Methylation profiling and validation of candidate tDMRs for identification of human blood, saliva, semen and vaginal fluid and its application in forensics

Natalie Naidoo
211535425

School of Life Sciences, Faculty of Science and Agriculture,
University of KwaZulu-Natal

School of Life Sciences
University of KwaZulu-Natal Westville
Private Bag X 54001
Durban
4000
South Africa

Tel: 031 260 3192 / 2029 **Email:** 211535425@stu.ukzn.ac.za

Master of Science - Genetics

Supervisor: Dr Meenu Ghai

(UKZN - Westville Campus)

Methylation profiling and validation of candidate tDMRs for identification of human blood, saliva, semen and vaginal fluid and its application in forensics

BY

NATALIE NAIDOO

2017

Submitted in fulfillment of the Academic Requirements for the degree of Master of Science (MSc) in the Discipline of Genetics, School of Life Sciences, College of Agriculture, Engineering and Science, University of KwaZulu-Natal (Westville Campus), Durban.

As the candidate's supervisor, I have approved this dissertation for submission.

Supervisor: Dr. Meenu Ghai

Signed: _____ Name: _____ Date: _____

PREFACE

The experimental work described in this dissertation was carried out in the School of Life Sciences, Discipline of Genetics, University of KwaZulu-Natal (Westville Campus), Durban, South Africa from February 2016 to December 2017, under the supervision of Dr. M. Ghai.

These studies represent original work performed by the author and have not otherwise been submitted in any form for any degree or diploma to any tertiary institution. Where use has been made of work of others, it is duly acknowledged in the text.

Supervisor: Dr. M. Ghai Signature: _____

COLLEGE OF AGRICULTURE, ENGINEERING AND SCIENCE

DECLARATION 1 – PLAGIARISM

I, ……... declare that

1. The research reported in this dissertation, except where otherwise indicated, is my original research.

2. This dissertation has not been submitted for any degree or examination at any other university.

3. This dissertation does not contain other person's data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

4. This dissertation does not contain other person's writing, unless specifically acknowledged as being sourced from other researchers. Where other research sources have been quoted, then:
    a) Their words have been rewritten but the general information attributed to them has been referenced;
    b) Where their exact words have been used, then their writing has been placed in italics inside quotation marks and referenced.

5. This dissertation does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed:

…….

Table of Contents

Contents	Page
Acknowledgements	i
Abbreviations	ii
Abstract	iii
Experimental Design	v
List of Figures	vi
List of Tables	xv
Chapter 1: Introduction	1
Chapter 2: Literature Review	4
2.1 Epigenetics	5
2.2 DNA methylation	5
2.3 Role of DNA methylation in gene regulation	7
2.4 Gene body methylation and gene expression	8
2.5 Tissue-specific differentially methylated regions (tDMRs)	9
2.6 Identification of tDMRs	9
2.7 Genetic factors that influence DNA methylation	11
2.8 Environmental factors that influence DNA methylation	11
2.9 DNA methylation and its application in forensic science	12
2.9.1 Identification of artificial DNA	12
2.9.2 Parent of origin	13
2.9.3 Age estimation	14
2.9.4 Differentiation between monozygotic/dizygotic twins	16
2.9.5 Human body fluid identification	17
2.10 Stability of DNA methylation markers	21
2.11 Techniques used for DNA methylation analysis	22
2.11.1 Affinity enrichment	22
2.11.1.1 Methyl-CpG immunoprecipitation (MCIp)	23
2.11.1.2 Methyl-DNA Immunoprecipitation (MeDIP)	23
2.11.2 Endonuclease digestion	24

2.11.2.1 Methylation sensitive restriction enzyme 25
2.11.2.2 Restriction Landmark Genomic Scanning (RLGS) 25
2.11.3 Chemical Modification 26
2.11.3.1 Methylation-sensitive single nucleotide primer extension (Ms-SNuPE) 26
2.11.3.2 Bisulfite sequencing (BS) 27
2.11.3.3 Methylation Specific Polymerase Chain Reaction (MSP) 28
2.12 Rationale of the study 32
2.13 Genes of interest studied in the present research 33
2.13.1 *ZNF282* gene (*Zinc finger protein 282*) 33
2.13.2 *PTPRS* gene (*Protein tyrosine phosphatase, receptor S*) 34
2.13.3 *HPCAL1* gene (*Hippocalcin like 1*) 35
2.14 Hypothesis, objectives and aims 36
2.14.1 Hypothesis 36
2.14.2 Objectives 36
2.14.3 Aims 36
2.14.4 Key Question 36

Chapter 3: Validation of tissue-specific differential DNA methylation in gene-specific tDMRs by methylation specific PCR and bisulfite sequencing 37
Abstract 38
3.1 Introduction 39
3.2 Materials and methods 47
3.2.1 Ethics approval 47
3.2.2 Sample collection, storage of samples and DNA extraction 47
3.2.3 Bisulfite conversion 47
3.2.4 Selection of candidate genes and primer design 48
3.2.5 MSP reaction 55
3.2.6 BS protocol 55
3.2.6.1 BS PCR reaction 55
3.2.6.2 Cloning and sequencing of BS reaction products 56
3.2.7 Bisulfite sequencing data analysis and DNA methylation profiling 56

3.2.8 Statistical analysis 56
3.3 Results 58
3.3.1 Methylation profile using MSP 58
3.3.1.1 ZNF282 tDMR 58
3.3.1.2 PTPRS 1 tDMR 60
3.3.1.3 PTPRS 2 tDMR 62
3.3.1.4 HPCAL1 tDMR 64
3.3.2 Methylation profiling using BS 68
3.3.2.1 Methylation profiles of body fluid using the ZNF282 tDMR 68
3.3.2.2 Methylation profiles of body fluids using the PTPRS 1 tDMR 70
3.3.2.3 Methylation profiles of body fluid using the PTPRS 2 tDMR 72
3.3.2.4 Methylation profiles of body fluid using the HPCAL1 tDMR 74
3.4 Discussion 75
3.4.1 Methylation profiling of candidate tDMRs by MSP and BS 76
3.4.1.1 ZNF282 tDMR 77
3.4.1.2 PTPRS 1 and PTPRS 2 tDMRs 77
3.4.1.3 HPCAL1 tDMR 78
3.5. Conclusions 79

Chapter 4: Analysis of the sensitivity and stability of methylation profiles of tDMR markers for human body fluid identification 80
Abstract 81
4.1 Introduction 82
4.2 Materials and methods 85
4.2.1 Ethics approval 85
4.2.2 Sample collection, storage of samples and DNA extraction 85
4.2.3 Bisulfite conversion 85
4.2.4 Primer design 86
4.2.5 MSP sensitivity 86
4.2.5.1 MSP PCR reaction 86
4.2.6 Forensic simulation study 87
4.2.7 BS protocol 88
4.2.7.1 BS PCR reaction 88

4.2.7.2 Cloning and sequencing of BS reaction products 88
4.2.8 Bisulfite sequencing data analysis and DNA methylation profiling 89
4.2.9 Statistical analysis 89
4.3 Results 91
4.3.1 Sensitivity test 91
4.3.2. Forensic simulation study 92
4.3.2.1 Methylation profile of body fluids using the ZNF282 tDMR analysed immediately after collection (t0 days) 92
4.3.2.2 DNA recovery after exposure to environmental conditions for t50 days 94
4.3.2.3 Effects of simulated forensic conditions on methylation profile of ZNF282 tDMR after t50 days 94
4.3.2.3.1 Methylation profile ZNF282 tDMR after placing all fluids outside on the ground (condition C) 94
4.3.2.3.2 Methylation profile of ZNF282 tDMR upon exposure to all conditions in vaginal fluid only 96
4.3.2.3.3 Methylation profiles of ZNF282 tDMR before (t0 days) and after exposure to forensic simulated conditions (t50 days) 99
4.4 Discussion 102
4.4.1 Sensitivity study 102
4.4.2 Forensic simulation 103
4.4.2.1 DNA recovery 103
4.4.2.2 Methylation profile of ZNF282 tDMR marker analysed immediately after collection (t0 days) for all body fluids 104
4.4.2.3 Methylation profile of ZNF282 tDMR analysed after t50 days of exposure to simulated forensic conditions 104
4.5 Conclusion 106

Chapter 5: General discussion and Conclusion 107
5.1 Purpose of study 108
5.2 Summary of findings 110
5.3 Conclusion and future work 113

References **114**
Appendix A **145**
Appendix B **150**
Appendix C **151**
Appendix D **157**
Appendix E **159**
Appendix F **160**
Appendix G **165**
Appendix H **166**
Appendix I **169**
Appendix J **174**
Appendix K **176**
Appendix L **178**

ACKNOWLEDGEMENTS

I sincerely would like to acknowledge the following people for the successful completion of this project:

My heartfelt thanks goes to my supervisor, Dr Meenu Ghai. Thank you for your wisdom, patience, guidance, motivation and support during the course of this research. I am grateful and appreciate that you have always encouraged me to grow and do better.

To Dr. A. J. Niehaus. Thank you for all your assistance and contribution to this study.

To Dr Oliver Zishiri and Dr Matthew Adeleke, for their constant enthusiasm and words of motivation.

I am grateful to the National Research Foundation for financial support.

Thanks to all the volunteers who participated in the present study.

I am immensely grateful to my parents Cyril and Roshni Naidoo. Thank you for your patience, support and blessings throughout the study.

To my significant other, Rishan Govender. Thank you for your continuous support, patience and encouragement.

To Farzeen Kader, for her assistance, constant motivation and support.

To my colleagues of the Genetics Department for their encouragement throughout the duration of this study.

Abbreviations

BS	Bisulfite sequencing
CODIS	Combined DNA Index System
DNMT	DNA methyltransferases
HPCAL1	Hippocalcin like 1
MCIp	Methyl-CpG immunoprecipitation
MeDIP	Methyl-DNA Immunoprecipitation
MSP	Methylation specific PCR
MS-RDA	Methylation sensitive represent difference analysis
MSRE-PCR	Methylation sensitive restriction enzyme PCR
Ms-SNuPE	Methylation-sensitive single nucleotide primer extension
PCR	Polymerase chain reaction
PTPRS	Protein tyrosine phosphatase, receptor S
RLGS	Restriction landmark genomic scanning
SNP	Single nucleotide polymorphism
STR	Short tandem repeat
tDMRs	Tissue specific differentially methylated regions
ZNF282	Zinc finger protein 282

Abstract

Identification of body fluids and tissues is an essential step in forensic investigation because it can be used as strong evidence in identifying suspects and victims. Currently in forensic investigations, catalytic, enzymatic and immunological techniques are used to identify body fluids, however, are limited due to lack of sensitivity and specificity. Hence, researchers are always on the lookout for novel methods that can be used to identify and analyse body fluids. Recently, DNA methylation-based markers have proven to be more sensitive and specific than conventional methods for body fluid identification. Genome-wide methylation studies have demonstrated that tissue specific differentially methylated regions (tDMRs) vary in methylation profiles in various cell types and tissues. The differences in methylation profiles of tDMRs can be targeted to be used as biomarkers to differentiate between body fluids and tissues. To date, only a few DNA methylation-based markers have been reported to identify body fluids. To enhance the specificity and robustness of DNA methylation-based identification, novel markers are required. Additionally, methylation-based markers require further interrogation, to evaluate the stability of their methylation profiles under simulated forensics conditions such as UV light, temperature, rain and microbes, which could cause DNA degradation and affect DNA recovery as well as the methylation status of body fluids. In a previous study, based on differential gene expression in blood, saliva, semen and vaginal fluid, gene body CpG islands were selected, in genes Zinc finger protein 282 (*ZNF282*), Protein tyrosine phosphatase, receptor S (*PTPRS*) and Hippocalcin like 1 (*HPCAL1*), that have potential tDMRs to differentiate between, blood, saliva, semen and vaginal fluid. It was proposed that differential gene expression could be possibly due to differences in methylation patterns. The present study was undertaken to establish the methylation status of potential tDMRs in target body fluids by using methylation specific PCR (MSP) and bisulfite sequencing (BS). In both MSP and BS, the methylation status of 3 genes *ZNF282*, *PTPRS* and *HPCAL1* were analysed in 10 samples of each body fluid. With MSP analysis the ZNF282 and PTPRS1 tDMR displayed semen-specific hypomethylation while HPCAL1 tDMR showed saliva-specific hypomethylation. The PTPRS 2 tDMR did not differentiate between any body fluids due to presence of methylation and unmethylation for all body fluids. With quantitative analysis by BS the ZNF282 tDMR showed statistically significant difference in overall methylation status between semen and all other body fluids as well as at individual CpG sites ($p < 0.05$). Therefore, ZNF282 tDMR has the potential to be used to be a semen-specific hypomethylated marker. However, no statistically significant difference in methylation profiles was observed for PTPRS 1 and PTPRS 2 tDMR between body fluids or at individual CpG sites

($p > 0.05$). The BS study showed that the tDMR for the *HPCAL1* gene displayed non-specific amplification therefore was not further analysed. Furthermore, a sensitivity and forensic simulation study was conducted to determine the stability of methylation profiles. To determine the lowest DNA concentration that can be evaluated with MSP, a sensitivity study was conducted using five-fold serial dilution (25, 20, 15, 10, 5, 1 ng) of blood DNA samples. Each DNA dilution was subjected to bisulfite modification, followed by amplification with ZNF282, PTPRS 1, PTPRS 2, and HPCAL1 primers. The results showed that the detection limits were 10 ng for ZNF282 tDMR, 5 ng for PTPRS 1, 15 ng for PTPRS 2, and 5 ng for HPCAL1 tDMR. Thus, it was concluded that a DNA concentration greater than 10 ng would yield successful results with MSP analyses. To evaluate whether environmental conditions has an effect on the stability of methylation profiles of the ZNF282 tDMR, five samples of each body fluid were subjected to five different forensic simulated conditions (dry at room temperature, wet in an exsiccator, outside on the ground, sprayed with alcohol and sprayed with bleach) for 50 days. Following the 50 days, vaginal fluid showed highest DNA recovery under all conditions while semen had least DNA quantity. Under outside on the ground condition, all body fluids except semen showed decrease in methylation level, however, significant decrease in methylation level was observed for saliva. A statistical significant difference was observed for saliva and semen ($p < 0.05$) in the outside on the ground condition. No differences in methylation level were observed for the ZNF282 tDMR under all conditions for vaginal fluid samples. Thus, ZNF282 tDMR is stable under environmental insults and can be used as reliable semen-specific hypomethylated marker. The analysis of tDMRs represents a unique, efficient and reliable technique that can be used to differentiate between human body fluids. In the future, identification and validation of new tDMRs based markers as well as determining methylation differences in other forensically relevant body fluids will be beneficial for forensics applications.

Experimental Design

In order to achieve the stated objectives, the present research was divided into the relevant chapters described below. A precise description of each phase will be described in the chapters which will include all rationale, challenges experienced and deviations from reported methods.

Chapter One

This chapter entails an introduction to the study.

Chapter Two

This chapter entails a detailed explanation on the various aspects of this study.

Chapter Three

This chapter involved the determination of methylation profile of candidate tDMRs in blood, saliva, semen and vaginal fluid using methylation specific PCR (MSP) and Bisulfite sequencing (BS). Specific primers targeting the gene body CpG islands were designed for methylation profiling.

Chapter Four

This chapter investigated the sensitivity of MSP method and stability of methylation profile of tDMRs in blood, saliva, semen and vaginal fluid under simulated forensic conditions.

Chapter Five

The final chapter, includes the general discussion and conclusion, which provides an overview of the main objectives and findings described in each of the chapters of the dissertation. Possible limitations are acknowledged and a scope for future improvement and development is provided.

List of Figures

Figure 2.1: Basic overview of DNA methylation. A methyl group (CH_3) attaches covalently to a cytosine residue (Vidaki *et al*., 2013). **6**

Figure 2.2: Regulation of gene expression by DNA methylation. (A) Absence of methyl groups in the CpG island promoter permits binding of transcription factors which facilitates transcription initiation (B) Presence of methyl groups hinders binding of transcription factors which prevents gene expression (Lim and Maher, 2010). **8**

Figure 2.3: Diagrammatic representation of immunoprecipitation techniques. Genomic DNA is first fragmented by sonication and then undergoes denaturation. The DNA is then separated into two fractions, namely the input DNA and immunoprecipitated DNA(IP). Primary antibody (anti-methylcytosine) and methyl-CpG binding protein (MBD) fusion protein are incubated with single-stranded DNA in MeDIP and MCIp, respectively. The methylated DNA–antibody complex is captured by a secondary antibody that is specific to the primary antibody (Thu *et al*., 2010). **24**

Figure 2.4: Schematic representation of bisulfite conversion. In methylated DNA, non-methylated cytosine is converted to uracil (blue U) and methylated cytosine remain the same (underlined CG) after bisulfite treatment (Hernández *et al*., 2013). **26**

Figure 2.5: A diagrammatic representation of Ms-SNuPE. Target DNA undergoes bisulfite treatment and amplification, which is followed by the SNuPE assay. SNuPE products are then analysed by electrophoresis. Methylated cytosine (represented by red-mC) and non-methylated cytosine (represented by green C) at a CpG site. Non- methylated cytosine is reproduced as T (blue) following PCR amplification. Arrows at the 3- or 5'- represent PCR primers that amplify the bisulfite treated target DNA. Arrows at S1, S2 or S3 are representative the oligonucleotide primers for the Ms-SNuPE analysis (Gonzalgo and Liang, 2007). **27**

Figure 2.6: A diagrammatic representation of cloned based bisulfite sequencing. Genomic DNA is bisulfite treated, followed by PCR amplification. PCR products are ligated into a vector. The ligated DNA product is then transformed into competent bacterial cells and grown in non-selective liquid culture. A selective media is used to plate transformed cells which are grown overnight. Blue/white selection method is used for selection of (white) colonies which contain the PCR amplicons as inserts. A PCR reaction followed by agarose gel electrophoresis is used to determine if selected positive colonies have the correct insert size. Positive amplicons are purified and sequenced. Sequenced results are used to establish methylation status. The rows represent a single clone and the columns individual CpG dinucleotides within the sequenced region of interest. Black dots represent methylated CpG sites and white dots represent unmethylated CpG sites (Huang *et al*., 2013a). **29**

Figure 2.7: MSP primer design for methylated and unmethylated DNA. The DNA sequence (left) is modified based on the methylation status. The modified sequence is used to design the forward primer by replacing Us with Ts (Hernández *et al.,* 2013). **30**

Figure 2.8: A diagrammatic representation of methylation specific polymerase chain reaction. In MSP, two sets of primer pairs are designed to target the methylated cytosine (black circles) and non-methylated cytosine (white circles) of the bisulfite converted genomic DNA. After PCR amplification, the amplicons are analysed by agarose gel electrophoresis. M represents methylated, U represents non-methylated and L represents the DNA ladder (Huang *et al*., 2013b). **30**

Figure 2.9: Chromosomal location of the *ZNF282* gene. **33**

Figure 2.10: Chromosomal location of the *PTPRS* gene. **34**

Figure 2.11: Chromosomal location of the *HPCAL1* gene. **35**

Figure 3.1: Target CpG island of *ZNF282* gene for primer design a) MSP: Visual representation of the location of the primers designed within the CpG island (blue colored area). Two primer sets are indicated by the purple boxes (MF1- methylated forward primer and MR1- methylated reverse primer) and green boxes (UF1- unmethylated forward primer and UR1- unmethylated reverse primer). Horizontal red line indicates the input sequence and vertical red lines represent CpG sites. b) BS: Visual representation of the location of the primers designed out of the CpG island. Primers are indicated by red boxes (F1- forward primer and R1- reverse primer). c) UCSC genome browser view of chromosomal location of the target CpG island within the gene indicated by the green bar and the position of the CpG island in relation to the introns and exons indicated by the blue horizontal line and box.

49

Figure 3.2: Target CpG island of *PTPRS* (PTPRS 1) gene for primer design a) MSP: Visual representation of the location of the primers designed within the CpG island (blue colored area). Two primer sets are indicated by the purple boxes (MF1- methylated forward primer and MR1- methylated reverse primer) and green boxes (UF1- unmethylated forward primer and UR1- unmethylated reverse primer). Horizontal red line indicates the input sequence and vertical red lines represent CpG sites. b) BS: Visual representation of the location of the primers designed out of the CpG island. Primers are indicated by red boxes (F1- forward primer and R1- reverse primer). c) UCSC genome browser view of chromosomal location of the target CpG island within the gene indicated by the green bar and the position of the CpG island in relation to the introns and exons indicated by the blue horizontal line and box

50

Figure 3.3: Target CpG island of *PTPRS* (PTPRS 2) gene for primer design a) MSP: Visual representation of the location of the primers designed within the CpG island (blue colored area). Two primer sets are indicated by the purple boxes (MF1- methylated forward primer and MR1- methylated reverse primer) and green boxes (UF1- unmethylated forward primer and UR1- unmethylated reverse primer). Horizontal red line indicates the input sequence and vertical red lines represent CpG sites. b) BS: Visual representation of the location of the primers designed. Primers are indicated by red boxes (F1- forward primer and R1- reverse primer). c) UCSC genome browser view of chromosomal location of the target CpG island within the gene indicated by the green bar and the position of the CpG island in relation to the introns and exons indicated by the blue horizontal line and box.

51

Figure 3.4: Target CpG island of *HPCAL1* gene for primer design a) MSP: Visual representation of the location of the primers designed within the CpG island (blue colored area). Two primer sets are indicated by the purple boxes (MF1- methylated forward primer and MR1- methylated reverse primer) and green boxes (UF1- unmethylated forward primer and UR1- unmethylated reverse primer). Horizontal red line indicates the input sequence and vertical red lines represent CpG sites. b) BS: Visual representation of the location of the primers designed. Primers are indicated by red boxes (F1- forward primer and R1- reverse primer). c) UCSC genome browser view of chromosomal location of the target CpG island within the gene indicated by by the green bar and the position of the CpG island in relation to the introns and exons indicated by the blue horizontal line and box. **52**

Figure 3.5: Flowchart showing the overview of the BS protocol. **57**

Figure 3.6: MSP based methylation profile of blood for the ZNF282 tDMR primer set. a) Products with primers specific for methylated cytosine (210 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- blood DNA samples. Lane 12- No template control. Lane 13- Universal Methylated Human DNA Standard (Zymo Research). b) Products with primers specific for unmethylated cytosine (210 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- blood DNA samples. Lane 12- No template control. **58**

Figure 3.7: MSP based methylation profile of vaginal fluid for the ZNF282 tDMR primer set. a) Products with primers specific for methylated cytosine (210 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- vaginal fluid DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (210 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- vaginal fluid DNA samples. Lane 12- No template control. **59**

Figure 3.8: MSP based methylation profile of semen for the ZNF282 tDMR primer set. a) Products with primers specific for methylated cytosine (210 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- semen DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (210 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- semen DNA samples. Lane 12- No template control **59**

Figure 3.9: MSP based methylation profile of saliva for the ZNF282 tDMR primer set. a) Products with primers specific for methylated cytosine (210 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- saliva DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (210 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- saliva DNA samples. Lane 12- No template control.

60

Figure 3.10: MSP based methylation profile of blood for the PTPRS 1 tDMR primer set. a) Products with primers specific for methylated cytosine (216 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- blood DNA samples. Lane 12- No template control. Lane 13- Universal Methylated Human DNA Standard (Zymo Research). b) Products with primers specific for unmethylated cytosine (215 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- blood DNA samples. Lane 12- No template control.

61

Figure 3.11: MSP based methylation profile of vaginal fluid for the PTPRS 1 tDMR primer set. a) Products with primers specific for methylated cytosine (216 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- vaginal fluid DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (215 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- vaginal fluid DNA samples. Lane 12- No template control.

61

Figure 3.12: MSP based methylation profile of semen for the PTPRS 1 tDMR primer set. a) Products with primers specific for methylated cytosine (216 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- semen DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (215 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- semen DNA samples. Lane 12- No template control.

62

Figure 3.13: MSP based methylation profile of saliva for the PTPRS 1 tDMR primer set. a) Products with primers specific for methylated cytosine (216 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- saliva DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (215 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- saliva DNA samples. Lane 12- No template control.

62

Figure 3.14 MSP based methylation profile of blood for the PTPRS 2 tDMR primer set. a) Products with primers specific for methylated cytosine (168 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- blood DNA samples. Lane 12- No template control. Lane 13- Universal Methylated Human DNA Standard (Zymo Research) b) Products with primers specific for unmethylated cytosine (169 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- blood DNA samples. Lane 12- No template control. **63**

Figure 3.15: MSP based methylation profile of vaginal fluid for the PTPRS 2 tDMR primer set. a) Products with primers specific for methylated cytosine (168 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- vaginal fluid DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (169 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- vaginal fluid DNA samples. Lane 12- No template control. **63**

Figure 3.16: MSP based methylation profile of semen for the PTPRS 2 tDMR primer set. a) Products with primers specific for methylated cytosine (168 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- semen DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (169 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- semen DNA samples. Lane 12- No template control. **64**

Figure 3.17: MSP based methylation profile of saliva for the PTPRS 2 tDMR primer set. a) Products with primers specific for methylated cytosine (168 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- saliva DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (169 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- saliva DNA samples. Lane 12- No template control. **64**

Figure 3.18: MSP based methylation profile of blood for the HPCAL1 tDMR primer set. a) Products with primers specific for methylated cytosine (113 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- blood DNA samples. Lane 12- No template control. Lane 13- Universal Methylated Human DNA Standard (Zymo Research). b) Products with primers specific for unmethylated cytosine (114 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- blood DNA samples. Lane 12- No template control **65**

Figure 3.19: MSP based methylation profile of vaginal fluid for the HPCAL1 tDMR primer set. a) Products with primers specific for methylated cytosine (113 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- vaginal fluid DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (114 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- vaginal fluid DNA samples. Lane 12- No template control. **65**

Figure 3.20: MSP based methylation profile of semen for the HPCAL1 tDMR primer set. a) Products with primers specific for methylated cytosine (113 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- semen DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (114 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- semen DNA samples. Lane 12- No template control. **66**

Figure 3.21: MSP based methylation profile of saliva for the HPCAL1 tDMR primer set. a) Products with primers specific for methylated cytosine (113 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- saliva DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (114 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- saliva DNA samples. Lane 12- No template control. **66**

Figure 3.22: Body fluid specific methylation of the ZNF282 tDMR. Each row indicates a single clone of bisulfite PCR products and each column indicates individual CpG site in the region of interest. Different methylation states of the CpG sites are indicated by colours (blue represents unmethylated; red represents methylated; white represents unknown). **69**

Figure 3.23: Body fluid specific methylation of the PTPRS 1 tDMR. Each row indicates a single clone of bisulfite PCR products and each column indicates individual CpG site in the region of interest. Different methylation states of the CpG sites are indicated by colours (blue represents unmethylated; red represents methylated; white represents unknown). **71**

Figure 3.24: Body fluid specific methylation of the PTPRS 2 tDMR. Each row indicates a single clone of bisulfite PCR products and each column indicates individual CpG site in the region of interest. Different methylation states of the CpG sites are indicated by colours (blue represents unmethylated; red represents methylated; white represents unknown).

73

Figure 4.1: Flowchart showing the overview of the BS protocol.

90

Figure 4.2: Evaluation of senstivity of MSP reaction. Products with primers specific for methylated cytosine. a) HPCAL1 tDMR b) PTPRS 1 tDMR c) PTPRS 2 tDMR d) ZNF282 tDMR. Lane 1 (a, b, c, d): 100bp ladder (Thermo scientific), Lane 2 – Lane 7 (a, b, c, d): The five-fold dilutions of blood with DNA concentrations (25, 20, 15, 10, 5, to 1 ng) and Lane 8: No template control.

91

Figure 4.3: Bisulfite sequencing results for the ZNF282 tDMR marker at t0 days. Each row indicates a single clone of bisulfite PCR products and each column indicates individual CpG site in the region of interest. Different methylation states of the CpG sites are indicated by colours (blue represents unmethylated; red represents methylated; white represents unknown). Overall methylation results for all 8 CpG sites were: 100% for blood, 100% for saliva, 0% for semen and 98% for vaginal fluid.

93

Figure 4.4: Bisulfite sequencing results of the ZNF282 tDMR marker for outside on the ground (condition C). Each row indicates a single clone of bisulfite PCR products and each column indicates individual CpG site in the region of interest. Different methylation states of the CpG sites are indicated by colours (blue represents unmethylated; red represents methylated; white represents unknown). Overall methylation results for all 8 CpG sites were: 96% for blood, 72% for saliva, 0% for semen and 94% for vaginal fluid.

96

Figure 4.5: Bisulfite sequencing results for the ZNF282 tDMR in vaginal fluid under all forensic simulation conditions: dry at room temperature (condition A), wet in an exsiccator (condition B) outside on the ground (condition C), sprayed with alcohol (Condition D) and sprayed with bleach (Condition E). Each row indicates a single clone of bisulfite PCR products and each column indicates individual CpG site in the region of interest. Different methylation states of the CpG sites are indicated by colours (blue represents unmethylated; red represents methylated; white represents unknown). Overall methylation results for all 8 CpG sites in vaginal fluid samples were: 81% for condition A, 98% for condition B, 94% for condition C, 98% for condition D and 98% for condition E.

98

List of Tables

Table 3.1: Current methods used for body fluid identification (Gaensslen, 1983; Greenfield and Sloan, 2003; Shaler, 2002; Spalding, 2003; Virkler and Lednev, 2009). **39**

Table 3.2: Methylation markers for identification of blood reported to date. **42**

Table 3.3: Methylation markers for identification of menstrual blood reported to date. **42**

Table 3.4: Methylation markers for identification of saliva reported to date. **43**

Table 3.5: Methylation markers for identification of vaginal fluid reported to date. **43**

Table 3.6: Methylation markers for identification of semen reported to date. **44**

Table 3.7: Genomic information for the three genes with candidate tDMRs characterized in this study. **53**

Table 3.8: BS primers designed for analysis of methylation in candidate tDMRs. **53**

Table 3.9: MSP primers designed for analysis of methylation in candidate tDMRs. **54**

Table 3.10: Universal Methylated Human DNA Standard for assessment of efficiency of bisulfite conversion of DNA. **54**

Table 3.11: Summary of results for all four candidate tDMRs using MSP. **67**

Table 3.12: Overview of bisulfite sequencing results for body fluids analysed using BISMA for the ZNF282 tDMR. **70**

Table 3.13: Pairwise comparison of total DNA methylation for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below. **70**

Table 3.14: Overview of bisulfite sequencing results for body fluids analysed using BISMA for the PTPRS 1 tDMR. **72**

Table 3.15: Pairwise comparison of total DNA methylation for PTPRS 1 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below. **72**

Table 3.16: Overview of bisulfite sequencing results for body fluids analysed using BISMA for the PTPRS 2 tDMR. **74**

Table 3.17: Pairwise comparison of total DNA methylation for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below. **74**

Table 4.1: Composition of human body fluids (Altman and Katz, 1961; Li, 2008; Spinrad, 1994). **83**

Table 4.2: MSP primers for sensitivity analysis. **86**

Table 4.3: BS primer for ZNF282 tDMR for forensic simulation study. **86**

Table 4.4: Sample collection and preparation of human body fluids for forensic validation according to the recommendations by Setzer *et al*. (2008). **87**

Table 4.5: Number of samples subjected to five different environmental conditions for two different time periods. The 20 samples consist of five blood, five saliva, five semen and five vaginal fluid samples. **88**

Table 4.6: Overview of bisulfite sequencing results for body fluids analysed using BISMA for the ZNF282 tDMR at t0 days. 2 **9**

Table 4.7: Pairwise comparison of overall DNA methylation for ZNF282 tDMR marker in blood, saliva, semen and vaginal fluid at t0 days. The *p*-values obtained from the Fisher's Exact test are indicated below. **93**

Table 4.8: Overview of bisulfite sequencing results analysed using BISMA for body fluids exposed to condition C (outside on the ground) for the ZNF282 tDMR marker after t50 days. **95**

Table 4.9: Pairwise comparison of total DNA methylation for ZNF282 tDMR for Condition C in blood, saliva, semen and vaginal fluid after t50 days. The *p*-values obtained from the Fisher's Exact test are indicated below. **95**

Table 4.10: Overview of bisulfite sequencing results for vaginal fluid analysed using BISMA for the ZNF282 tDMR marker under all forensic simulation conditions: dry at room temperature (condition A), wet in an exsiccator (condition B) outside on the ground (condition C), sprayed with alcohol (Condition D) and sprayed with bleach (Condition E). **97**

Table 4.11: Pairwise comparison of total DNA methylation for ZNF282 tDMR for all conditions in vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below. **97**

Table 4.12: Pairwise comparison of total DNA methylation for ZNF282 tDMR marker under Condition C compared between t0 days and t50 days in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below. **100**

Table 4.13: Pairwise comparison of total DNA methylation for ZNF282 tDMR under all forensic simulation conditions compared between t0 days and t50 days in vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below. **100**

Table 4.14: Overall change in DNA methylation level in each body fluid after t50 days of subjection to five forensic simulation conditions (A- dry at room temperature, B- wet in an exsiccator and C- outside on the ground, D- sprayed with alcohol and E- sprayed with bleach). Up or down arrows indicate increase or decrease in DNA methylation level and a dash indicates no change in methylation level. **100**

Table 4.15: Change in DNA methylation at each CpG site in each body fluid after t50 days of subjection to five forensic simulation conditions (A- dry at room temperature, B- wet in an exsiccator and C- outside on the ground, D- sprayed with alcohol and E- sprayed with bleach). Up or down arrows indicate increase or decrease in DNA methylation level and a dash indicates no change in methylation level.

101

Chapter 1

Introduction

Epigenetics involves heritable changes in gene activity and expression, without alteration of the DNA sequence. In the human body every cell has a unique epigenome which is a result of histone modification, chromatin remodelling and DNA methylation, all of which have a role in the regulation of gene expression (Ng and Gurdon, 2008). DNA methylation is a well characterized epigenetic modification which is essential for normal development and plays an important role in genome imprinting and X–chromosome inactivation, among other functions. (Sant *et al*., 2012). DNA methylation is the attachment of a methyl group to 5'-cytosine pyrimidine ring by DNA methyltransferases. In mammals, DNA methylation occurs at CpG sites, however not all cytosine residues are methylated. CpG islands are regions that are rich in CpG sites.

DNA methylation patterns are tissue-specific and there are regions of DNA in the genome that display differential methylation. Regions found in chromosomes which display differences in methylation pattern depending on the tissue or cell type are known as tissue-specific differentially methylated regions (tDMRs) (An *et al.*, 2013; Choi *et al.*, 2014; Lee *et al*., 2012; Rakyan *et al*., 2008). tDMRs are broadly distributed in intragenic and intergenic regions which include both CpG islands, and non-CpG island regions (Deaton and Bird, 2011; Song *et al*., 2009) as well as CpG shores (Irizarry *et al*., 2009). A study by Eckhardt *et al.* (2006) demonstrated that methylation status at tDMRs is stable and specific, therefore can be used as excellent markers for tissue identification.

Every person has a unique DNA profile which can be used to identify individuals that are involved in criminal investigations by analysing body fluids that are found at a crime scene which enables individuals of interest to be questioned or to eliminate suspects (Frumkin *et al*., 2010; Orphanou, 2015). Blood, vaginal fluid, semen and saliva are the most common body fluids found at crime scenes and vary in their appearance. It is possible to distinguish blood from other body fluids, however, differentiation between vaginal fluid, semen and saliva is difficult as they all appear creamy white or colorless (Orphanou, 2015). Identifying the source of biological material that is found at crime scenes provide important information in forensic investigations (Budowle and van Daal, 2009; Fleming and Harbison, 2010; Gaensslen, 1983; Juusola and Ballantyne, 2003). Currently, there are presumptive tests which are used as screening tests and confirmatory tests which are used for absolute identification of body fluids (Frumkin *et al*., 2011; Virkler and Lednev, 2009). These tests are based on catalytic, enzymatic, immunological or visual testing, thus depend on the stability of the target molecule to detect

the presence or activity of the specific protein found in the tissue of interest (Frumkin *et al*., 2011). Although these tests have been useful in forensic analyses, there are many limitations associated, such as low specificity due to cross reactions with other molecular species or tissues, lack of sensitivity, time consuming, use a large amount of sample and proteins are less stable than DNA (Frumkin *et al*., 2011; Haas *et al*., 2009). Differentiating between body fluids using RNA expression differences is possible (Bauer and Patselt, 2002; Juusola and Ballantyne, 2003; Juusola and Ballantyne, 2005; Nussbaumer *et al*., 2006; Setzer *et al*., 2008; Zubakov *et al*., 2008), however RNA molecules are not stable for long periods of time. Identifying DNA methylation differences in human tissues and body fluids is emerging as a new method for body fluid discrimination. DNA is stable and methylation analysis methods are compatible with short tandem repeat (STR) analysis (Lee *et al*., 2016b).

A difference in methylation patterns is a unique characteristic displayed by tDMRs, thus suggesting that tDMRs can be used as candidate biomarkers in forensic research. Many researchers have successfully used tDMRs to differentiate between body fluids and tissues (An *et al.,* 2013; Frumkin *et al.,* 2011; Lee *et al.,* 2012; Madi *et al.,* 2012; Park *et al.,* 2014).

Identification of human body fluids using DNA methylation methods in forensic application is relatively new. Thus, the present study endeavoured to validate potential gene-specific tDMRs that have been previously identified in the human genome, for differentiation of four body fluids (blood, saliva, semen and vaginal fluid). Validation of the candidate tDMRs will enable the tDMRs to be used as novel markers for the identification and differentiation of human body fluids in forensic casework. Sensitivity of the methods used for methylation analysis will also be tested to determine the minimum amount of DNA required for a reliable result. Environmental factors such as climate conditions (sunlight, rain, humidity) and non-climate conditions (diet, lifestyle, growth of microorganisms) may alter DNA methylation patterns. Thus, determining whether the methylation profiles of tDMRs are stable to identify human body fluids when exposed to environmental conditions will also be beneficial.

There are many tDMRs that are present in the mammalian genome, thus, the identification and /or validation of novel body fluid tDMRs is expected to spur the development of DNA methylation markers database for medical and forensic applications.

Chapter 2

Literature Review

2.1 Epigenetics

A British embryologist and geneticist, Conrad Hal Waddington was the first person to introduce the term epigenetics in the early 1940s (Hall, 2012; Slack, 2002). He defined epigenetics as "the branch of biology which studies causal interactions between genes and their products, which bring the phenotype into being" (Waddington, 1942). However, today the definition of epigenetics has developed to: the heritable changes in gene expression and function without altering the DNA sequence (Dupont *et al*., 2009; Galm *et al*., 2006; Wu and Morris, 2001). Epigenetics includes functional modifications such as, histone modification, chromatin structuring, non–coding RNAs and DNA methylation (Collins *et al*., 2011; Sijen, 2015). These modifications play a role in regulation of gene activity and expression during development and differentiation or in response to environmental influences (Jaenisch and Bird, 2003; Loscalzo and Handy, 2014; Sant *et al*., 2012). Epigenetic patterns are preserved during cell division and are inherited across generations as DNA is inherited from generation to generation (Sen *et al*., 2015). Epigenetic changes are influenced by environmental and genetic factors which result in the alteration of gene expression. It has been observed that an individual's epigenetic pattern may modify over time due to response to environmental factors, such as diet, smoking and consumption of alcohol or genetic factors, such as mutations (Dupont *et al*., 2009; Kader and Ghai, 2015). All cells in an organism contain the same genetic information, however epigenetic mechanisms result in differential gene expression profiles in all tissues and cells (Moore *et al*., 2012). In this study, the focus will be on DNA methylation.

2.2 DNA methylation

DNA methylation is a biochemical process that occurs in both prokaryotic and eukaryotic organisms which involves a methyl group ($-CH_3$) attaching to the 5' position of a cytosine pyrimidine ring (Figure 2.1). DNA methylation is involved in cellular processes such as regulation of gene expression, cellular development, genomic imprinting and X chromosome inactivation (Moore *et al*., 2012; Smith and Meissner, 2013). DNA methylation patterns change during development, for instance, during early embryogenesis methylation patterns in the genome are removed and later on re-established (Paulsen and Ferguson-Smith, 2001; Sant *et al*., 2012). In a DNA sequence, a cytosine nucleotide is followed by a guanine nucleotide which forms the CpG dinucleotide. In mammals, ~60-90% of CpG dinucleotides are methylated and only 15% are found in clusters known as CpG islands (Ghosh *et al*., 2010). CpG islands are >200 bp long, have a GC content that is greater than 50% and have an observed to expected ratio of CpG $\geq$ 0.6 (Bird, 1986; Espada and Esteller, 2010; Gardiner-Garden and Frommer,

1987; Zhang *et al*., 2015). In CpG islands, the frequency of CpG sites is 10 times greater than the average distribution of CpGs (Wang *et al*., 2016). CpG sites that are found outside CpG islands are generally methylated, however, majority of CpG sites found within CpG islands are unmethylated. CpG islands account for 1−2% of the genome and are mostly found in promoter and exonic regions (Larsen *et al*., 1992; Shen *et al*., 2007). Most studies focus on methylation patterns in CpG islands found in promoter regions have shown that methylation patterns differ from other regions in the genome, suggesting a specific biological role for these promoter CpG islands (Jones, 2012).

DNA methyltransferases (DNMT) are enzymes that catalyse DNA methylation, namely, DNMT1, DNMT3A and DNMT3B. DNMT1 maintains methylation after each cellular DNA replication cycle and DNMT3A and DNMT3B catalyse *de novo* methylation to establish new DNA methylation patterns in early developmental stages (Fernandez *et al*., 2012; He *et al*., 2011). The fourth DNA methyltransferase known as DNA methyltransferase 3 like (DNMT3L), shares homology and has similar activity to DNMT3A and DNMT3B. However, DNMT3L does not function in the same way as there are no amino acids present for methyltransferase activity (Borghese *et al*., 2012). The DNMT3L may be involved in establishment of maternal genomic imprints by stimulation of *de novo* methylation by DNA cytosine methyltransferases 3 alpha (Chédin *et al*., 2002).

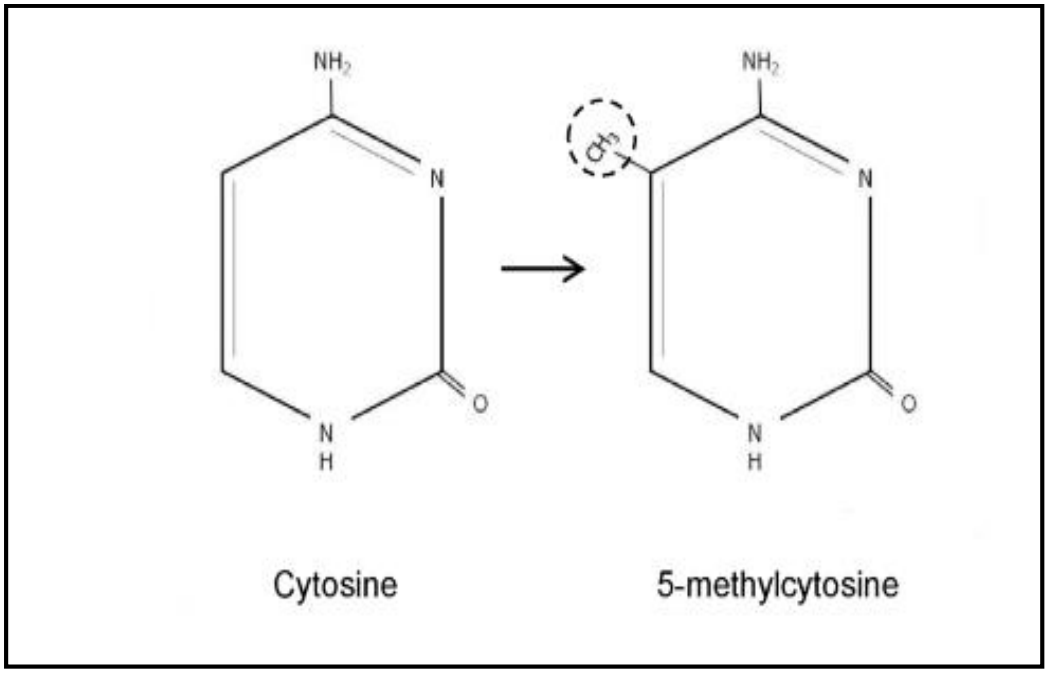


Figure 2.1: Basic overview of DNA methylation. A methyl group (CH_3) attaches covalently to a cytosine residue (Vidaki *et al*., 2013).

2.3 Role of DNA methylation in gene regulation

The genetic basis of gene regulation has been studied across tissues and populations and has suggested that genetic variations are associated with gene regulation and primarily occur in close proximity to the transcription start sites (Bell *et al*., 2011; Stranger *et al*., 2007). Epigenetic modifications such as DNA methylation and the interaction with transcription factors may play a critical role in the regulation of gene expression in humans (Hellman and Chess, 2007; Lister *et al*., 2009; Messerschmidt *et al*., 2014; Rivenbark *et al*., 2012). Gene expression may be directly or indirectly controlled by the DNA methylation levels at specific loci which results in silencing of genes or activation of genes (Antunes *et al*., 2016a).

Promoter regions that are highly methylated have weak binding capabilities with transcription factors which results in a decrease in expression of the genes (Kapoor *et al*., 2005; van Eijk *et al*., 2012). DNA methylation has the ability to directly prevent the transcription factors from binding resulting in change of chromatin structure which limits the accessibility of the transcription factors to the gene promoter (Figure 2.2) (Lim and Maher, 2010). Methyl-CpG-binding protein 2 (MePC2) is a protein that binds to methylated DNA through a methyl-CpG-binding domain (MBD) (Bird, 2002). The MBD-protein complex facilitates transcriptional repression which results in deacetylated repressive chromatin structure (Bird, 2002). Environmental and genetic factors play a role in gene expression which results in varied methylation patterns, however, these factors are not completely understood (Tammen *et al*., 2013; van Eijk *et al*., 2012; Wagner *et al*., 2014).

Martino and Saffery. (2015) examined the association between DNA methylation and gene regulation. In this study, the gene expression was separated into low expression and high expression. The methylation levels at all CpG sites in highly expressed genes and low expressed genes were determined using the IlluminaHumanMethylation450K.db annotation package. It was observed that CpGs in genes that had high expression levels have a greater unmethylated area (0-20% methylated) and CpGs in genes that had low expression levels have a greater methylated area (80-100% methylated). The association between DNA methylation and the transcription start site was also examined. It was found that within 200bp of the transcription start site (near the promoter) and the first exon, there was an inverse correlation between DNA methylation and gene expression (Martino and Saffery, 2015).

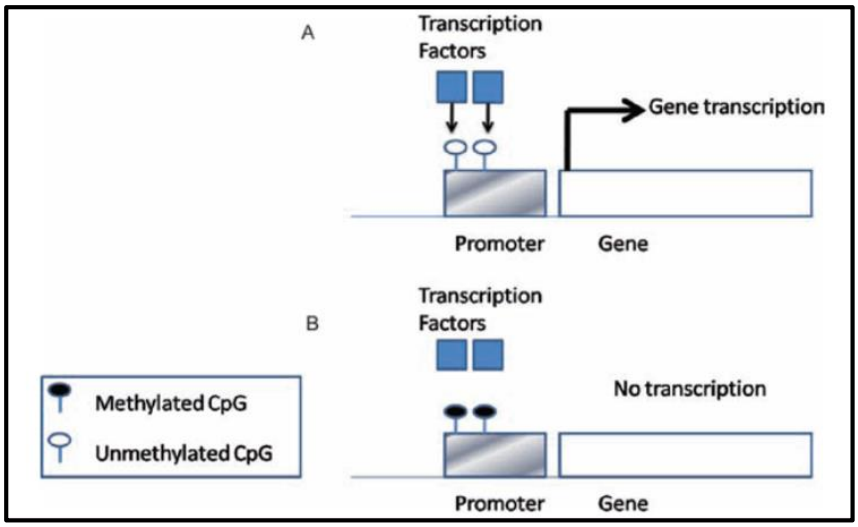


Figure 2.2: Regulation of gene expression by DNA methylation. (A) Absence of methyl groups in the CpG island promoter permits binding of transcription factors which facilitates transcription initiation (B) Presence of methyl groups hinders binding of transcription factors which prevents gene expression (Lim and Maher, 2010).

2.4 Gene body methylation and gene expression

The transcribed portions of genes are known as gene bodies which display hypermethylated patterns in comparison to the rest of the genome (Aran *et al*., 2010). DNA methylation has contradicting effects on promoters and gene bodies known as the DNA methylation paradox. According to the DNA methylation paradox, DNA methylation has an inverse correlation with gene expression in promoters (Jones, 1999; Jjingo *et al*., 2012; Klose and Bird, 2006) whereas in gene bodies there is a positive correlation with gene expression (Ball *et al*., 2009; Laurent *et al*., 2010). Gene expression patterns are maintained through DNA methylation (Aran *et al*., 2010) which occurs in CpG islands located in gene promoter sites as well as gene bodies which are located away from promoter sites (Ball *et al*., 2009). Research by Neri *et al*. (2017) indicates that in embryonic stem cells, intragenic methylation may serve to confine transcription initiation to canonical promoters thereby inhibiting production of aberrant transcripts. In vertebrates, invertebrates and plants it has been observed that the methylation profile of gene bodies are conserved which suggests that it plays an important role in multicellular organisms (Feng *et al*., 2010; Suzuki and Bird, 2008; Zemach *et al*., 2010). It has been suggested that transcription elongation is stimulated by DNA methylation and that the trimethylated histone H3 at lysine 36 (H3K36me3 - has an association with elongation) possibly recruits DNMTs (Hahn *et al*., 2011) thus, despite the fact that majority of gene bodies are CpG poor and highly methylated, transcription elongation still pursues (Jones, 2012). Whole-genome studies have

demonstrated that the exons in gene bodies are hypermethylated in comparison to the introns and transitions in the degree of methylation that occur at exon–intron boundaries. Suggesting that methylation may play a role in regulation of splicing (Laurent *et al*., 2010). In a study by Hellman and Chess. (2007) methylation patterns of the gene body in X chromosomes were analysed using Affymetrix 500000 (500K) single-nucleotide polymorphism (SNP) mapping array. An allele-specific analysis was performed along the human X chromosome in which the active X chromosome displayed more than twice as much allele-specific methylation in comparison to the inactive X chromosome. Majority of the methylation was observed in the gene body sites (Hellman and Chess, 2007).

2.5 Tissue-specific differentially methylated regions (tDMRs)

DNA methylation patterns are established via cellular differentiation as a result all cells and tissues have a unique DNA methylation profile (Bird, 2002; Nagase and Ghosh, 2008). Epigenetics studies have identified specific regions found within the genome which have varying methylation patterns depending on the cell type (Slieker *et al*., 2013; Song *et al*., 2009). These specific regions are known as tissue-specific differentially methylated regions (tDMRs) and have been observed in the human genome (Antunes *et al*., 2016b; Igarashi *et al*., 2008; Kitamura *et al*., 2007; Song *et al*., 2009). tDMRs are found in non-CpG island, CpG island promoters, intragenic and intergenic regions (Choufani *et al*., 2011). It is believed that tDMRs that are found at CpG islands are involved in development and tissue differentiation. These tDMRS generate cell or tissue-specific methylation patterns (hypo-/hypermethylation) (Illingworth *et al*., 2008; Kader and Ghai, 2015). As DNA fingerprints are unique, several studies have shown that DNA methylation profiles are specific to tissues and cells and can be used for identification of different cells or tissues in organisms (Choi *et al*., 2014; Christensen *et al*., 2009; Day *et al*., 2013; Ohgane *et al*., 2008; Rakyan *et al*., 2008). Therefore, the methylation profiles observed in tDMRs can be used to differentiate between the tissue and cells.

2.6 Identification of tDMRs

A minor but significant proportion of CpG islands are differentially methylated between various tissues and cell types (Eckhardt *et al*., 2006; Illingworth *et al*., 2008; Shiota *et al*., 2002). Characterisation of these differences may determine the function of tissue-specific CpG island methylation pattern (Irizarry *et al*., 2009; Ladd-Acosta *et al*., 2007). Many studies have

identified tDMRs by comparing DNA methylation patterns of various cells and tissues (Christensen *et al.,* 2009; Day *et al.,* 2013; Rakyan *et al.,* 2008).

To study the difference in methylation patterns, Slieker *et al*. (2013) developed an algorithm based on 450K methylation data to identify tDMRs. Genome-wide data was obtained for four peripheral tissues (buccal swabs, hair follicles, saliva and blood) and six internal tissues (pancreas, liver, omentum, muscle, spleen and subcutaneous fat) using Illumina 450K DNA methylation chips. The criteria for identification of tDMRs was that CpG sites had to display an average methylation difference of >10% between the target tissue and other tissues, should contain a minimum of three differentially methylated CpG sites with an inter-CpG distance that is ≤ 1 kb and interrupted by a maximum of three non-differentially methylated positions. Using the developed algorithm, 3 533 putative tDMRs in the peripheral tissues were identified and 5 382 putative tDMRs in the internal tissues were identified (Slieker *et al*., 2013).

In a study by Ghosh *et al*. (2010), 12 samples were collected from different brain tissues and non-brain tissues of seven autopsy patients that were tissue donors to determine methylation patterns. Brain tissues were collected from basal ganglia, the cerebellum and the frontal and temporal lobes of the cerebrum and non-brain tissues were collected from kidney, lung, prostate, spleen, stomach and thyroid. Restriction landmark genomic scanning (RLGS) was used to generate each patient's methylation profile for all tissues. A total of 33 RLGS profiles were generated and 34 differentially methylated CpG islands were identified. A total of 141 RLGS spots showed tissue-specific methylation differences within all tissues of at least a single patient. From the 141 spots, 82 of the spots were present or partially methylated in one tissue of each patient. A cluster analysis of the RLGS data showed that the brain tissue clustered together the strongest and were distinct from non-brain tissues which showed limited clustering (Ghosh *et al*., 2010).

In a study by Lokk *et al.* (2014), tissue-specific DNA methylation patterns were studied in 17 somatic tissues of post-mortem humans. The different somatic tissue used were: abdominal and subcutaneous adipose tissue, bone, joint cartilage, yellow and red bone marrow, coronary and splenic artery, abdominal and thoracic aorta, gastric mucosa, lymph node, tonsils, bladder, gall bladder, medulla oblongata, and ischiatic nerve. A total of 14 441 tDMRs were identified in which 11 242 (77.8%) were found in genes. From the 11 242 tDMRs, 4688 tDMRs (41.7%) were found in gene promoter regions in which 35.6% were found in CpG island regions and 6 554 (58.3%) were found in gene body regions in which 44.1% were found in CpG island

regions. Majority of the tDMRs were found in the gene body regions which may be an indication of alternate promoters being present (Maunakea *et al*., 2010). The number of tDMRs that were identified varied significantly between tissues that have different functions. The tonsils showed the highest number of hypermethylated tDMRs and red and yellow bone marrow showed that highest number of hypomethylated tDMRs. Genes that indicate hypomethylation in some tissues are at most times associated with a tissue-specific function. For example, genes that were detected in arteries as being hypomethylated are known as mediators of blood vessel development and for morphogenesis. Hypomethylated genes were also detected in tonsils and were found to be associated with immune response and leukocyte activation. Thus, this study has shown that hypomethylation of genes is associated with tissue-specific function (Lokk *et al.,* 2014).

2.7 Genetic factors that influence DNA methylation

DNA methylation patterns can be altered by genetic factors which can affect transcription and phenotypic variation (Feil and Fraga, 2012; Murrell *et al*., 2004; Schübeler, 2015). Studies have identified that the DNA methylation state of a locus depends on the DNA sequence of an individual (Bock *et al*., 2006; Chandler *et al*., 1987; Silva and White, 1988). The degree to which genetic variation affects DNA methylation is not known. Furthermore, the extent to which the variation in DNA methylation causes variation in gene expression in individuals is also unknown (Bell *et al*., 2011; Schübeler, 2015). Examining methylation patterns in families and twins enables the examination of the effects that genetic variation has on methylation patterns (Bjornsson *et al*., 2008; Gertz *et al*., 2011; Kaminsky *et al*., 2009), however environmental influences may also effect methylation patterns (Feil and Fraga, 2012). Studies on monozygotic and dizygotic twin pairs by Kaminsky *et al.* (2009) and Ollikainen *et al*. (2010) have demonstrated that epigenetic differences are more prominent in dizygotic twin pairs compared to monozygotic twin pairs.

2.8 Environmental factors that influence DNA methylation

Epigenetic processes regulate the interaction between the environment and the genome (Bonasio *et al*., 2010; Feil and Fraga, 2012). The establishment and maintenance of DNA methylation can be altered by environmental factors such as diet and nutrition, life experiences and age, however, the extent to which these factors alter DNA methylation is yet to be determined (Jirtle and Skinner, 2007; van Dongen *et al*., 2014). This may be an indication that alterations in DNA methylation status is possibly the step between the environment and human

diseases. Monozygotic twin studies have investigated the effects of age on epigenetic differences and have demonstrated that changes in DNA methylation patterns are associated with age (Christiansen *et al*., 2016; Fraga *et al*., 2005; Martino *et al*., 2013; Wong *et al*., 2010). Dietary and nutrition factors have an inevitable impact on epigenetic patterns (Herceg, 2007). For example, a study by Jirtle and Skinner. (2007) has shown that changes in folate diets can alter DNA methylation. However, determining the relationship between diet and epigenetic patterns is difficult as these factors show subtle changes over time. Studies have also shown that unfavourable living conditions such as famine, chemical and environmental pollutants or smoking can have an impact on an individual's DNA methylation pattern without altering the genetic sequence (Lee and Pausova, 2013; Tobi *et al*., 2009).

2.9 DNA methylation and its application in forensic science

DNA analysis provides an important tool in solving cases in forensic, such as identifying victims from disasters or crimes, exonerating innocent people that are convicted and identification of suspected perpetrators (Dumache *et al*., 2016). In particular DNA from biological material such as blood, semen, saliva, vaginal fluid and urine are commonly found at crime scenes and serve as evidence which provides important information (Kader and Ghai, 2015). It is essential that methods used to analyse biological material are not destructive because most times a minute amount of evidence is available for analysis. There are methods available for identifying biological material, however the sensitivity and specificity of the methods are limited. Analysis of DNA methylation profiles provide an alternative method to the current methods used for forensic applications and shows great potential because it does not use a large amount of sample, there is no additional destruction of the samples as DNA isolated for short tandem repeat (STR) typing can be used for methylation analysis (Frumkin *et al*., 2011).

2.9.1 Identification of artificial DNA

Detection of artificial DNA is important as any individual with the understanding of DNA and equipment would be able to produce specific DNA fragments with a genetic profile. The artificially synthesized DNA can be combined with actual human tissue and planted at crime scenes. Artificial DNA produces a typical profile after forensic analysis.

A study by Frumkin *et al*. (2010), amplified five loci, one was a reference Combined DNA Index System (CODIS) locus (FGAref) and the other four were non-CODIS loci (NT18, ADD6, MS53, SW14). After polymerase chain reaction (PCR) amplification of the natural

DNA which was extracted from blood and saliva and artificial DNA extracted from hair using organic extraction, the artificial DNA was indicated by amplification of the reference locus and no amplification of the non-CODIS loci. Natural DNA was indicated when all five loci were amplified and the FGAref loci was amplified in both the natural and artificial DNA. The amplified loci were sequenced and these results showed the artificial DNA was unmethylated compared to the natural DNA, thus the artificial and natural DNA can be distinguished (Frumkin *et al*., 2010).

A study by Wang *et al*. (2015) used a methylation sensitive restriction enzyme assay to distinguish artificial and natural DNA from blood samples. Artificial blood was created to determine whether it can be distinguished from natural blood. Peripheral blood samples were collected from three healthy participants which included two males and one female. Seven loci were included in this study (L1-L7): L2-L6 contained recognition sites for the methylation sensitive restriction enzyme *HhaI* and L1 and L7 were used as positive controls as these loci do not have a recognition site for *HhaI*. In the artificial blood sample L3 and L6 loci displayed unmethylated patterns and in the natural blood sample L4 and L5 loci show patterns of methylation (Wang *et al*., 2015). Initially, DNA STR profiles were generated for the natural and artificial bloodstain samples. The DNA STR profiles demonstrated that the artificial blood could not be distinguished from the natural blood.

2.9.2 Parent of origin

Parentage testing and individual identification can be analysed using DNA methylation markers that are specific to imprinted loci because alleles from maternal and paternal origin are differentially imprinted at these loci (Gršković *et al*., 2013; Nakayashiki *et al*., 2009). Parent of origin is a specific expression that is also known as genomic imprinting which is a result of maternal and paternal alleles from a subset of genes being expressed differentially (Bartolomei and Ferguson-Smith, 2011; Hata *et al*., 2002; Reik and Walter, 2001). There is a subset of genes which are known as imprinted genes: *H19*, *IGF2R*, *P57*, *SNRPN*, *PEG 1* and *PEG3*. The *H19*, *IGF2R* and *P57* genes are expressed when inherited from the maternal genome and the *SNRPN*, *PEG 1* and *PEG 3* genes are expressed when inherited from the paternal genome (Hata *et al*., 2002). These imprinted genes are involved in developmental processes such as fetal growth, placenta function, regulation of embryonic development and maternal behaviours (Barlow, 1995; Jaenisch, 1997; Moore and Haig, 1991; Tilghman, 1999). Differentially methylated regions in the maternal and paternal alleles of the imprinted genes

have a molecular basis for regulation of the imprinted genes expression, thus, it is believed that DNA methylation controls genomic imprinting in mammals (Ferguson-Smith and Surani, 2001; Neumann and Barlow, 1996; Xie *et al*., 2012).

A methylation sensitive restriction enzyme analysis was done by Lefebvre *et al*. (1997) to determine whether the CpG island of the *PEG 1* imprinted gene displayed a parental-specific methylation pattern. It was observed that in embryos and undifferentiated stem cells the *PEG 1* gene cytosine residues were partially methylated. The embryos used had carried a targeted mutation at the *PEG 1* locus, thus the partial promoter methylation pattern showed a strict parent of origin specific differential methylation. The paternal allele showed unmethylated patterns and the maternal allele showed complete methylation patterns at the CpG sites (Lefebvre *et al*., 1997).

A study by Zhao *et al*. (2005) determined the parental origin of alleles by using parent of origin specific DNA methylation markers. An imprinted single nucleotide polymorphism (SNP) (locus rs220028) was used for this analysis. In order to distinguish between the maternal and paternal allele, a mutagenically separated PCR and SNP typing technique was developed and performed on 18 heterozygous children to analyse one imprinted SNP. The imprinted SNP was detected specifically on the maternal allele. Thus, parent of origin alleles can be detected using parent of origin specific DNA methylation markers (Zhao *et al*., 2005).

2.9.3 Age estimation

Aging is a developmentally regulated process that is in part mediated by epigenetic modifications, such as DNA methylation which is involved in cellular senescence (Marciniak-Czochra *et al*., 2009; Schellenberg *et al*., 2011). Determining the age of an individual can provide essential leads in identifying unknown persons in criminal investigations, disaster victim identification as well as cases in which there is identity fraud (Zubakov *et al*., 2016). Research has shown that aging is associated with DNA damage and telomere shortening (Fraga and Esteller, 2007; Koch and Wagner, 2011; Schellenberg *et al*., 2011). An individual's age can be determined by analysing osteal markers such as bones and teeth. However, this would only be possible in cases in which a skeleton exists (Thevissen *et al*., 2012). During the aging process an individual's methylation profile is continually changing, therefore, the age of a biological trace can be analysed using DNA methylation (Gršković *et al*., 2013). Studies have observed that the CpG sites that are hypermethylated or hypomethylated form a linear correlation of aging in terms of DNA methylation (Christensen *et al*., 2009; Meißner *et al*.,

1997). Determining the age of an unknown biological trace would be beneficial for personal identification and forensic cases (Koch and Wagner, 2011).

A study by Bocklandt *et al.* (2011) identified CpG methylation in saliva using high density genome-wide screening. The saliva samples were obtained from 34 identical male twins between the ages of 21-55. A total of 88 CpG sites were identified in or close to 80 genes that are involved in age-related diseases. From the 88 CpG sites, 10 sites were previously reported to be correlated with age in whole blood and in isolated $CD4^+$ and $CD14^+$ (Rakyan *et al*., 2010). This study went one step further and included individuals that were between the ages of 18-70 and consisted of 31 men and 29 women for identification of methylation patterns of promoter sites of three genes, namely, EDAR associated death domain (*EDARADD*), target of myb1 (chicken)-like 1 (*TOM1L1*) and neuronal pentraxin II (*NPTX2*). The methylation patterns observed at the three promoter sites showed a linear correlation with age over a period of five decades, thus, an individual's age could be predicted with an accuracy of 5.2 years (Bocklandt *et al.*, 2011).

A multi-tissue predictor of age was developed by Horvath (2013) to allow the estimation of DNA methylation age of most tissue and cell types. The predictor was developed using 7 844 samples from 51 healthy tissue and cells types from 82 individuals by Illumina DNA methylation array datasets. There were 353 CpG sites that formed the age clock. From the 353 CpG sites, 193 CpG sites showed a positive correlation with age and 160 CpG sites showed a negative correlation with age. The accuracy of the age predictor was tested using three accuracy measures: Pearson correlation coefficient, median error and average age acceleration. An analysis using the age predictor was performed on 5 826 cancer samples from 32 individual cancer datasets. There was a large correlation between DNA methylation age and patient age in some of the cancer tissue, however the correlation between DNA methylation and chronological age was not strong. Every cancer tissue showed a significant age acceleration with a mean age acceleration of 36.2 years. Age acceleration was also determined in genes that had mutations. The TP53 mutation showed a lower age acceleration in five different cancer types (acute myeloid leukemia, breast cancer, ovarian cancer, uterine corpus endometrioid and lung squamous cell carcinoma). The *TP53* gene showed significant mutation in four out of the 13 cancer datasets. It was concluded that the model was able to predict the age of most tissues and cell types and performs well in heterogeneous tissues (Horvath, 2013).

Another age estimation-based study Zbieć-Piekarska *et al.* (2015) collected blood stains from 45 individuals between the ages of one-81 years. The samples were stored at room conditions for five, 10 and 15 years and thereafter, the *ELOVL2* locus was analysed by bisulfite treatment and pyrosequencing. Interestingly, methylation of the locus was proven to provide a powerful and reliable estimation of age in the forensic samples, ranging between 60–78% (Zbieć-Piekarska *et al*., 2015).

2.9.4 Differentiation between monozygotic/dizygotic twins

Monozygotic twins can be used to determine how environmental factors play a role in determining differences in phenotypes and complex disease (Fraga *et al*., 2005; Ribel-Madsen *et al.*, 2012). Monozygotic twins have the same genetic material and their methylation patterns at birth are nearly identical, however, as they each are exposed to environmental factors, their methylation patterns may differ (Bocklandt *et al*., 2011; Martino *et al*., 2013; Ye *et al*., 2013).

In a study by Fraga *et al*. (2005) epithelial skin cells, intra-abdominal fat, and skeletal muscle biopsies were collected from eight caucasian twin pairs. It was observed that in the older monozygotic twin pairs, these three tissues showed epigenetic differences as a result of different lifestyles and not spending their lives together. Thus, the DNA methylation patterns of different tissues in monozygotic twins are distinct due to different lifestyles and may explain some of the differences in phenotypes and the differential frequency or onset of common diseases (Fraga *et al*., 2005). A DNA methylation study using the Infinium Human-Methylation450 BeadChip platform was conducted on buccal swabs from 10 monozygotic twin pairs and five dizygoyic twin pairs from birth to the age of 18 months. It was evident that large epigenetic changes occur during the first 18 months (Martino *et al*., 2013).

Coolen *et al*. (2011) performed a quantitative mass spectrometric assay on 128 pairs of monozygotic twins and 128 pairs of dizygotic twins to determine DNA methylation patterns in four imprinted control region genes (*IGF2, H19, KCNQ1* and *GNAS*) and a single non-imprinted gene (*RUNX1*). It was observed that the DNA methylation patterns between each twin group was similar. The monozygotic twin pairs and the dizygotic twin pairs both displayed great variation in the DNA methylation patterns in the *H19* and *IGF2* genes. The *RUNX1* gene differed as DNA methylation levels were low. A non-parametric Spearman correlation test was performed to determine the association of DNA methylation levels at various CpG sites within an individual. The *H19* showed a strong association and the *RUNX1* showed a random association to DNA methylation. This suggests that within individuals the degree of DNA

methylation at specific CpG sites is a function of the degree of methylation of neighbouring sites. A Kruskal-Wallis one-way analysis of variance test (ANOVA) and a post hoc Dunn's test was performed to determine the DNA methylation difference in monozygotic and dizygotic twin pairs. It was observed that there was a significant difference between monozygotic and dizygotic twin pair methylation patterns. In the *H19* gene, monozygotic twin pairs showed less methylation differences as compared to dizygotic twin pairs (Coolen *et al*., 2011).

In a study by Li *et al*. (2011) a high throughput epigenome microarray analysis known as Illumina's HumanMethylation27 BeadChip was performed to identify epigenetic differences in adult monozygotic twin pairs. There were 377 CpG sites selected that showed variation in the methylation patterns which suggests that the CpG sites could be used as a biomarker to distinguish between monozygotic twins. This study demonstrates the potential of using epigenetic markers to differentiate monozygotic twin pairs (Li *et al*., 2011).

2.9.5 Human body fluid identification

In my present study, the focus is on human body fluid identification by methylation analysis. In forensic investigations, biological samples such as body fluids that are obtained from crime scenes have proven to provide important information that can link the evidence and the crime (Forat *et al*., 2016; Lee *et al*., 2012). The DNA extracted from the body fluid can be useful in reconstruction of the crime scene by determining the type and origin of the biological sample as well as identifying the donor of the biological sample (Lee *et al*., 2012). For example, if semen was present at a crime scene this could indicate sexual assault.

The first DNA methylation-based assay to identify forensic tissue was done by Frumkin *et al*. (2011) in which methylation sensitive restriction enzyme PCR (MSRE-PCR) was used to differentiate body fluids such as semen, blood, saliva, vaginal fluid, skin, urine and menstrual blood. In this study, the potential of the DNA methylation-based assays was observed by analysing tissue identification at 15 loci that show differentially methylated patterns. A methylation ratio was calculated for each locus to distinguish the biological fluids. The methylation ratio in semen samples for L91762 /L68346 was low. In all other body fluids for L91762/L68346, the methylation ratio was high. The methylation ratio of L76138/L26688 was high in semen and epidermis samples and low in blood and saliva. The study concluded that L76138/L26688 and L91762/L68346 could precisely identify semen and skin epidermis sample respectively (Frumkin *et al*., 2011).

A study by Lee *et al*. (2012) generated DNA methylation profiles to determine differential DNA methylation for body fluid identification using bisulfite sequencing. Five tDMRs (tDMRs for *USP49*, *DACT1*, *PFN3*, *PRMT2* and *HOXA4* genes) were used to generate methylation profiles for saliva, semen, vaginal fluid, menstrual blood and blood. The *USP49* and *DACT1* gene tDMRs were selected as semen-specific markers and tDMRs for genes *PFN3*, *PRMT2* and *HOXA4* could be used as markers for blood. The tDMR for *USP49* and *DACT1* genes showed hypomethylation in majority of the semen samples and hypermethylation in all other body fluid samples. Chi-square test or Fisher's exact test was performed to determine pairwise comparison of methylation at each marker. The results showed that for USP49 and DACT1 markers, semen differed from all other body fluids. Therefore, these tDMRs can be used to identify semen samples positively. The HOXA4 tDMR showed high levels of methylation in female saliva and in blood, however, the methylation patterns did not differ enough to differentiate the fluids accurately (Lee *et al*., 2012). In the PFN3 tDMR, hypermethylation patterns were observed for majority of the body fluids. Statistical analysis showed that vaginal fluid DNA methylation could be significantly differentiated from all other body fluids, thus, PFN3 tDMR has the potential to distinguish vaginal fluid. Hypomethylation patterns were observed for the PRMT2 tDMR in majority of semen samples, however, hypermethylation patterns were observed in menstrual blood and vaginal fluid. This marker has the potential to distinguish semen from other body fluids as well as distinguish menstrual blood and vaginal fluid from other body fluids based on the observed methylation patterns.

A study by An *et al.* (2013) used three tDMRs that were specific for semen (DACT1, USP49 and PRMT2) to determine the effect of age on DNA methylation patterns using bisulfite sequencing. Samples included: blood, saliva and semen. The semen samples were obtained from young and old men and this analysis showed that three tDMRs showed unmethylated patterns in most of the samples and methylation patterns in blood and saliva samples. Chi-square test or Fisher's exact test was performed to determine pairwise comparison of methylation at each marker. There was no significant difference observed in all body fluids for DACT1 and USP49 tDMRs, however, the PRMT2 tDMR showed significant differences for six CpG sites (1, 2, 3, 15, 17 and 19) in blood samples between young and elderly men (An *et al.*, 2013).

A set of four markers that differentiated between blood, saliva, semen and epithelial cells: C20orf117, ZC3H12D, BCAS4, and FGF7 were used to demonstrate differential methylation

in blood, saliva and semen (Madi *et al*., 2012). Two of the markers have been previously reported by Eckhardt *et al*. (2006), C20orf117 identified hypermethylation patterns in white blood cells and BCAS4 identified hypermethylation patterns in sperm. The ZC3H12D and FGF7 markers were used to identify semen and epithelial cells, respectively. In this study the C20orf117 marker demonstrated that the blood samples were hypermethylated in comparison to the other body fluids that were examined which is in agreement with results observed in the study conducted by Eckhardt *et al.* (2006). The ZC3H12D marker was successful in differentiating the semen samples from all other tissue that were examined. All semen samples showed hypomethylation patterns and all other tissue samples showed hypermethylation patterns. Based on the study by Eckhardt *et al.* (2006) the BCAS4 marker was expected to display hypermethylation patterns in semen compared to the other tissues, however, the study by Madi *et al*. (2012) showed that this marker was hypermethylated in saliva. The FGF7 marker showed hypermethylation patterns in semen relative to the other tissues which displayed hypomethylation patterns. Based on the results observed in this study and with further analyses, the ZC3H12D and C20orf117 epigenetic markers show great potential for identifying semen and blood, respectively (Madi *et al*., 2012).

In a recent study by Vidaki *et al*. (2016) identification and validation of 11 potential tDMRs, previously identified in methylation analysis (Rakyan *et al*., 2008; Rakyan *et al*., 2010) of blood, buccal cells and semen, was analysed in order to identify markers for blood, saliva and semen using bisulfite pyrosequencing. The markers selected for saliva and blood did not display a significant methylation difference to be proposed as saliva and blood markers due to some CpGs showing great inter-individual variation. The markers, cg04382920 and cg11768416 that were selected for identification of semen were successful in differentiation of semen from saliva and blood. These markers were further validated using a larger sample size and showed a high sensitivity (detected with 50 pg) and stability. Therefore, DNA methylation analysis holds a promising future for detection of body fluids (Vidaki *et al*., 2016).

A study by Ma *et al*. (2013) screened and identified tDMRs for forensic casework. DNA samples were collected from blood, saliva, semen and vaginal fluid. The DNA was pooled for each body fluid in an equal mixture and differentially methylated fragments were isolated using methylation sensitive represent difference analysis (MS-RDA). There were six tDMRs obtained that were blood-specific. Sequenom mass array quantitative analysis of methylation was used to quantify the methylation levels in the tDMRs. Two of the six tDMRs showed

hypomethylation patterns in blood compared to other body fluids and the remaining four tDMRs showed hypermethylation patterns in blood compared to other body fluids. A Chi-square test was performed for pairwise comparisons which showed a statistical difference between blood and other body fluids. Therefore, the differential methylation patterns of tDMRs makes it possible to distinguish between blood and other body fluids (Ma *et al*., 2013).

A study by Park *et al*. (2014) investigated novel DNA markers for the identification of blood, saliva, semen and vaginal fluid using Illumina Human Methylation 450K bead array. There were 16 samples collected for each body fluid. Using the methylation data, 2986 hypermethylated or hypomethylated regions that were specific for each type of body fluid were selected. Eight CpG sites were selected as being novel and forensically relevant DNA methylation markers, namely, cg26107890 and cg20691722 for saliva, cg06379435 and cg08792630 for blood, cg01774894 and cg14991487 for vaginal secretions and cg23521140 and cg17610929 for semen were identified. Pyrosequencing was used to analyse the eight novel markers in 80 body fluid samples. It was detected that each marker exhibited great sensitivity and specificity for identification of the target body fluid. This study suggests that the eight novel DNA methylation markers have the potential to differentiate body fluids in forensic casework (Park *et al*., 2014).

Antunes *et al*. (2016b) examined the PFN3A locus using pyrosequencing to decipher if it could be used to differentiate vaginal epithelia from other body fluids. The results showed that vaginal epithelia were able to be discriminated from other body fluids using the PFN3A primers using as little as 5ng of DNA. The study went a step further and performed mixture analysis to simulate sexual assault cases and to determine if the primers were specific to humans and higher primates. In the mixture analysis, DNA samples were selected from blood, semen and vaginal epithelia were mixed in ratios 75:25, 50:50 and 25:75. The results showed that when the DNA content of blood decreased from vaginal epithelia, the percentage methylation decreased and this was also observed for blood and semen mixtures. In vaginal epithelia and semen mixtures, when DNA content decreased in vaginal epithelia compared to semen, the percentage methylation also decreased. To test the specificity of the primers to humans, DNA from cat, mouse, chicken, cow, dog, bacterial pool (*Pseudomonas aeruginosa, Enterococcus faecali, Staphylococcus aureus,* and *Escherichia coli*), horse, chimpanzee, orangutan, and gorilla was isolated. The results showed that the PFN3A primers were specific to humans.

Thus, the PFN3A locus can be used to differentiate vaginal epithelia in forensic cases (Antunes *et al*., 2016b).

Silva *et al*. (2016) examined 15 samples of different body fluids blood, saliva, semen, menstrual blood, nasal epithelia and vaginal epithelia in a validation test using pyrosequencing. Three markers were used for identification of body fluids: ZC3H12D which is semen specific, cg06379435 which is specific for blood and BCAS4 which is specific for saliva. Majority of the CpG sites displayed a significant difference in methylation levels when being compared to different body fluids. A one-way ANOVA test was used to determine a statistical difference in methylation values across markers. The BCAS4 marker showed no significant difference ($p > 0.05$) in methylation level at CpG 5 when saliva and vaginal epithelia samples were compared and the cg06379435 marker showed no significant difference ($p > 0.05$) in methylation level at CpG 2 when blood and menstrual blood samples were compared. However, most of the CpG sites for each marker showed different average methylation values between the target body fluid and all other body fluid. A reproducibility test was performed on eight samples of three body fluids (blood, saliva and semen) in two different laboratories in which the pyrosequencing data was analysed and compared to determine the reliability of the results for the markers. A t-test analysis was performed and it was found that there was no significant difference when methylation values were compared for each CpG site in either laboratory (Silva *et al*., 2016).

2.10 Stability of DNA methylation markers

In order to use DNA methylation markers in forensic investigations, the markers are required to be extensively validated using different methylation techniques (Vidaki *et al*., 2016). Validation will involve testing the accuracy, sensitivity and the applicability in cold case samples, for identification of forensically relevant body fluid (Forat *et al*., 2016).

Three different forensic environmental influences were simulated to analyse 75 body fluid (peripheral blood, menstrual blood vaginal fluid, saliva and sperm) samples to determine the stability of methylation levels (Forat *et al*., 2016). The three different environmental conditions were: dry at room temperature, wet in an exsiccator and outside on the ground. All samples were analysed immediately after collection and analysed one, two, three and six months after storage. The main difference in methylation levels were caused by humidity and majority of samples remained stable after six months of storage. All body fluid samples in the dry at room temperature condition remained stable for all markers throughout the study (Forat *et al.*, 2016).

An *et al*. (2013) exposed blood, saliva, semen, menstrual blood, and vaginal fluid samples to an environmental condition which was ambient temperature in the shade for a period of 75 days to determine whether DNA methylation profiles could be established using damaged DNA. It was observed that once DNA was extracted from all body fluids after the 75-day period, all body fluids were successfully analysed except saliva samples.

To assess the robustness and stability of the markers: ZC3H12D, BCAS4, and cg06379435 to detect semen, blood and saliva, respectively, Silva *et al*. (2016) simulated natural DNA fragmentation by heating genomic DNA extract samples in deionized water at 95°C for 10, 15, 20, and 25 minutes. The study went one step further to analyse DNA from three simulated case-type samples which were blood on cotton fabric, semen on cotton fabric and a saliva swab from the lid of a coffee drink. The results displayed that under the degradation condition, all markers were stable and primers successfully amplified the target sequence. Additionally, there were no significant differences observed in methylation levels of degraded samples when compared to those samples that were not exposed to degradation conditions. The three simulated case-type samples were successfully modified, amplified and pyrosequenced. Thus, each marker was able to successfully amplify the body fluid it was selected for: the ZCH12D marker showed low levels on methylation in the semen samples and high methylation in blood and saliva, the cg06379435 marker showed higher methylation levels in blood compared to semen and saliva and the BCAS4 marker showed higher methylation levels saliva compared to semen and blood.

2.11 Techniques used for DNA methylation analysis

Several techniques have been developed to specifically identify DNA methylation patterns. There are three broad categorizations of methods that enable the identification of methylation profiles which are: immunoprecipitation, endonuclease digestion and chemical modification of cytosine residue based methods (Silva *et al*., 2016).

2.11.1 Affinity enrichment

The affinity enrichment also known as immunoprecipitation techniques involve specific interactions between proteins and methylated DNA (Jacinto *et al*., 2008). These methods are based on using methyl-CpG binding protein 2 (MeCP2 and MBD2) or antibodies that bind to methylated DNA (Figure 2.3) (Laird, 2010). The two most widely used immunoprecipitation methods are, methyl-CpG immunoprecipitation (MCIp) and methylated DNA immuno-precipitation (MeDIP). These methods are simple, there is no digestion or bisulfite treatment of the genomic DNA and the results of the analysis are relatively easy to deduce.

2.11.1.1 Methyl-CpG immunoprecipitation (MCIp)

The methyl-CpG immunoprecipitation (MCIp) method uses an immobilized recombinant protein which consists of a methyl binding domain fused to human immunoglobulin G (IgG) (Gebhard *et al*., 2006a; Gebhard *et al*., 2006b; Schmidl *et al*., 2009). The antibody-fused protein binds to the methylated CpG regions on double stranded DNA (Sonnet *et al*., 2013). Initially, double stranded DNA is digested and denatured, followed by incubation with sephrose beads (Schilling and Rehli, 2007). Unbound DNA is removed using centrifugation. Next the beads with bound DNA are washed with increasing salt concentrations to elute fragments that contain methylated CpG. The methylation status is analysed based on the salt fractionation (Gebhard *et al*., 2006a; Klose *et al*., 2005). Therefore, fragments that have a low methylation status will be a result of low salt fractions and fragments that have a high methylation status will be a result of high salt fractions.

2.11.1.2 Methyl-DNA Immunoprecipitation (MeDIP)

An immunocapturing technique that is based on direct immunoprecipitation of methylated DNA is known as methyl-DNA Immunoprecipitation (MeDIP) which uses monoclonal antibody specific for 5-methylcytidine (5mC) to purify methylated DNA (Weber *et al*., 2005). The genomic DNA is sheared through sonication resulting in the production of random sized fragments that range between 300-600bp. The fragmented DNA is denatured into single-stranded DNA as the antibody has a greater affinity with this form of the methylated DNA. This is followed by immunoprecipitation, in which proteins are coupled with standard or magnetic beads, incubated with the anti-5mC antibody and several wash steps to remove the excess antibodies. The methylated status of the immunoprecipitated DNA can be analysed using PCR, in which primers are designed to target a specific gene of interest. Immunoprecipitation techniques require an antibody that is of high quality. An antibody that does not have a high quality will result in an inadequate DNA-protein enrichment (Bulyk, 2006; Minard *et al*., 2009). This method however has low sensitivity and specificity, and requires high quality of DNA.

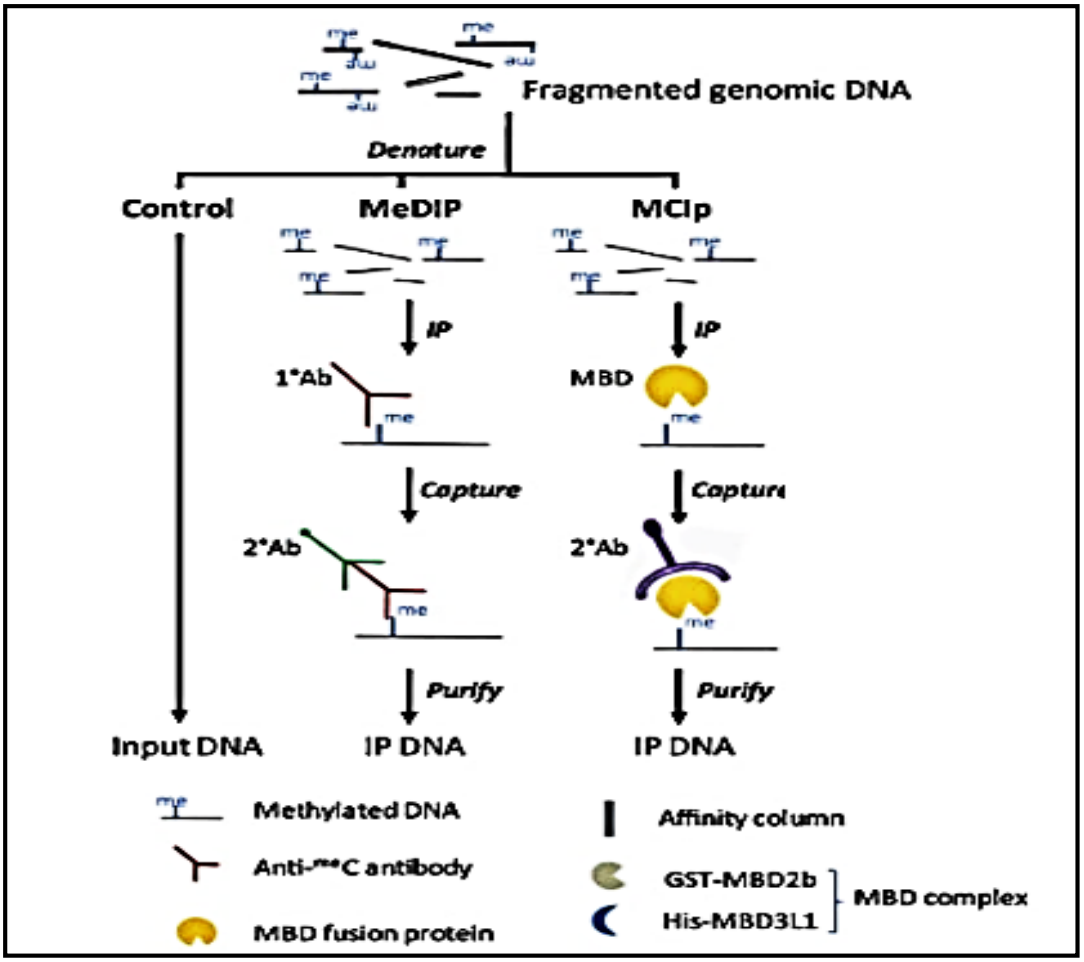


Figure 2.3: Diagrammatic representation of immunoprecipitation techniques. Genomic DNA is first fragmented by sonication and then undergoes denaturation. The DNA is then separated into two fractions, namely the input DNA and immunoprecipitated DNA(IP). Primary antibody (anti-methylcytosine) and methyl-CpG binding protein (MBD) fusion protein are incubated with single-stranded DNA in MeDIP and MCIp, respectively. The methylated DNA–antibody complex is captured by a secondary antibody that is specific to the primary antibody (Thu *et al*., 2010).

2.11.2 Endonuclease digestion

Restriction enzymes can be used to provide information on methylation patterns. Each sequence specific restriction enzyme that is used in the endonuclease digestion method has a DNA methyl transferase that prevents DNA from being restricted at the recognition site of the enzyme by methylation of the bases in the restriction site (Laird, 2010). There are some restriction enzymes that are sensitive and are unable to cleave a DNA sequence when the recognition site lies at a methylated CpG. Thus, endonuclease digestion method is used to determine DNA methylation patterns (Gršković *et al*., 2013). There are also enzymes that are not sensitive to methylation and will cleave methylated DNA (Bird, 1986). This method

depends on the presence of a recognition site at a specific CpG, thus, the enzyme used in this method will only cleave at the CpG which is not methylated.

2.11.2.1 Methylation sensitive restriction enzyme

The most widely used endonuclease digestion method is methylation sensitive restriction enzyme-polymerase chain reaction (MSRE-PCR) (Melnikov *et al*., 2005). In this method, a restriction enzyme such as *HpaII* is used, because it has an isoschizomer or neoschizomer that is not inhibited by methylated CpGs (Laird, 2010). Thus, this enzyme leaves the methylated sites intact (Gršković *et al*., 2013; Hua *et al*., 2011; Sijen, 2015). The digested DNA is subjected to PCR analysis using primers that are fluorescently labelled followed by the amplicons being separated using capillary electrophoresis. To confirm that all unmethylated DNA is cleaved, the extracted DNA undergoes extensive restriction enzyme digestion. It is not possible to determine the amount of methylation that is present at these CpG sites using this method, however, it does give information on the presence or absence of methylation. This is a simple and robust method, however, the analysis and interpretation is hindered by differences in the enzyme activity or in case of incomplete digestion. The major limitation that is associated with this method is the dependency on specific recognition sites at the CpG of interest for the restriction enzymes that are used (Melnikov *et al*., 2005).

2.11.2.2 Restriction Landmark Genomic Scanning (RLGS)

Restriction Landmark Genomic Scanning (RLGS) is a quantitative method that is used to determine the methylation status of CpG islands in a genome (Costello *et al*., 2000). This method involves two-dimensional separation of radioactively labelled genomic DNA. The extracted genomic DNA is digested with methylation sensitive restriction enzymes such as *NotI*, which cleaves at methylated regions usually in CpG islands (Costello *et al*., 2002; Laird, 2010). The digested DNA is radioactively labelled at the cleaved ends and is followed by a second digestion with secondary enzymes, like, *EcoRV* (Dhingra *et al*., 2014). Once the second digestion is completed the digested DNA is electrophoresed in an agarose tube gel, where, the DNA is digested by a third restriction enzyme which is followed by electrophoresis in a perpendicular direction to the first, separation on a polyacrylamide gel and then auto-radiographed (Costello *et al*., 2002). The RLGS profile that is obtained shows the copy number and the methylation status of the CpG islands. Although this method is sensitive it uses a large amount of DNA, is time consuming and labour intensive (Dhingra *et al*., 2014; Laird, 2010).

2.11.3 Chemical Modification

Bisulfite treatment or conversion is a chemical modification of cytosine residues developed by Frommer *et al*. (1992). It is an efficient approach that provides qualitative and quantitative results for DNA methylation analysis. In this method, the extracted DNA is treated with high concentrations of sodium bisulfite (Figure 2.4) in which cytosine nucleotides that are unmethylated are converted to uracil and methylated cytosine remain unchanged in genomic DNA sequences. Several studies have used sodium bisulfite treatment to determine DNA methylation profiles based on CpG dinucleotides that are methylated or unmethylated (Brena *et al*., 2006; Kristensen., *et al* 2008; Yang *et al*., 2006). Bisulfite conversion is generally followed by PCR and analysis of the PCR products which enables one to distinguish between the methylated and unmethylated DNA (Shen and Waterland, 2007; Leontiou *et al*., 2015). When the gene of interest is amplified during PCR amplification the PCR product will result in a thymine being present for an unmethylated cytosine and a cytosine being present at a methylated cytosine location. For my study the focus will be on techniques that use chemical modification.

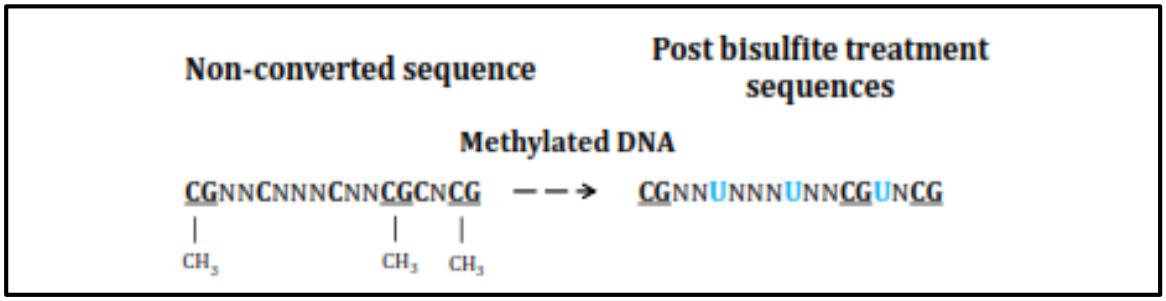


Figure 2.4: Schematic representation of bisulfite conversion. In methylated DNA, non-methylated cytosine is converted to uracil (blue U) and methylated cytosine remain the same (underlined CG) after bisulfite treatment (Hernández *et al*., 2013).

2.11.3.1 Methylation-sensitive single nucleotide primer extension (Ms-SNuPE)

The methylation-sensitive single nucleotide primer extension (Ms-SNuPE) is a quantitative and fast technique to determine differences in methylation patterns at specific CpG sites on DNA that has undergone bisulfite conversion which is followed by single nucleotide primer extension (Figure 2.5) (Gonzalgo and Jones, 1997; Gonzalgo and Liang, 2007; Tost *et al*., 2003). Specific primers bind upstream from the nucleotide position of interest, following incorporation of a labelled nucleoside triphosphate which results in termination of the reaction.

The ratio of unmethylated and methylated cytosine can be determined by running the PCR products that are generated on a denaturing polyacrylamide gel electrophoresis (Dhingra *et al*., 2014). Ms-SNuPE based methods have been demonstrated to be sensitive, fast and robust. Multiple CpG sites can be analysed in a single reaction by using a multiplex strategy to determine the quantity of methylation, therefore, the use of restriction enzymes is eliminated (Dhingra *et al*., 2014).

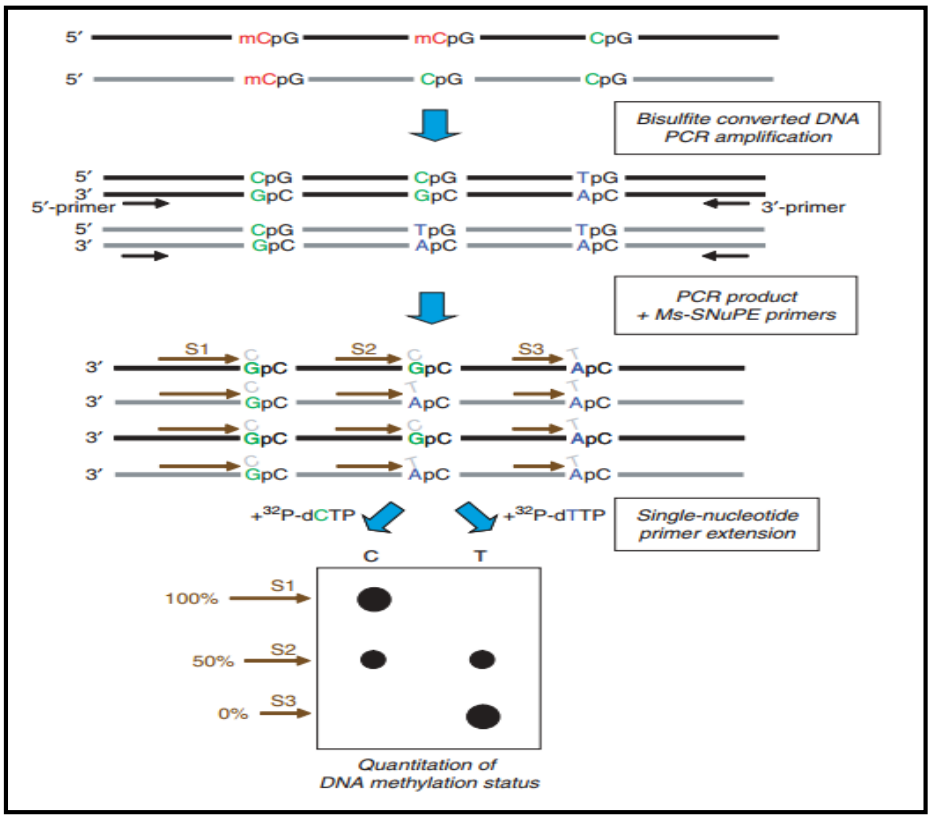


Figure 2.5: A diagrammatic representation of Ms-SNuPE. Target DNA undergoes bisulfite treatment and amplification, which is followed by the SNuPE assay. SNuPE products are then analysed by electrophoresis. Methylated cytosine (represented by red- mC) and non-methylated cytosine (represented by green C) at a CpG site. Non- methylated cytosine is reproduced as T (blue) following PCR amplification. Arrows at the 3- or 5'- represent PCR primers that amplify the bisulfite treated target DNA. Arrows at S1, S2 or S3 are representative the oligonucleotide primers for the Ms-SNuPE analysis (Gonzalgo and Liang, 2007).

2.11.3.2 Bisulfite sequencing (BS)

The gold standard method for analysis of methylation patterns post bisulfite conversion is bisulfite sequencing (BS) (Frommer *et al*., 1992). This method has improved DNA methylation research in which the presence and absence of methylation is identified. The converted DNA is amplified at a specific region using PCR primers that do not overlap at CpG sites resulting in amplification of methylated and unmethylated alleles. The primers that are designed for BS

should not have any CpG sites within the sequence to avoid discrimination between methylated or unmethylated DNA and the primer pair should span the maximum number of CpG sites (Li and Dahiya, 2002). The resulting PCR amplicons can be analysed in two ways: using direct sequencing or clone based sequencing (Hernández *et al*., 2013). In the direct sequencing method, the PCR amplicons are sequenced using a method such as Sanger sequencing which is followed by analysis of methylation patterns. This method is very effective in determining the amount of methylation at each CpG site and also determines the effect of DNA methylation on the expression of the gene of interest (Parrish *et al*., 2012). Although the direct sequencing method is not time consuming it is unable to distinguish methylation patterns between 5-Methylcytosine and 5-Hydroxymethylcytosine (Parrish *et al*., 2012). The products that are sent for direct sequencing may produce unambiguous results and is more difficult to analyse compared to clone based sequencing. In the cloned based sequencing method (Figure 2.6) the PCR amplicons are ligated into a cloning vector followed by transformation into competent cell (Shen and Waterland, 2007). Agar plates are then used to grow the antibiotic resistant colonies which are then selected individually followed by growth in Luria-Bertani (LB) medium. The plasmid DNA is then isolated and sequenced. The bisulfite sequencing method can be quantitative if a large number of clones are sequenced. A single allele is represented by each clone, thus, providing allele-specific methylation information. This method is widely used in studies for site specific methylation or allele-specific methylation due to the fact that it is straightforward and generates detailed methylation information. However, this method is expensive if there is a large sample size and is very labour intensive (Shen and Waterland, 2007). After bisulfite treatment the DNA is prone to degradation which affects the detection limit and the complementary strand is lost as a result of cytosine being converted to uracil (Dhingra *et al*., 2014).

2.11.3.3 Methylation Specific Polymerase Chain Reaction (MSP)

Herman *et al*. (1996) was first to use the methylation specific polymerase chain reaction (MSP) method. MSP is a sensitive and fast method used to determine methylation patterns by amplifying the methylated and unmethylated alleles at a region of interest (Figure 2.8) (Herman *et al*., 1996; Hernández *et al*., 2013). Following bisulfite conversion, a PCR is carried out using two sets of primers that are designed using software such as Methprimer (Li and Dahiya, 2002). The two sets of primers either amplify methylated or unmethylated alleles (Shen and Waterland, 2007). MSP is a beneficial method as it is very sensitive and can detect one methylated allele in more than one thousand unmethylated alleles (Shen and Waterland, 2007).

This technique can also be used on DNA samples that have a low quantity and quality as well as paraffin embedded samples (Herman *et al*., 1996; Dhingra *et al*., 2014). Primer design is crucial in MSP to ensure that reliable results are produced (Figure 2.7). The MSP method is able to recognise CpG sites without using methylation sensitive restriction enzymes, thus, the number of CpG sites that can be analysed is increased (Herman *et al*., 1996). In methods that use methylation sensitive enzymes false positive results may be produced due to partial digestion, however, MSP overcomes this limitation. The BS method is time consuming due to cloning and sequencing can take days to complete, whereas, MSP can be done within 1-2 days (Ohashi, 2002).

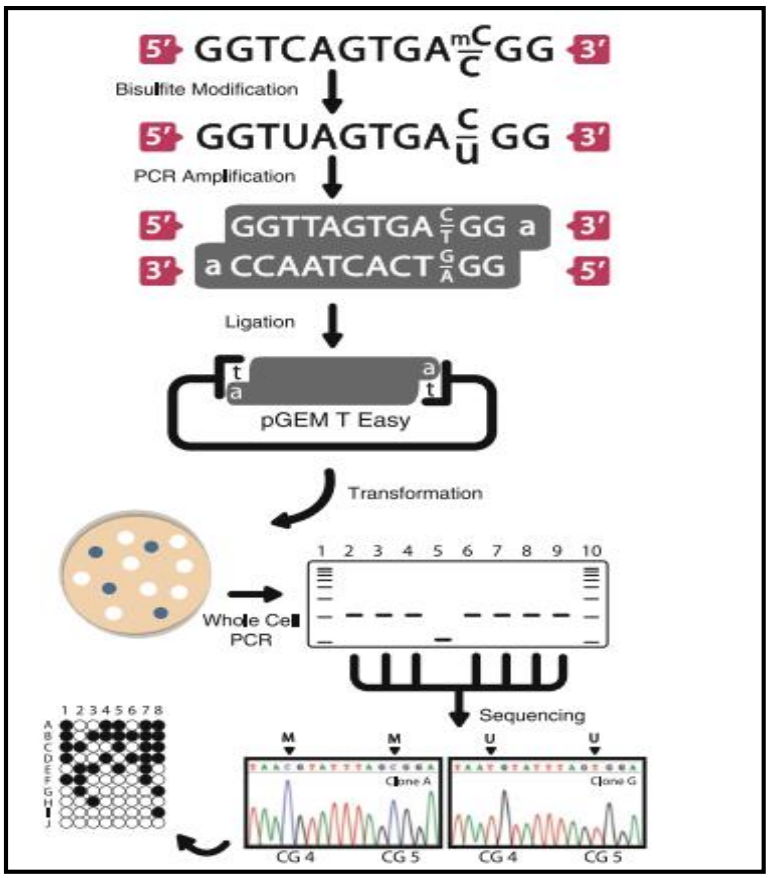


Figure 2.6: A diagrammatic representation of cloned based bisulfite sequencing. Genomic DNA is bisulfite treated, followed by PCR amplification. PCR products are ligated into a vector. The ligated DNA product is then transformed into competent bacterial cells and grown in non-selective liquid culture. A selective media is used to plate transformed cells which are grown overnight. Blue/white selection method is used for selection of (white) colonies which contain the PCR amplicons as inserts. A PCR reaction followed by agarose gel electrophoresis is used to determine if selected positive colonies have the correct insert size. Positive amplicons are purified and sequenced. Sequenced results are used to establish methylation status. The rows represent a single clone and the columns individual CpG dinucleotides within the sequenced region of interest. Black dots represent methylated CpG sites and white dots represent unmethylated CpG sites (Huang *et al*., 2013a).

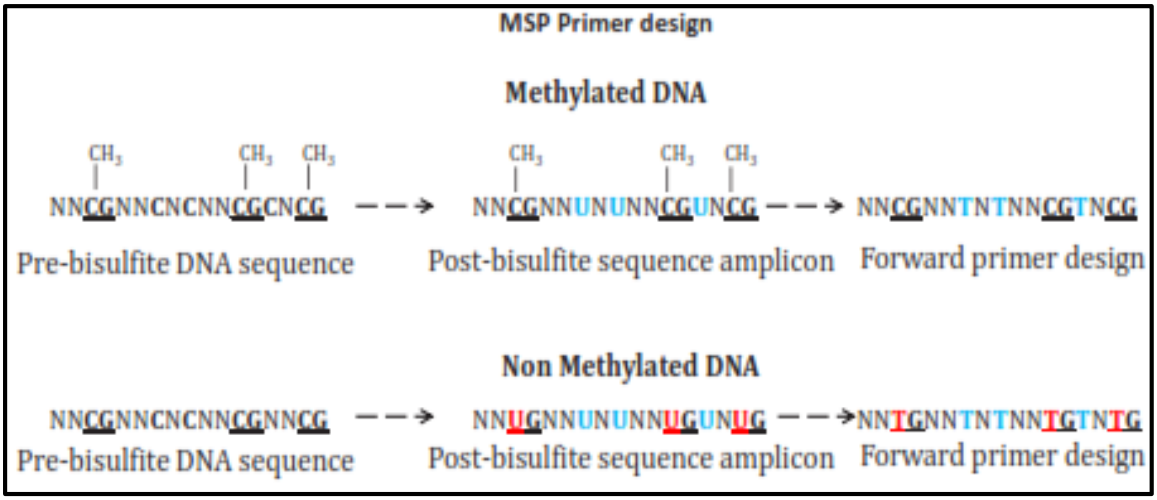


Figure 2.7: MSP primer design for methylated and unmethylated DNA. The DNA sequence (left) is modified based on the methylation status. The modified sequence is used to design the forward primer by replacing Us with Ts (Hernández *et al.*, 2013).

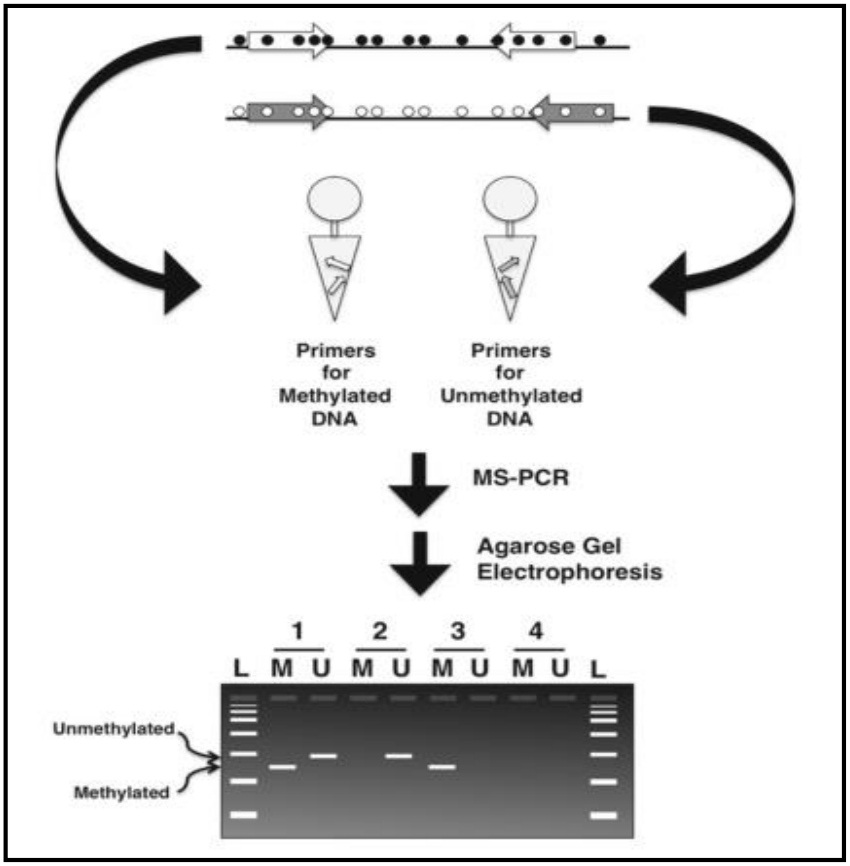


Figure 2.8: A diagrammatic representation of methylation specific polymerase chain reaction. In MSP, two sets of primer pairs are designed to target the methylated cytosine (black circles) and non-methylated cytosine (white circles) of the bisulfite converted genomic DNA. After PCR amplification, the amplicons are analysed by agarose gel electrophoresis. M represents methylated, U represents non-methylated and L represents the DNA ladder (Huang *et al.*, 2013b).

Combining the above mentioned techniques with DNA microarrays or high throughput sequencing enables genome-wide profiling of DNA methylation (Bock, 2012; Zilberman and Henikoff, 2007). The design and technology that is used to develop microarrays influences their applicability in methylation analysis, for example, Illumina arrays are best suited for DNA that has been bisulfite converted and other arrays such as short oligonucleotide arrays and long oligonucleotide arrays are better suited for restriction enzyme and affinity based techniques. In the Illumina arrays, two primers (one primer binds to methylated cytosine and the other primer binds to a converted sequence) that are fluorescently labelled are used in a PCR reaction with the bisulfite converted DNA followed by the Illumina assay (Bibikova *et al*., 2006; Fan *et al*., 2006). The ratio of the PCR products are established using Illuminas SEntrix Array Matrix bead array platform which can assay up to 1536 sites in 96 samples in a single experiment. This method is able to provide quantitative evaluation of specific cytosines in many samples, thus, it is suitable to compare a set of known methylated loci in many cell lines to determine the methylation profiles (Yousefi *et al*., 2013). However, the method has low coverage in comparison to other array based method. Development of a large set of selective primers is necessary, thus, limiting its usefulness for *de novo* genomes (Zilberman and Henikoff, 2007). The short oligonucleotide arrays use photolithographic technology in which millions of probes per chip are used to achieve very high feature density (Dalma-Weiszhausz *et al*., 2006). Short probes that are made up of 25 nucleotides and the chips are designed so that it only hybridizes to one sample at a time. Arrays have an excellent specificity, however, it has low sensitivity and high random signal variation due to the length of the probes (Kreil *et al*., 2006). To ensure statistical significance each sample is hybridized in triplicate. An adaptive photolithographic method is employed to produce the long oligonucleotide arrays (Nuwaysir *et al*., 2002). The probes consist of 60 oligonucleotides and have a dual channel in which two samples can be labelled with different fluorescent dyes and hybridized to a single chip. This is advantageous because it reduces the need for replicates. The longer probes provide a good balance between specificity, sensitivity and noise (Kreil *et al*., 2006).

DNA methylation can also be analysed by a sequencing-by-synthesis fluorescent nucleotide based system that can produce approximately 400 000 reads of over 100 bases in a single run (Madi *et al*., 2012; Park *et al*., 2014). Pyrosequencing produces great amounts of sequencing information, quicker and at a more cost effective rate than conventional Sanger sequencing, which eliminates the cloning steps. This technique can be used as an alternative to DNA microarrays. High throughput sequencing provides a quantitative measure of methylation

(Kottaridi *et al*., 2015). In comparison to array based techniques which provide a relative measure, biases that affect hybridization such as sequence composition do not have an effect on high throughput sequencing. However, a high quality reference sequence is required because the short reads that are produced by this technique is difficult to assemble *de novo*. These techniques are very informative, however, are also expensive and labor intensive and most laboratories are not capable of performing these techniques (Zilberman and Henikoff, 2007).

2.12 Rationale of the study

Determination of the type and origin of body fluids can provide important clues for crime scene reconstruction by linking donors and actual crime acts. Presumptive and confirmatory methods used for identification of body fluids include, catalytic, immunological and enzymatic tests. These methods have an increased risk of cross contamination and possess low sensitivity as well as low specificity (Hernández *et al*., 2013). In addition, there is no single method for all body fluids which results in extra sample consumption and makes samples incompatible with subsequent DNA STR typing. Approaches based on RNA expression offers high specificity however, the stability of RNA molecules is a problem (Akutsu *et al*., 2015; Vidaki *et al*., 2016). Recent studies have demonstrated that human tissues and forensically relevant body fluids can be identified based on their tissue-specific differential methylation pattern. In human genomes, tDMRs display varying methylation profiles according to cell or tissue type. DNA methylation analysis is compatible with STR analysis as DNA extracted for STR typing can also be used for methylation analysis (Watanabe *et al*., 2016). tDMRs show a great potential for identification of human body fluids in forensic investigations, however limited tDMRs have been verified so far (An *et al*., 2012; Choi *et al.,* 2014; Fu *et al*., 2015). At the University of KwaZulu-Natal, Kader. (2015, MSc Genetics dissertation) identified novel candidate gene based tDMRs for semen, saliva, blood and vaginal fluid identification based on differential gene expression in body fluids and surrogate tissues of body fluids. The present study aims to validate the methylation profile of candidate tDMRs and assess their suitability by testing their stability under simulated forensic conditions. The validated tDMR markers will add to the growing list of epigenetic markers for body fluid identification as the candidate tDMRs targeted in present study have not been reported previously.

2.13 Genes of interest studied in the present research

2.13.1 *ZNF282* gene (*Zinc finger protein 282*)

The *ZNF282* gene is also known as *HUB1*. It is a protein coding gene that contains nine exons and is located on chromosome seven on the long arm (q) on band three, sub-band six, sub-sub-band one (7q36.1) (Figure 2.9) (https://www.ncbi.nlm.nih.gov/gene/8427; ZNF282, 2004). The genomic size of the *ZNF282* gene is 30 787 bases and the protein size is 671 amino acids with a molecular mass of 74 295 Da (http://www.genecards.org/cgi-bin/carddisp.pl?gene=znf282; Fishilevich *et al*., 2016). The proteins that are encoded by the *ZNF282* gene belong to the krueppel C2H2-type zinc finger protein family (Okumura *et al*., 1997). The proteins found in *ZNF282* gene bind to the U5 repressive element (U5RE) in the T-cells leukemia virus type I long terminal repeat in humans (Yeo *et al*., 2014). A study by Yu *et al*. (2013) showed that *ZNF282* gene functions as an Era co-activator by an interaction with oestrogen receptor in breast cancer cells. There are two transcript variants that have been reported for *ZNF282* gene. The proteins recognize a sequence (5'-TCCACCCC-3') as its core motif and exerts a strong repressive effect on HTLV-I LTR-mediated expression. Other known functions include: DNA and zinc-ion binding and regulation of transcription (http://www.uniprot.org/uniprot /Q9UDV7; Breuza *et al*., 2016). According to the Genotype-Tissue Expression Consortium (GTEx), this gene is highly expressed in the prostate showing 18.1 RPKM (reads per kilobase of transcripts per million mapped reads) based on the RNA expression. In the vagina and cervix, the GTEx showed 10.1 RPKM and 11.5 RPKM respectively for RNA expression. In the salivary glands the GTEx showed 9.7 RPKM for RNA expression (https://www.proteinatlas.org/ENSG00000170265-ZNF282/tissue; Uhlén *et al*., 2015). There was no available expression data on peripheral blood. Based on the tissue-specific (Tissue-specific Gene Expression and Regulation) gene expression profiles of the TiGER database, ZNF282 is highly expressed in the cervix (http://bioinfo.wilmer.jhu.edu/tiger/db_gene /ZNF282- index.html; Liu *et al*., 2008).

Figure 2.9: Chromosomal location of the *ZNF282* gene.

2.13.2 *PTPRS* gene (*Protein tyrosine phosphatase, receptor S*)

The *PTPRS* gene is also known as *R-PTP-S, PTPSIGMA AND R-PTP-sigma*. It is a protein coding gene that contains 41 exons and is located on chromosome 19 on the short arm (p) on band one, sub-band three and sub-sub-band three (19p13.3) (Figure 2.10) **(**https://www.ncbi.nlm.nih.gov/gene/5802; PTPRS, 2004). The genomic size of the *PTPRS* gene is 182 309 bases and has a protein size of 1 948 amino acids with a molecular mass of 217 041 Da (http://www.genecards. org/cgi-bin/carddisp.pl?gene=ptprs; Fishilevich *et al*., 2016). The proteins encoded by the *PTPRS* gene belong to the protein tyrosine phosphatase (PTP) family. PTPs are identified as signalling molecules that are involved in regulation of various cellular processes, such as cell growth, differentiation, mitotic cycle and oncogenic transformation. Other known functions of *PTPRS* gene include, interaction with LAR-interacting protein (LIP.1), transmembrane receptor protein tyrosine phosphatase activity, cell adhesion, the development of spinal cord, cerebellum, cerebral cortex and synaptic transmission (http://www.uniprot.org/uniprot /Q13332; Breuza *et al*., 2016). There are four alternatively spliced transcript variants that have been reported which encode unique proteins (PTPRS, 2004). The *PTPRS* gene is expressed in all tissue excluding placenta and liver tissue. According to the GTEx, this gene is highly expressed in the cerebellum showing an average 31.2 RPKM (reads per kilobase of transcripts per million mapped reads) based on the RNA expression. In the vagina and cervix, the GTEx showed 20.8 RPKM and 23.5RPKM respectively for RNA expression. In the prostate and salivary glands, the GTEx showed 23.1 RPKM and 10 RPKM respectively for RNA expression (https://www.proteinatlas.org /ENSG00000105426-PTPRS/tissue; Uhlén *et al*., 2015). There was no expression data available on peripheral blood. Based on the tissue-specific (Tissue-specific Gene Expression and Regulation) gene expression profiles of the TiGER database, *PTPRS* is highly expressed in the small intestine (http://bioinfo.wilmer.jhu.edu/tiger/ db_gene/PTPRS-index.html; Liu *et al*., 2008).

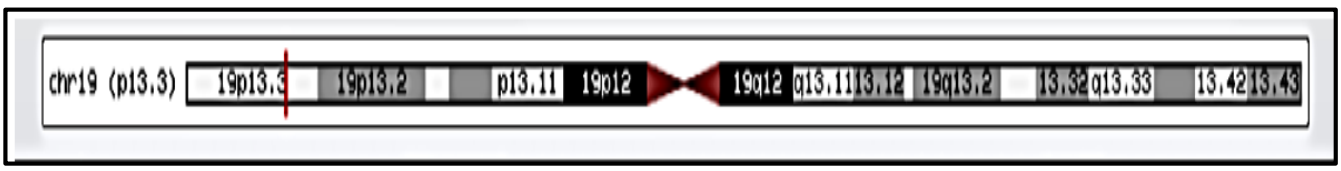


Figure 2.10: Chromosomal location of the *PTPRS* gene.

2.13.3 *HPCAL1* gene (*Hippocalcin like 1*)

The *HPCAL1* gene is also known as *calcium-binding protein (BDR1), hippocalcin-like protein 1 (HLP2)* and visinin-like protein 3 *(VILIP-3).* It is a protein coding gene that contains 17 exons and is located on chromosome two on the short arm (p) on band two, sub-band five, sub-sub-band one (2p25.1) (Figure 2.11) (https://www.ncbi.nlm.nih.gov/gene/?term=hpcal1; HPCAL1, 2004). The genomic size of the *HPCAL1* gene is 124 729 bases and has a protein size of 193 amino acids with a molecular mass of 22 313 Da (http://www.genecards.org/Search/Keyword?Query String=hpcal1; Fishilevich *et al*., 2016). There are eight alternatively spliced transcript variants that have been reported that encode the same protein (https://www.ensembl.org/Homosapiens/Gene/Summary?db=core;g=ENSG00000115756;r=2:10302889-10427617; Aken *et al*., 2016). The proteins that are encoded by the *HPCAL1* gene belong to the neuron-specific calcium-binding proteins family that are found in the retina and brain. The *HPCAL1* gene could possibly be involved in calcium-dependent regulation of rhodopsin phosphorylation and may be relevant in neuronal signalling in the central nervous system. Other known functions include, calcium-ion binding (http://www.uniprot.org/uniprot/?query=hpcal1&sort=score; Breuza *et al*., 2016). The *HPCAL1* gene is highly expressed in the Purkinje cells of the cerebellum (https://www.proteinatlas.org/ENSG00000115756-HPCAL1/tissue; Uhlén *et al*., 2015). The expression of *HPCAL1* gene occurs mainly in the cerebellum during late development (Spilker *et al*., 2000). According to the GTEx, this gene is highly expressed in the cerebellum showing 87.3 RPKM based on the RNA expression. In the vagina and cervix, the GTEx showed 6.4 RPKM and 11 RPKM respectively for RNA expression. In the prostate and salivary glands, the GTEx showed 6.5 RPKM and 5 RPKM respectively for RNA expression (Uhlén *et al*., 2015). There was no available data on peripheral blood. If the expression of *HPCAL1* gene is increased it could result in an increase of ERK2 activation and ERK2 protein expression (Spilker and Braunewell, 2003). Based on the tissue-specific (Tissue-specific Gene Expression and Regulation) gene expression profiles of the TiGER database, ZNF282 is highly expressed in the cervix (http://bioinfo.wilmer.jhu.edu/tiger/db_ gene/HPCAL1-index.html; Liu *et al*., 2008).

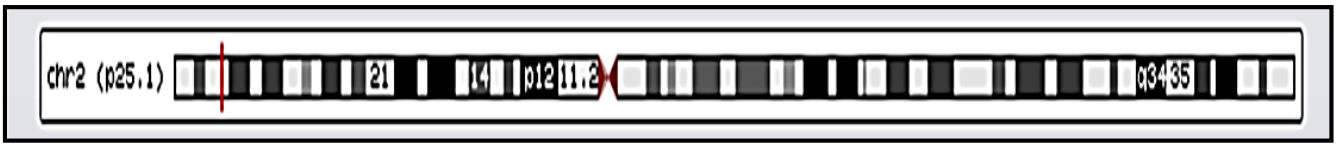


Figure 2.11: Chromosomal location of the *HPCAL1* gene.

2.14 Hypothesis, objectives and aims

2.14.1 Hypothesis:

- The candidate tDMRs for genes *ZNF282*, *PTPRS* and *HPCAL1* will be differentially methylated in human blood, saliva, semen, and vaginal fluid.
- The validated tDMRs markers will maintain a stable methylation profile under forensic simulation conditions.

2.14.2 Objectives:

- Methylation profiling of candidate tDMRs in blood, saliva, semen and vaginal fluid by Methylation Specific Polymerase Chain Reaction (MSP).
- Methylation profiling of candidate tDMRs in blood, saliva, semen and vaginal fluid by Bisulfite Sequencing (BS).
- Sensitivity testing of candidate tDMRs by Methylation Specific Polymerase Chain Reaction (MSP).
- Methylation profiling of candidate tDMR in blood, saliva, semen and vaginal fluid exposed to dry, wet and outside conditions (forensic simulation) by Bisulfite Sequencing (BS).
- Methylation profiling of candidate tDMR in blood, saliva, vaginal fluid and semen exposed to bleach and alcohol (forensic simulation) by Bisulfite Sequencing (BS).

2.14.3 Aims:

- Methylation profiling and validation of candidate tDMRs for identification of human body fluids.
- To analyse the effect of external environmental factors on methylation profile of validated tDMR markers.

2.14.4 Key Question:

- Do the candidate tDMRs present a differential methylation profile to differentiate between human blood, saliva, semen and vaginal fluid?
- Is the methylation profile of identified tDMRs markers stable under simulated forensic conditions?

Chapter 3

Validation of tissue-specific differential DNA methylation in gene-specific tDMRs by methylation specific PCR and bisulfite sequencing

Abstract

Identification of body fluids found at crime scenes is a critical step in forensic investigations. DNA methylation analysis is emerging as a reliable approach to differentiate between body fluids. Regions known as tissue-specific differential methylated regions (tDMRs) in the human genome display differential methylation patterns in various tissues or cells, hence, tDMRs can be targeted to differentiate between body fluids. To date only a few tDMRs markers have been developed for body fluid differentiation. In a previous study at the University of KwaZulu-Natal, based on differential gene expression in blood, saliva, semen and vaginal fluid candidate tDMRs in genes *ZNF282*, *PTPRS* and *HPCAL1* were identified to differentiate between blood, saliva, semen and vaginal fluid. It was proposed that differential gene expression could be possibly due to differences in methylation patterns. The present study was undertaken to examine the methylation profile of candidate tDMRs by methylation specific PCR (MSP) and bisulfite sequencing (BS) analysis. CpG islands in the gene body were targeted to design primers for both MSP and BS. The ZNF282 tDMR displayed hypomethylation in semen and hypermethylation in blood, saliva and vaginal fluid with both MSP and BS analysis. Statistical analysis by Fisher's exact test showed significant difference in overall methylation status and at individual CpG sites between semen and all other body fluids ($p < 0.05$) for the ZNF282 tDMR. Therefore, the ZNF282 tDMR could be successfully used as a semen-specific hypomethylation marker. The PTPRS 1 tDMR displayed hypermethylation for blood, vaginal fluid and saliva. In semen for PTPRS 1 tDMR both methylation and unmethylation was observed. Therefore, PTPRS 1 tDMR has the potential to be used as a semen-specific marker. However, BS analysis could not differentiate semen from other body fluids due to a fewer number of clones that were analysed. Both MSP and BS analysis showed no significant difference in overall methylation pattern and at individual CpG sites between body fluid for PTPRS 2 tDMR ($p > 0.05$) thus PTPRS 2 cannot be used to differentiate body fluids. The HPCAL1 tDMR showed hypomethylation in saliva and hypermethylation in other body fluids by MSP, however, the results could not be validated by BS due to non-specific amplification with bisulfite PCR primers. The present study reports methylation profiling and validation of novel gene based tDMRs which will add to the growing list of methylation markers for identification of forensically relevant body fluids. Future research would aim to re-analyse PTPRS 1 and HPCAL1 tDMR on a larger sample size for significant methylation differences.

3.1 Introduction

In forensic investigations, human body fluids are amidst the most significant types of evidence that are found at crime scenes (Greenfield and Sloan, 2003). Analysis of body fluids present at a crime scene have the potential to influence the outcome of the case by assisting in identification of victims and perpetrators, reconstruction of the crime scene events and can provide a link between the evidence and the crime committed (Akutsu *et al*., 2015; Virkler and Lednev, 2009; Lee *et al*., 2012). Body fluids are not always visible to the naked eye and are sometimes indistinguishable from other fluids or substances.

Presumptive and confirmatory tests are used to identify body fluids found at crime scenes (An *et al*., 2012; Frumkin *et al*., 2011; Gršković *et al*., 2013; Sinelnikov *et al*., 2013). Presumptive tests are used as screening tests to identify specific proteins, however, these tests are limited in their specificity, sensitivity, proteins are susceptible to degradation and inactivity (Casey and Price, 2010; Frumkin *et al*., 2011; Lin *et al*., 2016; Shaler, 1981; Virkler and Lednev, 2009). Confirmatory tests are used for absolute identification of body fluids or tissue. The presumptive and confirmatory tests commonly used are summarized in Table 3.1.

Majority of these tests rely on colour changing reactions which are difficult to interpret when coloured extracts or samples are required to be identified (Haas *et al*., 2009; Lindenbergh *et al*., 2012). Present tests are destructive for specific constituents that are found in the various body fluids. Body fluids that are found at crime scenes need to be identified accurately without destroying or degrading the biological sample, thus a technique that is able to do this as well as use a small amount of sample will prove to be advantageous.

Table 3.1: Current methods used for body fluid identification (Gaensslen, 1983; Greenfield and Sloan, 2003; Shaler, 2002; Spalding, 2003; Virkler and Lednev, 2009).

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Presumptive tests	Alternate light source (ALS), luminol test, benzidine test and Kastle-Meyer test	Alternate light source (ALS -based on the activity of amylase)	Seminal acid phosphatase (SAP) and Florence test	Periodic Acid-Schiff (PAS-based on glycogenated epithelial cells)
Confirmatory tests	Microscope tests or chromatographic technique	None	Microscopic identification and prostate-specific antigen (PSA)	None

Although body fluids and tissue can be identified using mRNA profiling (Bauer and Patzelt, 2002; Juusola and Ballantyne, 2003; Lindenbergh *et al*., 2012; Zubakov *et al*., 2008; Zubakov *et al*., 2010), RNA has a shorter half-life and is less stable than DNA (An *et al*., 2012; Phang *et al*., 1994). MicroRNAs (miRNAs) have a greater stability and longer half-life compared to mRNA and has better characteristics for identifying forensic biological material (Bail *et al*., 2010; Hanson *et al*., 2009; Sato *et al*., 2009; Zubakov *et al*., 2010). Identification of body fluids using mRNA and miRNA provides valuable information however these methods require a large amount of sample, are labour intensive and time consuming (Fleming and Harbison, 2010; Zubakov *et al*., 2010).

Analysis of differential DNA methylation patterns in cells and tissue is a promising method for identification of forensically relevant body fluids because DNA is more stable than RNA and proteins (An *et al*., 2013; Choi *et al*., 2014; Gršković *et al*., 2013; Lee *et al*., 2012; Lee *et al*., 2015; Lee *et al*., 2016a; Li *et al*., 2012; Silva *et al*., 2016; Park *et al*., 2016). DNA methylation analyses is beneficial because it has higher specificity, can be used in conjunction with the current forensic DNA protocols, thus, has the ability to be used in old cases in which only the DNA extracts are available (An *et al*., 2012).

DNA methylation occurs in CpG dinucleotides which are enriched in DNA regions known as CpG islands (Bird, 1986; Feltus *et al*., 2003). CpG islands are defined as: having approximately 200 base pairs, a GC content greater than 50% and the observed/expected CpG ratio must be greater than 60% (Derks *et al*., 2004; Gardiner-Garden and Frommer, 1987). CpG islands are associated with gene promoters and gene bodies (Deaton and Bird, 2011; Maunakea *et al*., 2010). Gene body methylation has a positive correlation with gene expression, conversely, methylation at promoter regions has an inverse correlation (Ball *et al*., 2009; Goll and Bestor, 2005; Klose and Bird, 2006; Laurent *et al*., 2010; Lister *et al*., 2009). CpG islands extend downstream of transcript regions which explains the high CpG density at the exon 5' end of the transcript flanked with the transcription start site (Choi, 2010). Genome-wide studies have shown that exons are more highly methylated in comparison to introns as well as transitions in the methylation status that occur at the exon-intron boundaries, this suggests that the methylation status plays a role in regulation of splicing (Laurent *et al*., 2010). Therefore, DNA methylation has an alternate function in gene bodies (Jones, 2012).

There are several regions in the human genome that display different methylation patterns depending on the cell or tissue type known as tissue-specific differentially methylated regions (tDMRs) (Kitamura *et al*., 2007; Slieker *et al*., 2013; Song *et al*., 2009). tDMRs are broadly distributed in intragenic and intergenic regions which include both CpG islands, and non-CpG islands regions (Song *et al*., 2009) as well as CpG shores (Irizarry *et al.,* 2009). Since methylation patterns of tDMRs vary between tissues and body fluids, tDMRs have the potential to be used as biomarkers to differentiate between human body fluids based on the methylation patterns (Ghosh *et al*., 2010; Lokk *et al*., 2014).

Recent studies have identified tDMRs either within a gene or at other genomic locations. Genome-wide methylation studies have also identified differentially methylated CpG sites which can distinguish between body fluids. A study by Lee *et al*. (2012) used five tDMRs with multiple CpG sites within genes *DACT1, USP49, HOXA4, PFN3,* and *PRMT2* for differentiation of body fluids by bisulfite sequencing (BS). Semen-specific hypomethylation patterns were observed for DACT1 and USP49. The HOXA4, PFN3, and PRMT2 tDMRs showed varying degrees of methylation in each body fluid. Other studies have also identified tDMRs for body fluid identification, BCAS4 (saliva marker), FGF7 and ZC3H12D (semen markers), PFN3A (vaginal epithelia marker) all have multiple CpG sites and C20orf117 (blood marker) and cg0969411 (menstrual blood marker) show differential methylation at a single CpG site (Antunes *et al.,* 2016b; Forat *et al*., 2016; Madi *et al*., 2012). Table 3.2 to 3.6 depicts tDMRs and CpG sites reported to date for body fluid identification.

Sodium bisulfite conversion methods are widely used to distinguish between methylated and unmethylated cytosine residues (Clark *et al*., 1994; Derks *et al*., 2004; Frommer *et al*., 1992; Herman *et al*., 1996). In the conversion reaction, DNA is treated with sodium bisulfite which results in unmethylated cytosine residues being converted to uracil and methylated cytosine residues remain unaffected in single-stranded DNA (Sasaki *et al*., 2003). The methylation level can be determined using polymerase chain reaction (PCR) based techniques such as, methylation specific PCR (MSP) and bisulfite sequencing (BS) which was used in the present study.

Table 3.2: Methylation markers for identification of blood reported to date.

tDMR/CpG site (cg represents a single CpG site and chr represents a chromosomal location)	Status	Proposed fluid of specificity	Reference
cg06379435	High	Blood	Lee *et al*., 2016a
cg08792630			
cg06379435	High	Blood	Watanabe *et al*., 2016
cg26285698	Low	Blood	Forat *et al*., 2016
cg03363565	High		
chr22:37705345- 37705493	Low	Blood	Fu *et al*., 2015
chr22:29708295- 29708444	High	Blood	
cg08792630	High	Blood	Park *et al*., 2014
cg06379435			
cg241244443	High	Blood	Lin *et al*., 2016
cg01607849			
C20orf117	High	Blood	Madi *et al*., 2012; Eckhardt *et al*., 2006
cg06379435	High	Blood	Lee *et al*., 2015
cg01543184			

Table 3.3: Methylation markers for identification of menstrual blood reported to date.

tDMR/CpG site (cg represents a single CpG site and chr represents a chromosomal location)	Status	Proposed fluid of specificity	Reference
cg18069290	High	Menstrual blood	Lee *et al*., 2016a
cg09696411			
cg0969411	High	Menstrual blood	Forat *et al*., 2016
chr6:4160458-41604650	Low	Menstrual blood	Fu *et al*., 2015

Table 3.4: Methylation markers for identification of saliva reported to date.

tDMR/CpG site (cg represents a single CpG site)	Status	Proposed fluid of specificity	Reference
cg09652652-2d	High	Saliva	Lee *et al*., 2016a
cg21597595	High	Saliva	Forat *et al*., 2016
cg15227982			
cg26107890	High	Saliva	Park *et al*., 2014
cg20691722			
BCAS4 tDMR	High	Saliva	Madi *et al*., 2012
cg09107912	High	Saliva	Lin *et al*., 2016
cg16732616			
cg09652652	High	Saliva	Lee *et al*., 2015

Table 3.5: Methylation markers for identification of vaginal fluid reported to date.

tDMR/CpG site (cg represents a single CpG site)	Status	Proposed fluid of specificity	Reference
cg09765089-231d	High	Vaginal fluid	Lee *et al.,* 2016a
cg26079753-7d			
cg14991487	High	Vaginal fluid	Forat *et al*., 2016
cg03874199			
cg14991487	High	Vaginal fluid	Park *et al*., 2014
cg01774894			
PFN3 tDMR	Low	Vaginal Fluid	An *et al*., 2013; Choi *et al.,* 2014
cg25416153	High	Vaginal fluid	Lin *et al.,* 2016
cg09765089			
cg09765089	High	Vaginal fluid	Lee *et al.,* 2015
cg26079753			

Table 3.6: Methylation markers for identification of semen reported to date.

tDMR (tDMR with cg represents a single CpG site)	Status	Proposed fluid of specificity	Reference
cg17610929	High	Semen	Lee *et al.*, 2016a
cg26763282-138d			
cg04382920	Low	Semen	Vidaki *et al*., 2016
cg11768416			
cg22407458	Low	Semen	Forat *et al*., 2016
cg05656364	High		
cg17610929	High	Semen	Park *et al*., 2014
cg23521140			
DACT1 tDMR	Low	Semen	An *et al*., 2013; Choi *et al.*, 2014; Lee *et al.*, 2012
USP49 tDMR	Low	Semen	An *et al.*, 2013; Choi *et al*., 2014; Lee *et al.*, 2012
L81528 tDMR	High	Semen	Choi *et al.*, 2014; Frumkin *et al.*, 2011
PRMT2 tDMR	Low	Semen	Lee *et al*., 2012
$\frac{L91762}{L68346}$ Methylation Ratio	Low	Semen	Frumkin *et al*., 2011
ZC3H12D tDMR	Low	Semen	Madi *et al*., 2012
cg05261336	High	Semen	Lin *et al*., 2016
cg17610929			
cg17610929	High	Semen	Lee *et al*., 2015
cg26763284			
cg17621389			

BS is the gold standard method used for analysis of DNA methylation (Li and Tollefsbol, 2011; Lee *et al*., 2012; Hernández *et al*., 2013; Pabinger *et al*., 2016) because primer design for bisulfite PCR is simple and it provides detailed quantitative methylation information for the region of interest by determining methylation analysis of individual CpG sites (Huang *et al*., 2013a). The bisulfite PCR products can be analysed by direct sequencing or by sub-cloning into plasmids and selecting random clones for sequencing to obtain the methylation status of CpG sites within DNA molecules (Chen *et al*., 2010; Chhibber and Schroeder, 2008; Hernández *et al*., 2013; Huang *et al*., 2013a; Krueger *et al*., 2012). Direct sequencing is effective in determining the amount of methylation at individual CpG sites and the effect of DNA methylation on the expression of the gene of interest (Kristensen and Hansen, 2009; Parrish *et al*., 2012). In this study, the focus will be on BS of clones as individual clones are able to provide quantitative information. Once the clones are sequenced, the sequencing results can be analysed using online software to determine the methylation status. BS is the only method other than high throughput bisulfite sequencing methods that provides allele-specific methylation status (Huang *et al*., 2013a).

MSP is a qualitative technique that is used to detect the presence or absence of methylation (Hernández *et al*., 2013; Huang *et al*., 2013b; Zhao and Bapat, 2016). MSP employs two different sets of methylation specific primers for amplification of the sequence of interest. One primer set is known as the unmethylated primers which amplify sodium bisulfite converted DNA in the unmethylated state and the second primer set known as methylated primers which amplify sodium bisulfite converted DNA in the methylated state (Derks *et al*., 2004; Liu and Maekawa, 2003). MSP is a sensitive technique as it is able to detect one methylated allele in a background of thousands of unmethylated alleles. The method is also inexpensive and analyses the methylation status of each CpG site present in a CpG island, thereby making it efficient and applicable for high-throughput analysis for clinical samples (Fraga and Esteller, 2002; Herman *et al*., 1996). MSP has been used successfully in previous research as shown by Esteller *et al.* (1999) and Herman *et al*. (1996) for identification of aberrant methylation in squamous cell lung carcinoma and detection of hypermethylation patterns in tumor suppressor genes, respectively. Other studies have also used MSP to determine methylation analysis for genes that have not been previously reported and diseases (Draht *et al*., 2016; Lan *et al*., 2014; Okazaki *et al*., 2011).

With the recent interest in DNA methylation profiling for body fluid identification, novel tDMRs need to be identified to develop a large database of methylation markers. A previous study by Kader. (2015, MSc Genetics dissertation) at the University of KwaZulu-Natal, identified candidate tDMRs in genes, *ZNF282, PTPRS* and *HPCAL1* based on differential gene expression. Genes which were differentially expressed in at least one of the body fluids were selected. CpG islands within gene bodies were identified and targeted for primer design for validation studies. Thus, the present study centres mainly on the validation of the candidate tDMRs in four forensically relevant body fluids, namely, blood, saliva, semen and vaginal fluid. The objective was to design primers for MSP and BS to obtain methylation profiles using developed primers and determine whether the body fluids display differential DNA methylation profiles. Methylation differences among the four human body fluids would lead to development of additional tDMRs for body fluid identification.

3.2 Materials and methods

3.2.1 Ethics approval

The study was conducted according to the methods specified by the Biomedical Research Ethics Committee (BREC) at the University of KwaZulu-Natal. The BREC reference number is BE187/16 (sub-study of BE221/14) (Appendix A). Samples were collected from volunteers after a written consent was signed and a questionnaire answered by each participant (Appendix A).

3.2.2 Sample collection, storage of samples and DNA extraction

Whole blood was collected in 4 ml EDTA tubes from 10 volunteers at Lab 24 laboratory, Mount Edgecombe, Durban. Blood samples were collected and transported to the Genetics Laboratory in ice boxes with frozen ice packs. Vaginal fluid was collected from 10 volunteers at the King Dinizulu Hospital in Durban using sterile cotton swabs (Dry Swab, Lasec). Semen and saliva samples were collected from 10 individuals. Freshly ejaculated semen was collected in plastic cups from volunteers at the King Dinizulu Hospital and University of KwaZulu-Natal, Westville campus. Saliva was collected from University of KwaZulu-Natal, Westville in 50 mL microcentrifuge tubes. All samples were immediately processed once collected, with the exception of vaginal fluid and semen which were deposited on sterile cotton swabs and air dried at room temperature prior to DNA isolation.

DNA was extracted from each aliquot of blood, saliva and each swab of semen and vaginal fluid using Quick-gDNA MiniPrep kit (Zymo Research, Irvine, California, United States) according to manufacturer's instructions. Extracted DNA was quantified using a NanoDrop (NanoDrop™ 2000 Spectrophotometer, Waltham, Massachusetts, United States). The concentration of blood, saliva, semen and vaginal fluid of each volunteer is listed in Table 1, Appendix B. All DNA samples were stored at -20ºC after quantification.

3.2.3 Bisulfite conversion

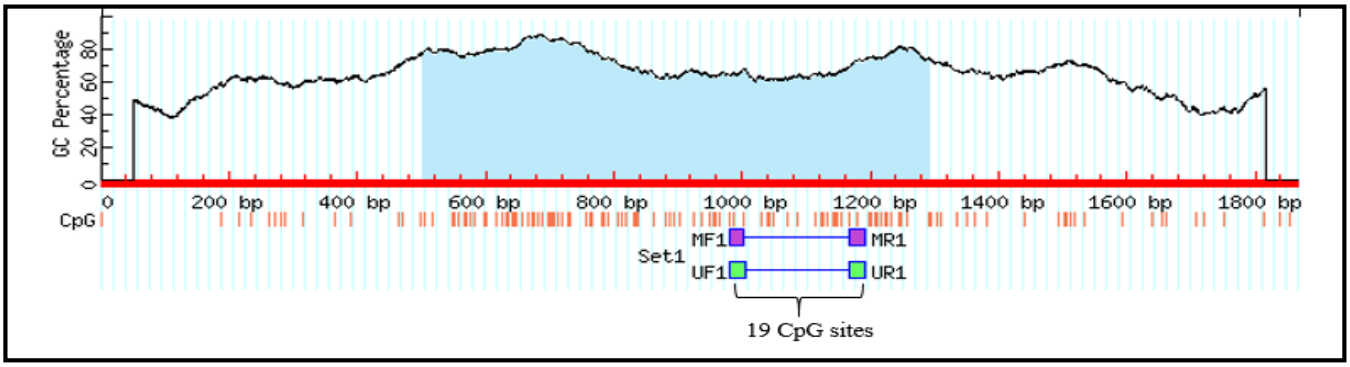
The DNA methylation sites on the extracted DNA has to be preserved when a standard PCR amplification is run. Therefore, the genomic DNA was treated with sodium bisulfite in order to convert the unmethylated cytosines to uracil and methylated cytosine remain the same (Silva *et al*., 2016) using the EZ DNA Methylation Gold Kit (Zymo Research, Irvine, California, United States). In this method the converted uracil is replaced by thymine during the PCR amplification process. The bisulfite modified DNA was amplified by site specific PCR primers designed to amplify the bisulfite modified target regions.

3.2.4 Selection of candidate genes and primer design

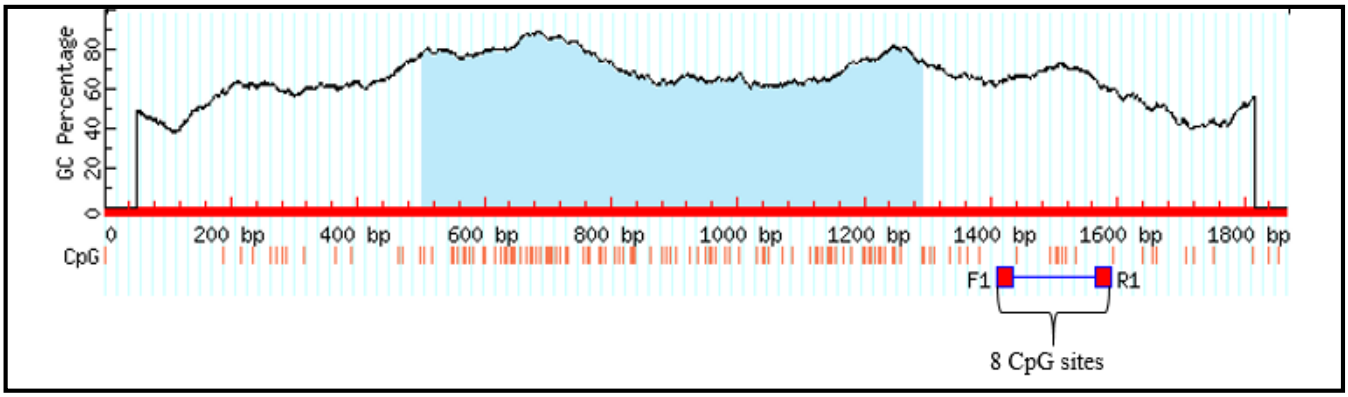
In a previous study by Kader. (2015, MSc Genetics dissertation) genes with differential gene expression in human tissues and body fluids were identified. It was proposed that differential gene expression in forensically relevant body fluids could be due to differences in methylation pattern in the target genes. Three genes namely, *ZNF282*, *PTPRS* and *HPCAL1* that displayed differential gene expression in forensically relevant body fluids viz. blood, semen, saliva and vaginal fluid were selected. One CpG island within each gene was selected for primer design. Except for the *PTPRS* gene, for which two CpG islands (referred to as PTPRS 1 and PTPRS 2) were selected for comparison of methylation patterns within a gene.

Two methods, MSP and BS were used for methylation analysis. Primers targeting the CpG islands for the genes, *ZNF282*, *PTPRS*, and *HPCAL1* were designed using the MethPrimer program (http://www.urogene.org/methprimer/). The amplicon sizes and annealing temperatures of the primers for BS and MSP are presented in Table 3.8 and Table 3.9. BiSearch is a primer design and search tool that used to search various genomes with primers that have been designed to prevent non-specific PCR products. In Appendix C, the primers that were designed in MethPrimer and tested in Bisearch are presented. The number of CpG sites analysed by the MSP and BS primer set is shown in Figure 3.1a and Figure 3.1b for ZNF282 tDMR, Figure 3.2a and Figure 3.2b for PTPRS 1 tDMR, Figure 3.3a and Figure 3.3b for PTPRS 2 tDMR and Figure 3.4a and Figure 3.4b for HPCAL1 tDMR. The chromosomal location of the target CpG island within each gene is located in Figure 3.1c (*ZNF282*), Figure 3.2c (*PTPRS*- PTPRS 1), Figure 3.3c (*PTPRS* –PTPRS 2) and Figure 3.4c (*HPCAL1*). The target region amplified by the primers will be referred to as candidate tDMRs in the rest of the chapter, for ease of understanding. Table 3.7 shows the genomic information for each gene that is studied in the present research.

a)



b)



c)

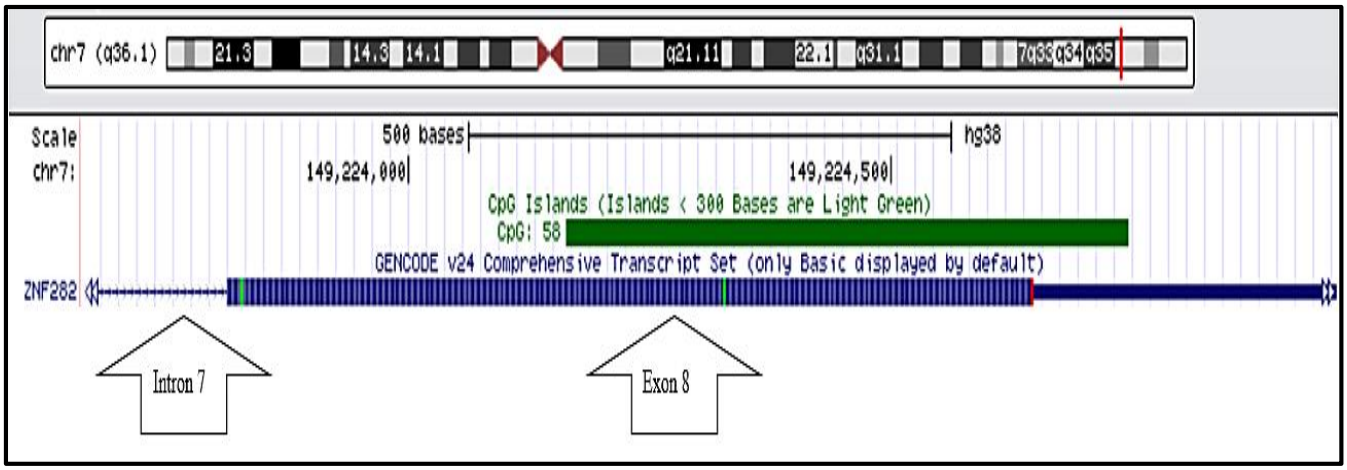
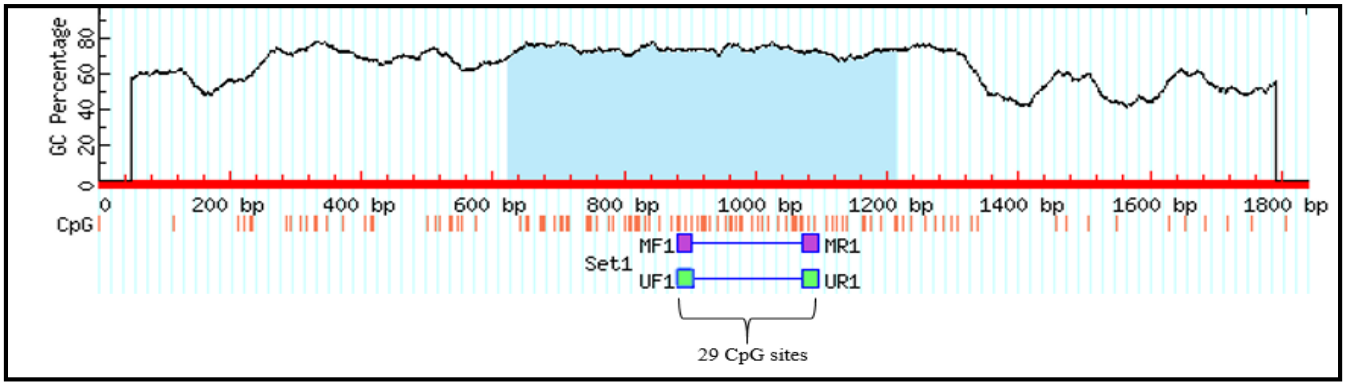
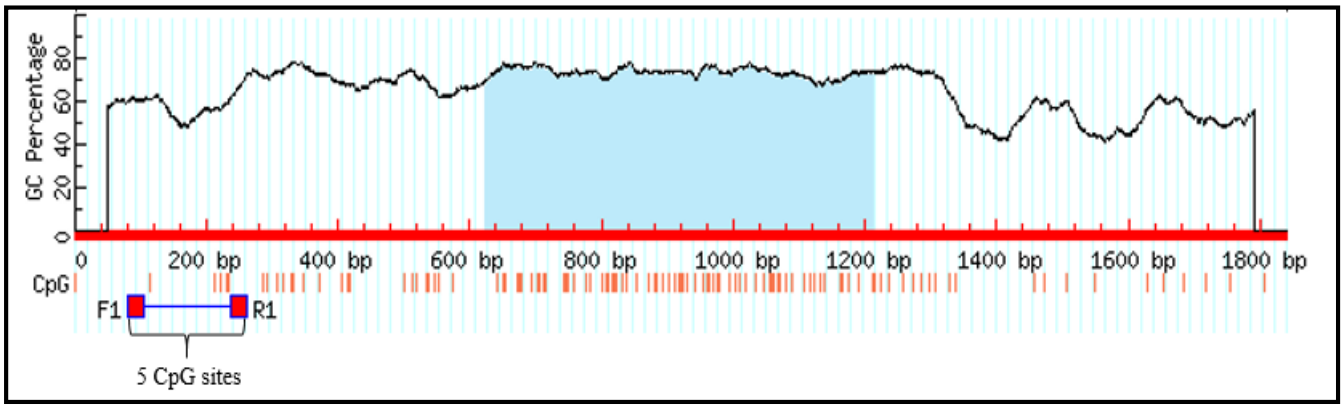


Figure 3.1: Target CpG island of *ZNF282* gene for primer design a) MSP: Visual representation of the location of the primers designed within the CpG island (blue colored area). Two primer sets are indicated by the purple boxes (MF1- methylated forward primer and MR1- methylated reverse primer) and green boxes (UF1- unmethylated forward primer and UR1- unmethylated reverse primer). Horizontal red line indicates the input sequence and vertical red lines represent CpG sites. b) BS: Visual representation of the location of the primers designed out of the CpG island. Primers are indicated by red boxes (F1- forward primer and R1- reverse primer). c) UCSC genome browser view of chromosomal location of the target CpG island within the gene indicated by the green bar and the position of the CpG island in relation to the introns and exons indicated by the blue horizontal line and box.

a)



b)



c)

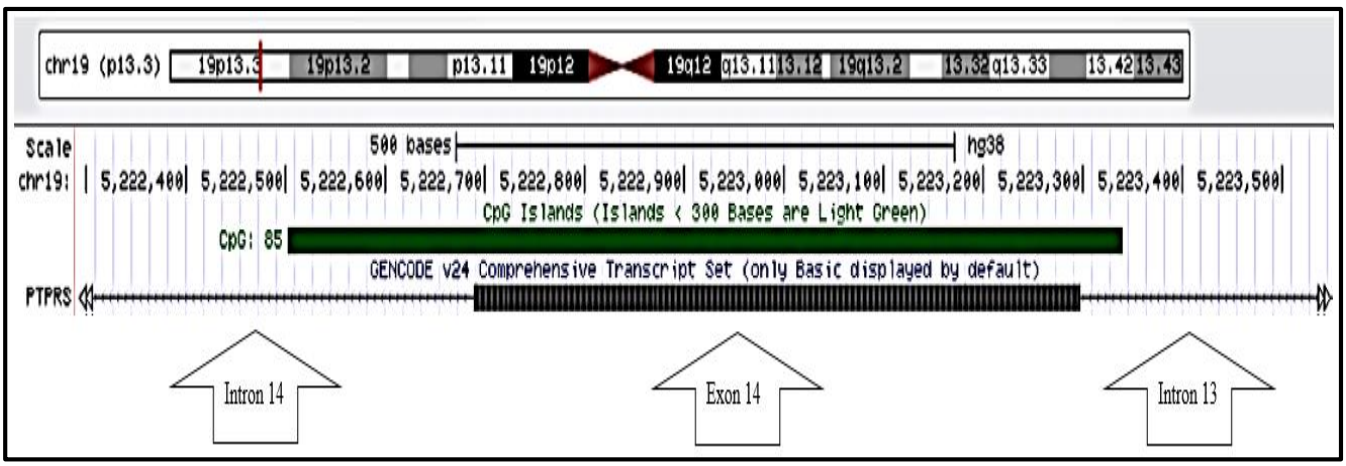
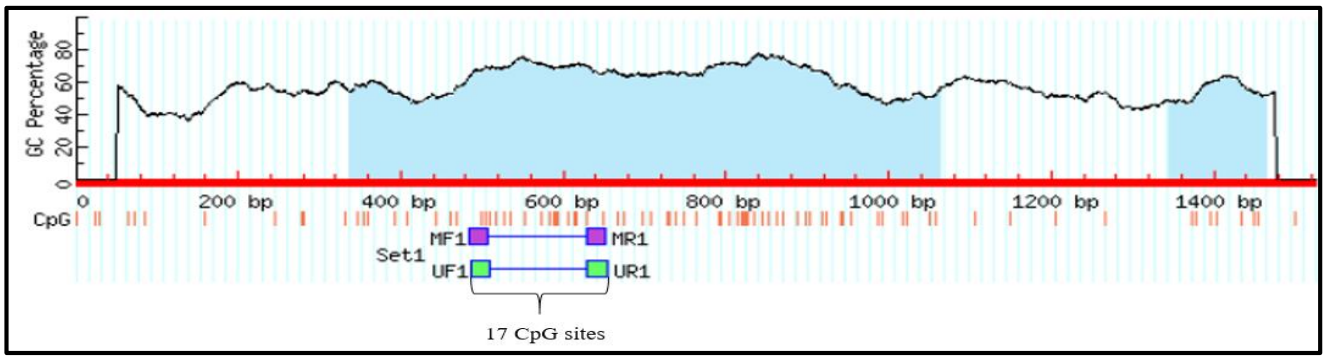
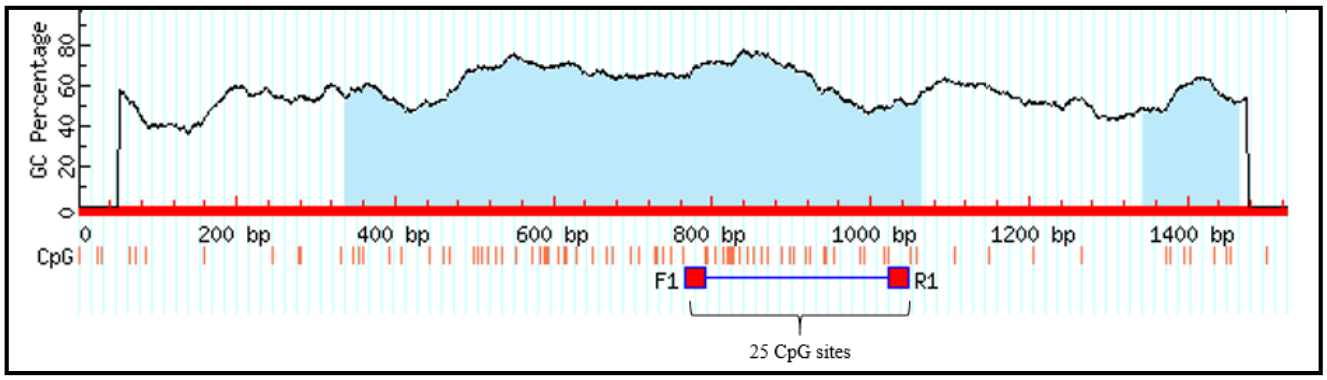


Figure 3.2: Target CpG island of *PTPRS* (PTPRS 1) gene for primer design a) MSP: Visual representation of the location of the primers designed within the CpG island (blue colored area). Two primer sets are indicated by the purple boxes (MF1- methylated forward primer and MR1- methylated reverse primer) and green boxes (UF1- unmethylated forward primer and UR1- unmethylated reverse primer). Horizontal red line indicates the input sequence and vertical red lines represent CpG sites. b) BS: Visual representation of the location of the primers designed out of the CpG island. Primers are indicated by red boxes (F1- forward primer and R1- reverse primer). c) UCSC genome browser view of chromosomal location of the target CpG island within the gene indicated by the green bar and the position of the CpG island in relation to the introns and exons indicated by the blue horizontal line and box.

a)



b)



c)

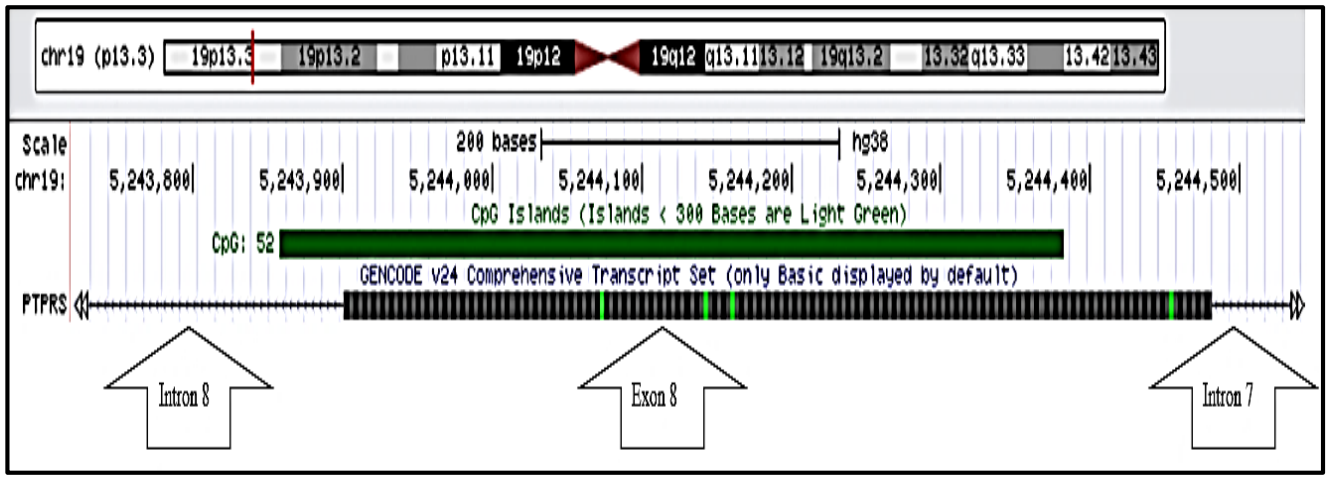
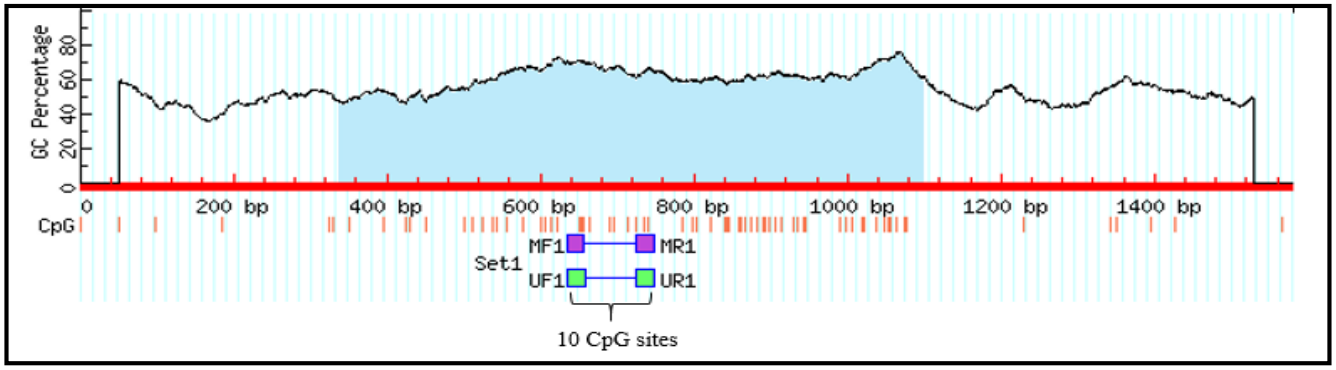
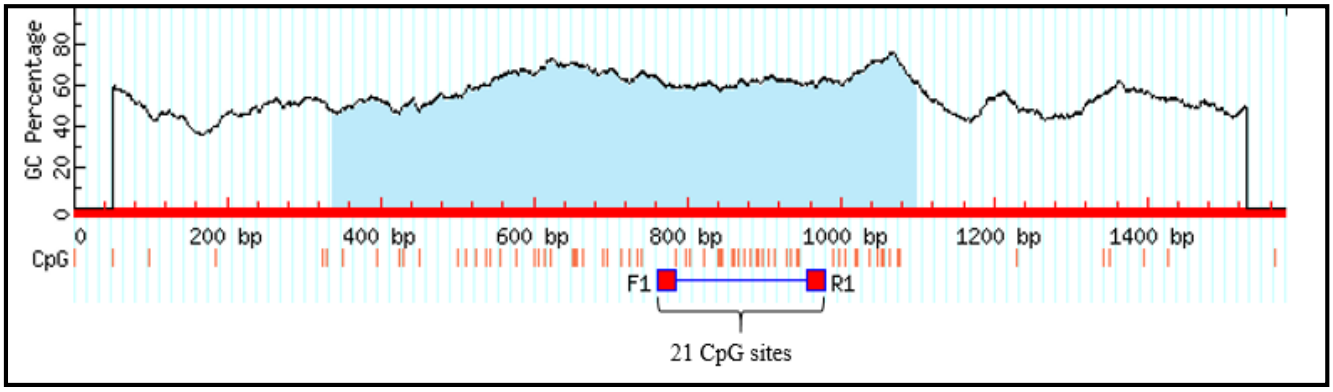


Figure 3.3: Target CpG island of *PTPRS* (PTPRS 2) gene for primer design a) MSP: Visual representation of the location of the primers designed within the CpG island (blue colored area). Two primer sets are indicated by the purple boxes (MF1- methylated forward primer and MR1- methylated reverse primer) and green boxes (UF1- unmethylated forward primer and UR1- unmethylated reverse primer). Horizontal red line indicates the input sequence and vertical red lines represent CpG sites. b) BS: Visual representation of the location of the primers designed. Primers are indicated by red boxes (F1- forward primer and R1- reverse primer). c) UCSC genome browser view of chromosomal location of the target CpG island within the gene indicated by the green bar and the position of the CpG island in relation to the introns and exons indicated by the blue horizontal line and box.

a)



b)



c)

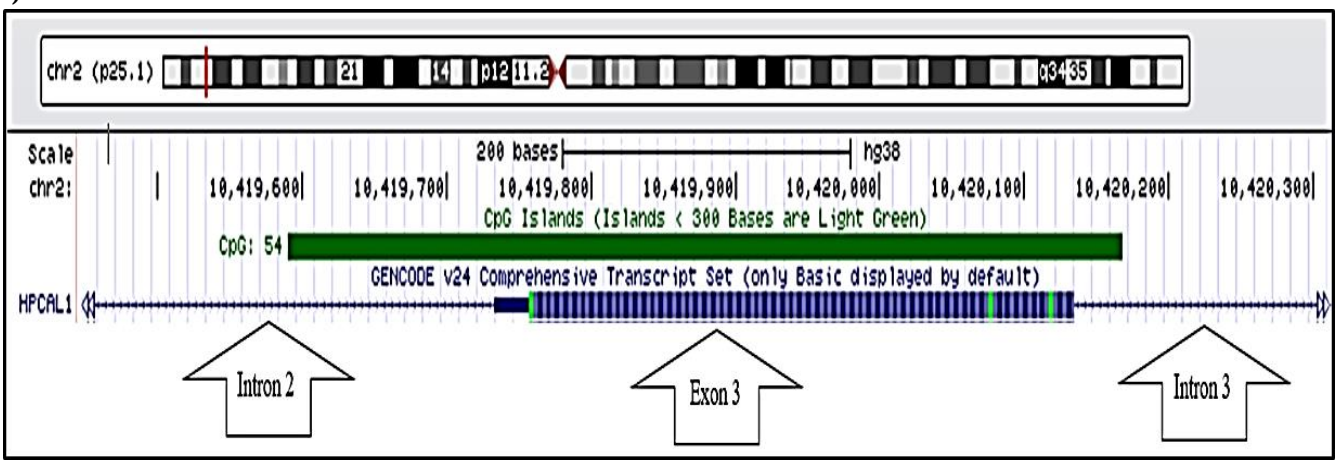


Figure 3.4: Target CpG island of *HPCAL1* gene for primer design a) MSP: Visual representation of the location of the primers designed within the CpG island (blue colored area). Two primer sets are indicated by the purple boxes (MF1- methylated forward primer and MR1- methylated reverse primer) and green boxes (UF1- unmethylated forward primer and UR1- unmethylated reverse primer). Horizontal red line indicates the input sequence and vertical red lines represent CpG sites. b) BS: Visual representation of the location of the primers designed. Primers are indicated by red boxes (F1- forward primer and R1- reverse primer). c) UCSC genome browser view of chromosomal location of the target CpG island within the gene indicated by by the green bar and the position of the CpG island in relation to the introns and exons indicated by the blue horizontal line and box.

Table 3.7: Genomic information for the three genes with candidate tDMRs characterized in this study.

	HPCAL1	***PTPRS* (PTPRS 1)**	***PTPRS* (PTPRS 2)**	***ZNF282***
Chromosomal Location	2	19	19	7
Genomic coordinates of gene (UCSC Genome Browser assembly -Human GRCh38/hg38)	10559217-10560794	5222013-5223851	5243370-5244893	148920473-148922337
Genomic coordinates of CpG island (UCSC Genome Browser assembly -Human GRCh38/hg38)	MSP: 10419591-10420168 BS: 10419852-10420069	MSP: 5222292-5223550 BS: 5222081-5222264	MSP: 5243721-5244544 BS: 5244124-5244407	MSP: 149223664-149224961 BS: 149224792-149224971
CpG islands number	54	85	52	58
Number of CpG sites	MSP: 10 BS: 21	MSP: 29 BS: 5	MSP: 17 BS: 25	MSP: 19 BS: 8

Table 3.8: BS primers designed for analysis of methylation in candidate tDMRs.

tDMR		**Primer Sequence**	**Amplicon Size (bp)**	**Annealing Temperature(°C)**
ZNF282	Forward	GGATTTTTTAGATTTGTTTGGTTTG	180	46
	Reverse	TCCCTAATAACTTCCCCTAATAACC		
PTPRS 1	Forward	TGATTTTGTTTGTTTGTTTGTTTGT	184	47
	Reverse	AACCTCTCTTCTACTCTCTCCCTACTC		
PTPRS 2	Forward	TTTTTTTATTGTATAATTATGGTGGT	283	43
	Reverse	ACAACCCAAATCCTATATCCTATTAC		
HPCAL1	Forward	AGGGTTTTTTTAAGGATTGTTTTAT	218	45
	Reverse	AACCCACTTAAACTTCTACTCCAACT		

Table 3.9: MSP primers designed for analysis of methylation in candidate tDMRs.

tDMR		Primer Sequence	Amplicon Size (bp)	Annealing Temperature (ºC)
PTPRS 1	[1]MF	CGTACGGTTATCGTGTATTTGAC	216	49
	[2]MR	AACCACGTATATATTCCGACTTACG		
	[3]UF	TGTATGGTTATTGTGTATTTGATGA	215	44
	[4]UR	ACCACATATATATTCCAACTTACAAC		
PTPRS 2	MF	ATGTAAAGAATTAAGTCGTACGGTC	168	46
	MR	CGTAAACAACCTACTAAAAAACGAA		
	UF	GTAAAGAATTAAGTTGTATGGTTGG	169	45
	UR	CACCATAAACAACCTACTAAAAAACAA		
HPCAL1	MF	GTTTTGTAGGTGTAGTCGTCGTC	113	48
	MR	AACTCGTAATCGATAAACTCCGTA		
	UF	GTTTTGTAGGTGTAGTTGTTGTTGT	114	43
	UR	CAACTCATAATCAATAAACTCCATA		
ZNF282	MF	GGAGTGCGAGAAGATTTATAGTC	210	48
	MR	TCCTTAAACGACTCCTTATAACGAA		
	UF	GGAGTGTGAGAAGATTTATAGTTGT	210	44
	UR	TCCTTAAACAACTCCTTATAACAAA		

[1]methylated forward, [2]methylated reverse, [3]unmethylated forward, [4]unmethylated reverse

Table 3.10: Universal Methylated Human DNA Standard for assessment of efficiency of bisulfite conversion of DNA.

---ggagtga aggaggcca**CG**ggcaagt**CG** ccctga**CG**caga**CG**ctccac cagggc**CG**CG **CG**ct**CG**c**CG**t c**CG**ccacata c**CG**ct**CG**tag tatt**CG**tgct cagcct**CG**ta gtgg**CG**cctga**CG**t**CGCG**tt **CGCG**ggtagc ta**CG**atgagg **CG**g**CG**acaga ccaggcacag ggccccat**CG** ccctc

3.2.5 MSP reaction

PCR amplification was carried out using a ZymoTaq™ DNA Polymerase Kit (Zymo Research, Irvine, California, United States) in a 25 µL reaction volume containing 12 µL of ZymoTaq PCR Buffer, 0.25 µL of dNTP mix (25 µM of each dNTP), 0.15 µM of each primer (Inqaba Biotec), 0.2 µL of ZymoTaq DNA Polymerase (5U/µL) and 1 µL of bisulfite converted DNA. A Universal Methylated Human DNA Standard (Zymo Research, Irvine, California, United States) was bisulfite converted using EZ DNA Methylation Gold Kit (Zymo Research, Irvine, California, United States) and was used as a positive control for complete conversion (Table 3.11). The original sequence for the methylated standard is given in Table 3.11. The PCR was conducted in a BIORAD T100™ Thermal Cycler (Bio-Rad Laboratories, Hercules, California, United States) under the following conditions: 95ºC for 10 minutes, 30 cycles of 95ºC for 30 seconds, varying annealing temperatures (Table 3.9) for 30 seconds, and 72ºC for 60 seconds and a final extension of 72ºC for 7 minutes. The amplified products were analysed by agarose gel electrophoresis using 2% agarose gel stained with 10mg/ml ethidium bromide at 60 volts for 60 minutes.

3.2.6 BS protocol

3.2.6.1 BS PCR reaction

DNA samples (10) of the same body fluids across individuals were pooled together such that 100 ng of DNA from each participant was included. The pooled DNA was modified by bisulfite treatment, following the conversion protocol PCR amplification was carried out using a ZymoTaq™ DNA Polymerase Kit in a 50 µL reaction volume containing 24 µL of ZymoTaq PCR Buffer, 0.5 µL of dNTP mix (25 µM of each dNTP), 0.15 µM of each primer (Inqaba Biotec) and 0.4 µL of DNA Polymerase (5U/µL) and 1 µL of bisulfite converted DNA. The PCR was conducted on a BIORAD T100™ Thermal Cycler under the following conditions: 95ºC for 10 minutes, 30 cycles of 95ºC for 30 seconds, varying annealing temperatures (Table 3.8) for 30 seconds, and 72ºC for 60 seconds and a final extension of 72ºC for 7 minutes. An aliquot of the amplified products was analysed by agarose gel electrophoresis using 2% agarose gel stained with 10mg/ml ethidium bromide at 60 volts for 60 minutes to confirm the presence of the amplicon.

3.2.6.2 Cloning and sequencing of BS reaction products

The remaining amplified products from the BS PCR reaction were then run on a 1% agarose gel and the bands were excised and purified using a Thermo scientific GeneJet Gel Extraction Kit according to the manufacturers protocol. The purified amplified products were ligated and cloned into pJET1.2 vector using the Thermo Scientific CloneJet PCR cloning Kit according to the manufacturers protocol. Following ligation, an aliquot of the ligation reaction was transformed using prepared chemically competent cells (*Escherichia coli* - DH5α). The transformed cells were then grown overnight on LB/amp plates. Between five to 10 positive clones were isolated from each body fluid. The plasmids were purified with the Thermo Scientific GeneJet Plasmid Miniprep Kit according to the manufacturers protocol. The purified plasmids were sent to Inqaba biotec for sequencing and this was followed by analysis on online software Bisulfite Sequencing DNA Methylation Analysis (BISMA) (http://services.ibc.uni-stuttgart.de/BDPC/BISMA/). To ensure that the cloning process was efficient a positive and negative control was also plated on LB/amp plates. The positive control was positive samples from trial runs and the negative control was DH5α cells. Figure 3.5 shows the overview of the cloning process.

3.2.7 Bisulfite sequencing data analysis and DNA methylation profiling

The BISMA web application was used for analysis of bisulfite sequencing results. The methylation results were compiled based on the methylation status results from each body fluid sample. To determine the methylation levels of candidate markers in the different body fluid samples, sequencing data were aligned against *in silico*-converted genomic reference sequences using BISMA. The output files were compiled using Bisulfite Sequencing Data Presentation and Compilation (BDPC) web application (http://biochem.jacobs-university.de/BDPC/) to derive information and compare results from each marker in the different body fluid samples (Rohde *et al*., 2008).

3.2.8 Statistical analysis

To evaluate the methylation profiles of candidate tDMRs and to determine whether they could be used to distinguish individual body fluids, statistical analyses was carried out using the web application, QUantification tool for Methylation Analysis (QUMA) (http://quma.cdb.riken.jp/). Pairwise comparisons of methylation were made for each marker and CpG locus using Chi-square test or Fisher's exact test as appropriate. Differences were statistically significant when *p*-values were less than 0.05.

Figure 3.5: Flowchart showing the overview of the BS protocol.

3.3 Results

3.3.1 Methylation profile using MSP

From a total of 40 body fluid samples collected (10 samples each of blood, vaginal fluid, semen and saliva) the following results were obtained for each candidate tDMR primer set.

3.3.1.1 ZNF282 tDMR

Blood was methylated at target CpG sites for ZNF282 tDMR as amplification was observed with methylated primers in seven out of 10 samples. No amplification was observed with the unmethylated primers (Figure 3.6a and b). Vaginal fluid presented similar results, with methylation in seven samples and no unmethylation in any of the sample (Figure 3.7a and b). In saliva, all samples were methylated (10) and two of these samples also showed unmethylation (Figure 3.9a and b). Semen showed contrasting results when compared to other body fluids as most of the samples were unmethylated (five) and only one sample showed methylation (Figure 3.8a and b). The positive control (Universal Methylated Human DNA Standard) showed amplification at 210 bp thus confirming complete bisulfite conversion of genomic DNA (Figure 3.6a). In Table 3.11 is the summarized results for the ZNF282 tDMR.

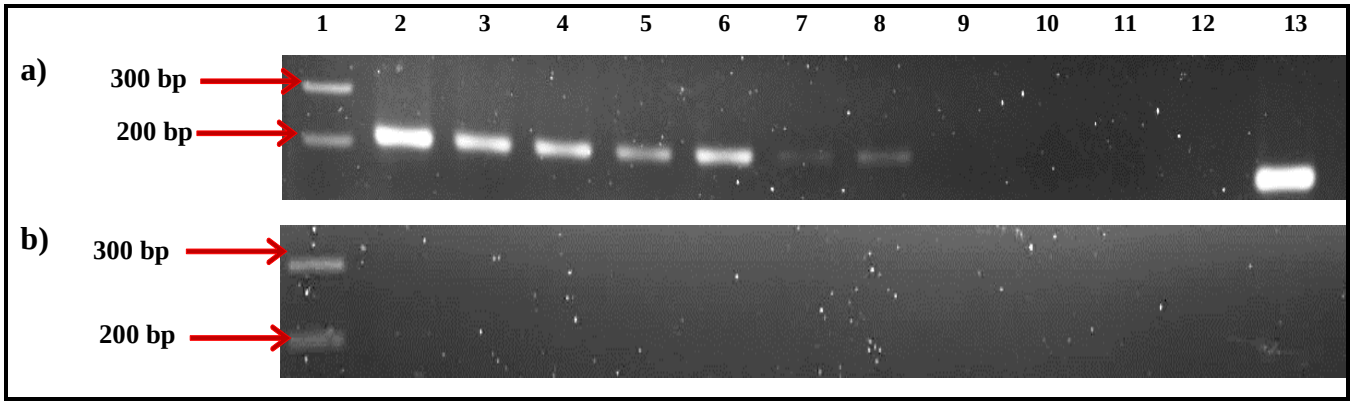


Figure 3.6: MSP based methylation profile of blood for the ZNF282 tDMR primer set. a) Products with primers specific for methylated cytosine (210 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- blood DNA samples. Lane 12- No template control. Lane 13- Universal Methylated Human DNA Standard (Zymo Research). b) Products with primers specific for unmethylated cytosine (210 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- blood DNA samples. Lane 12- No template control.

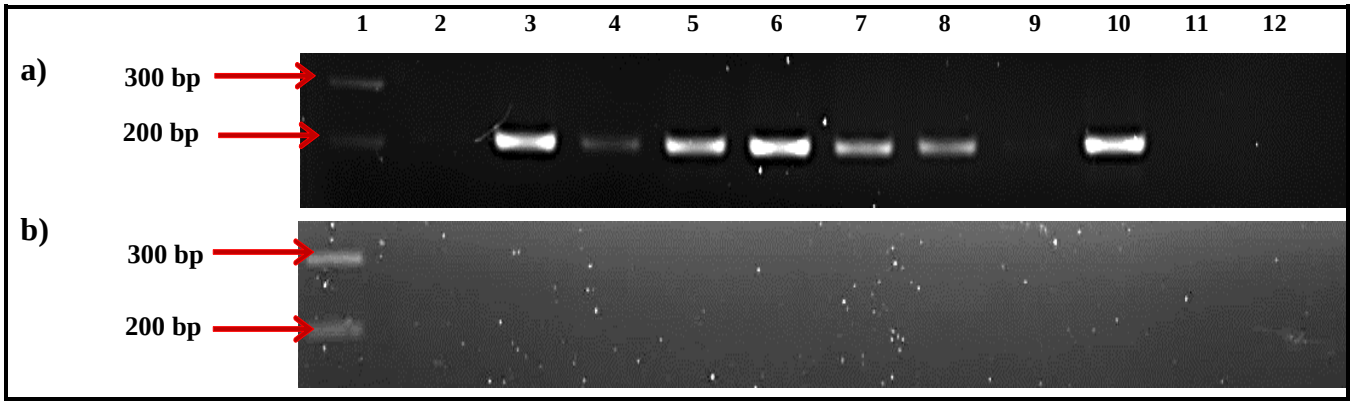


Figure 3.7: MSP based methylation profile of vaginal fluid for the ZNF282 tDMR primer set. a) Products with primers specific for methylated cytosine (210 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- vaginal fluid DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (210 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- vaginal fluid DNA samples. Lane 12- No template control.

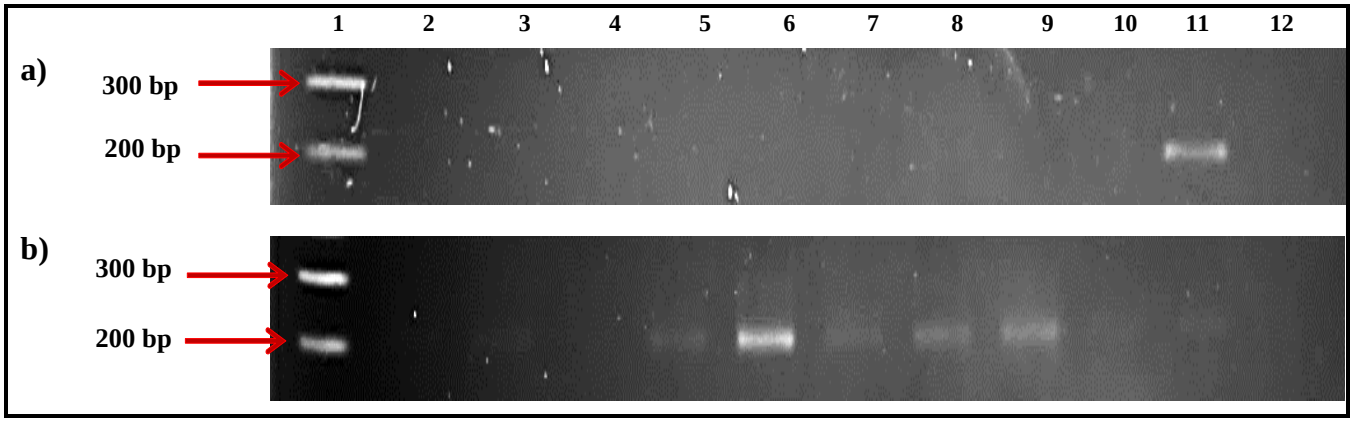


Figure 3.8: MSP based methylation profile of semen for the ZNF282 tDMR primer set. a) Products with primers specific for methylated cytosine (210 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- semen DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (210 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- semen DNA samples. Lane 12- No template control.

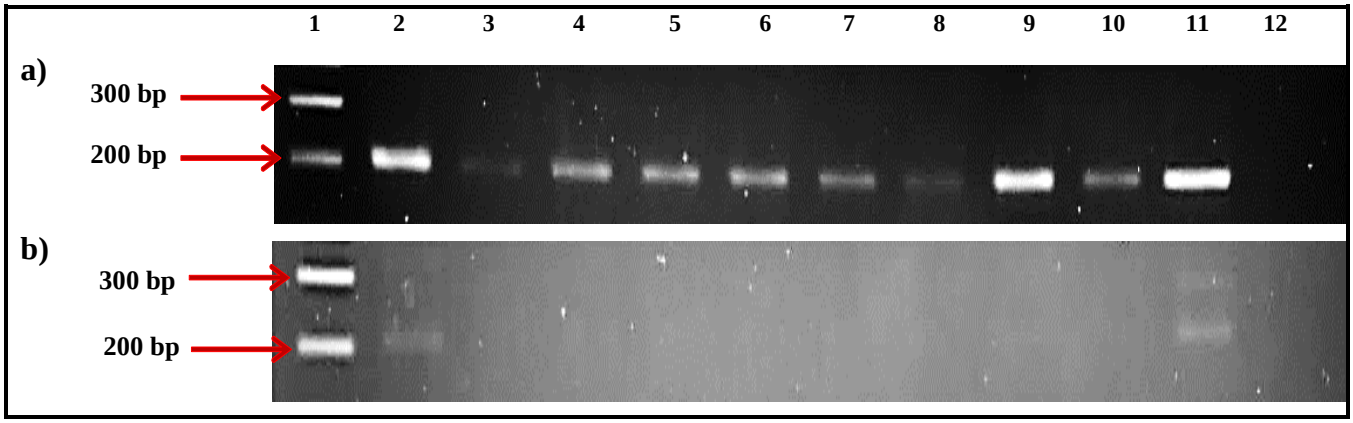


Figure 3.9: MSP based methylation profile of saliva for the ZNF282 tDMR primer set. a) Products with primers specific for methylated cytosine (210 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- saliva DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (210 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- saliva DNA samples. Lane 12- No template control.

3.3.1.2 PTPRS 1 tDMR

For the PTPRS 1 tDMR amplification with the methylated primer set was observed in nine blood samples (Figure 3.10a), nine vaginal fluid samples (Figure 3.11a) and in all 10 saliva samples (Figure 3.13a), however, with the unmethylated primer set no amplification was seen for blood (Figure 3.10b), vaginal fluid (Figure 3.11b) and saliva (Figure 3.13b) in all 10 samples in each fluid. Semen showed amplification for both methylated and unmethylated primer sets. Most of the samples were unmethylated (nine) with six samples also showing methylation (Figure 3.12a and Figure 3.12b). The positive control (Universal Methylated Human DNA Standard) showed amplification at 216 bp thus confirming complete bisulfite conversion (Figure 3.10a). In Table 3.11 is the summarized results for the PTPRS 1 tDMR.

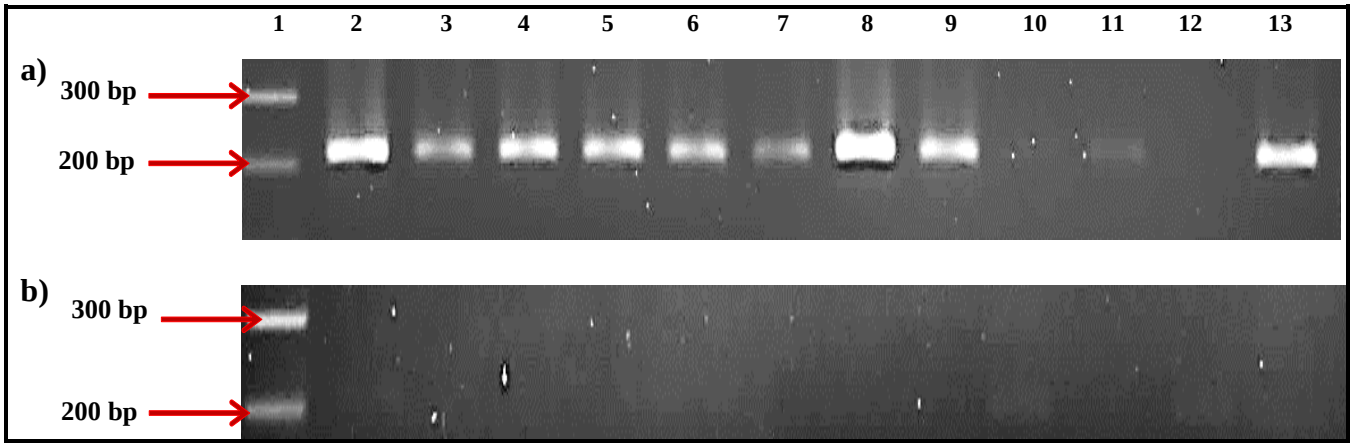


Figure 3.10: MSP based methylation profile of blood for the PTPRS 1 tDMR primer set. a) Products with primers specific for methylated cytosine (216 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- blood DNA samples. Lane 12- No template control. Lane 13- Universal Methylated Human DNA Standard (Zymo Research). b) Products with primers specific for unmethylated cytosine (215 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- blood DNA samples. Lane 12- No template control.

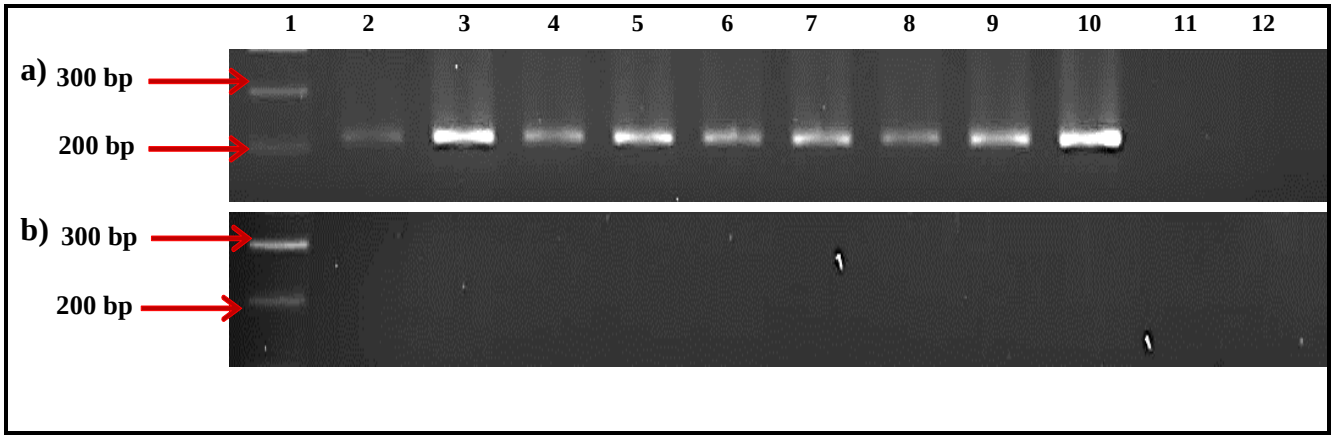


Figure 3.11: MSP based methylation profile of vaginal fluid for the PTPRS 1 tDMR primer set. a) Products with primers specific for methylated cytosine (216 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- vaginal fluid DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (215 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- vaginal fluid DNA samples. Lane 12- No template control.

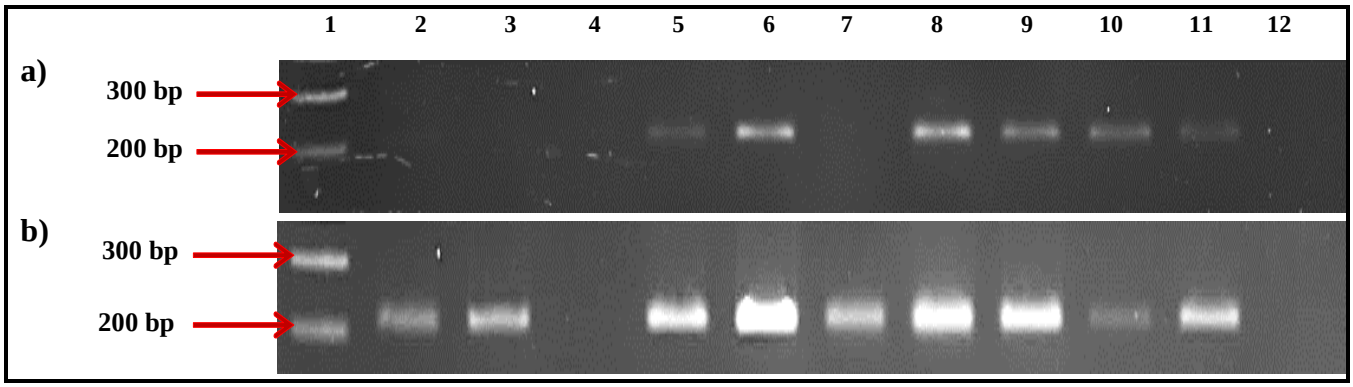


Figure 3.12: MSP based methylation profile of semen for the PTPRS 1 tDMR primer set. a) Products with primers specific for methylated cytosine (216 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- semen DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (215 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- semen DNA samples. Lane 12- No template control.

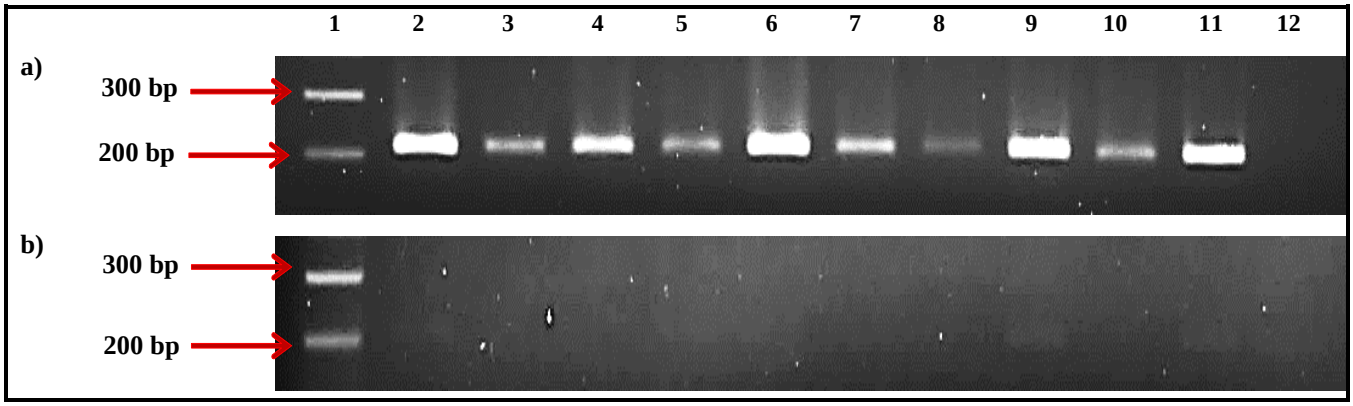


Figure 3.13: MSP based methylation profile of saliva for the PTPRS 1 tDMR primer set. a) Products with primers specific for methylated cytosine (216 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- saliva DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (215 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- saliva DNA samples. Lane 12- No template control.

3.3.1.3 PTPRS 2 tDMR

All four body fluids displayed amplification for the PTPRS 2 tDMR with methylated and unmethylated primers. For blood, methylation and unmethylation was observed in nine and nine samples, respectively (Figure 3.14a and b). Vaginal fluid displayed similar results with methylation and unmethylation in nine and eight samples, respectively (Figure 3.15a and b). Methylation was observed in semen for five samples and unmethylation in seven samples (Figure 3.16a and b). For saliva, methylation was observed in eight samples (Figure 3.17a),

however five samples with unmethylated primers showed amplification (Figure 3.17b). Complete bisulfite conversion was confirmed with the positive control (Universal Methylated Human DNA Standard) showing amplification at 168 bp (Figure 3.14a). In Table 3.11 is the summarized results for the PTPRS 2 tDMR.

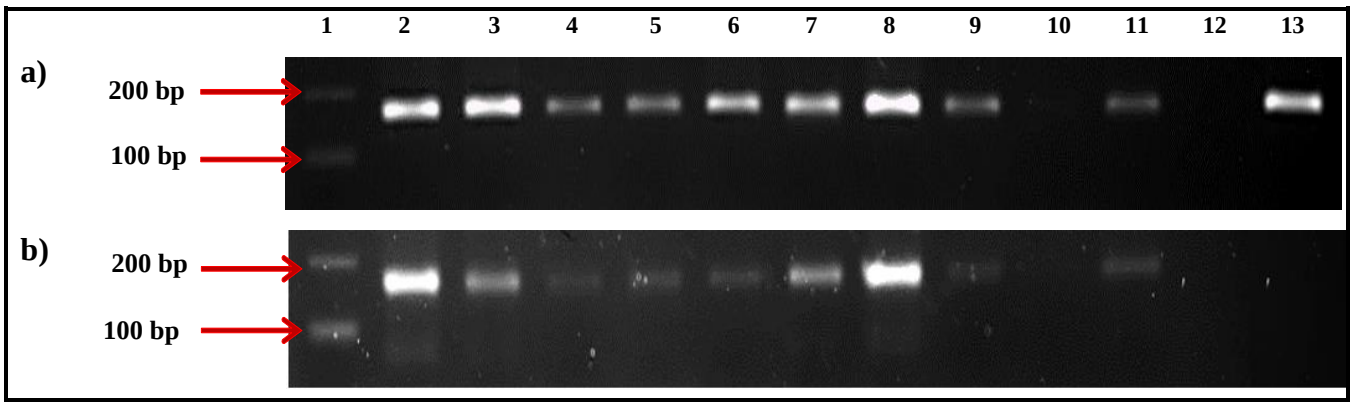


Figure 3.14: MSP based methylation profile of blood for the PTPRS 2 tDMR primer set. a) Products with primers specific for methylated cytosine (168 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- blood DNA samples. Lane 12- No template control. Lane 13- Universal Methylated Human DNA Standard (Zymo Research) b) Products with primers specific for unmethylated cytosine (169 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- blood DNA samples. Lane 12- No template control.

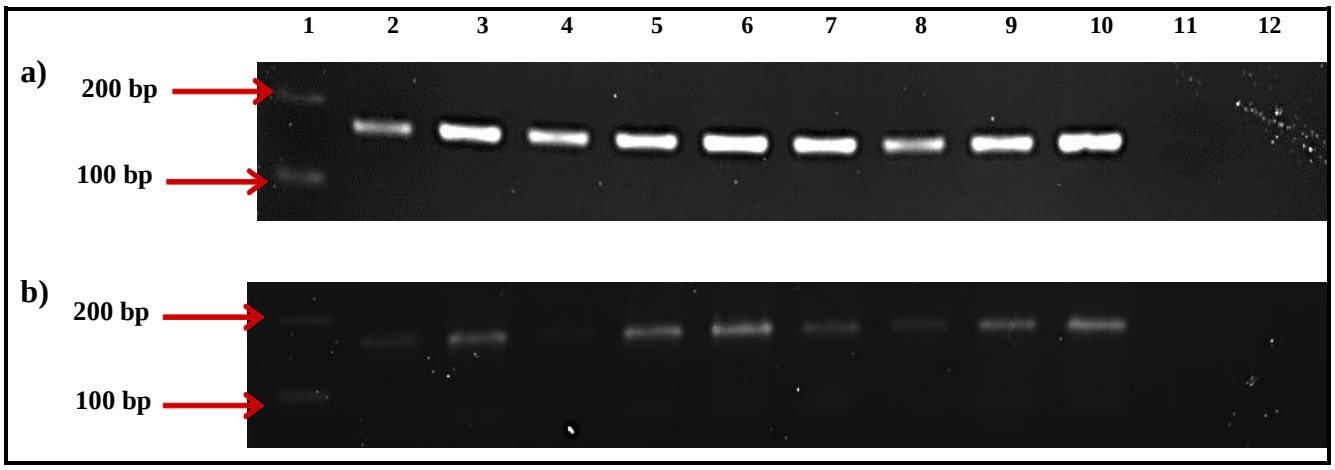


Figure 3.15: MSP based methylation profile of vaginal fluid for the PTPRS 2 tDMR primer set. a) Products with primers specific for methylated cytosine (168 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- vaginal fluid DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (169 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- vaginal fluid DNA samples. Lane 12- No template control.

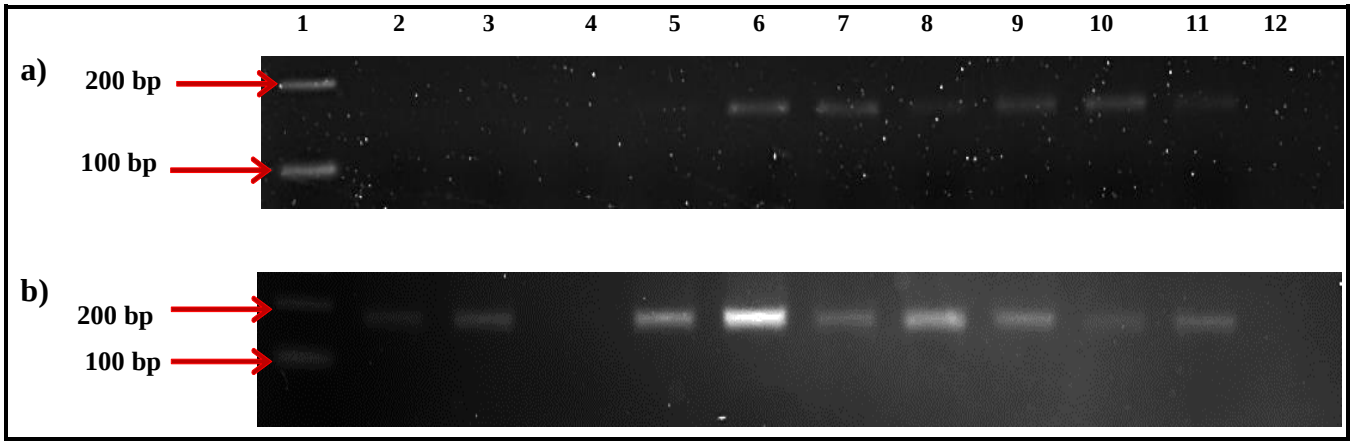


Figure 3.16: MSP based methylation profile of semen for the PTPRS 2 tDMR primer set. a) Products with primers specific for methylated cytosine (168 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- semen DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (169 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- semen DNA samples. Lane 12- No template control.

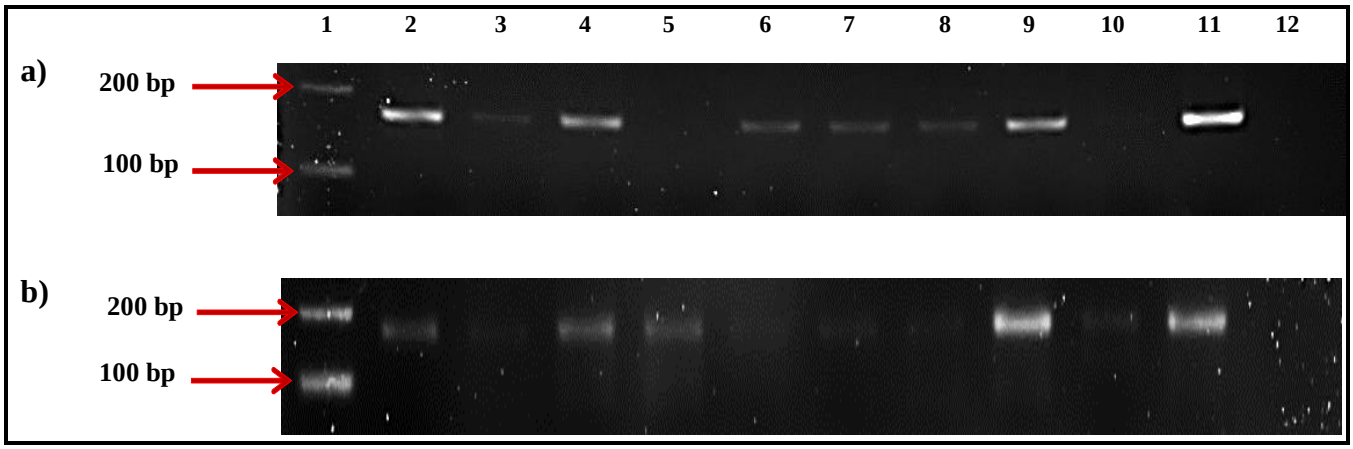


Figure 3.17: MSP based methylation profile of saliva for the PTPRS 2 tDMR primer set. a) Products with primers specific for methylated cytosine (168 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- saliva DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (169 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- saliva DNA samples. Lane 12- No template control.

3.3.1.4 HPCAL1 tDMR

Blood was methylated at target CpG sites for HPCAL1 tDMR as amplification was observed for all 10 samples (Figure 3.18a) and non-specific was observed in two samples and absence of amplification in eight samples with unmethylated primers (Figure 3.18b). For vaginal fluid, nine samples showed methylation (Figure 3.19a) and non-specific amplification was observed in two samples and eight sample did not amplify with unmethylated primers (Figure 3.19b). For semen, methylation was observed for eight samples (Figure 3.20a) and only two samples

displayed unmethylation (Figure 3.20b). All 10 saliva samples did not produce consistent amplification with methylation primers (Figure 3.21a). However, all 10 saliva samples displayed unmethylation (Figure 3.21b). The positive control (Universal Methylated Human DNA Standard) confirmed complete bisulfite conversion with amplification at 113 bp (Figure 3.18a). In Table 3.11 is the summarized results for the HPCAL1 tDMR.

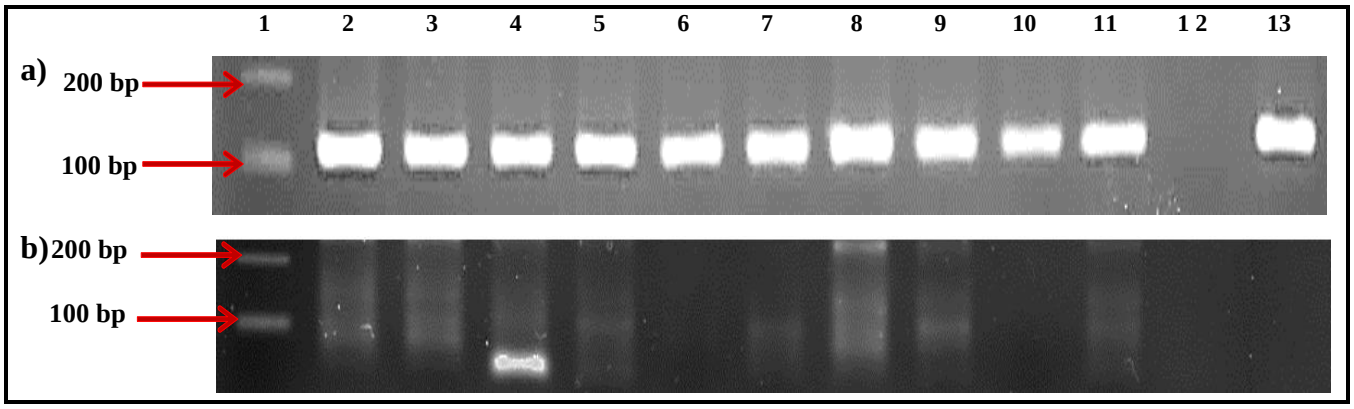


Figure 3.18: MSP based methylation profile of blood for the HPCAL1 tDMR primer set. a) Products with primers specific for methylated cytosine (113 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- blood DNA samples. Lane 12- No template control. Lane 13- Universal Methylated Human DNA Standard (Zymo Research). b) Products with primers specific for unmethylated cytosine (114 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- blood DNA samples. Lane 12- No template control.

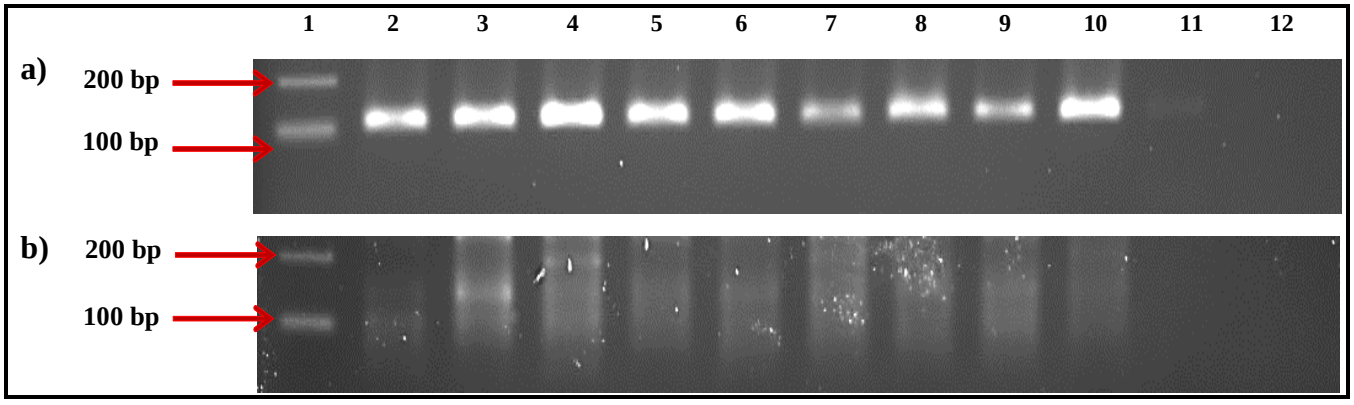


Figure 3.19: MSP based methylation profile of vaginal fluid for the HPCAL1 tDMR primer set. a) Products with primers specific for methylated cytosine (113 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- vaginal fluid DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (114 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- vaginal fluid DNA samples. Lane 12- No template control.

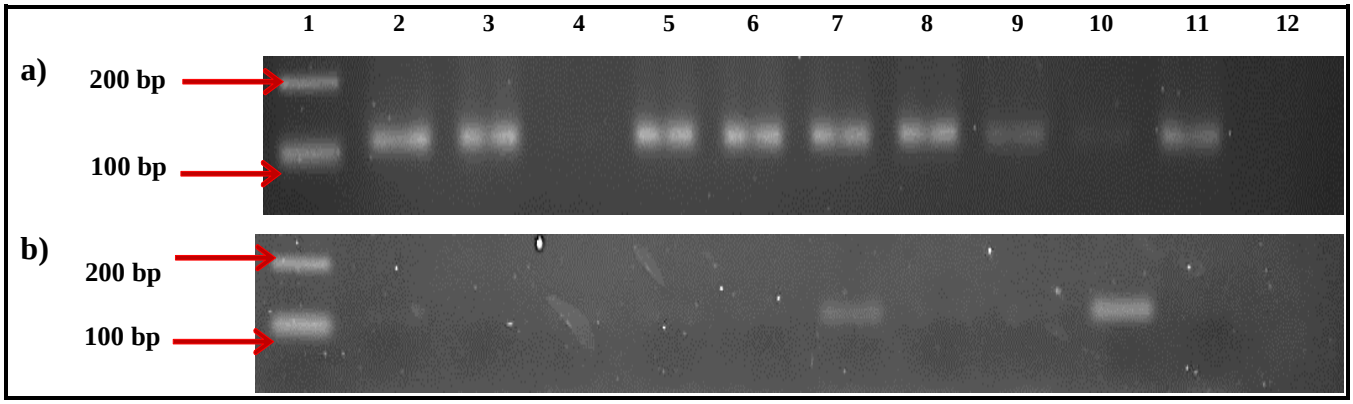


Figure 3.20: MSP based methylation profile of semen for the HPCAL1 tDMR primer set. a) Products with primers specific for methylated cytosine (113 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- semen DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (114 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- semen DNA samples. Lane 12- No template control.

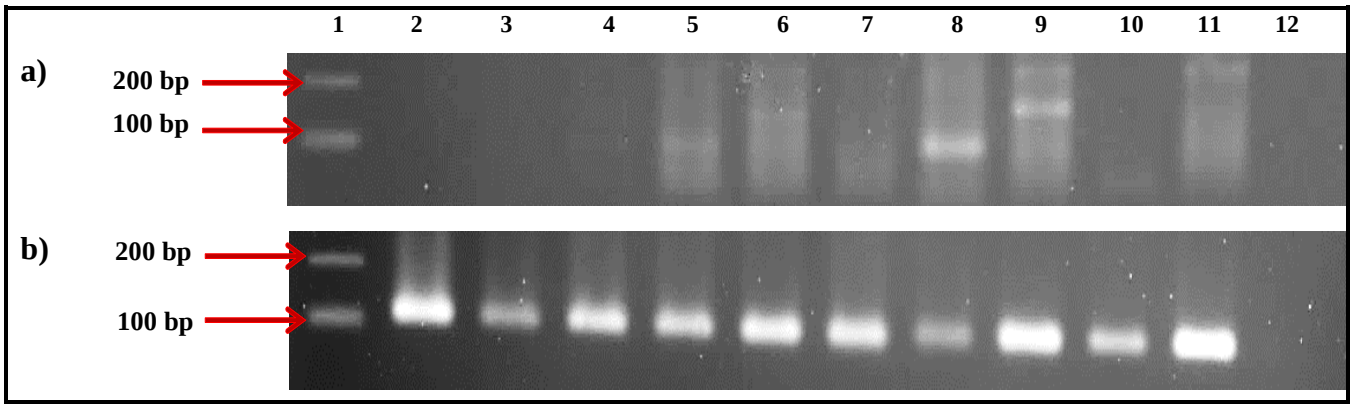


Figure 3.21: MSP based methylation profile of saliva for the HPCAL1 tDMR primer set. a) Products with primers specific for methylated cytosine (113 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- saliva DNA samples. Lane 12- No template control. b) Products with primers specific for unmethylated cytosine (114 bp). Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 11- saliva DNA samples. Lane 12- No template control.

Table 3.11: Summary of results for all four candidate tDMRs using MSP.

			Methylated primer set		Unmethylated primer set	
tDMR		**No. of samples**	**Methylated**	**Not amplified**	**Unmethylated**	**Not amplified**
ZNF282	Blood	10	7	3	0	10
	Vaginal fluid	10	7	3	0	10
	Semen	10	1	9	5	5
	Saliva	10	10	0	2	8
PTPRS 1	Blood	10	9	1	0	10
	Vaginal fluid	10	9	1	0	10
	Semen	10	6	4	9	1
	Saliva	10	10	0	0	10
PTPRS 2	Blood	10	9	1	9	1
	Vaginal fluid	10	9	1	8	2
	Semen	10	5	5	7	3
	Saliva	10	8	2	5	5
HPCAL1	Blood	10	10	0	2 (non-specific)	8
	Vaginal fluid	10	9	1	2 (non-specific)	8
	Semen	10	8	2	2	8
	Saliva	10	2 (non-specific)	8	10	0

3.3.2 Methylation profiling using BS

A total of 10 DNA samples of each body fluid were pooled for BS analysis. A study by Byun *et al.* (2009) demonstrated that DNA methylation profiles are similar between the same tissue from different individuals rather than different tissue of the same individual. Thus, DNA was pooled to obtain an estimation of the average methylation levels. The following results were obtained for each candidate tDMR using the BS analysis. The BISMA methylation analysis program was used to establish DNA methylation profiles and analyse BS sequencing data. The Chi-square and Fisher's exact tests were used for statistical analysis and was carried out in the QUMA program.

3.3.2.1 Methylation profiles of body fluid using the ZNF282 tDMR

The ZNF282 tDMR primer set targeted eight CpG sites. A total of 40 clones were sent for sequencing (10 clones for each body fluid), however only 33 clones could be analysed as these clones met the criteria of BISMA program which is a threshold sequence identity percentage of 90% or above. Overall hypermethylation (95%) was displayed in all clones for blood, with three clones (f, g and h) showing unmethylation (5%) at CpG 7, CpG 8 and CpG 3, respectively (Figure 3.22). For saliva, hypermethylation (98%) was displayed in all clones, with one clone (h) showing unmethylation at CpG 8 (Figure 3.22). In vaginal fluid, hypermethylation (97%) was observed, with two clones (e and h) showing an unexpected cytosine nucleotide at CpG 2 and unmethylation and CpG 6, respectively (Figure 3.22). Conversely, in semen for the ZNF282 tDMR, hypomethylation (99%)was observed, with one clone (a) showing methylation (1%) at CpG 4 (Figure 3.22). In Table 3.12 is the summarized results for the ZNF282 tDMR. Statistical analysis showed that methylation patterns between semen and all other body fluids (blood, vaginal fluid and saliva) were significantly different ($p < 0.05$) (Table 3.13). Methylation patterns at individual CpG sites were also significantly different ($p < 0.05$) (Appendix D: Table 1-8) at all eight CpG sites for semen compared to all other body fluids for the ZNF282 tDMR.

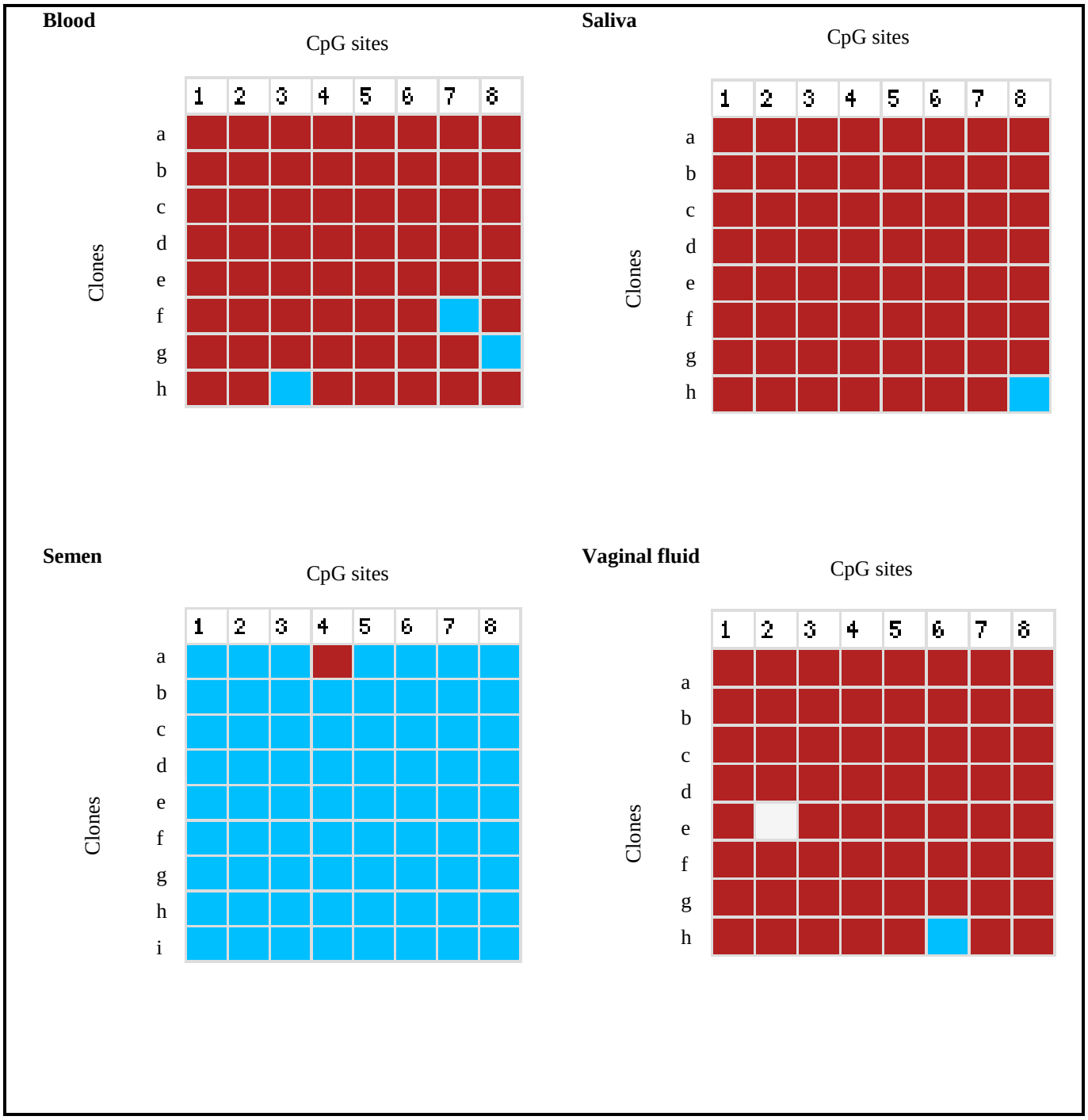


Figure 3.22: Body fluid specific methylation of the ZNF282 tDMR. Each row indicates a single clone of bisulfite PCR products and each column indicates individual CpG site in the region of interest. Different methylation states of the CpG sites are indicated by colours (blue represents unmethylated; red represents methylated; white represents unknown).

Table 3.12: Overview of bisulfite sequencing results for body fluids analysed using BISMA for the ZNF282 tDMR.

Body Fluid	Number of clones sequenced	Overall Methylation (%)	Overall Unmethylation (%)	Unknown CpG sites (%)
Blood	8	95	5	0
Saliva	8	98	2	0
Semen	9	1	99	0
Vaginal fluid	8	97	1.5	1.5

Table 3.13: Pairwise comparison of total DNA methylation for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	Saliva	Semen	Vaginal fluid
Blood	1			
Saliva	0.3380	1		
Semen	0.0000	0.0000	1	
Vaginal fluid	0.6182	1.0000	0.0000	1

3.3.2.2 Methylation profiles of body fluids using the PTPRS 1 tDMR

The PTPRS 1 tDMR primer set targeted five CpG sites. A total of 40 clones were sent for sequencing (10 clones for each body fluid), however only 14 clones were analysed for all body fluids using the BISMA program. The threshold sequence identity percentage was 90%, thus the clones that were not analysed were below the threshold. All saliva and vaginal fluid clones displayed 100% methylation (completely methylated) (Figure 3.23). For blood, hypermethylation (80%) was observed and one clone (b) displayed unmethylation (20%) at CpG 1 and CpG 2 (Figure 3.23). For semen, hypermethylation (93%) was observed with two clones (e and f) displaying unmethylation (7%) at CpG 1 and CpG 2 respectively (Figure 3.23). In Table 3.14 is the summarized results for the PTPRS 1 tDMR.Statistical analysis showed that methylation patterns between blood, vaginal fluid, semen and saliva were not significantly different from each other ($p > 0.05$) (Table 3.15) and that methylation patterns at individual CpG sites were not significantly different ($p > 0.05$) (Appendix E: Table 1-5) for the PTPRS 1 tDMR.

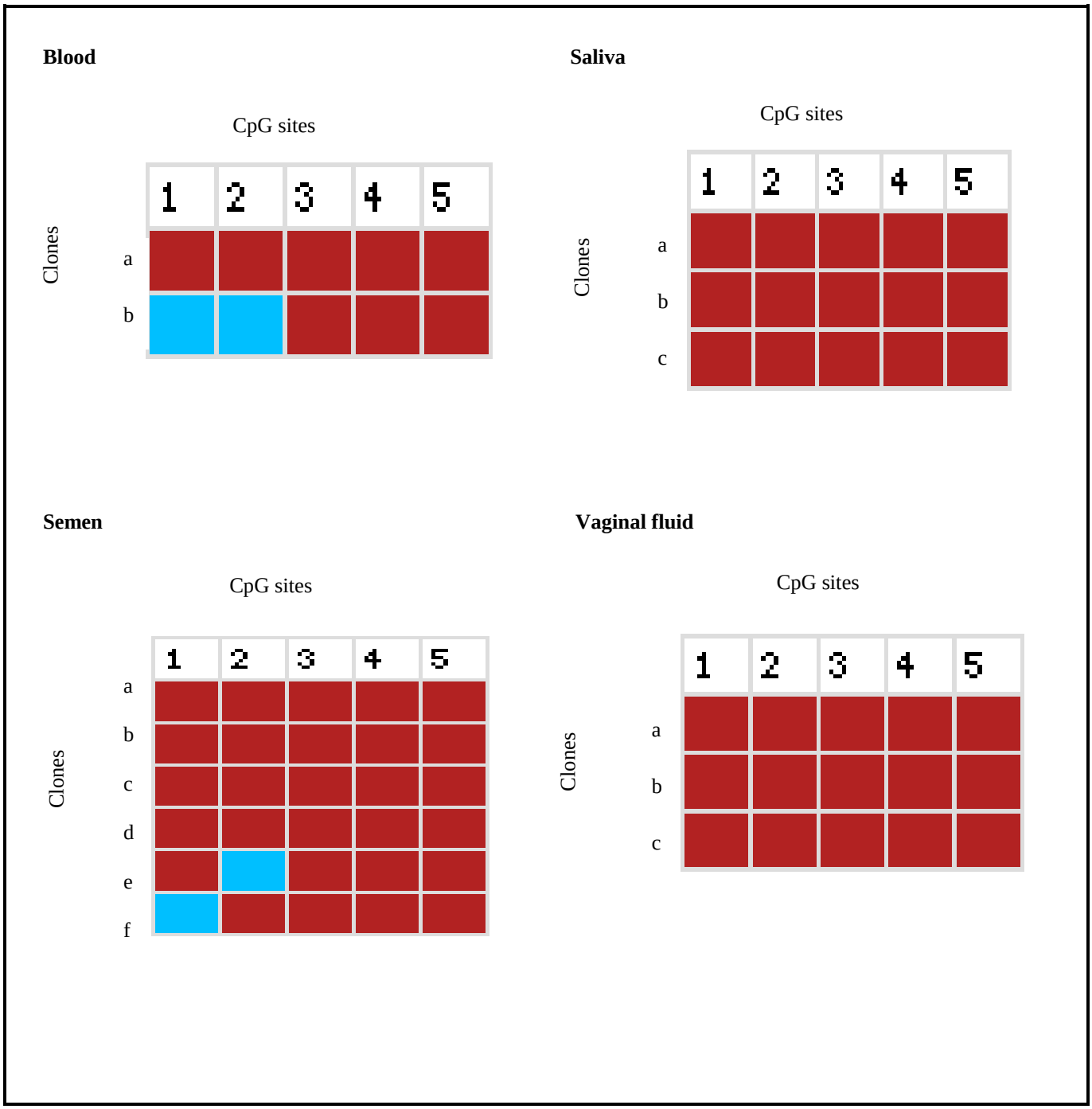


Figure 3.23: Body fluid specific methylation of the PTPRS 1 tDMR. Each row indicates a single clone of bisulfite PCR products and each column indicates individual CpG site in the region of interest. Different methylation states of the CpG sites are indicated by colours (blue represents unmethylated; red represents methylated; white represents unknown).

Table 3.14: Overview of bisulfite sequencing results for body fluids analysed using BISMA for the PTPRS 1 tDMR.

Body Fluid	Number of clones sequenced	Overall Methylation (%)	Overall Unmethylation (%)	CpG sites unknown (%)
Blood	2	80	20	0
Saliva	3	100	0	0
Semen	6	93	7	0
Vaginal fluid	3	100	0	0

Table 3.15: Pairwise comparison of total DNA methylation for PTPRS 1 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	Saliva	Semen	Vaginal fluid
Blood	1			
Saliva	0.4000	1		
Semen	0.2860	1.0000	1	
Vaginal fluid	0.4000	1.0000	0.5000	1

3.3.2.3 Methylation profiles of body fluid using the PTPRS 2 tDMR

The PTPRS 2 tDMR primer set targeted 25 CpG sites. A total of 20 clones were sent for sequencing (five clones for each body fluid), however only 14 clones were analysed for all body fluids using the BISMA program. The threshold sequence identity percentage was 90%, thus the clones that were not analysed were below the threshold. In blood, hypermethylation (96%) was observed in all clones, however one clone (b) displayed a cytosine nucleotide at an unexpected position in CpG 3 and unmethylated at CpG 4 (Figure 3.24). In all clones for saliva, hypermethylation (98%) was displayed and two clones (b and d) showed an unexpected cytosine at CpG 3 and an unmethylation at CpG 23 (Figure 3.24). Semen displayed hypermethylation (71%) with one clone (b) showing an unexpected cytosine at CpG 4 and unmethylation at CpG 21 (Figure 3.24). In clone c in semen for PTPRS 2 tDMR, CpG 1-15, 17, 18, 21-23 all displayed unmethylation. In vaginal fluid, hypermethylation (95%) was observed in all clones, with clones (a, d and e) displaying an unexpected cytosine at CpG 6-CpG 8 and two other clones displaying unmethylation at CpG 4, CpG 10 and CpG 22, respectively. In Table 3.16 is the summarized results for the PTPRS 2 tDMR.Statistical analysis showed that methylation patterns between blood, vaginal fluid, semen and saliva were not statistically different from each other ($p > 0.05$) (Table 3.17) and that methylation patterns at

individual CpG sites were not significantly different ($p > 0.05$) (Appendix F: Table 1-25) for the PTPRS 2 tDMR

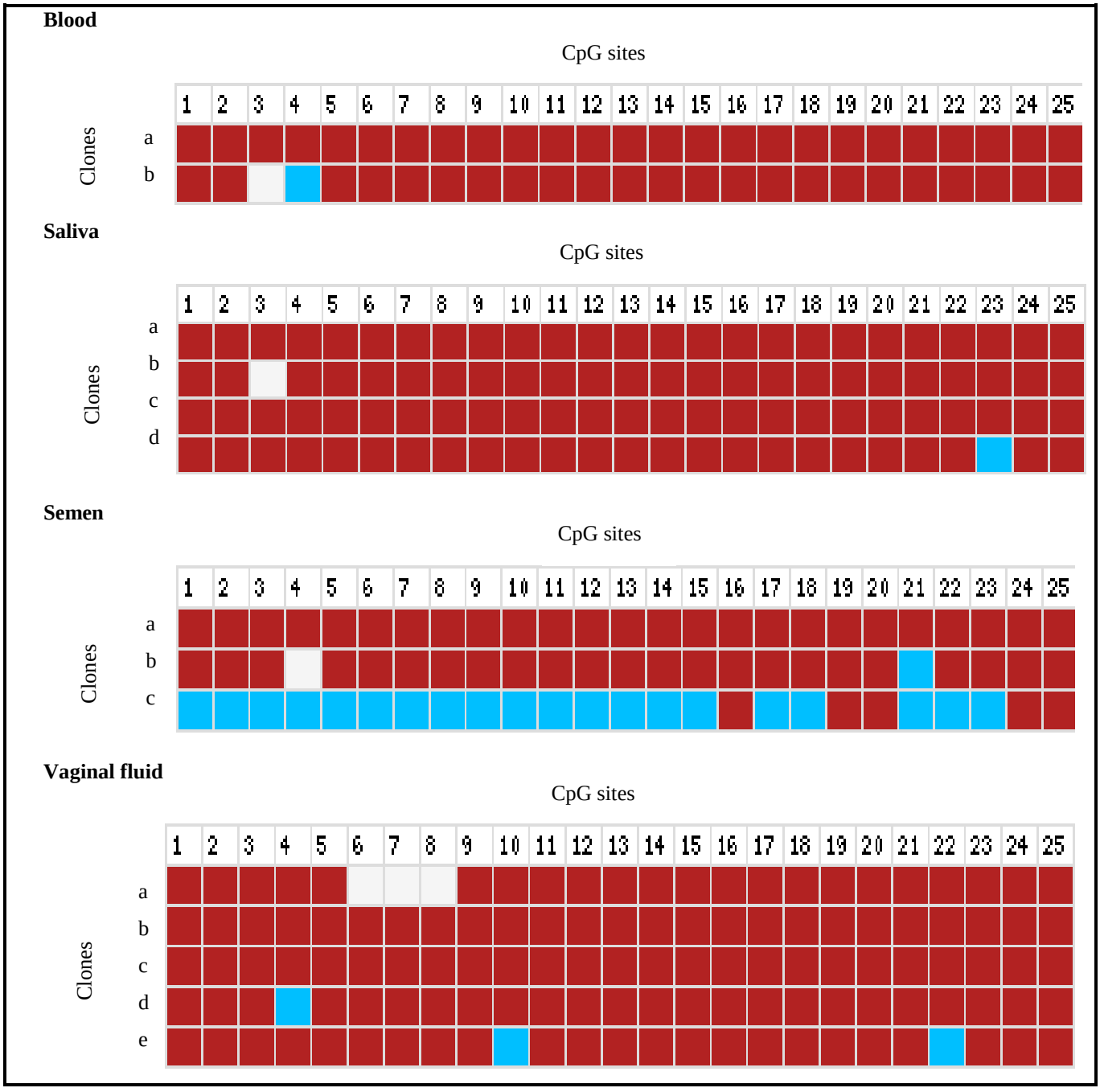


Figure 3.24: Body fluid specific methylation of the PTPRS 2 tDMR. Each row indicates a single clone of bisulfite PCR products and each column indicates individual CpG site in the region of interest. Different methylation states of the CpG sites are indicated by colours (blue represents unmethylated; red represents methylated; white represents unknown).

Table 3.16: Overview of bisulfite sequencing results for body fluids analysed using BISMA for the PTPRS 2 tDMR.

Body Fluid	Number of clones sequenced	Overall Methylation (%)	Overall Unmethylation (%)	CpG sites unknown (%)
Blood	2	96	2	2
Saliva	4	98	1	1
Semen	3	71	28	1
Vaginal fluid	5	95	2.5	2.5

Table 3.17: Pairwise comparison of total DNA methylation for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	Saliva	Semen	Vaginal fluid
Blood	1			
Saliva	0.6000	1		
Semen	1.0000	0.4860	1	
Vaginal fluid	0.7140	1.0000	0.5710	1

3.3.2.4 Methylation profiles of body fluid using the HPCAL1 tDMR

There were no sequencing results obtained for the HPCAL1 tDMR as non-specific amplification was observed for bisulfite PCR in all body fluids. The gel results for bisulfite are shown in Appendix G.

3.4 Discussion

Analysis of methylation status at tDMRs is emerging as a promising method of body fluid identification in forensics. However, so far a limited number of tDMR based markers have been verified (Fu *et al*., 2015). The aim of the present study was to analyse differential DNA methylation in candidate tDMRs for identification of four forensically relevant body fluids: blood, saliva, semen and vaginal fluid. The analysis of these tDMRs will lead to development of novel DNA methylation markers for forensic purposes.

The candidate genes were selected from a previous study (Kader, 2015, MSc Genetics dissertation) based on differential gene expression in blood, saliva, semen and vaginal fluid. Genes overexpressed in the target body fluids were identified from Tissue–specific Gene Expression and Regulation (TiGER) database (Appendix H: Figure 1-3) and gene body CpG islands, spanning an exon/intron region were selected for methylation analysis. It was proposed that differential gene expression could be associated with differential methylation. Several studies have provided evidence of correlation between DNA methylation and gene expression (Kulis *et al.,* 2013; Schilling and Rehli, 2007). Differential expression of genes in different tissues may be regulated by DNA methylation. While promoter hypermethylation is associated with gene repression, gene body methylation is more prevalent and is associated with increased gene expression (Jones, 2012). It has been proposed that gene body methylation possibly inhibits spurious intragenic transcription which slows efficient transcriptional elongation (Wierstra, 2008). The relationship between gene body methylation and gene expression has been described as non-monotonic and bell-shaped (Jjingo *et al*., 2012). This means that genes that are highly or lowly expressed have a low level of gene body methylation and genes that are mid-level expressed have the highest level of gene body methylation (Jjingo *et al*., 2012). Tissue-specific methylation occurs largely within CpG islands found in gene bodies and CpG island shores (Davies *et al*., 2012; Deaton *et al*., 2011; Irizarry *et al*., 2009; Maunakea *et al*., 2010).

As gene expression information was available for blood but not for saliva, semen and vaginal fluid, surrogate tissues representative of each body fluid of interest were chosen. The criteria for selection of surrogate tissue was: 1) it must have expression data associated with it, 2) it should be the main source of the fluid component in question. Hence, cervix was selected as the surrogate tissue for vaginal fluid and prostate and testis were selected as surrogate tissue for semen. Genes which were specifically up regulated in the tissue of interest were chosen.

The *ZNF282* gene showed overexpression in cervix, with median expression in blood, low expression in prostate but no expression in testis (Appendix H: Figure 1). The *PTPRS* gene showed higher expression in prostate, median expression in blood, low expression in testis and no expression in cervix (Appendix H: Figure 2). The *HPCAL1* gene was overexpressed in cervix, median expression in testis and blood, and low expression in prostate (Appendix H: Figure 3). No expression information was available for salivary glands at TiGER database, however, differential expression in any one fluid in comparison to other was considered adequate to select the genes.

In this study, it was proposed that the candidate tDMRs for genes *ZNF282*, *PTPRS* and *HPCAL1* would successfully differentiate between blood, saliva, semen and vaginal fluid. Hence, they were selected for methylation analysis by BS and MSP. To the best of our knowledge these candidate tDMRs have not been previously reported.

3.4.1 Methylation profiling of candidate tDMRs by MSP and BS

A MSP was designed to differentiate between methylated and unmethylated cytosines in four candidate tDMRs in genes, *ZNF282*, *PTPRS* and *HPCAL1*. MSP is a cost effective technique that enables detection of methylation in small quantities of DNA and allows detection of methylation and unmethylation simultaneously of a single sample (Esteller *et al*., 1999; Herman *et al*., 1996). DNA is first denatured to create single-stranded DNA and then modified with sodium bisulfite followed by PCR amplification using two pairs of primers, with one pair specific for methylated DNA and the other unmethylated DNA. Primer design for both MSP and BS is important for efficient analysis of methylation patterns. For maximum discrimination between methylated and unmethylated sites, MSP primers contain a minimum of one CpG site at the 3' end as well as many CpG sites within the primer sequences (Li and Dahiya, 2002).

However, as MSP is a qualitative method, therefore, BS was used for further validation to confirm and quantify body fluid specific differential DNA methylation of the candidate tDMRs used in the present study. BS is the gold standard method of methylation analysis because it is a quantitative method which provides detailed information on the methylation status of individual CpG sites for the region of interest (Huang *et al*., 2013a). The sodium bisulfite modified DNA is PCR amplified using primers that are specific for the modified DNA but do not contain any CpG sites in their sequence. The resulting PCR product is sequenced directly or after cloning (Li and Dahiya, 2002).

3.4.1.1 ZNF282 tDMR

With MSP analysis, the ZNF282 tDMR was successful in differentiating semen from blood, saliva and vaginal fluid because of the hypomethylation pattern observed in semen and hypermethylation pattern observed in all other body fluids examined. Observation of amplification by unmethylated primers in semen and amplification by methylated primers in all other body fluid indicates that all eight CpG sites in semen were hypomethylated and all CpG sites in all other body fluids were hypermethylated.

BS analysis confirmed the results that were obtained by MSP. Semen showed hypomethylation in all clones (nine) sequenced and all other body fluids showed hypermethylation pattern. Statistical analysis showed that there was no difference observed between all clones for each body fluid ($p > 0.05$). Based on the methylation data, the statistical analysis indicated that blood, saliva and vaginal fluid were not significantly different form each other however, semen differed from all other body fluids ($p < 0.05$) at each of the eight CpG sites as well as for overall methylation. Thus, ZNF282 tDMR can be successfully used as a novel semen-specific hypomethylation marker.

Previous studies have also reported semen-specific hypomethylation markers, such as, cg22407458 (Forat *et al*., 2016), PRMT (Lee *et al*., 2012), ZC3H12D (Madi *et al*., 2012), cg04382920 (SEU1) and cg11768416 (SEU2) (Vidaki *et al*., 2016). A study by Lee *et al.* (2012) used MSP and BS analysis to identify tDMRs in two genes *DACT1* and *USP49* which were able to differentiate semen from blood, saliva, menstrual blood and vaginal fluid.

3.4.1.2 PTPRS 1 and PTPRS 2 tDMRs

Two CpG islands were targeted in the *PTPRS* gene and were referred to as PTPRS 1 and PTPRS 2 tDMR. The PTPRS 1 tDMR differentiated semen from blood, vaginal fluid and saliva by MSP analysis because both hypermethylation and hypomethylation was observed in semen and only hypermethylation was seen in all other body fluids. Observation of amplification by both methylated and unmethylated primers indicated that some target CpG sites were methylated and some were unmethylated. Presence of both methylation and unmethylation at the same site could be due to allele-specific methylation which arises from parental alleles having different methylation patterns (Meaburn *et al*., 2010; Shoemaker *et al*., 2010). Previous studies have identified semen-specific hypermethylated markers cg05261336 (Lin *et al*., 2016) and cg17610929 (gene *ACCN4*) and cg23521140 (gene *ACCN4*) (Park *et al*., 2014) and cg17621389 and cg26763284 (Lee *et al*., 2015) and hypomethylated markers mentioned above.

The PTPRS 2 tDMR displayed both hypermethylation and hypomethylation patterns for all body fluids. This result is a reflection on heterogeneity of methylation and cell types within a tissue (Song *et al*., 2005). Therefore, PTPRS 2 tDMR was unable to differentiate between body fluids as no unique methylation pattern in any one body fluid was observed.

Based on the result of the MSP study, it was expected that the PTPRS 1 tDMR would be able to distinguish semen from all other body fluids and that PTPRS 2 tDMR would not differentiate between body fluids. However, BS analysis did not show any significant difference in overall methylation for both PTPRS 1 and PTPRS 2 tDMRs as well as at individual CpG sites (five and 25, respectively) ($p > 0.05$). There was no significant difference observed between all clones for each body fluid ($p > 0.05$) for both tDMRs. A probable explanation for no differentiation between all body fluids for PTPRS 1 tDMR could be because only a few clones met the sequence identity threshold (90%) set by BISMA analysis tool. Therefore, analyzing more clones could result in confirmation of methylation differences between body fluids. Though MSP indicates PTPRS 1 tDMR to be a semen-specific marker, further evaluation using more clones is required.

3.4.1.3 HPCAL1 tDMR

The HPCAL1 candidate tDMR displayed hypermethylation in blood, semen and vaginal fluid, conversely saliva samples displayed hypomethylation at the target CpG sites. Thus, the tDMR could be a candidate saliva-specific hypomethylation marker. Previous studies have identified saliva-specific hypermethylation markers: cg21597595 (named *Spei1*) (Forat *et al*., 2016), cg09652652-2d (Lee *et al*., 2015), cg09107912 (gene *FNDC1*) and cg16732616 (gene *DMRTA2*) (Lin *et al*., 2016), BCAS4 (Madi *et al*., 2012) and cg26107890 (gene *SLC12A8*) and cg20691722 (gene *SOX2OT*) (Park *et al*., 2014). However, the previously reported markers just targeted one CpG site while HPCAL1 tDMR targets 10 CpG sites, thus confirming more reliability.

The HPCAL1 tDMR could not be further analysed using BS due to non-specific amplification that was consistently observed in PCR reactions. During PCR amplification of bisulfite converted DNA, all cytosine nucleotides that are not methylated are converted to thymine nucleotides, thus the DNA sequence predominantly consists of adenine, guanine and thymine (approximately 50%). Therefore, non-specific amplification and low specificity of hybridization with primers may be due to polymerase slippage because DNA sequences are adenine and thymine rich (Arányi *et al*., 2006; Tusnády *et al*., 2005).

3.5. Conclusions

The present study analysed methylation profiles of four candidate tDMRs in genes *ZNF282*, *PTPRS* (PTPRS 1 and PTPRS 2 tDMRs) and *HPCAL1*. The objective was to target the identified tDMRs to design PCR primers for efficient identification of semen, saliva, blood and vaginal fluid by MSP and BS. Among the tDMRs examined using MSP, ZNF282 tDMR and PTPRS 1 tDMR showed potential to be used as semen-specific markers and HPCAL1 as a saliva-specific marker. Since MSP is a qualitative method, BS was used for further validation to quantify and confirm body fluid specific differential DNA methylation of the candidate tDMR markers. BS results confirmed that the ZNF282 tDMR showed semen-specific differential DNA methylation from all other body fluids and can be used as a biomarker in forensic applications. However, PTPRS 1 and HPCAL 1 tDMR require further evaluation using a bigger sample size and sequencing of more clones. Future research will involve tissue-specific DNA methylation profiling to identify and validate more markers for blood, saliva, semen and vaginal fluid. Furthermore, the development of multiplex methods to analyse body fluids with numerous markers in one reaction (Vidaki *et al*., 2016). In addition, multiplexing reduces the time and amount of starting DNA required which is important in a forensics setting where DNA is either scarce or degraded (Lee *et al*., 2015). Further investigation on the sensitivity of methods and stability of candidate tDMRs methylation profiles used for differentiation between body fluids is necessary to allow efficient use for forensic purposes.

Chapter 4

Analysis of the sensitivity and stability of methylation profiles of tDMR markers for human body fluid identification

Abstract

The use of tissue-specific differentially methylated regions (tDMRs) is emerging as a promising method for body fluid identification. In the previous chapter, methylation specific PCR (MSP) and bisulfite sequencing (BS) was used to validate differential methylation in candidate tDMRs for identification of blood, saliva, semen and vaginal fluid. The ZNF282 tDMR was selected as a suitable marker for semen identification based on consistent hypomethylation patterns with both MSP and BS analysis. However, to allow successful use of methylation markers the stability of their methylation profile needs to be tested under environmental conditions found commonly at crime scenes and the detection limit of methylation methods should also be evaluated. Thus, the present study was undertaken to determine the sensitivity of methylation analysis methods and to analyse the stability of methylation profile of ZNF282 tDMR marker under forensic simulation conditions. For the sensitivity test, DNA from blood was five-fold serially diluted from 25 ng to 1 ng and tested by MSP. DNA concentrations greater than 10 ng provided reliable detection. For the stability study blood was deposited on a piece of cloth and saliva, semen and vaginal fluid were deposited on cotton swabs and kept at room temperature, wet in an exsiccator, outside on the ground, sprayed with alcohol and bleach for 50 days and compared to samples that were not exposed to any environmental conditions (0 days). Semen displayed maximum DNA degradation under all conditions and vaginal fluid was most stable. Blood, saliva and semen provided results only for outside on the ground. Blood, saliva, semen and vaginal fluid provided results for outside on the ground condition. Overall a decrease in methylation level was observed for all body fluids, except, semen. However, saliva and semen displayed a significant difference ($p < 0.05$) in methylation level for outside on the ground. Vaginal fluid provided results in all conditions (at room temperature, wet in an exsiccator, outside on the ground, sprayed with alcohol and bleach). No significant change in methylation profile of ZNF282 tDMR marker was observed after 50 days under simulated conditions. Thus, ZNF282 tDMR maintained differential methylation profile under environmental insults and can be used as a novel marker for semen identification. Future studies will include analysis of methylation levels of methylation markers under diverse simulated conditions and for different human ethnic groups.

4.1 Introduction

In forensic investigations, genetic fingerprinting is an efficient method used to identify suspects, however, analysis of biological material that is found at crime scenes can be used to determine the course of events that have occurred as well as link the evidence found to the crime committed (An *et al*., 2013; Choi *et al*., 2014; Lee *et al*., 2012; Ma *et al*., 2013). Most times there is very little biological material found at a crime scene and identifying the biological material in a non-destructive manner is essential for further analysis (Kader and Ghai, 2015). Additionally, challenges arise when body fluids are exposed to various environmental conditions which affect the stability and integrity of the DNA.

The most common forensically relevant body fluids found at crime scenes are: blood, saliva, semen and vaginal fluid. The composition of each fluid is unique, hence, current methods identify body fluids based on the components of each body fluid. The main components of human body fluids such as, blood, saliva, semen and vaginal fluid are shown in Table 4.1 (Gaensslen, 1983; Greenfield and Sloan, 2003; Virkler and Lednev, 2009). There are several components that are common in more than one body fluid, however the difference in contribution of the component makes tests effective. For example, amylase is present in a large quantity in saliva compared to smaller quantities in semen and vaginal fluid. Other examples are, the difference in the ratio of citrate to lactate present in semen and vaginal fluid. Urea is a major component in urine but it is also found in semen and sweat, however, it is used to identify urine due to the high concentration in urine (Virkler and Lednev, 2009). Thus, using DNA molecules as an alternate source of identification for body fluids will be beneficial since it is a stable molecule.

Analysis of DNA is a standard forensic method for investigating and solving various criminal cases. In majority of cases a limited amount of biological material is found and/or DNA that is highly degraded resulting in fragmented DNA (Lee *et al*., 2013). Current methods that are used for identification of biological material require DNA that is intact and of good quality (Diegoli *et al*., 2012). Extracting essential information is a difficult task especially when the samples have been damaged (Alvarez-Cubero *et al*., 2012). Thus, it is important that the methods used to identify body fluids are definitive, regardless of the factors that may affect the stability of the DNA. The successful recovery of DNA and the extent to which the DNA is degraded depends on two main factors, time and environmental conditions (Fondevila *et al*., 2008). A study by Raymond *et al*. (2009) showed that environmental influences have a greater effect on

DNA degradation in comparison to the time of exposure. Environmental influences such as temperature, humidity, pH and soil chemistry could modify the rate of degradation (Burger *et al*., 1999; Fondevila *et al*., 2008; Forat *et al*., 2016; Hall and Ballantyne, 2004).

Table 4.1: Composition of human body fluids (Altman and Katz, 1961; Li, 2008; Spinrad, 1994).

Blood	**Saliva**	**Semen**	**Vaginal fluid**
-Hemoglobin	-Amylase	-Acid phosphatase	-Acid phosphatase
-Fibrinogen	-Lysozyme	-Prostate-specific antigen	-Lactic acid
-Erythrocytes	-Mucin	-Spermatozoa	-Citric acid
-Albumin	-Buccal epithelial cells	-Choline	-Urea
-Glucose	-Thiocyanate	-Spermine	-Vaginal peptidase
-Immunoglobulins	-Potassium	Semenogelin	-Glycogenated epithelial
	-Bicarbonate	-Zinc	cells
	-Phosphorus	-Citric acid	-Acetic acid
	-Glucose	-Lactic acid	-Pyridine
	-Immunoglobulins	-Fructose	-Squalene
		-Urea	-Immunoglobulins
		-Ascorbic acid	
		-Immunoglobulins	

Forensic biological material that are exposed to various physical, biochemical and microbiological factors at crime sites, degrade or destroy the biological components that are used for body fluid or tissue identification (Forat *et al*., 2016). DNA molecules exposed to high temperatures results in a faster degradation rate and low temperatures result in reduction of the rate of degradation (Höss *et al*., 1996; Willerslev *et al*., 2004). At high temperatures, the presence of microorganisms are greater and chemical reactions are sped up. Also, high temperature and humidity levels increase the effects of UV radiation on DNA damage (Foran, 2006). Dry environments decrease the rate of degradation, however, the presence of water would increase the rate of degradation as moisture encourages the growth of bacteria and also provides the substrate with hydrolytic enzymes (Gill-King, 1997). Furthermore, in dry conditions the hydrolytic and oxidative damage of DNA is decreased. The presence of microorganisms as well as their metabolites can lead to the complete degradation of DNA (Burger *et al*., 1999). An environment that has a pH that is low results in DNA as well as protective matter (bone and teeth) being destroyed (Burger *et al*., 1999). A study by Ambers *et al*. (2014) and Raymond *et al*. (2009) found that DNA recovery of samples left in outside surfaces decreased significantly overtime, whereas if samples are found in a cool and dark, low

traffic environment or dry state that DNA recovery is stable. Thus, environmental conditions play a role in the quality and quantity of DNA recovered from body fluids.

Identification of tissue-specific differentially methylated regions (tDMRs) markers seems to be a beneficial method for the identification of body fluids due to high specificity and sensitivity and its compatibility with short tandem repeat (STR) typing. Forensic researchers have been putting in an effort to identify and develop new tDMR markers for identification of body fluids. Even though DNA is stable when exposed to exogenous conditions, very little is known about the stability of DNA that is methylated. Hence, determining the stability of methylation profiles of tDMRs for human body fluids requires further research. Development of novel tDMR markers for accurate and reliable identification and differentiation of human body fluids is essential. In the present chapter, the sensitivity of four tDMRs (ZNF282, PTPRS 1, PTPRS 2 and HPCAL1) were analysed and the stability of the ZNF282 tDMR was analysed under simulated forensic conditions. In the previous study, the ZNF282 tDMR was validated as a semen-specific hypomethylation marker by both methylation specific PCR (MSP) and bisulfite sequencing (BS). The other tDMRs (PTPRS 1, PTPRS 2 and HPCAL1) were not found to be specific to a single body fluid. To ensure the efficacy of these tDMR markers, the present study was performed to determine the sensitivity and stability of using tDMRs for identification of human body fluids and its applicability in forensic casework.

4.2 Materials and methods

4.2.1 Ethics approval

The study was conducted according to the methods specified by the Biomedical Research Ethics Committee (BREC) at the University of KwaZulu-Natal. The BREC reference number is BE187/16 (sub-study of BE221/14). Samples were collected from volunteers after a written consent was signed and a questionnaire answered by each participant (Appendix A).

4.2.2 Sample collection, storage of samples and DNA extraction

Whole blood was collected in 4 ml EDTA tubes from 20 volunteers at Lab 24 laboratory, Mount Edgecombe, Durban. Blood samples were collected and transported to the Genetics Laboratory in ice boxes with frozen ice packs. Vaginal fluid was collected from 20 volunteers at the King Dinizulu Hospital, Durban using sterile cotton swabs (Dry Swab, Lasec). Semen and saliva samples were collected from 10 volunteers. Freshly ejaculated semen was collected in plastic cups from volunteers at the King Dinizulu Hospital and University of KwaZulu-Natal, Westville campus. Saliva was collected from University of KwaZulu-Natal, Westville in 50 mL microcentrifuge tubes.

DNA was extracted from blood, saliva, semen and vaginal fluid using Quick-gDNA MiniPrep kit (Zymo Research, Irvine, California, United States) according to manufacturer's instructions. Extracted DNA was quantified using a NanoDrop (NanoDrop™ 2000 Spectrophotometer, Waltham, Massachusetts, United States). Following quantification, isolated DNA samples were stored frozen at -20°C until processing (to maintain DNA integrity and arrest microbial growth). The concentration of blood, saliva, semen and vaginal fluid is listed in Table 1-Table 5, Appendix I.

4.2.3 Bisulfite conversion

The DNA methylation sites on the extracted DNA has to be preserved when a standard PCR amplification is run. Therefore, the genomic DNA was treated with sodium bisulfite in order to convert the unmethylated cytosines to uracil and methylated cytosine remain the same (Silva *et al*., 2016) using the EZ DNA Methylation Gold Kit (Zymo Research, Irvine, California, United States). In this method the converted uracil is replaced by thymine during the PCR amplification process. The bisulfite modified DNA was amplified by site specific PCR primers designed to amplify the bisulfite modified target regions.

4.2.4 Primer design

Primers that were used are presented in Table 4.2 for MSP reactions and Table 4.3 for BS reactions

Table 4.2: MSP primers for sensitivity analysis.

tDMR		Primer Sequence	Amplicon Size(bp)	Annealing Temperature (ºC)
PTPRS 1	[1]MF	CGTACGGTTATCGTGTATTTGAC	216	49
	[2]MR	AACCACGTATATATTCCGACTTACG		
PTPRS 2	MF	ATGTAAAGAATTAAGTCGTACGGTC	168	46
	MR	CGTAAACAACCTACTAAAAAACGAA		
HPCAL1	MF	GTTTTGTAGGTGTAGTCGTCGTC	113	48
	MR	AACTCGTAATCGATAAACTCCGTA		
ZNF282	MF	GGAGTGCGAGAAGATTTATAGTC	210	48
	MR	TCCTTAAACGACTCCTTATAACGAA		

[1]methylated forward, [2]methylated reverse

Table 4.3: BS primer for ZNF282 tDMR for forensic simulation study.

tDMR		Primer Sequence	Amplicon Size (bp)	Annealing Temperature(ºC)
ZNF282	Forward	GGATTTTTTAGATTTGTTTGGTTTG	180	46
	Reverse	TCCCTAATAACTTCCCCTAATAACC		

4.2.5 MSP sensitivity

Sensitivity testing was carried out using a five-fold serial dilution of DNA (25, 20, 15, 10, 5,1 ng) to evaluate the minimum quantity required to obtain the predictable body fluid methylation profile. Three different blood samples were used for sensitivity testing.

4.2.5.1 MSP PCR reaction

PCR amplification was carried out using a ZymoTaq^{TM} DNA Polymerase Kit (Zymo Research, Irvine, California, United States) in a 25 µL reaction volume containing 12 µL of ZymoTaq PCR Buffer, 0.25 µL of dNTP mix (25 µM of each dNTP), 0.15 µM of each primer (Inqaba Biotec), 0.2 µL of ZymoTaq DNA Polymerase (5U/µL) and 1 µL of bisulfite converted DNA. The PCR was conducted in a BIORAD T100^{TM} Thermal Cycler (Bio-Rad Laboratories, Hercules, California, United States) under the following conditions: 95ºC for 10 minutes, 30

cycles of 95ºC for 30 seconds, varying annealing temperatures for 30 seconds (Table 4.2), and 72ºC for 60 seconds and a final extension of 72ºC for 7 minutes. The amplified products were analysed by agarose gel electrophoresis using 2 % agarose gel stained with 10mg/ml ethidium bromide at 60 volts for 60 minutes.

4.2.6 Forensic simulation study

To simulate forensic conditions, blood, saliva, semen and vaginal fluid were exposed to five types of environmental influences: dry at room temperature (condition A), wet in an exsiccator (condition B), outside on the ground (condition C), sprayed with alcohol (condition D) and sprayed with bleach (condition E). Saliva, semen and vaginal fluid were kept on cotton swabs and blood on a piece of cloth for each condition (Table 4.4). A total of 20 samples, 5 samples each of blood, saliva, semen and vaginal fluid were analysed after t0 days (samples not exposed to any environmental insults and analysed immediately after sampling). A total of 100 samples, 25 samples each of blood, saliva, semen and vaginal fluid were analysed (five samples each for every condition) for the forensic simulation study over t50 days. Five replicates for each body fluid and condition was prepared. The replicate DNA samples of the same body fluids in the same condition were pooled together such that 100 ng of DNA sample was included. The pooled DNA was modified by bisulfite treatment. The initial analysis was done immediately after sample collection and DNA was isolated from 20 samples (t0 days). Subsequently, DNA was isolated from 100 samples which were tested after t50 days of storage in five types of environmental influences. In Table 4.5 is the summary of the number of samples subjected to five different environmental conditions for two different time periods. This was done to test the effect of environmental insults on methylation profile of each body fluid to determine the extent of DNA degradation, subsequent DNA recovery and methylation profile.

Table 4.4: Sample collection and preparation of human body fluids for forensic validation according to the recommendations by Setzer *et al*. (2008).

Body Fluid	Collection method	Amount of Body Fluid utilized (µl)	Type of material on which each body fluid was applied
Saliva	Sterile 50ml centrifuge tube	50	Cotton swab
Blood	Sterile EDTA collection tube	50	0.5cm X 0.5cm Pieces of plain cloth
Semen	Sterile Plastic cups	50	Cotton swab
Vaginal Fluid	Sterile cotton swabs	50	Cotton swab

Table 4.5: Number of samples subjected to five different environmental conditions for two different time periods. The 20 samples consist of five blood, five saliva, five semen and five vaginal fluid samples.

Environmental conditions	Time (days)	Number of samples	Time (days)	Number of samples per body fluid
None	0	20	50	0
Dry at room temperature	-	-	50	20
Wet in an exsiccator	-	-	50	20
Outside on the ground	-	-	50	20
Sprayed with alcohol	-	-	50	20
Sprayed with bleach	-	-	50	20
TOTAL		**20**		**100**

4.2.7 BS protocol

4.2.7.1 BS PCR reaction

PCR amplification was carried out using a ZymoTaq™ DNA Polymerase Kit in a 50 µL reaction volume containing 24 µL of ZymoTaq PCR Buffer, 0.5 µL of dNTP mix (25 µM of each dNTP), 0.15 µM of each primer (Inqaba Biotec) and 0.4 µL of DNA Polymerase (5U/µL) and 1 µL of bisulfite converted DNA. The PCR was conducted on a BIORAD T100™ Thermal Cycler under the following conditions: 95ºC for 10 minutes, 30 cycles of 95ºC for 30 seconds, 46ºC for 30 seconds, and 72ºC for 60 seconds and a final extension of 72ºC for 7 minutes. An aliquot of the amplified products (five microlitres) was analysed by agarose gel electrophoresis using 2 % agarose gel stained with 10mg/ml ethidium bromide at 60 volts for 60 minutes to confirm the presence of the amplicon.

4.2.7.2 Cloning and sequencing of BS reaction products

The remaining amplified products was then run on a 1% agarose gel and the bands were excised and purified using a Thermo scientific GeneJet Gel Extraction Kit according to the manufacturers protocol. The purified products were ligated and cloned into pJET1.2 vector using the Thermo Scientific CloneJet PCR cloning Kit according to the manufacturers protocol. Following ligation, an aliquot of the ligation reaction (five microlitres) was transformed using prepared chemically competent cells (*Escherichia coli* - DH5α). The transformed cells were then grown overnight on LB/amp plates. Five positive clones were isolated from each body

fluid in each condition. The plasmids were purified with the Thermo Scientific GeneJet Plasmid Miniprep Kit according to the manufacturers protocol. The purified plasmids were sent to Inqaba biotec for sequencing and this was followed by analysis on online software Bisulfite Sequencing DNA Methylation Analysis (BISMA) (http://services.ibc.uni-stuttgart.de/BDPC/BISMA/). To ensure that the cloning process was efficient a positive and negative control was also plated on LB/amp plates. The positive control was positive samples from trial runs and the negative control was DH5α cells. Figure 4.1 shows the overview of the cloning process.

4.2.8 Bisulfite sequencing data analysis and DNA methylation profiling

The BISMA web application was used for analysis of bisulfite sequencing results. The methylation results were compiled based on the methylation status results from each body fluid sample. To determine the methylation levels of candidate markers in the different body fluid samples, sequencing data were aligned against *in silico*-converted genomic reference sequences using BISMA. The output files were compiled using Bisulfite Sequencing Data Presentation and Compilation (BDPC) web application (http://biochem.jacobs-university.de/BDPC/) to compile the derived information and compare results from each marker in the different body fluid samples (Rohde *et al*., 2008).

4.2.9 Statistical analysis

To evaluate the differences in methylation under forensic simulation conditions, statistical analyses was carried out using the web application, QUantification tool for Methylation Analysis (QUMA) (http://quma.cdb.riken.jp/). Pairwise comparisons of methylation were made for each marker and CpG locus using Chi-square test or Fisher's exact test as appropriate Differences were statistically significant when *p*-values were less than 0.05.

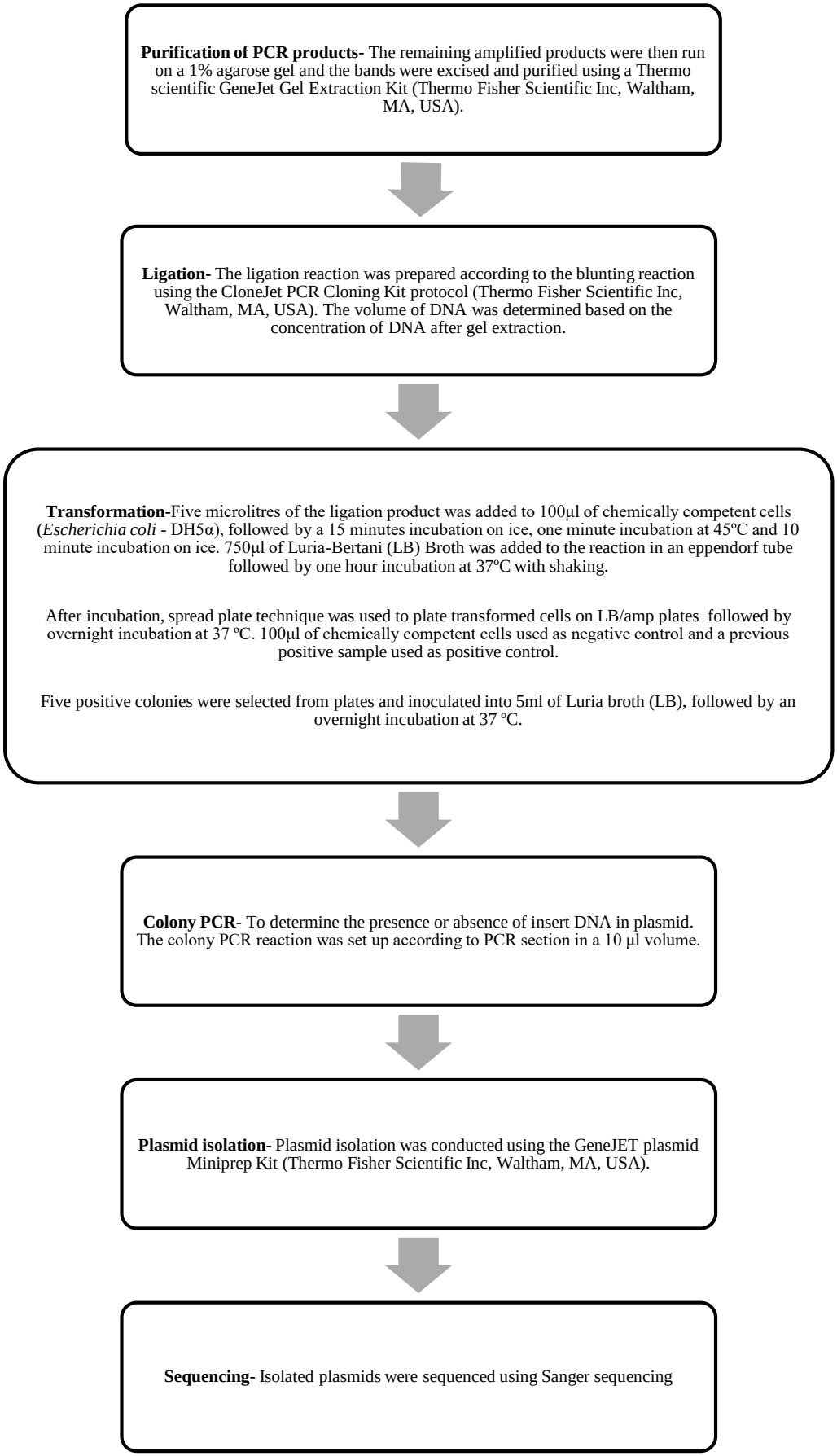


Figure 4.1: Flowchart showing the overview of the BS protocol.

4.3 Results

4.3.1 Sensitivity test

The sensitivity study was done to determine the minimum quantity of DNA with which accurate and reliable results could be obtained for methylation analysis. One of the four body fluids (blood) was selected for the sensitivity test and MSP was selected as method of methylation analysis as it is sensitive and less time consuming. DNA from blood was five-fold serially diluted from 25, 20, 15, 10, 5, to 1 ng. All dilutions were bisulfite modified.

The detection limit with HPCAL1 and PTPRS 1 tDMR was 5 ng/µl (Figure 4.2a and 4.2b). PTPRS 2 showed amplification with minimum of 15 ng DNA (Figure 4.2c). And for ZNF282 tDMR marker (Figure 4.2d) the detection limit was 10 ng. The MSP reaction was done in triplicate and similar results were obtained.

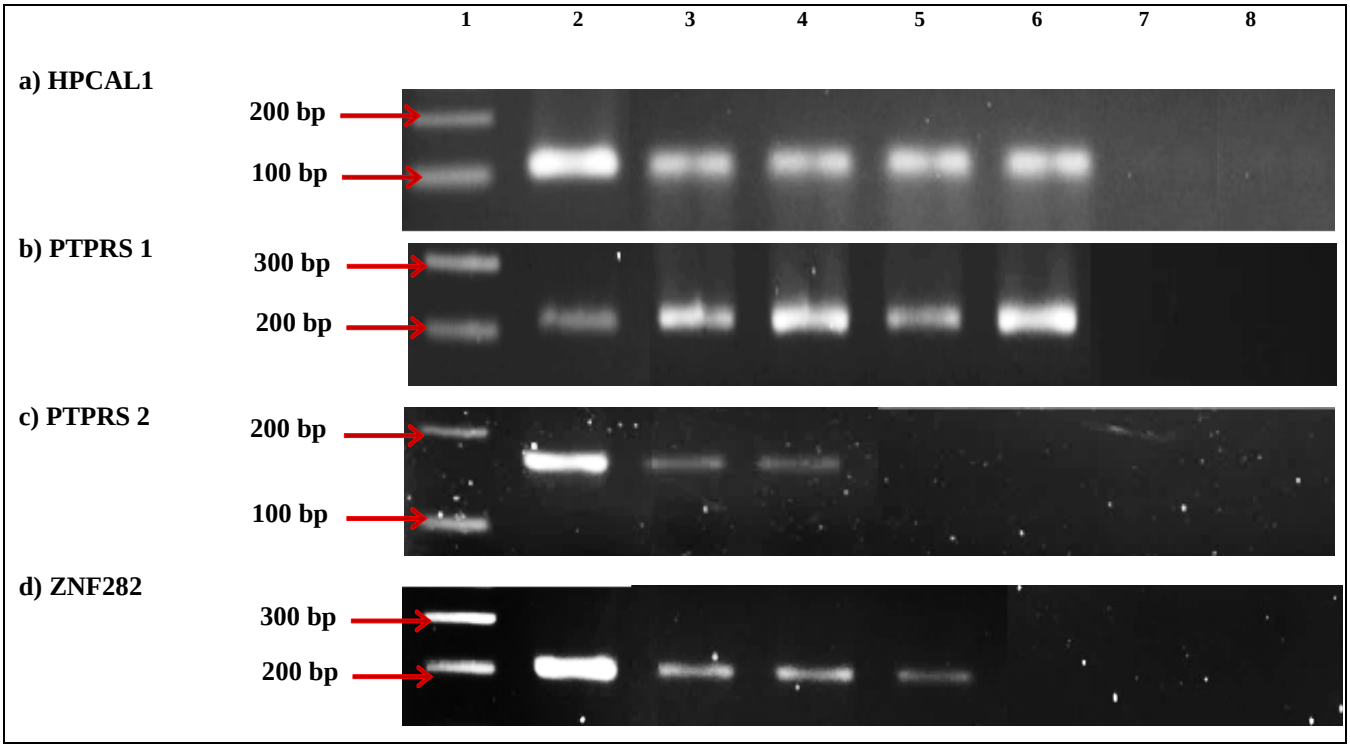


Figure 4.2: Evaluation of senstivity of MSP reaction. Products with primers specific for methylated cytosine. a) HPCAL1 tDMR b) PTPRS 1 tDMR c) PTPRS 2 tDMR d) ZNF282 tDMR. Lane 1 (a, b, c, d): 100bp ladder (Thermo scientific), Lane 2 – Lane 7 (a, b, c, d): The five-fold dilutions of blood with DNA concentrations (25, 20, 15, 10, 5, to 1 ng) and Lane 8: No template control.

4.3.2. Forensic simulation study

A forensic simulation study was conducted to determine the robustness and stability of the methylation profile of the ZNF282 tDMR marker in human blood, saliva, semen and vaginal fluid under five different simulated environmental conditions: dry at room temperature (condition A), wet in an exsiccator (condition B), outside on the ground (condition C), sprayed with alcohol (condition D) and sprayed with bleach (condition E). Two factors that could influence DNA methylation patterns were controlled in this study: age (time of exposure) and simulated environmental conditions.

4.3.2.1 Methylation profile of body fluids using the ZNF282 tDMR analysed immediately after collection (t0 days)

The ZNF282 tDMR primer set targets eight CpG sites. A total of 20 clones were sequenced (five clones for each body fluid) and sequencing results were analysed for all body fluids using the BISMA program. Overall hypermethylation (100%) was displayed for blood and saliva (Figure 4.3). For vaginal fluid, hypermethylation (98%) was observed, with one clones (e) showing an unexpected cytosine nucleotide at CpG 2 (Figure 4.3). Conversely, in semen, as expected, hypomethylation (100%) was observed (Figure 4.3). Table 4.6 shows the summary of bisulfite sequencing results for body fluids analysed using BISMA for the ZNF282 tDMR at t0 days. Statistical analysis showed that overall methylation patterns between semen and all other body fluids (blood, vaginal fluid and saliva) were significantly different ($p < 0.05$) (Table 4.7). Methylation patterns at individual CpG sites were also significantly different ($p < 0.05$) at all eight CpG sites for semen compared to all other body fluids for the ZNF282 tDMR (Appendix J: Table 1-8).

Table 4.6: Overview of bisulfite sequencing results for body fluids analysed using BISMA for the ZNF282 tDMR at t0 days.

Body Fluid	Number of clones that produced results	Overall Methylation (%)	Overall Unmethylation (%)	Unknown CpG sites (%)
Blood	5	100	0	0
Saliva	5	100	0	0
Semen	5	0	0	0
Vaginal fluid	5	98	0	2

Table 4.7: Pairwise comparison of overall DNA methylation for ZNF282 tDMR marker in blood, saliva, semen and vaginal fluid at t0 days. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.0008	0.0008	1	
Vaginal fluid	1.0000	1.0000	0.0008	1

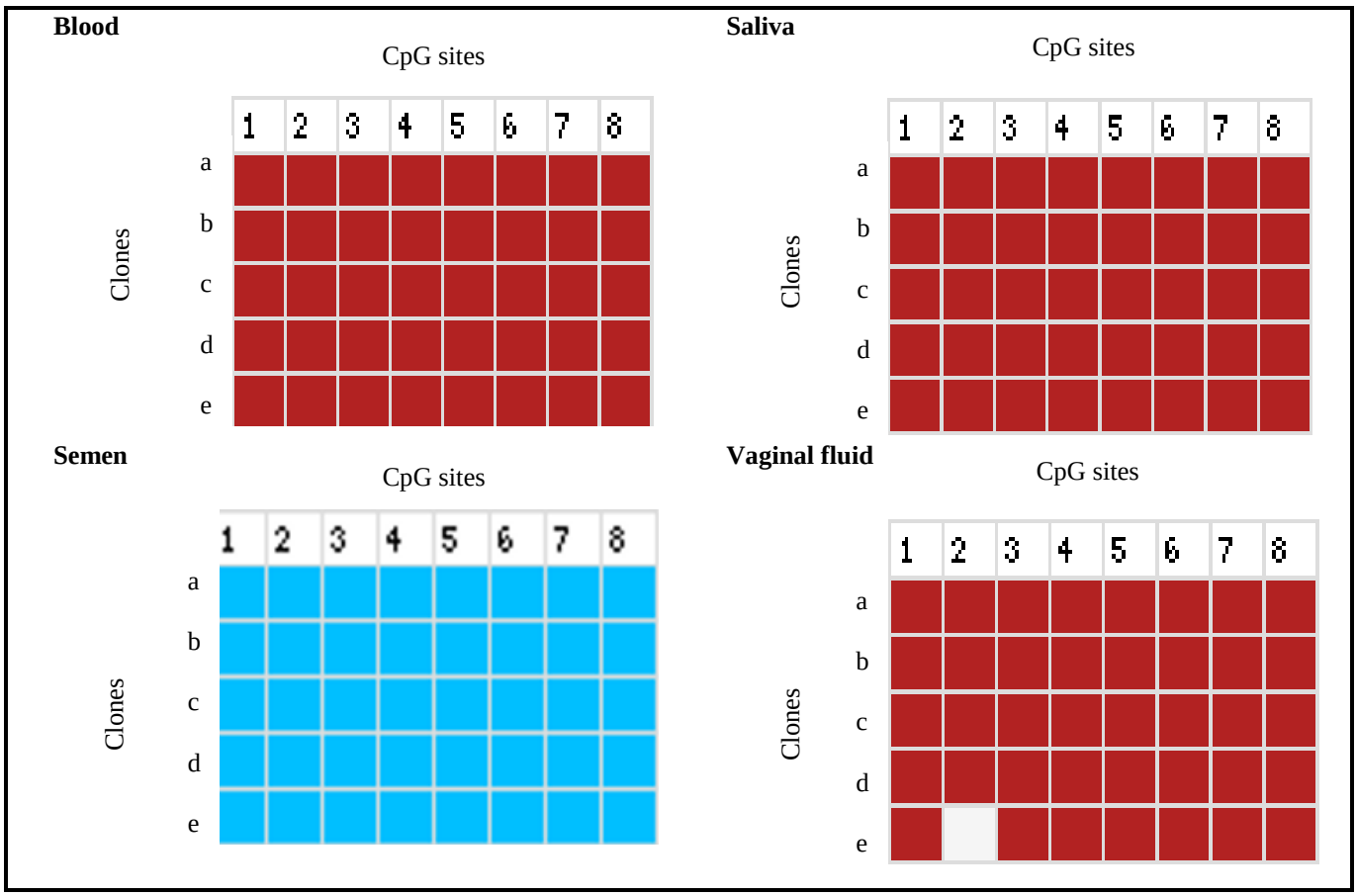


Figure 4.3: Bisulfite sequencing results for the ZNF282 tDMR marker at t0 days. Each row indicates a single clone of bisulfite PCR products and each column indicates individual CpG site in the region of interest. Different methylation states of the CpG sites are indicated by colours (blue represents unmethylated; red represents methylated; white represents unknown). Overall methylation results for all 8 CpG sites were: 100% for blood, 100% for saliva, 0% for semen and 98% for vaginal fluid.

4.3.2.2 DNA recovery after exposure to environmental conditions for t50 days

Recovery of DNA from different conditions were compared. All four body fluids were degraded when kept in all five conditions for t50 days as compared to t0 days. The DNA concentrations show on average an overall decrease for all four body fluids after t50 days (Appendix I: Table 2-5) compared to the DNA concentrations of body fluids at t0 days (Appendix I: Table 1).

After t50 days, vaginal fluid provided the highest DNA concentration (332.4 ng/µl) (Appendix I: Table 5) and semen provided the lowest DNA concentration (0.2 ng/µl) (Appendix I: Table 4). The DNA concentrations of blood, saliva and semen in conditions A, B, D and E were quite low (Appendix I Table 2-4) and following bisulfite conversion and PCR reactions no amplification was obtained, thus could not be further processed for cloning and sequencing. For condition C (outside on the ground), all body fluids provided a good concentration of DNA, where, blood provided the highest DNA concentration (84.3ng/µl) (Appendix I: Table 2) and saliva provided the lowest DNA concentration (5.5 ng/µl) (Appendix I: Table 3). In most conditions, vaginal fluid samples provided the highest DNA concentrations (Appendix I: Table 5).

4.3.2.3 Effects of simulated forensic conditions on methylation profile of ZNF282 tDMR after t50 days

4.3.2.3.1 Methylation profile ZNF282 tDMR after placing all fluids outside on the ground (condition C)

A total of 20 clones were sequenced (five clones for each body fluid), however only 13 clones produced results to be analysed using the BISMA software. The threshold sequence identity percentage was 90%, thus the clones that were not analysed were below the threshold. Blood samples showed an overall hypermethylation (96%) in all clones, with one clone (a) showing an unexpected cytosine nucleotide at CpG 3 (Figure 4.4). For saliva, hypermethylation (72%) was observed, with one clone showing unmethylation at all eight CpG sites and another clone showing unmethylation only at CpG 3 (Figure 4.4). An overall methylation of 94% was observed in vaginal fluid, with one clone showing an unexpected cytosine nucleotide at CpG 3 (Figure 4.4). Conversely, hypomethylation (100%) was observed in semen (Figure 4.4). Table 4.8 shows the summary of bisulfite sequencing results for body fluids exposed to condition C (outside on the ground) for the ZNF282 tDMR after t50 days. Statistical analysis showed a

significant difference in methylation patterns in saliva and semen when compared to all other body fluids ($p > 0.05$) (Table 4.9).

Table 4.8: Overview of bisulfite sequencing results analysed using BISMA for body fluids exposed to condition C (outside on the ground) for the ZNF282 tDMR marker after t50 days.

Body Fluid	Number of clones that produced results	Overall Methylation (%)	Overall Unmethylation (%)	Unknown CpG sites (%)
Blood	3	96	0	4
Saliva	4	72	28	0
Semen	4	0	100	0
Vaginal fluid	2	94	0	6

Table 4.9: Pairwise comparison of total DNA methylation for ZNF282 tDMR for Condition C in blood, saliva, semen and vaginal fluid after t50 days. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	Saliva	Semen	Vaginal fluid
Blood	1			
Saliva	0.0000	1		
Semen	0.0000	0.0002	1	
Vaginal fluid	1.0000	0.0008	0.0000	1

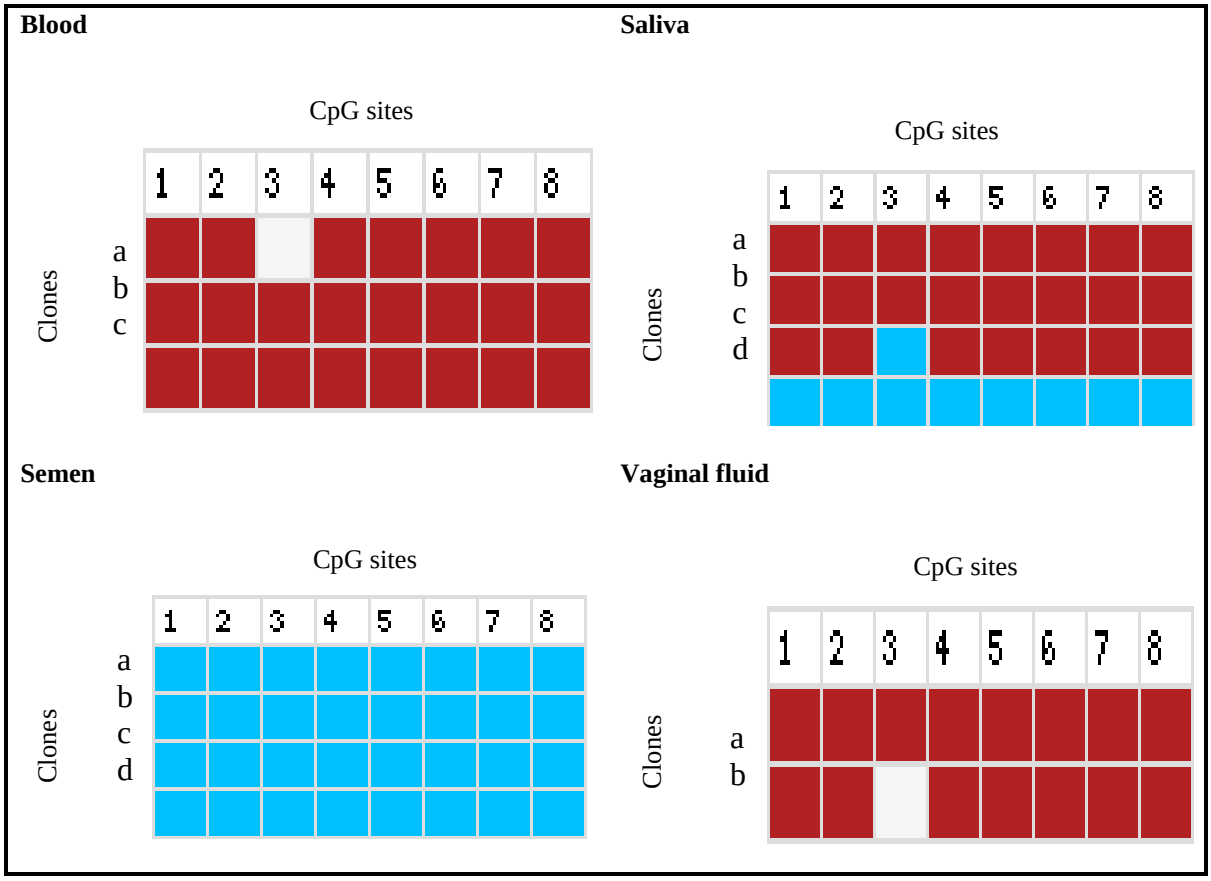


Figure 4.4: Bisulfite sequencing results of the ZNF282 tDMR marker for outside on the ground (condition C). Each row indicates a single clone of bisulfite PCR products and each column indicates individual CpG site in the region of interest. Different methylation states of the CpG sites are indicated by colours (blue represents unmethylated; red represents methylated; white represents unknown). Overall methylation results for all 8 CpG sites were: 96% for blood, 72% for saliva, 0% for semen and 94% for vaginal fluid.

4.3.2.3.2 Methylation profile of ZNF282 tDMR upon exposure to all conditions in vaginal fluid only

From a total of 25 sequenced clones, only 21 clones produced results to be analysed using the BISMA software. The threshold sequence identity percentage was 90%, thus the clones that were not analysed were below the threshold. The overall methylation for condition A (dry at room temperature) was 81%, with three clones (c and d) showing unmethylation at CpG 2 and CpG 6 and another clone (b) showing an unexpected cytosine nucleotide at CpG 5 and CpG 8 (Figure 4.5). For condition B (wet in an exsiccator), hypermethylation was observed (98%) with one clone showing an unexpected cytosine nucleotide at CpG 3 (clone d) (Figure 4.5). Condition C (outside on the ground) showed that there was hypermethylation (94%), with one clone showing an unexpected cytosine nucleotide at CpG 3 (clone b) (Figure 4.5). For condition

D (sprayed with alcohol), the clones displayed hypermethylation (98%) and one clone (clone e) showed unmethylation at CpG 3 (Figure 4.5). For condition E (sprayed with bleach), hypermethylation (98%) was displayed for all clones and an unexpected cytosine nucleotide at CpG 3 (clone e) (Figure 4.5). Table 4.10 shows the summary of bisulfite sequencing results for vaginal fluid exposed to all conditions for the ZNF282 tDMR after t50 days. Statistical analysis showed that methylation patterns of vaginal fluid in all conditions were not statistically different from each other ($p > 0.05$) (Table 4.11) and that methylation patterns at individual CpG sites were also not significantly different ($p > 0.05$) for the ZNF282 tDMR for vaginal fluid in all conditions (Appendix L: Table 1-8).

Table 4.10: Overview of bisulfite sequencing results for vaginal fluid analysed using BISMA for the ZNF282 tDMR marker under all forensic simulation conditions: dry at room temperature (condition A), wet in an exsiccator (condition B) outside on the ground (condition C), sprayed with alcohol (Condition D) and sprayed with bleach (Condition E).

Body Fluid	Number of clones that produced results	Overall Methylation (%)	Overall Unmethylation (%)	Unknown CpG sites(%)
Condition A	4	81	13	6
Condition B	5	98	0	2
Condition C	2	94	0	6
Condition D	5	98	2	0
Condition E	5	98	2	0

Table 4.11: Pairwise comparison of total DNA methylation for ZNF282 tDMR for all conditions in vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Vaginal fluid A	Vaginal fluid B	Vaginal fluid C	Vaginal fluid D	Vaginal fluid E
Vaginal fluid A	1				
Vaginal fluid B	0.1110	1			
Vaginal fluid C	0.4839	1.0000	1		
Vaginal fluid D	0.1934	1.0000	1.0000	1	
Vaginal fluid E	0.1934	1.0000	1.0000	1.0000	1

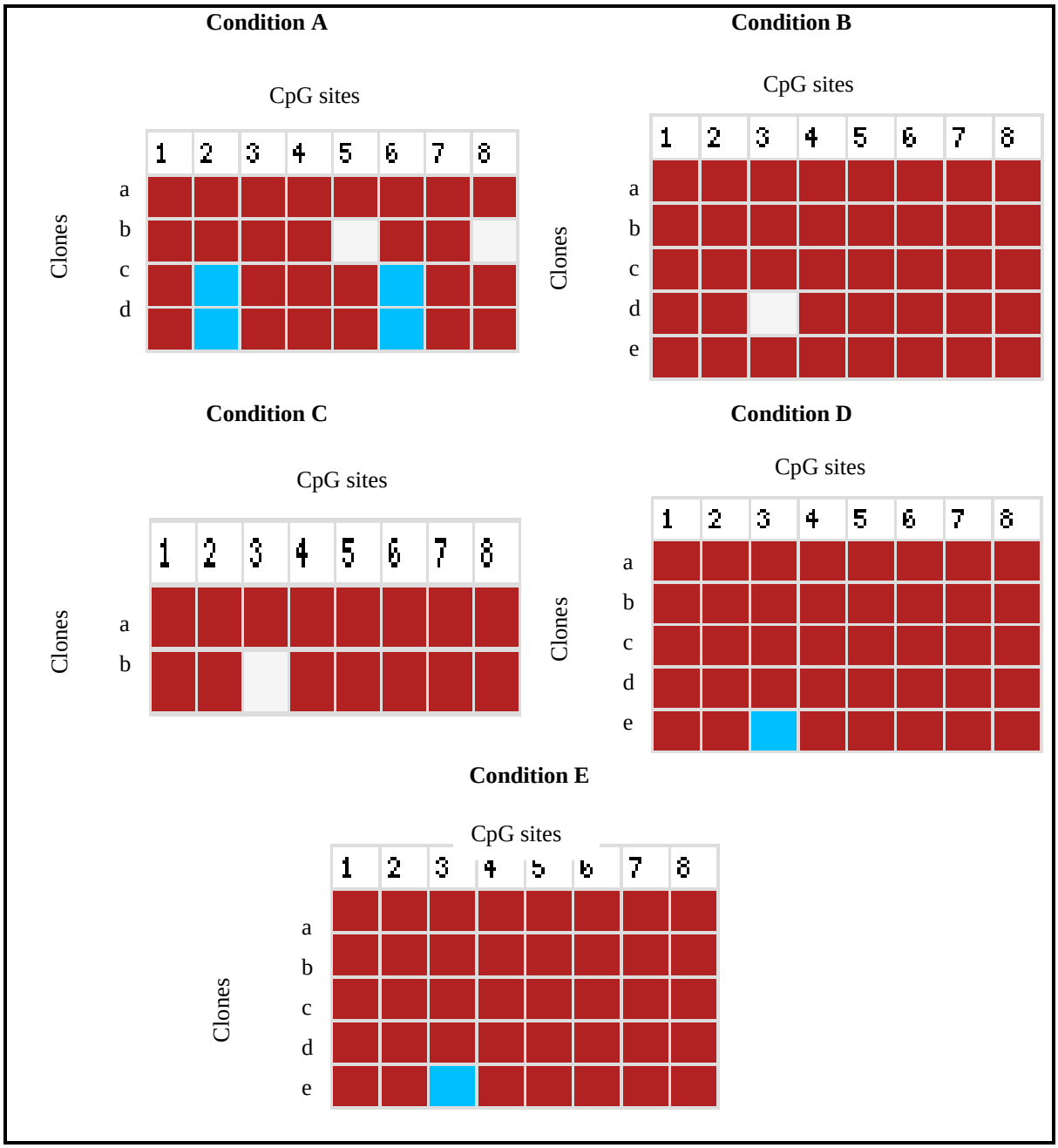


Figure 4.5: Bisulfite sequencing results for the ZNF282 tDMR in vaginal fluid under all forensic simulation conditions: dry at room temperature (condition A), wet in an exsiccator (condition B) outside on the ground (condition C), sprayed with alcohol (Condition D) and sprayed with bleach (Condition E). Each row indicates a single clone of bisulfite PCR products and each column indicates individual CpG site in the region of interest. Different methylation states of the CpG sites are indicated by colours (blue represents unmethylated; red represents methylated; white represents unknown). Overall methylation results for all 8 CpG sites in vaginal fluid samples were: 81% for condition A, 98% for condition B, 94% for condition C, 98% for condition D and 98% for condition E.

4.3.2.3.3 Methylation profiles of ZNF282 tDMR before (t0 days) and after exposure to forensic simulated conditions (t50 days)

A comparison between t0 days methylation profiles and t50 days methylation profiles of forensic simulated conditions was conducted to establish whether there was a difference in the degree of methylation for the ZNF282 tDMR marker. The methylation results in Figure 4.3 were compared to the methylation results in Figure 4.4 and Figure 4.5. Statistical analysis showed that methylation patterns between all body fluids at t0 days and t50 days for condition C (outside on the ground) were not statistically different from each other ($p > 0.05$) (Table 4.12) and that methylation patterns at individual CpG sites were also not significantly different ($p > 0.05$) for the ZNF282 tDMR for all body fluids (Appendix K: Table 1-8).

The vaginal fluid methylation results in Figure 4.3 were compared to the vaginal fluid methylation results in all conditions (Figure 4.5). Statistical analysis showed that methylation patterns between the vaginal fluid at t0 days and t50 days for all conditions (dry at room temperature (condition A), wet in an exsiccator (condition B), outside on the ground (condition C), sprayed with alcohol (condition D) and sprayed with bleach (condition E) was not statistically different from each other ($p > 0.05$) (Table 4.13) and that methylation patterns at individual CpG sites were also not significantly different ($p > 0.05$) for the ZNF282 tDMR between the vaginal fluid not exposed to any condition and vaginal fluid in all conditions (Appendix L: Table 9-16).

For condition C, methylation level decreased in all body fluids, however, semen methylation level remained the same after t50 days (Table 4.14). For all vaginal fluid samples in conditions A and C the methylation level decreased, however, methylation levels in conditions B, D and E was stable (Table 4.14). The methylation status of CpG sites for all body fluids in condition C showed that there was a decrease in methylation at every CpG site (8 CpG sites) for saliva, however, vaginal fluid and blood showed CpG 3 showed a decrease, while semen was stable at all CpG sites (8 CpG sites) (Table 4.15). The methylation status of vaginal fluid samples in all condition remained stable at most CpG sites, except, CpG 2, CpG 3, CpG 5, CpG 6 and CpG 8 (Table 4.15)

Table 4.12: Pairwise comparison of total DNA methylation for ZNF282 tDMR marker under Condition C compared between t0 days and t50 days in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	t0 days			
	Blood	Saliva	Semen	Vaginal fluid
t50 days	0.4910	0.2360	1.0000	1.0000

Table 4.13: Pairwise comparison of total DNA methylation for ZNF282 tDMR under all forensic simulation conditions compared between t0 days and t50 days in vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	t0 days				
	Vaginal fluid A	**Vaginal fluid B**	**Vaginal fluid C**	**Vaginal fluid D**	**Vaginal fluid E**
t50 days	0.1670	1.0000	1.0000	1.0000	1.0000

Table 4.14: Overall change in DNA methylation level in each body fluid after t50 days of subjection to five forensic simulation conditions (A- dry at room temperature, B- wet in an exsiccator and C- outside on the ground, D- sprayed with alcohol and E- sprayed with bleach). Up or down arrows indicate increase or decrease in DNA methylation level and a dash indicates no change in methylation level.

Body fluid	**Vaginal fluid**					**Saliva**	**Semen**	**Blood**
Condition	**A**	**B**	**C**	**D**	**E**	**C**	**C**	**C**
Overall methylation	↓	-	↓	-	-	↓	-	↓

Table 4.15: Change in DNA methylation at each CpG site in each body fluid after t50 days of subjection to five forensic simulation conditions (A- dry at room temperature, B- wet in an exsiccator and C- outside on the ground, D- sprayed with alcohol and E- sprayed with bleach). Up or down arrows indicate increase or decrease in DNA methylation level and a dash indicates no change in methylation level.

Body fluid	**Vaginal fluid**					**Saliva**	**Semen**	**Blood**
Condition	**A**	**B**	**C**	**D**	**E**	**C**	**C**	**C**
CpG 1	-	-	-	-	-	↓	-	-
CpG 2	↓	↑	↑	↑	↑	↓	-	-
CpG 3	-	↓	↓	↓	↓	↓	-	↓
CpG 4	-	-	-	-	-	↓	-	-
CpG 5	↓	-	-	-	-	↓	-	-
CpG 6	↓	-	-	-	-	↓	-	-
CpG 7	-	-	-	-	-	↓	-	-
CpG 8	↓	-	-	-	-	↓	-	-

4.4 Discussion

Identifying body fluids that are recovered from crime scenes is an important factor for determining the outcome of criminal investigations. Using tDMRs as epigenetic markers is proposed as a new method to determine the origin and source of biological material that are found at crime scenes. The CpG sites located within a tDMR have unique methylation patterns in various tissues and cells, thus can be used to differentiate between body fluids. For tDMR markers to efficiently differentiate between body fluids, it is important to determine the limits of DNA quantities that can be detected by methylation analysis as well as the stability of the analytical substrate (Forat *et al*., 2016; Silva *et al*., 2016). The objective of the present study was to analyse potential tDMRs for genes *ZNF282*, *PTPRS* and *HPCAL1* using MSP to assess the minimum quantity of DNA required for methylation analysis and to determine the effect of environmental conditions on the stability of methylation profile of ZNF282 tDMR marker.

4.4.1 Sensitivity study

DNA methylation based tissue identification methods require a small amount of DNA in order to obtain DNA profiles. Biological fluids at crime scenes are mostly found in either low amounts, are contaminated or degraded resulting in isolation of poor quantity and quality of DNA. Thus, a sensitivity study was conducted to determine the minimum amount of DNA needed to obtain accurate results using MSP.

Based on analysis with all four candidate tDMR markers, it can be concluded that DNA samples with a concentration greater than 10 ng would be ideal for analysis of methylation status because of the loss of DNA due to bisulfite conversion (Grunau *et al*., 2001). Also, PCR bias often occurs at the amplification stage, especially when working with bisulfite converted DNA that have low DNA concentrations. As a result, amplification is more efficient at unmethylated sites, however in certain instances an inverse bias can occur and amplification is more efficient at methylated site (Moskalev *et al*., 2011). A study by Park *et al*. (2014) performed a pyrosequencing analysis on varying DNA concentrations (500, 100, 50, 15, 10, or five ng) following bisulfite conversion. The results showed that DNA that had a concentration between 500 ng to 10 ng was successfully used in pyrosequencing analysis. However, a study by Madi *et al.* (2012) showed that bisulfite converted DNA with a concentration as low as five ng was able to be analysed using pyrosequencing.

4.4.2 Forensic simulation

In the previous study, the ZNF282 tDMR was identified as a semen-specific hypomethylation marker. The present study aimed to examine the stability of methylation profile of ZNF282 tDMR marker under simulated forensic conditions and to determine the effect of simulated forensic influences on DNA stability and recovery. The DNA of body fluids that were analysed immediately (t0 days and not exposed to any simulated conditions) was used as a control and compared to DNA of body fluids that were exposed to the five different environmental influences for t50 days: dry at room temperature (condition A), wet in an exsiccator (condition B) outside on the ground (condition C), sprayed with alcohol (condition D) and sprayed with bleach (condition E).

4.4.2.1 DNA recovery

DNA degradation was observed in all four body fluids that were kept in all five conditions for t50 days. The highest DNA concentration was observed for vaginal fluid samples (the highest was 332.4 ng/µl) and semen samples provided the lowest yield (as low as 0.2 ng/µl). In blood, condition A, B, D and E had a maximum effect on DNA degradation and for saliva and semen all conditions resulted in DNA degradation after t50 days of exposure. However, for vaginal fluid, all conditions had minimal effect on DNA degradation. DNA degradation has been reported in similar studies in which DNA was exposed to various environmental conditions over various time periods (Ambers *et al*., 2014; Parker, 2011). Ambers *et al*. (2014) examined the effects of three environmental conditions: UV radiation, temperature and humidity on DNA recovery and stability of bloodstains. The DNA in the bloodstains that were exposed to the temperature and humidity conditions remained quite durable and resistant to damage, probably due to being in a dry state. However, the bloodstains exposed to UV radiation varied in the degree of damage to the DNA. A possible explanation for this observed difference could be due to the composition of blood and that certain constituents of blood may absorb some UV radiation to provide a protective barrier to the DNA. Also, exposure to UV radiation produce photoproducts and cause DNA strand breakage, decreasing the chances of obtaining profiles (Hall and Ballantyne, 2004). The study by Parker. (2011) showed that body fluids exposed to microbes, UV radiation, moisture and high temperatures does effect the stability of DNA, however, dry conditions are favourable as the degradation process is reduced. Hence, exposure of body fluids to environmental influences such as temperature, humidity, chemicals and microbes as well as time of exposure, has an effect on the DNA recovery (Fondevila *et al*., 2008; Raymond *et al*., 2009; Ostojic and Wurmbach, 2017).

4.4.2.2 Methylation profile of ZNF282 tDMR marker analysed `immediately after collection (t0 days) for all body fluids

Body fluids at t0 days were not subjected to any environmental simulation. As expected, methylation pattern for semen was significantly different from all other body fluids ($p < 0.05$). This difference was a result of semen showing hypomethylation and all other body fluids showing hypermethylation. Semen-specific hypomethylated markers have been previously reported: DACT1 and USP49 (An *et al*., 2013; Choi *et al.*, 2014; Lee *et al.*, 2012), ZC3H12D (Madi *et al*., 2012) and cg04382920 and cg11768416 (Vidaki *et al*., 2016).

4.4.2.3 Methylation profile of ZNF282 tDMR analysed after t50 days of exposure to simulated forensic conditions

Endogenous and exogenous nucleases that are present in host cells and environment, respectively may result in DNA damage and modifications which affect the analysis of body fluids (Alaeddini *et al*., 2010). This can lead to inhibition of amplification processes. Additionally, when body fluids are exposed to environmental influences for long periods of time, the DNA in most instances will contain many lesions resulting in DNA damage and fragmentation, which affect DNA methylation analysis (Jun, 2010). Due to environmental impacts, the analytical CpG methylation state or the whole PCR template may be affected and result in changes in the methylation status.

In forensic simulation studies, methylation profiles under all conditions was obtained only for vaginal fluid. The remaining body fluids (blood, semen and saliva) yielded low DNA concentrations under conditions A (dry at room temperature), B (wet in and exsiccator), D (sprayed with alcohol) and E (sprayed with bleach) and hence, did not yield PCR product after bisulfite conversion and could not be further analysed.

Similar forensic simulation study conducted by Silva *et al*. (2016) showed that there were no significant differences observed in the methylation level of the samples exposed to environmental conditions compared to samples that were not exposed environmental conditions. However, a study by Forat *et al*. (2016) showed that the methylation status of samples exposed to humid environment conditions was effected but did not inhibit correct identification of body fluids.

All four body fluids provided methylation results for condition C (outside on the ground) only. The simulation study showed that methylation data for t50 days differed from methylation data

after t0 days. The overall methylation level for each body fluid exposed to environmental influences and analysed after t50 days compared to body fluid samples analysed immediately after sampling (t0 days), showed an overall decrease in methylation level of saliva, vaginal fluid and blood after exposure to condition C whereas, semen methylation status remained the same as for t0 days.

Statistical analysis showed that there was a significant difference observed in the overall methylation level of saliva and semen in condition C compared to blood and vaginal fluid ($p < 0.05$) in condition C. Individual CpG site statistical analysis showed a significant difference between blood and semen but not the other body fluids ($p < 0.05$). All vaginal fluid samples under all conditions showed that there was no significant difference in the overall methylation status ($p > 0.05$) and at individual CpG sites showed no significant difference ($p > 0.05$).

Any changes in the methylation status that occurred in body fluids after t50 days may be a result of spontaneous deamination in the natural environment of cells due to bacteria and fungi (Garvin *et al*., 2013; Forat *et al*., 2016). Other factors that may influence changes in methylation status are temperature, humidity, rain and chemicals (Meakin and Jamieson, 2013). Environmental influences can also result in DNA degradation and therefore can possibly inhibit DNA-based tissue identification. This can lead to erroneous tissue identification because methylated loci that are not amplified due to degradation can be mistaken to be unmethylated (Frumkin *et al*., 2011).

Even though the methylation levels changed after exposure to condition C (outside on the ground), semen still displayed 100% hypomethylation and blood, saliva and vaginal fluid >72% methylation. This further supports that the ZNF282 tDMR marker is a stable semen-specific hypomethylation marker. To the best of our knowledge this tDMR marker has not been previously identified and reported.

Furthermore, a comparison analysis was done between t0 days and t50 days for all body fluids. Statistical comparisons were made between methylation profiles obtained at t0 days and after t50 days. The statistical comparison of body fluids not exposed to environmental influences to all body fluids in condition C and vaginal fluid under all conditions for the ZNF282 tDMR, showed that there were no significant differences in the methylation status. This indicates that methylation status remains fairly stable when exposed to environmental influences.

4.5 Conclusion

The expected outcome of the research outlined in the present study, was to establish the sensitivity of methylation profiles for ZNF282, PTPRS 1, PTPRS 2 and HPCAL1 tDMRs using MSP and to determine the stability of methylation profile for the ZNF282 tDMR under forensically simulated conditions. DNA methylation analysis using bisulfite treatment methods (MSP and BS) is able to identify body fluids from low concentration DNA as well as degraded DNA. Thus, DNA methylation profiling for differentiation between body fluids and tissues appears to be a promising method for forensic applications. DNA methylation based methods are compatible with STR typing and can be incorporated into the current forensic workflow (Lee *et al*., 2016b).

In the future, DNA methylation analysis studies using different body fluid samples will be useful to identify novel DNA methylation markers which will add to the current markers that have been identified. This will enable the development of an efficient database of markers which can be used globally for discrimination between forensically relevant body fluids (Frumkin *et al*., 2011; Park *et al*., 2014). Furthermore, tDMRs that have been identified should be tested on larger sample sizes that include individuals of varying ages and ethnic groups as well as in a mixture of body fluids. Other forensic simulations such as the effect of microbes and longer exposure periods should also be investigated. This will improve the accuracy and reliability of using tDMR markers for body fluid identification.

Chapter 5
General discussion and Conclusion

5.1 Purpose of study

In most crime scenes, human body fluids that are recovered provide strong evidence that can be used to identify perpetrators and victims (An *et al*., 2012; Choi *et al*., 2014; Lee *et al*., 2012). Several types of body fluids are found at crime scenes; including saliva, blood, semen, vaginal fluid, menstrual blood, but also skin samples and urine (Sijen, 2015). Identifying the origin of the source and distinguishing between different types of body fluids may play a crucial role in linking evidence and the crime.

Currently in forensic casework, conventional presumptive and confirmatory tests are used to identify human body fluids. These tests are based on catalytic, immunological and enzymatic assays which identify specific components or cells in body fluids (Frumkin *et al*., 2011; Virkler and Lednev, 2009). The assays use a large amount of sample which is not readily available at crime scenes. Furthermore, they lack sensitivity and specificity which is crucial in forensic investigations, are prone to false positive results, and moreover, there are different tests for each body fluid, thus large amount of precious evidence is used (Haas *et al*., 2009; Virkler and Lednev, 2009). To overcome the limitations associated with presumptive and confirmatory tests, RNA-based methods were developed for identification of body fluids (Hanson *et al*., 2009; Juusola and Ballantyne, 2007; Setzer *et al*., 2008; Zubakov *et al*., 2008), however, RNA is unstable, prone to degradation and uses a large amount of samples. Short tandem repeat (STR) DNA typing is routinely employed for the identification of people using DNA from body fluids found at crime scenes. However, STR DNA typing cannot be used to determine the sources and origin of various body fluids (Xu *et al*., 2014). Thus, using DNA methylation-based markers have been suggested as being a valuable method for identification of body fluids as it is more stable and less susceptible to environmental influences (Xu *et al*., 2014). DNA is easily amplified by standard PCR techniques thus facilitating the use of minute amounts of samples. Every cell in the human genome has specific regions in DNA sequences that differ in DNA methylation patterns (Slieker *et al*., 2013; Song *et al*., 2009). These specific regions are known as tissue-specific differentially methylated regions (tDMRs) and can be useful in identification of body fluids (An *et al*., 2012; Choi *et al*., 2014). However, despite these advantages, thus far a limited number of tDMR markers have been validated for body fluid differentiation.

Considering the importance of identifying body fluids in forensic investigations, the present study aimed to investigate whether DNA methylation patterns of potential tDMRs can be used to differentiate between four forensically relevant human body fluids; namely blood, saliva,

semen and vaginal fluid. Blood samples were collected from 20 volunteers at Lab 24 laboratory, Mount Edgecombe, Durban. Vaginal fluid samples were collected from 20 volunteers at the King Dinizulu Hospital, Durban. Semen and saliva samples were collected from 10 volunteers at the King Dinizulu Hospital and University of KwaZulu-Natal, Westville campus. The potential tDMRs included in this study were located in three genes, namely, *ZNF282*, *PTPRS* and *HPCAL1*. Primers were designed for each gene at one CpG island, however, for the *PTPRS* gene, two CpG islands were targeted (referred to as PTPRS 1 and PTPRS 2) for comparison of methylation patterns within the gene. These gene-based tDMRs were identified as potential candidates for body fluid identification in a previous study conducted at University of KwaZulu-Natal, Westville (Kader, 2015, MSc Genetics dissertation) based on differential gene expression.

For profiling DNA methylation of the potential tDMRs in four body fluids; namely blood, saliva, semen and vaginal fluid, methylation specific PCR (MSP) and bisulfite sequencing (BS) techniques were employed. MSP is a simple, sensitive and cost effective qualitative method that determines the presence or absence of methylation (Herman *et al*., 1996; Hernández *et al*., 2013; Huang *et al*., 2013b). It is also able to produce methylation profiles for DNA samples that have low concentrations and that are of low quality. Furthermore, CpG sites are analysed without using methylation restriction enzymes (Herman *et al*., 1996; Dhingra *et al*., 2014). However, since the method does not provide quantitative results, BS was used to validate the methylation profiles using pooled DNA from each body fluid. The BS method is useful for site-specific or allele-specific methylation analysis (Huang *et al*., 2013a; Shen and Waterland, 2007). Although BS is limited by the cost and time taken to perform the method (Hernández *et al*., 2013) it still remains the gold standard method for methylation analysis that provides quantitative information (Huang *et al*., 2013a) since it facilitates detailed information at every CpG site in the target region.

In addition, since only minute quantities of DNA are generally recovered from crime scenes, the study set out to determine the minimum amount of DNA required to produce a stable methylation profile. Furthermore, DNA quality from crime scenes are usually compromised. Hence, the effects of environmental influences on DNA were examined to determine if the methylation profiles of the candidate tDMRs would remain stable.

5.2 Summary of findings

The first objective of the study was to establish methylation profiles for ZNF282, PTPRS 1, PTPRS 2 and HPCAL1 tDMRs using MSP and thereafter, BS. Based on the observed methylation profiles, MSP identified two semen-specific markers (ZNF282 tDMR and PTPRS 1 tDMR) and one saliva-specific marker (HPCAL1 tDMR).

The ZNF282 tDMR displayed hypermethylation for blood, saliva and vaginal fluid, but hypomethylation in semen only. The BS PCR reaction followed by sequencing, showed similar results to MSP for the ZNF282 tDMR, as there was significant difference between methylation profiles of semen and all other body fluids. Thus, based on the BS results, it was concluded that the ZNF282 tDMR can be used as a semen-specific hypomethylation marker. Using the same methods as the present study, Lee *et al*. (2012) identified two semen-specific markers (DACT1 and USP49) that displayed hypomethylation patterns in semen and hypermethylation patterns in blood, menstrual blood, saliva and vaginal fluid. Studies by Choi *et al*. (2014) and An *et al*. (2013) also observed that DACT1 and USP49 gene based markers could differentiate semen from all other body fluids. Other semen-specific markers which have been identified include cg22407458 (Forat *et al*., 2016), PRMT2 (Lee *et al.,* 2012), ZC3H12D (Madi *et al*., 2012) and cg04382920 and cg11768416 (Vidaki *et al*., 2016), all of which are hypomethylated in semen when compared to other fluids. Semen-specific hypermethylation markers include single CpG site based markers: L81528 (Choi *et al.*, 2014; Frumkin *et al.*, 2011), cg05656364 (Forat *et al*., 2016), cg26763284 and cg17621389 (Lee *et al*., 2015), cg26763282-138d and cg17610929 (Lee *et al.*, 2016a) cg05261336 (Lin *et al*., 2016) and cg23521140 (Park *et al*., 2014). However, most of these markers target just one CpG site which is not reliable and are yet to be validated, whereas in the present study the MSP results for ZNF282 tDMR were validated and reproduced using BS.

The MSP analysis for the PTPRS 1 tDMR displayed hypermethylation for blood, saliva and vaginal fluid, however, both hypermethylation and hypomethylation was observed for semen. However, BS was unable to distinguish semen from all other body fluids for the PTPRS 1 tDMR. The PTPRS 2 tDMR displayed both hypermethylation and hypomethylation patterns for all body fluids using MSP. Thus, no unique methylation patterns were observed to distinguish body fluids. Furthermore, the BS analysis was not able to differentiate between body fluids for both PTPRS 1 and PTPRS 2 tDMRs. Thus, the PTPRS 1 and PTPRS 2 tDMRs are not reliable markers for body fluid identification and requires further interrogation, perhaps

using a larger sample size for BS analysis with inclusion of more clones for sequencing or targeting different CpG islands of the gene.

The MSP study also identified a potential saliva-specific marker. For blood, semen and vaginal fluid, the HPCAL1 tDMR, displayed hypermethylation. However, saliva displayed hypomethylation patterns. While reliable and reproducible saliva-specific markers are not common, the *BCAS4* gene was identified as a potential saliva-specific hypermethylation marker by Madi *et al.* (2012). Further validation studies were conducted by Silva *et al.* (2016) by testing the sensitivity, species specificity, effect of degradation and mixed samples on the BCAS4 marker. The BCAS4 marker showed robust and reliable results. Similarly, other studies have also identified potential saliva-specific hypermethylated markers, such as, cg21597595 and cg15227982 (Forat *et al*., 2016), cg09652652 (Lee *et al*., 2015), cg09652652-2d (Lee *et al*., 2016a), cg09107912 (gene *FNDC1*) and cg16732616 (gene *DMRTA2*) (Lin *et al*., 2016) and cg26107890 (gene *SLC12A8*) and cg20691722 (gene *SOX2OT*) (Park *et al*., 2014). However, in the present study BS was unable to validate the HPCAL1 tDMR as a saliva-specific marker due to non-specific amplification for all body fluids. Therefore, the MSP results could not be confirmed by BS. Design of new primers for BS PCR reactions could be an option for future analysis.

Biological material found at crime scenes may be found in small quantities and exposed to environmental insults (Forat *et al*., 2016). Thus, the second objective of the study was to perform a sensitivity study to determine the lowest concentration of DNA required to obtain methylation profiles and the forensic simulation study was conducted to determine whether environmental influences have an effect on DNA methylation stability and DNA recovery.

For the sensitivity study, a five-fold serial dilution of DNA (25, 20, 15, 10, 5,1 ng) was prepared to determine the minimum amount required to obtain methylation profiles by MSP. The results determined that successful amplification can be obtained with a DNA concentration of 10 ng and more using the MSP method. Other studies that have also conducted sensitivity tests were also able to generate good methylation profiles at lower DNA concentrations (Antunes *et al*., 2016a; Madi *et al.*, 2012; Lee *et al*., 2015; Lin *et al*., 2016; Silva *et al*., 2016). However, these studies used different methods such as, pyrosequencing, multiplex SNaPshot or multiplex methylation sensitive restriction enzyme PCR.

Since MSP and BS showed that the ZNF282 tDMR marker was semen-specific, a forensic simulation study was conducted to determine if similar methylation profile could be obtained from body fluids exposed to the environment insults. Body fluids were subjected to five different simulated environment conditions, namely, dry at room temperature, wet in an exsiccator, outside on the ground, sprayed with alcohol and sprayed with bleach for 50 days (t50 days). BS was used to analyse the methylation status of the ZNF282 tDMR in blood, saliva, semen and vaginal fluid. Body fluids not exposed to any environment insults (t0 days) were analysed immediately after sample collection. As expected, semen displayed hypomethylation patterns which differentiated from all other body fluids that displayed hypermethylation patterns. After t50 days, environmental conditions had maximum detrimental effect on DNA recovery from semen, saliva and blood, however, high DNA quantity and quality was obtained from vaginal fluid. Methylation profile of saliva, semen, blood and vaginal fluid was obtained only for the outside on the ground condition. A significant difference in methylation profiles was observed for saliva and semen only. Only vaginal fluid methylation profile for ZNF282 could be analysed under all five environmental conditions. Vaginal fluid showed no significant difference in methylation under different conditions. Furthermore, no significant difference in methylation levels was observed between body fluids at t0 days and t50 days. However, the ZNF282 tDMR maintained hypomethylation patterns in semen as compared to hypermethylation patterns in other body fluids under environmental insults. Studies by Frumkin *et al*. (2011) and Silva *et al*. (2016) showed that even if DNA is degraded and subjected to environmental insults such as varying temperatures, there is no significant difference in the methylation profiles of markers. A study by Forat *et al*. (2016) exposed body fluids to three different conditions (dry on room temperature, humid/wet in an exsiccator and outside on the ground) and demonstrated that the most significant change in methylation profiles was observed in humid conditions. Methylation profiles of samples left outside were stable after a six-month period (Forat *et al*., 2016). Similar results were obtained in the present study, as duration and exposure of samples to the outside environment did not affect the methylation levels of the ZNF282 marker.

Future work should involve developing multiplex assays which will require minute amount of sample and the examination of multiple body fluids in a single reaction (Choi *et al.*, 2014). However, there is still much research required prior to using these developed methods routinely in forensic laboratories (Choi *et al*., 2014). Identification of more tDMRs using various body fluids will be advantageous in building a database of the current tDMRs as well as making the

methylation-based analysis more robust for body fluid identification. This will pave the way for the application of differential DNA methylation as an efficient and reliable method for identification of body fluids in forensic investigations.

5.3 Conclusion and future work

The ZNF282 tDMR was found to be a semen-specific hypomethylation marker. To our knowledge, the ZNF282 tDMR has not previously been reported for identification and differentiation of forensically relevant body fluids. Future research would include: examining the identified marker on a sample size that is larger and more diverse, comprising individuals of different ethnic groups, age groups, diverse geographical locations, as well as comparing methylation profiles of the marker obtained from healthy and diseased individuals. Additionally, exposure to the environment for longer and different time periods such as 30 days, 60 days, 90 days, and exposure to more simulated forensic conditions, such as different types of surfaces, UV radiation, a range of temperatures and soil microbes. Further validation can be conducted by testing tDMR marker in mixture analysis, artificial samples and non-human species. Additionally, research should be conducted to establish if variable DNA methylation affects normal or disease functional phenotypes in every tissue type (Igarashi *et al*., 2008). The results of such rigorous analysis of the tDMR marker using these conditions may enhance and compliment the methylation data obtained in the present study.

References

Aken, B., Ayling, S., Barrell, D., Clarke, L., Curwen, V., & Fairley, S. et al. (2016). The Ensembl gene annotation system. *Database*, *2016*, baw093.

Akutsu, T., Kitayama, T., Watanabe, K., & Sakurada, K. (2015). Comparison of automated and manual purification of total RNA for mRNA-based identification of body fluids. *Forensic Science International: Genetics*, *14*, 11-17.

Alaeddini, R., Walsh, S., & Abbas, A. (2010). Forensic implications of genetic analyses from degraded DNA—A review. *Forensic Science International: Genetics*, *4*(3), 148-157.

Altman, P., & Katz, D. (1961). *Blood and other body fluids*. Washington DC: Federation of American Societies for Experimental Biology.

Alvarez-Cubero, M., Saiz, M., Martinez-Gonzalez, L., Alvarez, J., Eisenberg, A., Budowle, B., & Lorente, J. (2012). Genetic Identification of Missing Persons: DNA Analysis of Human Remains and Compromised Samples. *Pathobiology*, *79*(5), 228-238.

Ambers, A., Turnbough, M., Benjamin, R., King, J., & Budowle, B. (2014). Assessment of the role of DNA repair in damaged forensic samples. *International Journal of Legal Medicine*, *128*(6), 913-921.

An, J., Choi, A., Shin, K., Yang, W., & Lee, H. (2013). DNA methylation-specific multiplex assays for body fluid identification. *International Journal of Legal Medicine*, *127*(1), 35-43.

An, J., Shin, K., Yang, W., & Lee, H. (2012). Body fluid identification in forensics. *BMB Reports*, *45*(10), 545-553.

Antunes, J., Silva, D., Balamurugan, K., Duncan, G., Alho, C., & McCord, B. (2016a). High-resolution melt analysis of DNA methylation to discriminate semen in biological stains. *Analytical Biochemistry*, *494*, 40-45.

Antunes, J., Silva, D., Balamurugan, K., Duncan, G., Alho, C., & McCord, B. (2016b). Forensic discrimination of vaginal epithelia by DNA methylation analysis through pyrosequencing. *Electrophoresis*, *37*(21), 2751-2758.

Aran, D., Toperoff, G., Rosenberg, M., & Hellman, A. (2010). Replication timing-related and gene body-specific methylation of active human genes. *Human Molecular Genetics*, *20*(4), 670-680.

Arányi, T., Váradi, A., Simon, I., & Tusnády, G. (2006). The BiSearch web server. *BMC Bioinformatics*, *7*(1), 431.

Bail, S., Swerdel, M., Liu, H., Jiao, X., Goff, L., Hart, R., & Kiledjian, M. (2010). Differential regulation of microRNA stability. *RNA*, *16*(5), 1032-1039.

Ball, M., Li, J., Gao, Y., Lee, J., LeProust, E., & Park, I. et al. (2009). Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nature Biotechnology*, *27*(4), 361-368.

Barlow, D. (1995). Gametic Imprinting in Mammals. *Science*, *270*(5242), 1610-1613.

Bartolomei, M., & Ferguson-Smith, A. (2011). Mammalian Genomic Imprinting. *Cold Spring Harbor Perspectives in Biology*, *3*(7), a002592-a002592.

Bauer, M., & Patzelt, D. (2002). Evaluation of mRNA Markers for the Identification of Menstrual Blood. *Journal of Forensic Sciences*, *47*(6), 15560J.

Bell, J., Pai, A., Pickrell, J., Gaffney, D., Pique-Regi, R., & Degner, J. et al. (2011). DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biology*, *12*(1), R10.

Bibikova, M., Lin, Z., Zhou, L., Chudin, E., Garcia, E., & Wu, B. et al. (2006). High-throughput DNA methylation profiling using universal bead arrays. *Genome Research*, *16*(3), 383-393.

Bird, A. (1986). CpG-rich islands and the function of DNA methylation. *Nature*, *321*(6067), 209-213.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes and Development*, *16*(1), 6-21.

Bjornsson, H., Sigurdsson, M., Fallin, M., Irizarry, R., Aspelund, T., & Cui, H. et al. (2008). Intra-individual Change Over Time in DNA Methylation With Familial Clustering. *JAMA*, *299*(24), 2877-2883.

Bock, C. (2012). Analysing and interpreting DNA methylation data. *Nature Reviews Genetics*, *13*(10), 705-719.

Bock, C., Paulsen, M., Tierling, S., Mikeska, T., Lengauer, T., & Walter, J. (2006). CpG Island Methylation in Human Lymphocytes Is Highly Correlated with DNA Sequence, Repeats, and Predicted DNA Structure. *Public Library of Sciences Genetics*, *2*(3), e26.

Bocklandt, S., Lin, W., Sehl, M., Sánchez, F., Sinsheimer, J., Horvath, S., & Vilain, E. (2011). Epigenetic Predictor of Age. *Public Library of Sciences One*, *6*(6), e14821.

Bonasio, R., Tu, S., & Reinberg, D. (2010). Molecular Signals of Epigenetic States. *Science*, *330*(6004), 612-616.

Borghese, B., Santulli, P., Héquet, D., Pierre, G., de Ziegler, D., Vaiman, D., & Chapron, C. (2012). Genetic Polymorphisms of DNMT3L Involved in Hypermethylation of Chromosomal Ends Are Associated with Greater Risk of Developing Ovarian Endometriosis. *The American Journal of Pathology*, *180*(5), 1781-1786.

Brena, R., Auer, H., Kornacker, K., Hackanson, B., Raval, A., Byrd, J., & Plass, C. (2006). Accurate quantification of DNA methylation using combined bisulfite restriction analysis coupled with the Agilent 2100 Bioanalyzer platform. *Nucleic Acids Research*, *34*(3), e17-e17.

Breuza, L., Poux, S., Estreicher, A., Famiglietti, M., Magrane, M., & Tognolli, M. et al. (2016). The UniProtKB guide to the human proteome. *Database*, *2016*, bav120.

Budowle, B., & van Daal, A. (2009). Extracting evidence from forensic DNA analyses: future molecular biology directions. *Biotechniques*, *46*(5), 339-350.

Bulyk, M. (2006). DNA microarray technologies for measuring protein–DNA interactions. *Current Opinion in Biotechnology*, *17*(4), 422-430.

Burger, J., Hummel, S., Herrmann, B., & Henke, W. (1999). DNA preservation: A microsatellite-DNA study on ancient skeletal remains. *Electrophoresis*, *20*(8), 1722-1728.

Byun, H., Siegmund, K., Pan, F., Weisenberger, D., Kanel, G., Laird, P., & Yang, A. (2009). Epigenetic profiling of somatic tissues from human autopsy specimens identifies tissue- and individual-specific DNA methylation patterns. *Human Molecular Genetics*, *18*(24), 4808-4817.

Casey, D., & Price, J. (2010). The sensitivity and specificity of the RSID™-saliva kit for the detection of human salivary amylase in the Forensic Science Laboratory, Dublin, Ireland. *Forensic Science International*, *194*(1-3), 67-71.

Chandler, L., Ghazi, H., Jones, P., Boukamp, P., & Fusenig, N. (1987). Allele-specific methylation of the human c-Ha-ras-1 gene. *Cell*, *50*(5), 711-717.

Chédin, F., Lieber, M., & Hsieh, C. (2002). The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proceedings of The National Academy of Sciences*, *99*(26), 16916-16921.

Chen, P., Cokus, S., & Pellegrini, M. (2010). BS Seeker: precise mapping for bisulfite sequencing. *BMC Bioinformatics*, *11*(1), 203-208.

Chhibber, A., & Schroeder, B. (2008). Single-molecule polymerase chain reaction reduces bias: Application to DNA methylation analysis by bisulfite sequencing. *Analytical Biochemistry*, *377*(1), 46-54.

Choi, A., Shin, K., Yang, W., & Lee, H. (2014). Body fluid identification by integrated analysis of DNA methylation and body fluid-specific microbial DNA. *International Journal of Legal Medicine*, *128*(1), 33-41.

Choi, J. (2010). Contrasting chromatin organization of CpG islands and exons in the human genome. *Genome Biology*, *11*(7), R70.

Choufani, S., Shapiro, J., Susiarjo, M., Butcher, D., Grafodatskaya, D., & Lou, Y. et al. (2011). A novel approach identifies new differentially methylated regions (DMRs) associated with imprinted genes. *Genome Research*, *21*(3), 465-476.

Christensen, B., Houseman, E., Marsit, C., Zheng, S., Wrensch, M., & Wiemels, J. et al. (2009). Aging and Environmental Exposures Alter Tissue-Specific DNA Methylation Dependent upon CpG Island Context. *Public Library of Sciences Genetics*, *5*(8), e1000602.

Christiansen, L., Lenart, A., Tan, Q., Vaupel, J., Aviv, A., McGue, M., & Christensen, K. (2016). DNA methylation age is associated with mortality in a longitudinal Danish twin study. *Aging Cell*, *15*(1), 149-154.

Clark, S., Harrison, J., Paul, C., & Frommer, M. (1994). High sensitivity mapping of methylated cytosines. *Nucleic Acids Research*, *22*(15), 2990-2997.

Collins, L., Schönfeld, B., & Chen, X. (2011). The Epigenetics of Non-coding RNA. *Handbook of Epigenetics*, 49-61.

Coolen, M., Statham, A., Qu, W., Campbell, M., Henders, A., & Montgomery, G. et al. (2011). Impact of the Genome on the Epigenome Is Manifested in DNA Methylation Patterns of Imprinted Regions in Monozygotic and Dizygotic Twins. *Public Library of Sciences One*, *6*(10), e25590.

Costello, J., Frühwald, M., Smiraglia, D., Rush, L., Robertson, G., & Gao, X. et al. (2000). Aberrant CpG-island methylation has non-random and tumour-type–specific patterns. *Nature Genetics*, *24*(2), 132-138.

Costello, J., Smiraglia, D., & Plass, C. (2002). Restriction landmark genome scanning. *Methods*, *27*(2), 144-149.

Dalma-Weiszhausz, D., Warrington, J., Tanimoto, E., & Miyada, C. (2006). The Affymetrix GeneChip platform: an overview. *Methods in Enzymology*, *410*, 3-28.

Davies, M., Volta, M., Pidsley, R., Lunnon, K., Dixit, A., & Lovestone, S. et al. (2012). Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biology*, *13*(6), R43.

Day, K., Waite, L., Thalacker-Mercer, A., West, A., Bamman, M., & Brooks, J. et al. (2013). Differential DNA methylation with age displays both common and dynamic features across human tissues that are influenced by CpG landscape. *Genome Biology*, *14*(9), R102.

Deaton, A., & Bird, A. (2011). CpG islands and the regulation of transcription. *Genes and Development*, *25*(10), 1010-1022.

Deaton, A., Webb, S., Kerr, A., Illingworth, R., Guy, J., Andrews, R., & Bird, A. (2011). Cell type-specific DNA methylation at intragenic CpG islands in the immune system. *Genome Research*, *21*(7), 1074-1086.

Derks, S., Lentjes, M., Hellebrekers, D., de Bruïne, A., Herman, J., & van Engeland, M. (2004). Methylation-specific PCR unraveled. *Analytical Cellular Pathology*, *26*(5-6), 291-299.

Dhingra, T., Mittal, K., & Sarma, G. (2014). Analytical Techniques for DNA Methylation – An Overview. *Current Pharmaceutical Analysis*, *10*(1), 71-85.

Diegoli, T., Farr, M., Cromartie, C., Coble, M., & Bille, T. (2012). An optimized protocol for forensic application of the PreCR™ Repair Mix to multiplex STR amplification of UV-damaged DNA. *Forensic Science International: Genetics*, *6*(4), 498-503.

Draht, M., Smits, K., Jooste, V., Tournier, B., Vervoort, M., & Ramaekers, C. et al. (2016). Analysis of RET promoter CpG island methylation using methylation-specific PCR (MSP), pyrosequencing, and methylation-sensitive high-resolution melting (MS-HRM): impact on stage II colon cancer patient outcome. *Clinical Epigenetics*, *8*(1), 44.

Dumache, R., Ciocan, V., Muresan, C., & Enache, A. (2016). *Molecular Genetics and its Applications in Forensic Sciences, Forensic Analysis - From Death to Justice*. Retrieved from: https://www.intechopen.com/books/forensic-analysis-from-death-to-justice/molecular-genetics-and-its-applications-in-forensic-sciences.

Dupont, C., Armant, D., & Brenner, C. (2009). Epigenetics: Definition, Mechanisms and Clinical Perspective. *Seminars in Reproductive Medicine*, *27*(05), 351-357.

Eckhardt, F., Lewin, J., Cortese, R., Rakyan, V., Attwood, J., & Burger, M. et al. (2006). DNA methylation profiling of human chromosomes 6, 20 and 22. *Nature Genetics*, *38*(12),

Espada, J., & Esteller, M. (2010). DNA methylation and the functional organization of the nuclear compartment. *Seminars in Cell and Developmental Biology*, *21*(2), 238-246.

Esteller, M., Sanchez-Cespedes, M., Rosell, R., Sidransky, D., Baylin, S., & Herman, J. (1999). Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. *Cancer Research*, *59*(1), 67-70.

Fan, J., Gunderson, K., Bibikova, M., Yeakley, J., Chen, J., & Wickham Garcia, E. et al. (2006). [3] Illumina Universal Bead Arrays. *Methods in Enzymology*, *410*, 57-73.

Feil, R., & Fraga, M. (2012). Epigenetics and the environment: emerging patterns and implications. *Nature Reviews Genetics*, *13*(2), 97-109.

Feltus, F., Lee, E., Costello, J., Plass, C., & Vertino, P. (2003). Predicting aberrant CpG island methylation. *Proceedings of The National Academy of Sciences*, *100*(21), 12253-12258.

Feng, S., Cokus, S., Zhang, X., Chen, P., Bostick, M., & Goll, M. et al. (2010). Conservation and divergence of methylation patterning in plants and animals. *Proceedings of The National Academy of Sciences*, *107*(19), 8689-8694.

Ferguson-Smith, A., & Surani, M. (2001). Imprinting and the Epigenetic Asymmetry Between Parental Genomes. *Science*, *293*(5532), 1086-1089.

Fernandez, A., Huidobro, C., & Fraga, M. (2012). De novo DNA methyltransferases: oncogenes, tumor suppressors, or both?. *Trends in Genetics*, *28*(10), 474-479.

Fishilevich, S., Zimmerman, S., Kohn, A., Iny Stein, T., Olender, T., & Kolker, E. et al. (2016). Genic insights from integrated human proteomics in GeneCards. *Database*, *2016*, baw030.

Fleming, R., & Harbison, S. (2010). The development of a mRNA multiplex RT-PCR assay for the definitive identification of body fluids. *Forensic Science International: Genetics*, *4*(4), 244-256.

Fondevila, M., Phillips, C., Naverán, N., Cerezo, M., Rodríguez, A., & Calvo, R. et al. (2008). Challenging DNA: Assessment of a range of genotyping approaches for highly degraded forensic samples. *Forensic Science International: Genetics Supplement Series*, *1*(1), 26-28.

Foran, D. (2006). Relative Degradation of Nuclear and Mitochondrial DNA: An Experimental Approach. *Journal of Forensic Sciences*, *51*(4), 766-770.

Forat, S., Huettel, B., Reinhardt, R., Fimmers, R., Haidl, G., Denschlag, D., & Olek, K. (2016). Methylation Markers for the Identification of Body Fluids and Tissues from Forensic Trace Evidence. *Public Library of Sciences One*, *11*(2), e0147973.

Fraga, M., & Esteller, M. (2002). DNA methylation: a profile of methods and applications. *Biotechniques*, *33*(3), 632-649.

Fraga, M., & Esteller, M. (2007). Epigenetics and aging: the targets and the marks. *Trends in Genetics*, *23*(8), 413-418.

Fraga, M., Ballestar, E., Paz, M., Ropero, S., Setien, F., & Ballestar, M. et al. (2005). From The Cover: Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of The National Academy of Sciences*, *102*(30), 10604-10609.

Frommer, M., McDonald, L., Millar, D., Collis, C., Watt, F., & Grigg, G. et al. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proceedings of The National Academy of Sciences*, *89*(5), 1827-1831.

Frumkin, D., Wasserstrom, A., Budowle, B., & Davidson, A. (2011). DNA methylation-based forensic tissue identification. *Forensic Science International: Genetics*, *5*(5), 517-524.

Frumkin, D., Wasserstrom, A., Davidson, A., & Grafit, A. (2010). Authentication of forensic DNA samples. *Forensic Science International: Genetics*, *4*(2), 95-103.

Fu, X., Wu, J., Wang, J., Huang, Y., Hou, Y., & Yan, J. (2015). Identification of body fluid using tissue-specific DNA methylation markers. *Forensic Science International: Genetics Supplement Series*, *5*, e151-e153.

Gaensslen, R. (1983). *Sourcebook in forensic serology, immunology, and biochemistry* (pp. 149-154). Washington, D.C.: U.S. Dept. of Justice, National Institute of Justice.

Galm, O., Herman, J., & Baylin, S. (2006). The fundamental role of epigenetics in hematopoietic malignancies. *Blood Reviews*, *20*(1), 1-13.

Gardiner-Garden, M., & Frommer, M. (1987). CpG Islands in vertebrate genomes. *Journal of Molecular Biology*, *196*(2), 261-282.

Garvin, A., Holzinger, R., Berner, F., Krebs, W., Hostettler, B., & Lardi, E. et al. (2013). The forensiX Evidence Collection Tube and Its Impact on DNA Preservation and Recovery. *Biomed Research International*, *2013*, 1-7.

Gebhard, C., Schwarzfischer, L., Pham, T., Andreesen, R., Mackensen, A., & Rehli, M. (2006a). Rapid and sensitive detection of CpG-methylation using methyl-binding (MB)-PCR. *Nucleic Acids Research*, *34*(11), e82-e82.

Gebhard, C., Schwarzfischer, L., Pham, T., Schilling, E., Klug, M., Andreesen, R., & Rehli, M. (2006b). Genome-Wide Profiling of CpG Methylation Identifies Novel Targets of Aberrant Hypermethylation in Myeloid Leukemia. *Cancer Research*, *66*(12), 6118-6128.

Gertz, J., Varley, K., Reddy, T., Bowling, K., Pauli, F., & Parker, S. et al. (2011). Analysis of DNA Methylation in a Three-Generation Family Reveals Widespread Genetic Influence on Epigenetic Regulation. *Public Library of Sciences Genetics*, *7*(8), e1002228.

Ghosh, S., Yates, A., Frühwald, M., Miecznikowski, J., Plass, C., & Smiraglia, D. (2010). Tissue-specific DNA methylation of CpG islands in normal human adult somatic tissues distinguishes neural from non-neural tissues. *Epigenetics*, *5*(6), 527-538.

Gill-King, H. (1997). *Chemical and ultrastructural aspects of decomposition, in: M.H. Haglund, W.D. Sorg (Eds.), Forensic Taphonomy: The Postmortem Fate of Human Remains, Robert B. Stern* (pp. 93-108). Boca Raton, Florida: CRC Press.

Goll, M., & Bestor, T. (2005). Eukaryotic cytosine methyltransferases. *Annual Review of Biochemistry*, *74*(1), 481-514.

Gonzalgo, M., & Jones, P. (1997). Rapid quantitation of methylation differences at specific sites using methylation-sensitive single nucleotide primer extension (Ms-SNuPE). *Nucleic Acids Research*, *25*(12), 2529-2531.

Gonzalgo, M., & Liang, G. (2007). Methylation-sensitive single-nucleotide primer extension (Ms-SNuPE) for quantitative measurement of DNA methylation. *Nature Protocols*, *2*(8), 1931-1936.

Greenfield, A., & Sloan, M. (2003). *Identification of biological fluids and stains, in: S.H. James, J.J. Nordby (Eds.), Forensic Science: An Introduction to Scientific and Investigative Techniques* (pp. 261-278). Boca Raton, Florida: CRC Press.

Gršković, B., Zrnec, D., Vicković, S., Popović, M., & Mršić, G. (2013). DNA methylation: the future of crime scene investigation?. *Molecular Biology Reports*, *40*(7), 4349-4360.

Grunau, C., Clark, S., & Rosenthal, A. (2001). Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Research*, *29*(13), e65-e65.

Haas, C., Klesser, B., Maake, C., Bär, W., & Kratzer, A. (2009). mRNA profiling for body fluid identification by reverse transcription endpoint PCR and realtime PCR. *Forensic Science International: Genetics*, *3*(2), 80-88.

Hahn, M., Wu, X., Li, A., Hahn, T., & Pfeifer, G. (2011). Relationship between Gene Body DNA Methylation and Intragenic H3K9me3 and H3K36me3 Chromatin Marks. *Public Library of Sciences One*, *6*(4), e18844.

Hall, A., & Ballantyne, J. (2004). Characterization of UVC-induced DNA damage in bloodstains: forensic implications. *Analytical and Bioanalytical Chemistry*, *380*(1), 72-83.

Hall, B. (2012). Evolutionary Developmental Biology (Evo-Devo): Past, Present, and Future. *Evolution: Education and Outreach*, *5*(2), 184-193.

Hanson, E., Lubenow, H., & Ballantyne, J. (2009). Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs. *Analytical Biochemistry*, *387*(2), 303-314.

Hata, K., Okano, M., Lei, H., & Li, E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development*, *129*(8), 1983-1993.

He, X., Chen, T., & Zhu, J. (2011). Regulation and function of DNA methylation in plants and animals. *Cell Research*, *21*(3), 442-465.

Hellman, A., & Chess, A. (2007). Gene Body-Specific Methylation on the Active X Chromosome. *Science*, *315*(5815), 1141-1143.

Herceg, Z. (2007). Epigenetics and cancer: towards an evaluation of the impact of environmental and dietary factors. *Mutagenesis*, *22*(2), 91-103.

Herman, J., Graff, J., Myohanen, S., Nelkin, B., & Baylin, S. (1996). Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proceedings of The National Academy of Sciences*, *93*(18), 9821-9826.

Hernández, H., Tse, M., Pang, S., Arboleda, H., & Forero, D. (2013). Optimizing methodologies for PCR-based DNA methylation analysis. *Biotechniques*, *55*(4), 181-197.

Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biology*, *14*(10), R115.

Höss, M., Jaruga, P., Zastawny, T., Dizdaroglu, M., & Paabo, S. (1996). DNA Damage and DNA Sequence Retrieval from Ancient Tissues. *Nucleic Acids Research*, *24*(7), 1304-1307.

HPCAL1 [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2004 – [24 August 2016]. Available from: https://www.ncbi.nlm.nih.gov/gene

Hua, D., Hu, Y., Wu, Y., Cheng, Z., Yu, J., Du, X., & Huang, Z. (2011). Quantitative methylation analysis of multiple genes using methylation-sensitive restriction enzyme-based quantitative PCR for the detection of hepatocellular carcinoma. *Experimental and Molecular Pathology*, *91*(1), 455-460.

Huang, Z., Bassil, C., & Murphy, S. (2013a). Bisulfite Sequencing of Cloned Alleles. *Methods in Molecular Biology*, *1049*, 83-94.

Huang, Z., Bassil, C., & Murphy, S. (2013b). Methylation-Specific PCR. *Methods in Molecular Biology*, *1049*, 75-82.

Igarashi, J., Muroi, S., Kawashima, H., Wang, X., Shinojima, Y., & Kitamura, E. et al. (2008). Quantitative analysis of human tissue-specific differences in methylation. *Biochemical and Biophysical Research Communications*, *376*(4), 658-664.

Illingworth, R., Kerr, A., DeSousa, D., Jørgensen, H., Ellis, P., & Stalker, J. et al. (2008). A Novel CpG Island Set Identifies Tissue-Specific Methylation at Developmental Gene Loci. *Public Library of Sciences Biology*, *6*(1), e22.

Irizarry, R., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., & Onyango, P. et al. (2009). The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nature Genetics*, *41*(2), 178-186.

Jacinto, F., Ballestar, E., & Esteller, M. (2008). Methyl-DNA immunoprecipitation (MeDIP): Hunting down the DNA methylome. *Biotechniques*, *44*(1), 35-43.

Jaenisch, R. (1997). DNA methylation and imprinting: why bother?. *Trends in Genetics*, *13*(8), 323-329.

Jaenisch, R., & Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genetics*, *33*(3s), 245-254.

Jirtle, R., & Skinner, M. (2007). Environmental epigenomics and disease susceptibility. *Nature Reviews Genetics*, *8*(4), 253-262.

Jjingo, D., Conley, A., Yi, S., Lunyak, V., & Jordan, I. (2012). On the presence and role of human gene-body DNA methylation. *Oncotarget*, *3*(4), 462-474.

Jones, P. (1999). The DNA methylation paradox. *Trends in Genetics*, *15*(1), 34-37.

Jones, P. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*, *13*(7), 484-492.

Jun, S. (2010). DNA Chemical Damage and Its Detected. *International Journal of Chemistry*, *2*(2), 261-265.

Juusola, J., & Ballantyne, J. (2003). Messenger RNA profiling: a prototype method to supplant conventional methods for body fluid identification. *Forensic Science International*, *135*(2), 85-96.

Juusola, J., & Ballantyne, J. (2005). Multiplex mRNA profiling for the identification of body fluids. *Forensic Science International*, *152*(1), 1-12.

Juusola, J., & Ballantyne, J. (2007). mRNA Profiling for Body Fluid Identification by Multiplex Quantitative RT-PCR. *Journal of Forensic Sciences*, *52*(6), 1252-1262.

Kader, F. (2015). Identification of Tissue-specific Differential Methylation in Human Body Fluids and its Potential Application in Forensics. MSc Genetics dissertation, University of KwaZulu-Natal, South Africa.

Kader, F., & Ghai, M. (2015). DNA methylation and application in forensic sciences. *Forensic Science International*, *249*, 255-265.

Kaminsky, Z., Tang, T., Wang, S., Ptak, C., Oh, G., & Wong, A. et al. (2009). DNA methylation profiles in monozygotic and dizygotic twins. *Nature Genetics*, *41*(2), 240-245.

Kapoor, A., Agius, F., & Zhu, J. (2005). Preventing transcriptional gene silencing by active DNA demethylation. *FEBS Letters*, *579*(26), 5889-5898.

Kitamura, E., Igarashi, J., Morohashi, A., Hida, N., Oinuma, T., & Nemoto, N. et al. (2007). Analysis of tissue-specific differentially methylated regions (TDMs) in humans. *Genomics*, *89*(3), 326-337.

Klose, R., & Bird, A. (2006). Genomic DNA methylation: the mark and its mediators. *Trends in Biochemical Sciences*, *31*(2), 89-97.

Klose, R., Sarraf, S., Schmiedeberg, L., McDermott, S., Stancheva, I., & Bird, A. (2005). DNA Binding Selectivity of MeCP2 Due to a Requirement for A/T Sequences Adjacent to Methyl-CpG. *Molecular Cell*, *19*(5), 667-678.

Koch, C., & Wagner, W. (2011). Epigenetic-aging-signature to determine age in different tissues. *Aging*, *3*(10), 1018-1027.

Kottaridi, C., Koureas, N., Margari, N., Terzakis, E., Bilirakis, E., & Pappas, A. et al. (2015). A Pyrosequencing Assay for the Quantitative Methylation Analysis ofGALR1in Endometrial Samples: Preliminary Results. *Biomed Research International*, *2015*, 1-6.

Kreil, D., Russell, R., & Russell, S. (2006). Microarray oligonucleotide probes. *Methods in Enzymology*, *410*, 73-98.

Kristensen, L., & Hansen, L. (2009). PCR-Based Methods for Detecting Single-Locus DNA Methylation Biomarkers in Cancer Diagnostics, Prognostics, and Response to Treatment. *Clinical Chemistry*, *55*(8), 1471-1483.

Kristensen, L., Mikeska, T., Krypuy, M., & Dobrovic, A. (2008). Sensitive Melting Analysis after Real Time- Methylation Specific PCR (SMART-MSP): high-throughput and probe-free quantitative DNA methylation detection. *Nucleic Acids Research*, *36*(7), e42-e42.

Krueger, F., Kreck, B., Franke, A., & Andrews, S. (2012). DNA methylome analysis using short bisulfite sequencing data. *Nature Methods*, *9*(2), 145-151.

Kulis, M., Queirós, A., Beekman, R., & Martín-Subero, J. (2013). Intragenic DNA methylation in transcriptional regulation, normal differentiation and cancer. *Biochimica Et Biophysica Acta - Gene Regulatory Mechanisms*, *1829*(11), 1161-1174.

Ladd-Acosta, C., Pevsner, J., Sabunciyan, S., Yolken, R., Webster, M., & Dinkins, T. et al. (2007). DNA Methylation Signatures within the Human Brain. *The American Journal of Human Genetics*, *81*(6), 1304-1315.

Laird, P. (2010). Principles and challenges of genome-wide DNA methylation analysis. *Nature Reviews Genetics*, *11*(3), 191-203.

Lan, V., Ha, N., Uyen, N., Duong, N., Huong, N., & Thuan, T. et al. (2014). Standardization of the methylation-specific PCR method for analyzing BRCA1 and ER methylation. *Molecular Medicine Reports*, *9*(5), 1844-1850.

Larsen, F., Gundersen, G., Lopez, R., & Prydz, H. (1992). CpG islands as gene markers in the human genome. *Genomics*, *13*(4), 1095-1107.

Laurent, L., Wong, E., Li, G., Huynh, T., Tsirigos, A., & Ong, C. et al. (2010). Dynamic changes in the human methylome during differentiation. *Genome Research*, *20*(3), 320-331.

Lee, H., An, J., Jung, S., Oh, Y., Lee, E., & Choi, A. et al. (2015). Genome-wide methylation profiling and a multiplex construction for the identification of body fluids using epigenetic markers. *Forensic Science International: Genetics*, *17*, 17-24.

Lee, H., Jung, S., Lee, E., Yang, W., & Shin, K. (2016a). DNA methylation profiling for a confirmatory test for blood, saliva, semen, vaginal fluid and menstrual blood. *Forensic Science International: Genetics*, *24*, 75-82.

Lee, H., Lee, S., & Shin, K. (2016b). Forensic DNA methylation profiling from evidence material for investigative leads. *BMB Reports*, *49*(7), 359-369.

Lee, H., Park, M., Choi, A., An, J., Yang, W., & Shin, K. (2012). Potential forensic application of DNA methylation profiling to body fluid identification. *International Journal of Legal Medicine, 126*(1), 55-62.

Lee, K., & Pausova, Z. (2013). Cigarette smoking and DNA methylation. *Frontiers in Genetics, 4*.

Lee, S., Crouse, C., & Kline, M. (2013). *Optimizing storage and handling of DNA extracts. Forensic DNA analysis: Current practices and emerging technologies* (pp. 19-38). Boca Raton, Florida: CRC Press.

Lefebvre, L., Viville, S., Barton, S., Ishino, F., & Surani, M. (1997). Genomic Structure and Parent-of-Origin-Specific Methylation of Peg1. *Human Molecular Genetics, 6*(11), 1907-1915.

Leontiou, C., Hadjidaniel, M., Mina, P., Antoniou, P., Ioannides, M., & Patsalis, P. (2015). Bisulfite Conversion of DNA: Performance Comparison of Different Kits and Methylation Quantitation of Epigenetic Biomarkers that Have the Potential to Be Used in Non-Invasive Prenatal Testing. *Public Library of Sciences One, 10*(8), e0135058.

Li, C., Zhang, S., Que, T., Li, L., & Zhao, S. (2011). Identical but not the same: The value of DNA methylation profiling in forensic discrimination within monozygotic twins. *Forensic Science International: Genetics Supplement Series, 3*(1), e337-e338.

Li, L., & Dahiya, R. (2002). MethPrimer: designing primers for methylation PCRs. *Bioinformatics, 18*(11), 1427-1431.

Li, L., Choi, J., Lee, K., Sung, H., Park, S., & Oze, I. et al. (2012). DNA Methylation in Peripheral Blood: A Potential Biomarker for Cancer Molecular Epidemiology. *Journal of Epidemiology, 22*(5), 384-394.

Li, R. (2008). *Forensic biology*. Boca Raton, Florida: CRC Press.

Li, Y., & Tollefsbol, T. (2011). DNA methylation detection: Bisulfite genomic sequencing analysis. *Methods in Molecular Biology, 791*, 11-21.

Lim, D., & Maher, E. (2010). DNA methylation: a form of epigenetic control of gene expression. *The Obstetrician & Gynaecologist*, *12*(1), 37-42.

Lin, Y., Tsai, L., Lee, J., Liu, K., Tzen, J., Linacre, A., & Hsieh, H. (2016). Novel identification of biofluids using a multiplex methylation-specific PCR combined with single-base extension system. *Forensic Science, Medicine, and Pathology*, *12*(2), 128-138.

Lindenbergh, A., de Pagter, M., Ramdayal, G., Visser, M., Zubakov, D., Kayser, M., & Sijen, T. (2012). A multiplex (m)RNA-profiling system for the forensic identification of body fluids and contact traces. *Forensic Science International: Genetics*, *6*(5), 565-577.

Lister, R., Pelizzola, M., Dowen, R., Hawkins, R., Hon, G., & Tonti-Filippini, J. et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, *462*(7271), 315-322.

Liu, X., Yu, X., Zack, D., Zhu, H., & Qian, J. (2008). TiGER: a database for tissue-specific gene expression and regulation. *BMC Bioinformatics*, *9*(1), 271.

Liu, Z., & Maekawa, M. (2003). Polymerase chain reaction-based methods of DNA methylation analysis. *Analytical Biochemistry*, *317*(2), 259-265.

Lokk, K., Modhukur, V., Rajashekar, B., Märtens, K., Mägi, R., & Kolde, R. et al. (2014). DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. *Genome Biology*, *15*(4), r54.

Loscalzo, J., & Handy, D. (2014). Epigenetic Modifications: Basic Mechanisms and Role in Cardiovascular Disease (2013 Grover Conference Series). *Pulmonary Circulation*, *4*(2), 169-174.

Ma, L., Yi, S., Huang, D., Mei, K., & Yang, R. (2013). Screening and identification of tissue-specific methylation for body fluid identification. *Forensic Science International: Genetics Supplement Series*, *4*(1), e37-e38.

Madi, T., Balamurugan, K., Bombardi, R., Duncan, G., & McCord, B. (2012). The determination of tissue-specific DNA methylation patterns in forensic biofluids using bisulfite modification and pyrosequencing. *Electrophoresis*, *33*(12), 1736-1745.

Marciniak-Czochra, A., Stiehl, T., & Wagner, W. (2009). Modeling of replicative senescence in hematopoietic development. *Aging*, *1*(8), 723-732.

Margulies, M., Egholm, M., Altman, W., Attiya, S., Bader, J., & Bemben, L. et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, *437*(7057), 376-380.

Martino, D., & Saffery, R. (2015). Characteristics of DNA methylation and gene expression in regulatory features on the Infinium 450k Beadchip. *Biorxiv*.

Martino, D., Loke, Y., Gordon, L., Ollikainen, M., Cruickshank, M., Saffery, R., & Craig, J. (2013). Longitudinal, genome-scale analysis of DNA methylation in twins from birth to 18 months of age reveals rapid epigenetic change in early life and pair-specific effects of discordance. *Genome Biology*, *14*(5), R42.

Maunakea, A., Nagarajan, R., Bilenky, M., Ballinger, T., D'Souza, C., & Fouse, S. et al. (2010). Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature*, *466*(7303), 253-257.

Meaburn, E., Schalkwyk, L., & Mill, J. (2010). Allele-specific methylation in the human genome. *Epigenetics*, *5*(7), 578-582.

Meakin, G., & Jamieson, A. (2013). DNA transfer: Review and implications for casework. *Forensic Science International: Genetics*, *7*(4), 434-443.

Meißner, C., von Wurmb, N., & Oehmichen, M. (1997). Detection of the age-dependent 4977 bp deletion of mitochondrial DNA. *International Journal of Legal Medicine*, *110*(5), 288-291.

Melnikov, A., Gartenhaus, R., Levenson, A., Motchoulskaia, N., & Levenson, V. (2005). MSRE-PCR for analysis of gene-specific DNA methylation. *Nucleic Acids Research*, *33*(10), e93-e93.

Messerschmidt, D., Knowles, B., & Solter, D. (2014). DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes and Development*, *28*(8), 812-828.

Minard, M., Jain, A., & Barton, M. (2009). Analysis of epigenetic alterations to chromatin during development. *Genesis*, *47*(8), 559-572.

Moore, L., Le, T., & Fan, G. (2012). DNA Methylation and Its Basic Function. *Neuropsychopharmacology*, *38*(2013), 23-38.

Moore, T., & Haig, D. (1991). Genomic imprinting in mammalian development: a parental tug-of-war. *Trends in Genetics*, *7*(2), 45-49.

Moskalev, E., Zavgorodnij, M., Majorova, S., Vorobjev, I., Jandaghi, P., Bure, I., & Hoheisel, J. (2011). Correction of PCR-bias in quantitative DNA methylation studies by means of cubic polynomial regression. *Nucleic Acids Research*, *39*(11), e77-e77.

Murrell, A., Heeson, S., Cooper, W., Douglas, E., Apostolidou, S., & Moore, G. et al. (2004). An association between variants in the IGF2 gene and Beckwith-Wiedemann syndrome: interaction between genotype and epigenotype. *Human Molecular Genetics*, *13*(2), 247-255.

Nagase, H., & Ghosh, S. (2008). Epigenetics: differential DNA methylation in mammalian somatic tissues. *FEBS Journal*, *275*(8), 1617-1623.

Nakayashiki, N., Takamiya, M., Shimamoto, K., Aoki, Y., & Hashiyada, M. (2009). Investigation of the methylation status around parent-of-origin detectable SNPs in imprinted genes. *Forensic Science International: Genetics*, *3*(4), 227-232.

Neri, F., Rapelli, S., Krepelova, A., Incarnato, D., Parlato, C., & Basile, G. et al. (2017). Intragenic DNA methylation prevents spurious transcription initiation. *Nature*, *543*(7643), 72-77.

Neumann, B., & Barlow, D. (1996). Multiple roles for DNA methylation in gametic imprinting. *Current Opinion in Genetics and Development*, *6*(2), 159-163.

Ng, R., & Gurdon, J. (2008). Epigenetic inheritance of cell differentiation status. *Cell Cycle*, *7*(9), 1173-1177.

Nussbaumer, C., Gharehbaghi-Schnell, E., & Korschineck, I. (2006). Messenger RNA profiling: A novel method for body fluid identification by Real-Time PCR. *Forensic Science International*, *157*(2-3), 181-186.

Nuwaysir, E., Huang, W., Albert, T., Singh, J., Nuwaysir, K., & Pitas, A. et al. (2002). Gene Expression Analysis Using Oligonucleotide Arrays Produced by Maskless Photolithography. *Genome Research*, *12*(11), 1749-1755.

Ohashi, H. (2002). Methylation-Specific PCR. *PCR Cloning Protocols*, *192*, 91-97.

Ohgane, J., Yagi, S., & Shiota, K. (2008). Epigenetics: The DNA Methylation Profile of Tissue-Dependent and Differentially Methylated Regions in Cells. *Placenta*, *29*, 29-35.

Okazaki, R., Ootsuyama, A., Yoshida, Y., & Norimura, T. (2011). Establishment of Methylation-Specific PCR for the Mouse p53 Gene. *Molecular Biology International*, *2011*, 1-4.

Okumura, K., Sakaguchi, G., Naito, K., Tamura, T., & Igarashi, H. (1997). HUB1, a novel Kruppel type zinc finger protein, represses the human T cell leukemia virus type I long terminal repeat-mediated expression. *Nucleic Acids Research*, *25*(24), 5025-5032.

Ollikainen, M., Smith, K., Joo, E., Ng, H., Andronikos, R., & Novakovic, B. et al. (2010). DNA methylation analysis of multiple tissues from newborn twins reveals both genetic and intrauterine components to variation in the human neonatal epigenome. *Human Molecular Genetics*, *19*(21), 4176-4188.

Orphanou, C. (2015). A bioanalytical approach to forensic body fluid identification & age determination. Doctoral Dissertation, Staffordshire University.

Ostojic, L., & Wurmbach, E. (2017). Analysis of fingerprint samples, testing various conditions, for forensic DNA identification. *Science & Justice*, *57*(1), 35-40.

Pabinger, S., Ernst, K., Pulverer, W., Kallmeyer, R., Valdes, A., & Metrustry, S. et al. (2016). Analysis and Visualization Tool for Targeted Amplicon Bisulfite Sequencing on Ion Torrent Sequencers. *Public Library of Sciences One*, *11*(7), e0160227.

Park, J., Kim, J., Seo, E., Bae, D., Kim, S., & Lee, H. et al. (2016). Identification and evaluation of age-correlated DNA methylation markers for forensic use. *Forensic Science International: Genetics*, *23*, 64-70.

Park, J., Kwon, O., Kim, J., Yoo, H., Lee, H., & Woo, K. et al. (2014). Identification of body fluid-specific DNA methylation markers for use in forensic science. *Forensic Science International: Genetics*, *13*, 147-153.

Parker, C. (2011). The Relative Recoverability Of Dna And Rna Profiles From Forensically Relevant Body Fluid Stains. *Electronic Theses and Dissertations*, 1948.

Parrish, R., Day, J., & Lubin, F. (2012). Direct Bisulfite Sequencing for Examination of DNA Methylation with Gene and Nucleotide Resolution from Brain Tissues. *Current Protocols in Neuroscience*, *Chapter 7*(Unit 7, 24).

Paulsen, M., & Ferguson-Smith, A. (2001). DNA methylation in genomic imprinting, development, and disease. *The Journal of Pathology*, *195*(1), 97-110.

Phang, T., Shi, C., Chia, J., & Ong, C. (1994). Amplification of cDNA via RT-PCR Using RNA Extracted from Postmortem Tissues. *Journal of Forensic Sciences*, *39*(5), 1275-1279.

PTPRS [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2004 – [24 August 2016]. Available from: https://www.ncbi.nlm.nih.gov/gene

Rakyan, V., Down, T., Maslau, S., Andrew, T., Yang, T., & Beyan, H. et al. (2010). Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Research*, *20*(4), 434-439.

Rakyan, V., Down, T., Thorne, N., Flicek, P., Kulesha, E., & Graf, S. et al. (2008). An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). *Genome Research*, *18*(9), 1518-1529.

Raymond, J., van Oorschot, R., Gunn, P., Walsh, S., & Roux, C. (2009). Trace evidence characteristics of DNA: A preliminary investigation of the persistence of DNA at crime scenes. *Forensic Science International: Genetics*, *4*(1), 26-33.

Reik, W., & Walter, J. (2001). Genomic imprinting: parental influence on the genome. *Nature Reviews Genetics*, *2*(1), 21-32.

Ribel-Madsen, R., Fraga, M., Jacobsen, S., Bork-Jensen, J., Lara, E., & Calvanese, V. et al. (2012). Genome-Wide Analysis of DNA Methylation Differences in Muscle and Fat from Monozygotic Twins Discordant for Type 2 Diabetes. *Public Library of Sciences One*, *7*(12), e51302.

Rivenbark, A., Stolzenburg, S., Beltran, A., Yuan, X., Rots, M., Strahl, B., & Blancafort, P. (2012). Epigenetic reprogramming of cancer cells via targeted DNA methylation. *Epigenetics*, *7*(4), 350-360.

Rohde, C., Zhang, Y., Jurkowski, T., Stamerjohanns, H., Reinhardt*, R., & Jeltsch, A. (2008). Bisulfite sequencing Data Presentation and Compilation (BDPC) web server a useful tool for DNA methylation analysis. *Nucleic Acids Research*, *36*(5), e34.

Sant, K., Nahar, M., & Dolinoy, D. (2012). DNA Methylation Screening and Analysis. *Methods in Molecular Biology*, *889*, 385-406.

Sasaki, M., Anast, J., Bassett, W., Kawakami, T., Sakuragi, N., & Dahiya, R. (2003). Bisulfite conversion-specific and methylation-specific PCR: a sensitive technique for accurate evaluation of CpG methylation. *Biochemical and Biophysical Research Communications*, *309*(2), 305-309.

Sato, F., Tsuchiya, S., Terasawa, K., & Tsujimoto, G. (2009). Intra-Platform Repeatability and Inter-Platform Comparability of MicroRNA Microarray Technology. *Public Library of Sciences One*, *4*(5), e5540.

Sauer, E., Reinke, A., & Courts, C. (2016). Differentiation of five body fluids from forensic samples by expression analysis of four microRNAs using quantitative PCR. *Forensic Science International: Genetics*, *22*, 89-99.

Schellenberg, A., Lin, Q., Schüler, H., Koch, C., Joussen, S., & Denecke, B. et al. (2011). Replicative senescence of mesenchymal stem cells causes DNA-methylation changes which correlate with repressive histone marks. *Aging*, *3*(9), 873-888.

Schilling, E., & Rehli, M. (2007). Global, comparative analysis of tissue-specific promoter CpG methylation. *Genomics*, *90*(3), 314-323.

Schmidl, C., Klug, M., Boeld, T., Andreesen, R., Hoffmann, P., Edinger, M., & Rehli, M. (2009). Lineage-specific DNA methylation in T cells correlates with histone methylation and enhancer activity. *Genome Research*, *19*(7), 1165-1174.

Schübeler, D. (2015). Function and information content of DNA methylation. *Nature*, *517*(7534), 321-326.

Sen, A., Heredia, N., Senut, M., Land, S., Hollocher, K., & Lu, X. et al. (2015). Multigenerational epigenetic inheritance in humans: DNA methylation changes associated with maternal exposure to lead can be transmitted to the grandchildren. *Scientific Reports*, *5*(14466), 1-10.

Setzer, M., Juusola, J., & Ballantyne, J. (2008). Recovery and Stability of RNA in Vaginal Swabs and Blood, Semen, and Saliva Stains. *Journal of Forensic Sciences*, *53*(2), 296-305.

Shaler, R. (1981). A multi-enzyme electrophoretic system for the identification of seminal fluid from postmortem specimens. *The American Journal of Forensic Medicine and Pathology*, *2*(4), 315-322.

Shaler, R. (2002). *Modern forensic biology, in: R. Saferstein (Ed.), Forensic Science Handbook* (pp. 529-546). Upper Saddle River, NJ: Prentice Hall.

Shen, L., & Waterland, R. (2007). Methods of DNA methylation analysis. *Current Opinion in Clinical Nutrition and Metabolic Care*, *10*(5), 576-581.

Shen, L., Kondo, Y., Guo, Y., Zhang, J., Zhang, L., & Ahmed, S. et al. (2007). Genome-Wide Profiling of DNA Methylation Reveals a Class of Normally Methylated CpG Island Promoters. *Public Library of Sciences Genetics*, *3*(10), 2023-2036.

Shiota, K., Kogo, Y., Ohgane, J., Imamura, T., Urano, A., & Nishino, K. et al. (2002). Epigenetic marks by DNA methylation specific to stem, germ and somatic cells in mice. *Genes to Cells*, *7*(9), 961-969.

Shoemaker, R., Deng, J., Wang, W., & Zhang, K. (2010). Allele-specific methylation is prevalent and is contributed by CpG-SNPs in the human genome. *Genome Research*, *20*(7), 883-889.

Sijen, T. (2015). Molecular approaches for forensic cell type identification: On mRNA, miRNA, DNA methylation and microbial markers. *Forensic Science International: Genetics*, *18*, 21-32.

Silva, A., & White, R. (1988). Inheritance of allelic blueprints for methylation patterns. *Cell*, *54*(2), 145-152.

Silva, D., Antunes, J., Balamurugan, K., Duncan, G., Alho, C., & McCord, B. (2016). Developmental validation studies of epigenetic DNA methylation markers for the detection of blood, semen and saliva samples. *Forensic Science International: Genetics*, *23*, 55-63.

Sinelnikov, A., Kalinina, A., Old, J., Boonlayangoor, P., & Reich, K. (2013). Evaluation of Rapid Stain IDentification (RSID™) Reader System for Analysis and Documentation of RSID™ Tests. *Applied Sciences*, *3*(3), 624-635.

Slack, J. (2002). Conrad Hal Waddington: the last Renaissance biologist?. *Nature Reviews Genetics*, *3*(11), 889-895.

Slieker, R., Bos, S., Goeman, J., Bovée, J., Talens, R., & van der Breggen, R. et al. (2013). Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450k array. *Epigenetics and Chromatin*, *6*(1), 26.

Smith, Z., & Meissner, A. (2013). DNA methylation: roles in mammalian development. *Nature Reviews Genetics*, *14*(3), 204-220.

Song, F., Mahmood, S., Ghosh, S., Liang, P., Smiraglia, D., Nagase, H., & Held, W. (2009). Tissue-specific differentially methylated regions (TDMR): Changes in DNA methylation during development. *Genomics*, *93*(2), 130-139.

Song, F., Smith, J., Kimura, M., Morrow, A., Matsuyama, T., Nagase, H., & Held, W. (2005). Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. *Proceedings of The National Academy of Sciences*, *102*(9), 3336-3341.

Sonnet, M., Baer, C., Rehli, M., Weichenhan, D., & Plass, C. (2013). Enrichment of Methylated DNA by Methyl-CpG Immunoprecipitation. *Methods in Molecular Biology*, *971*, 201-212.

Spalding, R. (2003). *Identification and characterization of blood and bloodstains, in: S.H. James, J.J. Nordby (Eds.), Forensic Science: An Introduction to Scientific and Investigative Techniques* (pp. 181-220). Boca Raton: CRC Press.

Spilker, C., & Braunewell, K. (2003). Calcium–myristoyl switch, subcellular localization, and calcium-dependent translocation of the neuronal calcium sensor protein VILIP-3, and comparison with VILIP-1 in hippocampal neurons☆. *Molecular and Cellular Neuroscience*, *24*(3), 766-778.

Spilker, C., Richter, K., Smalla, K., Manahan-Vaughan, D., Gundelfinger, E., & Braunewell, K. (2000). The neuronal EF-hand calcium-binding protein visinin-like protein-3 is expressed in cerebellar Purkinje cells and shows a calcium-dependent membrane association. *Neuroscience*, *96*(1), 121-129.

Spinrad, P. (1994). *The RE/search guide to body fluids,RE/Search Publications*. San Francisco.

Stranger, B., Nica, A., Forrest, M., Dimas, A., Bird, C., & Beazley, C. et al. (2007). Population genomics of human gene expression. *Nature Genetics*, *39*(10), 1217-1224.

Suzuki, M., & Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nature Reviews Genetics*, *9*(6), 465-476.

Tammen, S., Friso, S., & Choi, S. (2013). Epigenetics: The link between nature and nurture. *Molecular Aspects of Medicine*, *34*(4), 753-764.

Thevissen, P., Kaur, J., & Willems, G. (2012). Human age estimation combining third molar and skeletal development. *International Journal of Legal Medicine*, *126*(2), 285-292.

Thu, K., Pikor, L., Kennett, J., Alvarez, C., & Lam, W. (2010). Methylation analysis by DNA immunoprecipitation. *Journal of Cellular Physiology*, *222*(3), 522-531.

Tilghman, S. (1999). The Sins of the Fathers and Mothers: Genomic Imprinting in Mammalian Development. *Cell*, *96*(2), 185-193.

Tobi, E., Lumey, L., Talens, R., Kremer, D., Putter, H., & Stein, A. et al. (2009). DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Human Molecular Genetics*, *18*(21), 4046-4053.

Tost, J., Schatz, P., Schuster, M., Berlin, K., & Gut, I. (2003). Analysis and accurate quantification of CpG methylation by MALDI mass spectrometry. *Nucleic Acids Research*, *31*(9), e50-e50.

Tusnády, G., Simon, I., Váradi, A., & Arányi, T. (2005). BiSearch: primer-design and search tool for PCR on bisulfite-treated genomes. *Nucleic Acids Research*, *33*(1), e9-e9.

Uhlén, M., Fagerberg, L., Hallstrom, B., Lindskog, C., Oksvold, P., & Mardinoglu, A. et al. (2015). Tissue-based map of the human proteome. *Science*, *347*(6220), 1260419-1260419.

van Dongen, J., Ehli, E., Slieker, R., Bartels, M., Weber, Z., & Davies, G. et al. (2014). Epigenetic Variation in Monozygotic Twins: A Genome-Wide Analysis of DNA Methylation in Buccal Cells. *Genes*, *5*(2), 347-365.

van Eijk, K., de Jong, S., Boks, M., Langeveld, T., Colas, F., & Veldink, J. et al. (2012). Genetic analysis of DNA methylation and gene expression levels in whole blood of healthy human subjects. *BMC Genomics*, *13*(1), 636.

Vidaki, A., Daniel, B., & Court, D. (2013). Forensic DNA methylation profiling—Potential opportunities and challenges. *Forensic Science International: Genetics*, *7*(5), 499-507.

Vidaki, A., Giangasparo, F., & Syndercombe Court, D. (2016). Discovery of potential DNA methylation markers for forensic tissue identification using bisulphite pyrosequencing. *Electrophoresis*, *37*(21), 2767-2779.

Virkler, K., & Lednev, I. (2009). Analysis of body fluids for forensic purposes: From laboratory testing to non-destructive rapid confirmatory identification at a crime scene. *Forensic Science International*, *188*(1-3), 1-17.

Waddington, C. (1942). The epigenotype. *Endeavour*, *1*, 18-20.

Wagner, J., Busche, S., Ge, B., Kwan, T., Pastinen, T., & Blanchette, M. (2014). The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts. *Genome Biology*, *15*(2), R37.

Wang, Y., Liu, T., Xu, D., Shi, H., Zhang, C., Mo, Y., & Wang, Z. (2016). Predicting DNA Methylation State of CpG Dinucleotide Using Genome Topological Features and Deep Networks. *Scientific Reports*, *6*(1).

Wang, Z., Zhou, D., Zhang, S., Bian, Y., Hu, Z., & Zhu, R. et al. (2015). Logical Framework of Forensic Identification: Ability to Resist Fabricated DNA. *Molecular Biotechnology*, *57*(11-12), 1030-1037.

Watanabe, K., Akutsu, T., Takamura, A., & Sakurada, K. (2016). Evaluation of a blood-specific DNA methylated region and trial for allele-specific blood identification from mixed body fluid DNA. *Legal Medicine*, *22*, 49-53.

Weber, M., Davies, J., Wittig, D., Oakeley, E., Haase, M., Lam, W., & Schübeler, D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nature Genetics*, *37*(8), 853-862.

Wierstra, I. (2008). Sp1: Emerging roles—Beyond constitutive activation of TATA-less housekeeping genes. *Biochemical and Biophysical Research Communications*, *372*(1), 1-13.

Willerslev, E., Hansen, A., & Poinar, H. (2004). Isolation of nucleic acids and cultures from fossil ice and permafrost. *Trends in Ecology & Evolution*, *19*(3), 141-147.

Wong, C., Caspi, A., Williams, B., Craig, I., Houts, R., & Ambler, A. et al. (2010). A longitudinal study of epigenetic variation in twins. *Epigenetics*, *5*(6), 516-526.

Wu, C., & Morris, J. (2001). Genes, Genetics, and Epigenetics: A Correspondence. *Science*, *293*(5532), 1103-1105.

Xie, W., Barr, C., Kim, A., Yue, F., Lee, A., & Eubanks, J. et al. (2012). Base-Resolution Analyses of Sequence and Parent-of-Origin Dependent DNA Methylation in the Mouse Genome. *Cell*, *148*(4), 816-831.

Xu, Y., Xie, J., Cao, Y., Zhou, H., Ping, Y., & Chen, L. et al. (2014). Development of Highly Sensitive and Specific mRNA Multiplex System (XCYR1) for Forensic Human Body Fluids and Tissues Identification. *Public Library of Sciences One*, *9*(7), e100123.

Yang, I., Park, I., Jang, S., Shi, L., Ku, H., & Park, S. (2006). Rapid quantification of DNA methylation through dNMP analysis following bisulfite-PCR. *Nucleic Acids Research*, *34*(8), e61-e61.

Ye, K., Beekman, M., Lameijer, E., Zhang, Y., Moed, M., & van den Akker, E. et al. (2013). Aging as Accelerated Accumulation of Somatic Variants: Whole-Genome Sequencing of Centenarian and Middle-Aged Monozygotic Twin Pairs. *Twin Research and Human Genetics*, *16*(6), 1026-1032.

Yeo, S., Yun Ha, S., Ji Yu, E., Lee, K., Kim, J., & Kim, S. (2014). ZNF282 (Zinc finger protein 282), a novel E2F1 co-activator, promotes esophageal squamous cell carcinoma. *Oncotarget*, *5*(23), 12260-12272.

Yousefi, P., Huen, K., Schall, R., Decker, A., Elboudwarej, E., & Quach, H. et al. (2013). Considerations for normalization of DNA methylation data by Illumina 450K BeadChip assay in population studies. *Epigenetics*, *8*(11), 1141-1152.

Yu, E., Kim, S., Kim, M., Seo, W., Song, K., & Kang, M. et al. (2013). SUMOylation of ZFP282 potentiates its positive effect on estrogen signaling in breast tumorigenesis. *Oncogene*, *32*(35), 4160-4168.

Zbieć-Piekarska, R., Spólnicka, M., Kupiec, T., Makowska, Ż., Spas, A., & Parys-Proszek, A. et al. (2015). Examination of DNA methylation status of the ELOVL2 marker may be useful for human age prediction in forensic science. *Forensic Science International: Genetics*, *14*, 161-167.

Zemach, A., McDaniel, I., Silva, P., & Zilberman, D. (2010). Genome-Wide Evolutionary Analysis of Eukaryotic DNA Methylation. *Science*, *328*(5980), 916-919.

Zhang, W., Spector, T., Deloukas, P., Bell, J., & Engelhardt, B. (2015). Predicting genome-wide DNA methylation using methylation marks, genomic position, and DNA regulatory elements. *Genome Biology*, *16*(1), 14.

Zhao, F., & Bapat, B. (2016). *The role of methylation-specific PCR and associated techniques in clinical diagnostics.* (pp. 155-173). Cambridge, Massachusetts: Academic Press.

Zhao, G., Yang, Q., Huang, D., Yu, C., Yang, R., Chen, H., & Mei, K. (2005). Study on the application of parent-of-origin specific DNA methylation markers to forensic genetics. *Forensic Science International*, *154*(2-3), 122-127.

Zilberman, D., & Henikoff, S. (2007). Genome-wide analysis of DNA methylation patterns. *Development*, *134*(22), 3959-3965.

ZNF282 [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2004 – [24 August 2016]. Available from: https://www.ncbi.nlm.nih.gov/gene

Zubakov, D., Boersma, A., Choi, Y., van Kuijk, P., Wiemer, E., & Kayser, M. (2010). MicroRNA markers for forensic body fluid identification obtained from microarray screening and quantitative RT-PCR confirmation. *International Journal of Legal Medicine*, *124*(3), 217-226.

Zubakov, D., Hanekamp, E., Kokshoorn, M., van IJcken, W., & Kayser, M. (2008). Stable RNA markers for identification of blood and saliva stains revealed from whole genome expression analysis of time-wise degraded samples. *International Journal of Legal Medicine*, *122*(2), 135-142.

Zubakov, D., Liu, F., Kokmeijer, I., Choi, Y., van Meurs, J., & van IJcken, W. et al. (2016). Human age estimation from blood using mRNA, DNA methylation, DNA rearrangement, and telomere length. *Forensic Science International: Genetics*, *24*, 33-43.

Appendix A

UNIVERSITY OF
KWAZULU-NATAL
INYUVESI
YAKWAZULU-NATALI

RESEARCH OFFICE
BIOMEDICAL RESEARCH ETHICS ADMINISTRATION
Westville Campus
Govan Mbeki Building
Private Bag X 54001
Durban
4000
KwaZulu-Natal, SOUTH AFRICA
Tel: 27 31 2604769 - Fax: 27 31 260-4609
Email: BREC@ukzn.ac.za
Website: http://research.ukzn.ac.za/ResearchEthics/BiomedicalResearchEthics.aspx

08 February 2017

Ms N Naidoo (211535425)
Discipline of Genetics
School of Life Sciences
naidoo.n0308@gmail.com

Dear Ms Naidoo

Protocol: Methylation, profiling and validation of candidate tDMRs for identification of human blood, saliva, semen and vaginal fluid and its application in forensics.
Degree: MSc - Genetics
BREC reference number: BE187/16 *(sub-study of BE221/14)*

RECERTIFICATION APPLICATION APPROVAL NOTICE

Approved:	22 March 2017
Expiration of Ethical Approval:	23 March 2018

I wish to advise you that your application for Recertification dated 25 January 2017 for the above protocol has been noted and approved by a sub-committee of the Biomedical Research Ethics Committee (BREC) for another approval period. The start and end dates of this period are indicated above.

If any modifications or adverse events occur in the project before your next scheduled review, you must submit them to BREC for review. Except in emergency situations, no change to the protocol may be implemented until you have received written BREC approval for the change.

This approval will be ratified by a full Committee at its next meeting taking place on 14 March 2017.

Yours sincerely

Mrs A Marimuthu
Senior Administrator: Biomedical Research Ethics

Information Sheet and Consent to Participate in Research - Volunteers

Date: February 2016 - December 2017
Good day
My name is Natalie Naidoo from UKZN – Westville, School of Life Sciences (Genetics).
Contact Number: 031 260 8617. Email address: (naidoo.n0308@gmail.com)
You are being invited to participate in a research study. The title of the study is **"Methylation Profiling and validation of candidate tDMRs for identification of human blood, saliva, semen and vaginal fluid and its application in forensics".** This research study is towards the Masters of Science Degree of the Principal Investigator Ms. Natalie Naidoo.

The main purpose of this study is to evaluate a new test that will hopefully be used one day by the police, to easily identify body fluids such as, saliva, blood, semen and vaginal fluid found at a crime scene. The test uses genetic material (DNA) to determine whether a specimen is a body fluid. If this test works as well as we hope, then your participation will be seen as to have contributed towards helping the police solving many crimes.
To participate in this study, we will need you to provide us with the following specimen ……... We will also ask you to answer a few easy questions about yourself such as your age, weight, etc. we will not record any information, for example, name and hospital number, that will make it possible for anyone, including ourselves, to identify you in future. No one will be able to link you with the questionnaire, the specimen or the results. You may choose to not answer our questions. You will not be penalized in any way.
Your specimen will be taken to the UKZN Genetics department in Westville and stored in a freezer at -20˚C. genetic material will be taken from the specimen and this will be used to evaluate the new test. Everything will be stored until the study has been completed, after which all specimens will be destroyed. Your specimen and information will no circumstances be used for any other purpose than for this study.
The study will provide no direct benefits to participants. There are **no risks involved if you participate in this study**. Participation in this research is voluntary and you may withdraw participation at any point. In the event of refusal/withdrawal of participation you will not incur any penalty or loss.
All your details will be kept confidential not revealed to anyone else. We will provide you with a code name such as Volunteer 1, and this code will be written onto the tubes used for the sample collection as well as the accompanying body fluid data collection form.
This study has been ethically reviewed and approved by the UKZN Biomedical Research Ethics Committee (Approval Number: BE187/16 (sub-study of BE221/14)).
In the event of any problems or concerns/questions you may contact the researcher or the UKZN Biomedical Research Ethics Committee, contact details as follows:

BIOMEDICAL RESEARCH ETHICS ADMINISTRATION
Research Office, Westville Campus
Govan Mbeki Building
Private Bag X 54001
Durban
4000
KwaZulu-Natal, SOUTH AFRICA
Tel: 27 31 2604769 - Fax: 27 31 2604609
Email: BREC@ukzn.ac.za

INFORMED CONSENT

I (Full Name/s and Surname……) have been informed about the study entitled "Methylation Profiling and validation of candidate tDMRs for identification of human blood, saliva, semen and vaginal fluid and its application in forensics" by …….

I understand the purpose and procedures of the study.

I have been given an opportunity to ask questions about the study and have had answers to my satisfaction.

I declare that my participation in this study is entirely voluntary and that I may withdraw at any time without affecting any treatment or care that I would usually be entitled to.

If I have any further questions/concerns or queries related to the study I understand that I may contact the researcher at UKZN – Westville, School of Life Sciences (Genetics). Contact Number: 031 260 8617. Email address: Natalie Naidoo (naidoo.n0308@gmail.com).

If I have any questions or concerns about my rights as a study participant, or if I am concerned about an aspect of the study or the researchers then I may contact:

BIOMEDICAL RESEARCH ETHICS ADMINISTRATION
Research Office, Westville Campus
Govan Mbeki Building
Private Bag X 54001
Durban
4000
KwaZulu-Natal, SOUTH AFRICA
Tel: 27 31 2604769 - Fax: 27 31 2604609
Email: BREC@ukzn.ac.za

_____ _____
Signature of Participant **Date**

_____ _____
Signature of Witness **Date**
(Where applicable)

_____ _____
Signature of Translator **Date**
(Where applicable)

Questionnaire- Please tick the appropriate boxes where necessary.

1. Age

   Below 30 ☐

   Between 30 – 50 ☐

   Above 50 ☐

2. Sex

   Male ☐

   Female ☐

3. Weight

   Below 40 kg ☐

   Between 40 – 70 kg ☐

   Above 70 kg ☐

4. Do you smoke?

   Yes ☐

   No ☐

5. Do you consume alcohol?

   Yes ☐

   No ☐

6. Have you ever had a blood transfusion?

   Yes ☐

   No ☐

7. Do you suffer from any chronic illnesses? If yes, please elaborate.

   ……

   ……

8. Do you suffer from any cardiovascular illnesses? If yes, please elaborate.

   ……

   ……

9. Have you undergone any medical procedures? If yes, please elaborate.

   ……

   ……

10. Do you take any medication? This includes steroids. If yes, please elaborate.

    ……

    ……

11. Your assistance is well appreciated. Kindly note that the present study is not for diagnostic purposes. All information disclosed will be kept confidential.

Appendix B

Table 1: Concentration of DNA obtained from blood, vaginal fluid, semen and saliva.

Sample ID	Nucleic Acid Concentration
Blood 1	16,1 ng/µl
Blood 2	21,5 ng/µl
Blood 3	21,3 ng/µl
Blood 4	23,5 ng/µl
Blood 5	25,7 ng/µl
Blood 6	15,9 ng/µl
Blood 7	39,7 ng/µl
Blood 8	24 ng/µl
Blood 9	53,5 ng/µl
Blood 10	22,9 ng/µl
Blood 11	25 ng/µl
Vaginal fluid 1	101,7 ng/µl
Vaginal fluid 2	494,8 ng/µl
Vaginal fluid 3	54,3 ng/µl
Vaginal fluid 4	124,4 ng/µl
Vaginal fluid 5	179,4 ng/µl
Vaginal fluid 6	60,3 ng/µl
Vaginal fluid 7	206,7 ng/µl
Vaginal fluid 8	44,6 ng/µl
Vaginal fluid 9	64,2 ng/µl
Vaginal fluid 10	38,3 ng/µl
Semen 1	26,2 ng/µl
Semen 2	10,6 ng/µl
Semen 3	12,8 ng/µl
Semen 4	25,7 ng/µl
Semen 5	11,4 ng/µl
Semen 6	19,2 ng/µl
Semen 7	19,8 ng/µl
Semen 8	16,3 ng/µl
Semen 9	15,8 ng/µl
Semen 10	17,8 ng/µl
Saliva 1	14,6 ng/µl
Saliva 2	36,5 ng/µl
Saliva 3	36,1 ng/µl
Saliva 4	182,3 ng/µl
Saliva 5	187,5 ng/µl
Saliva 6	79 ng/µl
Saliva 7	107,8 ng/µl
Saliva 8	21,1 ng/µl
Saliva 9	54,2 ng/µl
Saliva 10	64 ng/µl

Appendix C

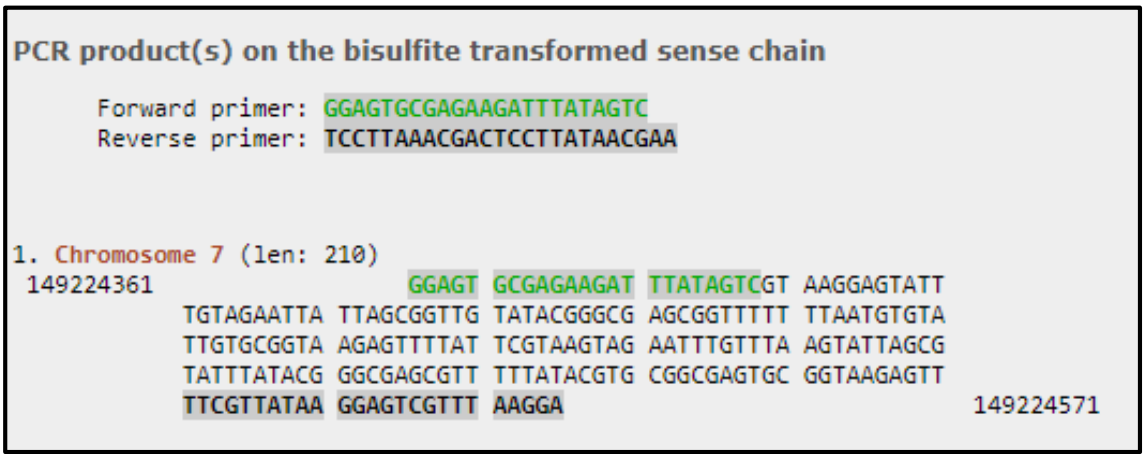


Figure 1: BiSearch results depicting specificity of MSP primers for methylated cytosine of ZNF282 tDMR.

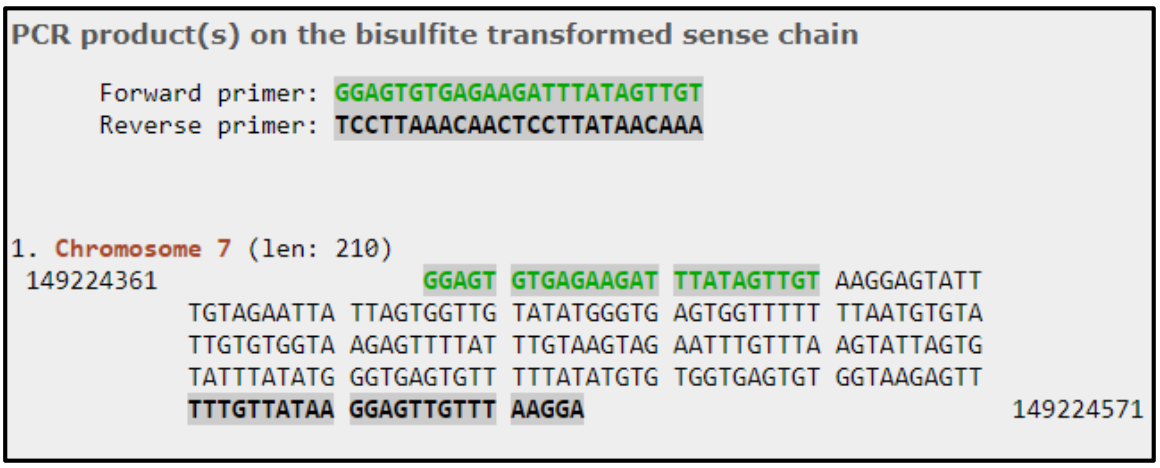


Figure 2: BiSearch results depicting specificity of MSP primers for unmethylated cytosine of ZNF282 tDMR.

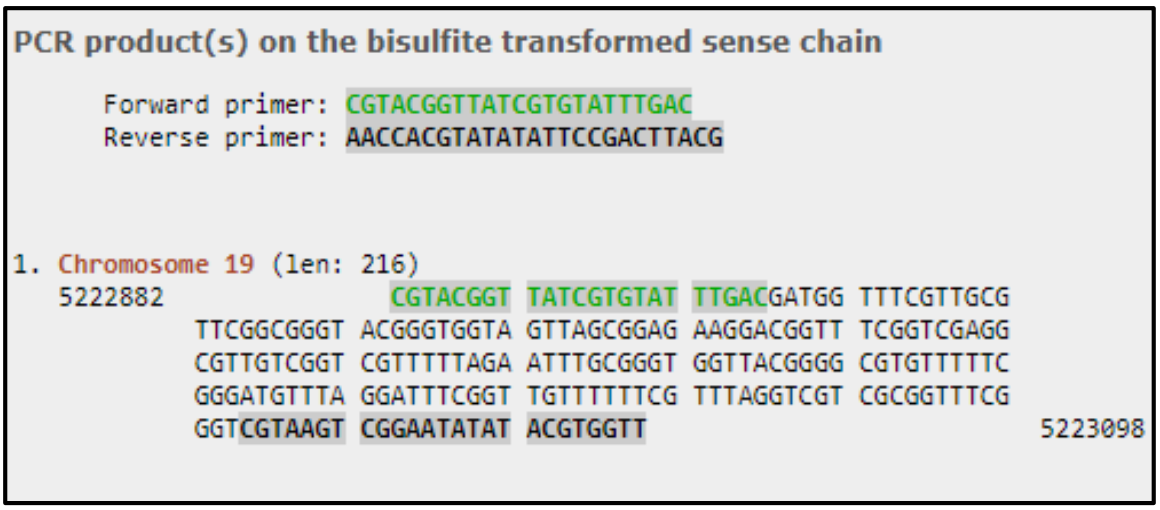


Figure 3: BiSearch results depicting specificity of MSP primers for methylated cytosine of PTPRS 1 tDMR.

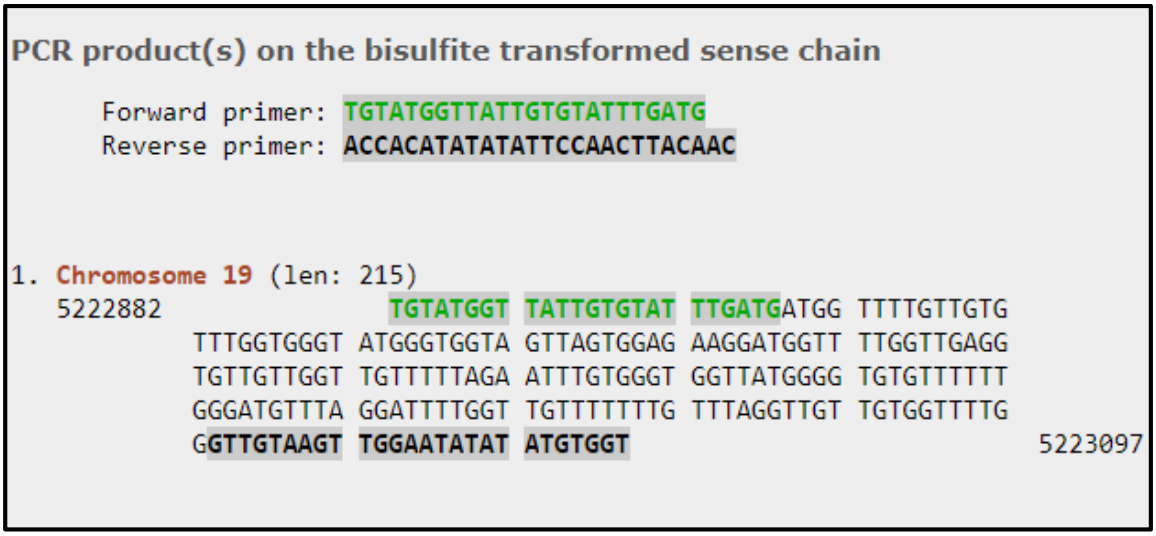


Figure 4: BiSearch results depicting specificity of MSP primers for unmethylated cytosine of PTPRS 1 tDMR.

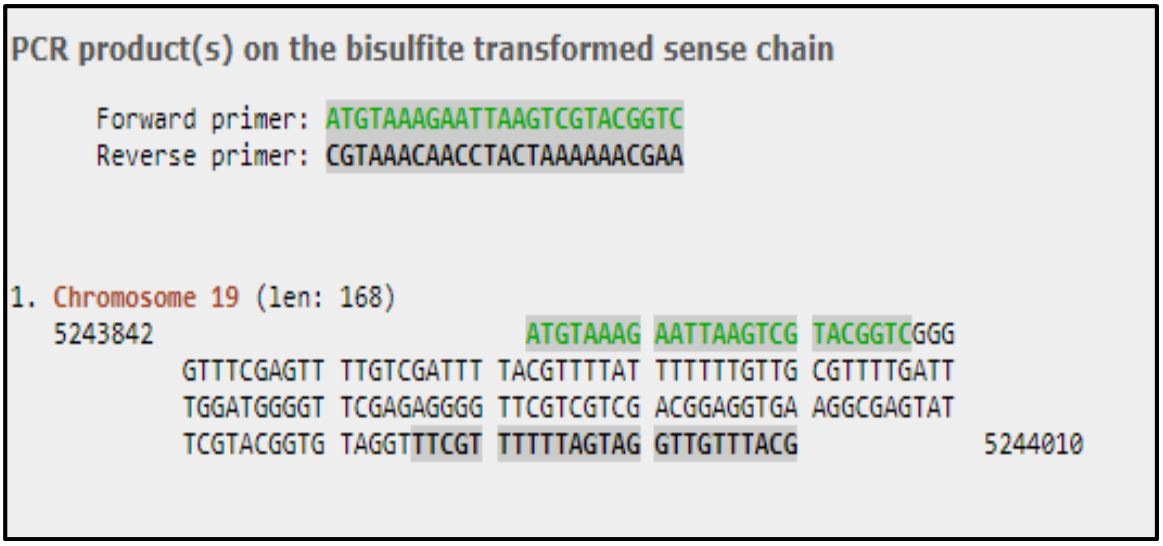


Figure 5: BiSearch results depicting specificity of MSP primers for methylated cytosine of PTPRS 2 tDMR.

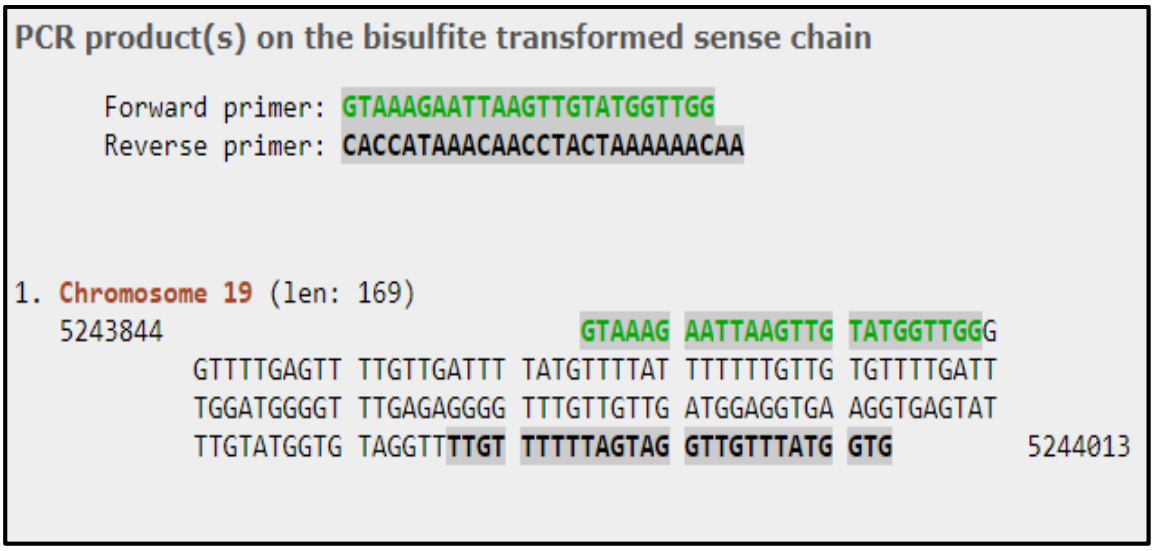


Figure 6: BiSearch results depicting specificity of MSP primers for unmethylated cytosine of PTPRS 2 tDMR.

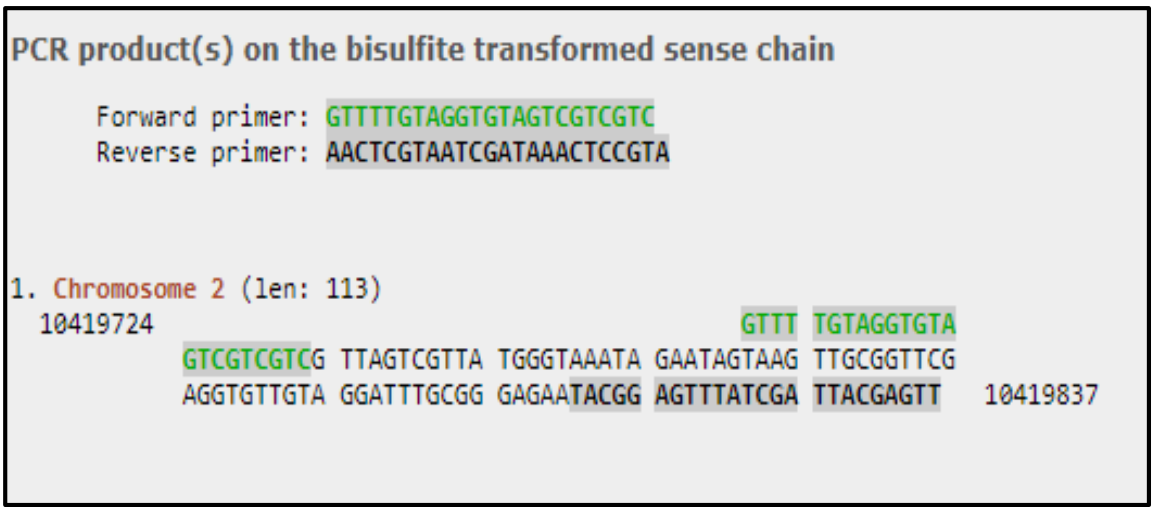


Figure 7: BiSearch results depicting specificity of MSP primers for methylated cytosine of HPCAL1 tDMR.

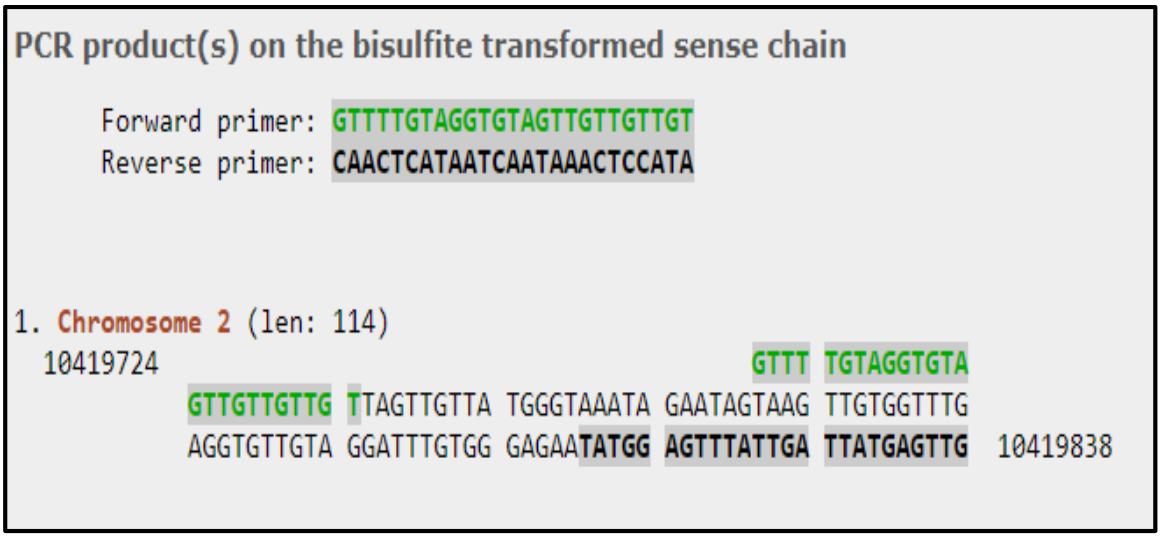


Figure 8: BiSearch results depicting specificity of MSP primers for unmethylated cytosine of HPCAL1 tDMR.

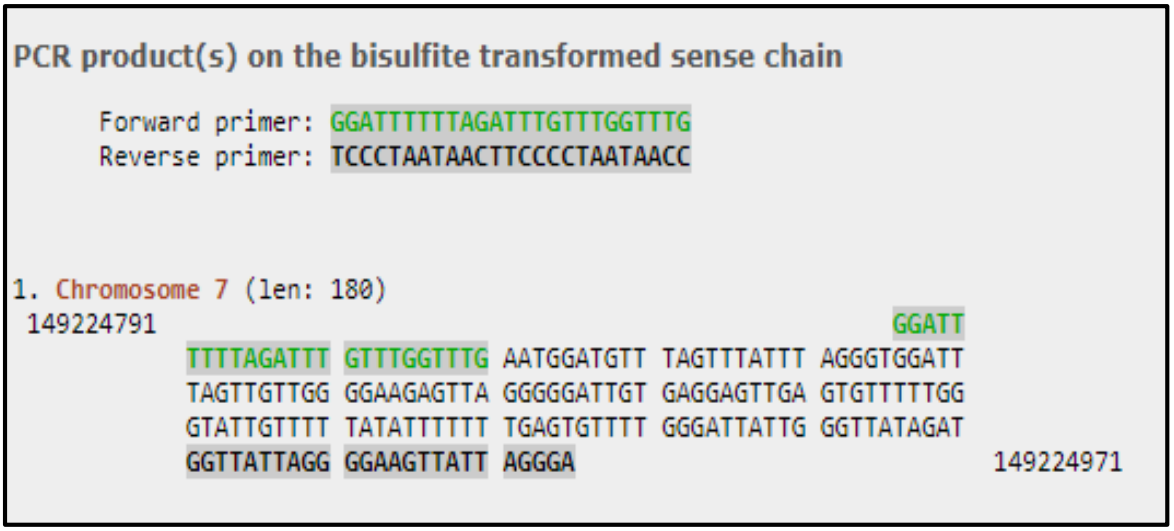


Figure 9: BiSearch results depicting specificity of BS primers designed for ZNF282 tDMR.

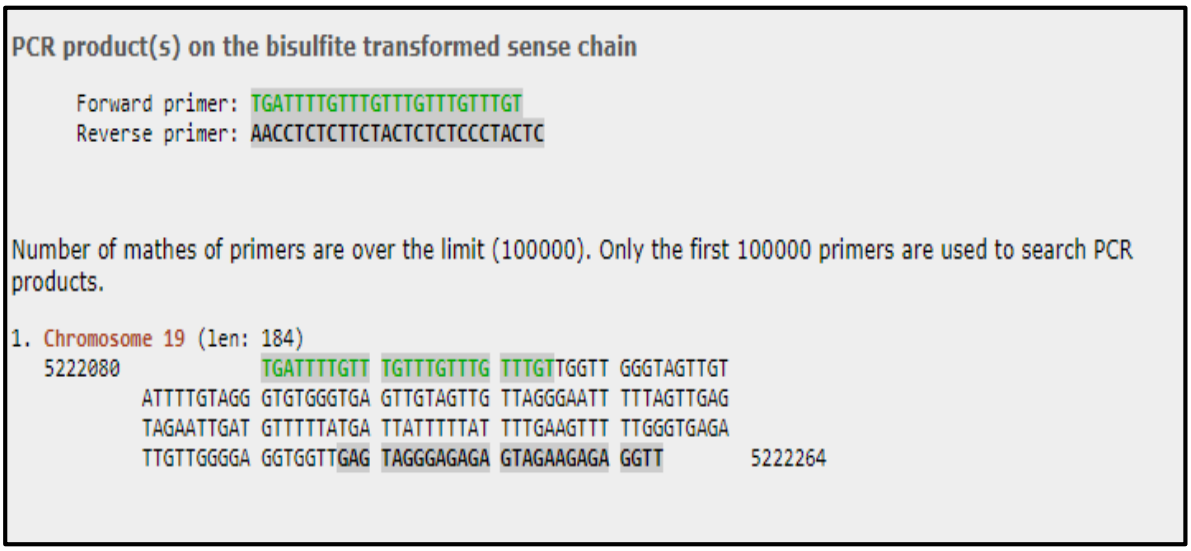


Figure 10: BiSearch results depicting specificity of BS primers designed for PTPRS 1 tDMR.

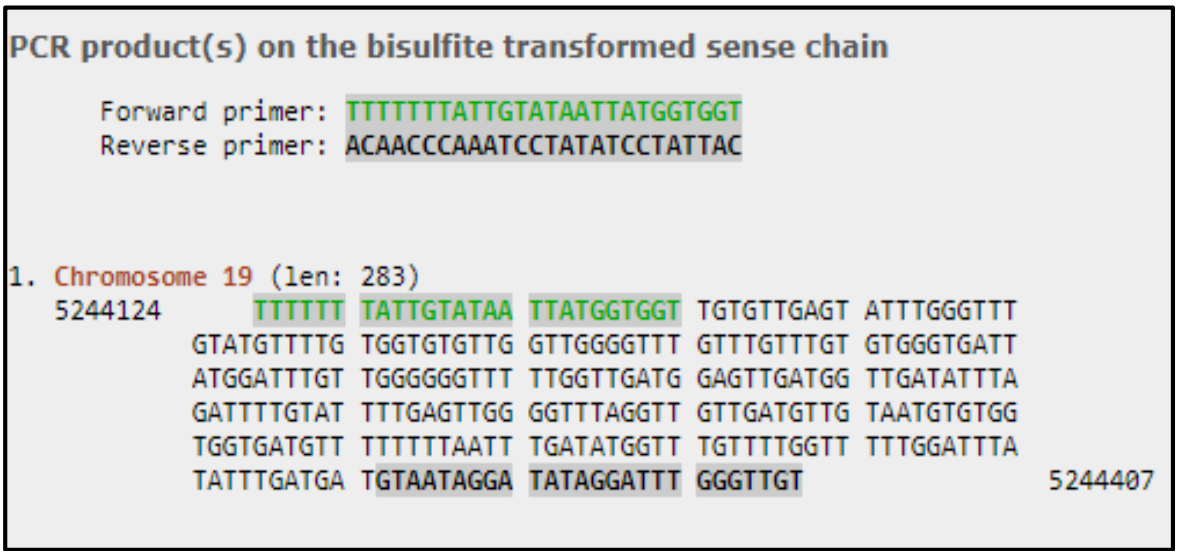


Figure 11: BiSearch results depicting specificity of BS primers designed for PTPRS 2 tDMR.

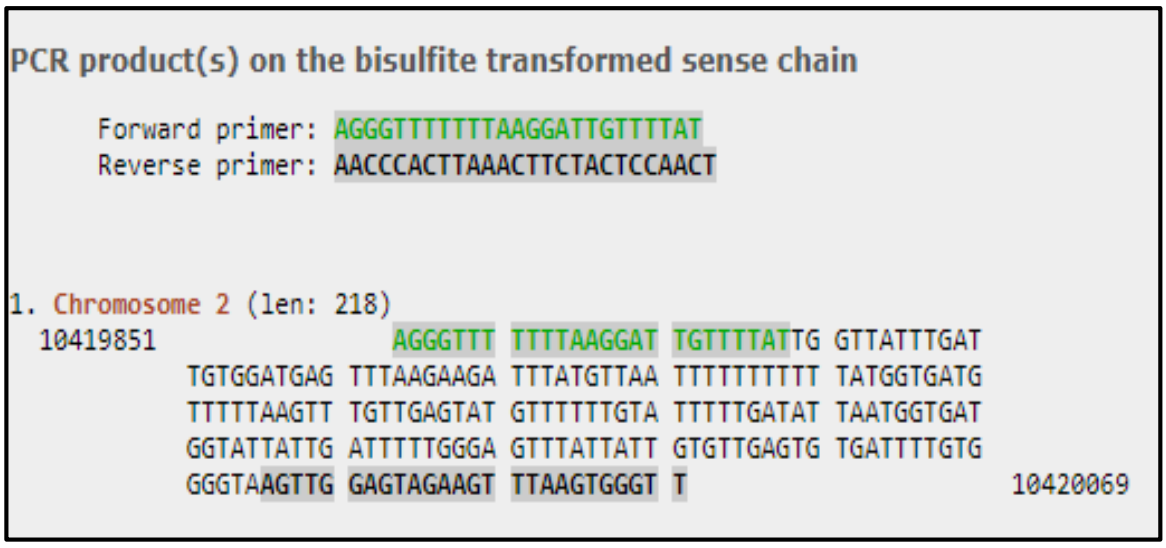


Figure 12: BiSearch results depicting specificity of BS primers designed for HPCAL1 tDMR.

Appendix D

Table 1: Pairwise comparison of total DNA methylation at CpG 1 for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.0001	0.0000	1	
Vaginal fluid	1.0000	1.0000	0.0001	1

Table 2: Pairwise comparison of total DNA methylation at CpG 2 for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.0001	0.0000	1	
Vaginal fluid	1.0000	1.0000	0.0001	1

Table 3: Pairwise comparison of total DNA methylation at CpG 3 for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	0.4667	1		
Semen	0.0009	0.0000	1	
Vaginal fluid	1.0000	1.0000	0.0001	1

Table 4: Pairwise comparison of total DNA methylation at CpG 4 for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.0014	0.0004	1	
Vaginal fluid	1.0000	1.0000	0.0014	1

Table 5: Pairwise comparison of total DNA methylation at CpG 5 for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.0001	0.0000	1	
Vaginal fluid	1.0000	1.0000	0.0001	1

Table 6: Pairwise comparison of total DNA methylation at CpG 6 for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.0001	0.0000	1	
Vaginal fluid	1.0000	0.4667	0.0009	1

Table 7: Pairwise comparison of total DNA methylation at CpG 7 for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p* -values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	0.4667	1		
Semen	0.0009	0.0000	1	
Vaginal fluid	1.0000	1.0000	0.0001	1

Table 8: Pairwise comparison of total DNA methylation at CpG 8 for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.0009	0.0004	1	
Vaginal fluid	1.0000	1.0000	0.0001	1

Appendix E

Table 1: Pairwise comparison of total DNA methylation at CpG 1 for PTPRS 1 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	0.3077	1		
Semen	1.0000	1.0000	1	
Vaginal fluid	0.3077	1.0000	1.0000	1

Table 2: Pairwise comparison of total DNA methylation at CpG 2 for PTPRS 1 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	0.0769	1		
Semen	0.5301	0.4706	1	
Vaginal fluid	0.0769	1.0000	0.4706	1

Table 3: Pairwise comparison of total DNA methylation at CpG 3 for PTPRS 1 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	0.3333	1		
Semen	1.0000	1.0000	1	
Vaginal fluid	0.3077	1.0000	1.0000	1

Table 4: Pairwise comparison of total DNA methylation at CpG 4 for PTPRS 1 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	1.0000	1.0000	1	
Vaginal fluid	1.0000	1.0000	1.0000	1

Table 5: Pairwise comparison of total DNA methylation at CpG 5 for PTPRS 1 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	1.0000	1.0000	1	
Vaginal fluid	1.0000	1.0000	1.0000	1

Appendix F

Table 1: Pairwise comparison of total DNA methylation at CpG 1 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.2000	1

Table 2: Pairwise comparison of total DNA methylation at CpG 2 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.2000	1

Table 3: Pairwise comparison of total DNA methylation at CpG 3 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.2000	1

Table 4: Pairwise comparison of total DNA methylation at CpG 4 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	0.3333	1		
Semen	1.0000	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.4000	1

Table 5: Pairwise comparison of total DNA methylation at CpG 5 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.2000	1

Table 6: Pairwise comparison of total DNA methylation at CpG 6 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.2000	1

Table 7: Pairwise comparison of total DNA methylation at CpG 7 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.2000	1

Table 8: Pairwise comparison of total DNA methylation at CpG 8 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.2000	1

Table 9: Pairwise comparison of total DNA methylation at CpG 9 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.2000	1

Table 10: Pairwise comparison of total DNA methylation at CpG 10 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.4000	1

Table 11: Pairwise comparison of total DNA methylation at CpG 11 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.2000	1

Table 12: Pairwise comparison of total DNA methylation at CpG 12 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.2000	1

Table 13: Pairwise comparison of total DNA methylation at CpG 13 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.2000	1

Table 14: Pairwise comparison of total DNA methylation at CpG 14 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.2000	1

Table 15: Pairwise comparison of total DNA methylation at CpG 15 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.2000	1

Table 16: Pairwise comparison of total DNA methylation at CpG 16 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	1.0000	1.0000	1	
Vaginal fluid	1.0000	1.0000	1.0000	1

Table 17: Pairwise comparison of total DNA methylation at CpG 17 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.2000	1

Table 18: Pairwise comparison of total DNA methylation at CpG 18 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.2000	1

Table 19: Pairwise comparison of total DNA methylation at CpG 19 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	1.0000	1.0000	1	
Vaginal fluid	1.0000	1.0000	1.0000	1

Table 20: Pairwise comparison of total DNA methylation at CpG 20 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	1.0000	1.0000	1	
Vaginal fluid	1.0000	1.0000	1.0000	1

Table 21: Pairwise comparison of total DNA methylation at CpG 21 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.2000	1

Table 22: Pairwise comparison of total DNA methylation at CpG 22 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.2000	1	
Vaginal fluid	1.0000	1.0000	0.4000	1

Table 23: Pairwise comparison of total DNA methylation at CpG 23 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.3333	0.4000	1	
Vaginal fluid	1.0000	1.0000	0.2000	1

Table 24: Pairwise comparison of total DNA methylation at CpG 24 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	1.0000	1.0000	1	
Vaginal fluid	1.0000	1.0000	1.0000	1

Table 25: Pairwise comparison of total DNA methylation at CpG 25 for PTPRS 2 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	1.0000	1.0000	1	
Vaginal fluid	1.0000	1.0000	1.0000	1

Appendix G

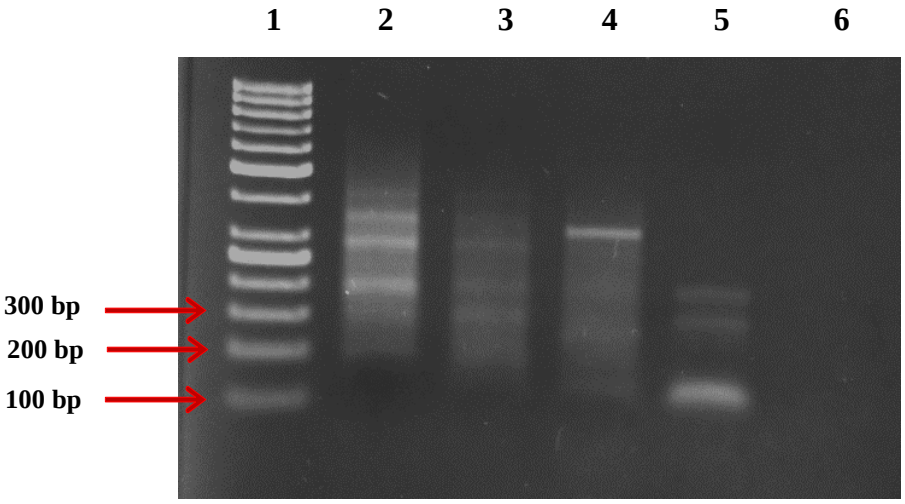


Figure 1: BS PCR reaction of all four body fluids for the HPCAL1 tDMR primer set. Products with primers showing non-specific amplification. Lane 1 – 100 bp ladder (Thermo Scientific). Lane 2 – Lane 5- blood, saliva, semen and vaginal fluid DNA samples, respectively. Lane 6- No template.

Appendix H

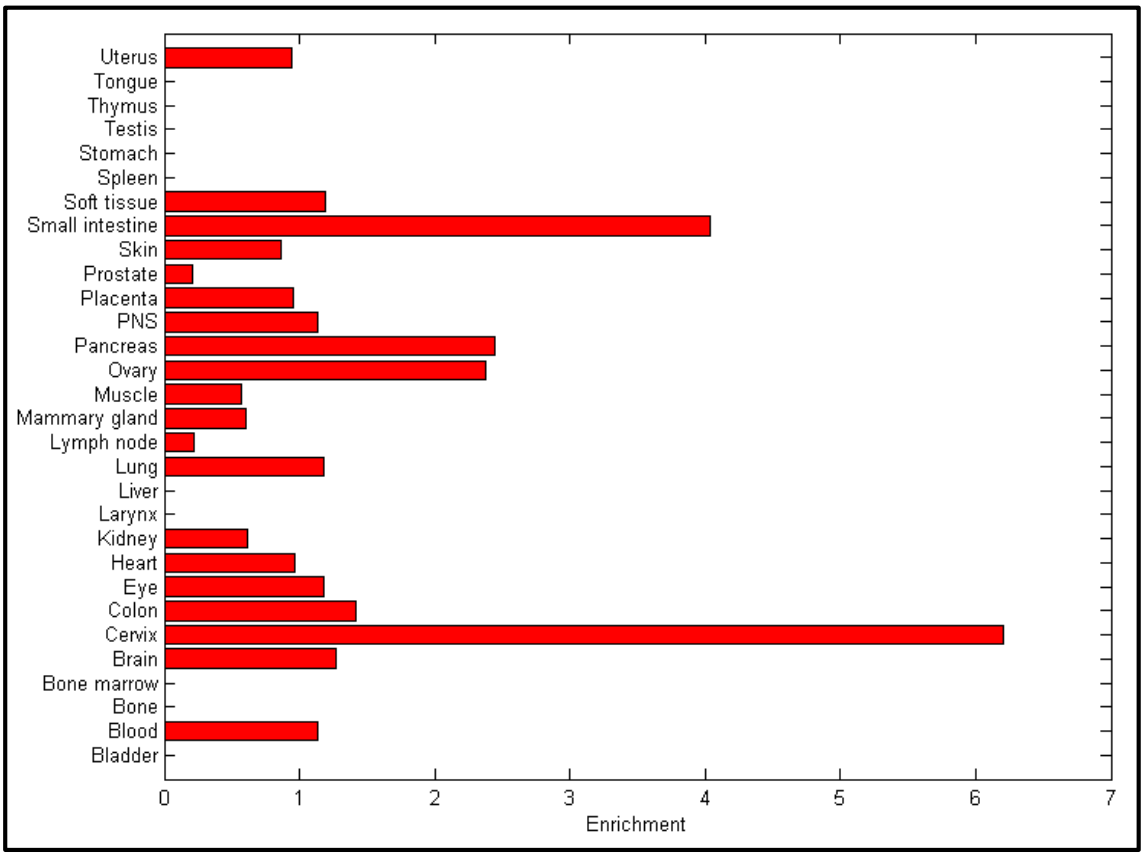


Figure 1: Expressed sequence tag (EST) profile of *ZNF282* gene in Tissue-specific Gene Expression and Regulation (TiGER) database.

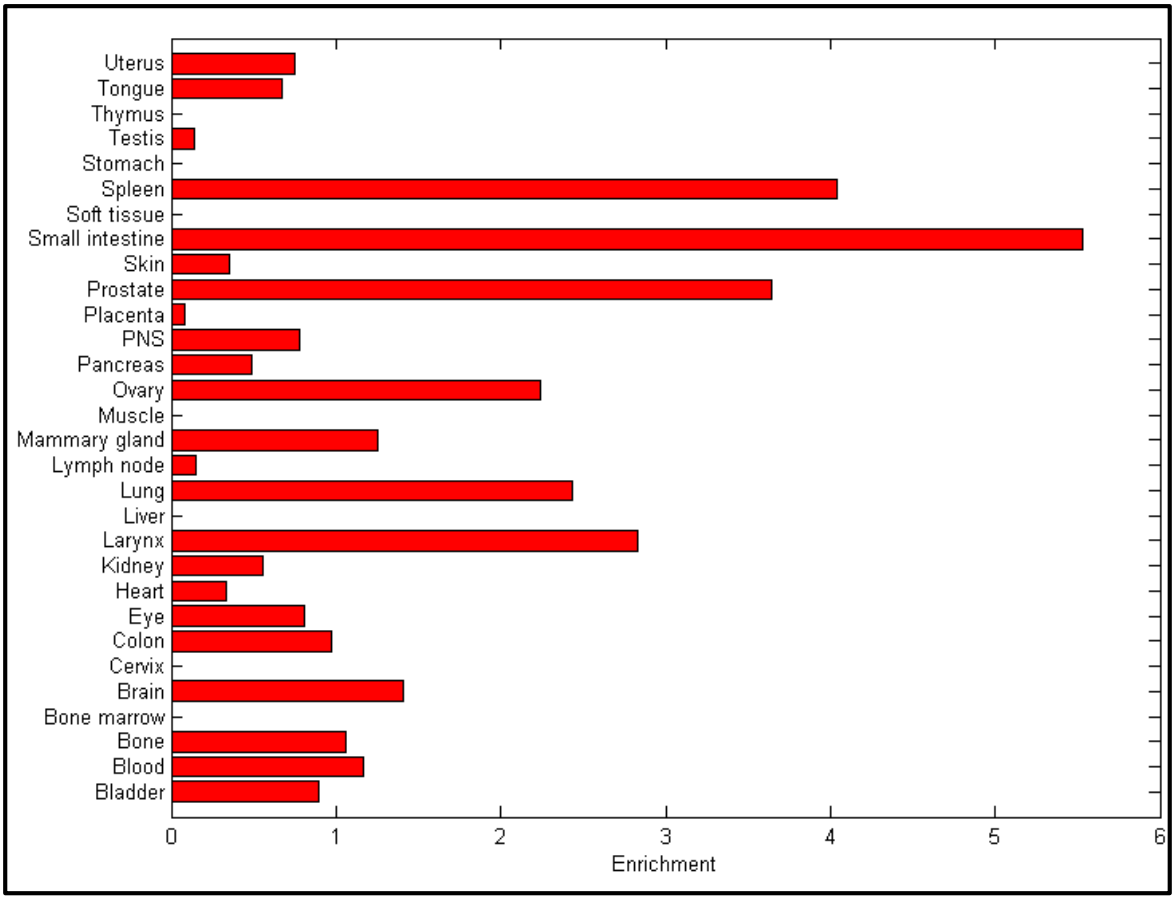


Figure 2: Expressed sequence tag (EST) profile of *PTPRS* gene in Tissue-specific Gene Expression and Regulation (TiGER) database.

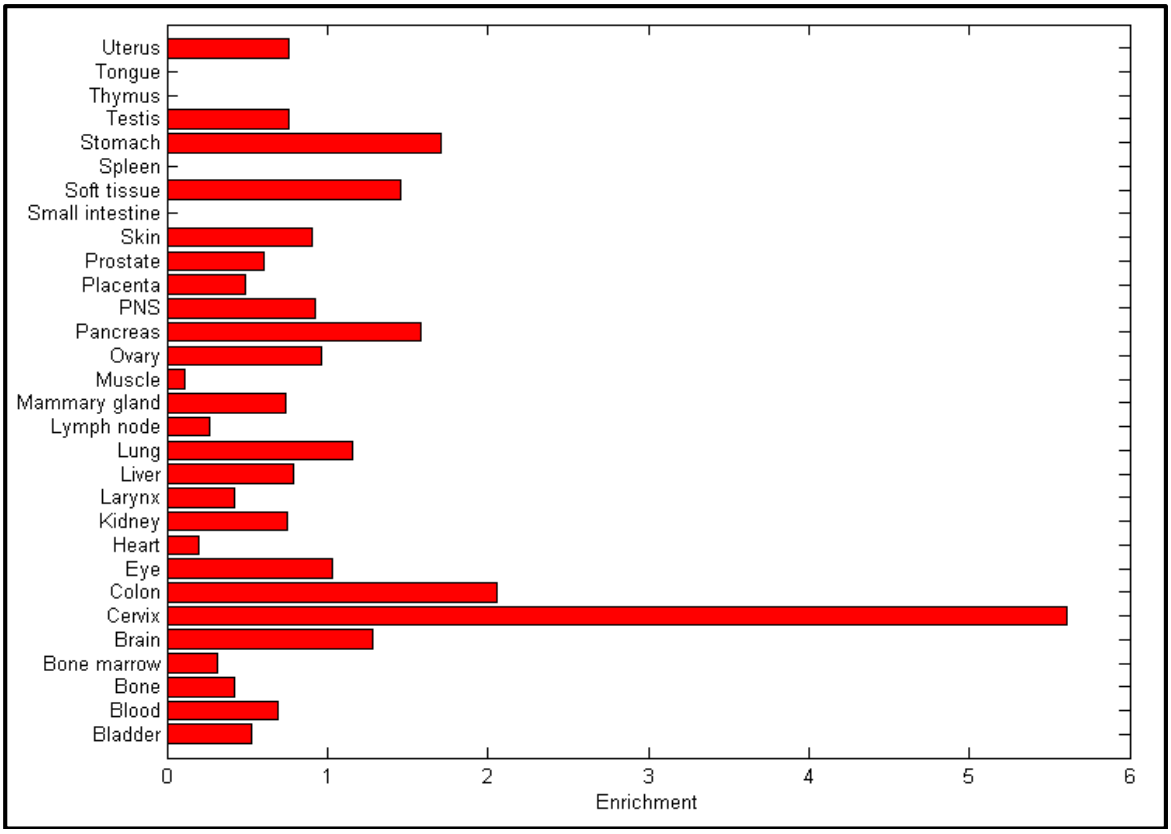


Figure 3: Expressed sequence tag (EST) profile of *HPCAL1* gene in Tissue-specific Gene Expression and Regulation (TiGER) database.

Appendix I

Table 1: DNA concentrations of five samples analysed for each body fluid not exposed to environmental insults (t0 days).

Sample ID	Nucleic Acid Concentration
Blood 1	14 ng/µl
Blood 2	15,3 ng/µl
Blood 3	41,1 ng/µl
Blood 4	28,1 ng/µl
Blood 5	35,6 ng/µl
Vaginal fluid 1	205,3 ng/µl
Vaginal fluid 2	168,6 ng/µl
Vaginal fluid 3	155,3 ng/µl
Vaginal fluid 4	163,9 ng/µl
Vaginal fluid 5	197,1 ng/µl
Semen 1	26,2 ng/µl
Semen 2	10,6 ng/µl
Semen 3	25,7 ng/µl
Semen 4	11,4 ng/µl
Semen 5	19,2 ng/µl
Saliva 1	36,5 ng/µl
Saliva 2	182,3 ng/µl
Saliva 3	187,5 ng/µl
Saliva 4	107,8 ng/µl
Saliva 5	54,2 ng/µl

Table 2: DNA concentrations of five samples analysed for blood after t50 days.

	Sample ID	Nucleic Acid Concentration
Condition A	Blood 1	3.5 ng/µl
	Blood 2	4.8 ng/µl
	Blood 3	3 ng/µl
	Blood 4	4.4 ng/µl
	Blood 5	4 ng/µl
Condition B	Blood 1	5.1 ng/µl
	Blood 2	4.7 ng/µl
	Blood 3	4.5 ng/µl
	Blood 4	5.5 ng/µl
	Blood 5	3.5 ng/µl
Condition C	Blood 1	12.6 ng/µl
	Blood 2	84.3 ng/µl
	Blood 3	10.3 ng/µl
	Blood 4	8.5 ng/µl
	Blood 5	13.2 ng/µl
Condition D	Blood 1	5.6 ng/µl
	Blood 2	3.1 ng/µl
	Blood 3	1.7 ng/µl
	Blood 4	4.1 ng/µl
	Blood 5	2.2 ng/µl
Condition E	Blood 1	16.2 ng/µl
	Blood 2	2.4 ng/µl
	Blood 3	3.3 ng/µl
	Blood 4	3.9 ng/µl
	Blood 5	4.6 ng/µl

Table 3: DNA concentrations of five samples analysed for saliva after t50 days.

	Sample ID	Nucleic Acid Concentration
Condition A	Saliva 1	1.4 ng/µl
	Saliva 2	14.9 ng/µl
	Saliva 3	6.5 ng/µl
	Saliva 4	9.6 ng/µl
	Saliva 5	12.3 ng/µl
Condition B	Saliva 1	3.1 ng/µl
	Saliva 2	12.1 ng/µl
	Saliva 3	10.8 ng/µl
	Saliva 4	4.5 ng/µl
	Saliva 5	9.7 ng/µl
Condition C	Saliva 1	6.4 ng/µl
	Saliva 2	10.6 ng/µl
	Saliva 3	11.6 ng/µl
	Saliva 4	5.5 ng/µl
	Saliva 5	9.4 ng/µl
Condition D	Saliva 1	1.8 ng/µl
	Saliva 2	9.7 ng/µl
	Saliva 3	13.9 ng/µl
	Saliva 4	14.3 ng/µl
	Saliva 5	12.3 ng/µl
Condition E	Saliva 1	1.9 ng/µl
	Saliva 2	1.5 ng/µl
	Saliva 3	6 ng/µl
	Saliva 4	4.5 ng/µl
	Saliva 5	3.6 ng/µl

Table 4: DNA concentrations of five samples analysed for semen after t50 days.

	Sample ID	Nucleic Acid Concentration
Condition A	Semen 1	1 ng/µl
	Semen 2	0.2 ng/µl
	Semen 3	0.3 ng/µl
	Semen 4	0.8 ng/µl
	Semen 5	2 ng/µl
Condition B	Semen 1	1.3 ng/µl
	Semen 2	1.3 ng/µl
	Semen 3	0.6 ng/µl
	Semen 4	2.2 ng/µl
	Semen 5	0.5 ng/µl
Condition C	Semen 1	13.3 ng/µl
	Semen 2	16 ng/µl
	Semen 3	12.5 ng/µl
	Semen 4	15.9 ng/µl
	Semen 5	16.8 ng/µl
Condition D	Semen 1	0.7 ng/µl
	Semen 2	1.6 ng/µl
	Semen 3	1.4 ng/µl
	Semen 4	0.4 ng/µl
	Semen 5	0.8 ng/µl
Condition E	Semen 1	0.5 ng/µl
	Semen 2	0.8 ng/µl
	Semen 3	0.5 ng/µl
	Semen 4	1 ng/µl
	Semen 5	2 ng/µl

Table 5: DNA concentrations of five samples analysed for vaginal fluid after t50 days.

	Sample ID	Nucleic Acid Concentration
Condition A	Vaginal fluid 1	143.8 ng/µl
	Vaginal fluid 2	35 ng/µl
	Vaginal fluid 3	130 ng/µl
	Vaginal fluid 4	62.3 ng/µl
	Vaginal fluid 5	102.6 ng/µl
Condition B	Vaginal fluid 1	156 ng/µl
	Vaginal fluid 2	44.4 ng/µl
	Vaginal fluid 3	120.8 ng/µl
	Vaginal fluid 4	52.6 ng/µl
	Vaginal fluid 5	113 ng/µl
Condition C	Vaginal fluid 1	12.7 ng/µl
	Vaginal fluid 2	12.8 ng/µl
	Vaginal fluid 3	11.2 ng/µl
	Vaginal fluid 4	10.5 ng/µl
	Vaginal fluid 5	13.5 ng/µl
Condition D	Vaginal fluid 1	165.9 ng/µl
	Vaginal fluid 2	1.3 ng/µl
	Vaginal fluid 3	0.7 ng/µl
	Vaginal fluid 4	142.8 ng/µl
	Vaginal fluid 5	50.6 ng/µl
Condition E	Vaginal fluid 1	3 ng/µl
	Vaginal fluid 2	332.4 ng/µl
	Vaginal fluid 3	171.6 ng/µl
	Vaginal fluid 4	1.4 ng/µl
	Vaginal fluid 5	200.6 ng/µl

Appendix J

Table 1: Pairwise comparison of total DNA methylation at CpG 1 for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.0079	0.0079	1	
Vaginal fluid	1.000	1.000	0.0179	1

Table 2: Pairwise comparison of total DNA methylation at CpG 2 for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.0079	0.0079	1	
Vaginal fluid	1.0000	1.0000	0.0179	1

Table 3: Pairwise comparison of total DNA methylation at CpG 3 for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.0079	0.0079	1	
Vaginal fluid	1.0000	1.0000	0.0179	1

Table 4: Pairwise comparison of total DNA methylation at CpG 4 for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.0079	0.0079	1	
Vaginal fluid	1.0000	1.0000	0.0179	1

Table 5: Pairwise comparison of total DNA methylation at CpG 5 for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.0079	0.0079	1	
Vaginal fluid	1.0000	1.0000	0.0179	1

Table 6: Pairwise comparison of total DNA methylation at CpG 6 for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.0079	0.0079	1	
Vaginal fluid	1.0000	1.0000	0.0179	1

Table 7: Pairwise comparison of total DNA methylation at CpG 7 for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	1.0000	1		
Semen	0.0079	0.0079	1	
Vaginal fluid	1.0000	1.0000	0.0179	1

Table 8: Pairwise comparison of total DNA methylation at CpG 8 for ZNF282 tDMR in blood, saliva, semen and vaginal fluid. The *p*-values obtained from the Fisher's Exact test are indicated below.

	Blood	**Saliva**	**Semen**	**Vaginal fluid**
Blood	1			
Saliva	0.4444	1		
Semen	0.0079	0.0079	1	
Vaginal fluid	1.0000	1.0000	0.0179	1

Appendix K

Table 1: Pairwise comparison of DNA methylation at CpG 1 of the ZNF282 tDMR for comparison of t0 days to t50 days in Condition C (outside on the ground). The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below.

	t0 days			
	Blood	**Saliva**	**Semen**	**Vaginal fluid**
t50 days	1.0000	0.2857	1.0000	1.0000

Table 2: Pairwise comparison of DNA methylation at CpG 2 of the ZNF282 tDMR for comparison of t0 days to t50 days in Condition C (outside on the ground). The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below.

	t0 days			
	Blood	**Saliva**	**Semen**	**Vaginal fluid**
t50 days	1.0000	0.2857	1.0000	1.0000

Table 3: Pairwise comparison of DNA methylation at CpG 3 of the ZNF282 tDMR for comparison of t0 days to t50 days in Condition C (outside on the ground). The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below.

	t0 days			
	Blood	**Saliva**	**Semen**	**Vaginal fluid**
t50 days	1.0000	0.2857	1.0000	1.0000

Table 4: Pairwise comparison of DNA methylation at CpG 4 of the ZNF282 tDMR for comparison of t0 days to t50 days in Condition C (outside on the ground). The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below.

	t0 days			
	Blood	**Saliva**	**Semen**	**Vaginal fluid**
t50 days	1.0000	0.2857	1.0000	1.0000

Table 5: Pairwise comparison of DNA methylation at CpG 5 of the ZNF282 tDMR for comparison of t0 days to t50 days in Condition C (outside on the ground). The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below.

	t0 days			
	Blood	**Saliva**	**Semen**	**Vaginal fluid**
t50 days	1.0000	0.2857	1.0000	1.0000

Table 6: Pairwise comparison of DNA methylation at CpG 6 of the ZNF282 tDMR for comparison of t0 days to t50 days in Condition C (outside on the ground). The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below.

	t0 days			
	Blood	**Saliva**	**Semen**	**Vaginal fluid**
t50 days	1.0000	0.2857	1.0000	1.0000

Table 7: Pairwise comparison of DNA methylation at CpG 7 of the ZNF282 tDMR for comparison of t0 days to t50 days in Condition C (outside on the ground). The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below.

			t0 days	
	Blood	**Saliva**	**Semen**	**Vaginal fluid**
t50 days	1.0000	0.2857	1.0000	1.0000

Table 8: Pairwise comparison of DNA methylation at CpG 8 of the ZNF282 tDMR for comparison of t0 days to t50 days in Condition C (outside on the ground). The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below.

			t0 days	
	Blood	**Saliva**	**Semen**	**Vaginal fluid**
t50 days	0.4444	0.2857	1.0000	1.0000

Appendix L

Table 1: Pairwise comparison of DNA methylation at CpG 1 of the ZNF282 tDMR for vaginal fluid in all conditions. The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below. A= dry at room temperature, B= wet in an exsiccator, C= outside on the ground, D= sprayed with alcohol and E= sprayed with bleach.

	Vaginal fluid A	**Vaginal fluid B**	**Vaginal fluid C**	**Vaginal fluid D**	**Vaginal fluid E**
Vaginal fluid A	1				
Vaginal fluid B	1.0000	1			
Vaginal fluid C	1.0000	1.0000	1		
Vaginal fluid D	1.0000	1.0000	1.0000	1	
Vaginal fluid E	1.0000	1.0000	1.0000	1.0000	1

Table 2: Pairwise comparison of DNA methylation at CpG 2 of the ZNF282 tDMR for vaginal fluid in all conditions. The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below. A= dry at room temperature, B= wet in an exsiccator, C= outside on the ground, D= sprayed with alcohol and E= sprayed with bleach.

	Vaginal fluid A	**Vaginal fluid B**	**Vaginal fluid C**	**Vaginal fluid D**	**Vaginal fluid E**
Vaginal fluid A	1				
Vaginal fluid B	0.3333	1			
Vaginal fluid C	1.0000	1.0000	1		
Vaginal fluid D	0.2857	1.0000	1.0000	1	
Vaginal fluid E	0.2857	1.0000	1.0000	1.0000	**1**

Table 3: Pairwise comparison of DNA methylation at CpG 3 of the ZNF282 tDMR for vaginal fluid in all conditions. The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below. A= dry at room temperature, B= wet in an exsiccator, C= outside on the ground, D= sprayed with alcohol and E= sprayed with bleach.

	Vaginal fluid A	**Vaginal fluid B**	**Vaginal fluid C**	**Vaginal fluid D**	**Vaginal fluid E**
Vaginal fluid A	1				
Vaginal fluid B	1.0000	1			
Vaginal fluid C	1.0000	1.0000	1		
Vaginal fluid D	1.0000	1.0000	1.0000	1	
Vaginal fluid E	1.0000	1.0000	1.0000	1.0000	1

Table 4: Pairwise comparison of DNA methylation at CpG 4 of the ZNF282 tDMR for vaginal fluid in all conditions. The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below. A= dry at room temperature, B= wet in an exsiccator, C= outside on the ground, D= sprayed with alcohol and E= sprayed with bleach.

	Vaginal fluid A	**Vaginal fluid B**	**Vaginal fluid C**	**Vaginal fluid D**	**Vaginal fluid E**
Vaginal fluid A	1				
Vaginal fluid B	1.0000	1			
Vaginal fluid C	1.0000	1.0000	1		
Vaginal fluid D	1.0000	1.0000	1.0000	1	
Vaginal fluid E	1.0000	1.0000	1.0000	1.0000	1

Table 5: Pairwise comparison of DNA methylation at CpG 5 of the ZNF282 tDMR for vaginal fluid in all conditions. The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below. A= dry at room temperature, B= wet in an exsiccator, C= outside on the ground, D= sprayed with alcohol and E= sprayed with bleach.

	Vaginal fluid A	**Vaginal fluid B**	**Vaginal fluid C**	**Vaginal fluid D**	**Vaginal fluid E**
Vaginal fluid A	1				
Vaginal fluid B	1.0000	1			
Vaginal fluid C	1.0000	1.0000	1		
Vaginal fluid D	1.0000	1.0000	1.0000	1	
Vaginal fluid E	1.0000	1.0000	1.0000	1.0000	1

Table 6: Pairwise comparison of DNA methylation at CpG 6 of the ZNF282 tDMR for vaginal fluid in all conditions. The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below. A= dry at room temperature, B= wet in an exsiccator, C= outside on the ground, D= sprayed with alcohol and E= sprayed with bleach.

	Vaginal fluid A	**Vaginal fluid B**	**Vaginal fluid C**	**Vaginal fluid D**	**Vaginal fluid E**
Vaginal fluid A	1				
Vaginal fluid B	0.3333	1			
Vaginal fluid C	1.0000	1.0000	1		
Vaginal fluid D	0.2857	1.0000	1.0000	1	
Vaginal fluid E	0.2857	1.0000	1.0000	1.0000	1

Table 7: Pairwise comparison of DNA methylation at CpG 7 of the ZNF282 tDMR for vaginal fluid in all conditions. The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below. A= dry at room temperature, B= wet in an exsiccator, C= outside on the ground, D= sprayed with alcohol and E= sprayed with bleach.

	Vaginal fluid A	**Vaginal fluid B**	**Vaginal fluid C**	**Vaginal fluid D**	**Vaginal fluid E**
Vaginal fluid A	1				
Vaginal fluid B	1.0000	1			
Vaginal fluid C	1.0000	1.0000	1		
Vaginal fluid D	1.0000	1.0000	1.0000	1	
Vaginal fluid E	1.0000	1.0000	1.0000	1.0000	**1**

Table 8: Pairwise comparison of DNA methylation at CpG 8 of the ZNF282 tDMR for vaginal fluid in all conditions. The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below. A= dry at room temperature, B= wet in an exsiccator, C= outside on the ground, D= sprayed with alcohol and E= sprayed with bleach.

	Vaginal fluid A	**Vaginal fluid B**	**Vaginal fluid C**	**Vaginal fluid D**	**Vaginal fluid E**
Vaginal fluid A	1				
Vaginal fluid B	1.0000	1			
Vaginal fluid C	1.0000	1.0000	1		
Vaginal fluid D	1.0000	1.0000	1.0000	1	
Vaginal fluid E	1.0000	1.0000	1.0000	1.0000	1

Table 9: Pairwise comparison of DNA methylation at CpG 1 of the ZNF282 tDMR for all conditions in vaginal fluid for comparison between t0 days and t50 days. The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below. A= dry at room temperature, B= wet in an exsiccator, C= outside on the ground, D= sprayed with alcohol and E= sprayed with bleach.

	t0 days				
	Vaginal fluid A	**Vaginal fluid B**	**Vaginal fluid C**	**Vaginal fluid D**	**Vaginal fluid E**
t50 days	1.0000	1.0000	1.0000	1.0000	1.0000

Table 10: Pairwise comparison of DNA methylation at CpG 2 of the ZNF282 tDMR for all conditions in vaginal fluid for comparison between t0 days and t50 days. The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below. A= dry at room temperature, B= wet in an exsiccator, C= outside on the ground, D= sprayed with alcohol and E= sprayed with bleach.

	t0 days				
	Vaginal fluid A	**Vaginal fluid B**	**Vaginal fluid C**	**Vaginal fluid D**	**Vaginal fluid E**
t50 days	0.4000	1.0000	1.0000	1.0000	1.0000

Table 11: Pairwise comparison of DNA methylation at CpG 3 of the ZNF282 tDMR for all conditions in vaginal fluid for comparison between t0 days and t50 days. The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below. A= dry at room temperature, B= wet in an exsiccator, C= outside on the ground, D= sprayed with alcohol and E= sprayed with bleach.

	t0 days				
	Vaginal fluid A	**Vaginal fluid B**	**Vaginal fluid C**	**Vaginal fluid D**	**Vaginal fluid E**
t50 days	1.0000	1.0000	1.0000	1.0000	1.0000

Table 12: Pairwise comparison of DNA methylation at CpG 4 of the ZNF282 tDMR for all conditions in vaginal fluid for comparison between t0 days and t50 days. The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below. A= dry at room temperature, B= wet in an exsiccator, C= outside on the ground, D= sprayed with alcohol and E= sprayed with bleach.

	t0 days				
	Vaginal fluid A	**Vaginal fluid B**	**Vaginal fluid C**	**Vaginal fluid D**	**Vaginal fluid E**
t50 days	1.0000	1.0000	1.0000	1.0000	1.0000

Table 13: Pairwise comparison of DNA methylation at CpG 5 of the ZNF282 tDMR for all conditions in vaginal fluid for comparison between t0 days and t50 days. The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below. A= dry at room temperature, B= wet in an exsiccator, C= outside on the ground, D= sprayed with alcohol and E= sprayed with bleach.

	t0 days				
	Vaginal fluid A	**Vaginal fluid B**	**Vaginal fluid C**	**Vaginal fluid D**	**Vaginal fluid E**
t50 days	1.0000	1.0000	1.0000	1.0000	1.0000

Table 14: Pairwise comparison of DNA methylation at CpG 6 of the ZNF282 tDMR for all conditions in vaginal fluid for comparison between t0 days and t50 days. The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below. A= dry at room temperature, B= wet in an exsiccator, C= outside on the ground, D= sprayed with alcohol and E= sprayed with bleach.

	t0 days				
	Vaginal fluid A	**Vaginal fluid B**	**Vaginal fluid C**	**Vaginal fluid D**	**Vaginal fluid E**
t50 days	1.0000	1.0000	1.0000	1.0000	1.0000

Table 15: Pairwise comparison of DNA methylation at CpG 7 of the ZNF282 tDMR for all conditions in vaginal fluid for comparison between t0 days and t50 days. The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below. A= dry at room temperature, B= wet in an exsiccator, C= outside on the ground, D= sprayed with alcohol and E= sprayed with bleach.

	t0 days				
	Vaginal fluid A	**Vaginal fluid B**	**Vaginal fluid C**	**Vaginal fluid D**	**Vaginal fluid E**
t50 days	1.0000	1.0000	1.0000	1.0000	1.0000

Table 16: Pairwise comparison of DNA methylation at CpG 8 of the ZNF282 tDMR for all conditions in vaginal fluid for comparison between t0 days and t50 days. The *p*-values obtained from the Chi-square test or Fisher's exact test are indicated below. A= dry at room temperature, B= wet in an exsiccator, C= outside on the ground, D= sprayed with alcohol and E= sprayed with bleach.

	t0 days				
	Vaginal fluid A	**Vaginal fluid B**	**Vaginal fluid C**	**Vaginal fluid D**	**Vaginal fluid E**
t50 days	1.0000	1.0000	1.0000	1.000	1.0000