Human Papillomavirus 16 Variants May Be Identified by E6 Gene Analysis

Nerea Fontecha\textsuperscript{a}  Miren Basaras\textsuperscript{a}  Elixabete Arrese\textsuperscript{a}  Silvia Hernáez\textsuperscript{b}  Daniel Andía\textsuperscript{c}  Ramon Cisterna\textsuperscript{a, b}

\textsuperscript{a}Immunology, Microbiology and Parasitology Department, University of the Basque Country, Leioa, and \textsuperscript{b}Clinical Microbiology and Infection Control Department and \textsuperscript{c}Department of Obstetrics and Gynecology, Basurto University Hospital, Bilbao, Spain

\textbf{Abstract}

\textbf{Aims:} The aims of the study were (1) to characterize the genetic variability of human papillomavirus (HPV) genotype 16 in the \textit{E6} region when this genotype is present in multiple infection samples, (2) to assess the prevalence of variants in our region and (3) to analyze the relationship between variants, patients' ages and pathology. \textbf{Methods:} The Clinical Microbiology and Infection Control Department analyzed samples which were positive for genotype 16 and other genotypes from 2007 to 2013. Variants were assigned to European, Euro-German, Asian, Asian-American or African lineage by sequence analysis. The relationship among variants, age and different types of lesion was studied. \textbf{Results:} In HPV-16 sequence analysis, the European variant was detected in 85.10\% of samples, the Asian-American in 7.80\%, the African in 4.25\% and the Euro-German in 2.83\% of specimens. Sequence genetic variability showed 16 nucleotide substitutions. Moreover, non-European variants were mainly found in old women and in isolates from high-grade squamous intraepithelial lesions since European variants were mainly detected in negative cytologies. \textbf{Conclusion:} Multiple infections may take effect on nucleotide substitution and the appearance of recombinant samples. Single gene analysis makes it impossible to detect recombination which has a great influence on drug response and vaccine strategies. Thus, \textit{E6} gene analysis would be enough to identify HPV-16 intratypic variants but not to confirm the results.
ity. These variants show geographical and ethnic distribution; thereby, the Asian-American variant is mostly found in Central and South America, the Asian variant is principally detected in southeast Asia, the African in Africa and the European is the most prevalent variant in all other regions excluding Africa [6–8].

The HPV-16 genome contains 2 oncogenes, E6 and E7. The loss of regulation in these two genes is the cause of intraepithelial neoplasia development [9, 10]. Moreover, as it has been described by other authors, nucleotide modifications in this region may be related with the presence of more oncogenic variants [11].

Nowadays, multiple HPV genotype infections are being found in a great number of patients, and this may interfere in variant determination and could have an impact on HPV-16 nucleotide substitutions [12–14].

Most of the worldwide studies based their classification on the analysis of one genomic region (usually the L1 region) in single infection samples. This work examines multiple HPV infection samples by analyzing HPV-16 E6 oncogene genetic variability.

The aims of the present study were (1) to describe the genetic variability of HPV-16 in the E6 gene in multiple HPV infection samples and study whether it is enough to classify intratypic variants, (2) to assess the prevalence of papillomavirus variants in our region, (3) to analyze the relationship between these variants, patient age and type of lesion.

Materials and Methods

Study Population and Specimen Collection

The Clinical Microbiology and Infection Control Department at Basurto University Hospital (Basque Country, North of Spain) analyzed samples with clinical manifestations of HPV-associated infections from 2007 to 2013. Till HPV presence was analyzed, all samples were stored in their own standard transport medium at −20°C till amplification.

Samples from women were classified by pathologists into 5 groups following the Bethesda system: (1) negative (no lesion was found), (2) atypical squamous cells of undetermined significance, (3) low-grade squamous intraepithelial lesion, (4) high-grade squamous intraepithelial lesion and (5) squamous-cellular carcinoma. With respect to samples from men, they could be classified as first group according to the Bethesda system (negative) or as condyloma acuminatum, which was added as a sixth group.

Molecular genotyping was performed using a Linear Array HPV Genotyping Test kit (Roche Molecular Diagnostics).

HPV-16 is the most prevalent genotype among high-grade lesion patients. Moreover, infections with more than 1 HPV genotype are often found in HPV-positive patients. Thus, in the present study, samples which were positive for both HPV-16 and other genotypes (multiple HPV infection) were analyzed (144 out of 2,085 total samples). All patients gave a written and informed consent prior to their inclusion in the study. All procedures followed were approved by the appropriate Ethics Committee related to our institutions and complied with the guidelines and ethical standards for experimental investigation with human subjects of the Helsinki Declaration of 1975, as revised in 2000.

Genomic Viral DNA Extraction

DNA extraction was executed using a QIAamp DNA mini kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. Extracted DNA was eluted with 100 μl AE buffer and stored at −20°C till amplification.

PCR Amplification and Sequencing

E6 gene amplification was performed by type-specific primers. These E6-specific primers were designed conforming to the HPV-16 genome prototype sequence (GenBank accession No. K02718).

PCR amplification was completed in 30 μl reaction mixture containing 10× PCR buffer, 25 mmol/l MgCl₂, 25 mmol/l of each deoxynucleoside, 100 pmol/l of each primer, 5 μl of template DNA and 2.5 units of Taq DNA polymerase (Qiagen). The thermal profile commenced with a preheating of 95°C for 15 min, followed by 40 cycles with 55°C as annealing temperature and concluded with a final extension at 72°C for 10 min.

The presence of PCR fragments was verified by that of a 503-bp specific band in agarose gel (2%), and subsequently amplification products were sequenced using the Big Dye Terminator Cycle Sequencing kit (Applied Biosystems) conforming to the manufacturer’s instructions. A designed E6-specific forward primer was used for sequencing.

Genetic Variability and Phylogenetic Analysis

HPV sequences were aligned and compared to the HPV-16 prototype sequence which belongs to the European lineage (GenBank accession No. K02718), using Chromas Lite Sequence Alignment Editor v2.1.1 and Clustal W (http://www.genome.jp/tools/clustalw/). Sequences were designated to a branch as a result of their similarity to HPV-16 known variant sequences which belong to the Asian-American lineage (GenBank accession No. AF402678, AF402678, AFY68579), Asian (FJ006723), African (AF472508, AF472509, AF536380, AF536380), Euro-German (AF536179) and the European lineage (AY685680, AY685681, AY685684, EU118173). Phylogenetic trees were constructed using the maximum likelihood method implemented in MEGA software version 5 [15].

HPV-16 Variants versus Patient Age and Clinical Lesions

Patient age and type of lesion were correlated with each sample variant. These parameters were analyzed for their statistically significant associations using the χ² test.

Results

Patient Characteristics

Among the analyzed 2,085 samples, 144 patient specimens were positive for HPV-16 and other HPV genotypes (multiple HPV infection); 128 were from women, while we only detected HPV-16 multiple infections in 13...
men. The women’s average age was 30.84 ± 8.09 years and the men’s mean age was 43.50 ± 11.30 years at the time of sampling.

**Relationship between Variants and Lesion/Age**

The 144 studied samples were analyzed to determine each specimen variant. A total of 120 (85.10%) samples were described as European variant, 11 (7.80%) as Asian-American, 6 (4.25%) as African and 4 (2.83%) as Euro-German variant.

With regard to lesion development, in women, it was observed that samples with lesions were more frequent among non-European lineage specimens than among European variants (66.67 and 33.33%, respectively; fig. 1). This relation between isolate branch and type of lesion in female specimens showed statistical significance (p < 0.05).

On the other hand, in male specimens, only 1 sample out of 13 presented a low-grade lesion; the other 12 specimens were classified by the pathologist as normal. Even so, the number of samples from men was too low to perform any statistical analysis.

According to the age average of patients and the intratypic variant, among women, it was found that European

![Table 1. Most important nucleotide sequence variations in the E6 gene among HPV-16 isolates in European variant specimens](image)

**Fig. 1.** HPV-16 variant distribution in women according to pathology: E6 gene-related variants; the non-European variant includes Asian-American and African variants. Lesion: low-grade squamous intraepithelial lesion, high-grade squamous intraepithelial lesion and invasive cancer.
variants were present in younger women than non-European variants, the average age being 29.95 ± 7.43 years among women with the European variant and 36.50 ± 11.60 and 41.33 ± 5.68 years among women with Asian-American and African variant infection, respectively. The correlation between younger age and European lineage isolates was determined to be statistically significant as much as the relation between older age and non-European branch samples (p < 0.05). On the other hand, regarding men, sample number was very low, and thus it was impossible to perform statistical analysis.

E6 Gene Nucleotide Variations

E6 gene genomic analysis showed important nucleotide substitutions in 16 nucleotide positions. One deletion (A104del) and one nucleotide variation (T350G, L83V) were found in all lineages. In contrast, the A131G nucleotide variation was specific to Euro-German isolates (table 1). Moreover, 4 of these nucleotide substitutions were specific to both Asian-American and African branches: G145T, T286A, A289G and C335T (detected in all Asian-American and African isolates; table 2). Furthermore, Asian-American branch-specific variation (A532G) and two African lineage-specific nucleotide substitutions (T109C and G132C) were detected. Seven of these nucleotide variations are nonsynonymous substitutions and lead to amino acid modification.

Phylogenetic Analysis

Phylogenetic analysis of the E6 gene showed that European and non-European (Asian-American and African) branches formed related nodes (fig. 2). Thus, European and non-European variants were clustered independently. Moreover, we observed that Euro-German specimens gathered together with European lineage samples. On the other hand, although non-European branches were clustered in a separate group, Asian-American and African specimens were in separate subgroups inside the non-European cluster.

Discussion

HPV-16 is the main cause of cervical cancer and the most prevalent HPV genotype worldwide [1, 16, 17]. In our area, the Basque Country (Northern Spain), it is the most commonly found virus type in women with abnormal cervical cytology [14]. On the other hand, HPV intratypic variants showed a different geographical distribution, the European variant being the most prevalent one in Europe, followed by African and Asian-American branch isolates [7, 18]. Variant distribution in this study showed concordance with these previous studies with the European variant the most prevalent (85.10%) followed by Asian-American (7.80%), African (4.25%) and Euro-German lineage isolates (2.83%).

Table 2. Most important nucleotide sequence variations in the E6 gene among HPV-16 isolates in non-European variants (Asian-American and African branch)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>E</td>
<td>K02718</td>
</tr>
<tr>
<td>Reference sequences</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>AA</td>
<td>AF402678</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td>FJ006723</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A131G</td>
<td>AF472508</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FJ006723</td>
<td>AF473509</td>
</tr>
<tr>
<td>Number of specimens</td>
<td>3</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>AA</td>
<td>KJ152735</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>A</td>
<td>.</td>
<td>Del</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>AA</td>
<td>KJ152736</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.</td>
<td>.</td>
<td>Del</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>AA</td>
<td>KJ152737</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>A</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>AA</td>
<td>KJ152738</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>C</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>AA</td>
<td>KJ152739</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.</td>
<td>.</td>
<td>Del</td>
<td>C</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>AA</td>
<td>KJ152740</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>AF</td>
<td>KJ152741</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>AF</td>
<td>KJ152742</td>
</tr>
</tbody>
</table>
On the other hand, differences in oncogenic potential of these intratypic variants have been described. HPV-16 variants are genetically distinct, and therefore they show different pathological risks [16, 19]. Some authors have proved that non-European variants, particularly the Asian-American, present a greater probability to cause cervical cancer or high-grade squamous intraepithelial lesions than European branch isolates [19]. Our results show agreement with previous works considering that the relationship between non-European variants and high-grade lesion development in women showed statistical significance (p < 0.05).

Furthermore, we studied the association between variants and the patient’s average age. It was observed that European variants were more commonly detected in younger women whereas non-European lineage isolates were mainly detected in older women, showing a statistical significance (p < 0.05).

Previous studies have shown that some nucleotide substitutions have remarkable importance in pathogenicity [11, 20–22]. In the present study, 16 nucleotide variations were detected. Most of these nucleotide substitutions have previously been described, for instance A131G, G145 and C335T which modified the p53 tumor suppressor protein binding site and degradation. Moreover, 2 of these nucleotide variations (T350G and A104Del) were frequently detected in high-grade lesion specimens in European branch isolates which showed their high relevance. Previous studies have related T350G and A104Del nucleotide substitutions with virus persistence and risk of cervical neoplasia development and have proposed that T350G nucleotide variation is an additional risk factor for persistent infection and high-grade lesion progression [11].

Infections with more than 1 HPV genotype are currently being detected in a great number of patients [12, 23]. In this study all samples had multiple HPV-16 infections, and this could affect the appearance of nucleotide substitutions. Moreover, although most authors have analyzed only one genomic region to identify HPV-16 variants, it would probably not be enough for an appropriate variant identification since it may be impossible to detect recombination [24, 25]. Furthermore, Angulo et al. [26] have proved that E6 is one of the genes with the highest recombination value in the HPV genome and, as recombination, is a great evolutionary mechanism that could have an effect on pharmacogenomics and vaccine strategies.

In conclusion, HPV-16 E6 gene analysis makes it possible to correlate patient age average, nucleotide substitutions, intratypic variants and lesion development; therefore it is enough to identify HPV-16 intratypic variants. However, single gene analysis makes it impossible to detect recombination which has a high impact on the influence of genetic variation on the response to drugs and vaccine improvement. Thus, more than one genomic region should be analyzed in variant classification so as to detect recombinant samples.

Fig. 2. HPV-16 E6 gene phylogenetic tree. The maximum likelihood method was used (MEGA software version 5).
Acknowledgments

This work was supported by the Department of Industry (S-PC11BF002 project) and Department of Health (project No. 200811058) from the Basque Government and by the University of the Basque Country (EHU13/04). N.F.’s research staff contract was supported by the University of the Basque Country (PIC 73/14).

References


Disclosure Statement

All authors declare no potential conflicts of interest.