Characterization of the Methylation Patterns in Human Papillomavirus Type 16 Viral DNA in Head and Neck Cancers

Il-Seok Park1,2, Xiaofei Chang1, Myriam Loyo1, Gaosong Wu1, Alice Chuang1, Myoung Sook Kim1, Young Kwang Chae3, Sofia Lyford-Pike1, William H. Westra4, John R. Saunders5, David Sidransky1,6, and Sara Isabel Pai1,6

Abstract

Human papillomavirus (HPV) type 16 can integrate into the host genome, thereby rendering the viral coding genes susceptible to epigenetic modification. Using bisulfite genomic sequencing, we determined the methylation status of all 110 CpG sites within the viral epigenome in advanced stage III/IV HPV-16–associated head and neck cancers. We found that the viral genome was hypomethylated in the majority of head and neck cancers, in particular within the viral regulatory region, long control region (LCR), which controls transcription of the E6 and E7 oncogenes. The hypomethylation status of LCR correlated with detectable levels of E6 and E7 expression, which suggests that the tumors may still be dependent on these viral oncogenes to maintain the malignant phenotype. In addition to the methylation status of LCR, we report other potential factors which may influence intratumoral E6 and E7 expression including viral copy number and integration site. We were able to detect the viral epigenetic alterations in sampled body fluids, such as serum and saliva, which correlated with the changes observed in the primary tumors. Because viral epigenetic changes occur in the setting of viral integration into the human genome, the detection of methylated HPV genes in the serum and/or saliva may have diagnostic potential for early detection strategies of viral integration and assessment of risk for cancer development in high-risk individuals. Our findings also support continued targeting of the E6 and/or E7 antigens through various vaccine strategies against HPV-associated cancers. Cancer Prev Res; 4(2); 207–17. ©2011 AACR.

Introduction

Epigenetics describes the regulation of gene expression through heritable changes in DNA methylation and chromatin structure. DNA methylation can impact the transcription of genes by either physically impeding the binding of transcriptional proteins to the gene and/or by changing the chromatin structure to repress transcription. We now know that epigenetics plays an important role in tumorigenesis in mammals (1). DNA methylation occurs in cytosines (5-methylcytosine) that precede guanines in dinucleotide CpG sites. CpGs are asymmetrically distributed into CpG-poor regions and dense regions called “CpG islands,” which are located in the promoter regions. These CpG islands are usually unmethylated in normal cells, whereas the sporadic CpG sites in the rest of the genome are generally methylated (2). Methylation of CpG islands in promoter regions is often associated with gene silencing, and aberrant DNA methylation can occur in cancers, leading to the silencing of tumor suppressor genes (1).

Several oncogenic viruses can integrate within the host genome and become susceptible to modification by the host epigenetic machinery and, at times, can utilize the machinery to regulate its own viral gene expression. One such virus is the human papillomavirus (HPV). HPV type 16 (HPV-16) is the most common virus type associated with cervical and head and neck cancer and is present in greater than 90% of HPV-associated head and neck squamous cell carcinomas (HPV-HNSCC; ref. 3). After viral entry into a cell, episomal HPV-16 DNA can integrate into the host genome with resultant deletion of noncritical and regulatory viral genes. Late genes (L1 and L2) and some early genes (E1 and E2) are commonly deleted, and the viral oncogenes E6 and E7 are often the only open reading frames consistently expressed in cancer cell lines (4) and in primary HPV-associated cancers (5).
To evaluate the role of epigenetic alterations in HPV-induced carcinogenesis, we mapped all potential DNA methylation sites in the HPV-16 epigenome in patients with head and neck cancer. Because the long control region (LCR) is a key regulatory site for viral gene expression, we focused our analysis in this region and evaluated the methylation status of key inhibitory and activating transcriptional sites. Furthermore, we correlated the E6 and E7 expression levels to the methylation status of LCR, and also evaluated other potential factors which may influence E6 and E7 expression such as viral copy number and integration status. We explored whether epigenetic alterations within the viral epigenome could be detected in body fluids, such as serum and saliva, which can have diagnostic significance for early detection strategies of virus integration as well as assessment of risk for cancer development in healthy individuals infected with the virus.

Materials and Methods

Cell lines, tissues, sera, and saliva

Cervical cancer cell lines, CaSki (~600 copies of integrated HPV-16 DNA per genome) and SiHa (2 copies of integrated HPV-16 DNA per genome) were purchased from ATCC (American Type Culture Collection) and cultured as per manufacturer’s instructions. Primary tumor tissue, serum, and saliva were collected from patients with HPV-positive, oropharyngeal squamous cell carcinoma (OPSCC) as determined by in situ hybridization with an HPV-16 specific probe. Primary tumor was obtained from 22 patients with advanced stage III/IV HPV-associated OPSCC. Tissue was microdissected to separate tumor from stromal elements, yielding at least 80% cancer cells. Tissue blocks were stained with hematoxylin and eosin, and tumor areas were subsequently outlined, cut by a head and neck pathologist, and processed for DNA and RNA extraction. RNA was available and extracted from nine patients. Matched saliva samples were available from 9 patients prior to any therapy. Oral rinsing was performed by gargling twice in 20 mL of saline. Matched serum samples were available from 10 patients prior to any therapy. This study was approved by the Institutional Review Board of the Johns Hopkins University and written informed consent was obtained from all patients.

DNA and RNA extraction

Tissue samples were centrifuged and digested in a solution of SDS and proteinase K at 50°C overnight. Genomic DNA (gDNA) was isolated by phenol/chloroform extraction, and precipitated in ethanol. The DNA pellet was resuspended in TE buffer (EDTA, 2.5 mmol/L; Tris-HCl, 10 mmol/L) and stored at −20°C. Total RNA was extracted from frozen tumor tissue using Trizol (Invitrogen). The RNA integrity was assessed by agarose gel electrophoresis.

Bisulfite treatment

Bisulfite conversion of gDNA was performed using the EpiTect Bisulfite Kit (Qiagen) as per manufacturer’s recommendations. This bisulfite-modified DNA was subsequently resuspended in 120 μL of TE buffer and stored at −80°C until use.

Bisulfite sequencing

The modified DNA was amplified in the form of 18 amplicons (Supplementary Table 1). Bisulfite-modified gDNA was amplified by PCR using 10× buffer [166 mmol/L (NH₄)₂SO₄, 670 mmol/L Tris Buffer (pH 8.8), 67 mmol/L MgCl₂, 0.7% 2-mercaptoethanol, 1% DMSO] and primer sets that were designed to recognize DNA bases after bisulfite treatment. The conditions for PCR amplifications were as follows: a 5-minute incubation at 95°C was followed by 45 cycles of 1 minute at 95°C, 1 minute at 54°C, and 2 minutes at 72°C. A 7-minute elongation step at 72°C completed the PCR amplification. For amplification of some amplicons, touchdown PCR was performed as follows: a 5-minute 95°C incubation step was followed by 2 cycles of 1 minute at 95°C for denaturation, 1 minute at 66°C for annealing, and 1 minute at 72°C for elongation. The annealing temperature was decreased by 2°C, and two PCR cycles were run each time until the annealing temperature was 56°C. The PCR was run for 35 cycles, with an additional 7 minutes at 72°C at the completion of the cycles for further elongation. PCR products were gel extracted (Qiagen) and sequenced with forward and reverse primers using the ABI BigDye cycle sequencing kit (Applied Biosystems).

Methylation-specific PCR

Bisulfite-treated DNA was amplified with methylation- and unmethylation-specific (UnMSP) primer sets for each individual gene. Primer sequences are shown in Supplementary Table 2. PCR reactions were performed for 35 cycles of 95°C for 30 seconds, 58°C for 30 seconds, and 72°C for 1 minute.

Quantitative PCR

To quantify the viral load of HPV-16, gDNA from patients was used for real-time PCR with primers and probes sets specific for the E6 and E7 region of HPV-16 DNA. PCR for β-actin was performed in parallel to normalize the input DNA. gDNA from the CaSki cell line was used to develop standard curves for the HPV-16 viral load, as it has been previously characterized to harbor 600 copies of HPV-16 DNA per genome equivalent. gDNA from CaSki cells was serially diluted to the following concentrations: 50, 5, 0.5, 0.05, and 0.005 ng. A standard curve was developed for β-actin which has 2 copies per genome equivalent, using the same serial dilutions. All samples were run in triplicate. Taqman Fast Universal PCR Master Mix was used according to the manufacturer’s instructions (Applied Biosystems). Primer and probe sequences are shown in Supplementary Table 3.

To evaluate the viral integration status, we used a previously described real-time quantitative PCR assay (6). E2 and E6 primers and probes were synthesized using published sequences (6). The final primer and probe
concentrations were 0.3 and 0.1 μmol/L, respectively, in a total volume of 10 μL. A standard curve was obtained by amplification of a 10-fold dilution series of 0.3 to 0.0003 ng of a HPV-16 plasmid clone (with a ratio of E2:E6 = 1:1). At least 3 water controls were included in each run. All experiments were performed twice in duplicate with similar ratios. The integration status of HPV-16 DNA was assessed by comparing the levels of detected HPV-16 E2 and E6 genes and was expressed as an E2:E6 ratio. Ratios of E2:E6 of less than 1 indicate the presence of both integrated and episomal forms. The ratio of E2:E6 represents the amount of the episomal form in relation to the integrated form.

Quantitative RT-PCR
To examine the mRNA expression of E6 and E7, cDNA was synthesized from 1 μg of total RNA which was isolated from patients’ samples, treated with DNase I, and subsequently cleaned as recommended by the manufacturer (Qiagen). For cDNA preparation, 1 μg of total RNA was transcribed with random hexamers and oligodT using the Superscript II reverse transcriptase (Invitrogen). The same amount of RNA was processed without reverse transcriptase (RT) in the cDNA synthesis to assure that no gDNA was amplified in the reaction. cDNA from the CaSki cell line was used to develop standard curves for the E6 and E7 transcripts. β-Actin was used as a loading control. All samples were run in triplicate. To examine the expression levels of E6 and E7, the mRNA levels of E6 and E7 were separately divided by the viral load within the tumor and multiplied by 10,000.

Statistics
Pearson’s product–moment correlation coefficients and Spearman’s rank correlation coefficients were used to assess the linearity and the rank association between methylation and expression levels. P value of 0.05 was used to assess the significance of the association.

Results
The HPV-16 epigenome is hypomethylated in advanced stage HPV-related head and neck cancers
The genome and epigenome map of HPV-16 DNA is shown in Figure 1A and B. We identified 110 CpG sites within the HPV-16 viral methylome. We evaluated the methylation status of all 110 sites in two established invasive cervical cancer cell lines (CaSki and SiHa) and in 22 patients with advanced stage III/IV HPV-16–associated OPSCC by performing bisulfite-sequencing analysis.

We found that the CaSki cell line harbored dense CpG methylation throughout the entire HPV-16 epigenome. Specifically, 94% of the genome was methylated and 5% of the genome was unmethylated in the E1 and E2 regions. In contrast, SiHa was methylated in 35% of the viral epigenome and 62% of the genome was unmethylated (Fig. 2).

We evaluated the methylation status of the HPV-16 viral epigenome in 22 primary advanced stage head and neck cancers which consisted of one patient with stage III disease and 21 patients with stage IV disease. Interestingly, we found that the methylation pattern in primary head and neck cancers was comparable with SiHa. Specifically, 12 of the 22 (54.5%) advanced stage OPSCC were unmethylated in greater than 50% of the viral DNA CpG sites with areas lacking methylation clustered in the E2, LCR, and E6 regions. Five of the 22 (22.7%) OPSCC were unmethylated in less than 10% of the HPV-16 epigenome with areas lacking methylation clustered within LCR (Fig. 2).

The LCR region is preferentially hypomethylated in advanced stage head and neck cancers
The LCR contains the origin of replication (ori), the E6 enhancer as well as the E2-binding site (E2BS; refs. 6–8). Because this region is critical in transcriptional regulation of the viral genome, we were interested in further evaluating the methylation status within this region. In the HPV
cell lines, LCR was 100% unmethylated in SiHa and 100% methylated in CaSki. In advanced stage head and neck cancers, we found that 21 of 22 (95.5%) of the advanced stage tumors were unmethylated at or greater than 50% of the CpG sites within LCR. Interestingly, one tumor (sample 3), which was methylated in greater than 90% of the whole viral genome, was unmethylated in 40% of the CpG sites within LCR, suggesting preferential hypomethylation within this viral gene regulatory region (Fig. 3A).

Both the E6 enhancer and the E2BS within LCR regulate transcription of the viral oncoproteins, E6 and E7. Therefore, we evaluated whether there was preferential methylation in either of these sites within LCR. The SiHa cell line was 100% unmethylated in both the E6 enhancer and the E2BS and the CaSki cell line was 100% methylated in both of these sites. We found that 19 of 22 (86%) of the advanced stage III/IV OPSCC were unmethylated in greater than 50% of the CpG sites within the E6 enhancer region (Fig. 3B). Furthermore, 74% (14/19) of these tumors were 100% unmethylated in this region. For the E2-binding region, 16 of 22 (73%) of the tumors were unmethylated in greater than 50% of the CpG sites and 88% (14/16) of these tumors were 100% unmethylated within this region. Both the E6 enhancer and E2-binding region were unmethylated in greater than 50% of the CpG sites in 80% of the advanced stage cancers.

Nine of 22 (41%) of the OPSCC demonstrated methylation in greater than 50% of the CpG sites at either the E6 enhancer region or the E2BS (Fig. 3B and C). Interestingly, the tumors which were heavily methylated in the E2BS (samples 3, 16, 5, 17, 19, and 4) were unmethylated in the E6 enhancer region and those tumors which were heavily methylated in the E6 enhancer region (samples 2, 7, and 9) were unmethylated in the E2BS. Our results demonstrate that the E6 enhancer region and/or the E2BS are hypomethylated in the majority of advanced stage OPSCC.
The methylation status within LCR correlates with E6 and E7 expression levels

LCR contains the E6 enhancer and the E2BS both of which regulate viral E6 and E7 expression. Therefore, we evaluated whether the methylation status of LCR correlated with the expression levels of E6 and E7 in the cell lines as well as 9 primary tumor samples for which RNA was available. E6 and E7 expression levels were normalized to β-actin. The CaSki cell line, which was 100% methylated in LCR, expressed 16 copies of E6 per viral load (Fig. 4A and B). SiHa, which was 100% unmethylated in LCR, expressed 361 copies of E6 per viral load. For the primary head and neck tumor samples, the majority of the cancers demonstrated a methylation pattern comparable with SiHa and were unmethylated in greater than 50% of the CpG sites in LCR. E6 expression levels in the primary samples ranged from 7 to 2,220 copies of E6 per viral load. In 78% (7 of 9) of these tumors, E6 expression ranged between 51 to 282 copies per viral load which is a higher level than CaSki but lower than SiHa (Fig. 4B). One patient (15) expressed 2,220 copies of the E6 per viral load and one sample (14) expressed 7 copies of the E6 per viral load. Interestingly, both of these patient samples were unmethylated in 93% to 100% of the CpG sites in LCR. Because the E2BS also regulates E7 expression, we performed the same analysis with E7 expression and found similar correlations between methylation status and expression (Fig. 4A and Supplementary Fig. 1).

Although there was a trend between methylation status of LCR and E6 and E7 expression, there was no statistical significance using Pearson correlation coefficients or Spearman’s rank correlation coefficients. The lack of statistical significance may be attributed to the low number of available samples (N = 9) which could be evaluated.
However, other factors may also be contributing to E6 and E7 expression and we explored these other potential confounding factors.

**There is a variable viral load within head and neck cancers**

It has been previously published that the CaSki cell line harbors between 500 and 600 copies of integrated HPV-16 DNA in more than 11 chromosomal sites of the hyperdiploid genome. However, the CaSki cell line has only one active papillomaviral transcriptional center per cell that maps to a low tandem copy integration site at chromosome 14. SiHa has 2 integrated copies of HPV-16 DNA at the 13q21 locus of the homologous chromosomes and both viral copies are transcriptionally active. On the basis of this observation, we evaluated the viral load in the primary cancers to determine if the viral copy number could influence the expression levels of E6 and E7 in the primary tumors. The CaSki cell line served as our reference for calculating the viral copy number in our tumors. In our study, we found that CaSki contained 613 copies of HPV-16 DNA per genome and SiHa 6 copies of HPV-16 DNA per genome. 78% (7/9) of the primary head and neck cancers contained integrated viral copy numbers which ranged between 24 and 283 copies of HPV-16 DNA per genome (Supplementary Fig. 2). Outliers included patient sample 15 which harbored 3 copies of HPV-16 DNA and patient sample 14 which harbored 1,866 integrated viral copies.

**A mixture of episomal and integrated forms of the virus is present in OPSCC**

HPV DNA can exist in multiple forms within the genome, including integrated and episomal forms. Therefore, we assessed the ratio of integrated to episomal forms of the viral DNA to determine if this factor could affect E6 and E7 expression, independent of methylation status. Current assays measuring HPV-16 integration are based on quantification with real-time PCR of HPV-16 E6 relative to E2 DNA because the E2 gene is often lost during the viral integration process (6). Therefore, detection of a greater quantity of HPV-16 E6 compared with E2 suggests the presence of integrated HPV-16 DNA. We were able to assess the viral integration status in 20 of 22 of the OPSCC cases based on the availability of the tumor samples (Fig. 5). We found there were 8 of 20 (40%) tumors which contained an E2:E6 ratio greater than 1.0, suggesting the predominance of the episomal forms of the virus, within the tumor (Fig. 5). There were 7 of 20 (35%) tumors with an E2:E6 less than 0.5, suggesting predominance of the integrated forms of the virus. Despite the lack of episomal forms of the virus within these tumors, there was significant hypomethylation detected within LCR (Figs. 3 and 5), suggesting that there is preferential hypomethylation within this region in integrated forms of the viral DNA.

**Methylation status within LCR is detected in serum and saliva**

We evaluated whether the HPV-16 DNA methylation status of LCR was detectable in the serum and saliva of OPSCC patients. Therefore, we performed methylation-specific PCR (MSP) with primers designed in LCR of HPV-16 DNA. Using these primers, we were able to detect a methylated allele in CaSki and an unmethylated allele in SiHa. An unmethylated DNA allele was observed in all of the serum and saliva samples from OPSCC patients tested except for case 7 in which greater than 50% of the CpG sites were methylated in this region (Figs. 6 and 3). Within the primary tumor 3, LCR was methylated in 60% of the CpG sites and both unmethylated and methylated alleles were detected in the serum and saliva using this assay (Figs. 6 and 3). Our results demonstrate that the LCR methylation status within the primary tumor is detectable in the saliva and serum of advanced stage HPV-16–associated OPSCC and can serve as a biomarker for viral DNA integration.

**Discussion**

The HPV-16 genome consists of an 8-kb circular, double-stranded DNA, which encodes 6 early expressed regulatory genes (E1, E2, E4–E7) and 2 late expressed structural genes, L1 and L2. These two sets of viral coding regions are separated by an LCR that contains the ori, P97, and non-coding cis-elements such as the E2BS, E6 enhancer, and promoter regions (7–9). Binding of the E2 protein to the E2BS inhibits transcription factors from docking onto LCR and, thus, represses the transcription of the viral oncogenes, E6 and E7 (10, 11). Expression of the E6 and E7 proteins are essential to HPV-mediated transformation due...
to their binding and inhibition of the cellular gatekeepers, p53 and Rb proteins, which can result in dysregulation of a variety of cellular processes, including cell-cycle control (12). The integration of the viral DNA into the chromosome of mammalian cells, as well as certain genetic or epigenetic alterations of the viral genome, can lead to loss of E2-mediated inhibition of E6 and E7 expression, which is a key event in papillomaviral carcinogenesis.

Integration of the viral DNA into the host genome, although requisite for carcinogenesis, makes it vulnerable to modification by the host’s methylation machinery. This phenomenon is highlighted when cultured primary human

Figure 5. Assessment of viral integration. The ratio of E2 to E6 was measured for 20 HPV-16–associated OPSCC using real-time quantitative PCR. An E2:E6 ratio of 0.5 indicates a mixture of integrated and episomal forms of detected HPV DNA. Values greater than 1 indicate predominance of the episomal form. Numbers represent the patient samples.

Figure 6. MSP analysis of the HPV-16 DNA. The methylation status of LCR in serum (A) and saliva (B) was examined by MSP after PCR with MSP and UnMSP primers. Bisulfite-treated DNA from CaSki and SiHa were used as controls. β-Actin was used to confirm the integrity of the bisulfite-treated DNA. Tumor sample 3 contained both methylated and unmethylated forms of viral DNA which was detected in this assay in both the serum and saliva. Tumor sample 7 contained greater than 70% methylation within LCR which was detected as a single methylated band in the serum and saliva. C, CaSki cell line; S, SiHa cell line. Numbers represent the patient samples.
foreskin keratinocytes are transfected with HPV-16 DNA (13, 14). In the preimmortal keratinocytes, the episomal HPV-16 DNA remains in an unmethylated state. However, the immortal descendant cells with integrated HPV-16 DNA acquire a densely methylated viral genome. Several hypotheses exist to explain this observation both in vitro and in vivo. Some groups suggest that the host cell may actively be silencing the foreign genome as a cellular defense mechanism to inhibit the expression of nonself proteins which can disrupt normal cellular function. Other groups have suggested that the HPV-16 genome may not be a passive spectator in this process, but might actively participate by recruiting DNA methyltransferases (DNMT) and/or histone deacetylases (HDAC) via the viral oncoprotein, E7, to strategically regulate viral gene expression during the viral life cycle to evade immune recognition by the host (15, 16).

In human cancers, methylation of oncogenic viral DNA may occur due to a combination of these factors. DNA methylation impacts the transcription of genes by either physically impeding the binding of transcriptional proteins to the gene and/or altering the chromatin structure to repress transcription. We know that alterations in methylation patterns play an important role in tumorigenesis in mammals (1). A hallmark of cancer is a paradoxical aberration of normal DNA methylation patterns, with a global loss of DNA methylation that coexists with regional hypermethylation of certain genes. For example, the hypermethylation of tumor suppressor genes has attracted significant attention, and DNA methylation inhibitors are being tested as potential anticancer agents. However, emerging data also suggest that hypomethylation can play a role in activating pro-oncogenic genes which may be required for metastasis and invasion. It has been proposed that hypermethylation and hypomethylation in cancer are independent processes, which target different cellular programs at different stages of cancer development. Therefore, evaluating the methylation status at specific sites, as opposed to a more global genomic assessment, may yield more valuable and predictive information regarding cancer progression.

This concept is exemplified in studies performed with the HPV-16 methylome in cervical lesions. Fernandez and colleagues evaluated the DNA methylome of the HPV-16 virus in a collection of human cervical samples at different progressive stages of disease (17). They found that the DNA methylome evolved from an unmethylated to a highly methylated state in association with disease progression, from asymptomatic healthy carriers, through chronically infected tissues and premalignant lesions, to the development of invasive cervical cancer. Thus, the progression to cancer was associated with increasing numbers of methylated CpGs and increasing proportions of methylated HPV-16 genomes, which has been supported by other studies (18–20).

However, other groups have evaluated the methylation status of specific regions within the HPV-16 methylome, which has resulted in conflicting data. Badal and colleagues evaluated 81 patients with HPV-16–associated cervical lesions and they found that the LCR and E6 genes of HPV-16 DNA were hypermethylated in 52% of asymptomatic smears, 21.7% of precursor lesions, and 6.1% of invasive carcinomas (18). Bisulfite modification and sequencing analysis revealed that in most of the HPV-16 genomes of the CaSki cell line and asymptomatic patients, all 11 CpG dinucleotides that overlap with the enhancer and the promoter regions in LCR were methylated, whereas in the SiHa cells and cervical lesions, the same subset of CpGs remained unmethylated (18). These results suggest that neoplastic transformation may be suppressed by CpG methylation, whereas hypomethylation seems to correlate with neoplastic progression. In contrast, Kalantari and colleagues evaluated 115 cervical samples to establish the methylation patterns of 19 CpG dinucleotides within LCR and the L1 gene by bisulfite modification and sequencing and reported that methylation of most sites was highest in carcinomas and methylation was lowest in dysplasia (19).

The discrepancy between these studies may be attributed to the differing methylation sites evaluated within LCR. LCR has binding sites for both inhibitory and activating transcription factors. For example, transcription starts at the E6 promoter, P97, which is regulated by one binding site for Sp1 and two binding sites for the inhibitory HPV-encoded E2 protein (21, 22). The activity of P97 is stimulated by an enhancer with binding sites for several cellular factors including AP1 (23), NF1 (24), and the progesterone receptor (25). Other factors which have been reported to be able to bind the LCR region are HIV tat-1 (26), YY1 (27), Octa (28), and TEF-2 (28). In addition, two specifically positioned nucleosomes can form over the enhancer and promoter regions (29) to repress transcription when they are modified by HDACs. Depending upon which of these sites one analyzes, correlations between methylation status and tumor progression might be directly or inversely related.

Because of the wide variance in the literature regarding HPV-associated cervical lesions, we were interested in characterizing the methylation status of the HPV-16 viral genome in HPV-associated head and neck cancers (HPV-HNSCC). We performed bisulfite sequencing analysis of all 110 CpG sites within the HPV-16 genome in 22 patients with advanced stage III/IV HPV-16–associated HNSCC. In contrast to the cervical lesions, we found that the majority of these advanced stage patients had significant hypomethylation of the viral epigenome, especially within LCR which is a regulatory region for viral oncogene expression.

Because of the various regulatory sites within LCR, we further investigated this control region by focusing on the methylation status of the inhibitory E2BS as well as the activating E6 enhancer region. We found that the E6 enhancer region was hypomethylated in advanced stage head and neck cancers, which facilitates binding of transcription factors to this region and, thus, allows for overexpression of the viral oncogenes. Interestingly, we found that the E2BS was also significantly hypomethylated.
Previous studies reported that methylation of CpG dinucleotides within the binding site of HPV-16 E2 protein in LCR can directly inhibit the binding of E2 to the cognate DNA sequences in vitro; therefore, hypomethylation in this region would facilitate binding of E2 to the E2BS and, subsequently, inhibit E6 and E7 expression, which is contrary to what one would expect in advanced stage cancers. However, an explanation for our findings might be found when reviewing the events which occur upon viral integration into the host genome. Specifically, the E2 gene is often disrupted or lost upon viral DNA integration. Without the inhibitor influence of the E2 protein, the selective pressure to methylate this site is lost, which is consistent with our findings of hypomethylation of this region in the primary tumors (30).

We found a trend between the methylation status of LCR, E6 enhancer, and E2BS and levels of E6 and E7 expression within the tumors. However, a statistical significance was not found which could be attributed to the low number of tumor samples available for the analysis. In addition, other potential factors may influence E6 and E7 expression levels including the integrated viral copy number. Experimental transformation of keratinocytes with HPV-16 reveals a consistent tendency toward reducing the number of transcriptionally active HPV genomes to 1 or 2 during the passages and reactivation of the silent viral copies can occur in the presence of the DNA methylation inhibitor 5-azacytidine (14). The silencing of the redundant copies may be advantageous as well as critical in clonal selection during carcinogenesis. One could speculate that limiting the number of actively transcribed viral oncogenes can prevent the accumulation of excess genomic instability of the host genome that could affect the growth and survival of the transformed cell. Our findings suggest this exact phenomenon may be occurring in vivo. We found that the majority of head and neck tumors harbored between 24 and 283 integrated viral copies per genome and expressed between 51 and 282 copies of E6 per viral load and 5 and 249 copies of E7 per viral load. However, we found one patient who harbored 1,866 copies of HPV-16 DNA, but expressed only 7 copies of E6 per viral load and 5 copies of E7 per viral load. In contrast, a patient who harbored 3 copies of HPV-16 DNA expressed 2,220 copies of E6 per viral load and 1,530 copies of E7 per viral load. We can hypothesize that cancers which contain a low number of viral genomes must keep them in an unmethylated form for transcribed viral oncogenes to maintain cell survival. Support for this hypothesis is found in a study by Cohen and colleagues, which evaluated the viral load in 35 HPV-16–associated HNSCC. They found that the patients with the highest viral loads had an improved overall and disease-free survival (31). This finding is counterintuitive; however, based on our study, we propose that those tumors with the highest viral loads could have significant silencing of the redundant copies through methylation and correspondingly low E6 and E7 expression levels, which could account for improved survival rates.

An alternative factor which may influence E6 and E7 expression, in addition to viral copy number and viral methylation status, may be the site of integration of the viral genome, which may be permissive to either high or low levels of viral gene transcription. Over 200 selected HPV-16 and HPV-18 integration sites have been reported. These are widely distributed across the genome; however, there seems to be preferential integration near common fragile sites (CFS), specific chromosomal loci that are particularly prone to forming double-strand breaks (32–34). It has been reported that CaSkii harbors between 500 and 600 integrated copies of HPV-16 DNA. CaSkii was found to be significantly methylated with only one active papillomaviral transcriptional center per cell, which maps to a low tandem copy integration site at chromosome 14. SiHa is significantly unmethylated and has 2 integrated copies of HPV-16 DNA at the 13q21 locus of the homologous chromosomes and both viral copies are transcriptionally active. Therefore, for patients with low viral copy numbers, such as our patient with 3 viral copies, it may have integrated into a site which is permissive for the high levels of E6 and E7 expression.

A limitation to our study is that the current methodology does not allow the selective detection of the methylation pattern for the transcriptionally active copies of the viral genome among the inactive ones. Rather, our assay reflects the cumulative status of all of the viral genomes, both active as well as inactive, within the tumor. Despite this limitation, our data demonstrate that we are able to detect certain regions within the viral epigenome that are more likely to be methylated or unmethylated and there is a trend toward hypomethylation of LCR with corresponding detectable levels of E6 and E7 expression in advanced stage OPSCC. Another limitation to the study is the lack of evaluation of early stage disease which would allow us to determine how the HPV-16 DNA methylation status may evolve with progression of disease and/or stage. Because HPV-HNSCC localizes to the tonsil and base of tongue which are areas of the head and neck that are more difficult to routinely evaluate without a directed physical examination, HPV-HNSCC patients are typically diagnosed after lymphatic spread to the cervical nodes, resulting in a diagnosis at an advanced stage of disease. Thus, we were not able to evaluate the HPV-16 methylation status of early stage I or II HPV-HNSCC because patients rarely present with these early stage lesions and tissue was not available for analysis.

The implications of our findings are several folds. We report that HPV-related head and neck cancers have regulated mechanisms of methylation of the viral epigenome. Our findings demonstrate that LCR is preferentially hypomethylated in the majority of advanced stage HPV-16–associated head and neck cancers. Hypomethylation of LCR corresponds to detectable expression levels of E6 and E7 within head and neck cancers, supporting the feasibility of targeting these antigens for novel
immunotherapeutic strategies (35–36). We demonstrate the feasibility of detecting methylated HPV genes in body fluids such as serum and saliva. Because viral epigenetic changes occur only in the setting of viral integration into the human genome, the detection of methylated HPV genes in the serum and/or saliva may help to identify the presence of viral integration as compared with episomal forms and can serve as a biomarker for HPV integration and may potentially allow assessment of risk for cancer development in high-risk individuals. Further studies which distinguish preferential sites of methylation within the HPV epigenome may be relevant and complement current assays for HPV integration.

References


Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Grant Support

NIH P50 CA19032 (W.H. Westra, D. Sidransky, S.J. Pai) and The Milton J. Dance, Jr., Head and Neck Center at the Greater Baltimore Medical Center (J.R. Saunders, S.J. Pai).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received June 30, 2010; revised December 9, 2010; accepted December 16, 2010; published online February 3, 2011.


Characterization of the Methylation Patterns in Human Papillomavirus Type 16 Viral DNA in Head and Neck Cancers

Il-Seok Park, Xiaofei Chang, Myriam Loyo, et al.


Updated version Access the most recent version of this article at: http://cancerpreventionresearch.aacrjournals.org/content/4/2/207

Supplementary Material Access the most recent supplemental material at: http://cancerpreventionresearch.aacrjournals.org/content/suppl/2011/04/14/4.2.207.DC1

Cited articles This article cites 34 articles, 19 of which you can access for free at: http://cancerpreventionresearch.aacrjournals.org/content/4/2/207.full.html#ref-list-1

Citing articles This article has been cited by 3 HighWire-hosted articles. Access the articles at: /content/4/2/207.full.html#related-urls

E-mail alerts Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.