T.C.

CUKUROVA UNIVERSITY

INSTITUTE OF ADDICTION AND

FORENSIC SCIENCES

DEPARTMENT OF FORENSIC SCIENCES

FORENSIC AGE PREDICTION BASED ON DNA METHYLATION USING NEXT GENERATION SEQUENCING

LILIAN DORIS MUTESI

MASTER OF FORENSIC SCIENCE

JUNE 2023

DECLARATION

I **Mutesi Lilian Doris,** declare that this research is my original compilation and has never been presented for the award of a degree in any University.

Signature ……. Date……

This dissertation has been approved for submission by the academic Supervisor below.

Prof. Dr. Ayşe SERIN

Institute of Forensic Science and Addiction

Cukurova University, Turkey.

Signature……...Date…….

ACKNOWLEDGEMENTS

My sincere appreciation goes to my supervisor Prof. Dr. Ayşe SERIN for her invaluable supervision mentorship, patient, guidance, and enthusiastic encouragement extended to me in making this study a success.

I sincerely appreciate the invaluable advice, continuous support and guidance and extended to me from; Prof. William Bazeyo, Prof. Ireeta, Dr. Lukande, ACP Mubiru Andrew, Serrute Charles, Dr. Kisitu Jaffar and Kobuyonjo Zainab. "Their immense knowledge and plentiful experience have encouraged me in all the time of my academic research and daily life.

I also take this honour to thank my course mates, students of MSc. Forensic Science for the unity, coordination, friendship, and a cherished time spent together throughout the course of the study.

My gratitude extends to Cukurova University, Makerere University and Uganda Police Force for the funding opportunity to undertake my studies at the Institute of Forensic Science and Addiction.

My appreciation also goes out to my family and friends for their encouragement and support all through my studies.

DEDICATION

This dissertation is dedicated to Prof Mete Korkut Gulmen who was our source of inspiration, May His Soul Rest in Eternal Peace.

I also dedicate this dissertation to Prof William Bazeyo who has been a constant source of support and encouragement accorded during the implementation of the work that led to this dissertation.

Lastly, I dedicate this dissertation to my husband for encouraging and supporting me to pursue my dreams and finish this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS AND ABBREVIATIONS

A, adenine

APM, Age Prediction Model

AGEs, advanced glycation end products

AAR, Aspartic acid Racemization

C1orf132, chromosome 1 open reading frame 132; MIR29B2CHG*

C, Cytosine

CGIs, CpG islands

Chr, chromosome

CH3, methyl group

5C, fifth carbon position

Cm or 5mC, methylated cytosine or 5-methylcytosine

CpG, cytosine-phosphate-guanine

DNAm, DNA methylation

DNMT, DNA methyltransferases

EDARADD, EDAR Associated Death Domain*

ELOVL2, ELOVL Fatty Acid Elongase 2*

FHL2, Four and A Half LIM Domains 2*

FASE, Forensic Anthropology Society of Europe

G, Guanine

MAD, mean absolute deviation.

MAE, mean absolute error.

MPS, massively parallel sequencing

MS-HRM, methylation-sensitive high-resolution melting

mtDNA, Mitochondrial DNA

N, number of samples

KLF14, Kruppel like factor 14*

PCR, Polymerase Chain Reaction,

PDE4C, phosphodiesterase 4C, cAMP specific*

R, correlation coefficient beta, β; Pearson correlation

r, Spearman correlation coefficient

RMSE, Root mean square error

SE, standard error

sjTREC, Signal joint T-cell receptor excision circles

STR, short tandem repeats

SBE, Single-base extension

T, Thymine

TRIM59, tripartite motif containing 59*

U, Uracil

vs., versus

ÖZET

Yeni Nesil Dizileme ile DNA Metilasyonuna Dayalı Adli Yaş Tahmini

DNA metilasyon tekniğine dayalı adli yaş tahmini çalışmaları son dönemlerde artış göstermektedir. Bu çalışmada 4-99 yaş arası 25 kişiden alınan yanak içi sürüntü örneğinde yeni nesil dizileme platformu ile 12 CpG bölgesinde (EDARADD, ELOVL2 RASSF5, MIR29B2CHG, FHL2, SMC4, TRIM59, LDB2, ELOVL2-AS1, CNTNAP2, ARHGAP22 ve CCDC102B) yaşa bağlı metilasyon seviyeleri araştırılmıştır. CpG bölgelerinin seçimi diğer araştırma grupları tarafından yayınlanan kısa listelere dayanmaktadır. Bir yaş tahmin modeli geliştirmek için doğrusal regresyon ve çoklu doğrusal regresyon kullanılmıştır. İstatistiksel değerlendirmede 10 CpG bölgesi (EDARADD, ELOVL2 RASSF5, MIR29B2CHG, FHL2, SMC4, LDB2, CNTNAP2, ARHGAP22 ve CCDC102B) için p değeri 0,02976 olup, istatistiksel olarak anlamlıdır. 5,1 yıllık Ortalama Mutlak Sapma ile kronolojik yaş ile tahmin edilen yaş arasında güçlü bir korelasyon tespit edilmiştir (R^2=0,736; düzeltilmiş R^2=0,5199). Çapraz doğrulanmış modelde, RASSF5 (p = 0.033), MIR29B2CHG (p = 0.034) ve SMC4 (p = 0.0323) değişkenleri (CpG siteleri), %95 güven aralığında öngörücü değişken (Yaş) üzerinde önemli bir etkiye sahiptir. Bu eğitim veri seti ile çoklu R^2 değeri 0,9878 (düzeltilmiş R^2=0.9332) ve 2,3 yıllık Ortalama Mutlak Sapma ile kronolojik yaş ile tahmin edilen yaş arasında güçlü bir pozitif korelasyon gözlenmiştir. Bisülfit dizileme ile elde edilen metilasyon düzeylerinin doğruluğunu değerlendirmek için bu çalışmaya %0, %10, %25, %50, %75 ve %100 metilasyon olmak üzere yedi metilasyon standardı dahil edilmiştir. Beklenen metilasyon ile gözlenen metilasyon arasında güçlü bir korelasyon olduğunu gösteren 0.9887 R^2 değeri elde edilmiştir. Bu sonuçlar metilasyon paternlerini ölçmede tekniğin güvenilirliğini ve kesinliğini göstermiştir.

Anahtar kelimeler: DNA metilasyonu, bisülfid dizileme, yaş tayini, adli bilimler

ABSTRACT

Forensic Age Prediction Based on DNA Methylation Using Next Generation Sequencing

Age estimation basing on DNA methylation technique has been on increase in the past years. In this study, we assessed age-associated methylation for 12 CpG sites, (EDARADD, ELOVL2 RASSF5, MIR29B2CHG, FHL2, SMC4, TRIM59, LDB2, ELOVL2-AS1, CNTNAP2, ARHGAP22 and CCDC102B) in 25 buccal samples of living individuals aged 4 -99 years using Next Generation sequencing platform. Selection of the CpG sites used in our model was based on short-lists published by other research groups. Linear regression and multiple linear regression were used to develop an age prediction model. A multiple linear regression validation model with 10 CpG sites (EDARADD, ELOVL2 RASSF5, MIR29B2CHG, FHL2, SMC4, LDB2, CNTNAP2, ARHGAP22 and CCDC102B) revealed a p-value of 0.02976 which is statistically significant. R^2 of 0.736 was obtained indicating a strong correlation between chronological age and predicted age with a Mean Absolute Deviation (MAD) of 5.1 years between chronological age and predicted age and an adjusted R2 of 0.5199. In the cross validated model, the variables (CpG sites) RASSF5 ($p = 0.033$), MIR29B2CHG ($p = 0.034$) and SMC4 ($p = 0.0323$) had a significant influence on the predictor variable (Age) at a 95% confidence level. The Multiple R square value of 0.9878 with this training set of data was obtained indicating a strong positive correlation between predicted age with a Mean Absolute Deviation (MAD) of 2.3 years between chronological age and predicted age and an adjusted R square value of 0.9332. Seven methylation standards of 0%,10%, 25%, 50%, 75%and 100% percentage methylation were included in this study in order to evaluate the accuracy of methylation levels obtained by bisulphite sequencing. R square value of 0.9887 was obtained indicating a strong correlation between expected methylation and observed methylation. This demonstrated the reliability and precision of the technique in measuring methylation patterns.

Keyword: DNA methylation, bisulfite sequencing, age determination, forensic sciences

CHAPTER I

INTRODUCTION

1.1 Background to the study

The study of forensic biology focuses on recovering information from DNA evidence recovered from crime scenes and DNA profiling is the most preferred method for suspect identification. Body fluids such as blood, saliva and semen are the commonest biological evidence retrieved from crime scenes. Proof of identity of the donor of this evidence can be attained through short tandem repeat (STR) profiling. However, in certain situations DNA profiling may not lead to a match with any of the samples in the DNA database and identification of the suspect may not be achieved from the various biological materials collected from crime scenes. With the introduction of age prediction techniques, supplementary information about the donor, such as chronological age can be used to acquire information important to criminal, legal, or anthropological investigations (Schmeling et al. 2007) hence very significant in police investigations (Gršković et al. 2017).

Unidentified remains of the missing persons, victims of mass disasters or casualties in war areas are usually encountered and are often challenging to identify if their surviving relatives cannot be traced (Wright et al. 2015). Mass disasters comprise of a variety of samples that require detailed analysis within a limited time frame. Age estimation from DNA could therefore be used as a screening tool to hasten the process of identification or as supporting data in complicated identifications.

Through analysing the age of a person from the remains or biological traces left behind a missing person or the donor of the traces at the crime scene can be identified (Agudelo et al. 2016). This could possibly help narrow down large groups of suspects or define target groups of interest hence reducing on the number of searches for suspects, especially in cases where there is no eyewitness (Vidaki et al. 2017) and legal hearings could be supported by the inference of age from DNA samples obtained from persons whose age is in under inquiry, such as the probable age of a refugee or the punishment appropriate to young offenders (Schmeling et al. 2016).

Forensic age determination was originally based on the examination of bones and teeth (Baccino et al.2006). Morphological examinations of the pubic symphysis, sacropelvic surface, fourth rib and clavicle from bones or root translucency from teeth are useful for age determination at age at death. It is very key to note that whatever the method of evaluation used, the major problem of use of skeletal features is their lack of informativeness for the old people since, after 65 years of age, the accuracy of age estimation from bones and teeth reduces markedly. The prediction of age from the skeletal remains of children and adolescents can be furthermore determined from the stage of dental or skeletal development (Lewis ME and Flavel A. 2006). In these cases, the level of mineralization of the dentition is the most ideal method. Additionally, assessment of the stage of development of the third molar is important in determining whether a person is under or over the age of 18 years (Lucas et al. 2017). Although it is compelling to apply this technique to the morphological examination of undocumented living subjects of juveniles, the associated use of X-radiation is a major problem that involves ethical issues.

Molecular approaches for age prediction have increasingly been reported in the past five years. These are due to the gradual modifications to biomolecules that happen during the life cycle of an individual because of the process of aging. Five main groups of molecular modifications that have been evaluated include, mitochondrial DNA deletions, advanced glycation end-products (AGEs), shortening of telomeres, asignal-joint T-cell receptor excision circles (herein sjTRECs) and aspartic acid racemization (AAR).

Changes in epigenetic patterns are found to occur either throughout the process of development or childhood/adulthood and have a key impact on the human aging process (Pal S and Tyler JK. 2016). These modifications are categorised under the term “epigenetic clock” or “epigenetic drift” (Jones et al. 2015) and they are due to the continuous interactions between genetics and environment that occur during a person´s lifetime. The epigenetic clock is a directional marvel caused by non-stochastic series of events and encompasses specific regions of the genome linked with age (Pal S and Tyler JK. 2016).

The cumulative variations created by the epigenetic clock and epigenetic drift during a person’s lifespan can be measured and therefore, will provide information on the

age of an individual. This measure will be referred to from now on as the epigenetic age. For forensic investigations, the epigenetic age should aim to be as close as possible to the chronological age. Chronological age is referred to as the time elapsed since birth (Benayoun et al. 2015) and is different from the biological age, which is a measure of the aging process in the individual associated to life expectancy and influenced by factors that are not based on the passage of time alone. Variations between chronological and biological age is mainly affected by the existence of health disorders. Epigenetic age acceleration is most likely to be associated with age-related diseases and death (Chen et al. 2016, Perna et al. 2016). Environmental exposure, such as diet and smoking also have an effect to epigenetics (Lee and Pausova 2013). Epigenetic modification forms include; DNA methylation, chromatic structuring, and histone modification.

Aging is a complex process influenced by several factors such as lifestyle, genetics, and environment. It involves several molecular modifications in cells that accrue over a person's lifetime. These include, gene expression alterations (D. Glass et al. 2013), chemical modifications (A. Pilin et al. 2007), and alterations at the DNA level (X. Ou et al. 2011, A. Tsuji et al. 2002). Previous studies have identified Methylation sites in the human genome that are highly correlated with age and can be used for the development of age estimation algorithms.

The DNA methylation outcome is proposed as a good potential to backup conventional STR profiling in cases of forensics. Most of the forensic cases where undamaged skeletal remains are available, age determination can be conducted effectively by anthropological measurements and calculations as well as cross-referencing with medical histories or records. DNA methylation markers for tissue identification at specific age associated CpG sites have been recommended as the most informative biomarkers for predicting the age of an unidentified donor. Chemical modification of DNA methylation mainly affects cytosines, followed by guanines in a 5′-3′ direction in the DNA double helix, subsequently resulting in the addition of a methyl group (-CH3) to their 5′ carbon (C5). These 5′-3′ CG methylation sites in DNA are referred to as "CpG" dinucleotides, which are mainly methylated in the human genome (Ehrlich et al. 1982). Unmethylated CpGs (CpG islands) are usually found in groups of 300–3000 bp with high CG density (>55% CG content), mainly positioned at the promoter of housekeeping genes (Antequerra and Bird 1993; Espada and Esteller 2010).

It was observed that specific CpG positions in the genome show systematic modifications during a person's lifespan, with progressive increases or decreases in methylation levels.

Several studies in forensic science have reported that DNA methylation analysis is the most dependable method for age prediction (Dias et al. 2020). This allowed the identification of several CpG markers in various tissues and body fluids, with high correlation between DNA methylation and chronological age, potential significant in forensic analysis (Zbieć-Piekarska et al. 2015). Various potential CpG markers that displayed high associations with chronological age were identified. Age-associated CpG sites have been identified in genes such as ELOVL2, C1orf132, TRIM59, FHL2, ASPA, KLF14, PDE4C, EDARADD, NOX4, and TTC7B (Cho et al. 2017).

To deduce the chronological age, several age estimation models have been established based on data generated using different DNA methylation technologies. These include EpiTYPER R (Xu et al., 2015; Freire-Aradas et al., 2016; Zubakov et al., 2016), massively parallel sequencing (MPS) (Naue et al., 2017; Vidaki et al., 2017; Aliferi et al., 2018), pyrosequencing (Weidner et al., 2014; Bekaert et al., 2015; Zbiec-Piekarska ´ et al., 2015), and SNaPshotTM (Lee et al., 2015; Hong et al., 2017; Jung et al., 2019) systems.

Although all these methods have been successfully applied in age estimation, massively parallel sequencing emerges to be the best method of choice for forensic applications as it is highly sensitive, with single base resolution and great multiplexing abilities. Additionally, the MPS technology has been successfully applied to several aspects of forensic analysis (Jaenisch R and Bird A. 2003, Johannessen et al. 2006) and DNA methylation-based age prediction techniques have been previously performed on the MPS platform suggesting potential preliminary results, with one technique reporting a mean average deviation (MAD) of 4.4 years between the biological and the chronological age of the donors when using 16 markers and a second reporting a MAD of 3.2 years with both 13 or a sub-selection of 4 markers (Illumina. 2017).

Next Generation Sequencing is a high-throughput technology that sequences numerous target-specific regions in parallel and quantitatively detects DNA methylation levels by the ratio of sequence read coverage among the target regions (Richards et al., 2018). NGS technology applies sequencing by synthesis where a fluorescently labelled

reversible terminator is imaged as every dNTP that is added, and then cleaved to allow amalgamation of the next base. Ion Torrent technology is based on semiconductor sequencing. Each time a nucleotide is added onto the growing DNA strand, a proton is released and results into a change in pH which is measured as a change in electrical conductivity. NGS is the most precise technology for DNA methylation analysis while allowing for multiplexing. However, the cost of equipment and reagents is high which limits its applicability in most forensic laboratories.

Buccal swabs are relevant sources of DNA, however few studies have developed Age Prediction Models using this tissue type (Bekaert et al., 2015b; Eipel et al., 2016; Naue et al., 2018; Jung et al., 2019; Koop et al., 2020). The specific cellular composition of buccal swabs could influence age predictions based on tissue specific CpGs, for instance, buccal swabs can show different DNAm patterns as a consequence of the different amounts of leukocytes and epithelial cells they contain (Thiede et al., 2000; Vidaki et al., 2016) and this can be reflected in the accuracy of developed models (Vidaki et al., 2016; Jung et al., 2019).

1.2 Problem statement

In forensic science, accurate determination of the age of a victim or suspect from biological evidence can aid in criminal investigations by facilitating the investigating officer to narrow down a search. Forensic DNA analyses have allowed science experts to obtain unique genetic profiles of persons from DNAs extracted from these specimens, which can then be utilised in the identification of the suspect. Nevertheless, when the suspect cannot be identified, determination of physical characteristics such as gender and age become very important in the investigation (D. Glass et al. 2013).

The conventional way of age prediction in forensic science is mainly through morphology of skeletal remains however, this is can only be utilised in cases where the skeletal remains are available and comparatively intact and often requires experienced forensic experts. In certain cases where biological evidence such as, saliva, blood and other tissues are recovered crime scenes, you find when the perpetrator has already run away after committing a crime, in such situations, especially when any direct comparisons with DNA or fingerprint databases cannot be linked and are unable to provide a definitive

match, the need for an accurate age prediction method based on biological material emerges. However, the main drawback of use of skeletal features is their lack of informativeness for the elderly above 65 years of age and the lack of accuracy of age inference from bones and teeth. Therefore, there is need to develop an accurate age prediction model to establish the donors of DNA traces recovered from biological evidence at crime scenes to aid in investigations.

1.3 General objective

The main aim of this study is to develop an Age Prediction Model based on the previously known and validated age-associated genes ELOVL2, EDARADD, FHL2, PDE4C, C1orf132, KLF14, TRIM59 NPTX2, ASPA, GRIA2, ITGA2B and PENK by evaluating the correlation between DNAm levels and age in buccal samples from Ugandan population using Next generation sequencing technology.

1.3.1 Specific objectives

i). To develop an Age Prediction Model for buccal swabs from living individuals aged 3 to 99 years.
ii). To evaluate the accuracy of Next generation sequencing methodology by comparing the estimated chronological age and biological age.
iii). To evaluate DNA methylation patterns and to estimate the chronological age of the donors of the buccal samples.
iv). To ascertain the most accurate age associated CpG gene for age estimation.

1.3.2 Research questions

i). What is the correlation between DNA methylation levels and age in buccal samples from Ugandan population using Next generation sequencing technology?
ii). What is the mean absolute deviation (MAD) or mean absolute error (MAE) of the estimated age based on DNA methylation by Next generation sequencing using data from buccal samples?
iii). How accurate is the age prediction method based on DNA methylation using Next generation sequencing?

iv). What is the most accurate age associated CpG gene for age estimation?

1.4 Justification of the study

The accurate estimation of chronological age of individuals based on DNA methylation will be a breakthrough in forensic investigative and will prove to be a valuable tool to aid in police investigations in Uganda. In cases where no suspect is identified but biological evidence is recovered from a crime scene and no eyewitness is available, determination of the age of the donor of the biological evidence left behind by the perpetrator will potentially narrow down the search for the number of suspects and this will speed up the process of investigations.

Accurate age estimation will also aid in solving judicial, criminal and civil problems, such as immigration cases where the identity and age of individuals are not certain, cases of juveniles regarding questions of adoption, in cases of rape and defilement which require determination of criminal responsibility and also in civil cases of pensionable age for old adults lacking documents (Cunha et al. 2009, Franklin et al. 2015, Parson, 2018; Nuzzolese and Di Vella, 2019).

This research will be the first study in Uganda to predict age basing on DNA methylation using next generation sequencing in individuals aged 3yrs to 98yrs and will add knowledge to the international literature regarding age estimation by Next Generation sequencing.

1.5 Conceptual Framework

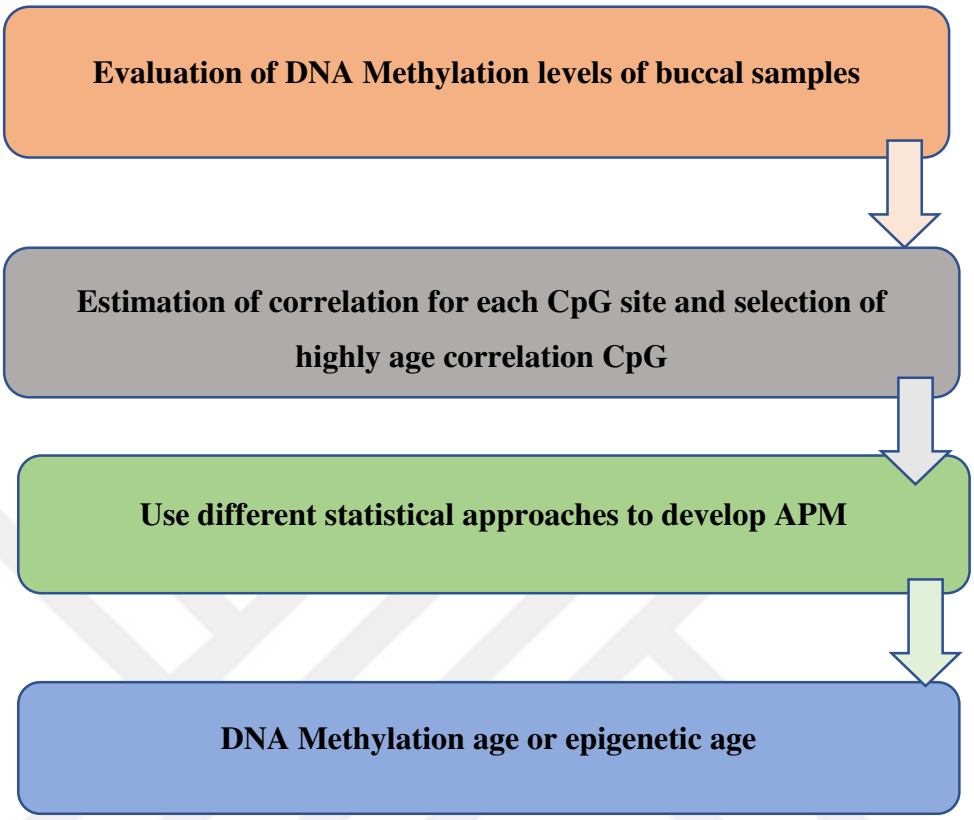


Figure 1: Conceptual framework for age estimation.

CHAPTER II

LITERATURE REVIEW

2.1 Initial Approaches for Age Estimation

Forensic age determination was traditionally based on the examination of bones and teeth. Different methods were suggested that examined adult skeletal remains (Baccino et al.2006), these included morphological examination of the sacropelvic surface, pubic symphysis, fourth rib and clavicle from bones, or root translucency from teeth (Haeusler et al. 2016). However, the major challenge of use of skeletal features is their lack of informativeness for the elderly since, after 65 years of age, the accuracy of age inference from bones and teeth reduces markedly. The determination of age from the skeletal remains of children and adolescents can be furthermore determined from the stage of dental or skeletal development by looking at the level of mineralization of the dentition. Determination of the stage of emergence of the third molar is also relevant in determining whether a person is below or above the age of 18 year (Lucas et al. 2017). Although it is compelling to apply this technique to the morphological analysis of undocumented living subjects of young age, the associated use of X-radiation is a major drawback that entails ethical issues. For these reasons, other methods that use nonionizing imaging techniques such as Magnetic Resonance Imaging (MRI) are currently being explored (De Tobel et al. 2017).

2.2 Molecular approaches for age estimation

These are based on the gradual modifications to biomolecules that occur during the lifespan of an individual because of the aging process. Five major groups of molecular modification have been assessed, these include; mitochondrial DNA deletions, Aspartic Acid Racemization (AAR), Shortening of telomeres, Advanced glycation end-products (AGEs), and Signal-joint T-cell receptor excision circles (herein sjTRECs). The analyses of mitochondrial DNA deletions, Signal-joint T-cell receptor excision circles and shortening of telomeres are DNA-based, while those of Aspartic Acid Racemization (AAR) and Advanced glycation end-products (AGEs) are protein-based methods.

2.2.1 MtDNA deletions technique

This method measures the accumulation of mtDNA deletions that increase as a person ages. The correlation between mtDNA deletions and age is based on the process of energy production in mitochondria from the respiratory chain mechanism. Oxidation of glucose and lipids in the mitochondria to produce ATP discharges free radicals also called reactive oxygen species and causes damage in the mtDNA leading to building up of deletions. In specific, the 4977-nucleotide deletion has been explored as a candidate for age prediction. However, technical problems and a lack of accuracy have hindered its application in forensic cases (Meissner et al.1999).

2.2.2 Shortening of telomeres technique

This method examines the characteristics of telomeres forming the structures positioned at the end of the chromosomes that consist of up to numerous thousands of tandem repeats of the sequence TTAGGG. Shortening of the telomeres happens in every single cell after each mitotic division. To manage this, cells can stimulate the expression of the enzyme telomerase, which incorporates repeat sequences to the end of the telomeres. Nevertheless, this process reduces efficacy with the passage of time. Although age-prediction tests require to measure telomere shortening, just like mtDNA deletions, technical reasons lead to inaccurate results with a general absence of reproducibility.

2.2.3 AGEs technique

It examines AGEs that form a heterogeneous group of compounds formed from the glycation of proteins. They accumulate during the life cycle of an individual in various tissues and have been found to create the frequently observed colour changes (yellow-to-brown) measurable using a colorimeter or spectrophotometer. The degree of colour changes can be used to deduce the age of an individual. Nevertheless, a lack of standardized procedures and the high heterogeneity of this group of biomolecules have hampered its use for age determination (Pilin et al. 2007).

2.2.4 AAR technique

Racemization is a chemical process that converts pure enantiomers (L or D) to a mixture of both. In mammals, exclusively L-amino acids are amalgamated during protein

synthesis. As the individual ages, racemization occurs, and the level of D-amino acid enantiomers starts to increase. AAR is based on the chromatographic measure of D-aspartic acid and this technique is specifically applicable to dentine as an optimal tissue for analysis. This is the most precise age-estimation technique with an estimation error of + or -3 years (Ohtani S and Yamamoto. 2010). Even though the accuracy of the method is high, the technique is damaging. Additionally, exposure of the teeth to high temperatures presents bias to the results, particularly in the identification of burned bodies.

2.2.5 sjTRECs technique

This technique analyses levels of sjTRECs, provides the most feasible molecular technique based on the occurrence of a particular species of DNA in the blood. The sjTRECs detected are circularized DNA molecules that originate from T-lymphocyte DNA rearrangement. An age-correlated decline in the degree of (delta)Rec-(psi)J(alpha) sjTREC occurs during the individual's lifetime. Although the subsequent accuracy of the test is reasonable the prediction error is approximately plus or minus nine years. The technique is limited to blood samples and disorders affecting the immune system and this can possibly affect the predictions made (Zubakov et al. 2010).

These methods, based on the gradual changes of biomolecules due the aging process, have been acknowledged as promising tools for age However, the lack of accuracy and other technical problems have limited the use of such methods in forensic casework (Meissner and Ritz-Timme, 2010, Zolotarenko et al. 2019).

Significant advances have been accomplished in developing new DNA tests to deduce the biological age from a biological trace. The change of emphasis that enabled these developments was a shift in focus from characteristics of DNA and proteins toward epigenetic alterations to DNA, exploiting increased knowledge of human epigenetic patterns from clinical studies. This shift has introduced the door to the identification and development of age-correlated markers that have facilitated robust age-prediction models for forensic use.

2.3 Epigenetic age estimation

The term epigenetics refers to the heritable alterations in gene function that cannot be described by DNA sequence changes (Deans and Maggert 2015; Felsenfeld 2014) which regulates the individual phenotype by modulating the expression and the activity of genes (Armstrong, Epigenetics, Garland Science, New York, 2013; Pinel et al. 2018). Epigenome is the set of chemical changes to the DNA that alter gene expression. Epigenetic modifications control when and how the genes are turned on or off which modulates the protein production in certain cells. There are different epigenetic modification types, and these include DNA methylation, chromatic structuring and histone modification.

Aging is a natural and inevitable process in human life characterized by a gradual loss of physiological integrity due to the accumulation of cellular impairment throughout an individual's lifetime (Jung et al. 2015) and its rates are influenced by heredity, environment, disease and lifestyle. Aging is a complex process associated with numerous molecular changes in cells that accrue over a person's lifetime including epigenetic patterns.

Alterations in epigenetic patterns are found to occur during a person's development. The major difference is that during embryonic and germ cell development, modifications to the epigenetic landscape are biologically reprogrammed and important for lineage determination and cellular identity (Cantone I and Fisher AG. 2013) however, in adult somatic cells, this change of epigenetic marks indicates age-associated deleterious events (Talens et al. 2012). Therefore, epigenetic modifications have a major effect on the human aging process (Pal S and Tyler JK. 2016). These modifications are grouped under the term "epigenetic clock" or "epigenetic drift" (Jones et al. 2015) and are due to the continuous interactions between genetics and environment that occur during the person's lifetime.

The epigenetic clock is a directional phenomenon produced by non-stochastic events and involves particular regions of the genome correlated with age (Pal S and Tyler JK. 2016). In contrast, epigenetic drift involves genome wide variations caused by stochastic events and consequently leads to unpredictable and inconsistent variations in

the epigenome between aging individuals. Although the underlying molecular pathways of epigenetic drift are still uncertain (Li Y et al. 2016), it is known that these differences can be caused by variations in lifestyle characteristics, including smoking habits (Tsaprouni et al. 2014), alcohol intake, sports activity, (Hagerty et al. 2016), physical activity (Barres et al. 2012), and diet (Adaikalakoteswari et at. 2015).

The increasing changes created by the epigenetic clock and epigenetic drift during a person's life cycle can be measured and, consequently, will be informative of the biological age of an individual and not the chronological age. This measure is called the epigenetic age. For forensic investigations, the epigenetic age should aim to be equal to the chronological age as closely as possible. Chronological age is the time that has elapsed since birth, while biological age refers to the measure of the aging process in the individual associated with life expectancy and influenced by factors that are independent of the passage of time alone related to changes in the organism with lifetime. Chronological age is associated with clinical conditions and diseases related to age (Horvath and Raj, 2018). As a result of genetics, epigenetics, and environmental influences as lifestyle choices and diseases, biological age can differ significantly from chronological age. Hence, with the continuous growth of an individual or exposure to various environments, occurrence of health disorders the variation between the biological and chronological ages increases (Baccino et al. 2013). In such a way, epigenetic age acceleration is commonly observed to be correlated with age-related diseases and mortality (Chen et al. 2016, Perna et al. 2016).

According to research, younger individuals demonstrate lower values of MAD between biological and chronological ages, suggesting a high accuracy in the APMs as compared to older ages (Horvath, 2013; Zbieć-Piekarska et al. 2015b; Bekaert et al. 2015a; Hamano et al. 2016; Thong et al. 2017; Naue et al. 2017;). The observed lower accuracy of the APMs with aging suggests a higher differences between biological and chronological ages of individuals with increase of age. This can be related to the increase of specific modifications in DNAm patterns of each individual with aging as a result of the stochastic or environmental factors, being accepted as the epigenetic drift contribution (Tan et al. 2016; Freire-Aradas et al. 2017).

2.4 DNA Methylation

DNA methylation is an epigenetic process that involves the transfer of a methyl (CH3) group onto the C5 position of the cytosine to form a 5-methylcytosine(5cm) or methylated cytosine (Cm) (Schübeler, 2015; Jiang and Guo, 2020). This reaction is catalyzed by DNA methyltransferase (DNMT) enzyme, the three active DNMT enzymes, namely DNMT1, DNMT3A and DNMT3B, modifies cytosine followed by a guanine residue, commonly known as CpG dinucleotide (Bird, 2002; Jurkowska et al. 2011; Miranda and Jones, 2007; Suzuki and Bird, 2008). DNMT1 has de novo as well as maintains methyltransferase activity during cell division, and DNMT3A and DNMT3b are related with de novo methylation(Lyko, 2018).

In mammals, DNAm occurs usually in dinucleotides CpG (5′-CpG-3′cytosine-phosphate-guanine). However, a smaller amount of DNAm can occur in non CpG sites (CpH: H = A, C, T), being detected in embryonic stem cells and neurons (He and Ecker, 2015). The human genome comprises of about 28 million CpG dinucleotides, not evenly distributed along the genome, which may be methylated or non-methylated. Majority of the genome are methylated, however 20-40% of CpG sites are non-methylated (Deaton and Bird, 2011; Schübeler, 2015; Ciccarone et al. 2018; Field et al. 2018).

Usually, unmethylated CpG dinucleotides are situated at clusters with high frequency of CpGs (CpG-rich regions) defined as CpG islands (CGIs) (Vidaki et al. 2013; Kader and Ghai, 2015). CGIs are considered as ≥200 bp long, having ≥50% of GC content and Obs CpG/Exp CpG ≥0.60 (Sarda and Hannenhelli, 2018). CGIs can be situated in promoter regions, intergenic or intragenic regions (Illingworth et al., 2010). Around 60% of the CGIs are present in gene promoters and are normally unmethylated being correlated with gene expression of many genes (Larsen et al. 1992; Freire-Aradas et al. 2017). The remaining 40% are located at “orphan CpG islands”, along the intergenic and intragenic regions (Illingworth et al. 2010; Zampieri et al. 2015; Freire-Aradas et al. 2017; Sarda and Hannenhelli, 2018; Ciccarone et al. 2018). Some orphan CGIs may possibly have a role as an alternative promoter, being linked to tissue-specific gene expression (Illingworth et al. 2010; Sarda and Hannenhelli, 2018).

DNAm has been currently associated with gene expression and gene silencing by employing proteins involved in gene repression or by preventing the binding of transcription factor(s) to DNA. Through the effect of gene transcription, DNAm patterns are important for various mechanisms in an organism such as development and differentiation, imprinting and X-chromosome inactivation and controls tissue-specific gene transcription among others (Deaton and Bird, 2011; Moore et al. 2013; Freire-Aradas et al. 2017; Sarda and Hannenhelli, 2018).

It has been observed that several CpGs located at different genes, such as ELOVL2, FHL2, NPTX2, EDARADD, ASPA, GRIA2, ITGA2B, PDE4C, PENK, CCDC102B, C1orf132, TRIM59 and KLF14, are highly correlated with chronological age (Freire-Aradas et al. 2017; Goel et al. 2017; Zolotarenko et al. 2019). Subsequently, a number of age prediction models (APMs) based on methylation information of these sites have been established allowing to accurate age predictions. These APMs capture chronological age but are also the indication of the biological age (Chen et al. 2016; Zhang et al. 2017a; Field et al. 2018; Horvath and Raj, 2018; Nwanaji-Enwerem et al. 2018). Actually, variation between chronological and biological ages are affected by age-related conditions as frailty (Breitling et al. 2016) and age-related diseases, such as cancer (Horvath, 2013; Ambatipudi et al. 2017; Dugué et al. 2018) and Parkinson's disease (Horvath and Ritz, 2015).

2.5 Approaches used for the development of APMs based on DNAm

Developing Age Prediction Models involves several steps, and these include (Goel et al. 2017).

i) The assessment of DNAm levels at different CpGs from a particular tissue type.
ii) The evaluation of correlation coefficient values between every CpG site and chronological age, enabling the selection of the highly age associated CpGs from the studied tissue.
iii) the use of multivariate regression analyses through various statistical approaches, to build the most accurate Age Prediction Model.

In recent years, various APMs have been developed using various CpGs, different tissue types and statistical analyses or DNAm technologies (Goel et al., 2017; Freire-Aradas et al. 2017, 2020; Horvath and Raj, 2018; Liu et al. 2019; Bergsma and Rogaeva, 2020). Various technologies exploring DNAm levels in the last decade include, EpiTyper mass spectrometry (Garagnani et al. 2012;Xu et al. 2015; Giuliani et al. 2016; Freire-Aradas et al. 2016, 2018), SNaPshot (Lee et al. 2015; Hong et al. 2017, 2019; Jung et al. 2019), massively parallel sequencing (MPS) (Vidaki et al. 2017; Naue et al. 2017, 2018; Aliferi et al. 2018), methylation-sensitive high resolution melting (MS-HRM) (Hamano et al. 2016, 2017) and microarray hybridization technology (Bocklandt et al. 2011; Hannum et al. 2013; Horvath, 2013; Florath et al. 2014). However, pyrosequencing was the most frequently used (Weidner et al. 2014; Bekaert et al. 2015a, 2015b; Zbieć-Piekarska et al. 2015a, 2015b; Eipel et al. 2016; Park et al. 2016; Cho et al. 2017; Thong et al. 2017; Spólnicka et al. 2017; Daunay et al. 2019; Pfeifer et al. 2020).

Majority of these approaches employ sodium bisulfite conversion, which facilitates the deamination of unmethylated cytosine bases to uracil and leaves 5-methylcytosine. The method of bisulfite sequencing is considered to be the "gold standard" technique for DNAm analysis being a quantifiable, qualitative and effective method (Li and Tollefsbol, 2011; Patterson et al. 2011). Normally, for forensics analysis it is acknowledged as the method of choice (Vidaki et al. 2013). Most established APMs are based on a few CpGs correlated with age, extensively distributed throughout the genome, majorly identified, and chosen from DNAm array technology. These APMs have been established basing on blood samples (mainly from living individuals) (Bekaert et al. 2015a; Zbieć-Piekarska et al. 2015a, 2015b; Huang et al. 2015; Park et al. 2016; Hamano et al. 2016; Cho et al. 2017; Naue et al. 2017, 2018; Spólnicka et al. 2017; Thong et al. 2017; Alghanim et al. 2017; Jung et al. 2019; Daunay et al. 2019; Pfeifer et al. 2020), but other body tissues and fluids have also been investigated, including saliva (Alghanim et al. 2017; Hong et al. 2017, 2019; Jung et al. 2019), semen (Lee et al. 2015, 2018), teeth (Bekaert et al. 2015a; Giuliani et al. 2016; Márquez-Ruiz et al. 2020), buccal swabs (Bekaert et al. 2015b; Eipel et al. 2016; Naue et al. 2018; Jung et al. 2019; Pfeifer et al. 2020) and bone (Naue et al. 2018; Gopalan et al. 2019; Lee et al. 2020). Moreover, few studies have investigated the stability of DNAm levels in bloodstains (Zbieć-Piekarska et al. 2015a; Huang et al. 2015; Thong et al. 2017). However, the selection of universal age

markers can be perplexing for the development of accurate APMs due to the inter-tissue inconsistency of DNAm patterns or tissue specificity of age-associated DNAm changes (Freire-Aradas et al. 2017; Jung et al. 2017; Naue et al. 2018; Slieker et al. 2018; Jung et al. 2019). The suggested age correlated CpGs in each tissue cannot be applied for age estimation in other tissue types, before being tested in those tissues. Nevertheless, some markers can be promising for age estimation across different tissue types, and DNAm alterations of various age-dependent markers have been applied in recent studies to develop multi-tissue APMs (Horvath, 2013; Alsaleh et al. 2017; Jung et al. 2019). In otherwards, the documentation of universal age biomarkers that could retain high correlation values with chronological age across different tissue types, with the same accuracy in APM and not being affected by other factors, could be a challenging task (Koch and Wagner, 2011; Horvath, 2013).

2.6 DNAm age prediction models (APMs)

In earlier years, research on DNAm analysis for forensic age estimation have increased in number, using a variety of partially overlapping genes, different tissues and different techniques. Many studies in forensic age estimation have been described using different methodological approaches to evaluate DNAm levels that are considered to be the highly promising age estimation biomarkers. Various APMs were developed in several particular tissue types, basically in blood samples from living individuals, and others were recommended across multiple tissues. Similar age estimation accuracies were constantly obtained based on these different methodological approaches using different tissue types, various markers, or different statistical methods. The ELOVL2 gene has been extensively investigated, showing to be the most accurate age prediction marker so far. Using many diverse sources of DNA, such as, blood, saliva, teeth, buccal cells, brain, bone, liver and old bloodstains, ELOVL2 CpGs shows a higher rates of age correlation. In otherwards, the development of multi-tissue APMs has risen. Nevertheless, it is essential to identify the best multi-tissue or tissue-specific age markers. It is necessary to determine and choose the specific markers that can show the same accuracy in various types of samples using the same methodology for development forensic casework. Subsequently, DNAm levels can become an accurate and reliable method to be applied for several contexts in forensic age prediction.

Majority of the forensic APMs have been developed based on blood DNA samples of healthy, living individuals, but other body tissues are now also being explored. For example, four studies focused on the possible DNAm changes in blood samples from deceased individuals, which can be an applicable source of DNA for forensic age prediction models (Bekaert et al. 2015a; Hamano et al. 2016; Naue et al. 2018; Pfeifer et al. 2020). Similarly, various studies have used tooth samples (Bekaert et al. 2015a; Giuliani et al. 2016; Márquez-Ruiz et al. 2020) and bone samples (Horvath et al. 2015, 2018; Naue et al. 2018; Gopalan et al. 2019; Lee et al. 2020) for DNAm age prediction. Teeth and bones are very important sources of DNA being, frequently, the last evidence of any person in cases of forensic identification. Considering the previous review addressing the highly age associated genes for the development of APMs based on DNAm levels in various tissue types, several genes were identified for the present study: the seven earlier known and validated age correlated genes include; ELOVL2, EDARADD, FHL2, PDE4C, C1orf132/MIR29B2CHG, KLF14 and TRIM59.

In the highest of this list, being one of the greatest investigated age estimation markers, is the ELOVL2 (ELOVL Fatty Acid Elongase 2) gene, positioned on the chromosome 6p24.2. CpG sites from the ELOVL2 gene has been frequently included in various APMs using different tissue types such as blood, bones buccal swabs, saliva and tooth samples (Garagnani et al. 2012; Bekaert et al. 2015a, 2015b; Zbieć-Piekarska et al. 2015a, 2015b; Giuliani et al. 2016; Park et al. 2016; Freire-Aradas et al. 2016; Hamano et al. 2016, 2017; Cho et al. 2017; Thong et al. 2017; Spólnicka et al. 2017; Bacalini et al. 2017: Naue et al. 2017, 2018; Jung et al. 2019; Gopalan et al. 2019; Daunay et al. 2019; Márquez-Ruiz et al. 2020; Lee et al. 2020; Pfeifer et al. 2020).

The remaining genes, including FHL2 (Four And A Half LIM Domains 2), positioned on chromosome 2q12.2, EDARADD (EDAR Associated Death Domain), positioned on chromosome 1q42.3, PDE4C (phosphodiesterase 4C, cAMP specific), positioned on chromosome 19p13.11, C1orf132 (chromosome 1 open reading frame 132), positioned on chromosome 1q32.2, TRIM59 (Tripartite motif containing 59), located on chromosome 3q25.33, and KLF14 (Kruppel-like factor 14), positioned on chromosome 7q32.3, are other promising age associated genes that have been also investigated in various studies using different biological samples including blood, buccal swabs and

saliva (Garagnani et al. 2012; Florath et al. 2014; Zbieć-Piekarska et al. 2015b; Hamano et al. 2016; Giuliani et al. 2016; Freire-Aradas et al. 2016; Cho et al. 2017; Thong et al. 2017; Spólnicka et al. 2017; Bacalini et al. 2017; Jung et al. 2019; Daunay et al. 2019; Pfeifer et al. 2020; Koop et al. 2020). In specific, FHL2, EDARADD and PDE4C genes were included in several studies for DNAm age estimation using tooth samples (Bekaert et al. 2015a; Giuliani et al. 2016; Márquez-Ruiz et al. 2020) and TRIM59, KLF14, C1orf132 andFHL2 genes revealed its valuable usefulness for DNAm age prediction in bones (Naue et al. 2018; Gopalan et al. 2019; Lee et al. 2020).

2.7 NGS technique and DNA methylation

NGS is a high-throughput technology that sequences many target-specific regions in parallel and quantitatively detects DNA methylation levels by the ratio of sequence read coverage among targets (Richards et al. 2018). NGS technology employs sequencing by synthesis where a fluorescently labeled reversible terminator is imaged as each dNTP addition, and then cleaved to permit incorporation of the next base. Ion Torrent technology is based on semiconductor sequencing. Each time a nucleotide is added to a growing DNA strand, a proton is released and the consequent change in pH is measured as an alteration in electrical conductivity.

NGS is the most precise technology for DNA methylation analysis while allowing for multiplexing. However, the high cost of equipment and reagents is a limiting factor for most forensic laboratories.

CHAPTER III

MATERIALS AND METHODS

3.1. Material list

3.1.1. Chemicals

Proteinase K solution, Lysis buffer, Ultra-pure Dithiothreitol, Prepfiler Automate Forensic DNA extraction kit, Absolute ethanol, TE buffer, Quant trio DNA quantification kit, Methyl code Bisulphite conversion kit, EpigenDx methylation standards, , Ion AmpliSeq Kit for Chef DL8, Ion Ampliseq 12CpG panel ,Exo-SAP purification Kit, Ion 530 chip kit, Ion 540 Kit-Chef (2 sequencing runs per initialization).

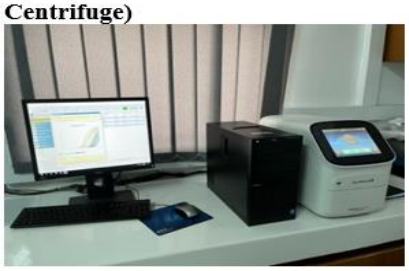
3.1.2. Plastic and other materials

Cotton swabs with wooden handle, Rnase free microfuge tubes 1.5ml, filter tips(10,20,50,200,300 and 1000),reagent troughs, Abgene 96 well PCR plate, 2.2ml 96 deep well plate, micro amp optical 8 tube strip with attached optical cap , prepfiler 96 well processing plate, Micro amp optical semi-skirted PCR Plate, disposable bags, gloves,

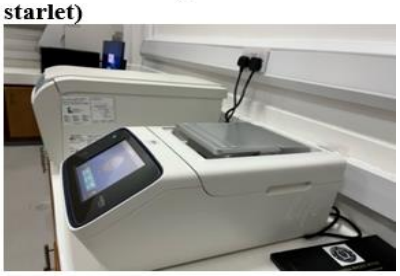
3.1.3. Equipment

Lysis (Thermo-shaker and Centrifuge)

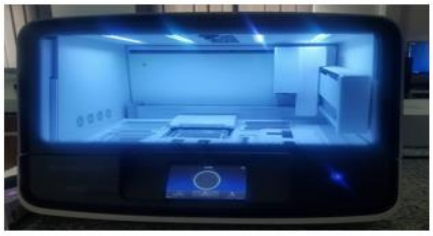
DNA extraction(Hamilton ID starlet)



DNA quantification(Quanti- studio 5)



Bisulphite conversion(Proflex PCR instrument)



Library preparation (Ion -chef)

Sequencing(Ion Gene Studio S5)

3.2. Methods

3.2.1. Sample Collection

This study was operated under ethical approval by College of Veterinary Medicine ,Animal Resources and Biosecurity, Makerere University, Research Ethics Committee (COVAB.MSC1.2023) in regard to buccal sample collection for DNA methylation analysis. Buccal samples were collected from a total of 25 donor volunteers aged between 4 and 99 years using sterile buccal swab sticks. Prior to sampling, full informed consent was acquired from the donors or their parents for the cases of under-aged individuals. All samples were unconnected to any disease study, and information was not collected regarding medical history of the donors in the effort to create an inclusive, unbiased dataset that was representative of the general population.

A sterile swab stick was gently rubbed against the inner cheek (buccal mucosa) for about 30 seconds. On both cheeks to increase the chances of obtaining a sufficient sample. The swab was then airdried and packed in a swab stick container and labeled with the participants name and date of birth.

3.2.2. DNA Standards of known methylation

To assess the sensitivity of the method, seven pre-mixed methylation standards (EpigenDx, Massachusetts, USA) for methylation levels of 0%, 5% ,10%, 25%, 50%, 75% and 100% at a concentration of 50ng/µl were used.

3.2.3. DNA Extraction

Before DNA extraction lysis was done, a master mix containing 300µL of prepfiler lysis buffer, 7µL of proteinase k and 3µL of DTT was added to every sample in an Eppendorf tube. The sample was then loaded on a thermo-shaker set at 70°C and speed of 900rps for 40 minutes. The samples were then removed from the thermo-shaker, centrifuged for 2 minutes at 12,000-14,000rpm and after the lysate removed from the substrate.

Genomic DNA was then extracted from the lysates using an automate ID starlet instrument (Thermofisher Scientific, Massachusetts, USA) in combination with the Prepfiler automate DNA extraction kit for all the samples. A work list was created for all samples, the run was started ,run parameters were defined , magnetic particles prepared and loaded onto the instrument, samples were put into tube careers and loaded onto the robotic, all required plastics, wash buffers and elution buffer were also loaded. After loading all the lab wear run was started and the pipetting process begun. Once run was complete a message was displayed(run successful), samples were removed and subsequently stored at -20°C.

3.2.4. DNA Quantification

Quantification of DNA extracts was conducted using the Quantifiler® Trio DNA Quantification kit in combination with the Quanti studio 5 instrument, both provided by Thermo-Fisher Scientific (Massachusetts, USA). For every DNA sample 1µL was added

into a master mix containing 2.5µL of primer mix and 2.5µL of master mix and put into the well of the plate, the plate was then loaded onto the instrument, HID Real-Time PCR Analysis Software v1.3 was started, run setup and run parameters defined (experimental name, defined samples, assigned targets and samples, assigned samples to wells and run started. Once the run was complete, analysis setting was applied to all samples and quantification results were then displayed ready for sodium bisulphite conversion. The manufacturer's guidelines were followed throughout the protocol.

3.2.5. Sodium Bisulfite Conversion

Treatment with sodium bisulfite was employed for the conversion of unmethylated cytosines to uracils in the DNA samples. A total of 100ng of DNA from each sample or standard was converted using the Methyl code Bisulfite conversion kit (Thermo-Fisher Scientific (Massachusetts, USA). CT conversion reagent was prepared and 130µl of the prepared CT Conversion Reagent was added to 20µl the DNA , the samples were then loaded onto the thermal cycler and run the program at 98°C for 10 minutes (DNA denaturation),64°C for 2.5 hours (Bisulfite conversion) and 4°C. Once the run was complete , the samples were unloaded, and the converted DNA was eluted in 10µL of the elution buffer. Eluates were processed immediately according to the manufacturer's guidelines. The estimated DNA recovery following bisulfite conversion by means of this chemistry has been calculated as 52% and therefore the final concentration of the eluate is predicted at approximately 2.6ng/µL.

Preparing the

3.2.6. Age-associated CpG sites

This study was based on 12 CpG sites previously described by Vidaki et al. 2017. These were considered due to the good amplification efficiency and good overall contribution to the prediction accuracy. Information on the location and gene association of the CpGs is displayed in Table 1. Primers used in this study were of the same design as those originally described by Vidaki et al. 2017 (Table 1).

Table 1: Designed bisulfite PCR assays

CpG site	Associated Genes	Primer Sequence (5'-3')	Amplicon length (bp)
cg16867657		F AGGGGYGTAGGGTAAGTGAGG	
cg21572722	ELOVL2		308
cg24724428		R AACAAAACCATTTCCCCCTAATAT	
cg06639320		F GTTTTTGGGATTAGGTAGAGATTT	
cg22454769	FHL2		165
cg24079702		R TTTATTTACCAAAACTCCTTTCTTC	
cg09809672	EDARADD	F GGTTTGATTTTGGTTAGATAATTAG	148
		R AAAAACTTTAATACCTCTCCCCATC	
cg22796704	ARHGAP22	F GGATTTAGGGGTAGGTAGAATTTGT	148
		R TCTAAACTAAACTTAACCACCTTCC	
cg08128734	RASSF5	F ATTTTGGGTATTTGGAAGGTATTT	189
		R TCCCAATTAAAACCAAAAATAAAAA	
cg17372101	CNTNAP2	F GTTTTAAAGTAGGTTAAGAAGTGGGAGT	124
		R AAAACAAAAAATATCCCTAAATTTCCT	
cg08160331	KLHL35	F TATTAAGAGGTAGTATTAAAAGATGATGAA	231
		R CTTACTTCCTAAAAAAAATAAAAAC	
cg10501210	MIR29B2CHG	F AAGAAGGTGAGAAAGATAGAGTATTTATAT	210
		R TAAAAAATTTAATAAAACCAAATTCTAAAA	
cg19283806	CCDC102B	F GGGTTATAAGTTTTGTTTTGATGAAGT	171
		R AATAAATTTCTCCTTAAACAATCCC	
cg07553761	SMC4, TRIM59	F GTGGTTTGGGGGAGAGGT	86
		R CCAAATAAAAAATAATTCCTCAAAAAC	
cg08262002	LDB2	F TTTTGGGTATTGAGTGAGGTATAGG	110
		R ACCATTCATACATTCTAACAAAACC	

3.2.7. Amplification of the bisulfite-converted DNA

Diluted the bisulfite converted DNA to 15 µL with Nuclease-free Water. DNA was added to the barcode plate for library preparation and loaded the plate on the instrument, the reagents were thawed and instrument prepared, Ion AmpliSeq™ 2X primer pools were added to positions A and B of the reagent cartridge, installed reagents and consumables on the Ion Chef™ instrument. Entered panel specific amplification parameters (amplicons per pool was 384 and target amplification cycles were 26 followed by 4 minutes extension time at 40°C) , performed automated library preparation and pooling of barcoded sample libraries on the Ion Chef™ Instrument. Once the run was completed, unloaded the library pools for sequencing and cleaned up the Ion Chef™

instrument. The Ion Ampliseq kit for Chef DL8 was used following manufacturers protocol.

3.2.8. Post-PCR Purification

Following amplification, samples were purified using the Exo-SAP purification Kit by Thermofisher Scientific, Massachusetts, USA in order to remove unincorporated primer residues according to manufacturer's protocol. 5µL of a post-PCR reaction product was mixed with 2µL of ExoSAP- IT reagent for a combined 7µL reaction volume. Incubated the mixture at 37°C for 15 minutes to degrade remaining primers and nucleotides and then incubated the mix at 80°C for 15 minutes to inactivate ExoSAP-IT™ reagent. The PCR was now ready to use in DNA sequencing. Treated PCR products were stored at –20°C until use.

3.2.9. Templating , chip loading and sequencing

The purified libraries were prepared for templating. Templating was carried out on the Ion Chef instrument using the Ion 530™ Kit and sequencing of the libraries was performed using the Ion Gene Studio™ S5 Sequencing instrument (Thermofisher Scientific, Massachusetts,,USA). Created a planned run in the Torrent Suite™ Software by entering run parameters for templating and sequencing the samples, diluted the library samples, thawed the reagents cartridge at room temperature for 45 minutes before use, then loaded the Ion Chef™ Instrument with sequencing chips, reagents, and sample libraries. After loading the Ion chef instrument, started the Ion Chef run by following the step-by-step instructions on the Ion Chef™ Instrument home touchscreen. Once the Ion Chef Run was completed, unloaded the prepared sequencing chips from the Ion Chef™ instrument and loaded the prepared sequencing chips on a sequencer one at a time and followed the instructions on the touchscreen to start the sequencing run.

3.2.10. Data Analysis and Normalisation

Analysis of the FASTQ files were indexed with the Burrows-Wheeler Aligner (BWA), Sequence Alignment/Map (SAMtools) software. The methylation percentages for each CpG site were calculated using the following formula methylation percentage = (total methylated C / total C) 100. All the analyses were carried out using inhouse python

scripts. The CpG sites were identified based on the following regions of the human chromosome(Table 2).

Table 2: Chromosomal location and genetic information on the 12 CpG sites employed in this study

CpG site	chromosome	chromosome Regions
EDARADD	chr1	chr1', 236557586, 236648230
RASSF5	chr1	chr1', 206680864, 206762780
MIR29B2CHG	chr1	chr1', 207974863, 207996048
FHL2	chr2	chr2', 105974169, 106054986
SMC4	chr3	chr3', 160117438, 160152741
TRIM59	chr3	chr3', 160153291, 160167574
LDB2	chr4	chr4', 16503164, 16900268
ELOVL2	chr6	chr6', 10980992, 11044538
ELOVL2_AS1	chr6	chr6', 11043991, 11079377
CNTNAP2	chr7	chr7', 145813893, 148118090
ARHGAP22	chr10	chr10', 49654079, 49864310
CCDC102B	chr18	chr18', 66382453, 66722426

3.2.11. Statistical analysis

The statistical analyses were performed using GraphPad Prism Version 5, Rstudio & R free ware Version 4.30 (Packages; MASS, QSARdata, caret, latticeExtra, ggcorrplot). Linear regressions were used to analyse relationships between methylation levels and chronological age at each single CpG site. Using the regression coefficients from the age correlated CpG sites, we predicted age of individuals for each individual site. The same ten highest age correlated CpG sites were selected for subsequent analysis using multiple linear regression to build a final age prediction model (APM). The p values, R square values and mean absolute deviation between chronological and predicted ages (MAD) were calculated. A validation of the age prediction model was performed in a test set of 11 buccal samples from living individuals and cross validation using the training set of 11 individuals. The MAD value, p value and R square values were obtained for the prediction model.

3.2.12. Ethical considerations

The study was approved by the Research and ethics committee of College of Veterinary Medicine, Animal Resources and Biosecurity, Makerere University (COVAB.MSC1.2023).

CHAPTER IV

RESULTS

4.1. Massive parallel sequencing data analysis

A total number of 10,010,234 reads were generated by the NGS platform, and a total number of 5,869,462 reads were mapped to the genomic reference (hg19 Lambda) after trimming off the adaptors, ambiguous 1bases, low-quality fragments and duplicated sequences. The coverage of each sample at every CpG was calculated by the following equations: coverage = (mapped reads x average length)/the genomic target region. The coverage of 25 samples and seven methylation standards were calculated.

Table 3: Coverage Analysis Report

Barcode Name	Sample	Mapped Reads	On Target	SampleID	Mean Depth	Uniformity
IonCode_0101	AEY23S001	116,920	0.87%	0.00%	0.282	94.57%
IonCode_0102	AEY23S002	141,998	4.75%	0.00%	1.181	77.80%
IonCode_0103	AEY23S003	88,003	0.45%	0.00%	0.106	97.93%
IonCode_0104	AEY23S004	67,087	0.40%	0.00%	0.074	98.54%
IonCode_0105	AEY23S005	85,413	0.52%	0.00%	0.11	97.86%
IonCode_0106	AEY23S006	68,477	51.07%	0.00%	73.67	35.49%
IonCode_0107	STD 1	9,250	0.48%	0.00%	0.061	98.84%
IonCode_0108	STD 2	29,095	0.14%	0.00%	0.006	99.89%
IonCode_0109	AEY23S007	69,516	37.46%	0.00%	11.43	15.47%
IonCode_0110	AEY23S008	119,817	22.72%	0.00%	15.48	26.65%
IonCode_0111	AEY23S009	83,150	1.05%	0.00%	0.233	95.49%
IonCode_0112	AEY23S0010	56,200	1.29%	0.00%	0.186	96.41%
IonCode_0113	AEY23S011	45,222	0.75%	0.00%	0.073	98.57%
IonCode_0114	AEY23S012	10,052	0.17%	0.00%	0.004	99.91%
IonCode_0115	STD 3	20,305	0.20%	0.00%	0.009	99.83%
IonCode_0116	STD 4	44,623	0.09%	0.00%	0.005	99.90%
IonCode_0117	AEY23S013	93,845	0.88%	0.00%	0.214	95.85%
IonCode_0118	AEY23S014	109,213	12.24%	0.00%	6.6	19.34%
IonCode_0119	AEY23S015	90,587	11.59%	0.00%	5.23	16.89%
IonCode_0120	AEY23S016	286,989	0.77%	0.00%	0.375	92.87%
IonCode_0121	AEY23S017	175,751	24.28%	0.00%	90.62	23.80%
IonCode_0122	AEY23S018	49,928	1.53%	0.00%	0.133	97.39%
IonCode_0123	STD 5	28,973	5.55%	0.00%	0.452	91.34%
IonCode_0124	STD 6	33,339	0.12%	0.00%	0.006	99.89%
IonCode_0125	AEY23S019	310,408	68.48%	0.00%	474.1	76.62%
IonCode_0126	AEY23S020	199,134	51.60%	0.00%	16.82	6.57%
IonCode_0127	AEY23S021	88,672	3.36%	0.00%	0.7	86.66%
IonCode_0128	AEY23S022	197,624	45.06%	0.00%	38.08	9.37%
IonCode_0129	AEY23S023	20,311	3.26%	0.00%	0.093	98.18%
IonCode_0130	AEY23S024	19,260	14.72%	0.00%	0.971	81.93%
IonCode_0131	AEY23S025	48,003	5.16%	0.00%	1.414	75.50%
IonCode_0132	STD 7	107,568	53.75%	0.00%	135.2	19.68%

Table 4: CpG Methylation results

Sample ID	Age	Total	EDARADD	RASSF5	MIR29B2CHG	FHL2	SMC4	TRIM59	LDB2	ELOVL2	ELOVL2-AS1	CNTNAP2	ARHGAP22	CCDC102B
0105_1and2_82	6	6.52	24.12	24.87	24.78	18.13	16.31	13.33	21.22	17.31	21.99	18.34	23.24	19.97
0102_1and2_82	4	6.26	24	24.54	23.45	19.8	14.42	15.01	18.28	20.44	22.54	18.79	20	17.68
0132_3and4_84	STD 7	7.24	26.94	19.71	23.52	19.23	14.44	20.54	19.62	23.99	23.91	16.21	24.94	17.77
0127_3and4_84	14	5.15	18.75	22.52	28.26	17.12	12.9	13.95	22.01	16.67	17.2	16.53	18.45	15.95
0110_1and2_82	9	6.32	29.32	20.83	22.02	22.16	14.06	19.23	19.07	19.86	19.53	19.37	22.39	17.57
0120_3and4_84	99	4.43	19.4	18.5	24.42	17.73	14.29	13.11	18.2	17.84	18.46	18.03	18.61	14.75
0117_3and4_84_2	28	6.06	16.67	20.34	21.7	24.86	12.06	21.9	19.91	21.79	24.11	18.16	22.41	17.92
0121_3and4_84	34	9.55	21.51	23.75	35	20	12	28.12	20.54	13.43	10	19.09	19.74	17.29
0111_1and2_82	26	6.73	20.34	15.04	22.41	23.89	14.29	23.08	21.03	13.13	18.57	18.33	20.06	11.76
0126_3and4_84	50	4	0	0	0	0	0	0	13.51	0	0	17.86	16.67	13.51
0116_1and2_82	STD4	3.4	24.6	19.27	24.84	22.11	14.76	15.74	17.89	19.66	25.44	16.83	21.38	17.03
0103_1and2_82	36	6.2	24.71	22.95	24.14	19.12	16.67	16.22	18.14	21.35	23.78	18.53	21.03	17.05
0104_1and2_82	50	5.61	24.29	23.88	34.88	26.87	19.57	12.2	25.75	19.64	26.09	19.11	23.12	15.25
0109_1and2_82	27	4.77	0	0	0	0	0	0	0	0	0	34.43	14.29	12.5
0108_1and2_82	STD 2	1.62	19.95	14.14	22.96	20.2	13.01	12.4	14.79	15.27	22.7	17.3	22	13.47
0119_3and4_84	30	7.16	21.69	17.33	30.43	26.67	6.25	18.18	19.79	16.67	15.62	20.45	24.37	16.09
0128_3and4_84	13	10.2	0	14.29	0	28.57	33.33	0	17.86	33.33	0	18.64	7.14	22.58
0118_3and4_84	50	6.65	22.06	12.12	27.59	22.73	14.29	12.05	20.73	17.8	19.05	17.39	20.39	18.17
0129_3and4_84	41	6.75	31.82	31.82	38.46	13.64	0	25	16.93	11.11	15.38	21.6	25	18.88
0114_1and2_82	55	6.4	15.91	18.18	28.57	14.63	19.44	15	21.05	7.89	36.11	16.89	19.74	11.4
0124_3and4_84	STD 6	6.21	25.98	21.31	21.76	22.09	13.12	14.77	18.48	18.73	23.74	16.44	21.51	16.31
0113_1and2_82	39	6.99	21.28	20.97	30.41	18.82	13.07	15.85	19.46	19.46	13.64	18.77	20.06	14.96
0123_3and4_84	STD 5	8.16	23.58	20.13	23.1	20.78	15.46	15.54	19.95	25.9	24.9	17.36	22.81	17.79
0106_1and2_82	45	7.25	13.64	43.24	40	35.14	20.69	0	17.02	26.47	17.24	19.34	17.2	21.68
0101_1and2_82	4	5.83	27.5	21.58	24.2	19.1	16	16.51	17.67	19.8	22.22	17.49	22.01	14.46
0131_3and4_84	93	10.04	15.25	23.21	25	30.36	22.22	9.68	19.93	10.42	13.33	16.3	28.67	17.22
0130_3and4_84	75	6.93	20	30.95	19.35	30.95	25.81	16	16.67	20	22.58	23.54	20.9	15.62
0107_1and2_82	STD 1	1.3	15.15	15.15	33.93	16.67	3.57	8.33	12.99	12.5	19.64	14.88	8.82	12.99
0122_3and4_84	65	7.11	26.42	19.23	39.13	21.15	20.59	10.53	20.77	15.91	20.59	18.44	19.49	18.27
0112_1and2_82	43	6.79	28.17	18.31	27	20.42	20.54	17.71	20.57	17.16	25.89	18.56	20.26	19.53
0125_3and4_84	27	7.98	20	32.56	34.62	6.98	25.93	12	19.49	27.03	14.81	19.32	24.48	13.33
0115_1and2_82	STD 3	2.2	19.51	19.06	23.96	21.82	12.53	15.32	14.95	20.56	22.01	16.64	21.77	15.96

4.2. Correlation between DNA methylation and age

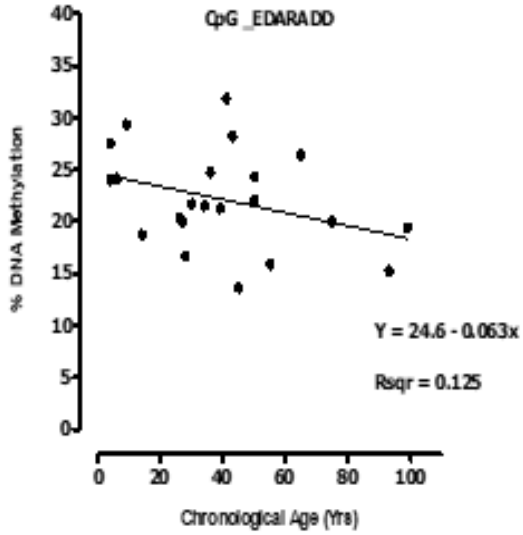
Age associated methylation measurements were obtained by NGS for each of the 12 CpG sites (*EDARADD, ELOVL2 RASSF5, MIR29B2CHG, FHL2, SMC4, TRIM59, LDB2, ELOVL2-AS1, CNTNAP2, ARHGAP22 and CCDC102B)* in 32 buccal samples from living individuals, aged 4–99 years old.

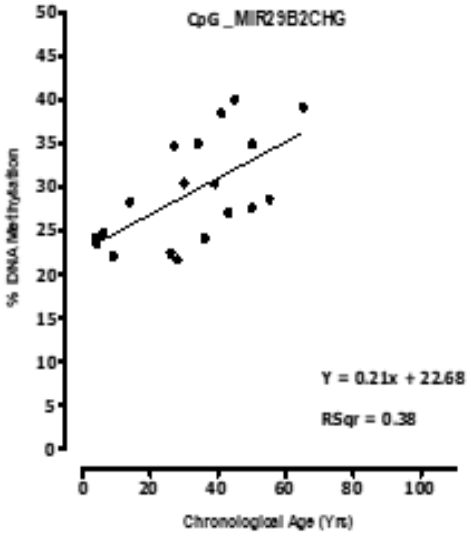
Correlation between CpG methylation of all the twelve CpG sites was then assessed using linear regression analysis in all the 25 samples excluding the 7 Methylation standards used. Simple linear regression revealed strong correlations (R^2>0.3) between CpG methylation and age for only CpG site *MIR29B2CHG (*(R^2>0.38). Relatively low correlation (R^2>0.3) was revealed between CpG Methylation and age for CpG sites *SMC4, TRIM59, FHL2, EDARADD, ELOVL2, ELOVL2-AS1, RASSF5, CCDC102B, ARHGAP22, CNTNAP2* and *LDB2* with R^2 values of*;* 0.17, 0.14, 0.13, 0.125, 0.12, 0.06, 0.025, 0.02, 0.004, 0.001 and 0.0036 respectively which were less than 0.3 indicating a weak relationship between CpG Methylation and age **(**Table 2**).**

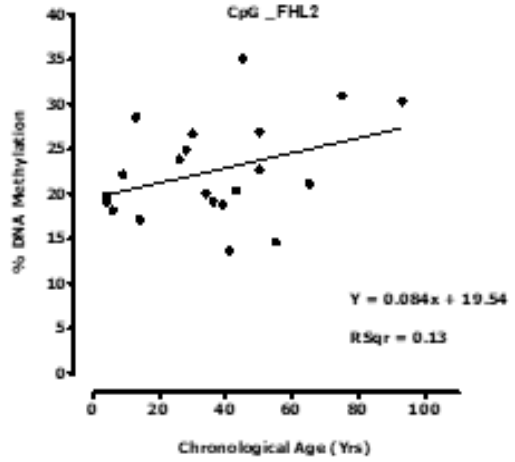
Methylation of CpG sites in *TRIM59, EDARADD, ELOVL2 and CCDC102B was* negatively correlated with age, whereas positive correlation was observed between methylation and age for CpG sites *MIR29B2CHG, SMC4, FHL2, ELOVL2-AS1 and RASSF5.* No correlation was observed between methylation with age for CpG sites *LDB2, ARHGAP22* and *CNTNAP2* (Figure 2)**.**

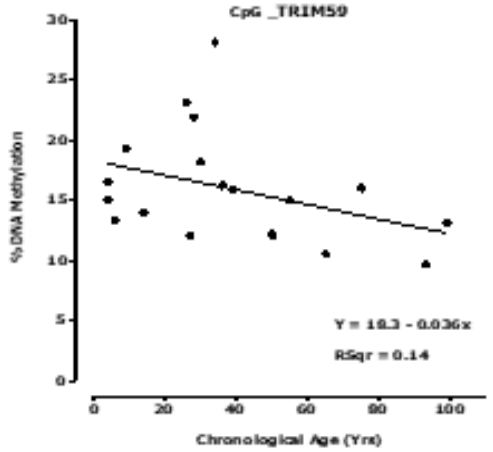
Table 5: Regression statistics for 12 CpG sites: *EDARADD, ELOVL2 RASSF5, MIR29B2CHG, FHL2, SMC4, TRIM59, LDB2, ELOVL2-AS1, CNTNAP2, ARHGAP22 and CCDC102B*.

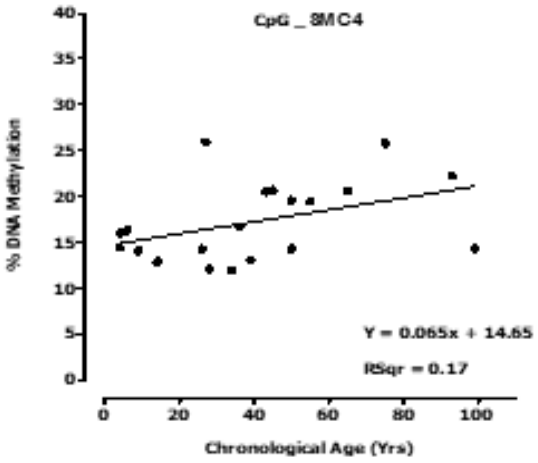
CpG	R Square
EDARADD	0.125
ELOVL2	0.12
RASSF5	0.025
MIR29B2CHG,	0.38
FHL2	0.13
SMC4	0.17
TRIM59	0.14
LDB2	0.0036
ELOVL2-AS1	0.06
CNTNAP2	0.001
ARHGAP22	0.004
CCDC102B	0.02

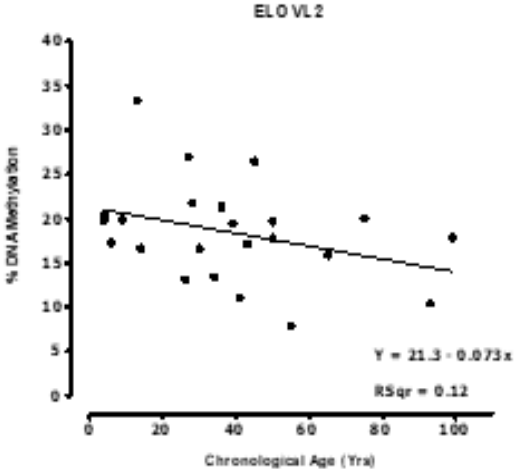

CpG _EDARADD
% DNA Methylation
Chronological Age (Yrs)
Y = 24.6 - 0.063x
Rsqr = 0.125


CpG _MIR29B2CHG
% DNA Methylation
Chronological Age (Yrs)
Y = 0.21x + 22.68
RSqr = 0.38


CpG _FHL2
% DNA Methylation
Chronological Age (Yrs)
Y = 0.084x + 19.54
RSqr = 0.13


CpG _TRIM59
% DNA Methylation
Chronological Age (Yrs)
Y = 18.3 - 0.036x
RSqr = 0.14


CpG _SMC4
% DNA Methylation
Chronological Age (Yrs)
Y = 0.065x + 14.65
RSqr = 0.17


ELOVL2
% DNA Methylation
Chronological Age (Yrs)
Y = 21.3 - 0.073x
RSqr = 0.12

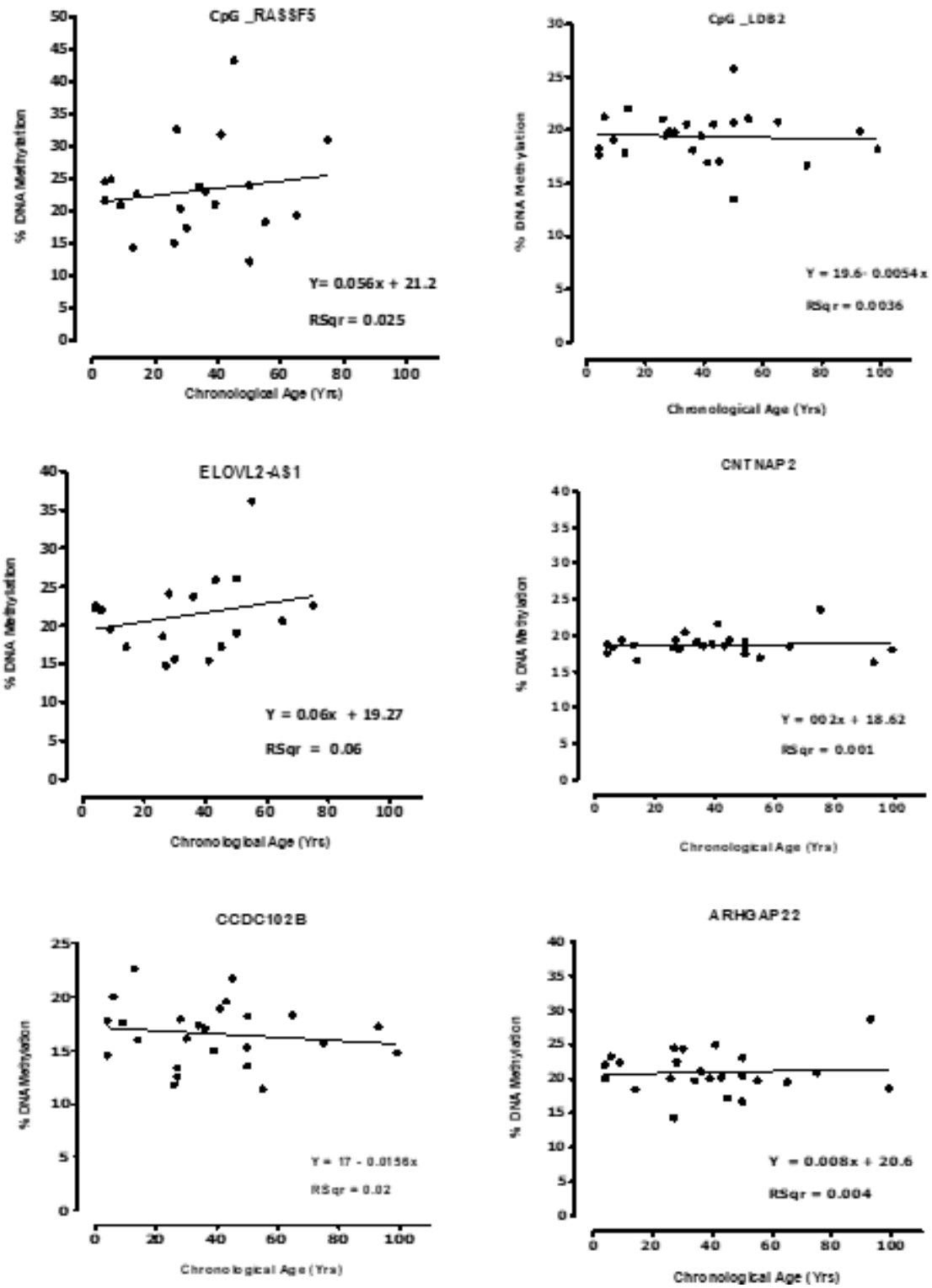


Figure 2: Correlation between Methylation and age at various CpG sites(*EDARADD, ELOVL2 RASSF5, MIR29B2CHG, FHL2, SMC4, TRIM59, LDB2, ELOVL2-AS1, CNTNAP2, ARHGAP22 and CCDC102B*)

4.3. Prediction of Chronological age

The methylation values of 25 samples of all the 12 CpG sites excluding the 7 standards were put in a multiple Linear regression model. None of the p - values obtained was less than 0.05 implying that none of the variables (CpG methylation) had a significant influence on the predictor variable (Age). The overall prediction model was not statistically significant (S- statistic P-value: 0.2758). However, the R^2 value for the prediction model showed a strong correlation between Methylation and age $R^2 = 0.9476$ (Table 3).

Table 6: Regression statistics of the age prediction model with the all the age predictors: *EDARADD, ELOVL2 RASSF5, MIR29B2CHG, FHL2, SMC4, LDB2, CNTNAP2, ARHGAP22 and CCDC102B.* No p-value was less than 0.05(P<0.05) and R^2 was 0.9476.

CpG	P -Value	R Square
EDARADD	0.224	
ELOVL2	0.474	
RASSF5	0.163	
MIR29B2CHG,	0.268	
FHL2	0.882	
SMC4	0.106	
TRIM59	0.811	
LDB2	0.674	
ELOVL2-AS1	0.505	
CNTNAP2	0.676	
CCDC102B	0.682	
ARHGAP22	0.283	
Prediction model	**0.2758**	**0.9476**

4.4. Multiple Linear Regression Model validation

The methylation values of all the 25 samples for the 10 CpG sites were all put in a multiple linear regression model. Data cleaning was then done to remove correlating variables, troublesome data points, define the model's applicability domain and fill in missing data points. The correlation threshold was set at 0.6, CpG sites (TRIM59, LDB2

and ELOVL2.AS1) were removed from the model and outlying data points (0109_1and2_82; 0126_3and4_84; 0131_3and4_84; 0120_3and4_84) were also removed.

This model assessed the statistical significance via p –values at various confidence levels of the remaining CpG sites (*EDARADD, ELOVL2 RASSF5, MIR29B2CHG, FHL2, SMC4, LDB2, CNTNAP2, ARHGAP22 and CCDC102B)* of methylation with respect to age. Overall, a p-value of 0.02976 was obtained which is significant $p < 0.05$) in assessing the predictor variable Age. The measure of goodness of Fit (R values) of the correlation between the model variables (methylation of CpG sites) and the predictor (Age) were also assessed and R^2 of 0.736 and adjusted R^2 of 0.5199 were obtained indicating a strong correlation between chronological age and predicted age. The overall prediction model was highly significant (S- statistic P-value: 0.02976) (Table 4, Figure. 3).

In order to use this model to estimate the age, we formulated an equation that takes into account the variables (CpG methylation) and the respective coefficients i.e., Age = (-0.7805*EDARADD) + (-1.2396*RASSFS) + (1.8686*MIR29B2CHG) + (0.6544*FHL2) + (1.8871*SMC4) + (-1.7493*ELOVL2) + (7.3904*CNTNAP2) + (0.1689*ARHGAP22) + (0.3195*CCDC102B) + (-133.9523). The resulting multiple linear regression model was able to predict donor age with a Mean Absolute Deviation (MAD) of 5.1 years.

Table 7: Regression statistics of the age prediction model with the 9 best age predictors:*EDARADD, ELOVL2 RASSF5, MIR29B2CHG, FHL2, SMC4, CNTNAP2, ARHGAP22 and CCDC102B.*

CpG	P -Value	R Square
EDARADD	0.04288	
RASSF5	0.11064	
MIR29B2CHG,	0.00982	
FHL2	0.32653	
SMC4	0.02100	
CNTNAP2	0.02220	
CCDC102B	0.84720	
ARHGAP22	0.93165	
ELOV2	0.10180	
Prediction model	**0.02976**	**0.736**

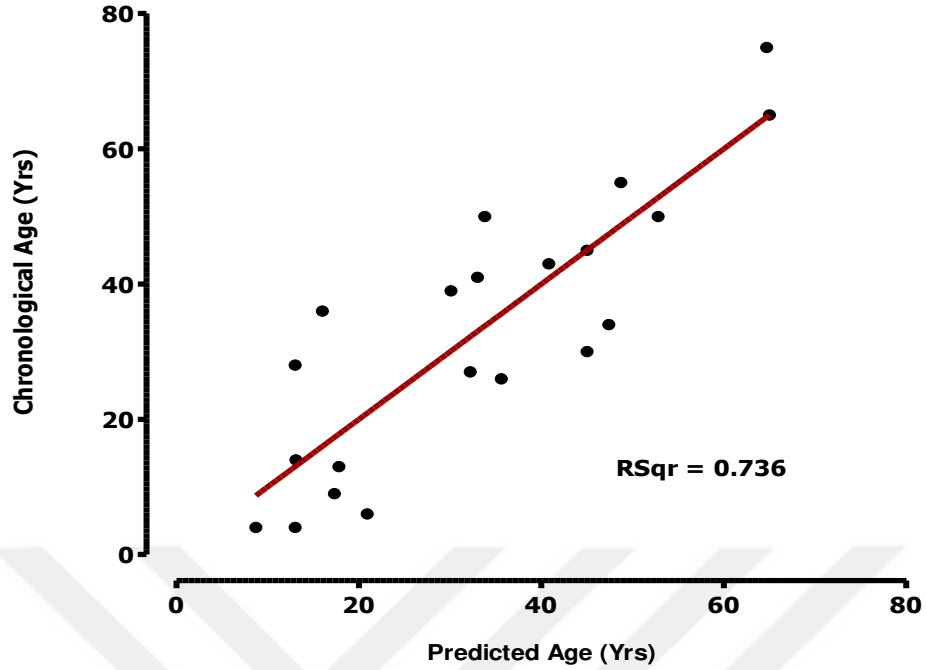


Figure 3: Chronological age versus predicted age for 9 CpG sites were selected for a multiple Linear regression model to predict age. A strong correlation was found between predicted and chronological age (R^2 = 0.736) with a Mean Absolute Deviation (MAD) of 5.1 years. The overall prediction model was highly significant (S-test P-value: 0.02976).

4.5. Cross validation of the Age-Prediction Model

A cross validation of the age prediction model was performed to assess performance of the model on future data sets during which the data set was split randomly into: A training set (used for building an MLR prediction model) and a test set (used to assess the performance of the model). Subsequently, this prediction model is used to predict the age of the removed samples. In this cross validated model, the variables (CpG sites) RASSF5 (p = 0.033), MIR29B2CHG (p = 0.034) and SMC4 (p =0.0323) had a significant influence on the predictor variable (Age) at a 95% confidence level (Table 5). The Multiple R squared and Adjusted R Squared values of the model with this training set of data were 0.9878 and 0.9332 respectively indicating a strong positive correlation between predicted and chronological age (Figure 4). The formulated equation for this age prediction model was Age = (-5.9659*EDARADD) + (-3.6187*RASSFS) + (2.2901*MIR29B2CHG) + (-2.1996*FHL2) + (5.9165*SMC4) + (2.1448*ELOVL2) + (7.1935*CNTNAP2) + (6.8003*ARHGAP22) + (1.3321*CCDC102B) + (-196.2118).

The resulting cross validation regression model was able to predict donor age with a Mean Absolute Deviation (MAD) of 2.3 years.

Table 8: Regression statistics of the cross-validation age prediction with the 9 best age predictors: *EDARADD, ELOVL2 RASSF5, MIR29B2CHG, FHL2, SMC4, CNTNAP2, ARHGAP22 and CCDC102B.*

CpG	P -Value	R Square
EDARADD	0.0388	
RASSF5	0.0606	
MIR29B2CHG,	0.0340	
FHL2	0.1438	
SMC4	0.0323	
CNTNAP2	0.0962	
CCDC102B	0.4206	
ARHGAP22	0.2126	
ELOV2	0.2515	
Prediction model	0.05354	0.9878

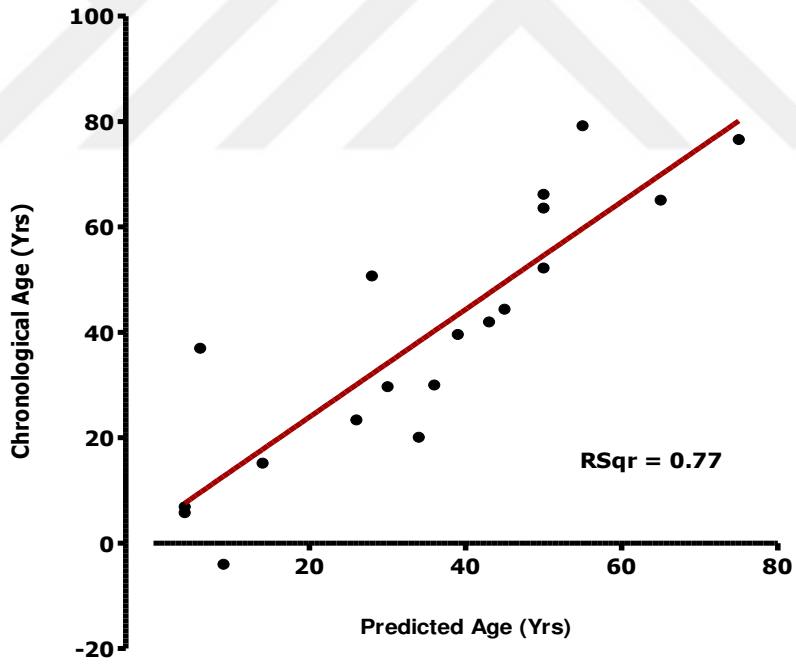


Figure 4: Chronological age versus predicted age for 9 CpG sites were selected for a multiple Linear regression model to predict age. A strong correlation was found between predicted and chronological age (R^2 = 0.9878) with a Mean Absolute Deviation (MAD) of 2.3 years. The overall prediction model was highly significant (S-test P-value: 0.05354).

In addition, to evaluate the accuracy of methylation levels obtained by bisulphite sequencing, we detected the total methylation for each of the seven methylation standards of 0%,10%, 25%, 50%, 75%and 100%. A positive correlation was obtained between total DNA methylation and percentage methylation standards. Bisulphite sequencing resulted in DNA methylation levels that bore a significant linear relationship to expected total methylation levels (Figure 6).

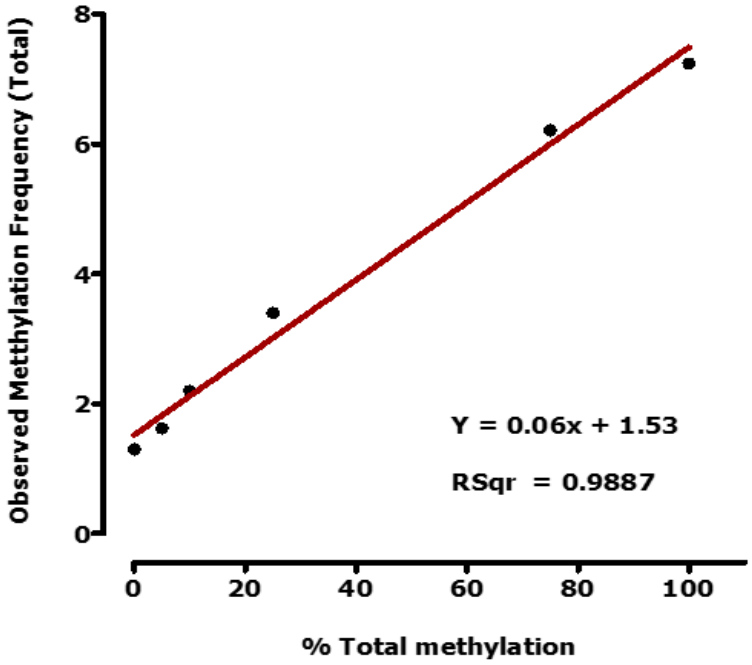


Figure 5: Relationship between % DNA methylation standards and observed total Methylation.

CHAPTER V

DISCUSSIONS AND STUDY LIMITATIONS

5.1. Discussion

In this study, we assessed age-associated methylation for 12 CpG sites, (*EDARADD, ELOVL2 RASSF5, MIR29B2CHG, FHL2, SMC4, TRIM59, LDB2, ELOVL2-AS1, CNTNAP2, ARHGAP22 and CCDC102B*) in 32 buccal samples of living individuals aged 4 -99 years using Next Generation sequencing platform (Table 2). CpG sites which were shown to be highly associated with chronological age in our study were put in a model to predict age. The resulting multiple linear regression model was able to predict donor age with a mean error of 2.3 years demonstrating that the model was accurate. These findings are in consistent with the previous reports by Bekaert et al. (2015b), Eipel et al. (2016), and Jung et al. (2019).

Selection of the CpG sites used in our model was based on short-lists published by other research groups. CpG sites were selected based on their overall rankings in multiple studies and their absolute difference in methylation between young and old individuals. There is a considerable variation in the rankings of age associated CpG sites produced by different studies The reason why not all studies share the same ranking lists is most probably due to the differences in population age ranges, methods, and statistical techniques used. This demonstrates that the selection of CpG sites to build a prediction model is a complex task. In addition, it is important to choose CpG sites that show a high degree of absolute methylation differences as these will result in an increased accuracy of the prediction model.

Aging and DNA methylation are complex processes influenced by multiple factors, which include; difference in sample type used for analysis, genetic and epigenetic variations between different populations and ethnic groups, environmental factors such as exposure to toxins, diet, lifestyle, and socioeconomic factors, individual characteristics, such as sex, genetic makeup, and epigenetic variations, the overall health

status of an individual, chronic diseases and conditions associated with aging, such as cardiovascular disease, diabetes, and cancer, Age-related diseases. The interplay between these factors is still an active area of research. Understanding how these various factors interact can provide valuable insights into the underlying mechanisms of aging and methylation.

Correlation between CpG methylation of all the twelve CpG sites (Table 2) was then assessed using simple linear regression analysis. Simple linear regression revealed a strong correlation (R2>0.3) between CpG methylation and age for only CpG site *MIR29B2CHG* (R^2>0.38) indicating that there is a moderate relationship between the two variables age and CpG methylation. However, low correlation between CpG Methylation and age was revealed for other CpG sites (R^2 < 0.3) (*SMC4, TRIM59, FHL2, EDARADD, ELOVL2, ELOVL2-AS1, RASSF5, CCDC102B, ARHGAP22, CNTNAP2* and *LDB2)* indicating a weak relationship between age and CpG methylation. This could have been because of the small sample size, data quality, material matrix and other confounding variables that were not considered which might not have been considered in the analysis.

A multiple linear regression model built with all the data points from the 12 CpG sites and the 25 samples, it was observed that none of the p values was less than 0.05 between the variables (CpG methylation) and the predictor variable (Age) for all the 12 CpG sites indicating that the observed relationship is not statistically significant implying that the evidence is insufficient to establish a significant relationship between age and CpG methylation. However, an R square value of was obtained indicating a very strong relationship between age and DNA methylation and suggests that age is a highly influential factor in determining DNA methylation patterns. One aspect for the p values above 0.05 could have been due to the small sample size, material matrix hence there is a need for further analysis, validation, and consideration of other factors that may affect DNA methylation.

Upon data clean up to remove correlating variable and outlying data points, 10CpG sites; *EDARADD, ELOVL2 RASSF5, MIR29B2CHG, FHL2, SMC4, LDB2, CNTNAP2, ARHGAP22 and CCDC102B* had a p value less than 0.05 indicating that they were highly associated with chronological age and statistically significant at various

confidence levels and were therefore considered in the age prediction model. The overall prediction model explains 95% of the variation in age and DNA methylation of (S-statistic P-value: 0.02976, R square : 0.736) was obtained indicating a statistically significant relationship between chronological age and predicted age. This finding suggests that the predictive model used captures some meaningful information associated with age and was able to predict donor age with a mean error of 5.1 years. Prior research has suggested that methylation patterns can be influenced by age, with certain genes showing changes in methylation levels as individuals grow older (Bekaert et al., 2015a; Zbieć- Piekarska et al., 2015a, 2015b; Hamano et al., 2016; Pfeifer et al., 2020). The correlation between DNA methylation and age for *EDARADD, ELOVL2 RASSF5, MIR29B2CHG, FHL2, SMC4, TRIM59, LDB2, ELOVL2-AS1, CNTNAP2, ARHGAP22 and CCDC102B*) has previously been reported by several research groups and aligns with existing knowledge supporting the notion that age is associated with methylation changes (Aliferi et al; 2022, Hong et al; 2017, Reynolds et al.,2015, Koch, C. M., & Wagner, W. (2011), Horvath, S. (2013), Aliferi et al., 2018). In contrast to previously published results, we did not observe a significant correlation between predicted age and chronological age for CpG sites TRIM59 and ELOVL2.AS1, TRIM59 and ELOVL2.AS1 hence not further analyzed.

Cross validation of the age prediction model was done to assess performance on future data sets during which the data set was split randomly into training set and test set. In this cross validated model, the variables (CGp sites) RASSF5, MIR29B2CHG and SMC4 were shown have a significant influence on the predictor variable (Age) at a 95% confidence level (Table 5) and a strong positive correlation between predicted and chronological age (Figure 4). The data from these cross validations demonstrated that the prediction model was very robust and highly reproducible. These findings are in consistent with the previous reports by Bekaert et al. (2015b), Eipel et al. (2016), and Jung et al. (2019).

Furthermore, a significant linear relationship obtained from the methylation standards used and the expected outcome provide strong evidence supporting the accuracy and consistency of the bisulfite sequencing technique in capturing methylation levels. This demonstrates the reliability and precision of the technique in measuring

methylation patterns obtained versus expected methylation of known quantities of methylated to unmethylated DNA with R square value of 0.9887 indicating a strong correlation between expected methylation and observed methylation.

5.2. Study limitations

The sample size used in the study might have influenced the reliability, accuracy and generalizability of the results as this could introduce variability and increase the risk of errors. This reduces the statistical power of the study, making it more challenging to detect true effects or relationships.

No information regarding the sampled individuals' health, genetic history and lifestyle were collected, as this information would not be available to a 'standard' forensic scenario where the primary focus is on analyzing and interpreting the physical evidence related to a crime. The absence of such information can limit the ability to fully understand and interpret the observed DNA methylation patterns or their relationship with age.

Limiting the sample tissue type to only buccal samples, it is known that different tissue types have different age-related DNA methylation patterns. In the context of developing multi-tissue APMs, it is essential to include predictor variables from various tissue types, such as blood, saliva, bones, and teeth. Incorporating multiple tissues allows for a more comprehensive representation of age-related DNA methylation patterns and improves the accuracy of age estimation.

NGS platforms offers several advantages, such as high-throughput sequencing and the ability to capture comprehensive DNA methylation profiles, however , the high cost of sample preparation, sequencing, and data analysis was a limiting factor to the sample size used in this study.

CHAPTER VI

CONCLUSIONS AND RECOMMENDATIONS

6.1. Conclussion And Reccomendations

In conclusion, the study demonstrated a strong correlation between DNA methylation and age, supporting the existing understanding of an association between these variables. However, further evaluation is required to assess the accuracy and validity of the predictive model used for age estimation. To determine the true accuracy and reliability of the predictive model, additional evaluations, validations, and comparisons with other age estimation methods are necessary.

The size and representativeness of the dataset used for analysis play a vital role in determining the robustness and generalizability of the results. Larger and more diverse datasets tend to provide more reliable and generalizable findings. Therefore, it is essential to consider the larger sample size and more diverse sample sets, including individuals from all age ranges in order to improve the reliability, accuracy, and applicability of these methods.

It is also important to consider the fact that the correlation coefficient is specific to the dataset and population under study. It is crucial to assess whether the observed correlation holds true in other populations or if there are specific characteristics of the dataset that might limit its generalizability.

The presence of outliers or influential data points can have a significant impact on the correlation coefficient and the overall analysis as they can distort the relationship between variables and affect the strength of the correlation. It is important to identify and understand the impact of such observations on the overall correlation and consider their potential effects on the validity and generalizability of the findings.

Future research should aim to explore additional variables and potential confounding factors such as environmental exposures, genetic variations, lifestyle factors,

and underlying health conditions to enhance our understanding of the relationship between age and DNA methylation. This knowledge can contribute to the development of more accurate age prediction models that can provide valuable insights into the complex interplay between these variables.

DNA methylation patterns can vary across different tissues, which means that Age Prediction Models (APMs) developed for one tissue may not be directly applicable to other types of tissues. Therefore, it is crucial to test the development of APMs for various tissue types as this will enable accurate age prediction based on tissue-specific DNA methylation patterns, further advancing our understanding of age-related changes in different tissues.

Forensic samples collected from crime scenes often have low quantities of DNA and can be of poor quality, which poses significant challenges for DNA methylation analysis. One of the major hurdles in DNA methylation analysis is the requirement for bisulfite conversion, but results in DNA loss and degradation and this can reduce the amount of available DNA for analysis, potentially compromising the accuracy and reliability of DNA methylation age estimation.

Additionally, the limited quantity and quality of forensic samples may also result in incomplete bisulfite conversion, leading to inaccurate methylation measurements. Therefore, it is necessary to consider alternative approaches or technologies that require smaller amounts of DNA or are less sensitive to DNA degradation.

7. REFERENCES

1. Adaikalakoteswari A, Finer S, Voyias PD, McCarthy CM, Vatish M, Moore J, Smart-Halajko M, Bawazeer N, Al-Daghri NM, McTernan PG, et al.: Vitamin B12 insufficiency induces cholesterol biosynthesis by limiting s-adenosylmethionine and modulating the methylation of SREBF1 and LDLR genes; Clin Epigenetics 7:14; 2015.
2. Alghanim H, Antunes J, Silva DSBS, Alho CS, Balamurugan K, McCord B. Detection and evaluation of DNA methylation markers found at SCGN and KLF14 loci to estimate human age. Forensic Sci Int Genet. 2017; 31:81-88. DOI: 10.1016/j.fsigen.2017.07.011.
3. Aliferi A, Ballard D, Gallidabino MD, Thurtle H, Barron L, Syndercombe Court D. DNA methylation-based age prediction using massively parallel sequencing data and multiple machine learning models. Forensic Sci Int Genet. 2018; 37:215-226. https://doi.org/10.1016/j.fsigen.2018.09.003.
4. Alsaleh, H.; McCallum, N.A.; Halligan, D.L.; Haddrill, P.R. A multi-tissue age prediction model based on DNA methylation analysis. Forensic Sci. Int. Genet. Suppl. Ser. 2017, 6, e62–e64.
5. Álvarez-Dios J, Ballard D, Court DS, Branicki W, Carracedo Á, Lareu MV. "A Comparison of Forensic Age Prediction Models Using Data from Four DNA Methylation Technologies." ("Wojciech BRANICKI | Professor | Professor | Jagiellonian University ...") Front Genet. 2020;19(11):932.
6. Antequera, F. and A. Bird, Number of CpG islands and genes in human and mouse. Proceedings of the National Academy of Sciences of the United States of America, 1993. 90(24): p. 11995-11999.
7. Armstrong, L. (2013). Epigenetics; Garland Science: New York. *NY, USA*.
8. Bacalini MG, Deelen J, Pirazzini C, De Cecco M, Giuliani C, Lanzarini C, Ravaioli F, Marasco E, van Heemst D, Suchiman HED, Slieker R, Giampieri E, Recchioni R, Marcheselli F, Salvioli S, Vitale G, Olivieri F, Spijkerman AMW, Dollé MET, Sedivy JM, Castellani G, Franceschi C, Slagboom PE, Garagnani P. Systemic Age-Associated DNA Hypermethylation of ELOVL2 Gene: In Vivo and In Vitro Evidences of a Cell Replication Process. J Gerontol A Biol Sci Med Sci. 2017;72(8):1015-1023. doi: 10.1093/gerona/glw185.
9. Baccino E, Cunha E, Cattaneo C. Aging the dead and the living. Elsevier. 2013;42-48.
10. Baccino E, Schmitt A: Determination of adult age at death in the forensic context; In Schmitt A, Cunga E, Pinheiro J (Eds): Forensic Anthropology and Medicine: Complementary Sciences from Recovery to Cause of Death; Humana Press: Totiwa, NJ; 2006.
11. Baccino E, Ubelaker DH, Hayek L-AC, Zerilli A. Evaluation of Seven Methods of Estimating Age at Death from Mature Human Skeletal Remains. J Forensic Sci. 1999;931-936.
12. Baccino, E., Sinfield, L., Colomb, S., Baum, T. P., & Martrille, L. (2014). The two-step procedure (TSP) for the determination of age at death of adult human remains in forensic cases. *Forensic science international*, *244*, 247-251.

13. Barres R, Yan J, Egan B, Treebak JT, Rasmussen M, Fritz T, Caidahl K, Krook A, O'Gorman DJ, Zierath JR: Acute exercise remodels promoter methylation in human skeletal muscle; Cell Metab 15:405; 2012.
14. Cantone I, Fisher AG: Epigenetic programming and reprogramming during development; Nat Sructural Mol Biol 20:282; 2013. ("Transgenerational Epigenetic Contributions to Stress Responses ... - PLOS")
15. Barricarte A, Fernandez AM, Travis RC, Vineis P, Muller DC, Riboli E, Gunter M, Romieu I, Herceg Z. DNA methylome analysis identifies accelerated epigenetic ageing associated with postmenopausal breast cancer susceptibility. Eur J Cancer. 2017; 75:299–307.
16. Bartel DP, Lee R, Feinbaum R. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function Genomics: The miRNA Genes. 2004; 116:281–97.
17. Bartel DP. Review MicroRNAs: Target Recognition and Regulatory Functions. 2009;215–33.
18. Bekaert B, Kamalandua A, Zapico SC, Van de Voorde W, Decorte R. A selective set of DNAmethylation markers for age determination of blood, teeth and buccal samples. Forensic Sci Int. Genet Suppl Ser 2015b;5: e144-e145.
19. "Bekaert B, Kamalandua A, Zapico SC, Van de Voorde W, Decorte R. Improved age determination of blood and teeth samples using a selected set of DNA methylation markers." ("Download .nbib") Epigenetics. 2015a; 10:922–930.
20. Bekaert, B., Kamalandua, A., Zapico, S. C., Van de Voorde, W., & Decorte, R. (2015). Improved age determination of blood and teeth samples using a selected set of DNA methylation markers. *Epigenetics*, *10*(10), 922-930.
21. Bekaert, B., Kamalandua, A., Zapico, S. C., Van De Voorde, W., and Decorte, R. (2015). "Improved age determination of blood and teeth samples using a selected set of DNA methylation markers." ("A selective set of DNA-methylation markers for age determination of ...") Epigenetics 10, 922–930. doi: 10.1080/ 15592294.2015.1080413
22. Bergsma T, Rogaeva E. DNA Methylation Clocks and Their Predictive Capacity for Aging Phenotypes and Healthspan. Neurosci Insights. 2020;21(15):2633105520942221
23. Bocklandt S, Lin W, Sehl ME, Sanchez FJ, Sinsheimer JS, Horvath S, Vilain E. Epigenetic predictor of age. PLoS One. 2011;6(6):1-6.
24. Bocklandt, S., Lin, W., Sehl, M., Sánchez, F., Sinsheimer, J., Horvath, S., et al. (2011). Epigenetic predictor of age. PLoS One 6:e14821. doi: 10.1371/journal. pone.0014821
25. "Breitling LP, Saum KU, Perna L, Schottker B, Holleczek B, Brenner H. Frailty is associated with the epigenetic clock but not with telomere length in a German cohort." ("Comparative validation of three DNA methylation algorithms of ageing ...") Clin Epigenetics. 2016; 8:21. doi: 10.1186/s13148-016-0186-5.
26. Catalanotto C, Cogoni C, Zardo G. MicroRNA in control of gene expression: An overview of nuclear functions. Int J Mol Sci. 2016;17(10).

27. Chen BH, Marioni RE, Colicino E, Peters MJ, Ward-Caviness CK, Tsai PC, Roetker NS, Just AC, Demerath EW, Guan W, Bressler J, Fornage M, Studenski S, Vandiver AR, Moore AZ, Tanaka T, Kiel DP, Liang L, Vokonas P, Schwartz J, Lunetta KL, Murabito JM, Bandinelli S, Hernandez DG, Melzer D, Nalls M, Pilling LC, Price TR, Singleton AB, Gieger C, Holle R, Kretschmer A, Kronenberg F, Kunze S, Linseisen J, Meisinger C, Rathmann W, Waldenberger M, Visscher PM, Shah S, Wray NR, McRae AF, Franco OH, Hofman A, Uitterlinden AG, Absher D, Assimes T, Levine ME, Lu AT, Tsao PS, Hou L, Manson JE, Carty CL, LaCroix AZ, Reiner AP, Spector TD, Feinberg AP, Levy D, Baccarelli A, van Meurs J, Bell JT, Peters A, Deary IJ, Pankow JS, Ferrucci L, Horvath S. DNA methylation-based measures of biological age: metaanalysis predicting time to death. Aging (Albany NY). 2016;8(9):1844-1865. doi: 10.18632/aging.101020.

28. Chen BH, Marioni RE, Colicino E, Peters MJ, Ward-Caviness CK, Tsai P, Roetker NS, Just AC, Demerath EW, Guan W, et al.: DNA methylation-based measures of biological age: Meta-analysis predicting time to death; Aging (Albany, NY) 8:1; 2016.
De Tobel J, Hillewig E, Verstraete K: Forensic age estimation based on magnetic resonance imaging of third molars: converting 2D staging into 3D staging; Ann Hum Biol 44:121; 2017.

29. Chiti E, Maiese A, Turillazzi E, Spinetti I, Area C, Chiara OS, et al. MicroRNAs: an update of applications in forensic medicine. 2020;1–30.

30. Cho S, Ge J, Seo SB, Kim K, Lee HY, Lee SD. Age estimation via quantification of signal-joint T cell receptor excision circles in Koreans. Leg Med. 2014;16(3):135-8. doi: 10.1016/j.legalmed.2014.01.009.

31. Cho S, Jung SE, Hong SR, Lee EH, Lee JH, Lee SD, Lee HY. Independent validation of DNAbased approaches for age prediction in blood. Forensic Sci Int Genet. 2017;29:250-256. doi: 10.1016/j.fsigen.2017.04.020.

32. Christoph Meissner, Stefanie Ritz-Timme, Molecular pathology and age estimation, Forensic Science International, Volume 203, Issues 1–3,2010, Pages 34-43, ISSN 0379-0738,https://doi.org/10.1016/j.forsciint.2010.07.010.(https://www.sciencedirect.com/science/article/pii/S0379073810003488)

33. "Ciccarone F, Tagliatesta S, Caiafa P, Zampieri M. DNA methylation dynamics in aging: how far are we from understanding the mechanisms?" ("Heterochromatin: an epigenetic point of view in aging - PMC") Mech Ageing Dev. 2018;174:3-17.

34. Correia Dias, H.; Cordeiro, C.; Corte Real, F.; Cunha, E.; Manco, L. Age estimation based on DNA methylation using blood samples from deceased individuals. J. Forensic Sci. 2019, 65, 465–470.

35. Correia Dias, H.; Cordeiro, C.; Pereira, J.; Pinto, C.; Corte Real, F.; Cunha, E.; Manco, L. DNA methylation age estimation in blood samples of living and deceased individuals using a multiplex SNaPshot assay. Forensic Sci. Int. 2020, 311, 110267.

36. Correia Dias, H.; Corte Real, F.; Cunha, E.; Manco, L. DNA methylation age estimation from human bone and teeth. Aust. J. Forensic Sci. 2021, Ahead-of-print. 1–14.

37. Correia Dias, H.; Cunha, E.; Corte Real, F.; Manco, L. Age prediction in living: Forensic epigenetic age estimation based on blood samples. Leg. Med. 2020, 47, 101763.

38. Courts C, Madea B. Micro-RNA – A potential for forensic science? §. Forensic Sci Int [Internet]. 2010;203(1–3):106–11. Available from: http://dx.doi.org/10.1016/j.forsciint.2010.07.002

39. D. Glass, A. Vinuela, M.N. Davies, A. Ramasamy, L. Parts, D. Knowles, A.A. Brown, A.K. Hedman, K.S. Small, A. Buil, E. Grundberg, A.C. Nica, P. Meglio, F.O. Nestle, M. Ryten, U.K.B.E. c. the, T. c. the Mu, R. Durbin, M.I. McCarthy, P. Deloukas, E.T. Dermitzakis, M.E. Weale, V. Bataille, T.D. Spector, Gene expression changes with age in skin, adipose tissue, blood and brain, Genome Biol. 14 (7) (2013) R75.

40. "Daunay A, Baudrin LG, Deleuze JF, How-Kit A. Evaluation of six blood-based age prediction models using DNA methylation analysis by pyrosequencing." ("Centenarians consistently present a younger epigenetic age than their ...") Sci Rep. 2019;9(1):8862. doi: 10.1038/s41598-019-45197-w.

41. Deans C, Maggert KA. What do you mean, "epigenetic"? Genetics. 2015 Apr;199(4):887-96. doi: 10.1534/genetics.114.173492. PMID: 25855649; PMCID: PMC4391566

42. Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Develop. 2011; 25:1010–1022.

43. Dugué P-A, Bassett JK, Joo JE, Jung C-H, Wong EM, Margarita Moreno-Betancur M, Schmidt D, Makalic E, Li S, Severi G, Hodge AM,Buchanan DD, English DR, Hopper JL, Southey MC, Giles GG and Milne RL. DNA methylation-based biological aging and cancer risk and survival:Pooled analysis of seven prospective studies. Int J Cancer. 2018; 142:1611–1619.

44. Ehrlich, M., et al., Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. Nucleic Acids Research, 1982. 10(8): p. 2709-2721.

45. Eipel M, Mayer F, Arent T, Ferreira MRP, Birkhofer C, Gerstenmaier U, Costa IG, Ritz-Timme S, Wagner W. Epigenetic age predictions based on buccal swabs are more precise in combination with cell type-specific DNA methylation signatures. Aging. 2016;8(5):1034-48.

46. Etheridge A, Lee I, Hood L, Galas D, Wang K. Extracellular microRNA: A new source of biomarkers. Mutat Res - Fundam Mol Mech Mutagen [Internet]. 2011;1–6. Available from: http://dx.doi.org/10.1016/j.mrfmmm.2011.03.004

47. evaluation of age-correlated DNA methylation markers for forensic use. Forensic Sci Int Genet. 2016; 23:64-70.

48. Felsenfeld, G. (2014). A brief history of epigenetics. *Cold Spring Harbor perspectives in biology*, *6*(1), a018200.

49. Ferreira MT, Coelho C, Makhoul C, Navega D, Gonçalves D, Cunha E, Curate F. New data about the 21st Century Identified Skeletal Collection (University of Coimbra, Portugal). Int J Legal Med. 2020. doi: 10.1007/s00414-020-02399-6.
50. Ferreira MT, Vicente R, Navega D, Gonçalves D, Curate F, Cunha E. A new forensic collection housed at the University of Coimbra, Portugal: The 21st century identified skeletal collection.Forensic Sci Int. 2014;245:202.e1-5. doi: 10.1016/j.forsciint.2014.09.021.
51. Field AE, Robertson NA, Wang T, Havas A, Ideker T, and Adams PD. DNA Methylation Clocks in Aging: Categories, Causes, and Consequences. Molecular Cell. 2018; 71:882-895.
52. Fiorito G, McCrory C, Robinson O, Carmeli C, Rosales CO, Zhang Y, Colicino E, Dugué PA, Artaud F, McKay GJ, Jeong A, Mishra PP, Nøst TH, Krogh V, Panico S, Sacerdote C, Tumino R, Palli D, Matullo G, Guarrera S, Gandini M, Bochud M, Dermitzakis E, Muka T, Schwartz J,
53. Fleckhaus, J., Freire-Aradas, A., Rothschild, M. A., & Schneider, P. M. (2017). "Impact of genetic ancestry on chronological age prediction using DNA methylation analysis." ("Impact of genetic ancestry on chronological age ... - ScienceDirect") *Forensic Science International: Genetics Supplement Series*, *6*, e399-e400.
54. Florath I, Butterbach K, Müller H, Bewerunge-Hudler M, Brenner H. Cross-sectional and longitudinal changes in DNA methylation with age: an epigenome-wide analysis revealing over 60 novel age-associated CpG sites. Hum Mol Genet. 2014;23 (5):1186-1201.
55. Freire-Aradas A, Phillips C, Girón-Santamaría L, Mosquera- Miguel A, Gómez-Tato A, Casares de Cal MÁ, Álvarez-Dios J, Lareu MV. Tracking age-correlated DNA methylation markers in the young. Forensic Sci Int Genet. 2018; 36:50–59. https:// doi.org/10.1016/J.FSIGEN.2018.06.011.
56. Freire-Aradas A, Phillips C, Lareu MV. Forensic individual age estimation with DNA: from initial approaches to methylation tests. Forensic Sci Rev. 2017; 29:121–144.
57. Freire-Aradas A, Phillips C, Lareu MV. "Forensic individual age estimation with DNA: From initial approaches to methylation tests." ("Age Estimation with DNA: From Forensic DNA Fingerprinting to Forensic ...") Forensic Sci Rev. 2017 Jul;29(2):121-144. PMID: 28691915
58. Freire-Aradas A, Phillips C, Mosquera-Miguel A, Girón-Santamaría L, Gómez-Tato A, Casares de Cal M, Álvarez-Dios J, Ansede-Bermejo J, Torres-Español M, Schneider PM, Pośpiech E, Branicki W, Carracedo Á, Lareu MV. Development of a methylation marker set for forensic age estimation using analysis of public methylation data and the Agena Bioscience EpiTYPER system. Forensic Sci Int Genet. 2016; 24:65-74.
59. Freire-Aradas A, Pośpiech E, Aliferi A, Girón-Santamaría L, Mosquera-Miguel A, Pisarek A, Ambroa-Conde A, Phillips C, Casares de Cal MA, Gómez-Tato A, Spólnicka M, Woźniak A,
60. Freire-Aradas, A., Phillips, C., & Lareu, M. V. (2017). "Forensic individual age estimation with DNA: from initial approaches to methylation tests." ("Age Estimation with DNA: From Forensic DNA Fingerprinting to Forensic ...") *Forensic science review*, *29*(2).

61. Freire-Aradas, A., Phillips, C., and Lareu, M. V. (2017). Forensic individual age estimation with DNA: from initial approaches to methylation tests. Forens. Sci. Rev. 29, 121–144.

62. Freire-Aradas, A., Phillips, C., Girón-Santamaría, L., Mosquera-Miguel, A., Gómez-Tato, A., Casares de Cal, M. Á, et al. (2018). Tracking age-correlated DNA methylation markers in the young. Foren. Sci. Int. Genet. 36, 50–59. doi: 10.1016/j.fsigen.2018.06.011

63. Freire-Aradas, A., Phillips, C., Mosquera-Miguel, A., Girón-Santamaría, L., Gómez-Tato, A., Casares De Cal, M., et al. (2016). "Development of a methylation marker set for forensic age estimation using analysis of public methylation data and the agena bioscience EpiTYPER system." ("Age Estimation with DNA: From Forensic DNA Fingerprinting to Forensic ...") Forens. Sci. Int. Genet. 24, 65–74. doi: 10.1016/j.fsigen.2016.06.005

64. Frontiers in Genetics | www.frontiersin.org 11 August 2020 | Volume 11 | Article 932 fgene-11-00932 August 17, 2020 Time: 16:35 # 12 Freire-Aradas et al. Comparison of DNA Methylation Technologies

65. Garagnani P, Bacalini MG, Pirazzini C, Gori D, Giuliani C, Mari D, Di Blasio AMD, Gentilini D, Vitale G, Collino S, Rezzi S, Castellani G, Capri M, Salvioli S, and Franceschi C. Methylation of ELOVL2 gene as a new epigenetic marker of age. Aging Cell. 2012; 11:1132–1134.

66. Giuliani C, Cilli E, Bacalini MG, Pirazzini C, Sazzini M, Gruppioni G, Franceschi C, Garagnani P, Luiselli D. Inferring chronological age from DNA methylation patterns of human teeth. Am J Phys Anthropol 2016;159(4):585-95. doi: 10.1002/ajpa.22921.

67. Goel N, Karira P, Garg VK. Role of DNA methylation in human age prediction, Mech. Ageing Dev. 2017;166:33–41.

68. Gopalan S, Gaige J, Henn BM. DNA methylation-based forensic age estimation in human bone. BioRxiv. 2019. https://doi.org/10.1101/801647.

69. Haeusler M, Haas C, Losch S, Moghaddam N, Villa IM, Walsh S, Kayser M, Seiler R, Ruehli F, Janosa M, et al.: Multidisciplinary identification of the controversial Freedom Fighter Jorg Jenatsch, assassinated 1639 in Chur, Switzerland; PLoS One 11(12): e0168014; 2016.

70. Hagerty SL, Bidwell LC, Harlaar N, Hutchison KE: An exploratory association study of alcohol use disorder and DNA methylation; Alcohol Clin Exp Res 40:1633; 2016.

71. Hamano Y, Manabe S, Morimoto C, Fujimoto S, Tamaki K. Forensic age prediction for saliva samples using methylation-sensitive high-resolution melting: exploratory application for cigarette butts. Sci Rep. 2017;7(1):10444. doi: 10.1038/s41598-017-10752-w.

72. Hamano, Y., Manabe, S., Morimoto, C., Fujimoto, S., Ozeki, M., & Tamaki, K. (2016). Forensic age prediction for dead or living samples by use of methylation-sensitive high -resolution melting. *Legal Medicine*, *21*, 5-10.

73. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, Klotzle B, Bibikova M, Fan J-B, Gao Y, Deconde R, Chen M, Rajapakse I, Friend S, Ideker T, Zhang K. Genome-wide methylation profiles reveal quantitative views of human aging

rates. (“CSIRO PUBLISHING | Animal Production Science”) Mol Cell. 2013; 49:359–367.

74. Hanson EK, Lubenow H, Ballantyne J. Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs. Anal Biochem [Internet]. 2009;387(2):303–14. Available from: http://dx.doi.org/10.1016/j.ab.2009.01.037.

75. "He Y, Ecker JR. Non-CG methylation in the human genome." (“Dynamic DNA Methylation Regulates Levodopa-Induced Dyskinesia”) Annu Rev Genomics Hum Genet. 2015; 16:55.

76. Hong, S. R., Jung, S. E., Lee, E. H., Shin, K. J., Yang, W. I., and Lee, H. Y. (2017). DNA methylation-based age prediction from saliva: high age predictability by combination of 7 CpG markers. Foren. Sci. Int. Genet. 29, 118–125. doi: 10.1016/j.fsigen.2017.04.006

77. Horvath S, Oshima J, Martin GM, Lu AT, Quach A, Cohen H, Felton S, Matsuyama M, Lowe D, Kabacik S, Wilson JG, Reiner AP, Maierhofer A, Flunkert J, Aviv A, Hou L, Baccarelli AA, Li Y, Stewart JD, Whitsel EA, Ferrucci L, Matsuyama S, Raj K. Epigenetic clock for skin and blood cells applied to Hutchinson Gilford Progeria Syndrome and ex vivo studies. Aging. 2018;10(7):1758-1775.

78. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nat Rev Genet. 2018; 19:371–384. https://doi.org/10.1038/s41576-018-0004-3.

79. Horvath S, Ritz BR. Increased epigenetic age and granulocyte counts in the blood of Parkinson's disease patients. Aging (Albany NY). 2015;7(12):1130-42. doi: 10.18632/aging.100859.

80. Horvath S, V. Mah V, Lu AT, Woo JS, Choi OW, Jasinska AJ, Riancho JA, Tung S, Coles NS, Braun J, Vinters HV, Coles LS. (“Download .nbib”) The cerebellum ages slowly according to the epigenetic clock. Aging. 2015; 7:294–306.

81. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013;14(10):R115.

82. Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome biology*, *14*(10), 1-20.

83. Horvath, S. (2013). DNA methylation age of human tissues and cell types. Genome Biol. 14:R115.

84. Horvath, S., & Raj, K. (2018). "DNA methylation-based biomarkers and the epigenetic clock theory of ageing." (“DNA Methylation Clocks and Their Predictive Capacity for Aging ...”) *Nature Reviews Genetics*, *19*(6), 371-384.

85. Huang Y, Yan J, Hou J, Fu X, Li L, Hou Y. Developing a DNA methylation assay for human age prediction in blood and bloodstain. Forensic Sci Int Genet. 2015; 17:129-136 http://dx.doi.org/10.1016/j.fsigen.2015.05.007.

86. Illingworth R, Kerr A, Desousa D, Jørgensen H, Ellis P, Stalker J, Jackson D, Clee C, Plumb R, Rogers J, Humphray S, Cox T, Langford C, Bird A. A novel CpG island set identifies tissuespecific methylation at developmental gene loci. PLoS Biol. 2008; 6: e22.

87. Illingworth RS, Gruenewald-Schneider U, Webb S, Kerr ARW, James KD, Turner DJ, Smith C, Harrison DJ, Andrews R, Bird AP. Orphan CpG Islands Identify Numerous Conserved Promoters and the Mammalian Genome. PLoS Genetics. 2010;6(9): e1001134.
88. Illumina: MethylationEPIC Beadchip; https://www.illumina.com/products/by-type/microarray-kits/infinium-methylation-epic.html (Accessed May 25, 2017). Jaenisch R, Bird A: Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals; Nat Genet 33 Suppl:245; 2003. roles in metabolism; Prog Lipid Res 45:237; 2006.
89. Jiang S, Guo Y. Epigenetic Clock: DNA Methylation in Aging. Stem Cells Int. 2020; 2020:1047896. doi: 10.1155/2020/1047896.
90. Johannessen M, Moler S, Hansen T, Moens U, Van Ghelue M: The multifunctional roles of the four-and-a-half-LIM only protein FHL2; Cell Mol Life Sci 63:268; 2006. ("TAB182 aggravates progression of esophageal squamous cell ... - Nature") Jones MJ, Goodman SJ, Kobor MS: DNA methylation and healthy human aging; Aging Cell 14:924; 2015.
91. Johansson, Å, Enroth, S., and Gyllensten, U. (2013). Continuous aging of the human DNA methylome throughout the human lifespan. PLoS One 8:e67378. doi: 10.1371/journal.pone.0067378
92. Jones, P. A. (2012). "Functions of DNA methylation: islands, start sites, gene bodies and beyond." ("[PDF] Functions of DNA methylation: islands, start sites, gene bodies ...") Nat. Rev. Genet. 13, 484–492. doi: 10.1038/nrg3230
93. Jung M, Pfeifer GP: Aging and DNA methylation; BMC Biol 13:7; 2015. Lewis ME, Flavel A: Age assessment of child skeletal remains in forensic contexts; In Schmitt A, Cunha E, Pinheiro J (Eds): Forensic Anthropology and Medicine: Complementary Sciences from Recovery to Cause of Death; Humana Press: Totiwa, NJ; 2006.
94. Jung SE, Lim SM, Hong SR, Lee EH, Shin KJ, Lee HY. DNA methylation of the ELOVL2, FHL2, KLF14, C1orf132/MIR29B2C, and TRIM59 genes for age prediction from blood, saliva, and buccal swab samples. ("Age estimation using DNA methylation technique in forensics: a ...") Forensic Sci Int Genet. 2019; 38:1-8.
95. Jung SE, Shin KJ, Lee HY. DNA methylation-based age prediction from various tissues and body fluids. BMB Rep. 2017;50(11):546-553.
96. Jung, S. E., Lim, S. M., Hong, S. R., Lee, E. H., Shin, K. J., and Lee, H. Y. (2019). "DNA methylation of the ELOVL2, FHL2, KLF14, C1orf132/MIR29B2C, and TRIM59 genes for age prediction from blood, saliva, and buccal swab samples." ("DNA methylation of the ELOVL2, FHL2, KLF14, C1orf132/MIR29B2C, and ...") Foren. Sci. Int. Genet. 38, 1–8. doi: 10.1016/j.fsigen.2018.09.010
97. Jung, S.E.; Lim, S.M.; Hong, S.R.; Lee, E.H.; Shin, K.J.; Lee, H.Y. DNA methylation of the ELOVL2, FHL2, KLF14, C1orf132/MIR29B2C, and TRIM59 genes for age prediction from blood, saliva, and buccal swab samples. Forensic Sci. Int. Genet. 2019, 38, 1–8.

98. Kader F, Ghai M. DNA methylation and application in forensic sciences. Forensic Sci Int. 2015; 249:255–265.
99. "Kucharczyk K, Płoski R, Branicki W. Development of a forensically useful age prediction method based on DNA methylation analysis." ("An integrative module analysis of DNA methylation landscape in aging") Forensic Sci Int Genet. 2015b; 17:173–179.
100. Lagos-quintana M, Lagos-quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of Novel Genes Coding for Small Expressed RNAs. 2012;853(2001).
101. Larsen F, Gundersen G, Lopez R, Prydz H. CpG Islands as Gene Markers in the Human Genome." ("Mining of simple sequence repeats (SSRs) loci and development of ... - PLOS") Genomics. 1992; 13:1095-1107.
102. Lau NC, Lim LP, Weinstein EG, Bartel DP. An Abundant Class of Tiny RNAs with Probable Regulatory Roles in Caenorhabditis elegans. 2001;294(October):858–62.
103. Lee RC. An Extensive Class of Small RNAs in Caenorhabditis elegans. 2010;862(2001).
104. Lee, H. Y., Jung, S. E., Oh, Y. N., Choi, A., Yang, W. I., and Shin, K. J. (2015). Epigenetic age signatures in the forensically relevant body fluid of semen: a preliminary study. Foren. Sci. Int. Genet. 19, 28–34. doi: 10.1016/j.fsigen.2015. 05.014
105. Li Y, Tollefsbol TO. DNA methylation detection: Bisulfite genomic sequencing analysis. ("Power and sample size calculations for high-throughput sequencing-based ...") Methods Mol Biol. 2011; 791:11–21.
106. Li Y, Tollefsbol TO: Age-related epigenetic drift and phenotypic plasticity loss: Implications in prevention of age-related human diseases; Epigenomics 8:1637; 2016.
107. Li Y, Zhu J, Tian G, Li N, Li Q, Ye M, Zheng H, Yu J, Wu H, Sun J, Zhang H, Chen Q, Luo R, Chen M, He Y, Jin X, Zhang Q, Yu C, Zhou G, Sun J, Huang Y, Zheng H, Cao H, Zhou X, Guo S, Hu X, Li X, Kristiansen K, Bolund L, Xu J, Wang W, Yang H, Wang J, Li R, Beck S, Wang J, Zhang X. The DNA methylome of human peripheral blood mononuclear cells. PLoS Biol. 2010;8:e1000533.
108. Liu X, Jiao B, Shen L. The Epigenetics of Alzheimer's Disease: Factors and Therapeutic Implications. Front Genet. 2018; 9:579. doi:10.3389/fgene.2018.00579.
109. Liu Z, Leung D, Levine M. Comparative analysis of epigenetic aging clocks from CpG characteristics to functional Associations. BioRxiv. 2019;512483.
110. Lucas VS, McDonald F, Andiappan M, Roberts G: Dental age estimation--Root Pulp Visibility (RPV) patterns: A reliable Mandibular Maturity Marker at the 18-year threshold; Forensic Sci Int 270:98; 2017. ("Can root pulp visibility in mandibular first molars be used as an ...")

111. Lyko F. The DNA methyltransferase family: A versatile toolkit for epigenetic regulation. Nat Rev Genet. 2018;19(2):81-92. https://doi.org/10.1038/nrg.2017.80.
112. Man CH, Fung TK, Wan H, Cher CY, Fan A, Ng N, Ho C, Wan TSK, Tanaka T, Wai C, et al.: Suppression of SOX7 by DNA methylation and its tumor suppressor function in acute myeloid leukemia; Blood 125:3928; 2015.
113. Márquez-Ruiz AB, González-Herrera L, Luna JD, Valenzuela A. DNA methylation levels and telomere length in human teeth: usefulness for age estimation. ("Neues Verfahren der forensischen Altersschätzung - Springer") Int J Legal Med 2020:134(2):451- 459.
114. Márquez-Ruiz AB, González-Herrera L, Valenzuela A. Usefulness of telomere length in DNA from human teeth for age estimation. Int J Legal Med. 2018; 132:353–359. https://doi.org/10.1007/s00414-017-1595-5.
115. Meissner C, Von Wurmb N, Schimansky B, Oehmichen M: Estimation of age at death based on quantitation of the 4977-bp deletion of human mitochondrial DNA in skeletal muscle; Forensic Sci Int 105:115; 1999.
116. Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology. 2013; 38:23–38. https://doi.org/10.1038/npp.2012.112.
117. Naue, J., Hoefsloot, H. C. J., Mook, O. R. F., Rijlaarsdam-Hoekstra, L., van der Zwalm, M. C. H., Henneman, P., et al. (2017). Chronological age predictionbased on DNA methylation: massive parallel sequencing and random forest regression. ("Age estimation using DNA methylation technique in ... - SpringerOpen") Foren. Sci. Int. Genet. 31, 19–28. doi: 10.1016/j.fsigen.2017. 07.015
118. Naue, J., Hoefsloot, H. C., Mook, O. R., Rijlaarsdam-Hoekstra, L., van der Zwalm, M. C., Henneman, P., ... & Verschure, P. J. (2017). Chronological age prediction based on DNA methylation: massive parallel sequencing and random forest regression. *Forensic science international: genetics*, *31*, 19-28.
119. O'Brien J, Hayder H, Zayed Y, Peng C. Overview of microRNA biogenesis, mechanisms of actions, and circulation. Front Endocrinol (Lausanne). 2018;9(AUG):1–12.
120. Ohtani S, Yamamoto T: Age estimation by amino acid racemization in human teeth; J Forensic Sci 55:1630; 2010.
121. Pal S, Tyler JK: Epigenetics and aging; Sci Adv 2(7): e1600584; 2016.
122. Patterson K, Molloy L, Qu W, Clark S. DNA Methylation: Bisulphite Modification and Analysis." ("Technical considerations in PCR-based assay design for diagnostic DNA ...") J Vis Exp. 2011;(56): e3170:1- 9.
123. Perna L, Zhang Y, Mons U, Holleczek B, Saum K-U, Brenner H: Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort; Clin Epigenetics 8:64; 2016.
124. Pfeifer M, Bajanowski T, Helmus J, Poetsch M. Inter-laboratory adaption of age estimation models by DNA methylation analysis-problems and solutions. Int J Legal Med. 2020;134(3):953-961. doi: 10.1007/s00414-020-02263-7.
125. Pilin A, Pudil F, Bencko V: Changes in colour of different human tissues as a marker of age; Int J Legal Med 121:158; 2007. ("A novel strategy for forensic age prediction by DNA ... - Nature")

126. Pilin, F. Pudil, V. Bencko, Changes in colour of different human tissues as a marker of age, Int. J. Legal Med. 121 (2007) 158–162.
127. Pinel, C., Prainsack, B., & McKevitt, C. (2018). Markers as mediators: A review and synthesis of epigenetics literature. *Biosocieties*, *13*, 276-303.
128. RA, Elbaz A, Brenner H, Kee F, Voortman T, Probst-Hensch N, Lehtimäki T, Elliot P, Stringhini S, Vineis P, Polidoro S; BIOS Consortium; Lifepath consortium. Socioeconomic position, lifestyle habits and biomarkers of epigenetic aging: a multi-cohort analysis. ("Home - Professor Paolo Vineis - Imperial College London") Aging (Albany NY) 2019;11(7):2045-2070. doi: 10.18632/aging.101900.
129. Richards, R., Patel, J., Stevenson, K., and Harbison, S. A. (2018). Evaluation of massively parallel sequencing for forensic DNA methylation profiling. Electrophoresis 39, 2798–2805. doi: 10.1002/elps.201800086
130. Sarda S, Hannenhalli S. Orphan CpG islands as alternative promoters. Transcription. 2018;9(3):171-176.
131. Schübeler D. Function and information content of DNA methylation. Nature. 2015;517(7534):321-6. doi: 10.1038/nature14192.
132. Spólnicka M, Pośpiech E, Pepłońska B, Zbieć-Piekarska R, Makowska Ż, Pięta A, et al. DNA methylation in ELOVL2 and C1orf132 correctly predicted chronological age of individuals from three disease groups. ("Centenarians consistently present a younger epigenetic age than their ...") Int J Legal Med. 2017;132(1):1-11.
133. Spólnicka, M., Pośpiech, E., Pepłońska, B., Zbieć-Piekarska, R., Makowska, Ż., Pięta, A., ... & Branicki, W. (2018). "DNA methylation in ELOVL2 and C1orf132 correctly predicted chronological age of individuals from three disease groups." ("DNA methylation in ELOVL2 and C1orf132 correctly predicted ... - Springer") *International journal of legal medicine*, *132*, 1-11.
134. Talens RP, Christensen K, Putter H, Willemsen G, Christiansen L, Kremer D, Suchiman HED, Slagboom PE, Boomsma DI, Heijmans BT: Epigenetic variation during the adult lifespan: Cross-sectional and longitudinal data on monozygotic twin pairs; Aging Cell 11:694; 2012.
135. Tan, Q., Heijmans, B. T., Hjelmborg, J. V. B., Soerensen, M., Christensen, K., & Christiansen, L. (2016). Epigenetic drift in the aging genome: a ten-year follow-up in an elderly twin cohort. *International journal of epidemiology*, *45*(4), 1146-1158.
136. Thong Z, Chan XLS, Tan JYY, Loo ES, Syn CKC. Evaluation of DNA methylation-based age prediction on blood. Forensic Sci Int Genet Suppl Ser. 2017;6:e249-e251.
137. Thong, Z., Chan, X. L. S., Tan, J. Y. Y., Loo, E. S., & Syn, C. K. C. (2017). Evaluation of DNA methylation-based age prediction on blood. *Forensic Science International: Genetics Supplement Series*, *6*, e249-e251.
138. Tsaprouni LG, Yang TP, Bell J, Dick KJ, Kanoni S, Nisbet J, Vinuela A, Grundberg E, Nelson CP, Meduri E, et al.: Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation; Epigenetics 9:1382; 2014.

139. Tsuji, A. Ishiko, T. Takasaki, N. Ikeda, Estimating age of humans based on telomere shortening, Forensic Sci. Int. 126 (2002) 197–199. (“Select aging biomarkers based on telomere length and ... - Springer”) [5] C. Meissner, S. Ritz-Timme, Molecular pathology and age estimation, For

140. Turelli P, Castro-diaz N, Marzetta F, Kapopoulou A, Duc J, Tieng V, Quenneville S, Trono D: Interplay of TRIM28 and DNA methylation in controlling human endogenous retroelements Interplay of TRIM28 and DNA methylation in controlling human endogenous retroelements; Genome Res 24:1260; 2014.

141. Vidaki A, Kayser M. From forensic epigenetics to forensic epigenomics: broadening DNA investigative intelligence. Genome Biol. 2017; 18:238-250.

142. Vidaki, A., Ballard, D., Aliferi, A., Miller, T. H., Barron, L. P., and Syndercombe Court, D. (2017). DNA methylation-based forensic age prediction usingartificial neural networks and next generation sequencing. (“(PDF) DNA methylation-based forensic age prediction ... - ResearchGate”) Foren. Sci. Int. Genet. 28, 225–236. doi: 10.1016/j.fsigen.2017.02.009.

143. Vokonas PS, Just A, Hodge AM, Giles GG, Southey MC, Hurme MA, Young I, McKnight AJ, Kunze S, Waldenberger M, Peters A, Schwettmann L, Lund E, Baccarelli A, Milne RL, Kenny .

144. Weidner CI, Lin Q, Koch CM, Eisele L, Beier F, Ziegler P, Bauerschlag DO, Jöckel KH, Erbel R, Mühleisen TW, Zenke M, Brümmendorf TH, Wagner W. Aging of blood can be tracked by DNA methylation changes at just three CpG sites. Genome Biol. 2014;15(2):R24. doi: 10.1186/gb-2014-15-2-r24.

145. Weidner, C. I., Lin, Q., Koch, C. M., Eisele, L., Beier, F., Ziegler, P., et al. (2014). Aging of blood can be tracked by DNA methylation changes at just three CpG sites. Genome Biol. 15:R24.

146. Wright K, Mundorff A, Chaseling J, Forrest A, Maguire C, Crane DI: A new disaster victim identification management strategy targeting "near identification-threshold" cases: Experiences from the Boxing Day tsunami; Forensic Sci Int 250:91; 2015.

147. X. Ou, H. Zhao, H. Sun, Z. Yang, B. Xie, Y. Shi, X. Wu, Detection and quantification of the age-related sjTREC decline in human peripheral blood, Int. J. Legal Med. 125 (4) (2011) 603–608.

148. Xu, C., Qu, H., Wang, G., Xie, B., Shi, Y., Yang, Y., et al. (2015). "A novel strategy for forensic age prediction by DNA methylation and support vector regression model." (“A novel strategy for forensic age prediction by DNA methylation and ...”) Sci. Rep. 5:17788.

149. Zampieri M, Ciccarone F, Calabrese R, Franceschi C, Bürkled A, Caiafa P. Reconfiguration of DNA methylation in aging. Mech Ageing Dev. 2015; 151:60–70.

150. Zapico SC, Ubelaker DH. mtDNA Mutations and their role in aging, diseases, and forensic

151. Zbieć-Piekarska R, Spólnicka M, Kupiec T, Makowska Ż, Spas A, Parys-Proszek A, Kucharczyk K, Płoski R, Branicki W. Examination of DNA methylation status

of the ELOVL2 marker may be useful for human age prediction in forensic science." ("DNA methylation-based age estimation for adults and minors: considering ...") Forensic Sci Int Genet. 2015a; 14:161–167.

152. Zbiec-Piekarska, R., Spólnicka, M., Kupiec, T., Parys-Proszek, A., Makowska, Z., ´ Pałeczka, A., et al. (2015). "Development of a forensically useful age prediction method based on DNA methylation analysis." ("Development of a forensically useful age prediction method based on DNA ...") Foren. Sci. Int. Genet. 17, 173–179. doi: 10.1016/j.fsigen.2015.05.001

153. Zbieć-Piekarska, R., Spólnicka, M., Kupiec, T., Parys-Proszek, A., Makowska, Ż., Pałeczka, A., ... & Branicki, W. (2015). "Development of a forensically useful age prediction method based on DNA methylation analysis." ("Development of a forensically useful age prediction method based on DNA ...") *Forensic Science International: Genetics*, *17*, 173-179.

154. Zhang Y, Hapala J, Brenner H, Wagner W. Individual CpG sites that are associated with age and life expectancy become hypomethylated upon aging. Clin Epigenetics. 2017a; 9:9. doi: 10.1186/s13148-017-0315-9.

155. Zhang Y, Wilson R, Heiss J, Breitling LP, Saum KU, Schöttker B, Holleczek B, Waldenberger M, Peters A, Brenner H. DNA methylation signatures in peripheral blood strongly predict allcause mortality. Nat Commun. 2017b; 8:14617

156. Zhao C, Zhao M, Zhu Y, Zhang L, Zheng Z, Wang Q, et al. The persistence and stability of miRNA in bloodstained samples under different environmental conditions. Forensic Sci Int [Internet]. 2021; 318:110594. Available from: https://doi.org/10.1016/j.forsciint.2020.110594.

157. Zolotarenko AD, Chekalin EV, Bruskin SA. Modern Molecular Genetic Methods for Age Estimation in Forensics. Russ J Genet. 2019; 55:1460–1471.

158. Zolotarenko, A.D., Chekalin, E.V. & Bruskin, S.A. Modern Molecular Genetic Methods for Age Estimation in Forensics. *Russ J Genet* **55**, 1460–1471 (2019). https://doi.org/10.1134/S1022795419120147

159. Zubakov D, Liu F, Van Zelm MC, Vermeulen J, Oostra BA, Van Duijn CM, Driessen GJ, Van Dongen JJM, Kayser M, Langerak AW: Estimating human age from T-cell DNA rearrangements; Curr Biol 20: R970; 2010. ("Aging of blood can be tracked by DNA methylation changes at just three ...")

160. Zubakov, D., Liu, F., Kokmeijer, I., Choi, Y., van Meurs, J. B. J., van IJcken, W. F. J., et al. (2016). "Human age estimation from blood using mRNA, DNA methylation, DNA rearrangement, and telomere length." ("Evaluation of six blood-based age prediction models using DNA ... - Nature") Foren. Sci. Int. Genet. 24, 33–43. doi: 10.1016/j.fsigen.2016.05.014