

RESEARCH ARTICLE

Mutation Detection of E6 and LCR Genes from HPV 16 Associated with Carcinogenesis

Jessica P Mosmann*, Marina S Monetti, Maria C Frutos, Ana X Kiguen, Raul F Venezuela, Cecilia G Cuffini

Abstract

Human papillomavirus (HPV) is responsible for one of the most frequent sexually transmitted infections. The first phylogenetic analysis was based on a LCR region fragment. Nowadays, 4 variants are known: African (Af-1, Af-2), Asian-American (AA) and European (E). However the existence of sub-lineages of the European variant has been proposed, specific mutations in the E6 and LCR sequences being possibly related to persistent viral infections. The aim of this study was a phylogenetic study of HPV16 sequences of endocervical samples from Córdoba, in order to detect the circulating lineages and analyze the presence of mutations that could be correlated with malignant disease. The phylogenetic analysis determined that 86% of the samples belonged to the E variant, 7% to AF-1 and the remaining 7% to AF-2. The most frequent mutation in LCR sequences was G7521A, in 80% of the analyzed samples; it affects the binding site of a transcription factor that could contribute to carcinogenesis. In the E6 sequences, the most common mutation was T350G (L83V), detected in 67% of the samples, associated with increased risk of persistent infection. The high detection rate of the European lineage correlated with patterns of human migration. This study emphasizes the importance of recognizing circulating lineages, as well as the detection of mutations associated with high-grade neoplastic lesions that could be correlated to the development of carcinogenic lesions.

Keywords: HPV 16 - lineages - mutations - E6 - LCR - cervical cancer

Asian Pac J Cancer Prev, 16 (3), 1151-1157

Introduction

Human papillomavirus type 16 (HPV16) is the main etiological agent of cervical cancer, which is the second most common type of cancer in women worldwide (Muñoz et al., 2003). In a study based on the detection of circulating HPV genotypes in women from Córdoba-Argentina, HPV16 demonstrated to be the most frequent etiological agent (Venezuela et al., 2012).

The HPV16 genome has about 8000 base pairs and includes 8 protein-coding genes (L1, L2, E1, E2, E4, E5, E6 and E7) plus a non coding region LCR (Long Control Region) (Smith et al., 2011). According to the Papillomavirus Nomenclature Committee, a new HPV type is defined by a nucleotide sequence variation of more than 10% compared to an already known HPV type in the L1 open reading frame. Those types differing in 2 to 10% are considered subtypes, whereas intratype variants vary by 2% in the L1 region (Pande et al., 2008).

The genetic variability study started with the analysis of HPV16, as this is the most prevalent viral agent responsible for increased risk of developing cervical cancer. The HPV16 genome used as reference was

sequenced for the first time by Seedorf (Seedorf et al., 1985). The first worldwide study of HPV16 lineages using the LCR region was performed in 1993 by Ho (Ho et al., 1993), who reported 4 major variant lineages: European (E), Asian American (AA), African-1 (AF-1) and African-2 (AF-2). Smith proposed a sublineage within the European lineage, differentiating European prototype (Ep) and European Asian (Ea) because of the small difference (0.5-1%) in the sequences of European and Asian lineages. Their names are derived from the geographical origin of the populations in which they are more prevalent (Ho et al., 1993).

Previous studies have also shown that specific lineages may influence the persistence of HPV infection and the progression of cervical cancer precursor lesions (Londesboroug et al., 1996; Zehbe et al., 1998). HPV lineages may also affect virus assembly, immunological response, pathogenicity, p53 degradation, immortalization activity and transcription regulation (Pande et al., 2008).

L1 encodes the major structural protein of the viral particles that are expressed at ultimate stages of the viral production (Fernandes Brenna and Syrjanen, 2003) and differences in their nucleotide sequences could be linked

¹Instituto de Virología "Dr. J. M. Vanella", Facultad de Ciencias Médicas-Universidad Nacional de Córdoba, Córdoba, Argentina
*For correspondence: jesticamosman@hotmail.com

to changes in the host immune response (Picconi et al., 2002). E6 and E7 are the major oncoproteins and are involved in tumorigenesis. LCR has been demonstrated to be the most variable region of HPV16 and contains the early promoter and regulatory elements involved in viral DNA replication and transcription (Cornet et al., 2012). Several authors have evaluated the association of specific HPV16 mutations with viral persistence and the development of cervical intraepithelial neoplastic lesions (Londesborough et al., 1996; Bontkes et al., 1998; Villa et al., 2000; Cornet et al., 2013).

The aim of this study was to perform a phylogenetic study of HPV16 sequences in cervical lesions samples obtained from women of Córdoba city- Argentina, in order to determine the circulating lineages and analyze the presence of mutations related to malignancies.

Materials and Methods

Samples

The analyzed samples belonged to women with different grades of cervical lesions that were examined to discard HPV. A total of 432 endocervical samples were collected between August 2011 and July 2012 in Córdoba- Argentina; all of them were obtained by swab or cervical smears. All patients included in the study were adult females; those younger than 18 years of age were excluded.

This study was approved by the Institutional Committee of Ethics of Health "Oulton Romagosa" of Córdoba- Argentina, according to the ethical principles stated in the declaration of Helsinki.

DNA extraction

The samples were collected in 500 µL of phosphate-buffered saline (PBS). DNA was extracted using the commercial AxyPrep Body Fluid Viral DNA/RNA Extraction Kit (Axygen Scientific Inc., USA), in accordance with the manufacturer's instructions.

Analysis of L1, LCR and E6 genes of HPV16 by PCR. HPV16 L1, LCR and E6 specific PCRs were performed with the following specific primers: for HPV16 L1 with an amplicon size of 450 bp, we used degenerate primers MY09 and MY11; for LCR with an amplicon size of 530 bp, 5'-CAACACCTACTAATTGTGTGG-3' and 5'-AAATCGGTTTGCACACACCCATT- 3' (Yamada et al., 1997) and for E6 with an amplicon size of 476 bp, 5'-GAAACCGGTTAGTATAAAAGCAGAC- 3' and 5'-AGCTGGGTTTCTCTACGTGTTCT- 3' (Pande et al., 2008).

The products were detected by electrophoresis in 1.5% agarose gel using a U.V. transilluminator. The β-globin gene was used as DNA preservation marker. Negative samples were considered inadequate.

HPV-DNA positive samples were typed by Restriction Fragment Length Polymorphism (RFLP), according to Bernard2. Briefly, aliquots of the PCR products obtained using the degenerate consensus primers MY09-MY11, targeting a region of approximately 450 bp in length in the L1 ORF of the viral genome, were mixed with 7 different restriction enzymes (Bam HI, Dde I, Hae III,

Hinf I, Pst I, Rsa I and Sau III) in separate reactions. The digestion products were separated by electrophoresis in 3% agarose gel and the pattern obtained was compared with published data.

Sequencing

PCR products were purified by gel electrophoresis using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA, US) and subjected to direct nucleotide sequencing reaction in both directions by Macrogen, Inc. (Seoul, Korea).

Sequence analysis and phylogenetic tree construction. The sequences obtained were edited and prepared with MEGA 5 version software and aligned to the sequences downloaded from the GenBank. Relatedness of newly characterized sequences was assessed by analysis with 2.2.19 Basic Local Alignment Search Tool (BLAST).

For each dataset, the best fit model of nucleotide substitution was selected using jModelTest v 0.1.1. Posada (2008), assessed by the Akaike Information Criterion. The phylogenetic relationships of each dataset analyzed were evaluated by the Maximum Likelihood methodology with PhyML 3.0 (Guindon and Gascuel, 2003) setting the parameters obtained by the model selection program. The branch support was evaluated by non-parametric bootstrapping with 1000 pseudo-replicas. The trees were prepared for publication using Dendroscope program.

Gen Bank accession numbers

Reference sequence: K02718 (Seedorf et al., 1985). Lineages: European prototype (Ep): HQ644286, European Asian (Ea): HQ644251, African 1 (AF-1): HQ644238, African 2 (AF-2): AF472509 and Asian American (AA): HQ644255 (Smith et al., 2011).

The sequences obtained in this study were deposited in the GenBank under the following accession numbers: KC291255 to KC291269 for the E6 region, KC291270 to KC291277 for the L1 region and KC291278 to KC291292 for the LCR region.

Results

One hundred and thirty three out of 432 cervical samples (31%) tested positive for HPV; of these, 24 (18%) belonged to the HPV16 genotype. Fifteen samples were optimal for LCR and E6 sequences, and 8 for L1 region.

According to the cytological analysis, 6 samples showed low-grade squamous intraepithelial lesions (LSIL), 4 high-grade squamous intraepithelial lesions (HSIL), 1 squamous cell carcinoma (SCC) and 4 samples did not yield any data (Table 1).

Phylogenetic analysis

PhyML software pointed L1, LCR, E6 and E6-LCR as the best fit model for phylogenetic trees TPM3uf, TPM1uf+G, TPM2uf+G and TPM1uf+G, respectively.

In the phylogenetic tree of L1 sequences, the lineages percentage was 88%: E (57% Ep, 43% Ea) and 22%: AF-1. No samples AF-2 and AA lineages were found (data not shown).

In the phylogenetic tree of LCR sequences, the

lineages percentage was 86%: E (54% Ep, 46% Ea), 7%: AF-1 and 7%: AF-2. No samples of the AA lineage were detected (data not shown).

In the phylogenetic tree of E6 sequences, the lineages percentage was 86%: E (77% Ep, 33% Ea), 7%: AF-1 and 7%: AF-2. No samples of the AA lineage were found (data not shown).

In the phylogenetic tree of LCR-E6 sequences, the lineages percentage was 86%: E (77% Ep, 33% Ea), 7%: AF-1 and 7%: AF-2. No samples of the AA lineage were found (Figure 1- Table 1).

The European lineage was observed in 4 HSIL samples, 5 LSIL, 1 SCC and 3 samples without defined cytology; African-1 lineage in 1 sample without defined cytology and African-2 lineage in 1 LSIL sample (Table 1).

LCR sequence variation

The LCR of HPV16 type is the most variable region and presents a high number of nucleotide variations in regulatory binding sites of a number of cellular and/or viral transcription and regulatory factors, such as E2, YY1, AP1, Tef-1, NF1 and Oct-1. The most frequently observed

mutation was G7521A in 80% of the analyzed samples, located in one of the many YY1 binding sites. This change was detected in most of the Ep sublineage samples as well as in African lineages, but it was not detected in any of the samples of the Ea sublineage.

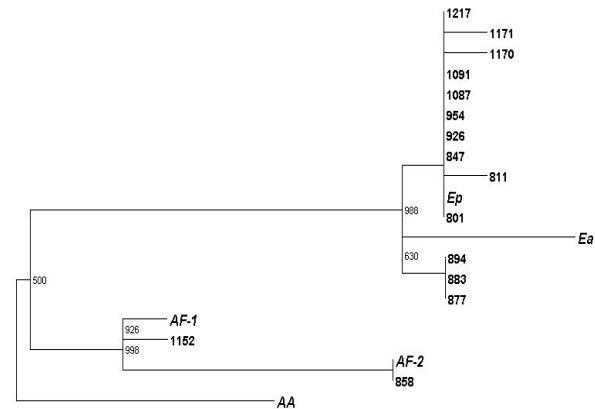


Figure 1. Phylogenetic Analysis of HPV 16 E6-LCR Sequences. The Bootstraps (≥ 50%) are Indicated in Internal Nodes

Table 1. Cytology, Lineage and Nucleotide Variation of HPV 16 Sequences

	E6			LCR					L1							
	1	2	3	7	7	7	7	7	6	6	6	6	6	6	6	6
Ref. K02718																
801																
801																
811																
847																
858																
877																
883																
894																
926																
954																
1087																
1091																
1152																
1170																
1171																
1217																

*The identification codes of the samples and reference sequences used are indicated on the left side, followed by the cytology and lineage data. The nucleotide position of E6, LCR and L1 in which variations were observed are written across the top. For each variant sequence, positions that do not vary from HPV reference sequence are marked with dashes

Table 2. HPV 16 Variants of E6 and L1 Region, according to Nucleotide Position and Altered Amino Acids

Gene	Position	Nucleotide		Amino Acid			Codon
		Prototype	Variant	Original	Altered		
E6	176	G	A	Aspartic acid	C-	Asparagine	P 59
	290	T	A	Cysteine	P	Serine	P 97
	350	T	G	Leucine	H	Valine	H 117
L1	6651	T	A	Serine	P	Arginine	C+ 2217
	6871	A	T	Threonine	P	Serine	P 2291
	6891	G	T	Arginine	C+	Serine	P 2297
	6908	C	T	Alanine	H	Valine	H 2303
	6970	A	T	Threonine	P	Serine	P 2324
	6994	G	A	Glutamic acid	C-	Lysine	C+ 2332
	6998	A	T	Lysine	C+	Methionine	H 2333

*C+, C- : charged amino acids; P: polar amino acids; H: hydrophobic amino acids

Other point mutations were detected: G7436C, T7450C, G7462C and G7502C (Table 1).

E6 sequence variation

The most frequently detected mutation was T350G, in 67% of the analyzed samples; it produces an amino acid change of leucine to valine (L83V). This mutation was detected in 100% of the Ep sublineage samples but was not detected in any of the samples of the Ea sublineage.

Other two point mutations have been detected G176A and T290A; they also produce amino acid changes, D59N and C97S, respectively. (Tables 1-2).

L1 sequence variation

In these sequences, 7 different point mutations were detected: T6651A, A6871T, G6891T, C6908T, A6970T, G6994A and A6998T. All the detected point mutations are responsible for amino acid changes: S2217R, T2991S, R2297S, A2303V, T2324S, E2332K and K2333M, respectively (Tables 1-2).

Discussion

Human Papillomavirus type 16 (HPV16) is the most prevalent oncogenic HPV type worldwide and the prevalence of its different lineages differs across geographical locations (zur Hausen 1991; Cento et al., 2009; Cornet et al., 2013).

Recent investigations have evaluated the association between specific HPV16 lineages and viral persistence, as well as with the development of high-grade cervical lesions (Xi et al., 1995; Londesboroug et al., 1996; Xi et al., 2007; Cornet et al., 2013).

When bearing in mind the geographical distribution of its lineages, it has been reported that the European lineage predominates in the ordinary European and American populations; the African 1 and 2 lineages predominate in Africa, the Asian lineage (emerging branch of the European lineage) predominates in Southern areas of Asia and the Asian-American lineage in Central and South America and Europe. The geographical distribution of HPV types and their specific lineages may be influenced by several factors, including founder effects, coevolution of HPV with human beings, human migration patterns and viral fitness measures such as transmissibility (Cornet et al., 2013).

The high detection of European lineage in this study is in agreement with previous studies performed in Argentina and other South American countries, such as the investigation by Yamada, Picconi, Tonón, Cornet and Argentina received many European immigrants during the first decades of the 20th Century; this fact may have led to the entrance this variant into the country (Yamada et al., 1997; Picconi et al., 2003; Tonón et al., 2007; Cornet et al., 2013).

Even though the number of samples analyzed in this study is relatively low (15 HPV16 positive samples from 432 women), it can be considered representative of our population, since the mortality rate due to cervical cancer in Córdoba province- Argentina, is 5.7%, different to other provinces of Argentina like Misiones (Tonón et al., 2007),

which has a mortality rate of 16.5% or Jujuy (Picconi et al., 2002) which has a mortality rate of 13.3%, according to data provided by the National Ministry of Health during the last years.

It is important to point out that we did find a correlation between the lineages detected by the analysis of the three regions; however, the percentage of the European lineage was greater in the analysis of L1 because 8 sequences of the HPV16 positive samples were analyzed, different than results from LCR and E6 sequences, in which all the samples were analyzed (N=15). Even though most of the studied samples presented great homology with the European sequences, some differences were detected at the time of classifying them by sublineages; in LCR, samples 883 and 877 grouped with sublineage Ep, and E6 grouped with Ea, respectively. The opposite occurred with samples 811, 847, 926, 1091 and 1217, which showed an association with sublineage Ea in LCR, as did E6 with Ep. These differences were not observed in those sequences that grouped with African lineages.

In order to get a better idea of which types of the lineages circulate in the studied population, we performed a complementary alignment of E6-LCR, which allowed detecting that most of sequences grouped with the European prototype sublineage which in part, is also in accordance with the origin of the persons that migrated to Argentina during early 20th Century, mainly to central regions of the country.

Many authors have suggested that some lineages of HPV16 show greater associations with cervical neoplasms (Londesboroug et al., 1996; Bontkes et al., 1998; Zehbe et al., 1998; Villa et al., 2000; Xi et al., 2007; Schiffman et al., 2010; Smith et al., 2011; Cornet et al., 2012). This would partially explain why some HPV16 infections progress to HSIL or cancer, while others do not (Picconi et al., 2003). Increasing the quality of DNA sequencing technology allowed a finer resolution of the different lineages and the association of non-European lineages with high grade lesions and cervical cancer was established (Xi et al., 2007; Smith et al., 2011). However, the underlying genetic details that make non-European lineages of HPV16 more carcinogenic are still unknown, and this important issue definitely deserves pursuit to identify the genetic basis of this association (Schiffman et al., 2010).

The expression of the virus genes is a complex process that involves viral and cellular transcription factors that have either stimulant or repressive effects (O'Connor et al., 1995). Deletion and point mutation experiments indicate that the activity of viral enhancers depends on several factors.

The LCR region contains several binding sites for the viral transcriptional regulatory protein E2 and cellular transcription factors, which modulate its function positively (AP1, Sp1, NF1, Oct-1, Tef-1) or negatively (YY1) (O'Connor et al., 1995; Bernad 2006). The transcription factor YY1 is a repressive agent responsible for silencing the E6 promoter transcription. The mutation that leads to a loss of this silencer was detected in patients with cancer, but was not found in asymptomatic carriers. This would explain the reason why although HPV infection is the most frequent sexually transmitted disease,

only a fraction of the women with HPV16 positive cervical precursor lesions evolve to cancer.

The analysis of the nucleotide changes found in cervical cancer isolates showed that the changes were mainly located within or close to YY1 binding sites, while in asymptomatic carriers these changes presented as a wide variety of transcription factor binding sites (Schmidt et al., 2001). The mutation most frequently observed in LCR was G7521A in the binding site of the transcription factor YY1 (Pande et al., 2008). In this study, in accordance to other authors (Yamada et al., 1997; Schmidt et al., 2001; Kammer et al., 2002; Pande et al., 2008; Shang et al., 2011), the mutation most frequently detected in the LCR region was G7521A, found in the African lineages and most European ones. Thus, only 20% of analyzed samples presented homology with the prototype sequence. In this analysis, the samples that presented G7521A mutation belonged mostly to low-grade squamous intraepithelial lesions (LSIL); this is an interesting and alarming fact since it indicates that these patients might be at higher risk of developing more severe lesions.

The mutations of the encoding genes could produce changes in the amino acid composition, which in turn result into protein changes of tertiary structures and alterations of their functions. *In vitro* studies of different HPV16 mutations have suggested some differences in the protein biological activity. T350G change constitutes the most frequent mutation in E6, a common such polymorphism within E (Cornet et al., 2013); it produces a replacement of leucine by valine, named L83V (Zuna et al., 2011). The amino acid 83 is located in the center of conserved amino acid S- L83V- YG, identical to those of genital high-risk oncogenic HPV types but different than those from the low risk types. The E6 protein presents transforming activity, which may be affected by amino acid changes in its functional sites (Andersson et al., 2000). E6 variants may differ in their immunogenicity by the generation of different peptides for the presentation of polymorphic HLA molecules to specific T cells (Zehbe et al., 2001). Several studies have reported the fact that the T350G mutation is associated to a rise in the risk of developing persistent infections or with progression of LSIL to HSIL (Cornet et al., 2013); however other studies have found opposite results (Zehbe et al., 1998). One possible hypothesis is that this phenomenon may be associated with the host cellular immune response. Different haplotypes of the major histocompatibility complex may recognize E6 variants with different efficiency scales (Zehbe et al., 1998).

The nucleotide variation with maximum percentage of detection for E6 sequences found in this study consisted on the modification found in position 350 (T350G- L83V), which was detected in most samples of the European lineages. These results agree with previous studies (Brady et al., 1990; Ho et al., 1991; Andersson et al., 2000; Kammer et al., 2002; Sicheo et al., 2012), which have described this modification as the most frequent one in the E6 sequence in E sublineage (Cornet et al., 2013). Noteworthy, we detected this mutation in all of the Ep sublineage samples, but in none of the Ea ones. This result is in accordance to findings reported by Pande (Pande et

al., 2008).

Tonón reported this mutation in 32% of their Ep sublineage samples but we detected it in 100% of our samples that corresponded to this sublineage. We consider that this difference could be due to the fact that Tonón studied Guarani-indian women. The Guarani-indians are no more than 1700 people who inhabit the rainforest of Misiones province (Northeast region of Argentina), concentrated in more than 20 small communities (Tonón et al., 2007).

Different from information published by other authors (Zehbe et al., 1998; Xi et al., 2007), most of the samples in which this modification was observed belonged to patients with low grade intraepithelial lesions. This could be due to the fact that, regardless of other risk factors, the samples were obtained before the progress of the lesions to malign processes, taking into consideration that the possibility of receiving appropriate health care affects not only the possibilities of early detection but also the chances of lesion progress. For these patients, the access to routine gynecological tests allows timely therapeutic decisions.

Nucleotide variation in L1 region are linked to changes of the host immune response, so that the nucleotide modifications can affect the efficiency of the L1 protein to self assemble and this variability can lead to conformational changes within the relevant neutralization epitopes (Pande et al., 2008).

Differently to findings in LCR and E6, in the sequence analysis of L1 gene, the most frequent mutation, T6862C, observed in previous studies (Pande et al., 2008) was not detected in our study. However, other unreported mutations were found (T6651A, A6871T, G6891T, C6908T, A6970T, G6994A and A6999T); they produced changes of one amino acid to another, which can affect the structure or function of the L1 protein, with the ability of playing an important role in the immune response.

Table 1 shows that 9 samples (801, 811, 847, 926, 954, 1087, 1091, 1170 and 1217) presented the two mutations more frequently associated with progression of the lesions (G7521A in LCR and T350G in E6). This could be the largest circulation of this strain in Córdoba- Argentina, constituting a warning for the Public Health System.

In this study, we detected two other nucleotide changes in E6 sequences: G176A and T290A. Importantly, while changes in E6 resulted in amino acid changes, G176A variation produced a change of negative amino acids for polar amino acids, which could modify the protein biological activity (Table 2).

Other 4 nucleotide changes were detected in the LCR sequences, which are located in close transcription factor binding sites. G7436C is located near the YY1 binding site and T7450C is located close to the viral transcriptional regulatory protein E2. G7462C changes were detected close to the Tef-1 binding site and G7502C. All these changes, both in LCR and in E6, had not been previously reported by other authors. As observed, in the LCR sequences, almost 95% of the nucleotide changes already found was located inside or close to the binding sites of the transcription factors. Five of the studied samples contained 2 of these specific changes in different sites, which could affect the regulation (in a positive or negative manner)

of the HPV transcription, being capable to trigger the progression of the lesions.

These results are the first contribution to the field of molecular epidemiology of HPV16 in patients from Córdoba province- Argentina, indicating the importance of the detection and study of circulating lineages and the analysis of the changes of the nucleotide structure: This is significant not only in terms of molecular epidemiology, but also in the impact of these variations on Public Health Affairs when these results are correlated to the evolution of the disease towards malignant processes.

Acknowledgements

This work was supported by PIO-MincytCba.2011-2013. Ref.Res. MINCYT Cba.N°170/2011.

References

Andersson S, Alemi M, Rylander E, et al (2000). Uneven distribution of HPV16 E6 prototype and variant (L83V) oncoprotein in cervical neoplastic lesions. *Br J Cancer*, **83**, 307-10.

Bernad HU (2006). Phylogeny and taxonomy of papillomaviruses, In 'Papillomavirus Research', Ed Saveria campo m. Institute Of Comparative Medicine University Of Glasgow, UK, pp.11-7.

Bernard HU, Chan S, Manos M (1994). Identification and assessment of known and novel human papillomaviruses by polymerase chain reaction amplification, restriction fragment length polymorphisms, nucleotide sequence, and phylogenetic algorithms. *J Infect Dis*, **170**, 1077-85.

Bontkes H, Van Duin M, De Gruijl T, et al (1998). HPV16 infection and progression of cervical intra-epithelial neoplasia: analysis of HLA polymorphism and HPV16 E6 sequence variants. *Int J Cancer*, **78**, 166-71.

Brady C, Duggan-keen M, Davidson J, Varley J, Stern P (1990). Human papillomavirus type 16 E6 variants in cervical carcinoma: relationship to host genetic factors and clinical parameters. *J Gen Virol*, **80**, 3233-40.

Cento V, Ciccozzi M, Ronga L, Perno, Ciotti M (2009). Genetic diversity of human papillomavirus type 16 E6, E7 and L1 genes in Italian women with different grades of cervical lesions. *J Med Virol*, **81**, 1627-34.

Cornet I, Gheit T, Franceschi S, et al (2012). Human papillomavirus type 16 genetic variants: phylogeny and classification based on E6 and LCR. *J Virol*, **86**, 6855-61.

Cornet I, Gheit T, Iannacone M, et al (2013). HPV16 genetic variation and the development of cervical cancer worldwide. *Br J Cancer*, **108**, 240-4.

Fernandes Brenna S, Syrjanen K (2003). Regulation of cell cycles is a key importance in human papillomavirus (HPV) associated cervical carcinogenesis. *Sao Paulo Med J*, **121**, 128-32.

Guindon S, Gascuel O (2003). A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*, **52**, 696-704.

Ho L, Chan S, Burk R, et al (1993). The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *J Virol*, **67**, 6413-23.

Ho L, Chan S, Chow V, et al (1991). Sequence variants of human papillomavirus type 16 in clinical samples permit verification and extension of epidemiological studies and construction of a phylogenetic tree. *J Clin Microbiol*, **29**, 1765-72.

Kammer C, Tommasino M, Syrjanen S, et al (2002). Variants of the long control region and the E6 oncogene in European human papillomavirus type 16 isolates: implications for cervical disease. *Br J Cancer*, **86**, 269- 73.

Londesboroug P, Ho L, Terr G, et al (1996). Human Papillomavirus genotype as a predictor of persistence and development of high-grade lesions in women with minor cervical abnormalities. *Int J Cancer*, **69**, 364-8.

Muñoz N, Bosch F, de Sanjose S, et al (2003). Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med*, **348**, 518-27.

O'Connor M, Chan S, Bernad HU (1995). Transcription factor binding sites in the long control region of genital HPVs. In 'Human Papillomaviruses', Eds Myers G, Bernad HU, Delius H, et al. Los Alamos National Laboratory, Los Alamos N. Mex pp. 21-40.

Pande S, Jain N, Prusty B, et al (2008). Human Papillomavirus type 16 variant analysis of E6, E7, and L1 genes and long control region in biopsy samples from cervical cancer patients in north India. *J Clin Microbiol*, **46**, 1060-6.

Picconi M, Gronda J, Alonio L, et al (2002). Virus papiloma humano en mujeres quechuas jujeñas con alta frecuencia de cáncer de cuello uterino. Tipos virales y variantes de HPV16. *Medicina*, **62**, 209-20.

Picconi M, Alonio L, Sichero L, et al (2003). Human papillomavirus type-16 variants in Quechua aboriginals from Argentina. *J Med Virol*, **69**, 546-52.

Posada D (2008). jModelTest: Phylogenetic model averaging. *Mol Biol Evol*, **25**, 1253-6.

Schiffman M, Rodriguez A, Chen Z, et al (2010). A population-based prospective study of carcinogenic human papillomavirus (HPV) variant lineages, viral persistence, and cervical neoplasia. *Cancer Res*, **70**, 3159-69.

Schmidt M, Kedzia W, Gozdzicka-Jozefiak A (2001). Intratype HPV16 sequence variation within LCR of isolates from asymptomatic carriers and cervical cancers. *J Clin Virol*, **23**, 65-77.

Seedorf K, Krammer G, Durst M, Suhai S, Rowekamp W (1985). Human Papillomavirus type 16 sequence. *Virology*, **145**, 181-5.

Shang Q, Wang Y, Fang Y, et al (2011). Human papillomavirus type 16 variant analysis of E6, E7, and L1 genes and long control region in identification of cervical carcinomas in patients in northeast China. *J Clin Microbiol*, **49**, 2656-63.

Sichero L, Sobrinho J, Villa L (2012). Oncogenic potential diverge among human papillomavirus type 16 natural variants. *Virology*, **432**, 127-32.

Smith B, Chen Z, Reimers L, et al (2011). Sequence imputation of HPV16 genomes for genetic association studies. *PLoS ONE*, **6**, 1-8.

Tonón S, Basiletti J, Badano I, et al (2007). Human papillomavirus type 16 molecular variants in Guarani Indian women from Misiones, Argentina. *Int J Infect Dis*, **11**, 76-81.

Venezuela R, Kiguen A, Frutos M, Cuffini C (2012). Circulation of Human Papillomavirus (HPV) genotypes in women from Córdoba, Argentina, with squamous intraepithelial lesions. *Rev Inst Med Trop Sao Paulo*, **54**, 11-6.

Villa L, Sichero L, Rahal P, et al (2000). Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. *J Gen Virol*, **81**, 2959-68.

Xi L, Demers G, Koutsky L, et al (1995). Analysis of human papillomavirus type 16 variants indicates establishment of persistent infection. *J Infect Dis*, **172**, 747-55.

Xi L, Kiviati N, Wheeler C, et al (2007) Risk of cervical intraepithelial neoplasia grade 2 or 3 after loop electrosurgical excision procedure associated with human papillomavirus type 16 variants. *J Infect Dis*, **195**, 1340-4.

Yamada T, Manos M, Peto J, et al (1997). Human papillomavirus

- type 16 sequence variation in cervical cancers: a worldwide perspective. *J Virol*, **71**, 2463-72.
- Zehbe I, Wilander E, Delius H, Tommasino M (1998). Human papillomavirus 16 E6 variants are more prevalent in invasive cervical carcinoma than the prototype. *Cancer Res*, **58**, 829-33.
- Zehbe I, Tachezy R, Mytilineos J, et al (2001). Human papillomavirus 16 E6 polymorphisms in cervical lesions from different european populations and their correlation with human leukocyte antigen class II haplotypes. *Int J Cancer*, **94**, 711-6.
- Zuna E, Tuller E, Wentzensen N, et al (2011). HPV16 variant lineage, clinical stage, and survival in women with invasive cervical cancer. *Inf Ag Cancer*, **6**, 1-9.
- Zur Hausen H (1991) Human papillomaviruses in the pathogenesis of anogenital cancer. *Virol*, **184**, 9-13.