Identification of Human papillomavirus type 16 integration sites in high-grade precancerous cervical lesions

Mihaela Matovina[a], Ivan Sabol[a], Goran Grubišić[b], Nina Milutin Gašperov[a], Magdalena Grce[a]*

[a] Rudjer Boskovic Institute, Department of Molecular Medicine, Zagreb, Croatia

[b] University Hospital "Sestre milosrdnice", Clinic of Obstetrics and Gynecology, Zagreb, Croatia

***Corresponding author:**

Bijenicka cesta 54, HR-10002 Zagreb, Croatia

Phone: +385 1 456 1110

Fax: +385 1 456 1010

E-mail: grce@irb.hr

Type of Article: Original Research Report.

Preliminary results described in this manuscript were presented as posters at the 22nd International Papillomavirus Conference & Clinical Workshop, April 30 to May 6, 2005, Vancouver, BC, Canada, and the 23rd International Papillomavirus Conference & Clinical Workshop, September 1-7, 2006, Prague, Czech Republic, and as an oral presentation at the 12th World Congress on Advances in Oncology & 10th International Symposium on Molecular Medicine, October 11-13, 2007, Hersonisos, Crete, Greece.

Abstract

Objectives. Infection with oncogenic human papillomaviruses (HPV) is a prerequisite for the development of cervical cancer. In many cases of cervical cancer and all cervical cancer derived cell lines oncogenic HPV DNA is found to be integrated, indicating the importance of integration in disease development. In this study, 176 HPV 16 positive precancerous cervical lesions were analyzed for the physical state of viral genome to determine the sites of integration into a host cell DNA and to evaluate the incidence of the integration in different stages of cervical lesions.

Methods. The detection of integrated papillomavirus sequences (DIPS) method in combination with the amplification by polymerase chain reaction (PCR) of E1/E2 region was used to identify the physical state of HPV 16 genome. The site of integration within a host cell genome was determined by sequencing of DIPS unusual sized amplicons.

Results. The combined results of DIPS and E1/E2 PCR revealed the integration of HPV 16 DNA in 7.4% samples. The integration was found only in high grade cervical lesions indicating that it is a late event in disease progression. Sequencing of 11 DIPS amplicons revealed HPV DNA from six samples (54.5%) to be integrated in cellular genes (*VMP1*, *PVRL1*, *CHERP*, *CEACAM5*, *AHR*, *MRF-2*) and also 6 (54.5%) within the common fragile sites (CFS).

Conclusions. Although, the HPV integration is known to be a random event, this study indicates that HPV 16 integrates more than chance within or close to CFSs. As most of the genes affected by HPV 16 integration can be linked with some aspects of tumor formation, this indicates that the site of HPV DNA integration might play a role in the rate and the nature of tumor development.

Keywords: Human papillomaviruses (HPV); viral genome integration; common fragile sites (CFS); polymerase chain reaction (PCR); E1/E2 amplification; detection of integrated papillomavirus sequences (DIPS)

Introduction

Infection with oncogenic human papillomaviruses (HPV) is a prerequisite for the development of cervical cancer [1]. HPV DNA is detected in almost every (99.7%) cervical cancer sample analyzed [2] with HPV 16 being the most common type of HPV found in cervical cancer samples [3]. Cervical cancer develops through several precancerous stages, namely, cervical inthraepithelial neoplasia grades 1 to 3 (CIN 1-3). Cervical cancer usually develops through several decades and only in a small fraction of women infected with oncogenic HPV types, which indicates involvement of other factors, besides HPV infection. However, other risk factors involved in the development of cervical cancer remain unknown. The integration of viral DNA in the genome of human cells is considered as one of the steps that lead to the progression of the precancerous lesion into cancer. Moreover, there are several consequences of viral integration that could lead to tumor progression. It has been established that integration of viral DNA in the cellular genome provides selective growth advantage to cervical epithelial cells [4]. Upon integration viral E2 transcription repressor gene is often lost or disrupted leading to deregulation of E6 and E7 oncogene transcription [5-7]. In addition, viral transcription from integrated viral DNA can be deregulated as the result of cis-acting host sequences [8] or differences in chromatin structure at the promoter site [9], again leading to an increased viral oncogenes production. The recent studies suggest that integration may also have additional effect on the progression of precancerous lesions to cancer by insertional mutagenesis [10].

The site of integration within genome is random, although preferential integration in common fragile sites (CFS) have been proposed [11]. However, since these sites are prone to breaks and different kinds of genetic rearrangements the preferential integration within CFS is considered merely to be the consequence of their facile accessibility [12].

The goal of this study was to determine the frequency and the sites of integration within the human genome in the precancerous lesions of the cervix ranging from atypical and mild to high grade lesions.

Material and methods

Clinical samples

Cervical samples were collected from women attending the outpatient unit of the Clinic of Gynecology and Obstetrics of the University Hospital "Sestre Milosrdnice" (Zagreb, Croatia). Cervical swabs were taken from women with cytological abnormalities determined by Pap-smear for the purpose of HPV detection and genotyping, according to the Croatian protocol for the management of women with precancerous cervical lesions [13]. A subset (N = 176) of HPV 16 positive samples determined by the previously established protocol of HPV DNA detection and typing [14] was used in this study. Examinees were 17 to 50 years old (mean age 30 years). The cytological diagnosis of cervical specimens was determined according to the Croatian classification of cervical smears "Zagreb 2002" [15] that is in line with the "2001 Bethesda" system for reporting Pap test results. In the study group there were 39 (22.2%) women with atypical squamous cells of unknown significance (ASCUS), 32 (18.2%) with CIN 1 (also nominated as low-grade squamous intraepithelial lesions, i.e. LSIL), 46 (26.1%) with CIN 2 (classified as high-grade squamous intraepithelial lesions, i.e. HSIL), 48 (27.3%) with CIN 3 (also classified as HSIL), and 11 (6.3%) with an unknown diagnosis. The study was approved by the Ethical Board of Rudjer Boskovic Institute.

DNA isolation

Cervical swabs were resuspended in lysis buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA, pH 7.9, 0.5 % SDS) and treated overnight at 37°C with proteinase K (100 μg/ml).

Standard phenol-chloroform extraction and ethanol precipitation were used for DNA purification, and pelleted DNA was resuspended in 50-100 μL of tridistillated sterile water and stored at -20°C. In order to determine the quality and the quantity of isolated DNA, each DNA was analyzed by electrophoresis on 1% agarose gels stained with ethidium bromide and spectrophotometrically.

DIPS PCR

To determine the physical state of HPV 16 DNA and site of integration within the cell genome the DIPS (detection of integrated papillomavirus sequences) [16] PCR (polymerase chain reaction) based method with our modifications was used. Briefly, 0.6 μg of sample DNA was digested with *Sau3A I* or *Taq I* enzymes at 37°C by overnight incubation. Enzyme was inactivated by heating for 20 minutes at 65°C or 80°C for *Sau3A I* or *Taq I* enzyme, respectively.

After digestion, 50 pmol of adapter oligonucleotides AS-*Taq I* (17 bp sequence: CGCAACGTGTAAGTCTG-NH2) or AS-*Sau3A I* (19 bp sequence: GATCCAACGTGTAAGTCTG-NH2) annealed to AL oligonucleotide (45 bp: GGGCCATCAGTCAGCAGTCGTAGCCGGATCCAGACTTACACGTTG) were ligated to the digested sample DNA in a 20 μl reaction with 2 U T4 DNA ligase (Roche) and incubated at 16°C for one hour.

Seven downstream and 8 upstream HPV 16 type-specific primers were used for the linear amplification of HPV 16 DNA, (Table 1). PCR reaction was performed in 20 μl reactions, containing 1 μl ligated sample DNA, 2 μl 10X PCR buffer, 0.2 mM of each dNTP 2.0 mM $MgCl_2$, 0.25 μM of primer and 0.125 U Ex-Taq DNA polymerase (Takara). The relative positions of PCR primers on the HPV genome are presented on Figure 1. Following cycling conditions for amplification were used: denaturation for 2 minutes at 94°C, 25 cycles

of 30 second denaturation at 94°C, 45 second annealing at 60°C and 2 minutes elongation at 72°C increased by 1 second per cycle. The PCR reaction was followed by final elongation for 5 minutes at 72°C. The amplification was performed in Gene Amp PCR System 2400 (Applied Biosystems).

A second exponential PCR was performed using 16 nested HPV primers (Table 1) and an adapter specific primer, AP (sequence: GGCCATCAGTCAGCAGTCGTAG). PCR reaction was performed in 20 μl containing 1 μl product of the first reaction, 0.25 μM of HPV specific and adapter specific primers, with the remaining reagents being identical to the first reaction. PCR cycling conditions were the same as for the first reaction. For some samples it was impossible to determine the genome state so a second restriction assay with *Taq I* was performed. Linear amplification of HPV DNA containing restricted fragments was performed using 6 downstream (DIPS1/DIPS1N, DIPS2/DIPS2N, DIPS3/DIPS3N, DIPS8/DIPS8N, DIPS9/DIPS9N, DIPS10/DIPS10N) and 6 upstream primers (DIPS4/DIPS4N, DIPS5/DIPS5N, DIPS6/DIPS6N, DIPS7/DIPS7N, DIPS11/DIPS11N, DIPS12/DIPS12N) (Table 1 and Figure 1). The PCR reaction conditions and cycling parameters for both linear and exponential reactions remain the same. For extrachromosomal HPV 16 it is possible to predict the size of all produced PCR amplicons as positions of restriction enzyme cutting are known; *Sau3A I* cuts at nucleotide positions 525, 622, 871, 3480, 4364, 4523, 4541, 5073, 5237, 6154, 6954, 7017 and *Taq I* at position 505 for the referent HPV16 sequence and additionally at positions 2608 or 3118 for HPV 16 variants (Figure 1).

The PCR amplicons were resolved by agarose gel electrophoresis (1 or 2%, depending on the amplicon size); those deviating from the expected size were sequenced to determine if they are generated from integrated HPV genomes.

PCR amplification of E1/E2 gene of HPV 16 DNA

The E1/E2 PCR was performed according to Yoshinouchi et al. [17] and Lukaszuk et al. [18], who describe the protocols for E1 and E2 gene amplification, respectively. Briefly, 2 pairs of PCR primers were used: E1A (ATGTTACAGGTAGAAGGGCG; 5' nt 1293) with E1B (TGCTGCCTTTGCATTACTAG; 5' nt 1463) and E2A (TGCACCAACAGGATGTATAA; 5' nt 3066) with E2B (TCAACTTGACCCTCTACCAC; 5' nt 3193) amplifying 189 and 147 bp regions of the E1 and E2 genes, respectively. The reactions were performed in 20 μL, containing 100 ng sample DNA, 2 μL 10x PCR buffer, 0.2 mM each dNTP, 2.0 mM $MgCl_2$, 0.25 μM each primer and 0.125 U Taq DNA-polymerase. The following amplification conditions were used: denaturation for 10 minutes at 95°C, 35 cycles of 30 second denaturation at 95°C, 45 second annealing at 55°C and 1 minutes elongation at 72°C increased by 1 second per cycle. The PCR reaction was followed by final elongation for 5 minutes at 72°C. The amplicons were resolved by 2% gel electrophoresis.

Sequence analysis

The DIPS amplicon sizes different from expected were sequenced to determine the place of integration in the host cell genome. PCR product was excised from 1 to 2% agarose gel, depending on the size of the amplicon, and purified using QIAquick PCR Purification Kit according to the manufacturer's instructions (Qiagen GmbH). Five hundred ng of sample was sequenced with the forward and reverse primers. The BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems) was used. The products were analyzed with ABI PRISM 3100 Genetic Analyzer (Applied Biosystems) using Sequencing Analysis 5.1 program (Applied Biosystems). The obtained sequences were compared with sequences in the NCBI database (http://www.ncbi.nlm.nih.gov/) using BLAST program from the same website to determine if they contain human DNA sequences. Human sequences from the integrated

samples were further analyzed in NCBI-Map Viewer (http://www.ncbi.nlm.nih.gov/mapview/) or UCSC Genome Bioinformatics (http://genome.ucsc.edu/) to determine the actual chromosome and the region in which HPV 16 was integrated.

Statistical analysis

The results obtained by different methods were entered and processed manually into Microsoft Excel on a personal computer. The differences between study groups were evaluated by Chi-square (χ^2) test (GraphPad Prism version 4.00 for Windows, GraphPad Software, San Diego California USA), with p value under 0.05 being statistically significant.

Results

Using the DIPS method, 176 HPV 16 positive DNA samples from patients with different grades of cervical neoplasia were analyzed. Using *Sau3A I* restriction enzyme, it was impossible to determine the HPV genome state in all samples. Samples in which the analysis with *Sau3A I* enzyme failed or gave undetermined result were additionally analyzed with *Taq I* restriction enzyme digestion.

If the HPV genome is present in the extrachromosomal form in the sample, DIPS PCR should amplify DNA of expected size. Failure to amplify DNA or amplification of different fragment size than expected suggests that the HPV genome might be integrated. To verify this assumption such PCR products were sequenced. In 10 of 176 samples (5.7%), sequencing confirmed the presence of only HPV DNA, even though the PCR product was of an unusual size. It appears that those samples contained HPV 16 variants with internal rearrangements and/or gain or loss of *Sau3A I* or *Taq I* restriction site (data not shown). Using the described methods, the status of HPV genomes was determined in 146 of 176 (82.9%) samples, 8 of

which (5.5%) were shown to be integrated, 3 (2%) were shown to contain both extrachromosomal and integrated forms, and the remaining 135 (92.5%) had extrachromosomal HPV genomes.

Both *Sau3A I* and *Taq I* based DIPS assays failed to resolve the HPV genome status of 30 of 176 (17.1%) samples. In 23 of 30 samples, the amplification of large DIPS amplicons was unsuccessful due to unknown reasons, but most likely because of the fragmentation (degradation) of sample DNA as the short amplicons of E1/E2 PCR were successfully obtained from those samples. In addition, in 7 of 30 samples, only some of all target large DIPS amplicons were amplified and in 2 of those 7 samples there were amplicons suggesting integration of HPV 16 DNA. However, the sequencing of the DIPS amplicons supposedly containing human DNA sequences was unsuccessful, leaving those 7 samples with inconclusive result, also.

The unsuccessful amplification of E1 and/or E2 genes indicates whether those genes are disrupted because of the HPV genome integration into a host cell genome. In this study, the E1/E2 PCR analysis has shown the presence of integration in 9 of 176 (5.1%) samples. In 8 of 9 (88.9%) E2 region was disrupted, with the remaining one (11.1%) sample having E1 region disrupted.

The combined integration results obtained with both methods, DIPS and E1/E2 PCR reveal the integration in 13 of 176 (7.4%) samples. Table 2 shows the distribution of integrated or extrachromosomal genomes found in relation to the patient cytological diagnosis. There were significantly (χ^2=4.103, p=0.0428) more integration of HPV 16 genome in the study group with HSIL (CIN 2 and CIN 3) compared to the study group with LSIL (CIN 1) (Figure 2).

The discordance between the findings by the two methods, DIPS and E1/E2 PCR (Table 3) is due to the fact that E1/E2 PCR cannot distinguish between mixed and extrachromosomal viral genomes, and it cannot detect the integration event in samples where the loss of viral DNA is not within regions covered by the PCR primers used in this study. Thus, the three samples determined as mixed forms by DIPS PCR were detected as extrachromosomal form by E1/E2 PCR. In addition, one sample that was classified as integrated by DIPS PCR and extrachromosomal by E1/E2 PCR had a break in the viral DNA that was not in the regions amplified by E1/E2 PCR primers (data not shown). Finally, two samples classified as integrated by E1/E2 PCR did not give conclusive result with DIPS PCR.

The DIPS PCR method enabled the amplification of 11 (6.3%) fragments containing viral and cellular DNA. In 6 of those samples both ends of cellular/viral junction amplicon were successfully sequenced, while in 5 cases only one end of viral cellular/junction was resolved, so the information about affected HPV and cellular genes and scope of deletions was incomplete in those cases. Briefly, in 6 of 11 (54.5%) samples the integration was found to occur within cellular genes and also in 6 of 11 (54.5%) samples within CFS (Table 4).

The HPV 16 genome in the sample 2002299 was integrated in the human genome at the position 17q23 that is the site of *VMP1* gene, which codes for the transmembrane protein involved in cell apoptosis (= TMEM49; http://www.hprd.org/protein/15649).

In the sample 2002458, the HPV 16 genome was integrated in the human genome at the position 1q32.2 but no particular gene or CFS was determined in this region as only the 3'-end of integrated sequence was resolved.

The HPV 16 genome state in the sample 2002472 was mixed, extrachromosomal and integrated. Only the 5'-end of integrated sequence was resolved. The integration was located in the 17q12 region of the human genome. Query of the UCSC database revealed the position

of the integration to be 25 kb upstream of the *TCF2* gene, which codes for the TCF2 protein (http://www.hprd.org/protein/08926). However, since the 3' end of integrated sequence was not determined we cannot say for sure whether any gene was affected by this integration event.

The sample 2003194 contained the HPV 16 genome integrated at the position 11q23.3 of the human genome, specifically within the *PVRL1* (= nectin, = HSV entry mediator C) gene; the FRA11B and FRA11G sites being both within this region.

In the sample 2003217 HPV 16 DNA was integrated at the position 2q37.3 in the human genome, probably in the FRA2J site. The integration is not directly in any gene but several might be affected, namely *HES2* (10 kb downstream), *PER2* (20 kb downstream) and *ILKAP* (25-30 kb upstream).

The sample 2003379 contains HPV 16 DNA integrated at the position 19p13.11 in the human genome in the FRA19B site. This particular integration is disrupting the coding region of the *CHERP* (= *DAN16*, = *SCAF6*; http://www.hprd.org/protein/13051) gene.

In the sample 2003416, the integration of HPV 16 genome occurred at the location 19q13.2 in the FRA19B site and within the *CEACAM5* (= *CEA*, = *CD66e*; http://www.hprd.org/protein/00274) gene.

The sample 2003417 contains HPV 16 genome in both, extrachromosomal and integrated forms. At the site of integration (1p34.3) there are no known genes and no known CFS. The nearest CFS are FRA1C (1p31.1) and FRA1A (1p36).

The sample 2004077 contains HPV 16 DNA integrated at the position 7p21.1. The integration did not occur in the CFS but upon integration a part of *AHR* (http://www.hprd.org/protein/02596) gene is deleted.

The sample 2004377 showed HPV 16 DNA to be integrated in 10q21.2 position in the human genome in the FRA10C site. The integration occurred in the noncoding regions of the *MRF-2* (= *ARID5B*, = *FLJ21150*) gene.

The sample 2004410 also contained both, integrated and extrachromosomal forms of HPV 16 DNA. The integration was found to be at the position 11p14.3, near the FRA11D site (11p14.2), but there are no known genes in this region.

Discussion

Most of HPV genome integration studies focused on HPV 16 and 18, the most common HPV types found in cervical cancer. In all cell lines derived from cervical cancer tissues HPV DNA is integrated in the host cell genome and this is the same in the majority of cervical cancers. The frequency of integration for cervical cancer also differs among studies, but most are around 80% [19,20]. Different HPV types also have different rates of integration with HPV 18 being integrated in almost 100% of cases, while for HPV 16 reported frequency of integration ranged from 72-83% [21,22].

Even though some studies have tried to determine the rate of integration of HPV DNA in low grade cervical neoplasia, the results are discrepant, and the frequencies of integration range from 0% [20,23] to more than 50% [24]. Thus, the goal of this study was to determine the physical state of HPV 16 genomes and the frequency of integration in samples of different grades of precancerous cervical lesions.

The DIPS method detected integrated HPV 16 DNA in 8 samples, and both, extrachromosomal and integrated forms of HPV 16 DNA in 3 samples. All samples containing integrated DNA were diagnosed as HSIL (CIN 2 or 3), except one sample with unknown diagnosis, indicating that integration in LSIL (CIN1) is a very rare event.

The E1/E2 PCR amplifying parts of those viral genes found integrated forms of HPV 16 DNA in 9 (5.1%) samples, 8 of which showed E2 region to be disrupted with one sample having E1 region disrupted. Similar results were obtained by Watts [25] finding 72.7% of E2 gene disruption in HPV 16 positive cervical cancer samples.

The E1/E2 PCR method is simple and easy to perform but it cannot be used for the determination of the site of integration and it cannot distinguish between extrachromosomal and mixed viral forms. Furthermore, this method does not detect integrated copies where the loss of viral DNA is outside of regions covered with PCR primers that were used in this study. These two issues were the main cause of discrepancy in the results obtained with both methods, DIPS and E1/E2 PCR.

The DIPS PCR method is rather time consuming and laborious, however it allows accurate determination of the site of integration in the human genome. On the other hand, the method is not very suitable for determining the frequency of integration, since it did not give us any conclusive result in 17.1% cases. For that reason, the results of both methods, DIPS and E1/E2 PCR are presented herein.

These combined findings indicate that integration of HPV 16 DNA is not an early event as Peitsaro et al. [26] and Gallo et al. [24] suggest, them having found 50% integrated HPV forms in LSIL samples. In this study, no integration of HPV DNA in LSIL samples was found, which is in line with other studies [20,22]. In addition, the frequency of HPV DNA integration in HSIL samples (i.e. 12.8% of CIN 2 and 13% of CIN 3 samples) is similar to those obtained by Klaes et al. [20], 5% and 15% of CIN 2 and CIN 3 samples, respectively. The discrepancy of the results for CIN 2 samples might be attributed to the inability of clearly distinguishing between CIN 2 and CIN 3 grades, both of which have been classified as HSIL in the recent cytological classification [27].

Apart from the frequency of integration, the main goal of this study was to determine the sites of integration within human genome by using the DIPS PCR method. The HPV 16 integration found in this study was random as expected; it occurred once on chromosomes 2, 7 and 10 and twice on chromosomes 1, 11, 17 and 19, but on different parts of chromosomes (Table 4). According to the majority of studies, the integration of HPV DNA is a random event and the place of integration does not affect the progression of the disease [10]. However, in this study HPV 16 integration occurred in many cases (54.5%) within CFS. This result is in accordance with Thorland et al. [29] who found HPV 16 integrated in the CFS in 48% of their cases. Some authors consider that HPV integration in CFS, which are prone to breaks and different types of rearrangements, is the result of their greater accessibility [10]. Supporting this view is the fact that observed HPV integration near common oncogenes, like *myc*, *TP63*, *FANCC*, *hTERT* and others, was not shown to affect the expression of those genes.

Yu et al. [28] created a model for HPV integration *in vitro*. They observed integration in the CFS in 50% of cases, some of which were identical to the observed integration sites for HPV 16 and 18 in cervical samples. They also believe that site of integration does play a role in the disease progression, and that during persistent infection HPV is integrated in many places in the genome, some of which lead to different expression of HPV oncoproteins and/or cellular genes conferring proliferative advantage to those cells and influencing disease progressions. To validate those assumptions, a large set of clinical samples with integrated genomes should be assayed and the disease progression should be closely monitored.

An interesting result to note is that 6 (54.5%) of the 11 integrated samples were integrated within genes and that 3 additional cases were located 10-30 kb from genes. Most of the genes that were affected by integration in our study could be linked to the neoplastic progression in one way or another. By studying the literature we found that some of those

sites were identified as targets of HPV integration in previous studies, and some of the affected genes were linked to cervical, as well as other types of cancer.

Namely, 3 cases of HPV 16 integration within the fragile site FRA17B [10] have been previously found in cervical cancer specimens and one of them was in the *VMP1* gene [29]. Furthermore, 5 integrations of HPV 18 in cervical cancer samples have been reported in the same region (17q23) [10]. This indicates that HPV integrations within region 17q23 are more frequent than chance would suggest, and potential disruption of the gene involved in apoptosis might be implicated in cervical cancer progression.

The integration in 17q12 region was not in the fragile site, but was in proximity of the *TCF2* gene. As the 3'-end of the integrated fragment was not resolved it is possible that the *TCF2* gene is affected. The TCF2 protein is a transcription factor involved in the metabolism of nucleic acids and nucleotides. It has been extensively studied, and one study indicates that it could be involved in the development of ovarian cancer [30]. In addition, Wentzensen et al. [10] also reported HPV16 to be integrated in this region (17q12) in one cervical cancer sample.

The HPV 16 integration within *PVRL* gene might also have an impact on progression of cancer. The PVRL1 is a transmembrane protein involved in cellular adhesion and Herpes Simplex virus entry [31,32]. No HPV genome integrations have been reported in this region previously, but loss of heterozigosity (LOH) of 11q23 region is common in tumors, indicating the presence of an unknown tumor suppressor gene in this region [33-36]. Lazo [37] reports LOH of 11q23 region in 38% (73/188) of cervical cancer samples. Kersemaekers et al. [36] report finding a strong correlation between LOH of 11q23 region and invasive cervical cancer. This indicates that the 11q23 region might contain metastasis suppressor gene, and since the *PVRL1* gene affects cellular adhesion, it is a potential candidate for that role.

The integration in FRA2J (2q37.3 locus) was mapped in the proximity of the *ILKAP* gene, which codes for a serine/threonine phosphatase from the phosphatase family PP2C. The ILKAP protein interacts with integrin linked kinases (ILK), which regulate integrin cellular signaling, and it is also associated with Wnt signaling pathway [38,39]. There is no published data on the HPV integration in this locus of the human genome, but again, some data suggest that LOH of 2q might influence cervical cancerogenesis [40].

The integration in FRA19B (19p13.11 locus) was found to be directly disrupting the coding region of the *CHERP* gene, which codes for a ribonucleoprotein situated in the endoplasmic reticulum and the cytoplasm, and it functions as a negative regulator of cellular proliferation [41,42]. There are no published data regarding the HPV integration in the 19p13.11 locus.

The integration of HPV 16 in FRA19A (19q13.2 locus) was mapped to *CEACAM5* gene. The CEACAM5 protein is a transmembrane adhesion molecule that also has a role in immune responses and apoptosis [43]. Many tumors have an elevated expression of CEACAM5 and CEACAM6 proteins, which are important for adhesion, invasion and metastasis in tumors. By blocking these proteins with monoclonal antibodies it is possible to stop the metastasis process [44,45]. Two integration events in this gene have been published, HPV 16 in a cervical cancer specimen and HPV 18 in a CIN 3 sample [10].

The integration on the short arm of chromosome 1, at 1p34.3 was not mapped to any gene or fragile site, the nearest fragile sites being FRA1C (1p31.1) and FRA1A(1p36). However, one integration event in that region has already been published [10].

Another HPV 16 integration outside of CFS was mapped to the *AHR* gene (7p21.1 locus), which codes for a transcription factor involved in the metabolism of nucleic acids and nucleotides that regulates the cellular response to aromatic hydrocarbons. It has been

suggested that it also has a role in tumorgenesis and proliferation. The AHR protein represses *c-myc* [47]. Here again, the disruption of AHR function might influence cervical cancerogenesis.

The integration of HPV 16 in FRA10C (10q21.2 locus) was found to occur in the noncoding region of MRF-2 gene. This gene contains a DNA binding domain, and although its function is still unclear [48], it is known to be involved in the differentiation of smooth muscle [49].

In case of the sample containing both, integrated and extrachromosomal forms of HPV 16 DNA, the integration occurred at the position 11p14.3, in the proximity of the FRA11D. As there are no known genes in this region it is difficult to predict any changes due to the viral integration.

The findings of this study show that HPV 16 genome integration is not an early event in the progression of cervical disease, and that the preferential regions of integration are the common fragile sites. This study also shows HPV genomes to often be integrated in or near cellular genes. Since most of those genes could be involved in tumor progression, this suggests that the site of integration might play a role in the rate and the nature of development of cervical cancer. By investigating the sites of HPV DNA integration we might also get some insight about cellular genes that may be involved in the progression of cervical cancer and also potentially some other types of cancer. Further research is necessary to finally resolve the timing and preferential sites of HPV integration. The real-time PCR on large number of precancerous samples would easily resolve the timing issue, while the DIPS PCR analysis of samples with integrated viral genome would answer whether there are indeed preferential integration sites. In addition, since HPV integration status indicates high risk for development and progression of cervical cancer, it should be considered for a diagnostic approach. In this case also, a method based on real-time PCR that is less time consuming and informative for

both integrated and extrachromosomal states, as shown in this study to be common, would be more appropriate.

Conflict of interests

All the authors declare that there are no conflicts of interest.

Acknowledgements

The authors are especially grateful to Mrs. Jasminka Golubić Talić for her expert technical assistance and Dr. Helena Četković for her sequencing analysis. This study was supported by the Croatian Ministry of Science, Education and Sport (grant number 098-0982464-2510).

References

[1] IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Human Papillomaviruses. WHO Press 2007;90:1-689.

[2] Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJ, Peto J, Meijer CJ, Munoz N. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol 1999;189:12-9.

[3] Bosch FX, Manos MM, Muñoz N, Sherman M, Jansen AM, Peto J, Schiffman MH, Moreno V, Kurman R, Shah KV, IBSCC Study group. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. J Natl Cancer Inst 1995;87:796-802.

[4] Jeon S, Allen-Hoffmann BL, Lambert PF. Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells. J Virol 1995;69:2989-97.

[5] Romanczuk H, Howley P. Disruption of either the E1 or the E2 regulatory gene of human papillomavirus type 16 increases viral immortalization capacity. Proc Natl Acad Sci USA 1992;89:3159-63.

[6] Dowhanick JJ, McBride AA, Howley PM. Suppression of cellular proliferation by the papillomavirus E2 protein. J Virol 1995;69:7791-7.

[7] Jeon S, Lambert PF. Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of E6 and E7 mRNAs: implications for cervical carcinogenesis. Proc Natl Acad Sci USA 1995;92:1654-8.

[8] von Knebel Doeberitz M, Bauknecht T, Bartsch D, zur Hausen H. Influence of chromosomal integration on glucocorticoid-regulated transcription of growth-

stimulating papillomavirus genes E6 and E7 in cervical carcinoma cells. Proc Natl Acad Sci U S A. 1991;88:1411-5.

[9] Bechtold V, Beard P, Raj K. Human papillomavirus type 16 E2 protein has no effect on transcription from episomal viral DNA. J Virol 2003;77:2021-8.

[10] Wentzensen N, Vinokurova S, von Knebel Doeberitz M. Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract. Cancer Res 2004;64:3878-84.

[11] Thorland EC, Myers SL, Persing DH, Sarkar G, McGovern RM, Gostout BS, Smith DI. Human papillomavirus type 16 integrations in cervical tumors frequently occur in common fragile sites. Cancer Res 2000;60:5916-21.

[12] Smith DI, Huang H, Wang L. Common fragile sites and cancer. Int J Oncol 1998;12:187-96.

[13] Ljubojević N, Babić S, Audy-Jurković S, Ovanin-Rakić A, Jukić S, Babić D, Grubišić G, Radaković B, Ljubojević-Grgec D. Improved national Croatian dignostic and therapeutic guidelines for premalignant lesions of the uterine cervix with some cost-benefit aspects. Coll Antropol 2001;25:467-74.

[14] Grce M, Husnjak K, Bozikov J, Magdic L, Zlacki M, Lukac J, Fistonic I, Sikanic-Dugic N, Pavelic K. Evaluation of genital human papillomavirus infections by polymerase chain reaction among Croatian women. Anticancer Res 2001;21:579-84.

[15] Ovanin-Rakic A, Pajtler M, Stankovic T, Audy-Jurkovic S, Ljubojevic N, Grubisic G, Kuvacic I. [Classification of cervical cytological smears "Zagreb 2002". Modification of the classification "Zagreb 1990" and "NCI Bethesda 2001".] Gynaecol Perinatol 2003;12:148-53 [In Croatian].

[16] Luft F, Klaes R, Nees M, Durst M, Heilmann V, Melsheimer P, von Knebel Doeberitz M. Detection of integrated papillomavirus sequences by ligation-mediated PCR (DIPS-PCR) and molecular characterization in cervical cancer cells. Int J Cancer 2001;92:9-17.

[17] Yoshinouchi M, Hongo A, Nakamura K, Kodama J, Itoh S, Sakai H, Kudo T. Analysis by multiplex PCR of the physical status of human papillomavirus type 16 DNA in cervical cancers. J Clin Microbiol 1999;37:3514-7.

[18] Lukaszuk K, Liss J, Wozniak I, Emerich J, Wojcikowski C. Human papillomavirus type 16 status in cervical carcinoma cell DNA assayed by multiplex PCR. J Clin Microbiol 2003;41:608-12.

[19] Kalantari M, Blennow E, Hagmar B, Johansson B. Physical state of HPV16 and chromosomal mapping of the integrated form in cervical carcinomas. Diagn Mol Pathol 2001;10:46-54.

[20] Klaes R, Woerner SM, Ridder R, Wentzensen N, Duerst M, Schneider A, Lotz B, Melsheimer P, von Knebel Doeberitz M. Detection of high-risk cervical intraepithelial neoplasia and cervical cancer by amplification of transcripts derived from integrated papillomavirus oncogenes. Cancer Res 1999;59:6132-6.

[21] Badaracco G, Venuti A, Sedati A, Marcante ML. HPV16 and HPV18 in genital tumors: Significantly different levels of viral integration and correlation to tumor invasiveness. J Med Virol 2002;67:574-82.

[22] Cullen AP, Reid R, Campion M, Lorincz AT. Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasm. J Virol 1991;65:606-12.

[23] Hopman AH, Smedts F, Dignef W, Ummelen M, Sonke G, Mravunac M, Vooijs GP, Speel EJ, Ramaekers FC. Transition of high-grade cervical intraepithelial neoplasia to micro-invasive carcinoma is characterized by integration of HPV 16/18 and numerical chromosome abnormalities. J Pathol 2004;202:23-33.

[24] Gallo G, Bibbo M, Bagella L, Zamparelli A, Sanseverino F, Giovagnoli MR, Vecchione A, Giordano A. Study of viral integration of HPV-16 in young patients with LSIL. J Clin Pathol 2003;56:532-6.

[25] Watts K. Sequence variation and physical status of human papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia. Int J Cancer 2002;97:868-74.

[26] Peitsaro P, Johansson B, Syrjanen S. Integrated Human papillomavirus type 16 is frequently found in cervical cancer precursors as demonstrated by a novel quantitative Real-Time PCR technique. J Clin Microbiol 2002;40:886-91.

[27] Nayar R, Solomon D. Second edition of 'The Bethesda System for reporting cervical cytology' – atlas, website, and Bethesda interobserver reproducibility project. Cytojournal 2004;1:4.

[28] Yu T, Ferber MJ, Cheung TH, Chung TK, Wong YF, Smith DI. The role of viral integration in the development of cervical cancer. Cancer Genet Cytogenet 2005;158:27-34.

[29] Thorland EC, Myers SL, Gostout BS, Smith DI. Common fragile sites are preferential targets for HPV16 integrations in cervical tumors. Oncogene 2003;22:1225-37.

[30] Tsuchiya A, Sakamoto M, Yasuda J, Chuma M, Ohta T, Ohki M, Yasugi T, Taketani Y, Hirohashi S. Expression profiling in ovarian clear cell carcinoma: identification of

hepatocyte nuclear factor-1β as a molecular marker and a possible molecular target for therapy of ovarian clear cell carcinoma. Am J Pathol 2003;163:2503-12.

[31] Hoshino T, Sakisaka T, Baba T, Yamada T, Kimura T, Takai Y. Regulation of E-cadherin endocytosis by nectin through afadin, Rap1, and p120ctn. J Biol Chem 2005;280:24095-103.

[32] Krummenacher C, Baribaud F, Ponce de Leon M, Baribaud I, Whitbeck JC, Xu R, Cohen GH, Eisenberg RJ. Comparative usage of herpesvirus entry mediator A and nectin-1 by laboratory strains and clinical isolates of herpes simplex virus. Virology 2004;322:286-99.

[33] Hampton GM, Penny LA, Baergen RN, Larson A, Brewer C, Liao S, Busby-Earle RM, Williams AW, Steel CM, Bird CC, Stanbridge EJ, Evans GA. Loss of heterozygosity in cervical carcinoma: subchromosomal localization of a putative tumor-suppressor gene to chromosome 11q22-q24. Proc Natl Acad Sci USA 1994;91:6953-7.

[34] Bethwaite PB, Koreth J, Herrington CS, McGee JO. Loss of heterozygosity occurs at the D11S29 locus on chromosome 11q23 in invasive cervical carcinoma. Br J Cancer 1995;71:814-8.

[35] Mullokandov MR, Kholodilov NG, Atkin NB, Burk RD, Johnson AB, Klinger HP. Genomic alterations in cervical carcinoma: losses of chromosome heterozygosity and human papilloma virus tumor status. Cancer Res 1996;56:197-205.

[36] Kersemaekers AM, Hermans J, Fleuren GJ, van de Vijver MJ. Loss of heterozygosity for defined regions on chromosomes 3, 11 and 17 in carcinomas of the uterine cervix Br J Cancer 1998;77:192-200.

[37] Lazo P. The molecular genetics of cervical carcinoma. Br J Cancer 1999;80:2008-18.

[38] Tong Y, Quirion R, Shen SH. Cloning and characterization of a novel mammalian PP2C isozyme. J Biol Chem 1998;273:35282-90.

[39] Leung-Hagesteijn C, Mahendra A, Naruszewicz I, Hannigan GE. Modulation of integrin signal transduction by ILKAP, a protein phosphatase 2C associating with the integrin-linked kinase, ILK1. EMBO J 2001;20:2160–70.

[40] Narayan G, Pulido HA, Koul S, Lu XY, Harris CP, Yeh YA, Vargas H, Posso H, Terry MB, Gissmann L, Schneider A, Mansukhani M, et al. Genetic analysis identifies putative tumor suppressor sites at 2q35-q36.1 and 2q36.3-q37.1 involved in cervical cancer progression. Oncogene 2003;22:3489-99.

[41] Laplante JM, O'Rourke F, Lu X, Fein A, Olsen A, Feinstein MB. Cloning of human Ca2+ homoeostasis endoplasmic reticulum protein (CHERP): regulated expression of antisense cDNA depletes CHERP, inhibits intracellular Ca2+ mobilization and decreases cell proliferation. Biochem J 2000;348:189-99.

[42] O'Rourke FA, LaPlante JM, Feinstein MB. Antisense-mediated loss of calcium homoeostasis endoplasmic reticulum protein (CHERP; ERPROT213-21) impairs Ca2+ mobilization, nuclear factor of activated T-cells (NFAT) activation and cell proliferation in Jurkat T-lymphocytes. Biochem J 2003;373:133-43.

[43] Wirth T, Soeth E, Czubayko F, Juhl H. Inhibition of endogenous carcinoembryonic antigen (CEA) increases the apoptotic rate of colon cancer cells and inhibits metastatic tumor growth. Clin Exp Metastasis 2002;19:155-60.

[44] Baczynska D, Wietrzyk J, Madej J, Krop-Watorek A, Dabrowska A, Widerak K, Opolski A, Ugorski M. The tumorigenic potential of human CX-1 colon adenocarcinoma cells depends on carcinoembryonic antigen (CEACAM5) expression. Cell Mol Biol Lett 2003;8:471-86.

[45] Blumenthal RD, Hansen HJ, Goldenberg DM. Inhibition of adhesion, invasion, and metastasis by antibodies targeting CEACAM6 (NCA-90) and CEACAM5 (Carcinoembryonic Antigen). Cancer Res 2005;65:8809-17.

[46] Andersson P, McGuire J, Rubio C, Gradin K, Whitelaw ML, Pettersson S, Hanberg A, Poellinger L. A constitutively active dioxin/aryl hydrocarbon receptor induces stomach tumors. Proc Natl Acad Sci USA 2002;99:9990-5.

[47] Yang X, Liu D, Murray TJ, Mitchell GC, Hesterman EV, Karchner SI, Merson RR, Hahn ME, Sherr DH. The aryl hydrocarbon receptor constitutively represses c-myc transcription in human mammary tumor cells. Oncogene 2005;24:7869-81.

[48] Zhu L, Hu J, Lin D, Whitson R, Itakura K, Chen Y. Dynamics of the Mrf-2 DNA-binding domain free and in complex with DNA. Biochemistry 2001;40:9142-50.

[49] Watanabe M, Layne MD, Hsieh CM, Maemura K, Gray S, Lee ME, Jain MK. Regulation of smooth muscle cell differentiation by AT-rich interaction domain transcription factors Mrf2α and Mrf2β. Circ Res 2002;91:382-9.

Article précis

Affected cell host genes by the integrated HPV 16 sequences can be linked to different aspects of cervical tumorigenesis.

Table 1

Sequences of primers and their location on the HPV 16 genome

Linear PCR			Exponential PCR			Expected amplicon size			
Primer	5' nucleotide	5' → 3' Sequence	Primer	5' nucleotide	5' → 3' Sequence	Primer used	5' nucleotide	Target Sau3A I restriction site	Episomal HPV 16 amplicon size (bp)
DIPS1	901	ACGGGATGTAATGGATGGTTTTATG	DIPS1N	952	AGGGGATGCTATATCAGATGACGAG	DIPS1/DIPS1N	952	3480	2573
DIPS2	1293	ATGTTACAGGTAGAAGGGCG	DIPS2N	1336	AGTCAGTATAGTGGTGGAAGTG	DIPS2/DIPS2N	1336	3480	2189
DIPS3	1870	ACGCCAGAATGGATACAAAGACAAAC	DIPS3N	1942	ATGGTACAATGGGCCTACGATAATG	DIPS3/DIPS3N	1942	3480	1583
DIPS8	3028	GTGGACATTACAAGACGTTAGCCTTG	DIPS8N	3092	CATGGATATACAGTGGAAGTGCAG	DIPS8/ DIPS8N	3092	3480	433
DIPS9	3702	CGTCTACATGGCATTGGACAGG	DIPS9N	3767	GATAGTGAATGGCAACGTGACC	DIPS9/DIPS9N	3767	4521	804
DIPS9A**	3486	CCACATCTATACCTTCATATGC	DIPS9AN**	3523	CGTCTGTGTTTCTTCGGTGC	DIPS9A/DIPS9AN	3523	4521	642
DIPS10	5286	CCACTTTACATGCAGCCTCACC	DIPS10N	5385	CTGTACCCTCTACATCTTTATCAGG	DIPS10/DIPS10N	5385	6154	814
DIPS4	3302	ACCCGCATGAACTTCCCATAC	DIPS4N	3193	TCAACTTGACCCTCTACCAC	DIPS4/DIPS4N	3213	872	2386
DIPS4A**	3480	CTTGGTCGCTGGATAGTCG	DIPS4AN**	3463	CGTCTGTGTTTCTTCGGTGC	DIPS4A/DIPS4AN	3463	872	2591
DIPS5	3991	AGAGGCTGCTGTTATCCACAATAG	DIPS5N	3942	ATGTAGACACAGACAAAAGCAGC	DIPS5/DIPS5N	3942	3480	529
DIPS5A**	4361	CAGCAATAGTTTTGCCTTCAACC	DIPS5AN**	4340	CCTTAGGTATAATGTCAGGTGG	DIPS5A/DIPS5AN	4340	3480	905
DIPS6*	5130	GTACGCCTAGAGGTTAATGCTGG	DIPS6N*	5087	CCAAAAAGTCAGGATCTGGAGC	DIPS6/DIPS6N	NA	NA	NA
DIPS6A**	5021	AGCCAGACACCGGAAACC	DIPS6AN**	4979	GTTTAGTGGGAGTGGTTACAAAAGC	DIPS6A/DIPS6AN	4979	4541	483
DIPS7	6062	ATCCACACCTGCATTTGCTGC	DIPS7N	6037	GCACTAGCATTTTCTGTGTCATCC	DIPS7/DIPS7N	6043	5235	853
DIPS11	6956	GATCTTCTTTAGGTGCTGGAGG	DIPS11N	6893	CCTATAAGTATCTTCTAGTGTGCC	DIPS11/DIPS11N	6893	6154	784
DIPS12	379	GTATTGCTGTTCTAATGTTGTTCC	DIPS12N	240	GCAAAGTCATATACCTCACGTCG	DIPS12/DIPS12N	240	7017	1173

* This primer was not used in *Sau3A I* assay due to the presence of restriction site in the primer sequence

** These primers were not used in *Taq I* assay

***HPV 16 is cut only once by *Taq I* enzyme at position 505, however there are HPV 16 variants with additional *Taq I* sites at position 2608 and 311

Table 2

HPV genome status according to patient cytological diagnosis

Status	Unknown N (%)	ASCUS N (%)	CIN 1 N (%)	CIN2 N (%)	CIN 3 N (%)	HSIL N (%)	Total N
Episomal	10 (90.9)	38 (97.4)	32 (100)	40 (87)	43 (89.6)	83 (88.3)	163 (92.6)
Integrated	0 (0)	1 (2.6)	0 (0)	5 (10.9)	4 (8.3)	9 (9.6)	10 (5.7)
Mixed	1 (9.1)	0 (0)	0 (0)	1 (2.2)	1 (2.1)	2 (2.1)	3 (1.7)
Total	11 (6.3)	39 (22.2)	32 (18.2)	46 (26.1)	48 (27.3)	94 (53.4)	176 (100)

Table 3

Comparison of E1/E2 PCR and DIPS PCR in determining HPV 16 integration

		E1/E2 PCR		
		Episomal	Integrated	Total N (%)
DIPS PCR	Episomal	135	0	135 (76.7)
	Integrated	1	7	8 (4.5)
	Mixed	3	0	3 (1.7)
	ND	28	2	30 (17.1)
	Total N (%)	167 (94.9)	9 (5.1)	176 (100)

Table 4

Genomic sites of HPV 16 DNA integration

Sample Number	Diagnosis	HPV 16 DNA form	Deletion in HPV DNA		Cellular DNA deleted	Integration site	Gene	CFS
			5'-end in HPV DNA	3'-end in HPV DNA				
2002299	CIN 2	Integrated	3024 (E2)	1067 (E1)	~7.5 kb	17q23	*VMP1*	FRA17B
2002458	CIN 2	Integrated	ND	2383 (E1)	ND	1q32.2	/	/
2002472	CIN 2	Mixed	6633 (L1)	ND	ND	17q12	/	/
2003194	CIN 2	Integrated	2837 (E2)	2170 (E1)	9 bp	11q23.3	*PVRL1*	FRA11B and FRA11G
2003217	CIN 3	Integrated	3394 (E2)	1906 (E1)	13 bp	2q37.3	/	FRA2J
2003379	CIN 3	Integrated	3902 (E5)	2897 (E2)	~5 Mb	19p13.11	*CHERP*	FRA19B
2003416	CIN 3	Integrated	ND	2085 (E1)	ND	19q13.2	*CEACAM5*	FRA19A
2003417	CIN 3	Mixed	3705 (E2)	2309 (E1)	~3.4 kb	1p34.3	/	/
2004077	CIN 3	Integrated	~ 3440 (E2)	3404 (E2)	ND	7p21.1	*AHR*	/
2004377	CIN 2	Integrated	5544 (L2)	ND	ND	10q21.2	*MRF-2*	FRA10C
2004410	ND	Mixed	ND	ND	ND	11p14.3	/	FRA11D

Figure 1 - Position of *Sau3A I* and *Taq I* restriction sites and DIPS primers on the HPV 16 genome

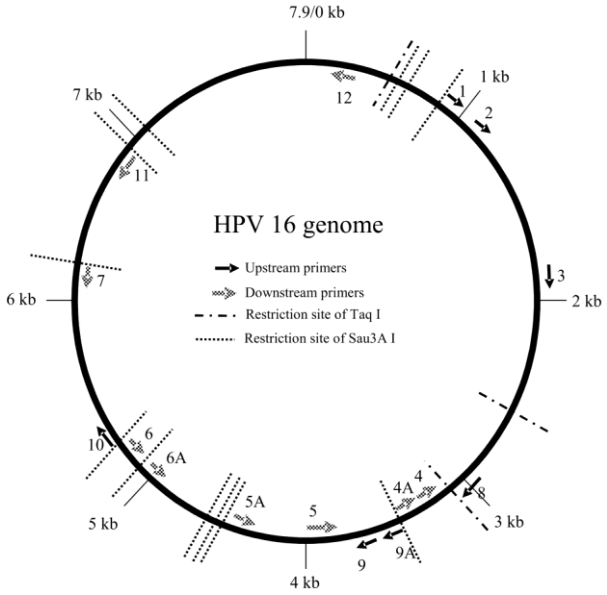

Figure 2 – Association of cervical lesions and HPV genome status