

**HUMAN PAPILLOMAVIRUS**  
**INDUCES EPIGENETIC**  
**REPROGRAMMING OF HUMAN**  
**KERATINOCYTES**

by

AFAF A H DIYAF

A thesis submitted to  
The University of Birmingham for the degree of  
DOCTOR OF PHILOSOPHY

Institute of Clinical Sciences  
College of Medical and Dental Sciences  
The University of Birmingham  
September 2020

UNIVERSITY OF  
BIRMINGHAM

**University of Birmingham Research Archive**

**e-theses repository**

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.



# ABSTRACT

In this thesis, I investigate the contribution of human papillomavirus (HPV) induced epigenetic changes to the process of cervical carcinogenesis. I used an *in vitro* progression model of cervical cancer involving primary human foreskin keratinocytes in conjunction with human cervical intraepithelial neoplasia cell line and different cervical cancer cell lines. The Primary Human Foreskin Keratinocyte from one donor transfected with episomal HPV16 and HPV18. Following propagation of PHFK HPV18 and PHFK HPV 16 cell lines, I have shown HPV-induced cytopathic changes in human keratinocytes using immunocytochemistry. I found that HPV-associated with changes in the expression of key epigenetic regulators that follow the HPV life cycle and the process of cervical carcinogenesis. Using PCR methods I showed that specific regions of HPV16 and HPV18 genome are prone to disruption during HPV integration. Overexpression of HPV E7 oncoprotein precedes the disruption of full length HPV E2 and induces expression and enzymatic activity of DNMT1 and DNMT3B. Using co-immunoprecipitation, I found that the endogenous HPV E7 oncoprotein interacts at physiological level with DNA methyltransferase 3B. I have also shown, using chromatin immunoprecipitation, that HPV associated with pattern of epigenetic switching from histone 3 lysine 27 trimethylation to DNA methylation. Using a genome wide methylation array, I found that HPV18 integration induces changes to the human DNA methylome similar to the changes expected in cervical neoplasia. Finally, I show that methylated tumour suppressor genes following HPV integration are reversibly demethylated with 5-azacitidine.

# DEDICATION

*To my beloved*

*Libya*

# ACKNOWLEDGEMENTS

I would like to thank my supervisor Professor Ciaran Woodman. His memory will be always with us. I would like to thank, Murray and Woodman group, past and present, both by supporting me in the hard times and encourage me to make this thesis possible. I also want to thank Dr Wenbin Wei and Dr Sally Robert and her group past and present. I would also thank Dr Laura O'Neill and Dr John Halsall who supported me in writing, and encouraged me to strive towards my goal. Words cannot express how grateful I am to my family. Very big thanks to my daughters Shahad, Shada and Shgf for being understanding. Most importantly, special thanks to the little man, Abdulmuhamin without his smile I could not write-up my thesis.

I am most grateful to Libya's Minister of Higher Education and Scientific Research who generously funded my project.

# Contents

<b>1</b>	<b>Introduction .....</b>	<b>2</b>
<b>1.1</b>	<b>Cervical neoplasia .....</b>	<b>2</b>
1.1.1	Histology of cervix .....	2
1.1.2	Risk factors of cervical neoplasia .....	5
1.1.3	Pre-invasive lesions of cervical cancer .....	5
1.1.4	Cervical screening programme .....	7
1.1.5	Cervical cancer .....	14
<b>1.2</b>	<b>Human papillomavirus .....</b>	<b>15</b>
1.2.1	General characteristics and classification of HPV .....	16
1.2.2	Biology and structure of HPV .....	19
1.2.3	Human papillomavirus life cycle .....	22
1.2.4	Integration of Human papillomavirus into human genome .....	28
1.2.5	HPV induced epigenetic changes .....	37
<b>1.3</b>	<b>Epigenetics .....</b>	<b>39</b>
1.3.1	DNA methylation .....	39
1.3.2	Histone methylation .....	51
1.3.3	Epigenetic reprogramming .....	54
1.3.4	Methylation markers of cervical carcinogenesis .....	57
1.3.5	Epigenetic therapy .....	60
<b>1.4</b>	<b>Aim .....</b>	<b>61</b>

---

<b>2</b>	<b>Materials and Methods .....</b>	<b>63</b>
<b>2.1</b>	<b>Cell culture .....</b>	<b>63</b>
2.1.1	Growing cervical cancer cell lines.....	63
2.1.2	Growth media and supplements .....	63
2.1.3	Cultivation of cervical cancer cell lines .....	63
2.1.4	Growing primary cell lines .....	67
2.1.5	Establishing PHFK cell lines.....	67
2.1.6	Cultivation of W12 and the PHFK following transfection with HPV .....	71
2.1.7	Investigating morphological changes of PHFK transfected with HPV18.....	71
2.1.8	Harvesting cells for RNA, DNA and/or protein .....	71
2.1.9	Cell counting.....	72
2.1.10	Cryopreservation .....	72
2.1.11	Recovery of frozen cells.....	72
2.1.12	Preparation of mycoplasma testing.....	73
2.1.13	Cell viability (cytotoxicity) assay .....	73
2.1.14	Demethylation treatment of cultured cells .....	74
2.1.15	U2OS Tet-On inducible expression of HPV16E7 .....	74
<b>2.2</b>	<b>DNA isolation and analysis.....</b>	<b>77</b>
2.2.1	DNA purification .....	77
2.2.2	Polymerase Chain Reaction (PCR) .....	78
2.2.3	DNMT3B gene copy number .....	81
2.2.4	Methylation analysis.....	82
2.2.5	Genome wide methylation analysis .....	89
<b>2.3</b>	<b>RNA detection and analysis .....</b>	<b>93</b>
2.3.1	RNA extraction.....	93

2.3.2	Assessment of RNA concentration .....	93
2.3.3	Assessment of RNA quality .....	93
2.3.4	Complimentary DNA synthesis (cDNA) .....	94
<b>2.4</b>	<b>Protein detection and analysis .....</b>	<b>106</b>
2.4.1	Protein extraction .....	106
2.4.2	Measuring protein concentration .....	106
2.4.3	Western blotting analysis .....	106
2.4.4	Cross-linked chromatin immunoprecipitation (X-ChIP) .....	111
2.4.5	Co-immunoprecipitation (Co-IP) .....	117
2.4.6	Immunocytochemistry .....	118
2.4.7	Direct measurement of DNA methyltransferase enzymatic activity .....	119
<b>3</b>	<b>Results I .....</b>	<b>109</b>
3.1	Introduction of Results I .....	109
3.2	<i>In vitro</i> experimental model system .....	110
3.3	Expression changes of DNA methyltransferases in <i>In vitro</i> experimental models	
system	.....	112
3.3.1	Expression profile of DNMTs in cervical cancer cell lines .....	113
3.3.2	Change in DNA methyltransferase expression in the W12 cell line .....	115
3.3.3	Expression profile of DNA methyltransferases in a PHFK HPV18 cell line .....	117
3.3.4	Expression profile of DNMT in a PHFK HPV16B cell line .....	123
3.3.5	Expression profile of DNMT in a of PHFK HPV16K cell line .....	128
3.3.6	Cellular localization of DNMT1 and DNMT3B proteins in PHFK HPV16B and PHFK HPV18 cell lines .....	131
3.3.7	Morphological changes of cultured PHFK transfected with HPV .....	134
3.3.8	Cytopathic changes of PHFK HPV16 and PHFK HPV18 cell lines .....	136

.....	139
<b>3.4 Expression changes of H3K27 methyltransferase and H3K27 demethylases in <i>In vitro</i> experimental models system</b>	<b>141</b>
3.4.1 Expression profile of the H3K27 methyltransferase and demethylases in cervical cancer cell lines and cervical cancer cell lines	141
3.4.2 Kinetic analysis of changes in expression of H3K27 methyltransferase and demethylases in a PHFK HPV18 cell line	143
3.4.3 Kinetic analysis of the changes in expression of the H3K27 methyltransferase and demethylases in a PHFK HPV16B cell line	146
3.4.4 Kinetic analysis of changes in the expression of the H3K27 methyltransferase and demethylases in a PHFK HPV16K cell line	146
<b>4 Results II</b>	<b>156</b>
4.1 Introduction of results II	156
4.2 Investigating the physical state of HPVs	157
4.2.1 Integrity assessment of full length HPV18E2 in an <i>in vitro</i> model	157
4.2.2 Integrity assessment of full length HPV16E2 in <i>in vitro</i> models	161
4.3 Kinetic changes of HPV16E7 and HPV18E7 protein expression in <i>in vitro</i> models	166
4.4 Correlation of DNMT3B gene copy numbers with mRNA and protein expression	170
4.5 HPV16E7 oncoprotein mediates the induction of DNMT3B expression	173
4.5.1 Optimization of the U2OS Tet-On inducible system	173
4.5.2 The HPV16E7 oncoprotein induces DNMT3B protein expression in the U2OS Tet-On inducible system	173
4.6 The human papillomavirus oncoprotein E7 targets DNMT3B	176
4.7 The E7 oncoprotein induces DNMT enzymatic activity	178
4.7.1 Kinetic changes of DNMT3B enzymatic activity in <i>in vitro</i> models	178

4.7.2	Kinetic of changes in DNMT1 enzymatic activity in <i>in vitro</i> models .....	180
<b>4.8</b>	<b>Overexpression of the E7 oncoprotein associated with changes in DNA methylation of the TSGs .....</b>	<b>182</b>
<b>5</b>	<b>Results III.....</b>	<b>190</b>
<b>5.1</b>	<b>Introduction of results III .....</b>	<b>190</b>
<b>5.2</b>	<b>Genome-wide DNA methylation profiling of the PHFK cell line following HPV18 integration .....</b>	<b>191</b>
5.2.1	Descriptive analysis of the Illumina HD 450K methylation array of the PHFK cell line following HPV18 integration.....	191
5.2.2	Validation of methylation changes following HPV18 integration predicted by the Illumina HD 450K methylation array .....	196
5.2.3	Analysis of the Illumina HD 450K methylation array reveals significant pathway changes following HPV18 integration .....	203
5.2.4	Illumina HD 450K methylation array analysis of human keratinocytes with integrated HPV18 reveals DNA methylation changes in genes involved in the processes of cervical neoplasia .....	214
5.2.5	<i>De novo</i> methylated genes following HPV18 episomal loss .....	217
<b>5.3</b>	<b>Investigation of possible HPV-induced epigenetic switching .....</b>	<b>221</b>
5.3.1	Chromatin immunoprecipitation (ChIP) assays .....	223
5.3.2	HPV 18 integration induces epigenetic switching of Slit-Robo pathway genes.....	228
5.3.3	HPV 18 integration associated with epigenetic switching of tumour suppressor genes.....	241
<b>5.4</b>	<b>Reversibility of DNA methylation of TSGs .....</b>	<b>251</b>
5.4.1	Cell viability following treatment with a demethylating agent .....	251
5.4.2	The demethylating agent induces down-regulation of epigenetic regulators. ....	253



5.5	Reversibility of DNA methylation of TSGs in the PHFK HPV18 cell line.....	256
5.5.1	Reversibility of DNA methylation of TSGs in a cervical cancer cell line .....	259
7	DISCUSSION.....	262
7.1	Cytopathic changes of PHFK following transfection with HPV .....	262
7.2	Disruption of HPV physical state associated with epigenetic reprogramming of DNA methyltransferases .....	266
7.3	The changes of DNA methylation pattern following HPV18 integration.....	272
7.4	Does integration of HPV18 associated with epigenetic switching? .....	275
7.5	Potential methylation markers for early HPV18 cervical neoplasia.....	278
7.6	The effect of demethylation in the PHFK HPV18 cell line.....	279
8	Appendices.....	284
8.1	Appendix I .....	284
8.2	Appendix II .....	289
8.3	Appendix III .....	297
9	References .....	300

## List of figures

FIGURE 1: THE PHYSIOLOGICAL CHANGES OF SQUAMOCOLUMNAR JUNCTION (SCJ) AND TRANSFORMATION ZONE (TZ) .....	4
FIGURE 2: FLOW CHART SUMMARISING THE INDICATION FOR REFERRAL TO COLPOSCOPY, ACCORDING TO THE UK SCREENING PROGRAMME. ....	12
FIGURE 3: CLASSIFICATION OF HUMAN PAPILLOMAVIRUSES.....	17
FIGURE 4: SHOWED AN ORGANIZATION OF OF HIGH RISK HPV GENOME. ....	20
FIGURE 5: HIGH-RISK HPV LIFE CYCLE IN EPITHELIUM CELLS. HPV INFECT AND REPLICATE IN A FULLY DIFFERENTIATING SQUAMOUS EPITHELIUM.....	25
FIGURE 6: IMPORTANT HPV INTEGRATION EVENTS IN CERVICAL CARCINOMAS. ....	35
FIGURE 7: A SCHEMATIC REPRESENTATION OF THE METHYLOME IN NORMAL AND CANCER CELLS. ....	40
FIGURE 8: THE STRUCTURE OF THE DNA METHYLTRANSFERASE FAMILY (DNMT).....	43
FIGURE 9: SCHEMATIC REPRESENTATION OF THE BIOCHEMICAL PATHWAYS FOR CYTOSINE METHYLATION. ....	43
FIGURE 10: TRANSCRIPTIONAL REGULATION OF DNMT GENE EXPRESSION BY PRB. ....	48
FIGURE 11: SCHEMATIC MODEL OF EPIGENETIC SWITCHING. ....	57
FIGURE 12: MODEL OF <i>DE NOVO</i> METHYLATION IN CANCER.....	58
FIGURE 13: FLOW DIAGRAM OF PHFK TRANSFECTION EXPERIMENT. ....	70
FIGURE 14: METHODS OF GENE-SPECIFIC METHYLATION ANALYSIS. ....	87
FIGURE 15: SCHEMATIC PRESENTATION OF ILLUMINA METHYLATION HD ASSAY. ....	94
FIGURE 16: BIOANALYZER RESULTS DEMONSTRATE THE HIGH INTEGRITY AND QUALITY OF RNA USED FOR Q-PCR EXPERIMENTS. ....	98
FIGURE 17: GENOMIC MAP SHOWING DNA METHYLTRANSFERASE PROBES LOCALIZATION. ....	103
FIGURE 18: GENOMIC MAP SHOWING DNMT 3-LIKE (DNMT3L) PROBE AND GENE LOCATION.....	104
FIGURE 19: GENOMIC MAP SHOWING H3K27ME3 METHYLTRANSFERASE AND DEMETHYLASE GENE AND PROBE LOCATIONS. THE ARROWS SHOW PROBE SITES. ....	106

FIGURE 20: WESTERN BLOT ANALYSIS SHOWING DNMT3B PROTEIN EXPRESSION.....	112
FIGURE 21: ALIGNMENT OF THE DIFFERENT STAGES OF CERVICAL CARCINOGENESIS .....	111
FIGURE 22: <i>IN VITRO</i> MODEL SYSTEM. ....	111
FIGURE 23: Q-PCR RESULTS DEMONSTRATING EXPRESSION OF DNA METHYLTRANSFERASES MRNA LEVEL RELATIVE TO BETA-ACTIN AS A REFERENCE GENE IN HPV-POSITIVE CANCER CELL LINES VERSUS HPV-NEGATIVE C33A CANCER CELL LINE. ....	114
FIGURE 24: Q-PCR RESULTS DEMONSTRATING DIFFERENTIAL EXPRESSION OF DNMT3L MRNA LEVEL RELATIVE TO BETA-ACTIN AS A REFERENCE GENE IN THE CERVICAL CANCER CELL LINES, SiHa, HeLa AND C33A VERSUS CA SKI (CONTROL).....	114
FIGURE 25: Q-PCR RESULTS DEMONSTRATING EXPRESSION OF DNA METHYLTRANSFERASES MRNA LEVEL RELATIVE TO BETA-ACTIN AS A REFERENCE GENE IN W12 P57 VERSUS W12 P12 (CONTROL). ....	116
FIGURE 26: WESTERN BLOT ANALYSIS AND QUANTITATIVE DENSITOMETRY OF DNMT3B PROTEIN EXPRESSION IN EARLY (W12 P11) AND LATE (W12 P57) W12 CELL L. ....	116
FIGURE 27: Q-PCR RESULTS DEMONSTRATING EXPRESSION PROFILE OF DNA METHYLTRANSFERASES MRNA LEVEL RELEVATIVE TO BETA-ACTIN AS A REFERENCE GENE IN PHFK HPV18 CELL LINE FROM PASSAGE 4 TO PASSAGE 14 (P14) VERSUS PHFK (CONTROL).....	120
FIGURE 28: Q-PCR RESULTS DEMONSTRATING EXPRESSION PROFILE OF DNMT3B MRNA LEVEL RELEVATIVE TO BETA-ACTIN AS A REFERENCE GENE IN PHFK HPV18 CELL LINE FROM PASSAGE 4 TO PASSAGE 14 (P14) VERSUS PHFK (CONTROL) USING DNMT3B (2) PROBE WHICH IS SPECIFIC TO ISOFORM 1.....	121
FIGURE 29: WESTERN BLOTTING ANALYSIS AND QUANTITATIVE DENSITOMETRY SHOWING DYNAMIC CHANGES IN DNMT1, DNMT3A AND DNMT3B PROTEIN EXPRESSION IN PHFK AND PHFK HPV18 CELLS FROM PASSAGE 4 TO PASSAGE 14.....	122
FIGURE 30: Q-PCR RESULTS DEMONSTRATING THE EXPRESSION PROFILE OF DNA METHYLTRANSFERASES MRNA LEVEL RELEVATIVE TO BETA-ACTIN AS REFERENCE GENE IN PHFK HPV16B CELL LINE FROM PASSAGE 4 UP TO PASSAGE 10 (P10) VERSUS PHFK (CONTROL). ....	125

FIGURE 31: Q-PCR RESULTS DEMONSTRATING THE EXPRESSION PROFILE OF DNMT3B MRNA (PROBE 2) LEVEL RELEATIVE TO BETA-ACTIN AS A REFERENCE GENE IN PHFK HPV16B VERSUS (CONTROL). .....	126
FIGURE 32: WESTERN BLOT ANALYSIS AND QUANTITATIVE DENSITOMETRY SHOWING CHANGES IN DNMT1, DNMT3A AND DNMT3B PROTEIN EXPRESSION IN PHFK AND A PHFK HPV16B CELL LINE PASSAGE 4 TO PASSAGE 10. ....	127
FIGURE 33: Q-PCR RESULTS DEMONSTRATING THE EXPRESSION PROFILE OF DNA METHYLTRANSFERASES MRNA LEVEL RELATIVE TO BETA-ACTIN AS REFERENCE GENE IN PHFK HPV16K CELL LINE PASSAGE 4 TO PASSAGE 10 (P10) VERSUS PHFK (CONTROL). ....	129
FIGURE 34: WESTERN BLOT ANALYSIS AND QUANTITATIVE DENSITOMETRY SHOWING CHANGES IN DNMT3B PROTEIN EXPRESSION IN PHFK AND PHFK HPV16K CELL LINE FROM PASSAGE 4 TO10, PHFK CALREGULIN WAS USED A LOADING CONTROL. ....	130
FIGURE 35: BOX PLOTS DEMONSTRATE QUANTIFICATION OF THE MEDIAN NUMBER OF CELLS EXPRESSING DNMT1 AND DNMT3B PROTEINS, ANALYSED BY TWO INDEPENDENT OBSERVERS. ....	133
FIGURE 36: PHASE CONTRAST MICROGRAPHS OF MONOLAYER CULTURE WITH GREEN FILTER TO SHOW MORPHOLOGICAL CHANGES OF THE PHFK HPV18 CELL LINE.. ....	135
FIGURE 37: IMMUNOCYTOCHEMISTRY DETECTION OF THE DNMT1 PROTEIN. INCREASED DNMT1 PROTEIN EXPRESSION WAS OBSERVED IN PHFK HPV16B AND PHFK HPV18 CELL LINES COMPARED TO PHFK. A CANCER CELL LINE (SiHA) WAS USED AS POSITIVE CONTROL AND H&E STAINING IS USED IN THE NEGATIVE CONTROL.....	137
FIGURE 38: IMMUNOCYTOCHEMISTRY DETECTION OF THE DNMT3B PROTEIN. INCREASED DNMT3B PROTEIN EXPRESSION WAS OBSERVED IN THE PHFK HPV16B AND PHFK HPV18 CELL LINES	138
FIGURE 39: CYTOPATHIC CHANGES OBSERVED IN PHFK TRANSFECTED WITH HPV16 AND HPV18..	139
FIGURE 41: Q-PCR RESULTS DEMONSTRATING EXPRESSION OF MRNA OF H3K27ME3 METHYLTRANSFERASE EZH2 AND H3K27ME3 DEMETHYLASES KDM6A AND KDM6B RELATIVE	

TO BETA-ACTIN AS A REFERENCE GENE IN DIFFERENT CERVICAL CANCER CELL LINE VERSUS C33A.....	142
FIGURE 42: Q-PCR RESULTS DEMONSTRATING EXPRESSION OF MRNA OF H3K27ME3 METHYLTRANSFERASE EZH2 AND H3K27ME3 DEMETHYLASES KDM6A AND KDM6B RELATIVE TO BETA-ACTIN AS A REFERENCE GENE IN W12 LATE PASSAGES (P57) VERSUS W12 P11 (CONTROL). ....	142
FIGURE 43: KINETIC ANALYSIS OF MRNA EXPRESSION CHANGES OF THE H3K27ME3 METHYLTRANSFERASE EZH2 AND THE DEMETHYLASES KDM6A AND KDM6B RELATIVE TO BETA-ACTIN AS A REFERENCE GENE USING Q-PCR IN PHFK HPV18 CELL LINE FROM PASSAGE 4 TO PASSAGE 14 VERSUS PHFK (CONTROL) .....	144
FIGURE 44: KINETIC ANALYSIS OF PROTEIN OF THE H3K27ME3 METHYLTRANSFERASE, EZH2 IN A PHFK AND HPV18 CELL LINE FROM PASSAGE 4 TO PASSAGE 14, USING WESTERN BLOTTING ANALYSIS AND QUANTITATIVE DENSITOMETRY.....	145
FIGURE 45: KINETIC ANALYSIS OF CHANGES IN MRNA EXPRESSION OF THE H3K27ME3 METHYLTRANSFERASE EZH2 AND DEMETHYLASES KDM6A AND KDM6B RELATIVE TO BETA-ACTIN AS A REFERENCE GENE USING Q-PCR IN PHFK HPV16B CELL LINE FROM PASSAGE 4 TO PASSAGE 10 VERSUS PHFK (CONTROL) .....	147
FIGURE 46: KINETIC ANALYSIS OF CHANGES IN PROTEIN EXPRESSION OF THE H3K27ME3 METHYLTRANSFERASE EZH2 IN A PHFK AND PHFK HPV16B CELL LINE FROM PASSAGE 4 TO PASSAGE 10, USING WESTERN BLOTTING ANALYSIS AND QUANTITATIVE DENSITOMETRY. ....	148
FIGURE 47: KINETIC ANALYSIS OF CHANGES IN MRNA EXPRESSION OF THE H3K27ME3 METHYLTRANSFERASE EZH2 AND THE DEMETHYLASES KDM6A AND KDM6B RELATIVE TO BETA-ACTIN AS REFERENCE GENE USING Q-PCR IN PHFK HPV16K CELL LINE FROM PASSAGE 4 TO PASSAGE 10 VERSUS PHFK (CONTROL). ....	149
FIGURE 48: TOPOGRAPHY OF OVERLAPPING PRIMERS I DESIGNED AND USED TO AMPLIFY THE FULL LENGTH HPV18 E2 GENE. HPV NUCLEOTIDES ARE NUMBERED ACCORDING TO THE WHOLE GENOME SEQUENCE OF HPV18.....	158

FIGURE 49: PCR RESULTS SHOW ASSESSMENT OF HPV18 E2 INTEGRITY.....	159
FIGURE 50: TOPOGRAPHY OF OVERLAPPING PRIMERS OF THE FULL LENGTH OF THE HPV16 E2 GENE. .....	162
FIGURE 51: PCR RESULTS SHOW ASSESSMENT OF HPV16 E2 INTEGRITY IN THE PHFK HPV16B CELL LINE.....	163
FIGURE 52: PCR RESULTS SHOW ASSESSMENT OF HPV16 E2 INTEGRITY IN THE PHFK HPV16K CELL LINE.....	164
FIGURE 53: EXPRESSION PROFILE OF HPV E7 ONCOPROTEIN AND P <sub>RB</sub> USING WESTERN BLOT ANALYSIS AND QUANTITATIVE DENSITOMETRY.....	166
FIGURE 54: Q-PCR ANALYSIS OF DNMT3B GENE COPY NUMBER IN THE <i>IN VITRO</i> MODEL SYSTEM, THE W12 CELL LINE AND CERVICAL CANCER CELL LINES.....	169
FIGURE 55: HPV16 E7-MEDIATED INDUCTION OF DNMT3B.....	172
FIGURE 56: THE HPV ONCOPROTEIN E7 ASSOCIATES WITH DNMT3B IN PHFK <i>IN VITRO</i> MODELS AND CERVICAL CANCER CELL LINES <i>IN VIVO</i> . EXPERIMENT. RESULTS SHOWN ARE TYPICAL OF 3 SIMILAR INDEPENDENT EXPERIMENTS. ....	174
FIGURE 57: RESULTS OF DNMT3B ENZYMATIC ACTIVITY ASSAY. ....	176
FIGURE 58: RESULTS OF A DNMT1 ENZYMATIC ACTIVITY ASSAY.....	178
FIGURE 59: QUANTIFICATION OF THE CpG METHYLATED SITES IN THE PROMOTER REGION OF THE DAPK1 AND VHL GENE BY PYROSEQ. ....	181
FIGURE 60: THE PERCENTAGE OF DIFFERENTIALLY METHYLATED CpG SITES IN THE PHFK HPV18 CELL LINE BEFORE (P7) AND AFTER (P14) HPV INTEGRATION.....	189
FIGURE 61: THE TOTAL NUMBER OF HYPERMETHYLATED AND HYPOMETHYLATED CpG SITES IN THE PHFK HPV18 CELL LINE FOLLOWING HPV INTEGRATION.....	189
FIGURE 62: ILLUMINA METHYLATION ARRAYS PROVIDE GENES AND CpG ISLANDS COVERAGE. ADAPTED FROM ILLUMINA WEBSITE. ....	189
FIGURE 63: AN ILLUMINA 450K DNA METHYLATION PORTRAIT OF THE DISTRIBUTION OF HYPERMETHYLATED CpGs IN THE PHFK HPV18 CELL LINE FOLLOWING HPV INTEGRATION. ....	190

FIGURE 64: AN ILLUMINA 450K DNA METHYLATION PORTRAIT OF THE DISTRIBUTION OF HYPOMETHYLATED CpGs IN THE PHFK HPV18 CELL LINE FOLLOWING HPV INTEGRATION. ..	191
FIGURE 65: CROSS-VALIDATION OF ILLUMINA HD 450K METHYLATION ARRAY AND QUANTITATIVE PYROSEQ WITH TWO STATISTICAL METHODS. ....	196
FIGURE 66: NEUROACTIVE LIGAND RECEPTOR INTERACTION PATHWAY IS ENRICHED WITH GENES SIGNIFICANTLY HYPERMETHYLATED FOLLOWING HPV18 EPISOMAL LOSS. ....	202
FIGURE 67: VENN DIAGRAM FOR GENE THAT ARE <i>DE NOVO</i> METHYLATED AT THEIR PROMOTERS. ....	214
FIGURE 68: CHIP OPTIMIZATION. CHIP ANALYSIS OF THE PHFK HPV18 (PASSAGE 7 AND 14) AND HeLa CELL LINES. ....	221
FIGURE 69: CHIP ANALYSIS SHOWING H3K27ME3, DNMT1 AND DNMT3B PROTEINS BINDING TO TWO CONSECUTIVE REGIONS AT THE ROBO1 PROMOTER.....	226
FIGURE 70: CHIP ANALYSIS SHOWING H3K27ME3, DNMT1 AND DNMT3B BINDING TO THREE CONSECUTIVE REGIONS AT THE ROBO2 PROMOTER.....	228
FIGURE 71: CHIP ANALYSIS SHOWING H3K27ME3, DNMT1 AND DNMT3B BINDING TO THREE CONSECUTIVE REGIONS AT THE ROBO3 PROMOTER.....	230
FIGURE 72: CHIP ANALYSIS SHOWING H3K27ME3, DNMT1 AND DNMT3B BINDING TO TWO CONSECUTIVE REGIONS OF THE SLIT1 PROMOTER. ....	232
FIGURE 73: CHIP ANALYSIS SHOWING H3K27ME3, DNMT1 AND DNMT3B BINDING TO TWO CONSECUTIVE REGIONS OF THE SLIT2 PROMOTER. ....	234
FIGURE 74: CHIP ANALYSIS SHOWING H3K27ME3 BINDING TO REGIONS OF THE SLIT3 PROMOTER. .....	239
FIGURE 76: CHIP ANALYSIS SHOWING DNMT3B BINDING TO REGIONS OF THE CALCA PROMOTER. .....	241
FIGURE 77: CHIP ANALYSIS SHOWING H3K27ME3, DNMT1 AND DNMT3B BINDING TO REGIONS OF THE CDKN2A PROMOTER. ....	243
FIGURE 78: CHIP ANALYSIS SHOWING DNMT1 BINDING TO REGIONS AT THE ESR1 PROMOTER.....	245

FIGURE 79: CELL VIABILITY OF HeLa AND PHFK HPV18, p14, CELL LINES CULTURED AND TREATED WITH TWO DIFFERENT DOSES OF 5-AZA-2'-DEOXYCYTIDINE. ....	250
FIGURE 80: ANALYSIS OF CHANGES IN THE MRNA EXPRESSION OF EPIGENETIC REGULATORS FOLLOWING TREATMENT WITH 5-AZA-2'-DEOXYCYTIDINE (5-AZA). ....	252
FIGURE 81: SUMMARY OF METHYLATION CHANGES OF CpGs AT THE PROMOTERS OF ESR1, VHL AND .....	262



## List of tables

TABLE 1: CERVICAL INTRAEPITHELIAL NEOPLASIA ACCORDING TO BETHESDA AND THE UK CLASSIFICATION.....	8
TABLE 2: KEY FUNCTIONAL DIFFERENCES BETWEEN THE HIGH- AND LOW- RISK TYPES OF HPV ONCOPROTEIN E6/E7. ....	26
TABLE 3: SUMMARY OF MAIN STRENGTHS AND WEAKNESSES OF DIFFERENT METHODS USED TO DETERMINE HPV PHYSICAL STATE.....	32
TABLE 4: EXAMPLES OF INTERACTIONS BETWEEN VIRAL ONCOPROTEINS AND DNA METHYLTRANSFERASES.....	37
TABLE 5: PROPERTIES OF CELL LINES USED IN THIS THESIS.....	66
TABLE 6: SUPPLEMENTS OF E-MEDIA.....	70
TABLE 7: CONTENTS OF 100X COCKTAIL.....	70
TABLE 8: LIST OF PCR PRIMERS SEQUENCES USED TO AMPLIFY E1/E2/E5 REGION OF HPV18.....	81
TABLE 9: RT-PCR THERMAL CYCLE PROGRAMME.....	83
TABLE 10: LIST OF PYROSEQUENCING PRIMERS USED IN THIS THESIS, AND THEIR ANNEALING TEMPERATURES (A.TM). FW, FORWARD. REV, REVERSE. SEQ, SEQUENCING PRIMERS.....	89
TABLE 11: MULTIPLEX RT-PCR REACTIONS VOLUMES. ....	100
TABLE 12: RT-PCR TAQMAN® PRIMER AND PROBE KITS.....	100
TABLE 13: PRIMARY ANTIBODIES USED IN WESTERN BLOTTING .....	110
TABLE 14: LIST OF PRIMERS DESIGNED AT PROMOTER REGIONS OF SLIT-ROBO PATHWAY GENES, USED FOR X-CHIP ANALYSIS. ....	115
TABLE 15: LIST OF PRIMERS DESIGNED AT PROMOTER REGIONS OF TUMOUR SUPPRESSOR GENES, USED FOR X-CHIP ANALYSIS. ....	116
TABLE 16: DESIGNATED REACTION MIXTURES OF DNMTs DIRECT ENZYMATIC ACTIVITY ASSAY ....	122
TABLE 17: GENES SELECTED FOR VALIDATION OF METHYLATION CHANGES PREDICTED BY ILLUMINA METHYLATION ARRAYS FOLLOWING HPV18 EPISOMAL LOSS. ....	195

TABLE 18: SUMMARY OF PYROSEQ RESULTS OF THE TEN SELECTED GENES. IDENTICAL CPGs FROM THE PYROSEQ AND ILLUMINA 450K ARRAY ARE HIGHLIGHTED. ....	195
TABLE 19: DNA METHYLATION OF GENES SELECTED FOR EXTERNAL VALIDATION OF THE ILLUMINA HD 450 K METHYLATION ARRAY IN THE CERVICAL CANCER CELL LINES HeLa, SiHa AND Ca Ski. ....	198
TABLE 20: OVERVIEW OF THE SIGNIFICANT FUNCTIONAL THEMES AND PATHWAYS WITHIN THE LIST OF THE MOST SIGNIFICANT HYPERMETHYLATED GENES FOLLOWING HPV18 INTEGRATION.. ....	201
TABLE 21: OFFICIAL SYMBOL OF HYPERMETHYLATED GENES ENRICHED TO THREE SIGNIFICANT BIOLOGICAL PATHWAYS .....	203
TABLE 22: OVERVIEW OF SIGNIFICANT FUNCTIONAL THEMES AND PATHWAYS WITHIN THE LIST OF THE MOST SIGNIFICANT HYPOMETHYLATED GENES FOLLOWING HPV18 INTEGRATION. ....	205
TABLE 23: THE MOST SIGNIFICANT HYPERMETHYLATED DMRs FOLLOWING HPV18 INTEGRATION.	207
TABLE 24: CLUSTERING OF HYPERMETHYLATED GENES FOLLOWING HPV18 INTEGRATION BY CHROMOSOMAL AND CYTOBAND LOCATION. ....	209
TABLE 25: CLUSTERING OF HYPOMETHYLATED GENES FOLLOWING HPV18 INTEGRATION BY CHROMOSOMAL AND CYTOBAND LOCATION. ....	209
TABLE 26: GENES KNOWN TO BE METHYLATED IN CERVICAL NEOPLASIA AND SIGNIFICANTLY HYPERMETHYLATED FOLLOWING HPV18 INTEGRATION. ....	211
TABLE 27: FREQUENCY WITH WHICH GENES THAT WERE DIFFERENTIALLY HYPERMETHYLATED IN P14 COMPARED TO P7 OF THE PHFK HPV18 CELL LINE ARE ALSO KNOWN TO BE METHYLATED IN CERVICAL NEOPLASIA .....	212
TABLE 28: THE <i>DE NOVO</i> METHYLATED GENES FOLLOWING HPV18 INTEGRATION. GENES ENRICHED WITH H3K27ME3 ARE HIGHLIGHTED. ....	215
TABLE 29: SUMMARY OF PREDICTED CPgs METHYLATION OF THE SLIT-ROBO PATHWAY FOLLOWING HPV18 INTEGRATION IN PHFK HPV18 CELL LINE BY THE ILLUMINA 450K HD METHYLATION ARRAY. ....	223

TABLE 30: SUMMARY OF CHIP EXPERIMENTAL RESULTS FOR P7 AND P14 OF THE PHFK HPV18 CELL LINE AND HE LA CELL LINE. ....	246
TABLE 31: SUMMARY OF PYROSEQ ANALYSIS SHOWS THE AVERAGE DNA METHYLATION CHANGES OF TSGs PHFK HPV18 CELL LINE P14. ....	254

## List of abbreviations

Abbreviations	Description
A.Tm	Annealing temperature
5-Aza-2'-deoxycytidine	5-aza
bp	base pairs
BSA	bovine serum albumin
°C	degrees Celsius
cDNA	complementary-DNA
CIN	cervical intraepithelial neoplasia
CpG	cytosine-phosphate-guanine
CT	cycle threshold
DMEM	Dulbeccco's modified Eagle's medium
DMSO	dimethyl sulphoxide
DMR	differentially methylated region
DNA	deoxyribonucleic acid
DNMTs	DNA methyltransferases
dNTP	deoxynucleotide-tri-phosphate
Doxycycline	dox
DiffScore	differential score
ECL	enhanced chemiluminescence system
ECM	extracellular matrix
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
FBS	foetal bovine serum
FCS	fetal calf serum
FDA	Food and drug administration
g	gram
GAPDH	Glyceraldehydephosphate dehydrogenase
H3	histone protein 3
HCL	hydrochloric acid
HES	human embryonic stem cells
HGCIN	high-grade cervical intraepithelial neoplasia
HPV	human papillomavirus
hrs	hours

HSIL	high- squamous intraepithelial lesion
IgG	immunoglobulin G
IHC	immunohistochemistry
IP	immunoprecipitation
KDa	kilodalton
l	litre
LCR	long control region
LSIL	low-grade squamous intraepithelial lesion
μ	micro
M	molar
min	minute
ml	millilitre
mRNA	messenger-RNA
ND10	nuclear domain 10
ng	nanogram
nm	nanometre
NHS	National Health Service
NIH J2 3T3	mouse fibroblasts cells
ORF	open reading frame
p	passage
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PHFK	Primary foreskin keratinocytes
PMID	PubMed identifier
PyroSeq	pyrosequencing
Rb	Retinoblastoma gene
RNA	ribonucleic acid
rpm	rotation per minutes
RT	room temperature
SAM	S-aenosyl-methionine
SCC	Squamous cell carcinoma

SD	Standard deviation
SDS	Sodium dodecyl sulphate
sec	Second
SFM	Serum free media
TBE	Tris-Borate-EDTA
TBS-T	Tris-buffered saline-Tween 20
TEMED	tetramethylethylenediamine
TNF	tumour necrosis factor
TSG	tumour suppresser gene
U	unit
UCSC	The University of California, Santa Cruz
URR	Upstream regulatory region
UV	Ultraviolet
V	voltage
VLP	virus-like particle
WHO	World Health Organization
X-ChIP	cross-linked chromatin immunoprecipitation

## CHAPTER 1

# INTRODUCTION

# 1 Introduction

Viral infection in humans can produce different consequences, from asymptomatic effects to induction of cancer. Some viruses can persist in their latent forms, causing no apparent changes to the infected cell such as herpes simplex virus. On the other end of the spectrum a member of the same family, EBV and Kaposi's sarcoma viruses, are often associated with cancer development. Several viruses contribute to different cancer types. For instance, human hepatitis virus is present in 80% of hepatocellular carcinomas (HCC), and the Epstein-Barr virus is associated with a third of Hodgkin's lymphoma cases (reviewed by Zur Hausen 2009).

Human papillomavirus (HPV) is present in 99.7% of cervical carcinomas. It has been reported that HPV like other oncoviruses use a diverse set of strategies to induce carcinogenesis (Mirabello et al., 2018), such as interactions with tumour suppressor genes like p53 and Rb (Münger et al., 1992; Elmore et al., 1997), sustaining proliferative signalling, evading growth suppressors, and causing genome instability (Mesri, Feitelson and Munger, 2014). HPVs can also interact with the host epigenetic machinery and drive gene silencing, affect cellular genes expression and thereby change biology of the host cell (McLaughlin-Drubin and Munger, 2013; Badal et al., 2003).

## 1.1 Cervical neoplasia

### 1.1.1 Histology of cervix

The cervical canal is formed from endocervix and ectocervix leading to the external orifice. The ectocervix is lined by the squamous epithelium while the endocervix is lined by the columnar epithelium. The junction between them is called the squamocolumnar junction (Bappa and Yakasai, 2013). The squamocolumnar junction of the cervix is situated in the region



of the external cervical orifice. At puberty, following oestrogen production the endocervical epithelium extends distally onto the ectocervix due to the acidic nature of the vaginal environment this columnar epithelium undergoes squamous metaplasia. The area between the original and newly formed squamocolumnar junction is called the transformation zone (Jacobson et al., 1999; Bappa and Yakasai, 2013). Most of cervical dysplasia originates at the transformation zone (Figure 1). During the active phase of squamous metaplasia, the cervical keratinocytes are more vulnerable to carcinogens (Arends, Buckley and Wells, 1998).

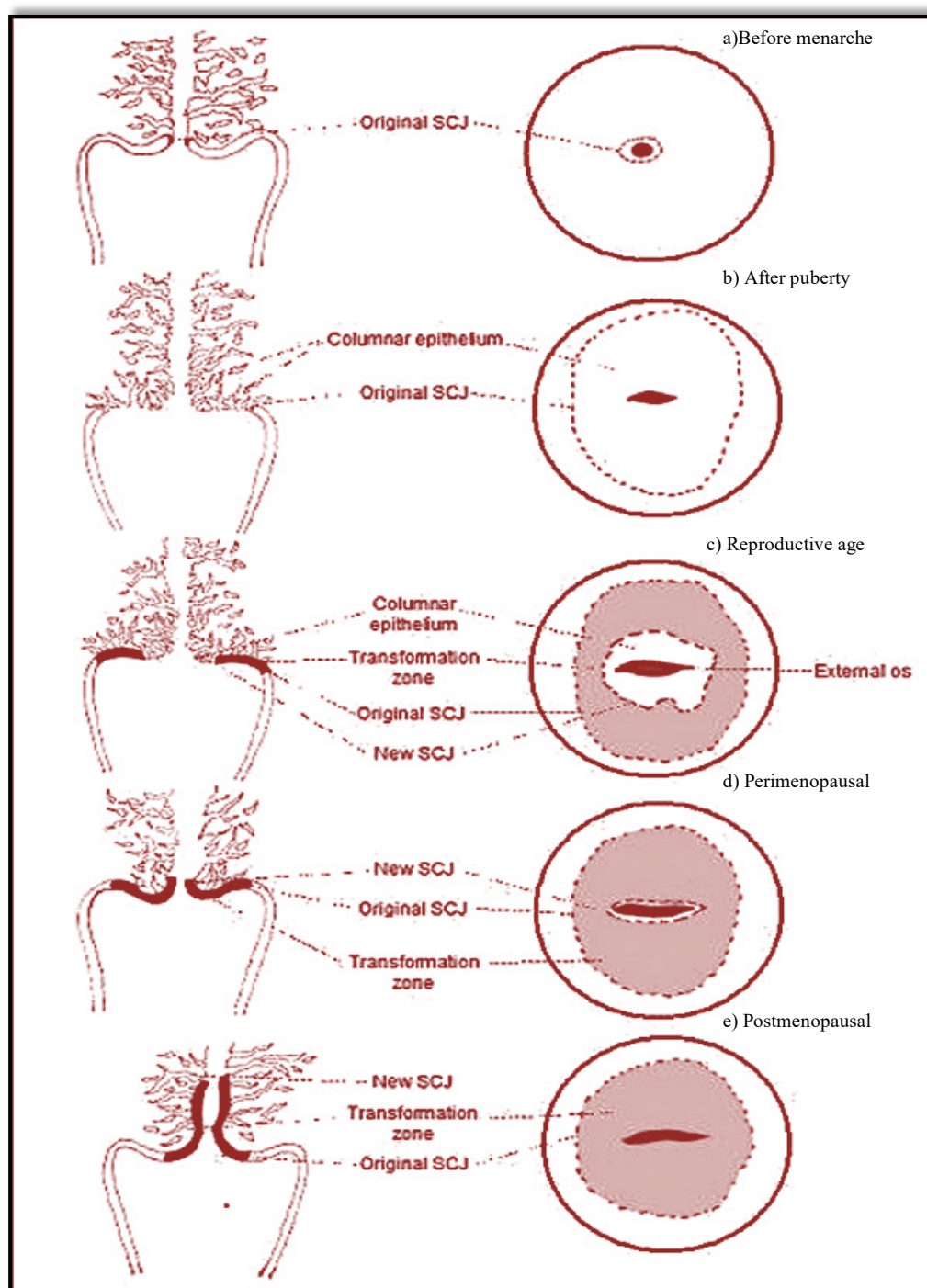


Figure 1: The physiological changes of squamocolumnar junction (SCJ) and transformation zone (TZ) at different age group. Adapted from (Bappa and Yakasai, 2013).

### **1.1.2 Risk factors of cervical neoplasia**

Although several risk factors have been identified, the single most important risk factor of cervical cancer is the persistence of HPV infection. Other risk factors include: early onset of sexual activity, multiple sexual partners, tobacco smoking, oral contraceptives, and low socioeconomic status (reviewed by Kyrgiou and Shafi, 2010). In addition, other sexually transmitted infections might play a role in the disease progression. In particular, immunocompromised women, such as women with human immunodeficiency virus (HIV), have a fivefold increase in risk for cervical cancer development (Rajkumar et al., 2006; Appleby et al., 2009).

### **1.1.3 Pre-invasive lesions of cervical cancer**

The long natural progression of cervical cancer displays a prolonged pre-cancerous phase. Pre-invasive lesions of the cervix are asymptomatic and are usually detected during routine cervical cytological screening. Pre-invasive disease is further subdivided into squamous lesions and glandular lesions reviewed by (Kyrgiou and Shafi, 2010).

#### **1.1.3.1 Squamous lesions**

The squamous lesion represents 80% of cervical neoplasia while adenocarcinomas account for most of the remaining cases (Berrington De González and Green, 2007). It has a well-defined, long pre-cancerous phase that is confined to the cervical epithelium before it goes on to invade through the basement membrane and into the stroma. Therefore, this has been called cervical intraepithelial neoplasia (CIN). However, the World Health Organization (WHO) has replaced the histopathological term of neoplasia with dysplasia. Therefore, Cervical Intra-epithelial Neoplasia (termed CIN I, CIN II and CIN III) has been replaced by the WHO

classification system, where it is now described as mild, moderate and severe dysplasia or carcinoma *in situ*.

#### **1.1.3.1.1 Natural history of cervical intraepithelial neoplasia**

Epidemiological studies have shown that most HPV infections are transient; following HPV infection, half of women will test negative at six months, and 90% at two years. The infection is cleared by the body's immune response. It is expected that dysplastic changes in the cervical epithelium will appear within 2–4 years in women with persistent HPV infection. Progression from one stage to a higher grade takes 3–5 years. It has been estimated that cervical neoplasia requires, on average, 15–20 years from the first detectable cytological abnormality to progress to invasive cancer. However, progression from cervical dysplasia to invasive cervical cancer can be swift (Holowaty et al., 1999).

With the introduction of screening programmes and definitive treatment of cervical pre-cancerous lesions, the natural history of cervical cancer has been largely interrupted, and accurate assessment of the rates of progression or regression has become difficult. In the absence of intervention, it was estimated that 60% of cases of low-grade dysplasia (CIN I) will regress spontaneously, while 10% will progress to high-grade dysplasia. Of these cases, 1% will progress to invasive disease. The remaining one-third of low-grade dysplasia will persist as low-grade lesions (Melnikow et al., 1998). On the other hand, one-third of cases of high-grade dysplasia (CIN II/III) progress to invasive disease, one-third persist, and one-third may still spontaneously regress. The rate of regression depends on many factors. The probability for a CIN lesion to progress to invasive disease increases with the severity of the dysplasia (CIN II/III). Another main factor is the age of the woman. (Woodman, Collins and Young, 2007).

### **1.1.3.2 Glandular lesions**

It is well documented that the epidemiology and incidence of glandular disease is changing. The incidence is becoming higher and more women under the age of 35 years have been affected. Glandular lesions have a more aggressive course than their squamous counterparts. The natural history and biology of squamous lesions have been well described. However, there is significant debate regarding the continuous nature of cervical glandular intraepithelial neoplasia, CGIN. Attempts have been made to mirror the range of cellular changes of CIN to glandular lesions. The proposed classification is cervical glandular intraepithelial neoplasia (CGIN I, II and III) (Gloor and Hurlimann, 1986).

### **1.1.4 Cervical screening programme**

The long pre-cancerous phase and the accessibility of the cervix make cervical cancer ideal for a screening programme. The cervical screening programmes in the UK have led to a reduction in the incidence of cervical cancer, with around a third of cases being diagnosed at a very early stage. Cervical screening saves approximately 5000 lives every year (Kyrgiou and Shafi, 2013). The UK cervical screening programme involves screening women between the ages of 25 to 49 every three years, and then every five years up to age of 64 (NHS cervical screening programme, 2013).

Three disciplines are involved in the final diagnosis and management of cervical dysplasia: the assessment of cytology, colposcopy and histological findings. However, cytological screening and confirmatory histological examinations fail to distinguish between the few abnormal smears which have a significant risk of progressing to invasive cancer, and the vast majority of abnormalities that will regress spontaneously. The overall sensitivity and specificity of cervical screening are 70% and 90%, respectively (Cuzick et al., 2003). The UK screening programme has used the term dyskaryosis to relate cytological findings of

cytoplasmic and nuclear abnormalities. The classification encompasses two main grades of dyskaryosis: low and high grade. Low-grade dyskaryosis provides a prediction for mild dysplasia, and high-grade dyskaryosis provides a prediction for moderate or severe dysplasia (NHS cervical screening programme, 2013). The Bethesda system (Table 1) has been used worldwide to report cervical cytological changes. The two main classifications are, low-grade squamous intra-epithelial lesions, (LSIL) (HPV and CIN I), high-grade SIL, HSIL (CIN II/ CIN III) and squamous cell carcinoma *in situ* (National Cancer Institute Workshop., 1989).

Table 1: Cervical intraepithelial neoplasia according to Bethesda and the UK classification.

Bethesda	Cytology in UK	Histology WHO	CIN
ASC-US/ASC-H	Borderline		
LSIL	Mild dyskaryosis	Mild dysplasia	CIN I
HSIL	Moderate dyskaryosis	Moderate dysplasia	CIN II
HSIL	Severe dyskaryosis	Severe dysplasia/ carcinoma in situ	CIN III

The glandular cytology is classified into four subcategories: Atypical glandular cells (AGC), Atypical glandular cells favouring neoplasticity, endocervical adenocarcinoma *in situ* (AIS) and adenocarcinoma. Any degree of glandular cell abnormality identified on a cervical smear is followed up by colposcopy and treatment (Mody et al., 2011).

#### 1.1.4.1 Cervical exfoliative cytology

Exfoliative cytology remains the mainstay of the cervical screening programme. One of the most frequent findings in cervical cytology from patients with low-grade CIN is HPV-induced cytological changes. In particular, the most common feature is the presence of koilocytes, which are reported in up to 70% of low-grade CIN (Bosch et al., 2008). Koilocytes are squamous cells that were first observed in 1955. They are characterized by enlarged hyperchromatic nuclei surrounded by a clear cytoplasmic zone. They have subsequently been

shown to be a feature of papillomavirus-induced cytopathic changes (Koss, 2012). The term ‘borderline nuclear abnormalities’ (BNA), in cervical cytology describes squamous cells that show both nuclear and cytoplasmic abnormalities that are less than low-grade dyskaryosis and do not show koilocytosis (NHS cervical screening programme, 2013). Of particular interest, the detection of koilocytes and cytomegaly can provide a good assurance that the cervical lesion is CIN I. Spindle cells are more frequent in CIN II and CIN III lesions (Martins et al., 2011).

There are different nuclear and cytoplasmic features of abnormal dyskaryotic cells. Nuclear abnormalities include increased nuclear size, hyperchromasia and abnormal chromatin pattern, whereas changes in cytoplasmic staining and vacuolation are examples of abnormal cytoplasmic features. The grading of squamous dyskaryosis is determined by the assessment of the nuclear to cytoplasmic ratio and nuclear abnormalities. A nuclear/cytoplasmic diameter ratio of < 50% is indicative of low-grade dyskaryosis. In high-grade dyskaryosis, the nucleus is enlarged and occupies between 50-100% of the cell. There may be irregularity in nuclear outline, multinucleation, and hyperchromasia. The dyskaryotic cells may clump together or occur singly (NHS cervical screening programme, 2016).

#### **1.1.4.2 Cervical smear**

Throughout the UK, the cytology performed for cervical screening has moved toward a newer sampling technique, liquid-based cytology (LBC). Liquid-based cytology has largely replaced the conventional Papanicolaou smear. LBC makes the sample much easier to read and allows a uniform spread of epithelial cells in a thin layer and with a clean background. It provides excellent, well-defined nuclear detail (NHS cervical screening programme, 2016). Additionally, the sample obtained by LBC can be used for HPV testing without the requirement of further samples and improves the accuracy and detection rate of abnormal smears (Hutchinson et al., 1999).

#### **1.1.4.3 Implementation of HPV testing in clinical practice**

The HPV test possesses a low positive predictive value in detecting dyskaryosis. Although the TOMBOLA (Trial Of Management of Borderline and Other Low grade Abnormal smears) found no increase in the detection rate of CIN when HPV testing adopted LBC, it showed that almost 50 % of women with low grade abnormal smears tested positive for high-risk HPV strains (TOMBOLA Group, 2009). Thus, in 2011, HPV testing was introduced into the cervical screening programme in England and Scotland, for the triage of women whose cervical smears showed borderline changes or mild dyskaryosis for further investigation by colposcopy (Figure 2).

Treatment of CIN is associated with a significant recurrence rate of approximately 10%. A test for whether the condition has been cured was implemented in England and Scotland in 2012. An HPV test is also performed to follow up with women post-treatment. Different methods have been used for HPV testing. The PCR method relies on detecting a highly conserved region of the L1 gene. The Hologic Cervista HPV method uses PCR, and is available for use in a clinical setting to detect 14 high-risk HPVs types, including HPV16 and HPV18 in cervical smears (Day et al., 2009). The APTIMAR HPV assay detects 14 high-risk HPV types more specifically, through detection of RNA expression of HPV oncogenes E6 and E7. The APTIMAR HPV assay is also available for use in the clinical setting (Dockter et al., 2009).

#### **1.1.4.4 Management of pre-invasive lesions**

Abnormal cervical smears require further investigation using colposcopic examination. The indications for referral to colposcopy are summarized in (Figure 2). Colposcopy is comprised of illumination and low-power magnification (ranging from 8-20x) of the cervix. It is deemed satisfactory when visualization of the entire squamocolumnar junction is accessible.



However, colposcopic examination of the lower genital tract continues to be subjective and prone to intra-observer variation. Moreover, it commonly produces inconclusive findings. In recent reports, only 39% of cases showed accurate histological agreement with colposcopic findings (Massad et al., 2009).

A conservative treatment modality for management of squamous cervical dysplasia involves either ablation or excision methods, and carries up to a 90% cure rate. The goal of treatment is to eradicate the entire transformation zone (TZ). Nowadays, most centres in the UK offer large loop excision of the transformation zone (LLETZ). However, there are other excision methods such as cone biopsy or hysterectomy. The ablation methods destroys the abnormal cells, either with laser therapy, cryotherapy, or cold coagulation (Kyrgiou and Shafi, 2010). Conservative management using excisional treatment methods may also be offered to women with CGIN lesions. However, management of CGIN is challenging; it is associated with minimal colposcopic changes and tends to extend a considerable distance into the endocervical canal. Thus, negative histological margins do not necessarily mean that the lesion has been completely excised. Consequently, careful follow up with endocervical cytology is required, and hysterectomy should also be considered in women who have completed childbearing (Kurian and al-Nafussi, 1999).

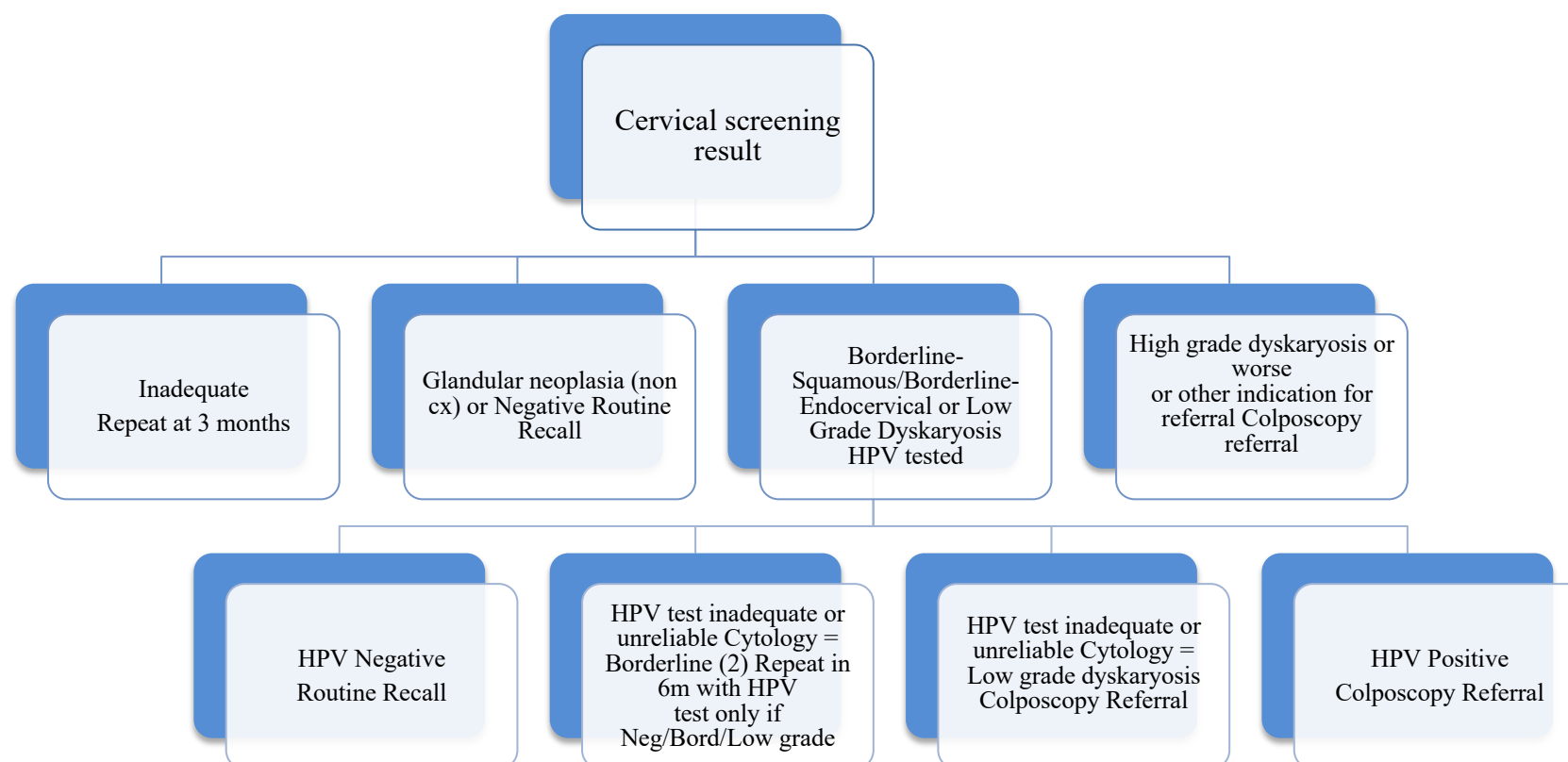


Figure 2: Flow chart summarising the indication for referral to colposcopy, according to the UK screening programme. Adopted from (NHS cervical screening programme 2016).

### **1.1.5 Cervical cancer**

#### **1.1.5.1 Epidemiology of cervical cancer**

Cervical cancer is the second most common female malignancy worldwide, after breast cancer, and seventh overall in cancers. There are an estimated half million newly diagnosed cases each year, leading to over 275,000 deaths worldwide per year.

Over 80% of cases of cervical cancer occur in the developing world. In the UK, approximately 1,000 women still die every year from cervical cancer, meaning that almost three women die each day from the disease. The incidence rate for cervical cancer has two peaks. The incidence is high among women in the age range of 30-40 years of age, and declines thereafter. However, it peaks again in women in their early 80's (CRUK website, 2003).

#### **1.1.5.2 Pathological subtypes of cervical cancer**

The most common pathological subtype of cervical cancer is squamous cell carcinoma, accounting for 70-80% of cervical cancer cases. Adenocarcinoma or adenosquamous carcinoma account for approximately 20%–30% of cases. Adenocarcinoma affects younger women, and has a more aggressive course and poorer prognosis than its squamous counterpart. Although in many developed countries, the incidence of cervical squamous cell carcinoma has been falling, the incidence of cervical adenocarcinoma is rising, despite an effective cervical screening programme. In fact, cytological screening was designed to detect cervical squamous lesions (Kyrgiou and Shafi, 2013).

#### **1.1.5.3 Management and prognosis of cervical cancer**

The management and prognosis of cervical cancer relies on clinical staging using the International Federation of Gynaecology and Obstetrics (FIGO) system. In early stages, the disease is managed either by surgery only or with adjunctive radiotherapy, whereas advanced

disease treatment relies on radiation or chemo-radiation (Kyrgiou and Shafi, 2013). The five-year survival rate in early stages is up to 95%, while that number is dramatically reduced for late stages to reach 5% (CRUK website, March 2014).

## **1.2 Human papillomavirus**

Papillomaviruses are a diverse group of strictly epitheliotrophic, non-enveloped, double-stranded, DNA viruses belonging to the family of papillomaviridae. Papillomaviruses infect more than 20 different mammalian species, as well as, birds, turtles, and snakes. Over 150 human papillomavirus types have been identified along with more than 60 animal papillomaviruses (Bernard et al., 2010). Classification of papillomaviruses is based on the DNA sequence of the L1 gene. The life cycles of the high- and low-risk HPV types have significant differences. In my thesis, I will focus primarily on the high-risk HPV cycle and will be abbreviated it as HPV.

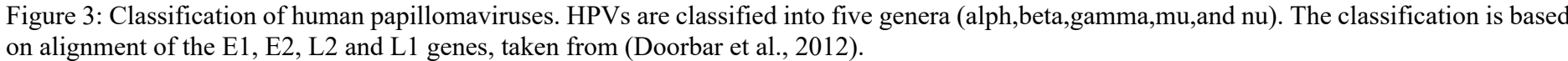
### **1.2.1 General characteristics and classification of HPV**

Human papillomaviruses are divided into five genera (Figure 3). Papillomaviruses are further subdivided into cutaneous and mucosal types. While the mucosal HPV types infect the epithelial linings of the anogenital region, upper digestive system, and upper respiratory tract, the cutaneous HPV types predominantly infect the skin reviewed by (Doorbar et al., 2012). The cutaneous HPV types, including alpha, beta and gamma, are not commonly associated with human cancers. The cutaneous alpha types are categorised as low-risk, primarily causing common warts (HPV2 and HPV57) and flat warts (HPV3 and HPV10). Most of HPV contained within the beta and gamma genera result only in asymptomatic infections (Bernard et al., 2010). However, in individuals with immune defects, these groups are occasionally found to be associated with human cancers. For example, HPV Beta types have been implicated in the development of non-melanoma skin cancers in patients with epidermodysplasia verruciformis, and in immunosuppressed patients. Beta types have also been associated with the appearance

of Bowen's disease and actinic keratosis, conditions arising from repeated excessive sun exposure, and occasionally progressing to squamous cell carcinoma.

A mucosal alpha genus is divided into a low-risk and a high-risk group. Low-risk alpha types cause benign genital warts that affect 1% of the sexually active population. The most common of these are HPV6 and HPV11. However, mucosal alpha types are also implicated in the development of rare, juvenile recurrent respiratory papillomatosis (JRRP), a serious condition affecting children born to condylomatous mothers. This condition progresses to cancer in approximately 5% of affected children (Shah et al., 1998).

To date, 12 of the high-risk, cancer-causing HPV types have been defined. They are: 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58 and 59 (Doorbar et al., 2015). High-risk HPV has been found in approximately 99.7% of cervical squamous cell carcinomas and 94-100% of cervical adenocarcinomas (Walboomers et al., 1999; Bosch et al., 2002). Although cervical infections with high-risk HPV are common in sexually active women, cervical cancer is a relatively rare disease. Approximately 70% of cervical cancers are associated with HPV types 16 and 18 (Walboomers et al., 1999). High risk HPVs are also associated with other anogenital cancers, including 70-100% of anal (Daling et al., 2004) and vulvar (van de Nieuwenhof et al., 2009) carcinomas, over 50% of penile (Heideman et al., 2007) and 60% of vaginal carcinomas (Daling et al., 2002). They are also implicated in non-genital cancers, such as in 15-35% of head and neck cancers (Gillison et al., 2000). Recent researches have also suggested HPV as a causative agent of breast (Lawson et al., 2009) and bladder (Jimenez-pacheco et al., 2012) cancers.



### 1.2.2 Biology and structure of HPV

Human papillomaviruses are small non-enveloped DNA viruses with a virion size of ~55 nm in diameter. They contain a double-stranded, circular DNA genome that is approximately 8 kilo base pairs in size. The HPV DNA genome is divided into an early region, late region, and a non-coding region (long control region LCR). The non-coding region contains two polyadenylation sites separating the early ( $A_E$ ) and late ( $A_L$ ) sites (Figure 4). The early region covers encodes six common open reading frames (E1, E2, E4, E5, E6 and E7) that translate into early proteins. The late region HPV genome represents approximately 40% of the virus genome, reviewed by (Fehrmann and Laimins, 2003; Doorbar et al., 2012). It lies downstream of the early region and encodes L1 and L2 ORFs that translate into two capsid proteins: a major (L1) and a minor (L2) protein (Danos, Katinka and Yaniv, 1982). The LCR region of HPV covers almost 10% of the HPV genome. It involves a segment of about 850 bp that has no protein-coding function. However, the late region of the HPV genome bears multiple transcription factor binding sites, which play an important role in regulation of RNA polymerase II-initiated transcription from early and late HPV promoters. It is also represents the origin of replication. In the papillomavirus genome, an ORF encoding polypeptide is usually spread across exons from different parts of the genome that combine into a full-length ORF through RNA splicing (Doorbar et al., 2012).

The characterization of new promoters and minor transcripts of high risk HPV has been done using HPV16 transcript mapping. The HPV16 genome contains two major promoters. The early promoter (p97) lies upstream of the E6 ORF. Whilst the late promoter (p670) lies within the E7 ORF region, the p97 and p670 promoters are responsible for early and late gene expression, respectively. The HPV-16 p97 promoter is equivalent to the p105 in HPV-18. The p97 promoter is potent, and is controlled by upstream cis-elements in the LCR. These cis-



elements, including four E2-binding sites (E2-BSs) regulate the transcription of p97 from undifferentiated basal stem cells to differentiated keratinocytes. Alternatively, the p670 promoter activity can be induced only in differentiated keratinocytes reviewed by (Fehrmann and Laimins, 2003; Doorbar et al., 2012).

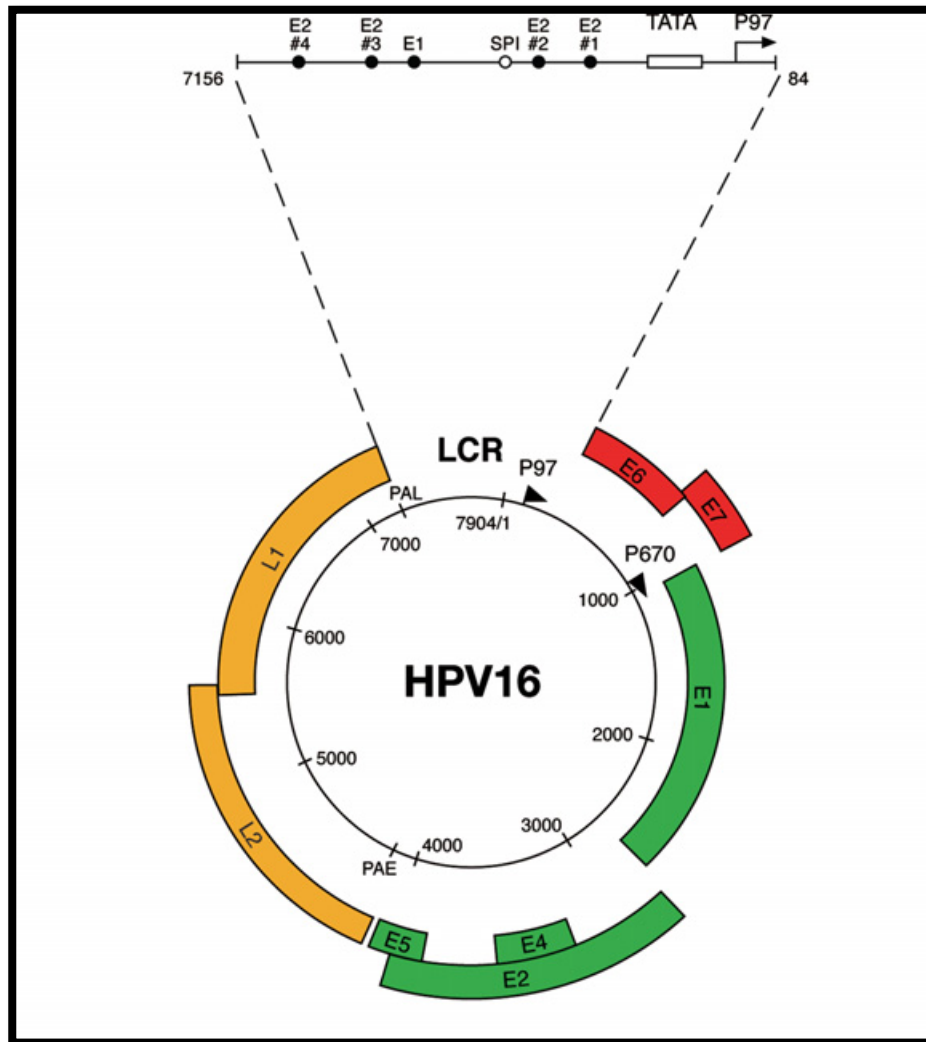


Figure 4: Showed an organization of of high risk HPV genome. The ORFs of the early (E1-E7) and late (L1-L2) genes, the upstream regulatory region (LCR), the two major promoters of HPV16 (P97 and P670) and the two polyadenylation sites (PAE and PAL). Taken from (Doorbar, 2006).

### **1.2.3 Human papillomavirus life cycle**

The HPV productive life cycle is exclusively intraepithelial. It is not associated with viraemia, inflammation, or cell death (Stanley, 2012). The life cycle of HPV is closely linked to the expression of the complete program of keratinocyte differentiation, (Figure 5).

#### **1.2.3.1 HPV attachment and entry to epithelial cells**

Papillomaviruses are the only viruses that initiate their infectious process at an extracellular site. Their internalization is a uniquely slow process requiring several hours, whereas other viruses require only a few minutes from cell binding to viral expression. Initial infection requires access of infectious particles to the basement membrane (BM). It has been suggested that, to be maintained, the virus must infect dividing epithelial stem cells (Organista-Nava et al., 2013). Late protein (L1) primarily mediates the binding to the epithelial attachment factors, heparan sulfate proteoglycans (HSPGs), of the BM. This binding induces a conformational change leading to the exposure of a site on L2, rendering the N-terminus of L2 susceptible to furin cleavage, reviewed by Aksoy et al., (2017). After L2 cleavage, the previously occluded region of L1 binds to an undetermined cell surface receptor on the edge of the epithelial cells that have migrated over the BM to repair the wound (Finnen et al., 2003; Bergant Marušič et al., 2012). Alpha 6 Integrin has been implicated in this process (Evander et al., 1997).

The virus enters the cell via endocytosis of clathrin-coated vesicles. Once HPV enter the cell, it transfers from the early endosome to the late endosome. In Lamp-2-positive late endosomes, L1 undergoes an uncoating process. The C-terminal peptide of L2 has strong membrane-penetrating activity, which facilitates escape from the late endosome. L2 and the genome remain in a complex, which travels through the cytoplasm and into the nucleus via microtubules (Kämper et al., 2006). Cell division is required for expression of the viral genome

in the nucleus (Pyeon et al., 2009). Once nuclear entry has happened, the L2–genome complex co-localizes in distinct, punctate, nuclear domains designated ND10 bodies. Localization at ND10 promotes transcription of the viral genome. Localization of viral genome in ND10 bodies is induced by changes in L2 protein composition (Day et al., 2004).

#### **1.2.3.2 HPV maintenance in the lower epithelial layers**

Initial HPV infection is followed by genome amplification, driven by low levels of early HPV protein expression: E1, E2, E6 and E7 in the basal layer. The E7 protein binds and degrades the members of Retinoblastoma protein (pRb) family, p105, p107 and p130, which control cell cycle entry in the basal layer. In addition, HPV E1 and E2 replication protein is required for HPV replication. However, HPV genome also depends on the host cellular machinery due to its limited coding capacity (Dao et al., 2006). Indeed, HPV only establishes infection on basal stem cells that have the ability to progress through the cell cycle (Pyeon et al., 2009). This process results in an increase in episome number by co-replication with the host genome of basal cells. The virus is maintained in the basal layer at around 200 copies per cell (De Geest et al., 1993). In order to maintain the viral episome as a low-copy number, E1 protein is expressed (Frattoni et al., 1996a) and E2 binds to the p97 promoter and regulates E6 and E7 expression. In addition, it is thought that the other viral early proteins (i.e., E1, E2, E4 and E5) are expressed prior to the onset of genome amplification in order to ensure maintenance of the viral episome at a low-copy number (Middleton et al., 2003).

#### **1.2.3.3 HPV amplification in the upper epithelial layers and role of HPV oncoprotein**

In normal uninfected epithelium, the suprabasal layer's cells exit the cell cycle and undergo a process of terminal differentiation. Both E6 and E7 stimulate cell cycle progression by associating with regulators of the cell cycle, and thus, normal terminal differentiation is

retarded (Sherman et al., 1996). However, there are differences between the functions of E6/E7 in high- and low- risk HPV types, (summarised in table 3).

During the productive infection in the upper epithelium, E7 has a limited ability to stimulate cell proliferation, and only a subset of cells in the parabasal layers are mitotically active. The Retinoblastoma protein (pRb) family prevents S-phase entry by associating with the E2F family of transcription factors. E7 binds to a member of the pRb protein family, p130, in the upper epithelial layers, and mediates its degradation in order to disrupt the E2F/p130 complex (Barrow-Laing, Chen and Roman, 2010). Furthermore, E7 also associates with other proteins that mediate cell proliferation, such as histone deacetylases (Longworth and Laimins, 2004) and the cyclin-dependent kinase inhibitors, p21 and p27 (Funk et al., 1997). In differentiating epithelial cells, the high levels of cyclin-dependent kinase inhibitors (p21<sup>cip1</sup> and p27<sup>kip1</sup>) can control E7 induction of the cell cycle by the formation of inactive complexes with E7 (Noya et al., 2001). Therefore, overexpression of E7 is required to overcome the blockade of S-phase entry mediated by high levels of p21/p27.

The presence of E6 is considered to be a predisposing factor in the development of HPV-associated cancers, as it allows the accumulation of chance errors in host cell DNA to go unchecked. The E6 protein can independently stimulate cell-mediated suprabasal cell proliferation through its C-terminal PDZ-ligand domain. Nevertheless, a site for protein kinase A phosphorylation within the PDZ domain-binding motif can negatively regulate the association of E6 with its PDZ domain-containing substrates (Nicolaidis et al., 2011; Kühne et al., 2000). Another characteristic of the E6 protein is its ability to mediate the degradation of p53 within the cell via its ability to stimulate its ubiquitination and proteasome-dependent degradation (Fu et al., 2010). E6 protein can also associate with other pro-apoptotic proteins,

such as Bak. Indeed, E6 proteins of both HPV16 and HPV 18 reduce Bak-induced apoptosis and stimulate its degradation (Thomas and Banks, 1999; Morgan et al., 2018).

Expression of E5 augments the activities of HPV oncogenes. In suprabasal cells HPV18 E5 is required for unscheduled DNA synthesis (Wasson et al., 2017). E5 expression activates epidermal growth factor receptors resulting in activation of MAP kinases, leading to uncontrolled proliferation (Crusius et al., 1997) and mediated upregulation of Met, a potent oncogene (Scott et al., 2018). Indeed, the ability of E5 to induce cellular hyperproliferation and tumour formation, reinforces it plays a key role in its oncogenic activity (Maufort et al., 2007).

#### **1.2.3.4 HPV genome packaging**

The completion of the HPV life cycle requires the expression of both minor (L2) and major (L1) coat proteins, and the exit of the cell from the cell cycle. HPV E2 recruits the minor coat protein to regions of replication, prior to the expression of major coat proteins and the assembly of the icosohedral capsid in the nucleus (Day et al., 1998). HPV maturation occurs in the most superficial layer of keratinocytes, which is an oxidizing environment due to loss of mitochondrial oxidative phosphorylation. This leads to the accumulation of disulphide bonds between the L1 proteins, and enables production of stable infectious virions (Finnen et al., 2003). Assembled virion particles contain 360 molecules of L1 protein arranged into 72 pentameric capsomeres, while L2 molecules occupy capsomeres at the 5-fold axis of symmetry (Buck et al., 2008). This virion assembly occurs in the nucleus (Cerqueira & Schiller 2017). It is thought that abundant E4 proteins assemble into amyloid fibres that disrupt keratin structure and contribute to virion release (Brown et al., 2006).

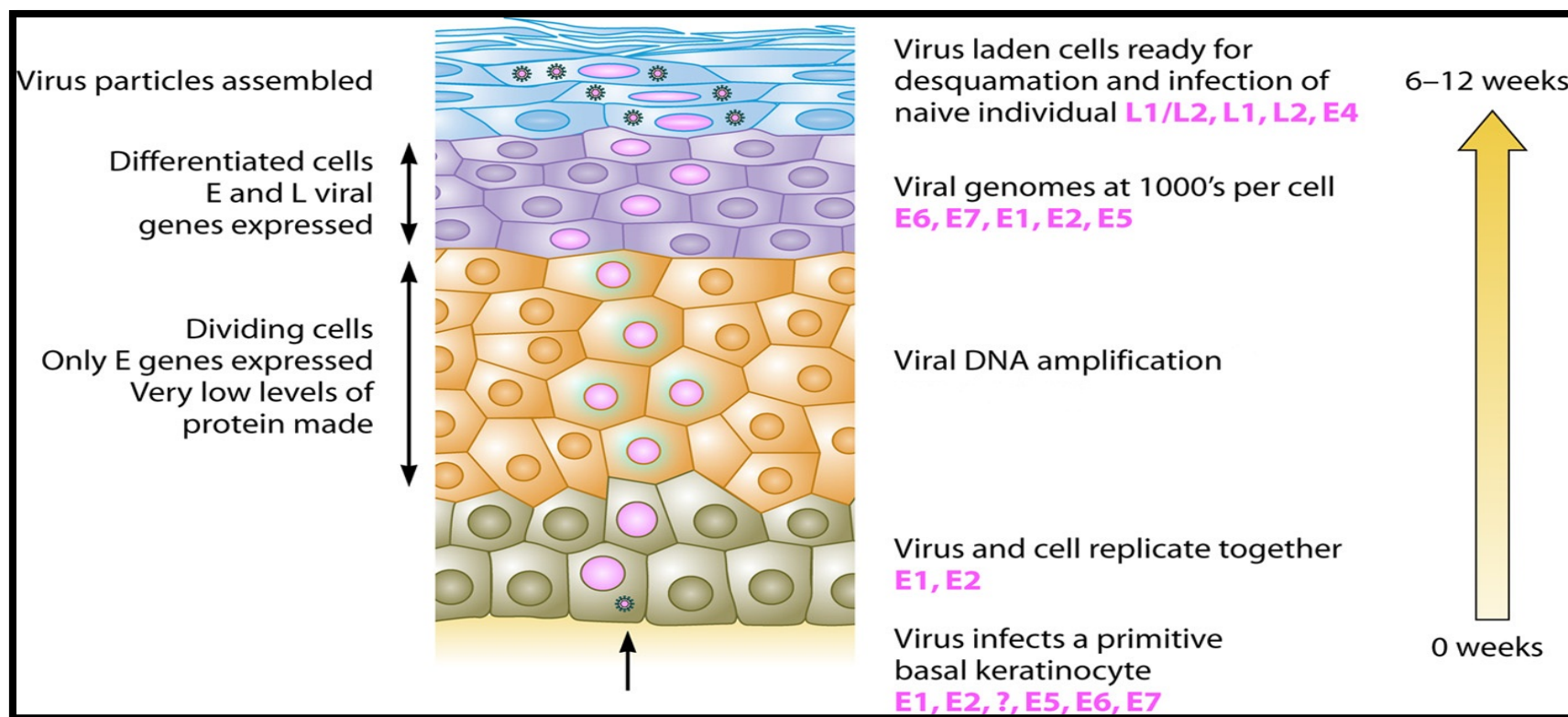


Figure 5: High-risk HPV life cycle in epithelium cells. HPV infect and replicate in a fully differentiating squamous epithelium. HPV infects a keratinocyte in the basal layer of the epithelium, where the virus and cell replicate together, and the viral copy number is maintained at around 50 to 100 copies. In the dividing cells compartment viral gene expression is very tightly controlled during this phase whilst E2 is expressed.. When the host cell begins to differentiate into a mature keratinocyte, HPV activates all of its genes to increase the viral genome copy number to the thousands. In the top epithelium layers, the L1 and L2 proteins are also expressed, thus viral genomes are encapsidated. The time taken from HPV infection to the generation of infectious virus particles is around 6-12 weeks. Adapted from (Stanley, 2012).

Table 2: Key functional differences between the high- and low- risk types of HPV oncoprotein E6/E7. Adapted from (Doorbar et al., 2012).

	<b>High-Risk Alpha</b>	<b>Low-Risk Alpha</b>
<b>E6</b>	Encodes E6 products	Does not encode E6 products
	Binding and degradation of p53 and specific PDZ domain proteins	Weaker binding, no degradation of p53 and no binding of PDZ domain proteins
	Interact with the E6AP ubiquitin ligase and inhibit p53 transactivation and acetylation	
	Inhibition of apoptosis	Unknown
	Bypassing of growth arrest following DNA damage	Normal growth arrest following DNA damage
	Inhibition of keratinocyte differentiation	Unknown
	Inhibition of interferon response	Weaker inhibition of interferon response
	Activation of signalling pathways such as Wnt	Unknown
	Telomerase activation	No activation
	c-myc activation	No activation
<b>E7</b>	Binding and degradation of pRb, p107, and p130	Weaker binding and no degradation of pRb, p107, and E2F
	Binding and no degradation of E2F and HDAC	Binding p130
	Binding of regulatory proteins, induction of cell cycle entry and DNA synthesis role in genome amplification	
	Induction of genome instability	No stimulation of instability
	Suppression of STAT-1 function	No suppression
	Immortalization and transformation function	No such function



#### **1.2.4 Integration of Human papillomavirus into human genome**

It is generally thought that the levels of E6 and E7 expression progressively increase from the early stage of cervical intraepithelial neoplasia toward cervical cancer. This is also thought to facilitate viral integration, which can further deregulate the expression of viral oncogenes, and predisposes the cell to cancer progression (Doorbar et al., 2012). Deregulated viral gene expression may be also driven by changes of oestrogen hormone (Gariglio et al., 2009), or epigenetic modifications such as viral DNA methylation (McLaughlin-Drubin et al. 2011).

The human papillomavirus life cycle is intimately linked to the differentiation state of keratinocytes. HPV requires the epithelial cell to enter the differentiation process in order for the productive life cycle to be completed (Frattini, Lim and Laimins, 1996). As HPV genomes do not encode for the enzymes involved in DNA synthesis, HPV relies on the induction of cellular proteins in differentiating keratinocytes for amplification of viral DNA and late promoter transcription. In the normal viral life cycle, HPV genomes exist in a low episomal copy number in basal keratinocytes. An intact E2 is required for episomal maintenance and plays an essential role in the control of viral replication and transcription (Pirsoo et al., 1996). HPV E2 activates the keratinocyte's differentiation marker, involucrin's promoter in keratinocytes (Hadaschik et al., 2003). The early protein, E2, tightly regulates oncogenes E6 and E7 with high-level expression, and does so only in the suprabasal layer of keratinocytes. Thus, expression of the viral oncogenes poses no carcinogenic threat, as it occurs within a compartment containing cells undergoing constant renewal. Therefore, for cervical neoplastic progression to occur, it possibly requires both quantitative and spatial deregulation of E2 expression, thus enabling a high level of viral oncogene expression throughout the epithelium (Dürst et al., 1992).

Although the HPV genome is retained in the episomal form in low-grade dysplastic lesions, it frequently becomes integrated in host-cell DNA in high-grade dysplastic lesions, as well as in HPV-associated cervical carcinomas, HPV-transformed human keratinocytes, and cervical cancer cell lines (Hopman et al., 2004; Park et al., 1997; Popescu and DiPaolo, 1990; Baker et al., 1987). However, the integrated form has also been detected in low-grade cervical intraepithelial neoplasia, CIN (Saunier et al., 2008). Integration is not a normal part of the productive HPV life cycle and it involves deletion of viral early genes that are essential for synthesis of an infectious virus (Baker et al., 1987). Thus, it represents an abortive infection that confers no apparent advantage to the virus. It is believed that the most common sequence of events in cervical carcinogenesis is the integrant-driven route (Pett and Coleman, 2007a). It has been reported that integration of HPV to cellular DNA occurs in up to 80% of HPV16 positive and 100% of HPV18 positive cervical carcinomas, reviewed by (Woodman et al., 2007).

A small percentage of cervical cancers do develop whilst the HPV remains in episomal form (Vinokurova et al., 2008). In these instances, cell transformation is suggested to be mediated by other mechanisms such as an elevation in the episome copy number which is accompanied by an increase in expression of viral oncogene (Gray et al., 2010). Reactivation of viral oncogenes E6 and E7 could also be mediated by DNA methylation of HPV E2-binding sites (Thain et al., 1996). DNA hypermethylation of HPV16 E2-binding sites has been observed in primary cervical carcinomas, (Bhattacharjee and Sengupta, 2006; Fernandez et al., 2009). In addition, DNA methylation of HPV18 LCR has also been found in several primary cell carcinomas but not HeLa cervical cell lines (Fernandez et al., 2009). It has been demonstrated that treatment with DNA demethylating agents induces recruitment of HPV16 E2 protein and

reduces HPV16 E6 and HPV16 E7 expression. However, the level of HPV18 E6 and HPV18 E7 was not modified by the treatment with DNA demethylating agents (Fernandez et al., 2009).

HPV integration has also been found in other neoplastic lesions. It seems to be related to the progression of a premalignant disease of vulva, Vulvar intraepithelial neoplasia (VIN). Studying the physical status of HPV in vulvar dysplasia showed integrated HPV in less than 40% of lesions with later stage vulvar dysplasia, VINIII (Hillemanns and Wang, 2006). However integrated HPV has also been reported in SCC of vulva (Ueda et al., 2004). In the other hand, most of HPV positive Head and Neck Squamous Cell Carcinoma, HNSCCs have integrated HPV (Hafkamp et al., 2003; Kim et al., 2007).

During integration, the continuity of the viral ring molecule is disrupted, most often within the E1 and/or E2 open reading frame. Therefore, part of E2 and the adjacent open reading frames (E4, E5 and part of L2) are regularly deleted after integration (Baker et al., 1987). These structural changes, which take place after viral integration, and the loss of E2, are responsible for aberrant and increased expression of E6 and E7 oncogenic proteins (McBride & Warburton 2017). Thus, integration facilitates persistent high-level expression of oncogenes. Expression of E6 and E7 accelerates malignant progression by interacting with a number of cellular proteins and by increasing genomic instability of the host cells. These events play a fundamental role in the progression from low-grade to high-grade cervical neoplasia, and are associated with the progression from polyclonal to monoclonal status in CIN (Hudelst et al., 2004).

HPVs integration sites are distributed randomly throughout the host chromosomes. However, there is a clear prediction for integration into a host's common fragile sites (Pett and Coleman, 2007a). An important question is whether integration occurs in common fragile sites as a consequence of their susceptibility to breakage, or because the host sequences provide a selective advantage when adjacent to HPV. The latter could be due to a host effect on viral

transcription and/or insertional mutagenesis caused by HPVs. However, studies suggest that common fragile sites are non-randomly distributed and prone to insertion of foreign DNA vectors at a similar frequency to HPV integrants in cervical cancers (Matzner, Savelyeva and Schwab, 2003; Pett and Coleman, 2007a).

#### **1.2.4.1 Methods of detecting HPV integration in cervical neoplasia**

There are two available strategies used to detect integrated HPVs. The first strategy depends on detecting virus host fusion transcripts. Thus, it detects transcriptionally-active viral integrants, using, for instance, RNA in situ hybridization (ISH) and amplification of HPVs oncogene transcripts (Klaes et al., 1999). The other strategy depends on detecting integrated viral DNA, regardless of its transcriptional status, using quantitative real-time PCR and Southern blotting. Whichever detection method is adopted, integrated HPV that is transcriptionally active is detected in most cervical carcinoma. However, there is a discrepancy in the frequency of viral integration reported in the early period of cervical carcinogenesis. This is possibly due to the differences in methods used. When the first strategy is employed, using the amplification of papillomavirus oncogene transcripts (APOT) technique, transcriptionally active, integrated HPVs are detected in 0% of low-grade squamous intra-epithelial lesions (LSIL), and in 15% of low-grade squamous intra-epithelial lesions (HSIL) (Klaes et al. 1999).

The different methods used to determine the physical state of HPV have several advantages and disadvantages (Table 3). Although, southern blotting is the most reliable method to estimate the overall prevalence of integrated forms, it requires a high concentration and a good quality of DNA that is usually impossible when DNA is obtained from exfoliated cervical cells. In the presence of a low human papillomavirus load, which is more likely the case following viral integration, the southern blotting method might fail to detect HPV integration. In addition, southern blotting is labour-intensive and requires the use of radioactive

material that raises a health and safety issue. Thus, using southern blotting to detect HPV integration is not suitable for use on large series of human specimens. In contrast, the PCR method requires a low concentration of DNA and is less labour-intensive (Collins et al., 2009). However, PCR methods might not distinguish a few HPV integrated forms in the presence of a large number of the episomal form. Failure to amplify full-length HPV E2 using PCR in the presence of the HPV E7 indicates HPV episomal loss. It has been shown that the detection of HPV episomal loss using PCR is comparable to the detection of the integrated form using southern blotting method on DNA obtained from women with CIN3 and invasive cervical cancer (Woodman, Collins and Young, 2007).

Table 3: Summary of main strengths and weaknesses of different methods used to determine HPV physical state.

Method	Advantages	Disadvantages	PMID
Estimation of physical state by real time PCR amplification of E2/E7 or disruption of full length E1/E2 sequences by PCR in the presence of E6/E7 sequences.	Requires low concentration of DNA Less labour-intensive	May not detect integrated form in the presence of high copy numbers of the episomal form	19401452
Southern blot hybridization	Reliably detects integrated from episomal forms	Requires a large concentration of DNA. Might not detect integrated forms in the presence of a low viral load. Uses radio-labeled probes	14973079
Amplification of Oncogene Transcript, APOT	Can obtain host/viral DNA sequence	Labour-intensive Expensive	24586376
Restriction Site PCR, RS-PCR	Can obtain host/viral DNA sequence	Labour-intensive Requires a large concentration of DNA	23824673
Detection of Integrated Papilloma virus, DIPS	Can obtain host/viral DNA sequence	Labour-intensive	11279600

#### 1.2.4.2 Transcriptional regulation of integrated HPV

The existence of the transcriptionally inactive HPV integrated forms in squamous intraepithelial lesions (SIL) indicates that important events are necessary for deregulated transcription to overcome host restraints to viral transcription. Viral gene expression in keratinocytes immortalized with HPV can be inhibited by tumour necrosis factor alpha (TNF $\alpha$ ) and interleukin 1 (Kyo et al., 1994). Indeed, resistance to such repression has been shown to be closely associated with tumorigenicity *in vivo* (Soto et al., 1999). Moreover, HPV integrants remain responsive to the transcriptional inactivation effects of E2 *in vivo*, suggesting regulation of integrated HPV by episomal HPV (Dowhanick, McBride and Howley, 1995). Observations of HPV *in vivo* reveal that the selection of cells containing integrated HPV with high levels of HPV oncogenes was consistently preceded by spontaneous and rapid loss of HPV episomes (Pett et al., 2006). Nevertheless, the concept that a complete loss of E2 expression is required for keratinocytes to be transformed from immortalised into a tumorigenic state is supported by the observation that most cervical carcinomas containing integrated HPV have no episomal form (Das et al., 1992) or methylation of the E2 binding site in the LCR. E2 binding site in LCR (Bhattacharjee and Sengupta, 2006). A study by Tine et al., (2004) suggested a natural selection in favour of cells that containing only one transcriptionally active site, while other integrated copies become transcriptionally inactive. For instance, the transcriptionally active template in Ca Ski cervical cancer cell line is located at a single-copy amongst up to 600 copies of viral DNA. However, in the SiHa cervical cancer cell line, two copies of HPV16 are transcriptionally active (Tine et al., 2004).

### 1.2.4.3 The significance of HPV integration

It is generally thought that the levels of E6 and E7 expression progressively increased, from the early stage of cervical intraepithelial neoplasia toward cervical cancer. This is also thought to facilitate viral integration, which can further deregulate the expression of viral oncogenes, and predisposes the cell to cancer progression (Doorbar et al., 2012). HPV integration into host cell DNA disrupts the E2 ORFs (Daniel et al., 1995), leading to over-expression of the viral oncoproteins E6 and E7. A high level of expression of the E6 and E7 oncogenes in basal epithelial cells inhibits terminal epithelial differentiation and induces high-level chromosomal instability that can drive progression towards the neoplastic phenotype by disruption of a number of key cellular proteins such as p53 and pRb (Münger et al., 2004) (Figure 6). In general, HPV integration into host DNA leads to the stable expression of transcripts encoding oncoproteins E6 and E7 and disruption in the HPV 3' untranslated region that contains an mRNA instability element (Jeon and Lambert, 1995). Critical cellular genes can also be affected by viral integration. Some of the genes disrupted by HPV integration site – myc, APM1, TP63 and TNFAIP2 – are known to be implicated in other cancers (Peter et al., 2006; Ferber et al., 2003). Furthermore, transcriptional regulation of human telomerase reverse transcriptase (hTERT) gene expression is inhibited in episomal infection by E2 (Lee et al., 2002), whereas HPV E6 activates hTERT and telomerase (Oh, Kyo and Laimins, 2001). Thus, HPV integration is a ‘one-shot’ mechanism for development of the malignant phenotype by induction of deregulated proliferation, cellular immortalization, and increased genomic instability (Pett and Coleman, 2007a). Such effects, which are restricted to high-risk HPV, but not low-risk HPV types, provide a biological explanation for the cancer risk associated with high-risk HPV types. However, HPV oncoproteins induce cellular changes by other important mechanism such as epigenetic modification.



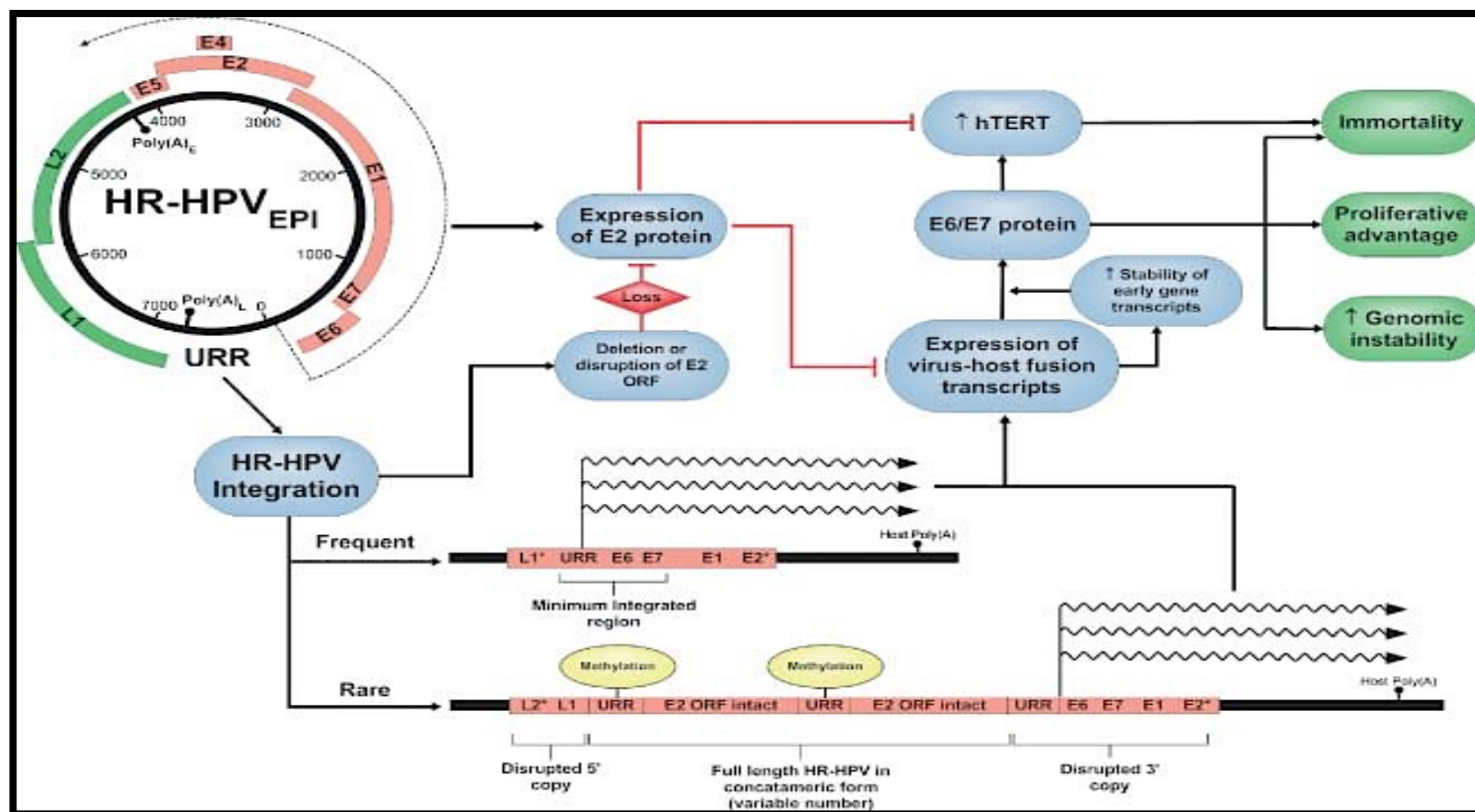


Figure 6: Important HPV integration events in cervical carcinomas. Typical integrants have a disruption of the open reading frame (ORF) for E2, leading to increased levels of E6 and E7 proteins. Taken from (Pett and Coleman, 2007b).

### 1.2.5 HPV induced epigenetic changes

Different oncogenic viruses have been shown to modulate the expression of key epigenetic regulators. Different viral oncoproteins interact with DNA methyltransferases, (summarised in table 4). HPV oncoproteins induce epigenetic changes in epithelial cells, a process which includes targeting of proteins involved in both DNA methylation and histone modification. The direct interaction and induction of the DNMT1 enzyme by the HPV16 E7 oncoprotein provides an explanation for the compelling evidence of DNA methylation changes in cervical carcinogenesis of both human cellular and viral genes (Burgers et al., 2007). In addition, HPV oncoproteins associate with several histone-modifying enzymes. In particular, the E7 oncoprotein interacts with histone deacetylases (HDACs) and histone acetyltransferases (HATs), including p300, pCAF, and SRC1 (McLaughlin-Drubin and Münger, 2009).

In order to establish long-term and persistent infection, and to complete the viral life cycle, HPV targets both epidermal differentiation and wound healing. Infact, it has been shown that PRC complexes regulate these two processes (Ezhkova et al., 2011; Mulder et al., 2012). Hence, PRC2 components are attractive targets for HPV oncoproteins. Exogenous expression of HPV16 E7 induces markedly high expression of the H3K27me3 demethylases, KDM6A and KDM6B. As a consequence, H3K27me3 levels are dramatically decreased. A similar reduction in H3K27me3 modifications was detected in premalignant lesions of cervix, or high-grade squamous intra-epithelial lesions (HSIL). PRC2 complexes control expression of the HOX gene. Indeed, aberrant HOX gene expression observed in E7 expressing cells may provide an explanation for the similar observation in cervical carcinoma cells, and points to the role of the E7 oncoprotein in disturbing HOX gene expression (McLaughlin-Drubin et al., 2011). Dysregulation of homeobox genes has shown to be involved in cervical cancer progression (Zhai et al., 2007).

More recently, high levels of EZH2 expression have been observed in primary human keratinocytes transfected with HPV16 E6/E7, HPV16-high-grade cervical intraepithelial neoplasia (CIN II/III), and HPV16-positive cervical carcinomas (Hyland et al., 2011). The expression of EZH2 was associated with proliferation of HPV-positive cervical cancer cell lines and contributes to their apoptotic resistance. Specifically, it is the E7 oncoprotein that targets and activate EZH2. Overexpression of EZH2 in HPV-positive cancer cells is shown to be required for their cellular proliferation and contributes to their apoptotic resistance (Holland et al., 2008).

Table 4: Examples of interactions between viral oncoproteins and DNA methyltransferases

<b>Virus</b>	<b>Oncoprotein</b>	<b>Interaction with DNMTs</b>	<b>Reference</b>
HPV16	E7	Binds and stimulate DNMT1 activity	(Burgers et al., 2007)
Adenovir	E1A	Binds and stimulate DNMT1 activity	(Burgers et al., 2007)
EBV	LMP1	Activates DNMTs 1, 3A and 3B	(Tsai et al., 2002)
KSHV	LANA	Activates DNMT3A	(Shamay et al., 2006)
HBV	HBx	Activates DNMT1	(Lee et al., 2005)
HBV	C	Activates DNMT1 and 3B	(Chen et al., 2013)

## 1.3 Epigenetics

Epigenetics refers to all mitotically and/or meiotically heritable changes in the pattern of gene expression that are not accompanied by changes in the primary nucleotide sequence (Berger et al., 2009). Epigenetic modifications include DNA methylation, covalent histone modification, and small non-coding RNA (Egger et al., 2004). Disruption of any of these interacting systems can lead to inappropriate expression or silencing of genes, resulting in epigenetic diseases. Aberrant epigenetic changes can occur at any time during tumour development. However, such changes occur mainly during the early stages of the neoplastic process (Fraga et al., 2004).

### 1.3.1 DNA methylation

In humans, DNA methylation involves the addition of a methyl group to the 5th position of cytosine by DNA methyltransferase to form 5-methylcytosine (5-mC). DNA methylation often imposed on cytosine that precedes guanine, these are called CpG dinucleotide. Approximately 1% of bases in a human somatic genome are methyl-cytosines, which represent approximately 80% of all CpG dinucleotides in the genome (Ehrlich et al., 1982). In normal cells CpGs are not randomly distributed in the genome; instead, there are CpG-rich regions known as CpG islands (Figure 7) CpG islands span around 1kb base pairs and display a high frequency of CpG dinucleotide. The CpG dinucleotide in CpG islands is approximately 10 times more frequent than the rest of DNA sequence (Cross et al., 1994). The human genome contains approximately 30.000 CpG islands. While, the CpG islands at gene promoters remain unmethylated, around 9000 CpG islands associated with gene bodies are frequently methylated (Antequera and Bird, 1993).

One of the main features of CpG islands is that they remain devoid of CpG methylation, whereas 70–80% of CpGs is methylated at the bulk of genome. However, CpG islands encompass the transcription start site (TSS) of approximately 60% of human genes. Another feature specific for CpG islands that encompass promoters is that they are marked with trimethylation of lysine regardless the expression status of gene (Mikkelsen et al., 2007; Guenther et al., 2007). Although the concept that CpG islands facilitate promoter function fits well with their presence of CpG islands at TSSs, but it is challenged by the fact that almost half of CpG islands in the human genome encompass either intergenic or intragenic regions and are remote from promoters (Illingworth et al., 2008). These remote CpG islands are associated with peaks of H3K4me3 mark and the de novo methylation during development affects these remote CpG islands mainly, while only few CpG islands at gene promoters are methylated. In contrast, in cancer, de novo methylation affects both CpG islands at promoters and remote CpG islands equally, with a strong preference for those marked by H3K27me3 (Cao et al., 2002; Müller et al., 2002).

DNA methylation at CpG dinucleotides plays a critical role in a diverse range of biological processes, such as the control of gene expression, embryonic development, silencing of transposable elements, X-chromosome inactivation in females, and defence against viral sequence (De Bonis et al., 2006). It is required for genomic imprinting where DNA hypermethylation at one of the two parental alleles of a gene ensure monoallelic expression (Li, Beard and Jaenisch, 1993), and it is also required for chromosomal stability through the hypermethylation of repetitive genomic sequences (Katargin et al., 2009). Furthermore, DNA methylation is implicated in the development of the immune system, brain function and behaviour (Jurkowska, Jurkowski and Jeltsch, 2011). Aberrant DNA methylation is associated with several human diseases, and, in particular, cancers (De Capoa et al., 2003).

The correlation between DNA methylation and gene silencing is well recognized. However, the mechanisms by which DNA methylation repress gene transcription is still not fully understood (Greenberg & Bourc'h 2019). The methylation pattern of CpG dinucleotides in mammalian genomes is not uniformly distributed. Nucleosome-associated DNA has a low level of methylated CpG dinucleotides compared to more accessible DNA sequences which is thought of as a direct repressive mechanism, where binding of some transcription factors (TFs) to their recognition motifs is prevented (Gaston & Fried 1995; Mann et al., 2013). By analysing binding of a large collection of TFs to DNA sequences Yin et al., (2017) showed that DNA methylation prevents transcription activation by reduced affinity of certain TF to their motifs when methylated compared with unmethylated.

There are five mammalian methyl-CpG domain-binding (MBD) methyl-CpG-binding protein 2 (MeCP2), MBD1-MBD4. MeCP2, MBD1, MBD2 and MBD4 (Baubec et al., 2013) proteins are known to have a preference for methylated cytosine. All MBD interact with and recruit histone deacetylase and nucleosome remodelling complexes, leading to heterochromatin formation and transcription repression (Ng et al., 1999). Furthermore, DNMT proteins play a role in heterochromatin formation by interaction with different chromatin modifiers and remodellers. For example, DNMT1 colocalize and directly interacts with the H3K9 histone methyltransferase on heterochromatic regions (Esteve et al., 2006) and *de novo* DNMTs interact with histone deacetylases to silence transcription (Fuks et al., 2001; Deplus et al., 2002).

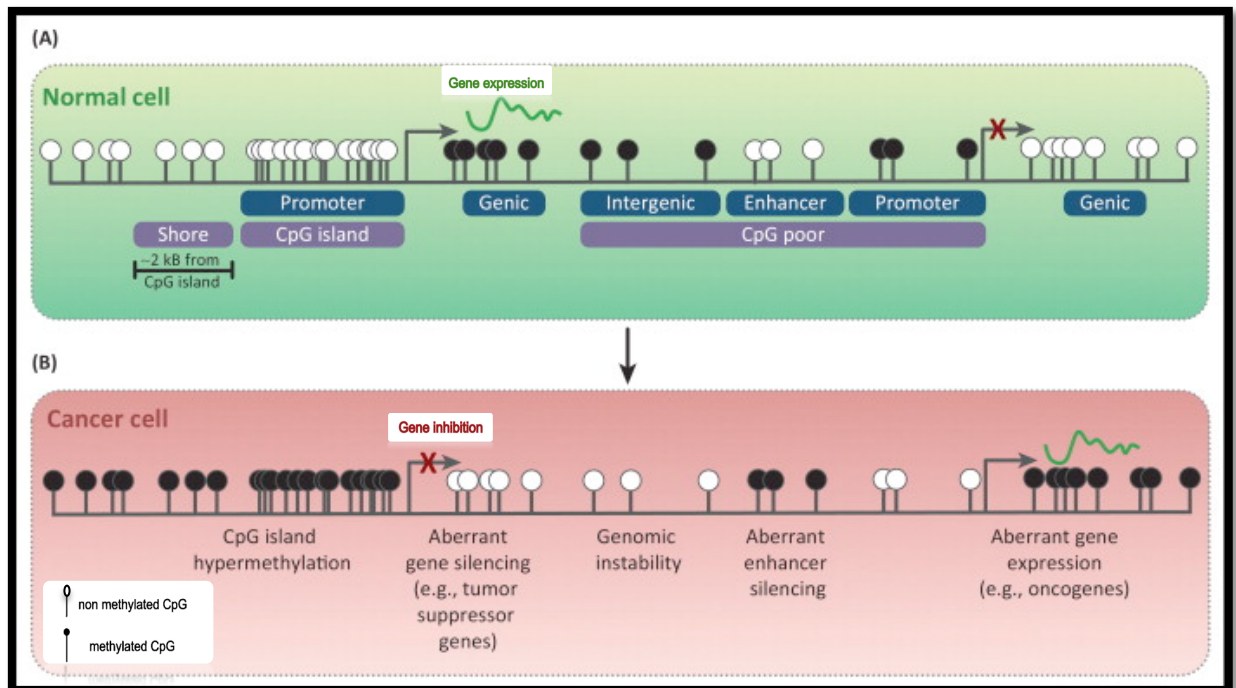


Figure 7: A schematic representation of the methylome in normal and cancer cells. (A) In normal cells, most methylated CpG sites (black circles) are outside of CpG islands such as gene body (genic) methylation, and CpG-poor regions. CpG-poor regions (intergenic & CpG-poor promoters), with the exception of enhancers, are typically methylated in normal cells. Hypomethylated CpG sites (White circles) are located in the promoter regions often in CpG islands. (B) In cancer cells, DNA hypermethylation of CpG islands results in aberrant gene silencing (e.g., of tumor suppressor genes). Concomitant DNA hypomethylation of CpG-poor promoters contributes to genomic instability and aberrant gene expression such as of oncogenes. Taken from (Stirzaker et al., 2014).

### 1.3.1.1 DNA methyltransferases

There are five Human DNMTs: DNMT1, DNMT2, DNMT3A, DNMT3B and DNMT3L. They form two main DNMTs sub-families: The maintenance methyltransferase (DNMT1), which maintain methylation of daughter DNA strands after DNA replication (Hermann, Gowher and Jeltsch, 2004), and *de novo* methyltransferases (DNMT3A, DNMT3B, and DNMT3L), which establish the methylation activity in mammalian cells (Okano et al., 1999). DNMT3A and DNMT3B are highly expressed in undifferentiated human embryonic stem cells (HES), and are expressed at low levels in adult somatic tissues and down-regulated after differentiation (Okano et al., 1998).

### 1.3.1.2 The structure of the DNA methyltransferase family

All DNMTs enzymes consist of an N-terminal regulatory domain and a C-terminal catalytic domain (Figure 8), with the exception of DNMT2 that lack the regulatory domain. The C-terminal section contains the active centre of the enzyme (Jurkowska, Jurkowski and Jeltsch, 2011). The catalytic domain contains six conservative motifs (I, IV, VI, VIII, IX and X), which represent active binding sites to S-adenosylomethionine and cytosine. However, DNMT3L is a catalytically inactive DNMT3 variant that lacks the C-terminal part of the catalytic domain as well as the N-terminal part of a regulatory domain. The N-terminal domain is responsible for localization of DNMTs in the nucleus (Reviewed by Lyko, 2018). The regulatory domains of DNMTs contain several specific subdomains of which the DMAP1 binding domain that mediates interactions of DNMT1 with DMAP1 and with HDAC2. The RFTS conserved subdomain is necessary for targeting of DNMT1 to replication foci and maintaining of DNA methylation. The CXXC subdomain is another conserved domain that responsible for binding of DNMT1 to unmethylated. DNMT1 possess another two BAH subdomain that their functions remain to be elucidated. DNMT3 enzymes also have conserved subdomain that play an



important role for DNMT3 and chromatin interactions such as the PWWP domain and ADD domain. The methylation reaction is initiated by the ‘base flipping’ mechanism that rotates the target base into the catalytic pocket of DNMT enzyme, followed by a nucleophilic attack on the C6 of the cytosine ring. This attack by conserved cysteine residue in the motif IV facilitated by a conserved glutamic acid residue in motif VI (Klimasauskas et al., 1994). DNA methyltransferase enzymes (DNMTs) use S-adenosyl-methionine (SAM) as the methyl donor to add a methyl group to the cytosine ring, forming methyl cytosine and S-adenosyl homocysteine (SAH) (Ramsahoye, Davies and Mills, 1996), shown in (Figure 9).

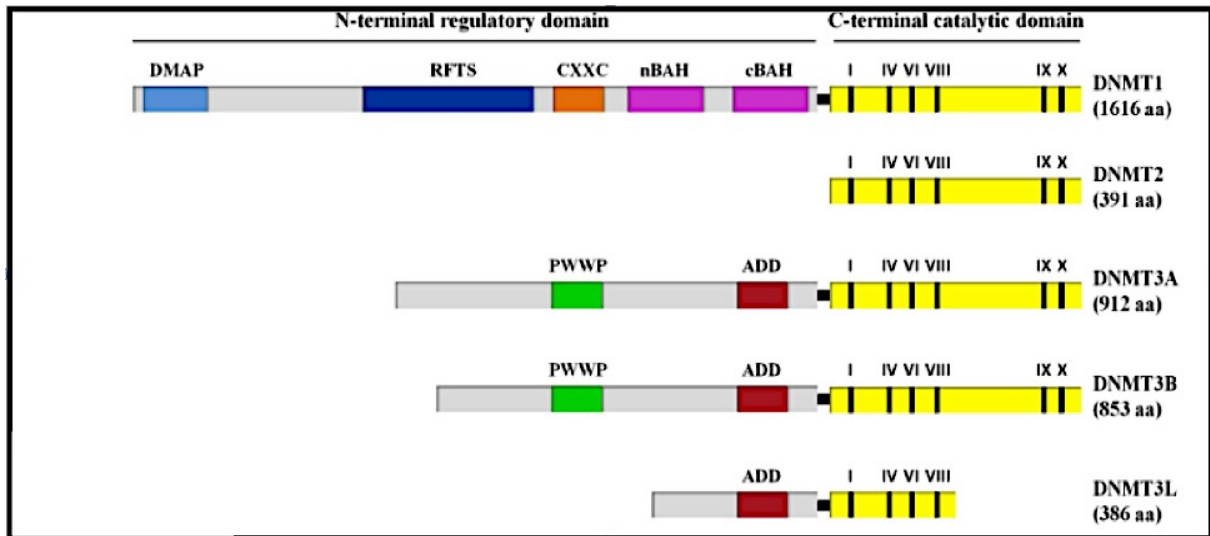


Figure 8: The structure of the DNA methyltransferase family (DNMT). A DNMT enzyme consists of a regulatory (N-terminal) domain and a catalytic (C-terminal) domain. The C-terminal domain contains six conservative motifs (I, IV, VI, VIII, IX and X), which represent active binding sites to S-adenosylomethionine and cytosine. The C-terminal domain of DNMTL lack IX and X motifs. The regulatory domain of DNMT1 enzymes divided into conserved subdomains: DMAP1, RFTS , CXXC and BAH domains. The regulatory domain of DNMT3 enzymes divided into conserved subdomains: PWWP and ADD domain, while DNMT2 lack the regulatory domain, Adapted from (Bestor, 2000).

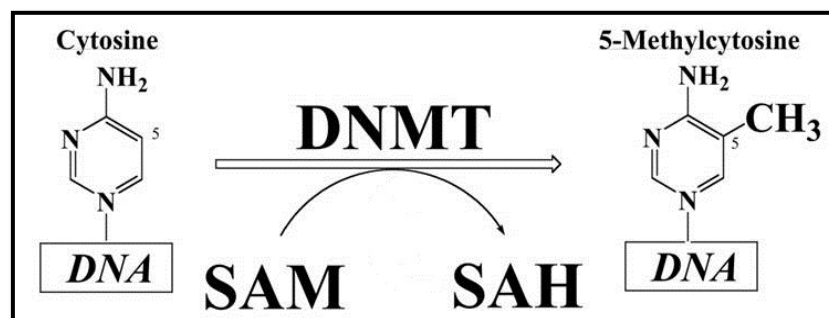


Figure 9: Schematic representation of the biochemical pathways for cytosine methylation. DNA methylation refers to the addition of a methyl group to the cytosine ring, a reaction catalyzed by DNA methyltransferase enzymes using S-adenosyl-methionine (SAM) as the methyl donor. Adapted from (Singal and Ginder, 1999).

#### 1.3.1.2.1 The maintenance methyltransferase

Human maintenance DNA methyltransferase (DNMT1) is a large enzyme, comprised of 1616 amino acids. DNMT1 was the first biochemically characterised DNA methyltransferase enzyme. Disruption of the DNMT1 gene in mice results in embryonic lethality, indicating that DNMT1 plays an essential role in embryonic development (Li, Bestor and Jaenisch, 1992). Interestingly, knockdown or disruption of the DNMT1 gene results in progressive cell death in human cancer cells (Chen et al., 2007) and immediate apoptosis in germ cell lines (Takashima et al., 2009). DNMT1 is highly expressed in proliferating cells, reaching a maximum during S phase of the cell cycle (Lee et al., 1996). However, it is only present at low levels in non-dividing cells (Robertson et al., 1999). DNMT1 has a preference for hemimethylated DNA by a factor of about 30–40-fold over unmethylated DNA (Goyal, Reinhardt and Jeltsch, 2006). This preference is due to an interaction of hemimethylated CpG sites with the active centre of the enzyme, as well as to methylation-specific conformational changes of the enzyme (Jurkowska et al., 2011). In addition to its main function as a maintenance DNA methyltransferase, it also displays *de novo* DNA methylation activity (Jair et al., 2006). DNMT1 has multiple DNA binding sites and interacts with many components of the epigenetic machinery, such as DNMT3A, DNMT3B (Kim et al., 2002), the histone deacetylases, and the retinoblastoma protein pRb (Robertson et al., 2000).

Several studies showed a link between DNMT1 and pRb pathway in human cancer. DNMT1 gene is transcriptionally regulated by Rb pathway. While overexpression of Rb leads to hypomethylation of the cellular through binding with E2F protein to the DNMT1 (Pradhan and Kim, 2002). Low expression of Rb is linked to induction of DNA hypermethylation of CpG islands (McCabe et al., 2005) (Figure 10). In addition, the retinoblastoma protein pRb binds to the regulatory domain of DNMT1 independently of its own phosphorylation status. This strong

binding inhibits DNMT1 methyltransferase activity by disruption of the DNMT1–DNA binary complex (Pradhan and Kim, 2002). Further to the epigenetic function of pRb, it has also been recently shown that pRb regulates DNA methylation by affecting DNMT1 protein stability (Shamma et al., 2013).

#### **1.3.1.2.2 The *de novo* methyltransferase family**

The DNMT3 family is comprised of DNMT3A, DNMT3B and DNMT3L. Whilst DNMT3A and DNMT3B are catalytically active enzymes, DNMT3L is an inactive enzyme (Aapola et al., 2000). Both DNMT3A and DNMT3B function mainly as *de novo* DNA methyltransferases. However, they also play a role in the maintenance of DNA methylation (Kim et al., 2002; Chen et al., 2003). Although, in mammals DNMT3A and DNMT3B predominantly methylate cytosine residues, both enzymes can also methylate non-CG sites such as CA,CT and CC (Gowher and Jeltsch, 2001). However, the biological significance of this remains unknown. Gene-specific non-CG methylation was reported in a study of lymphoma and myeloma cell lines (Malone et al., 2001) and a region of the human myogenic gene. The human myogenic gene is silenced in non-muscle cell lines by carrying several methylated cytosines within a CCTGG sequence (Franchina and Kay, 2000). It also reported that the p16 promoter contains non-CpG methylation in specimens from breast tumour compared to normal breast tissue (Woodcock et al., 1999). However, non-CpG methylation is rarely found in cancer. It has been shown that non-CpG methylation is prevalent in stem cell lines (Ramsahoye et al., 2000). Interestingly, in stem cell studies, non-CpG methylation is not generally present on both strands. It is concentrated more in actively transcribed genes bodies and it is methylation attributed to the DNMT3a (Lister et al., 2009; Ramsahoye et al., 2000). The presence of non-CpG methylation in human stem cell lines and its absence in the human fetal cell line suggests that non-CpG methylation may play a role in the pluripotent lineage (Lister et al., 2009).

Mice studies have revealed that disruption of both DNMT3A and DNMT3B genes is lethal. The DNMT3A knockout animals developed *in utero* but died at term, whereas the DNMT3B knockout animal embryos die *in utero* at day 9.5 with multiple developmental defects (Okano et al., 1999). Genetic disruption of DNMT3B induces aneuploidy of cancer cells and increases the number of novel chromosomal translocations (Karpf and Matsui, 2005). A rare autosomal disease develops as the result of mutations in the DNMT3B gene. Furthermore, mutations in the DNMT3B gene result in hypomethylation at pericentromeric regions (Xu et al., 1999), manifesting clinically as ICF syndrome, whose name is an abbreviation of the disease's main characteristics: Immunodeficiency, Centromere instability, and Facial abnormalities. Despite the sequence homology of DNMT3A and DNMT3B, knockout studies have demonstrated different phenotypes. Conversely, DNMT3L knockout animals have shown to be viable and do not display developmental defects (Okano et al., 1999). However, male mice lacking DNMT3L fail to produce mature sperm. Whereas both DNMT3A and DNMT3B are expressed in undifferentiated human embryonic stem (HES) cells, low expression is observed in differentiated cells. DNMT3L is specifically expressed in germ cells (Okano et al., 1999), and markedly induces *de novo* methylation activity of DNMT3A but not DNMT3B. This suggests that DNMT3L has a role as a general stimulatory factor for DNMT3A (Chedin, Lieber and Hsieh, 2002).

#### **1.3.1.3 Regulation of DNMT activity**

DNMTs are regulated by a variety of mechanisms. The molecular interaction of DNMTs highly regulate their activity for example a heterotetrameric complex is formed when two molecules of DNMT3A interact with two molecules of DNMT3L, this molecular interaction resulting in an increase in the affinity of DNMT3A for DNA and hence more effective DNA methylation (Jia et al., 2007). Post-translational modification is another mechanism that

regulates DNMTs activity. For example, phosphorylation of DNMT1 at serine 143 by the kinase AKT1 can protect DNMT1 from degradation (Estève et al., 2011). DNMTs activity is also regulated by alternative splicing. All catalytically active DNA methyltransferases are subject to alternative splicing. However, DNMT3B is the one having the most alternatively spliced isoforms (Robertson et al., 1999). These isoforms can be tissue-specific, for example, human testis expresses two additional isoforms, DNMT3B4 and DNMT3B5 (Robertson et al., 1999). They also vary in their enzymatic activity. Whilst DNMT3B3 and DNMT3B4 are catalytically inactive isoforms (Van Emburgh and Robertson, 2011), DNMT3B7 is catalytically active and overexpression of DNMT3B7 results in apparent methylation and enhanced tumorigenesis (Shah et al., 2010; Van Emburgh and Robertson, 2011). In addition, DNMTs can be regulated through changes in gene copy number. For example, over-expression of DNMT3B reported in solid tumours was associated with DNMT3B gene amplification (Esteller et al., 2011).

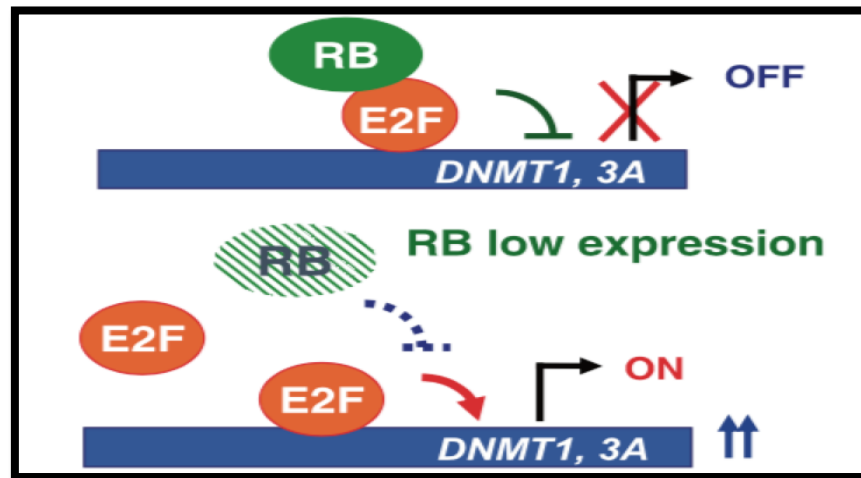


Figure 10: Transcriptional regulation of DNMT gene expression by pRB. RB binds to DNMT1 and DNMT3A promoters through binding to E2F1 protein. Low RB expression is associated with overexpression of DNMT1 and DNMT3A. Adapted from (Lin and Wang, 2014).

### **1.3.1.4 DNA methylation in cancer**

There are two DNA methylation changes that occur in cancer cells compared with normal cells: global demethylation and *de novo* methylation of select CpG islands. Genome-wide methylation technology has considerably expanded our understanding of the cancer methylome. The pattern of human cancer cells typically shows a global loss of intergenic DNA methylation with a simultaneous gain of methylation at localized gene-associated CpG islands. A recent study has revealed that, during carcinogenesis, aberrant promoter DNA methylation of certain biological pathways is shared across several cancers (Kim et al., 2012).

#### **1.3.1.4.1 DNA hypomethylation in cancer**

DNA hypomethylation at CpG dinucleotides was one of the first epigenetic changes discovered in human cancer (Feinberg and Vogelstein, 1983). Global DNA hypomethylation has been associated with several cancers including cervical carcinoma (De Capoa et al., 2003). Loss of DNA methylation in cancer is due primarily to the loss of methylation of repetitive sequences, coding regions, and introns. The degree of hypomethylation of genomic DNA increases progressively with the development of the neoplasm (Fraga et al., 2004). Different mechanisms have been proposed to explain the role of DNA hypomethylation in cancer development such as reactivation of transposable elements and generation of chromosomal instability. The loss of DNA methylation reactivates intragenomic endoparasitic DNA such as Alu (recombinogenic sequence) repeats and L1 (long interspersed nuclear elements) (Fraga et al., 2004). Transcription of these endoparasitic elements can drive translocation and further disrupts the genome, contributing to development of cancer. The other mechanisms that have been proposed to explain the role of DNA hypomethylation in cancer is demethylation of promoter regions which activate repressed genes such as the PAX gene involved in cell proliferation (Wu et al., 2005). Interestingly, progressive genomic hypomethylation of cervical



cancers activates the HPV genome and correlates with carcinogenic progression (Badal et al., 2003).

DNA methylation is generally considered a stable epigenetic event associated with gene repression, which may be alleviated through DNA replication. No direct DNA demethylating enzyme has been identified but this stable epigenetic modification could be modified by the ten eleven translocation protein (TET) (Tahiliani et al. 2009). The ten eleven translocation (TET) enzymes, TET1, TET2, and TET3 catalyze the sequential oxidation of 5-methylcytosine to 5-hydroxymethylcytosine, 5-formylcytosine and 5-carboxylcytosine and promote a pathway of active DNA demethylation (He et al. 2011; Ito et al. 2011) via base excision repair. Oxidation of 5-methylcytosine to 5-hydroxymethylcytosine may also result in passive demethylation during replication through a failure of DNMT1 to act on a substrate containing 5-hydroxymethylcytosine (Ji et al. 2014 ) Overexpression of TET1 contributes to aberrant hypomethylation and activation of oncogenic pathways in human solid cancer (Good et al., 2018).

#### **1.3.1.4.2 DNA hypermethylation in cancer**

Human cancer cells are characterized by localized, *de novo* aberrant hypermethylation of the CpG islands in tumour suppressor genes (TSGs) (Mizuno et al., 2001). This aberrant hypermethylation of CpG islands within promoter regions silences TSGs, leading to the development of the malignant phenotype (Mizuno et al., 2001). Since Retinoblastoma (Rb), the first TSG, was reported to be hypermethylated in cancer, aberrant promoter hypermethylation of TSGs has been found in virtually every tumour type. For instance, promoter hypermethylation of CpG islands of key TSGs such as CDH1, DAPK and RARB correlate with the repression of their transcription in various human cancers, including cervical carcinoma (Narayan et al., 2003). Hypermethylation of the CpG islands of promoter regions contributes

to development of the malignant phenotype through inactivation of genes involved in a variety of cellular mechanisms including the cell cycle, DNA repair, apoptosis, cell-to-cell interaction, and angiogenesis (Esteller, 2008). Tissue specific DNA methylation has been observed in both normal and cancer cells. It is estimated that 1% to 10% of CpG islands are aberrantly hypermethylated in individual tumours (Costello et al., 2000). However, a subset of CpG islands has found to be frequently methylated in different forms of tumour types (Ehrich et al., 2008).

### **1.3.2 Histone methylation**

Histones are chromatin-associated proteins that are involved in regulation of gene expression. A pair of each of the core histones (H2A, H2B, H3, and H4) makes up an octameric nucleosome, around which DNA winds. N-terminal tails of histones can be subject to covalent modification, including methylation, acetylation, phosphorylation and ubiquitylation. Histone modifications have important roles in several biological processes including gene expression (Kebede, Schneider and Daujat, 2015). Misregulation of histone modifications in particular histone methylation has been linked to a number of human diseases such as cancer (Fraga et al., 2005). Histone modifications can be involved in either transcriptional activation or repression. For instance, lysine acetylation is generally associated with transcriptional activation whereas lysine methylation correlates to transcriptional activation or repression depending upon the site and degree of methylation (Pelling et al., 2000).

Histone methylation occurs at both lysine and arginine residues. Lysines may be mono-, di- or tri-methylated (Heintzman et al., 2009), whereas arginines may be mono and symmetrical or asymmetrical di-methylated (Bannister and Kouzarides 2011). Numerous histone lysine methyltransferase and demethylases have been identified. The majority of histone lysine methyltransferases contain SET domain that harbours the enzymatic activity (Dillon et al.,

2005). Histone methylation considered as a reversible process since the discovery of first histone demethylase, H3K4 demethylase LSD1 (Shi et al., 2004). Almost all histone lysine demethylases possess a catalytic Jumonji domain (Whetstone et al., 2006). Histone lysine demethylase enzymes demethylate mono, di and tri methyl-lysine (Tsukada et al., 2006). Whilst H3K4, H3K36 and H3K79 has been shown to be associated with active gene expression, H3K9, H3K27me3 and H4K20 are associated with the epigenetic silencing (Jenuwein and Allis, 2001). The silencing mediated by tri-methylation of H3K27me3 is distinct from gene silencing by di-methylation at H3K9 in that it uniquely involves polycomb group proteins (PcG) (Cao and Zhang, 2004; Lachner and Jenuwein, 2002).

### **1.3.2.1 Histone 3 lysine 27 tri-methylation**

Histone 3 lysine 27 tri-methylation (H3K27me3) is involved in the regulation of X-chromosome inactivation and Hox gene expression (Zhai et al., 2007; Plath et al., 2003). Transcriptional repression of H3K27me3 is mediated by the family of chromatin regulating complexes that known as polycomb repressive complexes (PRCs). The mechanism of repression involves the formation of localised regions of silenced heterochromatin by the SET domain histone methyltransferase EZH2, as part of the PRC2 complex. PRC2 binds to the PRC1 complex, which contains a chromodomain protein (PC) (Schwartz and Pirrotta, 2008; Pietersen and van Lohuizen, 2008). PRC2 is composed of four core components, enhancer of zeste homologue 2 (EZH2), two WD40 domain proteins, EED and RBBP4 and suppressor of zeste 12 (SUZ12) (Müller et al., 2002). The enhancer of zeste, trithorax (SET) domain is the catalytic subunit of EZH2 and has methyltransferase activity towards H3K27 and lysine 26 of histone H1 (Kuzmichev et al., 2004). The histone methyltransferase activity of EZH2 enzyme is weak on its own. EZH2 and the other components of the PRC2 complex, such as EED and SUZ12, catalyse tri-methylation of the H3K27 (Czermin et al., 2002; Müller et al., 2002).

Demethylation of the H3K27me3 is performed by the histone demethylases, UTX/KMD6A and JMJD3/KDM6B, both containing a JmjC (Jumonji) catalytic domain that remove the methyl group (Agger et al., 2007). UTX is responsible for the activating histone H3 lysine 4 (H3K4) mark (Tie et al., 2014) that is suggesting a coordinated mechanism for transcriptional activation, in which the active H3K4 mark replaced repressive H3K27me3 methyl marks.

The amount and distribution of the H3K27me3 methyl mark can be pathologically altered by aberrant expression of histone methyltransferase and histone demethylases. Generally, the polycomb repressive complex (PRC) is highly active in cancer cells (Sparmann & van Lohuizen 2006). Studies suggest that EZH2 acts as an oncogene and its high level of overexpression has been documented in several malignancies including solid tumors (Bachmann et al., 2006) including human epithelial cancer (Kleer et al., 2003). The oncogenic activity of EZH2 has been linked to increased H3K27me3 levels and repression of tumor suppressor genes (Wang et al., 2012). As with EZH2, UTX behaves paradoxically, and both oncogenic and tumor suppressor effects are reported with UTX overexpression. Overexpression of UTX promotes proliferation and is associated with poor tumor prognosis (Kim et al., 2014). On the other hand, UTX expression is associated with tumor regression by silencing transcription factors (Choi et al., 2015).

### **1.3.3 Epigenetic reprogramming**

Epigenetic reprogramming refers to the erasure and remodelling of epigenetic changes like DNA methylation during mammalian development. However, it is also commonly observed in cancer, and it is suggested that it involves DNA methylation and PRCs. In fact, expression changes of a few transcription factors can reassign cell fate and induce pluripotent stem cells from adult somatic cells (Takahashi and Yamanaka, 2006).

### 1.3.3.1 Epigenetic switching

Epigenetic switching of developmentally significant genes from H3K27me3 to DNA methylation has already been shown in both stem and cancer cells (Figure 11). Indeed, this observation is consistent with the fact that genes that are reversibly repressed by H3K27me3 in human stem cells are more likely to be methylated in cancer (Gal-Yam et al., 2008; Widschwendter et al., 2007). It has been suggested that such epigenetic switching reduces epigenetic plasticity, thus resulting in the acquisition of a permanent heritable mark that silences key regulators (Kondo et al., 2008). Removal of the H3K27me3 mark is a prerequisite for epigenetic switching. Therefore, epigenetic reprogramming involves both the DNA methyltransferases and H3K27me3 demethylases.

DNA methylation provides highly stable silencing, whereas, easily reversible repression is induced by PRCs. Almost all of the polycomb target genes in embryonic stem cells are marked by a bivalent modification pattern, which involve both the activate H3K4me3 mark and the repressed H3K27me3 mark (Bernstein et al., 2006; Barski et al., 2007). It has been suggested that this bivalent action confers the potential for a gene to be driven either to its active or repressed state. Thus, genes in their active state also maintain the possibility of being readily repressed. However, genes that are repressed by this mechanism might revert to the active state during differentiation (Cedar and Bergman, 2009). The interplay between DNA cytosine methylation and histone modification ensures the stability of epigenetic states that are critical for development and differentiation. Alterations in these processes contribute to the establishment and maintenance of malignant phenotypes (Jones and Baylin, 2007). However, the evidence regarding an interaction between PRCs and DNA methylation is conflicting (Gal-Yam et al., 2008; Viré et al., 2006). Although, a direct physical connection between DNA methyltransferases and EZH2 has been suggested (Viré et al., 2006). There is a lack of DNA

hypermethylation at PRC-enriched regions in embryonic carcinoma cells (Ohm et al., 2007). Indeed knock down of EZH2 is associated with DNA hypomethylation in cancer cells (McGarvey et al., 2007). Although the increase in DNMT1 and DNMT3B expression in a variety of cancers correlates significantly with promoter hypermethylation (Rhee et al., 2002), the mechanism underlying aberrant hypermethylation of TSGs remains unclear. The presence of a stem-cell-like chromatin pattern predisposes genes to aberrant hypermethylation (Figure 12). Pluripotent stem cells rely on the polycomb-mediated tri-methylation of H3K27me3 to reversibly repress genes controlling differentiation. Of particular interest, de novo methylated genes in cancer have been shown to be marked by H3K27me3 in HES cells (Schlesinger et al., 2007), and the chromatin pattern of stem cells has been found to expose genes to DNA hypermethylation during the initiation of neoplasm formation (Ohm et al., 2007).

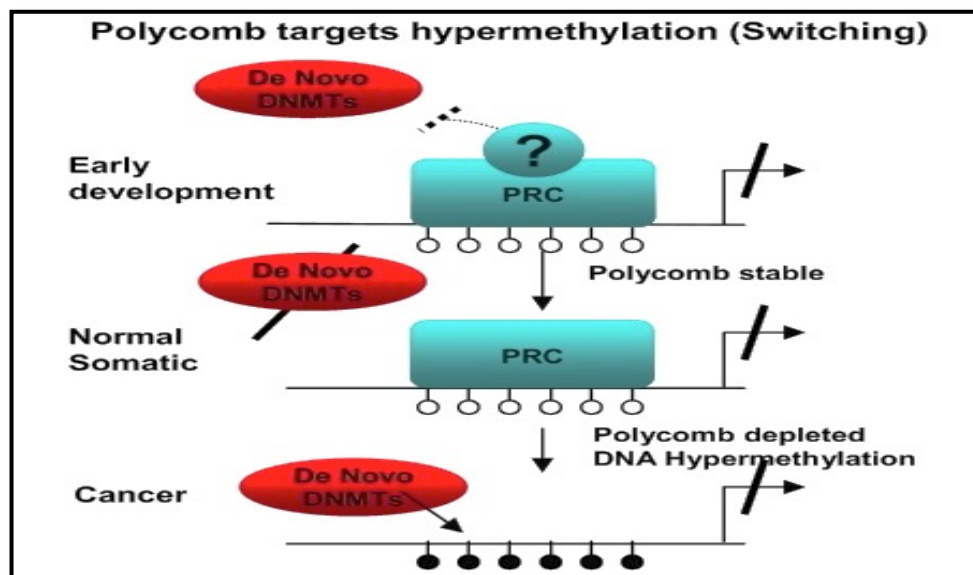


Figure 11: Schematic model of epigenetic switching. Arrows represent transcription; filled circles represent methylated CpG at promoters, whereas unmethylated CpG promoters are represented by empty circles. In cancer, reduction in polycomb repressive complexes are proceeded or followed by establishment of DNA hypermethylation. Taken from (Gal-Yam et al. 2008).

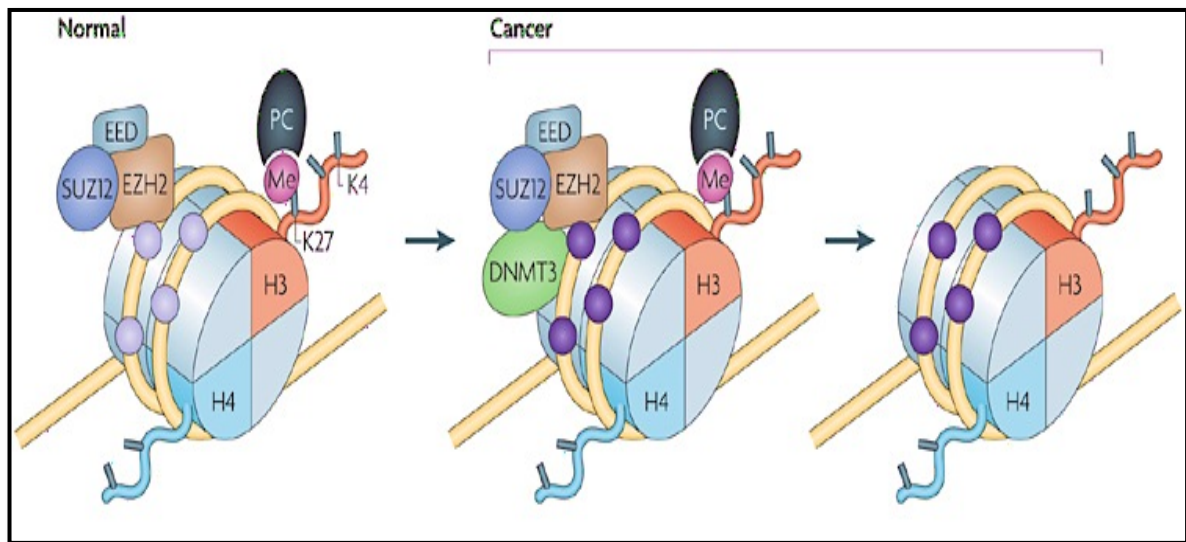


Figure 12: Model of *de novo* methylation in cancer. In normal cells, the genes have unmethylated CpG islands (light purple circles). Binding of the PRC2 complex (EZH2, SUZ12 and EED) to PRC1, which contains the chromodomain protein PC, maintains the histone H3K27me3 repressive modification. In cancer, the genes have methylated CpG islands (dark purple circles), possibly due to interaction between PRC2 and the DNMT methyltransferases. Taken from (Cedar and Bergman, 2009).



### **1.3.4 Methylation markers of cervical carcinogenesis**

The analysis of DNA methylation of gene promoters as a biomarker for cervical carcinogenesis seems to be promising. Recently, comprehensive analysis of whole-genome methylation patterns of HPV16 and HPV18 showed significant association between a diagnosis of CIN II or worse, and elevated DNA methylation levels in L1, L2, E2, and E4 ORFs (Mirabello et al., 2012; Wentzensen et al., 2012). Over the last decade, several studies have been investigating DNA methylation of human genes in cervical tissue. Most of the proposed methylation markers in cervical cancer were selected because their methylation was observed in other cancers. The variation in proposed candidate genes and in the level of methylation amongst the different studies for the same genes may be explained by the use of different techniques and their analysis of different promoter regions within the same gene.

In fact, some studies have investigated the functional involvement of proposed candidate genes in cervical carcinogenesis, or their HPV status. For instance, Steenbergen et al., (2004), studied the expression of CADM1 in relation to HPV status in cervical cancer cell lines. To study the relationship between altered methylation of a selective gene or a panel of genes in cervical carcinogenesis and HPV physical state and oncoprotein expression we have investigated the methylation status of cellular genes in primary human foreskin keratinocytes with HPV16 and HPV18 episomal forms (Leonard et al., 2012). We showed differential methylation of genes showed to be methylated in cervical neoplasia. Of particular interest, recent studies indicate that DNA methylation can be detected in cervical scrapings. Analysis of DNA methylation of cervical scrapings was found to closely reflect methylation patterns in cervical tissue (Reesink-Peters et al., 2004).

The detection of aberrant methylation in cervical smears several years prior to progression to invasive cervical carcinoma strongly suggests the implementation of methylation marker

analysis to treat women who test positive for high-risk HPVs (Wentzensen et al., 2009). Among several tumor suppressor genes (TSGs) that have been evaluated in cervical cancer, three genes (DAPK1, CADM1, and RARB) were consistently found methylated in cervical cancer (Wentzensen et al., 2010). Others, like Slit-Robo pathway genes are candidate tumour suppressor genes in a number of tumour types (Dallol et al., 2002b, 2002a), including cervical cancer (Narayan et al., 2006).

#### **1.3.4.1 Tumor suppressor genes in cervical carcinogenesis**

Wentzensen et al. (2009) reviewed several (TSGs) that were reported to be hypermethylated in cervical carcinogenesis. Progressive increase of promoter methylation of TSGs is common in cervical neoplastic progression. For instance, TSGs associated with the cell cycle, such as p16, CCNA1 and FHIT, showed increase methylation density at their promoters (Huang et al., 2011; Yang et al., 2010; Ki et al., 2008). Both CADM1 and E-cadherin, are TSGs involved in cell adhesion and their promoter hypermethylation was shown to reflect the severity of cervical dysplasia, and suggested to be methylation markers for the early detection of cervical cancer (Yang et al., 2010). Similarly, the frequency of promoter methylation of DAPK1, a TSG associated with apoptosis, cervical scrapings were found in, tissue specimens and in plasma samples and its methylation was increased with the severity of cervical dysplasia (Yang et al., 2010; Reesink-Peters et al., 2004). In addition, in cervical carcinogenesis methylation of TSGs of certain biological pathways has also been shown. Genes of the Wnt/ $\beta$ -catenin pathway and cell signalling pathways such as APC, and RASSF1A respectively are shown to be specifically methylated in adenocarcinomas but not squamous cell carcinoma (Wentzensen et al., 2009). Of particular interest, most of the Slit-Robo pathway genes have been shown to be candidate tumour suppressor genes in cervical cancer (Narayan et al., 2006).

#### **1.3.4.2 Role of the Slit-Robo pathway in cervical carcinogenesis**

The Slit-Robo pathway plays a crucial role during organ growth and development. The Slit-Robo interaction regulates cell migration, proliferation, apoptosis, adhesion, and angiogenesis in normal and tumour cells (Dickinson and Duncan, 2010). SLIT genes encode three homologous SLITs named SLIT1, SLIT2 and SLIT3. SLITs are large glycoproteins formed from N-terminal and C-terminal fragments (Simpson et al., 2000) and were identified as the ligand for the Robo receptor (Brose et al., 1999). The Roundabout (ROBO) gene encodes a family of transmembrane receptors (Seeger et al., 1993). They are subgroup of immunoglobulin characterized by the presence of IgG-like proteins and fibronectin type III repeats, a transmembrane portion and a cytoplasmic tail harbouring robo-specific motifs (Kidd et al., 1998). Recently, it has been shown that Slit-Robo genes modulate chemokines (Guan et al., 2003) which are number of genes that play role in tumour cell invasion and growth (Balkwill, 2004). The following chromosomal bands are ascertained for Slit-Robo genes; 3p12.3 (ROBO1 and ROBO2), 11q24.2 (ROBO3 and ROBO4), 4p15.3 (SLIT2), 10q24 (SLIT1) and 5q35 (SLIT3). Most of these regions have been frequently shown to be involved in loss of heterogeneity in number of tumours including cervical cancer (Mitra, Murty and Li, 1994; Pulido et al., 2000)

There is evidence that the Slit-Robo pathway is involved in tumourgenesis, either genetically by deletion of specific encoding region, or epigenetically by promoter hypermethylation. However, the SLIT2, ROBO1, and ROBO2 genes are not frequently targets of deletion in cervical cancer. In cervical neoplasia, SLIT1, SLIT2, SLIT3, ROBO1, and ROBO3 genes have been shown to harbour methylated CpG islands at their promoters (Narayan et al., 2006). In normal cervical epithelium, the expression of Slit-Robo in the basal and parabasal layer is high, whereas the intensity of expression is reduced with further

differentiation (Mitra et al., 2012). Although, SLIT genes and ROBO1 are expressed in normal cervical tissues, the ROBO3 gene is neither expressed in the normal cervix nor in cervical cancer cell lines (Narayan et al., 2006). Significantly reduced expression of Slit-Robo pathway genes was also revealed in primary cervical cancer (Narayan et al., 2006). The Slit-Robo pathway is inactivated through hypermethylation of promoter regions (Narayan et al., 2006). Therefore, Slit-Robo pathway genes are candidate tumour suppressor genes in cervical cancer. A high frequency of promoter hypermethylation of Slit-Robo pathway genes has recently been found in pre-cancerous cervical lesions (Narayan et al., 2006) and this hypermethylation increased significantly from pre-cancerous to early invasive cervical tumours. It has been shown that ROBO1 and ROBO2 promoter hypermethylation increased significantly from pre-cancerous to early invasive cervical tumours. Abrogation of the SLIT2-ROBO1 and SLIT2-ROBO2 signalling pathways plays a key role in the progression and initiation of cervical cancer (Mitra et al., 2012). This suggests the potential use of the ligand-receptor interactions of the Slit-Robo genes to as methylation markers in cervical neoplasia.

### **1.3.5 Epigenetic therapy**

Unlike the permanent nature of genetic changes, re-expression of genes silenced by DNA methylation offers strategies to induce the epigenetic changes via pharmacological manipulation. The main class of epigenetic therapy is DNA methylation inhibitors, of which 5-aza-2'-deoxycytidine and 5-Azacytidine have shown clinical activity against hematological cancers (Kantarjian et al., 2007). 5-aza-2'-deoxycytidine is incorporated into DNA, where azacytosine can substitute for cytosine and form Azacytosine-guanine dinucleotides which are recognized by DNA methyltransferase enzymes as a natural substrate. The DNA methyltransferases enzyme forms a covalent bond with the cytosine ring at the carbon-6 atom (Santi, Norment and Garrett, 1984a). The azacytosine block the bond by substitute carbon-5

with nitrogen and thus, prevent normal beta-elimination. This covalent bond between DNA and DNA methyltransferase blocks its enzymatic function. In addition, the covalent reaction affects the functionality of DNA and triggers DNA damage signalling, leading to the degradation of trapped DNA methyltransferases. As a consequence, this leads to the loss of methylation marks during DNA replication (Creusot and Christman, 1982; Santi et al., 1984a; Chen et al., 1991). Despite its ability to demethylate and reactivate expression of tumour suppressor gene (Chan et al., 2004), 5-azacytidine and its analogues showed very modest activity on solid tumours (Abele et al., 1987). In fact, it is more effective as demethylating agent at low concentrations but its toxic effects and poor clinical activity restrict their use against solid tumours (Issa et al., 2004). Hydralazine is another DNA methylation inhibitor and a known cardiovascular drug. Its use in treatment of patients with cervical cancer has shown its effectiveness in stimulating and re-expression of silenced tumour suppressor genes known to be found in cervical cancer, such as DAPK1, without affecting global DNA methylation (Zambrano et al., 2005).

## **1.4 Aim**

The overall aim of this thesis is to investigate the pattern of epigenetic changes induced by high-risk human papillomavirus, HPV16 and HPV18 in the early stages of HPV induced cervical neoplasia.

## CHAPTER 2

# **MATERIALS AND METHODS**

## **2 Materials and Methods**

### **2.1 Cell culture**

Illustrate of all cells used in this thesis with their HPV status, in (Table 5).

#### **2.1.1 Growing cervical cancer cell lines**

#### **2.1.2 Growth media and supplements**

Cell lines HeLa, SiHa, C33a, and Ca Ski were obtained from LGC promochen in partnership with ATCC. Cell lines HeLa, SiHa and C33a were maintained in DMEM while Ca Ski was maintained in RPMI. The media was supplemented with 2mmol/L of L-glutamine (GIBCO), 1% of (10,000U/mL) penicillin-streptomycin solution (Sigma-Aldrich Ltd, UK-Aldrich) and 10% fetal bovine serum (GIBCO).

#### **2.1.3 Cultivation of cervical cancer cell lines**

All the cancer cell lines used were adherent cells. Cells were maintained at 37°C in a humidified incubator with 5% CO<sub>2</sub>. Cells were grown in 10 cm culture dishes (IWAKI, Japan). Every 48 hrs the existing media were aspirated and replaced by fresh pre-warmed appropriate media to replenish nutrients and remove waste products. At confluence of 70-80% cells were either harvested or sub-cultured to reduce cell density and prevent inhibition of growth. Before media were aspirated cells were washed twice with 6-8 ml of 1 x phosphate buffered saline (PBS) to remove traces of media, which can interfere with the action of trypsin/EDTA, and then separated from the dishes by incubating them with 2 ml of 1 x trypsin/EDTA (GIBCO) for 5 min. Once the cells were separated they were pooled in a sterile universal container; 2 ml of appropriate pre-warmed media supplemented with 10% fetal bovine serum was added to suppress trypsin. Cell suspension was then pelleted at 1,000 rpm (eppendorf centrifuge 5415

R) for 5 min at room temperature. This centrifuge was always used unless otherwise indicated. The cell pellet then resuspended in fresh media to be plated out for new culture dishes at  $2 \times 10^5$  cells/dish for further cultivation.



Table 5: Properties of cell lines used in this thesis.

Cell line	Characteristics	HPV status	Information and properties
PHFK	Adherent	Negative	Untransfected primary human foreskin keratinocytes
PHFK HPV18	Adherent	HPV 18	PHFK from same donor. After transfected with episomal HPV18 passaged up to passage 14.
PHFK HPV16B	Adherent	HPV 16B	PHFK from same donor. After transfected with episomal HPV16B passaged up to passage 10.
PHFK HPV16K	Adherent	HPV 16 K	PHFK from same donor. After transfected with episomal HPV16K passaged up to passage 10.
W12	Adherent	HPV 16	Human cervical low grade lesion-derived cell line. Infected with HPV16. Early passages have episomal HPV16 and late ones have integrated HPV16 (Stanley et al., 1989; Alazawi et al., 2002)
HeLa	Adherent	HPV 18	Human cervical adenocarcinoma cell line, with 10-50 integrated HPV18 (Adey et al., 2013).
SiHa	Adherent	HPV 16	Human cervical cancer cell line with 1-2 copies of integrated HPV16 (Baker et al., 1987). Obtained from grade II squamous cell carcinoma. SiHa expresses p53 and pRb
Ca Ski	Adherent	HPV 16	Human cervical cancer cell line. Classified as Epidermoid carcinoma, established from a metastasis in the small bowel mesentery. Ca Ski contains 60-600 copies of integrated HPV16 with intact E2 ORF as well as sequences related to HPV18 (Baker et al., 1987).
C33A	Adherent	Negative	Human cervical cancer cell line, expresses p53 and pRb.
U2OS	Adherent	Negative	Human bone osteosarcoma.
J2 3T3 feeder cells	Adherent	Negative	J2 is the standard fibroblast cell line. Established from primary mouse embryonic fibroblast cells. The 3T3 refers to the abbreviation of 3-day, Transfer, inoculum $3 \times 10^5$ cells.

### **2.1.4 Growing primary cell lines**

Primary human foreskin keratinocytes (PHFK), established from the same donor and transfected with different episomal high risk HPVs, were used as *in vitro* models to study epigenetic changes after HPV transfection.

### **2.1.5 Establishing PHFK cell lines**

Generation of PHFK cell lines was established following methods described by Wilson et al. (2005). The PHFK used in this study were isolated from neonatal foreskin tissue of a single donor by Dr Sally Roberts. The Research Ethics Committee approval number for use of PHFK is 06/Q1701/45 and chief investigator is Dr Sally Roberts.

The PHFKs were maintained in serum-free keratinocyte growth medium (SFM, Invitrogen) and maintained in an incubator at 37°C with 5% CO<sub>2</sub> as adherent cells. Every 48 hrs the SFM was replaced. At 80% confluence, the media were removed and the cells were washed twice with PBS. One ml of 0.05% trypsin/EDTA was then added before cells were re-incubated at 37°C for 5 min. Trypsin was then inactivated by addition of trypsin neutralising solution (TNS, Invitrogen). Cells were pelleted by centrifugation at 500 rpm for 5 min before being resuspended with 15 ml SFM. The cell suspensions were then transferred to 10 cm tissue culture dish (IWAKI) or prepared for transfection at required density into a 5 cm dish.

#### **2.1.5.1 Preparing Feeder plates**

NIH 3T3 J2 mouse fibroblasts cells were cultured in (DMEM) HEPES modification (Sigma-Aldrich) supplemented with 10% of filtered 50 ml fetal bovine serum and filtered 10 ml of L-glutamine (GIBCO). When the cells reached 70-80% confluences they were separated from the dishes as shown in (Section 2.1.3). Cells at density of  $2 \times 10^6$  were then suspended with E-media in a 50 ml tube (Corning). Cell suspension was then irradiated using 50 gray of a

Caesium-137 source. Irradiated J2-3T3s cells were then plated out at  $2 \times 10^6$  cells per 10 cm dish in E-medium, (Table 5) and allowed to settle before the addition of PHFKs. Irradiated J2-3T3 were stored at 4°C and used for up to 5 days.

#### **2.1.5.2 Transfection of PHFK with HPV genome**

Transfection of the PHFKs with HPVs was undertaken by Dr Christothea M. Constandinous-Williams. In brief, prior to transfection, a plasmid containing the wild type HPV18 and two different variants of HPV16 (HPV16B or HPV16K) (Kirnbauer et al., 1993) genome were digested with EcoRI to release the HPV genomes. The genomes were re-circularised in the presence of T4 DNA ligase. Two variants of HPV16 were used. The only variation reported is a L1 sequence difference. L1 of clone 114/K (HPV16K) is considered to be a wild-type gene. While, clone 114/B (HPV16B), it assembles as efficiently as clone 114/K and has two amino acid changes. Therefore it is likely to be a natural wild-type variant and the effect of this difference on their oncogenic role remain unclear (Kirnbauer et al., 1993).

At passage 3 the PHFKs were co-transfected with the appropriate HPV genome and a plasmid containing a neomycin resistance gene. The transfected PHFK cells were then selected in G418 and serum containing E-medium for 8 days. When the clone pool was reaching approximately 2cm in size it was then transferred onto a plate with J2 3T3 fibroblast feeder cells in E-medium. Once they reached 70-80% confluence, the PHFKs were passaged into 10 cm tissue culture dishes sequentially. Three PHFK cell lines transfected with HPV18, HPV16B and HPV16K were generated shown in (Figure 13).

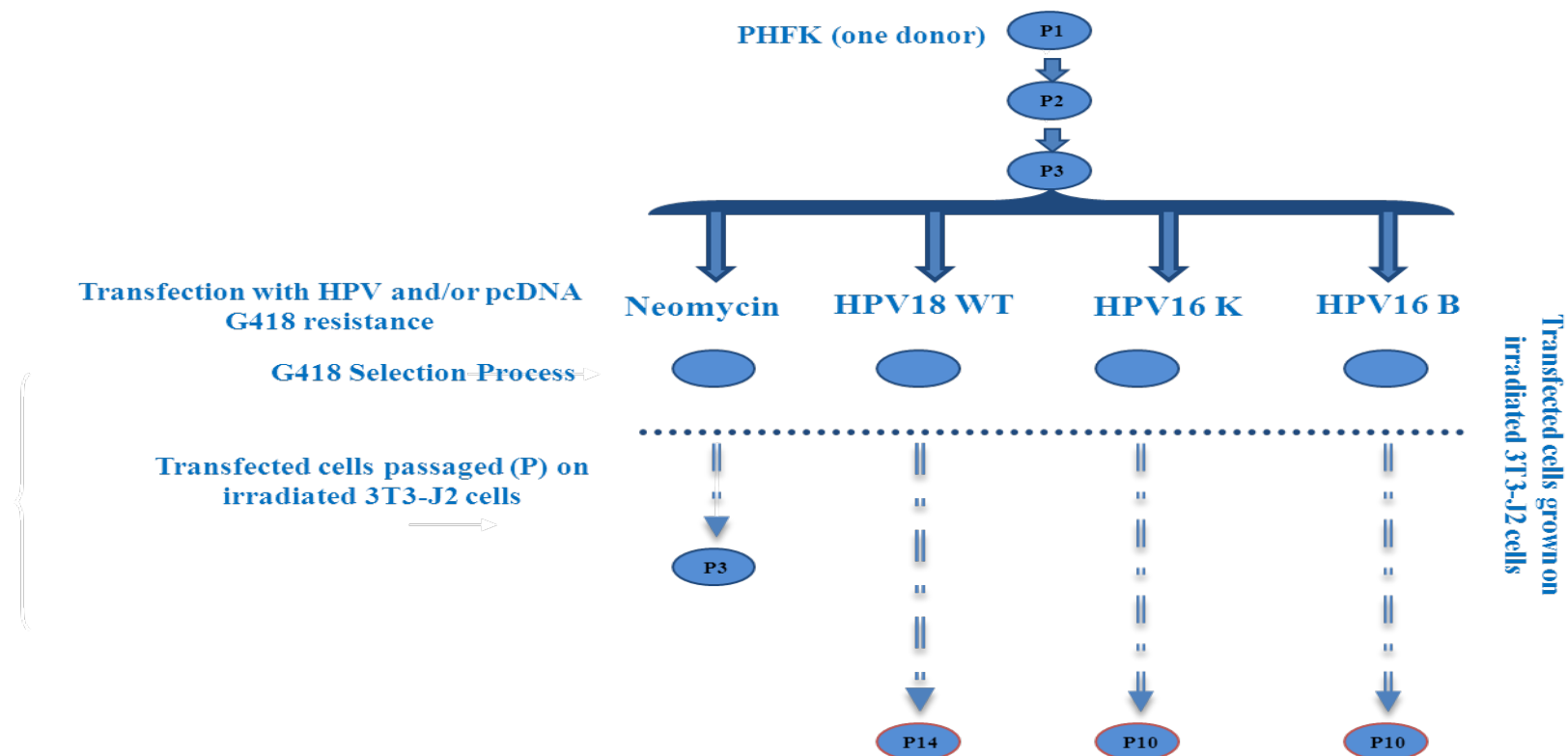


Figure 13: Flow diagram of PHFK transfection experiment. Primary foreskin keratinocyte cell lines which had been isolated from the same donor transfected with either episomal HPV16 or HPV18.

### 2.1.5.3 Growth media and supplements

The frozen aliquots of untransfected PHFK was maintained in defined keratinocyte serum-free media (GIBCO) for the first 24 hrs before growing in E-media; likewise the transfected PHFK and W12 cell line. The E-media was prepared using the ingredients in Table 6, the contents of 100 x cocktail shown in (Table 7) and filtered by 1.2  $\mu$ m filters (Corning Inc, USA). Prior to use 2.5 ml of epidermal growth factor (EGF, BD Biosciences) and 10 ml of 10 x L-glutamine (GIBCO) were added, and kept at 4°C.

Table 6: Supplements of E-media.

Supplements	Amount/ml	Supplier
DMEM+HEPES	1,200 ml	Sigma
Ham's F12	640 ml	Sigma
100 x Penicillin/Streptomycin	40 ml	PAA Laboratories
100 x Cocktail	20 ml	Sigma
100 x Hydrocortisone	2 ml	Sigma
1,000 x Cholera toxin	2 ml	ICN Biomedical
Inactivated Fetal bovine serum (5%)	100 ml	Hyclone®
10x L-Glutamine (5ng/ml)	10 ml	GIBCO

Table 7: Contents of 100x cocktail.

Supplements	Amount/ml
Adenine (0.18 M)	20 ml
Insulin (5mg/ml)	20 ml
Transferin (5mg/ml)	20 ml
5-Triiodo-L- thyronine (T3) (5mg/ml)	20 ml
PBS	120 ml

### **2.1.6 Cultivation of W12 and the PHFK following transfection with HPV**

The PHFK transfected with HPV and W12 cell lines were maintained at 37°C incubation with 5% CO<sub>2</sub>. Cells were grown in 10 cm culture dishes (IWAKI, Japan). Every 48 hrs the existing media was aspirated and replaced by fresh pre-warmed appropriate media to replenish nutrients and remove waste products and floating dead cells and irradiated J2s (feeder cells) were added every fifth day. At confluences of 70-80%, the cells were washed twice with 6-8 ml of 10 x PBS and then the feeder cells were separated from the dishes by incubating them with 2 ml of EDTA solution, the cells were then washed again with PBS and then primary cells were separated by adding 2 ml of 1 x Trypsin (GIBCO) for 5 min to be re-plated on new culture dishes at  $2 \times 10^5$  cells/ dish for further cultivation.

### **2.1.7 Investigating morphological changes of PHFK transfected with HPV18**

Both PHFKs passage 3 and PHFKs transfected with HPV18 passage 4 were cultured in 10 cm culture dishes (IWAKI, Japan) as described in the previous section, with and without feeder cells. At 70-80% confluence, pictures of cultured cells were taken while dishes over 37°C warm plate, using a Nikon microscope.

### **2.1.8 Harvesting cells for RNA, DNA and/or protein**

When the cells reached 70-80 % confluence they were harvested to extract RNA, DNA and/or protein. In order to harvest RNA, the dishes were placed in ice, as RNA is prone to rapid degradation by exogenous exonucleases. Following aspiration of the media, 6-8 ml of chilled PBS was used to wash the cells. Adherent cells were separated as appropriate and cell suspension was poured into a sterile falcon tube containing media and placed in ice. The cell suspension was then centrifuged at 1,000 rpm for 5 min at 4°C. The supernatant was discarded

and cell pellets were resuspended in 3 ml of chilled PBS. The homogenate suspension was then divided into three pre-labelled (DNA, RNA or protein) micro centrifuge tubes (Eppendorf, Inc. USA), 1 ml each. The tubes were directly re-centrifuged at 1,000 rpm for 5 min at 4°C. The supernatant was then discarded and the cell pellets were either freshly used or stored at -80°C.

### **2.1.9 Cell counting**

After the cells had been harvested, cell counting was undertaken. The haemocytometer (Sigma-Aldrich Ltd, UK) was prepared. The counting chambers were filled with approximately 10 µl of the cell suspension. The volume that can be contained in a large square is  $10^{-4}$  ml. Using 200 times magnification, the average of cells in four squares were counted giving an average number and multiplied by  $10^4$  ml. To give the overall number of cells in the sample, the number of cells was then multiplied by the total volume of cell suspension.

### **2.1.10 Cryopreservation**

Cells were harvested and pelleted as described previously before being re-suspended in fetal bovine serum at density of  $5-6 \times 10^6$  cells/ml were suspended in 1 ml of freezing media containing the cultural media, 10% fetal bovine serum and 10% of DMSO (Sigma-Aldrich Ltd, UK) or glycerol in case of primary foreskin keratinocytes transfected with HPV. The cell suspension was then transferred into 1.5 ml cryo-vials (Nunc, Denmark). The cryo-vials were then placed in a polystyrene container (Mr. Frosty, Nalgene, US) and stored at -80°C for 24 hrs to allow slow freezing of cells. The cryo-vials were then transferred to be stored in liquid nitrogen tanks (-160°C) for long-term storage.

### **2.1.11 Recovery of frozen cells**

The cells were thawed quickly by placing in cryo-vials containing 1 ml of frozen cells in a water bath at 37°C for 1-2 min to be rapidly defrosted as cryo-protectants such as DMSO

(Sigma-Aldrich Ltd, UK) are toxic at 4°C. The content was then transferred to 20 ml sterile universal tube diluted by adding pre-warmed appropriate complete culture media drop-wise slowly to allow the DMSO to diffuse out of the cells and therefore to minimise the toxic effect of cryo-protectants. The cells were then pelleted by centrifuging at 1,200 rpm for 5 min at room temperature. The media was then aspirated, discarded and replaced by 10 ml of fresh warm appropriate media. The cells were then examined microscopically and counted as shown in section 2.1.9 and seeded at  $1 \times 10^5$  cells per dish and cultured in the appropriate environment.

### **2.1.12 Preparation of mycoplasma testing**

Using the MycoAlert Mycoplasma Detection Kit (Lonza, Switzerland) all cell lines were periodically tested for the presence of mycoplasma, according to the manufacturer's protocol. The cells that tested positive were discarded, as the mycoplasma infection can obscure results and affect cell characteristics (Drexler and Uphoff, 2002).

### **2.1.13 Cell viability (cytotoxicity) assay**

Trypan blue (Sigma-Aldrich Ltd, UK) was used to determine the number of viable cells present in a cell suspension. Based on the principle that viable cells possess intact cell membranes that exclude trypan blue dye whereas dead cells do not, the total and viable cells were then counted. The cells were harvested as mentioned above and washed and centrifuged in PBS. An aliquot of cell suspension was then resuspended in 1 ml diluted PBS (as serum proteins stain with trypan blue and can produce misleading results). 10 µl of 0.4% trypan blue (stored in a dark bottle and filtered before use) and 10 µl of homogenate cell suspension were then mixed and incubated for 3 min at room temperature. Mixed cell suspension was counted as mentioned in Section 2.1.9, care was taken to count cells within 3 to 5 min as cell death and reduced viability counts can occur with longer incubation. The numbers of dead (stained) and



viable cells (unstained) were then recorded and the percentage of viable cells was calculated.

The results were expressed as the mean of three independent experiments.

Percentage of viable cells =  $\frac{\text{Number of viable cells per ml}}{\text{Number of viable + dead cells}} \times 100 \times 2 \times \text{sample dilution factor}$

#### **2.1.14 Demethylation treatment of cultured cells**

The cells were seeded on 10 cm culture dishes (IWAKI, Japan) as shown in section 2.1 and at 30-40% confluence they were treated with 5-aza-2'-deoxycytidine (Sigma-Aldrich Ltd, UK). The cells were treated with two different doses (Sova et al., 2006), 1  $\mu$ M and 5  $\mu$ M. Treatments were done in triplicate and control, for each dose. The 5-Aza-2'-deoxycytidine was prepared fresh by diluting with 0.1% DMSO (Sigma-Aldrich). The cells were treated for 72 hrs with 5-Aza-2'-deoxycytidine or 0.1% DMSO only, directly pipetting to the cultured media. Every 24 hrs both the cultured media and 5-Aza- 2'-deoxycytidine were replenished. On the following day after treatment, the cells were harvested as in section 2.1.8 and DNA and RNA extracted as shown in section 2.2.1 and 2.3.1 for further analysis.

#### **2.1.15 U2OS Tet-On inducible expression of HPV16 E7**

The expression of the tetA gene is strictly regulated by tetracycline responsive, Tet repressor (TetR). In the absence of an inducer, TetR binds to the operators tetO<sub>1</sub> and tetO<sub>2</sub>, prevents transcription of its own gene, tetR and of the resistance gene, tetA. Once tetracycline enters a cell, it binds TetR with high affinity leading to a conformational change in TetR resulting in its dissociation from tetO. The following rise of TetA and TetR expression leads to a rapid reduction of the cytoplasmic tetracycline concentration level, which shuts own expression of both genes off again. In the presence of tetracycline, TetA expression is fine-tuned so that export overcomes the slow uptake. This regulatory mechanism is the basis for application of TetR to selectively control the expression of a single gene (Berens and Hillen,

2003). Tetracycline-controlled transcriptional activation is a method used to induce expression of a gene of interest. This induction is reversibly turned on or off in the presence of the tetracycline or most commonly its derivative with a long half-life, doxycycline. This system has several other advantages over other regulated gene expression systems. It is extremely tight on/off regulation with high absolute expression, and rapid induction times. However, the maximal level of induction is usually not very high (Berens and Hillen, 2003).

#### **2.1.15.1 U2OS cell line**

U2OS is a human osteosarcoma cell line expressing wild type p53 and pRb, but lacking p16. It originally was known as 2T cell line. U2OS exhibits epithelial adherent morphology and viruses were not detected in the line during co-cultivation. The U2OS Tet-On system is formed from the U2OS cell line that serves as control and U2OS cell line that expresses tetracycline-regulated transactivator. Expression of HPV16 E7 is induced by the addition of doxycycline to the culture medium. The U2OS Tet-On system was a gift from Dr McLaughlin-Drubin (McLaughlin-Drubin, Margaret E. Crum, Christopher P. Münger, 2011)

#### **2.1.15.2 Maintenance of U2OS cells line**

The U2OS cell line is an adherent cell. It is maintained in DMEM supplemented with 10% Tet-Off system approved fetal bovine serum (Clontech) which ensure optimum induction with tetracycline inducible gene expression systems, 50 U/ml penicillin, 50 g/ml streptomycin and 250 g/ml G148 (Sigma). The cells were grown in an incubator at 37°C and with 5% CO<sub>2</sub>. Sub-culturing and freezing were like other cancer cell lines described in section 2.1.1.

#### **2.1.15.3 Determine the optimum dose of doxycycline**

A pilot study was undertaken to determine the optimum dose of doxycycline that causes the maximum induction of HPV16 E7 protein. The cells were seeded at  $2 \times 10^5$  cells in 10 cm

tissue culture dishes containing 10 ml of the appropriate medium and were allowed to grow for 48 hrs or 30-40% confluence. A different concentration (0, 0.5, 1.0 or 1.5 µg/ml) of freshly prepared doxycycline (Sigma-Aldrich) was added every 24 hrs for 3 days. After 24 hrs from the last dose, the cells were harvested and cell lysate was used to measure HPV16 E7 expression by western blotting. The dose of doxycycline (Dox) that induces maximum level of HPV16 E7 protein expression was then determined and used for required experiments.

#### **2.1.15.4 Investigation of DNMT3B level after induction of HPV16 E7 expression**

To investigate the possibility of HPV16 E7 induction would increase the expression of DNMT3B protein. U2OS inducible system and U2OS control were cultured for 48 hrs in triplicate and treated with 1µg/ml of doxycycline. After 24 hrs, cells were harvested and cell lysate was then subject to protein extraction. Induction of HPV16 E7 protein was first confirmed by western blotting. Investigation of DNMT3B level after induction of E7 protein was carried out a using western blotting.

## **2.2 DNA isolation and analysis**

### **2.2.1 DNA purification**

DNA extraction was undertaken by phenol-chloroform extraction.

#### **2.2.1.1 Lysis**

For every  $1 \times 10^7$  cells harvested, 400  $\mu$ l of lysis buffer was added. Lysis buffer was freshly prepared each time with 12.5  $\mu$ l of proteinase K (Roche), 40  $\mu$ l 10 x PCR buffer (Sigma), 2  $\mu$ l Tween 20 (Sigma) and 358  $\mu$ l nuclease-free water (Sigma). Cells were then vortexed and incubated overnight at 55°C.

#### **2.2.1.2 Phenol-chloroform extraction**

For each sample 400  $\mu$ l of phenol (Sigma) was added. Following vortexing and centrifugation at 13,000 rpm for 5 min the supernatant was carefully aspirated to a fresh tube, avoiding aspiration of both the interlayer and the organic phase. An equal volume (400  $\mu$ l) of phenol and chloroform (Sigma) was added to the sample, vortexed and centrifuged for 5 min. The supernatant was again carefully transferred to a fresh tube. To remove phenol one volume of chloroform (400  $\mu$ l) was then added. After the sample was vortexed and centrifuged at 13,000 rpm for 5 min, the upper aqueous layer was then transferred to a fresh tube to be precipitated in cold ethanol.

Prior to precipitation at -80°C overnight, the sample was vortexed with 1,000  $\mu$ l of 100% ethanol, 40  $\mu$ l volume of 3M sodium acetate (Sigma), and 1  $\mu$ l of glycogen (Invitrogen).

#### **2.2.1.3 Clean-up and Quantification**

The sample was firstly centrifuged at 13,000 rpm at 4°C for 30 min and the supernatant was carefully discarded. The sample was then centrifuged at 13,000 rpm at 4°C with 1,000  $\mu$ l

of ice-cold 70% ethanol for 30 min. After the supernatant was discarded, the sample was centrifuged at 13,000 rpm at 4°C, again for 10 min. To each sample, 500 µl of 70% ethanol was then added, and centrifuged at 13,000 rpm at 4°C for 10 min. Following this the pellet was air dried at 37°C for 30 min, and then resuspended in 50 µl of nuclease-free water. Extracted DNA samples were then stored at -20°C until use. The DNA was then quantified on a NanoDrop ND-1,000 Spectrophotometer (Thermo Scientific) according to the manufacturer's instructions. The ratio of optical density (OD) at 260 and 280 nm (A<sub>260</sub>/A<sub>280</sub>) of all the DNA samples were approximately 1.8.

### **2.2.2 Polymerase Chain Reaction (PCR)**

Fifty ng of DNA from each cell line was then amplified in a 25 µl volume using 12.5 µl of GoTaq Green Master Mix (Promega) and 12.5 pmol of each primer listed in (Table 8). Amplification was performed using a Px2 Thermal Cycler with the following cycle conditions: 95°C for 10 min, followed by 40 cycles of 95°C for 45 sec, annealing temperature for 1 min, 72°C for 2 min and a final extension of 72°C for 10 min. The PCR products were analysed using 2% agarose gel electrophoresis and ethidium bromide staining. Three independent experiments were performed.

#### **2.2.2.1 Designing PCR primers**

Human papillomaviruses' genomic sequences were obtained from the Addgene website ([www.addgene.org](http://www.addgene.org)). Human genomic sequences were obtained from the UCSC genome browser website ([www.ucsc.com](http://www.ucsc.com)). Primer 3 software version 0.4.0 ([www.primer3.com](http://www.primer3.com)) was used to design PCR primers. Desirable characteristics of the primer sequences were an annealing temperature of ~58 °C, a length of 18-25 bp and G+C content of ~50%. Product sizes

varied according to the region covered. All primers shown in (Table 8) were ordered from (AltaBiosciences, University of Birmingham, UK) and then optimised and validated.

### **2.2.2.2 Agarose gel electrophoresis**

Gel electrophoresis is a procedure that uses an electric field to enable migration of nucleic acid (DNA or RNA) molecules on the basis of their sizes; the smaller molecules move faster through the gel than the larger molecules. Nucleic acid molecules are negatively charged and move toward the anode at different rates. Agarose concentration is an important factor as it affects the migration speed. To visualise DNA or RNA bands, ethidium bromide dye is used. It fluoresces under UV light when intercalated into double-stranded DNA and RNA molecules.

### **2.2.2.3 Preparation of agarose gel**

To analyse molecules less than 1 Kb a 2% agarose gel was used: 4 g of agarose (Promega, USA) dissolved in 200 ml of 1 x a Tris-Borate-EDTA buffer solution (Sigma, UK). Different concentrations were made as required by estimated molecule size. After heating the solution in a microwave it was allowed to cool in a water bath for a few min before 4 µl 20mg/ml ethidium bromide (Sigma-Aldrich) was added. A gel casting tray was prepared by sealing the ends of the gel chamber with tape or appropriate casting system and an appropriate number of combs were placed in the gel tray. The cooled gel was then poured into the gel tray and allowed to cool for 30-60 min at room temperature. The combs were then gently removed and the gel placed in an electrophoresis chamber and covered with pre-prepared loading buffer (100 ml 10 x TBE buffer + 900 distilled water + 12.5 µl of 20 mg/ml ethidium bromide) until the gel and electrodes were completely covered. To allow estimation of band sizes within the samples, 13 µl of 100 base pair of DNA ladder (Promega, USA) with 2 µl of 1:5 dilution loading buffer (30% glycerol, 0.25% bromophenol blue) was mixed and pipetted into the first well in each row. An electric

voltage of 120 V was applied for the appropriate time indicated by estimated molecular size. A transilluminator (Bio-Rad) was used to visualise separated bands.

Table 8: List of PCR primers sequences used to amplify E1/E2/E5 region of HPV18. Sequences marked by \* obtained from (Collins et al., 2009).

Primer	Accession Number	Sequence	Nucleotide Sequence 5'-3'	A.Tm	bp
Long GAPDH fw	AY340484	187	GGGCTGCTCACATATTCTGG	55°C	980
Long GAPDH rev		1166	CCATTCCCCAGCTCTCATAC		
Short GAPDH rev*	AY340484	3701	GCTCAAGGGAGATAAAATTC	55°C	158
Short GAPDH fw*		3858	CGACCAAATCTAAGAGACAA		
HPV18 long E1/E2 fw	NC_001357	2507	ACGACCACGTGTTGGACATA	55°C	940
HPV18 long E1/E2 rev		3598	CCGTGTCGTCCTGGTACTG		
HPV18 E1/E2 set1 fw	NC_001357	2493	GTTAGATGATGCAACGACCAC	55°C	335
HPV18 E1/E2 set1 rev		2828	CGGTGTCTGCATCTTCCTCTT		
HPV18 E1/E2 set2 fw	NC_001357	2769	TTTTGAAAGGACATGGTCCAG	55°C	352
HPV18 E1/E2 set2 rev		2828	CATGTGTCTTGCAAGTGTCCA		
HPV18 E2 set1 fw	NC_001357	3086	ATACAAAACCGAGGATTGGA	54°C	303
HPV18 E2 set1 rev		3386	ACTTCCCACGTACCTGTGTT		
HPV18 E2 set2 fw*	NC_001357	3369	CGACCACGTGTTGGACATAC	55°C	390
HPV18 E2 set2 rev*		3739	GTAGCGGTTTTGTCCCATGT		
HPV18 E2/E5 set4 fw	NC_001357	3450	GACCTGTCAACCCACTTCT	54°C	397
HPV18 E2/E5 set4 rev		3994	ACATGGCAGCACACATACAT		
HPV18 E7 fw	NC_001357	76	GTTGACCTTCTATGTCACGA	55°C	151
HPV18 E7 rev		226	CAATTCTGGCTTCACACTTA		



### **2.2.3 DNMT3B gene copy number**

Copy number assays of gene of interest provide an accurate DNA quantification required for analysis of the gene copy number. After DNA samples were purified digestion of RNA with RNaseA, was carried out to ensure removal of RNA contamination that can cause inaccuracies in DNA concentration.

#### **2.2.3.1 Quantitative Real Time - polymerase chain reaction**

The reaction was set up in 96-well plates with four replicates for each sample and PHFK was used as a calibrator sample for PHFK transfected with HPV16B passage 10, HPV16K passage 7 or HPV18 passage 14. I also investigated W12 cell line and cervical cancer cell lines (C33a, HeLa, Ca Ski and SiHa). To allow for imprecision in pipetting that can affect the interpretation of the copy number results, 5% more mix was prepared. Each four replicates were made up of 100 µl in total with 80 ng of DNA made up to 45µl with nuclease free water, 50 µl of 2 x qBiomarker SYBR Mastermix (Qiagen) and 5µl of DNMT3B gene copy number assay (Qiagen) or reference gene copy number assay, MRef (Qiagen). Twenty-five µl reaction mix was added to each well. The plate was then sealed with optical adhesive film (Applied Biosystems) and placed in ABI 7500 standard real-time PCR system. Absolute quantification program was selected and the PCR system ran according to the manufacturer's instructions (Applied Biosystems, Foster City, CA) and was subjected to 40 thermal cycles (Table 9).

Table 9: RT-PCR thermal cycle programme

Stage	Cycle step	Repetitions	Temperature	Time
1	Enzyme activation	1	50°C	02min
2	Denature	1	95°C	10min
3	Denature	40	95°C	15sec
	Anneal/extend		60°C	01min
	Hold		04°C	

To verify PCR specificity a melting curve program was run. A dissociation curve was obtained for each well by ABI 7500 standard real-time software. Single peak only was accepted per reaction.

### 2.2.3.2 Analysing DNMT3B gene copy number

The calculated threshold cycle for each well was exported to a blank Excel spread sheet. Data were then used for qBiomarker copy number PCR assay data analysis web-based software (Qiagen). Samples that did not pass quality control were removed. Only statistically significant changes with  $p$  value  $\leq 0.05$  were considered.

### 2.2.4 Methylation analysis

Gene-specific methylation analysis can be characterised as genome wide or candidate gene approaches. Figure 14 summarises available methods reviewed by Shen & Waterland (2007). In my thesis, pyrosequencing was used to quantitatively measure DNA methylation at specific loci of interested genes. Studying genome wide methylation changes of PHFK transfected with HPV18 for passage 7 and passage 14 was done with Illumina Human Methylation 450K methylation array.

#### 2.2.4.1 Sodium bisulphite treatment of DNA

Bisulphite modification of genomic DNA converts all unmethylated cytosines to uracil while methylated cytosines are protected from deamination and sulphonation and are

unchanged (Wang, Gehrke and Ehrlich, 1980). Using bisulphite genomic sequencing (BGS) the methylated cytosines will be detected.

Bisulphite modified DNA is an uncomplimentary single-stranded DNA molecule and subsequent PCR analysis will result in replacement of the converted uracil to thymine. The EZ DNA Methylation-Gold Kit (Zymo research) was used according to the manufacturer's instructions. For every 500 ng of DNA, 130 µl CT conversions were added. The sample was then subjected to PCR; after denaturing at 98°C for 10 min, it was incubated for 2.5 hrs at 64°C followed by cooling to 4°C. The PCR product was applied to a Zymo-Spin™ IC Column and 600 µl binding buffer was added. The sample was then centrifuged at 13,000 rpm for 30 sec; the flow-through was discarded and 100 µl wash buffer was added. After the sample was centrifuged again for 30 sec, 200 µl of desulphonation buffer was added. The column was left to stand at room temperature for 20 min before it was centrifuged at 13,000 rpm for 30 sec. Following this the column was washed twice with 100 µl wash buffer, 10 µl elution buffer was added directly to the column twice and centrifuged at 13,000 rpm for 30 sec. The bisulphite modified DNA was stored at -20°C for up to 3 months.

#### **2.2.4.2 Pyrosequencing**

Pyrosequencing (PyroSeq) is a method of DNA sequencing which allows high-throughput analysis of CpG methylation quantitatively. This method has high reproducibility and provides an inbuilt internal control for the adequacy of bisulphite modification (Tost, Dunker and Gut, 2003). The DNA fragment of interest is immobilised and converted into single-stranded form, subsequently, a sequencing primer hybridises to it. A deoxynucleotide base is repeatedly incorporated into the growing DNA strand only if it is complementary to the base on the template strand. The amount of ATP produced determined by the luciferase assay.

The amount of light generated is proportional to the number of pyrophosphate molecules released, and detected using a luminometer.

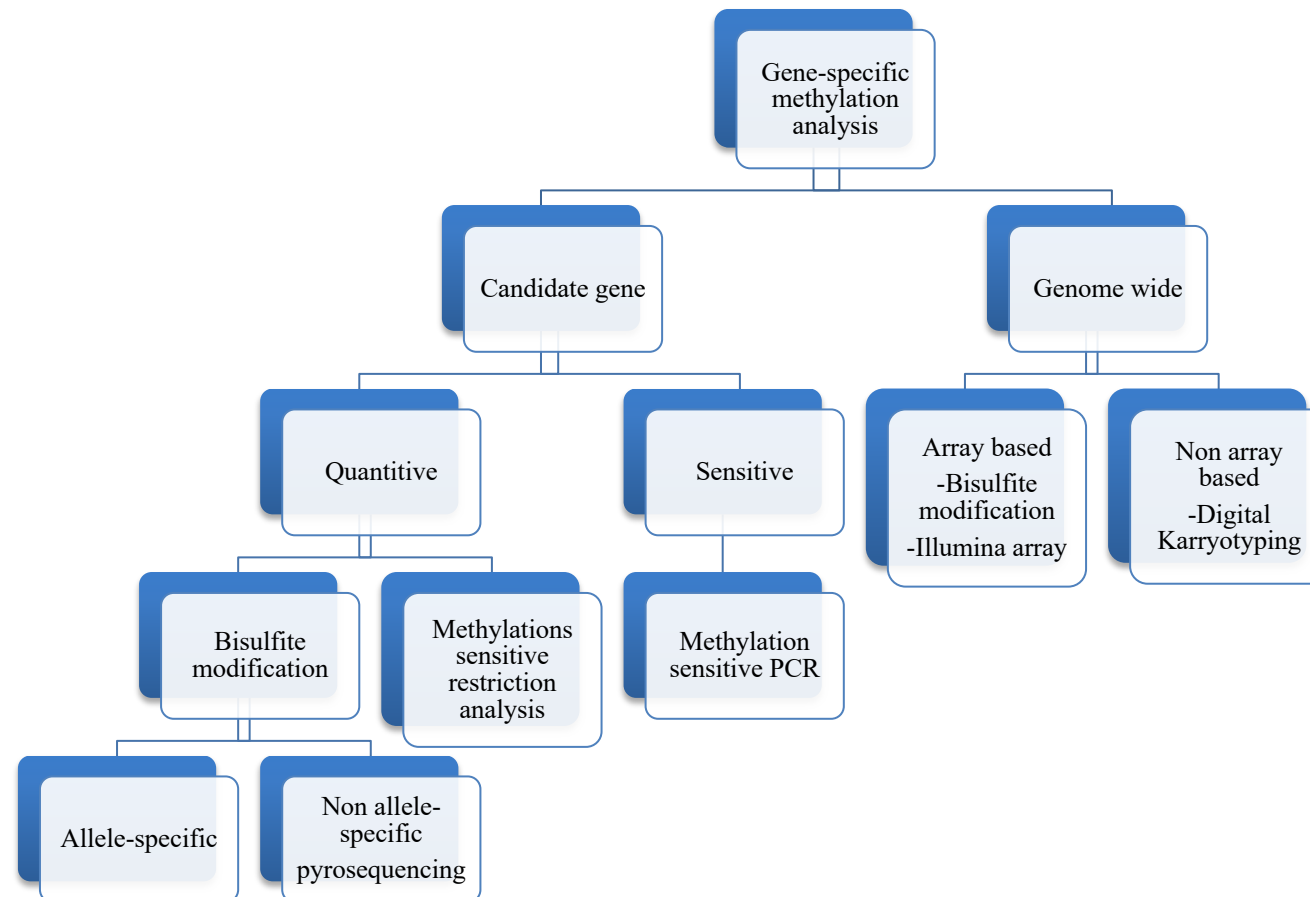


Figure 14: Methods of gene-specific methylation analysis.

#### **2.2.4.3 Pyrosequencing primer design**

Biotage PSQ Assay Design software version1.0 was used to design all PyroSeq primers. The genomic sequences of genes of interest were obtained from the illumine array and checked in a UCSC genome browser (<http://genome.ucsc.edu/index.html>). Sequences were bisulphate modified in silico, which means CpG dinucleotides were converted to YG, while the remaining cytosines (C) bases were converted to thymidine (T). Based on the bisulphite converted sequence, a biotinylated primer, non- biotinylated primer and a sequencing primer were generated. The genomic sequence was deposited in Biotage PSQ Assay Design software, which generates three primers: forward, reverse and sequencing primer. Only high-quality primers were accepted and ordered. All PyroSeq primers were optimised and validated (Biomers). Primer details are listed in Table 10.

#### **2.2.4.4 Pyrosequencing PCR**

PCR was performed in a total volume of 50 µl; five µl of bisulphite modified DNA, was amplified using 25 µl of Hotstarttaq Master Mix (Thermo Scientific), 5 pmol biotinylated primer and 10 pmol non-biotinylated primer. Amplification cycles were performed in a Thermal Cycler under the following conditions: 95°C for 15 min, followed by 40 cycles at 95°C for 30 sec, 45 sec of annealing temperature, 72°C for 30 sec, and final extension of 72°C for 10 min. Prior to commence PyroSeq, 5 µl of the PCR product was analysed using 2% of ethidium bromide staining agarose gel electrophoresis. Only if the expected size band was confirmed, was the PCR product taken forward for PyroSeq.

#### **2.2.4.5 Pyrosequencing reaction**

On a 96-well plate (Applied Biosystems), 40 µl of PCR product was bound to 3 µl streptavidin beads (Amersham) in 37 µl binding buffer (Qiagen). While the plate was left to shake for 5 min at 1,300 rpm a sequencing plate (Qiagen) was prepared with 15 pmol sequencing primer and 38.5 µl annealing buffer (Qiagen). The hedgehog vacuum probe was used to capture the beads from the 96-well plate at PyroMark Q96 Vacuum Prep Workstation (Qiagen). The beads were washed for 5 sec once in each of the following solutions; 70% ethanol, denaturation buffer (Qiagen) and wash buffer (Qiagen). Thereafter, the beads were released from the hedgehog probe onto the previously prepared sequencing plate and heated to 80°C for 5 min. Meanwhile, the plate was left to cool down for 5 min; the reagent cartridge (Qiagen) was loaded with the appropriate volume of PyroGold SQA reagents (enzyme, substrate and dNTPs; Biotage) and calculated by using the PyroMark ID software. The plate was set to run according to the manufacture's instruction on PyroGold SQA reagents (Qiagen). Data were analysed using the Pyro Q-CpG software (Qiagen).

Table 10: List of Pyrosequencing primers used in this thesis, and their annealing temperatures (A.Tm). fw, forward. rev, reverse. seq, sequencing primers

Primer	Sequence 5'-3,	Modification	A.Tm
CADM1 fw	TGGTTGAAAAGAAAGTAAAGAGAAT		52°C
CADM1 rev	CCTATCTCAAAAAAACCTCTCCTACTA	Biotin	
CADM1 seq	GGGGAGGGAGAGTGA		
CALCA fw	GGGTTTTAGTTTTATTGTGTTTGAG		52°C
CALCA rev	AACCCTAAAAAACTTCTTTCTTACCACTCT	Biotin	
CALCA seq	GTTTAGTAGAGGTAGGG		
CHFR fw	AGGGAGGTTTTAAGAGGGTGAAGTAG		52°C
CHFR rev	ACCTAAACCTTCCAAAACACT	Biotin	
CHFR seq	TTTTAAGAGGGTGAAGTAG		
DCC fw	GGTGGGGGTAGATAGATTATG		52°C
DCC rev	CAAACCAAACAATACAACCTCCA	Biotin	
DCC seq	GGGGTAGATAGATTATGG		
ESR1 fw	GGGGTATATAAGGTAGTATATTAGAGAAAG		52°C
ESR1 rev	ATCTACTAAAAAATAACCCAAAAAAACT	Biotin	
ESR1 seq	AGGTAGTATATTAGAGAAAGT		
PTPN6 fw	GGGGTTGTTATTTTATTGGTTTGG		52°C
PTPN6 rev	CTTCTCTTCCCTACTATACTCTAAAC	Biotin	
PTPN6 seq	ATTTTATTGGTTTGGGGT		
SFRP2 fw	GGTTAGGTTTTTTTGTGTTGTTTAAAG		52°C
SFRP2 rev	AAATTCTAACTACCCCCCTCCCTAAAT	Biotin	
SFRP2 seq	TTTTTGTTTGTTGTTTAAAGA		
SOX11 fw	GGTGGAGGTAAAGATAATAAAAAAAGGT		52°C
SOX11 rev	ATACATCAAAAAAACTTTCCACCTAC	Biotin	
SOX11 seq	GATAATAAAAAAAGGTAGAAATAT		
TP53AIP1 fw	AAGGTGGTTTTTATATTAAGTTAAAGTGT		52°C
TP53AIP1 rev	CCTTCAATCAATTTCAAAAAACTAACTA	Biotin	
TP53AIP1 seq	GGTTTTATTAAGTTAAAGTGTT		
TNFRSF10C fw	TGTTTGGAAGTGATTGTTGTAAG		52°C
TNFRSF10Crev	AAAAATTCCCACAACCTACACTAC	Biotin	
TNFRSF10C seq	AGTGATTGTTGTAAGTGATA		
RARB fw	GATTAGAGTAGGGTTTGTGTTGGGTAT		52°C
RARB rev	CAAATTCTCCTTCCAAATAAATACTTACAA	Biotin	
RARB seq	TCCAAATAAATACTTACAAAAAAC		
DAPK1 fw	GGTTAGGTTTTTTTGTGTTGTTTAAAG		52°C
DAPK1 rev	AAATTCTAACTCCCCCCTCCCTAAAT	Biotin	
DAPK1 seq	TTTTTGTTTGTTGTTTAAAGA		
VHL fw	GGTGGAGGTAAAGATAATAAAAAAAGGT		52°C
VHL rev	ATACATCAAAAAAAACTTTCCACCTAC	Biotin	
VHL seq	GATAATAAAAAAAGGTAGAATAT		



## **2.2.5 Genome wide methylation analysis**

The Human Methylation 450 Bead Chip allows interrogation of more than 485,000 methylation sites per sample at single-nucleotide resolution. It offers broad converge of methylation analysis by combining both Infinium I and Infinium II assay analysis. The Sequenom data were generated by the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics (funded by the Wellcome Trust grant reference 090532/Z/09/Z and MRC Hub grant G0900747 91070).

### **2.2.5.1 Illumina array methylation assay**

The protocol was undertaken according to the manufacturer's instruction; Figure 15 shows schematic presentation of illumina methylation HD assay. In brief, 500 ng of genomic DNA for each sample was bisulphite converted using Zymo EZ Methylation Kit. The bisulphite converted DNA were then denatured before being amplified. End-point fragmentation was used to fragment samples. The samples were then centrifuged at 4°C, before being resuspended in the hybridisation buffer. DNA samples were dispensed onto BeadChips and incubated in an Illumina Hybridization Oven overnight at 37°C. During hybridisation, each DNA locus specifically linked to one of over 500,000 bead types. For the Infinium I assays two beads per CpG locus; one bead corresponds to methylated cytosine, and another bead type to the unmethylated state. For the Infinium II assays only one bead type corresponding to each CpG locus. The beads were then washed before staining and extension step. The Illumina Scan System was then used to scan the BeadChip; this system uses a laser to excite the fluorophore of the single-base extension product on the beads. The scanner was set according to the

manufacturer's instructions. The image and data file were then analysed and quality control undertaken with Illumina GenomeStudio Software (Illumina, San Diego, CA).

### **2.2.5.2 Illumina HD 450K methylation array data analysis**

Analysis was undertaken by Illumina GenomeStudio Software (Methylation Module) and Microsoft Excel software package with advice from Dr Wenbin Wei, Senior Bioinformatician, School of Cancer Science, University of Birmingham. The Assay QC is performed with Illumina GenomeStudio Software, which also generates the final report with genotyping results. A quantitative measure of DNA methylation was assigned by Illumina GenomeStudio Software. This value was calculated from the ratio of fluorescent signals from the methylated alleles to the signals from both the unmethylated and methylated alleles. Upon transformation of the data, beta values ranged from 1 (completely methylated) to 0 (completely unmethylated) over the studied CpG sites.

Significantly differentiated sites were assigned by a differential score (DiffScore) for a probe, as computed by:  $\text{DiffScore} = 10 \cdot \text{sgn}(\beta_{\text{cond}} - \beta_{\text{ref}}) \log_{10} p$ . The DiffScore is a transformation of the  $p$ -value that provides directionality to the  $p$ -value based on the difference between the average signals in the reference group p7 versus the comparison group p14. Thus, for a DiffScore =  $\pm 13$ , the  $p$ -value = 0.05; for a DiffScore =  $\pm 22$ , the  $p$ -value = 0.01; and for a DiffScore =  $\pm 33$ , the  $p$ -value = 0.001. Differentially methylated sites were selected on the basis of the Delta value. The Delta value was calculated from differences in mean methylation changes between the two replicates, p14 and p7, of the PHFK HPV18 cell line samples. A CpG site was considered differentially methylated if the mean difference was  $\geq 10$ -20%. Only statistically significant sites with DiffScores  $\geq \pm 13$  and  $p$  values  $\leq 0.05$  were considered.

Microsoft 7 Excel software package was used for a detailed descriptive analysis. The Database for Annotation, Visualization and Integrated Discovery (DAVID) version 6.7 software was then used for functional annotations (Dennis et al., 2003).

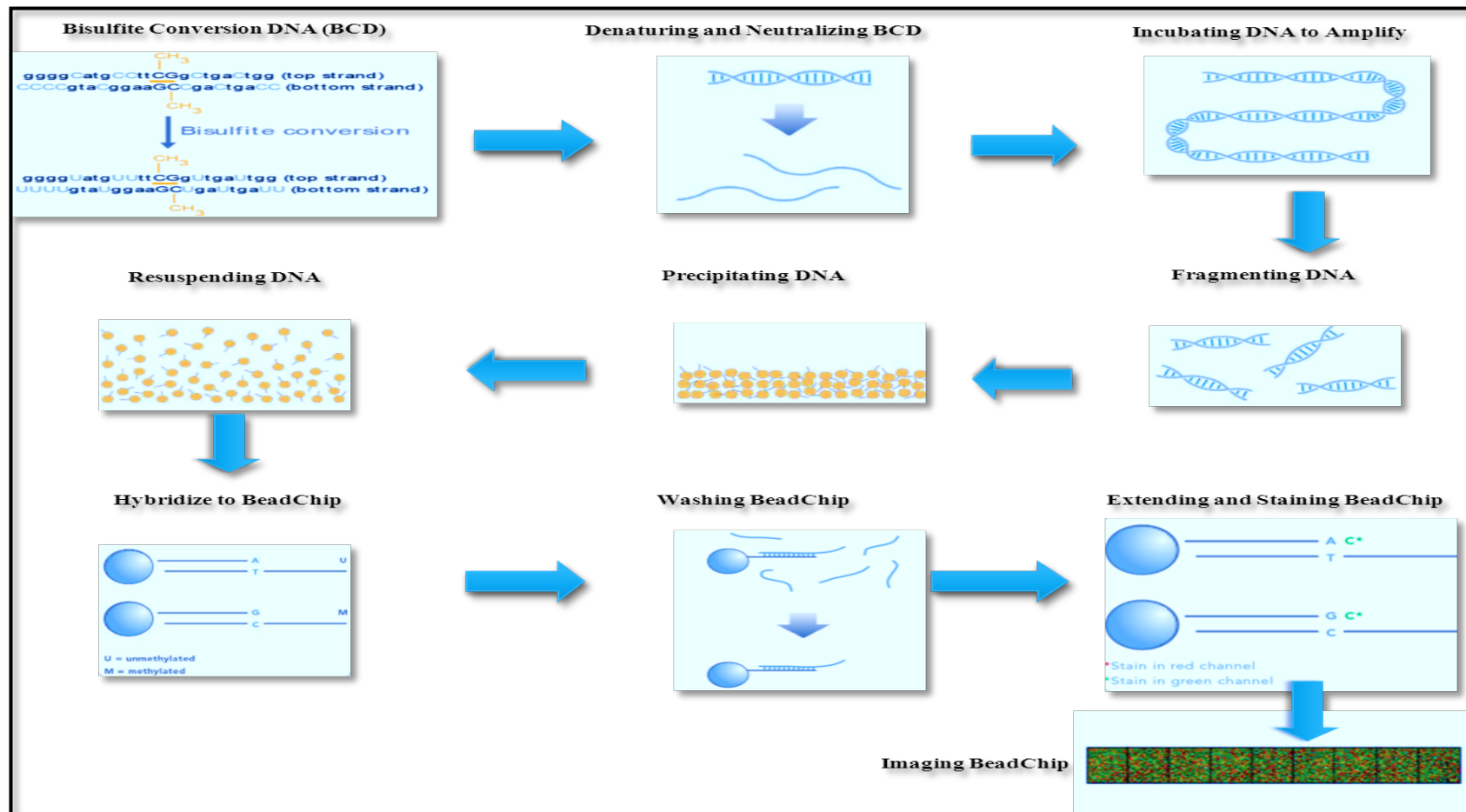


Figure 15: Schematic presentation of Illumina methylation HD assay. Adapted from (Illumina Infinium HD Assay methylation protocol guide, 2015).

## **2.3 RNA detection and analysis**

### **2.3.1 RNA extraction**

Total RNA extraction was carried out using the RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. Cells were lysed by incubation for 5 min at room temperature with 350 µl of lysis buffer. Cell lysates were then mixed with 350 µl of 70% ethanol before being transferred to spin columns placed in collection tubes. Following centrifugations at 8,000 rpm for 15 sec, the flow-through was discarded and 700 µl Buffer RW was added and spin columns were centrifuged again. The flow-through was discarded and the spin columns were transferred to fresh collection tubes before 500 µl of Buffer RPE was added. The spin columns were then centrifuged at 8,000 rpm for 15 sec and another 500 µl Buffer RPE was added. The columns were centrifuged at 8,000 rpm for 2 min and then transferred to microcentrifuge tubes. In order to dry, the membrane columns were centrifuged for 1 min at 13,000 rpm before being transferred to centrifuge tubes provided in the RNeasy kit to elute the RNA; 30 µl of DEPC water was pipetted directly onto the membranes and the columns were centrifuged for 1 min at 8,000 rpm.

### **2.3.2 Assessment of RNA concentration**

The concentration of the RNA was measured using a NanoDrop ND-1,000 Spectrophotometer (Thermo Scientific) according to the manufacturer's instructions. The ratio of optical density (OD) at 260 and 280 nm ( $A_{260/280}$ ) of all the RNA samples was approximately 2.0. The RNA stocks were then stored at -80°C.

### 2.3.3 Assessment of RNA quality

The Agilent 2,100 Bioanalyzer (Agilent Technologies, USA) was used to assess RNA quality according to the manufacturer's instructions. RNA 6,000 Nano LabChip Kit was used. RNA samples were diluted to 500 ng/ $\mu$ l. 2  $\mu$ L of RNA samples were submitted to Miss Simritpa Isihota, a laboratory technician to Professor John Arrand, School of Cancer Sciences, University of Birmingham. The RNA integrity number (RIN) software was used to automatically estimate the integrity of the RNA samples. Using this software, RNA integrity is not determined by the ratio of the ribosomal bands, but by the entire electrophoretic trace of the RNA sample. The output RIN is an integer number in the range of 1–10. An RIN of 1 is interpreted as completely degraded RNA samples while an RIN of 10 is an intact RNA sample. The RIN of my RNA samples was  $\geq 8$ . Figure 16 are examples of the Bioanalyzer results, which show electropherograms of blank sample, DNA ladder, and untransfected PHFK, P7 and P14 of PHFK HPV18, and HeLa cell lines. All q-PCR samples were run in triplicate and repeated three times. Expression of the endogenous Beta-actin control was always stable with changes in CT values within  $\pm 1$  SD. two-tailed paired nonparametric Mann-Whitney U test used to calculate fold change, which were considered significant if  $p \leq 0.05$ .

### 2.3.4 Complimentary DNA synthesis (cDNA)

#### 2.3.4.1 Reverse transcription of RNA

To analyse the expression of genes of interest the RNA was reverse transcribed to complimentary DNA (cDNA) using QuantiTect reverse transcription kit (Qiagen). The procedure was performed in a PCR hood. Working on ice, 2  $\mu$ l of 7 x gDNA Wipeout buffer was mixed with 500 ng of RNA sample made up to 12  $\mu$ l with RNase-free water in a 0.2 ml, PCR tube. The mixture was incubated for 2 min at 42°C. The tube was then quenched on ice

before a mixture of 4 µl of 5 x Quantiscript RT buffer, 1 µl of RT primer mix and 1 µl of Quantiscript Reverse transcriptase were added. The tube was then incubated for 30 min at 42°C followed by 3 min incubation at 95°C using a Px2 Thermal Cycler. All cDNA were then cleaned up before being stored at -20°C.

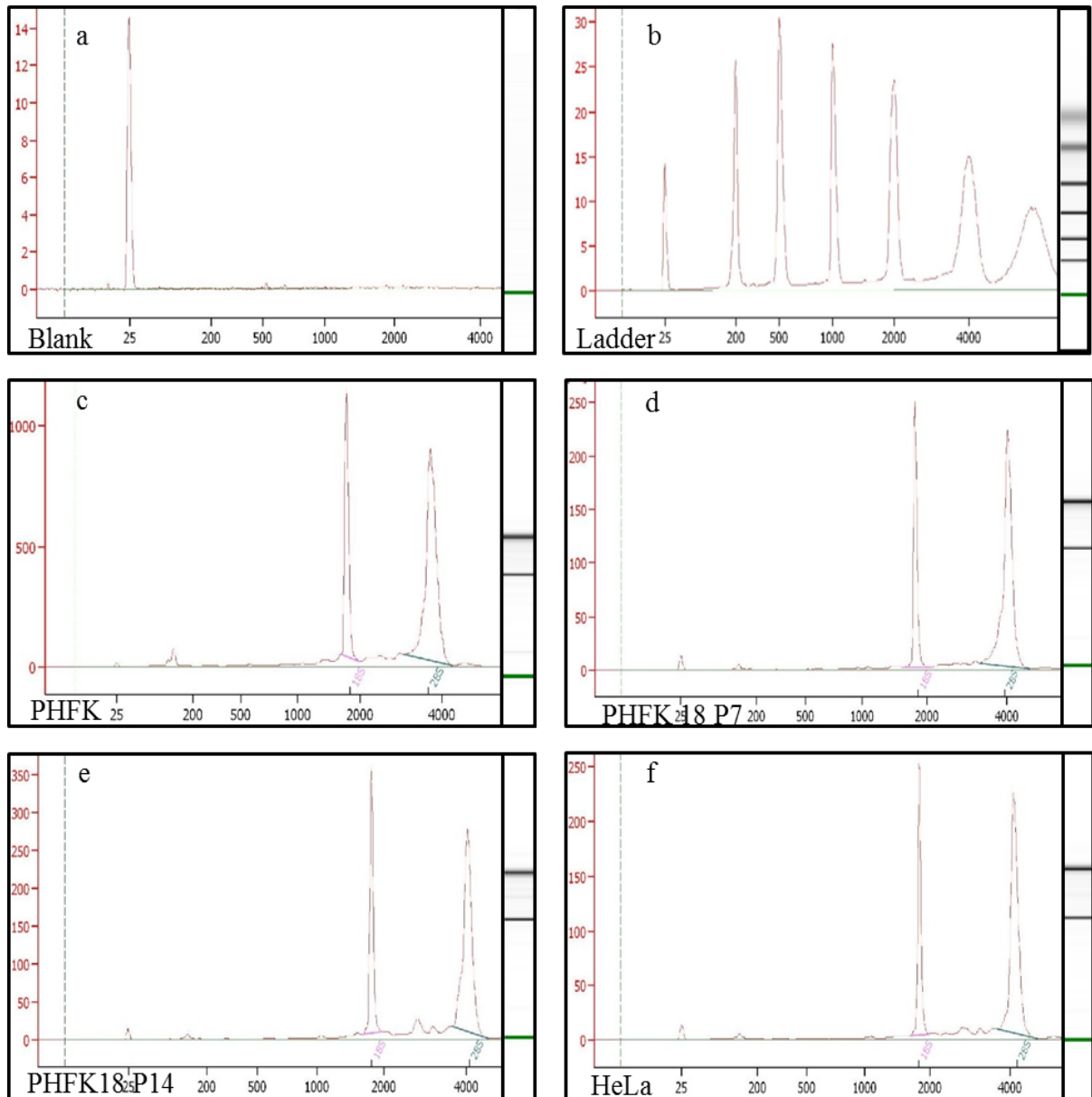


Figure 16: Bioanalyzer results demonstrate the high integrity and quality of RNA used for q-PCR experiments. Electropherograms present fluorescence on the y-axis, and the x-axis presents the run time in seconds. The two peaks of the 18S and 28S ribosomal RNAs are used to calculate RNA integrity number (RIN). Gel images from the Bioanalyzer are also shown at the right. a) A blank sample with no bands, b) a ladder with multiple bands of RNA fragments of known concentration. c) PHFK, RIN 9.6. d) PHFK HPV18 p7, RIN 9.9. e) PHFK HPV18 p14, RIN 10. f) HeLa, RIN 9.7.



### **2.3.4.2 Clean-up of complimentary DNA**

All cDNA were cleaned up before it was used to analyze of gene expression. I utilised MinElute Reaction Cleanup Kit (Qiagen) to clean up the cDNA. Firstly, each cDNA sample was transferred to a 1.5 ml tube. Before applying the sample to the MinElute column provided by the kit, 390 µl of buffer ERC was added and vortexed with the cDNA sample. The column was then centrifuged at 13,000 rpm for 1 min at room temperature. After discarding the flow-through, 750 µl of buffer PE was then pipetted directly to the column and centrifuged again. Prior to elute with 11 µl of DNase/RNase free water twice, the column was centrifuged with open lid at 13,000 rpm for 2 min to evaporate all ethanol in the buffer PE. The column was spined for 3 min at 13000 after elution and the concentration of clean single strand cDNA was measured using a NanoDrop ND-1,000 Spectrophotometer (Thermo Scientific) according to the manufacturer's instructions. The ratio of optical density (OD) at 260 and 280 nm ( $A_{260}/A_{280}$ ) of all the cDNA samples was required to be between 1.8 and 2.2. The purified cDNA was then stored at -20°.

### **2.3.4.3 Real Time - polymerase chain reaction**

Real time-PCR (RT-PCR) measures PCR amplification as it occurs. It is therefore more precise than standard PCR that measures the amount of accumulated PCR product during the plateau phase at the end of the PCR cycles. In contrast, RT-PCR data are collected during the exponential phase of the PCR. During this phase the quantity of PCR product is directly proportional to the amount of the present template. There are two types, semi-quantitative and quantitative PCR. The main difference between semi-quantitative and quantitative PCR is in the assay set up. The q-PCR requires a standard curve using a known number of target molecules in each step of the curve and real time cycler programme set as absolute quantification. I used this method to detect DNMT3B copy number in Section 2.2.3. Semi-

qPCR estimates gene expression relative to another transcript known to be stable and the program was set as relative quantification. RT-PCR using TaqMan chemistry requires a fluorogenic probe and specific PCR primers for the PCR reaction. The probe contains a quencher dye on the 3' end (TAMRA or MGB), and a reporter dye on the 5' end (FAM or VIC). The probes anneals downstream to one of the primer sites. Before the cleavage of the probe, the proximity of the quencher dye to the reporter dye decreases the amount of fluorescence emitted by the reporter dye. During PCR amplification, the 5' nuclease activity of Taq DNA polymerase separates the reporter dye from the quencher dye, which leads to increasing the reporter fluorescence signal. Hence this increase in fluorescence intensity is directly proportional to the number of PCR amplicons produced.

#### **2.3.4.4 Semi-quantitative real time-polymerase chain reaction**

Five ng of pre-cleaned cDNA was used per well; the total volume of reaction was 25 $\mu$ l, which was done in 96-well reaction plates (Applied Biosystems, US). All samples were run in triplicate to allow analysis by the delta CT method. Premixed primer sets and probe (Assay-on-demand, Applied Biosystems) were used to amplify target genes (Table 11). Primer details are shown in Table 12. Target genes and endogenous control gene (Beta-actin) were labelled with FAM and VIC respectively. The reaction was multiplex PCR, which means the target and endogenous primer genes were run at the same q-PCR plates but with different wells. DNase and RNase free water (Sigma-Aldrich) was used as a negative control and was run for each primer.

Table 11: Multiplex RT-PCR reactions volumes.

Target gene reaction	25µl(total volume)
cDNA	5ng
Taqman universal master mix	12.5µl
20 x primer & probe	1.25 µl
Water	X µl

After Taqman universal master mix was added to each sample the plates were then sealed with optically clear adhesive caps (Applied Biosystems). The plates then centrifuged at 2750 rpm for 5 min at room temperature to remove bubbles. The plate was then placed in ABI 7500 standard real-time PCR system according to the manufacturer's instructions (Applied Biosystems, Foster City, CA) and subject to 40 thermal cycles (Table 9).

Table 12: RT-PCR TaqMan® primer and probe kits. (Applied Biosystems). DNMT3B(2) assay from Qiagen.

Target gene	Probe fluorescent reporter dye	Product details
DNMT1	FAM	Hs00945899_m1
DNMT3A	FAM	Hs01003405-m1
DNMT3B(1)	FAM	Hs01003410_m1
DNMT3B(2)		PPH01054E
DNMT3L	FAM	Hs01081364_m1
EZH2	FAM	Hs00544830_m1
KDM6A	FAM	Hs00253500_m1
KDM6B	FAM	Hs00996325_g1

#### **2.3.4.5 Assay design and genes location of DNMTs**

DNMT genes are located on different chromosomes. In addition, alternative splicing results in multiple transcript variants encoding different isoforms. I chose probes to detect a maximum number of transcripts that were designed across exon-exon junctions. To increase the specificity of q-PCR experiments only probes with short amplicons were chosen. The DNMT1 gene is located on chromosome 19 on cytogenetic band p13.2 and has 97,942 bases. Two transcript splices encode 3 isoforms. I used a DNMT1 probe that spans the first exon and translated to two isoform a and b. The DNMT3A gene is located on chromosome 2 on cytogenetic band p23.3. Two isoforms of DNMT3A are produced by alternative splicing: isoform a and b. the DNMT3A protein localizes to the nucleus and the cytoplasm. The DNMT3A probe spans exons at the 3' end of transcript NM\_175629.2 and translates to isoform a, NP\_783328.1. The DNMT3B gene is located on chromosome 20q11.21 and has 46,972 bases. There are eight alternatively spliced transcript variants of the DNMT3B gene. The first DNMT3B probe (1) spans the exon at the 5' prime end of the NM\_001207055.1 transcript and translates to isoform 7, NP\_001193984.1. The second DNMT3B probe (2) of the NM\_006892.3 transcript translates to isoform1 NP\_008823.1, [Provided by RefSeq, May 2011]. The DNMT3-like (DNMT3L) gene is located on chromosome 21, cytogenetic band q22.3. It has 15,878 bases and translates into 2 isoforms. Figure 18 shows the DNMT3L probe location, which spans the exon-to-exon junction at the 3' prime end and translates into protein isoform 1.

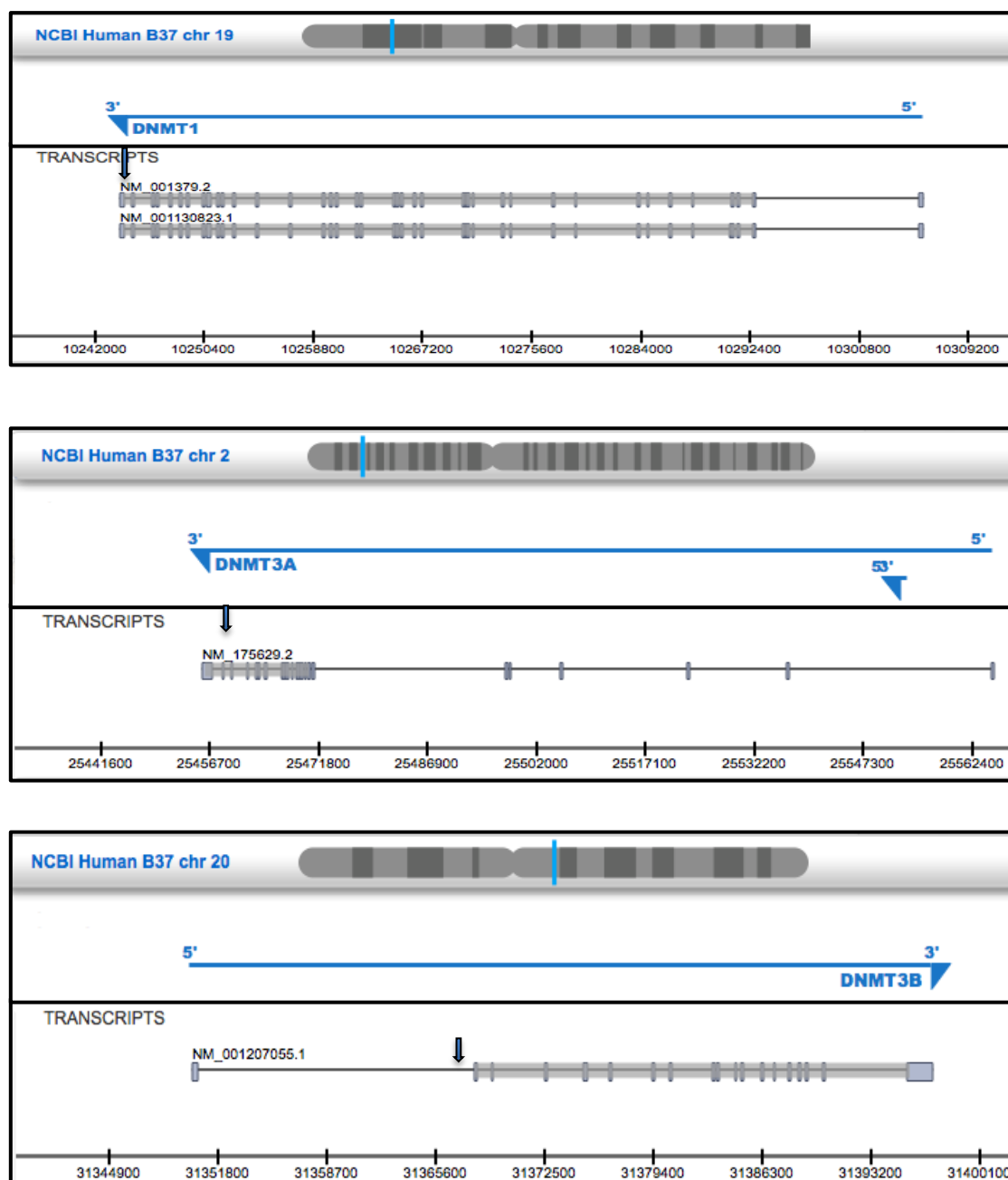


Figure 17: Genomic map showing DNA methyltransferase probes localization. Arrows show probe sites. The DNMT1 the probe spans the last exon at the 3' prime end of transcript NM\_001379.2. The DNMT3A probe spans the last exon at the 3' prime end of transcript NM\_175629.2. The DNMT3B probe spans the first exon at the 5' prime end of transcript NM\_001207055.1. This figure adapted from the Applied Biosystems website.

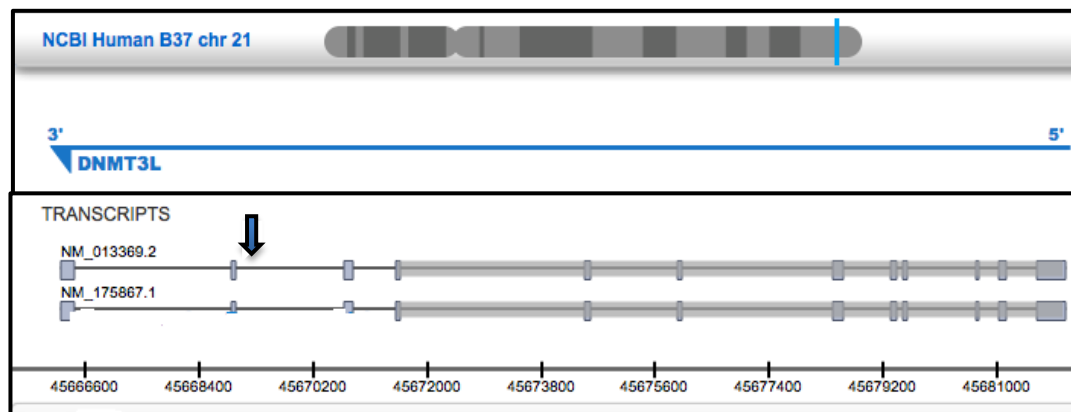


Figure 18: Genomic map showing DNMT 3-like (DNMT3L) probe and gene location. The arrow shows probe site. DNMT3L probe spans exons at the 3' prime end of transcript NM\_013369.2. This figure adapted from the Applied Biosystems website.

#### **2.3.4.6 Assay design and gene locations of H3K27me3 methyltransferase and the H3K27me3 demethylases**

Figure 19, illustrates the gene and probe localization of Enhancer of zeste homolog 2 (EZH2), Lysine (K)-specific demethylase 6 A (KDM6A) and Lysine (K)-specific demethylase 6 B (KDM6B). EZH2 is located at chromosome 7 q35-q36 with 76,940 bases. It can be alternatively spliced to give 5 isoforms. The EZH2 expression analysis probe spans the exon-to-exon junction of the NM\_152998.2 reference sequence and translates to isoform b, NP\_694543.1 (Figure 19) KDM6A and KDM6B are from the UTX family. KDM6A is located at chromosome Xp11.2 with 239,425 bases and encodes protein NP\_066963.2. The expression probe spans the exon-to-exon junction and translates to protein NP\_066963.2 (Figure 19). KDM6B is located at chromosome 17p13.1 with 14,897 bases. Two splice variants translate to isoforms a and b. the expression probe spans exons at the 3' end and translates to protein NP\_001073893.1 [Provided by RefSeq, May 2011] (Figure 19).

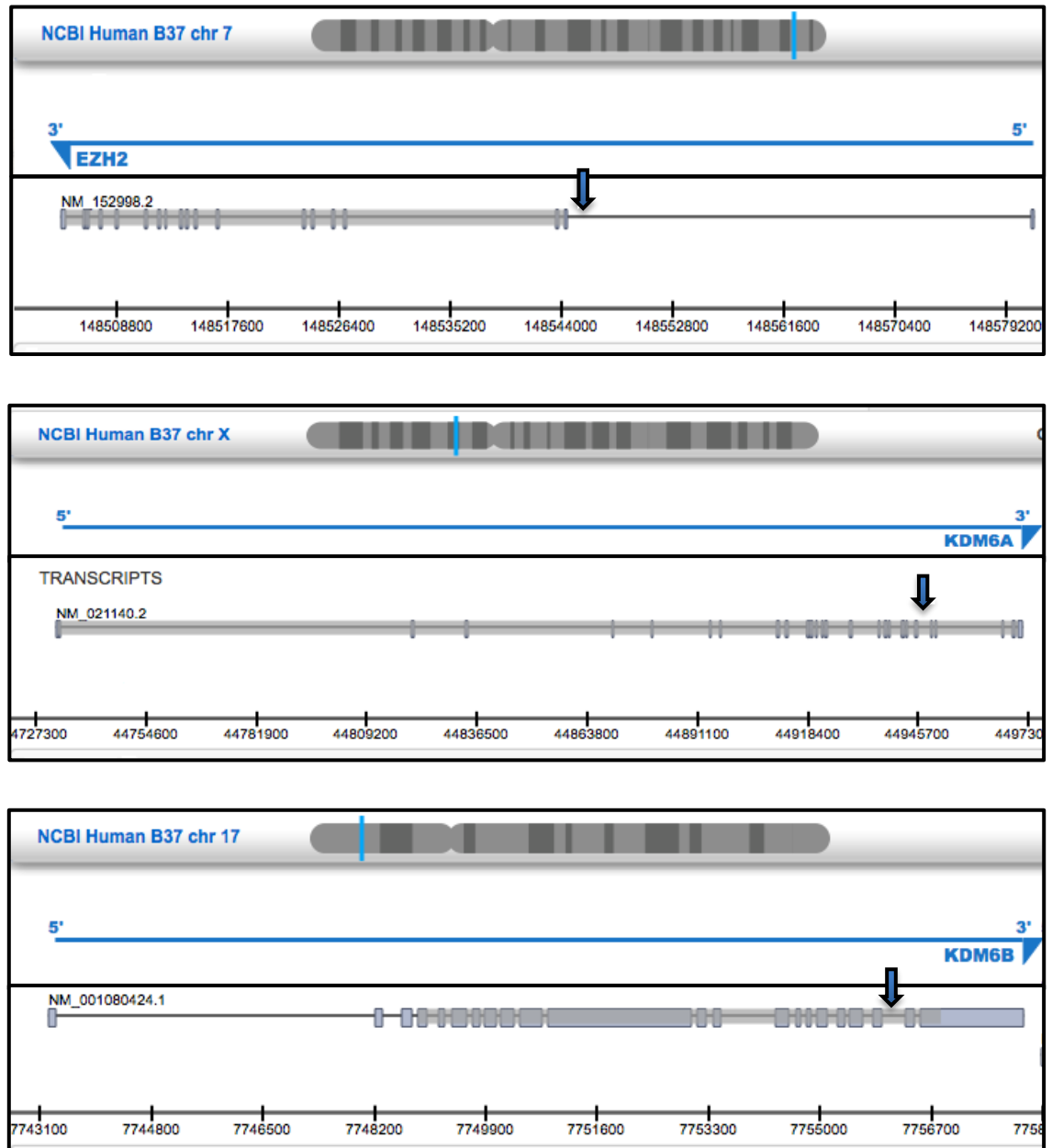


Figure 19: Genomic map showing H3K27me3 methyltransferase and demethylase gene and probe locations. The arrows show probe sites. The EZHZ2 probe spans the first exon at the 5' prime end of transcript NM\_152998.2. The KDM6A probe spans exons at the 3' prime end of transcript NM\_021140.2. KDM6B probe span exons at 3' primer end of transcript NM\_001080424.1. This figure is adapted from Applied Biosystems website.



**2.3.4.7 The comparative Ct method (ddCt) for semi-quantitation of gene expression**

The data were then analysed with 7,500 Fast System SDS software 1.3.1 (Applied Biosystems). The delta-delta ( $\Delta\Delta$ ) cycle threshold (Ct) method was used to calculate the fold change of target gene transcript or gene copy number in one cell type. Calculation of fold change is relative to an endogenous control that has been shown to have relatively constant expression or reference gene whose copy number is not affected by local genomic changes. The reference sample, which is the sample selected as a baseline for expression of target genes, is used to calculate the difference ( $\Delta\text{Ct}$ ) between the Ct values of target and housekeeping gene ( $\Delta\text{Ct} = \text{Target Ct} - \text{endogenous Ct}$ ). So  $\Delta\Delta\text{Ct}$  for each sample was set to a relative quantity (RQ) value of 1 and all other samples expressed as ratio of this. All RT-qPCR repeated three times and run in triplicate. To determine the statistical significance of differential gene expressions between two groups using RT-qPCR data, the gene expressions data were exported to GraphPad Prism, version 6.0h Software and calculation of expression changes was determined by two-tailed paired nonparametric Mann-Whitney U test. Expression changes were considered significant if  $p$  values  $\leq 0.05$ .

## **2.4 Protein detection and analysis**

### **2.4.1 Protein extraction**

Cells harvested for protein analysis were lysed by the addition of CytoBuster lysis buffer (Millipore). Typically, 50  $\mu$ l of lysis buffer with 2  $\mu$ L of 25 x Complete Protease Inhibitor Cocktail (Roche) was added to a cell pellet of  $1 \times 10^6$  cells. The cells were thoroughly mixed with lysis buffer and incubated for 5 min at room temperature. Cell lysis was then centrifuged at 13,000 rpm at 4°C for 5 min before transferring the cleared supernatant (cell extract) to a fresh tube and proceeding with analysis. Cell extract was then frozen at -20°C.

### **2.4.2 Measuring protein concentration**

Protein concentration was quantified using the Bio-Rad Protein Assay (Bio-Rad Laboratories). Five protein standards were prepared by diluting 1 mg/ml stock bovine serum albumin (BSA; Sigma-Aldrich Company Ltd., UK) to 0.1, 0.2, 0.3, 0.4 and 0.5 mg/ml concentrations with distilled water. An aliquot of each protein sample was diluted to 1:20 with distilled water. 10  $\mu$ l of each standard and diluted protein sample was plated out in duplicate into a 96-well plate (IWAKI). Bio-Rad Protein Assay Reagent was diluted to 1:5 in distilled water before 200  $\mu$ l was added to each well. The absorbance was read on a Bio-Rad 680 microplate reader at 595 nm. Protein concentration was then determined using a calibration curve, which was plotted using the five standards.

### **2.4.3 Western blotting analysis**

#### **2.4.3.1 Sodium dodecyl sulphate polyacrylamide gel electrophoresis**

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) gels were made in two phases; resolving and stacking gel. The concentration of the resolving gel was

determined according to the protein size. The prepared gel was poured immediately into mini-Protean 3 Bio-Rad apparatus. To give a uniform surface, isopropanol was pipetted onto the surface of the gel. Once the gel polymerised, the isopropanol was rinsed off with water. A 5 % stacking gel consisting of 30:1 acrylamide: bisacrylamide, 0.1% w/v SDS 0.1% w/v TEMED , 125 mM TRIS-HCL pH 6.8, and 0.1% w/v APS table was then poured on top and a suitable comb was placed and allowed the gel to polymerise at room temperature.

The appropriate protein concentrations of samples were diluted with 1:4 5 x Laemmli buffer (125 mM TRIS-HCl pH 6.8, 20% glycerol, 4% SDS, 10%  $\beta$ -mercaptoethanol, 0.004% bromophenol blue). The protein was then denatured by heating at 95-98°C for 5 min on a heat block and the sample contents were collected by brief centrifugation. The application of heat and SDS within the sample buffer ensures break down of di-sulphide bonds and disrupting secondary and tertiary peptide structures by applying a negative charge to the amino acids (Laemmli 1970). Following the collection of denaturised protein by brief centrifugations, samples were loaded into each well. To determine protein size, 10  $\mu$ l of See Blue Plus2 Pre-Stained Standard molecular weights were added to the first or last well. Gels were run in 1 x running buffer consisting of 30 g TRIS, 144 g glycine (Fisher), 10 g SDS in 10 L distilled water. Solubilised proteins were separated by electrophoresis at 120 V, 400 mA. The time was determined according to protein size.

#### **2.4.3.2 Protein transfer**

Proteins were transferred to a 0.45  $\mu$ m nitrocellulose transfer membrane (Protran BA85 membrane; Schleicher & Schuell UK Ltd., London, UK). Membrane, sponges and filter papers (Whatman® chromatography paper; Sigma-Aldrich Ltd., UK) were soaked for 5 min in transfer buffer (30 g Tris, 144 g glycine, 2 L methanol in 8 L distilled water) to ensure the complete surface was wet for the transfer step. A 'transfer sandwich' was set up into the X Cell II Blot

Module (Invitrogen Ltd., Paisley, UK), so that the gel and membrane were sandwiched and surrounded from both sides by two appropriately sized filter papers and three sponges. The transfer module was placed into the X Cell Sure lock Mini-Cell (Invitrogen Ltd., Paisley, UK), filled with transfer buffer and placed in ice to avoid overheating during the transfer. The blots were aligned to ensure protein moved from the negatively charged cathode core toward the positively charged anode onto the membrane. Proteins were transferred at 30 V for 2-3 hrs according to protein size, while the tank was placed on ice or in a cold room.

#### **2.4.3.3 Labelling of specific protein**

Prior to incubating the membrane with diluted primary antibody (Table 13), non-specific protein binding was blocked by incubating the membrane on a shaker in blocking solution (5% non-fat milk powder dissolved in Tween-tris buffered saline, T-TBS) for 1 hr at room temperature. T-TBS, consisting of 8.77 g NaCl and 1.21 g Tris in 1 L distilled water (pH adjusted to 7.6 using concentrated HCl) with 0.5 ml Tween 20 (Fisher Scientific). The membrane was incubated overnight at 4 °C in diluted primary antibody. After rinsing for 20 min in T-TBS with washes changed every 5 min, the membrane was incubated for 1 hr at room temperature in appropriate HRP-conjugated secondary IgG (DakoCytomation Ltd, Cambridgeshire, UK). Finally, the membranes were again rinsed as before ready for visualisation.

#### **2.4.3.4 Protein visualisation**

Antibody-protein complexes were detected using the enhanced chemiluminescence, ECL technique (Amersham Biosciences UK Ltd., Buckinghamshire, UK). Immediately prior to use, an equal amount of each ECL reagent was mixed in a universal tube. The membrane was

incubated in the ECL mixture for 1 min before being covered with clean film and placed immediately in a Hypercassette™ autoradiography cassette (Amersham Biosciences).

Table 13: Primary antibodies used in western blotting

Target protein	Approximate Protein size	Species	Primary antibody	Supplier	Dilution
DNMT1	183kDa	Mouse monoclonal	ab19904	Abcam	1:1,000
DNMT3A	102kDa	Mouse monoclonal	ab13888	Abcam	1:1,000
DNMT3B	110kDa	Rabbit polyclonal	HPA001595	Sigma	1:300
H3K27me3	17kDa	Rabbit polyclonal	ab6002	Abcam	1:1,000
KDM6A	154kDa	Rabbit polyclonal	SAB2102665	Sigma	1:400
KDM6B	177kDa	Rabbit polyclonal	ab85392	Abcam	1:300
EZH2	100kDa	Mouse monoclonal	07-689	Millipore	1:1,000
pRb	100kDa	Rabbit polyclonal	ab4788	Abcam	1:400
HPV18 E7	12kDa	Mouse monoclonal	ab100953	Abcam	1:1,000
HPV16 E7	20kDa	Goat polyclonal	sc-1587	Santa Cruz	1:100

#### 2.4.3.5 Autoradiography

In a dark room, a sheet of Hyperfilm™ (Amersham, Biosciences) was placed on top of the wrapped BioTrace NT membrane for the appropriate length of time. Hyperfilm™ was then developed in a Kodak X-OMAT 1,000 processor (Kodak Limited, UK). The actual bands sizes were observed to be different sometimes from the predicted proteins sizes. Almost all primary antibodies showed one band, but DNMT3B showed double band as per manufacture's blot, Sigma, Figure 20.

#### 2.4.3.6 Stripping membranes for reprobing with primary antibody

To facilitate re-probing of membranes with different primary antibodies after western blotting, Re-Blot Plus Mild Antibody Stripping Solution (Chemicon International, UK) was used. The membranes were washed in T-TBS for 20 min followed by gentle mixing in stripping solution for 15 min at room temperature. The membranes were further washed in T-TBS and

incubated for 1 hr in blocking solution (5% non-fat milk powder in T-TBS) before being re-incubated with primary antibody.

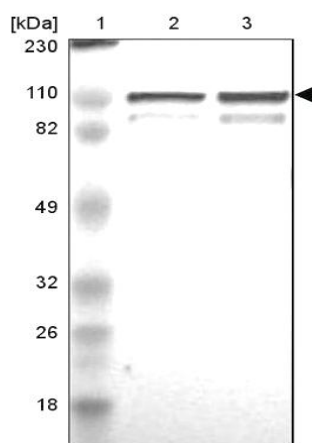


Figure 20: Western blot analysis showing DNMT3B protein expression represented by two bands at 110kDa using anti-DNMT3B antibody, the upper band represent DNMT3B protein Copied from, Sigma, Cat. No. HPA001595.

#### 2.4.3.7 Quantification of Western blot data

Signal intensities obtained from Western blot were quantified with ImageJ software (National Institutes of Health; <http://rsbweb.nih.gov/ij/>), as described previously (Medeiros et al., 2009; Chapman et al., 2006). At first images were transformed to 16-bit-type images and to eliminate background interference the threshold function was set to black and white. The rectangle tool was then used to draw the area for measurement surrounding the Western blot bands and the background subtracted. Lanes were then plotted and peak labelled. The proportion of the peak area, representing band, was calculated as a percentage of the total area under the curve and data was exported to Excel sheet, (Microsoft office package 7) for analysis. For data normalization, the control with the highest density value was used as normalizing control. The density values for both protein of interest and control of three independent western

blot experiments were divided by the highest density value to get a relative normalizing control value. Then, normalised protein of interest values were divide by the relative normalizing control values in their respective lanes. A two-tailed Students T-test was used to determine significant expression of protein of interest in any sample compared to control sample. Expression change was considered significant if p value of fold change was  $\leq 0.05$ .

#### **2.4.4 Cross-linked chromatin immunoprecipitation (X-ChIP)**

ChIP kit (Milipore, Temecula, California) was used according to the manufacturer's instructions. Two independent X-ChIP experiments were performed.

##### **2.4.4.1 Cross-linking**

All  $1 \times 10^6$  cells were resuspended in 500  $\mu$ l of ice-cold PBS containing protease inhibitors. Protease inhibitors were prepared with 1 mM phenylmethylsulphonyl fluoride (PMSF), 1  $\mu$ g/mL aprotinin and 1  $\mu$ g/mL pepstatin A. Protease inhibitors were added to PBS just prior to use because PMSF has a half-life of approximately 30 min in aqueous solutions. Native protein-DNA complexes were then cross-linked by adding formaldehyde at final concentration of 1% directly to PBS followed by gentle rotation at room temperature for 8 min. Cells were then pelleted by centrifuge for 4 min at 2,000 rpm at 4°C before being resuspended in 200  $\mu$ L of warm SDS Lysis buffer and incubated for 10 min on ice.

##### **2.4.4.2 DNA shearing**

Cells were sonicated in 200  $\mu$ L aliquots in ice lysate were sonicated using the Biorupter Sonicater (Diagenode) set to high power for 15 min twice with 30 sec off for passage 7 or passage 14 PHFK HPV18 cell line. The setting for HeLa cells was 10 min twice with 30 sec off. To visualise the shearing efficiency 4  $\mu$ l of supernatant was de-cross linked as shown in section 2.4.4.4 and the DNA was recovered by phenol/chloroform extraction. Orange G loading

dye with 1 µg of DNA per samples were then run in 1.5 % agarose gel. Immunoprecipitation was only performed if sonication was adequate. DNA shearing was required to be lengths between 200 bp and 1,000 bp.

#### **2.4.4.3 Immunoprecipitation (IP)**

Sheared lysates were pelleted for 10 min at 13,000 rpm at 4°C, and the supernatants transferred to new tubes. Sonicated cells were diluted by adding 1,800 µl of ChIP Dilution Buffer. One per cent of the diluted cell supernatant (~20 µL) was de-cross linked and kept to be used as the input material at the PCR. The diluted cells supernatant was then pre-cleared by adding 75 µL of Protein A agarose/Salmon Sperm DNA (50% Slurry), and rotated for 30 min at 4°C. After the collection of supernatant fraction, antibodies were added. To each sample, 2 µg of DNMT1 antibody (Abcam), DNMT3B antibody (Sigma) or H3K27me3 antibody (Abcam) was added to each tube. To serve as negative controls, samples were instead incubated with no antibodies or an irrelevant universal IgG/IgA rabbit antibody (Santa Cruz Biotechnology). Bead-antibody mixture was thereafter incubated on a rotator at 4°C for 2 hrs. The supernatant that contains unbound, non-specific DNA was then removed by gently pelleting agarose beads to 2,000 rpm at 4°C, 1min. The protein A agarose/antibody/protein complex was then washed with 1 mL of Low Salt, High Salt and LiClimmune complex wash buffer sequentially for 5 min once on a rotator. Lastly, it was washed twice with TE Buffer.

#### **2.4.4.4 Reversal of cross-links**

To de-crosslink, after capture the chromatin-antibody-bound beads, the supernatant was discarded, and 250 µl of freshly made elution buffer (0.1M NaHCO<sub>3</sub>, 1% SDS) was added to each tube. After a brief mix, I incubated tubes at room temperature for 15 min with rotation before re-incubation at 68°C overnight at 13,000 rpm with 50 µg/ml of proteinase K. The



immunoprecipitated DNA was then recovered by phenol/chloroform extraction and ethanol precipitation (Section 2.2.1.2). Following the clean-up, the pellet was then resuspended in 30 µl of nuclease-free water and stored at -20°C.

#### **2.4.4.5 Detection**

The immunoprecipitated DNA was then analysed by PCR. For PCR analysis, 50 ng of each immunoprecipitated DNA sample was amplified in a 25 µl volume using 5 pmol/µl primer mix (Alta Biosciences Ltd, Birmingham) and 12.5 µl of GoTaq Green Master Mix (Promega). Amplification was performed as described in as described in Section 2.2.2 for 30 cycles.

#### **2.4.4.6 Designing PCR primers for ChIP analysis**

Primer 3 software version 0.4.0 ([www.primer3.com](http://www.primer3.com)) was used to design ChIP primers. I designed all primers sequences spanning promoter regions of genes of interest. At least two consecutive regions starting from the upstream region of a gene of interest were designed. Primers were designed based on the following criteria: product size 100-200 bp in length, primer size between 18-24 bp and primer annealing temperature of ~58°C and primer size between 18-24 bp to be no more than between 18-24 bp in size and have an annealing temperature of ~58°C. Products were all 100-500 bp in length. All primers were ordered from (AltaBiosciences, University of Birmingham, UK) and then optimised and validated. Lists of primer sequences, their products' size and annealing temperatures are shown in Table 14 and Table 15.

Table 14: List of primers designed at promoter regions of Slit-Robo pathway genes, used for X-ChIP analysis. (fw = forward primer; rev = reverse primer).

Primer	Sequence 5'-3'	Products	A.Tm
ROBO1 reg I fw	GCGCACACTTGGACCTATTT	225 bp	56°C
ROBO1 reg I rev	AACCAAGCGAAAGGTGACAG		
ROBO1 reg II fw	CAGACAATTAAGCAGCTATGGA	220 bp	56°C
ROBO1 reg II rev	AACACTAGGCACGCAGCTTT		
ROBO2 reg I fw	AGCCTGGAGGAGCCACCTAT	287 bp	56°C
ROBO2 reg I rev	GTCCTGGAGAGCCGAGAGT		
ROBO2 reg II fw	ACTCCAGGCAGCTCTCATTC	300 bp	56°C
ROBO2 reg II rev	AAAGGGAGCGCTCAGTCC		
ROBO2 reg III fw	GCCTGGTGCACCTATCCTCAG	300 bp	56°C
ROBO2 reg III rev	TCTCCGGTTTCCAAAAGAGA		
ROBO3 reg I fw	GAACTTGTTTCGCGGACTCTC	300 bp	54°C
ROBO3 reg I rev	CCGTTGAGAGAGGGATCG		
ROBO3 reg II fw	GATATGGGATCCTGGGATGG	300 bp	56°C
ROBO3 reg II rev	ATGTGCTCGAAGCTCCAAAT		
ROBO3 reg III fw	CCTTAATTACGCCGTGATCG	300 bp	52°C
ROBO3 reg III rev	CCGTAAGCCCCTCCACAG		
SLIT1 reg I fw	ACACCCCCTTCCCTCTTCT	262 bp	56°C
SLIT1 reg I rev	TCCATCAATCGAAAAGCACA		
SLIT1 reg II fw	CGTGTGTGTGTGTGTGTGGTGT	300 bp	56°C
SLIT1 reg II rev	GCACCTTGCTCCTCCAAG		
SLIT2 reg I fw	CCAAGTTCATCCTTGGGAGA	300 bp	56°C
SLIT2 reg I rev	AAGGCAGTAGAGCCCACTCA		
SLIT2 reg II fw	CGTGAGTGAGCAGAGTCCAG	300 bp	56°C
SLIT2 reg II rev	TATTATGGCCCCGATCAAAG		
SLIT3 reg I fw	GGTCGCTCCAGCAAAGTTC	321 bp	56°C
SLIT3 reg I rev	GCTCGCTCTCTCCATTCACT		
SLIT3 reg II fw	AGTGCTCAGCTCAGTGCAAA	300 bp	56°C
SLIT3 reg II rev	AGGGAGGAGGCAGTAGAGGA		
SLIT3 reg III fw	CTCCTCCGCGCTAACTCC	399 bp	54°C
SLIT3 reg III rev	TGGAGGCCCACTCAGGAC		

Table 15: List of primers designed at promoter regions of tumour suppressor genes, used for X-ChIP analysis. \*Primer sequences of HOXA7E and HOXAF obtained from (Wu et al., 2008). Primer sequence of GAPDH obtained from (Collins et al., 2009). (fw = forward primer; rev = reverse primer).

Primer	Sequence 5'-3'	Products size	A.Tm
SOX11 I fw	GTGCCGAGGACTTTGCAACT	300 bp	56°C
SOX11 I rev	GTTGCTCTCCGCTTCCAA		
SOX11 II fw	GGCCTCCGAGTTAATAAAGGA	300 bp	56°C
SOX11 II rev	GTGGCTCCTACCGCAAGC		
SOX11 III fw	GACCCGGTCCTCGGTTAG	300 bp	58°C
SOX11 III rev	GGGGGAGAGGAGGCATGT		
ESR I fw	CAGCCTCCTGTCTACCGACT	200 bp	56°C
ESR I rev	GGAGTACCTACTTGAAGAGAAGA		
ESR II fw	CAGCCTCCTGTCTACCGACT	225 bp	56°C
ESR II rev	GAGTACCTACTTGAAGAGAAGATT		
CALCA I fw	CTCTGATCCAAGCCACCTC	150 bp	54°C
CALCA I rev	GGAGGAAAGGGAGTGAGAAAA		
CALCA II fw	CTCTGATCCAAGCCACCTC	200 bp	56°C
CALCA II rev	GAGAGGAGGAAAGGGAGTGAG		
CALCA III fw	CAAATCTAGGCCCCGTCAG	218 bp	52°C
CALCA III rev	TTCACAAATGACAGCATTGACA		
CDKNA2 I fw	GCCAACGCTGGCTCTGGCGA	278 bp	54°C
CDKNA2 I rev	AAGTCGCCCCAGGTTGGGTCT		
CDKNA2 II fw	CCAACGCTGGCTCTGGCGAGG	380 bp	58°C
CDKNA2 II rev	CCTGAAGTCGCCCCAGGTTGG		
CDH1 I fw	AACAAAAGAACTCAGCCAAGTG	200 bp	56°C
CDH1 I rev	ACGCCACTGAGAGGGGGTGC		
CDH1 II fw	CCCTCTCAGTGGCGTCGGAAC	300 bp	56°C
CDH1 II rev	AGACCTGCAGCAGCAGCAGCA		
HOXA7 reg E fw *	CTGGTGCAACATGGTGCAG	523 bp	52°C
HOXA7 reg E rev	CTGGTGCAACATGGTGCAG		
HOXA7 reg I fw*	GCCGACTTCTTGCTCCTTTG	494 bp	56°C
HOXA7 reg I rev	CTCCACGCAATGGCGCC		
GAPDH fw*	GCTCAAGGGAGATAAAATTC	158 bp	58°C
GAPDH rev	CGACCAAATCTAAGAGACAA		

#### **2.4.4.7 ChIP analysis**

ChIP PCR data needs to be normalised for sources of variability, including method of DNA purification, amount of chromatin used and efficiency of immunoprecipitation.

To normalise ChIP PCR data the ChIP signals are divided by input signals. The input sample represents certain per cent of total starting chromatin used in ChIP. Thus, fold enrichment calculated as ChIP/Input DNA. In these ChIP experiments 10% of starting chromatin used as input. It has been shown that densitometric analysis with the ImageJ software of PCR results qualitatively matched the real-time PCR results (Chapman et al., 2006; Medeiros et al., 2009). It therefore offers a viable alternative due to its practicality, simplicity and less resources and time used in preliminary calibration experiments.

##### **2.4.4.7.1 Densitometric analysis**

PCR Images were saved as TIFF files for analysis with ImageJ. Signal intensities obtained from PCR data of ChIP assays and PCR data from whole cell lysates (Input DNA) were quantified with ImageJ software (National Institutes of Health; <http://rsbweb.nih.gov/ij/>), as described previously (Medeiros et al., 2009; Chapman et al., 2006). At first images were transformed to 16-bit-type images and to eliminate background interference the threshold function was set to black and white. The rectangle tool was then used to draw the area for measurement surrounding the PCR bands and the same tool was also used to move to each band in the blot area. Lanes were then plotted and peak labelled. The proportion of the peak area, representing band, was calculated as a percentage of the total area under the curve and used for statistical analysis. Standard error of mean of two independent experiments were calculated.

## **2.4.5 Co-immunoprecipitation (Co-IP)**

### **2.4.5.1 Beads preparation**

100  $\mu$ l of protein agarose bead slurry (Sigma-Aldrich) was transferred to 1.5 ml Eppendorf tubes. For manipulation of the agarose beads, cut tip pipette tip were used to avoid disruption of the beads. The beads were then washed with PBS buffer three times before centrifugation at 2,000 rpm for 1 min. supernatant then aspirated and the beads were re-suspended in PBS buffer as 50% slurry.

### **2.4.5.2 Pre-clean**

Prior to immunoprecipitation 500  $\mu$ g of protein from each sample equalised with cold PBS buffer to 500  $\mu$ l and pre-cleared by adding 100  $\mu$ l of agarose bead-PBS slurry. After incubating on a rotating wheel at 4°C for 1 hr, the beads were pelleted by centrifuge for 1 min at 2,000 rpm. The supernatants were transferred to fresh Eppendorf microcentrifuge tubes and the bead pellets discarded.

### **2.4.5.3 Immunoprecipitation (IP)**

To each sample, 2  $\mu$ g of DNMT3B antibody (Sigma) was added. To serve as a negative control, lysate was instead incubated with an irrelevant universal IgG/IgA rabbit antibody (Santa Cruz, Biotechnology). Samples were then incubated with the antibody overnight on the rotating wheel at 4°C. The protein-complex was captured by adding 100  $\mu$ l of the pre-prepared 50% agarose bead slurry to each sample and incubated for 2 hrs on the rotating wheel at 4°C. The bound immune-complexes were then pelleted by being centrifuged for 1 min at 2,000 rpm and the supernatant were discarded. The beads were then washed three times with cold PBS buffer as described above. After the final spin the beads were resuspended in 50  $\mu$ l of 2x SDS-PAGE buffer before being boiled for 5 min at 95°C to elute the immune complexes from the beads.

The beads were then pelleted as above and supernatants was collected and subjected to SDS-PAGE and immunoblotting first with anti-DNMT3B to confirm success of IP and then with anti-HPV18 E7 or anti-HPV16 E7 (Table 13).

## **2.4.6 Immunocytochemistry**

### **2.4.6.1 Cytospin preparation**

After the cells were harvested and counted,  $2 \times 10^6$  cells were resuspended in 1 ml of PBS with 10% formal-saline (Genta Medical, York, UK) and stored at 4°C. Adhesive slides (Surgipath Europe, UK), cardboard filter and Cytoclips™ (Thermo Electron) were assembled with Cytofunnel® disposable sample chambers, according to the manufacturer's instructions and placed into the Shandon Cytospin centrifuge (Thermo Electron). 100 µl of fixed cells were aliquoted into the appropriate wells of Cytofunnels®. Following centrifugation at 1,000 rpm for 5 min, monolayers of cells were formed. The filters and Cytoclips were carefully removed and the slides were air dried for 10 min.

### **2.4.6.2 Blocking of endogenous peroxidase activity and antigen retrieval**

Endogenous peroxidase activity was blocked for 15 min by incubating slides in 3% hydrogen peroxide (Sigma-Aldrich, UK) in distilled water. The slides were then rinsed thoroughly in running tap water.

### **2.4.6.3 Agitated low temperature epitope retrieval (ALTER)**

The slides were incubated in 1 mM EDTA (pH 8.0), Tween-20 (0.1%) on a hotplate at 65°C overnight. A magnetic bar with the stirrer set to 600 rpm was used to achieve agitation. The slides were then cooled in running tap water for 5 min and then in 0.1% PBS-Tween-20 pH 7.6.

#### **2.4.6.4 Detection of antigen**

To minimise evaporation, all slides were placed in a metal microscope slide staining tray (Richardsons of Leicester Ltd, UK) and covered with a lid. The samples were washed with 0.1% PBS-Tween-20 for 5 min prior and after incubation with primary antibody. To each slide 100 µl of 1:1,000 DNMT1(Abcam) or 1:1,500 DNMT3B (Sigma, Aldrich) primary antibody diluted in 0.1% PBS- Tween-20 were applied and incubated for 1 hr at room temperature. Per each passage, three slides were prepared for DNMT1 or DNMT3B. For secondary detection, 2 drops of DAKO Envision universal secondary antibody (Dako, UK) were applied to each slide and incubated for 30 min at room temperature. This was followed by washing in PBS for 5 min.

#### **2.4.6.5 Visualisation and counterstaining**

Visualisation was carried out using the ImmPACT diaminobenzidine (DAB) (Vector Laboratories). To each slide, 100 µl of substrate system was applied for 30 sec to 1 min, which converted to an insoluble brown product. The slides were rinsed with distilled water and counterstained with Mayer's haematoxylin (Sigma-Aldrich) for 10 sec and followed by a further 5 min wash under running tap water. The slides were then dehydrated by placing in IMS and xylene for 5 min each before being mounted under cover slips with a drop of dibutyl phthalate xylene (DPX), mounting medium (Invitrogen) for microscopic examination.

#### **2.4.7 Direct measurement of DNA methyltransferase enzymatic activity**

The DNMT1 and DNMT3B enzymatic activity was measured using DNMT Direct Activity Kit (BPS, Bioscience). The kit provided a 96-well plate pre-coated with DNMT substrate. Cell lysate was prepared and measured with Bradford assay as described in sections 2.4.1 and 2.4.2.

#### **2.4.7.1 Determining optimal required protein concentration**

A range of different protein concentration of PHFK was used to optimise standard required protein concentration. Nuclear extracts of 0.5, 1, 5 and 10 ng were tasted against negative and positive control. Five ng of protein was considered the optimal concentration to activate DNMT activity.

#### **2.4.7.2 DNMTs enzymatic activity assay protocol**

Following hydration of the microwells by TBST buffer (1 x TBS, pH 8.0, containing 0.05% Tween-20), master mixture was added to designated wells as shown in (Table 16). The plate then incubated for 1 hr (DNMT1) and 2 hr (DNMT3B) before washing three times with TBST and adding 100 µl of blocking buffer to every well. Following shaking for of 10 min, the supernatant was removed by washing with TBST and anti-5-methylcytosine antibody was added and incubated for 1 hr at room temperature. Following the washing and incubation of the plate with blocking buffer as before, 1:1,000 diluted secondary HRP labelled antibody was added and left for 30 min at room temperature. A freshly prepared mixture of chemiluminescence was added to each well.

#### **2.4.7.3 Reading Chemiluminescence**

Chemiluminescence is the emission of light that results from a chemical reaction. Chemiluminescence detection does not require wave length selection as the method used is emission photometry and is not emission spectrophotometry. Immediately following the addition of chemiluminescence mixture the plate was read in a luminometer (Orion L Microplate, Geneflow Ltd, UK) according to the manufacturer's instructions. The plate reader was set for luminescence mode.



#### 2.4.7.4 Calculation of DNMT activity

All data were presented as relative luciferase units (RLU). Blank value was subtracted from all readings. All reactions were run in duplicate and repeated twice. Therefore, the calculations of enzymatic activities for each sample were an average of four readings. The following formula was used:

$$\text{DNMT activity} = \frac{\text{Average sample} - \text{Average blank}}{\text{Protein amount } (\mu\text{g}) + \text{incubation hr.}} \times 1,000$$

The difference of DNMT enzymatic activity between the mean of each sample was calculated. A t-test was used to obtain significant enzymatic activity changes. Changes in enzymatic activity were considered significant if the *p* value of the fold change was  $\leq 0.05$ .

Table 16: Designated reaction mixtures of DNMTs direct enzymatic activity assay

	Blank	Positive control	Substrate control	Test sample
DNMT(10 ng/ $\mu$ l)	-	20 $\mu$ l	20 $\mu$ l	20 $\mu$ l
DNMT assay buffer	12.5 $\mu$ l	12.5 $\mu$ l	12.5 $\mu$ l	12.5 $\mu$ l
400 $\mu$ M S-adenosylmethionine	5 $\mu$ l	5 $\mu$ l	-	5 $\mu$ l
Test activator	-	-	-	x $\mu$ l
H <sub>2</sub> O	32.5 $\mu$ l	12.5 $\mu$ l	17.5 $\mu$ l	12.5 $\mu$ l- x $\mu$ l
Total	50 $\mu$ l	50 $\mu$ l	50 $\mu$ l	50 $\mu$ l

## CHAPTER 3

# RESULTS I

# **HUMAN PAPILLOMAVIRUS ASSOCIATED WITH EXPRESSION CHANGES TO KEY EPIGENETIC REGULATORS**

**Aim of Chapter 3.**

To investigate changes in expression of epigenetic regulators associated with high-risk human papillomavirus.

**Objectives:**

1. To profile the expression changes of DNA methyltransferases, DNMT1, DNMT3A, DNMT3B and DNMT3L.
2. To profile the expression changes of the H3K27me3 methyltransferase, EZH2.
3. To profile the expression changes of the H3K27me3 demethylases, KDM6A and KDM6B.

The expression of these epigenetic regulators will be analysed at RNA and protein level using q-PCR and Western blot respectively, in a spectrum of cell lines to recapitulate cervical neoplastic disease:

- a. Primary human foreskin keratinocytes (PHFK) transfected with HPV16 and HPV18.
- b. Human cervical intraepithelial neoplasia cell line.
- c. Cervical cancer cell lines.

DNA methyltransferases, H3K27me3 methyltransferase and H3K27me3 demethylases are epigenetic regulators that control two important marks in cancer: DNA methylation and Tri methylation of H3K27me3 (H3K27me3me3). In this chapter I first describe the *in vitro* models that I have used in this thesis. Next, I present expression profiles of the DNA methyltransferases, DNMT1, DNMT3A, DNMT3B and DNMT3L at the RNA, protein and cellular levels in a spectrum of cell lines. I also demonstrate cytopathic changes of *in vitro* models. Finally, I show expression profiles of H3K27me3 methyltransferase and H3K27me3 demethylases in a spectrum of cell lines.

## 3 Results I

### 3.1 Introduction of Results I

Studies of cancer epigenome (Łuczak et al., 2012; Su et al., 2013) have highlighted the altered expression of DNMTs in advanced primary cervical cancer and cancer cell lines. This has raised many questions such as whether these changes apply to other key epigenetic regulators such as H3K27me3 methyltransferases and demethylases? Also does HPV status and type make a difference to the expression of these epigenetic regulators? Finally, do changes found early in precancerous cells and then inherited by neoplastic cells in the process of cervical carcinogenesis? To answer these questions I first investigated the expression changes of these epigenetic regulators, at the RNA level. Next I investigate the altered expression at the protein and cellular levels. *In vitro* HPV-transformed cell lines share complementation groups with cell lines derived from cervical cancers. It has been shown that common genetic defects are involved in both *in vitro* models and cervical carcinogenesis (Seagon, Dürst and Matthiasdärst, 1994). Therefore, the functional data obtained from HPV-transfected keratinocytes and cervical cancer cell lines can provide a progression model for cervical carcinogenesis whereby data can be mapped to the different stages of cervical neoplasia (Snijders et al., 2006).

### 3.2 *In vitro* experimental model system

In this study I used a progression model (Figure 21) that involves primary human foreskin keratinocytes (PHFK) *in vitro* model system in conjunction with W12 cell lines (human cervical low grade lesion) and different human cervical cancer cell lines: HeLa, SiHa, Ca Ski and C33a (Table 5) to mimic the multistep process that is required for the progression of HPV-induced cervical tumorigenicity from initial epithelial infection to dysplastic cellular transformation. The PHFK cell lines were obtained from the same donor. The PHFK HPV18 cell line was propagated over 10 passages, from passage 4 to passage 14. Two variants of HPV16-transfected cell lines; the PHFK 16K and PHFK 16B cell lines both had 7 passages, from passage 4 to passage 10 (section 2.1.5.2). A schematic representation of these cell lines is shown in Figure 22. Furthermore, I established another PHFK 16B cell line from the same donor, PHFK HPV16B (2). I also re-cultured the PHFK HPV18 cell line from passage 11 to passage 14 to confirm the physical state of HPV and for the methylation arrays. As a control, I used passage 3 PHFK that represented an untransfected cell line obtained from the same donor.

Different cell lines have different HBV status. W12 passage 11 represents an early passage that contains ~600 copies of HPV16 per cell in its episomal form. W12 passage 57 represent a late passage, containing ~2 copies of integrated HPV16 per cell (Stanley et al., 1989; Alazawi et al., 2002; Gray et al., 2010). Different cancer cell lines; HeLa cell lines with integrated HPV18 (Adey et al., 2013), two cell lines with HPV16; SiHa cell line, with two copies of integrated HPV16 per cell and Ca Ski cell line that harbours 600 copies of integrated HPV16 genomes per cell and retained E2 (Baker et al., 1987; Theelen et al., 2010). Lastly, C33A, an HPV-negative cervical cancer cell line.

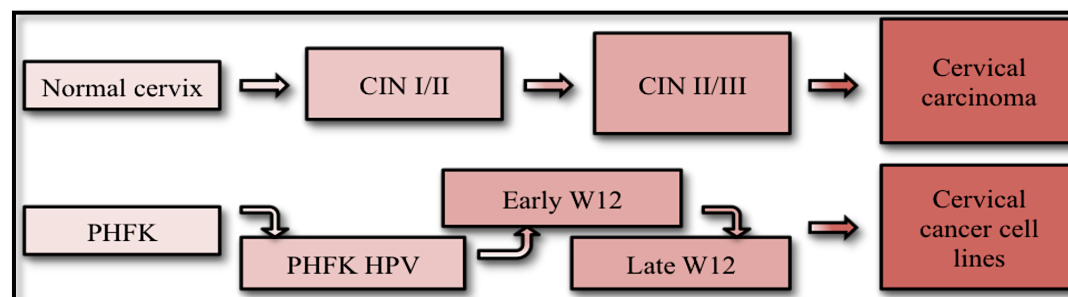


Figure 21: Alignment of the different stages of cervical carcinogenesis (upper part): normal cervix, CIN I /II/III and cervical carcinoma with the *in vitro* progression models of cervical cancer (lower part): PHFK, PHFK transfected with HPV, early and late W12 cell line (human cervical low grade lesion) and cervical cancer cell lines.

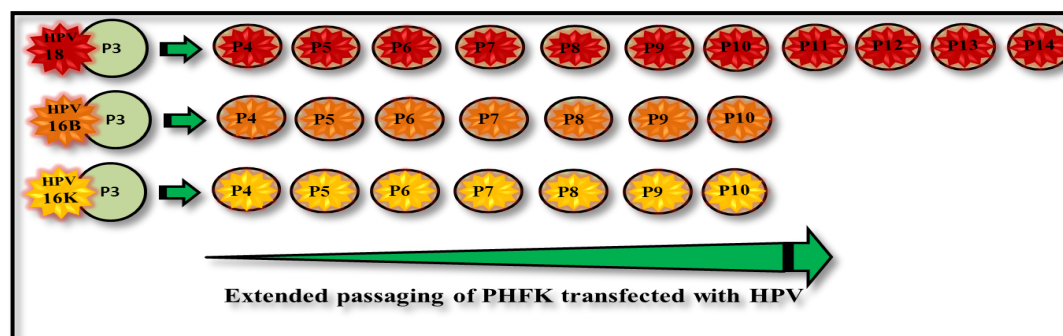


Figure 22: *In vitro* model system. P3 represents untransfected PHFK. The PHFK HPV18 cell line was propagated from passage 4 up to passage 14. PHFK transfected with two variants of HPV16: HPV16B cell line and the HPV16K cell line. The later two cell lines were propagated from passage 4 to passage 10.

### **3.3 Expression changes of DNA methyltransferases in *In vitro* experimental models system**

DNMT alternative splicing results in multiple transcript variants encoding different isoforms. Illustration of the different isoforms of DNMTs and maps of the DNMTs genes, those show the isoforms positions and highlight the transcripts covered by each probe in chapter 2 (Section 2.3.4.5). DNMT1 has two transcript splices encode 3 isoforms. I used a DNMT1 probe that translated to two isoform a and b. Two isoforms of DNMT3A are produced by alternative splicing: isoform a and b. I used DNMT3A probe that translates to isoform a. There are eight alternatively spliced transcript variants of the DNMT3B gene. The first DNMT3B probe (1) translates to isoform 7. The second DNMT3B probe (2) is specific to isoform1. DNMT3L gene translates into 2 isoforms. I used DNMT3L probe that translates into protein isoform 1. [Provided by RefSeq, May 2011]. I investigated the expression of DNMTs in *In vitro* experimental model system. I used HPV-negative C33A cancer cell line as control when expression of DNMTs level examined in cancer cell lines. I compared early W12 that harbour HPV16 episomal form with late passage of W12 with integrated HPV. I used PHFK as control for HPV transfected PHFK cell lines.



### 3.3.1 Expression profile of DNMTs in cervical cancer cell lines

First I studied the expression changes in mRNA expression using q-PCR in 4 cervical cancer cell lines, C33A, HeLa, SiHa and Ca Ski. All q-PCR experiments were run in triplicate and repeated three times. Two-tailed paired nonparametric Mann-Whitney U test was used to calculate significant fold changes with changes considered significant if  $p \leq 0.05$ . My results show differential expression of all the DNA methyltransferases tested. All cervical cancer cell lines show overexpression of DNMT1 compared to C33A cervical cancer cell line (Figure 23a). I demonstrate changes in the expression of *de novo* methyltransferases, DNMT3A and DNMT3B (Figure 23b &c). The HeLa cell line overexpressed DNMT3A compare to C33A. However, DNMT3B was not expressed at significantly higher levels in HeLa cells. In contrast, the Ca Ski cell line shows the opposite pattern since Ca Ski overexpressed DNMT3B compared to C33A cells and showed little change in the level of DNMT3A expression. On the other hand, the SiHa cell line, expressed the *de novo* methyltransferases at higher levels than C33A cells. I also used cervical cancer cell lines as a positive control for studying the PHFK *in vitro* model at a protein level in section 3.3.3.2 and 3.3.4.2 and I also examined DNMTs at a cellular level in section 3.3.6. Interestingly, DNMT3L is not expressed in SiHa cells and is expressed at very low level in HeLa and Ca Ski cells. However, C33A expresses a remarkably high level of DNMT3L mRNA (Figure 24).

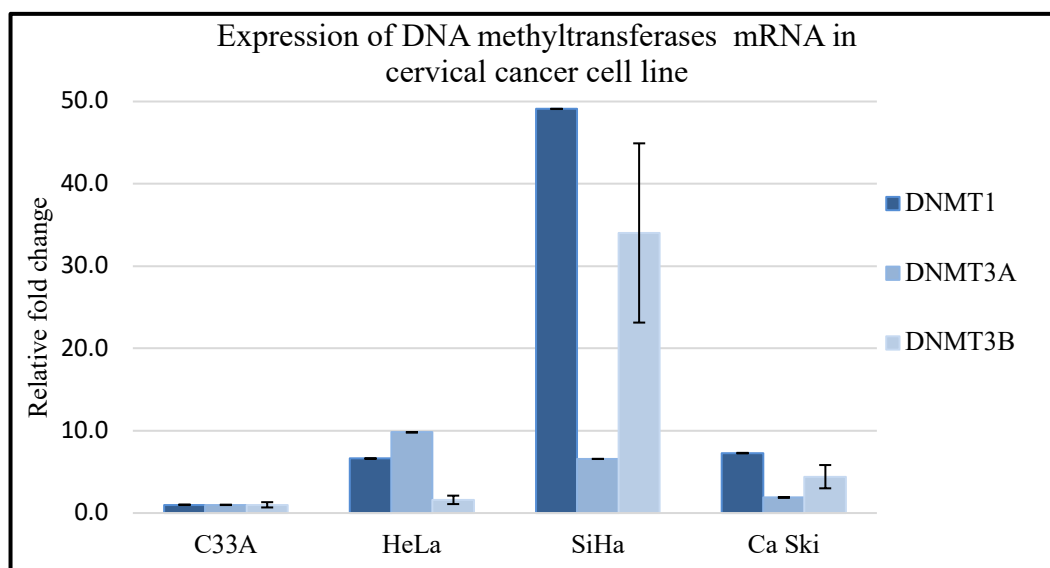


Figure 23: q-PCR results demonstrating expression of DNA methyltransferases mRNA level relative to Beta-actin as a reference gene in HPV-positive cancer cell lines versus HPV-negative C33A cancer cell line. a) Overexpression of DNMT1 in HeLa and SiHa and Ca Ski compared with C33A. b) overexpression of DNMT3A in HeLa and SiHa compared with C33A. c) SiHa and Ca Ski show a high level of DNMT3B expression. Results shown are typical of 3 similar experiments. Each experiment was run in triplicate. Error bars indicate standard deviations.

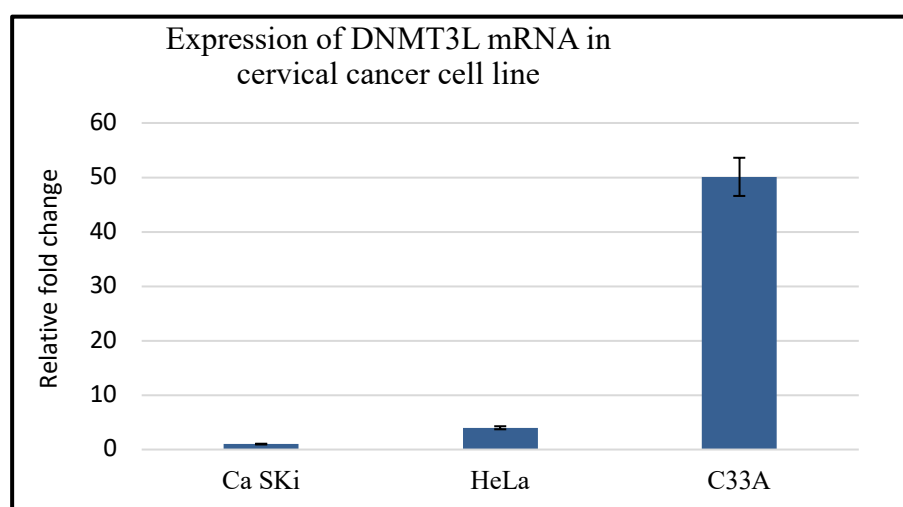


Figure 24: q-PCR results demonstrating differential expression of DNMT3L mRNA level relative to Beta-actin as a reference gene in the cervical cancer cell lines, SiHa, HeLa and C33A versus Ca Ski (control). Both Ca Ski and HeLa showed very low levels of expression. No expression was observed in the SiHa cell line (Data not shown). C33A cells showed very high levels. Results shown are typical of 3 similar experiments. Error bars indicate standard deviations.

### 3.3.2 Change in DNA methyltransferase expression in the W12 cell line

Next, I investigated the W12 cell line. I studied changes in expression using q-PCR. All q-PCR experiments were run in triplicate and repeated three times. Two-tailed paired nonparametric Mann-Whitney U test was used to calculate significant fold changes with changes considered significant if  $p \leq 0.05$ . As only a small number of W12 cells from an early passage were successfully harvested. Although this was enough to extract DNA and RNA, I was unable to get enough whole cell extract to investigate the protein expression of all DNMTs. However, I did examine DNMT3B protein expression level and was able to repeat this experiment three times. Figure 26 shows the DNMT3B with overexpression of DNMT3B protein in late W12 (W12 p57) compared to early W12 (W12 p11). Calregulin was used as a loading control and shows constant level. By analysing mRNA I found that all DNMTs were upregulated in W12 p57 compared to early W12, except DNMT3L, which was not expressed in W12 (Figure 25).

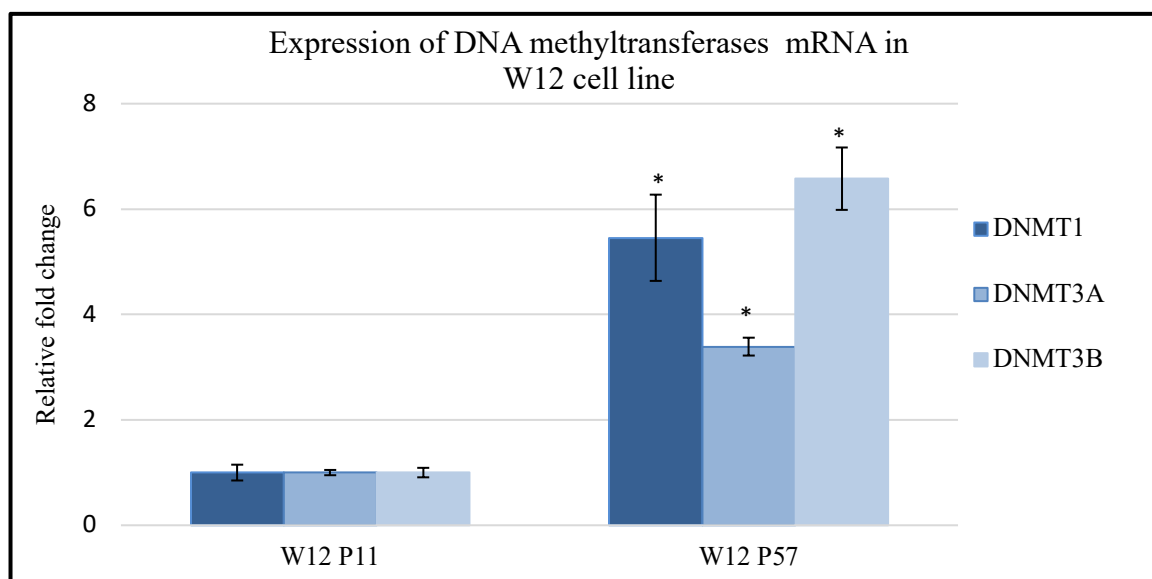


Figure 25: q-PCR results demonstrating expression of DNA methyltransferases mRNA level relative to Beta-actin as a reference gene in W12 P57 versus W12 P12 (control). This shows up-regulation of a) DNMT1, b) DNMT3A, and c) DNMT3B, in a late passage W12 cell line in comparison to an early passage. Results shown are typical of 3 similar experiments. Each experiment was run in triplicate.\* P value  $\leq 0.05$ , two-tailed paired nonparametric Mann-Whitney U test. Error bars indicate standard deviations.

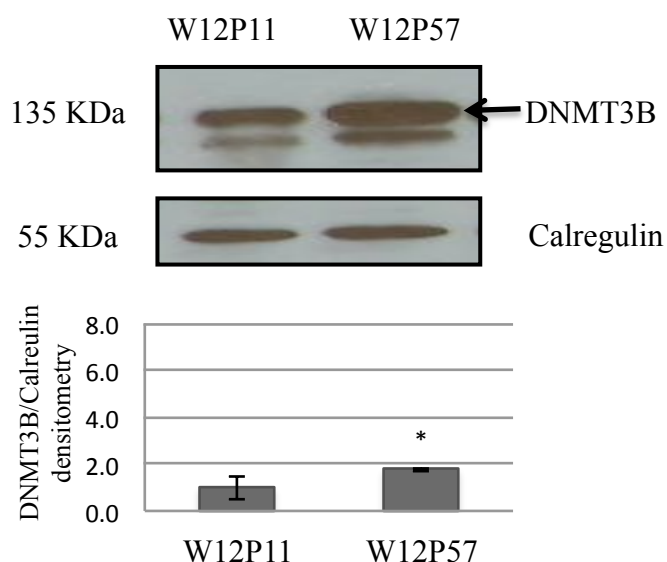


Figure 26: Western blot analysis and quantitative densitometry of DNMT3B protein expression in early (W12 p11) and late (W12 p57) W12 cell l. Results shown are typical of 3 similar experiments. Each experiment was run in triplicate.\* P value  $\leq 0.05$ , two-tailed paired nonparametric Mann-Whitney U test. Error bars indicate standard deviations.

### 3.3.3 Expression profile of DNA methyltransferases in a PHFK HPV18 cell line

In this section I demonstrate the sequential change in expression of DNMTs after transfection of PHFK with episomal HPV18. First, I show changes at the mRNA level using q-PCR. All q-PCR experiments were run in triplicate and repeated three times. Two-tailed paired nonparametric Mann-Whitney U test was used to calculate significant fold changes with changes considered significant if  $p \leq 0.05$ . I then show the dynamic changes in DNMT protein expression in the same PHFK HPV18 cell line. I also demonstrate changes to DNMT proteins at the cellular level using immunocytochemistry, which was performed using a cytopins technique, which was carried out using cells from selected passages and compared with both untransfected PHFK as a negative control and cervical cancer cell lines as a positive control.

#### 3.3.3.1 Kinetic analysis of changes in DNMT mRNA expression in a PHFK HPV18 cell line

I studied the changes in DNA methyltransferase expression overtime in PHFK transfected with HPV18. To do this I analysed the levels of DNMT mRNAs in PHFK transfected with HPV18 at each passage from p4 to p14 and compared this with the level of expression in PHFK before transfection, represented by passage 3 (

Figure 27). I show an up-regulation of both DNMT1 and DNMT3B in PHFK transfected with episomal HPV18. In contrast, DNMTL expression was not detected and DNMT3A, down-regulated (Figure 27b). However, that down-regulation was not significant. Interestingly, the up-regulation of DNMT1 and DNMT3B had the same patterns with two phases of overexpression compared with untransfected PHFK. Figure 27a, shows an immediate significant up-regulation of DNMT1 after transfection with episome HPV18. This increase

continued from passage 4 to passage 6. After a phase of quiescence during passages 9 and maintained until passage 13. DNMT3B mRNA expression was tested using transcription probes (1) and (2). Using probe (1) which is specific to isoform 7, I found an immediate transient up-regulation with another transient peak appears at passage 9 and passage 10 (

Figure 27c). Using probe (2) that is specific to isoform 1 (Figure 28) I observe mild transient significant up-regulation of DNMT3B at passage 5. Another up-regulation was then maintained until passage 14 and mimics the DNMT1 up-regulation pattern.

### **3.3.3.2 Kinetic analysis changes in DNMT protein expression in PHFK HPV18 cell line**

I used western blotting to study DNMT protein expression. Finding a good primary antibody that was sensitive enough to DNMT1 was difficult as the level of expression of DNMT1 in PHFK is known to be low (Sen et al., 2011). The calregulin protein was used as loading control. All western blots were repeated at least three times. The pattern of DNMT protein expression in most cases was similar to their messenger RNA (Figure 29). The DNMT3A protein showed no change in PHFK after transfection with HPV18 even after propagation of the cell line over many passages. The DNMT1 protein showed a low overall expression in PHFK, but after transfection a point of overexpression was detectable at passage 5. Then the next phase of overexpression of the DNMT1 protein started at passage 9.

The DNMT3B protein expression is represented in the western blotting by two bands. The expected size of DNMT3B protein using anti-DNMT3B antibody (Sigma) is approximately 110kDa, which is suggested to be the full length DNMT3B. Full length DNMT3B size is 96kDa, while DNMT3B7 is only 46 kDa (Ostler et al., 2007). The DNMT3B protein showed a high level of expression in PHFK. The increase in expression started at passage 5 and continued to passage 7 where the level rises by ten times compared to pre-transfection levels, before dropping by ~50% at passage 8 and passage 9 as shown by

quantitative densitometry analysis. At passage 10 the level doubled and maintained at that level. Overall, the DNMT1 and DNMT3B proteins showed two-phase pattern of overexpression. We note that the DNMT3B protein expression pattern differs from that of the mRNA which translate isoform 7, and mimicking the pattern of DNMT3B mRNA expression using probe (2) which translate full length DNMT3B protein, DNMT3B 1 isoform.

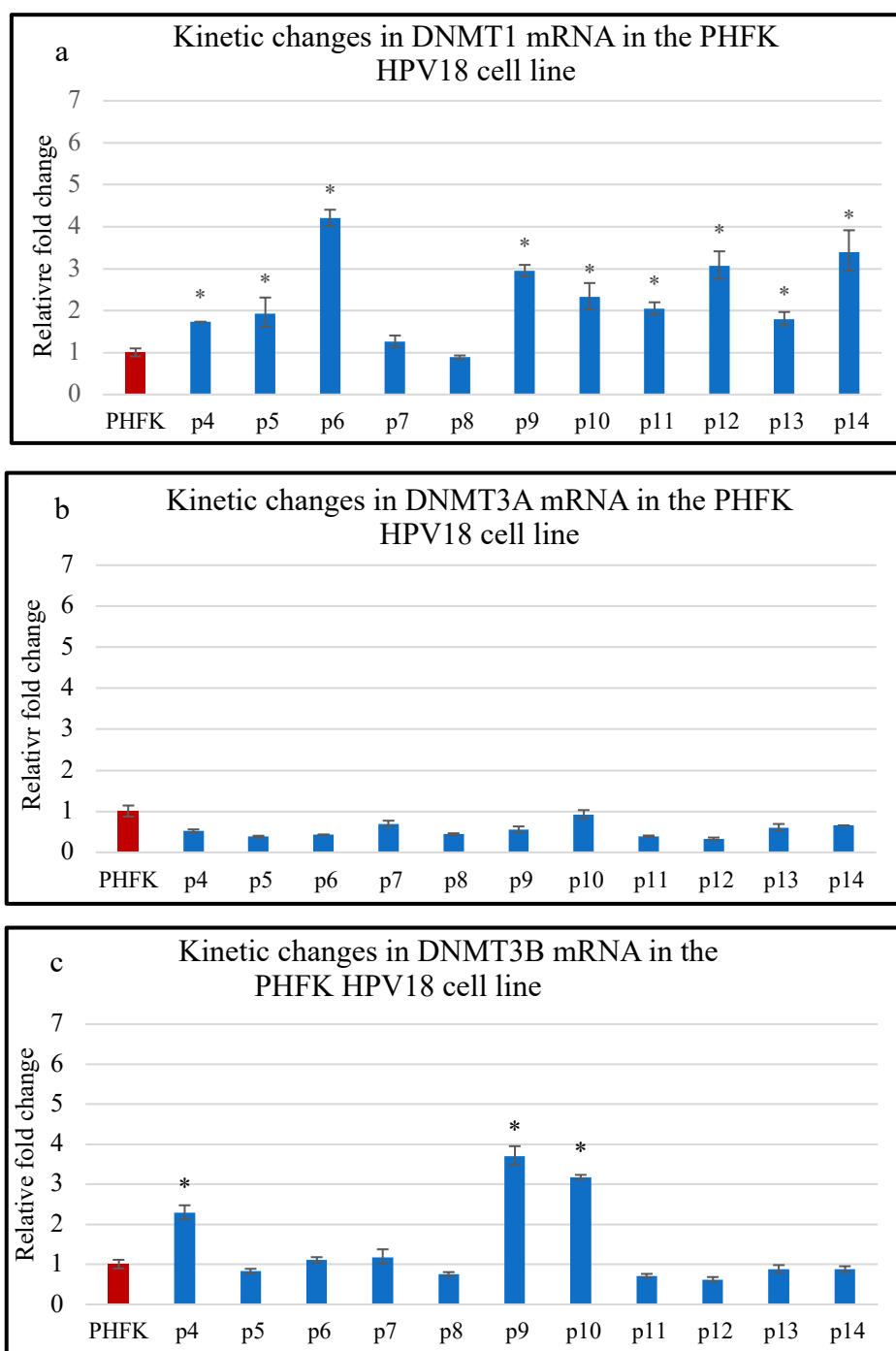


Figure 27: q-PCR results demonstrating expression profile of DNA methyltransferases mRNA level relative to Beta-actin as a reference gene in PHFK HPV18 cell line from passage 4 to passage 14 (p14) versus PHFK (control). a) Up regulation of DNMT1 mRNA. b) DNMT3A expression was not significantly changed. c) Up-regulation of DNMT3B using the DNMT3B (1) probe which translate to isoform 7. Results shown are typical of 3 similar experiments. Each experiment was run in triplicate.\* P value  $\leq 0.05$ , two-tailed paired nonparametric Mann-Whitney U test. Error bars indicate standard deviations.



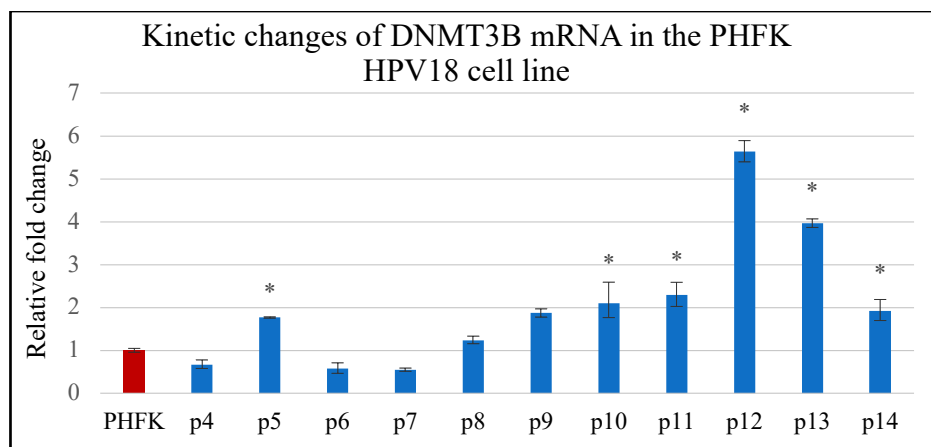


Figure 28: q-PCR results demonstrating expression profile of DNMT3B mRNA level relative to Beta-actin as a reference gene in PHFK HPV18 cell line from passage 4 to passage 14 (p14) versus PHFK (control) using DNMT3B (2) probe which is specific to isoform 1. Results shown are typical of 3 similar experiments. Each experiment was run in triplicate.\* P value  $\leq 0.05$ , two-tailed paired nonparametric Mann-Whitney U test. Error bars indicate standard deviations.

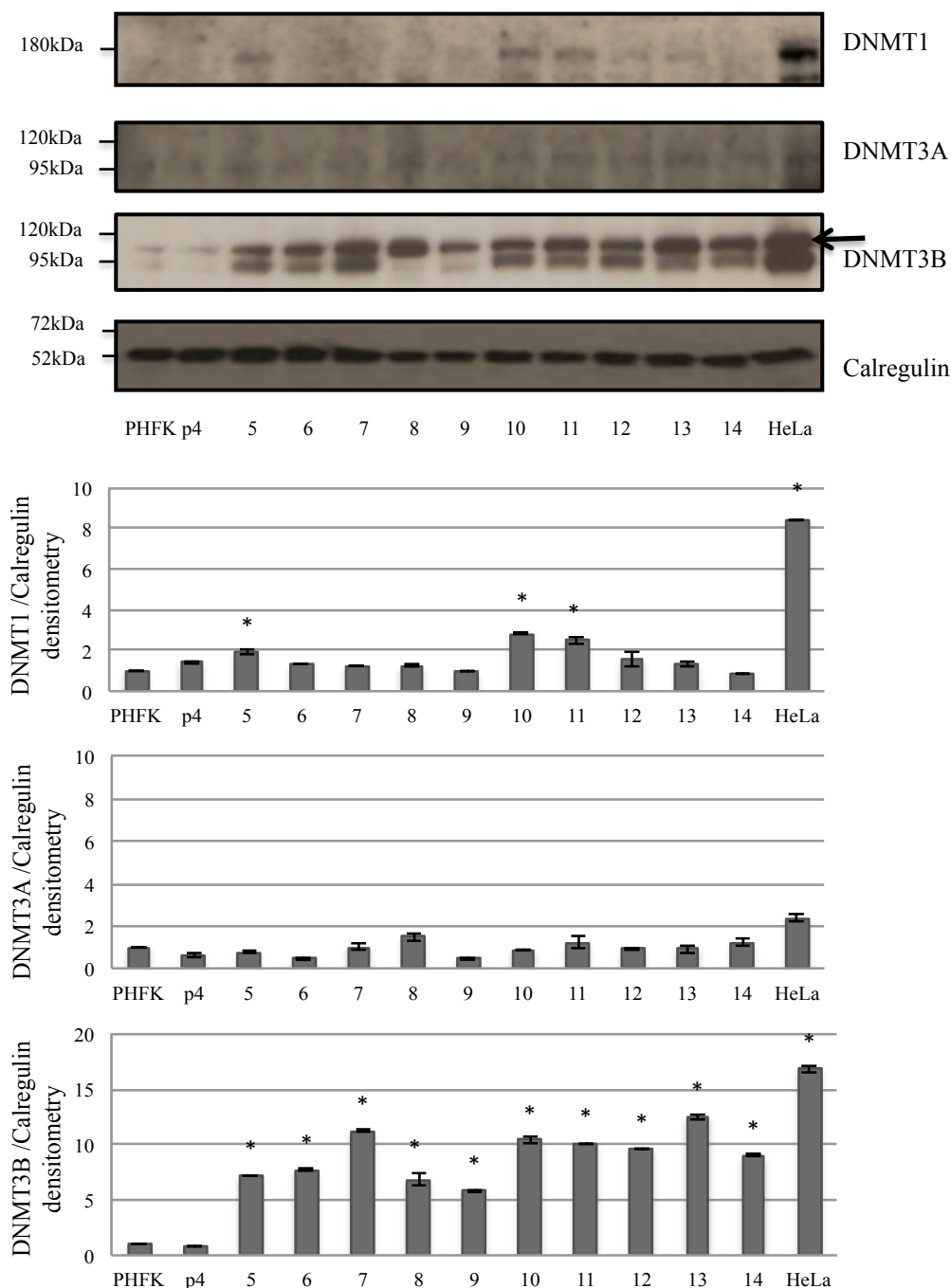


Figure 29: Western blotting analysis and quantitative densitometry showing dynamic changes in DNMT1, DNMT3A and DNMT3B protein expression in PHFK and PHFK HPV18 cells from passage 4 to passage 14. Calregulin was used as a loading control. Results shown are typical of 3 similar experiments. \*  $p \leq 0.05$  versus control (Student's  $t$  test). Error bars indicate standard deviations.

### **3.3.4 Expression profile of DNMT in a PHFK HPV16B cell line**

I next conducted similar experiments with HPV16B. I used qPCR to study the changes in DNMT mRNA expression over a time course after transfection and used western blotting to examine DNMT1 protein expression. I also investigated DNMT3B protein expression in the PHFK HPV16B cell line (2).

#### **3.3.4.1 Kinetic analysis of changes in DNMT mRNA expression in a PHFK HPV16B cell line**

DNMT3L expression was not detected and the level of DNMT3A mRNA did not change following transfection of PHFK with HPV16B (Figure 30b). Interestingly, as seen following HPV18 cell line, the same two-phase pattern of expression of DNMT1 and DNMT3B was observed in PHFK transfected with HPV16 cell line (Figure 30 a and c). DNMT1 mRNA showed a transient up-regulation at passage 5, after which mRNA expression returned to pre-transfection levels. The second surge of DNMT1 mRNA commenced at passage 9 and was maintained at passage 10, as shown in Figure 30 (a). The DNMT3B probe (1) which is specific to isoform 7, (discussed in section 2.3.4.5), showed the same pattern of expression as DNMT1, (Figure 30 c). With an up-regulation at passage 5 then another at passage 10. I also investigated DNMT3B mRNA expression using the DNMT3B probe (2), which is specific to isoform 1 (Figure 31). This is also had a peak at passage 5 and another at passage 9.

### **3.3.4.2 Kinetic analysis of changes in DNMT protein expression in a PHFK HPV16B cell line**

Both DNMT1 and DNMT3B show overexpression following transfection of HPV16B in two phases (Figure 32), which commenced at passage 5 and continued to passage 6. The second phase of over expression appears at passage 9 and 10. DNMT3A shows no marked changes of expression over all passages.

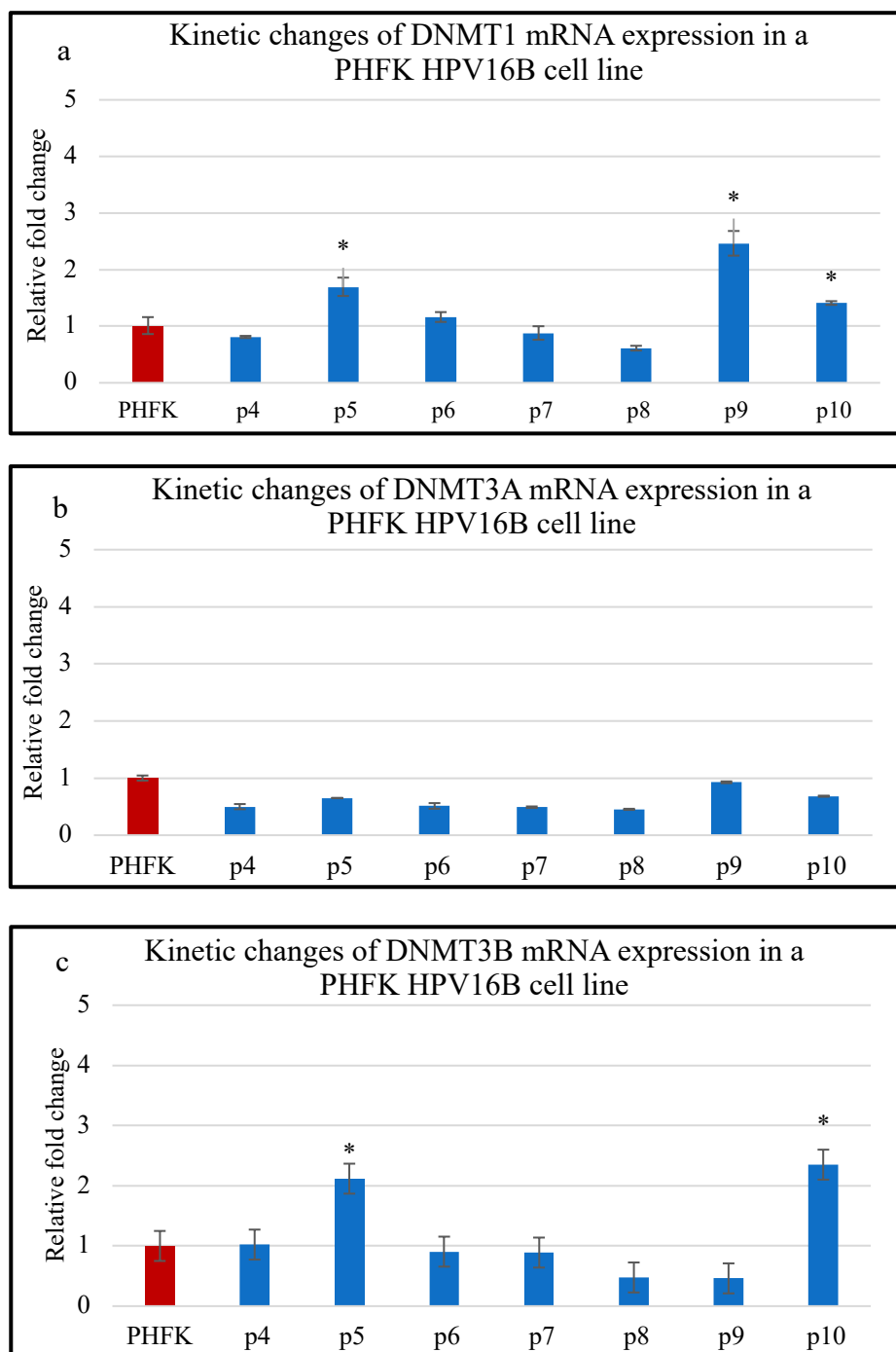


Figure 30: q-PCR results demonstrating the expression profile of DNA methyltransferases mRNA level relative to Beta-actin as reference gene in PHFK HPV16B cell line from passage 4 up to passage 10 (p10) versus PHFK (control). a) Up-regulation of DNMT1 mRNA. b) DNMT3A mRNA was not significantly changed. c) Up-regulation of DNMT3B using DNMT3B probe (1) that translate to DNMT3B isoform 7. Results shown are typical of 3 similar experiments. Each experiment was run in triplicate.\* P value  $\leq 0.05$ , two-tailed paired nonparametric Mann-Whitney U test. Error bars indicate standard deviations.

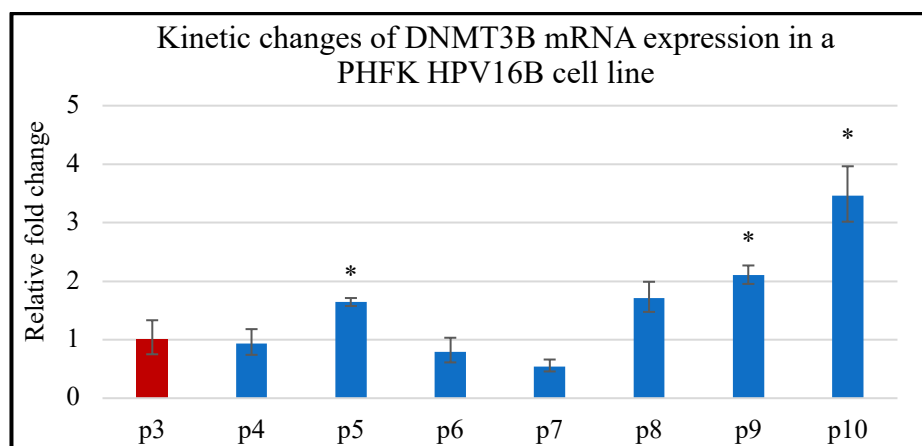


Figure 31: q-PCR results demonstrating the expression profile of DNMT3B mRNA (probe 2) level relative to Beta-actin as a reference gene in PHFK HPV16B versus (control). The DNMT3B expression probe (2) translates to DNMT3B isoform1. Results shown are typical of 3 similar experiments. Each experiment was run in triplicate.\* P value  $\leq 0.05$ , two-tailed paired nonparametric Mann-Whitney U test. Error bars indicate standard deviations.

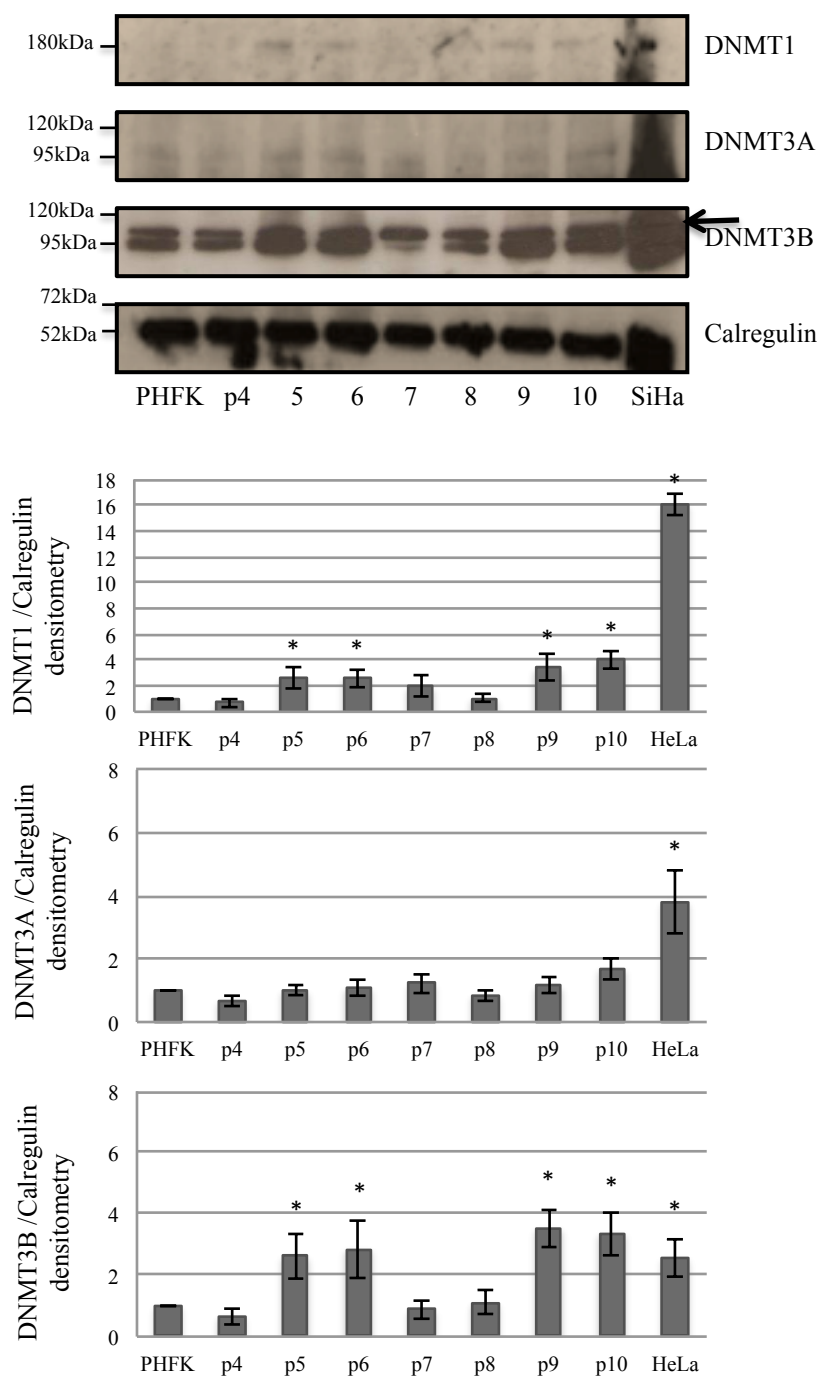


Figure 32: Western blot analysis and quantitative densitometry showing changes in DNMT1, DNMT3A and DNMT3B protein expression in PHFK and a PHFK HPV16B cell line passage 4 to passage 10. Calregulin was used as a loading control. Results shown are typical of 3 similar experiments. \*  $p \leq 0.05$  versus control (Student's  $t$  test). Error bars indicate standard deviations.

### **3.3.5 Expression profile of DNMT in a HPV16K-transfected PHFK cell line.**

Figure 33 shows the expression of DNMT1, DNMT3A and DNMT3B mRNA in PHFK HPV16K cell line. Like all the other cell lines that I have investigated, DNMT3L expression was not detected. DNMT3A expression did not change after transfection with HPV16K. Overall only one phase of up-regulation was seen for DNMT1 and DNMT3B. DNMT1 mRNA showed an immediate up-regulation following transfection. This significant surge reached a maximum level of six times the level of PHFK by passage 6. The up-regulation of DNMT for mRNA was maintained up to passage 10. For DNMT3B, mRNA expression was double that of the PHFK level at passage 6, and was maintained at this level until passage 9. I observed the same pattern at the protein level, (Figure 34) with overexpression of the DNMT3B protein at passage 6 that was maintained until passage 9.



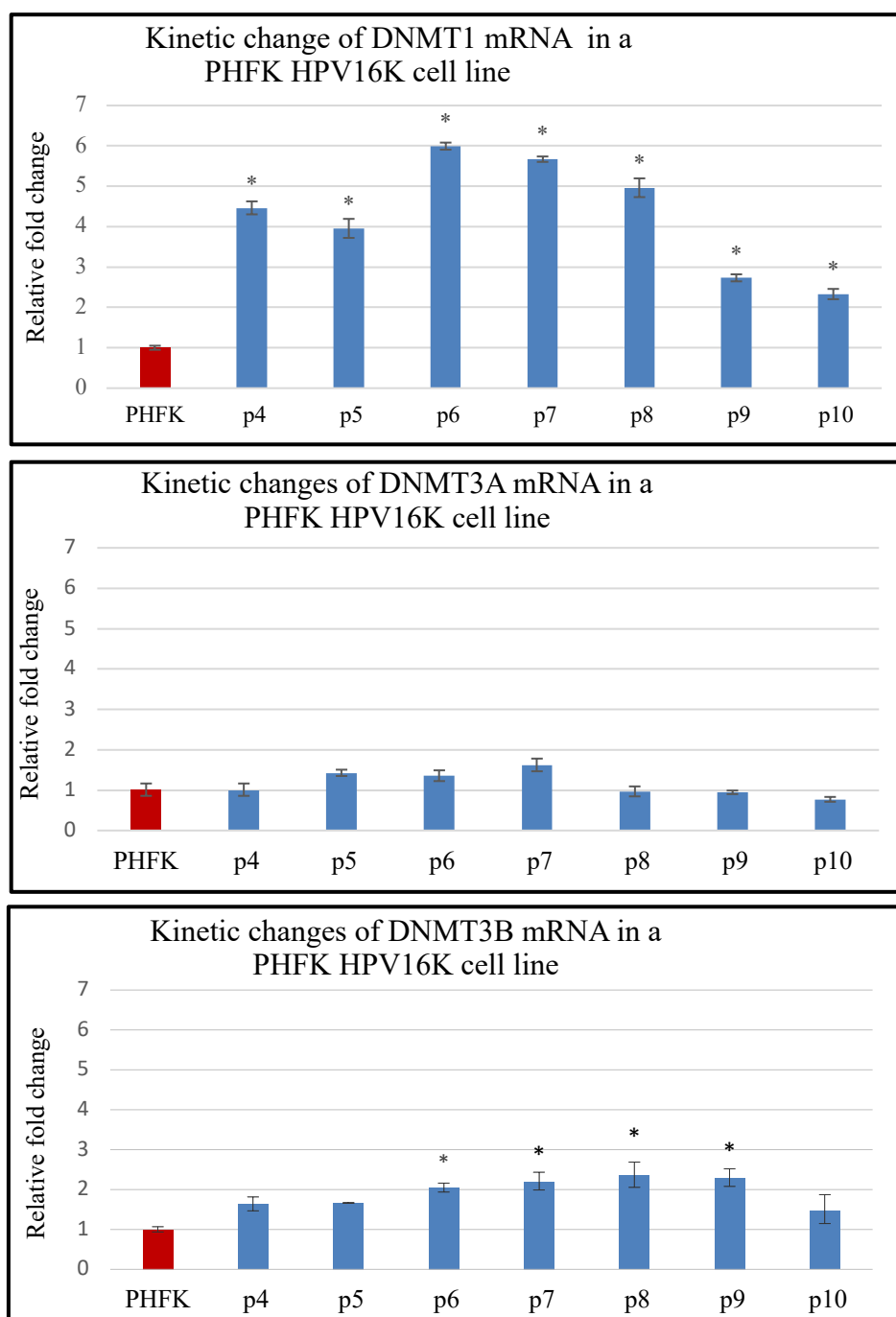


Figure 33: q-PCR results demonstrating the expression profile of DNA methyltransferases mRNA level relative to Beta-actin as reference gene in PHFK HPV16K cell line passage 4 to passage 10 (p10) versus PHFK (control). a) Up-regulation of DNMT1 mRNA. b) DNMT3A mRNA showed no changes in expression. c) Upregulation of DNMT3B mRNA. Results shown are typical of 3 similar experiments. Each experiment was run in triplicate.\*  $P$  value  $\leq 0.05$ , two-tailed paired nonparametric Mann-Whitney U test. Error bars indicate standard deviations.

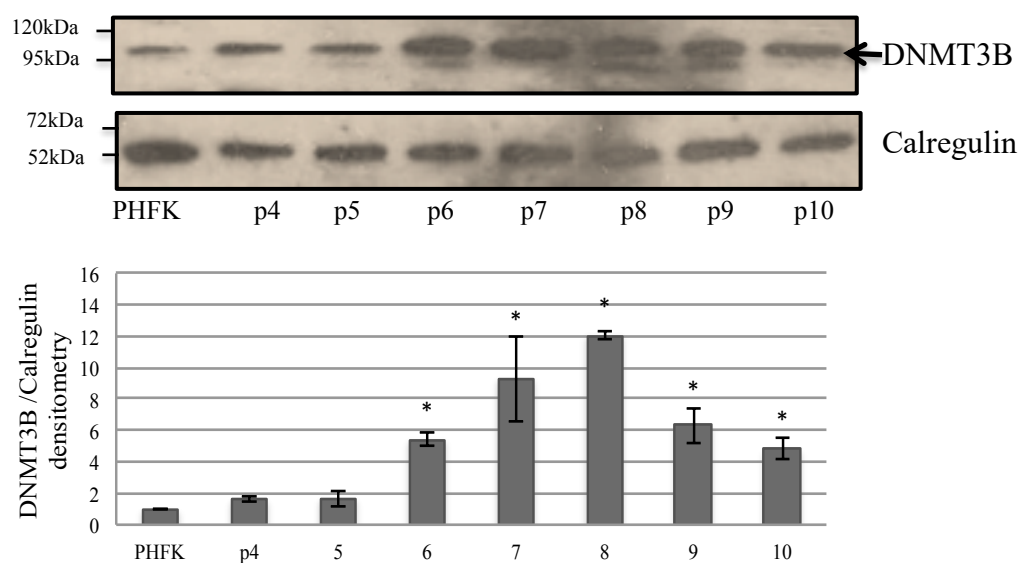


Figure 34: Western blot analysis and quantitative densitometry showing changes in DNMT3B protein expression in PHFK and PHFK HPV16K cell line from passage 4 to10, PHFK Calregulin was used a loading control. Results shown are typical of 3 similar experiments. \*  $p \leq 0.05$  versus control (Student's  $t$  test). Error bars indicate standard deviations.

### 3.3.6 Cellular localization of DNMT1 and DNMT3B proteins in PHFK

#### HPV16B and PHFK HPV18 cell lines

I further examined DNMT1 and DNMT3B expression changes at the cellular level using cytospin slides (Figure 37 and Figure 38). Immunocytochemical staining showed nuclear and cytoplasmic expression of both DNMT1 and DNMT3B proteins. The extent of staining was evaluated at 200x and 600x power magnification. DNMT1 and DNMT3B expression was greater in PHFK transfected with HPV18 and HPV16B compared with untransfected PHFK. Both DNMT1 and DNMT3B staining was more intense in the nucleus and cytoplasm in PHFK transfected with HPV18 cell line at passage 5 compared to untransfected PHFK which show negative nuclei and minimal cytoplasmic staining. I observed further intense nuclear staining at passage 14. Then, I analysed the intensity scoring. Slides were subjected to blind scoring performed by two independent pathologists. Immunocytochemistry staining was scored according to the intensity of the nuclear and cytoplasmic staining as follows: 0 (no staining or focal cytoplasmic and no nuclear staining); 1 (faint cytoplasmic or nuclear staining); 2 (moderate, smooth cytoplasmic or nuclear staining), or 3 (intense, granular cytoplasmic or nuclear staining). At least 300 cells were scored per slide and the percentage of positive cells was recorded. I considered cells positive for DNMT1 and DNMT3B expression if they had score of 2 or 3. In Figure 34 I show the median number of cells positive for staining compared to untransfected PHFK. Box and whisker plots were chosen to present the scoring data, as they are good for data analysing that are related to each other but collected from multiple sets by two independent observers.

The DNMT1 protein showed a low expression in untransfected PHFK, with only approximately 5% cells being positive for DNMT1 staining (Figure 35a). The percentage of p5HPV16B and p5 HPV18 cells positive for DNMT1 staining increased after transfection with

HPV16B and HPV18 positive cells representing ~35% and ~25% of the total cells at passage 5, respectively. The median number of cells positive for DNMT1 surged in the late passages of PHFK HPV16B and PHFK HPV18 cell lines to 45% of total cells. Cellular expression of DNMT3B shows the same overall pattern as DNMT1 protein expression (Figure 34b), with some differences in the median number of positive cells. Overall, these results indicate that DNMT3B and DNMT1 are expressed at cellular level to higher level in PHFK following transfection with HPV18 and HPV16B, with levels further raised in late passages.

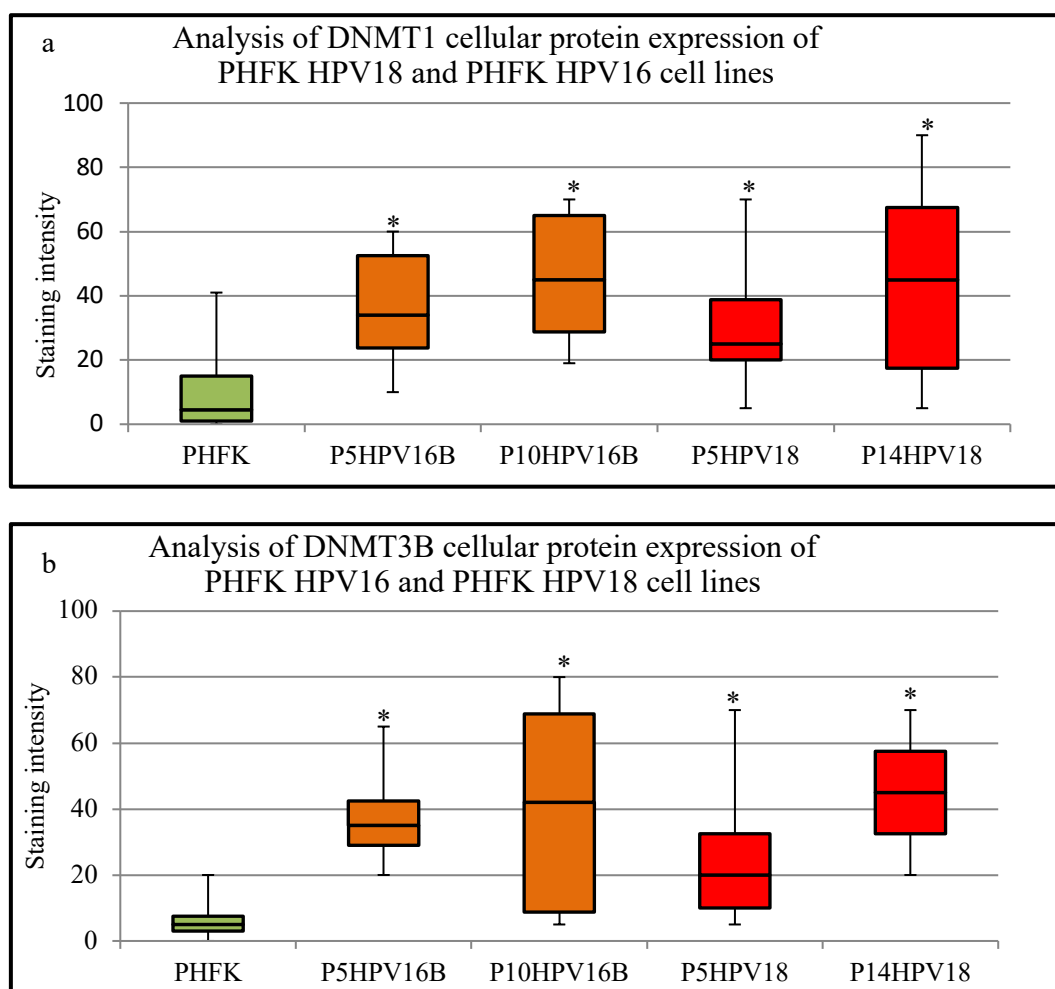


Figure 35: Box plots demonstrate quantification of the median number of cells expressing DNMT1 and DNMT3B proteins, analysed by two independent observers. The study involved analysis of slides from passage 5 and passage 10 of the PHFK HPV16B cell line, passage 5 and passage 14 of PHFK HPV18 cell line. PHFK are labelled green, cell lines with HPV16B are labelled orange and cell lines with HPV18 are labelled red. \*  $p \leq 0.05$  versus control (PHFK) two-tailed paired  $t$  test.

### **3.3.7 Morphological changes of cultured PHFK transfected with HPV**

Whilst culturing PHFK cell lines I noticed morphological changes in transfected PHFKs. I observed alterations in morphology in HPV-transfected PHFK compared with their untransfected counterparts (Figure 36). I present typical images of these cells, comparing passage 3 to passage 4 of PHFK HPV18 colonies. PHFK HPV18 colonies show a distinct cellular density and changes in cellular morphology, notably multinucleation and koilocytosis with cellular enlargement and perinuclear halos.

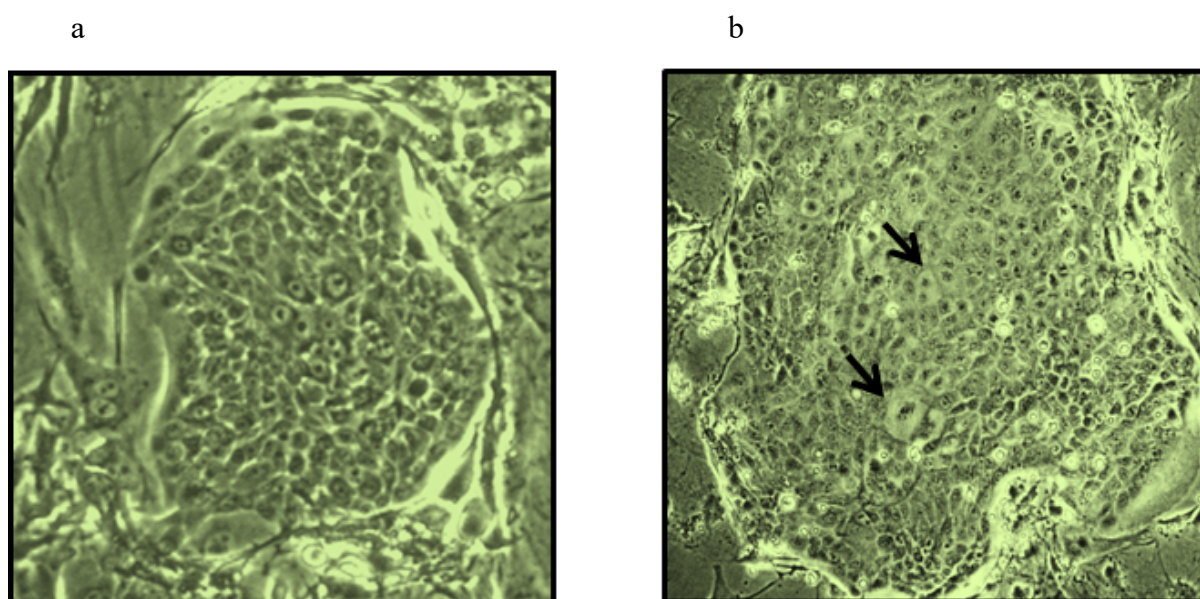


Figure 36: Phase contrast micrographs of monolayer culture with green filter to show morphological changes of the PHFK HPV18 cell line. a) Colony of passage 3 PHFK cultured over feeder cells, as compared with b) Colony of passage 4 of PHFK HPV18 cell line. Arrows show morphological changes: binucleation and koilocytosis, 40x magnification.

### **3.3.8 Cytopathic changes of PHFK HPV16 and PHFK HPV18 cell lines.**

Studying the morphological changes after transfection of PHFK cells with HPV18 and HPV16B. I observed changes at early and late passages. These include: koilocytosis, cellular enlargement, binucleation, multinucleation, spindle koilocytes, increased nuclear cytoplasmic ratios and other non-specific abnormal looking nuclei (Figure 39). Cells with morphological changes showed a higher level of staining of DNMT1 and DNMT3B than cells possessing normal morphology. In cancer cell lines, cellular changes were obvious such as irregular shaped nucleoli, increased nuclear to cytoplasmic ratios and abnormal mitotic figures. Cellular morphological changes were detected more in early than late passages in both PHFK HPV18 and PHFK HPV16 cell lines (Figure 39). I found that binucleation and cytomegaly were characteristic of early passages, whereas late passages showed more spindle shaped nucleoli. Increased nuclear to cytoplasmic ratios and other nuclear abnormalities such as tripolar, tetrapolar and irregular nuclear outlines were also observed in late passages, albeit infrequently.



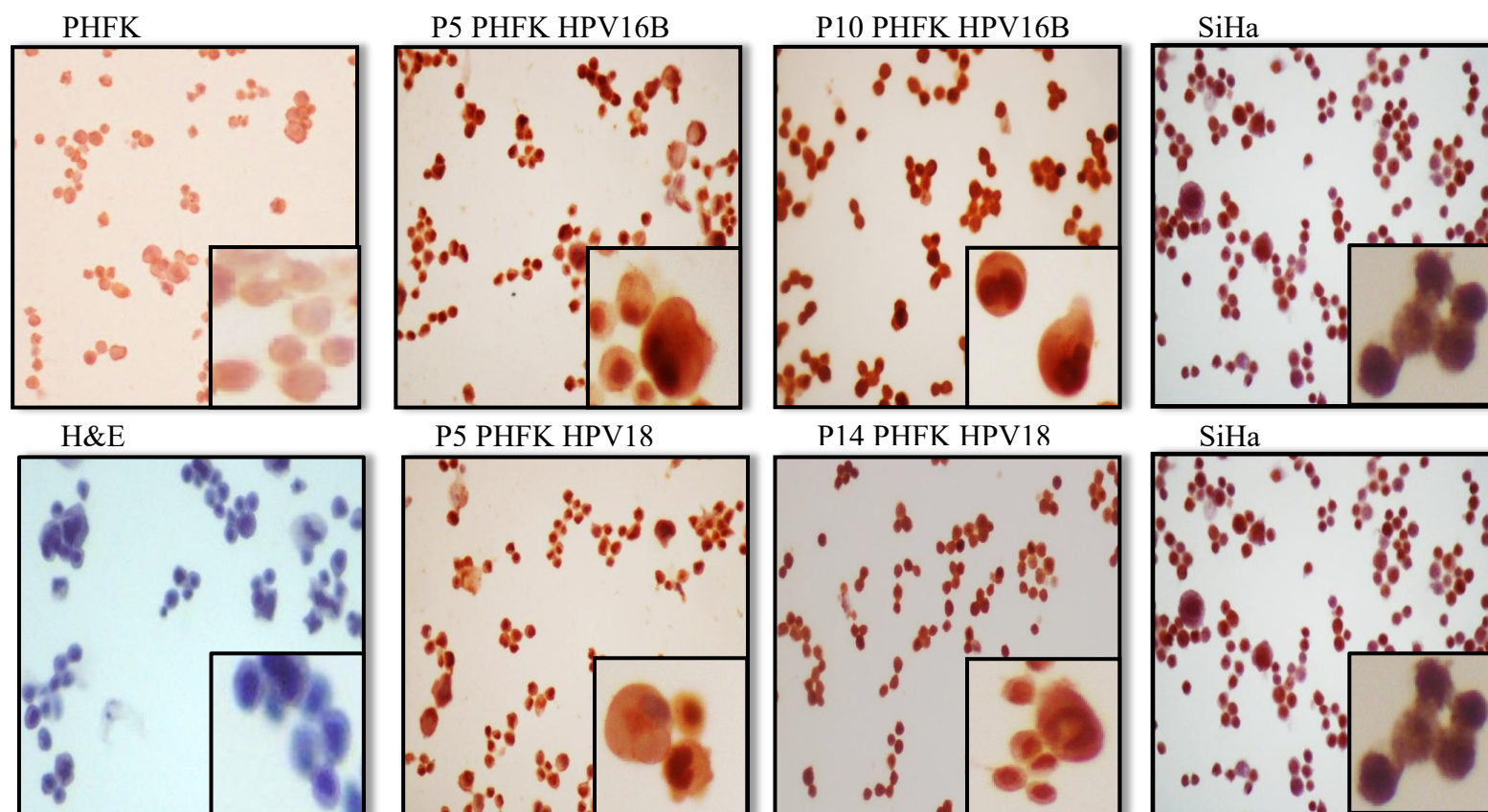


Figure 37: Immunocytochemistry detection of the DNMT1 protein. Increased DNMT1 protein expression was observed in PHFK HPV16B and PHFK HPV18 cell lines compared to PHFK. A cancer cell line (SiHa) was used as positive control and H&E staining is used in the negative control.

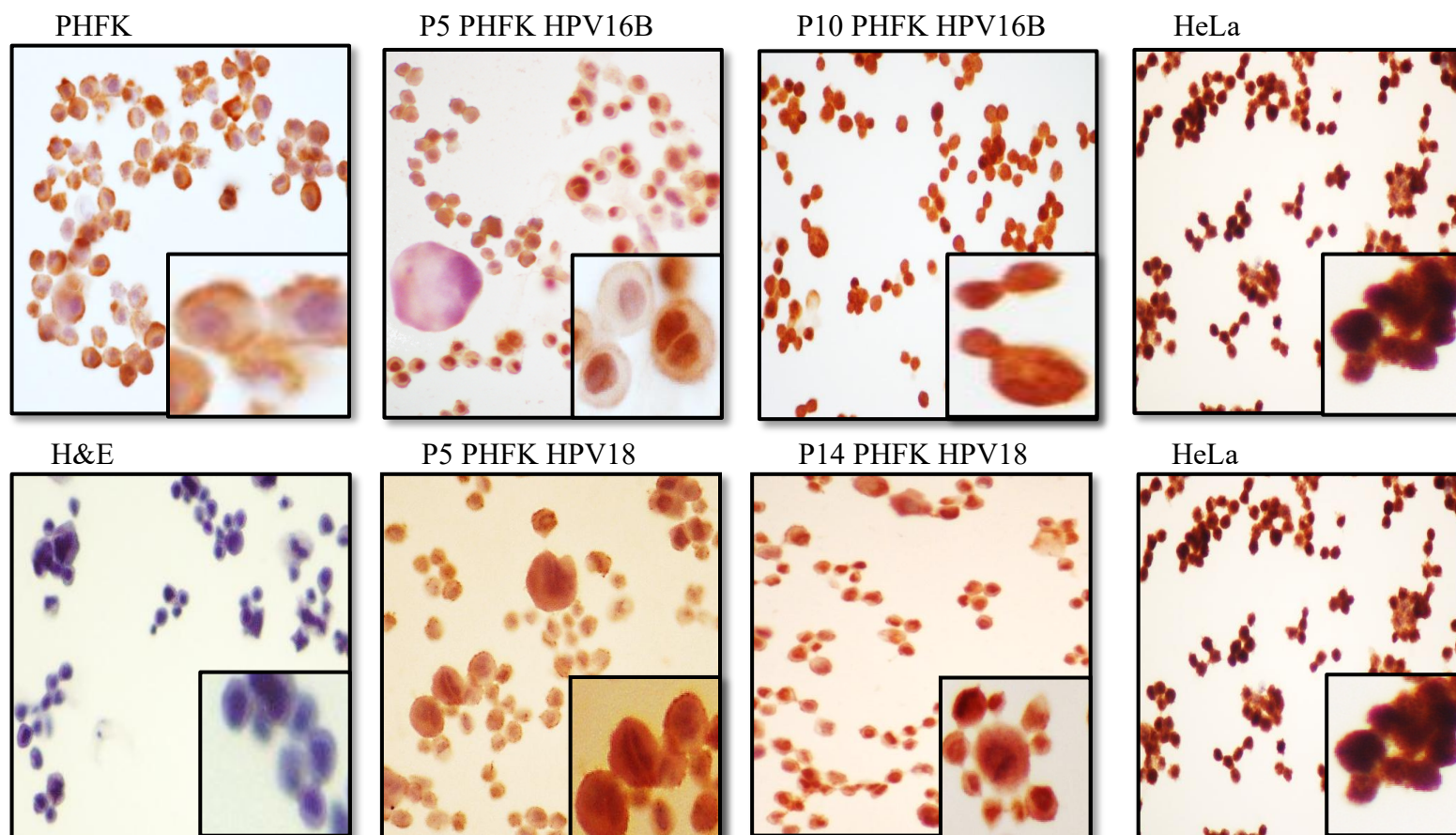


Figure 38: Immunocytochemistry detection of the DNMT3B protein. Increased DNMT3B protein expression was observed in the PHFK HPV16B and PHFK HPV18 cell lines compare to PHFK. A cancer cell line (SiHa) is used as positive control, and H&E staining in the negative control.

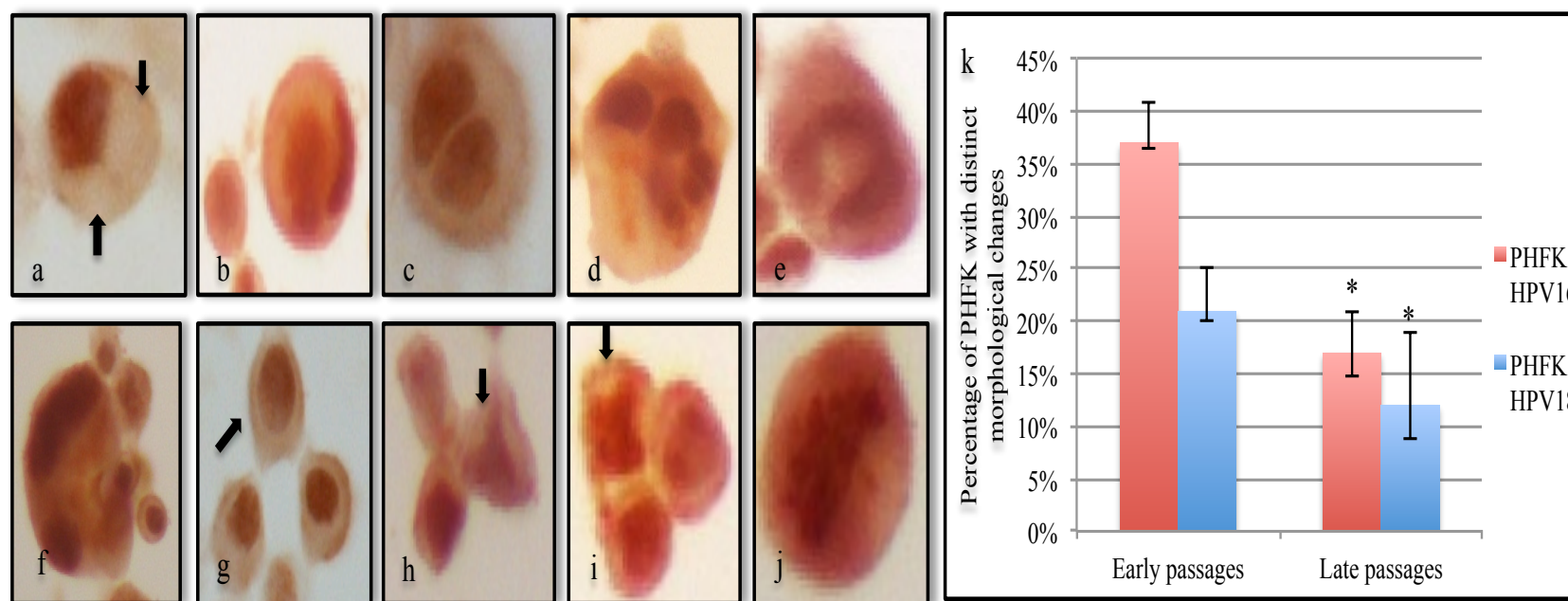


Figure 39: Cytopathic changes observed in PHFK transfected with HPV16 and HPV18. a, Koilocytosis, with an acentric nucleus and perinuclear vesicles indicated by black arrows. b, Cytomegaly. c, Binucleation. d, Multinucleation. e, Spindle shaped nucleus. f, Complex cell with cytomegaly and binucleation surrounded by cells with normal morphology. g, Cell with increased nuclear to cytoplasmic ratio indicated by arrow. h & i, Abnormal shaped nucleus; tripolar and tetrapolar nucleus. j, Cell with irregular nuclear outline. k, Histogram shows the percentage of PHFK showing cytopathic changes in early and late passages for both PHFK HPV16 and PHFK HPV18 cell lines, \*  $P$  value  $\leq 0.05$ , two-tailed paired  $t$  test.

In brief, in the above sections of this chapter I have shown that DNA methyltransferases DNMT1 and DNMT3B are overexpressed in PHFK after transfection with episomal HPV18 and HPV16B. Their pattern of expression in PHFK HPV18 and PHFK HPV16B cell lines shows two peaks: one immediately after transfection and another later peak after cell propagation. PHFK HPV16K cell lines showed different expression pattern with only one peak after transfection. I have also illustrated the different morphological changes observed in HPV transfected cell lines using cultured PHFK cells and immunocytochemistry.

This suggests that the levels of DNMTs rise throughout the process of cervical dysplasia, represented by W12 cell line, and continue to be high in different cervical cancer cell lines. Interestingly, I could not detect DNMT3L, neither in PHFK with or without HPV, nor in the W12 cell lines.

### **3.4 Expression changes of H3K27me3 methyltransferase and H3K27me3 demethylases in *In vitro* experimental models system**

In this section I investigate changes in the expression of H3K27me3 methyltransferase EZH2 and the H3K27me3 demethylases KDM6A and KDM6B in *In vitro* experimental models system. I profile expression changes at mRNA level using q-PCR. Then, I show the expression profile of EZH2 at the protein level using western blotting.

#### **3.4.1 Expression profile of the H3K27me3 methyltransferase and demethylases in cervical cancer cell lines and W12 cell lines**

SiHa expressed a higher level of EZH2 than HeLa and Ca Ski. However, a much more pronounced expression was observed, with the HPV-negative cell line, C33A (Figure 40). Noticeably, the H3K27me3 demethylases showed similar pattern of expression to EZH2, Overall, all HPV-positive cervical cancer cell lines expressed EZH2, KDM6A and KDM6B, although C33A expressed them at much higher levels. Figure 41 shows low levels of EZH2, KDM6A and KDM6B expression at an early passage (p11) with a marked up-regulation following cell propagation at late passage (p57). EZH2 levels were 12 times greater at p57 and KDM6A and KDM6B levels were remarkably higher, at ~20 and 30 times the level of expression observed at p11 respectively.

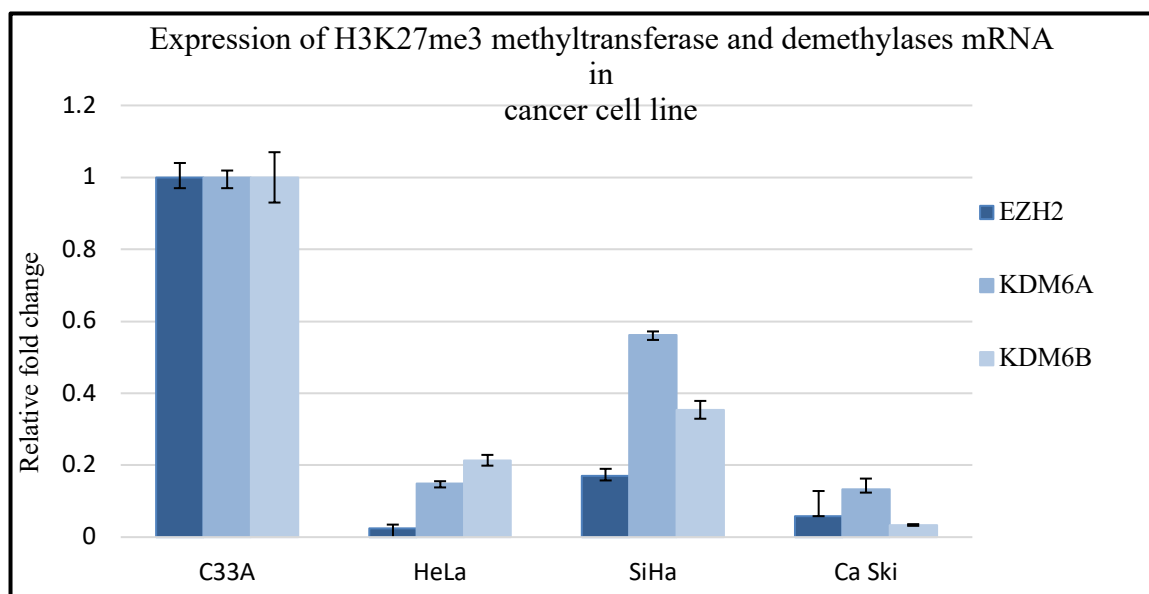


Figure 40: q-PCR results demonstrating expression of mRNA of H3K27me3 methyltransferase EZH2 and H3K27me3 demethylases KDM6A and KDM6B relative to Beta-actin as a reference gene in different cervical cancer cell line versus C33A. Results shown are typical of 3 similar experiments. Each experiment was run in triplicate. Error bars indicate standard deviations.

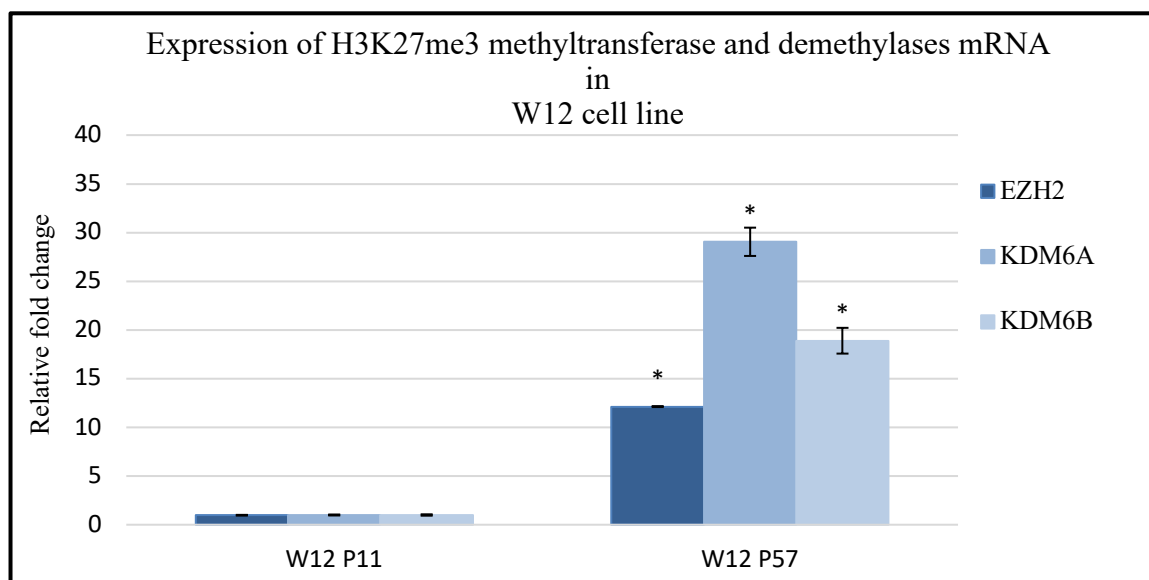


Figure 41: q-PCR results demonstrating expression of mRNA of H3K27me3 methyltransferase EZH2 and H3K27me3 demethylases KDM6A and KDM6B relative to Beta-actin as a reference gene in W12 late passages (P57) versus W12 P11 (control). Results shown are typical of 3 similar experiments. Each experiment was run in triplicate. Error bars indicate standard deviations.

### **3.4.2 Kinetic analysis of changes in expression of H3K27me3**

#### **methyltransferase and demethylases in a PHFK HPV18 cell line**

Figure 42 shows that following HPV transfection at passage 3, there was a gradual up-regulation of EZH2 mRNA that commenced at passage 6 and was maintained up to passage 14 followed by a second surge at passage 13. I also showed these changes at the protein level. I show gradual overexpression of EZH2 protein in PHFK after transfection with HPV18 commenced at passage 8 and maintained to passage 14. A higher level of EZH2 expression is observed in HeLa (Figure 43). Transfection of PHFK with HPV18 led a transient peak of KDM6A mRNA observed at passages 9 and 10 and changes observed in expression of KDM6B mRNA.

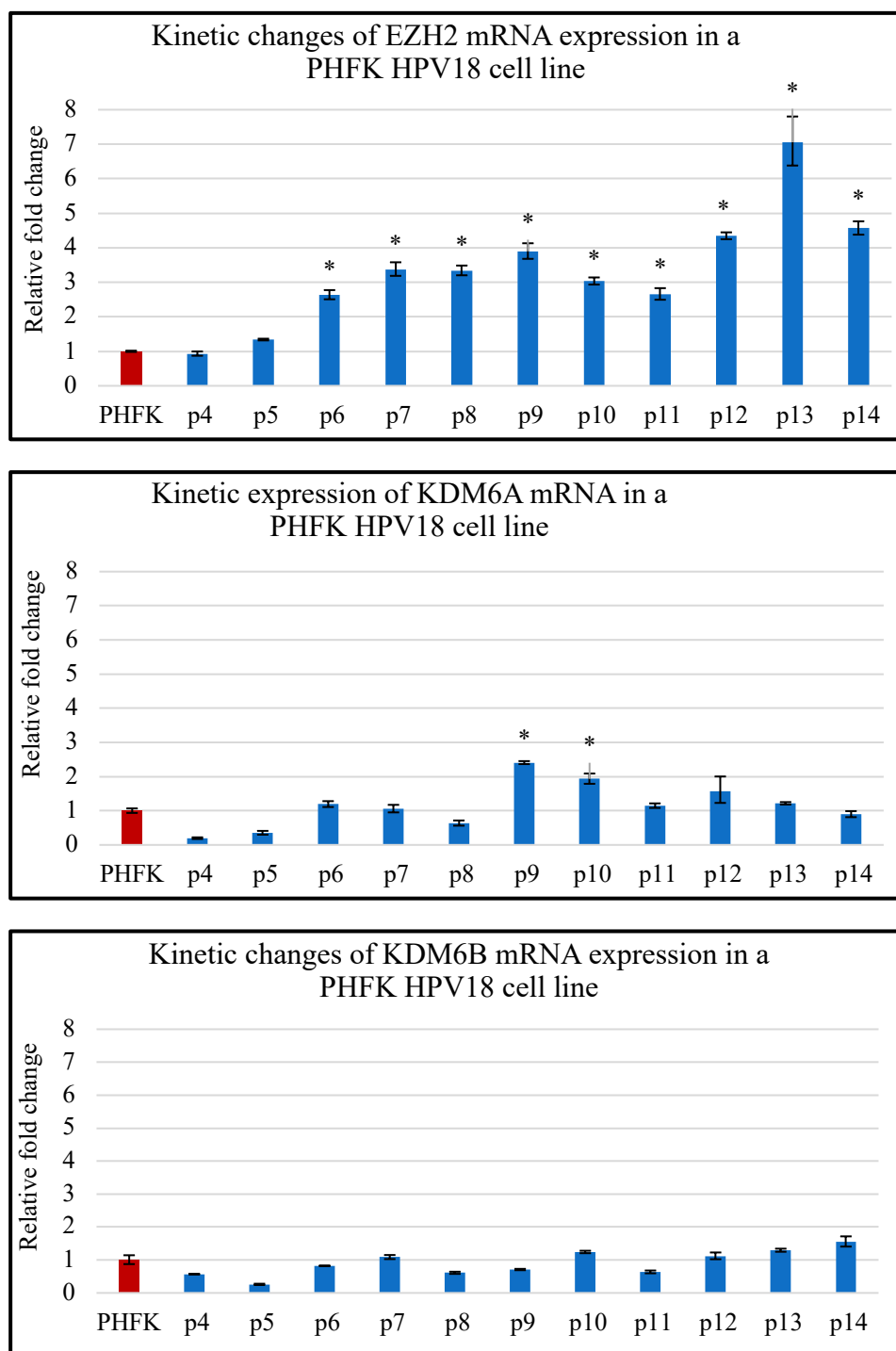


Figure 42: Kinetic analysis of mRNA expression changes of the H3K27me3 methyltransferase EZH2 and the demethylases KDM6A and KDM6B relative to Beta-actin as a reference gene using q-PCR in PHFK HPV18 cell line from passage 4 to passage 14 versus PHFK (control). Results shown are typical of 3 similar experiments. Each experiment was run in triplicate.\*  $p$  value  $\leq 0.05$ , two-tailed paired nonparametric Mann-Whitney U test. Error bars indicate standard deviations.



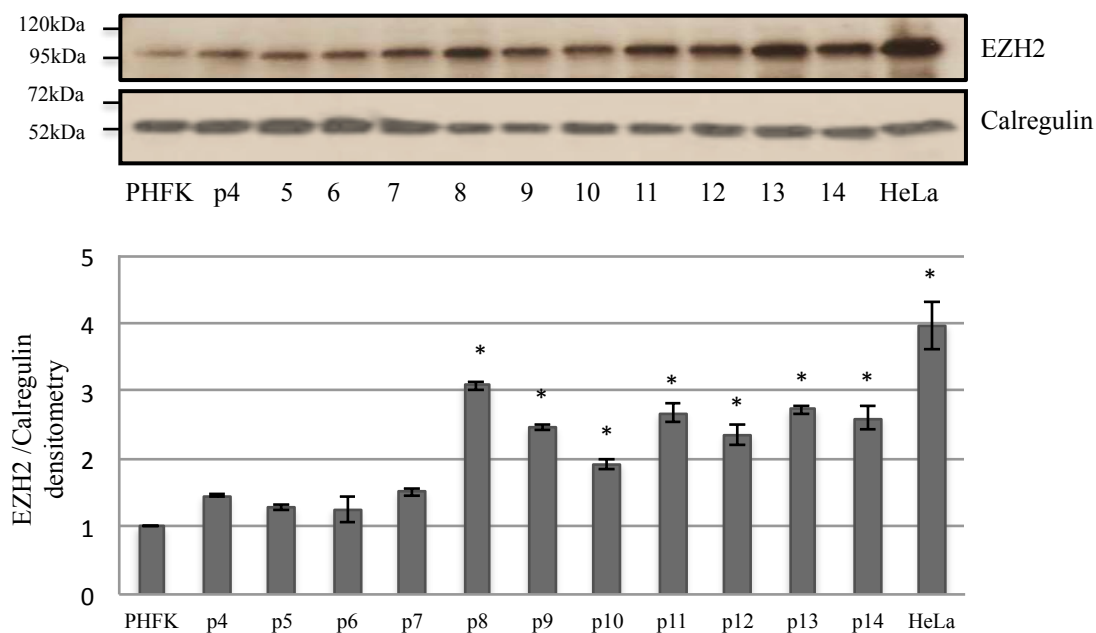


Figure 43: Kinetic analysis of protein of the H3K27me3 methyltransferase, EZH2 in a PHFK and HPV18 cell line from passage 4 to passage 14, using western blotting analysis and quantitative densitometry. Calregulin was used as a control. Results shown are typical of 3 similar experiments. \*  $p \leq 0.05$  versus control (Student's  $t$  test). Error bars indicate standard deviations.

### **3.4.3 Kinetic analysis of the changes in expression of the H3K27me3**

#### **methyltransferase and demethylases in a PHFK HPV16B cell line**

Interestingly, the pattern of EZH2 mRNA expression in PHFK transfected with HPV18 was the same in PHFK transfected with HPV16B (Figure 44). A gradual up-regulation commenced at passage 5 and was maintained with cell propagation. A much higher level was observed with the SiHa cell line (Figure 44). However, I observed no significant change in KDM6A mRNA expression as cells propagated. There was a transient non-significant reduction of KDM6B mRNA expression after transfection but expression levels returned to pre-transfection levels thereafter. A sudden up-regulation then appeared at passage 10 (Figure 44). Using western blotting I showed a gradual rise in EZH2 protein expression, commencing at passage 5 and which was maintained in later passages (Figure 45).

### **3.4.4 Kinetic analysis of changes in the expression of the H3K27me3**

#### **methyltransferase and demethylases in a PHFK HPV16K cell line**

PHFK transfected with HPV16K showed a unique pattern of expression compared to PHFK transfected with HPV18 and HPV16B. In Figure 46 I show an immediate and sustained up-regulation of EZH2 and KDM6A mRNA after transfection with HPV16K. I also show a gradual up-regulation of KDM6B that peaks at passage 6 and returned to pre-transfection levels at passage 8.

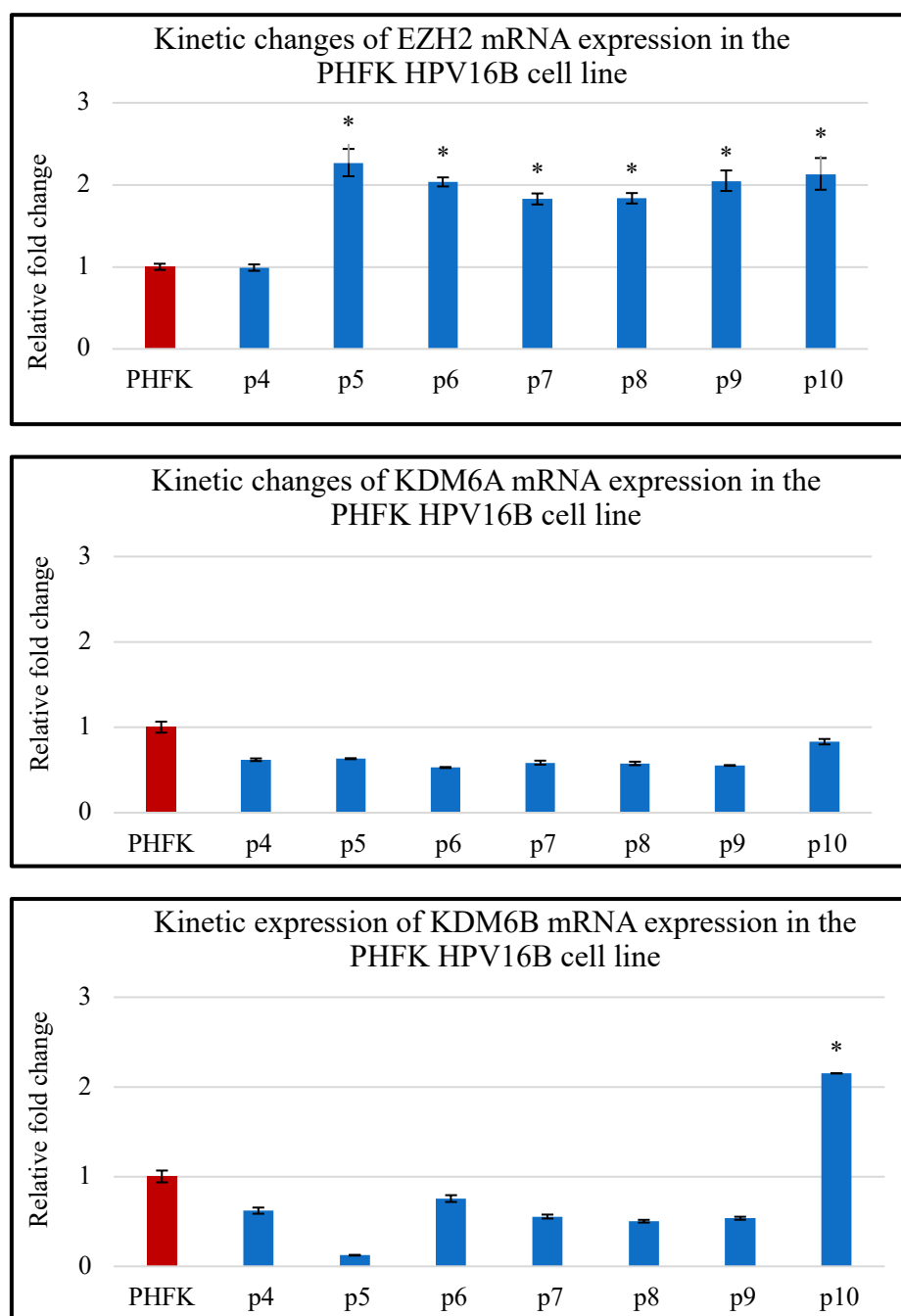


Figure 44: Kinetic analysis of changes in mRNA expression of the H3K27me3 methyltransferase EZH2 and demethylases KDM6A and KDM6B relative to Beta-actin as a reference gene using q-PCR in PHFK HPV16B cell line from passage 4 to passage 10 versus PHFK (control). Results shown are typical of 3 similar experiments. Each experiment was run in triplicate. \* P value  $\leq 0.05$ , two-tailed paired nonparametric Mann-Whitney U test. Error bars indicate standard deviations.

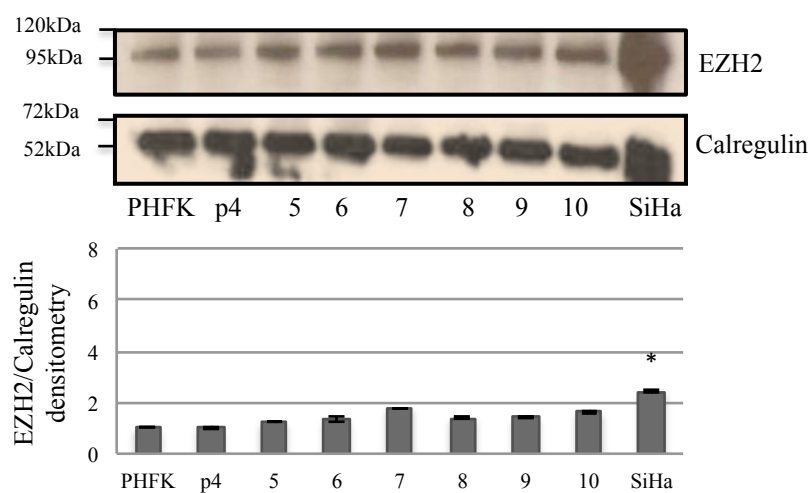


Figure 45: Kinetic analysis of changes in protein expression of the H3K27me3 methyltransferase EZH2 in a PHFK and PHFK HPV16B cell line from passage 4 to passage 10, using western blotting analysis and quantitative densitometry. Calregulin was used as

loading control. Results shown are typical of 3 similar experiments. \*  $p \leq 0.05$  versus control (Student's  $t$  test). Error bars indicate standard deviation.

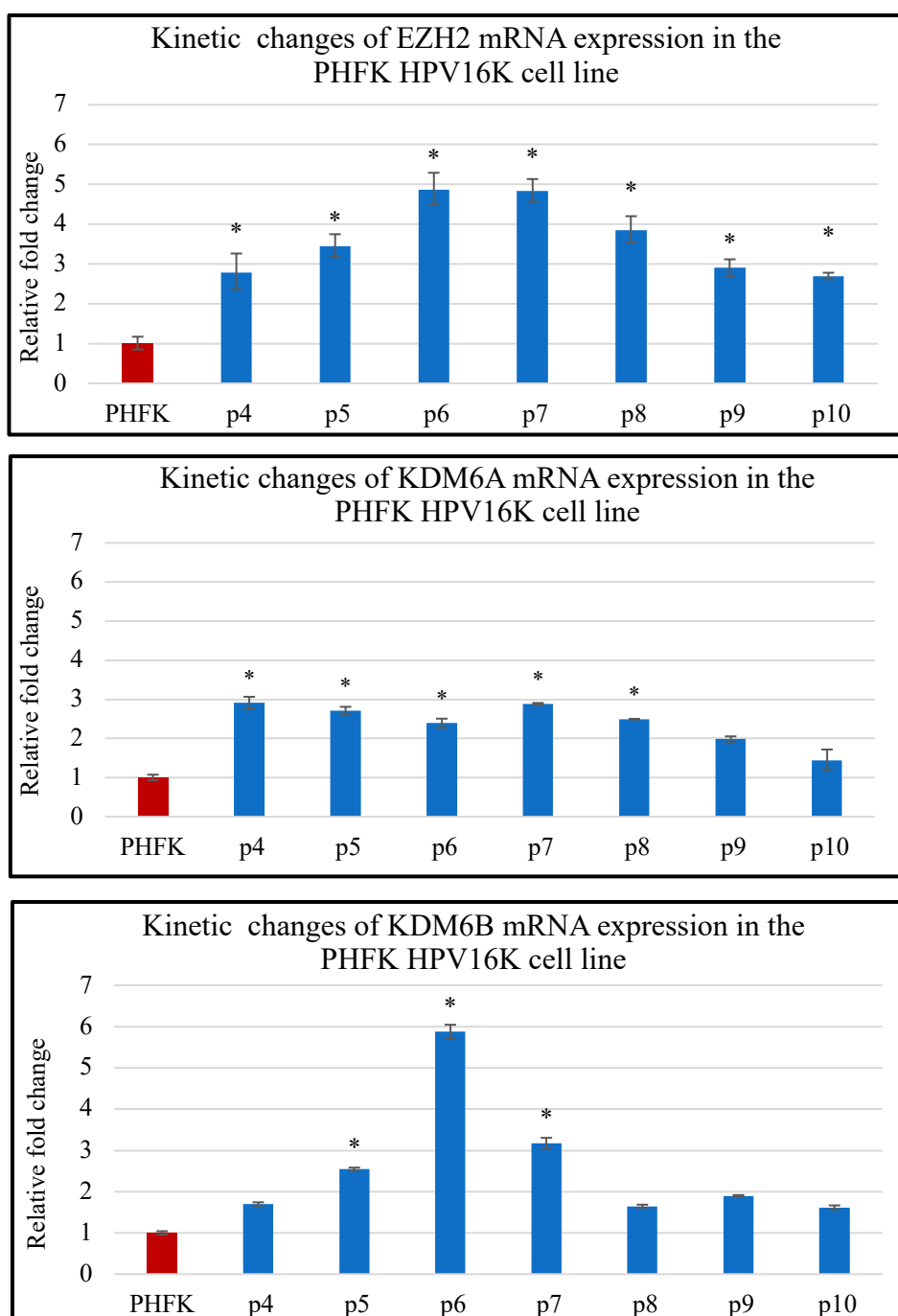


Figure 46: Kinetic analysis of changes in mRNA expression of the H3K27me3 methyltransferase EZH2 and the demethylases KDM6A and KDM6B relative to Beta-actin as reference gene using q-PCR in PHFK HPV16K cell line from passage 4 to passage 10 versus PHFK (control).. Results shown are typical of 3 similar experiments. Each experiment was run in triplicate. \*  $p$  value  $\leq 0.05$ , two-tailed paired nonparametric Mann-Whitney U test. Error bars indicate standard deviations.

This H3K27me3 expression data reveals that the level of EZH2 expression rose after transfection with HPV18, HPV16K and HPV16B. This level was sustained and is consistent with other data showing up-regulation of EZH2 at late passage of the W12 cell line. All cancer cell lines expressed EZH2. Both PHFK HPV18 and PHFK HPV16B cell lines showed a two-phase pattern of expression for KDM6A and KDM6B, similar to that seen for the DNMTs.

Result I, in summary the cytopathic changes associated with HPV infection *in vivo* also manifested in *in vitro* models. Interestingly, HPVs associated with changes in key epigenetic regulators mirror the changes observed in the process of cervical carcinogenesis. The pattern of changes in epigenetic regulators in my *in vitro* model suggested possibility of epigenetic switching.

## CHAPTER 4

### RESULTS II

# **HPV ONCOGENE E7 INDUCES EPIGENETIC REPROGRAMMING OF DNA METHYLTRANSFERASES**



**Aim of results II:**

To understand the mechanistic basis behind increased expression of DNA methyltransferases (DNMTs) in primary human foreskin keratinocytes (PHFK) transfected with high risk-HPVs.

**Objectives:**

1. To investigate the physical state of HPV in PHFK *in vitro* models.
2. To profile expression changes of the HPV E7 oncoprotein in PHFK *in vitro* models.
3. To investigate the DNMT3B gene copy number in untransformed and transformed cell lines harbouring HPV16 and HPV18.
4. To determine whether the HPV oncoprotein E7 can induce and interact with DNMT3B.
5. To study the enzymatic activity of DNMTs in PHFK *in vitro* models.
6. To determine whether the induction of DNMTs by HPVs can mediate changes in host DNA methylation of TSGs.

HPV integration into host cellular DNA disrupts viral integrity. Disruption mostly occurs at HPV E2 region. The subsequent loss of E2 and its negative feedback control on HPV oncogenes, promotes up-regulation of HPV oncogene E7, which itself induces destabilization of the pocket protein pRb. In the first section of this chapter I investigate the changes in the physical state (episomal vs. integrated) of HPV16K, HPV16B and HPV18 in my *in vitro* models.

Then, I study the expression changes of HPV oncoprotein E7 in relation to HPV physical state and investigate its activity by profiling the expression of pRb.

## 4 Results II

### 4.1 Introduction of results II

HPV integration into human genomes converts its circular genome into a linear DNA in which the HPV E6/E7 oncogenes always remain intact (Schwarz et al., 1985). In contrast, integration of HPV DNA is usually associated with disruption or complete deletion of HPV E1/E2 gene. As a consequence, abrogation of the E2-mediated repression of E6/E7 transcription may occur (Bechtold, Beard and Raj, 2003). Different methods have been used to determine the physical state of HPV. The PCR method was used to determine the sites of HPV integration in SiHa cells and late W12 cell lines (Pett et al. 2004a) which were originally identified as harbouring integrated HPV by southern blotting (Baker et al. 1987). Both methods showed disruption of HPV16 DNA to be at nucleotides 3132 to 3384 and at nucleotides 3732 to 4791 for SiHa and late W12 cell lines, respectively. Furthermore, Collins et al. (2009) study used a PCR-based approach to determine the disruption site of HPV E2 in HeLa to be at a similar position as that determined by the DIPS-PCR method (Luft et al., 2001) with both showing disruption of HPV18 DNA at nucleotides 3100 to 5736 (Luft et al., 2001; Collins et al., 2009). Thus, PCR-based methods were able to determine integration sites in both HPV16 and HPV18 (Pett et al. 2004a; Collins *et al.*, 2009). The oncoprotein HPV E7 overexpressed after viral integration. In fact HPV E7 oncoprotein is a major transforming protein capable of binding to and inducing degradation of retinoblastoma-associated protein, pRb (Helt and Galloway, 2001). This leads to the dissociation of the Rb/E2F complex, an event associated with G1 arrest, and potential apoptosis in response to DNA damage. This activation of the host DNA replication machinery is required for HPV DNA replication.

## 4.2 Investigating the physical state of HPVs

### 4.2.1 Assessing the integrity of full length HPV18 E2 in an *in vitro* model

Adapting an approach described previously by Collins et al. (2009), I used a PCR-based method and I designed a set of overlapping primers to assess the integrity of the full length HPV18 E2 gene (Figure 47). In this PCR method GAPDH and HPV E7 were used as loading and positive control, respectively. I show consistent amplification of both GAPDH and HPV18 E7 in all passages of PHFK HPV18 cell line (Figure 48 a). I was able to amplify a large segment (944 bp) of HPV18E1/E2 in all passages up to passage 12 where a gradual reduction of band intensity starts to appear. At passage 13 further reductions were noticed which was followed by a complete disappearance of HPV18E1/E2 by passage 14 (Figure 48b). This indicates that disruption of the HPV18E1/E2 region in my *in vitro* model commenced at passage 12 and was completed at passage 14. Then, I used five overlapping primers with a size of 300 to 400 base pair each to investigate the exact site of HPV18 E2 disruption (Figure 48b). The first two primer sets covered E1/E2 regions from 2494 to 3121. Another two primer sets designed to cover E2 region from 3086 to 3739. The last, primer set covered E2/E5 region from 3450 to 3994. I found amplification of HPV18 E1/E2 from nucleotides 2494 to 3121 in all passages of the PHFK HPV18 cell line. In contrast, I showed amplification of E2/E5 region from nucleotides 3086 to 3994 from passage 4 to passage 12. By passage 12 after which a gradual reduction in the intensity of the PCR band can be seen followed by a complete disappearance passage 14. Therefore, in the PHFK HPV18 cell line from passage 12 onwards there is a disruption in the continuity of HPV18 in E2/E5 region involved sequences, from 3086 to 3994. In the presence of the HPV18 E7 sequence this disruption indicates the reduction in the HPV18 episomal form commenced at passage 12, and was complete by passage 14. I used the HeLa cell line as a

control to validate my primer sets. I found no amplification of the long E1/E2 region. I found an intact E1/E2 region amplified by primer set 1 only but not using the other sets. This disruption of E2/E5 in HeLa involves the sequence from nucleotides 2828 to 3994 (Figure 48).

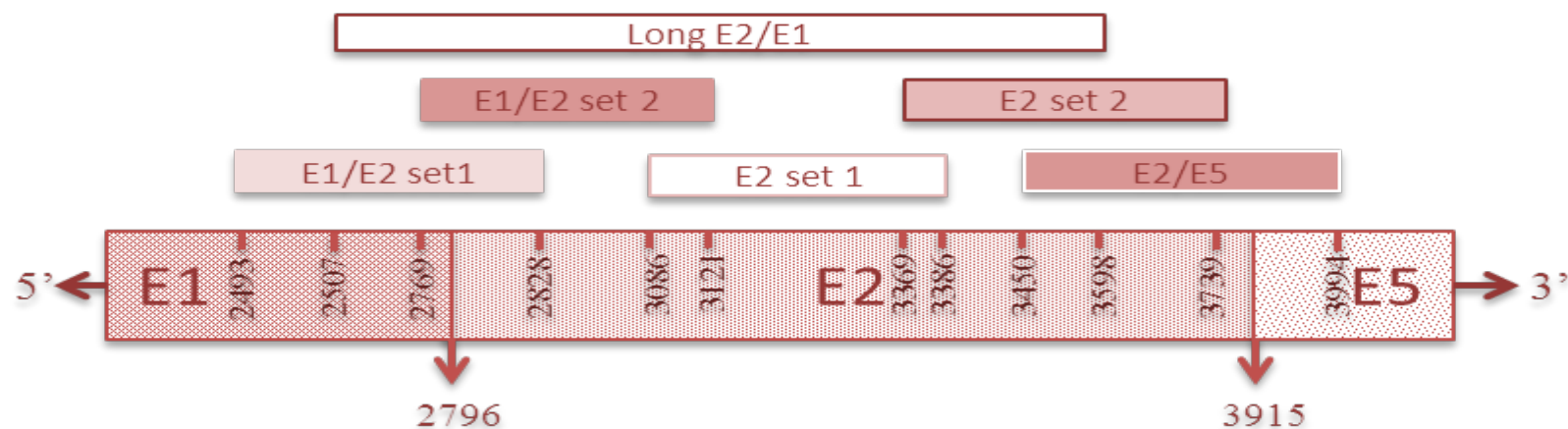


Figure 47: Topography of overlapping primers I designed and used to amplify the full length HPV18 E2 gene. HPV nucleotides are numbered according to the whole genome sequence of HPV18. The E2 set 2 primer was designed by Collins et al., 2009.

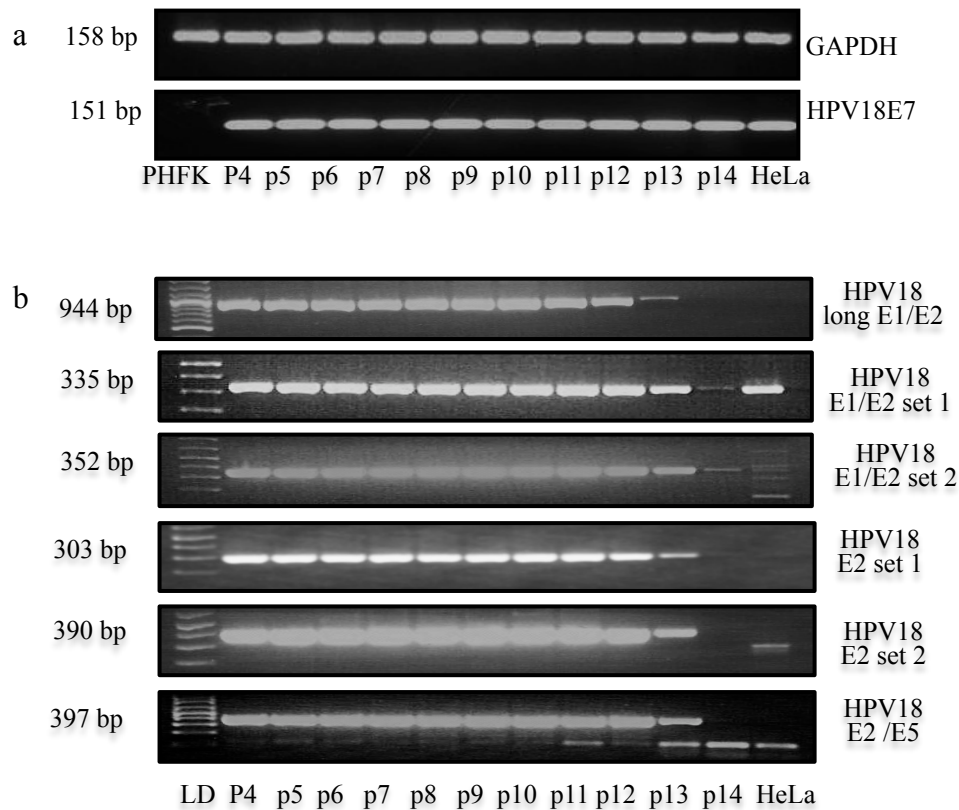


Figure 48: PCR results show assessment of HPV18 E2 integrity. a, Intact HPV18 E7 can be detected in all passages of the PHFKHPV18 cell line was used as a poitive control. b, Amplification of full length E2 region. Disruption of the E1/E2 region starts at passage 12 and is completely disrupted both by passage 14 and in HeLa cells. Five sets of PCR primers were used. In HeLa cells amplification was only visible with the E1/E2 set 1 primer. In PHFK HPV18 at passage14 amplification was only visible with the E1/E2 using set 1 and set 2 primers. Results shown are typical of 3 similar experiments.

### 4.2.2 Integrity assessment of full length HPV16 E2 in *in vitro* models

I used five sets of overlapping primers to assess the integrity of the full length HPV16 E2 gene (Collins et al., 2009) (Figure 49). First primer set cover the E1/E2 region (2701-3119), and three primer sets cover the E2 region from (3062-3785) and last primer set cover E2/E5 region (3710-3916).

#### 4.2.2.1 Integrity assessment of full length E2 in the PHFK HPV16B cell line

I found consistent amplification of HPV16 E6 and GAPDH in the PHFK HPV16B cell line from passage 4 to 10 (Figure 50a). I showed amplification of the complete HPV16 E2 gene, which has 1765 bp, in all passages. The PCR band for complete E2 at passage 10 showed less density compared to the previous passages (Figure 50b). These PCR results indicate that disruption of full length HPV16 E2 commenced at passage 10 in the PHFK HPV16B cell line. Further, I investigated the exact site of disruption using the five sets of E2 overlapping primers. The first and last primer sets that cover E1/E2 and E2/E5 regions showed amplification patterns similar to that showed to complete E2 primer, however, the E2/E5 sequence showed a more pronounced reduction, as reflected by a weak PCR band (Figure 50b). The other three primer sets that cover the E2 region showed equal amplification in all passages. in all passages.

#### 4.2.2.2 Integrity assessment of full length E2 in the PHFK HPV16K cell line

Studying the physical state of different HPV16 variants in the same PHFK *in vitro* models, gave me a clearer picture about how the different virus strains differ in their life cycles. I demonstrated intact HPV16 E6 in all passages (Figure 51a). The PHFK HPV16K cell line showed amplification of the full-length 1,765 bp sequences of E2 in passages 4 and 5 only, (Figure 51a). I found that full-length bands of HPV16 E2 gradually reduced from an early stage at passage 5 and completely disappeared by passage 6.



Further investigation of the site of disruption in Figure 51c showed complete failure to amplify the sequence from nucleotides 3710 to 3916 (primer set 5) from passage 6. In the other regions sequences were amplified equally at passages 4 and 5 then the intensity of the PCR bands gradually reduced. However, amplification was still visible in all passages with the other primer sets. These PCR results indicate an early disruption of full length HPV16 E2 in the E2/E5 regions. Thus in the PHFK HPV16K cell line episomal loss and the associated HPV integration occur early at passage 6.

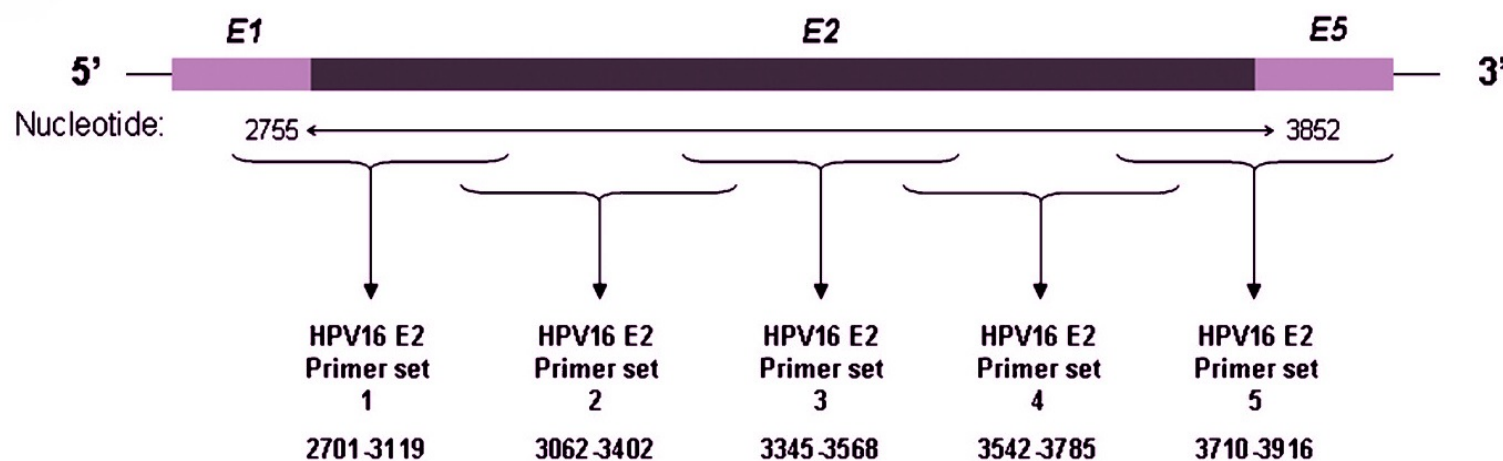


Figure 49: Topography of overlapping primers of the full length of the HPV16 E2 gene. HPV nucleotides are numbered according to the whole genome sequence of the HPV16, adapted from Collins et al., 2009.

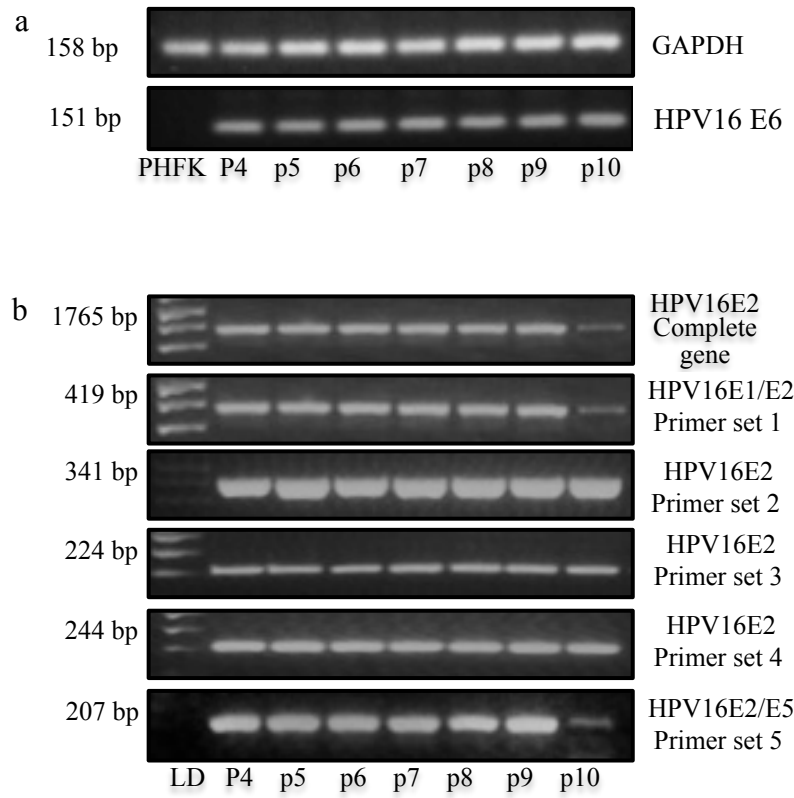


Figure 50: PCR results show assessment of HPV16 E2 integrity in the PHFK HPV16B cell line. a, Intact HPV16 E6 was seen in all passages. b, Amplification of the complete E2 region in all passages. Five sets of PCR primers were used and show disruption of the E1/E2 and E2/E5 regions that commenced at passage 10. DL, DNA Ladder. Passage (p). Results shown are typical of 3 similar PCR experiments.

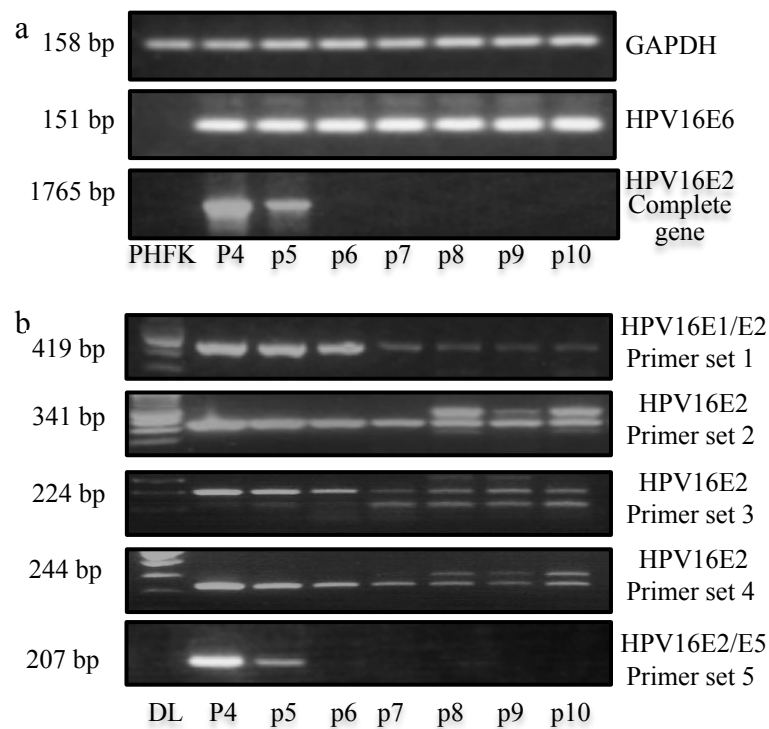


Figure 51: PCR results show assessment of HPV16 E2 integrity in the PHFK HPV16K cell line. a, Intact HPV16 E6 was seen in all passages. The complete E2 region gradually reduced amplification with a complete failure at passage 6. b, Five sets of PCR primers were used. They show a gradual reduction of the E2 region amplified by primer sets 1, 2, 3 and 4, and complete disappearance of the E2/E5 region commencing at passage 6. DL, DNA Ladder. Passage (p). Results shown are typical of 3 similar PCR experiments.

### 4.3 Kinetic changes of HPV16 E7 and HPV18 E7 protein

#### expression in *in vitro* models

As expected from the HPV life cycle, overexpression of the HPV E7 oncoprotein coincided with disruption of HPV E2. Using western blotting, I demonstrated a gradual overexpression of E7 oncoprotein in the PHFK HPV18 cell line (Figure 52 a). Interestingly, I found that a more marked overexpression of HPV18 E7 protein began at passage 10 prior to the appearance of disruption of full length HPV18 E2. HeLa was used as a positive control and showed a high expression of the HPV18 E7 protein. Calregulin was used as loading control, (Figure 52 b). In the PHFK HPV16B cell line, I showed expression of the HPV16 E7 protein only at passages 9 and 10. This again appeared prior to disruption of full length HPV16 E2. SiHa cell line showed a high level of HPV16 E7 expression.

I used western blotting to investigate the changes in expression level of pRb, which is a surrogate marker of E7 activity, in passages that expressed the E7 oncoprotein. Unfortunately, I was not able to investigate the kinetic changes of pRb in relation to E7 expression changes due to a shortage in cell extract for most of the passages. However, I was still able to investigate pRb protein expression in few passages. I showed a reduced level of pRb expression in the PHFK HPV18 and PHFK HPV16B cell lines at passages 12 and 10 respectively in relation to untransfected PHFK (Figure 52 c). These two passages showed increased expression of the HPV E7 oncoprotein.

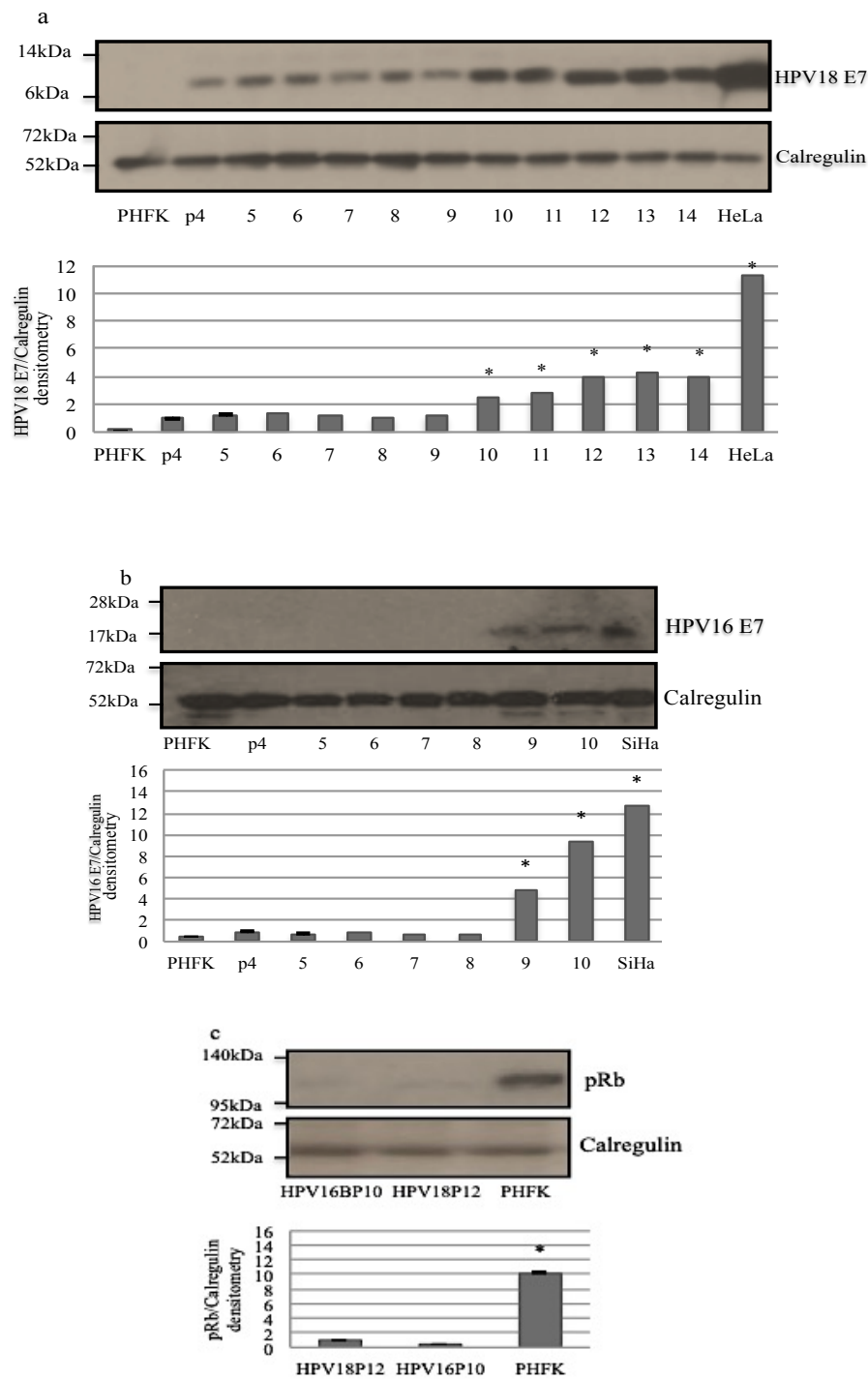


Figure 52: Expression profile of HPV E7 oncoprotein and pRb using western blot analysis and quantitative densitometry a, Gradual overexpression of HPV18 E7 oncoprotein in PHFK HPV18 cell line. b, Expression of HPV16 E7 oncoprotein, expressed in passages 9 and 10 PHFK HPV16B cell line. HeLa and SiHa cell lines were used as positive controls. c, pRb expression in PHFK HPV16B cell lines passage 10 and PHFK HPV18. Calregulin was used as a loading control. Passage (p). Results shown are typical of 3 similar experiments. \*  $p \leq 0.05$  versus control (Student's  $t$  test). Error bars indicate standard deviations.

In the next section I first examine DNMT3B gene copy number in my PHFK *in vitro* model and in W12 and cervical cancer cell lines, to search for a possible correlation with increased DNMTB expression. Next, I utilise a U2OS Tet-On inducible system to show the induction of DNMT3B expression by HPV16 E7. I also demonstrate the results of a co-immunoprecipitation (Co-IP) experiment showing an interaction between the E7 oncoprotein and DNMT3B. Then, I investigate DNMT1 and DNMT3B enzymatic activities in *in vitro* models. Lastly, I examine the methylation of tumour suppressor gene promoters, DAPK1 and VHL, before and after disruption of full-length E2.

## 4.4 Correlation of DNMT3B gene copy numbers with mRNA and protein expression

In Chapter 3 I showed overexpression of DNMT3B after transfection with HPVs. There were two patterns of expression in the *in vitro* models. In the first, expression peaked following transfection and the second showed an interesting two-peak pattern of expression in both PHFK HPV18 and PHFK HPV16B cell lines. I correlated the latter pattern with the PCR results of HPV E2 integrity and revealed that the second peak related to disruption of full-length E2 and oncoprotein E7 overexpression. This prompted me to determine whether there is an association between DNMT3B overexpression and possible DNMT3B gene copy number changes induced by E7 overexpression. I used q-PCR to study DNMT3B gene copy number in the PHFK *in vitro* model, and in W12 and cervical cancer cell lines.

My analysis showed no significant increase in DNMT3B gene copy number in the PHFK HPV16B cell line passage 10, PHFK HPV16K cell line passage 7 or PHFK HPV18 cell line passage 14 (Figure 53). These were the passages after E2 disruption. In contrast, I showed higher gene copy number in the late W12 cell line (Figure 53). Analysis of late passage W12 cells show the DNMT3B gene copy number increased by approximately four fold compared to the earlier passage. I also investigated the HPV-negative cell line C33A, which showed no increase in DNMT3B gene copy number. In the HPV-positive cell lines, in SiHa cells DNMT3B gene copy number increased significantly, while no changes were detected in both Ca Ski and HeLa cells. Overall, the DNMT3B gene copy number did not correlate with mRNA and protein overexpression in the *in vitro* models. It correlates however, with W12 and almost all cancer cell lines except the Ca Ski cell line, which showed no significant increase of DNMT3B gene copy number and yet a significant increase of its DNMT3B mRNA expression (Figure 32).



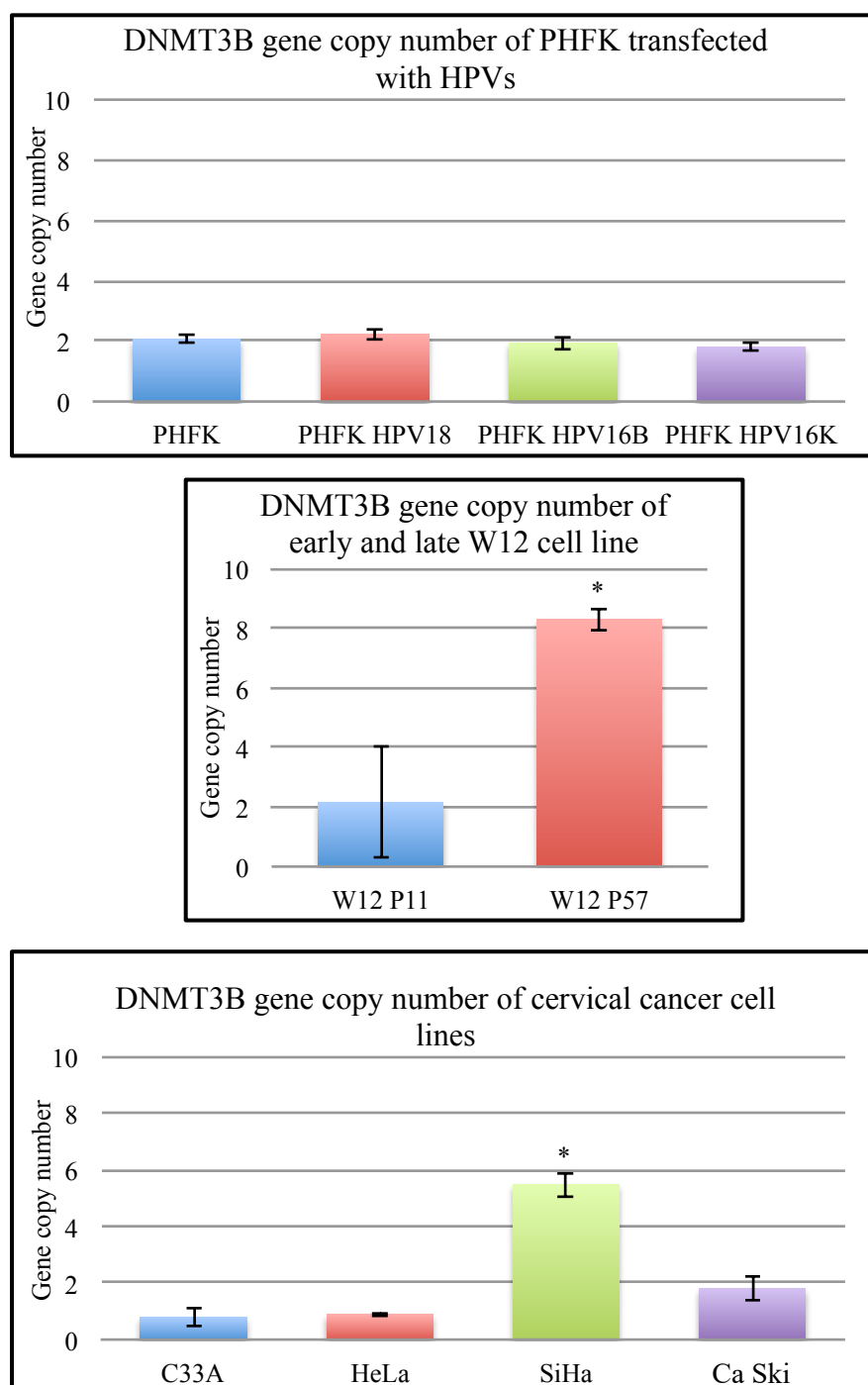


Figure 53: q-PCR analysis of DNMT3B gene copy number in the *in vitro* model system, the W12 cell line and cervical cancer cell lines. Only W12 P57 and SiHa cell line showed increases in DNMT3B gene copy number. Results shown are typical of 3 similar experiments. Each experiment was run in triplicate.\*  $p \leq 0.05$ . Error bars indicate standard deviations.

## **4.5 HPV16 E7 oncoprotein mediates the induction of DNMT3B expression**

My previous observations indicate an association between DNMT3B overexpression and overexpression of HPV E7. In Chapter 3 I showed an overexpression of DNMT3B1 and in this Chapter (section 4.3), I showed overexpression of HPV E7 oncoprotein, which precedes disruption of full length HPV E2. These results prompted me to study the possible effects of HPV integration and the subsequent overexpression of HPV E7 oncoprotein on the induction of DNMT3B overexpression in *in vitro* models. I used a U2OS Tet-On inducible system that has previously been shown (McLaughlin-Drubin, Margaret E. Crum, Christopher P. Münger, 2011) to induce HPV E7 to investigate its effect on another epigenetic regulator, H3K27me3.

### **4.5.1 Optimization of the U2OS Tet-On inducible system**

The effective doxycycline (dox) concentration required for induction in Tet-On systems is far below levels that are cytotoxic to cell culture. Using western blotting I showed HPV16 E7 protein expression of the U2OS Tet-On in response to three different doses of dox : 0.5, 1 and 1.5 µg/ml (Figure 54a). I determined the most effective concentration of dox needed to induce E7 to be 1 µg/ml as this showed the maximum E7 oncoprotein expression. U2OS without dox showed no expression of E7 oncoprotein. Loading control showed equal loading.

### **4.5.2 The HPV16 E7 oncoprotein induces DNMT3B protein expression in the U2OS Tet-On inducible system**

Having determined the optimum dose of doxycycline required for maximum E7 oncoprotein induction, this was associated with a more than 2-fold increase in the expression of the DNMT3B protein using western blot following the induction of HPV16 E7 with 1µg/ml

of dox. As expected, neither the U2OS Tet-On without dox, nor the U2OS-Ctrl showed HPV16 E7 expression, (Figure 54b). Loading controls show equal loading. These experiments were independently repeated three times.

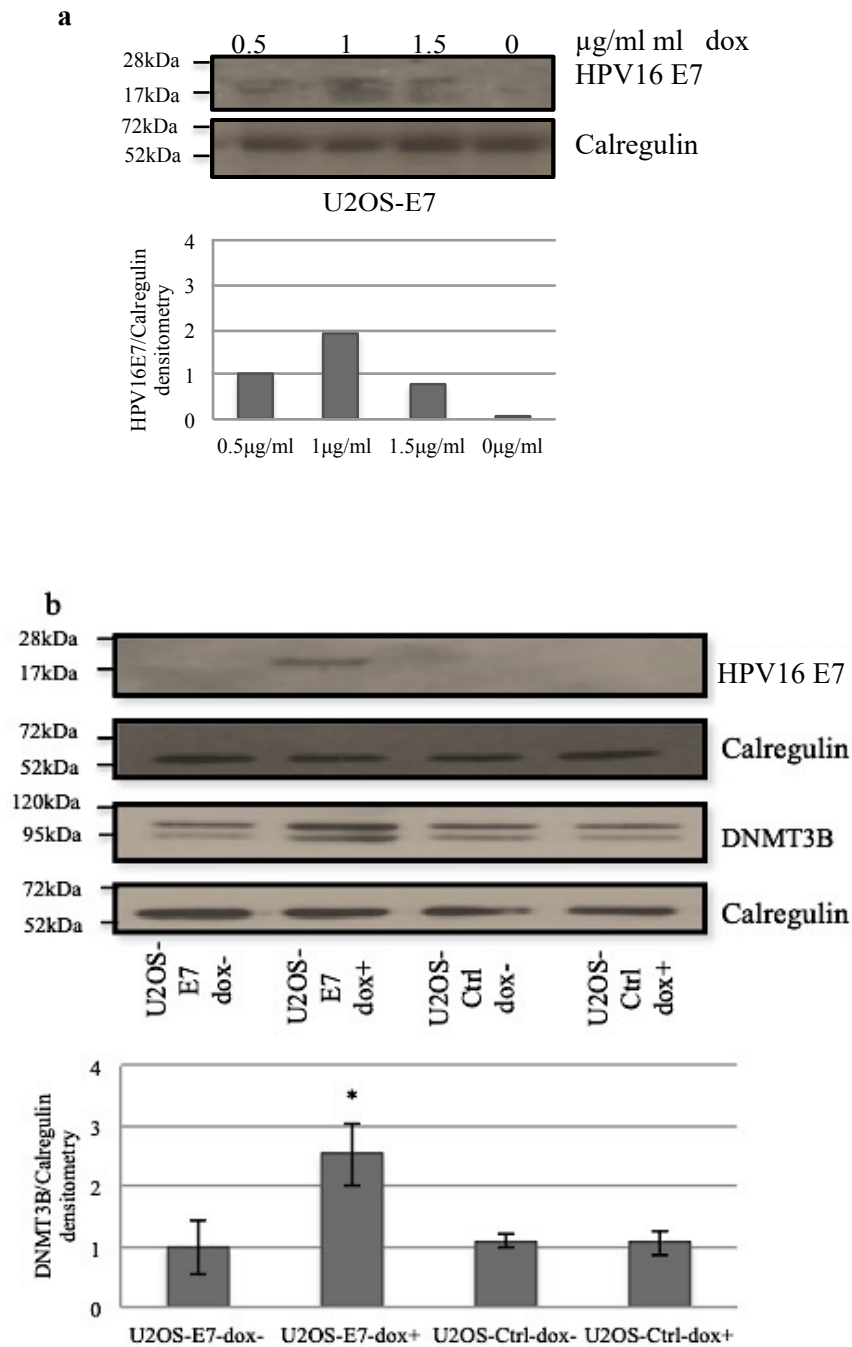


Figure 54: HPV16 E7-mediated induction of DNMT3B. a, Pilot study showing by western blot analysis and quantitative densitometry that maximum expression of HPV16 E7 was induced by 1  $\mu$ g/ml dox in U2OS Tet-On inducible cells b. Western blot analysis of HPV-E7 and western blot analysis and quantitative densitometry of DNMT3B protein expression and western blot analysis of E7 protein expression in the U2OS Tet-On system, results shown are typical of 3 independent experiments. \*  $p \leq 0.05$  versus control (Student's  $t$  test). Error bars indicate standard deviations.

## 4.6 The human papillomavirus oncoprotein E7 targets DNMT3B

In the previous section I showed that HPV E7 induces overexpression of DNMT3B. I next wanted to investigate whether this induction was due to a physical interaction between DNMT3B and E7 oncoprotein. It has been shown that E7 oncoproteins directly interact with DNMT1 and induce its overexpression (Burgers et al., 2007). I therefore investigated a possible interaction between HPV E7 oncoprotein and the DNMT3B protein using the gold standard method of investigating protein-protein interaction, co-immunoprecipitation.

To investigate a possible *in vivo* interaction between HPV E7 and DNMT3B at physiological levels, I used cellular extracts. I obtained cellular extracts from both PHFK HPV18 cells and PHFK HPV16B cells of passages 14 and 10, respectively. I also included cellular extracts from HeLa and SiHa cancer cell lines. I have previously showed that the chosen passages of these *in vitro* models and the cervical cancer cell lines both have increased expression DNMT3B and E7. Using western blotting I show confirmation of DNMT3B protein immunoprecipitation (Figure 55 a). Interestingly, (Figure 55 b) shows that the HPV oncoprotein E7 is associated with the *de novo* methyltransferase, DNMT3B in cervical cancer cell lines and in *in vitro* models. DNMT3B co-precipitated with endogenous HPV16 E7 and HPV18 E7 and did not precipitate with control IgG antibody, as shown with western blotting results bottom row (Figure 55 b).

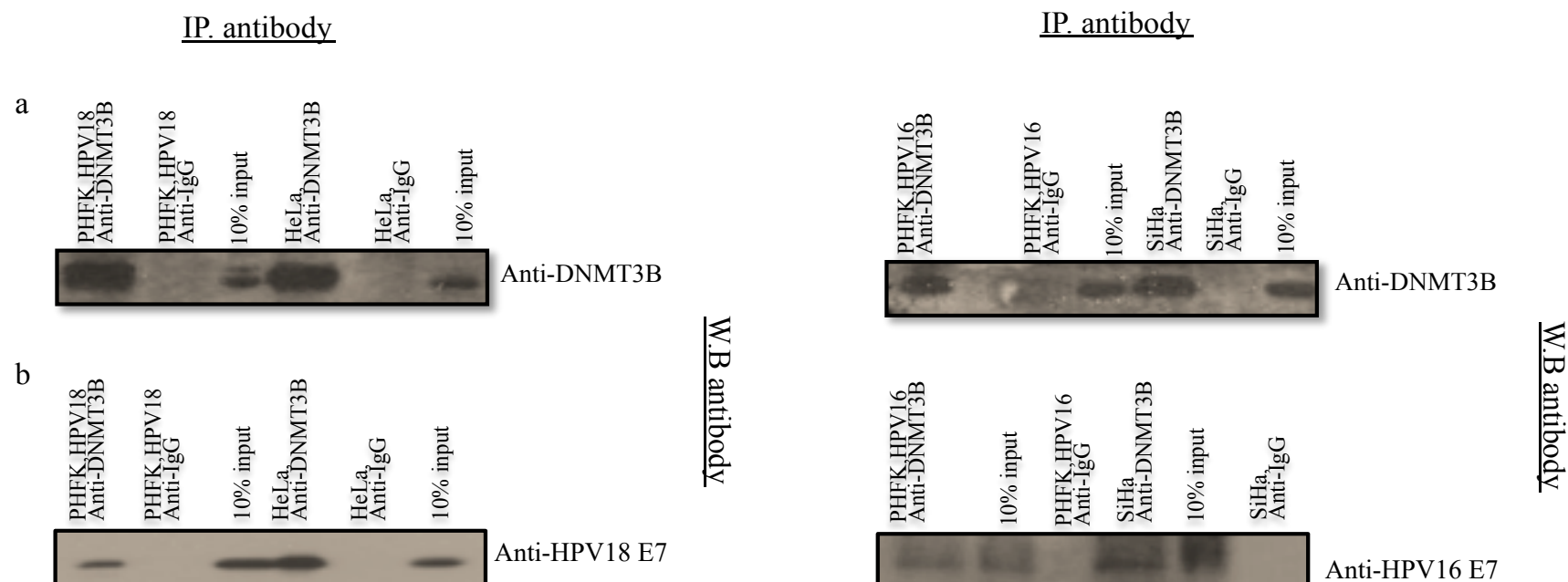


Figure 55: The HPV oncoprotein E7 associates with DNMT3B in PHFK *in vitro* models and cervical cancer cell lines *in vivo*. a) Immunoprecipitation of DNMT3B confirmed in PHFK HPV18, PHFK HPV16B and cervical cancer cell lines, HeLa and SiHa (upper part). b) DNMT3b co-immunoprecipitated with HPV18 E7 and HPV16 E7 proteins subjected to SDS-PAGE and Western blot analysis (lower part). 10% Input: represent 10% of the amount of protein used in the co-immunoprecipitation experiment. Results shown are typical of 3 similar independent experiments.

## 4.7 The E7 oncoprotein induces DNMT enzymatic activity

To determine whether overexpression of DNMTs in *in vitro* model associated with an increase of DNMTs enzymatic activity, I investigated the endogenous enzymatic activity of DNMT1 and DNMT3B in *in vitro* model, PHFK HPV18 cell line and PHFK HPV16B cell line kinetically. This investigation involved studying cellular extracts from both PHFK HPV18 and PHFK HPV16B cell lines after transfection and following disruption of full length E2.

### 4.7.1 Kinetic changes of DNMT3B enzymatic activity in *in vitro* models

Figure 56 shows that increased enzymatic activity of DNMT3B coincides with the pattern of increased expression of DNMT3B seen in the Chapter 3. This demonstrates that a more pronounced activity also correlates with increased expression of the E7 oncoprotein. Relative to untransfected PHFK, DNMT3B activity at passage 5 was three times greater in both PHFK HPV18 and PHFK HPV16B cell lines (Figure 56a and b). However, this increase was transient. At late passages this level markedly rose to ten times that of pre-transfection levels. The DNMT3B activity showed a period of quiescence preceding the overexpression of E7, demonstrated at passage 8 in both the *in vitro* models involved in my experiments.

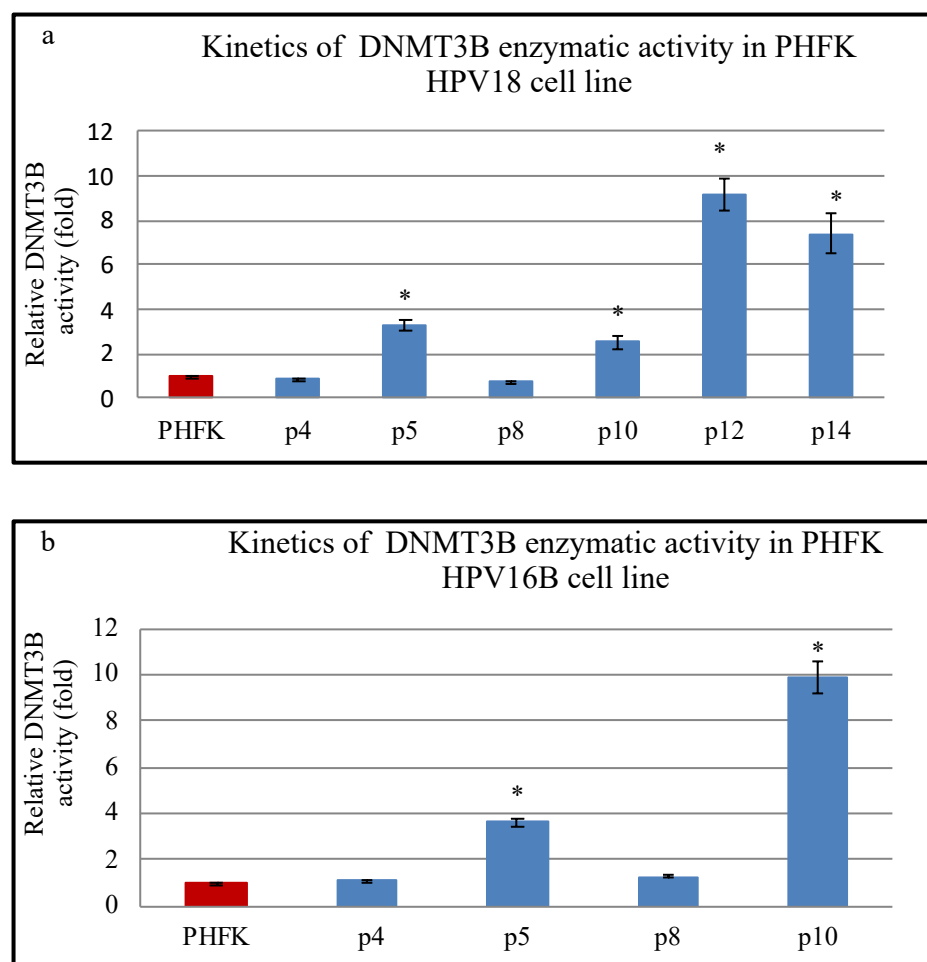


Figure 56: Results of DNMT3B enzymatic activity assay. a, Shows kinetic changes of DNMT3B enzymatic activity in PHFK HPV18 cell line, passage 5, 12 & 14 showed increased activity compare to PHFK. b, Shows kinetic changes of DNMT3B enzymatic activity in PHFK HPV16B cell lines, passage 5 & 10 showed increase of DNMT3B enzymatic activity compared to PHFK. Results are the average of two independent experiments. \* Indicates  $p \leq 0.05$ . Passage (p).



#### **4.7.2 Kinetic of changes in DNMT1 enzymatic activity in *in vitro* models**

My next question was whether expressed E7 can increase DNMT1 enzymatic activity in particular in human keratinocytes which are the natural host of high-risk HPVs. Interestingly, (Figure 57) shows the two peak pattern of enzymatic activity that follows the pattern of DNMT1 expression. It demonstrates a transient increase in DNMT1 enzymatic activity following transfection of PHFK with both HPV18 and HPV16B. A period of inactivity is then followed by a more stable rise. The later increase in enzymatic activity coincides with disruption of full length E2 in both the PHFK HPV18 and PHFK HPV16B cell lines (Figure 57a and b). Thus, endogenously expressed HPV16 E7 and HPV18 E7 oncoproteins in PHFK correlated with an increase in DNMT1 enzymatic activity in PHFK.

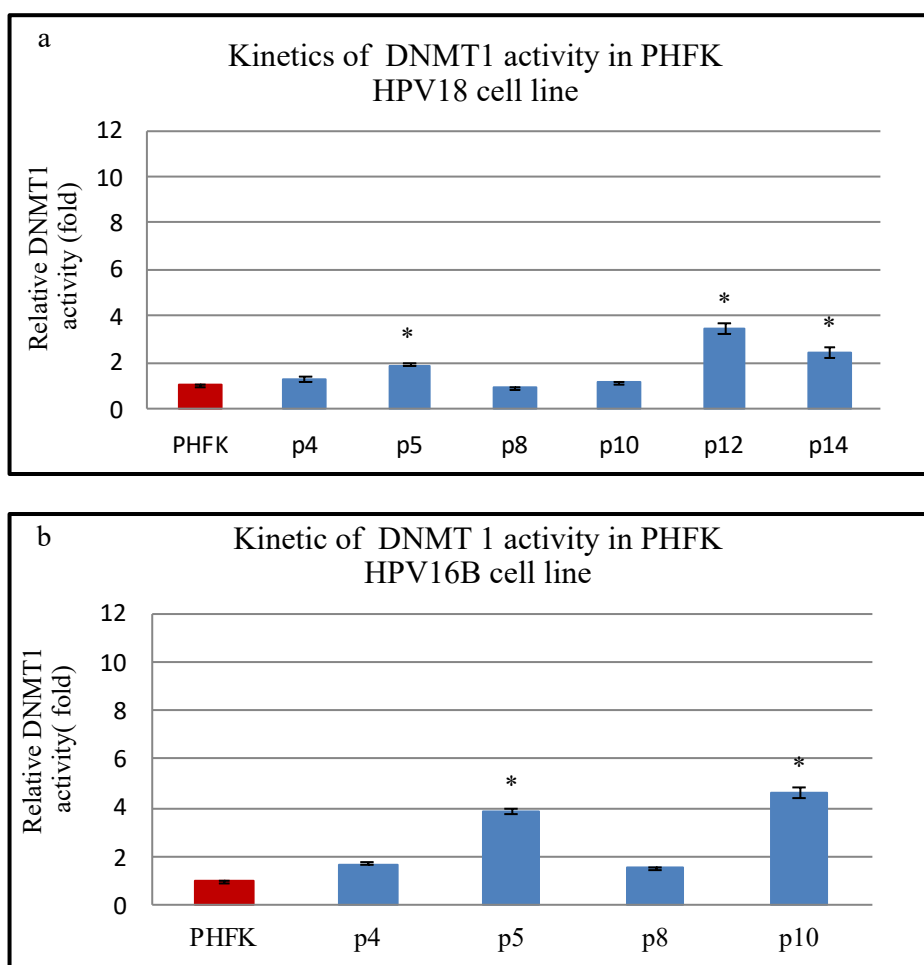


Figure 57: Results of a DNMT1 enzymatic activity assay. a, Shows kinetic changes of DNMT1 enzymatic activity in the PHFK HPV18 cell line. b, Shows kinetic changes in the DNMT1 enzymatic activity of PHFK HPV16B cell line. Results are the average of two independent experiments. \* Indicates  $p \leq 0.05$ . Passage (p).

## 4.8 Overexpression of the E7 oncoprotein associated with changes in DNA methylation of the TSGs

My next question was whether the induction and increase in DNMTs enzymatic activity is associated with aberrant DNA methylation of tumour suppressor gene (TSG) promoters? Hypermethylation of TSG promoters in cervical carcinogenesis have been extensively investigated over the last two decades. One of the TSG, which has been proposed as a candidate methylation marker for the early detection of cervical cancer, is the death-associated protein kinase (DAPK1) gene. The DAPK1 gene causes the apoptosis of transformed cells through activation of the p53 protein. Studies have found that progressive hypermethylation of the DAPK1 promoter correlates with the progression of cervical neoplasia (Feng et al. 2005). VHL is the Von Hippel-Lindau disease tumour suppressor. VHL is found on 3p25.3. It reported that this gene's promoter is hypermethylated in the early stage of cervical carcinogenesis. DNA hypermethylation can be detected in HSIL compared to normal cervical cells when cervical exfoliated cells are investigated (Feng et al., 2005). I evaluated DNA methylation of the DAPK1 and VHL gene promoters using the pyroSeq technique. This experiment involved studying the PHFK HPV18 cell line at passages 7 and 14. The amplified promoter region involves six and four CpG sites for DAPK1 and VHL, respectively.

The pyrogram of the DAPK1 promoter (Figure 58a), shows that the average methylation across the six investigated CpG sites increased from 12% to 27% between passage 7 and passage 14. Specifically, the last two CpG sites showed significant changes. The pyrogram shows that the methylation of one of the CpG sites increased from 7% to 15% and the other showed *de novo* methylation since methylation increased 0% to 5%. Figure 58b shows the pyrogram of the VHL gene. The average DNA methylation of the four CpGs increased from

51% to 67%. Of the three hypermethylated CpGs, methylation doubled in the last two by p14. These pyroSeq results present important preliminary evidence suggesting that aberrant methylation changes of TSG promoters can be induced early in the processes of carcinogenesis by HPV integration. Thus, this prompted me to investigate the genome-wide DNA methylation changes following overexpression of the HPV18 E7 oncoprotein, which are explored in detail in the next chapter.

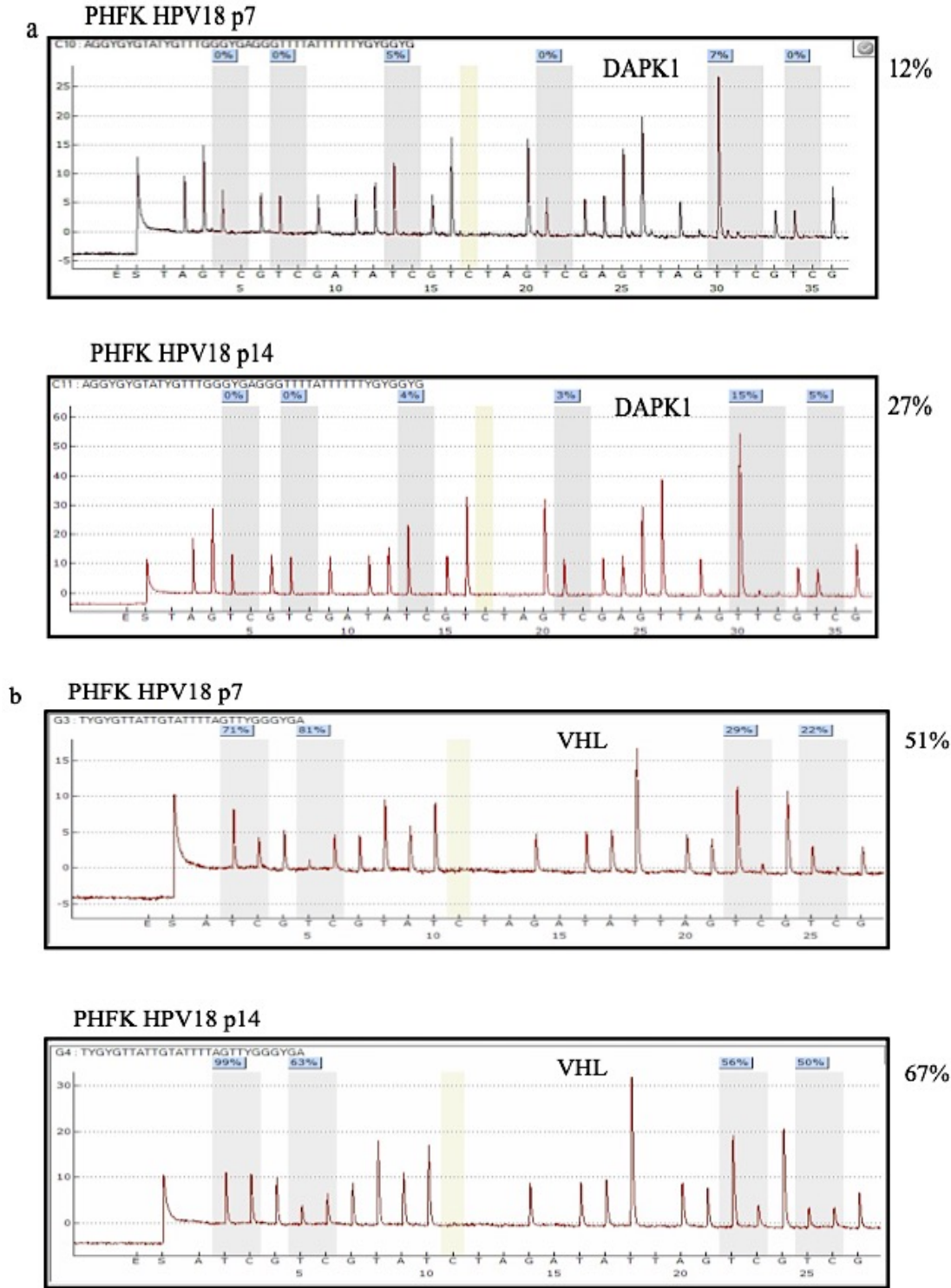


Figure 58: Quantification of the CpG methylated sites in the promoter region of the DAPK1 and VHL gene by pyroSeq. In the resulting pyrogram, the grey bars indicate the CpG sites that were analysed. Yellow bars indicate controls for bisulfide conversion. The X axis from left to right, indicates the order of reagent addition. The Y axis represent average DNA methylation. a. Pyrogram showed increased average methylation in PHFK HPV18 p7 compared with p 14 shown in b. Top: compare passage 7 and 14 of DAPK1. Bottom: compares passage 7 and 14 of VHL.

In summary, in the first section of this chapter I demonstrated the late disruption of HPV18 full length E2 at the E2/E5 region using the PCR method. Disruption of full length HPV18 E2 was followed by and coincided with over-expression of active E7 oncoprotein. I also showed a mixture of early and late disruption of HPV16 full length E2 in two cell lines. E2 disruption happens at the E1/E2/E5 region. I showed an active expression of HPV16 E7, which is followed and associated with E2 disruption.

In the second section I found no changes in DNMT3B gene copy number in my *in vitro* models, however I confirmed the induction of DNMT3B by HPV E7 oncoprotein. I also showed a physical interaction between DNMT3B and HPV E7 oncoprotein. Finally, I observed increased DNA methylation of two TSG, DAPK1 and VHL promoters following disruption of full length HPV18 E2 in PHFK HPV18 cell line

## CHAPTER 5

### RESULTS III

# **INTEGRATION OF HPV18 IS ASSOCIATED WITH CHANGES TO DNA METHYLOME OF HUMAN KERATINOYTES**

**Aim of results III:**

To explore the epigenetic changes following integration of high-risk human papillomavirus into human cellular DNA.

**Objectives:**

7. To investigate the association between HPV18 integration and changes in DNA methylation using a genome-wide approach (methylation array).
8. To determine the frequency of occurrence of changes in DNA methylation in cervical neoplasia following HPV18 integration.
9. To investigate the possible HPV18 integration-induced ‘epigenetic switching’, using both gene candidate and gene pathway approaches.
10. To study the effect of demethylating agents using an *in vitro* model, PHFK HPV18 cell line, following HPV integration.



In the first section of this chapter, I use a genome wide methylation array to analyse DNA methylation changes following HPV18 episomal loss. Firstly, I use the pyrosequencing method to validate the DNA methylation changes predicted by the methylation array. Then, I describe the global DNA methylation changes following HPV18 episomal loss and in more detail, I describe DNA methylation changes that impact on important biological pathways. I also explore DNA methylation changes within the regulatory regions of tumour suppressor genes and methylation markers of cervical cancer within the most significant hypermethylated genes following HPV18 integration. Lastly, I identify *de novo* methylated genes following HPV18 integration and determine their

## 5 Results III

### 5.1 Introduction of results III

Methylation markers of cervical cancer offer promise to enhance the effectiveness of current cervical screening programmes. A Phase III randomised controlled trial suggested that methylation marker analysis enables an earlier diagnosis of cervical dysplasia compared to current cervical screening methods (Verhoef *et al.*, 2014). We analysed genome-wide methylation differences following HPV infection using PHFK and PHFK HPV16 and PHFK HPV18 (Leonard *et al.*, 2012). Information about the specific epigenome-wide differences in DNA methylation based on HPV integration, and in particular HPV18 integration, remains unknown. In the previous chapter, I have shown the active overexpression of DNA methyltransferases following high-risk HPV episomal loss. Consequently, DNA methylation changes are expected. I have also demonstrated the hypermethylation of CpGs in the promoters of the tumour suppressor genes DABK1 and VHL following overexpression of the HPV oncoprotein and HPV18 episomal loss. In this chapter, the following questions will be addressed. Is the overexpression of the HPV oncoprotein associated with global changes in DNA methylation? And, if so, what is the pattern of these early changes?

PHFK *in vitro* models offer the possibility of investigating methylation changes very early following the expression of E7 transcripts (Figure 51). To answer my questions and investigate the DNA methylation changes following the integration of HPV18 into the human genome, I employed a genome-wide method, the Illumina HD 450K methylation array.

## 5.2 Genome-wide DNA methylation profiling of the PHFK cell line following HPV18 integration

The PHFK HPV18 cell line is the best available *in vitro* model for genome-wide DNA methylation studies. I have previously investigated changes in the expression of epigenetic regulators in this model, and I showed that overexpression of E7 commenced at p10 (Figure 51). There was a transient overexpression of DNMTs following HPV transfection, as I showed in chapter 3. We have shown transient changes in the human methylome following transfection of the PHFK HPV18 cell line (Leonard *et al.*, 2012). Following transfection, a period of quiescence preceded the HPV episomal loss, commencing at p7. Thus, I selected p7, in which the expression of mRNA DNMTs is similar to that of the PHFK cell line to ensure that any methylation changes will be due to HPV18 integration. In comparison to p7, the last passage of the PHFK cell line (p14) showed the loss of E2 and the overexpression of the E7 oncoprotein.

### 5.2.1 Descriptive analysis of the Illumina HD 450K methylation array of the PHFK cell line following HPV18 integration

Of the 484,968 CpGs studied, significant DNA methylation differences between p7 and p14 were observed at 8.7% of the sites (42,355 CpGs) (Figure 59). Interestingly, most of the observed DNA methylation changes following HPV18 episomal loss corresponded to CpG hypomethylation, 77.5% (23,276 CpGs), whilst hypermethylation accounted for 22.5% (9,540 CpGs). Following HPV18 integration, the number of significantly hypomethylated genes (12,254) was twice that of hypermethylated genes (6,026), (Figure 60). Illumina array coverage is targeted across gene regions from 5' to 3' end with sites in the promoter regions, 5'UTR, first exons, gene bodies, and 3'UTR. The promoter is the upstream region of DNA that initiates transcription, about 100–1000 bp long. It is covered in Illumina arrays by two regions upstream

from the transcription start site (TSS 1500 and TSS 200). TSS200 is the region from the transcriptional start site (TSS) to 200nt upstream, whereas TSS 1500 covers 200 to 1500nt upstream of these sites. This multiple-site approach is extended to CpG island regions with multiple sites within islands and island shores, as well as island shelves (Figure 61).

Approximately one-third of the CpG hypermethylation occurred in CpG islands (32%) and another third was in the proximal promoter (33.5%) (Figure 62). On the other hand, hypomethylation events mainly occurred in the body of the genes and open sea areas of the genome, including intergenic regions, with a value, altogether, of over 90% (Figure 63). For those hypomethylated CpGs observed in the proximal promoters (25.3%) of PHFK with integrated HPV18, less than 10% occurred in CpG islands (Figure 63). There was no association between the type of transcripts and DNA methylation gain or loss. Overall, the chromosomal locations of the hypermethylated and hypomethylated CpGs were equally distributed between autosomal chromosomes. The sex chromosome showed increased susceptibility to DNA methylation changes. Both X and Y chromosomes showed differential methylation changes of 8.5 and 13% of their CpGs, respectively. However, compared to the average differential methylation changes of autosomal chromosomes, these differences were not statistically significant ( $p = 0.2$ ).

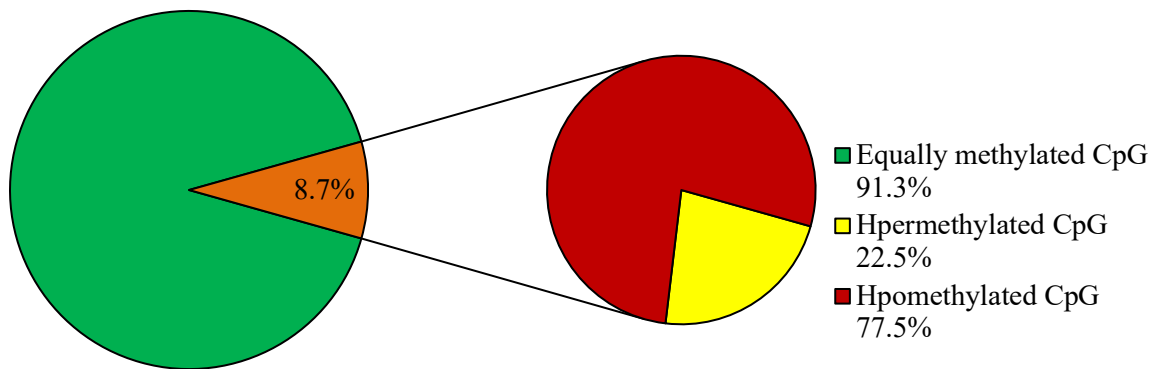


Figure 59: The percentage of differentially methylated CpG sites in the PHFK HPV18 cell line before (p7) and after (p14) HPV integration.

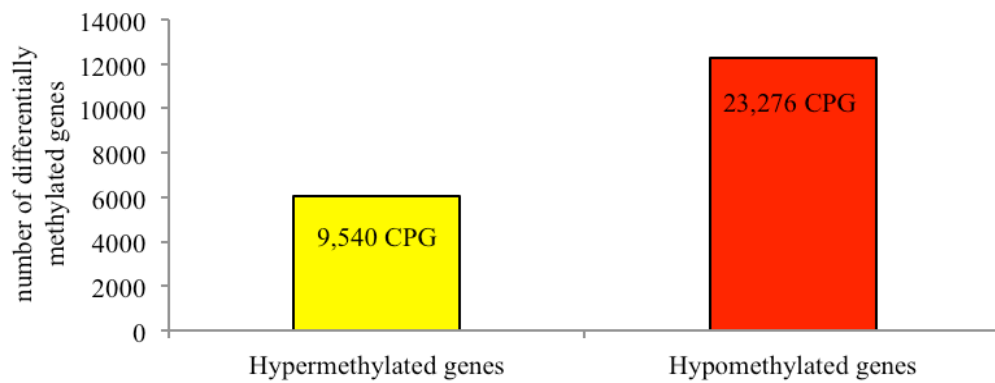


Figure 60: The total number of hypermethylated and hypomethylated CpG sites in the PHFK HPV18 cell line following HPV integration.  $p \leq 0.05$  and at a differential methylation of at least 10%.

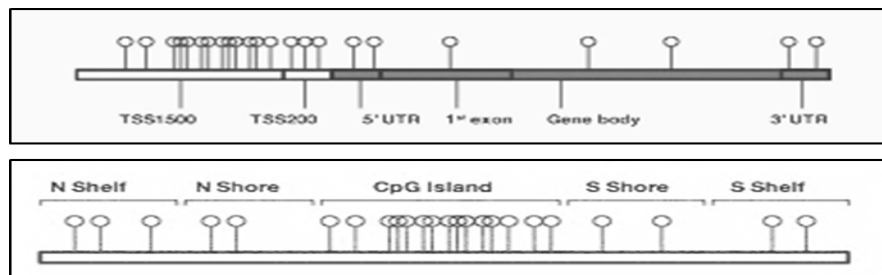


Figure 61: Illumina methylation arrays provide genes and CpG islands coverage. Adapted from Illumina website.

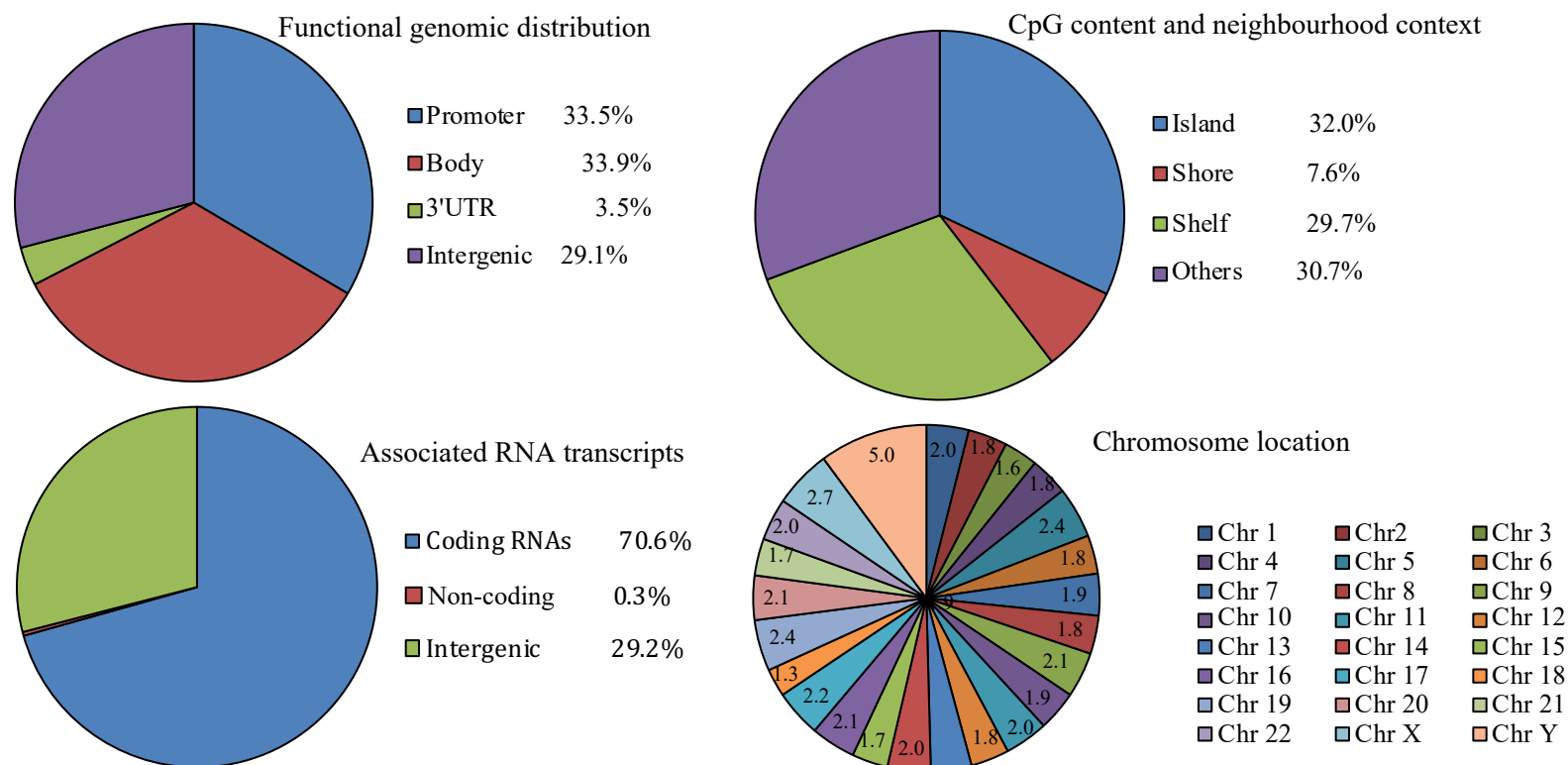


Figure 62: An Illumina 450K DNA methylation portrait of the distribution of hypermethylated CpGs in the PHFK HPV18 cell line following HPV integration. The distribution is shown according to their functional genomic distribution, CpG content and neighbourhood context, and associated RNA transcription and chromosome location (%).

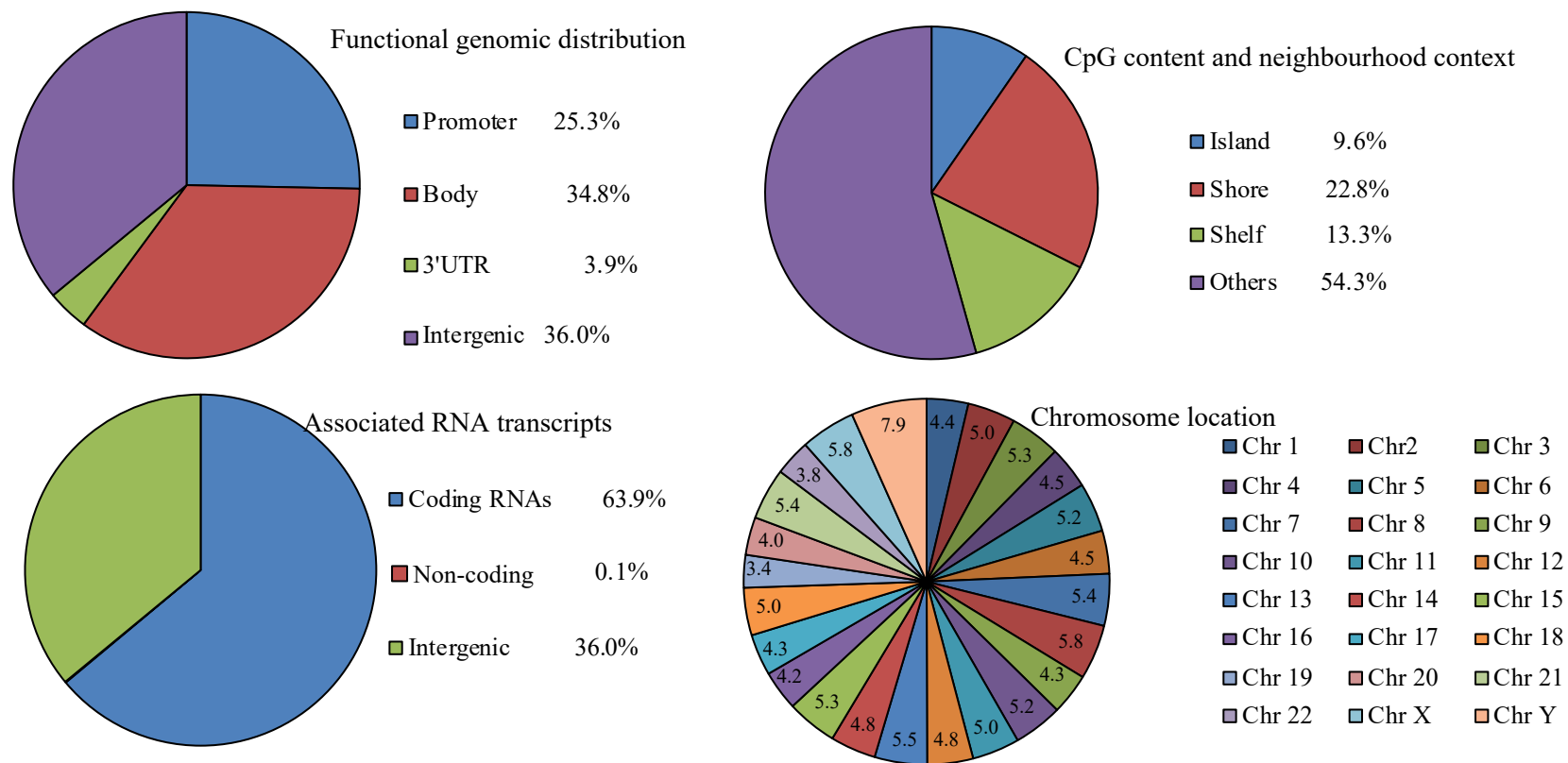


Figure 63: An Illumina 450K DNA methylation portrait of the distribution of hypomethylated CpGs in the PHFK HPV18 cell line following HPV integration. The distribution of is shown according to their functional genomic distribution, CpG content and neighbourhood context, and associated RNA transcription and chromosome location (%).

## **5.2.2 Validation of methylation changes following HPV18 integration predicted by the Illumina HD 450K methylation array**

### **5.2.2.1 Criteria for selecting candidate genes for validation**

In order for a candidate gene to be selected for validation, it should be both statistically and biologically significant. Thus, genes should fulfil at least three of the following criteria:

1. Demonstrated promoter hypermethylation between passage 7 and passage 14 of the PHFK HPV18 cell line.
2. Differentially methylated, i.e., a change in methylation of at least 10-20%.
3. A statistically significant methylation change ( $\text{DiffScore} \geq 13$  and  $p \leq 0.05$ ).
4. Known to be methylated at some time in the development of cervical neoplasia.

### **5.2.2.2 Cross-validation of pyrosequencing and the Illumina HD 450K methylation array**

As pyroSeq is accepted to be the gold standard for quantifying CpG methylation (Potapova et al., 2011). I selected 11 genes that fulfilled the above criteria (Table 17). Nine genes fulfilled all four required criteria, whilst the TP53AIP1 and DCC genes have not shown to be differentially methylated in cervical cancer. However, both have a biologically significant association with HPVs and cervical cancer. TP53AIP1 bind to and encodes an activator of p53 (Yuan et al., 2012), which is a tumour suppressor gene targeted by HPV (Doorbar et al. 2012). DCC is downregulated in cervical carcinoma; however, the mechanism of downregulation has not been explored (Saegusa and Okayasu, 1999). I designed high-quality primers that targeted the specific CpGs covered in the Illumina HD 450K array. However, the CHFR primer failed to amplify the required sequence and, therefore, the CHFR gene was excluded. Altogether, I compared the methylation level of 20 CpG sites in 10 genes in the PHFK HPV18 cell line p7



and p14 samples to quantitative pyroSeq methylation data. Table 18 shows a summary of the pyroSeq results, expressed as the percentage of methylation at each CpG, along with the average methylation across each region, for CpGs in the promoter regions of the 10 selected genes at p7 and p14. The pyrograms of the selected genes at both p7 and p14 are included in Appendix I. The average beta values of technical replicates of p7 and p14 were converted to absolute values to facilitate the comparison to the percentage of methylation from the pyroSeq results.

### 5.2.2.3 Analysis of cross-validation using a correlation coefficient method

The percentages of methylation were ranked in order of the percentage of CpGs that were methylated, as determined by the pyroSeq method. The correlation coefficient between pyroSeq and the Illumina methylation array data were then calculated. I compared the absolute methylation value of CpGs from the Illumina methylation array to either the percentage of identical CpGs or the average number of CpGs from the pyroSeq method. I found a good concordance between both methods. The results of comparing either methylation at identical CpGs, or at an average number of CpGs using these methods gave a Spearman's rank correlation coefficient  $\rho = 0.70$ ,  $p < 0.0005$  and Spearman's rank correlation coefficient  $\rho = 0.77$ ,  $p < 0.005$ , respectively (Figure 64a and b). However, the calculation of a correlation coefficient or a regression coefficient can be misleading because the scatter plots are dominated by an agglomeration of data points near the origin of the scatter plot. Thus, I used a more specific statistical analysis to compare the two methods which is the Bland-Altman plots method.

#### **5.2.2.4 Analysis of cross-validation using the Bland-Altman plots method.**

According to Altman & Bland (1983), when comparing differences between methods, it is preferable to plot the differences against the average to assess the magnitude of the disagreement (both error and bias). This type of statistical analysis, called ‘Limits of agreement’, is a very powerful way of displaying the results of a method comparison study. Thus, the difference between the two methods for comparing CpG methylation (i.e., conventional quantitative pyroSeq versus the Illumina HD 450K methylation array) was plotted against the mean of the two methods. The confidence limits, in which 95% of the differences are expected to lie, were also plotted, as it is the confidence intervals that show the spread of variation between methods. The limits were plotted as dotted horizontal lines of two standard deviations either side of the means. All data points fell inside the 95% limits of agreement (marked by the dotted lines). All the identical CpG sites targeted by the pyroSeq assay and the Illumina HD 450 K array were in concordance (Figure 64c and d). Thus, pyroSeq confirmed the array predictions. However, when the comparison was made between the average levels of CpG methylation of the regions analysed, the limit of agreement between the two methods was reduced. Approximately 15% of the investigated pyroSeq CpGs were outliers, which accounted for the disagreement with the representative CpGs from the Illumina method.

Table 17: Genes selected for validation of methylation changes predicted by Illumina methylation arrays following HPV18 episomal loss. The PIMD reference shows hypermethylation at some time during the process of cervical neoplasia for all except TP53AIP1 and DCC; for a DiffScore  $\geq \pm 13$  the  $p$ -value  $\leq 0.05$ , and for a DiffScore  $\geq \pm 22$ , the  $p$ -value  $\leq 0.01$ .

UCSC ref Gene	PMID	p7	p14	Delta $\beta$	DiffScore
CADM1	(Steenbergen et al., 2004)	0.1	0.2	0.1	17
CALCA	(Wisman et al., 2006)	0.1	0.2	0.1	17
ESR1	(Wisman et al., 2006)	0.3	0.4	0.1	19
PTPN6	(Kim et al., 2010)	0.3	0.5	0.2	21
RARB	(Narayan et al., 2003)	0.2	0.3	0.1	22
SFRP2	(Chung et al., 2009)	0.1	0.2	0.1	15
SOX1	(Lai et al., 2010)	0.2	0.3	0.1	26
TNFRSF10C	(Henken et al., 2007)	0.2	0.3	0.1	22
TP53AIP1	(Yuan et al., 2012)	0.4	0.5	0.1	15
DCC	(Rickinson, 1999)	0.1	0.2	0.1	18

Table 18: Summary of pyroSeq results of the ten selected genes. Identical CpGs from the pyroSeq and Illumina 450K array are highlighted. All numbers represent the percentage of methylation.

Sample	UCSC Gene Name	CpG 1	CpG 2	CpG 3	CpG 4	CpG 5	AVG methylation
PHFK HPV18 p7	CADM1	3	4				3
PHFK HPV18 p14		9	9				9
PHFK HPV18 p7	CALCA	0	5	0			2
PHFK HPV18 p14		5	3	2			3
PHFK HPV18 p7	DCC	3	3				3
PHFK HPV18 p14		5	7				6
PHFK HPV18 p7	ESR1	7	27	20	12		16
PHFK HPV18 p14		9	29	24	16		20
PHFK HPV18 p7	PTPN6	22	52	45	15	15	30
PHFK HPV18 p14		31	51	58	20	20	36
PHFK HPV18 p7	RARB	12	7	7	5	3	7
PHFK HPV18 p14		17	14	17	8	14	14
PHFK HPV18 p7	SFRP2	30	8	12	6		14
PHFK HPV18 p14		43	11	16	8		20
PHFK HPV18 p7	SOX1	3	4	3	4	4	3
PHFK HPV18 p14		4	9	8	7	7	7
PHFK HPV18 p7	TNFRSF10C	34	16	5	9		16
PHFK HPV18 p14		53	36	20	35		36
PHFK HPV18 p7	TP53AIP1	70	42	20			44
PHFK HPV18 p14		83	59	37			59

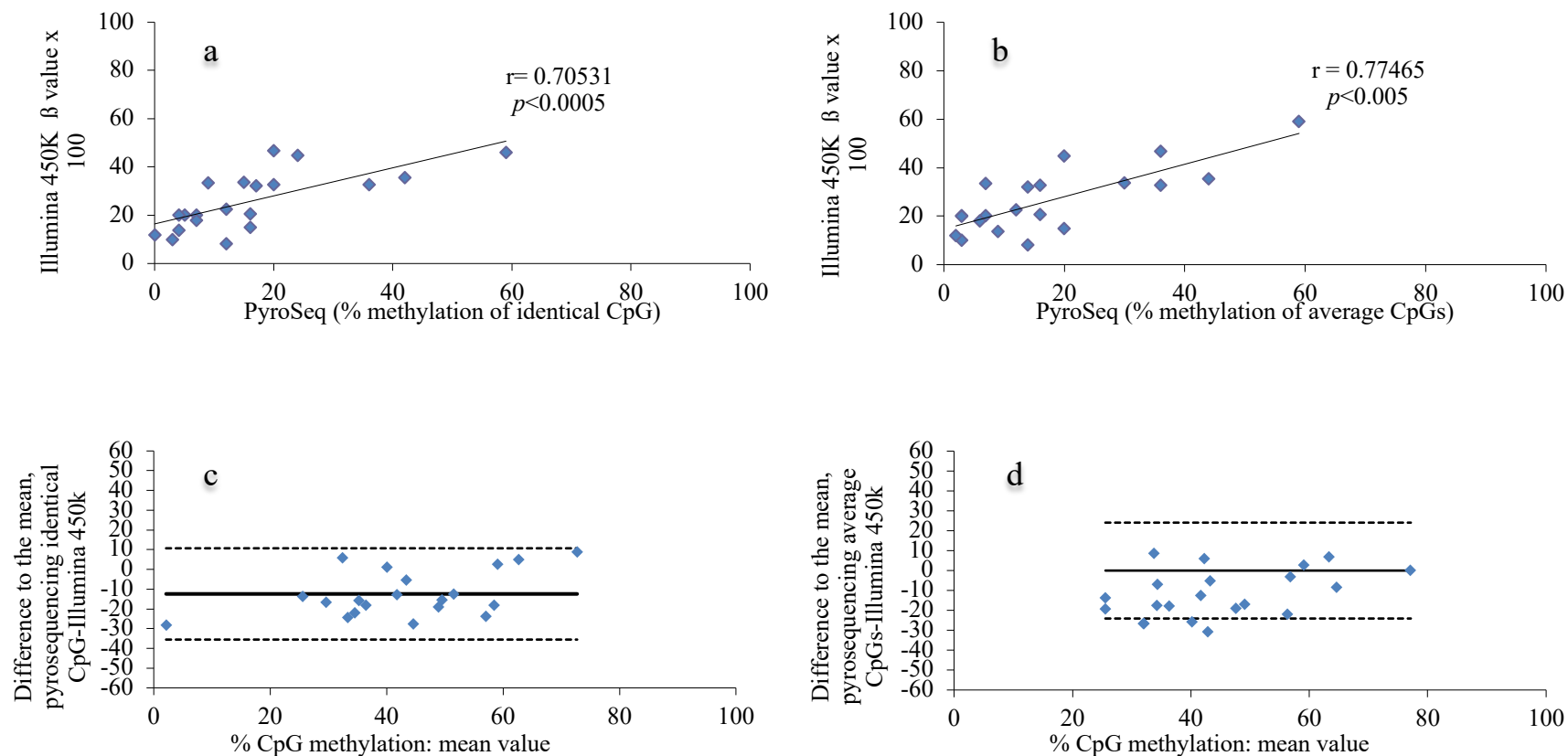


Figure 64: Cross-validation of Illumina HD 450K methylation array and quantitative pyroSeq with two statistical methods. 1. Scatter plots for absolute methylation values obtained from the Illumina HD 450K methylation array and the percentage of methylation of a) identical CpGs, b) average CpGs obtained from PyroSeq. 2.c) and d) show the corresponding Bland-Altman plots.

#### **5.2.2.5 External validation of genes predicted to be methylated by the Illumina HD 450K array.**

I further investigated the use of the pyroSeq DNA methylation to validate Illumina 450K array data in different cervical cancer cell lines. The investigation involved three cell lines, HeLa, SiHa and Ca Ski (Table 19). To obtain comparable data of the CpG methylation between different genes of the same cell line, I used DNA obtained from the same passage to investigate all chosen genes. The detailed pyrograms are shown in Appendix II. I observed intensive methylation of almost all CpGs of all the investigated genes. The average methylation level of CpGs in cervical cancer cell lines ranged from 48% to 96%. However, when I specifically examined CpGs using the Illumina HD 450K methylation array, the average methylation level rose to 66% to 100%. The former and latter methylations were for all genes investigated, with the exception of DCC. Two CpGs were investigated in the DCC promoter in both SiHa and Ca Ski cell lines. While Ca Ski showed a high level of methylation of both CpGs, 80 and 78%, the SiHa line exhibited low methylation levels of 18 and 21%, respectively. The HeLa line was not investigated due to the lack of sufficient DNA obtained from the same cell culture passage. Overall, these data validate the Illumina array and confirm that the included CpGs, particularly at promoter regions, are differentially methylated sites identified across tumour versus normal tissue. I also showed and for the first time significant promoter methylation of TP53AIP1 in the two cervical cancer cell lines.

Table 19: DNA methylation of genes selected for external validation of the Illumina HD 450 K methylation array in the cervical cancer cell lines HeLa, SiHa and Ca Ski. Representative Illumina CpGs are highlighted in blue.

Sample	UCSC Gene Name	CpG 1	CpG 2	CpG 3	CpG 4	CpG 5	AVG methylation
HeLa	CADM1	76	82				79
SiHa		87	89				88
HeLa	CALCA	96	100	78			91
Ca Ski		92	97	76			88
SiHa	DCC	18	21				20
Ca Ski		80	78				79
HeLa	ESR1	39	58	72	39		52
SiHa		42	66	84	49		60
HeLa	PTNT6	75	81	89	83	87	83
SiHa		57	68	90	85	77	78
Ca Ski		67	76	82	75	80	76
HeLa	RARB	96	89	95	90	88	92
SiHa		93	76	66	38	26	60
Ca Ski		94	81	66	35	24	60
HeLa	SFRP2	36	31	66	61		48
SiHa		76	35	84	32		57
Ca Ski		71	44	80	53		62
HeLa	SOX1	78	93	74	94	92	88
SiHa		93	99	91	97	100	96
Ca Ski		89	98	88	95	100	94
HeLa	TNFRSF10C	63	65	37	60		56
SiHa		88	100	74	93		89
Ca Ski		59	68	41	63		58
HeLa	TP53AIP1	88	66	80			78
SiHa		82	87	81			84

### **5.2.3 Analysis of the Illumina HD 450K methylation array reveals significant pathway changes following HPV18 integration**

To identify the common biological processes and molecular functions in which differentially methylated genes are involved, I performed functional annotation analyses separately for genes with significantly hypermethylated and hypomethylated CpG sites using DAVID bioinformatics software. The statistical significance of overlap between differentially methylated genes and a particular biological “theme” was measured by the geometric mean of the Expression Analysis Systematic Explorer (EASE) scores (modified Fisher’s exact test).

I used the following stringency criteria to determine the inclusion of differentially methylated CpGs in the functional annotation.

- 1) Differentially methylated in proximity to a promoter.
- 2) At least 10-20% methylation changes between p7 and p14.
- 3) Statistically significant changes, Diffscore  $\geq 22$  ( $p \leq 0.01$ ).
- 4) EASE score of  $p \leq 0.0001$ .

The DAVID analysis output revealed significant enrichments of genes with differentially methylated CpGs of different biological “themes”, according to the Gene Ontology (GO) terms, the Kyoto Encyclopedia of Genes and Genomes (KEGG), the PANTHER classification system and other descriptive terms related to biological functions or biochemical pathway information.

### 5.2.3.1 Hypermethylated genes following HPV18 integration are enriched for biological pathways involved in cervical neoplasia.

There were over three thousand CpGs considered to be hypermethylated at promoter regions with Diffscore  $\geq 13$  and  $p \leq 0.05$ . Applying more stringent criteria, in which  $p \leq 0.01$ , revealed 1,471 CpGs of 1,201 hypermethylated genes following HPV18 episomal loss. Hypermethylated genes were analysed by the DAVID software and a summary of the functional analysis output is shown in Table 20. The main topological domain affected by hypermethylated genes is the plasma membrane. Biological processes involved in signal transduction and signalling molecules interaction were enriched with highly methylated genes. Interestingly, those include two pathways known to be involved in cervical carcinogenesis, the Wnt signalling pathway and the cadherin signalling pathway. Neuroactive ligand receptor (Figure 65) and calcium signalling pathways also showed enrichment with genes carrying significantly hypermethylated CpGs.

The Wnt and cadherin signalling pathways encompass 139 and 181 genes, respectively. Approximately 45% of the former and 27% of the later pathway's genes showed hypermethylation following integration of HPV18 into human DNA (the Benjamini controlled FDR were  $1.20\text{E-}07$  and  $2.40\text{E-}13$ , respectively). Table 21, includes hypermethylated genes of the Wnt and cadherin signalling and neuroactive ligand receptor pathways. An examination showed that the molecular function of the majority of the hypermethylated gene sets identified encoded sequence-specific DNA binding proteins (Benjamini controlled FDR  $0.0019$ ), homeobox proteins (Benjamini controlled FDR  $7.3\text{E-}7$ ), and transcriptional regulators (Benjamini controlled FDR  $2.20\text{E-}03$ ).



Table 20: Overview of the significant functional themes and pathways within the list of the most significant hypermethylated genes following HPV18 integration. Only functional terms and pathways with  $p \leq 0.0001$  are listed.  $p$  values represent the EASE scores given by DAVID.

Biological process	Term	Count	P-Value	Benjamini
KEGG_PATHWAY	Neuroactive ligand-receptor interaction	31	1.30E-04	2.10E-02
KEGG_PATHWAY	Calcium signaling pathway	22	1.10E-03	8.50E-02
PANTHER_PATHWAY	Cadherin signaling pathway	48	2.50E-15	2.40E-13
PANTHER_PATHWAY	Wnt signaling pathway	62	2.40E-09	1.20E-07
Molecular function	Term	Count	P-Value	Benjamini
UP_SEQ_FEATURE	domain:Cadherin 5	40	1.70E-23	2.30E-20
UP_SEQ_FEATURE	domain:Cadherin 4	40	3.80E-22	3.40E-19
UP_SEQ_FEATURE	domain:Cadherin 3	40	3.80E-22	3.40E-19
UP_SEQ_FEATURE	domain:Cadherin 1	40	1.30E-21	8.70E-19
UP_SEQ_FEATURE	domain:Cadherin 2	40	1.30E-21	8.70E-19
UP_SEQ_FEATURE	domain:Cadherin 6	39	1.20E-27	3.20E-24
UP_SEQ_FEATURE	DNA-binding region:Homeobox	36	1.60E-09	7.30E-07
SP_PIR_KEYWORDS	Alternative splicing	521	4.60E-07	4.80E-05
SP_PIR_KEYWORDS	Transcription regulation	162	3.80E-05	2.20E-03
SP_PIR_KEYWORDS	DNA-binding	152	2.90E-05	1.90E-03
SP_PIR_KEYWORDS	Developmental protein	101	9.70E-14	5.10E-11
SP_PIR_KEYWORDS	Cell adhesion	61	1.90E-10	3.30E-08
SP_PIR_KEYWORDS	Differentiation	55	1.00E-06	8.80E-05
Cellular component	Term	Count	P-Value	Benjamini
GOTERM_CC_FAT	Plasma membrane	275	7.50E-06	3.50E-03
GOTERM_CC_FAT	Plasma membrane part	169	7.60E-05	8.80E-03
GOTERM_CC_FAT	Intrinsic to plasma membrane	106	2.00E-05	4.60E-03
SP_PIR_KEYWORDS	Glycoprotein	308	8.40E-05	4.00E-03

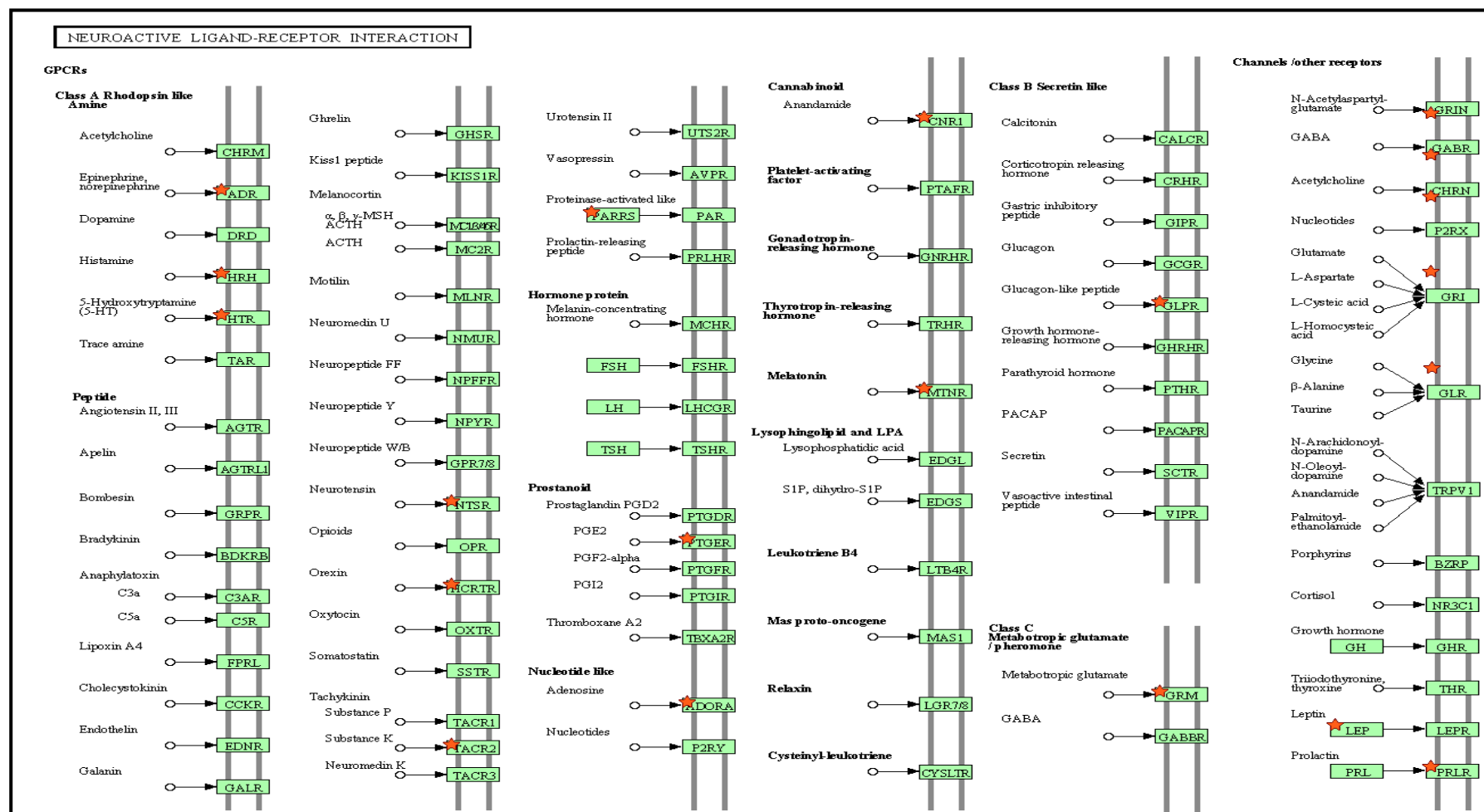


Figure 65: Neuroactive ligand receptor interaction pathway is enriched with genes significantly hypermethylated following HPV18 episomal loss. Red stars mark hypermethylated genes.

Table 21: Official symbol of hypermethylated genes enriched to three significant biological pathways

Wnt signalling pathway	Cadherin signalling pathway	Neuroactive ligand receptor
FYN	SMAD3	HTR2C
CDH20	AXIN2	HTR5A
CTNNA2	CDH20	HTR6
DCHS2	CTNNA2	ADORA1
FER	DCHS2	ADRA1A
FZD1	FZD1	ADRA1B
FZD5	FZD5	ADRA2C
FZD9	FZD9	CNR2
PCDH20	FRZB	GABRA2
PCDHA11	GNG3	GABRA4
PCDHA12	GNG4	GABRB3
PCDHA10,PCDHA13	GNG7	GABRD
PCDHA1, PCDHA4	HOXA4	GLP2R
PCDHA2	HOXB5	GRIA3
PCDHA3	HOXB6	GRIA4
PCDHA5	HOXC4	GRIN2A
PCDHA7	HOXC5	GRIN2D
PCDHA6, PCDHA8	HOXC6	GRM1
PCDHA9	NKD2	GRM8
PCDHB1	PCDH20	GLRA3
PCDHB12	PCDHA11	HRH1
PCDHB13	PCDHA12	HCRTR1
PCDHB15	PCDHA10,PCDHA13	LEP
PCDHB16	PCDHA1,PCDHA4	MTNR1B
PCDHB17	PCDHA2	NTSR1
PCDHB3	PCDHA3	NTSR2
PCDHB7	PCDHA5	PRLR
PCDHGA1	PCDHA7	PTGER1
PCDHGA10	PCDHA6,PCDHA8	PRSS3
PCDHGA11	PCDHA9	P2RX5
PCDHGA2	PCDHB1	TACR2
PCDHGA3	PCDHB12	
PCDHGA4	PCDHB13	
PCDHGA5	PCDHB15	
PCDHGA6	PCDHB16	
PCDHGA7	PCDHB17	
PCDHGA8	PCDHB3	
PCDHGA9	PCDHB7	
PCDHGB1	PCDHGA1	
PCDHGB2	PCDHGA10	
PCDHGB3	PCDHGA11	
PCDHGB4	PCDHGA2	
PCDHGB5	PCDHGA3	
PCDHGB6	PCDHGA4	
PCDHGB7	PCDHGA5	
PCDHGC4,PCDHGC3,	PCDHGA6	

### 5.2.3.2 Functional annotation of hypomethylated genes following HPV18 integration

Up to 6,786 CpGs were hypomethylated at promoter regions ( $\text{Diffscore} \geq 13$  and  $p \leq 0.05$ ). Applying more stringent criteria ( $\text{Diff score} \geq 22$  and  $p \leq 0.01$ ) revealed that there were 854 significantly hypomethylated CpGs in 724 genes submitted to the DAVID software for functional annotation. I provide a summary of the significant biological processes in Table 22. Significantly hypomethylated genes following HPV episomal loss were enriched in different topological domains, such as extracellular, cell membrane and the Golgi apparatus. The main biological process enriched among significantly hypomethylated genes ( $p \leq 0.01$ ) was immune response genes related to processes that include both responses to stimuli and receptor binding activity. The response to stimulus was evidenced by a significant enrichment of chemokine receptors (G protein coupled receptors) in hypomethylated genes (Benjamini controlled FDR  $1.50\text{E-}06$ ). Up to 45% (63) of the genes involved in chemokine signalling pathways were hypomethylated following HPV18 episomal loss. Examples of enriched receptor binding activities are glycoprotein and N-linked protein glycosylation pathways, with Benjamini controlled FDRs of  $3.10\text{E-}05$  and  $4.70\text{E-}04$ , respectively.

Table 22: Overview of significant functional themes and pathways within the list of the most significant hypomethylated genes following HPV18 integration. Only functional terms and pathways with EASE,  $p \leq 0.0001$  are listed.

Category	Term	Count	P-Value	Benjamini
SP_PIR_KEYWORDS	Cell membrane	134	3.30E-09	1.50E-06
SP_PIR_KEYWORDS	G-protein coupled receptor	63	5.20E-08	1.20E-05
SP_PIR_KEYWORDS	Transducer	65	1.20E-07	1.90E-05
SP_PIR_KEYWORDS	Glycoprotein	217	2.60E-07	3.10E-05
SP_PIR_KEYWORDS	Polymorphism	485	1.90E-06	1.80E-04
SP_PIR_KEYWORDS	Receptor	94	4.20E-06	3.30E-04
UP_SEQ_FEATURE	Topological domain: Extracellular	165	5.90E-11	1.20E-07
UP_SEQ_FEATURE	Topological domain: Cytoplasmic	188	2.20E-09	2.10E-06
UP_SEQ_FEATURE	Glycosylation site: N-linked (GlcNAc...)	208	7.20E-07	4.70E-04
UP_SEQ_FEATURE	Sequence variant	504	1.30E-06	6.60E-04
UP_SEQ_FEATURE	transmembrane region	233	1.10E-05	4.20E-03
GOTERM_CC_FAT	Plasma membrane	214	1.90E-11	6.50E-09

### **5.2.3.3 Identification and functional annotation of differentially methylated regions following HPV18 integration**

A differentially methylated region (DMR) is defined as a region in which there is an alteration of DNA methylation at multiple adjacent CpG sites. The most investigated regions were cancer-specific differentially methylated regions (CDMRs) and reprogramming-specific differentially methylated regions (RDMR). The Illumina HD 450K methylation array covers over 35,000 DMRs. Significantly ( $p \leq 0.01$ ) hypermethylated and hypomethylated DMRs following HPV18 episomal loss were 486 and 775, respectively. Functional annotation of hypermethylated DMRs revealed an enrichment of biological pathways similar to those of hypermethylated CpGs, the Wnt and cadherin signalling pathways. In addition, hypermethylated DMRs were enriched in biological processes, including genes encoding homeobox and DNA binding proteins. Most significantly, hypermethylated DMRs were clustered at chromosome 5 cytoband 5q31. A list of the most significant ( $\text{Diffscore} \geq 70$ ) DMRs is shown in Table 23.

Table 23: The most significant hypermethylated DMRs following HPV18 integration.

Gene_Symbol	Gene Name	DMR
C17orf107	Hypothetical protein LOC100130311	DMR
EVX2	Even-skipped homeobox 2	DMR
EYA4	Eyes absent homolog 4 (Drosophila)	CDMR
GNAS	GNAS complex locus	CDMR
GNG4	Guanine nucleotide binding protein (G protein), gamma 4	DMR
HOXA11	homeobox A11	RDMR
IL1RAPL2	Interleukin 1 receptor accessory protein-like 2	CDMR
ITGA4	Integrin, alpha 4 (antigen CD49D, alpha 4 subunit of VLA-4 receptor)	RDMR
KCTD12	Potassium channel tetramerisation domain containing 12	RDMR
LIG1	Ligase I, DNA, ATP-dependent	RDMR
LRBA	LPS-responsive vesicle trafficking, beach and anchor containing	RDMR
PAX3	Paired box 3	RDMR
PAX9	Paired box 9	RDMR
SLITRK1	SLIT and NTRK-like family, member 1	CDMR
SMAD3	SMAD family member 3	RDMR

#### **5.2.3.4 Hypermethylated genes and DMRs are clustered at known HPV integration sites**

It has been hypothesized that certain HPV integrations into host DNA alter genes near the site of integration and, thus, contribute, along with the viral oncoproteins, to the development of cervical tumours (Yu et al., 2005). I also investigated the possible clustering of differentially methylated genes at certain chromosomes or, more specifically, at certain chromosomal locations. Both hypermethylated genes and DMRs were clustered significantly ( $p \leq 0.01$ ) at chromosomes known to be insertion site of HPVs. Chromosomal site 5q31 is considered to be a hotspot region for hypermethylated genes and DMRs. The other hotspot regions for hypermethylation and hypomethylation are summarised in Table 24 and 25. Specifically chromosome 17, 5q31, 19q13.3 and 16p13.3 are hypermethylated hotspot regions. Hypomethylated regions were not mapped to known HPV integration sites. Of interest, the X chromosome showed significant clustering for both hypomethylated and hypermethylated genes.



Table 24: Clustering of hypermethylated genes following HPV18 integration by chromosomal and cytoband location.

Category	Term	Count	P-Value	Benjamini
Chromosome	5	90	7.90E-07	2.00E-05
Chromosome	19	99	3.40E-05	4.20E-04
Chromosome	17	81	6.00E-03	4.90E-02
Chromosome	X	71	3.30E-02	1.50E-01
cytoband	5q31	45	3.60E-38	2.50E-35
cytoband	16p13.3	18	8.20E-04	1.30E-01
cytoband	6p21.3	15	4.20E-02	7.20E-01
cytoband	19p13.2	13	1.00E-02	5.80E-01
cytoband	19q13.33	10	2.60E-03	2.60E-01
cytoband	Xp11.23	9	1.50E-03	1.80E-01

Table 25: Clustering of hypomethylated genes following HPV18 integration by chromosomal and cytoband location.

Category	Term	Count	P-Value	Benjamini
Chromosome	X	51	3.80E-03	9.00E-02
cytoband	21q22.3	9	3.20E-03	5.40E-01
cytoband	20q13.33	7	2.20E-03	6.50E-01
cytoband	9q22.3	3	1.40E-02	8.20E-01

#### **5.2.4 Illumina HD 450K methylation array analysis of human keratinocytes with integrated HPV18 reveals DNA methylation changes in genes involved in the processes of cervical neoplasia**

The results of filtered, significantly hypermethylated genes following HPV18 integration were compared to genes previously have reported to be methylated in cervical neoplasia. All of the 104 genes have reported to be methylated in cervical neoplasia were represented in the Illumina HD 450K array. Of these 104 genes, 32 have found to be significantly more methylated at p14 compared to p7 of the PHFK HPV18 cell line (Table 26). I examined how often genes previously reported to be methylated in cervical neoplasia were differentially hypermethylated following HPV integration (Table 27). This analysis revealed that, compared to genes that were not methylated after HPV18 integration, genes that were significantly methylated were 22 times more likely to be known to be methylated in cervical neoplasia, and this was highly statistically significant ( $p < 0.0001$ ). Thus, genes significantly hypermethylated following HPV integration were significantly enriched for genes known to be methylated in cervical neoplasia.

Table 26: Genes known to be methylated in cervical neoplasia and significantly hypermethylated following HPV18 integration.

UCSC Gene Ref	PMID	Remarks
APAF1	(Yang et al., 2010)	Methylation frequency increased with severity of cervical lesions.
APC	(Narayan et al., 2003)	Promoter was significantly more often hypermethylated in cervical cancers than in controls.
BLU	(Lai et al., 2007)	Gene methylation increased progressively; from LSIL, HSIL, to squamous cell carcinoma.
BRCA1	(Narayan et al., 2003)	High frequency of promoter methylation predict worse prognosis.
CADM1	(Steenbergen et al., 2004)	Promoter hypermethylation is a frequent event in the progression from high-grade CIN lesions to invasive cervical cancer.
CALCA	(Wisman et al., 2006)	Promoter was significantly more often hypermethylated in cancers than in controls.
CDH1	(Narayan et al., 2003)	High frequency of promoter methylation in more advanced stage of cervical carcinoma.
CHFR	(Banno et al., 2007)	Aberrant hypermethylation of CHFR in cervical cancer
DAPK1	(Narayan et al., 2003)	High frequency of promoter methylation in more advanced stage of cervical carcinoma.
DLC1	(Seng et al., 2007)	Promoter was frequently methylated in most cervical tumour cases.
ESR1	(Wisman et al., 2006)	Promoter was significantly more often hypermethylated in cancers than in controls.
HIC1	(Narayan et al., 2003)	Methylation frequency of increased with severity of cervical lesion.
HOXA11	(Apostolidou et al., 2009)	Methylation of HOXA11 could discriminate between HSIL and controls.
LHFPL4	(Wang et al., 2008)	Frequently methylated in cervical cancer.
LMX1A	(Lai et al., 2010)	The percentage of methylation is significantly higher in CIN3, compare with CIN1 or CIN2.
MLH1	(Kang et al., 2006)	Hypermethylation was significantly more frequent in cervical carcinoma.
MYOD1	(Widschwendter et al., 2004)	Promoter hypermethylation more frequently in advanced stage of cervical cancer
ONECUT1	(Lai et al., 2008)	Significantly more methylated in SCC than in normal control.
PAX1	(Lai et al., 2010)	Percentage of methylation is significantly higher in CIN3, than CIN1 or CIN2.
PTGS2	(Kang et al., 2005)	Methylated in cervical cancer.
PTPN6	(Kim et al., 2010)	Increase methylation with increasing severity of cervical squamous lesions.
RARB	(Narayan et al., 2003)	High frequency of promoter methylation predict worse prognosis.
ROBO1	(Narayan et al., 2006)	Promoter hypermethylation of the Slit-Robo genes occurs early in cervical tumour progression.

ROBO3	(Narayan et al., 2006)	Promoter hypermethylation of the Slit-Robo genes occurs early in cervical tumour progression.
RRAD	(Sova et al., 2006)	Aberrant methylation is present in a high proportion of cervical cancer but not in normal samples.
SFRP2	(Chung et al., 2009)	The frequency of <i>SFRP</i> genes promoter methylation increase with progression of cervical neoplasia from LISL to HSIL to SCC.
SLIT1	(Narayan et al., 2006)	Promoter hypermethylation of the Slit-Robo genes occurs early in cervical tumour progression.
SYK	(Feng et al., 2005)	Increase promoter methylation in CIN 2/3 and in the cervical cancer progressively.
TERT	(Iliopoulos et al., 2009)	Suggested to be molecular biomarker of cervical oncogenes progression.
TNFRSF10C	(Henken et al., 2007)	Methylation observed in early immortal cells with HPV
TP73	(Henken et al., 2007)	Methylation observed in early immortal cells with HPV
WT1	(Henken et al., 2007)	Testing for methylated WT1 is superior to the HPV test in detecting severe neoplasms CIN3

Table 27: Frequency with which genes that were differentially hypermethylated in p14 compared to p7 of the PHFK HPV18 cell line are also known to be methylated in cervical neoplasia

Change in methylation status		Methylated in cervical neoplasia		Odds ratio (95% CI)
		Yes	No	
Significantly more methylated in p14 vs p7	No n=476038	104	475934	1 (reference)
	Yes n=9540	32	9508	OR=22.3 95% CI:14.6 to 33.7 p <0.0001

### 5.2.5 *De novo* methylated genes following HPV18 episomal loss

From the 9,450 CpGs that showed significant hypermethylation, 2,363 CpGs were methylated at promoter regions. I explored genes that became *de novo* methylated at promoter regions following HPV 18 integration. Genes were considered to be *de novo* methylated if they fulfilled the following criteria.

1. Average beta of p7 is zero.
2. Average increase in beta of p14 methylation by at least 10-20%.
3. The methylation changes were statistically significant, DiffScore  $\geq 13$  and  $p \leq 0.05$ .

Interestingly, only 27 genes were significantly *de novo* methylated following HPV18 integration. From the *de novo* methylated genes, 18 genes were methylated at promoter regions (Figure 66a and Table 28). As epigenetic switching from H3K27me3m3 to DNA methylation is more likely to involve *de novo* methylated genes, I used the database of genes enriched with H3K27me3m3 in human embryonic cells (HEC) of Zhao et al. (2007) to extract potential gene candidates for epigenetic switching. Five genes matched all the above criteria and were simultaneously enriched with H3K27me3m3 in HEC; these genes are BTBD11, GABRA4, KCNA4, POU4F1 and ROBO3 (Figure 66b). The latter, ROBO3, is one of the genes frequently reported to exhibit differential methylation in cervical neoplasia (Table 26).

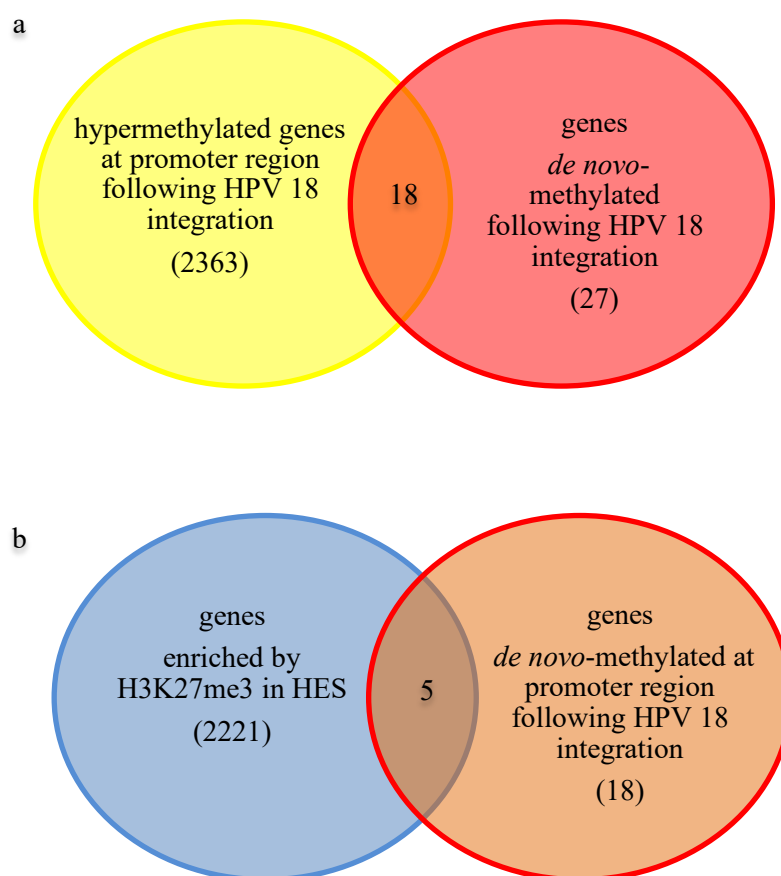


Figure 66: Venn diagram for gene that are *de novo* methylated at their promoters. (a) following HPV18 integration and (b) enriched by H3K27me3 in human embryonic cells.

Table 28: The *de novo* methylated genes following HPV18 integration. Genes enriched with H3K27me3 are highlighted. For a DiffScore  $\geq \pm 13$ , the  $p$ -value  $\leq 0.05$ ; for DiffScore  $\geq \pm 22$ , the  $p$ -value  $\leq 0.01$ ; and for DiffScore  $\geq \pm 33$ , the  $p$ -value  $\leq 0.001$ .

UCSC Gene Ref	Gene Group	Relation to CpG_Island	(p7)Avg Beta	(p14)Avg Beta	Diff Score
ABCG5	Body	Island	0.0	0.1	20
ASPG	TSS200	Island	0.0	0.1	15
BBS7	TSS200	Island	0.0	0.1	17
BTBD11	TSS1500	Island	0.0	0.1	24
C16orf13	Body	Island	0.0	0.1	13
CNIH2	TSS1500	Island	0.0	0.1	14
CYP1B1	Body	Island	0.0	0.1	34
DNAH9	TSS200	Island	0.0	0.1	21
FBXO7	1stExon	Island	0.0	0.1	16
FOXH1	Body	Island	0.0	0.1	21
GABRA4	TSS200	Island	0.0	0.1	23
GDF7	Body	Island	0.0	0.1	14
GLDN	1stExon	Island	0.0	0.1	14
KCNA4	TSS200	Island	0.0	0.1	21
MAMSTR	TSS1500	Island	0.0	0.1	39
RASIP1	TSS1500	Island	0.0	0.1	38
MTMR8	Body	N Shore	0.0	0.1	14
NR5A2	Body	Island	0.0	0.1	14
PENK	Body	Island	0.0	0.1	57
POU4F1	TSS200	Island	0.0	0.1	14
PSTPIP2	Body	Island	0.0	0.1	20
ROBO3	TSS200	Island	0.0	0.1	30
SLC6A1	5'UTR	Island	0.0	0.1	37
SOX11	1stExon	Island	0.0	0.1	20
ZDHHC9	TSS200	Island	0.0	0.1	14
ZFP28	1stExon	Island	0.0	0.1	23
ZSCAN18	1stExon	Island	0.0	0.1	18

In brief, an analysis of the Illumina 450K HD methylation array for the methylation changes in PHFK HPV18 cell line revealed that 8.7% of approximately 500,000 CpGs exhibited significant methylation changes. Three-quarters of the methylation changes were hypomethylated and one-quarter were hypermethylated CpG sites. Most of the hypermethylated changes occurred in proximal promoters and CpG islands. A significant number of known methylation markers of cervical cancer showed methylation following HPV integration. In addition, components of the Wnt pathway (involved in cervical carcinogenesis), and its regulators showed significant methylation changes in p14. In addition, significant differential methylation changes mapped to HPV integration sites. Only 18 genes showed significant *de novo* methylation at promoter regions, of which one is the ROBO3 gene, a methylation marker of cervical neoplasia and known to be enriched with H3K27me3.



### **5.3 Investigation of possible HPV-induced epigenetic switching**

Epigenetic switching locks the silencing of key pathways' regulators and tumour suppressor genes and eventually contributes to the host cell's abnormal growth potential. I showed in the chapter 2 that changes in the expression of epigenetic regulators of DNA methylation and H3K27me3 occurs following HPV episomal loss and the overexpression of the E7 oncoprotein. The changes showed a pattern consistent with epigenetic switching. Following overexpression of the E7 oncoprotein, both DNA methyltransferases and H3K27me3 demethylase were overexpressed. I also showed in the previous section DNA methylation changes in crucially important biological pathways following HPV18 integration. My hypothesis, following these observations, is that overexpression of HPV E7 oncoprotein triggers switching from H3K27me3 repression to stable DNA methylation repression.

In the second section of this chapter, I analyse ChIP assays for the DNMT1, DNMT3B proteins, and H3K27me3 methylation to search for gene binding in p7 and p14 of the PHFK HPV18 and HeLa cell lines. I also search for possible epigenetic switching following HPV18 episomal loss in the PHFK HPV18 cell line. Firstly, I validate the ChIP experiments. Then, I analyse data of Slit-Robo pathway genes and data of TSG separately. I show both PCR and a densitometric analysis obtained with ImageJ software.

### 5.3.1 Chromatin immunoprecipitation (ChIP) assays

The assay was performed in both PHFK HPV18 (p7 and p14) and HeLa cell lines. Having successfully established the ChIP technique with DNMT1, DNMT3B and H3K27me3 antibodies, I used the chromatin immunoprecipitation (ChIP) assay for the following:

- 1) To search for possible DNMT1 and DNMT3B binding sites in key TSG in PHFK HPV18 cell line.
- 2) To search for possible DNMT1 and DNMT3B binding sites in key TSG in HeLa cell line.
- 3) To determine whether there was an increase in DNMT1 and DNMT3B binding following HPV18 integration using p7 and p14 of PHFK HPV18 cell line.
- 4) To investigate possible epigenetic switching following HPV18 integration.

To analyse my results I demonstrate both PCR results and densitometry analysis of ChIP data obtained with ImageJ software of two independent ChIP experiments. I examined each gene promoter with overlapping primers. At least two regions were examined for each gene, covering ~1kb from the upstream region.

#### 5.3.1.1 Optimization of ChIP assays

To preclude non-specific enrichment by ChIP, two negative controls were used in each experiment, the IgG antibody and a no antibody control. To optimise the ChIP assay, I used the HeLa cell line. It has been shown that the DNMT1 and H3K27me3 proteins bind to the HOXA7 promoter region E, but not region F, in the HeLa cell line (Wu et al., 2008). I used HOXA7 regions E and F, and also GAPDH to optimise the ChIP assays. I performed the experiments twice in both the PHFK HPV18 (p7 and p14) and HeLa cell lines. Using antibodies specific for DNMT1, DNMT3B and H3K27me3, I showed, using PCR results, an enrichment of DNMT1

proteins in Region E of the HOXA7 gene, with no binding observed in Region F or GAPDH in the PHFK HPV18 (p7 and p14) and HeLa cell lines (Figure 67). There was also an enrichment of H3K27me3 in the HeLa cell line. Thus, the HeLa ChIP results reproduced the results from the paper published by Wu et al. (2008). Conversely, I found no enrichment of DNMT3B on HOXA in GAPDH in the PHFK HPV18 (p7 and p14) and HeLa cell lines.

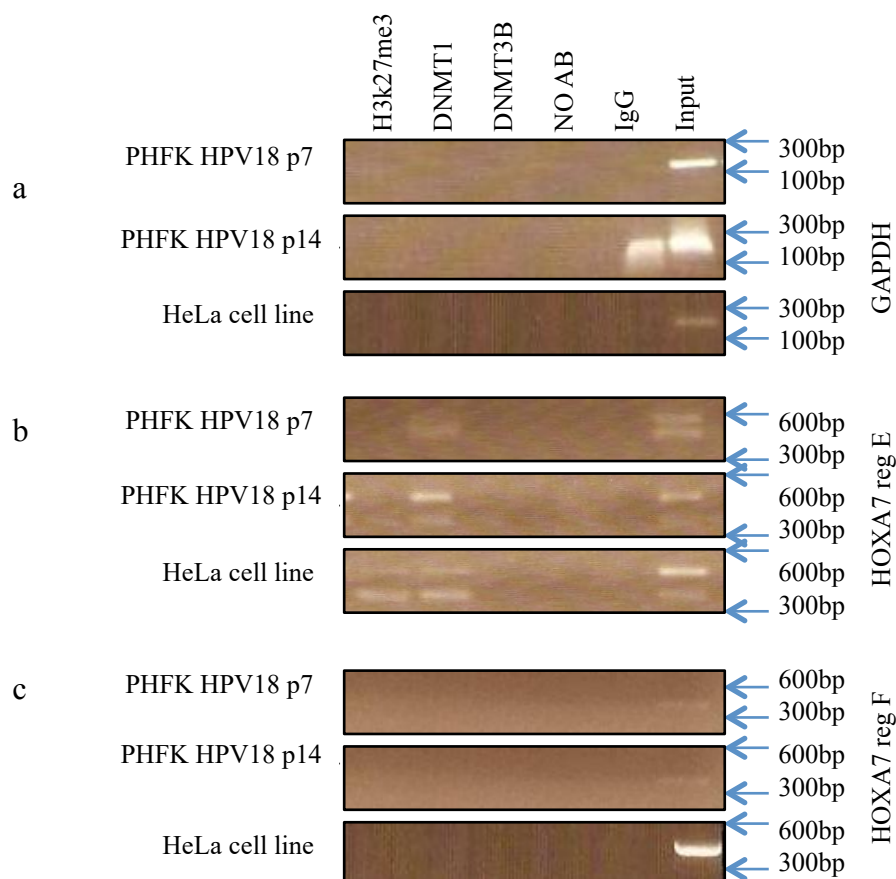


Figure 67: ChIP optimization. ChIP analysis of the PHFK HPV18 (passage 7 and 14) and HeLa cell lines. a, PCR reveals no enrichment of the H3K27me3 methyl mark, or DNMT1 and DNMT3B proteins in GAPDH b, PCR reveals enrichment of the DNMT1 proteins in region E of the HOXA7 promoter in the PHFK HPV18 (passage 7 and passage 14) and HeLa cell lines; it also shows enrichment of the H3K27me3 mark, relative to IgG, in region E of the HOXA7 promoter in HeLa cells. c, conversely, in region F no enrichment was observed.

### 5.3.1.2 Determination of candidate pathways and genes for ChIP experiments

Promoter hypermethylation of the Slit-Robo pathway genes increases significantly from precancerous to early invasive cervical tumours (Narayan et al., 2006). I have already showed that ROBO1, ROBO3 and SLIT1 become significantly methylated following HPV18 episomal loss, (Table 26). Some of the Slit-Robo pathway genes drew my attention, as all genes in the pathway are known to be involved in cervical carcinogenesis. I showed that ROBO3 undergoes *de novo* methylation following HPV18 integration (Table 28), and is enriched with H3K27me3 in HEC. In addition, all the SLIT genes, SLIT1, SLIT2 and SLIT3, are enriched with H3K27me3 in HEC. Therefore I included all genes of the Slit-Robo pathway, ROBO1, ROBO2, ROBO3, SLIT1, SLIT2 and SLIT3, in the ChIP assays. Slit-Robo pathway genes (ROBO1, ROBO3, SLIT1 and SLIT3) showed differential methylation changes following HPV18 integration. However, the others showed no methylation changes (Table 29). Of interest, two genes involved in the Rb-E2F proliferative pathway that are potentially manipulated by overexpression of HPV18 E7 oncoprotein, were also candidates for the ChIP assay. The first gene is SOX11, which is *de novo* methylated following HPV18 integration, and, thus, has the potential to undergo ‘epigenetic switching’. The other gene is CDKN2A, which is hypermethylated during cervical carcinogenesis (Gustavsson et al., 2010). I further investigated two potential genes for DNMT1 binding, ESR1 and CALCA, known as tumour suppressor genes (TSG), in cervical cancer and I have already showed that they were hypermethylated following HPV18 integration (Table 26).

Table 29: Summary of predicted CpGs methylation of the Slit-Robo pathway following HPV18 integration in PHFK HPV18 cell line by the Illumina 450K HD methylation array. The genes that show significant differential methylation are highlighted in blue. For a DiffScore =  $\pm 13$ , the  $p$ -value = 0.05; and for DiffScore =  $\pm 22$ , the  $p$ -value = 0.01.

UCSC ref Gene	Gene group	Relation to CpG Island	p7 Avg Beta	DiffSCORE
ROBO1	5'UTR	Island	0.1	24
ROBO1	TSS1500	S_Shore	0.3	19
ROBO2	TSS1500	Island	0.1	7.7
ROBO2	TSS1500	Island	0.0	6.3
ROBO2	TSS1500	Island	0.0	6.3
ROBO2	1stExon	S_Shore	0.0	5.2
ROBO3	TSS200	Island	0.0	30
ROBO3	Body	Island	0.1	48
ROBO3	Body	N_Shore	0.7	17
ROBO3	Body	N_Shelf	0.7	17
SLIT1	Body	N_Shelf	0.8	53
SLIT2	Body	Island	0.2	5.3
SLIT2	Body	Island	0.1	3.8
SLIT2	Body	Island	0.1	3.1
SLIT2	TSS200	Island	0.1	1.1
SLIT2	TSS200	Island	0.1	0.3
SLIT2	Body	Island	0.1	0.2
SLIT2	TSS200	Island	0.1	0.2
SLIT2	Body	S_Shore	0.6	0.0
SLIT2	TSS200	Island	0.9	0.1
SLIT2	TSS1500	Island	0.1	0.1
SLIT3	5'UTR	Island	0.0	3.9
SLIT3	TSS1500	Island	0.2	3.5
SLIT3	Body	S_Shore	0.3	38.0

### 5.3.2 HPV 18 integration induces epigenetic switching of Slit-Robo

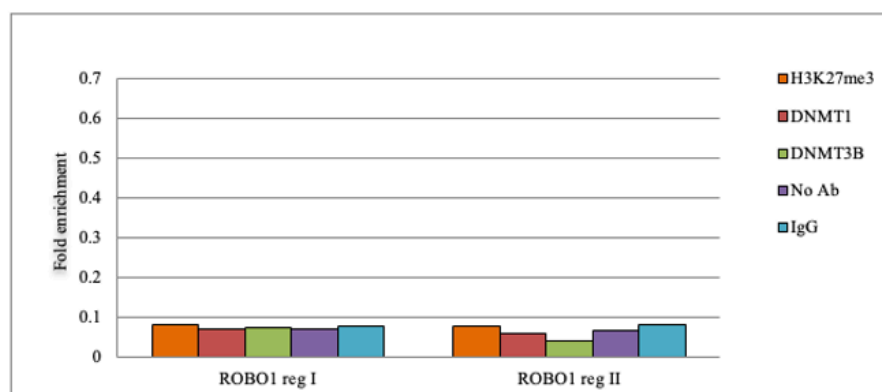
#### pathway genes

There was no enrichment of DNMT1 and DNMT3B proteins and H3K27me3 mark in ROBO1, ROBO2, SLIT1 and SLIT3 in the PHFK HPV18 cell line p7 (Figures 69-73). Enrichment of the H3K27me3 and DNMT1 proteins in the ROBO3 and SLIT2 promoters, respectively, were significantly greater than the binding of IgG antibody. Conversely, in the PHFK HPV18 cell line, p14, I detected a significant enrichment of DNMT1 at the promoters of ROBO1 and ROBO3 (Figure 69-70). In addition, I showed a significant enrichment of DNMT3B at the promoters of ROBO3 (Figure 70) and SLIT3 that represents a novel binding of DNMT3B in the PHFK HPV18 cell line. I found no enrichment of ROBO2 (Figure 69), SLIT1 (Figure 71) and SLIT2 (Figure 72) in any of the examined promoter regions. Unexpectedly, in the HeLa cell line, I detected only significant binding of DNMT1 at the promoter of ROBO1, relative to IgG. Although, I investigated an ~1Kb upstream region of SLIT3, I found no enrichment of DNMT1 or DNMT3B. That was also the case for HeLa cells, and both p7 and p14 of the PHFK HPV18 cell line (Figure 73).

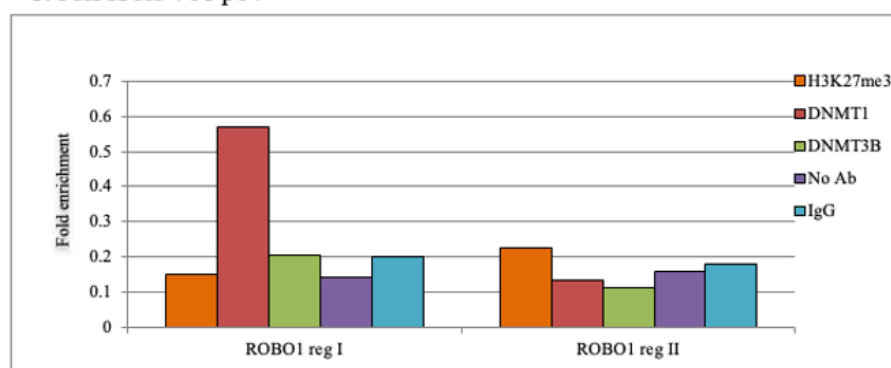
Interestingly, ROBO3 switched from H3K27me3 enrichment in p7, to enriched DNMT1 and DNMT3B in p14. I showed different binding regions for DNMT1, DNMT3B and H3K27me3 which suggests a possible promoter site preference. It is also of interest that I observed that all three proteins did not bind the ROBO3 promoter in the HeLa cell line. Additionally, I showed DNMT1 enrichment for SLIT2 in both p7 and p14 of the PHFK HPV18 cell line at the same promoter region (SLIT2 II), although the enrichment, relative to IgG, was reduced (Figure 72).



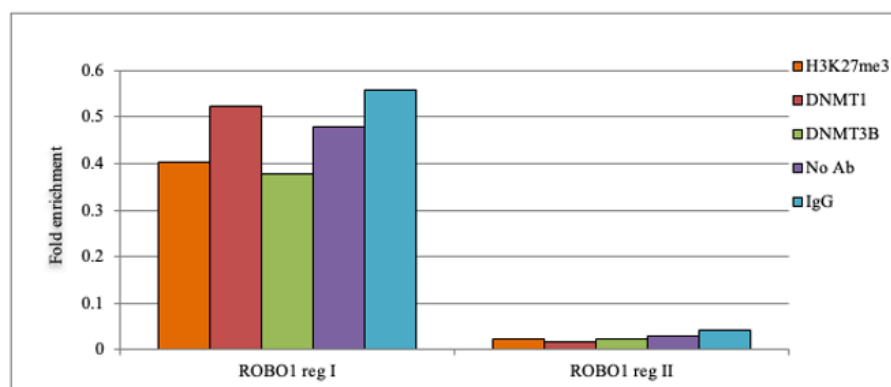
a: PHFK HPV18 p7



b: PHFK HPV18 p14



c: HeLa



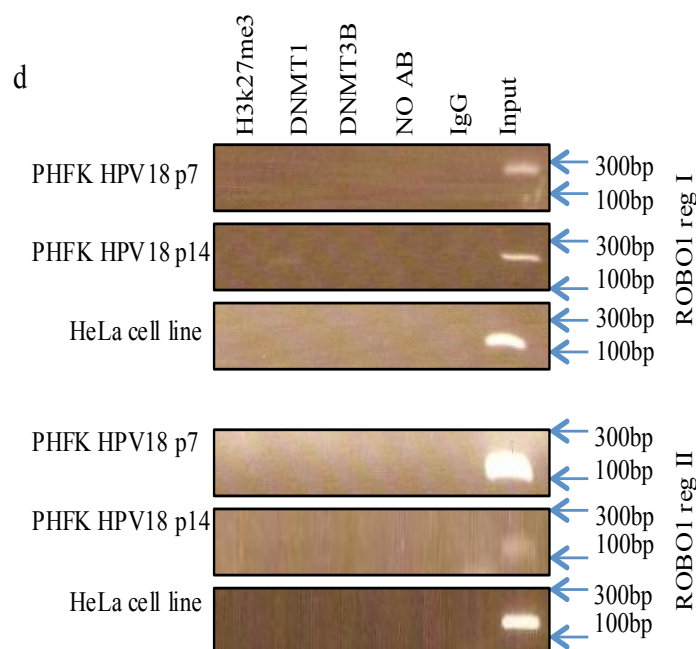
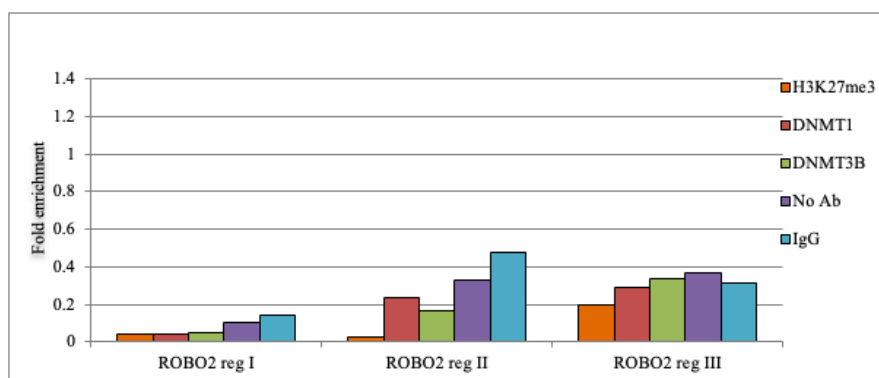
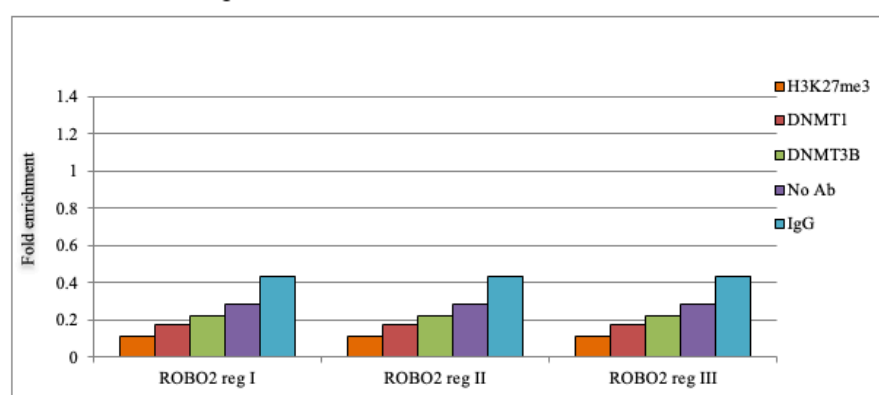


Figure 68: ChIP analysis showing H3K27me3, DNMT1 and DNMT3B proteins binding to two consecutive regions at the ROBO1 promoter. Figure a, b and c show densitometric analyses of ChIP data obtained with ImageJ software. Y-axis represents fold enrichment. ChIP experiments. a, no enrichments shown by densitometric analysis of ChIP data for the PHFK HPV18 cell line p7. b, demonstrate densitometric analysis of ChIP data of the PHFK HPV18 cell line at p14 shows significant enrichment of the DNMT1 protein at region I of the ROBO1 promoter. c, ChIP densitometry analysis of the HeLa cell line shows no enrichment. d, ChIP PCR results reveal an enrichment of the DNMT1 protein at region I of the ROBO1 promoter.

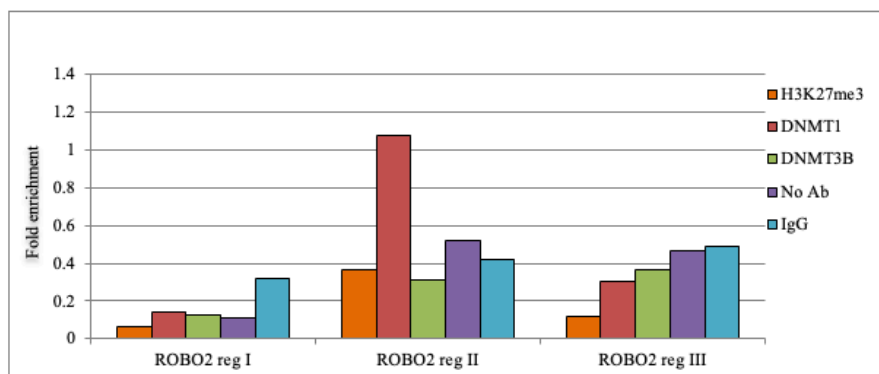
a: PHFK HPV18 p7



b: PHFK HPV18 p14



c: HeLa



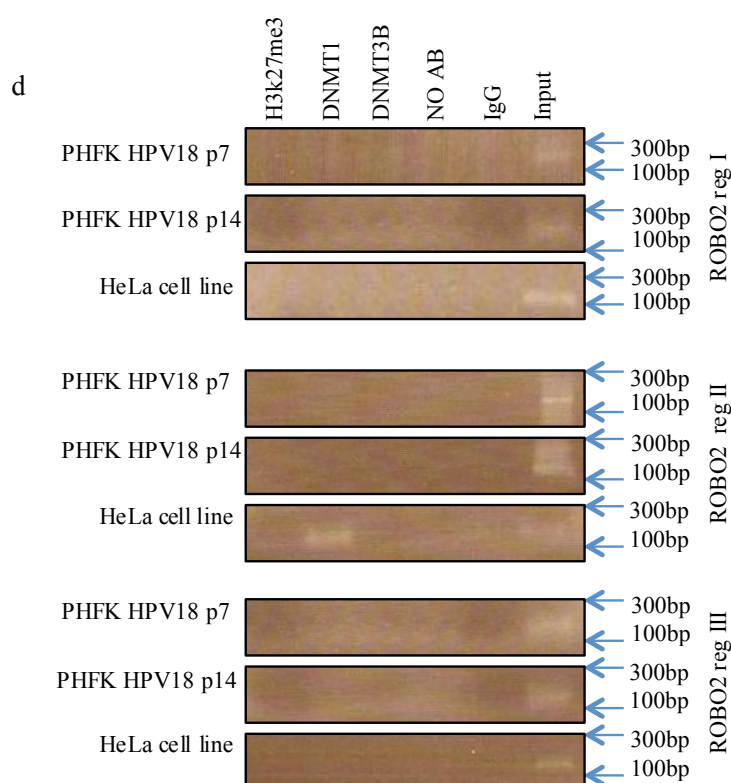
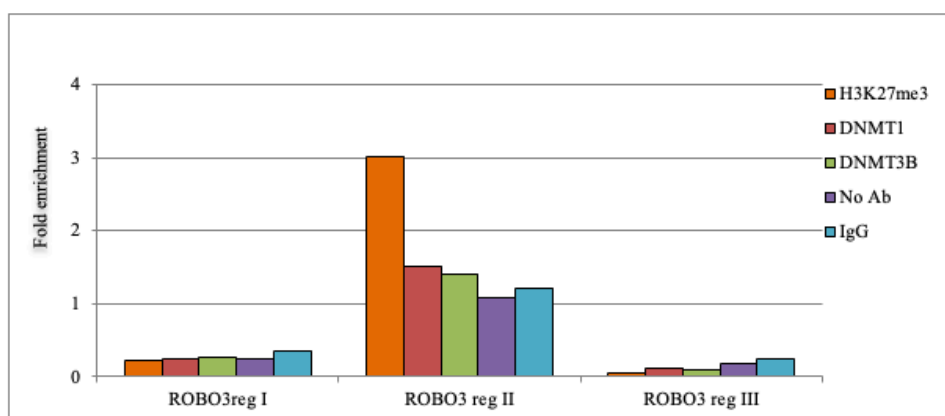
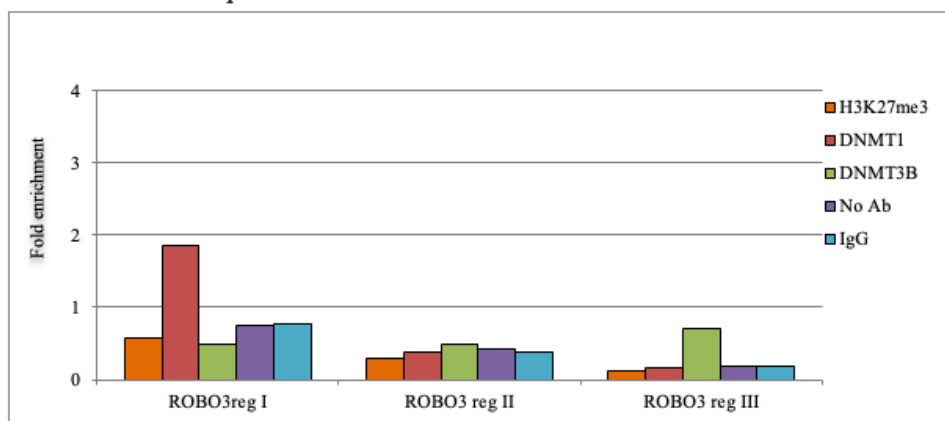


Figure 69: ChIP analysis showing H3K27me3, DNMT1 and DNMT3B binding to three consecutive regions at the ROBO2 promoter. Figure a, b and c show densitometric analyses of ChIP data obtained with ImageJ software. Y-axis represents fold enrichment. a, ChIP analysis of PHFK HPV18 p7. b, ChIP analysis of PHFK HPV18 p 14. c, ChIP densitometry of the HeLa cell line showing significant enrichment of the DNMT1 protein at region II of the ROBO2 promoter. d, ChIP PCR results reveal an enrichment of the DNMT1 protein at region II of the ROBO2 promoter in the HeLa cell line.

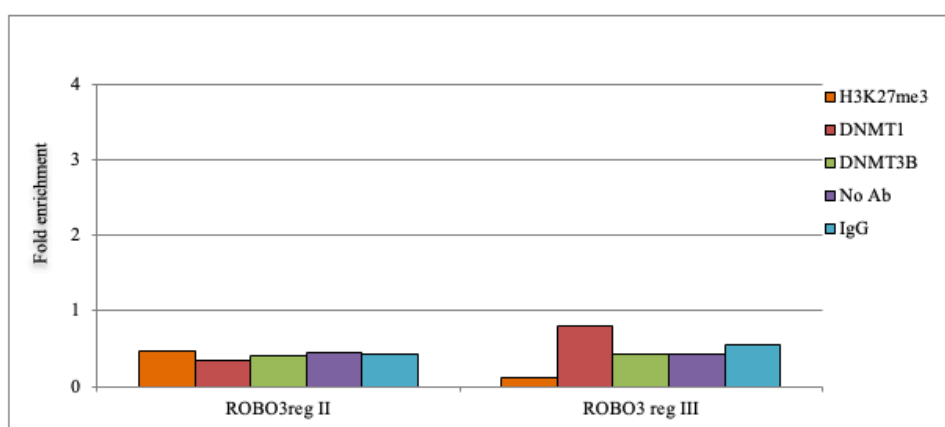
a: PHFK HPV18 p7



b: PHFK HPV18 p14



c: HeLa



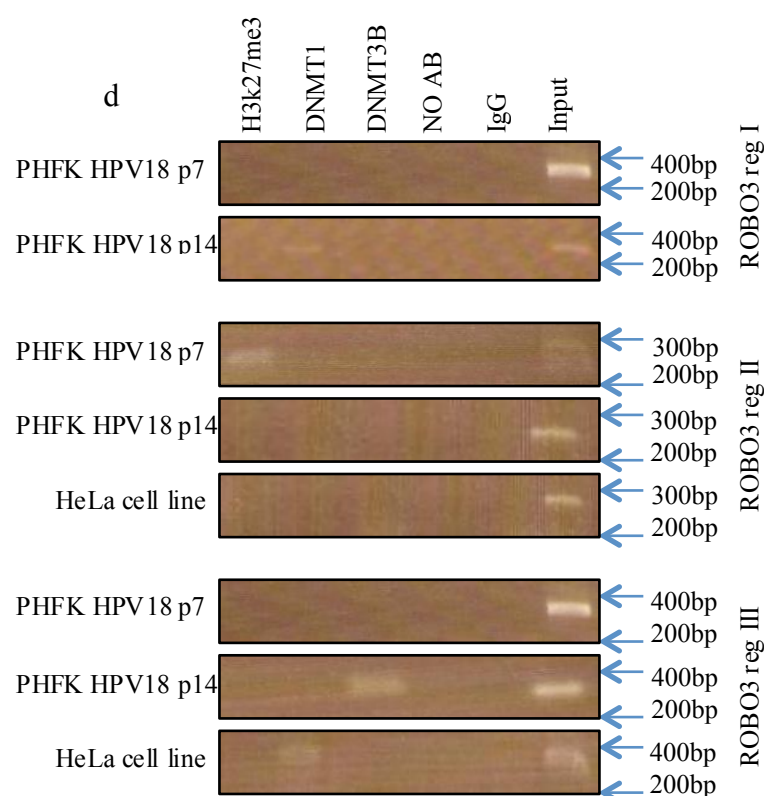
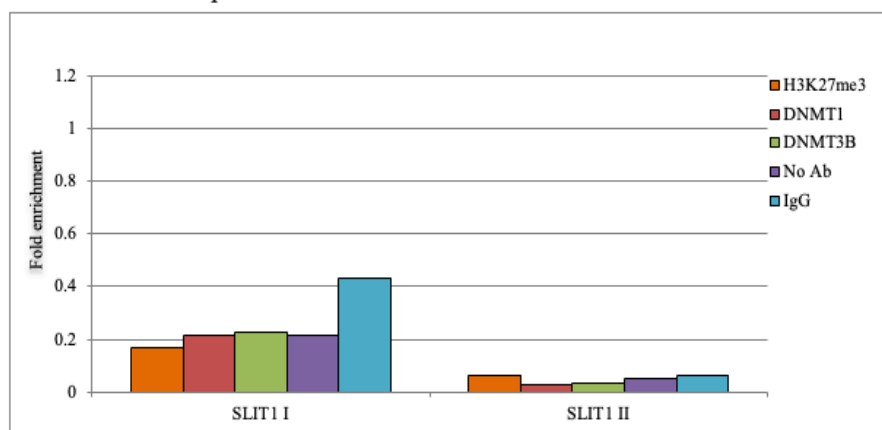
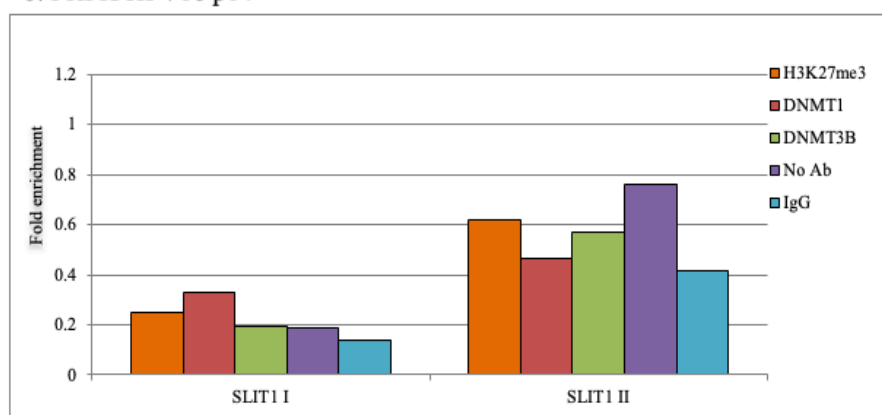


Figure 70: ChIP analysis showing H3K27me3, DNMT1 and DNMT3B binding to three consecutive regions at the ROBO3 promoter. Figure a, b and c show densitometric analyses of ChIP data obtained with ImageJ software. Y-axis represents fold enrichment. a, densitometry shows significant enrichment of H3K27me3 at region II of the ROBO3 promoter in PHFK HPV18 p7. b, shows a ChIP analysis of the PHFK HPV18 cell line p14. Significant enrichment is shown in region I of the ROBO3 promoter for the DNMT1 protein and in region III for the DNMT3B protein. c, ChIP densitometry analysis of the HeLa cell line shows a significant enrichment of the DNMT1 protein at region III of the ROBO3 promoter. Region I is not included due to inconsistent results of the two independent ChIP experiments. d, PCR of the binding of the H3K27me3, DNMT1 and DNMT3B proteins in regions I, II and III of the ROBO3 promoter of the PHFK HPV18 cell line, p 7 and p14, and the HeLa cell line.

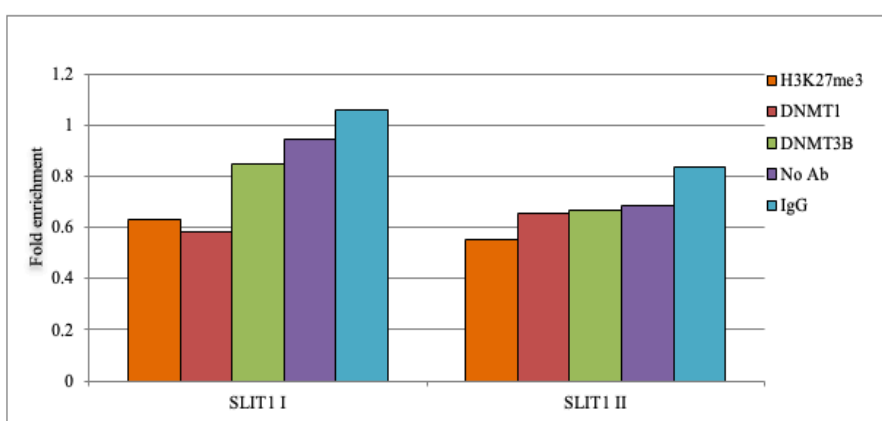
a: PHFK HPV18 p7



b: PHFK HPV18 p14



c: HeLa



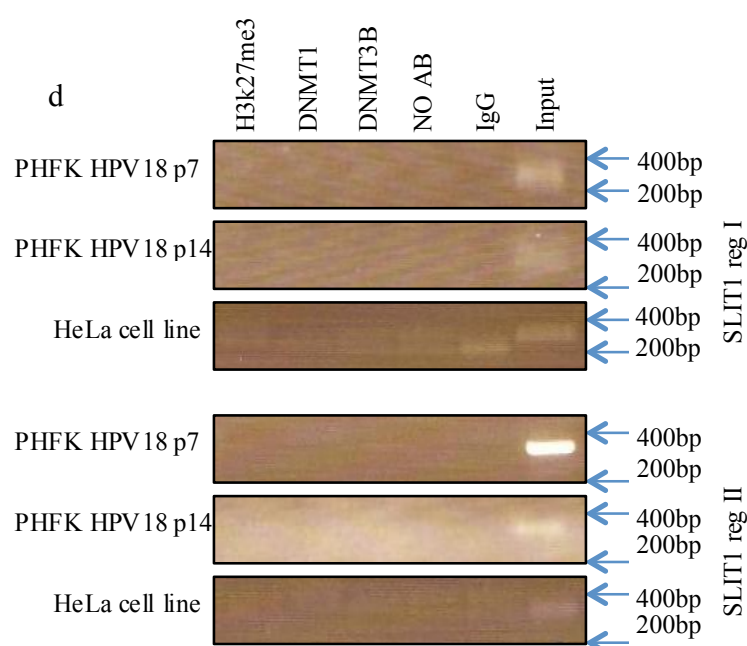
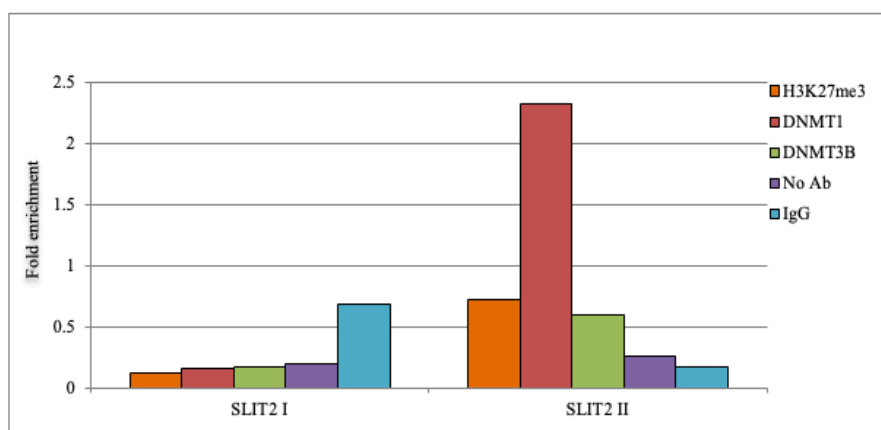


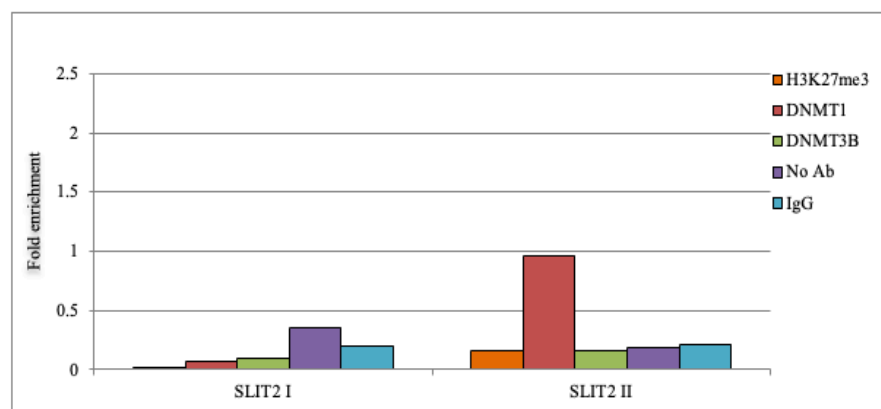
Figure 71: ChIP analysis showing H3K27me3, DNMT1 and DNMT3B binding to two consecutive regions of the SLIT1 promoter. Figure a, b and c show densitometric analyses of ChIP data obtained with ImageJ software. Y-axis represents fold enrichment. a, ChIP analysis of the PHFK HPV18 cell line, p7, b, ChIP analysis of the PHFK HPV18 cell line, p14. c, ChIP densitometry analysis of the HeLa cell line. No significant enrichment, relative to IgG, was observed in a, b or c. d, PCR of the binding of the H3K27me3, DNMT1 and DNMT3B proteins to regions I and II of the SLIT1 promoter of the PHFK HPV18 and HeLa cell lines.



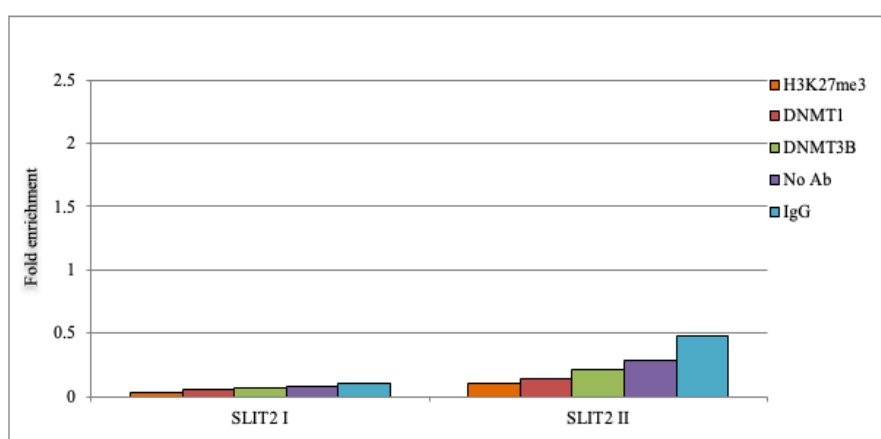
a: PHFK HPV18 P7



b: PHFK HPV18 P14



c: HeLa



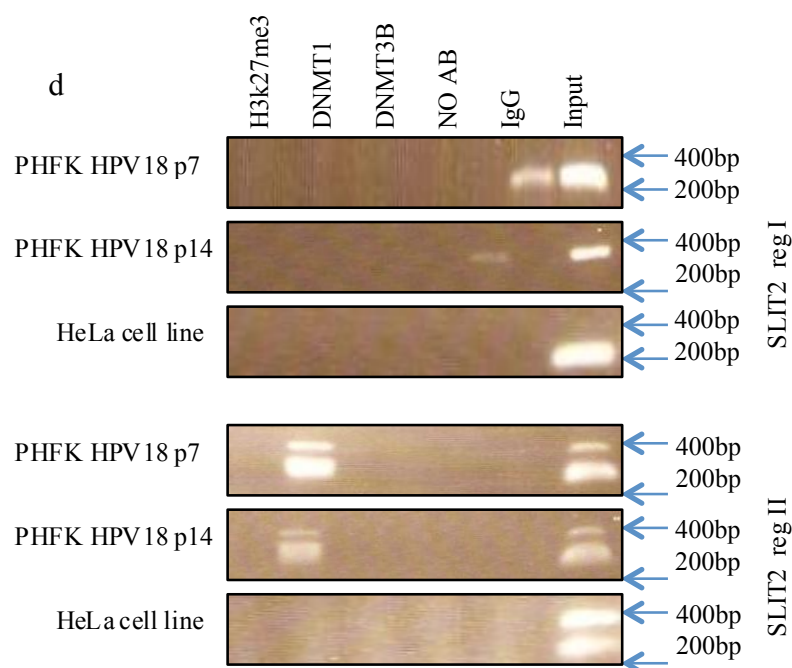
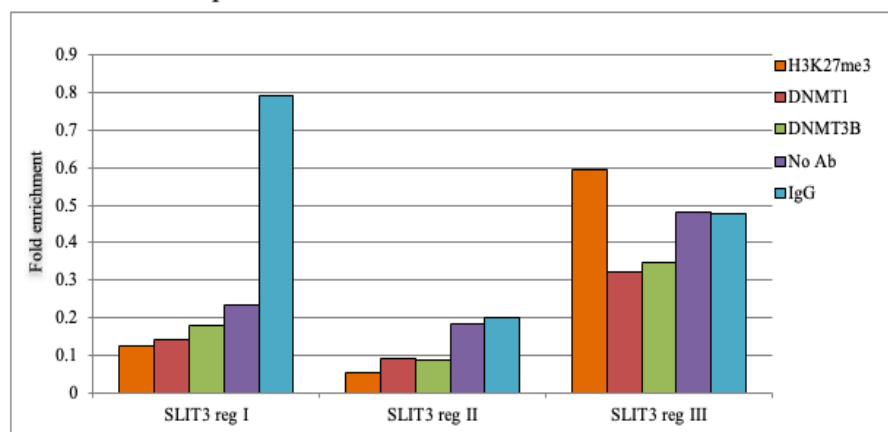
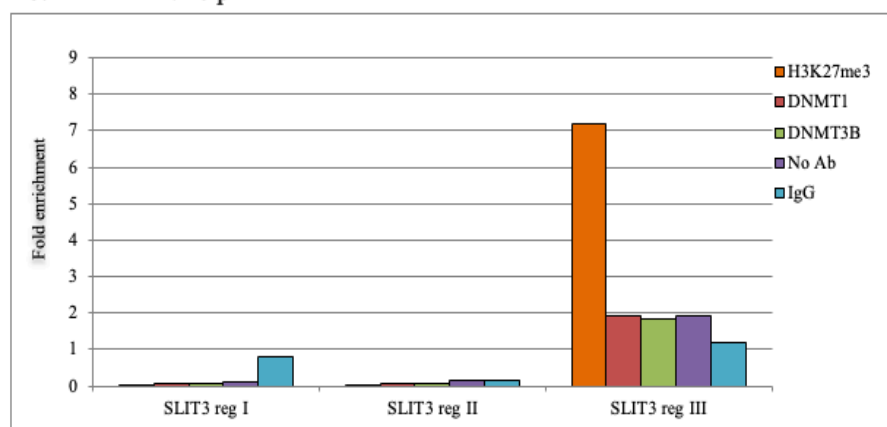


Figure 72: ChIP analysis showing H3K27me3, DNMT1 and DNMT3B binding to two consecutive regions of the SLIT2 promoter. Figure a, b and c show densitometric analyses of ChIP data obtained with ImageJ software. Y-axis represents fold enrichment. Error bars indicate the standard error for of mean of two independent a, ChIP analysis of the PHFK HPV18 cell line at p7 shows a significant enrichment of DNMT1 at region II of the SLIT2 promoter. b, ChIP analysis of the PHFK HPV18 cell line at p14 shows an enrichment of DNMT1 at region II of the SLIT2 promoter. c, ChIP densitometry analysis of the HeLa cell line shows no significant enrichment relative to IgG. d, PCR of the binding of the H3K27me3, DNMT1 and DNMT3B proteins to regions I and II of the SLIT2 promoter of the PHFK HPV18 cell line at p7 and p14 and the HeLa cell line.

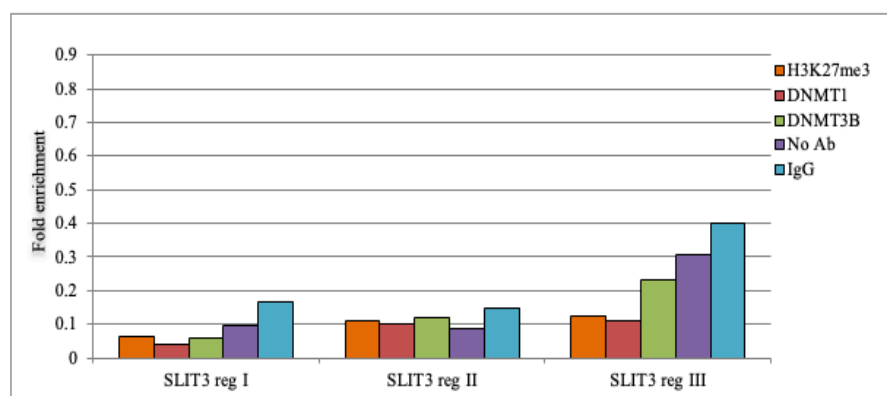
a: PHFK HPV18 p7



b: PHFK HPV18 p14



c: HeLa



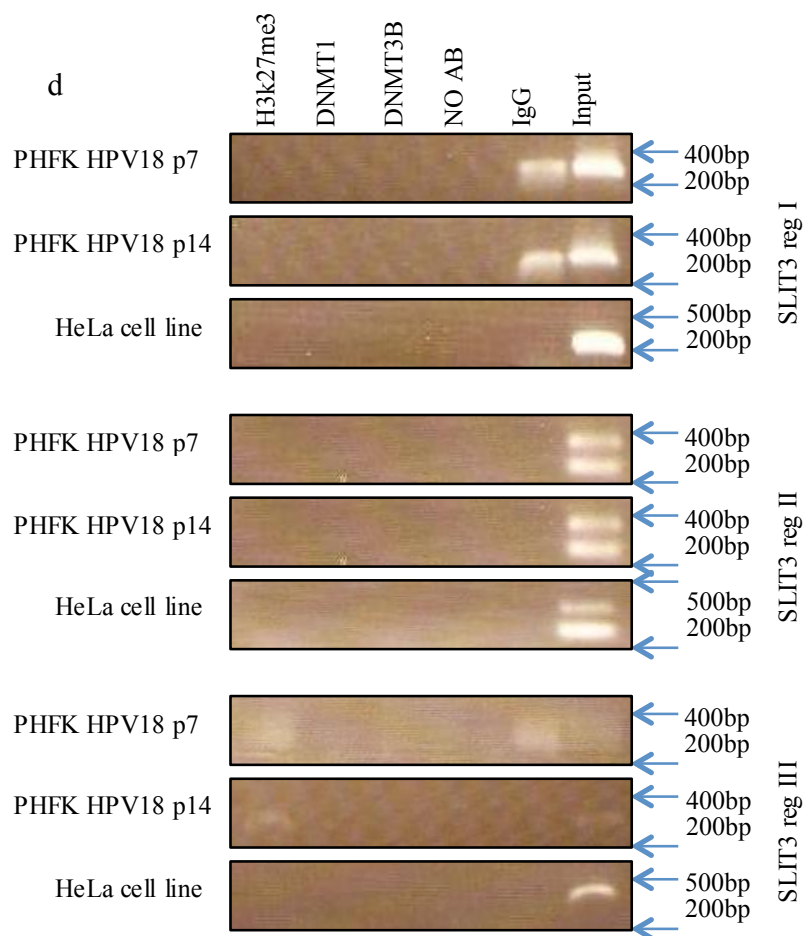
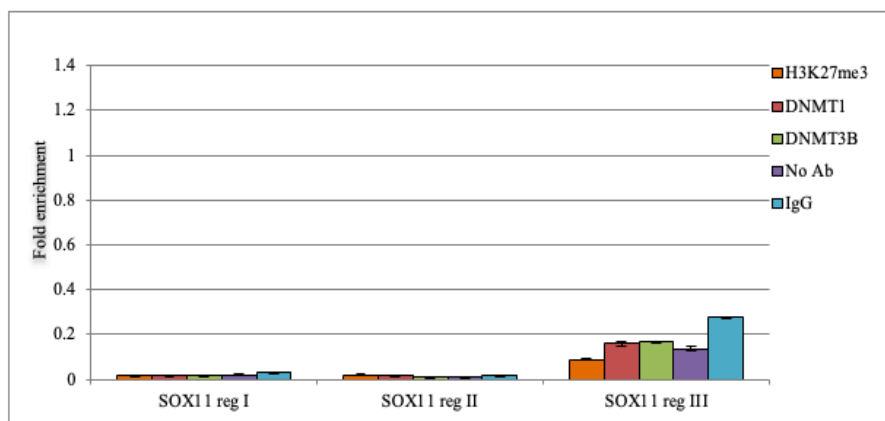


Figure 73: ChIP analysis showing H3K27me3 binding to regions of the SLIT3 promoter. Figure a, b and c show densitometric analyses of ChIP data obtained with ImageJ software. Y-axis represents fold enrichment. Error bars indicate the standard error for of mean of two independent a, ChIP analysis of the PHFK HPV18 cell line at p7 shows significant enrichment of H3K27me3 to SLIT3 III relative to IgG. b, ChIP analysis of the PHFK HPV18 cell line, p14, shows significant enrichment of H3K27me3 to the same region of the SLIT3 promoter. c, ChIP densitometry analysis of the HeLa cell line shows no significant enrichment relative to IgG. d, PCR of the binding of the H3K27me3, DNMT1 and DNMT3B proteins in regions I, II and III of the SLIT3 promoter of the PHFK HPV18 cell line, p7 and p14, and the HeLa cell line.

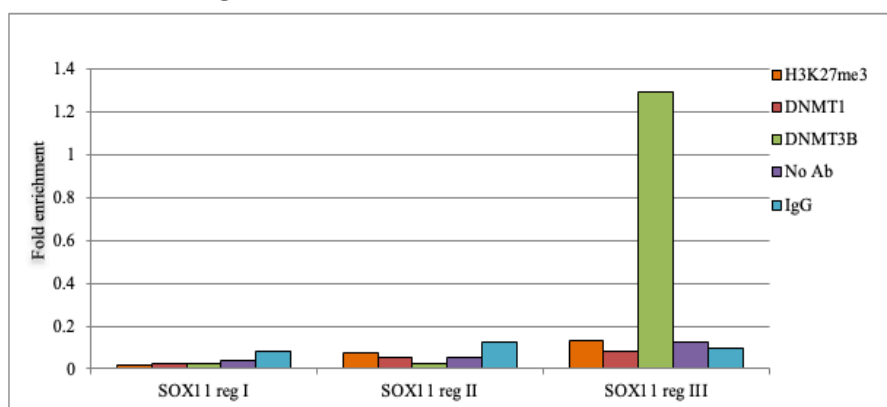
### 5.3.3 HPV 18 integration associated with epigenetic switching of tumour suppressor genes

Four tumour suppressor genes (TSG) were investigated, SOX11, CALCA, CDKN2A and ESR1. No enrichment of DNMT1, DNMT3B or H3K27me3 in the promoter regions of SOX11 and ESR1 was observed in the PHFK HPV18 cell line at p7. In contrast, we observe significant enrichment of H3K27me at promoter region I of CALCA (Figure 75), and promoter region II of CDKN2A (Figure 76). I showed in Figure 76 and 77 significant binding of DNMT1 at region II of the ESR promoter and region I of the CDKN2A promoter in p14 relative to that observed for the IgG and the no antibody controls. Interestingly, there were also significant enrichments of DNMT3B at region II of both CALCA (Figure 75) and SOX11 (Figure 74) relative to that observed for IgG. In the HeLa cell line, the binding of DNMT1 in region I within the CALCA promoter was significantly greater than that observed for the IgG and no antibody controls (Figure 75). However, no binding of the DNMT1, DNMT3B and H3K27me3 proteins to the promoter regions of SOX11 was observed. I also showed an enrichment of H3K27me3 in region I of the ESR1 promoter in HeLa cells, as demonstrated by the faint PCR band in (Figure 77). Nevertheless, densitometric analysis of the region II ESR1 ChIP data obtained with ImageJ software was not statistically significant. In HeLa cells, I showed significant enrichments of DNMT1, DNMT3B and H3K27me3 in region I of the CDKN2A promoter relative to that observed for the IgG and no antibody controls (Figure 76). I also showed in Figure 75 a significant enrichment of H3K27me3 in region II of the CDKN2A promoter. In my ChIP investigations, I included three overlapping primers covering over 1kb of the upstream region of the CALCA promoter. I excluded CALCA III primer results due to inconsistent results between the two independent ChIP experiments. A summary of these ChIP studies is presented on Table 30.

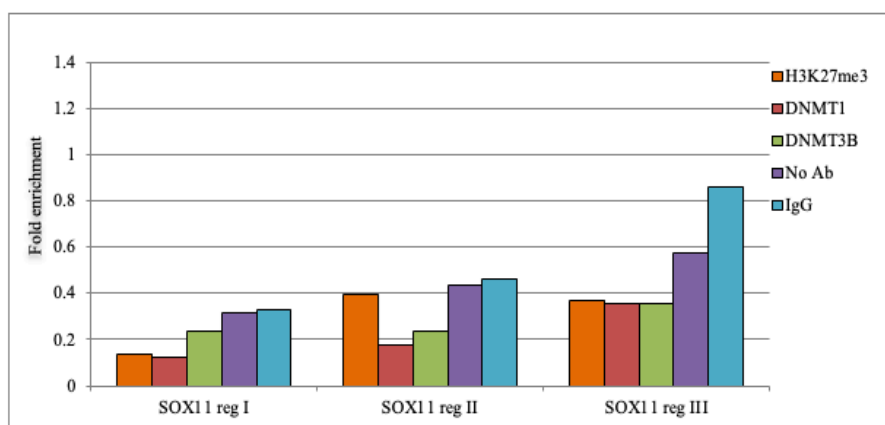
a: PHFK HPV18 p7



b: PHFK HPV18 p14



c: HeLa



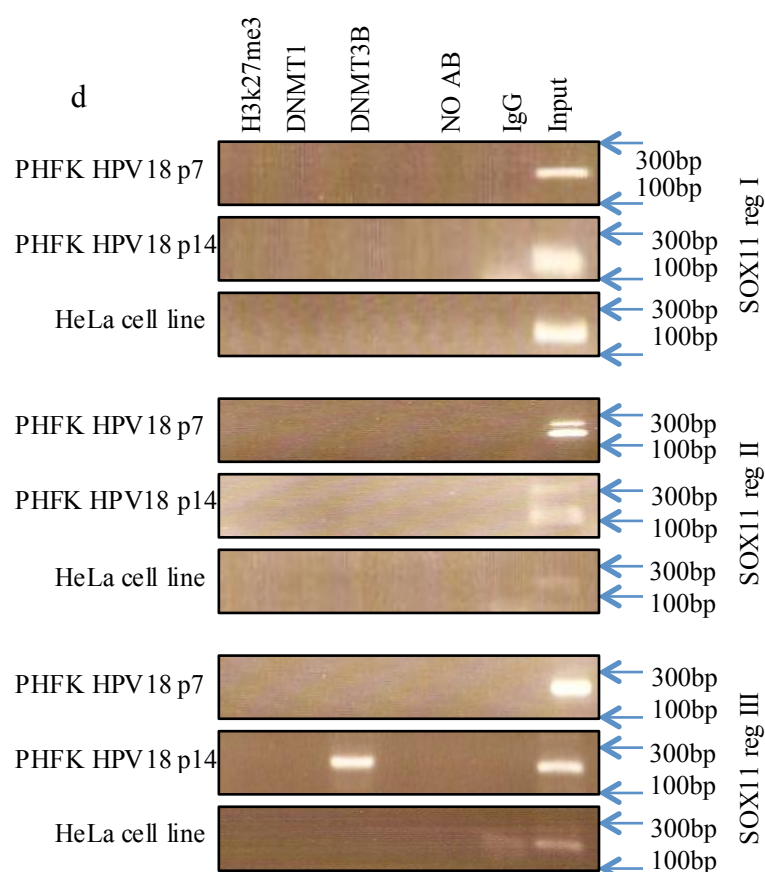
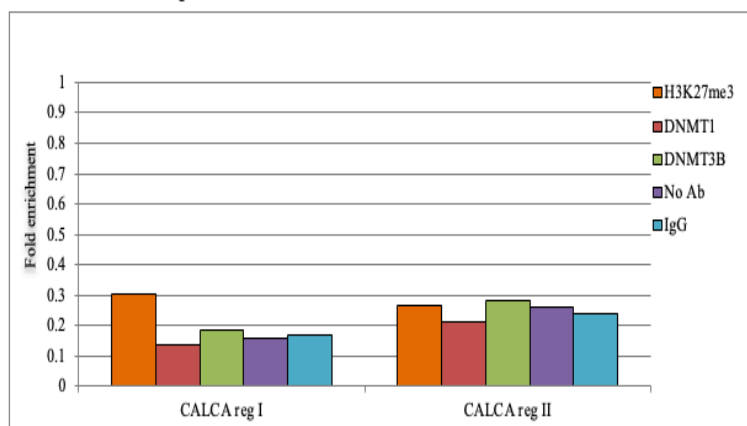
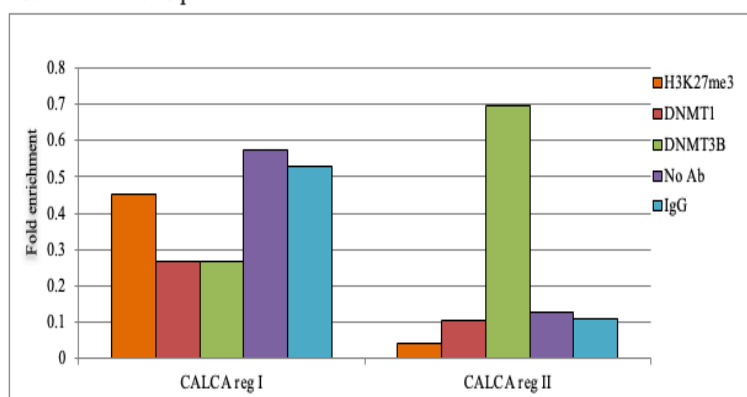


Figure 74: ChIP analysis showing DNMT3B binding to a region of SOX11 promoter. Figure a, b and c show densitometric analyses of ChIP data obtained with ImageJ software. Y-axis represents fold enrichment. Error bars indicate the standard error for of mean of two independent a, ChIP analysis of the PHFK HPV18 cell line at p7 shows no significant enrichment. b, ChIP analysis of the PHFK HPV18 cell line at p14 shows a significant enrichment of DNMT3B at region II of SOX11 relative to IgG. c, ChIP densitometry analysis of the HeLa cell line shows no significant enrichment of DNMT3B relative to IgG. d, PCR of the binding of H3K27me3, DNMT1 and DNMT3B proteins in regions I, II and III of the SOX11 promoter of the PHFK HPV18 cell line, p 7 and p14, and the HeLa cell line.

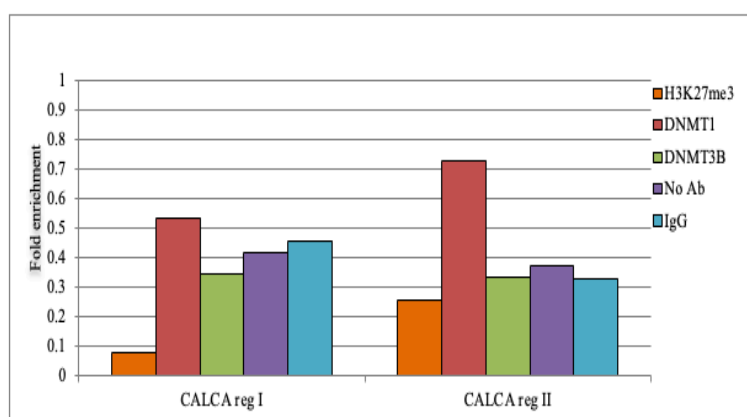
a: PHFK HPV18 p7



b: PHFK HPV18 p14



c: HeLa





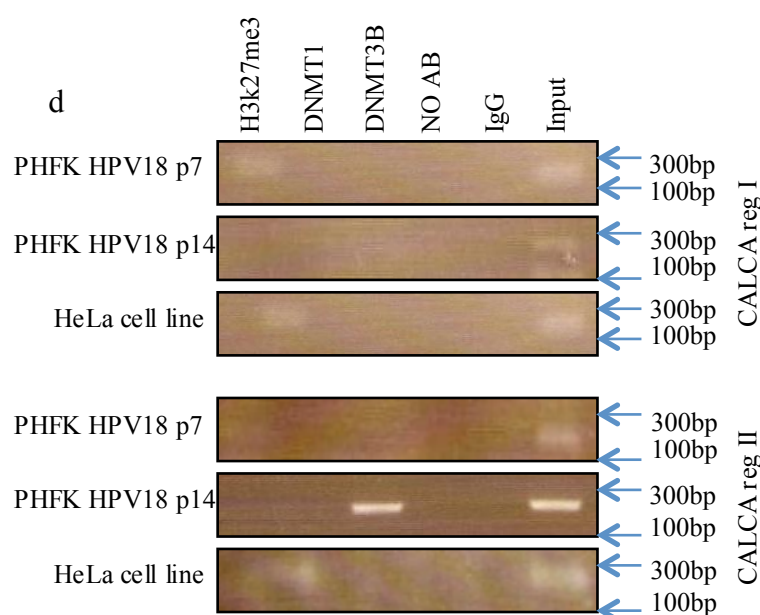
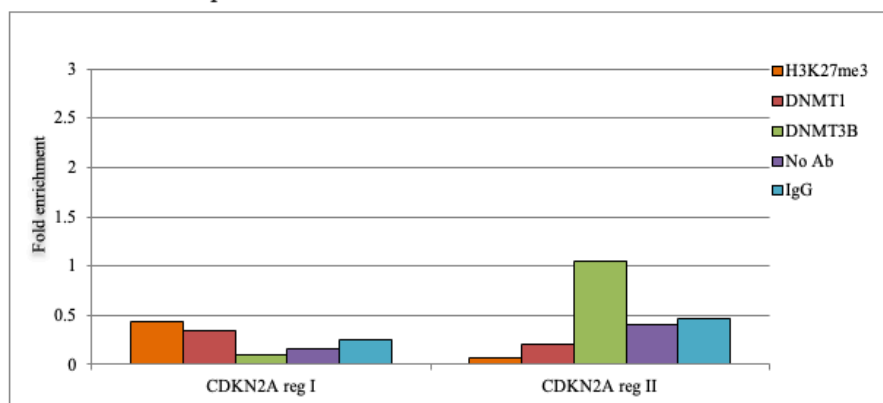
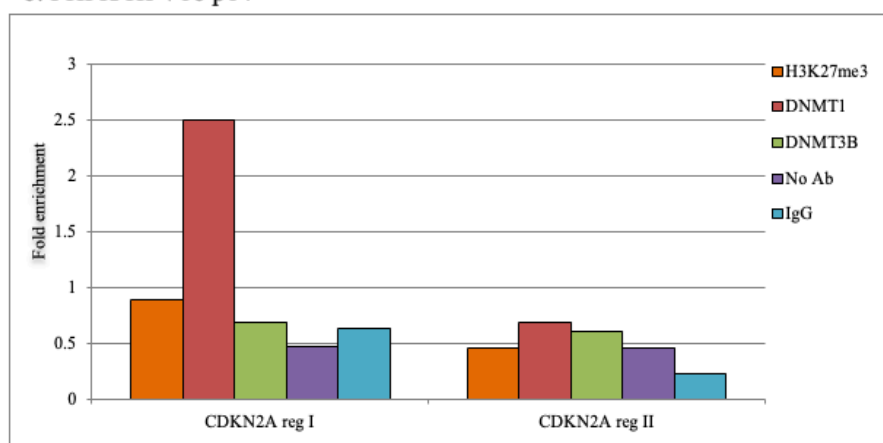


Figure 75: ChIP analysis showing DNMT3B binding to regions of the CALCA promoter. Figure a, b and c show densitometric analyses of ChIP data obtained with ImageJ software. Y-axis represents fold enrichment. Error bars indicate the standard error for of mean of two independent a, ChIP analysis of the PHFK HPV18 cell line at p7 shows significant enrichment of H3K27me3 at region I of CALCA relative to IgG. b, ChIP analysis of the PHFK HPV18 cell line, p14, shows significant enrichment of DNMT3B at region II of CALCA region II. c, ChIP densitometry analysis of the HeLa cell line shows enrichments of DNMT1 at regions I and II of CALCA. d, PCR of the binding of the H3K27me3, DNMT1 and DNMT3B proteins to regions I and II of the CALCA promoter of the PHFK HPV18 cell line, p7 and p14, and the HeLa cell line.

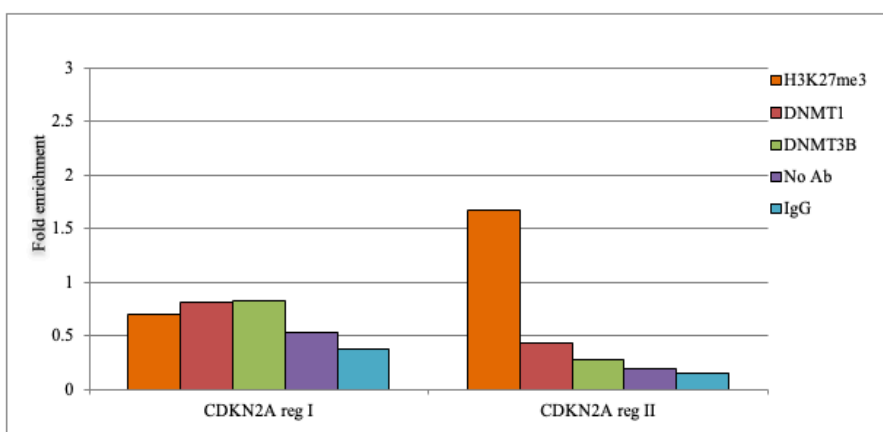
a: PHFK HPV18 p7



b: PHFK HPV18 p14



c: HeLa



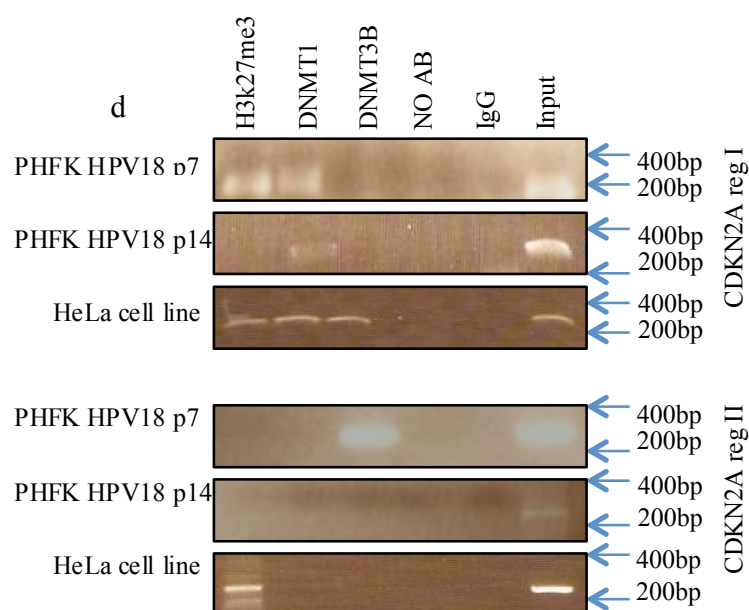
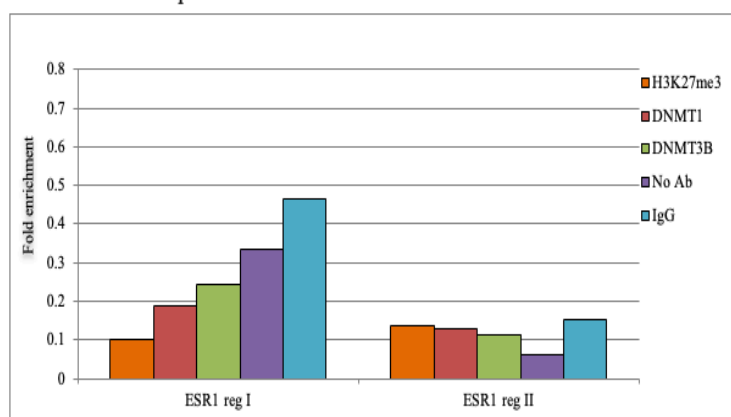
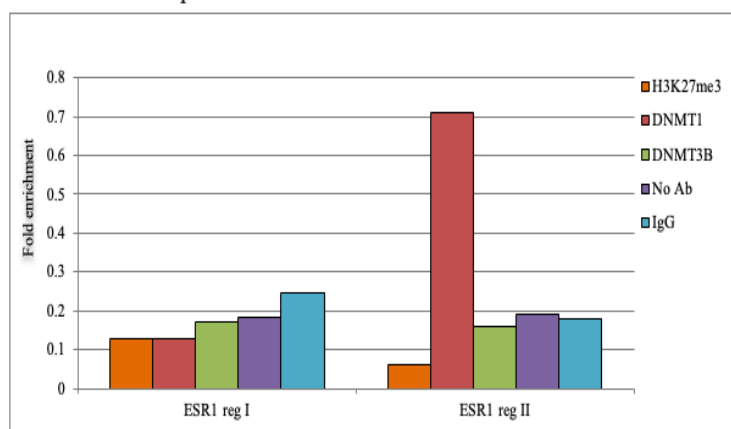


Figure 76: ChIP analysis showing H3K27me3, DNMT1 and DNMT3B binding to regions of the CDKN2A promoter. Figure a, b and c show densitometric analyses of ChIP data obtained with ImageJ software. Y-axis represents fold enrichment. Error bars indicate the standard error for of mean of two independent experiments a, ChIP analysis of the PHFK HPV18 cell line, p7, shows significant enrichments of H3K27me3 and DNMT1 at region I of CDKN2A relative to IgG. b, ChIP analysis of the PHFK HPV18 cell line, p14, shows a significant enrichment of DNMT3B at region II of CDKN2A. c, ChIP densitometry analysis of the HeLa cell line shows significant enrichments of H3K27me3 at regions I and II of CDKN2A. It also shows significant enrichments of DNMT1 and DNMT3B at region I of CDKN2A. d, PCR of the binding of H3K27me3, DNMT1 and DNMT3B proteins to regions I and II of the CDKN2A promoter of the PHFK HPV18, p7 and p14, and the HeLa cell line.

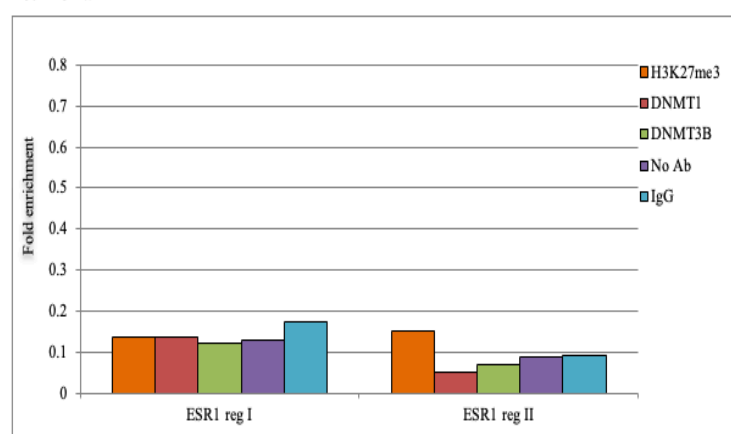
a: PHFK HPV18 p7



b: PHFK HPV18 p14



c: HeLa



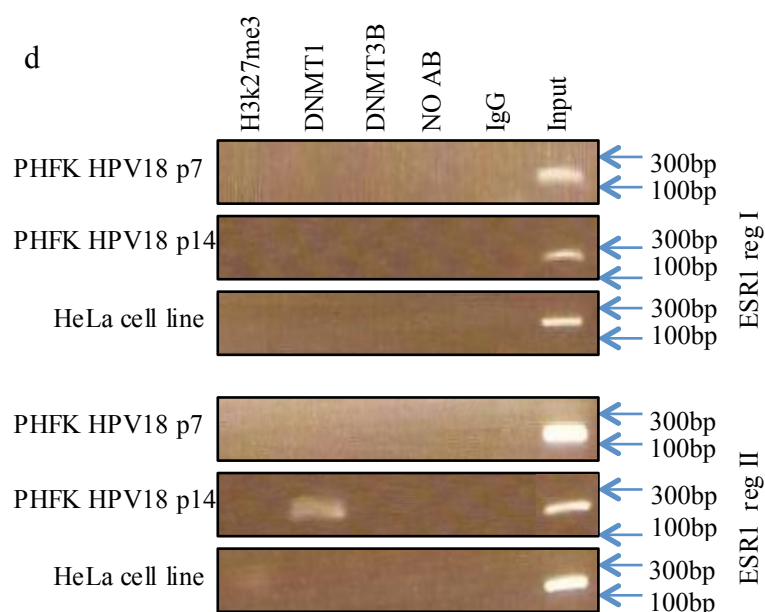


Figure 77: ChIP analysis showing DNMT1 binding to regions at the ESR1 promoter. Figure a, b and c show densitometric analyses of ChIP data obtained with ImageJ software. Y-axis represents fold enrichment. Error bars indicate the standard error for of mean of two independent a, ChIP analysis of the PHFK HPV18 cell line, p7, shows no significant enrichment. b, ChIP analysis of the PHFK HPV18 cell line, p14, shows a significant enrichment of DNMT1 at region II of ESR1 relative to IgG c, ChIP densitometry analysis of the HeLa cell line shows no significant enrichment. d, PCR of the binding of the H3K27me3, DNMT1 and DNMT3B proteins to regions I and II of the ESR1 promoter of the PHFK HPV18 cell line, p7 and p14, and the HeLa cell line.

Table 30: Summary of ChIP experimental results for p7 and p14 of the PHFK HPV18 cell line and HeLa cell line. H3K27me3, DNMT1 and DNMT3B indicate enrichment and hashes indicate no detected enrichment for the investigated promoter regions of the indicated genes.

Gene-symbol	PHFK HPV18 p7	PHFK HPV18 p14	HeLa
ROBO1	-	DNMT1	-
ROBO2	-	-	DNMT1
ROBO3	H3K27me3	DNMT1/DNMT3B	-
SLIT1	-	-	-
SLIT2	DNMT1	DNMT1	-
SLIT3	-	H3K27me3	-
SOX11	-	DNMT3B	-
CALCA	H3K27me3	DNMT3B	DNMT1
ESR	-	DNMT1	-
CDKNA2	DNMT3B	DNMT1	H3K27me3/DNMT1/DNMT3B

In summary, I demonstrated, in the previous section, the selective enrichment of DNMT1, DNMT3B and H3K27me3 at the promoters of key TSGs in p7 and p14 of PHFK HPV18 cell line and HeLa cell line. A subset of these genes showed evidence of epigenetic switching from H3K27me3 mediated silencing to DNA methylation following HPV18 integration using both pathway and candidate gene approaches.

## 5.4 Reversibility of DNA methylation of TSGs

The mechanism of action of 5-azacitidine is not fully understood. Azacitidine is a chemical analogue of the cytosine nucleoside. 5-azacitidine leads to inhibition of DNMTs, causing hypomethylation of DNA and the reversal of aberrant DNA methylation (Creusot, Acs and Christman, 1982) and at high doses, its incorporation into DNA leads to a covalent binding with DNA methyltransferase. Thus, it prevents DNA synthesis and blocks DNA methyltransferase function (Santi, Norment and Garrett, 1984b). Studies demonstrated that genes that have been silenced in cancer by DNA methylation could be activated after treatment with demethylating agents (Sova et al. 2006). This raises the question whether the DNA demethylating agent, 5-azacytidine or its congener 5-aza-2'-deoxycytidine has the same effect on non-cancerous cells. In this section, I study and compare the effect of exposure to 5-aza-2'-deoxycytidine on the HeLa and the PHFK HPV18 lines. I chose three genes (VHL, ESR1 and PTNP6) known to be significantly hypermethylated at promoter regions using the Illumina HD 450K methylation array and previously validated by pyroSeq. Following treatment of the PHFK HPV18 cell line, p14, and the HeLa cell line with two doses (low and high) of 5-aza-2'-deoxycytidine, I explore changes in the expression of epigenetic regulators. Then I study the changes in DNA methylation of the selected TSGs (VHL, ESR1 and PTNP6) by pyroSeq to validate the Illumina methylation array.

### 5.4.1 Cell viability following treatment with a demethylating agent

I showed the suppression of cell viability in a concentration-dependent manner in both HeLa cells and PHFK HPV18 cell line in response to 5-aza (Figure 78). I observed that 5-aza-2'-deoxycytidine was highly toxic to HeLa cells. Cell viability was reduced dramatically, from 85% to just over 15%, following treatment with 5-aza. Unexpectedly, the PHFK HPV18 cell



line showed mild toxicity and cell viability maintained more than than 40% at the highest dose (Figure 78).

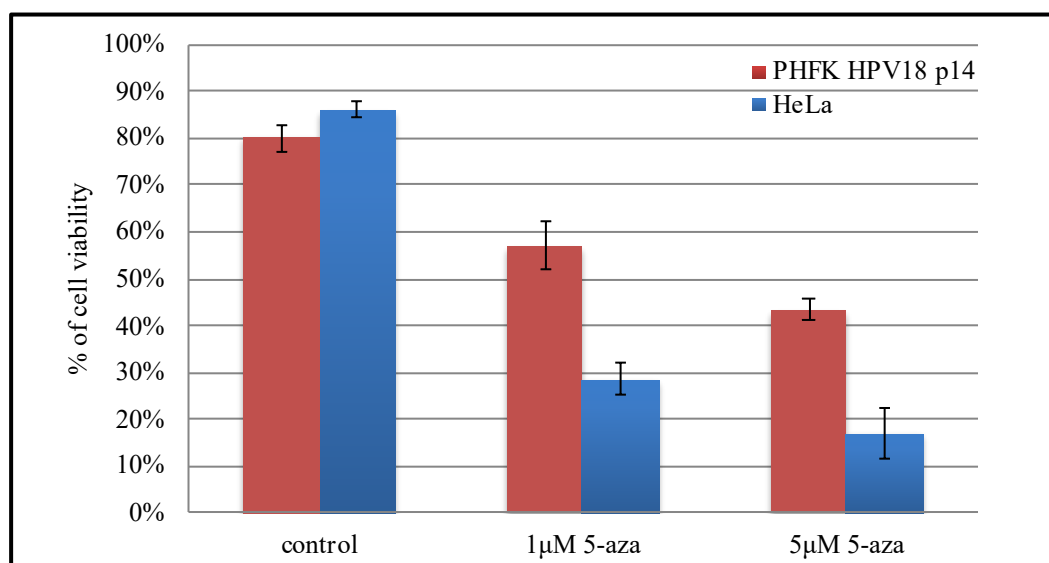


Figure 78: Cell viability of HeLa and PHFK HPV18, p14, cell lines cultured and treated with two different doses of 5-aza-2'-deoxycytidine. Following 72 hours of treatment, the Trypan blue cell viability assay was calculated manually. Standard error bars show the variation in cell viability across triplicate experiments. Y axis represents % of cell viability.

### **5.4.2 The demethylating agent induces down-regulation of epigenetic regulators.**

The demethylating agent 5-aza-2'-deoxycytidine (5-aza) binds to DNMT1, down regulating its expression and reducing its activity (Kaminskas et al., 2005). To ensure that the demethylation treatment worked, I analysed the expression of DNMT1 following treatment with two doses (1  $\mu$ M and 5  $\mu$ M) of 5-aza. It has been shown that treatment with 1  $\mu$ M and 5  $\mu$ M 5-aza effectively restores gene expression in several cervical cancer cell lines, including the HeLa cell line (Sova et al., 2006). I used qPCR to analyse changes in mRNA expression. I further investigated the expression of other epigenetic regulators, whose expression I previously showed was modulated following HPV18 integration. These epigenetic regulators are DNMT3A, DNMT3B, EZH2 and the H3K27me demethylases, KDM6A and KDM6B. I observed that the downregulation of DNMT1 required the higher dose of 5-aza in both HeLa and PHFK HPV18, p14, cell lines (Figure 79). I also demonstrated a remarkable downregulation of all epigenetic regulators with the high dose (5 $\mu$ M) of 5-aza in the HeLa cell line. However, no changes were observed with the low dose. On the other hand, the low dose of 5-aza downregulated the expression of EZH2, KDM6A and KDM6B in the PHFK HPV18 cell line, p14. The high dose had an effect on both DNMT1 and DNMT3B mRNA expressions levels. Interestingly, I observed no changes in DNMT3A expression following treatment with both doses in the PHFK HPV18 cell line, p14, despite the downregulation that was observed in HeLa cells with 5 $\mu$ M 5-aza.

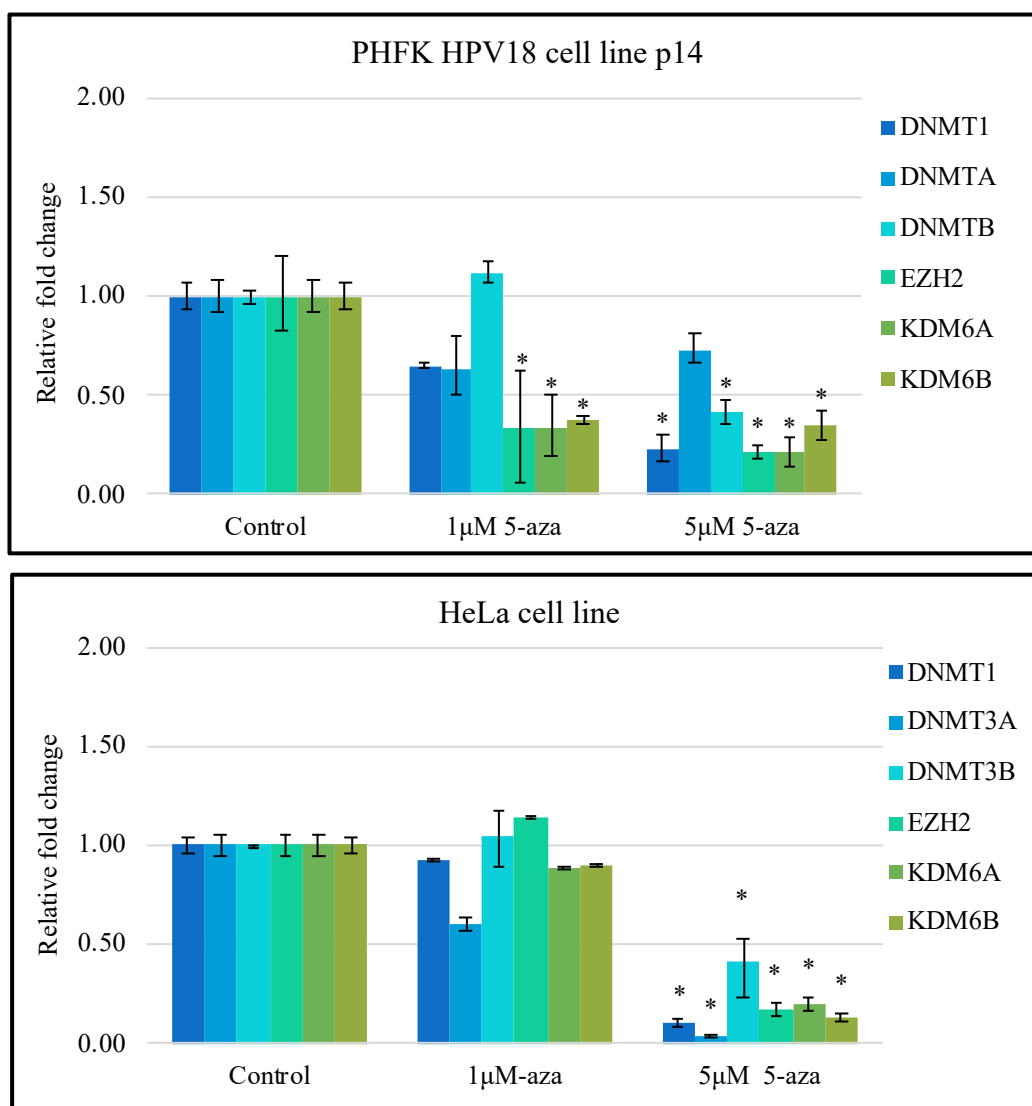


Figure 79: Analysis of changes in the mRNA expression of epigenetic regulators following treatment with 5-aza-2'-deoxycytidine (5-aza). The expression analysis includes DNMTs, EZH2 and H3K27me3 demethylases in both the PHFK HPV18, p14, and HeLa cell lines with two doses of 5-aza. Results shown are typical of 3 similar experiments. Each experiment was run in triplicate. \*  $P \leq 0.05$ , two-tailed paired nonparametric Mann-Whitney U test. Error bars indicate standard deviations. Y axis represent relative fold change.

## 5.5 Reversibility of DNA methylation of TSGs in the PHFK

### HPV18 cell line

Demethylation agents such as 5-aza-2'-deoxycytidine can restore gene expression by causing demethylation of genes silenced by hypermethylation (Kaminskas et al., 2005). I used pyrosequencing technique to investigate methylation changes in DNA obtained from cells treated with 5-aza, as it has shown the ability to downregulate its target gene, DNMT1. I showed that the loss of average methylation is more than 5% in almost all genes (VHL, ESR1 and PTNP6) with the two different doses of 5-aza (Table 31). The average changes in methylation at the ESR1 promoter dropped by 7% at the high dose, and almost all loss of methylation occurred in the second and forth CpGs, while no changes were observed in the other two CpGs examined (Figure 80). Interestingly these are the same CpGs that showed hypermethylation following integration at p14 compared with p7 (Table 32). More pronounced demethylation was seen in the VHL promoter. I demonstrated over a 20% methylation loss at three out of the four CpGs. The second CpG showed no changes with different doses of the demethylation agent. Remarkably, a 10% loss of methylation in PTNP6 was observed following treatment with low doses of 5-aza. Only one CpG, at position three of the five CpGs was responsible for the average loss of methylation. That CpG was responsible for the hypermethylation changes that were observed following HPV integration (Figure 80). Overall, a high dose of 5-aza demethylated the promoters of the three TSGs in the PHFK HPV18 cell line, p14. Of particular interest, hypermethylated CpGs following integration were the same CpGs that underwent demethylation following treatment with 5-aza (Figure 80). Thus, the methylation of CpGs in promoters following the integration of HPV18 into the human genome is a reversible process.

Table 31: Summary of pyroSeq analysis shows the average DNA methylation changes of TSGs PHFK HPV18 cell line p14. Hypomethylation of three TSGs (VHL, ESR1 and PTNP6) following treatment of the PHFK HPV18 cell line, p14, with two different doses, 1  $\mu$ M and 5  $\mu$ M, of 5-aza-2'-deoxycytidine (5-aza).

UCSC Gene Name	Sample	% AVG methylation
ESR1	PHFK HPV18 p14 C	20
	PHFK HPV18 p14 1 $\mu$ M 5-aza	15
	PHFK HPV18 p14 5 $\mu$ M 5-aza	13
PTPN6	PHFK HPV18 p14 C	36
	PHFK HPV18 p14 1 $\mu$ M 5-aza	26
	PHFK HPV18 p14 5 $\mu$ M 5-aza	32
VHL	PHFK HPV18 p14 C	67
	PHFK HPV18 p14 1 $\mu$ M 5-aza	45
	PHFK HPV18 p14 5 $\mu$ M 5-aza	44

Table 32: Summary of pyroSeq analysis shows the average DNA methylation changes of TSGs. Hypomethylation of three TSGs (VHL, ESR1 and PTNP6) following treatment of the HeLa cell line with two different doses, 1  $\mu$ M and 5  $\mu$ M, of 5-aza-2'-deoxycytidine (5-aza).

UCSC Gene Name	Sample	% AVG methylation
ESR1	HeLa	52
	HeLa 1 $\mu$ M 5-aza	41
	HeLa 5 $\mu$ M 5-aza	22
PTPN6	HeLa	83
	HeLa 1 $\mu$ M 5-aza	68
	HeLa 5 $\mu$ M 5-aza	50
VHL	HeLa	83
	HeLa 1 $\mu$ M 5-aza	82
	HeLa 5 $\mu$ M 5-aza	56

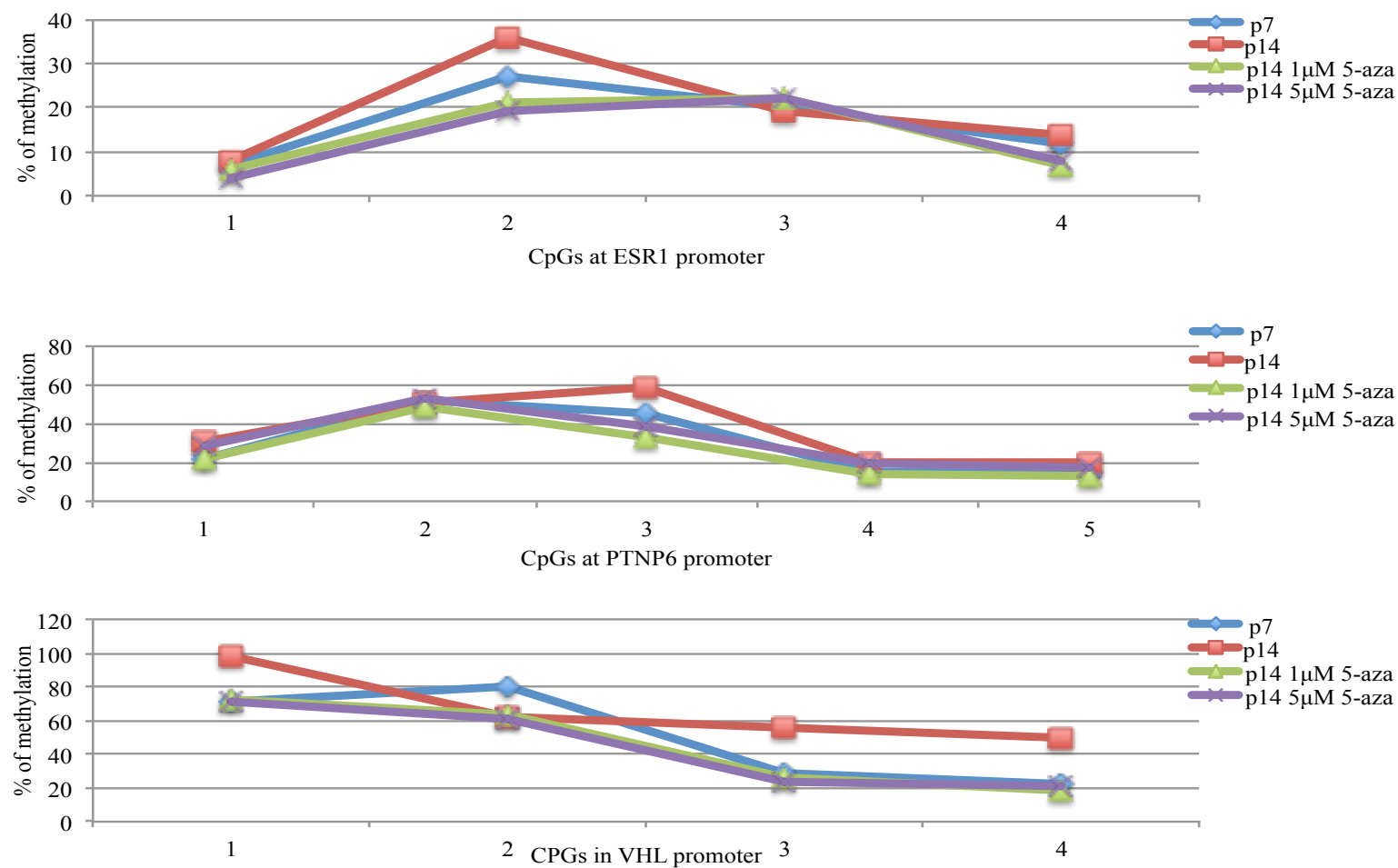


Figure 80: Summary of methylation changes of CpGs at the promoters of ESR1, VHL and PTNP6 following demethylation treatment of the PHFK HPV18 cell line, p14. Passage (p).

### **5.5.1 Reversibility of DNA methylation of TSGs in a cervical cancer cell line**

Here, I utilise the same genes and methods for the PHFK HPV18 cell line to investigate and study the reversibility of DNA methylation in the HeLa cell line. I show remarkable hypomethylation changes in TSGs in the HeLa cell line, with over a 25% loss of average methylation. The changes I noticed in ESR1 and PTNP6 methylation in the PHFK HPV18 cell line were dose dependent, (Figure 81). In contrast, VHL showed no response to 1  $\mu$ M 5-aza, and a loss of over 25 % of average methylation of the last two CpGs were observed with the high dose (Figure 81). The pyrogram of ESR1 includes four CpGs, of which the first two showed progressive loss of methylation with an increasing dose of demethylating agent. Specifically, the third and fourth CpGs were hypomethylated with 1 $\mu$ M and 5  $\mu$ M doses, respectively. The study of the pyrogram of the PTNP6 promoter showed a methylation loss in the first CpG in response to 1  $\mu$ M 5-aza, and no changes in the second CpG (Figure 81). The hypomethylation observed in the last three CpGs occurred mainly in response to 5  $\mu$ M 5-aza. Overall, the methylation changes in the three TSGs in the HeLa cell line were reversible.



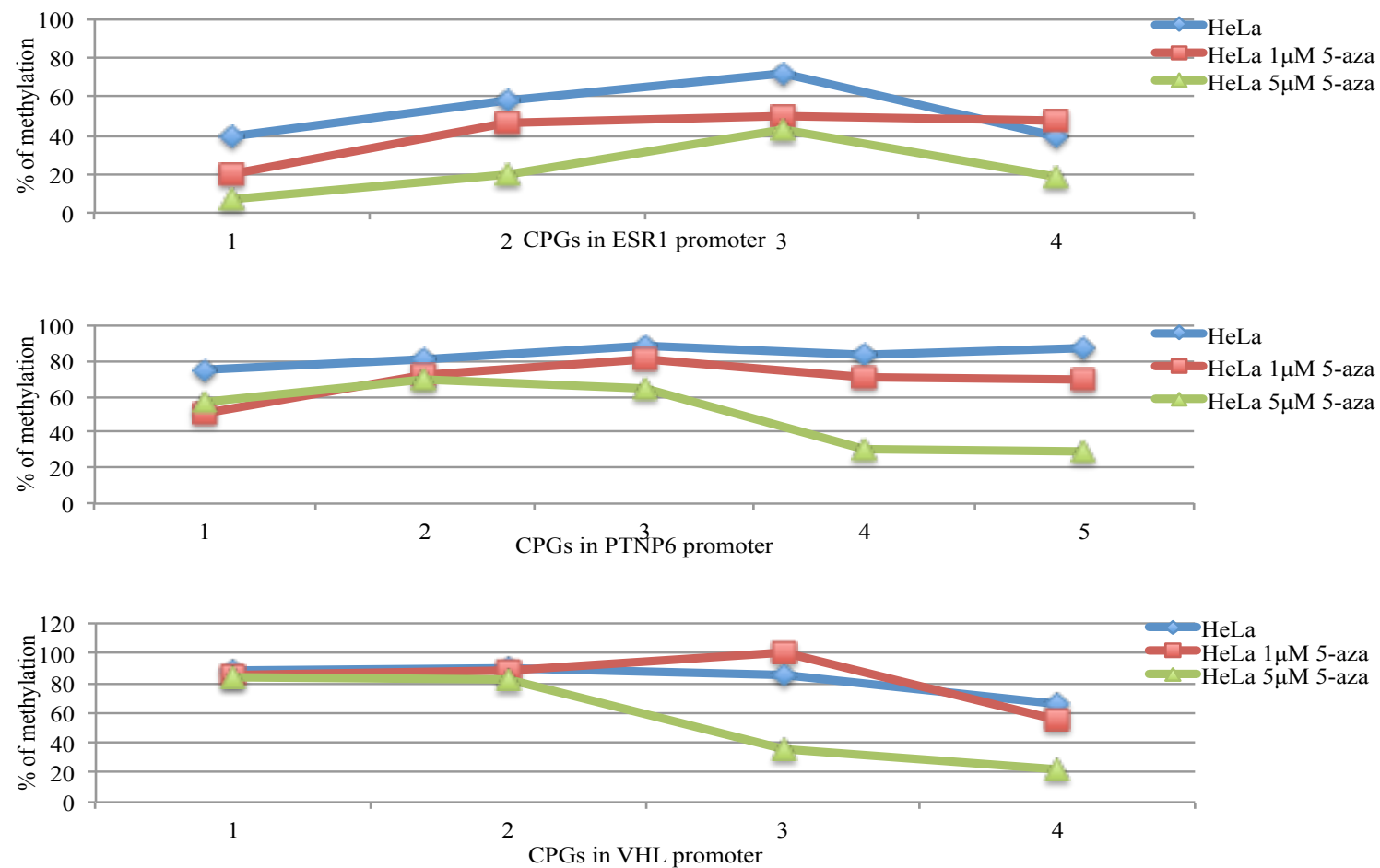


Figure 81: Summary of methylation changes of CpGs in the promoters of ESR1, VHL and PTNP6 following demethylation treatment of the HeLa cell line.

## CHAPTER 6

# **DISCUSSION**

## 6 DISCUSSION

### 6.1 Cytopathic changes of PHFK following transfection with HPV

Given the anatomical site preferences and epitheliotropic nature of the high-risk HPVs, primary human foreskin keratinocyte (PHFK) provides an appropriate model to study the early stages of HPV induced cervical carcinogenesis. The relevance of an *in vitro* system for studying HPV induced carcinogenesis depends on how closely the system reflects observed *in vivo* changes. Both cell lines derived from human foreskin keratinocytes and from the transformation zone of the cervix have been used as *in vitro* systems to study the interaction of HPV with its natural target cell (Dürst et al., 1987; Nakao et al., 1996). Considering nearly 90% of cervical neoplasia arise from the transformation zone, *in vitro* systems derived from the transformation zone of the cervix seem likely to represent an optimum model for studying HPV induced carcinogenesis. Nevertheless, the transformation zone consists of two distinct cells - columnar and squamous epithelial cells. In addition, squamous metaplasia formed in transformation zone by merging a mixed population of parental cells that may prevent a clear resolution of the effects of HPVs (Nakao et al., 1996).

In this thesis, PHFK obtained from the same neonate were transfected with three biologically different HPVs; HPV18 and two variants of HPV16B and HPV16K. These biological replicates indicated that the changes observed were virus-induced rather than human genetic variations. Thus, this *in vitro* system allowed me to investigate to a certain extent the differences between the two most important variants in HPV-induced cervical cancer; HPV16 and HPV18. Nevertheless, investigating the changes associated with each of these viruses in PHFK from different donors is vital to validate the results. The changes in PHFK growth properties observed in this thesis after transfection with episomal HPVs reflect alterations in the cytoskeleton and loss of the polygonal packaging that is characteristic of normal

keratinocytes. However, malignantly transformed cells are known to lose their polygonal packing and exhibit altered cytoskeletal organization (Fey and Penman, 1986). The morphologically altered phenotypes of PHFK after transfection with HPVs is thus a consequence of the transformed nature and changes of transfected keratinocytes to a more undifferentiated type rather than neoplastic changes (Pirisi et al., 1987). Indeed, cellular morphological assessment remains the main component in cervical screening programmes. Cytological changes are needed to identify women with HPV-associated lesions and offer them clinical management. As a consequence of HPV infection, HPV-infected keratinocytes show cytopathic changes such as koilocytosis, multinucleation and spindle koilocytes. Cytopathic changes provide significant discriminating power for detecting true HPV lesions from non-HPV lesions (Roteli-Martins et al., 2001).

The most reliable cytopathic evidence of clinical HPV infection is koilocytosis. It has been shown that cooperation between HPV early proteins E5 and E6 induces koilocyte formation in human cervical cells. Where koilocytosis is associated with meganuclei or abnormal mitoses, it is defined as koilocytotic atypia, which is an indicator of high-grade cervical intraepithelial lesions (Zioli et al., 1998). The most distinctive feature of koilocytes is multinucleation. In fact, multinucleation is multilobation as shown in Figure 82, in these cells HPV oncoproteins induce G2 checkpoint aberrations, resulting in chromosomal instability and as a consequence, the formation of a non-dividing, multilobated nucleus (Cho et al., 2005). Using DNA breakage detection-fluorescence in situ hybridization (DBD-FISH), it has been shown that HPV infection leads to chromatin changes in koilocytes (Cortés-Gutiérrez et al., 2010). On the other hand, the mechanism behind spindle shaped nuclei formation has not been thoroughly studied. Since the first discovery of koilocytosis as a pathognomonic morphological change of the HPV infection in 1955 (Koss, 2012) and the introduction of morphology-related

cervical screening in England in 1964, several measures have been taken to enhance the sensitivity of the cervical screening programme. The introduction of HPV testing to cervical screening has offered no increase in the detection rate of early stage of cervical neoplasia (TOMOLA group, 2009). The HPV testing aimed to detect the presence or absence of an HPV infection. Nevertheless, cervical smear programmes still rely heavily on cervical cytopathic changes to predict histological abnormalities. Observations in this thesis are in agreement with an *in vivo* study by Martins et al., (2011) reported spindle nuclei and other abnormal shaped nuclei at late passages of both PHFK HPV16 and PHFK HPV18 cell lines. This is consistent with observations that *in vivo*, koilocytosis and cytomegaly are more frequent in women with a histological diagnosis of early cervical neoplasia.

Interestingly, I found in the *in vitro* models some cytopathic changes that are expected following cervical HPV infection *in vivo*. These observations is in agreement with data from cervical smears and histology studies. The HPVs cytopathic effects manifest early in a productive HPV infection and keratinocytes exposed to HPV18 are reported to present less cytopathic changes compared with HPV16 infected cells (Kovacic et al., 2006). Furthermore, multinucleation and spindle koilocytosis were pathognomonic of early and late passages of PHFK, respectively (Martins et al., 2011). Meganuclei or an increased nuclear-to-cytoplasmic ratio is required to diagnose squamous intraepithelial lesions. Although not frequently detected in this *in vitro* model, meganuclei were observed in late passages, with no clear HPV type preference. Both nuclear abnormalities and raised nuclear-to-cytoplasmic ratios may create a strong suspicion of dyskaryosis. However, chromatin and/or nuclear abnormalities are a prerequisite for the diagnosis of a squamous intraepithelial lesion SIL in cervical smears (Cervical screening Programme, 2013). Thus, most of cytopathic changes associated with HPV infection *in vivo* are also manifested in this *in vitro* model.

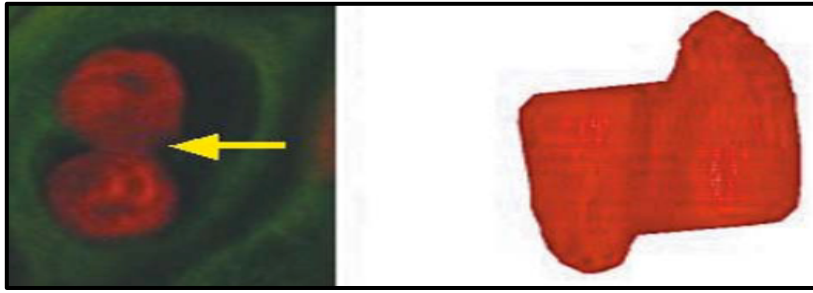


Figure 82: Binucleation is an artefact of conventional microscopy. Binucleation reveals a single, central, fused, bilobated nucleus. Surface and wire frame rendering showed bilobation connected by an anaphase bridge (yellow arrow). Figure adapted from Cho et al., (2005).

## **6.2 Disruption of HPV physical state associated with epigenetic reprogramming of DNA methyltransferases**

The transfection of primary human keratinocytes with episomal HPVs and followed by extended passaging that mimics the epithelial changes induced by each HPV type. These models allow changes that occur early in the HPV life cycle to be observed. However, PHFK transfected with full length HPV in tissue culture lose the episomal form and become fully transformed with extended passaging (Seagon et al., 1994). I investigated the physical state and the stability of the HPV episomal form, using PCR method that shows a disruption of full length HPV E2 that indicates HPV integration. The PCR method for detecting HPV integration has low sensitivity for detecting the integrated form in the presence of large numbers of episomal forms. The PHFK HPV18 cell line showed partial loss of full length HPV E2 commenced at passage 12 but full disruption happens only at passage 14. Whereas, PHFK HPV16 cell lines showed two different patterns. The first pattern was seen in the PHFK HPV16B cell line, which did not show complete disruption of E2. This partial loss of full length E2 at passage 10, suggests there is a mixture of viral genomes with both disrupted and intact HPVE2 at this point. The other pattern was seen with PHFK HPV16K cell line, where the full length E2 disruption was seen early at passage 7. However, the only variation reported between HPV16B and HPV16K is L1 sequence difference and the effect of this difference on their oncogenic role remain unclear (Kirnbauer et al., 1993).

The overlapping designed primers that amplify the HPVE1/E2/E5 allow me to study the approximate location of disrupted region in HPV genomes. I first validated the primers by investigating cancer cell lines, where genome breakpoints have been studied before with different methods. In HeLa, it is known that HPV integration disrupts the region from 3100-5736 nucleotides (Luft et al., 2001; Collins et al., 2009). The result in this thesis, demonstrated

loss of the HPV18 sequence from nucleotides 2828 to 3994 in the HeLa cell line. Similar to HeLa, the PHFK HPV18 cell line lost part of the HPV18 E2/E5 region, albeit a smaller sequence, from 3086 to 3994. The PCR method I employed in my investigation identified the disrupted regions in the HPV genome, but would not show the exact breakpoints. Thus, failure to amplify certain region of full length HPV E2 would not rule out a possibility of more than one breakpoint. In fact, studies using whole-genome sequencing technique have supported this suggestion. (Liang et al., 2014; Adey et al., 2013). Only three breaks have been identified in the HPV18 genome following integration to human DNA, with disruption observed at nucleotides sequences 829, 3274 and 3535 (Liang et al., 2014). Interestingly, using PCR I also show no amplification of the region that span the last two breakpoints observed. In this thesis however I did not characterise the region with nucleotide sequence number 829. In contrast, a study by Adey et al., (2013) that also employed whole genome sequencing to investigate integration of HPV18 into the human genome, showed three integration breakpoints of partial HPV18 genome at the HPV18E1/E2/E5 region. The HPV18 region, which failed to be amplified according to those breakpoints, would be the region from 3088 to 5736 nucleotides. Thus, collaboratively, my results build upon previously published data, to suggest that sequences within the HPV18 E2/E5 regions are specifically targeted when double stranded HPV18 DNA integrates into the human genome both *in vivo* showed in HeLa cell line and *in vitro* cell line manifested in PHFK HPV18 cell line.

Interestingly, the two PHFK HPV16 cell lines that showed either partial or complete loss of full length E2, where two regions failed to either partially or completely to amplify, which are from 2701-3119 and 3710-3916 respectively. This would suggest more than one breakpoint. The region that showed complete disruption in my PHFK HPV16 *in vitro* models, overlapped with the region disrupted in W12 cell line (from 3732 to 4791 nucleotides) as shown by Pett et



al., (2004). Thus, the breakpoint or points would be possibly anywhere from 3732-3916 in the HPV16 genome sequence. Interestingly, these results further suggest that a specific region of the HPV genome is prone to disruption *in vivo* and *in vitro*.

In agreement with previous research (Alazawi et al., 2002), I found that endogenous overexpression of HPV E7 oncoprotein preceded E2 disruption. In fact, Alazawi et al., (2002) found overexpression of HPV E7 oncoprotein was associated with detection of a mixture of episomal and integrated forms when assessed by southern blotting suggesting that an assay able to detect HPV E7 protein expression could predict HPV integration better than PCR methods. The E7 oncoprotein expressed early at passage 4 in PHFK HPV18. In contrast, only passages 9 and 10 expressed E7 in the PHFK HPV16B cell line (Figure 57). However, I experienced difficulty in finding a good anti-HPV16E7 antibody for immunoblotting. Thus, undetected expression of the HPV16E7 oncoprotein in early passages even at low levels was more likely due to weak signals, but it did not rule out the possibility of differences between HPV types. The expression level of HPV E7 protein does not correlate with its activity (Giarre et al., 2001). However, active HPV E7 protein promotes pRb degradation (Giarre et al., 2001). Thus, low pRb protein expression in PHFK HPV16 and PHFK HPV18 cell lines suggests active HPV E7 oncoprotein expression.

Expression profile of DNMTs in the cervical cancer progression models, showed elevated expression of DNMTs that parallel to HPV-induced cervical carcinogenesis. Elevated expression of DNMTs follow HPV transfection and a more sustained raise level of expression correlates with the overexpression of the HPV E7 oncoprotein that precede the disruption of full length HPV E2. High levels of DNMT1 expression was also seen in late stage of CIN that is represented by late passage W12 cells. DNMT1 expression remains elevated throughout the cervical cancer progression model. However, *de novo* DNMTs show stage or cell line

preference. The PHFK transfected with HPV, show no elevated expression of DNMT3A and significant expression of DNMT3B. in cancer cell lines, the HeLa cell line expressed DNMT3A mRNA and Ca Ski expressed DNMT3B mRNA. In contrast, late passage W12 cells behave like the SiHa cell line, expressing both expressed both *de novo* DNMTs, DNMT3A mRNA and DNMT3B mRNA.

A regulator of *de novo* DNA methylation, DNMT3L showed no expression in *in vitro* model. Untransfected keratinocytes and PHFK HPV cell lines showed no expression of DNMT3L mRNA. Nevertheless, HPV induced changes to *de novo* DNA methyltransferase DNMT3B, no effect was observed on its regulator, DNMT3L. Even in late passage W12 cell lines late passages, DNMT3L expression was not detected. In fact human normal cervical cells do not express DNMT3L, correlating with the DNA methylation of its promoter (Gokul et al., 2007). Investigating DNMT3L mRNA level in cervical cancer cell lines showed no expression of DNMT3L in SiHa and very low levels in the Ca Ski and HeLa cervical cell lines. Correlation between DNA methylation profile at DNMT3L promoter and DNMT3L mRNA expression were observed in human primary cervical cancer and cancer cell lines. Low levels of DNA methylation was observed in HeLa cells which showed higher level of expression of DNMT3L in comparison with SiHa where very low level of DNMT3L expression correlated with methylation of most of CpGs promoter (Gokul et al., 2007). Unexpectedly, I detected very high levels of DNMT3L expression in the HPV-negative, cancer cell line C33A. It is worth mentioning that DNMT3L is a biomarker of human embryonic but not somatic carcinoma (Minami et al., 2010). A key role of DNMT3L in carcinogenesis was suggested by the fact that ectopic overexpression of DNMT3L in HeLa cells showed substantial changes in the expression profile of genes important in cell cycle regulation and nuclear reprogramming (Gokul et al., 2007). DNMT3A was not expressed early in *in vitro* model for cervical carcinogenesis. In

PHFK transfected with HPVs, there was no change in expression of DNMT3A after transfection with either HPV16 or HPV18. This could be due to the probe I used to analyse DNMT3A expression since the transcript involved was specific to isoform a but not isoform b. However, using the same probe in the dyskaryotic W12 cell line increased DNMT3A expression and DNMT3A levels were raised in both HPV16 and HPV18 cervical cancer cell lines relative to C33A. Another explanation could be that DNMT3A is not expressed in PHFK HPV cell lines as DNMT3A and DNMT3B are both expressed in a tissue-specific manner and I found a strong up-regulation of DNMT3B instead (Watanabe et al., 2002).

I found that expression of DNMT1 and DNMT3B transiently increased after transfection, followed by a more stable rise at later passages of the PHFK cell models. We showed that the transient upregulation of DNMT1 and DNMT3B after transfection was associated with hypermethylation of many genes reported to be differentially methylated in cervical neoplasia (Leonard et al., 2012). However, this two-peak pattern was only seen in PHFK HPV18 and PHFK HPV16B cell lines. The PHFK HPV16K cell line showed only one peak of DNMT1 and DNMT3B overexpression. Since these cell lines all have the same genetic origin, this pattern of differences suggests viral changes to be the cause, with the most likely change being an altered viral physical state. Investigation of the HPV physical state in in vitro model systems and expression profile of HPVE7 oncoprotein facilitate the interpretation of the results in relation to any HPV life cycle events.

The expression profile of DNMTs showed a positive correlation between increase of DNMT1 and DNMT3B expression and enzymatic activity and the overexpression of HPVE7 in both PHFK HPV18 and PHFK HPV16 cell lines. Where the first peak of DNMT1 and DNMT3B expression and enzymatic activity followed the HPV infection, the second peak coincides with disruption of HPVE2 and precisely with overexpression of HPVE7 oncoprotein.

Whereas, the PHFK HPV16K cell line showed only one peak of DNMT1 and DNMT3B expression which would suggest early HPV integration that PCR fails to detect in the presence of a large excess of the episomal form of the virus. However, it seems that splice variants and the consequences of different isoforms with different activities and tissue preferences play a major role in explaining the differences between DNMT mRNA and protein expression (Robertson et al., 1999). Interestingly, there was a correlation between DNMT3B mRNA isoform 7 expression and HPV integration. DNMT3B mRNA isoform 7 overexpression was more stable and prominent in the second peak and it was also stable in the cell lines with a one-peak pattern. Specific expression of DNMT3B7 has been linked to cancer (Ostler et al., 2007). It has also been recently shown that overexpression of the DNMT3B7 isoform results in both hypermethylation and hypomethylation defects and promotes tumorigenesis (Shah et al., 2010). This makes it more likely that expression of DNMT3B7 is induced by HPV oncoprotein.

One of the genes shown to have numerical changes in tumour cells is DNA methyltransferase (DNMT3B). DNMT3B located in cytoband 20q11.21. Deletion of DNMT3B suppresses neoplastic cells (Lin et al., 2006), whilst DNMT3B overexpression promotes tumorigenesis by causing transcriptional silencing and hypermethylation of specific tumour suppressor genes (Linhart et al., 2007). The increased levels of DNMT3B expression in human cancer are suggested to be due to DNMT3B gene amplification (Esteller et al., 2011). The correlation between DNMT3B mRNA expression and gene copy number are conflicting. A study by Wilting et al., (2006) using genome-wide microarray-based comparative genomic hybridization, DNMT3B gene copy number and elevated expression of DNMT3B mRNA revealed high level of correlation of squamous cell cervical carcinoma (Wilting et al., 2006). In contrast, comparison between copy number and genes expression by cytoband 20q11 using microarrays of single nucleotide polymorphism showed no significant gain of gene copy

numbers, in spite of gene upregulation in cervical cancer cell lines. Generally, the correlation between gene copy number and changes in gene expression in cervical cancer cell lines is found to be poor. It is thought that other factors, such as epigenetic mechanisms, may influence gene expression within the genome segments (Vazquez-Mena et al., 2012). Although, the two studies used two different methods, validation of DNMT3B copy number by fluorescent in situ hybridization (FISH) were performed by Wilting et al., (2006). Furthermore, Wilting et al., (2006) investigated SiHa and cervical carcinoma, whilst Vazquez-Mena et al., (2012) involved only cervical cancer cell lines. However, cervical cancer cell lines usually retain their genetic properties (Soder et al., 1997). In this thesis I showed a high-level of amplification of DNMT3B gene copy number in SiHa and late passage W12 cell lines, which correlated with elevated DNMT3B mRNA expression. On the other hand, I found no correlation in PHFK transfected with HPV16 and HPV18 at late passages, and both HeLa and Ca Ski cell lines.

Indeed, the mechanism of DNMT1 overexpression in cell line transfected with E7 revealed in investigation by Burger et al., (2006). However, the mechanism of overexpression of DNMT3B in HPV expressing cell lines has not been investigated before. In agreement with previous published study (Burger et al., 2007), I show that increased DNMT1 expression in PHFK expressing HPV E7 is associated with an increase in its enzymatic activity. Specifically, these results indicate that the endogenous HPV16E7 oncoprotein induce DNMT3B and an increase of DNMT3B enzymatic activity which correlates well with the expression of E7 oncoprotein in both PHFK HPV18 and PHFK HPV16. Using Co-IP, I demonstrated that HPV16E7 and HPV18E7 oncoproteins interact with DNMT3B at physiological levels when they are co-expressed in untransformed and transformed human keratinocytes. However this result did not revealed any direct interaction, and further analysis to E7-DNMT3B protein complex is required. In particular, the known direct interaction of DNMT1 with HPV E7

(Burgers et al., 2007) and DNMT1 with DNMT3B (Kim et al., 2002) could suggest there are complex protein interactions.

### **6.3 The changes of DNA methylation pattern following HPV18 integration**

Results in this thesis revealed that the integration of HPV18 into PHFK induces a pattern of genome-wide methylation similar to that seen in neoplasia. Interestingly, almost all of chromosomal locations that predicted by Illumina HD 450K methylation array to be hotspot regions for hypermethylated genes and differentially methylated regions (DMRs), mapped to mapped to known integrated HPV sites. Specifically, chromosomal hotspot locations 5q31, 19q13.3 and 16p13.3 were mapped to known HPV integration sites (Kraus et al., 2008 ; Yu et al., 2005 ; Peter et al., 2006). Schmitz et al., (2012) determined HPV 16 and HPV18 integration sites of human cervical carcinoma using DIPS assay. Both HPV18 and HPV 16 cervical carcinoma shared chromosome 19 as integration site. In contrast, specifically both chromosome 17 and the X chromosome showed to be integration sites of HPV18 but not HPV16 cervical carcinoma. Actually, Illumina HD 450K methylation array predicted both chromosome 17 and the X chromosome to be hypermethylated site following HPV18 integration in PHFK.

Descriptive analysis of genes that undergo methylation changes following HPV18 integration showed hypomethylation of a substantial proportion of CpGs (77.5 %) that occupied the body of investigated genes. Indeed, cancer cells are characterised by global progressive hypomethylation compared to their normal counterparts (Esteller, 2008). The hypermethylated CpGs represented less than one-quarter of the observed hypermethylation changes following integration. In agreement with the changes predicted in cancer cells, almost all of the hypermethylated CpGs occurred in promoters and within CpG islands. In addition, significantly

methyated genes following HPV18 integration were over twenty times more likely to be methylated in cervical neoplasia. Nevertheless, studies of cervical carcinogenesis have proposed different methylation markers, although three genes (DABK1 and RARB and CADM1) were shown to be consistent hypermethylated (Mares et al 2013) in all studies. My results confirm the hypermethylation of these TSGs early in the processes of cervical carcinogenesis, following integration of HPV18 in the PHFK HPV18 cell line. I also validated the Illumina methylation array externally with cervical cell lines, and the validated genes were methylated following integration in both HPV18 and HPV16 cervical cancer cell lines. Using pyrosequencing, I show for the first time that both DCC and TP53AIP1 genes are hypermethylated in cervical cancer cell lines. Promoter hypermethylation of DCC in cervical cancer cell lines could explain the downregulation of DCC expression in cervical cancer that has previously shown by Saegusa & Okayasu (1999).

Following HPV18 E7 oncoprotein overexpression, the methylated genes were shown to comprise particular biological processes. Unexpectedly, I found that the pathways enriched in the PHFK HPV18 cell line, following HPV18 integration, are common pathways deregulated by DNA methylation across cancer types (Kim et al., 2012). Both the immune response and receptor binding processes are hypomethylated in most cancers and were shown to be enriched in hypomethylated genes following HPV integration. Furthermore, the hypermethylated pathways are homeobox, signal transduction and neuroactive pathways. In fact, a study highlighted the extensive differential methylation of regulators of the neuroactive pathway in tumour cells harbouring HPV (Colacino et al., 2013). Two of the main signals transduction pathways, Wnt and cadherin, showed significant hypermethylation of their genes. The Wnt pathway has an established role in cervical carcinogenesis and gene regulators of the cadherin signalling pathway have also been implicated in cancer. The hypermethylation of CpGs in the

promoters of cadherin genes, which I showed following HPV integration, could be responsible for their reduced expression in keratinocytes expressing the HPV E7 oncoprotein (Caberg et al., 2008). The cadherin signalling pathway plays an important role by allowing the immune system to mount a protective, cell-mediated response against transformed cells. Thus, it is possible that the HPV E7 oncoprotein induced an immunodeficiency via dysfunctional antigen presentation. Overall, genome-wide methylation analysis results showed that early in cervical neoplasia and specifically following HPV18 integration epigenetic changes could predict changes occurred later in cervical carcinoma.

## **6.4 Does integration of HPV18 associated with epigenetic switching?**

In the *in vitro* model systems and specifically in PHFK HPV18 cell line there was an overexpression of both DNMT1 and DNMT3b following cell propagation described as a second peak. Concomitant, with overexpression of DNMTs following HPV E7 overexpression, H3K27me3 methyltransferase and demethylases were also up regulated. Both H3K27me3 and DNA methylation are repressive marks in cancer and, unexpectedly, their regulators showed patterns consistent with epigenetic switching following HPV episomal loss and overexpression of HPV E7 oncoprotein. Using ChIP assays I showed epigenetic switching from H3K27me3 mark to DNA methylation in the late passage of PHFK HPV18 cell line.

In addition, to the changes I showed to DNMTs, in *in vitro* cervical cancer progression model. Expression profile of H3K27me3 methyltransferase showed that EZH2 overexpression was parallel to the cervical cancer progression model. I showed expression of EZH2 mRNA in cervical cancer cell lines and W12 cell lines. Expression of both EZH2 mRNA



and protein are more pronounced in cells expressing a high level of endogenous HPV16 E7 oncoprotein represented with late passages of PHFK HPV16 cell line and previously where represented with cells expressing ectopic HPV16E7 (McLaughlin-Drubin et al., 2011; Münger, 2011; Hyland et al., 2011b). However, I here also showed overexpression of endogenous EZH2 in cells expressing endogenous HPV18 E7. Concomitant with EZH2 overexpression, I also unexpectedly found overexpression of H3K27me3 demethylase *in vitro* models. Overexpression of both EZH2 and KDM6A has been reported previously in prostate cancer (Xiang et al., 2007) and it has been shown that H3K27me3 methyltransferase, EZH2 overexpression does not result in increased H3K27me3 methylation activity, instead EZH2 required to be incorporated into a PRC2 complex to methylate H3K27 (Kuzmichev et al., 2005; Margueron et al., 2009). In contrast, KDM6A and KDM6B demethylate the H3K27me3 mark (Agger et al., 2007). HPV infections suppressed both H3K27me3 demethylases, KDM6A and KDM6B expression, however following extended passaging KDM6A and KDM6B were overexpressed and mimicked DNMTs pattern of expression in both PHFK HPV16B and PHFK HPV18 cell lines following endogenous expression of HPVE oncoprotein. Two studies have already demonstrated overexpression of H3K27me3 demethylase in cells expressing HPV16 E6/E7 oncoproteins (McLaughlin-Drubin et al., 2011; Hyland et al., 2011). It is worth mentioning that KDM6A is located on X-chromosome. It was known previously as UTX and it escapes X-chromosome inactivation. My model system is established from male neonate foreskin. However, UTY represents the male counterpart, and is highly homologous with UTX and they both shared a structural relationship (Greenfield et al., 1998).

Interestingly, HPV-negative cancer cell lines showed different expression features of epigenetic marks from HPV-positive cell lines. H3K27me3 methyltransferase mRNA is remarkably overexpressed in HPV-negative cells. Of particular interest, I also found (Section

3.3.3.1) that DNMT3L is remarkably overexpressed in C33A cells, these two features are in common with embryonic stem cells. Collectively, these results would suggest that there is an epigenetic difference between HPV negative and HPV positive cervical cancer and worth further investigation.

In this thesis, ChIP allowed me to investigate 1 kb upstream of the gene promoter regions and enabled me to identify binding sites in the investigated genes. Nevertheless, I do not deem negative results as proof of the absence of protein-DNA binding, as I did not examine the entire promoter regions due to the inadequate DNA yield in the ChIP assays. Moreover, I could not calculate the significance of these ChIP results, as I was only able to perform these experiments twice. Investigating protein-DNA binding using ChIP techniques is time consuming and requires relatively large DNA quantities to identify all of the enriched DNA sequences. Alternatively, ChIP sequencing scans for genome-wide associations, which have high resolutions, can be used. However, this method is very expensive and requires expertise in analysing the data.

I investigated DNA binding to DNMT1, DNMT3B and H3K27me3 in the PHFK HPV18 cell line or in cervical cancer. Two genes showed a pattern of epigenetic switching. I did find DNMT3B binding into sites of TSGs promoter, which were confirmed by two independent experiments. Although my ChIP experiments lacked a positive control as could not find in literature any known binding sites for DNMT3B. DNMT3B enrichment was detected at the ROBO3, CALCA, SOX11 and CDKNA2 genes. The ChIP experiments results showed the enrichment of H3K27me3 at the promoter of ROBO3 in PHFK HPV18 p7 but not in p14. Thus ROBO3, a member of the Slit-Robo pathway, showed epigenetic switching from H3K27me3-mediated repression to both DNMT1 and DNMT3B following HPV18 integration. The CALCA CpG promoter showed DNA hypermethylation from p7 to p14 in PHFK HPV18 cell

line, as shown by the Illumina HD 450K methylation array and validated by pyroSeq (section 5.2.3). The ChIP assays demonstrated enrichments of H3K27me at p7, and DNMT1 on CALCA at p14 which could also indicate epigenetic switching. It has been shown that ROBO3 is methylated in high-grade squamous intraepithelial lesions (HSIL) and in cervical cancer (Narayan et al., 2006). Of particular interest, ROBO3 showed no enrichment of DNMT1 and DNMT3B in HeLa cell line, which might be explained by differences in the investigated regions or might indicate insignificant of this results. On the other hand, there was no correlation between promoter DNA methylation and expression of ROBO3 in normal cervical tissue and low-grade squamous intraepithelial lesions (LSIL), (Narayan et al., 2006). My results showed an enrichment of H3K27me3repression mark at ROBO3 promoter, which possibly explains the lack of expression and ROBO3 DNA promoter methylation early in cervical carcinogenesis. Thus, these results suggest that known methylation markers of cervical cancer that showed alterations in response to overexpression of E7 oncogenes, such as ROBO3 and CALCA, could be investigated as potential methylation markers for early cervical neoplasia.

## **6.5 Potential methylation markers for early HPV18 cervical neoplasia**

To identify methylation changes following early cervical neoplastic changes, a comprehensive analysis of genome-wide DNA methylation changes associated with HPV integration of primary human foreskin keratinocytes (PHFK) has been conducted by Illumina HD 450K methylation arrays. The disruption of full length HPV E2 and associated overexpression of oncogene E7 mimic the natural processes of cervical carcinogenesis. PHFK HPV18 cell line represents an ideal model to study methylation changes induced by this critical event in the HPV life cycle. However, only one human donor where investigated and the

Illumina array outcome data where results of re culturing frozen cells of PHFK HPV18 cell line cells again in triplicate. Thus, validating these investigations with another cell lines generated from different human donor will be vital before a final conclusion can be made on the significance of these results. Most studies have relied on investigating clinical materials, to identify novel methylation markers. However, that approach will not distinguish the methylation changes induced primarily by active transcripts of viral oncogene itself and those that are secondary to the accumulation of late critical stages of cervical carcinogenesis or other factors, such as age or other virally induced methylation.

The detection of aberrant methylation in cervical smears several years prior to progression to invasive cervical carcinoma, strongly suggests implementation of methylation markers analysis to triage women with high risk HPVs (Wentzensen et al., 2009). Among several host genes that have been evaluated as methylation markers in cervical cancer, some of the advocated TSGs lack the specificity and their promoter showed to be hypermethylation in different human cancers (Luczak & Jagodziński, 2006). These TSGs include APC, CDH1, DAPK1 and VHL and several others. I propose a panel of methylation markers for detection of early stages of cervical neoplasia. Only a panel of methylation markers that are *de novo* methylated can be used to define the DNA methylation signature (methylo type) of each type of human cancer (Esteller et al., 2001). Of almost 10,000 CpG sites showed significant hypermethylation in the PHFK HPV18 cell line, p14, only 18 genes were *de novo* methylated at promoter regions. Therefore, a panel of these genes could possibly be employed as a predictor of early methylation changes in cervical carcinogenesis. ROBO3, a member of the Slit-Robo pathway, is the only one of this group that has been implicated previously in cervical neoplasia (Narayan et al., 2006). However, it has been reported that genes methylated by H3K27me3 in human stem cells are prone to DNA hypermethylation in cancer (Gal-Yam et al. 2008).

Specifically, *de novo* methylated genes in cancer are frequently marked by H3K27me3 in human stem cells (Widschwendter et al. 2007). Therefore, most candidate genes proposed to be methylation markers of early cervical neoplasia from the analysis of the *de novo* methylated genes at promoters are BTBD11, GABRA4, KCNA4, POU4F1 and ROBO3. These genes have shown to be methylated by H3K27me3 in human stem cells. Hence, HPV reprograms both ROBO3 and CALCA by inducing epigenetic switching following integration. In addition, CALCA is another candidate gene to be added to the panel of methylation markers of early cervical neoplasia.

## **6.6 The effect of demethylation in the PHFK HPV18 cell line**

Cancer studies have shown that demethylating agents could reverse DNA methylation. However, the effect of demethylating agents on non-cancer cells is still not well known. Although 5-aza demethylates DNA by binding to DNMT1 and reducing its activity, the full mechanism of action of 5-aza is still unclear. Both cell lines I studied have shown high level of expression of DNMT1 and DNMT3B, which were downregulated following treatment with 5-aza. However, the PHFK HPV18 cell line expresses a low level of DNMT3A, and no response to both doses of 5-aza was observed. Indeed, it has been suggested that the high expression level of DNMT3B and DNMT3A is mechanistically linked to enhanced cellular sensitivity to 5-aza. Knockdown of DNMT3B expression results in substantial resistance to 5-aza (Beyrouthy et al., 2009; Oka et al., 2005). Paradoxically, in cancer cell lines increase of DNMT3B genes copy number and the overexpression of DNMT3B predicts resistance to 5-aza (Simó-Riudalbas et al., 2011). My study of gene copy number revealed no amplification of the DNMT3B gene in the in vitro model systems and the HeLa cell line did show any response to 5-aza. However, the study did not include cervical cancer cell lines, but instead focussed on breast, pancreatic, lung, and colorectal cancer cell lines were involved.

A clinical trial has precluded the use of 5-aza to treat solid tumours due to its modest antitumor effect and high toxicity (Abele et al., 1987). However, the trial did not include cases of cervical cancer. Recently, in vitro study of different cervical cancer cell lines reported the suppression of cancer cell proliferation and the reactivation of the apoptosis genes APAF-1 and RASSF1A in response to 5-aza (Yao et al., 2013). Genome-wide methylation array offer the possibility to study the overall pattern of tissue methylation, predict methylation markers and explore enrichment of different biological processes, whereas pyroSeq shows the pattern of methylation of multiple consecutive CpGs at promoter regions. However, it is time consuming and expensive to run. I only studied promoter regions of three genes following demethylation treatment. The reversibility of TSGs' methylation is possible at early changes of carcinogenesis. This finding might open the window to early treatment, particularly with the availability of demethylating agents such as hydralazine, which exhibits little toxicity. My observations revealed that only CpGs with at least 30% of methylation responded to the demethylation agent 5-aza. This is also the case for both the PHFK HPV in vitro model and the cancer cell line. The CpGs that were hypermethylated following HPV integration responded to demethylation and become hypomethylated. In addition, the response of epigenetic regulators to demethylation was dose dependent. Loss of DNA promoter methylation of the three TSGs also showed a dose-dependent effect. The percentage of DNA methylation of a particular CpG could possibly predict its sensitivity to demethylation. Studying the promoter region of VHL, I noticed no changes in DNA methylation of the first two CpGs in HeLa cells, despite a high level of methylation of over 80%. In response to treatment with 1  $\mu$ M 5-aza in the PHFK HPV18 cell line, I noticed a drop in methylation from almost 100% to 70%. However, despite increasing the dose five fold, no further changes were observed. This was also the case for other CpGs, for instance, the second CpG in PTNP6. Thus, it is possible that the percentage of DNA

methylation of a particular CpG could predict its sensitivity to demethylating agents, as well as at the specific level of treatment that would cause CpG methylation resistance to develop. A genome-wide analysis showed the ability of 5-aza to induce global changes in gene demethylation, yet only a limited number are significantly reactivated (Yang et al., 2012). New window into the epigenetic therapy of cervical cancer has been opened with clinical trials investigating the role of hydralazine, a demethylating agent with mild toxicity, as a treatment option in different cancer stages (Zambrano et al., 2005; Candelaria et al., 2007). The early methylated TSGs following integration, are reversibly demethylated (Chapter 5). It would be interesting to investigate the possible expression change of TSG following demethylation early in HPV induced cervical neoplasia. However, the time and resources were limited.

The integration of HPV into the human genome has been shown to be an important aetiological event in cervical carcinogenesis (Pett and Coleman, 2007). The HPV genome loses the episomal form and becomes integrated into host-cell DNA in high grade dysplastic lesions and HPV associated cervical carcinomas (Alazawi et al., 2002). The preliminary studies of this thesis using the in vitro model of cervical cancer, revealed that the loss of HPV episomal forms and overexpression of HPV E7 oncogenes reprogrammed the human keratinocyte and carried into a different cellular fate. Following this important event in the HPV life cycle, these cells showed morphological changes. There were expression changes of different key epigenetic regulators which were associated with methylation of different TSGs and epigenetic switching to a more stable repression mark. Genome wide methylation revealed changing in methylome pattern with a similarity to that expected in cervical cancer. Furthermore, some of proposed markers of cervical cancer showed methylation changes in human keratinocytes. Multiple genes showed methylation changes following of HPV integration, however, only 18 genes showed de novo methylation, they worth future work to investigate their methylation status in PHFK from

different donors. The investigations in these thesis however, have many limitations. Only one donor where available to be investigated, transfected cells before and after episomal loss were cultured in triplicates and harvested separately and submitted for the process of methylation arrays. It would be usful if these changes in TSGs and methylation markers further studied in other PHFK transfected with HPV 18 from other donors using the pyrosequencing technique.

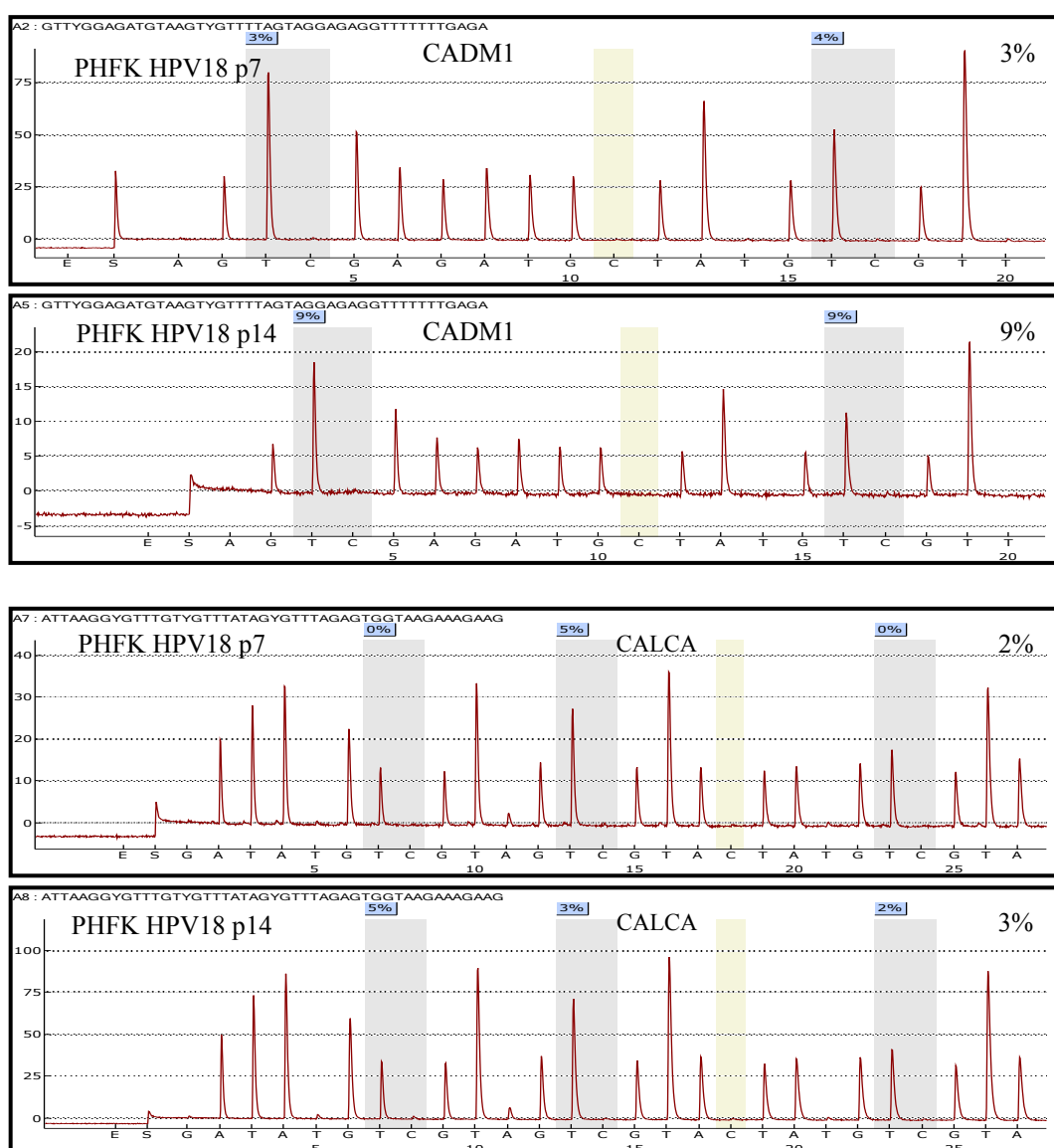


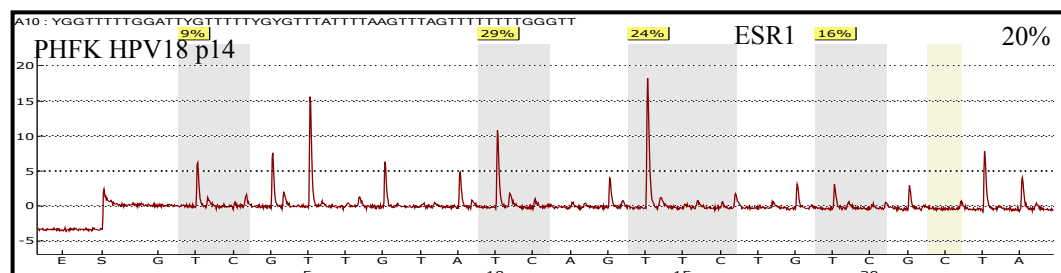
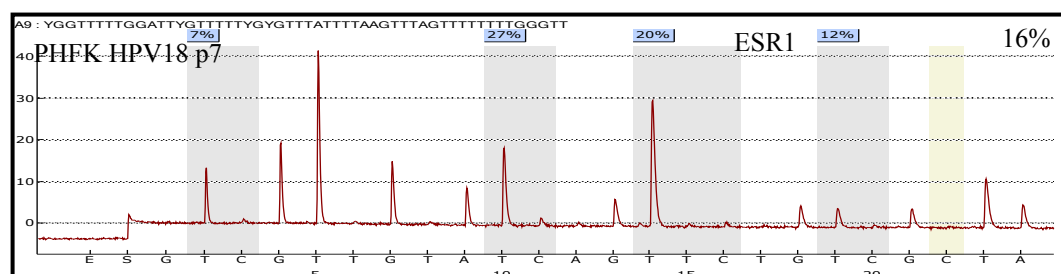
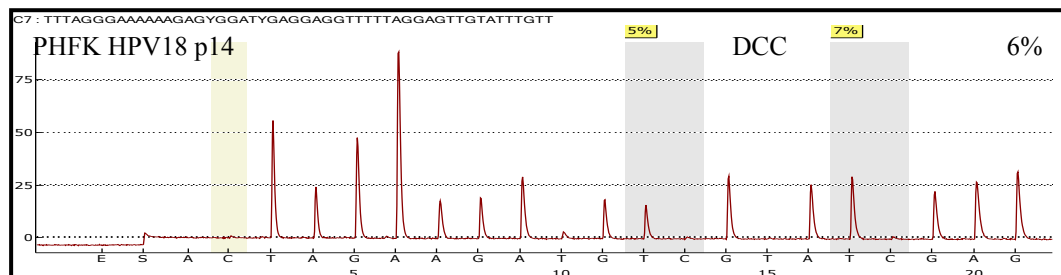
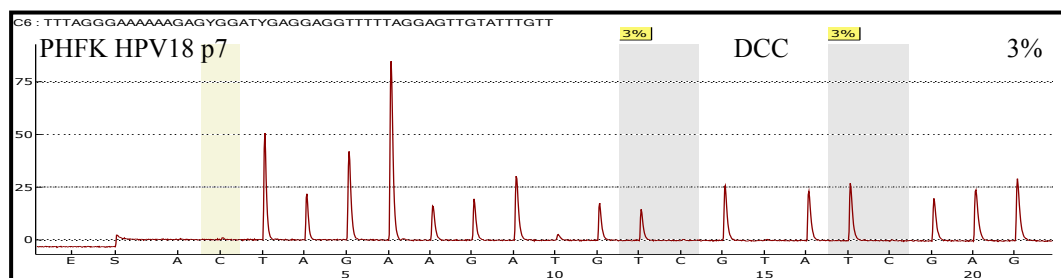
# APPENDICES

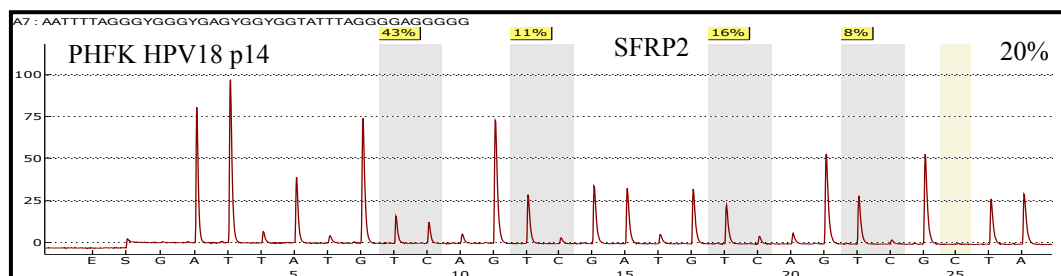
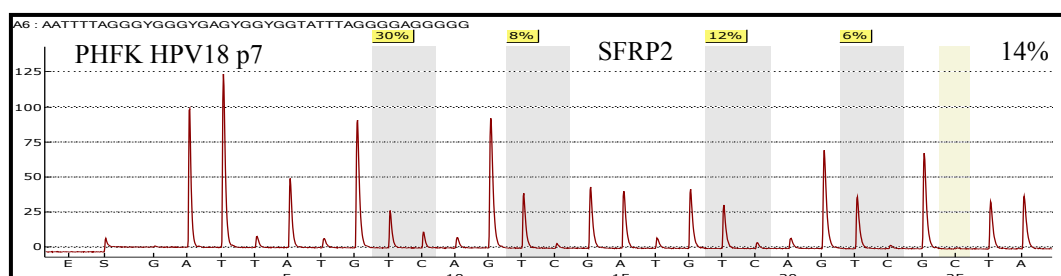
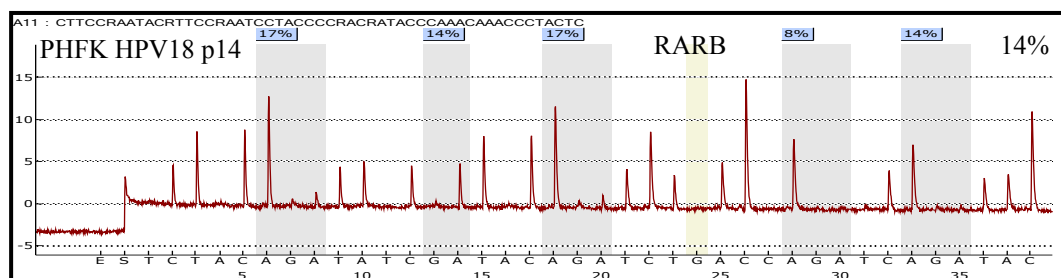
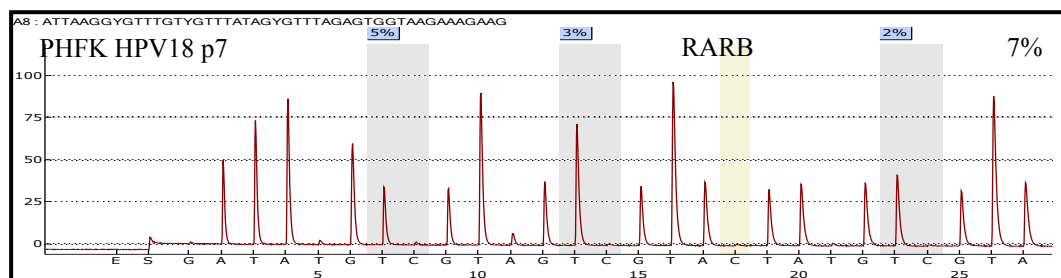
## 7 Appendices

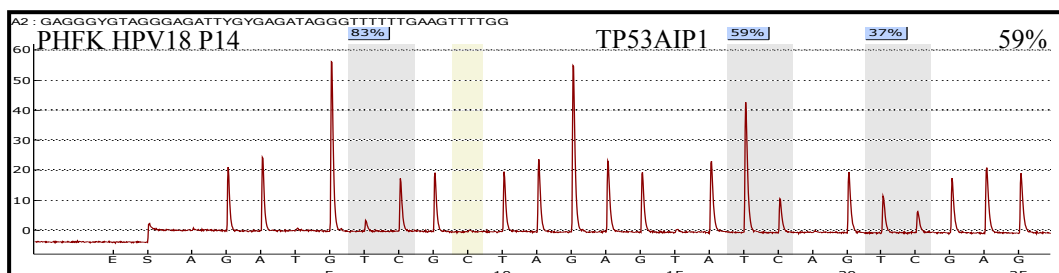
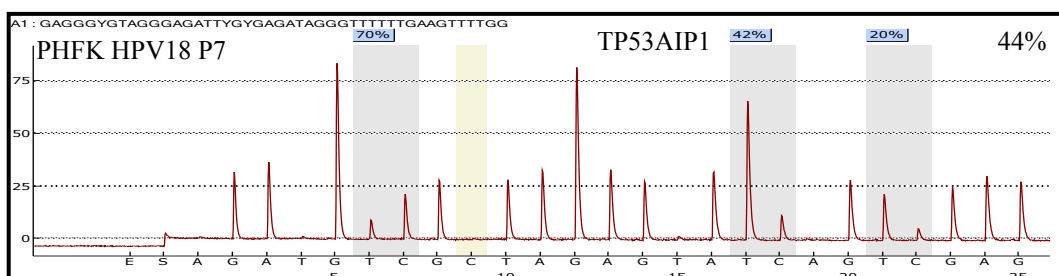
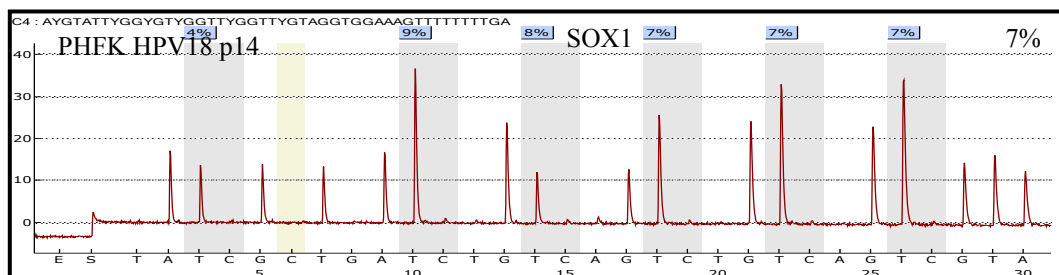
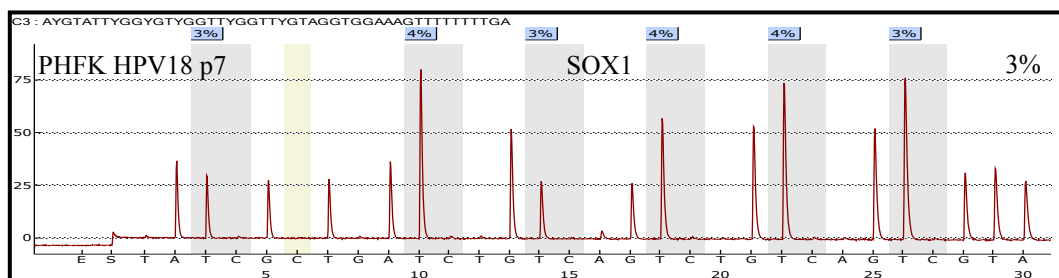
### 7.1 Appendix I

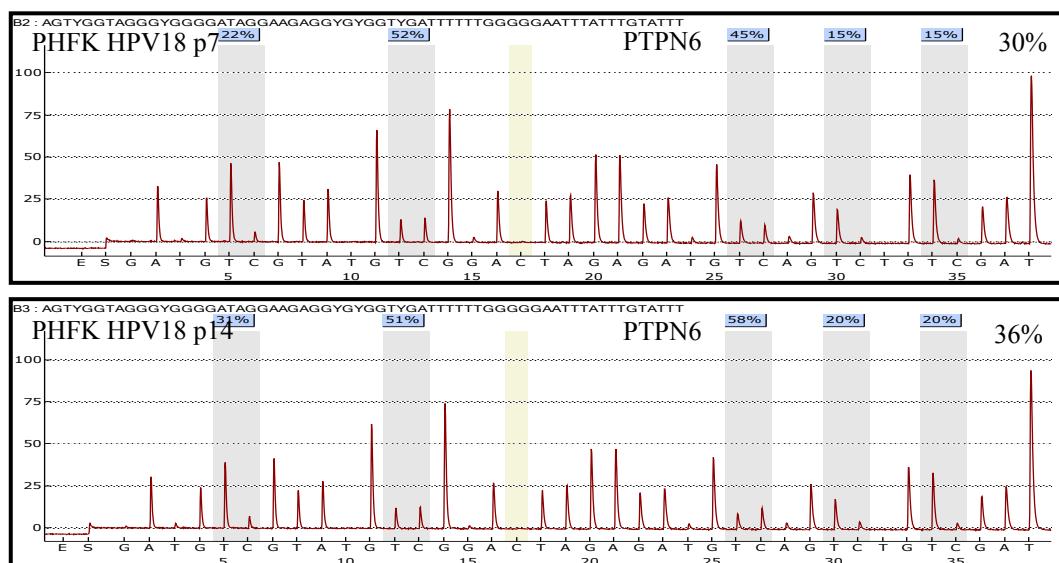
Pyrograms show CpGs methylation of the selected genes used for the validation of illumina HD 450 K methylation array. For each gene there are two pyrograms, the first for PHFK HPV18 p7 and the second for p14.





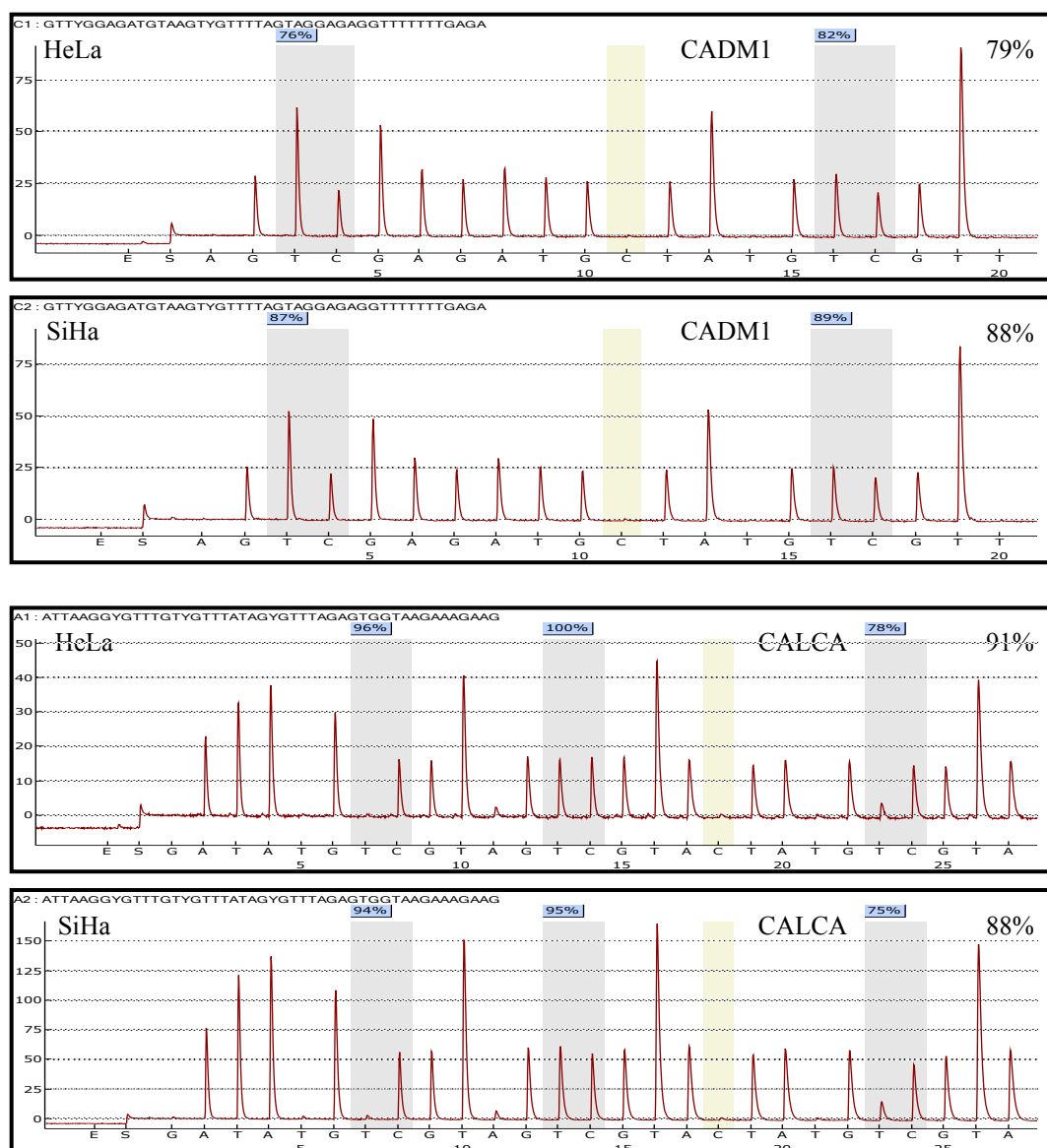


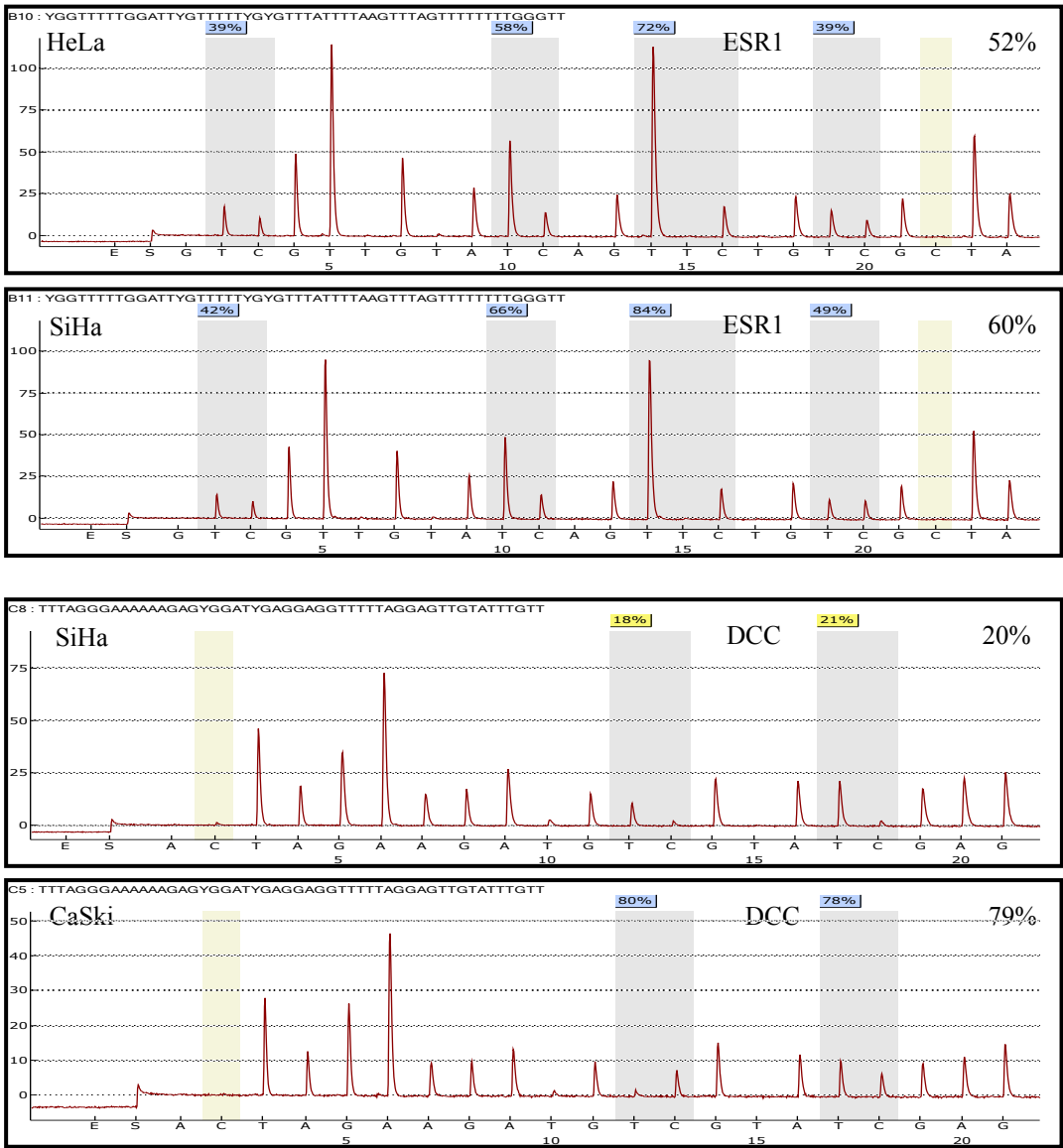




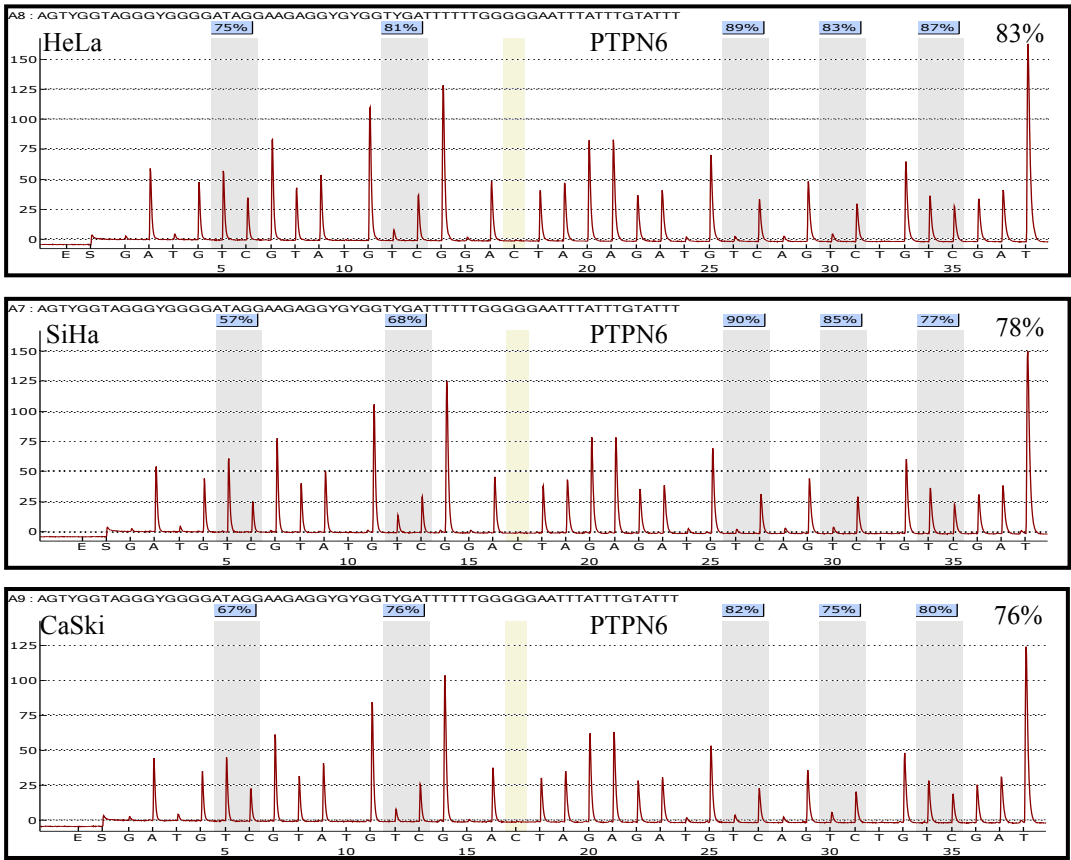
## 7.2 Appendix II

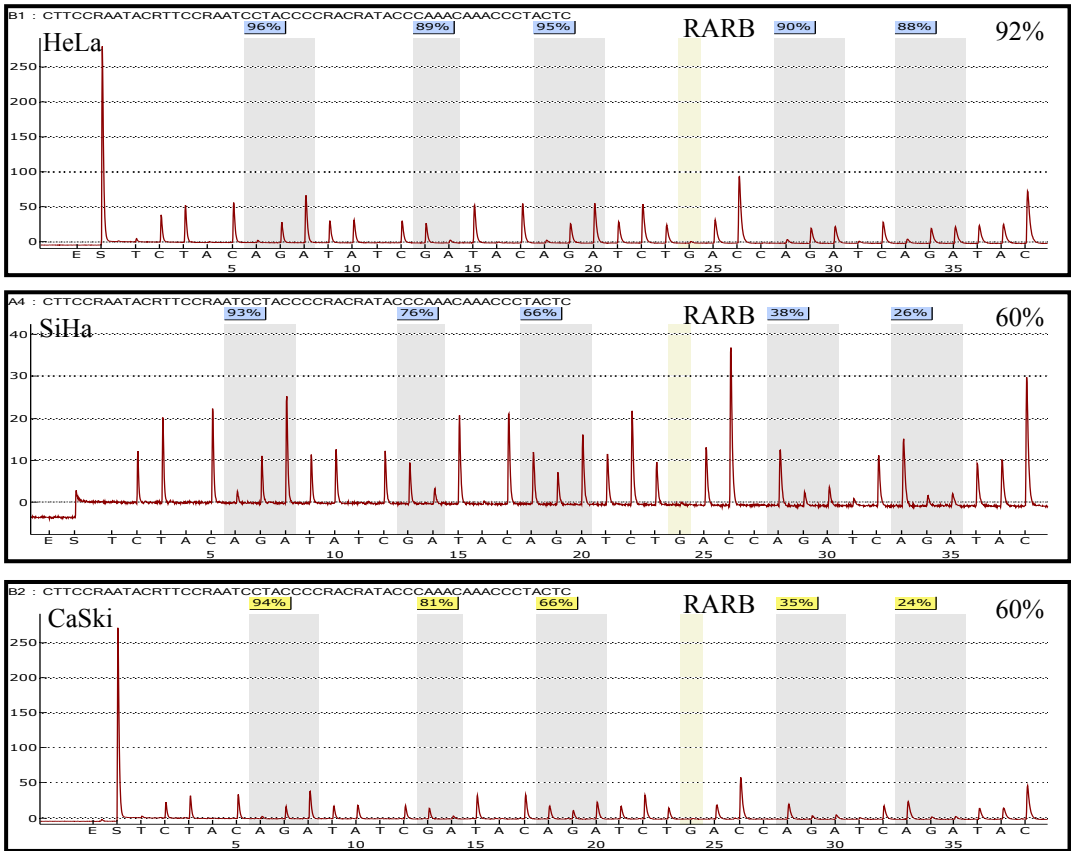
Pyrograms show CpGs methylation of the selected genes used for the external validation of illumina HD 450 K methylation array. This validation involved three cervical cancer cell lines, HeLa, SiHa and Ca Ski.

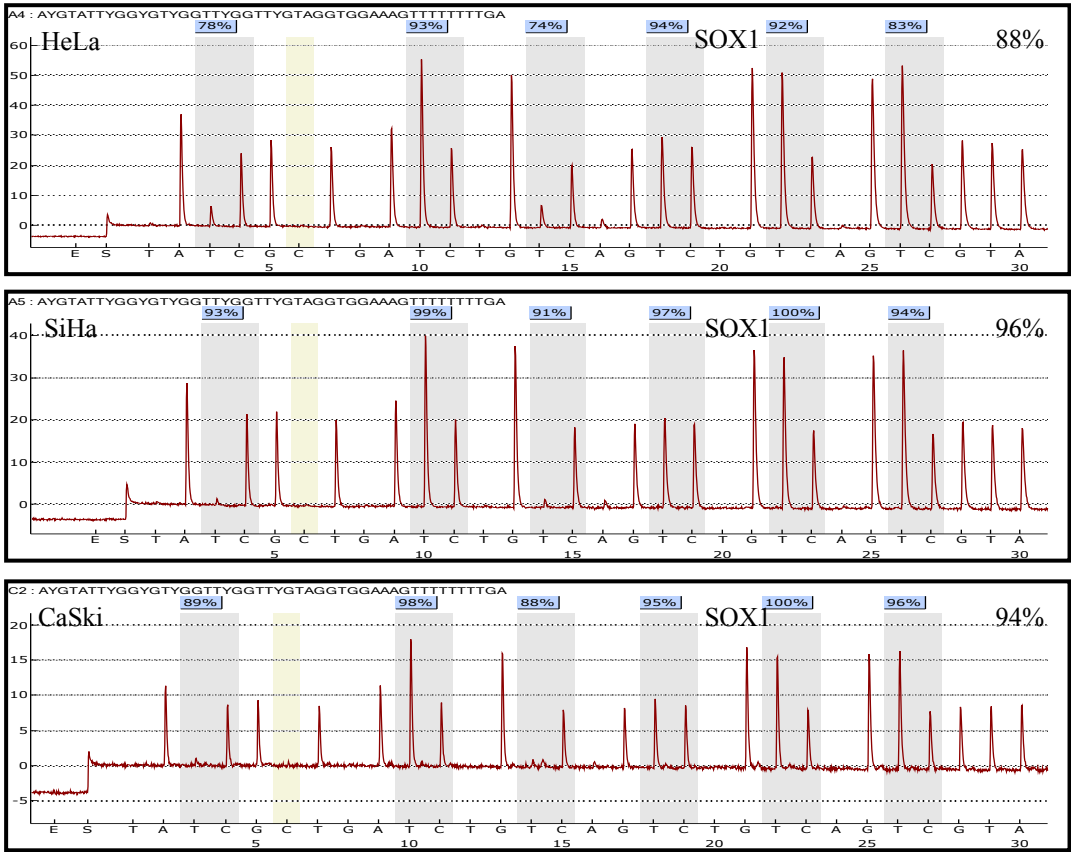


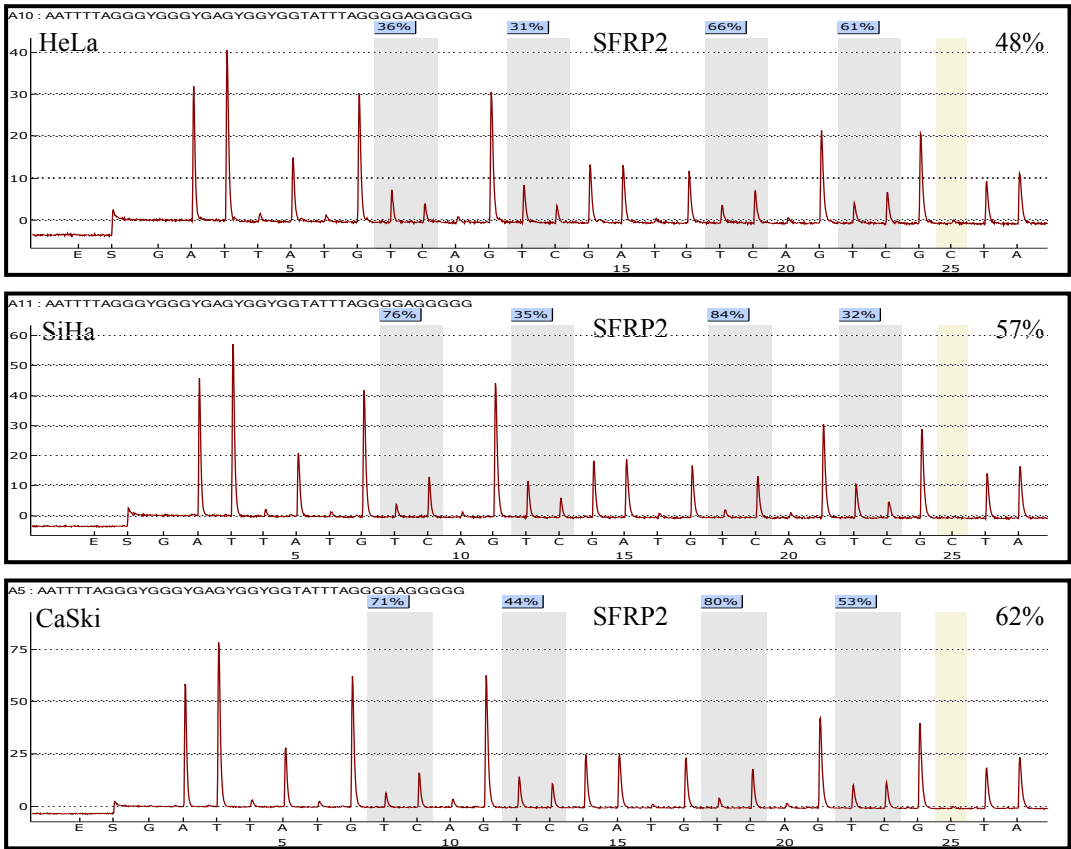


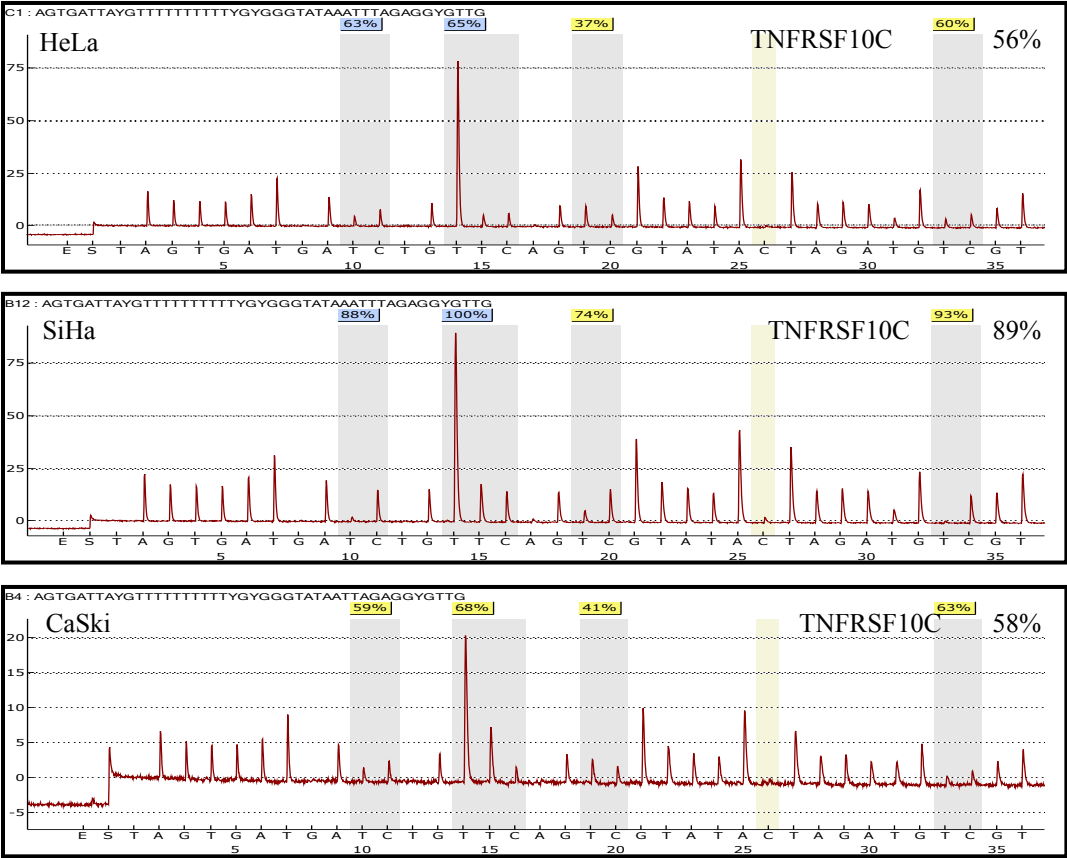


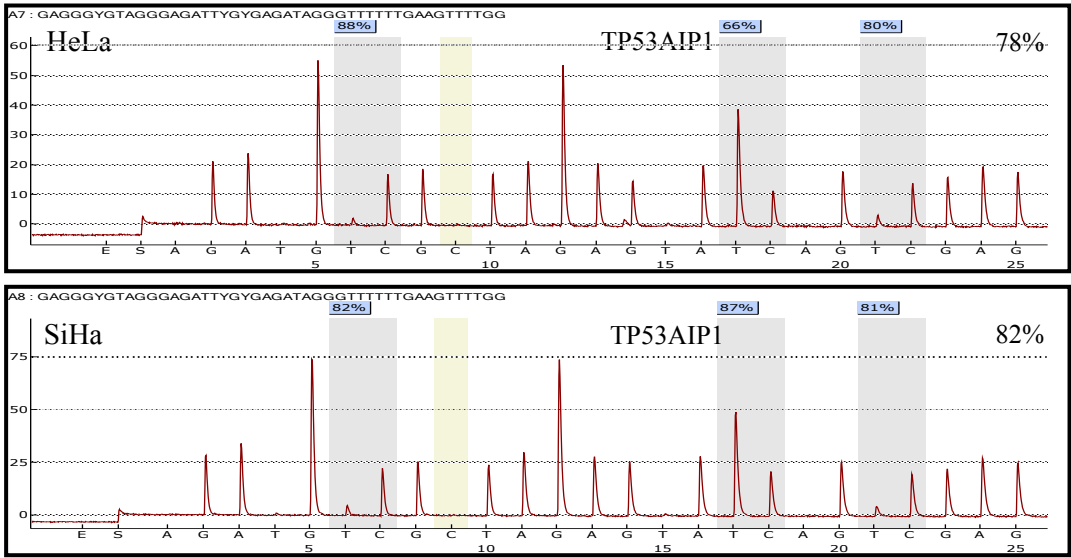






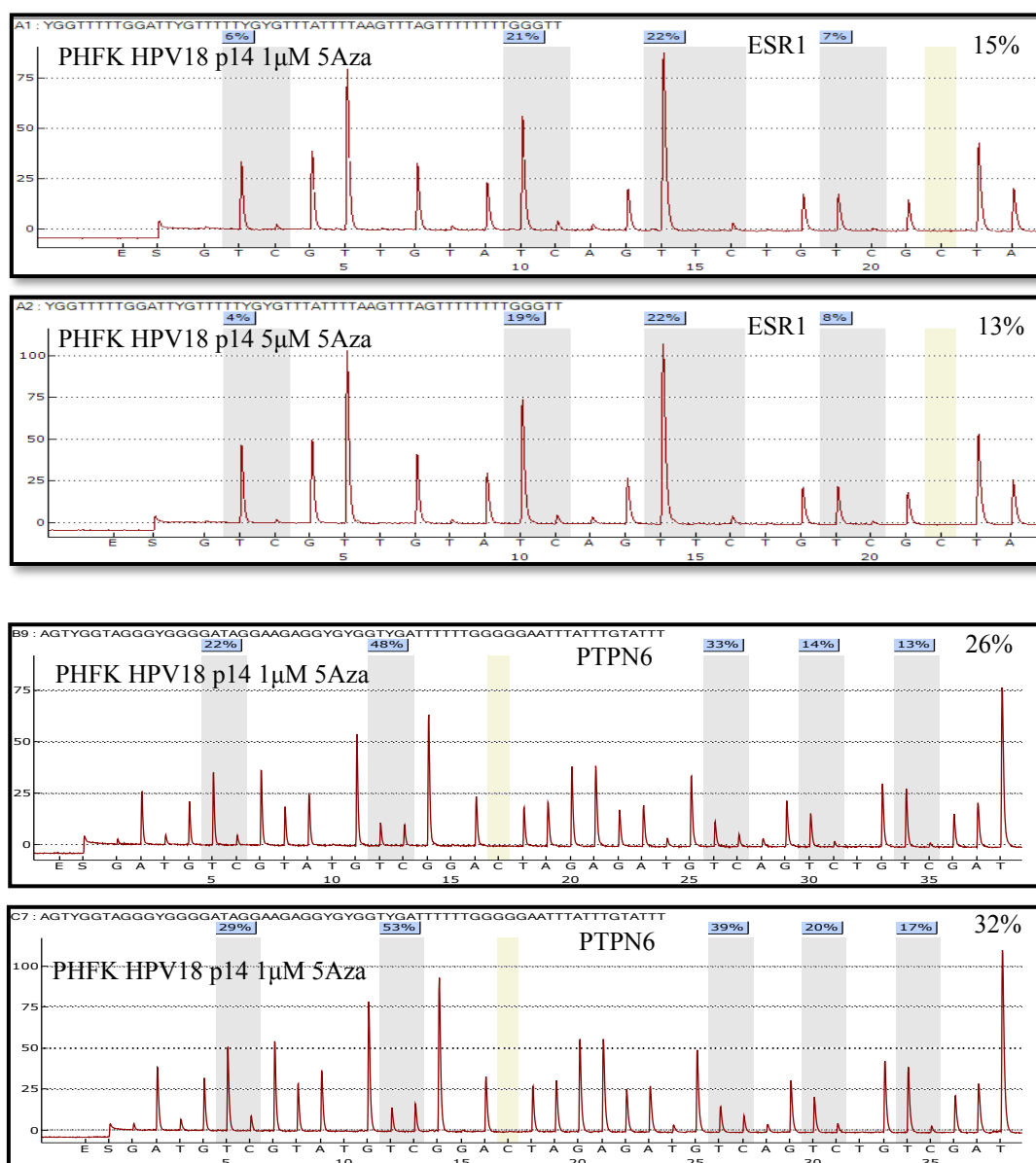


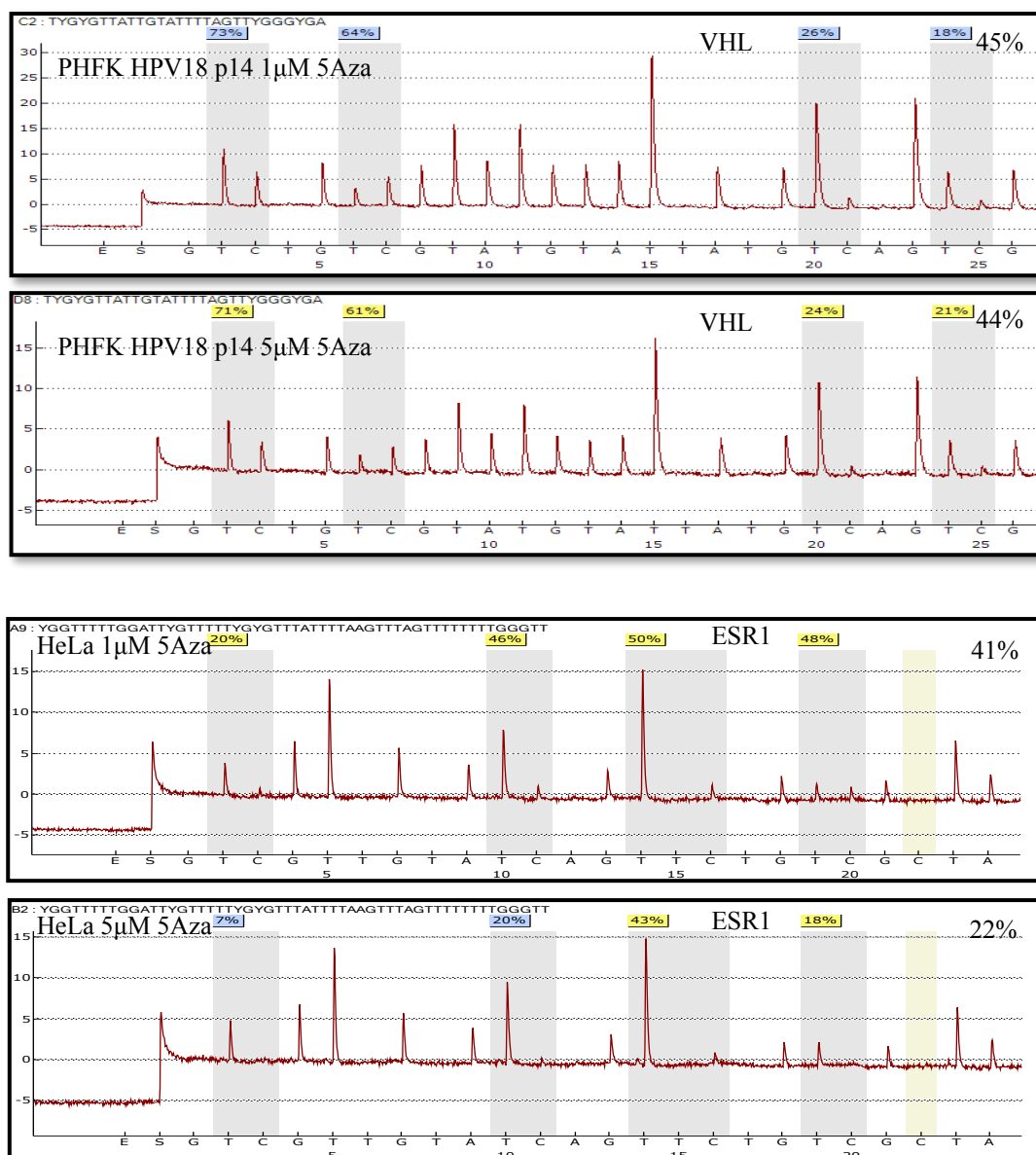




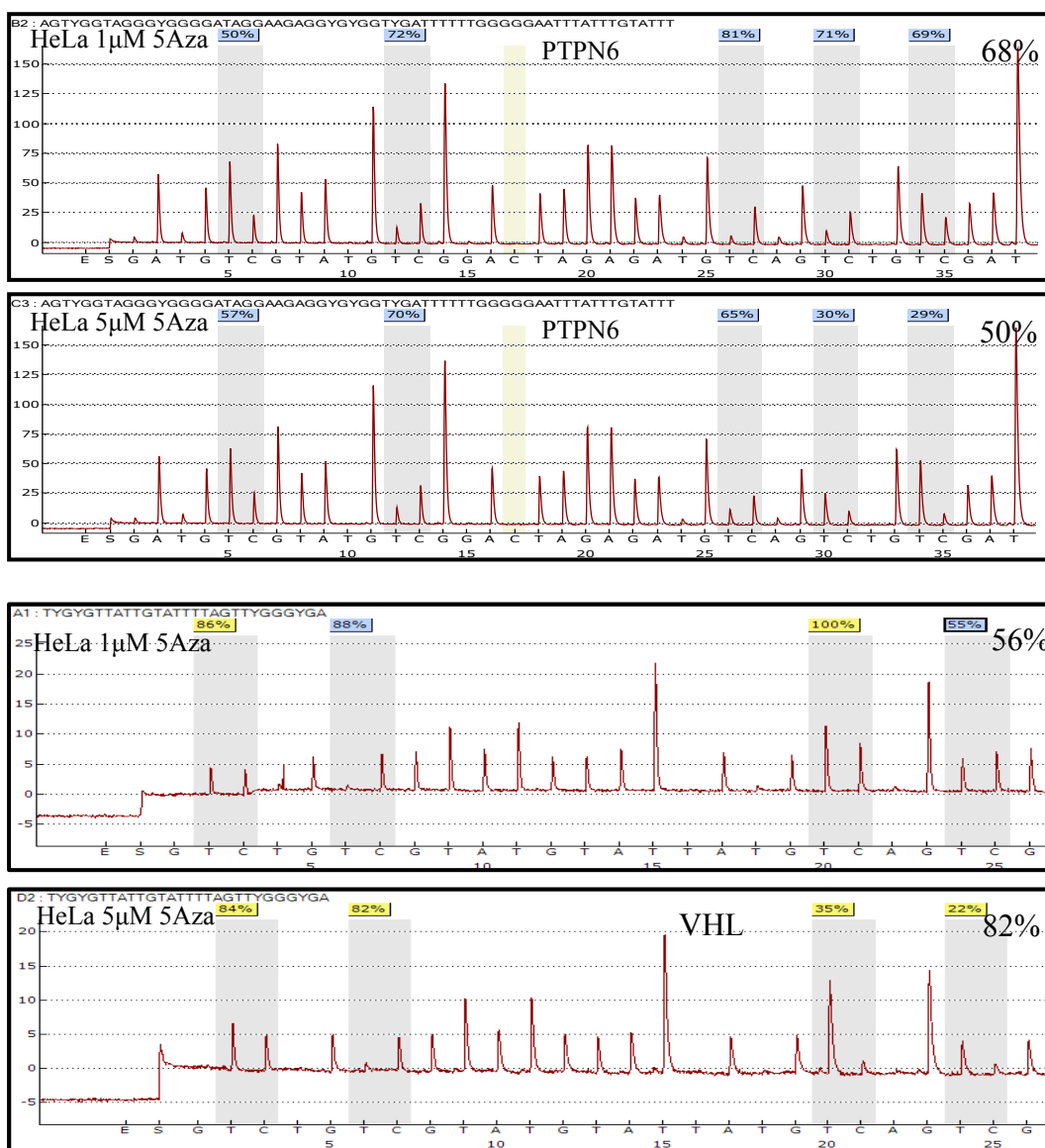
## 7.3 Appendix III

Pyrograms show CpGs methylation of three genes ESR1, PTPN6 and VHL in PHFK HPV18 p14 and HeLa cell lines following treatment with two doses of 5Aza. Results are typical of three independent experiments.









## REFERENCES

## 8 References

- Aapola, U., Kawasaki, K., Scott, H. S., Ollila, J., Vihinen, M., Heino, M., Shintani, A., Minoshima, S., Krohn, K., Antonarakis, S. E., et al. (2000). Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family. *Genomics*, 65 (3), pp.293–298.
- Abele, R., Clavel, M., Dodion, P., Brunsch, U., Gundersen, S., Smyth, J., Renard, J., Van Glabbeke, M. and Pinedo, H. M. (1987). The EORTC early clinical trials cooperative group experience with 5-aza-2'-deoxycytidine (NSC 127716) in patients with colo-rectal, head and neck, renal carcinomas and malignant melanomas. *European Journal of Cancer and Clinical Oncology*, 23 (12), pp.1921–1924.
- Adey, A., Burton, J. N., Kitzman, J. O., Hiatt, J. B., Lewis, A. P., Martin, B. K., Qiu, R., Lee, C. and Shendure, J. (2013). The haplotype-resolved genome and epigenome of the aneuploid HeLa cancer cell line. *Nature*, 500 (7461), pp.207–211.
- Agger, K., Cloos, P. a C., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A. E. and Helin, K. (2007). UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature*, 449 (7163), pp.731–734.
- Aksoy P, Gottschalk EY, Meneses PI (2017). Hpv entry into cells. *Mutat Res*, 772, pp.13–22.
- Alazawi, W., Pett, M., Arch, B., Scott, L., Freeman, T., Stanley, M. a and Coleman, N. (2002). Changes in cervical keratinocyte gene expression associated with integration of human papillomavirus 16. *Cancer research*, 62 (23), pp.6959–6965.
- Apostolidou, S., Hadwin, R., Burnell, M., Jones, A., Baff, D., Pyndiah, N., Mould, T., Jacobs, I. J., Beddows, S., Kocjan, G., et al. (2009). DNA methylation analysis in liquid-based cytology for cervical cancer screening. *International journal of cancer*, 125 (12), pp.2995–3002.
- Appleby, P., Beral, V., De González, A. B., Colin, D., Franceschi, S., Green, J., La Vecchia, C., Peto, J., Plummer, M., Sweetland, S., et al. (2009). Cervical carcinoma and sexual behavior: Collaborative reanalysis of individual data on 15,461 women with cervical carcinoma and 29,164 women without cervical carcinoma from 21 epidemiological studies. *Cancer Epidemiology Biomarkers and Prevention*, 18 (4), pp.1060–1069.
- Arends, M. J., Buckley, C. H. and Wells, M. (1998). Aetiology, pathogenesis, and pathology of cervical neoplasia. *Journal of clinical pathology*, 51, pp.96–103.
- Bachmann, I. M., Halvorsen, O. J., Collett, K., Stefansson, I. M., Straume, O., Haukaas, S. A., Salvesen, H. B., Otte, A. P. and Akslen, L. A. (2006). EZH2 expression is associated with high proliferation rate and aggressive tumor subgroups in cutaneous melanoma and cancers of the endometrium, prostate, and breast. *Journal of Clinical Oncology*, 24 (2), pp.268–273.
- Badal, V., Chuang, L. S. H., Tan, E. H.-H., Badal, S., Villa, L. L., Wheeler, C. M., Li, B. F. L. and Bernard, H.-U. (2003). CpG methylation of human papillomavirus type 16 DNA in cervical

- cancer cell lines and in clinical specimens: genomic hypomethylation correlates with carcinogenic progression. *Journal of virology*, 77 (11), pp.6227–6234.
- Baker, C. C., Phelps, W. C., Lindgren, V., Braun, M. J., Gonda, M. a and Howley, P. M. (1987). Structural and transcriptional analysis of human papillomavirus type 16 sequences in cervical carcinoma cell lines. *Journal of virology*, 61 (4), pp.962–971.
- Balkwill, F. (2004). Cancer and the chemokine network. *Nature reviews. Cancer*, 4 (7), pp.540–550.
- Bannister, A. J. and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Nature Publishing Group*, 21, pp.381–395.
- Banno, K., Yanokura, M., Kawaguchi, M., Kuwabara, Y., Akiyoshi, J., Kobayashi, Y., Iwata, T., Hirasawa, A., Fujii, T., Susumu, N., et al. (2007). Epigenetic inactivation of the CHFR gene in cervical cancer contributes to sensitivity to taxanes. *International journal of oncology*, 31 (4), pp.713–720.
- Bappa, L. A. and Yakasai, I. A. (2013). Colposcopy: the scientific basis. *Annals of African medicine*, 12 (2), Medknow Publications and Media Pvt. Ltd., pp.86–89.
- Barrès, R., Osler, M. E., Yan, J., Rune, A., Fritz, T., Caidahl, K., Krook, A. and Zierath, J. R. (2009). Non-CpG methylation of the PGC-1alpha promoter through DNMT3B controls mitochondrial density. *Cell metabolism*, 10 (3), pp.189–198.
- Barrow-Laing, L., Chen, W. and Roman, A. (2010). Low- and high-risk human papillomavirus E7 proteins regulate p130 differently. *Virology*, 400 (2), pp.233–239.
- Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell*, 129 (4), pp.823–837.
- Baubec, T., Ivánek, R., Lienert, F. & Schübeler, D (2013). Methylation-dependent and -independent genomic targeting principles of the MBD protein family. *Cell*, 153 (2), 480–492.
- Bechtold, V., Beard, P. and Raj, K. (2003). Human papillomavirus type 16 E2 protein has no effect on transcription from episomal viral DNA. *Journal of virology*, 77 (3), pp.2021–2028.
- Berens, C. and Hillen, W. (2003). Gene regulation by tetracyclines. Constraints of resistance regulation in bacteria shape TetR for application in eukaryotes. *European Journal of Biochemistry*, 270 (15), pp.3109–3121.
- Bergant Marušič, M., Ozbun, M. A., Campos, S. K., Myers, M. P. and Banks, L. (2012). Human papillomavirus L2 facilitates viral escape from late endosomes via sorting nexin 17. *Traffic (Copenhagen, Denmark)*, 13 (3), pp.455–467.
- Bernard, H., Burk, R. D., Chen, Z., Doorslaer, K. Van and Villiers, E. De. (2010). Classification of papillomaviruses ( PVs ) based on 189 PV types and proposal of taxonomic amendments. *Virology*, 401 (1), pp.70–79.

- Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*, 125 (2), pp.315–326.
- Berrington De González, A. and Green, J. (2007). Comparison of risk factors for invasive squamous cell carcinoma and adenocarcinoma of the cervix: Collaborative reanalysis of individual data on 8,097 women with squamous cell carcinoma and 1,374 women with adenocarcinoma from 12 epidemiological studies. *International Journal of Cancer*, 120 (4), pp.885–891.
- Bestor, T. H. (2000). The DNA methyltransferases of mammals. *Human molecular genetics*, 9 (16), pp.2395–2402.
- Bestor, T. H. (2005). Transposons reanimated in mice. *Cell*, 122 (3), pp.322–325.
- Beyrouthy, M.J., Garner, K.M., Hever, M.P., et al. (2009) High DNA methyltransferase 3B expression mediates 5-aza-deoxycytidine hypersensitivity in testicular germ cell tumors. *Cancer research*, 69 (24): 9360–6
- Bhattacharjee, B. and Sengupta, S. (2006). CpG methylation of HPV 16 LCR at E2 binding site proximal to P97 is associated with cervical cancer in presence of intact E2. *Virology*, 354 (2), pp.280–285.
- De Bonis, M. L., Cerase, A., Matarazzo, M. R., Ferraro, M., Strazzullo, M., Hansen, R. S., Chiurazzi, P., Neri, G. and D’Esposito, M. (2006). Maintenance of X- and Y-inactivation of the pseudoautosomal (PAR2) gene SPRY3 is independent from DNA methylation and associated to multiple layers of epigenetic modifications. *Human molecular genetics*, 15 (7), pp.1123–1132.
- Bosch, F. X., Burchell, A. N., Schiffman, M., Giuliano, A. R., de Sanjose, S., Bruni, L., Tortolero-Luna, G., Kjaer, S. K. and Muñoz, N. (2008). Epidemiology and natural history of human papillomavirus infections and type-specific implications in cervical neoplasia. *Vaccine*, 26 Suppl 1, pp.K1-16.
- Bosch, F. X., Lorincz, a, Muñoz, N., Meijer, C. J. L. M. and Shah, K. V. (2002). The causal relation between human papillomavirus and cervical cancer. *Journal of clinical pathology*, 55 (4), pp.244–265.
- Bouchard, J. (1989). Mechanism of action of 5-AZA-dC: Induced DNA hypomethylation does not lead to aberrant gene expression in human leukemic CEM cells. *Leukemia Research*, 13 (8), pp.715–722.
- Brown, D. R., Kitchin, D., Qadadri, B., Neptune, N., Batteiger, T. and Ermel, A. (2006). The human papillomavirus type 11 E1--E4 protein is a transglutaminase 3 substrate and induces abnormalities of the cornified cell envelope. *Virology*, 345 (1), pp.290–298.
- Buck, C. B., Cheng, N., Thompson, C. D., Lowy, D. R., Steven, A. C., Schiller, J. T. and Trus, B. L. (2008). Arrangement of L2 within the papillomavirus capsid. *Journal of virology*, 82 (11), pp.5190–5197.

- Burgers, W. a, Blanchon, L., Pradhan, S., de Launoit, Y., Kouzarides, T. and Fuks, F. (2007). Viral oncoproteins target the DNA methyltransferases. *Oncogene*, 26 (11), pp.1650–1655.
- Caberg, J.-H.D., Hubert, P.M., Begon, D.Y., et al. (2008) Silencing of E7 oncogene restores functional E-cadherin expression in human papillomavirus 16-transformed keratinocytes. *Carcinogenesis*, 29 (7): 1441–7
- Candelaria, M., Gallardo-Rincón, D., Arce, C., Cetina, L., Aguilar-Ponce, J. L., Arrieta, O., González-Fierro, a, Chávez-Blanco, a, de la Cruz-Hernández, E., Camargo, M. F., et al. (2007). A phase II study of epigenetic therapy with hydralazine and magnesium valproate to overcome chemotherapy resistance in refractory solid tumors. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*, 18 (9), pp.1529–1538.
- Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S. and Zhang, Y. (2002). Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. *Science*, 298 (5595), pp.1039–1043.
- Cao, R. and Zhang, Y. (2004). The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Current opinion in genetics & development*, 14 (2), pp.155–164.
- De Capoa, A., Musolino, A., Della Rosa, S., Caiafa, P., Mariani, L., Del Nonno, F., Vocaturo, A., Perrone Donnorso, R., Niveleau, A. and Grappelli, C. (2003). DNA demethylation is directly related to tumour progression: Evidence in normal, pre-malignant and malignant cells from uterine cervix samples. *Oncology Reports*, 10 (3), Spandidos Publications., pp.545–549.
- Cedar, H. and Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. *Nature reviews. Genetics*, 10 (5), pp.295–304.
- Cerqueira C, Schiller JT (2017). Papillomavirus assembly: an overview and perspectives. *Virus Res*, 23, pp.103–107.
- Cervical, N. H. S. and Programme, S. (2016). Achievable standards , Benchmarks for reporting , and Criteria for evaluating third edition including revised performance indicators Achievable standards , Benchmarks for reporting , and Criteria for evaluating cervical cytopathology Third edition.
- Chan, A. T. C., Tao, Q., Robertson, K. D., Flinn, I. W., Mann, R. B., Klencke, B., Kwan, W. H., Leung, T. W.-T., Johnson, P. J. and Ambinder, R. F. (2004). Azacitidine induces demethylation of the Epstein-Barr virus genome in tumors. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 22 (8), J Clin Oncol., pp.1373–1381.
- Chapman, M. D., Keir, G., Petzold, A. and Thompson, E. J. (2006). Measurement of high affinity antibodies on antigen-immunoblots. *Journal of immunological methods*, 310 (1–2), pp.62–66.
- Chedin, F., Lieber, M. R. and Hsieh, C.-L. (2002). The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proceedings of the National Academy of Sciences of the United States of America*, 99 (26), pp.16916–16921.

- Chen, C., Pan, D., Deng, A.-M., Huang, F., Sun, B.-L. and Yang, R.-G. (2013). DNA methyltransferases 1 and 3B are required for hepatitis C virus infection in cell culture. *Virology*, 441 (1), pp.57–65.
- Chen, L., MacMillan, A. M., Chang, W., Ezaz-Nikpay, K., Lane, W. S. and Verdine, G. L. (1991). Direct identification of the active-site nucleophile in a DNA (cytosine-5)-methyltransferase. *Biochemistry*, 30 (46), American Chemical Society., pp.11018–11025.
- Chen, T., Hevi, S., Gay, F., Tsujimoto, N., He, T., Zhang, B., Ueda, Y. and Li, E. (2007). Complete inactivation of DNMT1 leads to mitotic catastrophe in human cancer cells. *Nature genetics*, 39 (3), pp.391–396.
- Chen, T., Ueda, Y., Dodge, J. E., Wang, Z. and Li, E. (2003). Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Molecular and cellular biology*, 23 (16), pp.5594–5605.
- Cho, N.H., Kang, S., Hong, S., et al. (2005) Multinucleation of koilocytes is in fact multilobation and is related to aberration of the G2 checkpoint. *Journal of clinical pathology*, 58 (6): 576–82
- Choi, H., Park, J., Park, M., Won, H., Joo, H., Lee, C. H., Lee, J. and Kong, G. (2015). UTX inhibits EMT -induced breast CSC properties by epigenetic repression of EMT genes in cooperation with LSD 1 and HDAC 1 . *EMBO reports*, 16 (10), pp.1288–1298.
- Chung, M.-T., Sytwu, H.-K., Yan, M.-D., Shih, Y.-L., Chang, C.-C., Yu, M.-H., Chu, T.-Y., Lai, H.-C. and Lin, Y.-W. (2009). Promoter methylation of SFRPs gene family in cervical cancer. *Gynecologic oncology*, 112 (2), pp.301–306.
- Colacino, J. a, Dolinoy, D.C., Duffy, S. a, et al. (2013) Comprehensive analysis of DNA methylation in head and neck squamous cell carcinoma indicates differences by survival and clinicopathologic characteristics. *PloS one*, 8 (1): e54742
- Collins, S. I., Constandinou-Williams, C., Wen, K., Young, L. S., Roberts, S., Murray, P. G. and Woodman, C. B. J. (2009). Disruption of the E2 gene is a common and early event in the natural history of cervical human papillomavirus infection: a longitudinal cohort study. *Cancer research*, 69 (9), pp.3828–3832.
- Cortés-Gutiérrez, E.I., Dávila-Rodríguez, M.I., Fernández, J.L., et al. (2010) Koilocytes are enriched for alkaline-labile sites. *European Journal of Histochemistry*, 54 (4): 189–192
- Costello, J. F., Frühwald, M. C., Smiraglia, D. J., Rush, L. J., Robertson, G. P., Gao, X., Wright, F. A., Feramisco, J. D., Peltomäki, P., Lang, J. C., et al. (2000). Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nature genetics*, 24 (2), pp.132–138.
- Creusot, F., Acs, G. and Christman, J. K. (1982). Inhibition of DNA methyltransferase and induction of Friend erythroleukemia cell differentiation by 5-azacytidine and 5-aza-2'-deoxycytidine. *The Journal of biological chemistry*, 257 (4), pp.2041–2048.
- Cross, S. H., Charlton, J. A., Nan, X. and Bird, A. P. (1994). Purification of CpG islands using a methylated DNA binding column. *Nature Genetics*, 6 (3), pp.236–244.

- CRUK, (2014) Cervical cancer statistics. available at: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/cervical-cancer>
- Crusius, K., Auvinen, E. and Alonso, a. (1997). Enhancement of EGF- and PMA-mediated MAP kinase activation in cells expressing the human papillomavirus type 16 E5 protein. *Oncogene*, 15 (12), pp.1437–1444.
- Cuzick, J., Szarewski, A., Cubie, H., Hulman, G., Kitchener, H., Luesley, D., McGoogan, E., Menon, U., Terry, G., Edwards, R., et al. (2003). Management of women who test positive for high-risk types of human papillomavirus: the HART study. *Lancet*, 362 (9399), pp.1871–1876.
- Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A. and Pirrotta, V. (2002). Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell*, 111 (2), pp.185–196.
- Daling, J. R., Madeleine, M. M., Johnson, L. G., Schwartz, S. M., Shera, K. a, Wurscher, M. a, Carter, J. J., Porter, P. L., Galloway, D. a and McDougall, J. K. (2004). Human papillomavirus, smoking, and sexual practices in the etiology of anal cancer. *Cancer*, 101 (2), pp.270–280.
- Daling, J. R., Madeleine, M. M., Schwartz, S. M., Shera, K. a, Carter, J. J., McKnight, B., Porter, P. L., Galloway, D. a, McDougall, J. K. and Tamimi, H. (2002). A population-based study of squamous cell vaginal cancer: HPV and cofactors. *Gynecologic oncology*, 84 (2), pp.263–270.
- Dallol, A., Fernandes, N., Silva, D., Viacava, P., Minna, J. D., Bieche, I. and Maher, E. R. (2002a). SLIT2 , a Human Homologue of the Drosophila Slit2 Gene , Has Tumor Suppressor Activity and Is Frequently Inactivated in Lung and Breast Cancers SLIT2 , a Human Homologue of the Drosophila Slit2 Gene , Has Tumor Suppressor. pp.5874–5880.
- Dallol, A., Forgacs, E., Martinez, A., Sekido, Y., Walker, R., Kishida, T., Rabbitts, P., Maher, E. R., Minna, J. D. and Latif, F. (2002b). *Tumour speci ® c promoter region methylation of the human homologue of the Drosophila Roundabout gene DUTT1 ( ROBO1 ) in human cancers*. pp.3020–3028.
- Daniel, B., Mukherjee, G., Seshadri, L., Vallikad, E. and Krishna, S. (1995). Changes in the physical state and expression of human papillomavirus type 16 in the progression of cervical intraepithelial neoplasia lesions analysed by PCR. *Journal of General Virology*, 76 (10), pp.2589–2593.
- Danos, O., Katinka, M. and Yaniv, M. (1982). Human papillomavirus 1a complete DNA sequence: a novel type of genome organization among papovaviridae. *The EMBO journal*, 1 (2), pp.231–236.
- Dao, L. D., Duffy, A., Van Tine, B. A., Wu, S.-Y., Chiang, C.-M., Broker, T. R. and Chow, L. T. (2006). Dynamic localization of the human papillomavirus type 11 origin binding protein E2 through mitosis while in association with the spindle apparatus. *Journal of virology*, 80 (10), pp.4792–4800.



- Das, B. C., Sharma, J. K., Gopalakrishna, V. and Luthra, U. K. (1992). Analysis by polymerase chain reaction of the physical state of human papillomavirus type 16 DNA in cervical preneoplastic and neoplastic lesions. *The Journal of general virology*, 73 ( Pt 9), pp.2327–2336.
- Day, P. M., Baker, C. C., Lowy, D. R. and Schiller, J. T. (2004). Establishment of papillomavirus infection is enhanced by promyelocytic leukemia protein (PML) expression. *Proceedings of the National Academy of Sciences of the United States of America*, 101 (39), pp.14252–14257.
- Day, P. M., Roden, R. B., Lowy, D. R. and Schiller, J. T. (1998). The papillomavirus minor capsid protein, L2, induces localization of the major capsid protein, L1, and the viral transcription/replication protein, E2, to PML oncogenic domains. *Journal of virology*, 72 (1), pp.142–150.
- Day, S. P., Hudson, A., Mast, A., Sander, T., Curtis, M., Olson, S., Chehak, L., Quigley, N., Ledford, J., Yen-Lieberman, B., et al. (2009). Analytical performance of the Investigational Use Only Cervista HPV HR test as determined by a multi-center study. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology*, 45 Suppl 1, pp.S63-72.
- Dennis, G., Sherman, B. T., Hosack, D. a, Yang, J., Gao, W., Lane, H. C. and Lempicki, R. a. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome biology*, 4 (5), p.P3.
- Deplus R, Brenner C, Burgers WA, Putmans P, Kouzarides, T. et al (2002).. Dnmt3L is a transcriptional repressor that recruits histone deacetylase. *Nucleic Acids Res*, 30(17), 3831–3838.
- Dickinson, R. E. and Duncan, W. C. (2010).The SLIT / ROBO pathway : a regulator of cell function with implications for the reproductive system. *Reproduction* 139 (4), pp.697–704.
- Dillon, S. C., Zhang, X., Trievel, R. C. and Cheng, X. (2005). The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biology*, 6 (8), p.227.
- Dockter, J., Schroder, A., Hill, C., Guzinski, L., Monson, J. and Giachetti, C. (2009). Clinical performance of the APTIMA HPV Assay for the detection of high-risk HPV and high-grade cervical lesions. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology*, 45 Suppl 1, pp.S55-61.
- Doorbar, J. (2006). Molecular biology of human papillomavirus infection and cervical cancer. *Clinical science (London, England : 1979)*, 110 (5), pp.525–541.
- Doorbar, J., Quint, W., Banks, L., Bravo, I. G., Stoler, M., Broker, T. R. and Stanley, M. a. (2012). The biology and life-cycle of human papillomaviruses. *Vaccine*, 30 Suppl 5, Elsevier Ltd., pp.F55-70.
- Doorbar J, Egawa N, Griffin H, Kranjec C, Murakami I. (2015). Human papillomavirus molecular biology and disease association. *Rev Med Virol*, 25.
- Dowhanick, J. J., McBride, A. A. and Howley, P. M. (1995). Suppression of cellular proliferation by the papillomavirus E2 protein. *Journal of virology*, 69 (12), pp.7791–7799.

- Drexler, H. G. and Uphoff, C. C. (2002). Mycoplasma contamination of cell cultures: Incidence, sources, effects, detection, elimination, prevention. *Cytotechnology*, 39 (2), pp.75–90. [Online]. Available at: doi:10.1023/A:1022913015916.
- Dürst, M., Glitz, D., Schneider, A. and zur Hausen, H. (1992). Human papillomavirus type 16 (HPV 16) gene expression and DNA replication in cervical neoplasia: analysis by in situ hybridization. *Virology*, 189 (1), pp.132–140.
- Egger, G., Liang, G., Aparicio, A. and Jones, P. a. (2004). Epigenetics in human disease and prospects for epigenetic therapy. *Nature*, 429 (6990), pp.457–463.
- Ehrich, M., Turner, J., Gibbs, P., Lipton, L., Giovanneti, M., Cantor, C. and van den Boom, D. (2008). Cytosine methylation profiling of cancer cell lines. *Proceedings of the National Academy of Sciences of the United States of America*, 105 (12), pp.4844–4849.
- Ehrlich, M., Gama-Sosa, M. A., Huang, L. H., Midgett, R. M., Kuo, K. C., McCune, R. A. and Gehrke, C. (1982). Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic acids research*, 10 (8), pp.2709–2721.
- Elmore, L. W., Hancock, A. R., Chang, S. F., Wang, X. W., Chang, S., Callahan, C. P., Geller, D. A., Will, H. and Harris, C. C. (1997). Hepatitis B virus X protein and p53 tumor suppressor interactions in the modulation of apoptosis. *Proceedings of the National Academy of Sciences of the United States of America*, 94 (26), pp.14707–14712.
- Van Emburgh, B. O. and Robertson, K. D. (2011). Modulation of Dnmt3b function in vitro by interactions with Dnmt3L, Dnmt3a and Dnmt3b splice variants. *Nucleic acids research*, 39 (12), pp.4984–5002.
- Esteller, M. (2008). Epigenetics in cancer. *The New England journal of medicine*, 358 (11), pp.1148–1159.
- Esteller, M., Simó-Riudalbas L, Melo SA, (2011). DNMT3B Gene Amplification Predicts Resistance To DNA Demethylating Drugs. *Genes, Chromosomes & Cancer* 50(7):527-534.
- Estève, PO., Chang, Y., Samaranayake, M., Upadhyay, A. K., Horton, J. R., Feehery, G. R., Cheng, X. and Pradhan, S. (2011). A methylation and phosphorylation switch between an adjacent lysine and serine determines human DNMT1 stability. *Nature structural & molecular biology*, 18 (1), pp.42–48.
- Estève PO, Chin HG, Smallwood A, Feehery GR, Gangisetty O, Karpf AR, Carey MF, Pradhan S (2006). Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. *Genes Dev.* 20(22), 3089–3103.
- Evander, M., Frazer, I. H., Payne, E., Qi, Y. M., Hengst, K. and McMillan, N. A. (1997). Identification of the alpha6 integrin as a candidate receptor for papillomaviruses. *Journal of virology*, 71 (3), pp.2449–2456.
- Ezhkova, E., Lien, W., Stokes, N., Pasolli, H. A., Silva, J. M. and Fuchs, E. (2011). *trimethylation and are essential for hair follicle homeostasis and wound repair.* pp.485–498.

- Fehrmann, F. and Laimins, L. A. (2003). Human papillomaviruses: targeting differentiating epithelial cells for malignant transformation. *Oncogene*, 22 (33), Nature Publishing Group., pp.5201–5207.
- Feinberg, A. P. and Vogelstein, B. (1983). Hypomethylation of ras oncogenes in primary human cancers. *Biochemical and biophysical research communications*, 111 (1), pp.47–54.
- Feng, Q., Balasubramanian, A., Hawes, S. E., Toure, P., Sow, P. S., Dem, A., Dembele, B., Critchlow, C. W., Xi, L., Lu, H., et al. (2005). Detection of hypermethylated genes in women with and without cervical neoplasia. *Journal of the National Cancer Institute*, 97 (4), pp.273–282.
- Ferber, M. J., Montoya, D. P., Yu, C., Aderca, I., McGee, a, Thorland, E. C., Nagorney, D. M., Gostout, B. S., Burgart, L. J., Boix, L., et al. (2003). Integrations of the hepatitis B virus (HBV) and human papillomavirus (HPV) into the human telomerase reverse transcriptase (hTERT) gene in liver and cervical cancers. *Oncogene*, 22 (24), pp.3813–3820.
- Fernandez, A. F., Rosales, C., Lopez-Nieva, P., Graña, O., Ballestar, E., Ropero, S., Espada, J., Melo, S. a, Lujambio, A., Fraga, M. F., et al. (2009). The dynamic DNA methylomes of double-stranded DNA viruses associated with human cancer. *Genome research*, 19 (3), pp.438–451.
- Fey EG and Penman S. (1986). The morphological oncogenic signature. Reorganization of epithelial cytoarchitecture and metabolic regulation by tumor promoters and by transformation. *Review*, 1986;3:81-100..
- Finnen, R. L., Erickson, K. D., Chen, X. S. and Garcea, R. L. (2003). Interactions between papillomavirus L1 and L2 capsid proteins. *Journal of virology*, 77 (8), pp.4818–4826.
- Fraga, M. F., Ballestar, E., Villar-Garea, A., Boix-Chornet, M., Espada, J., Schotta, G., Bonaldi, T., Haydon, C., Ropero, S., Petrie, K., et al. (2005). Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nature Genetics*, 37 (4), pp.391–400.
- Fraga, M. F., Herranz, M., Espada, J. S., Ballestar, E., Paz, M. F., Ropero, S., Erkek, E., Bozdogan, O., Peinado, H., Niveleau, A., et al. (2004). *A Mouse Skin Multistage Carcinogenesis Model Reflects the Aberrant DNA Methylation Patterns of Human Tumors*.
- Franchina, M. and Kay, P. H. (2000). Evidence that cytosine residues within 5'-CCTGG-3' pentanucleotides can be methylated in human DNA independently of the methylating system that modifies 5'-CG-3' dinucleotides. *DNA and cell biology*, 19 (9), pp.521–526.
- Frattoni, M. G., Lim, H. B. and Laimins, L. a. (1996). In vitro synthesis of oncogenic human papillomaviruses requires episomal genomes for differentiation-dependent late expression. *Proceedings of the National Academy of Sciences of the United States of America*, 93 (7), pp.3062–3067.
- Fu, L., Van Doorslaer, K., Chen, Z., Ristriani, T., Masson, M., Travé, G. and Burk, R. D. (2010). Degradation of p53 by human Alphapapillomavirus E6 proteins shows a stronger correlation with phylogeny than oncogenicity. *PloS one*, 5 (9), pp.1–8.

- Fuks, F., Burgers, W. A., Godin, N., Kasai, M. & Kouzarides, T (2001). Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. *EMBO J*, 20, pp.2536–2544.
- Funk, J. O., Waga, S., Harry, J. B., Espling, E., Stillman, B. and Galloway, D. a. (1997). Inhibition of CDK activity and PCNA-dependent DNA replication by p21 is blocked by interaction with the HPV-16 E7 oncoprotein. *Genes & Development*, 11 (16), pp.2090–2100.
- Gal-Yam, E. N., Egger, G., Iniguez, L., Holster, H., Einarsson, S., Zhang, X., Lin, J. C., Liang, G., Jones, P. a and Tanay, A. (2008). Frequent switching of Polycomb repressive marks and DNA hypermethylation in the PC3 prostate cancer cell line. *Proceedings of the National Academy of Sciences of the United States of America*, 105 (35), pp.12979–12984.
- Gariglio, P., Gutiérrez, J., Cortés, E. and Vázquez, J. (2009). The role of retinoid deficiency and estrogens as cofactors in cervical cancer. *Archives of medical research*, 40 (6), pp.449–465.
- Gaston & Fried (1995). CpG methylation has differential effects on the binding of YY1 and ETS proteins to the bi-directional promoter of the Surf-1 and Surf-2 genes. *Nucleic Acids Res*, 23, pp.901–909.
- De Geest, K., Turyk, M. E., Hosken, M. I., Hudson, J. B., Laimins, L. A. and Wilbanks, G. D. (1993). Growth and Differentiation of Human Papillomavirus Type 31b Positive Human Cervical Cell Lines. *Gynecologic Oncology*, 49 (3), pp.303–310.
- Gillison, M. L., Koch, W. M., Capone, R. B., Spafford, M., Westra, W. H., Wu, L., Zahurak, M. L., Daniel, R. W., Viglione, M., Symer, D. E., et al. (2000). Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *Journal of the National Cancer Institute*, 92 (9), pp.709–720.
- Giarrè, M., Caldeira, S., Malanchi, I., et al. (2001) Induction of pRb degradation by the human papillomavirus type 16 E7 protein is essential to efficiently overcome p16INK4a-imposed G1 cell cycle Arrest. *Journal of Virology*, 75 (10): 4705-12.
- Good CR, Panjarian S, Kelly AD, Madzo J, Patel B, Jelinek J, Issa JJ(2018). (2018)TET1-Mediated Hypomethylation Activates Oncogenic Signaling in Triple-Negative Breast Cancer. *Cancer Res*, 78(15) pp.4126-4137.
- Gloor, E. and Hurlimann, J. (1986). Cervical intraepithelial glandular neoplasia (adenocarcinoma in situ and glandular dysplasia). A correlative study of 23 cases with histologic grading, histochemical analysis of mucins, and immunohistochemical determination of the affinity for four lectin. *Cancer*, 58 (6), pp.1272–1280.
- Gokul, G., Poli, R., Jain, M., Ramakrishna, G. and Khosla, S. (2007). UKPMC Funders Group UKPMC Funders Group Author Manuscript DNA Methylation Profile at the DNMT3L Promoter : *Time*, 2 (2), pp.80–85.

- Gokul, G., Ramakrishna, G. and Khosla, S. (2009). Reprogramming of HeLa cells upon DNMT3L overexpression mimics carcinogenesis. *Epigenetics : official journal of the DNA Methylation Society*, 4 (5), pp.322–329.
- Gowher, H. and Jeltsch, A. (2001). Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG [correction of non-CpA] sites. *Journal of molecular biology*, 309 (5), pp.1201–1208.
- Goyal, R., Reinhardt, R. and Jeltsch, A. (2006). Accuracy of DNA methylation pattern preservation by the Dnmt1 methyltransferase. *Nucleic acids research*, 34 (4), pp.1182–1188.
- Gray, E., Pett, M. R., Ward, D., Winder, D. M., Stanley, M. a, Roberts, I., Scarpini, C. G. and Coleman, N. (2010). In vitro progression of human papillomavirus 16 episome-associated cervical neoplasia displays fundamental similarities to integrant-associated carcinogenesis. *Cancer research*, 70 (10), pp.4081–4091.
- Greenfield, a, Carrel, L., Pennisi, D., et al. (1998) The UTX gene escapes X inactivation in mice and humans. *Human molecular genetics*, 7 (4), 737–42.
- Greenberg, M.V.C., Bourc'his, D (2019). The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol*, 20 (10), 590–607.
- Guan, H., Zu, G., Xie, Y., Tang, H., Johnson, M., Xu, X., Kevil, C., Xiong, W.-C., Elmetts, C., Rao, Y., et al. (2003). Neuronal repellent Slit2 inhibits dendritic cell migration and the development of immune responses. *Journal of immunology (Baltimore, Md. : 1950)*, 171 (12), pp.6519–6526.
- Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R. and Young, R. A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. *Cell*, 130 (1), pp.77–88.
- Gustavsson, E., Sernbo, S., Andersson, E., Brennan, D. J., Dictor, M., Jerkeman, M., Borrebaeck, C. A. and Ek, S. (2010). SOX11 expression correlates to promoter methylation and regulates tumor growth in hematopoietic malignancies. *Molecular cancer*, 9, p.187.
- Hadaschik, D., Hinterkeuser, K., Oldak, M., Pfister, H. J. and Smola-Hess, S. (2003). The Papillomavirus E2 protein binds to and synergizes with C/EBP factors involved in keratinocyte differentiation. *Journal of virology*, 77 (9), pp.5253–5265.
- Hafkamp, H. C., Speel, E. J. M., Haesevoets, A., Bot, F. J., Dinjens, W. N. M., Ramaekers, F. C. S., Hopman, A. H. N. and Manni, J. J. (2003). A subset of head and neck squamous cell carcinomas exhibits integration of HPV 16/18 DNA and overexpression of p16INK4A and p53 in the absence of mutations in p53 exons 5-8. *International journal of cancer*, 107 (3), pp.394–400.
- zur Hausen, H. (2009). Papillomaviruses in the causation of human cancers - a brief historical account. *Virology*, 384 (2), pp.260–265.

- He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, Song CX, Zhang K, He C, Xu GL. (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*. 333 (6047), pp. 1303-1307.
- Heideman, D. a M., Waterboer, T., Pawlita, M., Delis-van Diemen, P., Nindl, I., Leijte, J. a, Bonfrer, J. M. G., Horenblas, S., Meijer, C. J. L. M. and Snijders, P. J. F. (2007). Human papillomavirus-16 is the predominant type etiologically involved in penile squamous cell carcinoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 25 (29), pp.4550–4556.
- Heilmann, V. and Kreienberg, R. (2002). Molecular biology of cervical cancer and its precursors. *Current women's health reports*, 2 (1), pp.27–33.
- Heintzman, N. D., Hon, G. C., Hawkins, R. D., Kheradpour, P., Stark, A., Harp, L. F., Ye, Z., Lee, L. K., Stuart, R. K., Ching, C. W., et al. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*, 459 (7243), pp.108–112.
- Helt, A. and Galloway, D. A. (2001). Destabilization of the Retinoblastoma Tumor Suppressor by Human Papillomavirus Type 16 E7 Is Not Sufficient To Overcome Cell Cycle Arrest in Human Keratinocytes. *J Virol*, 75(15), pp.6737-47.
- Henken, F. E., Wilting, S. M., Overmeer, R. M., van Rietschoten, J. G. I., Nygren, a O. H., Errami, a, Schouten, J. P., Meijer, C. J. L. M., Snijders, P. J. F. and Steenbergen, R. D. M. (2007). Sequential gene promoter methylation during HPV-induced cervical carcinogenesis. *British journal of cancer*, 97 (10), pp.1457–1464.
- Hermann, a, Gowher, H. and Jeltsch, a. (2004). Biochemistry and biology of mammalian DNA methyltransferases. *Cellular and molecular life sciences : CMLS*, 61 (19–20), pp.2571–2587.
- Holland, D., Hoppe-Seyler, K., Schuller, B., Lohrey, C., Maroldt, J., Dürst, M. and Hoppe-Seyler, F. (2008). Activation of the enhancer of zeste homologue 2 gene by the human papillomavirus E7 oncoprotein. *Cancer research*, 68 (23), pp.9964–9972.
- Holowaty, P., Miller, A. B., Rohan, T. and To, T. (1999). Natural history of dysplasia of the uterine cervix. *Journal of the National Cancer Institute*, 91 (3), pp.252–258.
- Hopman, A. H. N., Smedts, F., Dignef, W., Ummelen, M., Sonke, G., Mravunac, M., Vooijs, G. P., Speel, E.-J. M. and Ramaekers, F. C. S. (2004). Transition of high-grade cervical intraepithelial neoplasia to micro-invasive carcinoma is characterized by integration of HPV 16/18 and numerical chromosome abnormalities. *The Journal of pathology*, 202 (1), pp.23–33.
- Huang, L.-W., Pan, H.-S., Lin, Y.-H., Seow, K.-M., Chen, H.-J. and Hwang, J.-L. (2011). P16 methylation is an early event in cervical carcinogenesis. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society*, 21 (3), pp.452–456.
- Hudelist, G., Manavi, M., Pischinger, K. I. D., Watkins-Riedel, T., Singer, C. F., Kubista, E. and Czerwenka, K. F. (2004). Physical state and expression of HPV DNA in benign and dysplastic cervical tissue: different levels of viral integration are correlated with lesion grade. *Gynecologic oncology*, 92 (3), pp.873–880.

- Hutchinson, M. L., Zahniser, D. J., Sherman, M. E., Herrero, R., Alfaro, M., Bratti, M. C., Hildesheim, A., Lorincz, A. T., Greenberg, M. D., Morales, J., et al. (1999). Utility of liquid-based cytology for cervical carcinoma screening: results of a population-based study conducted in a region of Costa Rica with a high incidence of cervical carcinoma. *Cancer*, 87 (2), pp.48–55.
- Hyland, P. L., McDade, S. S., McCloskey, R., Dickson, G. J., Arthur, K., McCance, D. J. and Patel, D. (2011). Evidence for alteration of EZH2, BMI1, and KDM6A and epigenetic reprogramming in human papillomavirus type 16 E6/E7-expressing keratinocytes. *Journal of virology*, 85 (21), pp.10999–11006.
- Iliopoulos, D., Oikonomou, P., Messinis, I. and Tsezou, A. (2009). Correlation of promoter hypermethylation in hTERT, DAPK and MGMT genes with cervical oncogenesis progression. *Oncology reports*, 22 (1), pp.199–204.
- Illingworth, R., Kerr, A., DeSousa, D., Jørgensen, H., Ellis, P., Stalker, J., Jackson, D., Clee, C., Plumb, R., Rogers, J., et al. (2008). A Novel CpG Island Set Identifies Tissue-Specific Methylation at Developmental Gene Loci. Liu, E. T. (Ed). *PLoS Biology*, 6 (1), p.e22.
- Illumina (2015) Illumina Infinium HD Assay methylation protocol guide. available at: [https://support.illumina.com/infinium\\_assays/infinium\\_hd\\_methylation/infinium-hd-methylation-assay-reference-guide-15019519-06.pdf](https://support.illumina.com/infinium_assays/infinium_hd_methylation/infinium-hd-methylation-assay-reference-guide-15019519-06.pdf)
- Issa, J.-P. J., Garcia-Manero, G., Giles, F. J., Mannari, R., Thomas, D., Faderl, S., Bayar, E., Lyons, J., Rosenfeld, C. S., Cortes, J., et al. (2004). Phase 1 study of low-dose prolonged exposure schedules of the hypomethylating agent 5-aza-2'-deoxycytidine (decitabine) in hematopoietic malignancies. *Blood*, 103 (5), pp.1635–1640.
- Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y (2011).. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, 333 (6047), pp. 1300-1303.
- Jacobs, J. J. L. and van Lohuizen, M. (2002). Polycomb repression: from cellular memory to cellular proliferation and cancer. *Biochimica et biophysica acta*, 1602 (2), pp.151–161.
- Jacobson, D. L., Peralta, L., Farmer, M., Graham, N. M. H., Wright, T. C. and Zenilman, J. (1999). Cervical ectopy and the transformation zone measured by computerized planimetry in adolescents. *International Journal of Gynecology and Obstetrics*, 66 (1), pp.7–17.
- Jaenisch, R. and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genetics*, 33 (S3), Nature Publishing Group., pp.245–254
- Jair, K.-W., Bachman, K. E., Suzuki, H., Ting, A. H., Rhee, I., Yen, R.-W. C., Baylin, S. B. and Schuebel, K. E. (2006). De novo CpG island methylation in human cancer cells. *Cancer research*, 66 (2), pp.682–692.
- Jeltsch, A., Nellen, W. and Lyko, F. (2006). Two substrates are better than one: dual specificities for Dnmt2 methyltransferases. *Trends in Biochemical Sciences*, 31 (6), pp.306–308.

- Jenuwein, T. and Allis, C. D. (2001). Translating the histone code. *Science (New York, N.Y.)*, 293 (5532), pp.1074–1080.
- Jeon, S. and Lambert, P. F. (1995). Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of E6 and E7 mRNAs: implications for cervical carcinogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 92 (5), pp.1654–1658.
- Ji D, Lin K, Song J, Wang Y. (2014). Effects of Tet-induced oxidation products of 5-methylcytosine on Dnmt1- and DNMT3a-mediated cytosine methylation. *Mol Biosyst*, 10(7), pp.1749-1752.
- Jia, D., Jurkowska, R. Z., Zhang, X., Jeltsch, A. and Cheng, X. (2007). Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature*, 449 (7159), pp.248–251.
- Jimenez-pacheco, A., Exposito-ruiz, M., Arrabal-polo, M. A. and Lopez-luque, A. J. (2012). Urological Oncology Meta-Analysis of Studies Analyzing the Role of Human Papillomavirus in the Development of Bladder Carcinoma. pp.240–247.
- Jones, P. a and Baylin, S. B. (2007). The epigenomics of cancer. *Cell*, 128 (4), pp.683–692.
- Jurkowska, R. Z., Jurkowski, T. P. and Jeltsch, A. (2011). Structure and function of mammalian DNA methyltransferases. *Chembiochem : a European journal of chemical biology*, 12 (2), pp.206–222.
- Kaminskas, E., Farrell, A. T., Wang, Y.-C., Sridhara, R. and Pazdur, R. (2005). FDA drug approval summary: azacitidine (5-azacytidine, Vidaza) for injectable suspension. *The oncologist*, 10 (3), AlphaMed Press., pp.176–182.
- Kämper, N., Day, P. M., Nowak, T., Selinka, H.-C., Florin, L., Bolscher, J., Hilbig, L., Schiller, J. T. and Sapp, M. (2006). A membrane-destabilizing peptide in capsid protein L2 is required for egress of papillomavirus genomes from endosomes. *Journal of virology*, 80 (2), pp.759–768.
- Kang, S., Kim, J. W., Kang, G. H., Lee, S., Park, N. H., Song, Y. S., Park, S. Y., Kang, S. B. and Lee, H. P. (2006). Comparison of DNA hypermethylation patterns in different types of uterine cancer: cervical squamous cell carcinoma, cervical adenocarcinoma and endometrial adenocarcinoma. *International journal of cancer*, 118 (9), pp.2168–2171.
- Kang, S., Kim, J. W., Kang, G. H., Park, N. H., Song, Y. S., Kang, S. B. and Lee, H. P. (2005). Polymorphism in folate- and methionine-metabolizing enzyme and aberrant CpG island hypermethylation in uterine cervical cancer. *Gynecologic oncology*, 96 (1), pp.173–180.
- Kantarjian, H., Oki, Y., Garcia-Manero, G., Huang, X., O'Brien, S., Cortes, J., Faderl, S., Bueso-Ramos, C., Ravandi, F., Estrov, Z., et al. (2007). Results of a randomized study of 3 schedules of low-dose decitabine in higher-risk myelodysplastic syndrome and chronic myelomonocytic leukemia. *Blood*, 109 (1), Blood., pp.52–57.



- Karpf, A. R. and Matsui, S. (2005). Genetic disruption of cytosine DNA methyltransferase enzymes induces chromosomal instability in human cancer cells. *Cancer research*, 65 (19), pp.8635–8639.
- Katargin, A. N., Pavlova, L. S., Kisseljov, F. L. and Kisseljova, N. P. (2009). Hypermethylation of genomic 3.3-kb repeats is frequent event in HPV-positive cervical cancer. *BMC medical genomics*, 2, p.30.
- Kebede, A. F., Schneider, R. and Daujat, S. (2015). Novel types and sites of histone modifications emerge as players in the transcriptional regulation contest. *FEBS Journal*, 282 (9), pp.1658–1674.
- Ki, K.-D., Lee, S.-K., Tong, S.-Y., Lee, J.-M., Song, D.-H. and Chi, S.-G. (2008). Role of 5'-CpG island hypermethylation of the FHIT gene in cervical carcinoma. *Journal of gynecologic oncology*, 19 (2), pp.117–122.
- Kidd, T., Brose, K., Mitchell, K. J., Fetter, R. D., Tessier-lavigne, M., Goodman, C. S. and Tear, G. (1998). *of the CNS Midline and Defines a Novel Subfamily of Evolutionarily Conserved Guidance Receptors*. 92, pp.205–215.
- Kim, G.-D., Ni, J., Kelesoglu, N., Roberts, R. J. and Pradhan, S. (2002). Co-operation and communication between the human maintenance and de novo DNA (cytosine-5) methyltransferases. *The EMBO journal*, 21 (15), pp.4183–4195.
- Kim, J.-H., Choi, Y. D., Lee, J. S., Lee, J. H., Nam, J. H. and Choi, C. (2010). Assessment of DNA methylation for the detection of cervical neoplasia in liquid-based cytology specimens. *Gynecologic oncology*, 116 (1), pp.99–104.
- Kim, J.-H., Sharma, A., Dhar, S. S., Lee, S.-H., Gu, B., Chan, C.-H., Lin, H.-K. and Lee, M. G. (2014). UTX and MLL4 coordinately regulate transcriptional programs for cell proliferation and invasiveness in breast cancer cells. *Cancer research*, 74 (6), pp.1705–1717.
- Kim, J. H., Karnovsky, A., Mahavisno, V., Weymouth, T., Pande, M., Dolinoy, D. C., Rozek, L. S. and Sartor, M. a. (2012). LRpath analysis reveals common pathways dysregulated via DNA methylation across cancer types. *BMC genomics*, 13 (1), BMC Genomics., p.526.
- Kim, S.-H., Koo, B.-S., Kang, S., Park, K., Kim, H., Lee, K. R., Lee, M. J., Kim, J. M., Choi, E. C. and Cho, N. H. (2007). HPV integration begins in the tonsillar crypt and leads to the alteration of p16, EGFR and c-myc during tumor formation. *International journal of cancer. Journal international du cancer*, 120 (7), pp.1418–1425.
- Kirnbauer, R., Taub, J., Greenstone, H., Roden, R., Dürst, M., Gissmann, L., Lowy, D. R. and Schiller, J. T. (1993). Efficient self-assembly of human papillomavirus type 16 L1 and L1-L2 into virus-like particles. *Journal of virology*, 67 (12), pp.6929–6936.
- Klaes, R., Woerner, S. M., Ridder, R., Lotz, B., Melsheimer, P. and Doeberitz, M. V. K. (1999). Detection of High-Risk Cervical Intraepithelial Neoplasia and Cervical Cancer by Amplification of Transcripts Derived from Integrated Papillomavirus Oncogenes *Cancer Res* pp.6132–6136.

- Kleer, C. G., Cao, Q., Varambally, S., Shen, R., Ota, I., Tomlins, S. A., Ghosh, D., Sewalt, R. G. A. B., Otte, A. P., Hayes, D. F., et al. (2003). EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proceedings of the National Academy of Sciences of the United States of America*, 100 (20), pp.11606–11611.
- Klimasauskas, S., Kumar, S., Roberts, R. J. and Cheng, X. (1994). HhaI methyltransferase flips its target base out of the DNA helix. *Cell*, 76 (2), Cell Press., pp.357–369.
- Kondo, Y., Shen, L., Cheng, A. S., Ahmed, S., Bumber, Y., Charo, C., Yamochi, T., Urano, T., Furukawa, K., Kwabi-addo, B., et al. (2008). Gene silencing in cancer by histone H3 lysine 27 trimethylation independent of promoter DNA methylation. *Nature Genetics*, 40 (6), pp.741–750.
- Koss, L. G. (2012). The 57th birthday of koilocytes. *Cancer cytopathology*, 120 (6), p.421.
- Kovacic, M. B., Castle, P., Herrero, R., Schiffman, M., Sherman, M. E., Wacholder, S., Rodriguez, A., Martha L Hutchinson, M. L., Bratti, C., Allan, h., (2006) Relationships of human papillomavirus type, qualitative viral load, and age with cytologic abnormality. *Cancer Res*, 15;66(20).
- Kraus, I., Driesch, C., Vinokurova, S., et al. (2008) The majority of viral-cellular fusion transcripts in cervical carcinomas cotranscribe cellular sequences of known or predicted genes. *Cancer research*, 68 (7): 2514–22
- Kühne, C., Gardiol, D., Guarnaccia, C., Amenitsch, H. and Banks, L. (2000). Differential regulation of human papillomavirus E6 by protein kinase A: conditional degradation of human discs large protein by oncogenic E6. *Oncogene*, 19 (51), pp.5884–5891.
- Kurian, K. and al-Nafussi, A. (1999). Relation of cervical glandular intraepithelial neoplasia to microinvasive and invasive adenocarcinoma of the uterine cervix: a study of 121 cases. *Journal of clinical pathology*, 52 (2), pp.112–117.
- Kuzmichev, A., Jenuwein, T., Tempst, P. and Reinberg, D. (2004). Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. *Molecular cell*, 14 (2), Mol Cell., pp.183–193.
- Kuzmichev, A., Margueron, R., Vaquero, A., Preissner, T. S., Scher, M., Kirmizis, A., Ouyang, X., Brockdorff, N., Abate-Shen, C., Farnham, P., et al. (2005). Composition and histone substrates of polycomb repressive group complexes change during cellular differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, 102 (6), pp.1859–1864.
- Kyo, S., Inoue, M., Hayasaka, N., Inoue, T., Yutsudo, M., Tanizawa, O. and Hakura, A. (1994). Regulation of early gene expression of human papillomavirus type 16 by inflammatory cytokines. *Virology*, 200 (1), pp.130–139.
- Kyrgiou, M. and Shafi, M. I. (2010). Colposcopy and cervical intra-epithelial neoplasia. *Obstetrics, Gynaecology & Reproductive Medicine*, 20 (5), Elsevier Ltd., pp.138–146.
- Kyrgiou, M. and Shafi, M. I. (2013). Invasive cancer of the cervix. *Obstetrics, Gynaecology & Reproductive Medicine*, 23 (11), pp.343–351.

- Lachner, M. and Jenuwein, T. (2002). The many faces of histone lysine methylation. *Current opinion in cell biology*, 14 (3), pp.286–298.
- LAEMMLI, U. K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, 227 (5259), pp.680–685.
- Lai, H.-C., Lin, Y.-W., Chang, C.-C., Wang, H.-C., Chu, T.-W., Yu, M.-H. and Chu, T.-Y. (2007). Hypermethylation of two consecutive tumor suppressor genes, BLU and RASSF1A, located at 3p21.3 in cervical neoplasias. *Gynecologic oncology*, 104 (3), pp.629–635.
- Lai, H.-C., Lin, Y.-W., Huang, R.-L., Chung, M.-T., Wang, H.-C., Liao, Y.-P., Su, P.-H., Liu, Y.-L. and Yu, M.-H. (2010). Quantitative DNA methylation analysis detects cervical intraepithelial neoplasms type 3 and worse. *Cancer*, 116 (18), pp.4266–4274.
- Lai, H.-C., Lin, Y.-W., Huang, T. H. M., Yan, P., Huang, R.-L., Wang, H.-C., Liu, J., Chan, M. W. Y., Chu, T.-Y., Sun, C.-A., et al. (2008). Identification of novel DNA methylation markers in cervical cancer. *International journal of cancer*, 123 (1), pp.161–167.
- Lavelle, D., Sauntharajah, Y. and DeSimone, J. (2008). DNA methylation and mechanism of action of 5-azacytidine. *Blood*, 111 (4), pp.2485–2485.
- Lawson, J. S., Glenn, W. K., Heng, B., Ye, Y., Tran, B., Lutze-Mann, L. and Whitaker, N. J. (2009). Koilocytes indicate a role for human papilloma virus in breast cancer. *British journal of cancer*, 101 (8), pp.1351–1356.
- Lee, D., Kim, H.-Z., Jeong, K. W., Shim, Y. S., Horikawa, I., Barrett, J. C. and Choe, J. (2002). Human Papillomavirus E2 Down-regulates the Human Telomerase Reverse Transcriptase Promoter. *J. Biol. Chem.*, 277 (31), pp.27748–27756.
- Lee, J.-O., Kwun, H. J., Jung, J. K., Choi, K. H., Min, D. S. and Jang, K. L. (2005). Hepatitis B virus X protein represses E-cadherin expression via activation of DNA methyltransferase 1. *Oncogene*, 24 (44), pp.6617–6625.
- Lee, P. J., Washer, L. L., Law, D. J., Boland, C. R., Horon, I. L. and Feinberg, A. P. (1996). Limited up-regulation of DNA methyltransferase in human colon cancer reflecting increased cell proliferation. *Proceedings of the National Academy of Sciences of the United States of America*, 93 (19), pp.10366–10370.
- Leonard, S. M., Wei, W., Collins, S. I., Pereira, M., Diyaf, A., Constandinou-Williams, C., Young, L. S., Roberts, S. and Woodman, C. B. (2012a). Oncogenic human papillomavirus imposes an instructive pattern of DNA methylation changes which parallel the natural history of cervical HPV infection in young women. *Carcinogenesis*, 33 (7), pp.1286–1293.
- Li, E., Beard, C. and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. *Nature*, 366 (6453), pp.362–365.
- Li, E., Bestor, T. H. and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, 69 (6), pp.915–926.
- Liang, W.S., Aldrich, J., Nasser, S., et al. (2014) Simultaneous characterization of somatic events and HPV-18 integration in a metastatic cervical carcinoma patient using DNA and RNA sequencing. *International journal of gynecological cancer*, 24 (2): 329–38

- Lim, E. H., Ng, S. L., Li, J. L., Chang, A. R., Ng, J., Ilancheran, A., Low, J., Quek, S. C. and Tay, E. H. (2010). Cervical dysplasia: Assessing methylation status (Methylight) of CCNA1, DAPK1, HS3ST2, PAX1 and TFPI2 to improve diagnostic accuracy. *Gynecologic Oncology*, 119 (2), pp.225–231.
- Lin, R. K. and Wang, Y. C. (2014). Dysregulated transcriptional and post-translational control of DNA methyltransferases in cancer. *Cell and Bioscience*, 4 (1), pp.1–11.
- Linhart, H.G., Lin, H., Yamada, Y., et al. (2007) Dnmt3b promotes tumorigenesis in vivo by gene-specific de novo methylation and transcriptional silencing. *Genes and development*, 21(23):3110-22
- Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., Nery, J. R., Lee, L., Ye, Z., Ngo, Q.-M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, 462 (7271), pp.315–322.
- Longworth, M. S. and Laimins, L. A. (2004). The Binding of Histone Deacetylases and the Integrity of Zinc Finger-Like Motifs of the E7 Protein Are Essential for the Life Cycle of Human Papillomavirus Type 31. *Journal of Virology*, 78 (7), pp.3533–3541.
- Luczak, M. W. and Jagodziński, P. P. (2006). The role of DNA methylation in cancer development. *Folia histochemica et cytobiologica / Polish Academy of Sciences, Polish Histochemical and Cytochemical Society*, 44 (3), pp.143–154.
- Łuczak, M. W., Roszak, A., Pawlik, P., Kędzia, H., Kędzia, W., Malkowska-Walczak, B., Lianeri, M. and Jagodziński, P. P. (2012). Transcriptional analysis of CXCR4, DNMT3A, DNMT3B and DNMT1 gene expression in primary advanced uterine cervical carcinoma. *International Journal of Oncology*, 40 (3), Spandidos Publications., pp.860–866.
- Luft, F., Klaes, R., Nees, M., Dürst, M., Heilmann, V., Melsheimer, P. and von Knebel Doeberitz, M. (2001). Detection of integrated papillomavirus sequences by ligation-mediated PCR (DIPS-PCR) and molecular characterization in cervical cancer cells. *International journal of cancer. Journal international du cancer*, 92 (1), pp.9–17.
- Lyko, F. (2018). The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. *Nature Reviews Genetics*, 19 (2), pp.81–92.
- Malone, C. S., Miner, M. D., Doerr, J. R., Jackson, J. P., Jacobsen, S. E., Wall, R. and Teitell, M. (2001). CmC(A/T)GG DNA methylation in mature B cell lymphoma gene silencing. *Proceedings of the National Academy of Sciences of the United States of America*, 98 (18), pp.10404–10409.
- Mann IK, Chatterjee R, Zhao J, He X, Weirauch MT, Hughes TR, Vinson C. CG methylated microarrays identify a novel methylated sequence bound by the CEBPB|ATF4 heterodimer that is active in vivo (2013). *Genome Res*, 23(6), pp.988-97.
- Margueron, R., Justin, N., Ohno, K., Sharpe, M. L., Son, J., Drury III, W. J., Voigt, P., Martin, S. R., Taylor, W. R., De Marco, V., et al. (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature*, 461 (7265), Nature Publishing Group., pp.762–767.

- Mares, A. (2013) Molecular markers in cervical screening – a promise for the future. *Romanian journal of laboratory medicine*, 21 (2): 231–239
- Martins, M. R., Sc, M., Kennia, N., Ribeiro, A. A., Zeferino, L. C., Ph, D., Dufloth, R. M. and Rabelo-santos, S. H. (2011). *Cytopathic Effects of Human Papillomavirus Infection and the Severity of Cervical Intraepithelial Neoplasia* : 40 (10), pp.871–875.
- Massad, L. S., Jeronimo, J., Katki, H. A. and Schiffman, M. (2009). The accuracy of colposcopic grading for detection of high-grade cervical intraepithelial neoplasia. *Journal of lower genital tract disease*, 13 (3), pp.137–144.
- Matzner, I., Savelyeva, L. and Schwab, M. (2003). Preferential integration of a transfected marker gene into spontaneously expressed fragile sites of a breast cancer cell line. *Cancer Letters*, 189 (2), pp.207–219.
- Maufort, J. P., Williams, S. M. G., Pitot, H. C. and Lambert, P. F. (2007). Human papillomavirus 16 E5 oncogene contributes to two stages of skin carcinogenesis. *Cancer research*, 67 (13), pp.6106–6112.
- McBride AA, Warburton A (2017). The role of integration in oncogenic progression of HPV-associated cancers. *PLoS Pathog*, 13 (4), pp.1006211.
- McCabe, M. T., Davis, J. N. and Day, M. L. (2005). Regulation of DNA methyltransferase 1 by the pRb/E2F1 pathway. *Cancer research*, 65 (9), pp.3624–3632.
- McGarvey, K. M., Greene, E., Fahrner, J. a, Jenuwein, T. and Baylin, S. B. (2007). DNA methylation and complete transcriptional silencing of cancer genes persist after depletion of EZH2. *Cancer research*, 67 (11), pp.5097–5102.
- McLaughlin-Drubin, Margaret E. Crum, Christopher P. Münger, K. (2011). Human papillomavirus E7 oncoprotein induces KDM6A and KDM6B histone demethylase expression and causes epigenetic reprogramming. *Proceedings of the National Academy of Sciences of the United States of America*, 108 (5), pp.2130–2135.
- McLaughlin-Drubin, M. E. and Munger, K. (2013). Biochemical and functional interactions of human papillomavirus proteins with polycomb group proteins. *Viruses*, 5 (5), pp.1231–1249.
- McLaughlin-Drubin, M. E. and Münger, K. (2009). Oncogenic activities of human papillomaviruses. *Virus research*, 143 (2), pp.195–208.
- Medeiros, R. B., Papenfuss, K. J., Hoiu, B., Coley, K., Jadrich, J., Goh, S.-K., Elayaperumal, A., Herrera, J. E., Resnik, E. and Ni, H.-T. (2009). Novel sequential ChIP and simplified basic ChIP protocols for promoter co-occupancy and target gene identification in human embryonic stem cells. *BMC biotechnology*, 9, pp.59.
- Melnikow, J., Nuovo, J., Willan, A. R., Chan, B. K. and Howell, L. P. (1998). Natural history of cervical squamous intraepithelial lesions: a meta-analysis. *Obstetrics and gynecology*, 92 (4 Pt 2), pp.727–735.
- Mesri, E. A., Feitelson, M. A. and Munger, K. (2014). Human viral oncogenesis: a cancer hallmarks analysis. *Cell host & microbe*, 15 (3), pp.266–282.

- Michalak, E. M., Burr, M. L., Bannister, A. J. and Dawson, M. A. (2019). The roles of DNA, RNA and histone methylation in ageing and cancer. *Nature Reviews Molecular Cell Biology*, 20 (10), Nature Publishing Group., pp.573–589.
- Middleton, K., Peh, W., Southern, S., Griffin, H., Sotlar, K., Nakahara, T., El-Sherif, A., Morris, L., Seth, R., Hibma, M., et al. (2003). Organization of human papillomavirus productive cycle during neoplastic progression provides a basis for selection of diagnostic markers. *Journal of virology*, 77 (19), pp.10186–10201.
- Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R. P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, 448 (7153), pp.553–560.
- Minami, K., Chano, T., Kawakami, T., et al. (2010a) DNMT3L is a novel marker and is essential for the growth of human embryonal carcinoma. *Clinical cancer research*, 16 (10): 2751–9
- Mirabello, L., Sun, C., Ghosh, A., Rodriguez, A. C., Schiffman, M., Wentzensen, N., Hildesheim, A., Herrero, R., Wacholder, S., Lorincz, A., et al. (2012). Methylation of human papillomavirus type 16 genome and risk of cervical precancer in a Costa Rican population. *Journal of the National Cancer Institute*, 104 (7), pp.556–565.
- Mirabello L., Clarke M. A., Nelson C. W., et al. (2018). The intersection of HPV epidemiology, genomics and mechanistic studies of HPV-mediated carcinogenesis. *Viruses*, 10(2):80.
- Mitra, A. B., Murty, V. V. V. S. and Li, R. G. (1994). Allelotype Analysis of Cervical Carcinoma. *Cancer Res.* pp.4481–4487.
- Mitra, S., Mazumder-Indra, D., Mondal, R. K., Basu, P. S., Roy, A., Roychoudhury, S. and Panda, C. K. (2012). Inactivation of SLIT2-ROBO1/2 pathway in premalignant lesions of uterine cervix: clinical and prognostic significances. *PloS one*, 7 (6), p.e38342.
- Mizuno, S., Chijiwa, T., Okamura, T., Akashi, K., Fukumaki, Y., Niho, Y. and Sasaki, H. (2001). Expression of DNA methyltransferases DNMT1, 3A, and 3B in normal hematopoiesis and in acute and chronic myelogenous leukemia. *Blood*, 97 (5), pp.1172–1179.
- Mody, D. R., Nayar, R. and Thrall, M. (2011). 2001 Bethesda System classification of glandular lesions on cervical cytology. *Monographs in clinical cytology*, 20, pp.5–14.
- Mosammaparast, N. and Shi, Y. (2010). Reversal of Histone Methylation: Biochemical and Molecular Mechanisms of Histone Demethylases. *Annual Review of Biochemistry*, 79 (1), pp.155–179.
- Mulder, K. W., Wang, X., Escribe, C., Ito, Y., Schwarz, R. F., Gillis, J., Sirokmány, G., Donati, G., Uribe-Lewis, S., Pavlidis, P., et al. (2012). Diverse epigenetic strategies interact to control epidermal differentiation. *Nature cell biology*, 14 (7), pp.753–763.
- Müller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., Wild, B., Miller, E. L., O'Connor, M. B., Kingston, R. E. and Simon, J. A. (2002). Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell*, 111 (2), pp.197–208.

- Münger, K., Baldwin, A., Edwards, K. M., Hayakawa, H., Nguyen, C. L., Owens, M., Grace, M. and Huh, K. (2004). Mechanisms of human papillomavirus-induced oncogenesis. *Journal of virology*, 78 (21), pp.11451–11460.
- Münger, K., Scheffner, M., Huibregtse, J. M. and Howley, P. M. (1992). Interactions of HPV E6 and E7 oncoproteins with tumour suppressor gene products. *Cancer surveys*, 12, pp.197–217.
- Nakao, Y., Yang, X., Yokoyama, M. (1996). Malignant transformation of human ectocervical cells immortalized by HPV 18: in vitro model of carcinogenesis by cigarette smoke. *Carcinogenesis*, 17,(3), p 577–583,
- Narayan, G., Arias-Pulido, H., Koul, S., Vargas, H., Zhang, F. F., Vilella, J., Schneider, A., Terry, M. B., Mansukhani, M. and Murty, V. V. (2003). Frequent promoter methylation of CDH1, DAPK, RARB, and HIC1 genes in carcinoma of cervix uteri: its relationship to clinical outcome. *Molecular cancer*, 2, p.24.
- Narayan, G., Goparaju, C., Arias-Pulido, H., Kaufmann, A. M., Schneider, A., Dürst, M., Mansukhani, M., Pothuri, B. and Murty, V. V. (2006). Promoter hypermethylation-mediated inactivation of multiple Slit-Robo pathway genes in cervical cancer progression. *Molecular cancer*, 5 (i), p.16.
- National Cancer Institute Workshop. (1989). The 1988 Bethesda System for reporting cervical/vaginal cytological diagnoses. *JAMA*, 262 (7), pp.931–934.
- Ng HH, Zhang Y, Hendrich B, Johnson CA, Turner BM, Erdjument-Bromage H, Tempst P, Reinberg D, Bird A (1999). MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat. Genet*, 23(1), pp.58–61.
- Nichol, J. N., Dupéré-Richer, D., Ezponda, T., Licht, J. D. and Miller, W. H. (2016). H3K27 Methylation: A Focal Point of Epigenetic Deregulation in Cancer. *Adv Cancer Res*, 131, pp.59–95.
- Nicolaidis, L., Davy, C., Raj, K., Kranjec, C., Banks, L. and Doorbar, J. (2011). Stabilization of HPV16 E6 protein by PDZ proteins, and potential implications for genome maintenance. *Virology*, 414 (2), pp.137–145.
- van de Nieuwenhof, H. P., van Kempen, L. C. L. T., de Hullu, J. a, Bekkers, R. L. M., Bulten, J., Melchers, W. J. G. and Massuger, L. F. a G. (2009). The etiologic role of HPV in vulvar squamous cell carcinoma fine tuned. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 18 (7), pp.2061–2067.
- Noya, F., Chien, W., Broker, T. R. and Chow, L. T. (2001). p21cip1 Degradation in Differentiated Keratinocytes Is Abrogated by Costabilization with Cyclin E Induced by Human Papillomavirus E7. *Journal of Virology* .75 (13), pp.6121–6134.
- Oh, S. T., Kyo, S. and Laimins, L. A. (2001). Telomerase activation by human papillomavirus type 16 E6 protein: induction of human telomerase reverse transcriptase expression through Myc and GC-rich Sp1 binding sites. *Journal of virology*, 75 (12), pp.5559–5566.

- Ohm, J. E., McGarvey, K. M., Yu, X., Cheng, L., Schuebel, K. E., Cope, L., Mohammad, H. P., Chen, W., Daniel, V. C., Yu, W., et al. (2007). A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. *Nature genetics*, 39 (2), pp.237–242.
- Okano, M., Bell, D. W., Haber, D. a and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, 99 (3), pp.247–257.
- Okano, M., Xie, S. and Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature genetics*, 19 (3), pp.219–220.
- Oka, M., Meacham, A.M., Hamazaki, T., et al. (2005) De novo DNA methyltransferases Dnmt3a and Dnmt3b primarily mediate the cytotoxic effect of 5-aza-2'-deoxycytidine. *Oncogene*, 24 (19): 3091–9
- Organista-Nava, J., Gómez-Gómez, Y. and Gariglio, P. (2013). Embryonic stem cell-specific signature in cervical cancer. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*.
- Ostler, K. R., Davis, E. M., Payne, S. L., Gosalia, B. B., Expósito-Céspedes, J., Le Beau, M. M. and Godley, L. a. (2007). Cancer cells express aberrant DNMT3B transcripts encoding truncated proteins. *Oncogene*, 26 (38), pp.5553–5563.
- Park, J. S., Hwang, E. S., Park, S. N., Ahn, H. K., Um, S. J., Kim, C. J., Kim, S. J. and Namkoong, S. E. (1997). Physical status and expression of HPV genes in cervical cancers. *Gynecologic oncology*, 65 (1), pp.121–129.
- Pelling, A. L., Thorne, A. W. and Crane-Robinson, C. (2000). A human genomic library enriched in transcriptionally active sequences (aDNA library). *Genome research*, 10 (6), Cold Spring Harbor Laboratory Press., pp.874–886.
- Peter, M., Rosty, C., Couturier, J., Radvanyi, F., Teshima, H. and Sastre-Garau, X. (2006). MYC activation associated with the integration of HPV DNA at the MYC locus in genital tumors. *Oncogene*, 25 (44), pp.5985–5993.
- Pett, M. and Coleman, N. (2007). Integration of high-risk human papillomavirus : a key event in cervical carcinogenesis ? *Journal of Pathology*, pp.356–367.
- Pett, M. R., Alazawi, W. O. F., Roberts, I., Dowen, S., Smith, D. I., Stanley, M. A. and Coleman, N. (2004). Acquisition of High-Level Chromosomal Instability Is Associated with Integration of Human Papillomavirus Type 16 in Cervical Keratinocytes. 9, pp.1359–1368.
- Pett, M. R., Herdman, M. T., Palmer, R. D., Yeo, G. S. H., Shivji, M. K., Stanley, M. A. and Coleman, N. (2006). Selection of cervical keratinocytes containing integrated HPV16 associates with episome loss and an endogenous antiviral response. *PNAS*.
- Pietersen, A. M. and van Lohuizen, M. (2008). Stem cell regulation by polycomb repressors: postponing commitment. *Current opinion in cell biology*, 20 (2), pp.201–207.



- Piirsoo, M., Ustav, E., Mandel, T., Stenlund, A. and Ustav, M. (1996). Cis and trans requirements for stable episomal maintenance of the BPV-1 replicator. *The EMBO journal*, 15 (1), pp.1–11.
- Pirisi, L., Yasumoto, S., Feller, M., Doniger, J., and DiPaolo, J.A. (1987). Transformation of human fibroblasts and keratinocytes with human papillomavirus type 16 DNA. *Journal of Virology*, 61(4): 1061–1066.
- Plath, K., Fang, J., Mlynarczyk-Evans, S. K., Cao, R., Worringer, K. A., Wang, H., de la Cruz, C. C., Otte, A. P., Panning, B. and Zhang, Y. (2003). Role of histone H3 lysine 27 methylation in X inactivation. *Science*, 300 (5616), pp.131–135.
- Popescu, N. C. and DiPaolo, J. A. (1990). Integration of human papillomavirus 16 DNA and genomic rearrangements in immortalized human keratinocyte lines. *Cancer research*, 50 (4), pp.1316–1323.
- Potapova, A., Albat, C., Hasemeier, B., Haeussler, K., Lamprecht, S., Suerbaum, S., Kreipe, H. and Lehmann, U. (2011). Systematic cross-validation of 454 sequencing and pyrosequencing for the exact quantification of DNA methylation patterns with single CpG resolution. *BMC biotechnology*, 11 (1), p.6.
- Pradhan, S. and Kim, G.-D. (2002). The retinoblastoma gene product interacts with maintenance human DNA (cytosine-5) methyltransferase and modulates its activity. *The EMBO journal*, 21 (4), EMBO Press., pp.779–788.
- Programme, N. cervical screening. (2016). Indication for referral to colposcopy. (January), pp.1–25.
- Pulido, H. A., Fakruddin, M. J., Chatterjee, A., Esplin, E. D., Belen, N., Martı, G., Posso, H., Evans, G. A. and Murty, V. V. V. S. (2000). Exclusion of PPP2R1B Gene as a Deletion Target in Cervical Cancer PPP2R1B Gene as a Deletion Target in Cervical Cancer 1. pp.6677–6682.
- Pyeon, D., Pearce, S. M., Lank, S. M., Ahlquist, P. and Lambert, P. F. (2009). Establishment of human papillomavirus infection requires cell cycle progression. *PLoS pathogens*, 5 (2), p.e1000318.
- Rajkumar, T., Cuzick, J., Appleby, P., Barnabas, R., Beral, V., Berrington De González, a., Bull, D., Canfell, K., Crossley, B., Green, J., et al. (2006). Cervical carcinoma and reproductive factors. *International Journal of Cancer*, 119 (5), pp.1108–1124.
- Ramsahoye, B. H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A. P. and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proceedings of the National Academy of Sciences of the United States of America*, 97 (10), pp.5237–5242.
- Ramsahoye, B. H., Davies, C. S. and Mills, K. I. (1996). DNA methylation: biology and significance. *Blood reviews*, 10 (4), pp.249–261.
- Reesink-Peters, N., Wisman, G. B. A., Jéronimo, C., Tokumaru, C. Y., Cohen, Y., Dong, S. M., Klip, H. G., Buikema, H. J., Suurmeijer, A. J. H., Hollema, H., et al. (2004). Detecting

- cervical cancer by quantitative promoter hypermethylation assay on cervical scrapings: a feasibility study. *Molecular cancer research : MCR*, 2 (5), pp.289–295.
- Rhee, I., Bachman, K. E., Park, B. H., Jair, K.-W., Yen, R.-W. C., Schuebel, K. E., Cui, H., Feinberg, A. P., Lengauer, C., Kinzler, K. W., et al. (2002). DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature*, 416 (6880), pp.552–556.
- Rickinson, A. B. (1999). Targeting human tumours with antigen-specific cytotoxic T-cells. *British journal of cancer*, 80 Suppl 1 (1–2), pp.51–56.
- Robertson, K. D., Ait-Si-Ali, S., Yokochi, T., Wade, P. A., Jones, P. L. and Wolffe, A. P. (2000). DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nature genetics*, 25 (3), pp.338–342.
- Robertson, K. D., Uzvolgyi, E., Liang, G., Talmadge, C., Sumegi, J., Gonzales, F. A. and Jones, P. A. (1999). The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors. *Nucleic acids research*, 27 (11), pp.2291–2298.
- Roteli-Martins, C.M., Alves, V. a, Santos, R.T., et al. (2001) Value of morphological criteria in diagnosing cervical HPV lesions confirmed by in situ hybridization and hybrid capture assay. *Pathology, research and practice*, 197 (10): 677–82.
- Saegusa, M. and Okayasu, I. (1999). DCC expression is related to mucinous differentiation apoptosis , cell proliferation and human papillomavirus infection in uterine cervical adenocarcinomas. 80, pp.51–58.
- Santi, D. V., Norment, A. and Garrett, C. E. (1984). Covalent bond formation between a DNA-cytosine methyltransferase and DNA containing 5-azacytosine. *Proceedings of the National Academy of Sciences*, 81 (22), pp.6993–6997.
- Saunier, M., Monnier-Benoit, S., Mauny, F., Dalstein, V., Briolat, J., Riethmuller, D., Kantelip, B., Schwarz, E., Mougin, C. and Pr  tet, J.-L. (2008). Analysis of human papillomavirus type 16 (HPV16) DNA load and physical state for identification of HPV16-infected women with high-grade lesions or cervical carcinoma. *Journal of clinical microbiology*, 46 (11), pp.3678–3685.
- Schlesinger, Y., Straussman, R., Keshet, I., Farkash, S., Hecht, M., Zimmerman, J., Eden, E., Yakhini, Z., Ben-Shushan, E., Reubinoff, B. E., et al. (2007). Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. *Nature genetics*, 39 (2), pp.232–236.
- Schmitz, M., Driesch, C., Beer-Grondke, K., Jansen, L., Runnebaum, I. B. and D  rst, M. (2012). Loss of gene function as a consequence of human papillomavirus DNA integration. *International journal of cancer. Journal international du cancer*, pp.1–10.
- Schwartz, Y. B. and Pirrotta, V. (2008). Polycomb complexes and epigenetic states. *Current opinion in cell biology*, 20 (3), pp.266–273.

- Schwarz, E., Freese, U. K., Gissmann, L., Mayer, W., Roggenbuck, B., Stremlau, A. and zur Hausen, H. (1985). Structure and transcription of human papillomavirus sequences in cervical carcinoma cells. *Nature*, 314 (6006), pp.111–114.
- Scott ML, Coleman DT, Kelly KC, Carroll JL, Woodby B et al (2018). Human papillomavirus type 16 E5-mediated upregulation of Met in human keratinocytes. *Virology*, 519, pp.1–11.
- Seagon, S., Dürst, M. and Matthiasdörst, S. (1994). Genetic Analysis of an in Vitro Model System for Human Papillomavirus Type 16-associated Tumorigenesis Genetic Analysis of an in Vitro Model System for Human Papillomavirus. *Cancer Res*, pp.5593–5598.
- Seeger, M., Tear, G., D Ferres-Marco, D., Goodman, C S. (1993) Mutations affecting growth cone guidance in Drosophila: genes necessary for guidance toward or away from the midline. *Neuron* (3):409-26.
- Sen, G. L., Reuter, J. A., Webster, D. E., Zhu, L. and Khavari, P. A. (2011). *DNMT1 Maintains Progenitor Function in Self-Renewing Somatic Tissue*. 463 (7280), pp.563–567.
- Seng, T. J., Low, J. S. W., Li, H., Cui, Y., Goh, H. K., Wong, M. L. Y., Srivastava, G., Sidransky, D., Califano, J., Steenbergen, R. D. M., et al. (2007). The major 8p22 tumor suppressor DLC1 is frequently silenced by methylation in both endemic and sporadic nasopharyngeal, esophageal, and cervical carcinomas, and inhibits tumor cell colony formation. *Oncogene*, 26 (6), pp.934–944.
- Shah, M. Y., Vasanthakumar, A., Barnes, N. Y., Figueroa, M. E., Kamp, A., Hendrick, C., Ostler, K. R., Davis, E. M., Lin, S., Anastasi, J., et al. (2010). DNMT3B7, a truncated DNMT3B isoform expressed in human tumors, disrupts embryonic development and accelerates lymphomagenesis. *Cancer research*, 70 (14), pp.5840–5850. [Online]. Available at: doi:10.1158/0008-5472.CAN-10-0847.
- Shamay, M., Krithivas, A., Zhang, J. and Hayward, S. D. (2006). Recruitment of the de novo DNA methyltransferase Dnmt3a by Kaposi's sarcoma-associated herpesvirus LANA. *Proceedings of the National Academy of Sciences of the United States of America*, 103 (39), pp.14554–14559.
- Shamma, A., Suzuki, M., Hayashi, N., Kobayashi, M., Sasaki, N., Nishiuchi, T., Doki, Y., Okamoto, T., Kohno, S., Muranaka, H., et al. (2013). ATM mediates pRB function to control DNMT1 protein stability and DNA methylation. *Molecular and cellular biology*, 33 (16), pp.3113–3124.
- Shen, L. and Waterland, R. a. (2007). Methods of DNA methylation analysis. *Current opinion in clinical nutrition and metabolic care*, 10 (5), pp.576–581.
- Sherman, L. and Schlegel, R. (1996). Serum- and calcium-induced differentiation of human keratinocytes is inhibited by the E6 oncoprotein of human papillomavirus type 16 . Serum- and Calcium-Induced Differentiation of Human Keratinocytes Is Inhibited by the E6 Oncoprotein of Human Papillomav. *Journal of Virology*, 70 (5).
- Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., Casero, R. A. and Shi, Y. (2004). Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1. *Cell*, 119 (7), Cell Press., pp.941–953.

- Simó-Riudalbas, L., Melo, S.A. and Esteller, M. (2011) DNMT3B gene amplification predicts resistance to DNA demethylating drugs. *Genes, chromosomes & cancer*, 50 (7): 527–34
- Simpson, J. H., Kidd, T., Bland, K. S. and Goodman, C. S. (2000). Short-range and long-range guidance by slit and its Robo receptors. Robo and Robo2 play distinct roles in midline guidance. *Neuron*, 28 (3), pp.753–766.
- Singal, R. and Ginder, G. D. (1999). DNA Methylation. *Blood*, 93 (12), pp.4059–4070.
- Snijders, P. J. F., Steenbergen, R. D. M., Heideman, D. a M. and Meijer, C. J. L. M. (2006). HPV-mediated cervical carcinogenesis: concepts and clinical implications. *The Journal of pathology*, 208 (2), pp.152–164.
- Soto, U., Das, B. C., Lengert, M., Finzer, P., zur Hausen, H. and Rösl, F. (1999). Conversion of HPV 18 positive non-tumorigenic HeLa-fibroblast hybrids to invasive growth involves loss of TNF-alpha mediated repression of viral transcription and modification of the AP-1 transcription complex. *Oncogene*, 18 (21), pp.3187–3198.
- Sova, P., Feng, Q., Geiss, G., Wood, T., Strauss, R., Rudolf, V., Lieber, A. and Kiviat, N. (2006). Discovery of novel methylation biomarkers in cervical carcinoma by global demethylation and microarray analysis. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 15 (1), pp.114–123.
- Sparmann, A. and van Lohuizen, M. (2006). Polycomb silencers control cell fate, development and cancer. *Nature reviews. Cancer*, 6 (11), pp.846–856.
- Stanley, M. A. (2012). Epithelial cell responses to infection with human papillomavirus. *Clinical microbiology reviews*, 25 (2), pp.215–222.
- Stanley, M. A., Browne, H. M., Appleby, M. and Minson, A. C. (1989). Properties of a non-tumorigenic human cervical keratinocyte cell line. *International journal of cancer*, 43 (4), pp.672–676.
- Steenbergen, R. D. M., Kramer, D., Braakhuis, B. J. M., Stern, P. L., Verheijen, R. H. M., Meijer, C. J. L. M. and Snijders, P. J. F. (2004). TSLC1 gene silencing in cervical cancer cell lines and cervical neoplasia. *Journal of the National Cancer Institute*, 96 (4), pp.294–305.
- Stirzaker, C., Taberlay, P. C., Statham, A. L. and Clark, S. J. (2014). Mining cancer methylomes: prospects and challenges. *Trends in genetics : TIG*, 30 (2), pp.75–84.
- Su, P.-H., Lin, Y.-W., Huang, R.-L., Liao, Y.-P., Lee, H.-Y., Wang, H.-C., Chao, T.-K., Chen, C.-K., Chan, M. W. Y., Chu, T.-Y., et al. (2013). Epigenetic silencing of PTPRR activates MAPK signaling, promotes metastasis and serves as a biomarker of invasive cervical cancer. *Oncogene*, 32 (1), Macmillan Publishers Limited., pp.15–26.
- Takahashi, K. and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, 126 (4), pp.663–676.
- Takashima, S., Takehashi, M., Lee, J., Chuma, S., Okano, M., Hata, K., Suetake, I., Nakatsuji, N., Miyoshi, H., Tajima, S., et al. (2009). Abnormal DNA methyltransferase expression in

- mouse germline stem cells results in spermatogenic defects. *Biology of reproduction*, 81 (1), pp.155–164.
- Thain, A., Jenkins, O., Clarke, A. R. and Gaston, K. (1996). CpG methylation directly inhibits binding of the human papillomavirus type 16 E2 protein to specific DNA sequences. *Journal of virology*, 70 (10), pp.7233–7235.
- Theelen, W., Reijans, M., Simons, G., Ramaekers, F. C. S., Speel, E.-J. M. and Hopman, A. H. N. (2010). A new multiparameter assay to assess HPV 16/18, viral load and physical status together with gain of telomerase genes in HPV-related cancers. *International journal of cancer. Journal international du cancer*, 126 (4), pp.959–975.
- Thomas, M. and Banks, L. (1999). Human papillomavirus (HPV) E6 interactions with Bak are conserved amongst E6 proteins from high and low risk HPV types. *The Journal of general virology*, 80 ( Pt 6), pp.1513–1517.
- Tie, F., Banerjee, R., Saiakhova, A. R., Howard, B., Monteith, K. E., Scacheri, P. C., Cosgrove, M. S. and Harte, P. J. (2014). Trithorax monomethylates histone H3K4 and interacts directly with CBP to promote H3K27 acetylation and antagonize polycomb silencing. *Development (Cambridge)*, 141 (5), pp.1129–1139.
- Tine, B. A. Van, Kappes, J. C., Banerjee, N. S., Knops, J., Lai, L., Steenbergen, R. D. M., Meijer, C. L. J. M., Snijders, P. J. F., Chatis, P., Broker, T. R., et al. (2004). *Clonal Selection for Transcriptionally Active Viral Oncogenes during Progression to Cancer*. 78 (20), pp.11172–11186.
- Tost, J., Dunker, J. and Gut, I. G. (2003). Analysis and quantification of multiple methylation variable positions in CpG islands by Pyrosequencing. *BioTechniques*, 35 (1), pp.152–156.
- Tsai, C.-N., Tsai, C.-L., Tse, K.-P., Chang, H.-Y. and Chang, Y.-S. (2002). The Epstein-Barr virus oncogene product, latent membrane protein 1, induces the downregulation of E-cadherin gene expression via activation of DNA methyltransferases. *Proceedings of the National Academy of Sciences of the United States of America*, 99 (15), pp.10084–10089.
- Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P. and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. *Nature*, 439 (7078), pp.811–816.
- Ueda, Y., Enomoto, T., Miyatake, T., Shroyer, K. R., Yoshizaki, T., Kanao, H., Ueno, Y., Sun, H., Nakashima, R., Yoshino, K., et al. (2004). Analysis of clonality and HPV infection in benign, hyperplastic, premalignant, and malignant lesions of the vulvar mucosa. *American journal of clinical pathology*, 122 (2), pp.266–274.
- Vazquez-Mena, O., Medina-Martinez, I., Juárez-Torres, E., et al. (2012) Amplified genes may be overexpressed, unchanged, or downregulated in cervical cancer cell lines. *PloS one*, 7 (3): e32667
- Verhoef, V. M. J., Bosgraaf, R. P., van Kemenade, F. J., Rozendaal, L., Heideman, D. a M., Hesselink, A. T., Bekkers, R. L. M., Steenbergen, R. D. M., Massuger, L. F. a G., Melchers, W. J. G., et al. (2014). Triage by methylation-marker testing versus cytology in women who

- test HPV-positive on self-collected cervicovaginal specimens (PROHTECT-3): a randomised controlled non-inferiority trial. *The lancet oncology*, 15 (3), pp.315–322.
- Vinokurova, S., Wentzensen, N., Kraus, I., Klaes, R., Driesch, C., Melsheimer, P., Kisseljov, F., Dürst, M., Schneider, A. and von Knebel Doeberitz, M. (2008). Type-dependent integration frequency of human papillomavirus genomes in cervical lesions. *Cancer research*, 68 (1), pp.307–313.
- Viré, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J.-M., et al. (2006). The Polycomb group protein EZH2 directly controls DNA methylation. *Nature*, 439 (7078), pp.871–874.
- Walboomers, J. M., Jacobs, M. V, Manos, M. M., Bosch, F. X., Kummer, J. A., Shah, K. V, Snijders, P. J., Peto, J., Meijer, C. J. and Muñoz, N. (1999). Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *The Journal of pathology*, 189 (1), pp.12–19.
- Wang, H., Albadine, R., Magheli, A., Guzzo, T. J., Ball, M. W., Hinz, S., Schoenberg, M. P., Netto, G. J. and Gonzalgo, M. L. (2012). Increased EZH2 protein expression is associated with invasive urothelial carcinoma of the bladder. *Urologic Oncology: Seminars and Original Investigations*, 30 (4), Elsevier Inc., pp.428–433.
- Wang, S. S., Smiraglia, D. J., Wu, Y.-Z., Ghosh, S., Rader, J. S., Cho, K. R., Bonfiglio, T. A., Nayar, R., Plass, C. and Sherman, M. E. (2008). Identification of novel methylation markers in cervical cancer using restriction landmark genomic scanning. *Cancer research*, 68 (7), pp.2489–2497.
- Wasson CW, Morgan EL, Muller M, Ross RL, Hartley M et al (2017). Human papillomavirus type 18 E5 oncogene supports cell cycle progression and impairs epithelial differentiation by modulating growth factor receptor signalling during the virus life cycle. *Oncotarget*, 8, pp.103581–103600.
- Watanabe, D., Suetake, I., Tada, T., et al. (2002) Stage- and cell-specific expression of Dnmt3a and Dnmt3b during embryogenesis. *Mechanisms of development*, 118 (1-2): 187–90
- Wentzensen, N., Sherman, M. E., Schiffman, M. and Wang, S. S. (2009). Utility of methylation markers in cervical cancer early detection: appraisal of the state-of-the-science. *Gynecologic oncology*, 112 (2), pp.293–299.
- Wentzensen, N., Sherman, M. E., Schiffman, M. and Wang, S. S. (2010). *Cancer*, 112 (2), pp.293–299.
- Wentzensen, N., Sun, C., Ghosh, A., Kinney, W., Mirabello, L., Wacholder, S., Shaber, R., LaMere, B., Clarke, M., Lorincz, A. T., et al. (2012). Methylation of HPV18, HPV31, and HPV45 genomes and cervical intraepithelial neoplasia grade 3. *Journal of the National Cancer Institute*, 104 (22), pp.1738–1749.
- Whetstine, J. R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M., et al. (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell*, 125 (3), Elsevier., pp.467–481.

- Widschwendter, A., Müller, H. M., Fiegl, H., Ivarsson, L., Wiedemair, A., Müller-Holzner, E., Goebel, G., Marth, C. and Widschwendter, M. (2004). DNA methylation in serum and tumors of cervical cancer patients. *Clinical cancer research: an official journal of the American Association for Cancer Research*, 10 (2), pp.565–571.
- Widschwendter, M., Fiegl, H., Egle, D., Mueller-Holzner, E., Spizzo, G., Marth, C., Weisenberger, D. J., Campan, M., Young, J., Jacobs, I., et al. (2007). Epigenetic stem cell signature in cancer. *Nature genetics*, 39 (2), pp.157–158.
- Wilson, R., Fehrman, F. and Laimins, L. A. (2005). Role of the E1  $\wedge$  E4 Protein in the Differentiation-Dependent Life Cycle of Human Papillomavirus Type 31 Role of the E1  $\wedge$  E4 Protein in the Differentiation-Dependent Life Cycle of Human Papillomavirus Type 31. *J Virol*, 79(20), pp. 13150–13165.
- Wilting, S.M., Snijders, P.J.F., Meijer, G. a, et al. (2006) Increased gene copy numbers at chromosome 20q are frequent in both squamous cell carcinomas and adenocarcinomas of the cervix. *The Journal of pathology*, 209 (2): 220–30
- Wisman, G. B. A., Nijhuis, E. R., Hoque, M. O., Reesink-Peters, N., Koning, A. J., Volders, H. H., Buikema, H. J., Boezen, H. M., Hollema, H., Schuurin, E., et al. (2006). Assessment of gene promoter hypermethylation for detection of cervical neoplasia. *International journal of cancer*, 119 (8), pp.1908–1914.
- Woodcock, D. M., Linsenmeyer, M. E., Doherty, J. P. and Warren, W. D. (1999). DNA methylation in the promoter region of the p16 (CDKN2/MTS-1/INK4A) gene in human breast tumours. *British journal of cancer*, 79 (2), pp.251–256.
- Woodman, C. B. J., Collins, S. I. and Young, L. S. (2007). The natural history of cervical HPV infection: unresolved issues. *Nature reviews. Cancer*, 7 (1), pp.11–22.
- Wu, H., Chen, Y., Liang, J., Shi, B., Wu, G., Zhang, Y., Wang, D., Li, R., Yi, X., Zhang, H., et al. (2005). Hypomethylation-linked activation of PAX2 mediates tamoxifen-stimulated endometrial carcinogenesis. *Nature*, 438 (7070), pp.981–987.
- Wu, X., Gong, Y., Yue, J., Qiang, B., Yuan, J. and Peng, X. (2008). Cooperation between EZH2, NSPc1-mediated histone H2A ubiquitination and Dnmt1 in HOX gene silencing. *Nucleic acids research*, 36 (11), pp.3590–3599.
- Xiang, Y., Zhu, Z., Han, G., et al. (2007) JMJD3 is a histone H3K27 demethylase. *Cell research*, 17 (10): 850–7
- Xu, G. L., Bestor, T. H., Bourc’his, D., Hsieh, C. L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J. J. and Viegas-Péquignot, E. (1999). Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature*, 402 (6758), pp.187–191.
- Yang, N., Nijhuis, E. R., Volders, H. H., Eijnsink, J. J. H., Lendvai, A., Zhang, B., Hollema, H., Schuurin, E., Wisman, G. B. A. and van der Zee, A. G. J. (2010). Gene promoter methylation patterns throughout the process of cervical carcinogenesis. *Cellular oncology, the official journal of the International Society for Cellular Oncology*, 32 (1–2), pp.131–143.

- Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., Das, P. K., Kivioja, T., Dave, K., Zhong, F., et al. (2017). Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science* 356 (6337).
- Yu, T., Ferber, M. J., Hong, T., Kwok, T., Chung, H., Fu, Y. and Smith, D. I. (2005). The role of viral integration in the development of cervical cancer. *Cancer genet cytogenet*, 158(1), pp.27–34.
- Yuan, L., Tian, C., Wang, H., Song, S., Li, D., Xing, G., Yin, Y., He, F. and Zhang, L. (2012). Apak competes with p53 for direct binding to intron 1 of p53AIP1 to regulate apoptosis. *EMBO reports*, 13 (4), pp.363–370.
- Zambrano, P., Segura-Pacheco, B., Perez-Cardenas, E., Cetina, L., Revilla-Vazquez, A., Taja-Chayeb, L., Chavez-Blanco, A., Angeles, E., Cabrera, G., Sandoval, K., et al. (2005). A phase I study of hydralazine to demethylate and reactivate the expression of tumor suppressor genes. *BMC cancer*, 5, BioMed Central., p.44.
- Zhai, Y., Kuick, R., Nan, B., Ota, I., Weiss, S. J., Trimble, C. L., Fearon, E. R. and Cho, K. R. (2007). Gene expression analysis of preinvasive and invasive cervical squamous cell carcinomas identifies HOXC10 as a key mediator of invasion. *Cancer Research*, 67 (21), pp.10163–10172.
- Zioli, M., Di Tomaso, C., Biaggi, E., Tepper, J., Piquet, M., Carbillon, L., Uzan, M. and Guettier, C. (1998). Virological and biological characteristics of cervical intraepithelial neoplasia grade I with marked koilocytotic atypia. *Human Pathology* 29(10): P1068-1073