Universidade de Aveiro Departamento de Biologia
Ano 2018/19

João Leonardo Azevedo Meireles Pereira

Assessement of *Postmortem* Intervals using DNA Methylation.

Estimação de Intervalos *Postmortem* usando Metilação do DNA.

DECLARAÇÃO

Declaro que este relatório é integralmente da minha autoria, estando devidamente referenciadas as fontes e obras consultadas, bem como identificadas de modo claro as citações dessas obras. Não contém, por isso, qualquer tipo de plágio quer de textos publicados, qualquer que seja o meio dessa publicação, incluindo meios eletrónicos, quer de trabalhos académicos.

Universidade de Aveiro Departamento de Biologia
Ano 2018/19

João Leonardo Azevedo Meireles Pereira

Assessement of *Postmortem* Intervals using DNA Methylation.

Estimação de Intervalos *Postmortem* usando Metilação do DNA.

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos exigidos para a obtenção de um Mestrado em Biologia Molecular e Celular, realizado sob a orientação científica de Luís Manuel Souto de Miranda, Professor Auxiliar em Regime Laboral do Departamento de Biologia da Universidade de Aveiro.

Agradecimentos

Gostaria de agradecer à minha mãe por me ajudar com tudo durante a minha vida, minha irmã por me ajudar a aprender a ajudar-me a mim próprio em tempos difíceis e aos meus amigos João Oliveira, Rafael Almeida e Joana Castro por me ajudarem com uma variedade de dificuldades durante o desenvolvimento deste trabalho e me manterem direito e focado enquanto escrevia este trabalho. Gostaria também de agradecer à doutora Helena Moreira por me dar ótimas noções sobre o mundo da pesquisa científica e por me orientar em duas grandes experiências de divulgação científica e ao professor Luís Souto por possibilitar a criação deste trabalho.

o júri

presidente	Professor Doutor António José de Brito Fonseca Mendes Calado, Professor Auxiliar, Universidade de Aveiro
	Professora Doutora Maria de Lourdes Gomes Pereira, Professora Associada C/ Agregação, Universidade de Aveiro
	Professor Doutor Luís Manuel Souto de Miranda, Professor Auxiliar em Regime Laboral, Universidade de Aveiro

Acknowledgments

I would like to thank my mother for helping me with everything during my life, my sister for helping me learn how to help myself in hard times and my friends João Oliveira, Rafael Almeida and Joana Castro for helping me with a variety of difficulties during the development of this work and keeping me well in track and focused while writing this work. I would also like to thank Doctor Helena Moreira for giving me great insights about the world of scientific research and guiding me through two great scientific divulgation experiences and Professor Luis Souto for enabling the creation of this work.

palavras-chave Intervalos *Postmortem*; Análise de Metilação de DNA; Epigenetica Forense.

resumo A estimativa do tempo de morte de um cadáver é uma parte absolutamente crucial de muitos casos forense. Um posicionamento preciso no tempo na sequência de eventos de um caso é uma informação que pode criar uma lista de suspeitos ou isentar suspeitos, alterando profundamente um caso inteiro. Atualmente, o intervalo *postmortem* e a hora da morte são avaliados por meio de várias avaliações físicas, metabólicas e físico-químicas durante a autópsia. Essas avaliações trazem pontos fortes, fracos e variações, e um procedimento mais estável e robusto seria inestimável para o avanço das ciências forenses. Um dos focos mais recentes para esse possível procedimento baseia-se numa das regulações epigenéticas do DNA, a adição de um grupo metilo a certas citosinas, conhecido como metilação do DNA. A metilação do DNA regula muitos aspetos do genoma, desde o silenciamento de genes expressos só em alguns tecidos, seções não codificantes do DNA, até a especialização de tecidos e uma resposta estável a longo prazo a estímulos do ambiente. Este trabalho é uma revisão sobre procedimentos, possíveis biomarcadores, possíveis variáveis e obstáculos no uso desta modificação regulatória do consideravelmente estável *postmortem* DNA.

Keywords *Postmortem* Intervals; DNA Methylation Analysis; Forensic Epigenetics.

Abstract Estimation of a corpse's time of death is an absolute crucial part of many forensic cases. An accurate time placement in a case's sequence of events is information that can create a suspect list or exempt suspects from it, deeply altering an entire case. Currently, *postmortem* intervals and time of death is assessed through several physical, metabolic and physiochemical evaluations during autopsy. These evaluations carry strengths, weaknesses and variations, and a more stable and robust procedure would be invaluable to the advancement of forensic sciences. One of the most recent focus for such a possible procedure relies in one of the epigenetic regulations of DNA, the addition of a methyl group to certain cytosines, known as DNA methylation. DNA methylation regulates many aspects of the genome, from silencing tissue specific genes, non-coding sections of DNA, up to tissue specialization and working as a stable long-term response to the environment. This work is a review on procedures, potential biomarkers, possible viability and obstacles in the way of using this regulatory modification of the considerably stable *port-mortem* DNA.

Index

Image Index

List of abbreviations

ACTB- Actin Beta
ADORA2A- Adenosine A2A receptor
AID- Activation-induced cytidine deaminase
APOBEC- Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like
ATP- Adenosine triphosphate
BDNF- Brain Derived Neurotrophic Factor
BER- Base excision repair
BiMP- Bisulfite methylation profiling
CDKN1A- Cyclin dependent kinase inhibitor 1A
cDNA- Complementary deoxyribonucleic acid
CHARM- Comprehensive high-throughput arrays for relative methylation
COBRA- Combined bisulfite restriction analysis
CODIS- Combined DNA index system
COMPARE-MS- Combination of methylated DNA precipitation and methylation-sensitive restriction enzymes
CRAT- Carnitine O-acetyltransferase
dATPαS- 2'-Deoxyadenosine-5'-(α-thio)-triphosphate
dCTP- Deoxycytidine triphosphate
ddNTP- Dideoxynucleotides triphosphate
dGTP- Deoxyguanosine triphosphate
DHPLC- Denaturing HPLC
DMH- Differential methylation hybridization
DNA- Deoxyribonucleic acid
DNMT- DNA methyltransferase
dNTP- Deoxyribonucleotide triphosphate
dTTP- Thymidine triphosphate
EDAR- Ectodysplasin A receptor
EDARADD- Ectodysplasin-A receptor-associated adapter protein
ELISA- Enzyme-linked immunosorbent assay
F3- Tissue factor 3
FASN- Fatty acid synthase
FASN-mRNA- Fatty acid synthase-messenger ribonucleic acid
GAPDH- Glyceraldehyde 3-phosphate dehydrogenase
GCR- Glucocorticoid receptor
GRCz11- Genome Reference Consortium Zebrafish Build 11
HCA- Hierarchical clustering analysis
HELP- HpaII tiny fragment enrichment by ligation-mediated PCR
HRM- High resolution melt
ICAM-1- Intercellular adhesion molecule 1

IFN-γ- Interferon gamma
iNOS- Inducible nitric oxide synthase
IP- Immunoprecipitated
LC-MS/MS- Liquid chromatography coupled with tandem mass spectrometry
LINE- Long interspersed nuclear elements
LINE-1- Long interspersed nuclear element
LTR- Long terminal repeats
LUMA- Luminometric methylation assay
MAD- Methylation amplification DNA chip
MALDI-TOF- Time-of-flight matrix-assisted laser desorption/ionization
MAPT- Microtubule-associated protein tau
MCA- Methylated CpG island amplification
MCAM- Methylated CpG island amplification microarray
MeDIP- Methylated DNA immunoprecipitation
MeDIP-PCR- Methylated DNA immunoprecipitation PCR
MeDIP-PCR/seq- Methylated DNA Immunoprecipitation PCR/Sequencing
MIRA- Methylated CpG island recovery assay
miRNA- Micro ribonucleic acid
MMASS- Microarray-based methylation assessment of single samples
mRNA- Messenger ribonucleic acid
MSCC- Methylation-sensitive cut counting
MS-FLAG- Methylation-specific fluorescent amplicon generation
MSNP- Methylation single-nucleotide polymorphism
MSP- Methylation-specific PCR
NTRK2- Neurotrophic receptor tyrosine kinase 2
OGG1- 8-oxoguanosine DNA glycosylase-1
ORF- Open reading frame
PCR- Polymerase chain reaction
PMAD- Promoter-associated methylated DNA amplification DNA chip
PMI- *Postmortem* Interval
qPCR- Real-time polymerase chain reaction
RAGE- Advanced glycation end products
RLGS- Restriction landmark genome scanning
RNA- Ribonucleic acid
RRBS- Reduced representation bisulfite sequencing
rRNA- Ribossomal ribonuceic acid
SAM- S-Adenosyl methionine
SINE- Short interspersed nuclear elements
SMART-MSP- Sensitive melting analysis after real-time methylation-specific PCR
snRNA- Small nuclear ribonucleic acid
STR- Short Tandem Repeats
TBP- TATA Binding protein

TFIIA- Transcription factor II A
TFIIB- Transcription factor II B
TFIID- Transcription factor II D
TFIIE- Transcription factor II E
TFIIF- Transcription factor II F
TFIIH- Transcription factor II H
TLR-2- Toll-like receptor 2
uhrf1- Ubiquitin like with PHD and ring finger domains 1
WGSGS- Whole genome shotgun bisulfite sequencing
WGSGS- Whole genome shotgun bisulfite sequencing

1. Forensic Sciences, an Introduction

Forensic science is a broad term applied to any and all science which is accepted by legal standards that can be used in criminal investigation in order to uphold the perpetrator accountable to the established criminal and civil law and current legislation.

The term forensic derives from a form roman trial [1] that when a dispute or criminal charge would be made, defense and prosecution would be presented to several individuals in the roman forums, giving the word its meaning of "pertaining to the forum" [2]. While romans had a strong judicial system, laws and enforcement bodies, one of the first written books on how to use medicine and entomology to help with crime solving, effectively being one of the first published forensic books, was *Xǐ-yuān lù* or *Collected Writings on the Washing Away of Wrongs,* written by Song Ci, during the reign of Shun Yu, in 1247 [3] **(Fig. 1).**

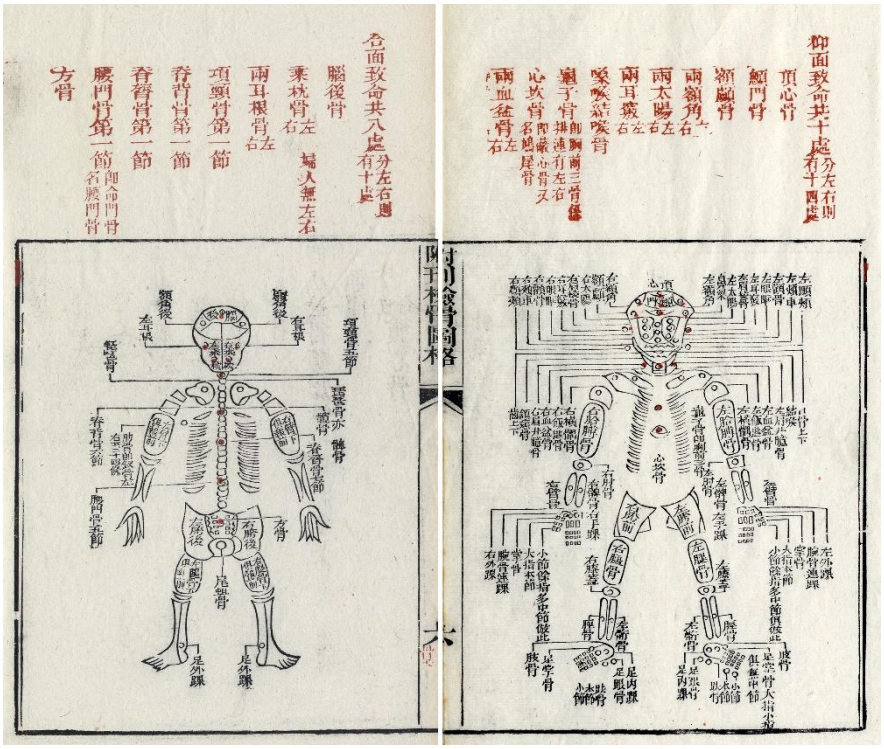

Fig. 1- A sample of Xǐ-yuān lù, detailing the human skeleton and its nomenclature, present in Sòng Cí: Xǐ-yuān lù jí-zhèng. Adapted from *https://upload.wikimedia.org/wikipedia/commons/thumb/1/18/Xiyuanlu_jizheng-1843-Bones.jpg/330px-Xiyuanlu_jizheng-1843-Bones.jpg*

While currently being a wide and encompassing field of study, in early human history, forensic sciences have been unreliable and often misused. In early criminal investigations and trials, confession and witness testimony were the most used methods of enforcing accountability, often leading to dubious results due to confessions being forced and witnesses unreliable.

In more recent times, several fields of forensic sciences have been well defined. Thanatology, from the Greek word *Thanatos*, meaning death or god of death in mythology [4], is the study of death and how the body and others are affected by it [5]. Forensic thanatology on the other hand is the study of every phenomenon cause by death through *postmortem* examination [6]. It encompasses the medico-legal autopsy, a form of specialized autopsy realized by a pathologist qualified in forensic pathology done with the intent of finding and registering any and all injuries in detail, so that they may be used in the court of law [7], on top of the regular objectives of a clinical autopsy. A medico-legal autopsy also has as objectives the identification of the body, the estimation of the time of death, the determination of the relevance of each injury to the cause of death, the prior presence of disease and its effects on the body, toxin and poison screening and presence of effects lingering from prior medical surgeries or interventions [7]. The amount of information about a case that can be acquired through a medico-legal autopsy is large, but it is limited in the scope of time available, because as more time passes, the finer details may be lost.

Another area of expertise that are used to glean information about the case on a bigger time frame is forensic entomology [8]. Whenever a body dies, in nature, it becomes a bounty of easy food, ripe for the taking. Allured by such an opportunity, a great variety of animals flock to the scene, most being scavengers, bacteria, fungi, with necrophagous insects being the most prominent, both in number and forensic importance [8]. Many of the insect species that gorge themselves in the decaying flesh also breed on it, and the timings of their life cycle and knowledge of the times taken for the species to settle in a body can be used to calculate the age of a body. Not only that but some of the species are only native to certain environments and habitats, causing their presence in a body found in a foreign area to the species to denounce the movement of the body *postmortem* [8].

Unfortunately, there will be cases where the body will not be found for many, many years, or even decades. By then, naught but the skeleton and hard objects like metal belongings of the victim may remain in its resting place. In these cases, the aid of forensic anthropology is required. Forensic anthropology encompasses several fields of study, such as anatomy, anthropology, archeology to some extent, and shares many points of interest with odontologists [9]. Forensic anthropologists work with remains that have almost to none flesh, such as skeletons, fully decomposed bodies and burnt to carbonized remains, to name a few [9], in order to provide those remains an identification, be it of ancestry, sex, age, both at time of death and current age, or all combined [9], information necessary for a criminal case. This is achieved through extensive knowledge of the morphology of the human skeleton and the dysmorphisms it exhibits between diverse ages, ancestries and both sexes.

These three fields of forensic sciences are used in conjunction with each other and many others to obtain several distinct bits of information about the case but all three can provide the very crucial and critical time of death and the time passed between its occurrence and the discovery, the *postmortem* interval.

1.1. *Postmortem* Intervals

Postmortem intervals refer to the time transpired between an organism's death and the time of its discovery. It's a critical piece of information necessary for the creation of a solid forensic analysis of a case, without which a case might go unsolved due to the impossibility of placing victim and perpetrator in the crime scene at the same time without reasonable doubt. Current techniques and tactics used to determine a PMI can range from social fields like time and location of last phone calls and text messages sent by the victim and passerby witness sightings to thanatological findings like body temperature and *rigor mortis* to more experimental protocols in development such as mRNA degradation [10], [11].

Most of these techniques are based on the effects of death on the body. Upon the irreversible cessation of the heartbeat, several changes start happening to the organism. Without blood being pumped and the onset of hypoxia beginning, cells begin to sustain themselves through anoxic metabolic pathways, until reserves are depleted and cell death ensues [7]. At a macro level, quickly after death the termination of circulation drains the blood from capillaries, creating a paleness called *pallor mortis* (**Fig. 2**) [12], setting in between fifteen to twenty minutes, and muscles lose all tone, signaling the primary flaccidity, but still staying responsive to internal and external stimuli [7].

Fig. 2- Example of *pallor mortis*, denoted by the pallidity of the skin. Adapted from "*http://www.doctoralerts.com/wp-content/uploads/2017/03/Pallor-Mortis.jpg*"

Following the primary flaccidity, muscle cells start building up lactic acid as a sub-product of the anoxic pathways, and in conjunction with low levels of ATP in the cell, the myosin fibers and the actin bind together, creating rigid structures, turning the muscles stiff, signaling the onset of *rigor mortis* (**Fig. 3**) [7]. The onset of *rigor mortis* occurs by stages, being noticed between one to four hours in the face and three to six hours in the limbs, reaching peak strength at eighteen hours, and once fully set, lasts for fifty hours [7].

Fig. 3- Example of *rigor mortis*, right foot not touching the ground defying gravity due to onset of stiffness. Adapted from "*https://www.semanticscholar.org/paper/Rigor-mortis-in-an-unusual-position%3A-Forensic-D%E2%80%99Souza-Harish/fbe0653e31dbd1c230ea0d7726e8e87640e60a79/figure/1*"

During these stages, *livor mortis* and *algor mortis* are also taking effect. *Livor mortis*, or *postmortem* hypostasis refers to the settling of blood to the bottom of the organism of whatever position it may be in, due to gravity and lack of agitation (**Fig. 4**) [13]. It starts to settle immediately upon death, with early visible signs of discoloration on the superior part and reddening of the inferior part starting to appear by the hour mark, being fully visible from two to four hours after death, stage during which external pressure will still cause blanching, and will be fully fixated by the nine to twelve hour mark [13].

Fig. 4- Example of *livor mortis* settling in, as blood drains to the bottom of the body. Adapted from "*https://www.semanticscholar.org/paper/Early-post-mortem-changes-and-stages-of-in-exposed-Goff/67107f4441924394c376c198800bf49357dbc23f/figure/1*"

Algor mortis pertains to the cooling of the body. As cells switch to anoxic pathways, the body loses the capacity to regulate its own temperature, and as the laws of thermodynamics state, a warm object in contact with a cooler one will transfer its energy until balance is achieved. As such, a body will eventually reach the temperature of its environment [7]. On a general case, this will take up to eighteen to twenty hours, becoming unreliable past such interval [13]. After these stages, putrefaction and decomposition start. The exact form differs between the environment and temperature body lays in. Under water, putrefaction occurs much slower than bodies present in open air [7] and may suffer *saponification*, a process of anaerobic hydrolysis of the fatty tissue (**Fig. 5**) [13].

Fig. 5- A body that underwent *saponification.* Adapted from "*https://www.semanticscholar.org/paper/Early-adipocere-formation%3A-a-case-report-and-review-Kumar-Monteiro/f5387caf2c343b39ba9f1342bc0610a7b0832455/figure/0*"

Under dry environments, a body may suffer *mummification*, turning the skin dry and leathery-like (**Fig. 6**) [7].

Fig. 6- Mummy of Ötzi the Iceman, a 3000-year-old naturally preserved mummy by the dry, icy conditions of the Similaun mountain. Adapted from "*https://media.nbcdfw.com/images/653*367/72125418.jpg*"

A combination of these markers can be used to calculate the time of death, and PMI, in cases of early body recovery, but in cases of periods larger than hours or days, whenever putrefaction and decomposition have settled in, the assessments starts to become ranges of possible PMIs, and the work of forensic entomologists may help greatly to refine the possibilities [7]. Their results are also obtained in ranges of possible PMIs, and the overlap is the key to such refinement.

Yet, for such a vital necessity, the field techniques or thanatological assessments sustain variations in sensitivity due to several factors such as the temperature, the regional climate, local humidity, depth of burial, exposure to the elements and animals, body weight, *etc*. [14], suffering greater and greater deviations and errors as time passes. Laboratorial analysis may be able to avoid these variables or take them into account when determining the PMI. Ideally, and most times required by court to assure validity, a battery of reliable PMI assessing techniques based on physical, metabolic, bacterial, physiochemical, and now, genetic processes with enough well researched values, means and confidence limits should be used in order to combat each techniques' shortcomings. With RNA degradation already under study to potentially be applied as a PMI estimator [10], [15], [16], the degradation of the epigenetic code of the DNA is another possible route of study.

2. Genetics and Epigenetics

Deoxyribonucleic acid, or DNA, is the basis for which cells rely on to store nearly all the critical genetic information necessary for the entirety of its mechanisms. It is a complex double helix chain macromolecule composed of a large sequence of 4 nucleotides: the purines guanine and adenine and the pyrimidines cytosine and thymine, in a myriad of combinations, strung together through phosphodiester bonds created by the phosphate groups present in the 2-deoxyribose sugar residue backbone present in all of the 4 nucleotides [17]. These nucleotides are themselves the basis for the genetic code. After transcription and during translation, sets of 3 nucleotides, with thymine being substituted with uracil in mRNA, are read by the ribosome as codons, which are correspondent to an amino acid to be added to the extending protein chain [18]. These sets of nucleotides that are to be translated into functional proteins through transcription and translation are stored within genes in the DNA, which are subjected to many forms of regulations. Upstream of the open reading frame (ORF), the promoter region which is necessary for gene expression exists [19]. It's a region that is divided into 2 sections, the core promoter directly upstream of the ORF and the proximal promoter directly upstream of the core promoter. The core promoter is a critical structure for gene expression, being recognized by general transcription factors, regulatory proteins that are absolutely required for gene expression, and it is a binding site for RNA polymerase I, which transcribes rRNA precursors, RNA polymerase II, which transcribes mRNA, snRNA and miRNA, and RNA polymerase III, which

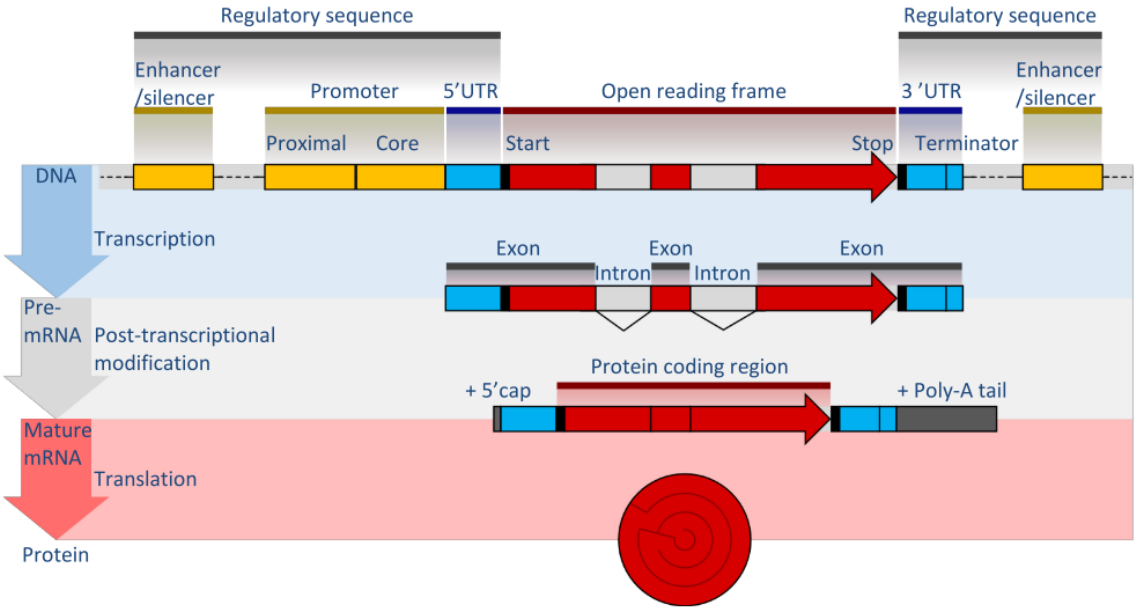

Fig. 7- A simplified flowchart from DNA information to protein, depicting cis-regulatory elements present is the gene structure such as the promoters, enhancers and silencers (top section), a pre-mRNA containing introns before being excised and lacking 5'cap and Poly-A tail (middle section), a fully matured mRNA without introns and protected by the 5'cap and Poly-A tail (bottom section) and a fully folded, functional protein (red circle). Adapted from T. Shafee and R. Lowe, "Eukaryotic and prokaryotic gene structure" vol. 4, no. 1, pp. 0–3, 2017.

transcribes tRNA and 5S rRNA [20]. RNA Polymerase II is a protein complex with 12 subunits whose binding to the DNA molecule is mediated by the presence of TATA binding protein (TBP), a TFIID subunit, that once bound to the promoter recruits TFIIB, which in turn helps the Polymerase II complex (RNA polymerase II and TFIIF) bind properly to the promoter. Once this happens, TFIIE and TFIIH bind to the complex and form the transcription preinitiation complex. This complex is now capable of unwinding and melting the double helix, point where the RNA Polymerase is capable of initiating elongations and all but TFIID leave the complex [21]. The proximal promoter is the area of the promoter that houses the specific transcription factor binding sites, responsible for the regulation of the affinity of the core promoter to the RNA Polymerases **(Fig. 7).**

Epigenetics pertain to a group of modifications made to the nucleotides and histones, affecting the overall structure and condensation of the DNA double helix. These modifications serve as another layer of genetic regulation on top of all molecules that affect the binding of the RNA polymerases to the promoters, and their binding affinity. Epigenetic processes are responsible for a wide array of important phenomena such as genomic imprinting, inactivation of the X chromosome and several tumor suppressing regulations [22] and also act as the long term response to environmental pressures and stresses suffered over a large period of time [23]. The epigenetic code is highly sensitive and susceptible to these changes and can act as a genetic fingerprint on top of the already discriminatory DNA, with enough discerning power to distinguish even monozygotic twins, who share the same genetic information and very similar environments for most of their early life and yet end up with noticeable differences [24].

2.1. Epigenetics: DNA Methylation

DNA methylation is a regulatory process where a methyl group is attached to, primarily, cytosine residues, although small amounts of cytosine hydroxymethylation and adenine methylation in certain unicellular eukaryotes can also be found. The process of DNA methylation is carried out by the DNA Methyltransferases enzyme family (DNMT), consisting of three core proteins: DNMT1, DNMT3a and DNMT3b.

Throughout the entirety of an organism's life, DNA methylation is kept as a housekeeping process. During the S phase of cell division [25], the phase on which DNA replication occurs, DNMT1 is recruited to the DNA replication sites through interactions with proliferating cell nuclear antigen and UHRF1. This mechanism is usually enough to maintain all imprinted DNA methylation. DNMT1 has higher affinity for hemi-methylated DNA [26], [27], being bound to new DNA strands during replication, which are constituted by a methylated template strand and an unmethylated nascent strand. The requirement of a partially methylated DNA molecule to use as a template highly reduces DNMT1's effectiveness as a *de novo* methyltransferase. When a complete DNA methylation wipe occurs, during gametogenesis and just before embryonic specification in the blastocyst, the *de novo* capable methyltransferases DNMT3a and DMNT3b are recruited instead. DNMT3b has been linked to earlier stages of embryo development and germline activity and DNMT3a

to later stages and post birth [27], maintaining and remaking the DNA methylation pattern during those stages. Without a template strand to guide the methylation, DNMT3 a and DNMT3b require regulation and support to prevent deviant methylation and correctly maintain repressed promoters. This help comes from many sources, one of them being the histone methyltransferases, enzymes responsible for the other facet of epigenetic regulation, histone methylation, acetylation, phosphorylation, and ubiquitination. These enzymes are named based on their function, starting by the kind of histone they act upon (H X), the base they act upon citing both residue and position and the number of methyl groups attached. CpG locations on or near transcription start sites are generally protected from DNMT3a and DNMT3b activity by the presence and activity of the transcription factor, H3K4 methyltransferases, the CXXC finger protein 1 (CFP1), also known as CpG-binding protein, CpG binding protein 1 and DNA demethylation related enzymes such as TET enzymes or thymidine DNA glycosylase (**Fig. 8**)[25]. Promoters that do require silencing are done so after transcription, starting with the recruitment of LSH, a chromatin remodeling enzyme. Afterwards, Linker histone H1, H3K9 methyltransferases, heterochromatin protein 1 which are proteins with high binding affinity to nucleosomes containing histone H3 methylated at lysine K9, and the *de novo* DNA methyltransferases are recruited into the area, in that order [25]. In germline *de novo* methylation, DNMT3b acts after repression factors like Transcription factor E2F6 are recruited to the site. Two other DNMT proteins exist, DNMT2 and DNMT3L, that do not function as cytosine methyltransferases.

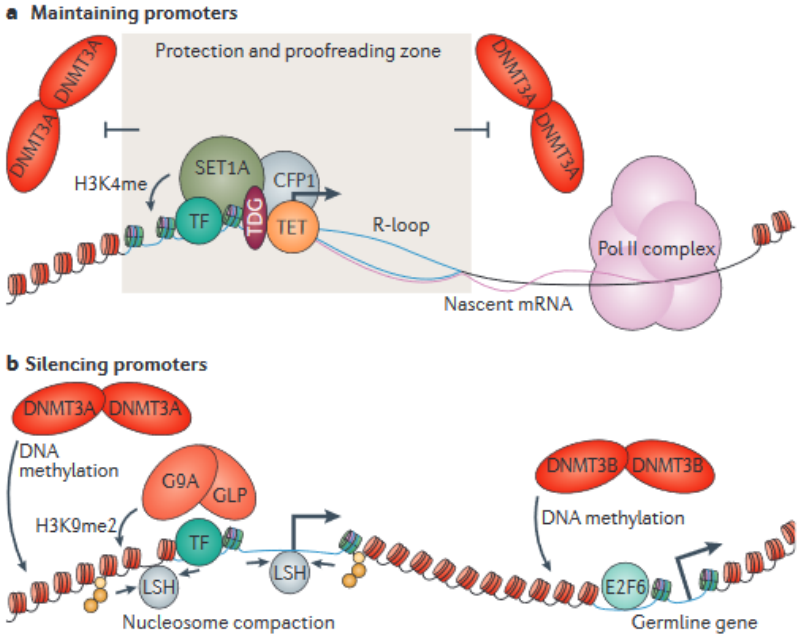

Fig. 8- Transcription start site's CpG site protection mechanisms. TF: Transcription Factor; TET: Ten-eleven Translocation enzyme; TDG: Thymine-DNA Glycosylase; SET1A: SET Domain Containing 1A Histone Lysine Methyltransferase; CFP1: CpG-binding protein (a). *De novo* methylation mechanism at promoters that are to be silenced. G9A is a H3K9 methyltransferase with GLP being an associated protein (b). Adapted from *"Z. D. Smith and A. Meissner, "DNA methylation: Roles in mammalian development" Nat. Rev. Genet., vol. 14, no. 3, pp. 204–220, 2013."*

However, most of the DNA expanses are not compromised of genes and their regulatory areas, the largest expanses of genetic code in the human genome are comprised of regions of non-coding DNA that also undergo regulatory methylation. The pericentromeric repeats located at the chromosome's centromere are crucial for correct chromosome alignment during mitosis. These areas contain two types of satellites: major satellites and minor satellites. DNMT3b acts on both methylation mechanisms alongside H3K9me3 Histone-lysine N-methyltransferase SUV39. The difference between both mechanisms lies in the necessity of the presence of LSH for methylation to occur in major satellites [25]. Long interspersed nuclear elements (LINE), short interspersed nuclear elements (SINE) and long terminal repeats (LTR) are vast sections of genetic code interspersed throughout the mammalian code and are kept under strict hypermethylation, with hypomethylation of some elements being linked to cancer [28]. The methylation of these areas of the genome are maintained by the same DNMT1 mechanism but suffer hemimethylation and hydroxymethylation at a higher rate than the rest of the genome, requiring DNMT3a intervention more frequently [25].

The reaction consists in the addition of a methyl group to the 5' carbon of the cytosine residues, creating 5'-methyl-cytosine, using S-Adenosyl methionine (SAM) as a donor (**Fig. 9).** For the reaction to occur, a specific CpG motif called CpG site is required for recognition and binding to the correct location. Areas of the genome that contains a high frequency of these motifs are denominated CpG Islands. Although the CpG motif is required for DNMT activity and 60% to 90% of CpG sites being methylated [29], CpG Islands, sections of DNA that range from 300bp to 3000bp with a 55%+ CG content present in the 5'-end promoter regions are mostly unmethylated [30], and silencing of these areas is usually done by histone methylation (H3K27) [25]. High DNA methylation levels in this area renders downstream genes silenced and diminishes their expression considerably [31], leaving histone methylation as a failsafe against incorrect DNMT activity [25].

If left unmaintained, and through influence of an organism's ambient, age, health and many other external pressures, 5'-methyl-cytosine may suffer oxidation, turning into 5'-hydroxymethyl-cytosine [30], which through DNA repair mechanisms suffers demethylation, either through deamination into a T:G mismatch with prompt repair to a C:G match or through deamination catalyzed by Activation-induced cytidine deaminase (AID)/Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like (APOBEC) deaminase families into 5'-Hydroxymethyl-uracil which is excised and repaired by the Base excision repair (BER) pathway leaving an unmethylated cytosine in that position.

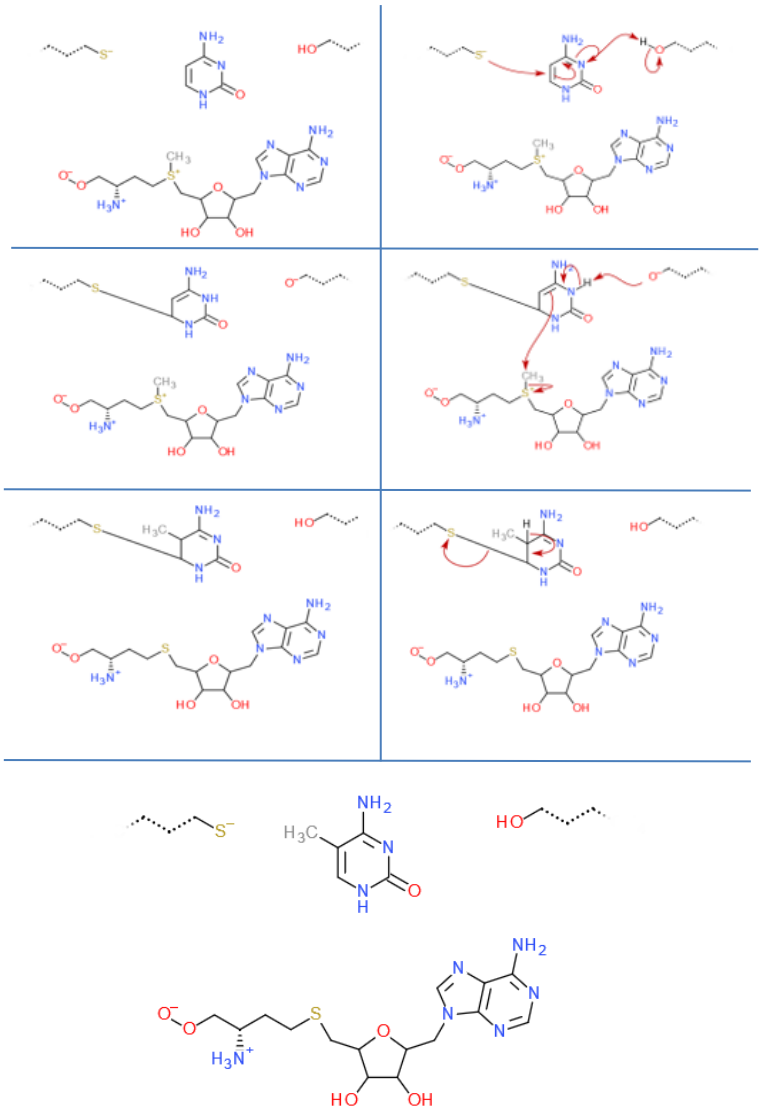

Fig. 9- Possible mechanism of the DNA Methylation reaction between a cytosine nucleotide (center top), SAM (center bottom) and the catalytic center of DNMT (top right and top left).

2.2. Genetics in Forensic Science

Presently, the genetic field of science holds only one, yet key, practice in the forensic scope: the DNA profiling. DNA profiling is the process of identifying a series of extremely variable areas of the genetic code in a sample of genetic material in order to create an individual profile that can be used in matching procedures. Most of the human genetic code is shared amongst the entire population as is with only some small variable components, yet these areas within themselves can contain tremendous amounts of variability, known as polymorphic markers [32].

Initially and currently, the main group of polymorphic markers used for DNA profiling is the STR, short tandem repeats. Short tandem repeats are repeats of 3, 4, 5 or

more nucleotides present throughout the DNA that are repeated multiple times back-to-back several times. The differentiating factor between STR loci is the number of such repeats, with an individual being able to possess up to two different numbers of repeats for every STR loci, each chromosome having its own "allele", inherited through the donating progenitor. STRs also work very alike genes when it comes to mutations and linkage. STR loci that are physically closer in a chromosome tend to stay together, as the chance of suffering separation through cross over is lower, meaning that variation in close loci is smaller than in loci far apart. In the same manner, STR can suffer mutations, albeit these may come not in the change of nucleotides, but at times in the number of overall repeats [33].

The number of possible repeats in a certain locus itself is not a very large number, but the number of combinations when accessing an entire array of loci quickly becomes large enough for a particular combination of different alleles to be accepted as a fingerprint (**Table 1**).

Table 1- Frequently used STR loci in DNA profiling accepted by the various databases. Adapted from *"E. Giardina,* "DNA Fingerprinting" *Brenner's Encycl. Genet. Second Ed., vol. 2, pp. 356–359, 2013."*

Locus (UniSTS)	*Chromosome location*	*Physical position GRCh37 assembly (Mb)*	*Category and repeat motif[a]*	*Allele range*
D1S1656 (58809)	1q42	Chr 1 230.905	Compound TAGA	8–20.3
TPOX (240638)	2p25.3 thyroid peroxidase,10th intron	Chr 2 1.493	Simple AATG	4–16
D2S441 (71306)	2p14	Chr 2 68.239	Compound TCTA/TCAA	8–17
D2S1338 (30509)	2q35	Chr 2 218.879	Compound TGCC/TTCC	10–31
D3S1358 (148226)	3p21.31	Chr 3 45.582	Compound TCTA/TCTG	6–26
FGA (240635)	4q31.3 alpha fibrinogen, third intron	Chr 4 155.509	Compound CTTT/TTCC	12.2–51.2
D5S818 (54700)	5q23.2	Chr 5 123.111	Simple AGAT	4–29
CSF1PO (156169)	5q33.1 c-*fms* proto-oncogene, sixth intron	Chr 5 149.455	Simple AGAT	5–17
SE33 (ACTBP2) (none reported)	6q14 beta-actin-related pseudogene	Chr 6 88.987	Complex AAAG	3–49
D7S820 (74895)	7q21.11	Chr 7 83.789	Simple GATA	5–16
D8S1179 (83408)	8q24.13	Chr 8 125.907	Compound TCTA/TCTG	6–20
D10S1248 (51457)	10q26.3	Chr 10 131.093	Simple GGAA	7–19
THO1 (240639)	11p15.5 tyrosine hydroxylase, first intron	Chr 11 2.192	Simple TCAT	3–14
vWA (240640)	12p13.31 von Willebrand factor, 40th intron	Chr 12 6.093	Compound TCTA/TCTG	10–25
D12S391 (2703)	12p13.2	Chr 12 12.450	Compound AGAT/AGAC	13–27.2
D13S317 (7734)	13q31.1	Chr 13 82.692	Simple TATC	5–17
Penta E	15q26.2	Chr 15 97.374	Simple AAAGA	5–32
D16S539 (45590)	16q24.1	Chr 16 86.386	Simple GATA	4–17
D18S51 (44409)	18q21.33	Chr 18 60.949	Simple AGAA	5.3–40
D19S433 (33588)	19q12	Chr 19 30.416	Compound AAGG/TAGG	5.2–20
D21S11 (240642)	21q21.1	Chr 21 20.554	Complex TCTA/TCTG	12–43.2
Penta D	21q22.3	Chr 21 45.056	Simple AAAGA	1.1–19
D22S1045 (49680)	22q12.3	Chr 22 37.536	Simple AAT	7–20

A battery of four and five nucleotide long repeats (tetranucleotide and pentanucleotide motifs, respectively) STRs that are completely delinked and suffer low rates of mutations are now used to create a DNA profile, with CODIS (Combined DNA Index System), the system used by the Federal Bureau of Investigation of the United States, using 20 loci as a standard [34], the UK's National DNA Database DNA-17's system using 17 loci

[35], and the European Interpol operating on a cross-referencing overlapping system from several contributing countries with different numbers of required loci [36]. Alongside individual identification, STR based DNA profiles are also used for kinship tests due to the hereditary nature of STRs. Such tests are conducted using a battery of STRs similar to DNA profile batteries, ranging from 15 to 23 different loci [37], yet, unlike DNA profiles, operates solely on percentual chances calculated upon the presence or absence of similar alleles of the tested loci checked between the donor, known familiars if existent and individual whose kinship is in question.

Outside of these individual identification and relation identification techniques that have been consolidated and accepted as viable evidence, all other applications of DNA, genetics and epigenetics in forensics are still in very early stages of study and development and/or need of applicability, such as the epigenetic determination of age and *postmortem* intervals in the first case and the use of epigenetic markers to distinguish the origins of DNA samples found in the crime scene, as coming from biological sources or created by laboratory procedures and planted in the local in the latter [30].

3. Model Organisms

3.1. Zebrafish

Zebrafish is a freshwater vertebrate from the *Cyprinidae* family hailing from the southeastern Himalayan region [38] with populations near Asiatic regions such as Pakistan, Bangladesh, India and Nepal [39]. The species was named by Francis Buchanan-Hamilton [40], a Scottish physician that spent many years studying the flora and in the southern India.

Zebrafish is a small fusiform fish with 5 longitudinal blue stripes that run from its branchial arch until the end of the caudal fin, with extra stripes running parallel on the pelvic fins. The color of the complementary stripes is different from males to females, prior ones having golden stripes and later ones possessing white to silvery stripes **(Fig. 10)**. Individuals can grow up to 6 centimeters with a mean size of 4 centimeters in laboratory conditions [41] and can survive up to 5 years with the average lifespan being roughly three and a half years.

Fig. 10- A male zebrafish (top, golden stripped) and a female zebrafish (bottom, silver stripped). Adapted from *http://www.brainn.org.br/zebrafish-um-peixe-com-vocacao-para-pesquisas-geneticas/*.

It's an organism with rising popularity as a model organism over the years to be used for a wide array of genetic studies. Albeit being further taxonomically from humans than mice, Zebrafish share many orthologue genes with humans, with 71.4% of human genes having at least one zebrafish orthologue, and of those, 47% have a one-to-one relationship with a zebrafish orthologue [42].

Zebrafish also provide some benefits over mice as model organisms through the sheer number of individuals that can be bred, matured, monitored and used for testing in a smaller time frame and laboratorial installations, allowing for far more iterations of the procedure for a much lower cost, providing an opportunity to amass much more data for data

driven studies while not sacrificing any quality or relevance. This is particularly useful when trying to amass sufficient data to create any sort of standard or even proof of concept.

In terms of DNA methylation, zebrafish possess a relatively simpler overall system for their epigenetic code, not needing genetic imprinting nor sex chromosome imprinting to produce viable individuals [43]. Despite that, the zebrafish's methylation machinery shares many important orthologs with mammalians, most importantly being DNMT1, DNMT2 and DNMT3 orthologues [43]. Zebrafish's DNMT1 and human's DNMT1 share a 73% similarity overall, with a 89% identity in the catalytic domain [43], as well as an *uhrf1* orthologue with 66% similarity [44]. Ubiquitin like with PHD and ring finger domains 1 (*uhrf1*) is a necessary protein for correct embryo development and tissue proliferation [44] as it binds to hemi-methylated DNA, recruiting DNMT1. Zebrafish's DNMT4 also possesses a 68% amino acid sequence identity with human DNMT3b, the protein responsible for many *de novo* methylations [43]. Human DNMT3a is also represented in the zebrafish's machinery by the orthologues DNMT6 and DNMT8, with an overall 77% identity and 89% catalytic domain [43]. The presence of these proteins ensures that zebrafish have *de-novo* and maintenance methylation akin to mammals, making them an appropriate organism to base early methylation related studies on.

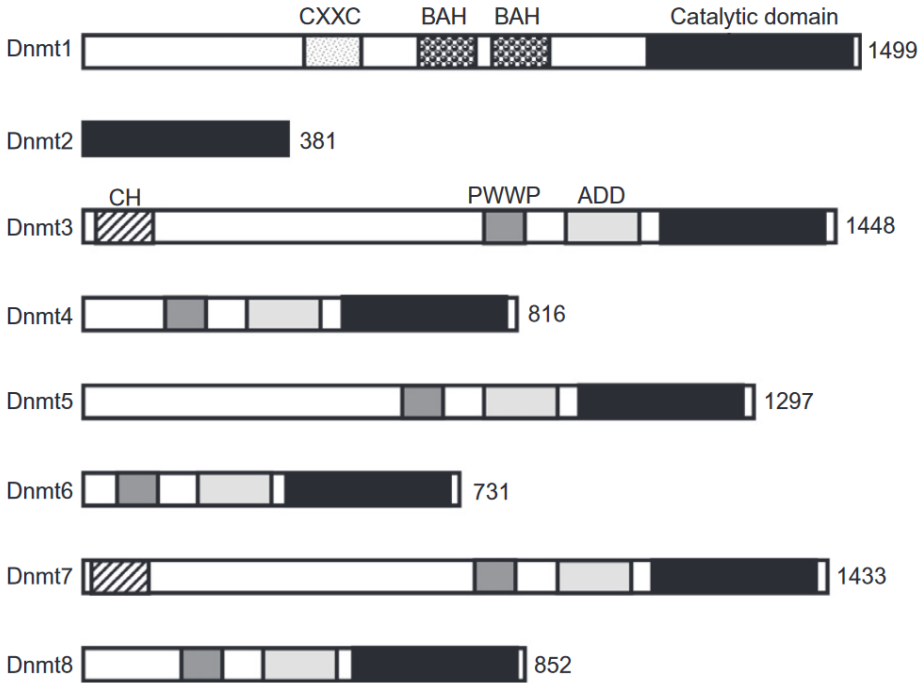

Fig. 11- Structure of the eight DNMT orthologues present in zebrafish. There is bromo-adjacent homology domain (BAH) on DNMT1, calponin homology (CH) on DNMT3 and DNMT7, cysteine-rich domain (CXXC) on DNMT1, conserved proline-tryptophan-tryptophan-proline domain (PWWP) throughout DNMT3 to DNMT8, and ATRX-DNMT3-DNMT3L domain (ADD) throughout DNMT3 to DNMT8. Adapted from *"N. Shimoda, K. Yamakoshi, A. Miyake, and H. Takeda, "Identification of a gene required for de novo DNA methylation of the zebrafish* no tail *gene" Dev. Dyn., vol. 233, no. 4, pp. 1509–1516, 2005."* [38]

3.2. Laboratory Mouse

Mus Musculus is one of the most widespread model organisms to date, being only rivaled by *Drosophila Melanogaster*. The laboratory mouse is a small furry rodent, with an average of 8.75 cm from snout to tail and an average tail length of 7.5 cm with a fur coat ranging from black to brown to white with a life expectancy of two to three years in average, with different strains affecting the range.

Fig. 12- A C57BL/6 mouse (left) and a BALB/c mouse (right). Adapted from *https://www.jax.org/*

The species began being inbred for scientific purposes in 1907 by Clarence Cook Little [45], with two years afterwards being obtained two mice with the recessive genes three of the characteristics in study. Descendants of these two mice where then bred over multiple generations, creating the first strain, dba [45]. From that strain, many more were created alongside others that were bred from wild type progeny (**Fig. 13**).

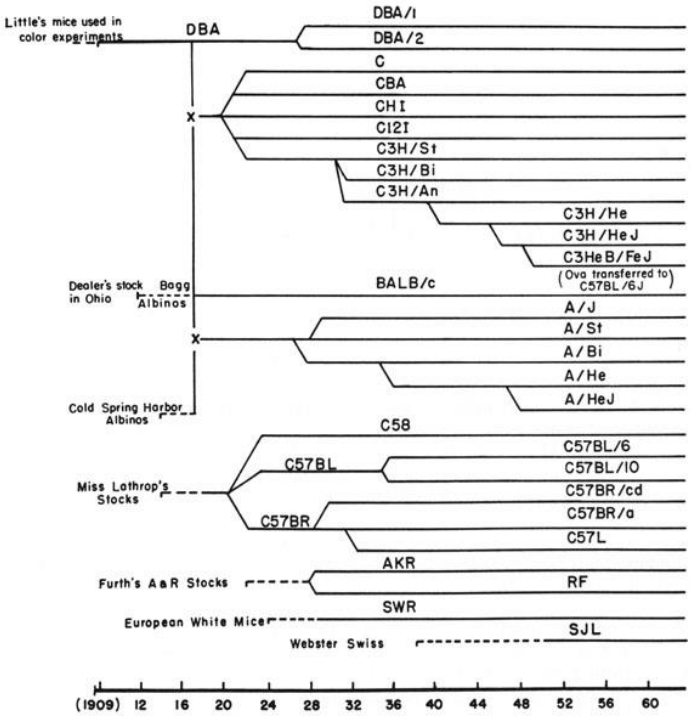

Fig. 13- Laboratory mouse strains tree from 1909 to 1960. Adapted from **"***Roscoe B. Jackson Memorial Laboratory,* Biology of the laboratory mouse, *2nd ed. New York: Dover Publications, Inc, 1966."*

From then on, many more strains were developed by many laboratories, many specifically created, either for just certain studies or for valued traits, with as of 2017 the total mice references and citations used totaling over 69000, with inbred strains being the most used at over 23000 references, in the PUBMED database [46]. From these numbers, the two most used and referenced strains where the C57BL/6 and the BALB/c strains (

Fig. 12). The strain C57BL/6 was the first one to have its entire genome sequenced, in 2012 and it continues being the preferred strain due to this fact and also because it's a generally stable strain that's also easy to breed, also being affected by less pathologies that its peers [46].

The laboratory mouse shares a more similar methylation mechanism to humans than the zebrafish, with the DNMT1 orthologues sharing a 76% protein identity, DNMT3a orthologues a 92% protein identity, DNMT3b orthologues a 76% protein identity and DNMT3L orthologues a 57% protein identity. These proteins share the same function as in humans, DNMT1 maintaining DNA methylation, DNMT3a and DNMT3b being responsible for *de novo* methylation, with the regulatory function of DNMT3L. Mice do possess a third *de novo* DNMT, named DNMT3c, evolved through DNMT3b duplication, responsible for the methylation of promoters of recent retrotransposons in the male germ line, being necessary for male fertility [47].

4. Candidate Biomarkers

4.1. Global DNA Methylation

Global DNA methylation is the first biomarker that should be taken into consideration in the context of a forensic epigenetic analysis as it can function both as a preliminary viability check for the sample and as an internal reference to contextualize all further data obtained. Global DNA methylation decays naturally, both *in vivo* and *postmortem*, at yet unknown rates *postmortem*. These rates are key to the overall process of determining the PMI of a body using epigenetics. This is due to two facts: either the technique requires input DNA normalization which renders the global DNA methylation as the only way to normalize the epigenetic information of a given sample to the standard ladder or the technique can accept various ranges of input DNA, which can severely boost or diminish overall DNA methylation values in specific biomarkers.

4.2. Housekeeping genes

One of the characteristics that is highly valued in a forensic science is its stability against adverse conditions, yet for the purpose of PMI estimation, stability of the methylation status throughout the organism's life would be key. Genes that keep their methylation stable as an organism ages, changes and suffers environmental stresses would provide a biomarker with less variance than other genes. Candidates for these kinds of genes would be housekeeping genes, genes that are absolutely necessary for the most basic functions of a cell, being expressed at basal levels in all cells of all tissues at all stages of development, even though the vast, vast majority of pathologies [48], albeit the amount of expression basal is might differ from author to author. Eisenberg *et al.* definition of housekeeping genes is to have a geometrical mean expression over 50 Reads Per Kilobase Million in RNA-seq and no tissue having expression deviation from the geometrical mean by a factor of two or superior (**Table 2**). As their expression is required, these genes may be in a permanent state of relative hypomethylation to no methylation, requiring a more sensitive technique for their study in the first case or be completely disregarded in the latter, or be used as a negative internal control for the standard.

Table 2- List of housekeeping genes. Data partially adapted from "*E. Eisenberg and E. Y. Levanon,* "Human housekeeping genes, revisited" *Trends Genet., vol. 29, no. 10, pp. 569–574, 2013.*", "*J. Pampel,* "Housekeeping genes" *2017. [Online]. Available: https://www.genomics-online.com/resources/16/5049/housekeeping-genes/. [Accessed: 04-Sep-2019]*" [49] and *"C. Zhang, Y. Q. Wang, G. Jin, S. Wu, J. Cui, and R. F. Wang*, "Selection of reference genes for gene expression studies in human bladder cancer using SYBR-green quantitative polymerase chain reaction" *Oncol. Lett., vol. 14, no. 5, pp. 6001–6011, 2017."* [50]

Symbol	Name	Function
Eisenberg *et al.*'s list of reference genes		
C1orf43	Chromosome 1 open reading frame 43	
CHMP2A	Charged multivesicular body protein 2A	Core component of the transport complex III (ESCRT III), which is involved in all ESCRT machinery mediated processes [51].
EMC7	ER membrane protein complex subunit 7	
GPI	Glucose-6-phosphate isomerase	Interconverts glucose-6-phosphate and fructose-6-phosphate. Core cellular respiration enzyme.
PSMB2	Proteasome subunit beta type 2	Component of the 20S core proteasome complex involved in the proteolytic degradation of most intracellular proteins [52].
PSMB4	Proteasome subunit beta type 4	Component of the 20S core proteasome complex involved in the proteolytic degradation of most intracellular proteins [53].
RAB7A	Member RAS oncogene family	Involved in all stages of endosomal maturation, all microtubule directed endosomal migration and positioning, endosome-lysosome transport, core protein in growth factor mediated cell signaling, nutrient-transporter mediated nutrient uptake, neurotrophin transport in the axons of neurons and lipid metabolism [54].

REEP5	Receptor accessory protein 5	
SNRPD3	Small nuclear ribonucleoprotein D3	Core component of the spliceosome, a nuclear ribonucleoprotein complex that catalyzes RNA splicing [55].
VCP	Valosin containing protein	Part of the AAA ATPase family of proteins. Performs necessary functions in protein degradation, intracellular membrane fusion, DNA repair and replication, regulation of the cell cycle, and activation of the NF-kappa B pathway [56].
VPS29	Vacuolar protein sorting 29 homolog	Component of the retromer cargo-selective complex, core functional component of retromer or respective retromer complex variants that prevents missorting of transmembrane cargo proteins into the lysosomal degradation pathway [57].
Other reference genes / housekeeping genes		
RRN18S	18S ribosomal RNA	18S rRNA is a component of the 40S small eukaryotic ribosomal subunit, constituting the structural RNA for the small component of eukaryotic cytoplasmic ribosomes. Proposed to be used as a reference gene in renal cell carcinoma [58].
RPS23	40S ribosomal protein S23	A protein that is a component of the 40S eukaryotic ribosomal subunit.
RPS13	40S ribosomal protein S13	A protein that is a component of the 40S eukaryotic ribosomal subunit.
ATP5B	β- ATP Synthase F1 Subunit	Beta subunit of the catalytic core of the mitochondrial membrane ATP synthase [59].
HSP90AB1	Heat Shock Protein 90kDa Alpha (Cytosolic), Class B Member	Molecular chaperone involved in maturation, structural maintenance and regulation of specific target proteins involved in cell cycle control and signal transduction [60].
S100A6	S100 Calcium Binding Protein A6	Acts as calcium sensor and modulator in cellular calcium signaling [61].
UBB	β-Ubiquitin	Ubiquitin is a highly conserved protein that has a major role in 26S proteasome mediated cellular proteins degradation. It has other roles in the maintenance of chromatin structure, regulation of gene expression and stress response [62].

POLR2A	RNA Polymerase II Subunit A	It's both the largest and catalytic component of RNA polymerase II, responsible for synthesizing mRNA precursors functional non-coding RNAs [63].
ACTB	β-Actin	Major component of the contractile apparatus and one of the two non-muscle cytoskeletal actins [64].
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	Catalyzes the reversible oxidative phosphorylation of glyceraldehyde-3-phosphate in the presence of inorganic phosphate and nicotinamide adenine dinucleotide (NAD). Core cellular respiration enzyme [65]. Proposed to be used as a reference gene in renal cell carcinoma [58].
PGK1	Phosphoglycerate kinase 1	Catalyzes the conversion of 1,3-diphosphoglycerate to 3-phosphoglycerate. Core cellular respiration enzyme [66].
PPIA	Peptidylprolyl isomerase A	Speeds up protein folding [67].
RPL13A	Ribosomal protein L13a	A protein that is a component of the 60S eukaryotic ribosomal subunit. Has a minor role in the repression of inflammatory genes as a component of the IFN-gamma-activated inhibitor of translation (GAIT) complex [68].
RPLP0	Ribosomal protein, large, P0	A protein that is a component of the 60S eukaryotic ribosomal subunit [69].
B2M	β-2-microglobulin	Component of the class I major histocompatibility complex involved in the presentation of peptide antigens to the immune system [70].
YWHAZ	Tyrosine 3-monooxygenase / tryptophan5-monooxygenase activation protein, zeta polypeptide	Highly conserved protein that mediates signal transduction by binding to phosphoserine-containing proteins [71]. Proposed to be used as a reference gene in renal cell carcinoma [58].
SDHA	Succinate dehydrogenase complex, subunit A, flavoprotein	Core catalytic subunit of succinate-ubiquinone oxidoreductase, part of a core cellular respiration enzyme [72].
TFRC	Transferrin receptor	Protein responsible for cellular iron intake through the receptor-mediated endocytosis of ligand-occupied transferrin receptor into specialized endosomes [73].
GUSB	β-Glucuronidase	Hydrolase that catalyzes the anabolism of glycosaminoglycans [74].

HMBS	Hydroxymethylbilane synthase	Third enzyme of the heme biosynthetic pathway catalyzing the head-to-tail condensation of four porphobilinogen molecules into the linear hydroxymethylbilane [75].
HPRT1	Hypoxanthine phosphoribosyltransferase 1	Catalyzes the conversion of guanine into guanosine monophosphate and hypoxanthine into inosine monophosphate, being a core enzyme of the purine salvage pathway [76].
TBP	TATA box binding protein	General transcription factor that is at the central enabling role in the DNA-binding multiprotein factor TFIID's function [77].

4.3. LINE-1 and repetitive DNA

LINEs, or long interspersed nuclear elements, are a form of repetitive DNA that constitutes roughly 21% of the human genetic code, that in conjunction with other form of repetitive DNA constitute half of entirety of the human genome [78]. These repetitive DNA section can be categorized firstly in two ways: sections that repeat themselves in succession, back to back, are denominated satellites and sections that repeat themselves throughout the entire genome in small patches are denominated interspersed repeats, that are derived from retrotransposons. These interspersed repeats can be subdivided even further by the presence or absence of a long terminal repeat (LTR), defining LTR retrotransposons and non-LTR retrotransposons. LINE-1s are the denomination of the latter, retrotransposon based mobile interspersed repeats without long terminal repeats of up to 6kb in length, with nearly 500000 copies in the human genome, 3000 to 4000 copies present in a non-truncated form and 30 to 100 copies being in their active retrotransposon form, constituting in total 17% of the human genome [78]. They are hypermethylated in normal cells, with constant maintenance of the methylation status through the pattern keeping mechanisms from DNMT1, but suffer accelerated demethylation, requiring the recruitment of TET and DNMT3a to remake the methylation pattern *de novo* [25]. As a consequence of the large portion of the genome that LINE-1s comprise alongside their rigorous hypermethylation status, LINE-1 methylation, alongside other similar repetitive elements like *Alu*, a non-LTR short interspersed nucleotide element (SINE) that comprises roughly 10% of the human genome, can be used as a partial substitute for, or accompaniment to, global methylation levels [28], [79].

Another form of repetitive DNA present well disseminated throughout the human genomes is the large tandemly repeated, generally short-sequenced, non-coding satellite DNA. These large areas of repetitive DNA are present mostly in the centromeres of the chromosomes and some are present in the heterochromatin [80]. In the human genome there are several note-worthy satellites: Satellite α, Satellite β, Satellite 1, Satellite 2 and Satellite 3. The most ubiquitous of them, Satellite α, is present in the centromeres of all of the chromosomes, with repeating patterns of 171bp monomers, with the number of different monomers tandemly repeating in a certain order being called the higher order repeat,

identifying the chromosome it belongs to [81]. As such, every chromosome has a different sequence for its Satellite α, but as the area spans over several mega bases of highly identical arrays, making up to 3% of the total human genome [82], it becomes hard to sequence using short sequence reads, appearing homogenous in such cases [81]. These areas of the genome do undergo DNA methylation and regulation, and, albeit direct causation is still under debate and study, methylation instability in the form of hypomethylation of the satellite DNA appears correlated with chromosome instability, breakage and creation of micronuclei, with normal levels of methylation being inhibitors of recombination between the highly similar repetitive areas of the pairing chromosomes [82]. On oncology based studies, hypomethylation of satellite DNA was linked to the genomic instability that is common occurrence in, or may even cause, the pathology [83]. Therefore, these areas undergo strict methylation regulation, and so, they are prime candidate target biomarkers that, unlike LINE-1s, are not wide-spread enough to be used as a global methylation surrogate but rather biomarkers tested for variability in the battery.

4.4. Individual target biomarkers

As for possible individual genes that suffer a sustained, stable rate of degradation of their methylation pattern in *postmortem* bodies exposed to natural elements, not much information has been gathered by the general scientific community. In regards to the creation of a PMI assessing battery, the exhaustive study and search for potential individual target biomarker candidates will constitute most of the work needed to be done in order to create the biomarker battery after the pilot, proof-of-concept, forensic focused studies have firmly demonstrated that such methodology is viable. Most of the information that can be foraged from the current epigenetic studies pertain to *premortem* changes in many genes, be it susceptibility to environmental pressures, linkage to passage of time *premortem*, and other factors, denoting individual genes that may possess initial methylation patterns at time of death too variable to provide reliable results in a standard. Tentative lists of genes that fall under this category are located under Chapter 6.3, Obstacles to Overcome, and Table 2.

5. Laboratorial Techniques

5.1. The Scope

The sheer number of techniques that were created with the sole purpose of, or that can be used to, study DNA Methylation is extensive (**Fig. 14**). They can be divided into three main groups: bisulfite conversion-based techniques; techniques that leverage affinity to methylation antibodies; restriction enzyme-based techniques. There is also a smaller group of broader quantitative and qualitative techniques that are used for other areas of expertise that can be used to measure methylation through small clever protocol tweaks. These three main groups contain a plethora of different options, far too many for the scope of this work, so a selection of procedures from each group fitting to the end goal of using epigenetics to determine PMI shall be explored.

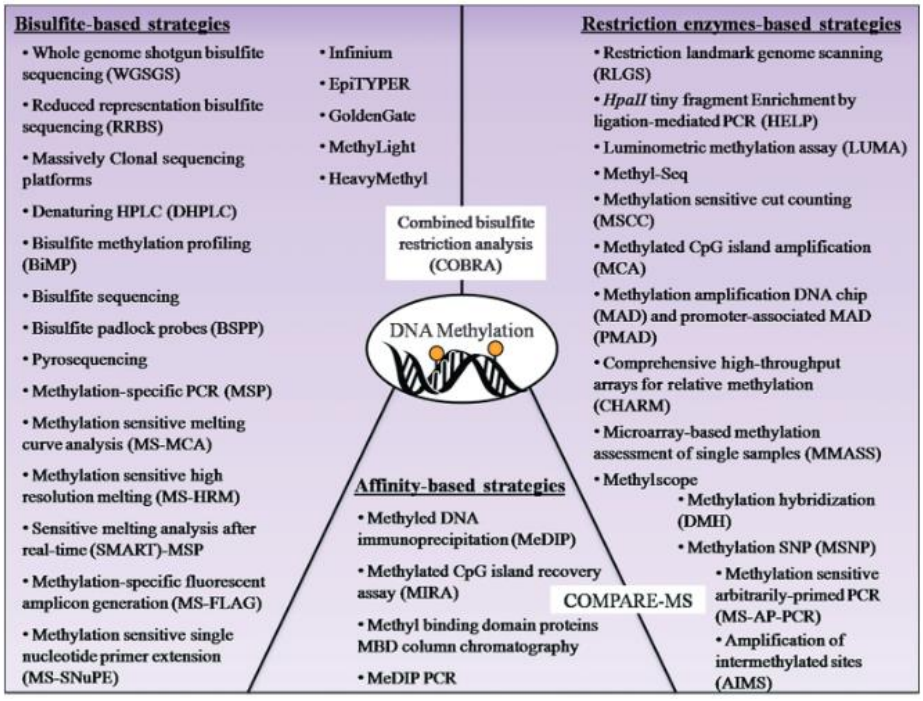

Fig. 14- A comprehensive list of the most used DNA methylation assessing techniques. Adapted from "*E. Olkhov-Mitsel and B. Bapat, "*Strategies for discovery and validation of methylated and hydroxymethylated DNA biomarkers*" Cancer Med., vol. 1, no. 2, pp. 237–260, 2012."* [84].

5.2. Bisulfite Conversion

Sodium bisulfite conversion consists on the mutation of unmethylated cytosine residues into uracil residues, and subsequent conversion to thymine if converted to cDNA, effectively translating epigenetic information into genetic information, making it an absolutely cornerstone technique for many epigenetic studies. Initially, bisulfite causes the sulphonation of an unmethylated cytosine on the 6' carbon, adding a $-SO_3$ group, turning it into a cytosinsulphonate. With a following hydrolytic deamination, cytosinsulphonate is converted into uracilsulphonate, an uracil residue with a $-SO_3$ group on his 6' carbon. Finally, a desulphonation returns uracilsulphonate to a normal uracil residue, thus completing the conversion [85] **(Fig. 15)**. The resulting DNA will reveal the presence of methylation in a certain cytosine residue if when matched with unconverted DNA no change has been caused and will reveal the lack of methylation when an uracil, or thymine if processed, is present in a cytosine location.

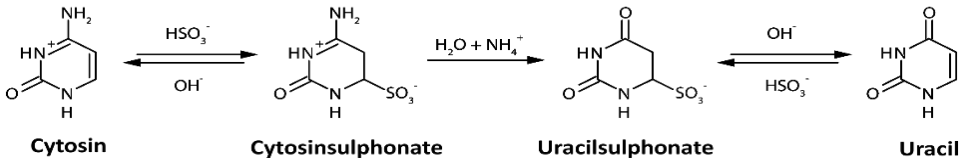

Fig. 15: Depiction of the four stages of a Bisulfite conversion of cytosine.

Many of the methylation accessing techniques require prior bisulfite conversion to enable them to obtain the pretended results. Methylation-Specific PCR, HRM Analysis, COBRA Methylation Assay and Bisulfite Pyrosequencing require the bisulfite conversion [86], [87]. Common usage bisulfite conversion kits from Qiagen, Analytik Jena and Zymo Research were tested by efficiency, with results ranging from 98.7% to 99.9% [88]. Despite high conversion rates, bisulfite conversion protocols are carried under very unfavorable conditions for DNA stability, resulting in longer conversion protocols having higher rates of DNA fragmentation [88].

5.3. PCR, qPCR and Methylation-Specific PCR (MSP)

Polymerase Chain Reaction (PCR) was one of the biggest breakthroughs in recent story for many fields of science. In the last 36 years, from biology to forensic sciences, since the technique was created and refined by Kary Mullis in 1983, the ability to amplify a sample of DNA quickly and efficiently has grown to a day-to-day basic necessity for many laboratories [89].

The technique consists on a simplified *in-vitro* recreation of the DNA replication that occurs on living organisms. It occurs in 3 main phases:

- Denaturation: A small portion of the DNA sample or complementary DNA created out of an originally RNA sample is heated enough to melt its double helix structure and create single stranded DNA.

- Annealing: The temperature is lowered just enough to allow the primers to bind to the DNA around the area of interest to elongate. The design of the primers is critical both to achieve total annealing of the primers at a temperature that does not allow for non-specific annealing between single stranded DNA and to achieve amplification of the specific genetic material of interest in the case of qualitative PCR.
- Extension: The temperature is adjusted towards the optimal catalytic conditions for the DNA polymerase in use, usually Taq DNA Polymerase, which will elongate the DNA chain from the primers using the single strand DNA as a base. By the end of this phase, the total amount of DNA present will have theoretically doubled. [90]

The reaction is run for several cycles which fall into three categories differentiated by the amount of new copies generated per cycle:

- Initiation phase: In this phase, each cycle produces small amounts of new copies as there are few copies to copy from.
- Exponential phase: In this phase, each cycle produces larger and larger amounts of copies.
- Plateau phase: In this phase, each cycle slows down the amounts of copies made due to the rising scarcity of nucleotides to fuel the DNA Polymerase, until copies cease to be created and the reaction ends.

While PCR may be a simple process nearly completely automated in current days, many tweaks have been done to the core protocol in order to use the technique as a method of obtaining data by itself, and not merely be used to obtain enough sample for posterior analysis. Real time PCR (qPCR) is one of these variants of the PCR technique capable of providing data about the quantity of replicated DNA present in each cycle in real time. Through the addition of fluorescent dyes like SYBR Green I that only fully activates when intercalated with double-stranded DNA, qPCR allows the total quantification of replicated sequences plotted against time. If short sequence-specific DNA or RNA sequences labelled with a fluorescent reporter and a quencher that is active until hybridization, known as hybridization probes, are instead added, qPCR allows for the quantification of the replication of precise DNA sequences in a sample containing several sequences plotted against time.

Methylation-specific PCR (MSP) is another tweak to the PCR protocol, taking basis on the regular qPCR technique, tailored to be sensitive towards methylation patterns within the DNA sample provided. This sensitivity is achieved through the usage of primers designed to be complementary to the sequence in the targeted area of interest post bisulfite conversion [22]. Since methylated cytosines are protected from the mutation induced by the bisulfite conversion, epigenetic information is transformed into sequence information that can be targeted with primers that will only anneal to the sample DNA if it had previously a certain methylation pattern. As such, methylation data is read through difference in Ct values. The downside of this technique is that the methylation locations accessed by it is limited by the numbers of primers available, and each addition of methylation location testing primer incurs higher costs.

5.4. HRM Analysis

The HRM analysis consists on the comparison between the melting temperature of a double strand of DNA to another, in order to either spot differences like point mutations between very similar sequences or group several sequences according to the variable in study, with DNA methylation being the one in focus on this work [91].

A typical HRM analysis starts with a PCR in order to amplify the samples of interest to be compared. Then, with the addition of intercalating fluorescent dyes to allow the process to be monitored in real time like qPCR and a slow and steady warming of the samples, the melting of the DNA is registered for each sample in the form of a graph depicting fluorescence against temperature, in a downward fashion, when normalized, due to the fact as more DNA denatures, more of the intercalating dyes are released and overall fluorescence diminishes. Then, the samples are compared based on the overall graphic and melting temperature, the temperature at which 50% of the double strands have melted. This is possible because the two different base pairs have different proprieties. The G-C base pairing is held by 3 hydrogen bonds while the A-T base pairing is held by 2, making a G-C base pairing heavy sequence more resistant to melting than a A-T heavy one [18]. Even single point mutations replacing one pairing with another can create a significant enough difference between the melting temperatures to be detectable and registered.

Paired with MSP, an HRM analysis can be done in the same protocol in order to further validate the obtained results. If a standard is created using DNA methylation for post mortem analysis, a quick HRM analysis can be used to group the standard ladder and place the sample within one of those groups, either authenticating the MSP results by grouping with the same range of PMI or denoting an error in the process.

5.5. ELISA Assay Technique

The ELISA assay technique, or enzyme-linked immunosorbent assay, is a plate-based assay technique that can be used to quantify a myriad of molecules ranging from peptides to macromolecules.

Albeit several types of ELISAs exist, every type has its foundation on the binding of antibodies to the antigen to quantify the intended molecule, depending merely on how direct the measurement is done. In direct capture ELISAs, the complex reactional mixture where the antigen to quantify resides is simply placed within the well, where the antigen is bound to. In indirect capture ELISAs, well-bound antibodies are selective to the antigen to quantify, binding it to the well [92].

Disregarding the capture type, the detection of the antigen to quantify can also be separated into two types. In direct detection, the wells are flooded with selective antibodies conjugated with a reporter which will bind to the antigens in a one-on-one fashion. In indirect detection, also known as Sandwich ELISA, firstly the wells are flooded with unconjugated antibodies that will bind to the antigen. Afterwards, the wells are flooded with antibodies conjugated with reporters that will bind to the first antibodies, with one of these first ones

being able to bind more than one conjugated antibody [92]. Finally, in both detection types, the wells are washed and then substrate is added to activate the reporters and allow quantification. As indirect detection allows more reporters per antigen, it suffers signal amplification, augmenting the technique's sensitivity, being currently the most robust option [92].

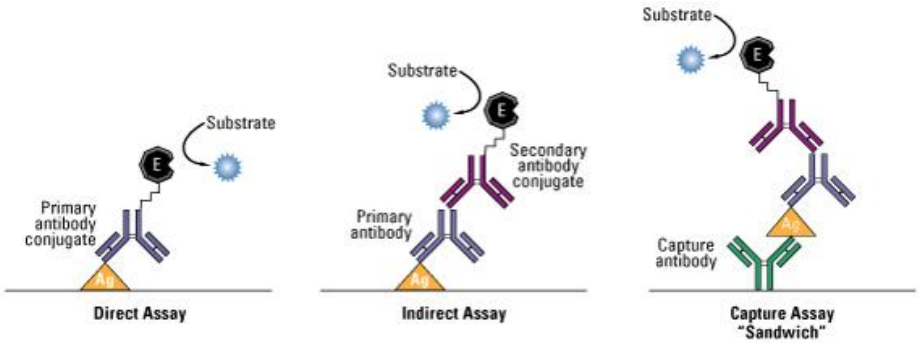

Fig. 16- Illustration of the two types of ELISA detection, direct and indirect detection, on a direct capture protocol (left and middle) and the *sandwich* Elisa diagram (right). Adapted from: *"ThermoFisher, "Overview of ELISA." [Online]. Available: https://www.thermofisher.com/pt/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/overview-elisa.html. [Accessed: 18-Jul-2019]".*

Antigens can be designed to bind to methyl groups, so when DNA samples are tested through ELISA based assays, methylation information is read and amplified directly from the source, although no gene specificity can be achieved without previous DNA cutting and only inserting the wanted genes into the reaction wells, adding more steps to the protocol that can damage the source material. Even so, ELISA based assays are a prime candidate for global methylation assessment, with relative ease and low cost.

5.6. Methylated DNA immuno-precipitation (MeDIP)

Methylated DNA immuno-precipitation is a laboratorial technique based on immunocapturing that allows for the enrichment of methylated DNA [93], [94]. The protocol consists on the cleavage of DNA either by restriction enzymes or sonication to the desired fragment length, followed by the separation into two samples, input DNA that is left as is and immunoprecipitated DNA that will undergo the rest of the protocol. The immunoprecipitated DNA then undergoes denaturation and immunoprecipitation by Anti-5-methylcytosine (5-mC) antibodies and is purified (**Fig. 17**) [93], [94]. The protocol itself does not produce any data results, instead it produces two different samples from the original one that can be put through PCR, qPCR and oligonucleotide microarray protocols for the data to be accessible [94].

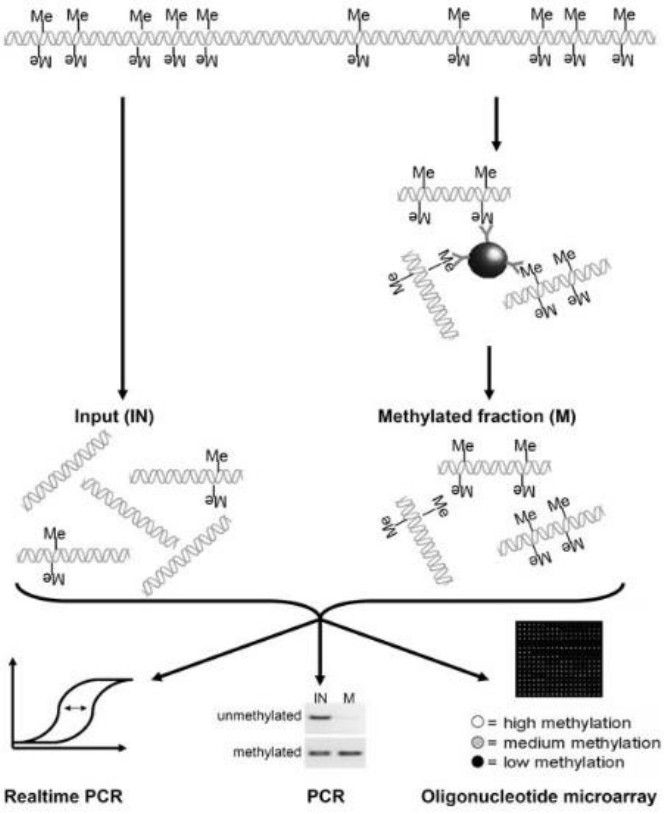

Fig. 17- The schematic of a MeDIP methylated DNA enrichment protocol. Adapted from *"F. Mohn, M. Weber, D. Schübeler, and T.-C. Roloff, "Methylated DNA Immunoprecipitation (MeDIP)" 2008."*

The data from qPCR and oligonucleotide microarrays are the most reliable for forensic usage, and covers the entire genome, with the latter one possessing an 80kb resolution [93]. In qPCR, the more methylated the original sample is, the closer the Ct value for the IP DNA will be in comparison to the Input DNA's Ct value. Paired with HRM analysis, this sequence of protocols could be a strong candidate for a fast, cheap and reliable protocol for PMI assessment, as DNA degradation and denaturation's effects on the initial steps of the protocol may be minimal.

5.7. Bisulfite Pyrosequencing

Bisulfite Pyrosequencing is a technique that allows the study of the methylation state of singular CpG locations at a quantitative level. It starts with a bisulfite conversion of the DNA sample, followed by a variation on the PCR protocol, modifying the primer complementary to the CpG location of interest, by adding biotin to the 5'-end of the primer. After PCR amplification, the biotin bound primer becomes the template for the pyrosequencing, that is isolated through denaturation of the double-stranded amplicon and filtration through streptavidin beads [95]. The high affinity between biotin and streptavidin separates the sample DNA and the amplified template primer. Then the technique follows a standard pyrosequencing protocol. The amplified template primer is incubated alongside:

DNA polymerase, present to allow the incorporation of dNTPs; ATP sulfurylase, that will catalyze the conversion of the pyrophosphate produced by the incorporation of dNTPs into ATP; luciferase, that will use the ATP as fuel to convert luciferin into oxyluciferin; apyrase, present to digest the remaining non-incorporated dNTPs; substrates needed for the enzymes, such as adenosine 5' phosphosulfate for ATP sulfurylase and luciferin for luciferase [87], [95]. After incubation, a solution of a single kind of dNTP is added to the reaction. DNA polymerase will try to incorporate them into the growing complementary strand and, if successful, release pyrophosphate that will trigger the cascading reaction and result in measurable light emission (**Fig. 18**). Since the light emitted is proportional to the amount of incorporated dNMT, the presence of light at the time of addition of a thymine will denounce a methylated cytosine in the original sample and presence of light at the time of cytosine addition will denounce an unmethylated cytosine. This is so because the template used for the pyrosequencing is a complementary strand to the original sample, possessing an adenosine on the CpG location in study for the first case and a guanine for the latter.

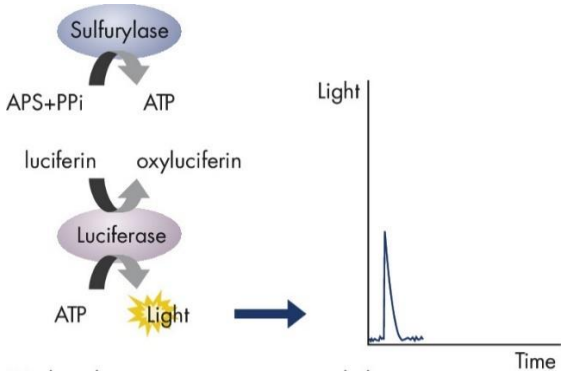

Fig. 18- The resulting pyrophosphate from the dNTP incorporation will allow the ATP sulfurylase to catalyze ATP, which is used by luciferase to catalyze luciferin into oxyluciferin, producing light. Apyrase present in the reaction will digest any unused dNTPs and ATP, resetting the reaction for the next addition of dNTP. Adapted from: *"QIAGEN, "Pyrosequencing Technology and Platform Overview" 2019. [Online]. Available: https://www.qiagen.com/us/service-and-support/learning-hub/technologies-and-research-topics/pyrosequencing-resource-center/technology-overview/. [Accessed: 16-Jul-2019]."*

The usage of single CpG Island methylation for forensic usage has not been studied, both in terms of identification and *postmortem* interval calculation most likely because its focus on small possible changes could be used more akin to SNPs in order to create a profile and not for bulk degradation analysis, but suffering the lack of enough polymorphism that SNPs provide and possible connection to lifestyle and health, making such small variances very unreliable. A comprehensive gene methylation assessment would provide a much more reliable source on information to be used.

5.8. HELP Methylation Assay

The HpaII tiny fragment Enrichment by Ligation-mediated PCR (HELP) assay technique cleaves the DNA using the HpaII restriction enzyme and the isoschizomer Msp1 restriction enzyme [96]. HpaII is an enzyme that cleaves unmethylated 5' – CCGG – 3' locations, while Msp1 cleaves any 5' – CCGG – 3' location regardless of methylation status. This creates two sets of cleaved DNAs differentiated by the methylation status. HpaII will not cleave methylated locations that Msp1 does, creating different length strands that will be used in the following ligation-mediated PCR as samples. It is necessary to note that for each cleavage site on any biomarker in stud, that total minus one different sets of primers must be designed that individually label each cleaved strand sample [97].

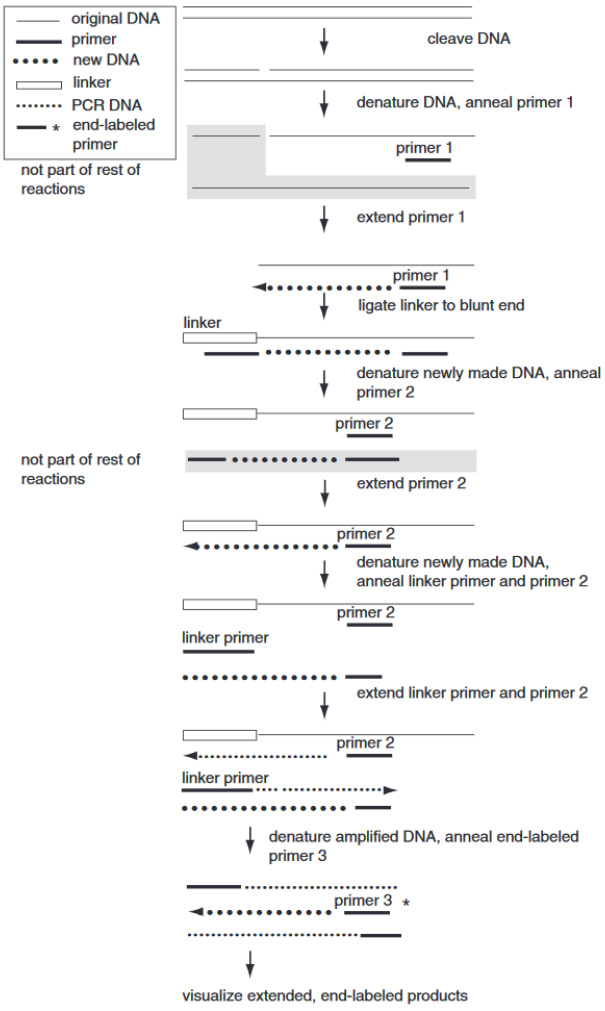

Fig. 19- Illustration of the protocol sequence of a HELP assay. Adapted from *"P. R. Mueller, B. Wold, and P. A. Garrity, "Ligation-Mediated PCR for Genomic Sequencing and Footprinting" in Current Protocols in Molecular Biology, 2004."*

Before resuming the standard PCR protocol, each cleaved strand sample must be converted into an amplifiable, labeled product. Firstly, the cleaved DNA is denatured, and the first primer is extended along the cleaved strand sample. Then the sample is 5'-end capped using a DNA ligase to add a linker. Now the capped sample can be used in a PCR using the second primer and a linker primer to exponentially duplicate the sample strand. Finally, on the last cycle, a labeled third primer is used for data analysis purposes (**Fig. 19**) [98].

This technique falls more on the qualitative analysis side than the quantitative side [97] and does not possess CpG island bias [99], although only 3.9% of all non-repeat CpG locations in the human genome fall under HpaII's recognized motif, severely limiting possible biomarkers [94]. Another hurdle for the usage of HELP for forensic protocols is that the original sample DNA must be intact [97]. DNA degradation may cause methylation-

unrelated cleavage points that skew the results, highly decreasing fidelity on longer PMIs, possibly making it unsuitable for forensic standards.

5.9. Luminometric Methylation Assay (LUMA)

The Luminometric Methylation Assay is another technique that is based on using the HpaII restriction enzyme and the isoschizomer Msp1 restriction enzyme, with the addition of EcoRI, a restriction enzyme that cleaves 5' – AATT – 3', as internal reference, followed by a pyrosequencing protocol [100]. The cleavage points of HpaII and Msp1 leave a 5' – CG – 3' overhang and the EcoRI cleavage points leaves an AATT overhang that are used in the pyrosequencing technique to obtain the results. In the first step of the pyrosequencing, dATPαS is added, filling the 5' – TT – 3' overhang on the extension process, giving out the first control signal. The second step consists in adding dCTP and dGTP to the sequence, filling all 5' – CG – 3' overhangs left by both HpaII and Msp1, giving out the sample results. Steps three and four are the addition of dTTP and another round of dCTP and dGTP. The first will finish filling the EcoRI overhang and provide the second and final internal reference results, while the second one is expected to not produce any signal, as all 5' – CG – 3' overhangs should be completely filled, serving as a protocol integrity check (**Fig. 20**)[100].

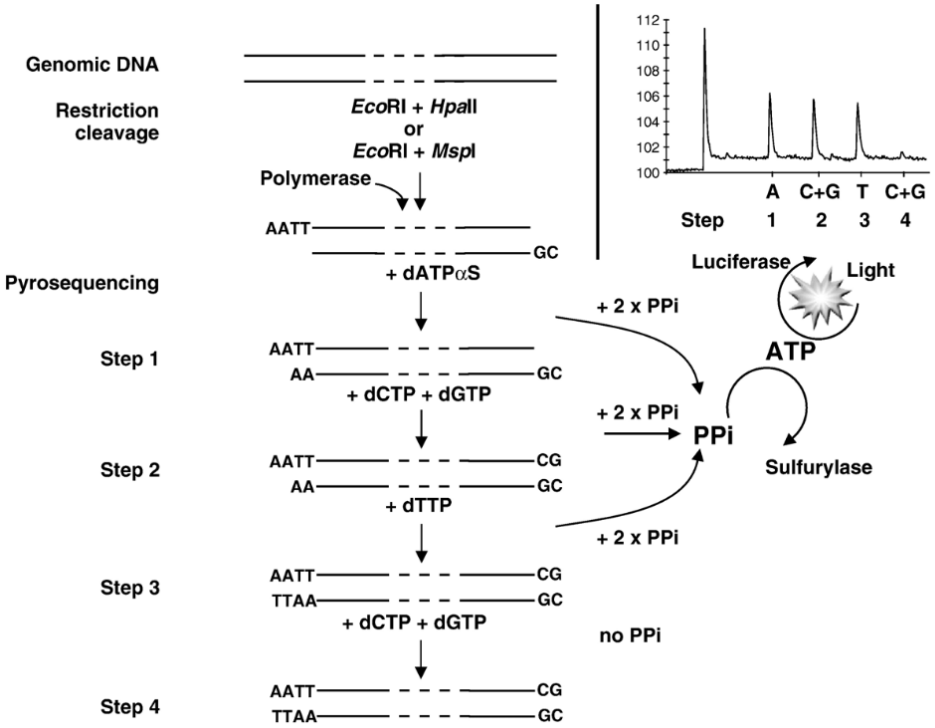

Fig. 20- Simplified visualization of a LUMA protocol. All data is retrieved in the form of a pyrogram (top right). Adapted from *"M. Karimi et al., "LUMA (LUminometric Methylation Assay)-A high throughput method to the analysis of genomic DNA methylation" Exp. Cell Res., 2006."*

Each sample is subjected to a HpaII/EcoRI and a Msp1/EcoRI treatment separately, with the results being given as a ratio between (HpaII/EcoRI) / (Msp1/EcoRI), becoming

HpaII/Msp1. This internal reference eliminates result variations depending on the amount of input DNA [100]. Another benefit of this technique above HELP assays is that the HpaII/EcoRI coefficient can be used in isolation, as it mirrors the HpaII/Msp1 coefficient, if the amount of available sample DNA is small, slashing the required amount in half (**Fig. 21**).

The LUMA assay is a prime candidate for a global methylation test on the entire genome for its high scalability and simple protocol, being a much more promising technique in the forensic areas than bisulfite pyrosequencing and HELP assays. The hurdle it may face when used in forensic sciences is that extensive DNA degradation might skew the results in the same way as in the HELP assay.

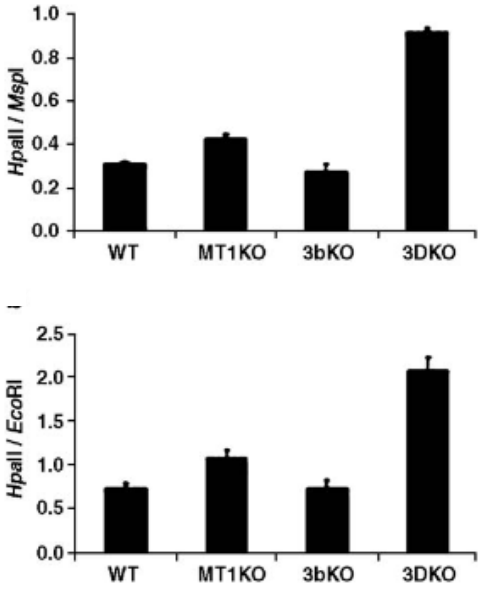

Fig. 21- LUMA DNA methylation analysis of four HCT116 cell lines, wild type (WT), DNMT1 knock-out (MT1KO), DNMT 3b knock-out (3bKO) and DNMT1 plus DNMT3 knock-out (3DKO), displaying HpaII/Msp1 coefficient on the upper graph and HpaII/EcoRI coefficient on the lower graph. Adapted from *"M. Karimi et al., "LUMA (Luminometric Methylation Assay)-A high throughput method to the analysis of genomic DNA methylation" Exp. Cell Res., 2006."*

5.10. COBRA Methylation Assay

COBRA methylation assays are yet another restriction-based protocol to determine methylation status. COBRA protocols start with a standard bisulfite conversion, followed by PCR amplification using methylation independent primers. After purification, *Bst*UI restriction is applied. Finally, through polyacrylamide gel electrophoresis, the data is analyzed (**Fig. 22**) [101].

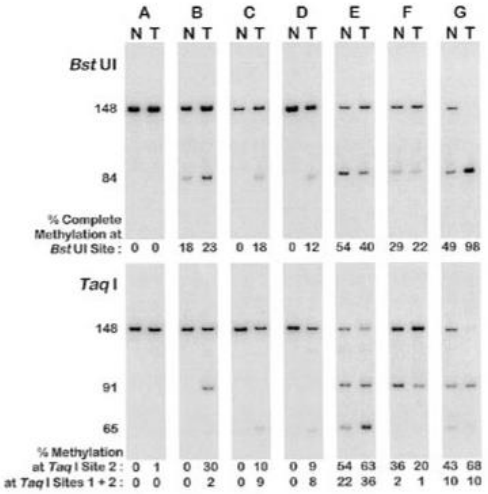

Fig. 22- An example of a final COBRA end product. The bisulfite conversion applied will create different restriction zones for *Bst*UI, resulting in two different molecular sized fragments that will separate under electrophoresis. The intensity of the stain shows not only the presence of methylation but its quantity.

5.11. Illumina Methylation Assay

The Illumina Infinium Methylation Assay is a bead-based assay technique that uses bisulfite conversion and beads containing primers for over 850,000 methylation sites [102], both for methylated and unmethylated variants in Infinium I and a single primer for each location in Infinium II that stops one base before the methylation location, in the Infinium MethylationEPIC Kit, alongside single-base extension using labeled ddNTPs in order to develop the data.

The workflow begins with a bisulfite conversion, followed by an amplification and fragmentation of the converted DNA, which is then incubated in the assay's chip. During incubation, DNA is denatured and bound to the primer beads, to which is then added labeled ddNTPs for the single base extension. The results come out as an array as seen in **Fig. 23**, which can be read by array scanners and have data automatically processed by Illumina's proprietary software [103].

Fig. 23- The resulting array of an Illumina Infinium Assay, each dot representing a distinct methylation point, with the possible results of green for unmethylated locations, red for methylated locations and yellow for locations with intermediate levels of methylation. Adapted from "Illumina, "Infinium HD Methylation Assay Protocol Guide (15019519)" no. November 2015."

The assay covers 99% of human Refseq genes, 95% of human CpG Islands, and ENCODE enhancer regions [104], at a single nucleotide resolution, across the entire genome [102], making it viable for both biomarker methylation assessment and inducing information about global at the same time.

5.12. LC-MS/MS

Liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) is a small molecule quantitative/qualitative combo technique that combines the separation power of liquid chromatography with the mass analysis power of mass spectrometry. Liquid chromatography is a technique that is able to separate many different components in a short amount of time by using a changing mobile phase against a stationary phase to elute the particles through several possible characteristics like particle size, particle charge, the particle's solute properties, *etc*. In this case, DNA that underwent complete lysis until nucleotides were completely separated, would then be separated and leave the elution tower at different times, with methylated cytosines leaving at a slightly different time than unmethylated cytosines.

Right after leaving the elution tower, the circuitry continues directly into the mass spectrometer. Right before entering the machine, the particles are ionized by an ion source,

required for the mass spectrometer to be able to magnetically separate the particles by mass per charge. Once separated, a single ionized particle is fragmented through Collision-induced dissociation, on which the particle is accelerated and collided with a neutral gas, such as noble gases. These fragments have a predictable mass per charge, meaning a cytosine would always produce a certain type of fragments while a methylated cytosine would produce slightly different particles. These fragments undergo another mass per charge separation and are then detected right after the first pass, creating the tandem effect. This amplifies the technique's sensitivity and ability to differentiate between very similar particles, a trait necessary due to the fact that the nucleotides that constitute the DNA create roughly similar fragments as well as the minor difference between methylated cytosines and unmethylated cytosines (**Fig. 24**) [105].

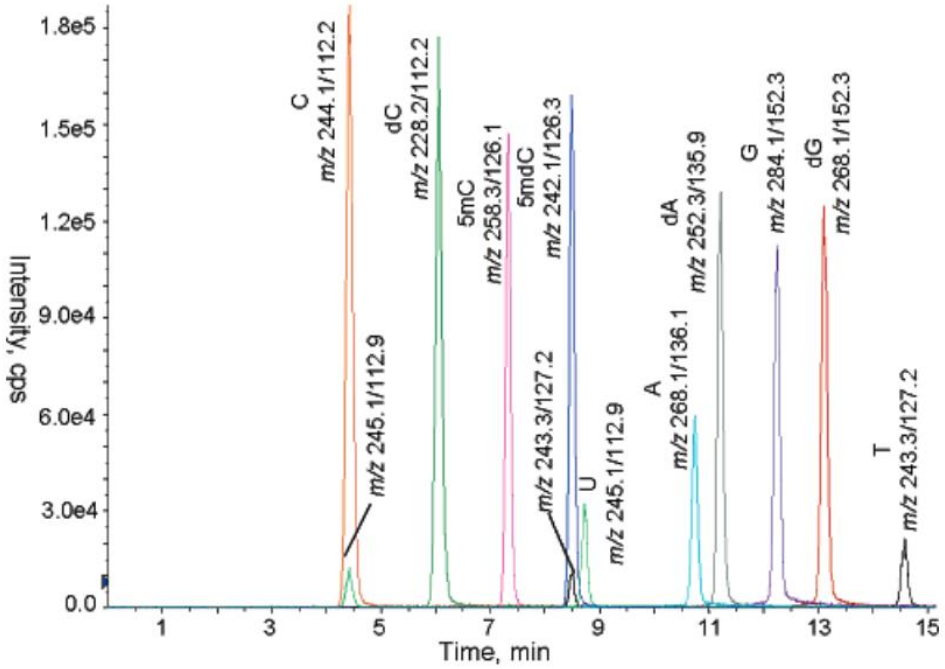

Fig. 24- LC-ESI-MS/MS chromatogram acquired by Song *et al.*'s pilot test obtained by testing 1ng of each nucleotide and variants through a standard LC-MS/MS protocol. Adapted from: **"***L. Song, S. R. James, L. Kazim, and A. R. Karpf*, "Specific method for the determination of genomic DNA methylation by liquid chromatography-electrospray ionization tandem mass spectrometry" *Anal. Chem., vol. 77, no. 2, pp. 504–510, 2005."*

5.13. Choosing a technique

With so many options at our disposal, each vastly different from the others, one must take into consideration the technique's strengths, weaknesses and the rigor needed for forensic usage.

Before delving into them, these techniques all have different applications. Techniques can excel in global methylation data, specific loci or a broad arrange of loci. LUMA, MeDIP, LC-MS/MS, ELISA based assays and COBRA assays are great for global DNA methylation data, while Bisulfite Pyrosequencing, MSP and Illumina methylation assays shine when it comes to providing gene specific methylation status, albeit Bisulfite Pyrosequencing could be used to mirror global methylation status as a surrogate if used to measure the methylation status of LINE-1. This means that to use global methylation as an internal standard for the PMI assessment, two techniques would be required, save for Bisulfite Pyrosequencing.

To make a choice, one must take into consideration the forensic scenario. Two possibilities exist in this case: all but small blood splatters were found and sampled, and it may be needed to know how old those splatters are to position the crime in a time frame; or a body was found, time of death is needed and there is no shortage of input DNA. When reaching an accredited laboratory for testing, the samples may be degraded in the case of a decomposing body recovery; in low quantity per sample, often under 10ng in the cases of blood splatters [106]; or in high quantities of different samples, many blood splatters around the body as an example, under the expectancy of a fast and precise turnaround of results. Many of the standard protocols for many of the techniques require DNA in both pristine quality and around the 1µg mark of quantity (**Table 3**), and, while in the presence of a body with ample DNA for testing all of the listed techniques are viable, very few provide solid data when provided little input DNA in the case of a missing body. The only technique that can be used with completely degraded DNA is the Liquid chromatography coupled with tandem mass spectrometry, as it requires no bisulfite conversion and relies on complete disassociation between all the DNA's nucleotides. It is also one of the very few global methylation accessing methodologies that doesn't require high quality DNA alongside, when considering the setting, high amounts of DNA to obtain acceptable levels of reliability. Methodologies that target specific loci, or an array of loci, generally can handle DNA with some degradation, some with very low quantities of DNA, being good news when considering the creation of a battery of biomarkers for forensic usage will require techniques that can target different loci with sparse, damaged DNA. Finally, considering that an internal sample control like global methylation may be required to normalize all the results, it seems to become more likely that more than one technique may be required for the process of PMI assessment, lest LINE-1 methylation perform acceptably as a surrogate or another way of normalizing the results within a certain technique is found viable.

One variable to also take into consideration when choosing a technique is the cost. This is highly subject to change, from technique to technique, from laboratory to laboratory, but when presented with a new emerging industry, in this case forensic epigenetic testing,

the prices tend to go down from the current research focused market, either through the establishment of grants, optimization of kit manufacturing when prompted with a higher product demand, increased competitivity, *etc.*, as seen by the example of the commercialization of whole genome sequencing as the next generation sequencing technology matured.

Table 3- Comparison between several DNA methylation accessing techniques. Data partially adapted from "*S. Kurdyukov and M. Bullock*, "DNA methylation analysis: Choosing the right method" *Biology (Basel)., vol. 5, no. 1, pp. 1–21, 2016."* [107] and "*E. Olkhov-Mitsel and B. Bapat, "*Strategies for discovery and validation of methylated and hydroxymethylated DNA biomarkers*" Cancer Med., vol. 1, no. 2, pp. 237–260, 2012."* [84].

Technique	Group	Amount of DNA Required	Quality of DNA Required
Global DNA methylation techniques (sequencing)			
Restriction landmark genomic scanning (RLGS)	Restriction enzyme digestion	~7µg [108]	High
HpaII tiny fragment Enrichment by Ligation-mediated PCR (HELP-seq)		1µg [96]	High
Methyl-seq		~3µg	High
Luminometric Methylation Assay (LUMA)		500ng to 250ng	High
Methylation-sensitive cut counting (MSCC)		5µg to 1µg	High
Methylated CpG island amplification (MCA-seq)		5µg [109]	High
Methylated DNA immuno-precipitation (MeDIP-seq)	Affinity-based	1µg to 100ng [110]	High
Methylated CpG island recovery assay (MIRA)		1µg to 1ng [111]	High
Whole genome shotgun bisulfite sequencing (WGSGS)	Bisulfite Conversion	10ng [112]	High

Reduced representation bisulfite sequencing (RRBS)	Bisulfite Conversion	~1µg	High
Denaturing HPLC (DHPLC)		100ng to 10ng [113]	Medium
Liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS)		1µg to 4ng	Low
Global DNA methylation techniques (microarray)			
HpaII tiny fragment Enrichment by Ligation-mediated PCR (HELP)	Restriction enzyme digestion	1µg [96]	High
Methylated CpG island amplification microarray (MCAM)		5µg [109]	High
Methylation amplification DNA chip (MAD)		1µg [114]	High
Promoter-associated methylation amplification DNA chip (PMAD)		0.5µg [115]	High
Comprehensive high-throughput arrays for relative methylation (CHARM)		3.5µg [116]	High
Microarray-based methylation assessment of single samples (MMASS)		2µg (v1); 1.2µg (v2) [117]	High
Differential methylation hybridization (DMH)		200ng to 100ng [118]	High
Methylation single-nucleotide polymorphism (MSNP)	Restriction enzyme digestion	600ng [110]	High
Methylated DNA immuno-precipitation (MeDIP)	Affinity-based	1µg to 100ng [110]	High

Methylated CpG island recovery assay (MIRA)	Affinity-based	3µg to 2µg [119]	High
Enzyme-linked immunosorbent assay (ELISA)		2µg to100ng	High
Bisulfite methylation profiling (BiMP)	Bisulfite Conversion	2µg to100ng [120]	High
Specific Loci Methylation techniques			
Bisulfite Sequencing	Bisulfite Conversion	~1µg [121]	Medium
Bisulfite Padlock Probes		~200ng [122]	Medium
Bisulfite Pyrosequencing		~1µg [123]	Medium
Methylation-specific PCR (MSP)		100ng to 50pg [124]	Medium
Sensitive melting analysis after real-time methylation-specific PCR (SMART-MSP)		1ng [125]	Medium
Methylation-specific fluorescent amplicon generation (MS-FLAG)		1µg [126]	Medium
Combined bisulfite restriction analysis (COBRA)	Bisulfite Conversion / Restriction enzyme digestion	2µg [127]	Medium
Methylated DNA immunoprecipitation PCR (MeDIP-PCR)	Affinity-based	1µg to 100ng [110]	Medium
Combination of methylated DNA precipitation and methylation-sensitive restriction enzymes (COMPARE-MS)	Affinity-based / Restriction enzyme digestion	100ng to 4ng [128]	Medium
Illumina Infinium Methylation Assay	Affinity-based	250ng	High

Many of the listed techniques and technologies have the possibility to undergo optimization for the low quantities of possible available DNA, but if taking into consideration the current standard protocols, the techniques most close to being able to adapt to such low input DNA quantities that might be available in the cases of missing body are: LUMA; MeDIP; MIRA; WGSGS; DHPLC; LC-MS/MS; PMAD; DMH; MSNP; ELISA

based assays; BiMP; Bisulfite Padlock Probes; Bisulfite Pyrosequencing; MSP; SMART-MSP; MeDIP-PCR; COMPARE-MS; Illumina Infinium Methylation Assays. In that group, those that can handle slightly degraded DNA in their current level of optimization are: DHPLC; LC-MS/MS; Bisulfite Padlock Probes; Bisulfite Pyrosequencing; MSP; SMART-MSP; MeDIP-PCR; COMPARE-MS. While LUMA and Infinium Methylation assays suffer slightly from DNA degradation, they are respectfully powerful techniques that should not be entirely disregarded. The Infinium assays cover incredibly large amounts of the human genome in a quick and succinct protocol, albeit being quite expensive per sample. LUMA is one of the few global methylation techniques capable of utilizing low amounts of input DNA that don't require expensive hardware to perform, in the form of a mass spectrometer, being a good candidate in lieu of the strong LC-MS/MS.

As the methodology is still in its early years of infancy, reaching a concrete decision on the best technique to use is difficult. The size of the biomarker battery, lack of optimization for low input DNA, price range changes, possible reaction kit developments, volume of samples per case and other variables yet unknown may cause discarded techniques to become viable. With current knowledge, the frontrunners for viability are LUMA, LC-MS/MS, MSP and SMART-MSP, Bisulfite Padlock Probes, MeDIP-PCR and Infinium Assays, in no particular order. Alongside techniques that use PCR, HRM analysis is also a front to be studied that can easily compliment them and provide further information.

6. The State of the Art

6.1. Lessons from the RNA based PMI assessment efforts

While designing a protocol to determine PMI through genetic information, one must take into consideration the molecule to work with. RNA and DNA are vastly different molecules with different properties that bring with them different problems that need to be solved before an all-variable encompassing standard can be created for forensic use.

The work of determining if the usage of DNA degradation in relation to PMIs can be traced to the decade of 1980s [129], with unfavorable results. In the following years, this trend did not alleviate, with research being nearly dropped in the early 2000s [130]. The research into the topic focused in flow cytometric evaluations, tried in several tissue samples, with one of the most recent being dental pulp [131]. Flow cytometry is a fluorescence-based technique that forces single cells through the analysis node one at a time, where a laser excites the fluorescent reporters the sample was stained with. Using reporters that can bind solely to the desired biomarker, flow cytometry can detect quantity in undamaged cells, or most commonly, used to analyze quantitively and qualitatively cells, if used cell specific reporters.

Even notwithstanding the poor to no linearity of the results obtained through such protocols [130], the decisive fact that made such usage impossible was that after 72h mass DNA denaturation settled in on most tissues, making any scalability of the protocol very unlikely. So, after no progress having been achieved in this front, efforts shifted to another form of genetic material.

Focus on RNA degradation as a tool for PMI estimation in a forensic setting was piloted by Martin Bauer *et al.* in 2003 [10]. In this study, fatty acid synthase-messenger RNA (FASN-mRNA), an mRNA tied to a housekeeping gene, therefore present across time and tissue, with a size of roughly 10 kb was compared between fifty freshly drawn blood samples, fifty blood samples drawn and 36 brain samples from cadavers with exact time of death known. As for the theory, the study opted to access the degradation from the 5’ end of the mRNA, using four distinct primers that annealed to it at areas roughly 1.5 kb apart, each requiring a different amount of the molecule to be intact in order to form fragments **(Fig. 25)**.

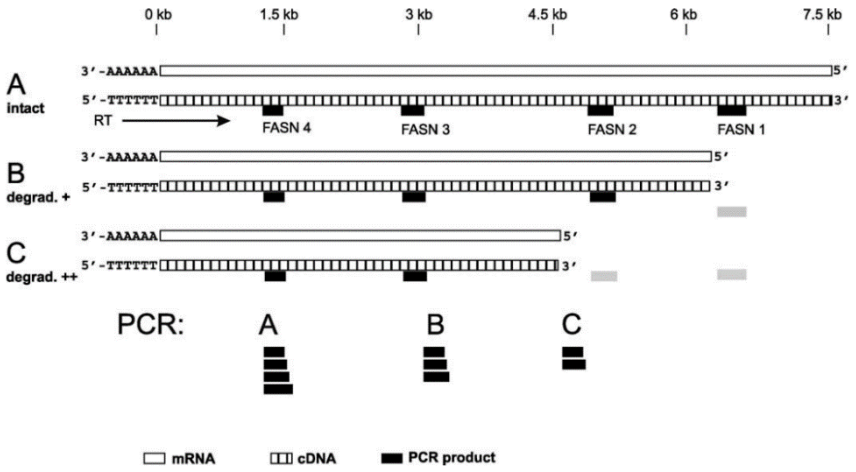

Fig. 25- Primer setup used to determine mRNA degradation. The further stages of overall molecular degradation, the less the number of fragments created by FASN 1, FASN 2 and FASN 3. The bigger the difference between the amount of FASN 4 fragments, that stood near the protected poly-A chain, away from the degraded 5' end, and each of the other 3 primers, the more advance the mRNA degradation was, in a quantifiable way. Adapted from *M. Bauer, I. Gramlich, S. Polzin, and D. Patzelt, "Quantification of mRNA degradation as possible indicator of postmortem interval - A pilot study" Leg. Med., vol. 5, no. 4, pp. 220–227, 2003.*

As for the workflow, it consisted a concise four step process of extraction, purification, amplification and analysis. The focal part of this study's laboratorial work was the qPCR. After the FASN-mRNA being converted into cDNA, each sample and each primer were amplified in separate tubes in the same qPCR, which negated possible discrepancies between results **(Fig. 26).** Such results also substantiated through ultra-thin polyacrylamide-gel electrophoresis.

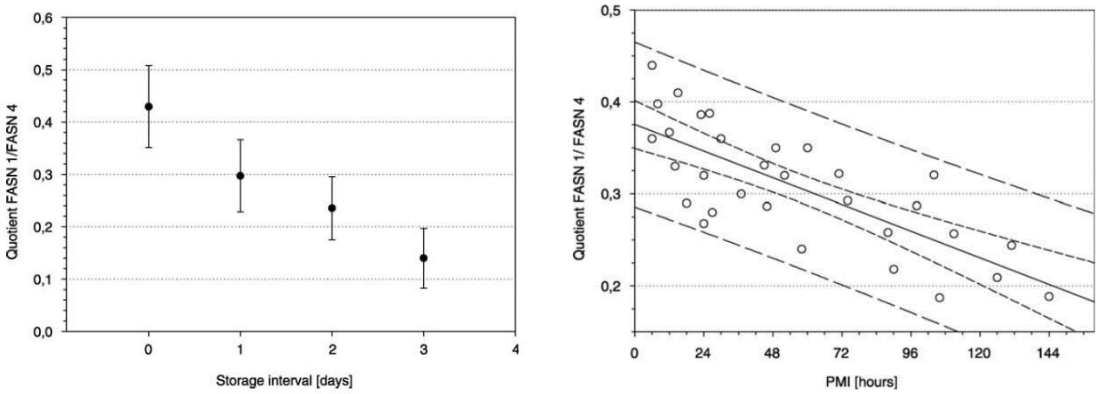

Fig. 26- The mean values for the quotient between FASN1 and FASN 4 from 10 blood samples taken from living individuals, ranging from immediate processing to three days of 4ºC storage (Left). The scatterplot plotting each postmortem's brain sample's ratio between FASN1 and FASN4 against time passed since death (Right). Adapted from *M. Bauer, I. Gramlich, S. Polzin, and D. Patzelt, "Quantification of mRNA degradation as possible indicator of postmortem interval - A pilot study" Leg. Med., vol. 5, no. 4, pp. 220–227, 2003.*

These results obtained from *postmortem* blood samples were but a fraction of the entire sampling, as 25 of the samples, half of the total drawn, were unable to produce any results due to the advanced stage of hemolysis. Although the successful samples produced a steady decrease of FASN1/FASN4 quotient, as seen in the left part of **Fig. 26**, with a reported strong correlation ($r = 0.808$; $P < 0.01$), the fact that half the samples did not produce usable results puts a strong damper on the usage of blood mRNA for forensic procedures. Also, when the quotients for more inner FASN primers is analyzed, the downward pattern diminishes, vanishing on the FASN3/FASN4 quotient **(Fig. 27),** thus limiting any possible PMI calculations to the FASN1/FASN4 quotient, or the quotient between the most distal and proximal areas relative to the protected poly-A chain in other mRNA molecules. Brain samples on the other hand provided a much more robust success rate in obtaining results but shared the same issue as blood sample results of losing the downward pattern the closer the FASN primers were. Another issue with the protocol is the fact that live blood samples had their mRNA degradation slowed by storing them at 4ºC instead of room temperature, meaning the 4 to 5 day period tested and window of scalability will become much shorter once the mRNA is allowed to degrade at its usual rate. This also points towards the temperature of the body being a heavy variable in mRNA degradation PMI estimation, and the variations it suffers due to the various *postmortem* stages a body goes through, as well as daily temperature cycles, may turn creating a standard very difficult.

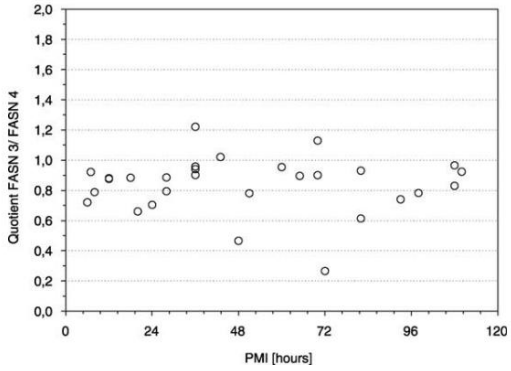

Fig. 27- The scatterplot denoting the FASN3/FASN4 quotient off all *postmortem* blood samples. Adapted from *M. Bauer, I. Gramlich, S. Polzin, and D. Patzelt, "Quantification of mRNA degradation as possible indicator of postmortem interval - A pilot study" Leg. Med., vol. 5, no. 4, pp. 220–227, 2003.*

The temperature induced error margins have been since worked on and in 2016, Ye-Hui *et al.* tackled this issue [16]. The study followed a similar laboratorial procedure as Bauer's one, extraction, purification, amplification and analysis but with a much heavier emphasis on the data analysis. It started with 216 rats, divided into three major groups *postmortem,* kept at 10ºC, 20ºC and 30ºC, with each major group being divided further by PMIs (0h, 1h, 3h, 6h, 12h, 24h, 36h, 48h, 72h, 96h, 120h, 144h), in which tissue samples from the liver and muscle were retrieved and used to create a mathematical model that could be used to determine PMI. Fifteen other rats were separated into five major groups *postmortem,* ranging from 10ºC to 30ºC in intervals of 5ºC, with liver and muscle samples taken at 10h, 60h and 110h *postmortem.* This set of rats were used as checks against the mathematical model to substantiate it. Twelve *postmortem* lung and muscle tissue samples were also obtained from humans at the time of autopsy, with clear time of death, PMI including time passed while under freezing conditions at the morgue, and temperature assumed to be the average of the crime scene before the transfer of the body.

Four different classes of RNA molecules were chosen to be potential targets for the mathematical model's variable or for reference, these being: 18S rRNA and 5S representing ribosomal RNAs, RPS29, ACTB and GAPDH standing in for messenger RNAs, let-7a, miR-195, miR-200c, miR-1 and miR-206 on behalf of micro RNAs, and the small nuclear RNA U6 .

After successful RNA extraction, isolation, and amplification by qPCR on all samples, all data was processed with the aid of geNorm, an algorithm that can be used to determine the most stable reference gene or genes from the qPCR results of the entire set of studied RNAs. The algorithm calculates the M-value, the relative stability of all appointed reference samples in the processed set in accordance with Vandesompele et al., Genome Biology, 2002, 'Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes' [132]. Analysis has shown that the housekeeping ACTB and GAPDH were the most unstable RNAs and most prone to degradation (**Fig. 28**), while 5S and RPS29 could be used as internal reference biomarkers for lung tissues and control markers for muscle tissues.

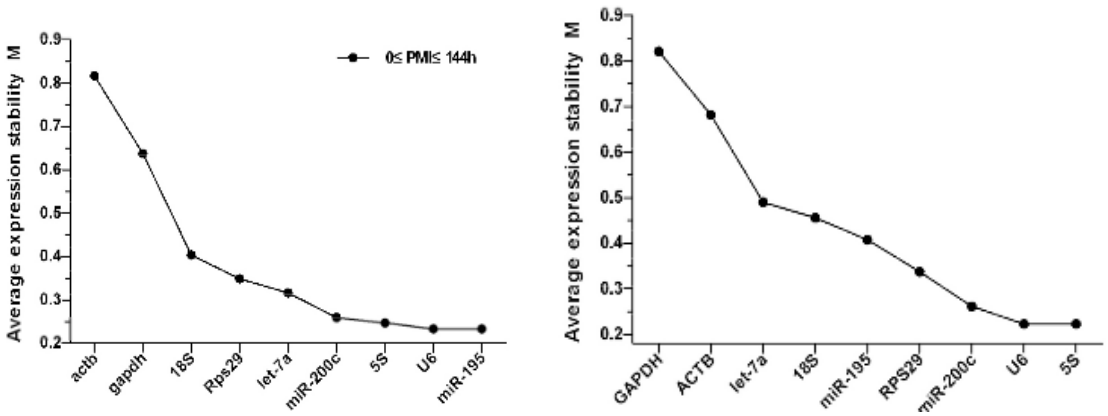

Fig. 28- Mean M-values in lung tissue samples from rats kept at 20ºC, from all PMIs tested (Left). Mean M-values in human *postmortem* lung tissue samples used for model validation (Right). ACTB and GAPDH possess the highest values in both sets of data, indicating that from all the tested biomarkers they're the ones most prone to bigger differentiation of degradation along PMIs. miR-195, miR-200c, 5S, U6 and RPS29 showed low M-values, being selected as reference biomarkers for the creation of the mathematical model. Adapted from *Y. H. Lv et al., "RNA degradation as described by a mathematical model for postmortem interval determination" J. Forensic Leg. Med., 2016.*

The mathematical model itself was constructed using the ΔCt values obtained in the qPCR. The Ct value is the number of PCR cycles a sample's fluorescence required to surpass the threshold set to ignore background fluorescence. This value is inversely proportional to the amount of initial target genetic sample, meaning the bigger the number of molecules in the initial sample, the lower the Ct value will be. In this study, the target genetic sample that can provide a successful duplication in a cycle is the non-degraded RNA molecules in the sample. Following the same logic, the lesser the overall RNA degradation is, the more molecules can provide a successful duplication and the lower the Ct is. The ΔCt values are the difference between the Ct values of the chosen target biomarkers, ACTB and GAPDH,

and the geometrical Ct values of the reference biomarkers, miR-195, miR-200c, 5S, U6 and RPS29 (**Fig. 29**). At 20ºC, the produced curves fit a nearly straight line with a positive slope, indicating a linear divergence between the Ct values of ACTB and the reference biomarkers, as expected. ACTB is more unstable, therefore will degrade much quicker and will generate a bigger Ct value much quicker than the reference biomarkers. At 30ºC the linearity isn't as strong, and the positive slope is lost in the quadratic and cubic curve fits at around 108 hours. This loss of a linear relation may indicate the break point where the model loses its effectiveness.

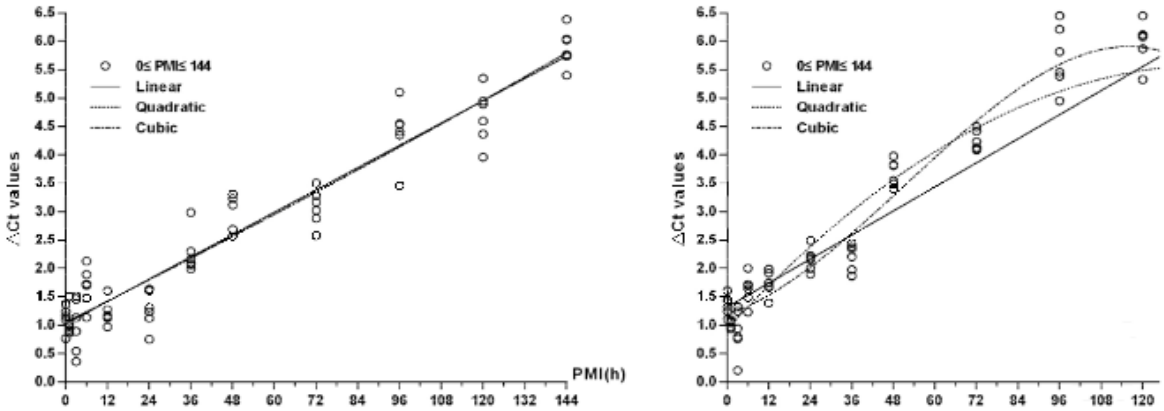

Fig. 29- Curve fit of *postmortem* rat lung tissue samples kept at 20ºC between ACTB ΔCt and PMI (Left). Curve fit of *postmortem* rat lung tissue samples kept at 30ºC between ACTB ΔCt and PMI (Right). Adapted from *Y. H. Lv et al., "RNA degradation as described by a mathematical model for postmortem interval determination" J. Forensic Leg. Med., 2016.*

No.	T(°C)	Real PMI (h)	Estimated PMI (h)				Mean	SD	Error rate (%)
			Lung		Muscle				
			actb	*gapdh*	*actb*	*gapdh*			
1	10	10	13.9	11.2	9.2	11.7	10.9	2.4	8.5
2		60	56.45	60.85	65.7	57.4	61.2	4.9	2.0
3		110	123.1	115.8	120.3	114.6	116.6	6.5	6.0
4	15	10	12.45	12.05	13.2	9.4	11.7	1.6	16.6
5		60	58.6	64.1	50.5	52.1	57.7	8.6	3.9
6		110	106.8	114.95	102.1	118.4	112.6	9.8	2.4
7	20	10	10	9.55	12.3	13.0	11.1	1.9	10.9
8		60	67.1	61.8	66.4	70.1	65.0	5.9	8.4
9		110	105.65	107.45	125.5	130.4	118.2	11.7	7.4
10	25	10	12.65	11.65	9.7	12.1	11.3	1.3	12.8
11		60	55.45	57.05	53.2	66.8	58.5	6.0	2.5
12		110	99.15	98.9	109.2	94.9	100.5	6.1	8.7
13	30	10	13.75	11.7	12.4	8.9	11.2	2.3	11.8
14		60	70.4	62.35	55.5	50.6	57.7	8.7	3.8
15		110	90.65	106.7	92.1	110.4	104.0	15.4	5.5
Total Error rate (%)			14.6	8.5	12.8	12.7		7.4	
			7.7		9.4				

Once the model was created, the samples stored at the 10ºC to 30ºC in 5ºC intervals were compared to it, with the results being displayed in **(Fig. 30)**. The results show that early on, at 10h there is the highest error rates at a mean of 12.12%, followed by lower error rates at 110h at a mean of 6% and having the lowest error rates at the 60h mark at a mean of 4.12%, with lung samples being more accurate than muscle samples by a difference of 1.7%. The previously noted

Fig. 30- Cross reference between the mathematical model and rat sample used for validation's results. Adapted from *Y. H. Lv et al., "RNA degradation as described by a mathematical model for postmortem interval determination" J. Forensic Leg. Med., 2016.*

loss of positive relation in the model explains the decreased accuracy at 110, while early on there might not be enough difference between the target and refrence biomarkers for the model to be accurate, meaning that an internal refrence normalized mathematical model might be restricted to a certain time frame for maximum accuracy.

The insight that the variation of the RNA models is great between temperatures but can be mitigated using stable biomarkers as internal references Ye-Hui's work provided is very key insight that will inevitably be required if a standard for forensic usage is to be created using DNA methylation. As many other variables might affect DNA methylation alongside temperature, the usage of a gene with a long term stability in its methylation that has suffered all the same variations as the more unstable target gene's epigenetic in order anchor the data may be the key to create a robust, all-encompassing methylation standard for forensic usage, even though if maximum accuracy may only be achieved in a certain timeframe depending on the combination of reference and target biomarkers.

6.2. Work in the Epigenetics Front

Work on the epigenetic side is a relatively more recent endeavor, at least on the forensics front. *Postmortem* DNA methylation status has been a hot topic of research in neurobiology and medical fields since the late 2000s. Deregulation of DNA methylation has been linked to major psychosis [133]–[135], hinted at a possible link in major depressive disorder [136] and sporadic amyotrophic lateral sclerosis [137]. While a many of these papers provide information on *postmortem* DNA methylation status, even if focusing on a few genes, none published yet focuses on measuring PMI directly using it.

Nonetheless, the basis for this field of work are being currently built by such studies. Much of the information given is crucial for the precision of the standard. Marie-Abele Bind *et al.* [138] studied DNA methylation on *in-vivo* blood samples from 777 elderly participants of the normative aging study, focusing on long interspersed nuclear element (LINE-1), *Alu* and tissue factor 3 (F3), intercellular adhesion molecule 1 (ICAM-1), toll-like receptor 2 (TLR-2), carnitine O-acetyltransferase (CRAT), 8-oxoguanosine DNA glycosylase-1 (OGG1), interferon gamma (IFN-γ), glucocorticoid receptor (GCR) and inducible nitric oxide synthase (iNOS) genes, measuring methylation at 1-5 CpG base positions on each of the promoter regions. Such genes were chosen due to their relation to pathways in turn related to adverse effects of high temperature and humidity in elderly populations, including such as coagulation and inflammation. Through bisulfite pyrosequencing, the methylation status of ICAM-1 over three weeks for each participant was plotted against temperature and humidity (**Fig. 31**). This graph shows that with the decrease of temperature, ICAM-1 suffers hypomethylation in the period of 3 weeks, and vice versa for the increase of temperature, and suffers hypomethylation under high humidity, with greater variation the greater the temperature is. The other studied genes also presented changes in methylation with variations in temperature and humidity. A decrease of temperature caused a CRAT hypermethylation, an iNOS and GCR hypomethylation after air pollution normalization and temperature increases caused a TLR-2 hypomethylation. Alongside that, higher relative

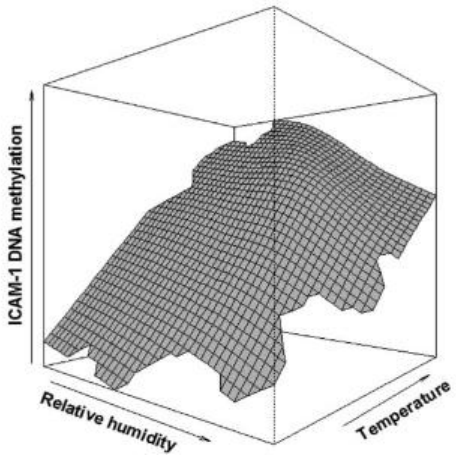

Fig. 31- The 3D representation of the 3-week methylation status of the ICAM-1 gene against humidity and temperature using a 2-covariate penalized thin plate spline. Adapted from *"M. A. Bind et al. "*Effects of temperature and relative humidity on DNA methylation" Epidemiology, *2014"*.

humidity produced LINE-1 hypomethylation and *Alu* hypermethylation, with effects being compounded by temperature changes as well. This study's results, albeit being taken from in-vivo samples, hint at the fact that epigenetic information on some genes may be more fluid than expected. Small but statistically sound variations occurred in a 3-week window period with changes as small as 5ºC per week, reinforcing the notion that Ye-Hui *et al.*'s [16] efforts to normalize temperature variations effects are necessary to be carried over and applied to epigenetic information.

As epigenetics regulate gene expression, different tissues will also possess different DNA methylation values at certain genes and may degrade at different rates. Barrachina *et al.* [139] studied DNA methylation degradation in *postmortem* brain tissues with different PMIs to a maximum of 20 hours and a maximum of 48 hours in storage, alongside several neurologic pathologies, ranging from several stages of Alzheimer's disease, Parkinson's disease and Dementia with Lewy's Bodies, in order to determine if *postmortem* delays in sample analyzing affected the validity of DNA methylation relations to such diseases. The study proceeded with 124 samples, from which 26 controls, 27 Alzheimer disease patients from stage III to stage VI, 10 argyrophilic grain disease patients, 6 patients with frontotemporal lobar degeneration linked to tau mutations, 4 patients with frontotemporal lobar degeneration with ubiquitin-immunoreactive inclusions, 3 patients with frontotemporal lobar degeneration with motor neuron disease, 3 Pick disease patients, 8 Parkinson disease patients and 20 patients with dementia with Lewy bodies. The methylation status data gathering was confined to the promoter regions of the receptor for advanced glycation end products (RAGE), adenosine A2A receptor (ADORA2A) and microtubule-associated protein tau (MAPT) genes, genes linked to the pathologies included in the sample population. The epigenetic information was studied through bisulfite conversion followed

by PCR amplification and time-of-flight matrix-assisted laser desorption/ionization (MALDI-TOF), a three-step mass spectrometry analysis protocol.

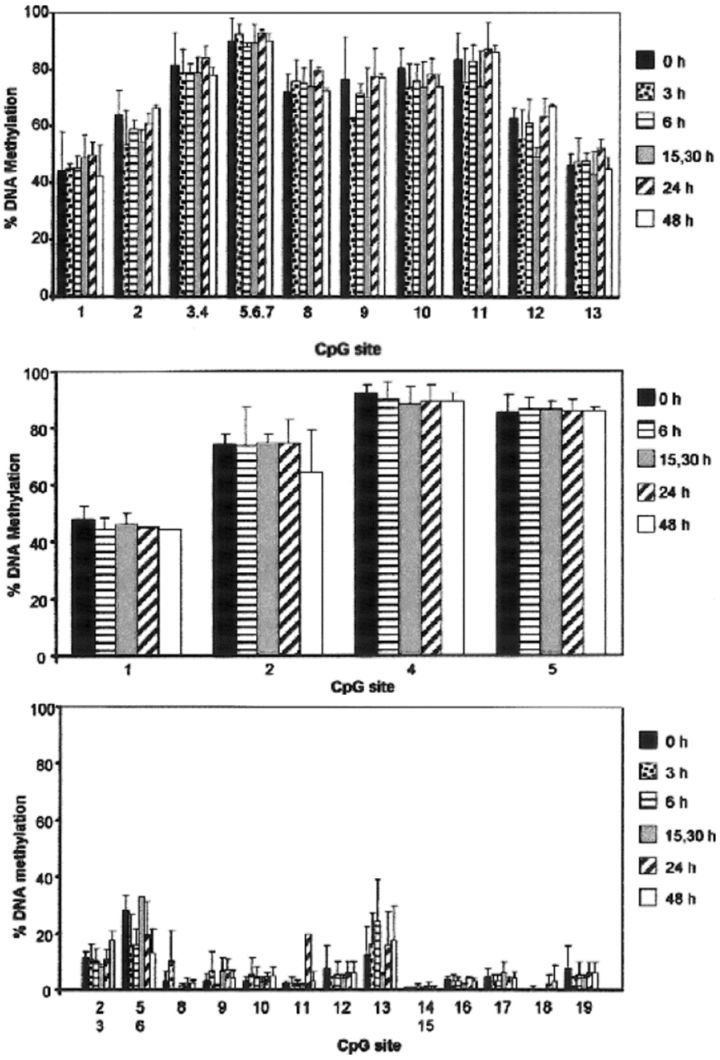

Fig. 32- Graph plotting the percentage of methylation in different CpG sites of the studied gene's promoter regions. All three cases, RAGE (top), ADORA2A (middle), and MAPT (bottom) show little to no variation between different PMIs. Adapted from *"M. Barrachina and I. Ferrer, "DNA methylation of Alzheimer disease and tauopathy-related genes in postmortem brain" J. Neuropathol. Exp. Neurol., vol. 68, no. 8, pp. 880–891, 2009."*.

The results from the tree studied locations showed stability across most CpG sites across all PMI (**Fig. 32**), either suggesting that these three genes may be stable candidate biomarkers or that the brain tissue is a stable enough candidate for sample collection. The first conclusion would have the issue that these three biomarkers are linked to pathologies, but the second could prove very useful. Monoranu *et al.* [140] continued the work on this front with brain samples from patients that died of causes unrelated to neurological pathologies and found out that methyltransferase activity, not DNA methylation itself, remains largely preserved along the PMIs before storage.

Ernst *et al.* [141] studied the effects that pH might have on *postmortem* brain tissue, selecting the rRNA gene and the neurotrophic receptor tyrosine kinase 2 (NTRK2) gene. Twenty patients' *postmortem* brain samples were used, ten from the hippocampus for rRNA testing and twenty from the frontal cortex for NTRK2 testing. The methylation status of these genes was studied through bisulfite conversion followed by PCR, and subsequent

integration in a pDrive cloning vector, cloned through competent *Escherichia coli* cells and sequenced. The results came as no relation between pH and rRNA methylation and negligent relation between pH and NTRK2, results that should be taken as merely preliminary due to the low number of samples and tested biomarkers (**Fig. 33**).

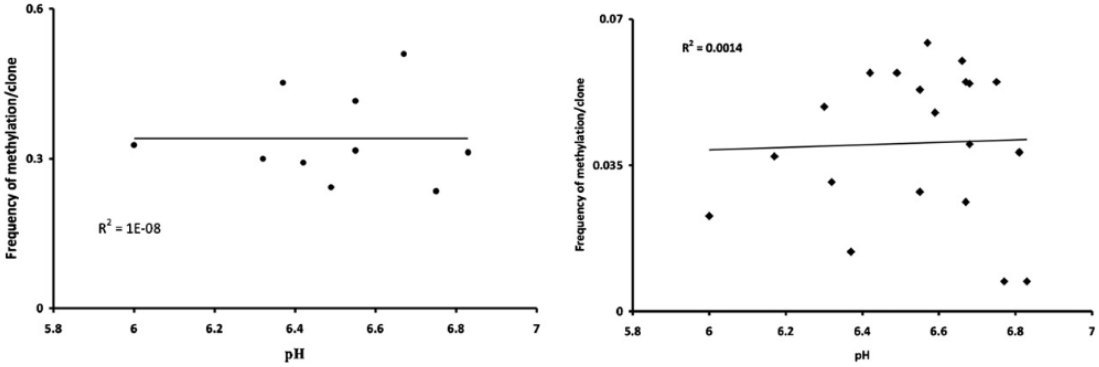

Fig. 33- Graphs plotting the frequency of methylation per clone against pH of rRNA (left) and NTRK2 (right). Adapted from *"C. Ernst, P. O. McGowan, V. Deleva, M. J. Meaney, M. Szyf, and G. Turecki, "The effects of pH on DNA methylation state: In vitro and post-mortem brain studies" J. Neurosci. Methods, vol. 174, no. 1, pp. 123–125, 2008."*

DNA methylation stability is another variable to take into consideration. The DNA molecule itself is quite resilient in some tissues, even remaining usable once only skeletal remains are left [106] and the methylation itself is stable for roughly 48h as stated previously. Large term stability in a decomposing body has not yet been studied, which is the crucial question that will solidify the viability of a DNA methylation PMI accessing methodology, but long-term stability in storage was. Li *et al.* [142] took several samples of blood taken up to 20 years prior to the study, stored in 4ºC, -20ºC and -80ºC. The samples were grouped as such: Group 1 contained six 20-year-old DNA samples stored at 4ºC; Group 2 contained the same samples, but diluted 13 years later to 10ng/µl, kept at 4ºC; Group 3 contained six whole blood samples from the same donators of Group 1, kept with anticoagulants and stored at -80ºC until DNA extraction 20 years after sample collection; Group 4 contained five samples collected from the same donators as the Group 1, but collected 19 years more recently proceeding a recall, with DNA stored for less than 3 months at -20ºC; Group 5 contained five samples equal to Group 4, but stored at 4ºC. Each sample had its methylation profiled through bisulfite conversion preceding reduced representation bisulfite sequencing (RRBS), with a total of 33 DNA libraries, five of them replications of Group 1, Group 2 and Group 3 samples for variance control. The resulting sequence data was matched and compared to the hg19 human genome reference and the methylation rate was calculated as the percentage of unmethylated cytosines in each CpG location, with each site being covered by at least 10 reads. The entirety of CpG methylation data was examined for five groups of interest: CpG Islands (CGI), CGI Shores, transcription start sites (TSS), promoters and gene bodies. All comparisons between each group of interest of each sample and each group was done using ComBat-adjusted hierarchical clustering analysis (HCA), adjustment required due to present

sequencing batch bias shown by initial methylation data, where samples from the same group would cluster together. When data from the subject from groups 1 to 3 were analyzed together, same individuals in different storage conditions, it was found that methylation levels were highly correlated (CGI:0.995; CGI Shore: 0.985; Gene Body: 0.973; Promoter: 0.982; TSS: 0.988) between the same individual in different groups, across all groups of interest, and also that different subjects in the same group had high correlation between all groups of interest (**Fig. 34**). These results suggest that even with different storage conditions, with variable temperatures and different extraction times, methylation profiles of each subject were kept similar after 20 years.

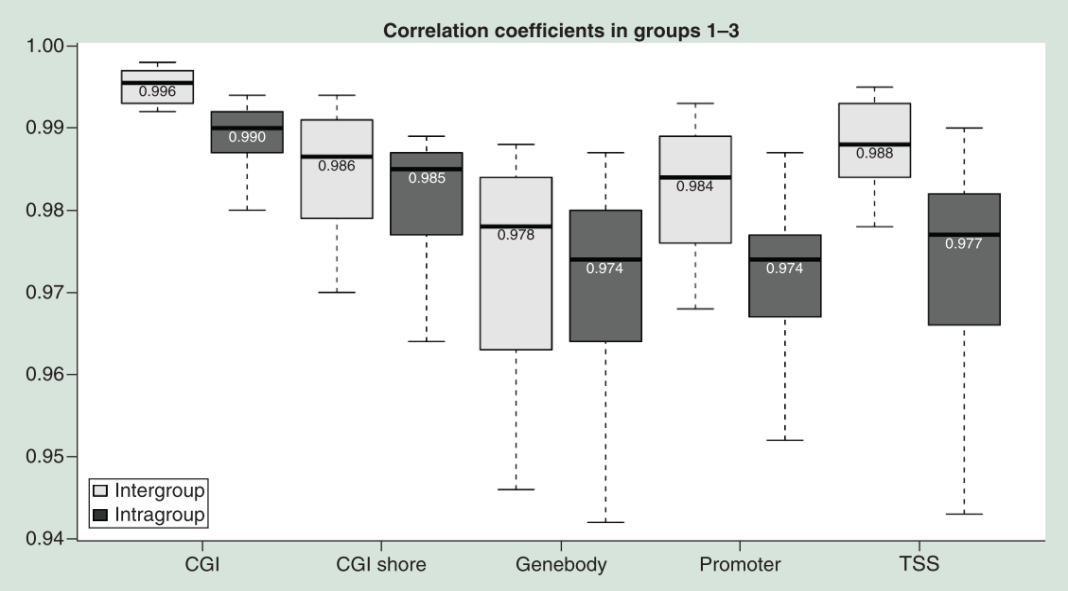

Fig. 34- Resulting boxplots using ComBat-adjusted methylation profiles of the DNA samples, presenting correlation coefficients for same subject in different groups (Intergroup) and different subjects in the same group (Intragroup), across all 5 groups of interest. Adapted from *"Y. Li, X. Pan, M. L. Roberts, P. Liu, and A. Theodore, "*Stability of global methylation profiles of whole blood and extracted DNA under different storage durations and conditions" *vol. 10, pp. 797–811, 2018.".*

When groups one through three where compared to groups four and five, samples with different storage times, it was also found that methylation levels were highly correlated between the same individual in different groups and between different individuals in the same groups (**Fig. 35**). These results suggest that methylation profiles don't change statistically significantly over periods of up to 20 years in storage when compared against

newer samples from the same individual, meaning that these samples are still valid samples that provide accurate data.

	CGI	CGI shore	Gene body	Promoter	TSS
G1 vs G2	0.91	0.98	0.21	0.85	0.87
G1 vs G3	0.91	0.98	0.49	0.87	0.87
G1 vs G4	0.91	0.98	0.72	0.85	0.87
G1 vs G5	0.91	0.98	0.21	0.85	0.87
G2 vs G3	0.91	0.98	0.72	0.85	0.87
G2 vs G4	0.91	0.98	0.72	0.96	0.87
G2 vs G5	0.91	0.98	0.96	0.85	0.87
G3 vs G4	0.91	0.98	0.98	0.85	0.87
G3 vs G5	0.91	0.98	0.72	0.85	0.87
G4 vs G5	0.91	0.98	0.75	0.87	0.87

Fig. 35- Benjamin–Hochberg-adjusted false-discovery rate of pair-wise Wilcoxon rank-sum test of correlation coefficients between methylation profiles of DNA samples amongst all groups. Adapted from "Y. Li, X. Pan, M. L. Roberts, P. Liu, and A. Theodore, "Stability of global methylation profiles of whole blood and extracted DNA under different storage durations and conditions" vol. 10, pp. 797–811, 2018.".

To study if the presence of chemicals for large periods of time affected the methylation profiles, groups three and four, samples stored in Acid-citrate-dextrose vacutainers, were tested against each other. Methylation profiles across different subjects in the same group of both groups showed high correlation again (CGI: 0.990 G3/ 0.990 G4; CGI Shore: 0.984 G3/ 0.981 G4; Gene Body: 0.977 G3/ 0.966 G4; Promoter: 0.975 G3/ 0.971 G4; TSS: 0.978 G3/ 0.974 G4). Such similarities between the methylation profiles suggest that the presence of anticoagulants does not interfere with DNA methylation, or that if it does, the changes in pattern are changed in a reliable way. Finally, it was tested if different anticoagulants had a different effect on methylation. Groups 4 and 5 had their samples stored with Acid-citrate-dextrose (G4) and EDTA (G5), for a period of three months. Here the data shows a slightly smaller correlation (G4- CGI: 0.988; CGI Shore: 0.974; Gene Body: 0.955; Promoter: 0.960; TSS: 0.970 / G5- CGI: 0.988; CGI Shore: 0.983; Gene Body: 0.975; Promoter: 0.967; TSS: 0.973). This implies that different anticoagulants may affect the DNA methylation profiles slightly in different ways, inferring that anticoagulants do interfere with the methylation profiles in reliable ways.

Li *et al.*'s findings have preliminarily found that methylation profiles are kept stable under controlled storage conditions for very long periods of time and diverse amounts of thawing and refreezing processes, yet the crux for the viability of an epigenetic PMI assessor is the stability under uncontrolled conditions. Bulla *et al.* [143] tackled most of the same points as Li *et al.,* with the addition on room temperature. Eight subjects donated 41 blood samples, kept under different storage conditions. Each subject had their 41 samples divided into three groups: Group 1 consisted in EDTA whole blood samples without any additives;

Group 2 consisted in EDTA whole blood samples with one quarter of DNAgardBlood solution per volume, added before storage; Group 3 consisted in EDTA whole blood samples with one quarter of DNAgardBlood solution per volume, added after storage, right after thawing, before samples were processed. For each group, various samples were stored at different temperatures and processed at different times (**Table 4**).

Table 4- Division of samples by group and time in storage. RT: Room Temperature. Data adapted from *"A. Bulla, B. De Witt, W. Ammerlaan, F. Betsou, and P. Lescuyer,* "Blood DNA Yield but Not Integrity or Methylation Is Impacted after Long-Term Storage" *Biopreserv. Biobank., vol. 14, no. 1, pp. 29–38, 2016."*

	Group 1 RT	Group 1 (4ºC)	Group 1 (-20ºC)	Group 1 (-80ºC)	Group 2 RT	Group 2 (-20ºC)	Group 2 (-80ºC)	Group 3 (-20ºC)	Group 3 (-80ºC)
0h	3								
24h	1	1							
48h	1	1			1				
7 days	1	1	1	1	1	1	1	1	1
14 days	1	1	1	1	1	1	1	1	1
6 months		1	1	1	1	1	1	1	1
12 months			1	1	1	1	1	1	1

From each sample, DNA quality, DNA yield and DNA methylation were assessed. DNA yield and quality were obtained through spectrofluorometry (Quant-iT PicoGreen dsDNA Assay Kit) and spectrophotometry (NanoDrop 3000). The three room-temperature, 0h Group 1 samples were used as control samples. DNA methylation was assessed through post-restriction enzyme MSP (Epitect Methyl II PCR Arrays Kit). The compared groups were Group 1 at -20ºC, Group 1 at 80ºC and Group 2 at room temperature. All statistical data was obtained through Kruskal–Wallis test accompanied by Dunn's multiple comparison subsequently. In the spectrophotometric assessment of DNA yield, the 3 Group 1 samples tested at 0h were used as reference, ranging from 10.9mg/mL to 24.6mg/mL. The data shows that: Group 1 samples had a quick DNA degradation at higher temperatures, rate decreasing alongside temperature, losing 70% of its total yield at room temperature by 2 weeks up to keeping 70% of its total yield at -80ºC at 12 months; Group 2 samples suffered a much lower DNA degradation rate across all times and temperatures, from keeping 60-80% of total yield at room temperature across time to stable 80% of total yield at -80ºC in up to 1 year; Group

3 samples had the best yields with stable 90% of total yield at both -20ºC and -80ºC in up to 1 year (**Fig. 36**).

			24 h	*48 h*	*7 days*	*14 days*	*6 months*	*12 months*
EDTA	RT	Mean	**82**	**77**	**49**	**33**		
		CV (%)	16	20	30	19		
EDTA	4°C	Mean	**87**	**81**	**76**	**66**	**63**	
		CV (%)	19	20	23	25	22	
EDTA	–20°C	Mean			**72**	**54**	**59**	**53**
		CV (%)			19	26	34	42
EDTA	–80°C	Mean			**68**	**51**	**81**	**73**
		CV (%)			20	23	16	29
DgB before	RT	Mean		**78**	**78**	**63**	**76**	**82**
		CV (%)		28	25	22	28	27
DgB before	–20°C	Mean			**79**	**63**	**49**	**36**
		CV (%)			29	18	33	47
DgB before	–80°C	Mean			**87**	**86**	**90**	**86**
		CV (%)			28	26	14	33
DgB after	–20°C	Mean			**91**	**78**	**93**	**92**
		CV (%)			20	31	26	26
DgB after	–80°C	Mean			**90**	**90**	**92**	**96**
		CV (%)			24	18	26	25

Fig. 36- DNA yield data obtained through spectrophotometry, relative to Group 1 0h control samples. CV: Coefficient of Variation; EDTA: Group 1; DgB before: Group 2; DgB after: Group 3. Adapted from *"A. Bulla, B. De Witt, W. Ammerlaan, F. Betsou, and P. Lescuyer,* "Blood DNA Yield but Not Integrity or Methylation Is Impacted after Long-Term Storage" *Biopreserv. Biobank., vol. 14, no. 1, pp. 29–38, 2016."*

In the spectrofluorometric assessment of DNA yield, the 3 Group 1 samples tested at 0h were used again as reference, with the yields ranging from 7.5mg/mL to 16.0mg/mL. The data follows a similar trend to the spectrophotometric one, with quick Group 1 loss of total yield rapidly in the higher storage temperatures and slowing down as the temperature fell.

			24 h	*48 h*	*7 days*	*14 days*	*6 months*	*12 months*
EDTA	RT	Mean	**84**	**81**	**49**	**40**		
		CV (%)	21	16	26	24		
EDTA	4°C	Mean	**89**	**88**	**76**	**68**	**86**	
		CV (%)	17	18	14	19	29	
EDTA	–20°C	Mean			**87**	**66**	**76**	**65**
		CV (%)			14	20	29	34
EDTA	–80°C	Mean			**77**	**63**	**90**	**83**
		CV (%)			14	20	35	27
DgB before	RT	Mean		**92**	**101**	**84**	**81**	**96**
		CV (%)		20	25	15	29	23
DgB before	–20°C	Mean			**94**	**79**	**56**	**38**
		CV (%)			23	18	36	99
DgB before	–80°C	Mean			**102**	**96**	**104**	**105**
		CV (%)			22	23	14	34
DgB after	–20°C	Mean			**98**	**92**	**95**	**85**
		CV (%)			19	26	25	33
DgB after	–80°C	Mean			**98**	**105**	**101**	**106**
		CV (%)			19	19	18	26

Fig. 37- DNA yield data obtained through spectrofluorometry, relative to Group 1 0h control samples. CV: Coefficient of Variation; EDTA: Group 1; DgB before: Group 2; DgB after: Group 3. Adapted from *"A. Bulla, B. De Witt, W. Ammerlaan, F. Betsou, and P. Lescuyer,* "Blood DNA Yield but Not Integrity or Methylation Is Impacted after Long-Term Storage" *Biopreserv. Biobank., vol. 14, no. 1, pp. 29–38, 2016."*

Group 2 and Group 3 showed higher percentages of total yield across all storage temperatures and times, being close to or statistically equivalent to the reference samples (**Fig. 37**).

Such results are needed to put the following methylation data in context (**Table 5**). Twenty-two genes were tested, across Group 1 samples tested at 0h after extraction that served as the reference values, Group 1 samples stored at -20ºC and -80ºC for 1 year and Group 2 samples stored at room temperature for 1 year. The table displays each gene's average methylated cytosine percentage and unmethylated cytosine percentage (Mean) and the standard deviation (SD). All genes appeared highly unmethylated, with the exception of GDF15, known as the Growth Differentiation Factor 15, a gene that codes a ligand of the TGF-beta [144], which appeared highly methylated. All the genes had the same levels of methylation percentages across groups and time hovering around the half percentile for the highest difference gaps. Between samples for each group, the standard deviation ranged a bit more, with genes like MLH1, which the encoded protein heterodimerizes with mismatch repair endonuclease PMS2, forming a part of the DNA mismatch repair system, the αMutL [145], showing merely 0.01% standard deviation across times and groups, while the gene Cyclin Dependent Kinase Inhibitor 1A (CDKN1A), which the encoded protein binds to and inhibits the activity of cyclin-cyclin-dependent kinase 2 complex, which in turn heavily regulates cell cycle progression at G1 [146], has shown standard deviations of up to 3.32% in Group 1 -80ºC storage. When merely comparing Group 1 at 0h with Group 2 and 1 year, the group that more, if not very, closely resembles what a forensic sample would experience both before and after collection from the scene of crime, the data commences to show that DNA methylation itself, for up to 1 year, doesn't suffer much degradation, being a rather stable biomarker, even if given that fact that the tested DNA on which it resided was artificially preserved by DNAgardBlood solution, making the DNA methylation analysis much easier, or that the number of tested genes was small.

Table 5- Average DNA methylation of the 22 tested genes across all eight subjects. T_0: Group 1 samples at 0h; EDTA: Group 1; DgB before: Group 2. Adapted from *"A. Bulla, B. De Witt, W. Ammerlaan, F. Betsou, and P. Lescuyer,* "Blood DNA Yield but Not Integrity or Methylation Is Impacted after Long-Term Storage" *Biopreserv. Biobank., vol. 14, no. 1, pp. 29–38, 2016."*

			T_0		*EDTA –20°C*		*EDTA –80°C*		*Dgb before RT*	
Gene	RefSeq		UM	M	UM	M	UM	M	UM	M
ATM	NM_000051	Mean	**99.95%**	**0.05%**	**99.98%**	**0.02%**	**99.97%**	**0.03%**	**99.95%**	**0.05%**
		SD	0.06%	0.06%	0.03%	0.03%	0.04%	0.04%	0.04%	0.04%
BNIP3	NM_004052	Mean	**98.07%**	**1.93%**	**97.82%**	**2.18%**	**98.12%**	**1.88%**	**98.52%**	**1.48%**
		SD	1.22%	1.22%	1.84%	1.84%	1.59%	1.59%	1.06%	1.06%
BRCA1	NM_007294	Mean	**99.62%**	**0.38%**	**99.53%**	**0.47%**	**99.55%**	**0.45%**	**99.28%**	**0.72%**
		SD	0.51%	0.51%	0.78%	0.78%	0.57%	0.57%	0.73%	0.73%
CCND1	NM_053056	Mean	**99.96%**	**0.04%**	**99.97%**	**0.03%**	**99.94%**	**0.06%**	**99.97%**	**0.03%**
		SD	0.05%	0.05%	0.02%	0.02%	0.05%	0.05%	0.05%	0.05%
CDKN1A	NM_000389	Mean	**97.93%**	**2.07%**	**97.69%**	**2.31%**	**97.56%**	**2.44%**	**98.36%**	**1.64%**
		SD	0.98%	0.98%	2.33%	2.33%	3.32%	3.32%	1.75%	1.75%
CSTB	NM_000100	Mean	**99.88%**	**0.12%**	**99.86%**	**0.14%**	**99.86%**	**0.14%**	**99.86%**	**0.14%**
		SD	0.06%	0.06%	0.12%	0.12%	0.13%	0.13%	0.15%	0.15%
CYP1A1	NM_000499	Mean	**99.97%**	**0.03%**	**99.97%**	**0.03%**	**99.96%**	**0.04%**	**99.91%**	**0.09%**
		SD	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.05%	0.05%
DNAJC15	NM_013238	Mean	**97.47%**	**2.53%**	**98.01%**	**1.99%**	**97.68%**	**2.32%**	**97.78%**	**2.22%**
		SD	4.29%	4.29%	3.54%	3.54%	3.93%	3.93%	3.76%	3.76%
GADD45A	NM_001924	Mean	**99.04%**	**0.96%**	**98.20%**	**1.80%**	**98.25%**	**1.75%**	**98.89%**	**1.11%**
		SD	0.40%	0.40%	2.31%	2.31%	2.81%	2.81%	1.70%	1.70%
GADD45G	NM_006705	Mean	**100.00%**	**0.00%**	**99.98%**	**0.02%**	**99.99%**	**0.01%**	**99.99%**	**0.01%**
		SD	0.00%	0.00%	0.02%	0.02%	0.01%	0.01%	0.01%	0.01%
GDF15	NM_004864	Mean	**0.87%**	**99.13%**	**0.95%**	**99.05%**	**1.17%**	**98.83%**	**1.18%**	**98.82%**
		SD	0.27%	0.27%	0.16%	0.16%	0.41%	0.41%	0.32%	0.32%
GPX3	NM_002084	Mean	**99.44%**	**0.56%**	**99.45%**	**0.55%**	**99.44%**	**0.56%**	**99.39%**	**0.61%**
		SD	0.29%	0.29%	0.38%	0.38%	0.33%	0.33%	0.32%	0.32%
INSIG1	NM_005542	Mean	**99.58%**	**0.42%**	**99.45%**	**0.55%**	**99.44%**	**0.56%**	**99.65%**	**0.35%**
		SD	0.24%	0.24%	0.65%	0.65%	0.87%	0.87%	0.45%	0.45%
MLH1	NM_000249	Mean	**99.99%**	**0.01%**	**99.99%**	0.01%	**99.99%**	**0.01%**	**99.99%**	**0.01%**
		SD	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
MSH2	NM_000251	Mean	**100.00%**	**0.00%**	**99.91%**	**0.09%**	**100.00%**	**0.00%**	**100.00%**	**0.00%**
		SD	0.00%	0.00%	0.23%	0.23%	0.00%	0.00%	0.00%	0.00%
PRDX2	NM_005809	Mean	**99.97%**	**0.03%**	**99.95%**	**0.05%**	**99.95%**	**0.05%**	**99.93%**	**0.07%**
		SD	0.03%	0.03%	0.05%	0.05%	0.04%	0.04%	0.04%	0.04%
PTGS2	NM_000963	Mean	**99.98%**	**0.02%**	**99.97%**	**0.03%**	**99.98%**	**0.02%**	**99.99%**	**0.01%**
		SD	0.01%	0.01%	0.04%	0.04%	0.02%	0.02%	0.01%	0.01%
RARA	NM_001024809	Mean	**99.37%**	**0.63%**	**99.30%**	**0.70%**	**99.34%**	**0.66%**	**99.23%**	**0.77%**
		SD	0.39%	0.39%	0.34%	0.34%	0.36%	0.36%	0.37%	0.37%
SCARA3	NM_182826	Mean	**99.69%**	**0.31%**	**99.48%**	**0.52%**	**99.55%**	**0.45%**	**99.76%**	**0.24%**
		SD	0.15%	0.15%	0.69%	0.69%	0.67%	0.67%	0.36%	0.36%
TP53	NM_000546	Mean	**99.87%**	**0.13%**	**99.84%**	**0.16%**	**99.89%**	**0.11%**	**99.65%**	**0.35%**
		SD	0.06%	0.06%	0.08%	0.08%	0.07%	0.07%	0.13%	0.13%
UBE2G2	NM_182688	Mean	**100.00%**	**0.00%**	**99.97%**	**0.03%**	**100.00%**	**0.00%**	**100.00%**	**0.00%**
		SD	0.00%	0.00%	0.04%	0.04%	0.00%	0.00%	0.01%	0.01%
XPC	NM_004628	Mean	**99.98%**	**0.02%**	**99.97%**	**0.03%**	**99.97%**	**0.03%**	**99.99%**	**0.01%**
		SD	0.01%	0.01%	0.06%	0.06%	0.04%	0.04%	0.02%	0.02%

On a crime scene with a missing body however, the sampled DNA can come from old, dried blood left behind during the offence, instead of being directly drawn and preserved immediately from a body. Hollegaard *et al.* tackled this angle somewhat by studying the viability of methylome screening using dried neonatal blood spot samples [147]. Two individuals had neonatal venous blood samples dried in pure cellulose collection cards, stored for 28 and 26 years until analysis (neoDBSS); venous blood drawn and analyzed at the ages of 25 and 23 (Reference); venous blood collected and dried on pure cellulose collection cards at the ages of 25 and 23 and stored at -20ºC for 3 years until analysis (RefDBSS). Methylome was studied through bisulfite conversion and Infinium HumanMethylation27, of which, from the neonatal dried blood spot samples, only 30ng of

DNA were used for the initial bisulfite conversion, close to the expected amount available in forensic samples with a missing body. Methylation data was obtained through fluorescent probe labeling, with a total of 27,578 probes divided as: Probe group A, red signal, representing unmethylated CpG locations; Probe group B, green signal, representing methylated CpG locations, being held against CpGenome Universal Methylated DNA (Millipore) as a positive control and whole-genome amplified eldest individual before bisulfite treatment as negative control. Each CpG location had a β-value attributed to it, obtained through averaging both probe groups signals, ranging from 0 to 100% methylation (**Fig. 38**).

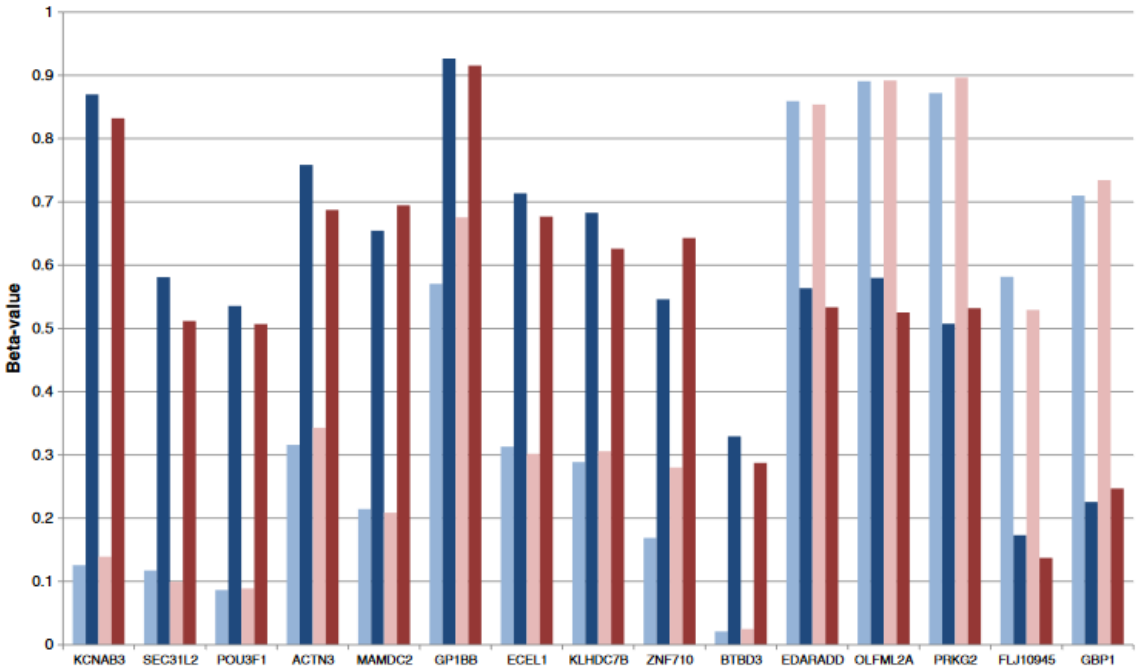

Fig. 38- Graph plotting β-values of the genes that most changed between neonatal dried blood spot samples (light blue and light red) and adulthood dried blood spot samples (dark blue and dark red). Adapted from "*M. V. Hollegaard, J. Grauholm, B. Nørgaard-Pedersen, and D. M. Hougaard,* "DNA methylome profiling using neonatal dried blood spot samples: A proof-of-principle study" *Mol. Genet. Metab., vol. 108, no. 4, pp. 225–231, 2013.*"

Following such, the two used DNA extraction protocols used, Extract-N-Amp Blood PCR kit (ENA) and ChargeSwitch Forensic DNA Purification kit (CS), were tested for performance (**Fig. 39**). A thinner plot cloud and a more centered plot cloud around the theoretical ideal line represents a higher level of correlation, meaning that DNA extracted through ChargeSwitch Forensic DNA Purification kit (CS) showed a lesser fidelity to the real data, caused by an insufficient amount of extracted DNA, and caused CS-extracted samples to be discarded as a reliable source of data.

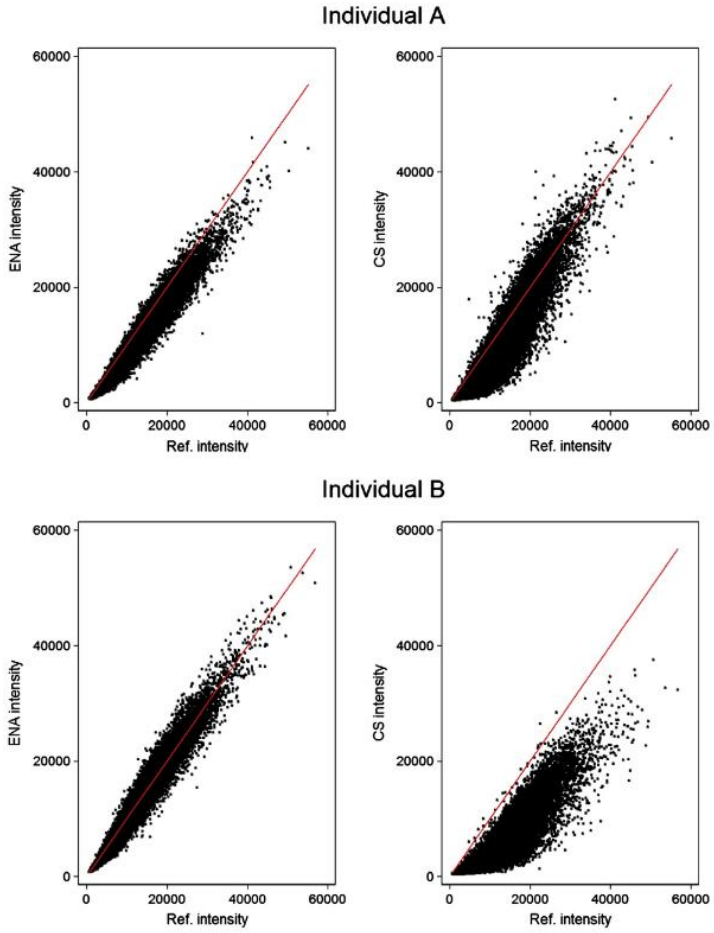

Fig. 39- Signal intensity from RefDBSS samples extracted with ENA (right) and CS (left) kits of individual A (eldest) and individual B (youngest) plotted against the Reference samples. Adapted from "*M. V. Hollegaard, J. Grauholm, B. Nørgaard-Pedersen, and D. M. Hougaard*, "DNA methylome profiling using neonatal dried blood spot samples: A proof-of-principle study" *Mol. Genet. Metab., vol. 108, no. 4, pp. 225–231, 2013.*"

Afterwards, it was tested if the RefDBSS samples stored for 3 years were an accurate representation of the real methylome, represented by the Reference samples. Both ENA and CS samples were tested against each individual's reference sample in a differential methylation analysis. ENA extracted RefDBSS of each individual presented no changes from the respective individual's Reference sample's methylome, indicating that they are equivalent in terms of DNA methylation, even after stored as dried blood samples for 3 years at -20ºC. In contrast, CS extracted RefDBSS presented several deviations from the correspondent Reference samples, indicating that the extraction process is once again unviable for such small amounts of available input DNA.

Table 6- Differential methylation analysis between reference samples and dried blood spot samples extracted by ENA and CS. Adapted from "*M. V. Hollegaard, J. Grauholm, B. Nørgaard-Pedersen, and D. M. Hougaard,* "DNA methylome profiling using neonatal dried blood spot samples: A proof-of-principle study" *Mol. Genet. Metab., vol. 108, no. 4, pp. 225–231, 2013.*"

			Hypermethylated			Hypomethylated		
Reference	Sample	Individual	No.	Δβ mean	Δβ range	No.	Δβ mean	Δβ range
Individual A	Ref.	B	402	0.4031	(0.2032; 0.7161)	25	−0.328	(−0.5275; −0.2226)
	refDBSS (ENA)	A	0	0	0	0	0	0
		B	345	0.3764	(0.2199; 0.6430)	29	−0.3344	(−0.5144; −0.2411)
	refDBSS (CS)	A	128	0.2974	(0.2058; 0.7882)	858	−0.2972	(−0.5178; −0.1585)
		B	1080	0.3112	(0.1681; 0.7949)	2513	−0.3119	(−0.7351; −0.1514)
Individual B	Ref.	A	25	0.328	(0.2226; 0.5275)	402	−0.4031	(−0.7161; −0.2033)
	refDBSS (ENA)	A	25	0.3156	(0.2109; 0.4947)	385	−0.4044	(−0.7059; −0.2112)
		B	0	0	0	0	0	0
	refDBSS (CS)	A	234	0.2823	(0.1762; 0.7975)	1439	−0.3205	(−0.7022; −0.1540)
		B	885	0.2604	(0.1560; 0.6718)	2552	−0.3131	(−0.6519; −0.1535)

As RefDBSS samples are faithful representations of the methylome, they were used as the referenced to which the neoDBSS were tested against. Again, the comparison was done with a differential methylation analysis (**Table 7**). When checked between the same individual, RefDBSS and neoDBSS samples showed high similarity. Individual A neoDBSS samples when compared to individual A RefDBSS, presented 19 hypermethylated and 31 hypomethylated CpG locations, with β-value means of roughly 0.33 and -0.35, translating to roughly 30% methylation difference in average on those locations, out of 27,578 locations. Similarly, individual B neoDBSS samples when compared to individual B RefDBSS, presented 35 hypermethylated and 16 hypomethylated CpG locations, with β-value means of roughly 0.34 and -0.37, translating to roughly yet again 30% methylation difference in average on those locations, out of 27,578 locations. When comparing both RefDBSS and neoDBSS of one individual to the RefDBSS of the other, its noted that the difference between hypermethylated and hypomethylated CpG locations mirrors the one present in the intra-individual analysis. Individual B neoDBSS and individual B RefDBSS samples when compared against individual A RefDBSS showed 89 hypermethylated and 37 hypomethylated CpG locations, with β-value means of roughly 0.37 and -0.34.

Table 7- Differential methylation analysis between ENA extracted RefDBSS samples and ENA extracted neoDBSS samples. Adapted from "*M. V. Hollegaard, J. Grauholm, B. Nørgaard-Pedersen, and D. M. Hougaard,* "DNA methylome profiling using neonatal dried blood spot samples: A proof-of-principle study" *Mol. Genet. Metab., vol. 108, no. 4, pp. 225–231, 2013.*"

			Hypermethylated			Hypomethylated		
Reference	Sample	Individual	No.	Δβ mean	Δβ range	No.	Δβ mean	Δβ range
refDBSS (A)	refDBSS (ENA)	A	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
		B	332	0.3766	(0.2007; 0.6328)	22	−0.3445	(−0.4700; −0.2333)
	neoDBSS (ENA)	A	19	0.3328	(0.2574; 0.4843)	31	−0.3586	(−0.7440; −0.2401)
		B	421	0.3685	(0.1994; 0.6701)	59	−0.3305	(−0.7304; −0.2203)
refDBSS (B)	refDBSS (ENA)	A	22	0.3445	(0.2333; 0.4700)	332	−0.3766	(−0.6328; −0.2007)
		B	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
	neoDBSS (ENA)	A	55	0.3203	(0.1885; 0.5112)	398	−0.3646	(−0.7063; −0.1976)
		B	35	0.3456	(0.2400; 0.5036)	16	−0.3731	(−0.6927; −0.2399)

Hollegaard *et al.*'s pilot work tentatively shows that blood samples stored in pure cellulose disks, without any blood preservatives nor DNA stabilizers, dried and kept at -20ºC for over 25 years were able to provide reliable, true to reference, global methylome status from very small amounts of input DNA. Hollegaard *et al.'s* work is, however, very small in scale, and as such, the results are a mere test to the concept of DNA methylation stability *postmortem*, albeit one passed successfully.

6.3. Obstacles to Overcome

6.3.1. DNA Methylation and Pathologies

One of the guidelines for the viability of a marker for forensic usage is that it must not be related to any pathology, as it may skew results and create doubt in the validity of the results. Epigenetic information regulates DNA expression and, as a result, irregularities in its maintenance or levels in many genes can lead to pathologies like cancer, or even be deregulated as collateral damage of a lifelong medication intake, has seen in the study performed by J. Mill *et al.* where it was discovered that a lifelong antipsychotic usage in schizophrenia patients led to a localized aberrant methylation in the MEK1 gene promoter region [133]. Many, many DNA methylation irregularities in certain genes haven been found in patients with psychological pathologies and cancers.

Some of the known deregulated genes linked to depression are the Brain Derived Neurotrophic Factor (BDNF) gene, a gene that codes an important protein responsible for promoting neuron survival, growth, maturation and maintenance, while maintaining activity in synapses, presenting increased methylation [148]–[151]. The Solute Carrier Family 6 Member 4 (SLC6A4), which codes an important serotonin membrane transporter protein that recycles it from the synaptic valley into the presynaptic neuron, is hypermethylated in

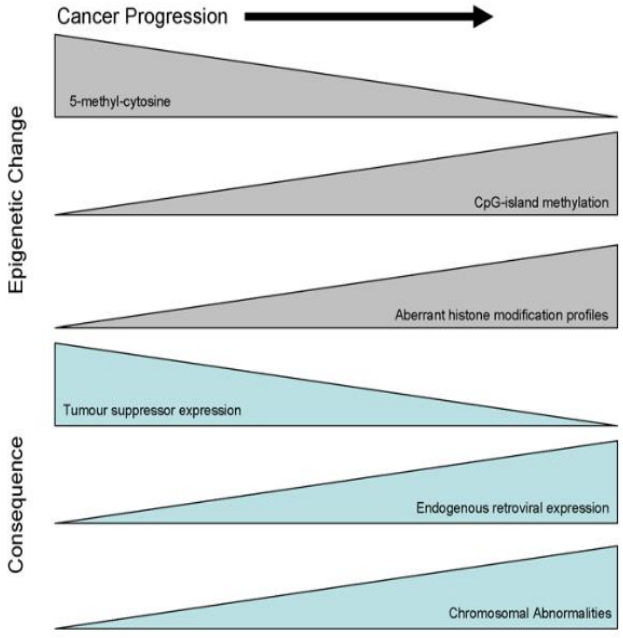

Fig. 40- Relation between cancer progression, DNA methylation, histone methylation and CpG island methylation. As cancer progresses in stages, there is an overall loss of DNA methylation, yet an hypermethylation of promoter dwelling CpG islands and an increase in incorrect histone methylation patterns. As a consequence of the latter, tumor suppressor genes may become silenced and retrotransposon based mobile interspersed repeats may become active, causing genome instability by inserting themselves in critical gene areas. Adapted from "*M. Hirst and M. A. Marra, "The International Journal of Biochemistry Epigenetics and human disease" Int. J. Biochem., vol. 41, pp. 136–146, 2009*."

patients with depression [152]–[154] and hypomethylated on patients undergoing antidepressant treatment [155], [156]. The Nuclear Receptor Subfamily 3 Group C Member 1 (NR3C1), which codes a glucocorticoid receptor, has inconsistent results, but has shown some slight tendency to be hypermethylated [157], [158] but there are also report of hypomethylation [159]. Currently, DNA hypomethylation in CpG dinucleotides, repetitive sequences, CpG islands hypermethylation, alongside general loss of methylation, is a well-researched marker for cancer (**Fig. 40**) [160].

Non-coding areas of the genome are also linked between aberrant methylation and pathologies. LINE-1 hypomethylation has been present in many forms of cancer [78]. Tandem repeats, like Satellite α, a structurally critical centromeric heterochromatin located satellite DNA present in all human chromosomes, similarly carry connections with cancer, with hypomethylation usually accompanying a bad prognosis [161], due to the fact hypomethylation of these areas causes genomic and structural instability. The presence of polymorphic tandem repeats in a gene's promoter region has likewise been found to create variation in its methylation levels [162], but no link between their presence, number of repeats and size has been linked to pathologies yet.

The possibility that nearly any DNA methylation deregulation, within or outside genes, be it coding, regulatory or structural expanses of DNA, can cause, be linked to or being a symptom of a wide range of pathologies will be something that will require further study to resolve, be it through the finding of methylation locations that are exempt of this issue or the finding of a way to offset variability caused by such cases, either through a larger battery of biomarkers or stronger internal standards.

6.3.2. DNA Methylation variation with aging

There are efforts to create a methodology that allows the age of a human to be derived from epigenetic information currently being made. Several genes already have been found to suffer hypomethylation at a stable rate with age, stable enough to create predictive models as accurate as 5 years [163]. Bocklandt *et al.* has determined 10 of such genes: Glycine Receptor Alpha 1, GLRA1; Neuronal Pentraxin 2, NPTX2; TSPY Like 5, TSPYL5; Solute Carrier Family 5 Member 7, SLC5A7; ATPase Phospholipid Transporting 8A2, ATP8A2; Leptin, LEP; Complexin 2, CPLX2; Beta-1,3-Galactosyltransferase 6, B3GALT6; CUGBP Elav-like family member 6, CELF6; Potassium Two Pore Domain Channel Subfamily K Member 12; KCNK12. Those 10 genes were part of an 88 gene pool found to have methylation status changes related with age, most with varying rates across time (**Table 8**).

Table 8- Gene symbols of the 88 genes found to have methylation status changes related with age. Adapted from "*S. Bocklandt et al.,* "Epigenetic Predictor of Age" *PLoS One, 2011."*

High ⟶ Strength of correlation with age ⟶ Low

Hypermethylation				Hypomethylation
KCNG3	IRF8	MCHR2	LRRC2	ASPA
NPTX2	SKIP	FAM19A4	NEF3	Bles03
GREM1	CPLX2	RGC32	SPAG6	EDARADD
VGF	POU4F2	PCDHGB7	GCM2	TCEA2
GRIA2	KLF14	FBLN2	IRXL1	ELN
PDE4C	GATA4	SLC15A3	NRN1	PGLYRP2
FLJ42486	TBX20	PCDHGB4	SMPD3	LGP1
ATP8A2	FLJ90650	LEP	ZNF671	TOM1L1
KCNK12	NEFH	GRIA2	HOXB4	LAG3
C10orf82	PCDH8	SHOX2	MOXD1	SLC14A1
ZIC1	ZNF540	LOC349136	TCL1A	CSNK1D
BRUNOL6	ADRB1	KCNC1	KCNA5	ACSS2
FLJ42486	ATP8A2	TSPYL5		C9orf66
ZNF667	VMP	PNMA2		CENTD3
TRIM58	GATA4	WNK2		HNRPL
ZNF540	GLRA1	ADRA2C		CDH5
B3GALT6	KCNC3	SLC5A7		ABHD14A
DCC	ZNF154	GATA4		LTBR
HTR7	LEP	BARHL2		RENBP

Even though DNA methylation suffers an overall small decline with age [30], these genes carry more variability that could diminish the methodology's accuracy, being poor choices for *postmortem* interval biomarkers, and any potential biomarker should be checked for correlation with age, but they could be included in the methodology's data retrieval process as they could provide more information about the victim's age for virtually no added cost or time to the procedure, while being kept separate from the PMI assessment battery. The methylation status of the EDARADD, TOM1L1, and NPTX2 genes are linear with aging for five decades [163], being good candidates for this potential side addition to the methodology in the future.

6.3.3. DNA Methylation variation due to the environment

DNA methylation is a very resilient form of DNA regulation that undergoes strict regulation itself, but it is not an ironclad set of rules. DNA methylation is a plastic layer of regulation that can function both as a long-term response to environmental pressures and stresses in the context of generations of a specie or as short term response in the context of the life span of a single individual [164]. There are many, many forms of environmental and

behavioral pressures that can influence the methylation patterns, both at the global level, gene level and CpG location level.

6.3.4. Validation

For a novel forensic DNA analysis methodology to be accepted and validated it must go through rigorous studying and testing. To be accepted by the Federal Bureau of Investigation of the United States of America, a methodology requires extensive knowledge of the genetic biomarker in question, any species specificity that may change data between model organism and human testing; extensive studies for the methodology's stability, sensitivity, reproducibility, precision and accuracy; have case-studies that validate the methodology's theory and practice; if a methodology is PCR based it also must provide reaction conditions, testing for differential and preferential amplification, effects of multiplexing, assessment of appropriate controls and product detection studies [165]. Accordingly, the European Network of Forensic Science Institutes has its own guidelines for methodology validation [166]. It is a multiple step process, with many different requirements for different types of methodologies, be them human based, automatic based, qualitative or quantitative, *etc*, much harder to pass successfully if the methodology is brand new and underlying data and science has not yet been standardized by standards approving organizations like International Organization for Standardization, European Committee for Standardization or American Society for Testing and Materials. The process begins with a submission of the validation plan, which requires extensive information about the method, such as:

- Measurement procedure, roughly outlying the basic premise of the methodology;
- The biomarkers in study;
- What property of the biomarker is measured and its measurement unit;
- The range of the measurements;
- The kind of samples the methodology accepts, including the type of matrixes;
- The intended usage of the results, that for forensic methodologies usually being court reports and intelligence useful for the police force;
- The precision of the methodology, with repeatability measured in RSDr5 %
- Percentage of methodology bias and possible methodology selectiveness;
- The robustness of the methodology and its uncertainty.
- If it is a non-modified standards approving organizations protocol, modified protocol or new method;
- If the methodology is quantitative, the extraction systems, the quantitation systems, the amplification systems, the reactional systems and the detection systems must pass, with provided sensitivity, stochastic and mixture studies, tests for repeatability, reproducibility, contamination assessment and tests with known samples.

Afterwards, a long process of peer-reviewing and standardization is conducted and after accreditation a methodology may be accepted as forensic evidence. The process requires large amounts of studies over enough independent laboratories to establish the foundation for the acceptance of the underlying scientific basis and large amounts of results that fall under the accepted ranges of reproducibility and repeatability. For a *postmortem* interval assessing methodology based on epigenetics to be created from the current availability of information as a starting point, a joint effort by many independent laboratories will be required to create such vast amounts of evidence required by the many international standards that try to regulate and intercommunicate forensic science, data and regulations.

7. Closing Thoughts

Epigenetic information has incredibly stable biomarkers throughout the entire life of an organism, and, as more recent studies have shown, epigenetic information is also incredibly stable after death. Both dried blood spots [147] and brain tissue [139]–[141] provide reliable, stable and pristine methylation data when stored at even room temperatures, even if the DNA itself where it is stored becomes very degraded and hard to work with [143].

Yet, even with the preliminary studies showing that its stability is suited for the for this kind of forensic usage, there are many variables that need to be accounted for before the work of creating a battery viable for PMI assessment. Firstly, the biomarkers used need to be thoroughly studied, not only for normal methylation levels in all stages of life and degradation rates *postmortem*, but also for possible changes it can suffer throughout an organism's life, as DNA methylation changes as a response to environmental stimuli, changes that are enough to even distinguish monozygotic twins [23], [24]. Ideally, biomarkers should be chosen taking in consideration what kind of response they produce to batches of stimuli, to make sure any major variation due to lifestyle choices, response to local pollution levels, *etc*, is kept to the minimal amount of biomarkers in the battery, as well as choosing biomarkers that are also affected by the least amount of these pressures. LINE-1 are preliminary prime candidates for PMI assessment, as they are large, spread out sections of DNA kept under high amounts of methylation and regulation, not directly linked to any protein that could function as a response to environmental stimuli. Other satellite DNA, like *Alu*, should also be the subject of scrutiny due to the striking similarities between itself and LINE-1. Individual gene biomarkers, on the other hand, will be heavily reliant on a case-by-case study, and only further study will show more candidates for a possible battery of biomarkers. In the terms of internal standards, the battery possesses global methylation as the normalization factor between samples and a large array of housekeeping genes to be used as negative controls. The battery could also incorporate extra biomarkers outside of the core PMI assessing battery like the EDARADD, TOM1L1 and NPTX2 genes, whose methylation levels have been linked to biological passage of time in a linear way, possibly adding age of the sample donor to the information the methodology can provide.

DNA availability is a big variable as well. Upon finding a body, there will be no shortage of organic matrixes to obtain DNA from, albeit its condition may be quite degraded if the body itself is already in late stages of decomposition. Blood splatters that may be spread around the crime scene where the victim was originally assaulted but no body being present however, provide much lower quantities of DNA and may be all that it may be available. Albeit blood splatters do not confirm death and may give different *postmortem* intervals after a possible body recovery, they do help to cement the time at which the confrontation may have occurred and if the death was caused during it, thus strengthening the timeline of events, and subsequently the investigation. Thus, it would be highly recommended that the techniques chosen to perform the PMI assessments are capable of handling the low quantities of DNA usually present in forensics whenever a body has not

yet been found, in order for the methodology to be able to provide the most amount of information in the largest amounts of cases in a standard form. The discrepancy between *postmortem* intervals between a retrieved corpse and a blood splatter obtained from a crime scene can be interpreted in various ways: if the PMIs don't match, there was an assault or struggle at the location of the blood splatters but the death did not occur during it, and the larger the gap, the more information about the case may be missing; if the PMI of the blood splatter matches when a suspect was seen at the scene but the PMI of the corpse doesn't, with enough of a gap, the suspect might possess an alibi for that time, or the victim was abandoned and emergency services were not contacted, denoting criminal intent; if the PMIs match but the body and samples were not recovered at the same location, it can be inferred that the body was moved *postmortem*, and both locations become tightly linked. As any information during investigation is precious, this take on the methodology is highly recommended.

While the amount of *postmortem* methylation data and its stability gathered by the scientific community is plentiful, the study of the key rates of degradation, both globally and of individual loci, are far and few, studies of methylation using old samples under poor storage conditions are extremely rare, and the forensic study of this particular use of methylation is in its early stages, but the preliminary research and the research done in other areas of expertise point towards a good prognosis to the epigenetic assessment of *postmortem* intervals.

8. Bibliography

[1] W. R. Trumble, *Shorter Oxford English Dictionary: Sixth Edition*, 6th ed. Oxford University Press, 2007.

[2] E. et al. Weizman, *Forensis The Architecture of Public Truth*, 1st ed. Sternberg Press and Forensic Architecture, Goldsmiths, University of London, London, 2014.

[3] D. Asen, "Song Ci (1186–1249), 'Father of World Legal Medicine': History, Science, and Forensic Culture in Contemporary China," *East Asian Sci. Technol. Soc.*, vol. 11, no. 2, pp. 185–207, 2017.

[4] L. M. Fonseca and I. Testoni, "The Emergence of Thanatology and Current Practice in Death Education," *OMEGA - J. Death Dying*, vol. 64, no. 2, pp. 157–169, 2012.

[5] David E. Balk and David K. Meagher, *Handbook of Thanatology: The Essential Body of Knowledge for the Study of Death, Dying, and Bereavement*, 2nd ed. Routledge, 2013.

[6] A. Franco, S. D. S. C. Mendes, F. F. Picoli, L. G. Rodrigues, and R. F. Silva, "Forensic thanatology and the pink tooth phenomenon: From the lack of relation with the cause of death to a potential evidence of cadaveric decomposition in dental autopsies — Case series," *Forensic Sci. Int.*, vol. 291, pp. e8–e12, 2018.

[7] J. Payne-James, R. Jones, S. B. Karch, and J. Manlove, *Simpson's Forensic Medicine, 13th Edition*, 13th Editi. CRC Press, 2011.

[8] M. L. Goff, *Forensic Entomology*, Second Edi. Elsevier Inc., 2009.

[9] D. Donlon, R. Lain, and J. A. Taylor, "Forensic anthropology," *Forensic Odontol. Princ. Pract.*, vol. 11, no. 1, pp. 336–354, 2016.

[10] M. Bauer, I. Gramlich, S. Polzin, and D. Patzelt, "Quantification of mRNA degradation as possible indicator of postmortem interval - A pilot study," *Leg. Med.*, vol. 5, no. 4, pp. 220–227, 2003.

[11] H. A. Elghamry, M. I. Mohamed, F. M. Hassan, D. S. Abdelfattah, and A. G. Abdelaal, "Potential use of GAPDH m-RNA in estimating PMI in brain tissue of albino rats at different environmental conditions," *Egypt. J. Forensic Sci.*, vol. 7, no. 1, pp. 0–5, 2017.

[12] A. T. Schäfer, "Colour measurements of pallor mortis," *Int. J. Legal Med.*, vol. 113, no. 2, pp. 81–83, 2000.

[13] M. L. Goff, "Early post-mortem changes and stages of decomposition in exposed cadavers," *Exp. Appl. Acarol.*, vol. 49, no. 1–2, pp. 21–36, 2009.

[14] D. L. Cockle and L. S. Bell, "Human decomposition and the reliability of a 'Universal' model for post mortem interval estimations," *Forensic Sci. Int.*, vol. 253, pp. 136.e1-136.e9, 2015.

[15] F. Sampaio-Silva, T. Magalhães, F. Carvalho, R. J. Dinis-Oliveira, and R. Silvestre, "Profiling of RNA Degradation for Estimation of Post Morterm Interval," *PLoS One*, vol. 8, no. 2, 2013.

[16] Y. H. Lv *et al.*, "RNA degradation as described by a mathematical model for postmortem interval determination," *J. Forensic Leg. Med.*, 2016.

[17] A. Ghosh and M. Bansal, "A glossary of DNA structures from A to Z," *Acta Crystallogr. - Sect. D Biol. Crystallogr.*, vol. 59, no. 4, pp. 620–626, 2003.

[18] J. M. Berg, J. L. Tymoczko, and L. Stryer, "Biochemistry," in *Biochemistry*, 5th ed., 2002, pp. 194–195.

[19] T. Shafee and R. Lowe, "Eukaryotic and prokaryotic gene structure," vol. 4, no. 1,

pp. 0–3, 2017.
[20] R. A. Young, “RNA Polymerase II,” *Annu. Rev. Biochem.*, vol. 60, no. 1, pp. 689–715, 1991.
[21] R. D. Kornberg, “Eukaryotic transcriptional control,” *Trends Biochem. Sci.*, vol. 24, no. 12, pp. 46–49, 1999.
[22] J. G. Herman, J. R. Graff, S. Myohanen, B. D. Nelkin, and S. B. Baylin, “Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands.,” *Proc. Natl. Acad. Sci.*, vol. 93, no. 18, pp. 9821–9826, 1996.
[23] A. Vidaki and M. Kayser, “From forensic epigenetics to forensic epigenomics: Broadening DNA investigative intelligence,” *Genome Biol.*, vol. 18, no. 1, pp. 1–13, 2017.
[24] J. E. Castillo-Fernandez, T. D. Spector, and J. T. Bell, “Epigenetics of discordant monozygotic twins: Implications for disease,” *Genome Med.*, vol. 6, no. 7, pp. 1–16, 2014.
[25] Z. D. Smith and A. Meissner, “DNA methylation: Roles in mammalian development,” *Nat. Rev. Genet.*, vol. 14, no. 3, pp. 204–220, 2013.
[26] A. Hermann, R. Goyal, and A. Jeltsch, “The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites,” *J. Biol. Chem.*, vol. 279, no. 46, pp. 48350–48359, 2004.
[27] M. Kulis and M. Esteller, “DNA Methylation and Cancer,” vol. 70, no. 10, 2010, pp. 27–56.
[28] D. J. Weisenberger *et al.*, “Analysis of repetitive element DNA methylation by MethyLight,” *Nucleic Acids Res.*, vol. 33, no. 21, pp. 6823–6836, 2005.
[29] M. Ehrlich *et al.*, “Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells,” *Nucleic Acids Res.*, vol. 10, no. 8, pp. 2709–2721, 1982.
[30] F. Kader and M. Ghai, “DNA methylation and application in forensic sciences,” *Forensic Sci. Int.*, vol. 249, pp. 255–265, 2015.
[31] A. Vidaki, B. Daniel, and D. S. Court, “Forensic DNA methylation profiling - Potential opportunities and challenges,” *Forensic Sci. Int. Genet.*, vol. 7, no. 5, pp. 499–507, 2013.
[32] E. Giardina, “DNA Fingerprinting,” *Brenner’s Encycl. Genet. Second Ed.*, vol. 2, pp. 356–359, 2013.
[33] H. Fan and J. Y. Chu, “A Brief Review of Short Tandem Repeat Mutation,” *Genomics, Proteomics Bioinforma.*, vol. 5, no. 1, pp. 7–14, 2007.
[34] “Combined DNA Index System (CODIS).” [Online]. Available: https://www.fbi.gov/services/laboratory/biometric-analysis/codis#CODIS-Brochure. [Accessed: 08-Mar-2019].
[35] The Crown Prosecution Service, “DNA-17 Profiling,” 2017. [Online]. Available: https://www.cps.gov.uk/legal-guidance/dna-17-profiling. [Accessed: 01-Oct-2019].
[36] Interpol, “DNA,” 2019. [Online]. Available: https://www.interpol.int/en/How-we-work/Forensics/DNA. [Accessed: 01-Oct-2019].
[37] K. L. O’Connor, “Interpretation of DNA Typing Results for Kinship Analysis,” 2011.
[38] N. W. Blackstone, “Phylogenetic Relationships of Danio Within the OrderCypriniformes:AFramework forComparative and Evolutionary Studies of a Model Species,” *J. Exp. Zool. B. Mol. Dev. Evol.*, vol. 306, no. 1, pp. 1–7, 2006.

[39] A. G. K. Menon, *Check List--fresh Water Fishes of India*, no. 175. Survey, 1999.
[40] F. Hamilton, *An account of the fishes found in the river Ganges and its branches* , vol. plates. Edinburgh : Printed for A. Constable and company; [etc., etc.], 1822.
[41] R. Spence, G. Gerlach, C. Lawrence, and C. Smith, "The behaviour and ecology of the zebrafish, Danio rerio," *Biol. Rev.*, vol. 83, no. 1, pp. 13–34, 2008.
[42] K. Howe *et al.*, "The zebrafish reference genome sequence and its relationship to the human genome," *Nature*, vol. 496, no. 7446, pp. 498–503, 2013.
[43] M. G. Goll and M. E. Halpern, "DNA methylation in zebrafish," *Prog. Mol. Biol. Transl. Sci.*, vol. 101, pp. 193–218, 2011.
[44] K. C. Sadler, K. N. Krahn, N. A. Gaur, and C. Ukomadu, "Liver growth in the embryo and during liver regeneration in zebrafish requires the cell cycle regulator, uhrf1," *Proc. Natl. Acad. Sci.*, vol. 104, no. 5, pp. 1570–1575, 2007.
[45] Roscoe B. Jackson Memorial Laboratory, *Biology of the laboratory mouse*, 2nd ed. New York: Dover Publications, Inc, 1966.
[46] M. Johnson, "Laboratory Mice and Rats." [Online]. Available: https://www.labome.com/method/Laboratory-Mice-and-Rats.html. [Accessed: 09-Aug-2019].
[47] J. Barau *et al.*, "The DNA methyltransferase DNMT3C protects male germ cells from transposon activity," *Science (80-.).*, vol. 354, no. 6314, pp. 909–912, 2016.
[48] E. Eisenberg and E. Y. Levanon, "Human housekeeping genes, revisited," *Trends Genet.*, vol. 29, no. 10, pp. 569–574, 2013.
[49] J. Pampel, "Housekeeping genes," 2017. [Online]. Available: https://www.genomics-online.com/resources/16/5049/housekeeping-genes/. [Accessed: 04-Sep-2019].
[50] C. Zhang, Y. Q. Wang, G. Jin, S. Wu, J. Cui, and R. F. Wang, "Selection of reference genes for gene expression studies in human bladder cancer using SYBR-green quantitative polymerase chain reaction," *Oncol. Lett.*, vol. 14, no. 5, pp. 6001–6011, 2017.
[51] UniProtKB, "UniProtKB - O43633 (CHM2A_HUMAN)," 2019. [Online]. Available: https://www.uniprot.org/uniprot/O43633. [Accessed: 04-Sep-2019].
[52] UniProtKB, "UniProtKB - P49721 (PSB2_HUMAN)," 2019. [Online]. Available: https://www.uniprot.org/uniprot/P49721. [Accessed: 04-Sep-2019].
[53] UniProtKB, "UniProtKB - P28070 (PSB4_HUMAN)," 2019. [Online]. Available: https://www.uniprot.org/uniprot/P28070. [Accessed: 04-Sep-2019].
[54] UniProtKB, "UniProtKB - P51149 (RAB7A_HUMAN)," 2019. [Online]. Available: https://www.uniprot.org/uniprot/P51149. [Accessed: 04-Sep-2019].
[55] National Center for Biotechnology Information, "SNRPN small nuclear ribonucleoprotein polypeptide N [Homo sapiens (human)] - Gene - NCBI," 2015. [Online]. Available: http://www.ncbi.nlm.nih.gov/gene/6638. [Accessed: 04-Sep-2019].
[56] National Center for Biotechnology Information, "VCP valosin containing protein [Homo sapiens (human)]," 2019. [Online]. Available: http://www.ncbi.nlm.nih.gov/gene/7415. [Accessed: 04-Sep-2019].
[57] UniProtKB, "UniProtKB - Q9UBQ0 (VPS29_HUMAN)," 2019. [Online]. Available: https://www.uniprot.org/uniprot/Q9UBQ0. [Accessed: 04-Sep-2019].
[58] V. Medina Villaamil *et al.*, "GAPDH, YWHAZ, and RRN18S as control reference genes for gene expression studies on renal cell carcinoma (RCC) formaldehyde-fixed paraffin-embedded (FFPE) tissue samples.," *J. Clin. Oncol.*, vol. 29, no.

7_suppl, p. 389, Mar. 2011.
[59] GeneCards, “ATP5F1B Gene (Protein Coding),” 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=ATP5F1B&keywords=ATP5B. [Accessed: 05-Sep-2019].
[60] GeneCards, “HSP90AB1 Gene (Protein Coding),” 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=HSP90AB1&keywords=HSP90AB1. [Accessed: 05-Sep-2019].
[61] GeneCards, “S100A6 Gene (Protein Coding),” 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=S100A6&keywords=S100A6. [Accessed: 05-Sep-2019].
[62] GeneCards, “UBB Gene (Protein Coding),” 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=UBB&keywords=UBB. [Accessed: 05-Sep-2019].
[63] GeneCards, “POLR2A Gene (Protein Coding),” 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=POLR2A&keywords=POLR2A. [Accessed: 05-Sep-2019].
[64] GeneCards, “ACTB Gene (Protein Coding),” 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=ACTB&keywords=ACTB. [Accessed: 05-Sep-2019].
[65] GeneCards, “GAPDH Gene (Protein Coding),” 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=GAPDH&keywords=GAPDH. [Accessed: 05-Sep-2019].
[66] GeneCards, “PGK1 Gene (Protein Coding),” 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=PGK1&keywords=PGK1. [Accessed: 05-Sep-2019].
[67] GeneCards, “PPIA Gene (Protein Coding),” 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=PPIA&keywords=PPIA. [Accessed: 05-Sep-2019].
[68] GeneCards, “RPL13A Gene (Protein Coding),” 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=RPL13A&keywords=RPL13A. [Accessed: 05-Sep-2019].
[69] GeneCards, “RPLP0 Gene (Protein Coding),” 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=RPLP0&keywords=RPLP0. [Accessed: 06-Sep-2019].
[70] GeneCards, “B2M Gene (Protein Coding),” 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=B2M&keywords=B2M. [Accessed: 06-Sep-2019].
[71] GeneCards, “YWHAZ Gene (Protein Coding),” 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=YWHAZ&keywords=YWHAZ. [Accessed: 06-Sep-2019].
[72] GeneCards, “SDHA Gene (Protein Coding),” 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=SDHA&keywords=SDHA. [Accessed: 06-Sep-2019].
[73] GeneCards, “TFRC Gene (Protein Coding),” 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=TFRC&keywords=TFRC. [Accessed: 06-Sep-2019].
[74] GeneCards, “GUSB Gene (Protein Coding),” 2019. [Online]. Available:

https://www.genecards.org/cgi-bin/carddisp.pl?gene=GUSB&keywords=GUSB. [Accessed: 06-Sep-2019].
[75] GeneCards, "HMBS Gene (Protein Coding)," 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=HMBS&keywords=HMBS. [Accessed: 06-Sep-2019].
[76] GeneCards, "HPRT1 Gene (Protein Coding)," 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=HPRT1&keywords=HPRT1. [Accessed: 06-Sep-2019].
[77] GeneCards, "TBP Gene (Protein Coding)," 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=TBP&keywords=TBP. [Accessed: 06-Sep-2019].
[78] Y. Baba *et al.*, "Long Interspersed Element-1 Methylation Level as a Prognostic Biomarker in Gastrointestinal Cancers," *Digestion*, vol. 97, no. 1, pp. 26–30, 2018.
[79] A. S. Yang, "A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements," *Nucleic Acids Res.*, vol. 32, no. 3, pp. 38e – 38, 2004.
[80] M. A. Garrido-Ramos, "Satellite DNA: An evolving topic," *Genes (Basel).*, vol. 8, no. 9, 2017.
[81] L. L. Sullivan, K. Chew, and B. A. Sullivan, "α satellite DNA variation and function of the human centromere," *Nucleus*, vol. 8, no. 4, pp. 331–339, 2017.
[82] Scelfo and Fachinetti, "Keeping the Centromere under Control: A Promising Role for DNA Methylation," *Cells*, vol. 8, no. 8, p. 912, 2019.
[83] P. A. Jones, "Functions of DNA methylation: Islands, start sites, gene bodies and beyond," *Nat. Rev. Genet.*, vol. 13, no. 7, pp. 484–492, 2012.
[84] E. Olkhov-Mitsel and B. Bapat, "Strategies for discovery and validation of methylated and hydroxymethylated DNA biomarkers," *Cancer Med.*, vol. 1, no. 2, pp. 237–260, 2012.
[85] T. O. Tollefsbol, "Epigenetics," in *Handbook of Epigenetics*, 2011, pp. 1–6.
[86] A. K. Rana, "Crime investigation through DNA methylation analysis: methods and applications in forensics," *Egypt. J. Forensic Sci.*, vol. 8, no. 1, pp. 1–17, 2018.
[87] QIAGEN, "Pyrosequencing Technology and Platform Overview," 2019. [Online]. Available: https://www.qiagen.com/us/service-and-support/learning-hub/technologies-and-research-topics/pyrosequencing-resource-center/technology-overview/. [Accessed: 16-Jul-2019].
[88] E. E. Holmes *et al.*, "Performance evaluation of kits for bisulfite-conversion of DNA from tissues, cell lines, FFPE tissues, aspirates, lavages, effusions, plasma, serum, and urine," *PLoS One*, vol. 9, no. 4, 2014.
[89] J. M. S. Bartlett and D. Stirling, "A Short History of the Polymerase Chain Reaction BT - PCR Protocols," J. M. S. Bartlett and D. Stirling, Eds. Totowa, NJ: Humana Press, 2003, pp. 3–6.
[90] J. Sambrook, *Molecular cloning : a laboratory manual*. Third edition. Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory Press, [2001] ©2001.
[91] T. K. Wojdacz and A. Dobrovic, "Methylation-sensitive high resolution melting (MS-HRM): A new approach for sensitive and high-throughput assessment of methylation," *Nucleic Acids Res.*, vol. 35, no. 6, 2007.
[92] ThermoFisher, "Overview of ELISA." [Online]. Available: https://www.thermofisher.com/pt/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-

methods/overview-elisa.html. [Accessed: 18-Jul-2019].
[93] M. Weber *et al.*, "Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells," *Nat. Genet.*, 2005.
[94] F. Mohn, M. Weber, D. Schübeler, and T.-C. Roloff, "Methylated DNA Immunoprecipitation (MeDIP)," 2008.
[95] C. F. Bassil, Z. Huang, and S. K. Murphy, "Bisulfite pyrosequencing," *Methods Mol. Biol.*, 2013.
[96] R. Shaknovich, M. E. Figueroa, and A. Melnick, "HELP (HpaII tiny fragment enrichment by ligation-mediated pcr) assay for dna methylation profiling of primary normal and malignant b lymphocytes," *Methods Mol. Biol.*, 2010.
[97] M. Suzuki and J. M. Greally, "DNA methylation profiling using HpaII tiny fragment enrichment by ligation-mediated PCR (HELP)," *Methods*. 2010.
[98] P. R. Mueller, B. Wold, and P. A. Garrity, "Ligation-Mediated PCR for Genomic Sequencing and Footprinting," in *Current Protocols in Molecular Biology*, 2004.
[99] C. J. Heuck *et al.*, "Myeloma Is Characterized by Stage-Specific Alterations in DNA Methylation That Occur Early during Myelomagenesis," *J. Immunol.*, 2013.
[100] M. Karimi *et al.*, "LUMA (LUminometric Methylation Assay)-A high throughput method to the analysis of genomic DNA methylation," *Exp. Cell Res.*, 2006.
[101] Z. Xiong and P. W. Laird, "COBRA: A sensitive and quantitative DNA methylation assay," *Nucleic Acids Res.*, vol. 25, no. 12, pp. 2532–2534, 1997.
[102] I. Illumina, "Infinium MethylationEPIC Kit." [Online]. Available: https://www.illumina.com/products/by-type/microarray-kits/infinium-methylation-epic.html. [Accessed: 02-Aug-2019].
[103] Illumina, "Infinium HD Methylation Assay Protocol Guide (15019519)," no. November, 2015.
[104] I. Illumina, "Understanding Epigenetic Changes." [Online]. Available: https://www.illumina.com/science/technology/beadarray-technology/infinium-methylation-assay.html. [Accessed: 02-Aug-2019].
[105] L. Song, S. R. James, L. Kazim, and A. R. Karpf, "Specific method for the determination of genomic DNA methylation by liquid chromatography-electrospray ionization tandem mass spectrometry," *Anal. Chem.*, vol. 77, no. 2, pp. 504–510, 2005.
[106] W. Parson, "Age estimation with DNA: From forensic DNA fingerprinting to forensic (Epi)genomics: A mini-review," *Gerontology*, vol. 64, no. 4, pp. 326–332, 2018.
[107] S. Kurdyukov and M. Bullock, "DNA methylation analysis: Choosing the right method," *Biology (Basel).*, vol. 5, no. 1, pp. 1–21, 2016.
[108] H. Okuizumi, T. Takamiya, Y. Okazaki, and Y. Hayashizaki, "Restriction landmark genome scanning," *Methods Mol. Biol.*, vol. 791, pp. 101–112, 2011.
[109] M. R. H. Estécio, P. S. Yan, T. H. M. Huang, and J. P. J. Issa, "Methylated CpG island amplification and microarray (MCAM) for high-throughput analysis of DNA methylation," *Cold Spring Harb. Protoc.*, vol. 3, no. 3, pp. 1–8, 2008.
[110] Active Motif, "MeDIP Catalog No. 55009MeDIP," 2011. [Online]. Available: https://www.activemotif.com/documents/1754.pdf. [Accessed: 18-Sep-2019].
[111] T. Rauch and G. P. Pfeifer, "Methylated-CpG island recovery assay: A new technique for the rapid detection of methylated-CpG islands in cancer," *Lab. Investig.*, vol. 85, no. 9, pp. 1172–1180, 2005.

[112] A. Adey and J. Shendure, “Ultra-low-input, tagmentation-based whole-genome bisulfite sequencing,” *Genome Res.*, vol. 22, no. 6, pp. 1139–1143, 2012.
[113] W. Liu, D. I. Smith, K. J. Rechtzigel, S. N. Thibodeau, and C. D. James, “Denaturing high performance liquid chromatography (DHPLC) used in the detection of germline and somatic mutations,” *Nucleic Acids Res.*, vol. 26, no. 6, pp. 1396–1400, 1998.
[114] I. Hatada *et al.*, “A microarray-based method for detecting methylated loci,” *J. Hum. Genet.*, vol. 47, no. 8, pp. 448–451, 2002.
[115] M. Fukasawa *et al.*, “Microarray analysis of promoter methylation in lung cancers,” *J. Hum. Genet.*, vol. 51, no. 4, pp. 368–374, 2006.
[116] C. Ladd-Acosta, M. J. Aryee, J. M. Ordway, and A. P. Feinberg, “Comprehensive High-Throughput Arrays for Relative Methylation (CHARM),” *Curr. Protoc. Hum. Genet.*, vol. 65, no. 1, pp. 20.1.1-20.1.19, Apr. 2010.
[117] A. E. K. Ibrahim *et al.*, “MMASS: An optimized array-based method for assessing CpG island methylation,” *Nucleic Acids Res.*, vol. 34, no. 20, 2006.
[118] P. S. Yan, D. Potter, D. E. Deatherage, T. H.-M. Huang, and S. Lin, “Differential Methylation Hybridization: Profiling DNA Methylation with a High-Density CpG Island Microarray BT - DNA Methylation: Methods and Protocols,” J. Tost, Ed. Totowa, NJ: Humana Press, 2009, pp. 89–106.
[119] T. A. Rauch and G. P. Pfeifer, “The MIRA Method for DNA Methylation Analysis,” *DNA Methylation Methods Protoc.*, vol. 507, no. 3, pp. 65–75, 2009.
[120] J. Reinders, C. D. Vivier, G. Theiler, D. Chollet, P. Descombes, and J. Paszkowski, “Genome-wide, high-resolution DNA methylation profiling using bisulfite-mediated cytosine conversion,” *Genome Res.*, vol. 18, no. 3, pp. 469–476, 2008.
[121] A. Fernandez-fernandez and M. Esteller, “Epigenome NoE - Protocol : DNA Methylation Analysis by Bisulfite Sequencing (BS) of genomic DNA with sodium bisulfite Method,” pp. 1–7, 2007.
[122] D. Diep, N. Plongthongkum, A. Gore, H. L. Fung, R. Shoemaker, and K. Zhang, “Library-free methylation sequencing with bisulfite padlock probes,” *Nat. Methods*, vol. 9, no. 3, pp. 270–272, 2012.
[123] J. Tost and I. G. Gut, “DNA methylation analysis by pyrosequencing,” *Nat. Protoc.*, vol. 2, no. 9, pp. 2265–2275, 2007.
[124] Epigentek Group Inc., “Methylamp ™ MS-qPCR Fast Kit Complete Solutions for Epigenetics.” pp. 1–6, 2019.
[125] L. S. Kristensen, T. Mikeska, M. Krypuy, and A. Dobrovic, “Sensitive melting analysis after real time-methylation specific PCR (SMART-MSP): High-throughput and probe-free quantitative DNA methylation detection,” *Nucleic Acids Res.*, vol. 36, no. 7, 2008.
[126] C. Bonanno, E. Shehi, D. Adlerstein, and G. M. Makrigiorgos, “MS-FLAG, a novel real-time signal generation method for methylation-specific PCR,” *Clin. Chem.*, vol. 53, no. 12, pp. 2119–2127, 2007.
[127] A. Bilichak and I. Kovalchuk, “The combined bisulfite restriction analysis (COBRA) assay for the analysis of locus-specific changes in methylation patterns,” *Methods Mol. Biol.*, vol. 1456, pp. 63–71, 2017.
[128] A. Sulewska *et al.*, “Detection of DNA methylation in eucaryotic cells,” *Folia Histochem. Cytobiol.*, vol. 45, no. 4, pp. 315–324, 2007.
[129] W. L. Perry, W. M. Bass, W. S. Riggsby, and K. Sirotkin, “The Autodegradation of Deoxyribonucleic Acid (DNA) in Human Rib Bone and Its Relationship to the Time

Interval Since Death," *J. Forensic Sci.*, 2015.
[130] B. Madea, "Is there recent progress in the estimation of the postmortem interval by means of thanatochemistry?," *Forensic Sci. Int.*, vol. 151, no. 2–3, pp. 139–149, 2005.
[131] S. C. Boy, H. Bernitz, and W. F. P. Van Heerden, "Flow cytometric evaluation of postmortem pulp DNA degradation," *Am. J. Forensic Med. Pathol.*, 2003.
[132] J. Vandesompele *et al.*, "Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.," *Genome Biol.*, 2002.
[133] J. Mill *et al.*, "Epigenomic Profiling Reveals DNA-Methylation Changes Associated with Major Psychosis," *Am. J. Hum. Genet.*, 2008.
[134] H. M. Abdolmaleky *et al.*, "Hypomethylation of MB-COMT promoter is a major risk factor for schizophrenia and bipolar disorder," *Hum. Mol. Genet.*, vol. 15, no. 21, pp. 3132–3145, 2006.
[135] D. R. Grayson *et al.*, "Reelin promoter hypermethylation in schizophrenia," *Proc. Natl. Acad. Sci.*, vol. 102, no. 26, pp. 9341–9346, 2005.
[136] S. Sabunciyan *et al.*, "Genome-wide DNA methylation scan in major depressive disorder.," *PLoS One*, vol. 7, no. 4, pp. 1–9, 2012.
[137] J. M. Morahan, B. Yu, R. J. Trent, and R. Pamphlett, "A genome-wide analysis of brain DNA methylation identifies new candidate genes for sporadic amyotrophic lateral sclerosis," *Amyotroph. Lateral Scler.*, vol. 10, no. 5–6, pp. 418–429, 2009.
[138] M. A. Bind *et al.*, "Effects of temperature and relative humidity on DNA methylation," *Epidemiology*, 2014.
[139] M. Barrachina and I. Ferrer, "DNA methylation of Alzheimer disease and tauopathy-related genes in postmortem brain," *J. Neuropathol. Exp. Neurol.*, vol. 68, no. 8, pp. 880–891, 2009.
[140] C. M. Monoranu *et al.*, "Methyl- and acetyltransferases are stable epigenetic markers postmortem," *Cell Tissue Bank.*, vol. 12, no. 4, pp. 289–297, 2011.
[141] C. Ernst, P. O. McGowan, V. Deleva, M. J. Meaney, M. Szyf, and G. Turecki, "The effects of pH on DNA methylation state: In vitro and post-mortem brain studies," *J. Neurosci. Methods*, vol. 174, no. 1, pp. 123–125, 2008.
[142] Y. Li, X. Pan, M. L. Roberts, P. Liu, and A. Theodore, "Stability of global methylation profiles of whole blood and extracted DNA under different storage durations and conditions," vol. 10, pp. 797–811, 2018.
[143] A. Bulla, B. De Witt, W. Ammerlaan, F. Betsou, and P. Lescuyer, "Blood DNA Yield but Not Integrity or Methylation Is Impacted after Long-Term Storage," *Biopreserv. Biobank.*, vol. 14, no. 1, pp. 29–38, 2016.
[144] GeneCards, "GDF15 Gene (Protein Coding)," 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=GDF15. [Accessed: 26-Sep-2019].
[145] GeneCards, "MLH1 Gene (Protein Coding)," 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=MLH1&keywords=MLH1. [Accessed: 26-Sep-2019].
[146] GeneCards, "CDKN1A Gene (Protein Coding)," 2019. [Online]. Available: https://www.genecards.org/cgi-bin/carddisp.pl?gene=CDKN1A&keywords=CDKN1A. [Accessed: 26-Sep-2019].
[147] M. V. Hollegaard, J. Grauholm, B. Nørgaard-Pedersen, and D. M. Hougaard, "DNA methylome profiling using neonatal dried blood spot samples: A proof-of-principle

study," *Mol. Genet. Metab.*, vol. 108, no. 4, pp. 225–231, 2013.
[148] K.-S. Na *et al.*, "Brain-derived neurotrophic factor promoter methylation and cortical thickness in recurrent major depressive disorder," *Sci. Rep.*, vol. 6, p. 21089, Feb. 2016.
[149] V. Januar, M.-L. Ancelin, K. Ritchie, R. Saffery, and J. Ryan, "BDNF promoter methylation and genetic variation in late-life depression," *Transl. Psychiatry*, vol. 5, p. e619, Aug. 2015.
[150] H. J. Kang *et al.*, "A longitudinal study of BDNF promoter methylation and depression in breast cancer," *Psychiatry Investig.*, 2015.
[151] Y. C. Chagnon, O. Potvin, C. Hudon, and M. Préville, "DNA methylation and single nucleotide variants in the brain-derived neurotrophic factor (BDNF) and oxytocin receptor (OXTR) genes are associated with anxiety/depression in older women ," *Frontiers in Genetics* , vol. 6. p. 230, 2015.
[152] A. M. Devlin, U. Brain, J. Austin, and T. F. Oberlander, "Prenatal Exposure to Maternal Depressed Mood and the MTHFR C677T Variant Affect SLC6A4 Methylation in Infants at Birth," *PLoS One*, vol. 5, no. 8, p. e12201, Aug. 2010.
[153] J. Zhao, J. Goldberg, J. D. Bremner, and V. Vaccarino, "Association between promoter methylation of serotonin transporter gene and depressive symptoms: A monozygotic twin study," *Psychosom. Med.*, 2013.
[154] H. Sugawara *et al.*, "Hypermethylation of serotonin transporter gene in bipolar disorder detected by epigenome analysis of discordant monozygotic twins," *Transl. Psychiatry*, vol. 1, p. e24, Jul. 2011.
[155] K. Domschke *et al.*, "Serotonin transporter gene hypomethylation predicts impaired antidepressant treatment response," *Int. J. Neuropsychopharmacol.*, 2014.
[156] S. Okada *et al.*, "The potential of SLC6A4 gene methylation analysis for the diagnosis and treatment of major depression," *J. Psychiatr. Res.*, 2014.
[157] P. A. Melas *et al.*, "Genetic and epigenetic associations of MAOA and NR3C1 with depression and childhood adversities," *Int. J. Neuropsychopharmacol.*, 2013.
[158] E. Conradt, B. M. Lester, A. A. Appleton, D. A. Armstrong, and C. J. Marsit, "The roles of DNA methylation of NR3C1 and 11β-HSD2 and exposure to maternal mood disorder in utero on newborn neurobehavior," *Epigenetics*, vol. 8, no. 12, pp. 1321–1329, 2013.
[159] K.-S. Na *et al.*, "Association between Glucocorticoid Receptor Methylation and Hippocampal Subfields in Major Depressive Disorder," *PLoS One*, vol. 9, no. 1, p. e85425, Jan. 2014.
[160] M. Hirst and M. A. Marra, "The International Journal of Biochemistry Epigenetics and human disease," *Int. J. Biochem.*, vol. 41, pp. 136–146, 2009.
[161] S. Fabris *et al.*, "Biological and clinical relevance of quantitative global methylation of repetitive DNA sequences in chronic lymphocytic leukemia," *Epigenetics*, vol. 6, no. 2, pp. 188–194, 2011.
[162] J. Quilez *et al.*, "Polymorphic tandem repeats within gene promoters act as modifiers of gene expression and DNA methylation in humans," *Nucleic Acids Res.*, vol. 44, no. 8, pp. 3750–3762, 2016.
[163] S. Bocklandt *et al.*, "Epigenetic Predictor of Age," *PLoS One*, vol. 6, no. 6, 2011.
[164] B. Angers, E. Castonguay, and R. Massicotte, "Environmentally induced phenotypes and DNA methylation: How to deal with unpredictable conditions until the next generation and after," *Mol. Ecol.*, vol. 19, no. 7, pp. 1283–1295, 2010.
[165] Federal Bureau of Investigation, "QUALITY ASSURANCE STANDARDS FOR

FORENSIC DNA TESTING LABORATORIES,” 2009.
[166] T. De Baere, B. Magnusson, G. O’donnell, W. Dmitruk, and D. Meuwly, “Guidelines for the single laboratory Validation of Instrumental and Human Based Methods in Forensic Science,” no. 001, pp. 1–31, 2014.