ARTICLE

Methylation of HPV18, HPV31, and HPV45 Genomes and Cervical Intraepithelial Neoplasia Grade 3

Nicolas Wentzensen,* Chang Sun,* Arpita Ghosh, Walter Kinney, Lisa Mirabello, Sholom Wacholder, Ruth Shaber, Brandon LaMere, Megan Clarke, Attila T. Lorincz, Philip E. Castle, Mark Schiffman, Robert D. Burk

Manuscript received November 25, 2011; revised August 16, 2012; accepted August 31, 2012.

*These authors contributed equally to this manuscript.

Correspondence to: Nicolas Wentzensen, MD, PhD, MS, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Blvd, Room 5024, Rockville, MD 20852-7234 (e-mail: wentzenn@mail.nih.gov) and Robert Burk, MD, Albert Einstein Medical College, 1300 Morris Park Ave, Bronx, NY 10461 (e-mail: burk@aecom.yu.edu).

Background Persistent infections with carcinogenic human papillomavirus (HPV) types are the necessary cause of cervical cancer. We recently demonstrated that the HPV16 genome is strongly methylated in cervical precancer compared with transient infections. However, the extent of methylation in other HPV types and its role in progression to cancer is poorly understood.

Methods We analyzed whole-genome methylation patterns of the three next most carcinogenic HPV genotypes: HPV31 (closely related to HPV16), and two other closely related types, HPV18 and HPV45. DNA was extracted from cervical cytology specimens from 92 women with precancer and 96 women infected with HPV31, HPV18, or HPV45, but who had no cytological or histological abnormalities. After bisulfite modification, genome-wide pyrosequencing was performed covering 80–106 sites. We calculated differences in median methylation, odds ratios, areas under the curve, and Spearman rank correlation coefficients for methylation levels between different sites. All statistical tests were two-sided.

Results For all three HPV types, we observed strongly elevated methylation levels at multiple CpG sites in the E2, L2, and L1 regions among women with cervical intraepithelial neoplasia grade 3 compared with women with transient infections. We observed high correlation of methylation patterns between phylogenetically related types. The highest areas under the curve were 0.81 for HPV31, 0.85 for HPV18, and 0.98 for HPV45. Differential methylation patterns in cervical intraepithelial neoplasia grade 3 patients with multiple infections suggest that methylation can clarify which of the infections is causal.

Conclusions Carcinogenic HPV DNA methylation indicates transforming HPV infections. Our findings show that methylation of carcinogenic HPV types is a general phenomenon that warrants development of diagnostic assays.

J Natl Cancer Inst 2012;104:1738–1749

Introduction

Persistent infections with carcinogenic human papillomaviruses (HPV) are the necessary cause of virtually all cervical cancers (1). HPV infections are very common in the population with a characteristic age distribution, peaking in women between 18–25 years and slowly declining with increasing age. While a large proportion of sexually active women acquire at least one HPV infection at some time, only a minority of infections persist and progress to true cervical cancer precursors, best defined as cervical intraepithelial neoplasia grade 3 (CIN3) (2). Over 30 years, approximately a third of CIN3 lesions invade and form cervical cancers (3). The transition from a productive infection to a transforming infection in CIN3 is characterized by a substantial change of viral gene expression, from high-level of expression of structural genes (L1

and L2) and low expression of viral oncogenes (E6 and E7) to high expression of oncogenes that interfere with cellular apoptosis and cell cycle regulation (4,5).

The determinants of and molecular basis for viral persistence, progression to CIN3, and invasion to cancer are poorly understood. Recently, we demonstrated that methylation of multiple CpG sites in various regions of the HPV16 genome is associated with prevalent CIN3 and with progression to CIN3 (6). HPV16 is the most carcinogenic HPV type, accounting for approximately 57% of cervical cancers worldwide (7), and is thought to progress faster from infection to cervical cancer than other types. To evaluate whether the association between viral methylation and CIN3 is a phenomenon specific to HPV16, or a general feature of carcinogenic HPV types, we studied whole-genome methylation

for three important carcinogenic types, namely HPV18, 31, and 45. HPV31 causes about 4% of all cervical cancers and is closely related genetically to HPV16, and both belong to the alpha-9 species. HPV18 is the second most carcinogenic type after HPV16 and accounts for 16% of all cervical cancers (7). HPV45 causes about 5% of cervical cancers and is closely related to HPV18; both are members of the alpha-7 species. For these three types, we compared methylation in women with CIN3 compared with those with infections without signs of progression and methylation levels in women infected with one versus multiple types of HPV.

Materials and Methods

Population

At all medical centers in Kaiser Permanente Northern California (KPNC), women aged 30 and older are tested for carcinogenic HPV using the Hybrid Capture 2 (HC2) test as an adjunct to Pap smears for primary cervical cancer screening. Women under 30 with a Pap smear of atypical cells of undetermined significance (ASC-US) are triaged for co-testing. The HPV Persistence and Progression Cohort (The PaP Cohort) is a repository of residual cervical specimens, placed into specimen transport medium (STM; Qiagen), from women who tested HC2 positive (8). Eligibility was defined as women 21 and older who had not opted out from having their specimen banked and tested for HPV-related biomarkers including HPV genotypes. De-identified age data and results of follow-up cytological and pathologic results of women in the PaP cohort were obtained from linked electronic health records. For this analysis, we sampled from women who tested HPV-positive in cervical cancer screening (total population of 34 850 women) and had HPV genotyping information available (a stratified random sample of 580 CIN3 case patients and 1500 women with <CIN2 histology from the total population of 34 850 women) (8). We defined as case patients women positive for HPV31, HPV18, or HPV45 with CIN3 detected at biopsy or loop electrosurgical excision procedure (LEEP) following the workup of an index HC2 positive/cytology positive test result between January 2007 and October 2008. We selected all CIN3s positive for the three HPV types. As control subjects, we selected women infected with HPV31, HPV18, or HPV45 but without reported histological or cytological abnormalities at a ratio of about 1 control subject per 1 case patient. Five CIN3 case patients had infections with two of the targeted HPV genotypes (four case patients with HPV18 and HPV31 co-infections and one case patient with a HPV31 and HPV45 co-infection). The final analytic group included 92 women with CIN3 and 96 women with <CIN2 histology and <ASC-US cytology positive for HPV31, HPV18, or HPV45 in HPV genotyping. The study has been approved by the Institutional Review Board at KPNC and by the Special Studies Institutional Review Board at the National Cancer Institute.

HPV DNA Testing

Hybrid Capture 2 (HC2; Qiagen Inc., Gaithersburg, MD), a DNA test for a pool of 13 carcinogenic HPV genotypes (HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68) was conducted on the banked STM specimen per the manufacturer's instructions as

part of routine cervical cancer screening, as described above. The MY09/M11 L1 degenerate primer PCR (MY09/11 PCR) method was used to test neutralized and frozen residual STM specimens (after HC2 testing) for HPV-specific genotypes, as described previously (8,9). Dot blot hybridization of the amplicons, using HPV genotype-specific oligonucleotide probes, was used to identify individual HPV genotypes.

HPV DNA Methylation Measurement

We sampled CpG sites from the whole HPV genomes. The CpG sites within the L1, L2, and E2 ORFs were surveyed because these regions showed a statistically significant association between HPV16 methylation and CIN3+ in our previous studies (6,10). For the other open reading frames and the upstream regulatory region (URR), at least one segment with CpG sites was sampled for methylation. Primers for PCR were designed using MethPrimer (<http://www.urogene.org/methprimer/index1.html>) and are listed in [Supplementary Table 1](#) (available online). We used the following reference sequences (GenBank IDs) as a denominator for CpG counts and for primer design: X05015 (HPV18), J04353 (HPV31), and X74479 (HPV45). All primers were purchased from Integrated DNA Technologies Inc (Coralville, IA). The extent of CpG methylation was determined by pyrosequencing (Qiagen, Germany) according to the manufacturers' recommended protocol. In brief, DNAs from cervicovaginal samples previously used for HPV typing were treated with freshly prepared bisulfate, using the EZ DNA methylation kit (Zymo Research, Orange, CA), according to the recommended protocol. After bisulfite modification, genome-wide pyrosequencing was performed covering 80–106 sites. Bisulfite treated DNA results in unmethylated C's being converted into U's, which are then converted to T's by Taq polymerase during PCR amplification; methylated C's remain unmodified. All segments were amplified by HotStart-IT Fidelity Taq DNA polymerase (United States Biochemicals, Cleveland, OH). All PCR reactions were performed with one primer containing biotin that was captured by streptavidin-coated beads, and annealed to a sequencing primer. Pyrosequencing was performed by DNA polymerase with sequential addition of each dNTP, and the released pyrophosphate was catalyzed by ATP sulfurylase, luciferase, and apyrase sequentially to generate light signals, which are detected by a charge-coupled device (CCD) camera on a PSQ96 system (Qiagen). Because the bioluminescent response is linear for the sequential addition of dNTPs, the signal intensity reflects the original methylation ratio (ie, percent methylation). Each PCR product was pyrosequenced by one or up to four sequencing primers ([Supplementary Table 1](#), available online). For HPV18, 67% of all sites had a 90% or higher pass rate, and 90% had a pass rate of 80% or higher. For HPV31, 91% of all sites had a 90% or higher pass rate, and 100% had a pass rate of 80% or higher. For HPV45, 28% of all sites had a 90% or higher pass rate, and 90% had a pass rate of 80% or higher.

Statistical Analysis

Median methylation levels in case patients and control subjects at each site were compared using the Mann-Whitney U test. A correlation matrix based on Spearman rank correlation coefficients of methylation levels at individual CpG sites was calculated,

using levels in cases only because many controls had very low levels leading to statistical imprecision. To compare correlation of methylation levels in HPV genes across genotypes, medians and interquartile ranges of correlation coefficients were computed for E6/7, E2, L2, L1, and URR. Methylation levels were categorized in tertiles for each site, defined by the distribution of methylation levels in control subjects. Specifically, logistic regression was used to calculate odds ratios (ORs) and 95% confidence intervals (CIs) based on a dichotomized result for methylation (the highest tertile vs middle and low tertiles combined), comparing the case patients and control subjects. ROC curves and areas under the curve with 95% CIs were calculated for distributions of methylation levels in case patients and control subjects. Exemplary pairs of sensitivity and specificity are presented based on cutoffs that achieved the highest Youden's index (Sensitivity+Specificity-1) with a minimal sensitivity of 0.8 and a minimal specificity of 0.6. Raw *P* values are presented in the text and tables with footnotes indicating *P* values significant after conservative Bonferroni correction. This presentation permits an assessment of the results without adjustment for multiple comparisons or with the most conservative correction for multiple comparisons. All analyses were performed in Stata 12 (StataCorp, College Station, TX).

Results

HPV31 Methylation in Case Patients and Control Subjects

A total of 88 women infected with HPV31, including 45 women with CIN3 (mean age 37, median age 35) and 43 women without cytological or histological abnormalities (mean age 37, median age 37) were included in the analysis. Eighty CpGs (out of 117 for HPV31) covering the complete genome were analyzed. The CpG sites analyzed included 4 of 10 CpGs from E6, 4 of 8 CpGs from E7, 9 of 25 CpGs from E1, 15 of 18 CpGs from E2, 22 of 25 CpGs from L2, 21 of 22 CpGs from L1 and 5 of 9 CpGs from the URR (Supplementary Table 2, available online). Except for one CpG site in E2 (nucleotide position 5502), methylation levels at all sites in case patients were equal to or higher than in control subjects (Figure 1). The highest levels in case patients were observed in the L2 and L1 regions. A site in L2 with the highest median methylation level (0.29 at genome position 5086) in control subjects also had the highest median methylation level (0.55 at genome position 5086) in CIN3 case patients; the highest median difference (0.35, Mann-Whitney *U* *P* value <0.00002, significant after Bonferroni correction) between case patients and control subjects was observed for site 5881 in L1 (Table 1).

The correlation between methylation levels at each site compared with all other sites within each gene revealed weak to moderate correlation within E6, E7, E1, and URR, while correlation was very high in parts of E2, all of L2 and all of L1. There was also high correlation of methylation levels across genes for E2, L2, and L1 (Figure 1: red shading in lower part of panel).

In logistic regression models, increased methylation was associated with increased risk of CIN3 at all but four sites with odds ratios ranging from 1.01 to 15.1. After adjusting for multiple comparisons, 33 sites in E1, E2, L2, and L1 were statistically significant (Bonferroni-adjusted level of significance *p* <0.00063) with ORs

ranging from 5.1 to 15.1 (Supplementary Table 2, available online). Areas under the curve were calculated to compare the discrimination of individual CpG sites between case patients and control subjects; 11 CpG sites in E1, E2, L2, and L1 had an AUC of at least 0.75 with a maximum AUC of 0.81 (position 5524) (Table 1). ROC curves are shown for three sites in E2 and L2 (Supplementary Figure 1, A–C, available online). An assay based on the detection of CpG methylation at site 4263 with a methylation threshold of ≥0.04 would have 82% sensitivity and 74% specificity for the detection of CIN3 among HPV31-positive women in this population.

HPV18 Methylation in Case Patients and Control Subjects

A total of 82 women infected with HPV18, including 40 women with CIN3 (mean age 38, median age 37), and 42 women without detectable cytological or histological abnormalities (mean age 36, median age 33) were included in the analysis. One hundred and six CpGs (out of 178 total for HPV18) were analyzed, including 4 of 16 CpGs from E6, 10 of 11 CpGs from E7, 5 of 36 CpGs from E1, 25 of 29 CpGs from E2, 33 of 39 CpGs from L2, 24 of 32 CpGs from L1 and 5 of 15 CpGs from the URR (Supplementary Table 3, available online). Methylation levels at all CpG sites except one (L1 nucleotide position 5502) were equal to or higher in case patients than in control subjects. Methylation levels in case patients were low (<0.04) in E6, E7, E1, and the URR, and much higher in E2, L2, and L1 (Figure 2). HPV18 methylation levels in E2, L2, and L1 were higher compared with the levels observed for HPV16 and HPV31. The highest median methylation level among control subjects was observed for a site in L2 (0.3 at genome position 4898). The highest median methylation level among CIN3 case patients was observed for a site in L1 (0.71 at genome position 6014); this site also had the highest median difference between case patients and control subjects (0.45, Mann-Whitney *U* *P* value 0.0009, Bonferroni-adjusted level of significance *p* <0.00047) (Table 2).

CpG methylation was highly correlated within individual HPV18 genes, with intra-gene correlation coefficients between 0.7 and 1.0, for sites within E6, E7, E1, E2, L2, L1, and URR. In addition, we observed two blocks of correlated methylation levels across the early genes E6-E2, and across the L2-L1 region (Figure 2).

In logistic regression models, increased methylation was associated with increased risk of CIN3 at all but one site with ORs ranging from 1.7–25.0. After adjusting for multiple comparisons, 38 sites in E2, L2, and L1 remained statistically significant (Bonferroni-adjusted level of statistical significance *p* <0.00047), with ORs ranging from 5.7–14.2 (Supplementary Table 3, available online). Sixteen CpG sites in E2, L2, and L1 had an AUC of at least 0.80, with a maximum AUC of 0.85 (position 7068) (Table 2). ROC curves are shown for three sites in L2 and L1 (Supplementary Figure 1, D–F, available online). As an example, an assay based on the detection of CpG methylation at site 7041 with a methylation threshold of ≥0.07 would have 85% sensitivity and 71% specificity for the detection of CIN3 among HPV18-positive women in this population.

HPV45 Methylation in Case Patients and Control Subjects

A total of 23 women infected with HPV45, including 12 women with CIN3 (mean age 42, median age 42) and 11 women without

HPV31

Figure 1. HPV31 methylation levels in case patients and control subjects. For each of the analyzed 80 CpG sites, the median percent methylation in case patients (red bars) and control subjects (grey bars) is shown (secondary y-axis). The black dots indicate Mann-Whitney U-test *P* values for median methylation differences between case patients and control subjects on an inverse logarithmic scale (primary y-axis). The HPV31 genomic region is shown on the x-axis. The correlation matrix

in the bottom part of the graph indicates the correlation of methylation levels at all sites across the HPV31 genome. **Black lines** indicate transitions between adjacent HPV genes. For each combination of two sites, the Spearman rho correlation coefficient is shown color-coded with a range from perfect correlation (1, dark red) to inverse correlation (-1, dark blue) as indicated by the color bar. All statistical tests were two-sided.

detectable cytological or histological abnormalities (mean age 37, median age 33) were included in the analysis. One hundred and five CpG sites (out of 160 total for HPV45) covering the complete genome were analyzed, including 6 of 13 CpGs from E6, 9 of 11 CpGs from E7, 4 of 32 CpGs from E1, 30 of 31 CpGs from E2, 24 of 32 CpGs from L2, 25 of 29 CpGs from L1 and 5 of 12 CpGs from the URR (Supplementary Table 4, available online). Similar to HPV18 and HPV31, the vast majority of CpG sites had higher methylation levels in cases than in controls (Supplementary Table 4, available online). In CIN3 cases, HPV45 showed the

highest methylation levels in E2, L2, and L1 of all three types included in this study (Figure 3).

The highest median methylation level for case patients and control subjects was observed for a CpG at genome position 4908 in L2 (0.33 in control subjects, 1.0 in case patients). The highest median difference between case patients and control subjects was observed for site 3580 in E2 (0.95, Mann-Whitney U test *P* value 0.0016).

The correlation pattern of HPV45 showed strong intra-gene correlations of methylation levels within E6, E7, E1, E2, L2, L1,

Table 1. HPV31 methylation in case patients and control subjects*

Region	Position	Median			Mann-Whitney U-test <i>P</i>	Univariate regression high HPV methylation vs low/medium methylation			ROC curve	
		Control subject (n=43)	Case patient (n=45)	Difference		OR	95% CI	<i>P</i>	AUC	95% CI
E1	2648	0.00	0.02	0.02	<.0001	15.07	5.05 to 44.94	<.0001	0.79	0.70 to 0.89
E2	3411	0.00	0.05	0.05	<.0001	10.24	3.75 to 27.95	<.0001	0.80	0.70 to 0.89
E2	3414	0.01	0.05	0.04	<.0001	10.33	3.77 to 28.32	<.0001	0.79	0.69 to 0.89
L2	4263	0.00	0.10	0.10	<.0001	13.09	4.68 to 36.58	<.0001	0.80	0.71 to 0.90
L2	4402	0.04	0.09	0.05	<.0001	11.55	4.26 to 31.31	<.0001	0.76	0.66 to 0.87
L2	5038	0.12	0.33	0.21	<.0001	6.71	2.61 to 17.27	<.0001	0.75	0.64 to 0.86
L2	5086	0.29	0.55	0.26	<.0001	5.96	2.35 to 15.13	.0002	0.75	0.64 to 0.86
L2	5524	0.06	0.23	0.17	<.0001	12.20	4.32 to 34.42	<.0001	0.81	0.71 to 0.90
L2	5530	0.02	0.10	0.08	<.0001	12.83	4.64 to 35.53	<.0001	0.79	0.69 to 0.89
L1	5659	0.20	0.40	0.20	<.0001	5.70	2.27 to 14.27	.0002	0.74	0.64 to 0.85
L1	5881	0.11	0.46	0.35	<.0001	8.06	3.05 to 21.28	<.0001	0.77	0.66 to 0.87
L1	6351	0.00	0.08	0.08	<.0001	11.19	3.70 to 33.97	<.0001	0.77	0.66 to 0.88

* Logistic regression was used to calculate odds ratios (ORs) and 95% confidence intervals (CIs) based on a dichotomized result for methylation (the highest tertile vs. middle and low tertiles combined), comparing the case patients and control subjects. ROC curves and areas under the curve with 95% CIs were calculated for distributions of methylation levels in case patients and control subjects. All CpG sites in HPV31 with at least 0.2 median difference, or at least 10.0 odds ratio (OR), or at least 0.80 area under the curve (AUC) are displayed. All displayed *P* values are below 0.00063, the level of statistical significance after correction for multiple infections using Bonferroni's correction. All statistical tests were two-sided.

and URR. Similar to HPV18, there were two blocks of correlated methylation sites across E6, E7, E1, and L2, L1 (Figure 3).

In logistic regression models, increased methylation was associated with increased risk of CIN3 at most sites with ORs ranging from 1.3–70.0 (Supplementary Table 4, available online), but *P* values were less significant in HPV45 because of the smaller numbers of case patients and control subjects. None of the sites showed a statistically significant association after adjusting for multiple comparisons (Bonferroni-adjusted level of significance $p < 0.00047$). The lowest *P* value was observed for site 7042 in L1 (OR 40.0, 95% CI 3.05 to 525, $p = 0.00498$).

Sixty-one CpG sites in E6, E2, L2, and L1 had an AUC of at least 0.80, with a maximum AUC of 0.98 (position 6990) (Table 3). ROC curves are shown for three sites in E2 and L1 (Supplementary Figure 1, G–I, available online). An assay based on the detection of CpG methylation at site 7042 at a cutoff of ≥ 0.3 would have 91% sensitivity and 100% specificity for the detection of CIN3 among HPV45-positive women in this population.

Correlation of Methylation Levels Within HPV Genes Across Genotypes

The visual presentation of the correlation matrices for HPV31, HPV18, and HPV45 shows notable differences across types, with higher similarity between HPV18 and HPV45 compared with HPV31 (Figures 1–3). We computed median Spearman rank coefficients and interquartile ranges for E6/E7, E2, L2, L1, and URR for each genotype (Table 4). While intra-gene correlation was consistently high across the three types for E2, L2, and L1, there were substantial differences in intra-gene correlation for the E6/E7 and URR regions. HPV18 and HPV45 showed similar (E6/E7) or slightly lower (URR) levels of correlation compared with the other genes. In contrast, Spearman rank correlation coefficients were much lower for HPV31 in E6/E7 and URR, confirming that the

methylation patterns were more similar between the two alpha-7 types compared with the alpha-9 type HPV31.

Methylation Levels in Single versus Multiple Infections

We reasoned that for case patients, detection of a single HPV type implied that the type caused the precancer; in contrast, in a case patient with multiple carcinogenic HPV types, the lesion might be attributable to any one of the types present. In consequence, the causal type for a CIN3 lesion would have altered methylation patterns, whereas the other, non-causal HPV infections would not. To test this notion, we evaluated the methylation levels in case patients and control subjects by infection status with multiple versus single carcinogenic HPV types. For HPV31, 12 case patients and 5 control subjects had multiple infections, whereas 32 case patients and 38 control subjects were infected with HPV31 only. For HPV18, 12 cases and 13 controls had multiple carcinogenic infections, whereas 28 case patients and 28 control subjects had infections with HPV18 only. Data for HPV45 were too sparse to analyze. For HPV31, several sites in E2, L2, and L1 had higher median methylation levels in women with single HPV31 infections, with one site in L1 reaching statistical significance (Supplementary Table 5, available online). Similarly, for HPV18, median methylation levels at ten sites in E2, L2, and L1, were found statistically significantly higher in case patients with single HPV18 infections compared with those with multiple infections (Supplementary Table 5, available online). We further stratified the analysis for HPV31 into case patients with single types ($n = 32$), multiple types including HPV16 ($n = 6$), and multiple types without HPV16 ($n = 6$), because HPV16 is the most carcinogenic type. There were not enough co-infections of HPV16 with HPV18 or HPV45 to extend the analysis beyond HPV31. With few exceptions, the methylation levels of HPV31 among case patients with multiple infections including HPV16 were lower

HPV18

Figure 2. HPV18 methylation levels in case patients and control subjects. For each of the analyzed 105 CpG sites, the median percent methylation in case patients (**red bars**) and control subjects (**grey bars**) is shown (secondary y-axis). The **black dots** indicate Mann-Whitney-U test *P* values for median methylation differences between case patients and control subjects on an inverse logarithmic scale (primary y-axis). The HPV18 genomic region is shown on the x-axis. The correlation matrix

in the bottom part of the graph indicates the correlation of methylation levels at all sites across the HPV18 genome. **Black lines** indicate transitions between adjacent HPV genes. For each combination of two sites, the Spearman rho correlation coefficient is shown color-coded with a range from perfect correlation (1, dark red) to inverse correlation (-1, dark blue) as indicated by the **color bar**. All statistical tests were two-sided.

compared with single infections and multiple infections without HPV16 (Figure 4).

Discussion

Whole-genome viral methylation patterns for HPV31, HPV18, and HPV45 vary significantly between women with CIN3 and women without histological or cytological evidence of high-grade CIN.

Consistent across HPV types 31, 18, 45, and 16 (6), four carcinogenic types accounting for over 80% of cervical cancers worldwide, high methylation levels in cases were observed for regions in E2, L2, and L1, whereas methylation was low in E6, E7, E1, and the URR. Although our previous findings for HPV16 could have been features specific to this unique carcinogenic type, our results here suggest that increased methylation at viral CpG sites is a general phenomenon associated with cellular transformation caused by carcinogenic HPV types.

Table 2. HPV18 methylation in case patients and control subjects*

Region	Position	Median			Mann-Whitney U-test <i>P</i>	Univariate regression high HPV methylation vs low/medium methylation			ROC curve	
		Control subject (n=42)	Case patient (n=40)	Difference		OR	95% CI	<i>P</i>	AUC	95% CI
E6	423	0.00	0.01	0.01	.0057	20.00	2.47 - 162.19	.0050	0.67	0.55 to 0.78
E6	477	0.00	0.01	0.01	.0013	20.00	2.47 to 162.19	.0050	0.69	0.58 to 0.80
E6	480	0.01	0.01	0.00	.0030	25.00	3.10 to 201.40	.0025	0.68	0.57 to 0.79
L2	4265	0.04	0.25	0.21	.0001	6.75	2.36 to 19.29	.0004	0.80	0.68to 0.92
L2	4268	0.12	0.32	0.21	.0005	7.36	2.50 to 21.33	.0002	0.78	0.67 to 0.90
L2	4505	0.03	0.30	0.27	<.0001	13.05	4.17 to 40.80	<.0001	0.85	0.75 to 0.94
L2	4641	0.06	0.28	0.23	.0081	5.40	1.80 to 16.03	.0024	0.75	0.61 to 0.89
L2	4659	0.07	0.27	0.20	.0003	6.18	2.05 to 18.60	.0012	0.76	0.63 to 0.90
L2	4680	0.18	0.46	0.29	.0003	7.95	2.50 to 25.16	.0004	0.77	0.64 to 0.90
L2	4737	0.14	0.49	0.35	.0001	8.33	2.70 to 25.76	.0002	0.78	0.66 to 0.89
L2	4898	0.30	0.50	0.20	.0038	3.81	1.45 to 10.00	.0067	0.70	0.57 to 0.82
L2	4901	0.19	0.40	0.21	.0013	8.38	2.94 to 23.81	<.0001	0.72	0.59 to 0.84
L2	4916	0.28	0.62	0.34	.0041	4.88	1.82 to 13.08	.0016	0.69	0.57 to 0.82
L2	4973	0.13	0.42	0.29	.0023	4.67	1.81 to 12.05	.0015	0.70	0.58 to 0.82
L2	5004	0.11	0.31	0.20	.0001	9.06	3.23 to 25.40	<.0001	0.75	0.64 to 0.87
L2	5045	0.11	0.39	0.28	.0002	5.42	1.99 to 14.77	.0010	0.76	0.64 to 0.87
L2	5084	0.11	0.34	0.23	.0008	4.88	1.82 to 13.09	.0016	0.73	0.61 to 0.85
L2	5114	0.14	0.41	0.27	.0023	4.91	1.83 to 13.20	.0016	0.71	0.58 to 0.83
L2	5126	0.20	0.42	0.22	.0016	4.91	1.83 to13.20	.0016	0.71	0.59 to 0.83
L2	5331	0.22	0.56	0.34	<.0001	6.25	2.36 to 16.55	.0002	0.80	0.71 to 0.90
L2	5333	0.18	0.48	0.30	<.0001	8.35	3.01 to 23.10	<.0001	0.81	0.71 to 0.90
L2	5367	0.19	0.56	0.37	.1851	4.67	0.53 to 40.88	.1642	0.70	0.40 to 0.99
L2	5406	0.22	0.43	0.21	.2429	2.63	0.30 to 23.00	.3835	0.67	0.38 to 0.97
L1	5585	0.03	0.26	0.23	.0034	4.14	1.65 to 10.41	.0025	0.72	0.61 to 0.84
L1	5625	0.19	0.50	0.31	.0041	3.63	1.40 to 9.10	.0063	0.68	0.56 to 0.80
L1	5642	0.10	0.33	0.23	.0017	4.85	1.90 to 12.50	.0011	0.70	0.58 to 0.82
L1	6014	0.26	0.71	0.45	.0009	5.03	1.96 to 12.91	.0008	0.71	0.60 to 0.83
L1	6019	0.14	0.53	0.39	<.0001	7.79	2.85 to 21.29	<.0001	0.79	0.69 to 0.89
L1	6038	0.10	0.37	0.27	<.0001	8.32	3.06 to 22.66	<.0001	0.77	0.66 to 0.88
L1	6142	0.18	0.56	0.38	<.0001	6.25	2.36 to 16.55	.0002	0.80	0.70 to 0.90
L1	6144	0.16	0.52	0.36	<.0001	9.85	3.45 to 28.12	<.0001	0.81	0.71 to 0.91
L1	6161	0.07	0.40	0.33	<.0001	11.85	3.98 to 35.25	<.0001	0.82	0.72 to 0.92
L1	6363	0.15	0.47	0.32	<.0001	5.68	2.18 to 14.77	.0004	0.78	0.67 to 0.88
L1	6366	0.10	0.43	0.33	<.0001	11.39	3.96 to 32.79	<.0001	0.81	0.71 to 0.92
L1	7011	0.01	0.24	0.23	<.0001	14.22	4.87 to 41.51	<.0001	0.84	0.75 to 0.93
L1	7038	0.05	0.28	0.23	<.0001	5.68	2.18 to 14.77	.0004	0.75	0.64 to 0.86
L1	7041	0.04	0.33	0.29	<.0001	13.69	4.57 to 41.06	<.0001	0.85	0.76 to 0.93
L1	7062	0.07	0.37	0.30	<.0001	12.89	4.28 to 38.77	<.0001	0.85	0.76 to 0.94
L1	7068	0.02	0.22	0.20	<.0001	11.49	3.85 to 34.24	<.0001	0.85	0.77 to 0.94
L1	7090	0.05	0.35	0.30	<.0001	6.03	2.27 to 16.02	.0003	0.81	0.72 to 0.91
L1	7110	0.08	0.45	0.37	<.0001	10.52	3.70 to 29.92	<.0001	0.83	0.73 to 0.92
L1	7116	0.07	0.36	0.29	<.0001	8.61	3.17 to 23.39	<.0001	0.83	0.74 to 0.92
L1	7122	0.03	0.29	0.26	<.0001	8.45	3.13 to 22.81	<.0001	0.81	0.72 to 0.91

* Logistic regression was used to calculate odds ratios (ORs) and 95% confidence intervals (CIs) based on a dichotomized result for methylation (the highest tertile vs. middle and low tertiles combined), comparing the case patients and control subjects. ROC curves and areas under the curve with 95% CIs were calculated for distributions of methylation levels in case patients and control subjects. All CpG sites in HPV18 with at least 0.2 median difference, or at least 10.0 odds ratio (OR), or at least 0.80 area under the curve (AUC) are displayed. Bold *P* values are below 0.00048, the level of statistical significance after correction for multiple infections using Bonferroni's correction. All statistical tests were two-sided.

Although the major methylation patterns were similar across types, we observed type-specific differences. Methylation levels were highest in case patients who were positive for HPV45, followed by HPV18 and HPV31. The methylation patterns and the correlation of methylation levels between different sites were more similar for HPV18 and HPV45 compared with HPV31. The similarities observed for methylation levels correspond to other general similarities between these two alpha-7 types. In addition to their

close evolutionary relatedness (11), HPV18 and HPV45 share biologic characteristics: they are associated with glandular lesions and cervical adenocarcinomas (1). Both types are found to have a higher rate of viral integration in cervical cancers compared with HPV16 (12). The methylation patterns observed for HPV31 are similar to the patterns we previously found for HPV16. HPV16 and HPV31 are closely related types from the alpha-9 species group (13). HPV16 and HPV31 are about 80% identical, and HPV18

HPV45

Figure 3. HPV45 methylation levels in case patients and control subjects. For each of the analyzed 106 CpG sites, the median percent methylation in case patients (**red bars**) and control subjects (**grey bars**) is shown (secondary y-axis). The **black dots** indicate Mann-Whitney U-test *P* values for median methylation differences between case patients and control subjects on an inverse logarithmic scale (primary y-axis). The HPV45 genomic region is shown on the x-axis. The correlation matrix

in the bottom part of the graph indicates the correlation of methylation levels at all sites across the HPV45 genome. **Black lines** indicate transitions between adjacent HPV genes. For each combination of two sites, the Spearman rho correlation coefficient is shown color-coded with a range from perfect correlation (1, dark red) to inverse correlation (-1, dark blue) as indicated by the **color bar**. All statistical tests were two-sided.

and HPV45 are 82% identical, whereas 16/31 and 18/45 are less than 70% identical. Moreover, there are 110 and 117 CpG sites in HPV16 and 31, respectively, but 178 and 160 CpG sites in HPV18 and 45, respectively.

We currently do not have functional data implicating a causal relationship between viral methylation and progression to CIN3. DNA methylation plays an important role in gene regulation. Transcriptional repression by DNA methylation can be mediated through methylation-sensitive transcription factors, proteins with DNA-methylation binding domains that block transcription factor

binding, and through dense chromatin packing of methylated chromosomal regions (14). While we observed high methylation levels in the L1 and L2 regions in women with CIN3, there was only low-level methylation in the URR and E6/E7 regions, which did not differ between case patients and control subjects. Thus, the genome regions involved in regulation of HPV oncogene expression remained largely unmethylated in CIN3. Several observations support a functional role of viral methylation in carcinogenesis. We observed strong associations between viral methylation and CIN3 for four different carcinogenic HPV types. If viral methylation

Table 3. HPV45 methylation in case patients and control subjects

Region	Position	Median			Mann-Whitney U-Test	Univariate regression high HPV methylation vs low/medium methylation			ROC curve	
		Control subject (n=11)	Case patient (n=12)	Difference		OR	95% CI	P	AUC	95% CI
E6	474	0.00	0.02	0.02	.0103	6.43	0.60 to 68.31	.1228	0.80	0.64 to 0.95
E6	537	0.01	0.01	0.01	.0155	15.75	1.42 to 174.25	.0246	0.77	0.56 to 0.95
E7	751	0.00	0.02	0.02	.0245	10.67	1.39 to 82.03	.0230	0.78	0.58 to 0.98
E2	2761	0.00	0.05	0.05	.0059	14.00	1.54 to 127.00	.0191	0.84	0.67 to 1.00
E2	2767	0.00	0.06	0.06	.0151	8.17	1.03 to 64.94	.0471	0.81	0.61 to 1.00
E2	2785	0.00	0.03	0.03	.0008	36.00	2.72 to 476.30	.0065	0.92	0.80 to 1.00
E2	2800	0.00	0.11	0.11	.0022		NA		0.90	0.76 to 1.00
E2	2815	0.00	0.04	0.04	.0017		NA		0.88	0.75 to 1.00
E2	3080	0.01	0.05	0.05	.0139	10.00	1.05 to 95.46	.0455	0.82	0.60 to 1.00
E2	3308	0.01	0.17	0.16	.0009	70.00	3.72 to 1317.00	.0046	0.91	0.77 to 1.00
E2	3337	0.00	0.26	0.26	.0029	NA			0.91	0.78 to 1.00
E2	3443	0.00	0.23	0.23	.0220	20.00	0.93 to 430.00	.0556	0.88	0.65 to 1.00
E2	3451	0.00	0.19	0.19	.0682	7.50	0.46 to 123.00	.1576	0.82	0.54 to 1.00
E2	3454	0.00	0.11	0.11	.0366		NA		0.80	0.56 to 1.00
E2	3467	0.00	0.13	0.13	.0366		NA		0.80	0.56 to 1.00
E2	3473	0.00	0.15	0.15	.0366		NA		0.80	0.56 to 1.00
E2	3502	0.01	0.85	0.85	.0007	18.67	1.56 to 223.00	.0207	0.96	0.86 to 1.00
E2	3505	0.02	0.87	0.85	.0006		NA		0.97	0.89 to 1.00
E2	3516	0.02	0.22	0.20	.0231	8.17	1.03 to 65.00	.0471	0.81	0.58 to 1.00
E2	3551	0.01	0.87	0.86	.0010	32.00	2.40 to 428.00	.0088	0.94	0.83 to 1.00
E2	3569	0.03	0.81	0.78	.0022	32.00	2.40 to 428.00	.0088	0.91	0.74 to 1.00
E2	3580	0.04	0.98	0.95	.0016	18.67	1.56 to 223.00	.0207	0.93	0.81 to 1.00
E2	3640	0.03	0.90	0.87	.0013	16.33	1.35 to 198.00	.0281	0.95	0.84 to 1.00
E2	3699	0.15	0.94	0.79	.0184	16.33	1.35 to 198.00	.0281	0.83	0.62 to 1.00
L2	4248	0.06	0.85	0.79	.0010	23.33	2.00 to 273.30	.0121	0.92	0.81 to 1.00
L2	4257	0.12	0.81	0.70	.0017	10.50	1.36 to 81.05	.0241	0.90	0.76 to 1.00
L2	4260	0.19	0.78	0.60	.0048	10.50	1.36 to 81.05	.0241	0.86	0.69 to 1.00
L2	4266	0.16	0.66	0.51	.0075	10.50	1.36 to 81.05	.0241	0.85	0.65 to 1.00
L2	4312	0.09	0.92	0.83	.0013	23.33	2.00 to 273.30	.0121	0.91	0.79 to 1.00
L2	4418	0.09	0.52	0.44	.0007	23.33	2.00 to 273.30	.0121	0.94	0.83 to 1.00
L2	4437	0.12	0.76	0.65	.0017	23.33	2.00 to 273.30	.0121	0.90	0.78 to 1.00
L2	4441	0.10	0.82	0.72	.0009	23.33	2.00 to 273.30	.0121	0.93	0.82 to 1.00
L2	4588	0.14	0.74	0.60	.0009	26.67	2.30 to 308.00	.0085	0.92	0.81 to 1.00
L2	4604	0.26	0.89	0.63	.0034	12.00	1.58 to 91.08	.0163	0.87	0.72 to 1.00
L2	4627	0.13	0.52	0.39	.2489	3.20	0.54 to 18.98	.2003	0.64	0.38 to 0.90
L2	4673	0.09	0.85	0.76	.0069	15.75	1.75 to 141.40	.0138	0.86	0.69 to 1.00
L2	4890	0.06	0.54	0.48	.0025	26.67	2.31 to 308.00	.0085	0.88	0.71 to 1.00
L2	4893	0.03	0.49	0.46	.0012	45.00	3.46 to 584.00	.0036	0.91	0.75 to 1.00
L2	4908	0.33	1.00	0.67	.0014	26.67	2.31 to 308.00	.0085	0.87	0.72 to 1.00
L2	4938	0.08	0.68	0.60	.0010	23.33	2.00 to 273.30	.0121	0.92	0.81 to 1.00
L2	5076	0.05	0.60	0.55	.0101	20.25	2.32 to 176.80	.0065	0.82	0.62 to 1.00
L2	5106	0.03	0.33	0.30	.0019	23.33	2.00 to 273.30	.0121	0.90	0.77 to 1.00
L2	5157	0.02	0.26	0.24	.0018	10.50	1.36 to 81.00	.0241	0.90	0.76 to 1.00
L1	5580	0.02	0.41	0.39	.0022	13.33	1.78 to 100.00	.0118	0.88	0.71 to 1.00
L1	5586	0.03	0.30	0.27	.0034	13.33	1.78 to 100.00	.0118	0.86	0.69 to 1.00
L1	5620	0.33	0.93	0.60	.0010	23.33	2.00 to 273.30	.0121	0.92	0.79 to 1.00
L1	5636	0.26	0.90	0.64	.0007	23.33	2.00 to 273.30	.0121	0.94	0.83 to 1.00
L1	5695	0.13	0.66	0.53	.0006	26.67	2.31 to 308.00	.0085	0.93	0.82 to 1.00
L1	5728	0.09	0.34	0.25	.0409	4.67	0.77 to 28.50	.0950	0.76	0.54 to 0.97
L1	5808	0.17	0.79	0.62	.0008	23.33	2.00 to 273.30	.0121	0.93	0.81 to 1.00
L1	5844	0.10	0.94	0.85	.0027	21.00	1.50 to 293.00	.0236	0.95	0.83 to 1.00
L1	6361	0.22	0.92	0.71	.0006	23.33	2.00 to 273.30	.0121	0.94	0.85 to 1.00
L1	6364	0.14	0.82	0.68	.0010		NA		0.92	0.81 to 1.00
L1	6628	0.14	0.62	0.49	.0048	23.33	2.00 to 273.30	.0121	0.86	0.70 to 1.00
L1	6874	0.11	0.53	0.43	.0003		NA		0.97	0.90 to 1.00
L1	6990	0.06	0.59	0.54	.0002		NA		0.98	0.93 to 1.00
L1	7015	0.08	0.65	0.57	.0002		NA		0.97	0.91 to 1.00
L1	7042	0.04	0.84	0.81	.0004	40.00	3.05 to 525.00	.0050	0.96	0.87 to 1.00

(Table continues)

Table 3 (Continued).

Region	Position	Median			Mann-Whitney U-Test	Univariate regression high HPV methylation vs low/medium methylation			ROC curve	
		Control subject (n=11)	Case patient (n=12)	Difference		OR	95% CI	P	AUC	95% CI
L1	7045	0.03	0.76	0.74	.0007	23.33	2.00 to 273.30	.0121	0.94	0.81 to 1.00
L1	7066	0.08	0.83	0.76	.0007	23.33	2.00 to 273.30	.0121	0.94	0.81 to 1.00
L1	7072	0.03	0.67	0.64	.0012	23.33	2.00 to 273.30	.0121	0.92	0.77 to 1.00
L1	7088	0.09	0.63	0.55	.0039	10.50	1.36 to 81.05	.0241	0.87	0.72 to 1.00
L1	7123	0.03	0.77	0.75	.0006	23.33	2.00 to 273.30	.0121	0.94	0.82 to 1.00
L1	7129	0.18	0.85	0.67	.0006	23.33	2.00 to 273.30	.0121	0.94	0.85 to 1.00
L1	7135	0.17	0.81	0.64	.0003		NA		0.97	0.91 to 1.00
URR	7573	0.00	0.00	0.00	.0126	15.75	1.42 to 174.00	.0246	0.78	0.60 to 0.95

* Logistic regression was used to calculate odds ratios (ORs) and 95% confidence intervals (CIs) based on a dichotomized result for methylation (the highest tertile vs middle and low tertiles combined), comparing the case patients and control subjects. ROC curves and areas under the curve with 95% CIs were calculated for distributions of methylation levels in case patients and control subjects. All CpG sites in HPV45 with at least 0.2 median difference, or at least 10.0 odds ratio (OR), or at least 0.80 area under the curve (AUC) are displayed. Bold *P* values are below 0.00048, the level of statistical significance after correction for multiple infections using Bonferroni's correction. NA indicates that there were no case patients with methylation levels in the lowest tertile. This generates a 0 in the two-by-two table, creating an infinite odds ratio for that site that cannot be computed. All statistical tests were two-sided.

Table 4. Correlation of methylation levels in HPV genes across genotypes*

Type	E6/E7		E2		L2		L1		URR	
	Median	IQR	Median	IQR	Median	IQR	Median	IQR	Median	IQR
31	0.36	0.28	0.75	0.64	0.70	0.25	0.77	0.24	0.30	0.08
18	0.77	0.09	0.67	0.20	0.73	0.20	0.62	0.32	0.59	0.11
45	0.87	0.10	0.72	0.39	0.65	0.56	0.80	0.35	0.62	0.47

* Median Spearman rank coefficients and interquartile ranges were computed for each genotype. IQR= interquartile range.

were merely a sign of epigenetic instability, we would expect a random distribution of methylated CpG sites throughout the HPV genome rather than the sharp contrast between URR-E1 methylation and E2-L1 methylation in case patients related to all three types (and HPV16 in our previous study). Although the data in this analysis are cross-sectional and do not allow us to analyze the temporal relationship between methylation and progression to CIN3, in our previous study, methylation of HPV16 was detected years before clinical detection of the case patients, suggesting that the temporality criterion of causality is fulfilled (6). It is possible that HPV methylation in the L1/L2 regions influences viral gene expression patterns toward increased expression of viral oncogenes, but functional studies are required to elucidate the causal relationship between methylation and viral gene expression. We cannot exclude that the increased methylation of viral genomes in CIN3 is merely an indirect indicator of alterations of the cellular methylation machinery. However, there is no evidence for general genome-wide methylation changes in CIN3 supporting this hypothesis (15).

We observed differences of methylation levels in women with CIN3 and infection with multiple HPV types compared with those with single type infections. For HPV31, we found higher methylation levels in women with single HPV31 infections compared with women who had a co-infection of HPV31 with HPV16. Multiple HPV infections are common and complicate attribution of carcinogenic types to cervical disease (16). Usually, one carcinogenic

type causes a CIN3, while other carcinogenic types may be because of independent, transient infections. Because HPV31 methylation levels were lower when HPV16 was concomitantly present, it is likely that HPV16 was the causal type in some of these CIN3 cases and HPV31 was present in an unrelated infection with lower methylation levels. If confirmed, measuring viral methylation could provide important insight into HPV genotype attribution, which currently can only be studied with complex tissue-based approaches. Precise estimates of HPV genotype attribution to cervical disease have translational importance for evaluating the impact of preventive HPV vaccines and to decide which genotypes should be included in HPV DNA screening tests.

The consistent findings of increased methylation levels in women with CIN3 for four important carcinogenic types suggest that HPV methylation may serve as a biomarker for detection of prevalent or incident cervical precancer. This finding may have important implications for cervical cancer screening approaches. Large randomized trials have proven that testing for carcinogenic HPV DNA can be effectively used as a primary screening test (17–19). Several countries are moving to replacing the Pap smear as primary screening test by HPV DNA assays. However, a major challenge is finding an efficient management strategy for HPV-positive women, because the vast majority of HPV infections are transient and will not cause precancer or cancer (2). A 10% risk of CIN3 over 2 years is currently considered to be the approximate threshold for referral to colposcopy (2).

Figure 4. Median methylation levels in case patients with single and multiple infections. For HPV31, median methylation levels are shown for the complete genome stratified by single infections (red line), multiple infections with HPV16 (green line), and multiple infections without HPV16 (orange line). Percent methylation is indicated on the y-axis, and CpG site positions are indicated on the x-axis. The panel below the graph shows

the Mann-Whitney U-test *P* values for median differences between case patients and control subjects in the three strata. The color scale ranges from low (dark green) to high (dark red) *P* values. The distribution of statistically significant *P* values across the HPV31 genome was similar for case patients with single HPV31 infections, and case patients with multiple infections without HPV16. All statistical tests were two-sided.

The range of best AUC values for individual types was from 0.81 for HPV31 to 0.98 for HPV45. Assuming a prevalence of CIN3 of about 5%–10% in oncogenic HPV-positive women and requiring sensitivity of 90% for an assay to triage HPV-positive women, the corresponding specificity values should be at least 50% and up to greater than 90%. With these performance characteristics, positive predictive values between 10%–50% can be achieved in a HPV triage setting. Thus, even based on AUC values from single sites, measuring HPV methylation can achieve clinical performance comparable to currently used tests to triage HPV-positive women. Because of the high correlation of CpG sites across HPV genomes, measuring a few discriminative sites per HPV genotype will provide a good representation of overall methylation status. Because HPV methylation can be measured from the same specimen used for the screening the HPV test, a reflex triage of HPV-positive women using a multitype methylation assay could serve as a specific and reliable marker to identify women who need to undergo further evaluations in colposcopy.

Few previous studies have reported data on HPV18 methylation in cervical carcinogenesis. High levels of methylation have been reported for a limited set of HPV18 L1 CpG sites in cervical cancer cell lines and in primary cervical cancers, but few precancerous samples have been evaluated previously with inconsistent results

(20–22). There is no previous data available for the association of HPV31 and HPV45 methylation with cervical precancer.

Our study had several strengths. The population was drawn from a large screening population that underwent co-testing with cytology and HPV DNA detection. We performed methylation analysis covering most of the viral genome in a large set of individuals using highly reliable state-of-the-art methylation assays with high reproducibility. As the PaP cohort study progresses over time, we will be able to study prediagnostic samples of incident cases. We are also currently expanding the assays to cover more carcinogenic genotypes.

This study also had some limitations. Despite the large sample size, the cases and controls available for HPV45 were limited. Although nested in a large cohort, our analysis at this point was purely cross-sectional, and it did not allow us to study prediagnostic methylation levels. However, the main translational applications, resolving type attribution and detecting CIN3 among HPV-positive women, do not require prospective testing. Because intermediate disease groups were not included in our study, some specificity measures may be inflated. However, our data strongly support development of an HPV methylation assay to detect cervical precancer. We acknowledge the limited availability of covariate data in our cohort, eg, on smoking and nutritional factors. While these factors have

been discussed to be related to DNA methylation, most previous studies have shown no or only marginal associations between HPV methylation at individual sites and various potential risk factors (23,24). Further, there are only few established risk factors for the progression from HPV infection to precancer, and they have relatively low ORs: the relative risk for progression is less than double for smoking. It is unlikely that confounding related to a weak risk factor will explain the strong associations observed in our analysis.

In conclusion, we demonstrate that HPV methylation is substantially different among women with HPV31, 18, 45 infections and CIN3 compared with women with the same genotype but without histological or cytological evidence of high-grade CIN. Methylation patterns were similar among closely related HPV genotypes. These findings may help improve our understanding of the molecular basis for viral persistence and increase our ability to identify women who are HPV-positive who are more likely to progress to cervical cancer. Measuring carcinogenic HPV DNA methylation could serve as a specific marker for infections with a high risk of cancer development.

References

- Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, Wacholder S. Human papillomavirus and cervical cancer. *Lancet*. 2007;370(9590):890–907.
- Schiffman M, Wentzensen N, Wacholder S, Kinney W, Gage JC, Castle PE. Human papillomavirus testing in the prevention of cervical cancer. *J Natl Cancer Inst*. 2011;103(5):368–383.
- McCredie MR, Sharples KJ, Paul C, Baranyai J, Medley G, Jones RW, et al. Natural history of cervical neoplasia and risk of invasive cancer in women with cervical intraepithelial neoplasia 3: a retrospective cohort study. *Lancet Oncol*. 2008;9(5):425–434.
- Doorbar J. Papillomavirus life cycle organization and biomarker selection. *Dis Markers*. 2007;23(4):297–313.
- Wentzensen N, von Knebel DM. Biomarkers in cervical cancer screening. *Dis Markers*. 2007;23(4):315–330.
- Mirabello L, Sun C, Ghosh A, Rodriguez AC, Schiffman M, Wentzensen N, et al. Methylation of the HPV16 genome is associated with disease progression in a prospective population-based cohort. *J Natl Cancer Inst*. 2012;104:556–565.
- Li N, Franceschi S, Howell-Jones R, Snijders PJ, Clifford GM. Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: variation by geographical region, histological type, and year of publication. *Int J Cancer*. 2011;128(4):927–935.
- Castle PE, Shaber R, LaMere BJ, Kinney W, Fetterma B, Poitras N et al. Human papillomavirus (HPV) genotypes in women with cervical precancer and cancer at Kaiser Permanente Northern California. *Cancer Epidemiol Biomarkers Prev*. 2011;20(5):946–953.
- Castle PE, Schiffman M, Gravitt PE, Kendall H, Fishman S, Dong H, et al. Comparisons of HPV DNA detection by MY09/11 PCR methods. *J Med Virol*. 2002;68(3):417–423.
- Sun C, Reimers LL, Burk RD. Methylation of HPV16 genome CpG sites is associated with cervix precancer and cancer. *Gynecol Oncol*. 2011;121(1):59–63.
- Chen Z, Desalle R, Schiffman M, Herrero R, Burk RD. Evolutionary dynamics of variant genomes of human papillomavirus types 18, 45, and 97. *J Virol*. 2009;83(3):1443–1455.
- Vinokurova S, Wentzensen N, Kraus I, Klaes R, Driesch C, Melsheimer P, et al. Type-dependent integration frequency of human papillomavirus genomes in cervical lesions. *Cancer Res*. 2008;68(1):307–313.
- Chen Z, Schiffman M, Herrero R, Desalle R, Anastos K, Segondy M, et al. Evolution and taxonomic classification of human papillomavirus 16 (HPV16)-related variant genomes: HPV31, HPV33, HPV35, HPV52, HPV58 and HPV67. *PLOS ONE*. 2011;6(5):e20183.
- Curradi M, Izzo A, Badaracco G, Landsberger N. Molecular mechanisms of gene silencing mediated by DNA methylation. *Mol Cell Biol*. 2002;22(9):3157–3173.
- Wentzensen N, Sherman ME, Schiffman M, Wang SS. Utility of methylation markers in cervical cancer early detection: appraisal of the state-of-the-science. *Gynecol Oncol*. 2009;112(2):293–299.
- Wentzensen N, Schiffman M, Dunn T, Zuna RE, Gold MA, Allen RA, et al. Multiple human papillomavirus genotype infections in cervical cancer progression in the study to understand cervical cancer early endpoints and determinants. *Int J Cancer*. 2009;125(9):2151–2158.
- Bulkman NW, Berkhof J, Rozendaal L, van Kemenade FJ, Boeke AJ, Bulk S, et al. Human papillomavirus DNA testing for the detection of cervical intraepithelial neoplasia grade 3 and cancer: 5-year follow-up of a randomised controlled implementation trial. *Lancet*. 2007;370(9601):1764–1772.
- Mayrand MH, Duarte-Franco E, Rodrigues I, Walter SD, Hanley J, Ferenczy A, et al. Human papillomavirus DNA versus Papanicolaou screening tests for cervical cancer. *N Engl J Med*. 2007;357(16):1579–1588.
- Naucle P, Ryd W, Tornberg S, Strand A, Wadell G, Elfgrén K, et al. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. *N Engl J Med*. 2007;357(16):1589–1597.
- Badal S, Badal V, Calleja-Macias IE, Kalantari M, Chuang LS, Li BF, et al. The human papillomavirus-18 genome is efficiently targeted by cellular DNA methylation. *Virology*. 2004;324(2):483–492.
- Fernandez AF, Rosales C, Lopez-Nieva P, Grana O, Ballestar E, Ropero S, et al. The dynamic DNA methylomes of double-stranded DNA viruses associated with human cancer. *Genome Res*. 2009;19(3):438–451.
- Turan T, Kalantari M, Calleja-Macias IE, Cubie HA, Cuschieri K, Villa LL, et al. Methylation of the human papillomavirus-18 L1 gene: a biomarker of neoplastic progression? *Virology*. 2006;349(1):175–183.
- Ding DC, Chiang MH, Lai HC, Hsiung CA, Hsieh CY, Chu TY. Methylation of the long control region of HPV16 is related to the severity of cervical neoplasia. *Eur J Obstet Gynecol Reprod Biol*. 2009;147(2):215–220.
- Xi LF, Jiang M, Shen Z, Hulbert A, Zhou XH, Lin YY, et al. Inverse association between methylation of human papillomavirus type 16 DNA and risk of cervical intraepithelial neoplasia grades 2 or 3. *PLOS ONE*. 2011;6(8):e23897.

Notes

The research was funded by the Intramural Research Program of the National Cancer Institute. The funding source had no influence on the design, conduct, and evaluation of the study.

Affiliations of authors: Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD (NW, AG, LM, SW, MC, MS); Department of Epidemiology and Population Health, at Albert Einstein College of Medicine, Bronx, NY (CS, RDB); Laboratory for Conservation and Utilization of Bio-Resources, and Key Laboratory of Microbial Diversity in Southwest China, Ministry of Education, Yunnan University, Kunming, P.R. China (CS); Division of Gynecologic Oncology, Kaiser Permanente Medical Care Program, Oakland, CA (WK); Care Management Institute, Kaiser Permanente, Oakland, CA (RS); Women's Health Research Institute, Division of Research, Kaiser Permanente, Oakland, CA (BL); Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary, University of London, London, UK (ATL); American Society of Clinical Pathology (PEC).