” and “old” individuals or individuals of various ages.

- 8-deoxyguanine is an oxidized base that is reported to accumulate with age (Wood *et al.* 2000). HPLC (high performance liquid chromatography) could be used to assess the level of this compound in samples from individuals of various ages.

- Whole genome transcriptional analyses (Murphy, CT 2006) of gene expression in particular tissue types, using microarray technology. This could be targeted towards "young" and "old" samples, matched for gender, ethnicity and other factors. This may lead to identification of gene expression changes in relation to the ageing process. For example, genes identified so far as expressed differently during the ageing process are: Notch2, H2AFY2 and CDC5L (Geigl *et al.* 2004), and many more are likely to be involved. The expression levels of a whole suite of genes could, in theory, is used to predict the age of an individual. However, because of variations in gene expression within cell types, one would have to make sure that a homogeneous sample of one cell type was studied. Furthermore, the likelihood of such an approach being useful in forensic investigations is very low due to crime scene samples not being likely to yield un-degraded mRNA for such analyses.

- Similarly protein arrays may be useful in determining the level or activity of particular proteins found in samples from "young" vs. "old" individuals. Post-translational modifications and their effects on protein activity would also have to be taken into account.

8.5 Conclusions

Although the overall goal of predicting the age of a person from his/her DNA for use in forensic investigations was not achieved, this study has yielded a body of valuable information that could be useful in a large number of fields such as studies of ageing, premature ageing, variability in telomere lengths in populations, effects of ethnicity and family inheritance of telomere length, which in turn could be related to longevity, as well as relationship of mitochondrial mutations and heteroplasmy to ageing. Several questions have been raised that warrant a whole set of further investigations, some of which are suggested above.

References

Achi, M V, *et al.* 2000, 'Telomere length in male germ cells is inversely correlated with telomerase activity', *Biology of Reproduction*, vol. 63, no. 2, pp. 591-8.

Alberts, B, *et al.* 1994, *Molecular Biology of the Cell*, 3 edn, Garland science publishing, Taylor & Francis group.

Alberts, B, *et al.* 2008, *Molecular Biology of the Cell*, 5 edn, Garland Science Publishing.

Allsopp, R C, *et al.* 1992, 'Telomere length predicts replicative capacity of human fibroblasts', *Proceedings of the National Academy of Sciences of the USA*, vol. 89, pp. 10114-8.

Alvarez, M, *et al.* 2006, 'The identification of newborns using messenger RNA profiling analysis', *Analytical Biochemistry*, vol. 357, no. 1, pp. 21-34.

Anderson, S, *et al.* 1981, 'Sequence and organization of the human mitochondrial genome', *Nature*, vol. 290, pp. 457-65.

Andrews, R M, *et al.* 1999, 'Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA', *Nature Genetics*, vol. 23, no. 2, p. 147.

Attardi, G 2002, 'Role of mitochondrial DNA in human aging', *Mitochondrion 2* pp. 27-37.

Baird, D M, *et al.* 2005, 'Telomere instability in the male germline ', *Human Molecular Genetics*, vol. 15, no. 1, pp. 45-51.

Baird, D M, *et al.* 2003, 'Extensive allelic variation and ultrashort telomeres in senescent human cells', *Nature Genetics*, vol. 33, no. 2, pp. 203-7.

Baumer, A, *et al.* 1994, 'Age related human mtDNA deletions: A heterogenous set of deletions arising at a single pair of directly repeated sequences', *American Journal of Human Genetics*, vol. 54, pp. 618-30.

Bekaert, S, *et al.* 2004, 'Telomere biology in mammalian germ cells and during development ', *Developmental Biology*, vol. 274, no. 1, pp. 15-30.

Blackburn, E H 1990, 'Telomeres and their synthesis', *Science*, vol. 249, pp. 489-90.

Blasco, M A 2007, 'The epigenetic regulation of mammalian telomeres', *Nature Reviews Genetics*, vol. 8, pp. 299-309.

Boukamap, P 2001, 'Ageing mechanisms: the role of telomere loss', *Clinical and Experimental Dermatology*, vol. 26, pp. 562-5.

Brautbar, A, *et al.* 2008, 'The mitochondrial 13513G>A mutation is associated with Leigh disease phenotypes independent of complex I deficiency in muscle ', *Molecular Genetics and Metabolism*, vol. 94, pp. 485-90.

Brierley, E J, *et al.* 1997, 'Mitochondrial involvement in the ageing process. Facts and controversies', *Molecular and Cellular Biochemistry*, vol. 174, pp. 325-8.

Butler, J, *et al.* 1998, 'Forensic applications of mitochondrial DNA', *Trends in biotechnology*, vol. 16, no. 4, pp. 158-62.

Butler, J M 2005, *Forensic DNA Typing Biology, Technology, and Genetics of STR Markers*, 2 edn, Biology, Technology, and Genetics of STR Markers, Elsevier Academic Press.

Butler, J M 2006, 'Genetics and Genomics of Core Short Tandem Repeat Loci Used in Human Identity Testing', *Journal of Forensic Sciences*, vol. 51, no. 2, pp. 253-65.

Butler, J M, *et al.* 1998, 'Forensic applications of mitochondrial DNA', *Trends in Biotechnology*, vol. 16, no. 4, pp. 158-62.

Butler, M G, *et al.* 1998, 'Comparison of chromosome telomere integrity in multiple tissues from subjects at different ages', *Cancer Genetics and Cytogenetics*, vol. 105, no. 2, pp. 138-44.

Calloway, C D, *et al.* 2000, 'The frequency of heteroplasmy in the HVII region of mtDNA differs across tissue types and increases with age', *American Journal of Human Genetics*, vol. 66, pp. 1384-97.
Cawthon, R M 2002, 'Telomere measurement by quantitative PCR', *Nucleic Acids Research*, vol. 30, no. 10, p. e47.
Cawthon, R M, *et al.* 2003, 'Association between telomere length in blood and mortality in people aged 60 years or older', *Lancet*, vol. 361, no. 9355, pp. 393-5.
Centre, S C U W R 2008, *Humpback whale telomere ageing research project*, 24/02/2008, <http://www.scu.edu.au/research/whales/index.php>.
Chan, C C W, *et al.* 2005, 'Mitochondrial DNA content and 4977 bp deletion in unfertilized oocytes ', *Molecular Human Reproduction*, vol. 11, no. 12, pp. 843-6.
Chan, S R, *et al.* 2004, 'Telomeres and telomerase', *Philosophical Transactions of the Royal Society B: Biological Sciences* vol. 29, no. 359, pp. 109-22.
Chang, E, *et al.* 1995, 'Telomere length and replicative aging in human vascular tissues', *Proceeding of National academic Science USA*, vol. 92, pp. 11190-4.
Cheng, S, *et al.* 1994, 'Complete mitochondrial genome amplification', *Nature Genetics*, vol. 7, pp. 350-1.
Chesnasis, J C 1992, *The demographic transition: stages, patterns and economic implications. A longitudinal study of sixty-seven countries covering the period 1720-1984*, Oxford University press.
Cohen, S B, *et al.* 2007, 'Protein composition of catalytically active human telomerase from immortal cells', *Science*, vol. 315, no. 5820, pp. 1850-3.
Collins, K, *et al.* 2002, 'Telomerase in the human organism', *Oncogene*, vol. 21, no. 4, pp. 564-79.
Coolbaugh-Murphy, M I, *et al.* 2005, 'Microsatellite instability (MSI) increases with age in normal somatic cells ', *Mechanisms of Ageing and Development* vol. 126, no. 10, pp. 1051-9.
Corral-Debrinski, M, *et al.* 1992, 'Mitochondrial DNA deletions in human brain: regional variability and increase with advanced age', *Nature Genetics*, vol. 2, no. 4, pp. 324-9.
Cortopassi, G A, *et al.* 1990, 'Detection of a specific mitochondria DNA deletion in tissues of older humans.' *Nucleic Acids Research*, vol. 18, no. 23, pp. 6927-33.
Database, N C I D 2008, *NCIDD - National Criminal Investigation DNA Database*, 22/02/2008, <http://www.crimtrac.gov.au/systems_projects/NationalCriminalInvestigationDNADatabaseNCIDD.html>.
Del-Bo, R, *et al.* 2003, 'High mutational burden in the mtDNA control region from aged muscles: a single-fiber study.' *Neurobiology of Aging*, vol. 2003 no. 6, pp. 829-38.
Dudek, R W 2006, *High-Yield Cell and Molecular Biology*, Second edn, Lippincott Williams & Wilkins. United States of America. .
Fahn, H J, *et al.* 1996, 'Age-related 4,977 bp deletion in human lung mitochondrial DNA', *American Journal of Respiratory and Critical care Medicine*, vol. 154, no. 4, pp. 1141-5.
Frenck, R W, *et al.* 1998, 'The rate of telomere sequence loss in human leukocytes varies with age ', *Proceedings of the National Academy of Sciences of the USA*, vol. 95, no. 10, pp. 5607-10.
Frudakis, T, *et al.* 2003, 'Sequences associated with human iris pigmentation ', *Genetics*, vol. 165, pp. 2071-83.
Garcia, C K, *et al.* 2007, 'Human diseases of telomerase dysfunction: insights into tissue aging', *Nucleic Acids Research*, vol. 35, no. 22, pp. 1-11.

Geigl, J B, *et al.* 2004, 'Analysis of gene expression patterns and chromosomal changes associated with aging ', *Cancer Research* vol. 64, pp. 8550-7.
Genomics, D P 2007, *DNA Print Genomics Forensic*, 21/09/2007, <http://www.dnaprint.com/welcome/productsandservices/forensics/>.
Gil, M E, *et al.* 2004, 'Real time quantitative PCR of telomere length', *Molecular Biotechnology*, vol. 27, no. 2, pp. 169-72.
Giles, R, *et al.* 1980, 'Maternal inheritance of human mitochondrial DNA', *Proceedings of National Academy of Science U S A*, vol. 77, no. 11, pp. 6715-9.
Giles, R E, *et al.* 1980, 'Maternal inheritance of human mitochondrial DNA', *Proceedings of National Academy of Science USA*, vol. 77, no. 11, pp. 6715-9.
Gillham, N W 1994, *Organelle genes and genomes*, Oxford University press, New York.
Graakjaer, J, *et al.* 2004, 'The relative lengths of individual telomeres are defined in the zygote and strictly maintained during life', *Aging Cell*, vol. 3 no. 3, pp. 97-102.
Graeber, M, *et al.* 1998, 'Recent developments in the molecular genetics of mitochondrial disorders', *Journal of the neurological sciences*, vol. 153, no. 2, pp. 251-63.
Green, R L, *et al.* 2005, 'Developmental validation of the quantifiler real-time PCR kits for the quantification of human nuclear DNA samples', *Journal of Forensic Sciences*, vol. 50, no. 4, pp. 809-25.
Greider, C W 1998, 'Telomerase activity, cell proliferation, and cancer', *Proceedings of the National Academy of Science USA*, vol. 95, no. 1, pp. 90-2.
Greider, C W 1999, 'Telomeres do D-loop - T- loop', *Cell*, vol. 97, pp. 419-22.
Greider, C W, *et al.* 2004, 'Tracking telomerase', *Cell*, vol. 116, no. 2, pp. 83-6.
Griffith, J D, *et al.* 1999, 'Mammalian telomeres end in a large duplex loop', *Cell*, vol. 97, no. 4, pp. 503-14.
Griffiths, A J F, *et al.* 2005, *Introduction to genetic analysis* 8edn, W.H Freeeman and company, New York.
Grimes, E A, *et al.* 2001, 'Sequence polymorphism in the human melanocortin 1 receptor gene as an indicator of the red hair phenotype.' *Forensic Science International*, vol. 122, pp. 124-9.
Harely, C B, *et al.* 1990, 'Telomeres shorten during ageing of human fibroblasts', *Nature*, vol. 345, pp. 458-60.
Harley, C B 1991, 'Telomere loss: mitotic clock or genetic time bomb?' *Mutation Research*, vol. 256, pp. 271-82.
Hattori, K, *et al.* 1991, 'Cardiomyopathy with mitochondrial DNA mutations', *American Heart Journal*, vol. 122, pp. 866-9.
Holland, M M, *et al.* 1999, 'Mitochondrial DNA sequence Analysis validation and use for forensic casework', *Forensic science Review*, vol. 11, no. 1, pp. 22-41.
Home Office, U K 2007, *The national DNA database*, 21/09/2007, <http://www.homeoffice.gov.uk/>.
Hudson, G, *et al.* 2006, 'Mitochondrial DNA polymerase- and human disease ', *Human Molecular Genetics*, vol. 15, no. 2, pp. R244-R52.
Investigation, F B o 2007, *CODIS combined DNA index system*, 20/09/2007, <http://www.fbi.gov/>.
Jacobs, H T, *et al.* 1998, 'The Mitochondrial Genetic System', in KK Singh (ed.), *Mitochondrial DNA mutations in Aging, Disease and cancer*, Springer- Verlag Berlin Heidelberg, New York, pp. 43-71.
Jeffreys, A J, *et al.* 1985, 'Hypervariable 'minisatellite' regions in human DNA', *Nature*, vol. 314, no. 6006, pp. 67-73.
Jobling, M A, *et al.* 2004, 'Encoded evidence: DNA in forensic analysis', *Nature Reviews. Genetics*, vol. 5, no. 10, pp. 739-51.

Kang, D, *et al.* 1998, 'Mitochondria are Essential for Energy Dependent Function of the cell', in K K singh (ed.), *Mitochondrial DNA Mutations in Aging, Disease and Cancer*, Springer- Verlag Berlin Heidelberg, New York, pp. 1-11.
Karlsson, A O, *et al.* 2007, 'Estimating human age in forensic samples by analysis of telomere repeats', paper presented to 22 nd Congress of the International Society for Forensic Genetics, Copenhagen, Denmark.
Katayama, M, *et al.* 1991, 'Deleted mitochondrial DNA in the skeletal muscle of aged individuals', *Biochem Int*, vol. 25, no. 1, pp. 47-56.
Khrapko, K, *et al.* 1999, 'Cell by cell scanning of whole mitochondrial genomes in aged human heart reveals a significant fraction of myocytes with clonally expanded deletions', *Nucleic Acids Research*, vol. 27, no. 11, pp. 2434-41.
Khrapko, K, *et al.* 2007, 'Mitochondrial DNA mutations and aging: a case closed?' *Nature Genetics*, vol. 39, pp. 445-6.
Kirkwood, T B 2005, 'Understanding the odd science of aging', *Cell*, vol. 120, pp. 437-47.
Kitagawa, T, *et al.* 1993, 'Rapid accumulation of deleted mitochondrial deoxyribonucleic acid in postmenopausal ovaries', *Biol Reprod*, vol. 49, no. 4, pp. 730-6.
Koosis, D J 1997, *Statistics - A self-teaching guide*, 4 edn, Wiley publishers.
Kopsidas, G, *et al.* 1998, 'An age associated correlation between cellular bioenergy decline and mtDNA rearrangements in human skeletal muscle', *Mutation Research*, vol. 421, pp. 27-36.
Kovalenko, S A, *et al.* 1998, 'Tissue specific Distribution of Multiple Mitochondrial DNA rearrangements during Human Aging', *Annals of the New York Academy of Sciences*, vol. 854, pp. 171-81.
Kufe, D W, *et al.* 2003, *Cancer Medicine*, BC Decker Inc, Hamilton.
Larsson, N-G, *et al.* 1995, 'Molecular Genetic Aspects of Human Mitochondrial Disorders', *Annual Review of Genetics*, vol. 29, pp. 151-78.
Lee, C, *et al.* 1997, 'Age-associated alterations of the mitochondrial genome', *Free Radical Biology & Medicine*, vol. 22, no. 7, pp. 1259-69.
Lewin, B 1999, *Genes VII*, Oxford University Press.
Lewis, R 1998, 'Telomere tales', *Bio Science*, vol. 48, pp. 981-5.
Lindsey, J, *et al.* 1991, 'In vivo loss of telomeric repeats with age in humans', *Mutation Research*, vol. 256, pp. 45-8.
Linnane, A, *et al.* 1990, 'Mitochondrial gene mutation: the ageing process and degenerative diseases.' *Biochem Int*, vol. 22, no. 6, pp. 1067-76.
Linnane, A W, *et al.* 1989, 'Mitochondrial DNA mutations as an important contributor to ageing and degenerative diseases', *The Lancet*, vol. 1, no. 8639, pp. 642-5.
Liu, V W S, *et al.* 1997, 'Quantitative Allele specific PCR: Demonstration of Age Associated Accumulation in Human Tissues of the A to G mutation at Nucleotide 3243 in Mitochondrial DNA', *Human Mutation*, vol. 9, pp. 265-71.
Liu, V W S, *et al.* 1998, 'Mutations in mitochondrial DNA accumulate differentially in three different human tissues during ageing', *Nucleic Acids Research*, vol. 26, no. 5, pp. 1268 - 75.
Lodish, H, *et al.* 2000, *Molecular Cell Biology*, 4 edn, W. H. Freeman & Co, New York.
Lovmar, L, *et al.* 2006, 'Multiple displacement amplification to create a long-lasting source of DNA for genetic studies', *Human Mutation*, vol. 27, no. 7, pp. 603-14.
Lowe, A L, *et al.* 2001, 'Inferring ethnic origin by means of an STR profile', *Forensic Science International*, vol. 119, no. 1, pp. 17-22.
Lutz-Bonengel, S, *et al.* 2007, 'Single lymphocytes from two healthy individuals with mitochondrial point heteroplasmy are mainly homoplasmic ', *International Journal Legal Medicine*, pp. 10.1007/s00414-007-0190-6.

Maeda, T, *et al.* 2009, 'Age related changes in subtelomeric methylation in the normal Japanese population', *The journals of gerontology series A*, vol. 64A, no. 4, pp. 426-34.
Meissner, C 2007, 'Mutations of mitochondrial DNA – cause or consequence of the ageing process? ', *Zeitschrift für Gerontologie und Geriatrie*, vol. 40, no. 5, pp. 325-33.
Meissner, C, *et al.* 2006, 'Tissue-specific deletion patterns of the mitochondrial genome with advancing age', *Experimental Gerontology*, vol. 41, no. 5, pp. 518-24.
Meißner, C, *et al.* 1997, 'Detection of the age-dependent 4977 bp deletion of mitochondrial DNA ', *International Journal Legal of Medicine*, vol. 110, pp. 288-91.
Meissner, C, *et al.* 1997, 'Detection of the age-dependent 4997bp deletion of mitochondrial DNA', *International Journal of Legal Medicine*, vol. 110, pp. 288-91.
Meissner, C, *et al.* 1999, 'Estimation of age at death based on quantitation of the 4977-bp deletion of human mitochondrial DNA in skeletal muscle ', *Forensic Science International*, vol. 105, pp. 115-24.
Melov, S, *et al.* 1995, 'Marked increase in the number and variety of mitochondrial DNA rearrangements in aging human skeletal muscle', *Nucleic Acids Research*, vol. 23, no. 20, pp. 4122–6.
Meyer, T D, *et al.* 2007, 'Paternal age at birth is an important determinant of offspring telomere length', *Human Molecular Genetics*, vol. 16, no. 24, pp. 3097-102.
Michikawa, Y, *et al.* 1999, 'Aging dependent large accumulation of point mutations in the human mtDNA control region for replication', *Science*, vol. 286, pp. 774-9.
A human mitochondrial genome database, http://www.mitomap.org/.
Mohamed, S A, *et al.* 2006, 'Mitochondrial DNA deletions and the aging heart ', *Experimental Gerontology*, vol. 41, pp. 508-17.
Moraes, C T, *et al.* 2002, 'Mitochondrial DNA Structure and Function', in AHV Schapira (ed.), *Mitochondrial Function and Dysfunction*, Elsevier Science, California USA, vol. 53, pp. 3-24.
Murdock, D G, *et al.* 2000, 'The age related accumulation of a mitochondrial DNA control region mutation in muscle, but not brain, detected by a sensitive PNA directed PCR clamping based method.', *Nucleic Acids Research*, vol. 28, no. 21, pp. 4350-5.
Murphy, C T 2006, 'Using whole-genome transcriptional analyses to identify molecular mechanisms of aging', *Drug Discovery Today: Disease Mechanisms* vol. 3, no. 1, pp. 41-6.
Murphy, S K, *et al.* 2003, 'Imprinting evolution and the price of silence', *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, vol. 25, no. 6, pp. 577-88.
Nagley, P, *et al.* 1998, 'Mitochondrial DNA mutations in Aging', in KK Singh (ed.), *Mitochondrial DNA mutations in Aging, Disease and cancer*, Springer- Verlag Berlin Heidelberg, New York, pp. 205-30.
Nakagawa, S, *et al.* 2004, 'Measuring vertebrate telomeres: applications and limitations', *Molecular Ecology*, vol. 13, no. 9, pp. 2523-33.
Nawrot, T S, *et al.* 2004, 'Telomere length and possible link to X chromosome', *Lancet*, vol. 14, no. 363 (9408), pp. 507-10.
Nekhaeva, E, *et al.* 2002, 'Clonally expanded mtDNA point mutations are abundant in individual cells of human tissues', *Proceeding of National academic Science USA*, vol. 99, no. 8, pp. 5521-6.
Neumann, A A, *et al.* 2002, 'Telomere maintenance and cancer -- look, no telomerase', *Nature Reviews Cancer*, vol. 2, no. 11, pp. 879-84.

Njajou, O T, *et al.* 2007, 'Telomere length is paternally inherited and is associated with parental lifespan', *Proceeding of National academic Science USA*, vol. 104, no. 29, pp. 12135-9.
Nordfjäll, K, *et al.* 2005, 'Telomere length and heredity: Indications of paternal inheritance ', *Proceedings of the National Academy of Sciences of the USA*, vol. 102, no. 45, pp. 16374-8
Oberdoerffer, P, *et al.* 2007, 'The role of nuclear architecture in genomic instability and ageing', *Nature Reviews. Molecular Cell Biology*, vol. 8, no. 9, pp. 692-702.
Ohki, R, *et al.* 2001, 'In Vitro Reconstitution of the end replication problem', *Molecular and Cellular Biology*, vol. 21, no. 17, pp. 5753-66.
Okuda, K, *et al.* 2002, 'Telomere length in the newborn', *Pediatric Research*, vol. 52, no. 3, pp. 377-81.
Ozawa, T 1997, 'Genetic and functional Changes in Mitochondria Associated with Aging', *Physiological Reviews*, vol. 77, no. 2, pp. 425-64.
Pak, J W, *et al.* 2003, 'Mitochondrial DNA mutations as a fundamental mechanism in physiological declines associated with aging', *Aging cell*, vol. 2, pp. 1-7.
Parson, W, *et al.* 1998, 'Population data for 101 austrian caucasian mitochondrial DNA d-loop sequences: application of mtDNA sequence analysis to a forensic case', *International Journal of Legal Medicine*, vol. 111, no. 3, pp. 124-32.
Perls, T T, *et al.* 1998, 'Why women live longer than men- what gives women the extra years?' *Scientific American*, vol. 2, no. Special, pp. 100-3.
Pisano, S, *et al.* 2008, 'Telomeric nucleosomes: forgotten players at chromosome ends ', *Cellular and Molecular Life Sciences* vol. 65, no. 22, pp. 3553-63.
Promislow, D E L 1994, 'DNA repair and the evolution of longevity: a critical analysis ', *Journal of Theoretical Biology* vol. 170, no. 3, pp. 291-300
QIAGEN 2006, *REPLI-g Mitochondrial DNA Handbook*.
Rasmussen, E M, *et al.* 2002, 'Sequencing strategy of mitochondrial HV1 and HV2 DNA with length heteroplasmy ', *Forensic Science International*, vol. 129, no. 3, pp. 209-13.
Rasmussen, L J, *et al.* 1998, 'Genetic Integrity of the Mitochondrial Genome', in KK Singh (ed.), *Mitochondrial DNA mutations in Aging, Disease and Cancer*, Springer-Verlag Berlin Heidelberg, New York, pp. 115-22.
Rattner, J B 1995, 'Centromeres and telomeres', in EE Bittar & N Bittar (eds), *Cellular Organelles*, Greenwich, Conn:JAI Press.
REPLI-g Mitochondrial DNA Handbook, 2006, QIAGEN
Robin, E, *et al.* 1988, 'Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells', *Journal of cell physiology*, vol. 136, no. 3, pp. 507-13.
Robin, E D, *et al.* 1988, 'Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells', *Journal of Cell Physiology*, vol. 136, no. 3, pp. 507-13.
Rufer, N, *et al.* 1999, 'Telomere fluorescence measurements in granulocytes and T lymphocyte subsets point to a high turnover of hematopoietic stem cells and memory T Cells in early childhood ', *The Journal of Experimental Medicine*, vol. 190, no. 2, pp. 157-68.
Saldanha, S N, *et al.* 2003, 'Assessment of telomere length and factors that contribute to its stability ', *European Journal of Biochemistry*, vol. 270, pp. 389-403.
Santoro, A, *et al.* 2006, 'Mitochondrial DNA involvement in human longevity', *Biochimica et Biophysica Acta*, vol. 1757, pp. 1388-99.
Scheffler, I E 2007, *Mitochondria*, Second edn, John Wiley & Sons, Hoboken, New Jersey.
Schmeling, A, *et al.* 2007, 'Age estimation', *Forensic Science International*, vol. 165, no. 2-3, pp. 178-81.

Shiels, P G, *et al.* 1999, 'Analysis of telomere lengths in cloned sheep', *Nature*, vol. 399, no. 6734, pp. 316-7.
Slagboom, P E, *et al.* 1994, 'Genetic determination of telomere size in humans: a twin study of three age groups.' *American Journal of Human Genetics*, vol. 55, no. 5, pp. 876-82.
Slijepcevic, P 2004, 'Telomere length measurement by Q-FISH ', *Methods in Cell Science*, vol. 23, no. 1-3, pp. 17-22.
SoftGenetics 2006, *DNA Variant Analysis Software*, 3.1 edn, SoftGenetics LLC, USA.
Solomon, E P, *et al.* 2008, *Biology*, 8 edn, Thomson Brooks/Cole.
Song, X, *et al.* 2005, 'Specific point mutations may not accumulate with aging in the mouse mitochondrial DNA control region', *Gene*, vol. 350, no. 2, pp. 193-9.
Souza-Pinto, N C d, *et al.* 2002, 'The Mitochondrial theory of aging: involvement of mitochondrial DNA damage and repair', in AHV Schapira (ed.), *Mitochondrial Function and Dysfunction*, Elsevier Science, California USA, vol. 53, pp. 519-34.
Spits, C, *et al.* 2006, 'Whole-genome multiple displacement amplification from single cells', *Nature Protocols*, vol. 1, pp. 1965-70.
Strachan, T, *et al.* 1999, *Human Molecular Genetics 2/e*, BIOS Scientific Publishers.
Suganuma, N, *et al.* 1993, 'Human ovarian aging and mitochondrial DNA deletion', *Horm Res*, vol. 39, no. Suppl 1, pp. 16-21.
Takasaki, T, *et al.* 2003, 'Age estimation in dental pulp DNA based on human telomere shortening', *International Journal of Legal Medicine*, vol. 117, no. 4, pp. 232-4.
Terman, A, *et al.* 2004, 'Aging as a catabolic malfunction', *International Journal of Biochemistry and Cell Biology*, vol. 36, no. 12, pp. 2365-75.
Theves, C, *et al.* 2006, 'Detection and quantification of the age-related point mutation A189G in the human mitochondrial DNA', *Journal of Forensic Sciences*, vol. 51, no. 4, pp. 865-73.
Trifunovic, A 2006, 'Mitochondrial DNA and ageing', *Biochimica et Biophysica Acta (BBA) - Bioenergetics* vol. 1757, no. 5-6, pp. 611-7.
Tsuji, A, *et al.* 2002, 'Estimating age of humans based on telomere shortening', *Forensic Science International*, vol. 126, no. 3, pp. 197-9.
Tully, G 2007, 'Genotype versus phenotype: human pigmentation ', *Forensic Science International: Genetics*, vol. 1, no. 2, pp. 105-10.
Unryn, B M, *et al.* 2005, 'Paternal age is positively linked to telomere length of children', *Aging Cell*, vol. 4, no. 2, pp. 97-101.
Vaziri, H, *et al.* 1993, 'Loss of telomeric DNA during aging of normal and trisomy 21 human lymphocytes.', *American Journal of Human Genetics*, vol. 52, no. 4, pp. 661-7.
Voccolia, V, *et al.* 2009, 'Mitochondrial remodelling in differentiating neuroblasts', *Brain Research*, vol. 1252, no. C, pp. 15-29.
Wang, Y, *et al.* 2001, 'Muscle specific mutations accumulate with aging in critical human mtDNA control sites for replication', *Proceedings of National Academy of Science USA*, vol. 98, no. 7, pp. 4022-7.
Wood, S G, *et al.* 2000, *Measurement of 8-Oxo-deoxyguanosine in lymphocytes, cultured cells, and tissue samples by HPLC with electrochemical detection*
Wurmb-Schwark, N v, *et al.* 2002, 'Quantification of human mitochondrial DNA in a real time PCR', *Forensic Science International*, vol. 126, pp. 34-9.
Yehezkel, S, *et al.* 2008, 'Hypomethylation of subtelomeric regions in ICF syndrome is associated with abnormally short telomeres and enhanced transcription from telomeric regions', *Human Molecular Genetics*, vol. 17, no. 18, pp. 2776-89.
Yen, T C, *et al.* 1992, 'Age-dependent 6kb deletion in human liver mitochondrial DNA', *Biochemistry International*, vol. 26, no. 3, pp. 457-68.

Yen, T C, *et al.* 1991, 'Ageing associated 5kb deletion in human liver mitochondrial DNA', *Biochemical and Biophysical Research Communications*, vol. 178, no. 1, pp. 124-31.
Youngren, K, *et al.* 1998, 'Synchrony in telomere length of the human fetus ', *Human Genetics*, vol. 102, no. 6, pp. 640-3.
Zhang, J, *et al.* 2003, 'Strikingly higher frequency in centenarians and twins of mtDNA mutation causing remodelling of replication origin in leukocytes ', *Proceedings of National Academy of Science U S A*, vol. 100, no. 3, pp. 1116-21.
Zhu, W, *et al.* 2005, 'Mitochondrial DNA mutations in breast cancer tissue and in matched nipple aspirate fluid ', *Carcinogenesis*, vol. 26, no. 1, pp. 145-52.

Appendix I

The Physical ID questionnaire (see below)

The physical ID questionnaire was compiled for the age prediction study as well as for the hair colour prediction study. After collecting the data from individuals on many characters, the answers were assigned numbers (see below) for simplicity in analysing on SPSS15**.**

Gender

Female – 1

Male - 2

Exercise

Mild - 1

Moderate - 2

Heavy – 3

Diet

Small - 1

Moderate - 2

Large – 3

Ethnicity

African Negroid - 1

European - 2

Oriental Asian - 3

South Asian - 4

Middle Eastern - 5

Native American – 6

Appendix II

The physical ID questionnaire data obtained from the 167 individuals were under the file name of categorical data in the compact disk provided. The data saved as both EXCEL and SPSS15 spread sheets.

Appendix III

The sequencing chromatograms and sequences of all 37 individuals of the 12 families were under the file name of sequencing data. The file sequencing data has two separate files for HV1 and HV2 sequences of each family.

Sample Code: _____

Researcher's initials _____

Date completed: ____ / _____ / ____
(Day / Month / Year)

General Information

1. Reason for interest in project:	**2. Date of Birth:** _____ / ____ / _____ (Day / Month / Year)	**3.Gender:** Male ☐ Female ☐

Pigmentation Characteristics

Natural hair colour:

	4. At present age	**5. At ~20 years of age**	**6. At ~5 years of age**
*Unknown	☐	☐	☐
*Dark blonde	☐	☐	☐
*Golden blonde	☐	☐	☐
*Light blonde	☐	☐	☐
*White blonde	☐	☐	☐
*Auburn	☐	☐	☐
*Copper red	☐	☐	☐
*Blonde-red highlights	☐	☐	☐
*Black	☐	☐	☐
*Dark Brown	☐	☐	☐
*Brown	☐	☐	☐
*Brown-red highlights	☐	☐	☐
*Brown-blonde highlights	☐	☐	☐
*Light brown	☐	☐	☐
*Grey	☐	☐	☐
*Other	☐	☐	☐

*(Please specify):

Overall Hair Colour: Black ☐ Dark Brown ☐ Light Brown ☐ Blonde ☐ Red ☐

7. Eye Colour:

*Unknown	☐
*Grey	☐
*Light blue	☐
*Dark blue	☐
*Aquamarine	☐
*Light green	☐
*Dark green	☐
*Hazel	☐
*Light brown	☐
*Brown	☐
*Dark brown	☐
*Other	☐

* (Please specify) :

Sample Code: _____

Researcher's initials _____

Date completed: ____ / _____ / ____
(Day / Month / Year)

Pigmentation Characteristics (Continued)

8. Skin colour

*Very fair ☐
*Fair ☐
*Ruddy ☐
*Olive ☐
*Brown-light ☐
*Brown ☐
*Brown-dark ☐
*Black
*Other
*(please specify):

9. Freckles on face

*No freckles ☐
*Light freckling ☐
*Medium freckling ☐
*Extensive freckling ☐

10. Freckles on body

*No freckles ☐
*Light freckling ☐
*Medium freckling ☐
*Extensive freckling ☐

11. Skin type

*Unknown ☐

*Type 1: Never tan always burn ☐
*Type 2: Tan slightly, always burn ☐
*Between Type 1-2 ☐

*Type 3: Tan well, sometimes burn ☐
*Type 4: Tan well, never burn ☐
*Between Type 3-4 ☐

Other information

12. Hair type:

* Unknown ☐
* Straight ☐
* Wavy ☐
* Loose curls ☐
* Very curly ☐
* Afro ☐

13. Balding

Yes ☐
No ☐

If yes go to 14 & 15

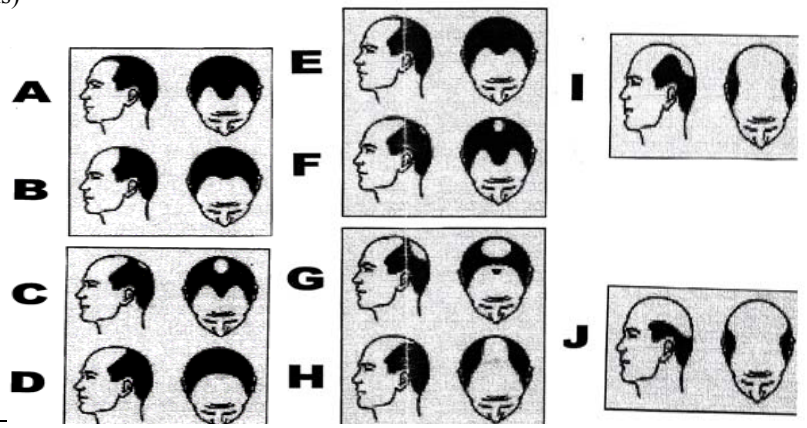
14. Age started balding (yrs)

Not applicable ☐
10-19 ☐
20-29 ☐
30-39 ☐
40-49 ☐
50-59
60+

15. Hair loss pattern
(refer to Balding Patterns)

A ☐
B ☐
C ☐
D ☐
E ☐
F ☐
G ☐
H ☐
I ☐
J ☐

Balding Patterns



Sample Code: _____

Researcher's initials _____

Date completed: ____ / ____ / ____
(Day / Month / Year)

Lifestyle

16. Smoking:

Non smoker ☐

Smoker ☐
(~ how many cigarettes per day) ___

17. Alcohol Consumption:

None ☐

Occasional ☐

Often ☐

18. Diet: (average amount per day)	**Small** (1serve)	**Moderate** (2-3 serves)	**Large** (4-5 serves)
Serves described below			
Green vegetables (~150 gm)	☐	☐	☐
Citrus fruit (1 whole fruit)	☐	☐	☐
Other fruit (1 whole fruit or ~150 gm)	☐	☐	☐
Starchy foods (e.g. rice, potatoes) (150 gm cooked weight)	☐	☐	☐
Beans/Pulses (150 gm cooked weight)	☐	☐	☐
Red meat (100 gm)	☐	☐	☐
Poultry (100 gm)	☐	☐	☐
Fish & Sea food (100 gm)	☐	☐	☐

19. Daily exercise:

(brisk walking, aerobics, weight training, pilates, etc.)

Mild (~20-25 min) ☐

Moderate (~30-50 min) ☐

Intensive (Over 2 hours) ☐

Sample Code: _____

Researcher's initials _____

Date completed: ____ / ____ / ____
(Day / Month / Year)

Background (please tick appropriate box for each)

11. Ethnicity*:	**Self**	**Biological Mother**	**Biological Father**	**Maternal Grandmother**	**Maternal Grandfather**	**Paternal Grandmother**	**Paternal Grandfather**
North European							
South European							
Middle Eastern							
South Asian							
Oriental Asian							
Pacific Islander							
Australian aboriginal							
African Negroid							
Native American							
Other : (Please specify any further factors of your biological ancestry that you think may be relevant or describe using your own words)							
12. Current Age or Age at death							

***North European:** France, Switzerland, Belgium, Netherlands, Denmark, Germany, Austria, Czech Rep., Slovakia, Hungary, Romania, Poland, Ukraine, Moldova, Belarus, Lithuania, Latvia, Estonia, West Russia, Iceland, Greenland (Denmark), United Kingdom, Sweden, Norway, Finland

***South European:** Portugal, Spain, Monaco, Italy, Malta, Slovenia, Croatia, Bosnia and Herzegovina, Serbia and Montenegro, Greece, F.Y.R.O.M (Macedonia), Bulgaria, Turkey.