INTRODUCTION

Human papillomavirus type 16 (HPV16) has been classified into Alphapapillomavirus-9 species belonging to Papillomaviridae family. It is the most common type detected in early and malignant cervical lesions in the world as well as in some HPV-related cancers of the vulva, vagina, penis, anus, and oropharynx (Crow, 2012; Garbuglia, 2014; Bruni et al., 2015).

Persistent infection with oncogenic human papillomavirus genotypes is essential, although not sufficient, to cause invasive cervical cancer and cervical intraepithelial neoplasia as its precursor (Walboomers et al., 1999; Muñoz et al., 2003).

The distribution of HPV genotypes in various populations and areas is different and the epidemiological studies suggest that certain HPV16 E6 variants could influence the persistence of infection, progression to pre-cancer and cancer development (Londesborough et al., 1996). A distinct HPV type is confirmed when the DNA sequence of the L1 ORF of the cloned viral genome differs from that of any other characterised type by at least 10% but isolates of the same HPV type are named as subtypes when the nucleotide sequences of the L1 ORF differ by less than 10% (de Villiers et al., 2004; Bernard et al., 2010). Cornet and coworkers (2012) proposed to define major variant lineages based on approximately 1.0% difference between full genomes of the same HPV type while variant sublineages by 0.5–0.9% differences. Variants of E6 HPV16 were divided into four distinct phylogenetic branches, the distribution of which varies geographically: lineage A – European-Asian, including the sublineages European (EUR) and Asian (As); lineage B – African 1 (Afr1) with Afr1a and Afr1b sublineages; lineage C – African 2 (Afr2) including Afr2a and Afr2b sublineages and lineage D – Asian American/North American including Asian American 1, Asian American 2 and North American 1 sublineages (Burk, 2013, Cornet et al., 2012).

Lee and coworkers (2008) proposed that nucleotide polymorphism within the E6 sequence of HPV16, rather than in other ORFs, influences the progression of cervical malignancy, and that additional factors are likely to play a significant role as well.

Integration of viral DNA into host cell genome is a frequent event in carcinogenesis. The HPV genome integration usually disrupts E2 gene, leading to overexpression of E6 and E7 oncogenes.

The aim of the present study was to identify HPV 16 E6 variants and to analyse the physical status of HPV16 genome (integrated and/or episomal form of viral DNA) among patients with cervical cancer in comparison to women with low-grade squamous intraepithelial lesions (LSIL).
MATERIALS AND METHODS

Specimens and DNA extraction. The study was performed on eighty HPV16 positive cervical cytobrush samples confirmed by the INNO-LiPA HPV genotyping assay (Innogenetics). All HPV isolates included in the present study came from the collection of samples obtained from women (aged 22–90 years, mean 42±15) with invasive cervical carcinoma (n=40) and low-grade squamous intraepithelial lesions – LSIL (n=40). The LSIL group was dominated by younger women (up to 40 years of age) and women over the age of 50 were predominant in the cervical cancer group (p<0.001). All samples were taken after cytological examination, but before any treatment. All cancer patients were treated at the Maria Sklodowska-Curie Memorial Cancer Centre and Institute of Oncology, Cracow Branch, Poland. The study was approved by the Ethics Committee of the Jagiellonian University.

DNA was isolated from cervical smears and purified using Genomic DNA Prep Plus kit (A&A Biotechnologies, Poland) according to the manufacturer’s instructions.

HPV 16 E6 variants identification. Determination of HPV16 E6 variants was performed using PCR with type-specific E6 primers according to Zuna and coworkers (2009) and followed by DNA sequencing.

The final PCR reaction contained 50 ng of the isolated DNA, 1x PCR Buffer II (Applied Biosystems, USA), 2 mM MgCl₂, 200 µM dNTPs, 2 µM each primer and 1 U AmpliTag Gold DNA polymerase (Applied Biosystems, USA). The specific primers, flanking the encoding E6 open reading frame (nt 83 – 559) comprised: 5’ – TGA ACC GAA ACC GGT TAG TA – 3’ and 5’ – CAT GCA ATG TAG GTG TAT CTC C – 3’. The PCR conditions were as follows: polymerase activation 3 min at 94°C, 40 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 1 min, and extension at 72°C for 2 min, and a final elongation at 72°C for 7 min. For every reaction, positive control using DNA extracted from SiHa cell line, as well as a negative control without template DNA (H₂O), were included.

The amplified DNA fragments were tested by electrophoresis in 2% agarose gel with ethidium bromide. The PCR products were sequenced by Genomed (Poland).

The ExoSAP-IT clean up kit (USB, USA) was used to purify the products and DNA sequences were obtained with an ABI 3730xl DNA Analyzer or 3130xl Genetic Analyzer (Applied Biosystems, USA) using the same primers as those used for PCR and BigDye™ Terminator V3.1 Cycle Sequencing Kits (Applied Biosystems, USA). The same forward and reverse primers were used separately in cycle sequencing in order to sequence both sense and antisense strands.

Nucleotide sequences of the complete E6 ORF from 80 studied samples were compared with the HPV16 E6 prototype sequence (GenBank KO2718). Table 1 shows variant analysis for the E6 gene. A total of 12 different intratypic variants were identified and were classified into two different lineages (A and D), including European (EUR) and the North-American 1 (NA1) sublineages.

Among the sequences belonging to the European sublineage, 26 out of 77 (34%) were identical with the prototype sequence (EUR-350T). Moreover, there were only 3 samples with the EUR-350T sequence with an additional single nucleotide changes, at position T137G (n=2), resulting in amino acid change (L12V), and at T109C without amino acid change (silent mutation), in one case. EUR-350T was present in 32.5% (13/40) of cervical cancer samples and in 16/40 (40%) of the LSIL group, but the differences were not statistically significant. The EUR-350G variants were the most frequently detected isolates in our population and represented 60% (48/80) of total HPV16 samples, with similar frequency in cervical cancer and LSIL groups. The NA1 sublineage was detected in 4% (3/80) of samples, only in the cervical cancer group. Distribution of HPV16 E6 variants according to different sublineages in LSIL and carcinoma groups was not statistically different (Pearson’s chi-squared, p=0.19).

Sequence analysis of 80 studied samples showed substitutions in 9 nucleotides located between positions 122 and 350 (G122A, T137G, G145T, G153T, G176A, G188A, T295G, C335T, T350G) leading to the amino acid change (E7K, L12V,Q14H, T17I, D25N, E29K, D64E, H78Y, L83V) while other nucleotide substitutions (T109C, C256T, T286A, A289G) were silent. Moreover, for the first time, unknown missense mutations G122A, C153T and G188A were recorded along with EUR-350G (Table 1). Only in two cases, the amino acid changes (Q14H and T17I) were non-conservative, while other changes were conservative. The most prevalent non-synonymous variant was L83V (T350G). It was found in 64% (51/80) of samples in both the European and NA1 sublineages.
either alone (47%, 24/51) or along with other nucleotide changes (53%, 27/51). The prevalence of T350G was similar in cervical cancer and LSIL groups and amounted to 67.5% and 60%, respectively (Pearson’s chi-squared, $p = 0.48$). Missense mutations in nucleotides other than EUR-350G were detected in individual samples. Among the synonymous variants T109C dominated (16%, 13/80).

In samples of the European lineage, frequencies of additional mutations (synonymous and non-synonymous) in variants with EUR-350T or EUR-350G sequences differed significantly and occurred in 10% (3/29) and 50% (24/48) of samples, respectively (Pearson’s chi-squared, $p = 0.0004$). Furthermore, additional mutations (excluding T350G) in these samples were observed more often in LSIL than in cervical cancer, in 47.5% and 21.6% respectively (Pearson’s chi-squared, $p = 0.02$) (Fig. 1). The number of all mutations, including three missense mutations and two silent ones, calculated per sample was highest in the NA1 sublineage variant. In European variants, only one or two nucleotide changes were observed. The average number of mutations per sample was not statistically significant when comparing cancer and LSIL patients (Mann–Whitney U test, $p = 0.29$). However, silent mutations were pointed out in the LSIL group significantly more often (Pearson’s chi-squared, $p = 0.003$) as shown in Fig. 1. A multivariate logistic regression analysis showed that the presence of synonymous mutations was correlated with a six-fold (OR 6.1; 95% CI 1.55 – 24.00; $p = 0.003$) increase in the probability of the occurrence of the LSIL diagnosis. The occurrence of mutation for the entire group (N=80) did not correlate with women’s age (Mann–Whitney U test, $p = 0.84$).

Table 1. HPV16 E6 variants and their lineage identified in 80 patients with cervical cancer and low–grade squamous intraepithelial lesions.

<table>
<thead>
<tr>
<th>Nucleotide position</th>
<th>HPV16 prototype</th>
<th>Lineage</th>
<th>Sublineage</th>
<th>Class/subclass</th>
</tr>
</thead>
<tbody>
<tr>
<td>109</td>
<td>T</td>
<td>A</td>
<td>EUR-p</td>
<td>EUR-350T</td>
</tr>
<tr>
<td>122</td>
<td>G</td>
<td>EUR-350T/109C</td>
<td>G</td>
<td>1</td>
</tr>
<tr>
<td>137</td>
<td>T</td>
<td>EUR-350/109C</td>
<td>T</td>
<td>2</td>
</tr>
<tr>
<td>145</td>
<td>G</td>
<td>EUR-350G/109C</td>
<td>G</td>
<td>4</td>
</tr>
<tr>
<td>153</td>
<td>C</td>
<td>EUR-350G/122A</td>
<td>A</td>
<td>1</td>
</tr>
<tr>
<td>176</td>
<td>G</td>
<td>EUR-350G/153T</td>
<td>G</td>
<td>2</td>
</tr>
<tr>
<td>188</td>
<td>G</td>
<td>EUR-350G/176A</td>
<td>T</td>
<td>1</td>
</tr>
<tr>
<td>256</td>
<td>C</td>
<td>EUR-350G/188A</td>
<td>A</td>
<td>0</td>
</tr>
<tr>
<td>286</td>
<td>T</td>
<td>EUR-350G/295G</td>
<td>G</td>
<td>3</td>
</tr>
<tr>
<td>289</td>
<td>G</td>
<td>EUR-350G/256T</td>
<td>G</td>
<td>4</td>
</tr>
<tr>
<td>295</td>
<td>C</td>
<td>EUR-350G/286</td>
<td>T</td>
<td>8</td>
</tr>
<tr>
<td>335</td>
<td>T</td>
<td>EUR-350G/295G</td>
<td>G</td>
<td>3</td>
</tr>
<tr>
<td>350</td>
<td>T</td>
<td>EUR-350G/295G</td>
<td>G</td>
<td>3</td>
</tr>
</tbody>
</table>

Figure 1. Distribution of additional nucleotide changes (without T350G) in the European variants of samples with LSIL and cervical carcinoma taking into account synonymous and non-synonymous mutations. The differences of mutation number between groups were significant (Pearson’s chi-squared, $p = 0.02$).
Detection of HPV16 physical status

The integrated HPV16 DNA was identified in 24 (30%) samples, mixed form (episomal and integrated in the same sample) in 45 (56.3%) and episomal form in 11 (13.7%) samples. Physical status of HPV16 genome related to the studied groups (cervical cancer and LSIL) was significant (Fig. 2). The integrated and mixed forms of HPV16 DNA were statistically more often found in cervical cancer than in the LSIL group (48.7% vs 37.5% respectively; V-chi-squared test, \( p=0.004 \)). Among the isolates containing the episomal form of HPV16 DNA, 91% were defined as LSIL.

In women with LSIL, the episomal form of HPV16 DNA was present in 25%, while the mixed form in 72% of cases. Only one material (2.5%) was demonstrated to have integrated HPV16 genome. The integrated viral genome was predominant in women with cervical cancer (23/40, 57.5%). The mixed form of HPV16 DNA was found in 40% of women with cancer, whereas the episomal form was shown only in one patient.

Association of HPV16 E6 polymorphisms with viral genome physical state

In NA1 variant mixed form of HPV16 DNA was identified. Figure 3 shows occurrence of EUR-350T and EUR-350G variants in relation to the physical status of HPV16 DNA. The EUR-350G variant dominated in mixed form while EUR-350T was slightly more often present as the integrated form. These differences were statistically significant (Pearson’s chi-squared test, \( df=2, p<0.0001 \)).

As regards the European variant, all mutations (non-synonymous as well as synonymous) were most frequently detected in mixed form compared to integrated form, 68.63% (35/51) and 25.49% (13/51), respectively. Only 3 samples with episomal forms of HPV16 genome contained the nucleotide changes (Pearson’s chi-squared test, \( p=0.0007 \)). In multivariate analysis, the distribution of the different variants depending on the form of the HPV16 genome and cytological diagnosis showed that in the samples of cervical cancer with the integrated form, EUR-350T prototype and EUR-350G without additional mutations were the most frequently observed. These results were statistically significant and amounted to, respectively: 43% (10/23); Pearson’s chi-squared 13.1 (df=2), \( p=0.001 \) and 47% (11/23); Pearson’s chi-squared 10.9 (df=2), \( p=0.004 \). For the other variants, there was no such relationship. Figure 4 demonstrates the average number of nucleotide changes present in materials assigned to particular groups depending on the form of HPV16 genome. The percentage of synonymous and non-synonymous mutations in three forms of HPV16 genome was shown in Fig. 5.

DISCUSSION

HPV16 is the most commonly encountered high risk type of HPV in Polish women with cervical cancer. Diversity of the HPV16 E6 sequence was investigated in women from southern Poland. For the first time, in the present study, we also analysed the relationship between the polymorphism of HPV16 E6 gene and the viral genomic physical status in this patient group. It is widely accepted that the distribution of HPV variants is related to geographical or population ethnicity distribution (Xi et al., 2006; de Araujo Souza et al., 2009; Cornet et al., 2013a). In the present study, of 80 HPV16 positive women, European lineage variants (both prototype and related variant) were dominant (96%) in both LSIL and cervical cancer group. In 3 (4%) cases, only in the cervical cancer group, North American sublineage was detected. Similar results were obtained in Slovenian patients with cervical cancer, where mainly European genomic variations of HPV16 were found and only 2/40 (5%) of the cases were associated with non-European HPV16 genomic variants – the Asian-American and African 2 branches (Vrtačnik Bokal et al., 2010). Several studies showed that non-European HPV16 genomic variants (lineage B, C, D) are associated with stronger oncogenic potential than the European line-
both Burroni and coworkers (2013) in Italian women with LSIL and Gudleviciene and coworkers (2015) in Lithuanian women with cervical cancer revealed a single case of the North-American 1 variant, which is similar to our findings in the population of women from the south of Poland.

The most common European polymorphism in E6 gene is a nucleotide substitution at position 350 (350G instead of 350T) that leads to a protein change in codon 83, leucine to valine (L83V). Notably, the same E-350G nucleotide variation is also a characteristic finding in all Asian-American and North-American sequences (Zuna et al., 2009). In some previous studies in Danish, Swedish and French populations, it was shown that this substitution is associated with an increased risk of cervical cancer (Andersson et al., 2000; Zehbe et al., 2001; Grodzki et al., 2006). However, the association was not consistently observed in other studies (Cornet et al., 2013b; Freitas et al., 2014). In our study, the EUR-350G pattern occurred at a similar rate, 67.5% in cervical cancers and 60% in low grade lesions and these differences were not statistically significant. Similar results in women with cervical cancer were observed in Slovenia (62.5%) (Vrtačnik Bokal et al., 2010) but a lower frequency (49%) of EUR 350G was reported across Europe in a recent meta-analysis (Tornesello et al., 2011). In our research the incidence of EUR-350G variant in the LSIL group was similar to the results of studies in Greece (64%) and in Italy (65%) but higher than in England (41%) (Tsakogiannis et al., 2013; Burroni et al., 2013; Marongiu et al., 2014). In a recent study, the International Agency for Research on Cancer (IARC) HPV Variant Study Group revealed, by examining 1121 HPV16 positive cervical cancer cases and 400 HPV16 positive controls worldwide, that E-350G is associated with an increased risk of developing cervical cancer in South and Central America but not in Europe nor Central Asia (Cornet et al., 2013a). Furthermore, EUR-350T infections were more likely to persist and progress to CIN3 than in case of EUR-350G in Denmark (Gheit et al., 2011). Also Cornet and coworkers (2013b) demonstrated an approximately twofold risk of HPV16 persistence in the presence of EUR-350T. European variants with the polymorphism in EUR-350T were revealed to persist more often than those containing EUR-350G (OR=1.6, 95% CI=0.8-3.4).

In the previously conducted studies (Yamada et al., 1997; Vrtačnik Bokal et al., 2010) additional nucleotide changes (T137G, G176A, T295G) occurring in EUR-350G were identified only in individual cases similarly as in our study. The frequencies of additional mutations (synonymous and non-synonymous) in EUR-350G variants were significantly higher than in EUR-350T in the present examination. These findings inspired us to try to assess the correlation between the presence of these additional mutations and clinical diagnosis and the physical status of HPV16 DNA. Additional mutations excluding T350G were observed.
in our samples more often in LSIL than in cervical cancer. Taking into consideration the relationship between the total number of nucleotide changes and the physical state of HPV16, we most frequently observed mutations in the mixed form of viral genome, while non-synonymous mutations were predominantly found in mixed and integrated states.

In our study, analysis of HPV16 E6 sequences showed three new (not found in the BLAST program) missense mutations associated with T350G in the LSIL. (C153T, G122A) and cancer (G188A) mutation (Table 1). It seems that there is a large variability in HPV viruses circulating in the population affecting women with a low grade of cervical lesions, while in women with cancer, this variability is smaller. In the previous study we confirmed that together with the progression of changes in the direction of tumorigenesis, in women with multiple genotype HPV infection, there is a natural selection leading to infection with one genotype and it is usually HPV16 (Szostek et al., 2008a). Perhaps an analogical selection takes place also in case of mutated HPV16 strains along with the persistent infection and development of cancer.

Integration of the HPV genome in the host cell DNA represents a key step in the progression of cervical lesions, as it allows the continuous overexpression of E6 and E7 oncoproteins and their transforming activities via their interaction with p53 and pRB, respectively (zur Housen, 2002). In the present study, the integrated HPV16 form was significantly more frequently detected in cervical cancer than in LSIL. (Pearson’s chi-squared 26.9, p<0.001). These results are consistent with other authors (Han et al., 2015; Shukla et al., 2014). In our previous study, the increase of frequency of the integrated HPV16 DNA with progression of dysplastic lesions was shown (Szostek et al., 2008b, Szostek et al., 2011). Some authors showed relatively high prevalence of mixed and integrated forms of HPV16 DNA in cervical epithelial cells obtained from women with low grade cervical lesions (Peisaro et al., 2002; Kumala et al., 2006; Huang et al., 2008). In our study, 72.5% (29/40) of samples with mixed form of HPV16 and one sample with integration of viral genome were detected in the LSIL group. In research of Huang and coworkers (2008) mixed form was detected in 83.3% (5/6) of samples while in Peisaro and coworkers (2002) even in 93.3% (14 z 15) but in CIN II group. Furthermore, Kumala and coworkers (2006) identified dominant mixed form among women with persistent Pap smear abnormalities as well as among women who cleared their Pap smear abnormalities, 50% and 51.4%, respectively. Our findings and the results of previous studies strongly indicate that the integration of HPV16 may occur in the early stages of neoplastic transformation of the cervix and could be a prognostic factor of cervical cancer.

We found a high incidence of the mixed form of HPV16 both in cervical cancer and in LSIL, but we observed that only in samples from cervical cancer with integrated form of HPV 16 genome two specific variants: EUR-350T and EUR-350G, without additional mutations dominated. However, it should be noted that the group analysed by us consisted of only 23 patients. It was also observed that non-synonymous and synonymous mutations were significantly more frequent in the mixed form of HPV 16 genome and in LSIL group.

In conclusion, the results of our study showed that, in the studied population of women from southern Poland, mainly European variants of HPV16 E6 could be found. The EUR-350G variant was very common in cervical cancer as well as in low grade lesions. Interestingly, the additional nucleotide changes occurred significantly more often in 350G variants than in 350T variants. Thus, we concluded that the presence of 350G mutation could affect the number of additional nucleotide changes appearing in the studied HPV16 E6 gene sequence, but this was mostly observed in LSIL.

Our study suggested that 350T and 350G variants without additional mutations promoted viral integration to the host genome which can lead to cancer development. Therefore, it seems that the study of HPV16 E6 polymorphism may have prognostic significance in long-term monitoring of women.

Acknowledgements

This research was supported by KNOW Leading National Research Centre 2012 - 2017.

REFERENCES


