CHAPTER 6

Host cell DNA methylation patterns during high-risk HPV-induced carcinogenesis reveal a heterogeneous nature of cervical precancer
ABSTRACT

Cervical cancer development following a persistent infection with high-risk human papillomavirus (hrHPV) is driven by additional host cell changes, such as altered DNA methylation. In previous studies, we have identified 12 methylated host genes associated with cervical cancer and precancer (CIN2/3). This study systematically analysed the onset and DNA methylation pattern of these genes during hrHPV-induced carcinogenesis using a longitudinal in vitro model of hrHPV-transformed cell lines (n = 14) and hrHPV-positive cervical scrapings (n = 113) covering various stages of cervical carcinogenesis. DNA methylation analysis was performed by quantitative methylation-specific PCR (qMSP) and relative qMSP values were used to analyse the data. The majority of genes displayed a comparable DNA methylation pattern in both cell lines and clinical specimens. DNA methylation onset occurred at early or late immortal passage, and DNA methylation levels gradually increased towards tumorigenic cells. Subsequently, we defined a so-called cancer-like methylation-high pattern based on the DNA methylation levels observed in cervical scrapings from women with cervical cancer. This cancer-like methylation-high pattern was observed in 72% (38/53) of CIN3 and 55% (11/20) of CIN2, whereas it was virtually absent in hrHPV-positive controls (1/26). In conclusion, hrHPV-induced carcinogenesis is characterised by early onset of DNA methylation, typically occurring at the pretumorigenic stage and with highest DNA methylation levels at the cancer stage. Host cell DNA methylation patterns in cervical scrapings from women with CIN2 and CIN3 are heterogeneous, with a subset displaying a cancer-like methylation-high pattern, suggestive for a higher cancer risk.

INTRODUCTION

Following a persistent infection with a high-risk (hr) type of human papillomavirus (HPV), additional genetic and epigenetic changes in the host cell genome are necessary for progression to cervical cancer\(^1\). Part of these host cell alterations are induced by expression of viral oncoproteins E6 and E7 and include DNA methylation of tumour suppressor genes\(^2\). Methylation of cytosines at CpG-sites in promoter regions can lead to gene silencing. The DNA methyltransferases (DNMTs) responsible for CpG methylation can be activated by both hrHPV E6 and E7. E7 can directly bind to and activate DNMT1, whereas E6 can upregulate DNMT1 via p53\(^3,4\). Conversely, silencing of E6 and E7 has been shown to reduce DNA methylation of tumour suppressor genes and to restore the transformed phenotype in cervical cancer cells\(^5\).

Increased DNA methylation levels of several (candidate) tumour suppressor genes are associated with cervical cancer and a subset of its high-grade precursor lesions, i.e. cervical intraepithelial neoplasia grade 2 and 3 (CIN2 and CIN3)\(^6,7\). Using both targeted and genome-wide approaches, we previously identified 12 genes, including \(\text{ANKRD18CP, C13orf18, EPB41L3 and JAM3}\)\(^8\); \(\text{SOX1 and ZSCAN1}\)\(^9,10\); \(\text{GHSR, SST and ZIC1}\)\(^11\); and \(\text{FAM19A4, PHACTR3 and PRDM14}\)\(^12\). These genes were considered promising biomarkers to identify hrHPV-positive women and/ or women with abnormal cytology at risk for cervical cancer\(^13-14\).

So far, the majority of these genes has been evaluated individually or in small sets in separate studies. Herein, we performed a systematic analysis of 12 genes on the same cohort to compare their DNA methylation onset and patterns during cervical carcinogenesis. To this end, we analysed a well characterised longitudinal in vitro model of primary keratinocytes immortalised by HPV16 or HPV18\(^15\), complemented with a large series of clinically annotated hrHPV-positive cervical scrapings covering various stages of cervical carcinogenesis. The consecutive passages of hrHPV-immortalised cells, which are characterised by an initial acquisition of telomerase activity, followed by a more transformed phenotype characterised by anchorage independent growth, were shown to closely mimic the progressive stages of cervical precancerous disease with respect to both genetic and epigenetic changes\(^15-17\).
**Materials and Methods**

**Cell lines**

Cells representing the various stages of hrHPV-induced transformation, consisted of (i) primary human foreskin keratinocytes (HFK) from three donors, (ii) HPV16- (FK16A and FK16B) and HPV18- (FK18A and FK18B) immortalised keratinocytes, including early (passages 32 to 52) and late (passage 129 to 156) passages and (iii) hrHPV-positive cervical cancer cell lines SiHa, HeLa and CaSki. HFKs, FK16A, FK16B, FK18A and FK18B were obtained and cultured as described before. Cervical cancer cell lines SiHa, HeLa and CaSki (American Type Culture Collection) were cultured as described previously. Cell lines were authenticated using the PowerPlex 16 System (Promega) and were negative for mycoplasma.

**Clinical samples and hrHPV testing**

A set of hrHPV-positive scrapings (n = 113) obtained from screening or gynaecologic outpatient populations was used, comprising scrapings from control women with normal cytology and/or without evidence of CIN2+ (n = 26) or from women who were histologically diagnosed with CIN2 (n = 20), CIN3 (n = 53) or (micro-invasive) cervical cancer ([mi]Ca; n = 14). HrHPV was detected by GP5+/6+ PCR enzyme immunoassay or HPV-Risk assay (Self-screen B.V.). This study followed the ethical guidelines of the Institutional Review Board of the VU University Medical Centre and University Medical Centre Groningen.

**DNA isolation, modification and qMSP**

DNA extraction, quality assessment, sodium bisulphite-treatment, (multiplex) qMSP conditions and DNA methylation ratio calculations were described previously for ANKRD18CP, C13orf18, EPB41L3 and JAM3; SOX1 and ZSCAN1; GHSR, SST and ZIC1; and FAM19A4, PHACTR3 and PRDM14. Quality control was routinely performed using a methylation-independent assay for the housekeeping gene ACTB.

**Data analyses**

The DNA methylation ratios of the 12 genes were relatively analysed to obtain equivalent values from each of the individual genes in the datasets of both the in vitro model and the cervical scrapings.

To analyse the onset of DNA methylation using the in vitro model representing the various stages of hrHPV-induced transformation, we set the highest DNA methylation ratio of each methylated gene at 100% and subsequently defined the time point of DNA methylation onset at the stage where DNA methylation was detected in at least 2 out of 4 cell lines above 5%.

In cervical scrapings, the Kruskal-Wallis test was first performed on square root-transformed DNA methylation ratios of each methylated gene to assess the differences in DNA methylation levels among disease categories. Following a significant result from the test, post hoc testing was performed using Wilcoxon rank sum test. The Bonferroni correction was subsequently applied for multiple testing, with a significance level of 0.05 (two-sided).

To obtain equivalent DNA methylation values in cervical scrapings, univariable logistic regression analysis was performed on the square root-transformed DNA methylation ratios of the 12 genes. For this analysis, scrapings from women with cervical cancer were categorised as cases and scrapings from hrHPV-positive control women (i.e. with normal cytology and/or without evidence of CIN2+) as controls. Subsequently, the logistic regression models were used to calculate the predicted probability (value range 0 to 1), representing the risk for an underlying cervical cancer. By this, equivalent values for the levels of DNA methylation were obtained for each of the individual genes allowing direct comparison of the genes. The average predicted probability of the 12 methylated genes was calculated for each sample. The lowest average predicted probability in cervical cancer was used as a threshold to define a cancer-like methylation-high pattern (i.e. threshold of ≥ 0.19). Hierarchical clustering was performed for the 12 methylated genes using the predicted probabilities.

All statistical analyses have been performed using R version 3.1.2.
RESULTS

Early onset and gradual increase in DNA methylation during hrHPV-induced transformation in vitro

DNA methylation patterns of the 12 genes (ANKRD18CP, C13orf18, EPB41L3, FAM19A4, GHSR, JAM3, PHACTR3, PRDM14, SST, ZIC1, SOX1 and ZSCAN1) were analysed in an in vitro model of hrHPV-induced transformation, using HFKs, consecutive passages of hrHPV-transformed cell lines reflecting the progressive stages of cervical precancerous disease and cervical cancer cell lines. In general, the 12 genes showed no DNA methylation in HFKs, a progressive increase in DNA methylation levels from early to late passages of hrHPV-immortalised keratinocytes and highest levels in the cervical cancer cell lines (Fig. 1). The onset of DNA methylation varied between the 12 genes from early immortal passages (ANKRD18CP, FAM19A4, GHSR, JAM3, PRDM14, SST and ZSCAN1) to late immortal passages (C13orf18, EPB41L3, PHACTR3, SOX1 and ZIC1) (Supplementary Fig. S1).

A cancer-like methylation-high pattern in a subset of cervical scrapings from women with CIN2/3

Next, DNA methylation patterns of the 12 genes in hrHPV-positive cervical scrapes, covering different underlying histology (no disease, CIN2, CIN3 and cancer) were analysed. Alike the in vitro model, a progressive and significant increase in DNA methylation levels towards cervical cancer was observed for all 12 genes ($P < 0.05$; Fig. 2).

The DNA methylation signature of the 12 genes depicted by predicted probabilities is shown in Fig. 3. All cancer cases ($n = 14$) had an average predicted probability of ≥ 0.19, which was further considered to be a cancer-like methylation-high pattern. Using this threshold, 72% (38 of 53) of CIN3 and 55% (11 of 20) of CIN2 showed a cancer-like methylation-high pattern. The remaining CIN2 and CIN3 lesions had predicted probabilities that were similar to those detected in the far majority (25 of 26) of hrHPV-positive controls, and were considered as methylation-low.

A few differences in DNA methylation patterns were observed between genes. ANKRD18CP and C13orf18 were the only 2 genes with low predicted probabilities in the single hrHPV-positive control with a cancer-like methylation-high pattern. Furthermore, C13orf18 showed either very low or very high predicted probabilities, whereas the other 11 methylated genes showed a gradual range of values. PHACTR3 had generally very low predicted probabilities in CIN2 and CIN3 lesions. This is also illustrated by unsupervised clustering, revealing a similar cluster for most genes, except for ANKRD18CP, C13orf18 and PHACTR3 (Supplementary Fig. S2).
Figure 2. DNA methylation level distributions in hrHPV-positive cervical scrapings across histological subgroups for 12 genes. Normal: hrHPV-positive controls (n = 26); CIN2 and CIN3: cervical intraepithelial neoplasia grade 2 (n = 20) and 3 (n = 53); (mi)Ca: micro-invasive cervical cancer (n = 14). * P < 0.05; ** P < 0.01; *** P < 0.001; NS: not significant.

Figure 3. DNA methylation signature of 12 genes in hrHPV-positive cervical scrapings. Predicted probabilities are shown for 12 genes in the different histological subgroups and coloured from green (predicted probability of 0) to red (predicted probability of 1). In each group, samples are ordered based on their average predicted probability (Avg P). The black line indicates the cut-off for a cancer-like methylation-high pattern at ≥ 0.19.
DISCUSSION

This study systematically analysed the onset and DNA methylation pattern of 12 genes during hrHPV-induced carcinogenesis using consecutive passages of hrHPV-immortalised cells, reflecting the progressive stages of cervical precancerous disease, complemented with clinically annotated specimens. A progressive increase in DNA methylation levels represented a general phenomenon during hrHPV-induced carcinogenesis. All genes showed increased DNA methylation levels during hrHPV-induced transformation in vitro and in clinical specimens proportional to disease severity. Highest DNA methylation levels were consistently seen in cervical cancer cell lines and cervical scrapings from women with cervical cancer. Using consecutive stages of hrHPV-transformed cells, we showed that the onset of DNA methylation occurred either at the early immortal stage or late immortal stage, when cells are still premalignant. In a series of hrHPV-positive cervical scrapings, we showed that the majority of genes displayed a comparable DNA methylation pattern with an onset of a cancer-like methylation-high pattern being detected at the CIN2/3 stage.

Interestingly, the hrHPV-positive scrapings from women with CIN2/3 displayed a heterogeneous DNA methylation pattern, in which three-quarters of the CIN3 samples and half of the CIN2 samples had a cancer-like methylation-high pattern. These results correspond to the finding that only a subset of CIN2/3 will progress to cancer over a long time period\(^\text{6,11}\). Previous studies on copy number changes and DNA methylation levels of only a few genes also showed a cancer-like profile in only a subset of CIN2/3 lesions\(^\text{27,28}\). These lesions were characterised by a preceding hrHPV infection of ≥ 5 years and considered as more advanced lesions, which have a high short-term progression risk to cancer. Conversely, the methylation-low CIN2/3 may be more likely to regress, knowing that approximately ~30% of CIN3 and ~40%-54% of CIN2 regress\(^\text{22,23,26}\).

Most of the genes evaluated in this study were reported to have a (putative) tumour suppressive function (i.e. C13orf18\(^\text{27}\), EPB41L3\(^\text{28-32}\), PRDM14\(^\text{33}\), SOX1\(^\text{34,35}\), SST\(^\text{36}\), ZIC1\(^\text{38,39}\) and ZSCAN4\(^\text{40}\)). The remaining genes with a currently unknown function await further study.

A few genes (ANKRD18CP, C13orf18 and PHACTR3) showed a different DNA methylation pattern compared with the majority of genes. This may play a role in different onset of DNA methylation during cervical carcinogenesis. DNMTs can be activated by both hrHPV E6 and E7\(^\text{41,42}\) and specific promoter sequences might be more prone to DNA methylation compared with others\(^\text{41,43}\), which may relate to the binding of either the polycomb repressor complex (PRC) 1 or 2\(^\text{44}\). We purposely retained these 3 genes in our analysis on DNA methylation patterns in order to obtain an unbiased representative result. In cervical scrapings, only one hrHPV-positive control showed a cancer-like methylation-high pattern with very high individual predicted probabilities for 10 of 12 methylated genes. However, no gynaecological diseases were identified in the follow-up data. Other unidentified underlying abnormalities may have resulted in increased DNA methylation levels. Considering a potential effect of age, i.e. 55 year for this control, DNA methylation levels of the markers tested were not or minimally affected by age in our previous studies using large cohorts\(^\text{8-11,45,46}\). Therefore, solely age is very unlikely to explain the cancer-like methylation-high pattern in this hrHPV-positive control.

A limitation is that the amount of dysplastic cells in the specimens is unknown and a potential effect of varying sample constitution cannot be fully excluded.

The genes evaluated in this study may well serve as objective molecular tools to improve cervical cancer screening, especially as triage test after primary hrHPV testing\(^\text{11,47}\). In several countries, including The Netherlands, hrHPV testing is replacing cytology as primary screening method. Compared with cytology, hrHPV testing has a higher sensitivity for CIN2+ detection\(^\text{16,48}\). However, its 3-5% lower specificity makes triage testing of hrHPV-positive women necessary to prevent over-referral and overtreatment. In the Dutch population-based HPV screening programme, cytology is included as triage method. However, cytology comes with some limitations, including its subjective nature and the required repeated cytology to ensure sufficient safety in the screening programme. Moreover, prior knowledge of hrHPV positivity influences cytology reading, which may result in an increase of false-positive referrals with simultaneously higher costs for the healthcare system\(^\text{10-42}\). Recent clinical validation studies in screening populations have shown that DNA methylation markers provide a good alternative for cytology\(^\text{14,33-37}\).

In conclusion, this study showed that hrHPV-induced carcinogenesis is characterised by increased DNA methylation, with onset typically occurring at the premalignant stage and highest DNA methylation levels at the cancer stage. Host cell DNA methylation patterns of the 12 genes are comparable and reveal the heterogeneous nature of cervical precancer, with a subset of CIN2 and CIN3 lesions displaying a cancer-like methylation-high pattern, suggestive for a higher risk of progression to cervical cancer.
ACKNOWLEDGEMENTS

We thank Annina van Splunter for excellent technical assistance.

REFERENCES


**SUPPLEMENTARY FIGURES**

**Supplementary Figure S1.** Onset of DNA methylation of 12 genes. Presented is relative DNA methylation calculated as percentage of the DNA methylation ratio, with the highest ratio per gene set at 100%. All values above 5% are shown in red. Onset of DNA methylation was determined at the stage with at least 2 out of 4 cell lines above 5%. HFK: human foreskin keratinocytes; Early: early immortal passages of FK16A/B and FK18A/B p30-p70; Late: late immortal passages of FK16A/B and FK18A/B > p70; CCL: cancer cell lines.

**Supplementary Figure S2.** Cluster dendrogram of the DNA methylation pattern of the 12 genes. Clustering is performed based on the predicted probabilities of the 12 methylated genes in 113 hrHPV-positive cervical scrapings using complete linkage and the Manhattan distance, which do not include a root. The Manhattan distance indicates the absolute distance between the markers.