Genetic Diversity of Human Papillomavirus Type 16 E6, E7, and L1 Genes in Italian Women With Different Grades of Cervical Lesions

Valeria Cento,[1] Massimo Ciccozzi,[2] Luigi Ronga,[1] Carlo Federico Perno,[1,3] and Marco Ciotti[1]*

[1]*Laboratory of Molecular Virology, University Hospital Tor Vergata, Rome, Italy*
[2]*Department of Infectious, Parasitic and Immunomediated Disease, Istituto Superiore di Sanita, Rome, Italy*
[3]*Department of Experimental Medicine and Biochemistry, Tor Vergata University, Rome, Italy*

High-risk human papillomaviruses (HPVs) are risk factors for the development of cervical cancer. HPV 16 is the most common type, being present in about 60% of cervical cancers worldwide. Previous studies have reported upon the association between HPV 16 E6 variants and increased risk of cervical intraepithelial neoplasia and invasive cervical cancer. In this study, the presence of HPV 16 polymorphisms in the E6, E7, and L1 genes was investigated in relation to the presence of high-grade lesions. Sequencing of the E6 gene revealed the presence of nucleotide mutations resulting in 15 amino acid changes. Of these, the G134D and C136R fall within the CXXC zinger finger domain important for p53 binding. In the E7 gene, four nucleotide variations were identified with two leading to the amino acid substitutions L15V and S31R. The L1 gene showed 13 nucleotide changes leading to 11 amino acid substitutions. Among these, the R364C and N367D are located at the base of the HI-loop of the L1 protein, considered to be the immunodominant epitope of HPV 16. No significant relationship between HPV 16 variants and high-grade lesions was found. Phylogenetic analysis showed that all the HPV 16 variants identified belonged to the European lineage, except one which was of the Asian-American lineage. The European-350G variant was detected most frequently (22 of 34, 64.7%). The study provides some new data on the genetic diversity of HPV 16 which may help to understand the oncogenic potential of the virus and to improve management of patients. *J. Med. Virol. 81:1627–1634, 2009.* © 2009 Wiley-Liss, Inc.

KEY WORDS: HPV16; phylogenetic analysis; E6/E7/L1 genes; HPV16 variants

INTRODUCTION

Human papillomavirus (HPV) is the recognized etiological agent of cervical cancer. Genital HPV infection is the most common sexually transmitted infection worldwide, and though most of those infected will clear the virus without developing clinically recognizable features. Only a few infected women will progress to cervical cancer [Baseman and Koustsky, 2005]. To date, more than 100 HPVs are recognized and about 40 can infect the ano-genital tract [de Villiers et al., 2004]. Based on epidemiological and phylogenetic studies, HPVs are classified as low-risk (LR) and high-risk (HR) types. LR types are associated commonly with benign lesions (warts and condylomata), while HR types are associated with high-grade lesions and cervical cancer [Munoz et al., 2006; zur Hausen, 1991]. The oncogenic potential of HPV is related to the transforming activity of the E6 and E7 oncoproteins, which are responsible for the functional inactivation of p53 and pRB, two key proteins for the control of the cell cycle [Crook et al., 1991; Dyson et al., 1989; Munger, 2002]. As a consequence, there is an inhibition of apoptosis and the entry of the cell into the S phase [Wise-Draper and Wells, 2008]. The deregulation of cell cycle control favors cell transformation and ultimately progression to cervical cancer.

HPV evolutionary analysis revealed the African origin of the virus [Alvarez-Salas and DiPaolo, 2007] and its dissemination throughout the world. Among the oncogenic HPVs, HPV 16 and 18 are the types detected most frequently in cervical cancer. These two types are responsible for nearly 80% of cervical tumors with HPV 16 accounting for as much as 60% of the total cases [Munoz et al., 2003]. For this reason, research efforts have concentrated on HPV 16 and its variants in order to understand the molecular basis underlying its oncogenic

potential [Ellis et al., 1995; Yamada et al., 1995, 1997; Zehbe et al., 1998a,b; Nindl et al., 1999; Ortiz et al., 2006]. Few data are available on the molecular epidemiology of HPV in Italy and even fewer about the distribution of different HPV 16 variants. HPV 16 intra-type variants have been grouped into six phylogenetic branches: European (E); Asian (As); African-1 (Af-1); African-2 (Af-2); Asian-American (AA) and North American (NA) [Yamada et al., 1995, 1997; Ortiz et al., 2006].

The objective of this study was to detect nucleotide polymorphisms within the *E6*, *E7*, and *L1* genes of HPV 16 and to relate them to the grade of histological lesions.

MATERIALS AND METHODS

Study Population

The present study includes 63 HPV 16 positive women admitted to Gynecology and Colposcopy Clinics of Tor Vergata University Hospital from January 2004 to March 2008. Punch biopsies were taken from 51 women with abnormal cytology. Twelve patients had a normal cytology, no biopsies were taken. Since the study has been performed on DNA extracted from eso/endocervical cells, analysis of the *E6*, *E7*, and *L1* genes was performed on all samples where sufficient material was still available: 34 samples for *E6* gene, 35 samples for *E7* gene, and 39 samples for *L1* gene. The mean age of the patients was 34 years (range 20–75; SD ± 10.5). Informed consent was obtained from all women enrolled in the study.

Amplification and Sequencing of HPV 16 E6, E7, and L1 Genes

The entire *E6* gene was amplified using the sense primer HPV 16 E6 (141–161 nt) 5′-CACAGTTATGCAC- AGAG-CTGC-3′ and the reverse primer HPV16 E6 (565–541 nt) 5′-TAGAGAAACCCAGCTGTAATCATGC-3′. Five microliters of DNA were added to 45 µl of reaction mix containing 1× PCR buffer, 200 µM dNTPs, 2 mM $MgCl_2$, 40 pmol of each primer, and 1 U of Taq Gold polymerase (Applied Biosystems, Monsa, Italy). The cycling conditions were as follows: 94°C for 10 min, 1 cycle; then, 94°C for 30 sec, 60°C for 30 sec, and 72°C for 30 sec, for 35 cycles using a 9700 GeneAmp instrument (Applied Biosystems, Foster City, CA).

HPV 16 *E7* and *L1* genes were amplified as already described [Ciotti et al., 2006a,b].

The PCR products were sequenced using the Genome Lab DTCS Quick Start Kit, (Beckman Coulter, Fullerton, CA) and run on a Beckman Coulter CEQ2000XL gene analyzer after column purification (Millipore, Billerica, MA). The obtained sequences were submitted to Genbank and aligned with the HPV 16 reference sequence [Seedorf et al., 1985]. The HPV 16 sequences were numbered according to the 1997 sequence database (Los Alamos National Laboratory, Los Alamos, NM) and variants designed according to Yamada et al. [1997].

Phylogenetic Analysis

Multiple sequence alignment of the full length *E6* and *E7* genes, as well as of the 450 bp fragment of the *L1* gene,

was performed using ClustalW Multiple Alignment software. Sequences were manually edited with the Bioedit software [Hall, 1999] and gaps removed from the final alignment. The best fitting nucleotide substitution model was tested with a hierarchical likelihood ratio test, using a neighbor-joining (NJ) tree with LogDet corrected distances as base tree [Swofford and Sullivan, 2003]. A maximum likelihood (ML) tree was then inferred with the selected model and ML-estimated substitution parameters. The heuristic search for the best tree was performed using an NJ tree as the starting tree and the Tree Bisection and Reconnection (TBR) branch-swapping algorithm. NJ trees were also obtained with ML estimated pair-wise distances using the best fitting nucleotide substitution model. Calculations were performed with PAUP* software (version 4.0) [Swofford and Sullivan, 2003].

Statistical support for specific clades in the phylogeny was obtained with the ML-based zero branch length test for the ML tree [Swofford and Sullivan, 2003], by bootstrapping (1,000 replicates) for the NJ trees. Trees were rooted by outgroup rooting. The nucleotide sequences obtained in this study have been deposited in Genbank under accession numbers FJ644954–FJ645061.

Statistical Analysis

Statistical analysis was performed using the SPSS software (version 17.0, www.spss.com). A χ^2 test using 4X2 contingency table was performed and P value < 0.05 was considered as statistically significant.

RESULTS

Histological Examination

The histological evaluation of the 51 cervical biopsies obtained from all women with abnormal cytology revealed the presence of chronic or acute cervicitis in 11 (21.5%) patients, cervical intraepithelial neoplasia grade 1 in 21 (41.2%) patients, cervical intraepithelial neoplasia grade 2–3 in 18 (35.3%) patients and carcinoma in situ in 1 (2%) patient. The median age was 36 years (mean 29.3 ± 6.9) for women with chronic or acute cervicitis, 30 years (mean 32.5 ± 10.0) for women with cervical intraepithelial neoplasia grade 1, 34 years (mean 36.9 ± 12.3) for women with cervical intraepithelial neoplasia grade 2–3, 46 years for woman with carcinoma *in situ*. No statistical significant relationship between patient's age and histological lesion was found.

Infection by HPV16 Variants and Histological Cervical Lesion

Variants of the *E6* gene were found to be more common than prototype in all grades of cervical lesion, while for *E7* and *L1* genes the prototype was more common than variants.

The European (E) E-350G was the most common E6 variant identified, being present in 10/11 of cervical intraepithelial neoplasia grade 1, 4/8 of cervical

intraepithelial neoplasia grade 2–3, and 1/1 of carcinoma in situ, against 8/14 of HPV 16 positive patients with normal cytology. It was not observed a statistically significant increase of E-350G with high-grade lesions ($P = 0.17$) when this mutation was found associated with other nucleotide changes.

Phylogenetic Analysis

The Maximum Likelihood trees of *E6*, *E7*, and *L1* genes (identified by the prefix "TV") are reported in Figures 1, 1S, and 2, respectively.

The 34 HPV 16 E6 positive strains were interspersed within clusters E-350G, E-350T, E-350G/E-350T, and AA. Of these, 10 sequences grouped into the E-350T cluster, whereas 7 isolates plus 14 isolates (not represented in the tree) identical to the TV155a sequence grouped into the E-350G cluster. Interestingly, the TV9a and TV212a sequences, closely related to each other,

clustered together although belonging to the E-350T and E-350G variants, respectively. The last sequence TV177a clustered into the Asian American (AA) clade, Figure 1.

The 35 HPV16 E7 positive strains clustered into two different clades: AA and E. Thirty-four sequences were interspersed into the European sub-type (E-350G/E-350T clade) not clearly distinguishably in net sub-clusters. The last one fell into the AA sub-type, Figure 1S.

The 39 HPV 16 L1 positive strains clustered into two clades. Thirty-eight sequences were interspersed into the European group (E-350G/E-350T), whereas the TV177a clustered into the AA clade, Figure 2.

All clades represented in the trees were statistically supported in the NJ (bootstrap values >70%) and ML (zero branch length test, *P*-value <0.001).

Overall, based on the phylogenetic analysis 62 patients were infected by the European strain, and 1 by the AA strain.

Fig. 1. Phylogenetic tree based on 34 sequences of the *E6* gene (420 bp). K02718 is the European reference sequence for HPV 16 DNA [Seedorf et al., 1985]. AF536179 sequence refers to a member of the E-131G subclass, while AF402678, AF472508, and AF472509 refer to sequences of the AA, Af-1, and Af-2 classes, respectively. The tree was rooted using the European reference sequence. Branch length was estimated using the F81 + G model. One asterisk (*) along the branches represent significant statistical support for the clade ($P < 0.001$) in the zero branch length test. The numbers above the branches indicate bootstrap values greater than 70%.
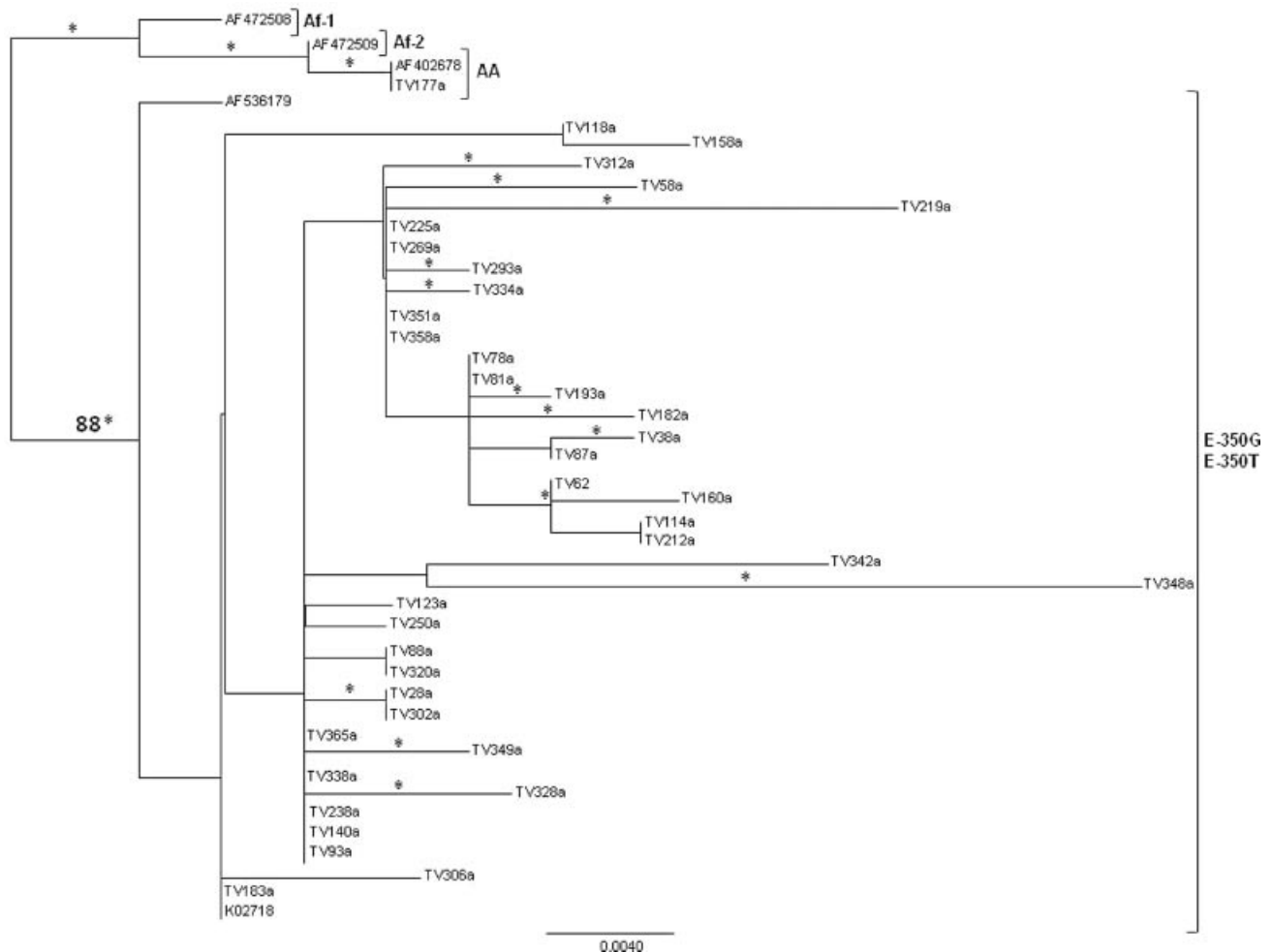
Fig. 2. Phylogenetic tree based on 39 sequences of the L1 gene fragment (450 bp). K02718 is the European reference sequence for HPV 16 DNA [Seedorf et al., 1985], AF536179 is a member of the E-131G subclass, while AF402678, AF472508, and AF472509 are members of the AA, Af-1, and Af-2 classes, respectively. The tree was rooted using the Af-1 reference sequence. Branch length was estimated using the HKY + G model. One asterisk (*) along the branches represent significant statistical support for the clade ($P < 0.001$) in the zero branch length test. The numbers above the branches indicate bootstrap values greater than 70%.

E6, E7, and L1 Sequence Variations

Nucleotide sequences of the complete E6 open reading frame from 34 patients were compared with the HPV 16 E6 reference sequence [Seedorf et al., 1985]. Sequence variations observed in HPV 16 E6 amplified products are summarized in Table I.

Of the sequences belonging to the E lineage, 6 out of 34 (17.6%) were identical to the prototype sequence (E-350T), while 4 out of 34 (11.8%) contained a single nucleotide (nt) change at position 204 (A to T), 205 (G to T), 504 (G to A), and 482 (A to C), respectively. All these changes were novel missense mutations, resulting in the amino acid (aa) substitution K34M, K34N, N127H, and G134D, respectively. The cervical lesions associated with these changes were cervical intraepithelial neoplasia grade 2–3, cervical intraepithelial neoplasia grade 1, cervical intraepithelial neoplasia grade 2–3 and negative histology, respectively. Only one sample presented

two nucleotide changes at positions 163 and 297, resulting in the amino acids substitutions Q20H and K65T, respectively. This sample presented a normal cytology.

The majority of samples (16 out of 34, 47.1%) contained HPV 16 E6 sequences belonging to E clade with a single nucleotide change (T to G) at nucleotide 350 (E-350G) replacing L83 with V. These patients presented grades of related cervical lesions ranging from negative (seven samples) to cervical intraepithelial neoplasia grade 1 (seven samples) and cervical intraepithelial neoplasia grade 2–3 (two samples). Furthermore, 6 sequences (17.6%) that presented sporadic nucleotide changes fell into the E-350G variant. In this group 10 novel nucleotide variations were identified, resulting in six different amino acid substitutions (Table I). Interestingly, the novel mutation T527C was recorded in association with 350G, which was previously reported associated with the T527A [Pande et al., 2008]. The related cervical lesions for all these samples are reported in Table I.

TABLE I. E6 and E7 Nucleotide Changes Identified in HPV 16 Isolates

Class		n° of samples (n = 40)	145*	163	166	189	203	204	205	297	319	335*	350*	433	477	482	502	504	508	509	521	523	527	532*	604	618*	654*	732*	789*	795*	846	Grade of related Cervical Lesion
																	E6									**E7**						
E-p (Reference-nt K02718)			G	A	A	A	A	A	G	A	A	C	T	G	T	A	C	G	A	T	T	C	T	A	T	A	C	T	T	T	T	
	Reference-aa		Q	Q	E		K	K	K	K	K	H	L		F	N		G	C	C	S				L	S						
	aa-positions		14	20	29		34	34	34	65	72	78	83		125	127		134	136	140	142				15	31						
	aa-mutations		H	H	G		Q	M	N	T	N	Y	V		C	H		D	R	S	P				V	R						
AA	AA-a	0	T	-	-	-	-	-	-	-	-	T	G	A	-	-	-	-	-	-	-	-	-	G	-	-	-	C	C	G	-	
		1	-	-	-	-	-	-	-	-	-	T	G	-	G	-	-	-	-	-	A	-	-	G	-	-	-	C	C	G	-	CIS
E	E-p	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	NEG (3); CIN2-3 (2)
		1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	ND	-	-	-	NEG
		1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	NEG
		1	-	T	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	NEG
	E-350T	1	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CIN1
		1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	T	-	-	-	-	-	CIN2-3
		1	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CIN2-3
	E-350G	12	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	NEG (3); CIN1 (7); CIN2-3 (2)
		1	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	ND	-	-	-	NEG
		1	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	ND	-	-	-	NEG
		1	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	NEG
		1	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	C	NEG
		1	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	NEG
		1	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	T	-	-	-	-	-	-	-	ND	-	-	-	NEG
		1	-	-	C	G	C	-	-	C	C	-	G	-	-	-	-	-	-	-	T	C	-	-	-	-	-	-	-	-	-	CIN1
		1	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	T	C	-	-	-	-	-	-	-	-	-	CIN1
		1	-	-	-	-	-	-	-	-	-	-	G	-	-	A	-	-	-	-	-	-	-	-	-	-	-	ND	-	-	-	CIN2-3
		1	-	-	-	-	-	-	C	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CIN2-3
	No sequence for E6	6																							-	-	-	-	-	-	-	NEG (2); CIN1 (1); CIN2-3 (3)

Nucleotide position in E6 and E7 genes is reported at the top of the table. Amino acid changes are reported under the corresponding amino acidic position. The classes of HPV 16 variants and isolates are written on the left [European Prototype (E-p), Asian-American (AA) and Europeans (E), with the E-350G variant]. For each variant, positions that do not vary from those of the HPV reference sequence are marked with dashes, whereas presence of nucleotide changes is indicated by the corresponding letter. ND, not determined; CIS, carcinoma in situ; NEG, negative; CIN1, cervical intraepithelial neoplasia grade 1; CIN2-3, cervical intraepithelial neoplasia grade 2–3. Nucleotide changes previously reported in other studies are marked with asterisk (*).

Analysis of the E6 sequence classified as AA showed a silent mutation at position 433 (A to G), a mutation at position 477 (T to G) and another one at position 521 (T to A) with the amino acid changes F to C and C to S, respectively. These changes do not allow to classify the sequence as a variant of the AA clade as described by Yamada [1997].

Analysis of the complete E7 open reading frame from 35 patients revealed 4 nucleotide variations in 3 (8.5%) patients, whereas the other 32 (91.5%) did not contain any nucleotide changes compared with the prototype [Seedorf, 1985], Table I. Of the identified changes, two were silent nucleotide changes and two were missense substitutions, resulting in the amino acid changes L15V and S31R. The silent mutation A618T was detected in a patient with a histological diagnosis of cervical intra-epithelial neoplasia grade 2–3, while the missense mutation T604G as well as a silent C654A and the missense T846C were associated with a diagnosis of chronic or acute cervicitis.

Nucleotide and amino acids changes of the 450 bp fragment of the *L1* gene are summarized in Table II. Of the 38 sequences belonging to the E lineage, a total of 13 different nucleotide variations were detected in 13 (33.3%) patients, while the remaining 25 (60.1%) showed the prototype sequence.

Of the 13 mutations detected, 3 (23.1%) lead to a silent mutation and 10 (76.9%) to a missense mutation.

Of the 9 mutations found independently, the 3 silent changes A6649T, C6854T (2 samples), and A6964G, were associated with a histological diagnosis of cervical intraepithelial neoplasia grade 2–3 and cervical intra-epithelial neoplasia grade 1, Table II. The other six independent mutations at 6659 nt (A to G); 6756 nt (A to T); 6768 nt (A to C); 6909 nt (A to T); 6962 nt (A to G), and 6990 nt (T to C) lead to the aa changes N341D, Q373L, Q377P, Q424L, K442E, and L451S, respectively. The related grade of cervical lesion for each variation is reported in Table II. The remaining four missense mutations were found together in two samples. The variations at 6650 nt (C to T) and 6695 nt (A to C), resulting in the aa changes R338C and T353P, were detected in a patient with cervical intraepithelial neo-plasia grade 2–3, whereas the variations at 6833 nt (T to A) and 6838 nt (G to C), resulting in the aa changes L399M and E400D, were detected in a patient with cervical intraepithelial neoplasia grade 1.

All the L1 sequence variations found in the analyzed samples were novel except the C6854T and A6695C [Yamada et al., 1997].

The sequence belonging to the AA lineage, detected in a patient with carcinoma in situ, showed a silent

TABLE II. Nucleotide Changes Within the 450 bp Fragment of L1 Gene

Class		n° of samples (n = 39)	6649	6650	6659	6695*	6721*	6756	6768	6803*	6833	6838	6854*	6865*	6909	6962	6964	6970*	6990	6994*	Grade of related Cervical Lesion
E-p (Reference- nt K02718)			A	C	A	A	G	A	A	A	T	G	C	C	A	A	A	C	T	G	
	Reference-aa			R	N	T		Q	Q	T	L	E			Q	K			L		
	aa-positions			338	341	353		373	377	389	399	400			424	442			451		
	aa-mutations			C	D	P		L	P	S	M	D			L	E			S		
AA	AA-a	0	-	-	-	C	A	-	-		T	-	T	T	-	-	-	T	-	A	
		1	-	-	-	C	-	-	-	-	T	-	T	T	‖						CIS
E	E-p	25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	NEG (8); CIN1 (10); CIN2-3 (7)
		1	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	NEG
		1	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	NEG
		1	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	NEG
		2	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	CIN1 (2)
		1	-	-	-	-	-	-	-	-	A	C	-	-	-	-	-	-	-	-	CIN1
		1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	CIN1
		1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	CIN1
		2	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CIN2-3 (2)
		1	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	CIN2-3
		1	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CIN2-3
		1	-	T	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CIN2-3

Nucleotide position within the L1 gene fragment is reported at the top of the table. Amino acid changes are reported under the corresponding amino acidic position. The classes of HPV 16 variants and isolates are written on the left [European Prototype (E-p), Asian-American (AA) and Europeans (E)]. For each variant, positions that do not vary from those of the HPV reference sequence are marked with dashes, whereas presence of nucleotide changes is indicated by the corresponding letter. CIS, carcinoma in situ; NEG, negative; CIN1, cervical intraepithelial neoplasia grade 1; CIN2–3, cervical intraepithelial neoplasia grade 2–3. Nucleotide changes previously reported in other studies are marked with asterisk (*).

nucleotide variation at position 6721 where a G replaced the A reported in the reference sequence for the AA HPV 16 variant.

DISCUSSION

HPV 16 is the most prevalent oncogenic HPV type worldwide. However, the prevalence of different HPV 16 variants differs across geographical locations, probably because of differences in the distribution of HLA alleles involved in the immune response against the virus [Brady et al., 1999; Andersson et al., 2000].

Intra-type variations of the *E6* gene seem to influence the oncogenic activity of the virus, affecting viral persistence and disease progression. Actually, mutations have been found in regions critical for p53 interaction and binding to HLA class I peptides [Zehbe et al., 1998b]. Similarly, intra-type HPV 16 amino acid variations in L1 protein may influence virus assembly or the evocation of a specific neutralizing immune response [Ellis et al. 1995; Hildesheim et al., 2001]. Therefore, identification of HPV 16 variants may be critical for the rational design of newer therapeutic and diagnostic interventions in cervical cancer, as well as for vaccine development strategies.

In this study on HPV genomic diversity, 63 clinical samples identified previously to contain HPV 16 genome were analyzed.

A Maximum Likelihood phylogenetic analysis of HPV 16 variants with respect to *E6*, *E7*, and *L1* genes has been conducted in order to study the local distribution of different HPV 16 lineages, according to the classification of Yamada et al. [1997]. The results showed an extremely high prevalence of the E-lineage (98.4%), with only one (1.6%) AA variant found. Phylogenetic analysis of the *E6* gene showed that two sequences (TV9a and TV212a) although belonging to the E-350T and E-350G variants, respectively, were closely related and clustered together due to the presence of the same mutation at 297 nt (A to C). When this mutation was removed from the analysis, the two sequences TV9a and TV212a grouped in the E-350T and E-350G clusters, respectively.

In Europe and in Italy [Villa et al., 2000; Tornesello et al., 2004], it was reported that the frequency of non-European variants, in particular of the AA variant, increases strikingly with severity of the lesion, suggesting that they could have a more relevant role in oncogenesis. Although the sample population contains only one case infected with the AA variant, the histological examination revealed the presence of a carcinoma in situ.

Nucleotide variations observed in *E6*, *E7*, and *L1* genes showed a high prevalence of the E-350G E6 variant (22/34, 64.7%) in all histological cervical lesions. This result is in agreement with previous reports [Ho et al., 1991; Andersson et al., 2000], but contrasts with the

results reported by other investigators who showed an increased prevalence of this variant in cervical cancer [Zehbe et al., 1998a; Xi et al., 1997].

Among the identified mutations, the N127H mutation is interesting. Since the N127 residue is considered necessary for the binding of the E6 protein to p53, a change at this position may affect the activity of the oncoprotein E6 [Wise-Draper and Wells, 2008]. The patient carrying this mutation had a diagnosis of cervical intraepithelial neoplasia grade 2–3.

Regarding the CXXC zinger finger domain, two novel mutations were identified in this region: the G134D located right before the CXXC sequence and the C136R where the cysteine residue of the CXXC motif was replaced by an arginine. These two mutations fall within a region involved in p53 binding. Whether these two mutations affect the E6 activity remains to be explored.

Analysis of the *E7* gene confirms that the E7 ORF is very stable. Variations occurring within the E7 ORF in most instances do not affect amino acids sequence. In other cases, amino acid changes are located in regions considered not relevant for the transforming activity of the protein [Phelps et al., 1992]. However, in this study, a change at position 31, replacing Ser with Arg was identified. This change should alter the phosphorylation of E7 protein by Casein Kinase II (CKII) [Wise-Draper and Wells, 2008].

Examination of the HPV 16 *L1* gene showed that 64.1% (25/39) of the samples contained the prototype sequence. The L1 protein represents the major capsid protein and contains both linear and conformational epitopes, that are responsible for the evocation of a type -specific, neutralizing, antibody-mediated immune response.

Sequence analysis revealed the presence of three non-conservative amino acid mutations in the HI-loop, one of the two regions of the L1 protein that constitute the immunodominant epitope of HPV 16 [Giroglou et al., 2001; Bishop et al., 2007]. The novel variations R364C and N367D are located at the base of the HI-loop, in a region whose structure is highly conserved among different HPV types. All samples containing this mutant protein were diagnosed as having a histological lesion ranging from cervical intraepithelial neoplasia grade 2–3 to carcinoma in situ. The substitution T379P was already described in HPV 16 Af-2 and AA.

In conclusion, some novel mutations were uncovered within the HPV 16 *E6* and *L1* genes encoding for amino acids located at positions likely to be important for the biological activity of the E6 protein and host immune recognition of L1 protein.

REFERENCES

Alvarez-Salas LM, DiPaolo JA. 2007. Molecular approaches to cervical cancer therapy. Curr Drug Discov Technol 4:208–219.

Andersson S, Alemi M, Rylander E, Strand A, Larsson B, Sällström J, Wilander E. 2000. Uneven distribution of HPV 16 E6 prototype and variant (L83V) oncoprotein in cervical neoplastic lesions. Br J Cancer 83:307–310.

Baseman JG, Koustsky LA. 2005. The epidemiology of human papillomavirus infections. J Clin Virol 32:16–24.

Bishop B, Dasgupta J, Klein M, Garcea RL, Christensen ND, Zhao R, Chen XS. 2007. Crystal structures of four types of human papillomavirus L1 capsid proteins: Understanding the specificity of neutralizing monoclonal antibodies. J Biol Chem 282:31803–31811.

Brady CS, Duggan-Keen MF, Davidson JA, Varley JM, Stern PL. 1999. Human papillomavirus type 16 E6 variants in cervical carcinoma: Relationship to host genetic factors and clinical parameters. J Gen Virol 80:3233–3240.

Ciotti M, Giuliani L, Ambrogi V, Ronci C, Benedetto A, Mineo TC, Syrjanen K, Favalli C. 2006a. Detection and expression of human papillomavirus oncogenes in non-small cell lung cancer. Oncol Rep 16:183–189.

Ciotti M, Coletti A, Giuliani L, Cappiello G, Syrjanen K, Favalli C. 2006b. The p53 codon 72 arg/arg homozygous women in Central Italy are at increased risk for HPV infections. Anticancer Res 26:3745–3748.

Crook T, Fisher C, Vousden KH. 1991. Modulation of immortalizing properties of human papillomavirus type 16 E7 by p53 expression. J Virol 65:505–510.

De Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H. 2004. Classification of papillomaviruses. Virology 324:17–27.

Dyson N, Howley PM, Munger K, Harlow E. 1989. The human papillomavirus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. Science 243:934–937.

Ellis JR, Keating PJ, Baird J, Hounsell EF, Renouf DV, Rowe M, Hopkins D, Duggan-Keen MF, Bartholomew JS, Young LS. 1995. The association of an HPV16 oncogene variant with HLA-B7 has implications for vaccine design in cervical cancer. Nat Med 1:464–470.

Giroglou T, Sapp M, Lane C, Fligge C, Christensen ND, Streeck RE, Rose RC. 2001. Immunological analyses of human papillomavirus capsids. Vaccine 19:1783–1793.

Hall TA. 1999. Bioedit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98 NT. Nucl Acids Symp Ser 41:95–98.

Hildesheim A, Schiffman M, Bromley C, Wacholder S, Herrero R, Rodriguez A, Bratti MC, Sherman ME, Scarpidis U, Lin QQ, Terai M, Bromley RL, Buetow K, Apple RJ, Burk RD. 2001. Human papillomavirus type 16 variants and risk of cervical cancer- J Natl. Cancer Inst 93:315–318.

Ho L, Chan SY, Chow V, Chong T, Tay SK, Villa LL, Bernard HU. 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 29:1765–1772.

Munger K. 2002. The role of human papillomaviruses in human cancers. Front Biosci 7:d641–d649.

Munoz N, Bosch FX, de Sanjose S, Herrero R, Castellsangue X, Shah KV, Snijders PJF, Meijer CJLM. 2003. Epidemiologic classification of human papillomavirus types associated with cervical cancer. N Engl J Med 348:518–527.

Munoz N, Castellsague X, de Gonzalez AB, Gissmann L. 2006. Chapter 1: HPV in the etiology of human cancer. Vaccine 24S3:S1–S10.

Nindl I, Rindfleisch K, Lotz B, Schneider A, Durst M. 1999. Uniform distribution of HPV 16 E6 and E7 variants in patients with normal histology, cervical intra-epithelial neoplasia and cervical cancer. Int J Cancer 82:203–207.

Ortiz M, Torres M, Muñoz L, Fernández-García E, Canals J, Cabornero AI, Aguilar E, Ballesteros J, Del Amo J, García-Sáiz A. 2006. Oncogenic human papillomavirus (HPV) type distribution and HPV typ16 E6 variants in two Spanish population groups with different levels of HPV infection risk. J Clin Microbiol 44:1428–1434.

Pande S, Jain N, Prusty BK, Bhambhani S, Gupta S, Sharma R, Batra S, Das BC. 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 46:1060–1066.

Phelps WC, Munger K, Yee CL, Barnes JA, Howley PM. 1992. Structure -function analysis of the human papillomavirus type 16 E7 oncoprotein. J Virol 66:2418–2427.

Seedorf K, Krammer G, Durst M, Suhai S, Rowekamp WG. 1985. Human papillomavirus type 16 DNA sequence. Virology 145:181–185.

Swofford D, Sullivan J. 2003. Phylogeny Inference based on parsimony and other methods with PAUP*. In: Salemi M, Vandamme AM, editors. The phylogenetic handbook—A practical approach to DNA and protein phylogeny. New York:Cambridge University Press. pp. 160–206.

Tornesello ML, Duraturo ML, Salatiello I, Buonaguro L, Losito S, Botti G, Stellato G, Greggi S, Piccoli R, Pilotti S, Stefanon B, De Palo G, Franceschi S, Buonaguro FM. 2004. Analysis of human papillomavirus type-16 variants in Italian women with cervical intraepithelial neoplasia and cervical cancer. J Med Virol 74:117–126.

Villa LL, Sichero L, Rahal P, Caballero O, Ferenczy A, Rohan T, Franco EL. 2000. Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. J Gen Virol 81:2959–2968.

Wise-Draper TM, Wells SI. 2008. Papillomavirus E6 and E7 proteins and their cellular targets. Front Biosci 13:1003–1017.

Xi LF, Koutsky LA, Galloway DA, Kuypers J, Hughes JP, Wheeler CM, Holmes KK, Kiviat NB. 1997. Genomic variation of human papillomavirus type 16 and risk for high grade cervical intraepithelial neoplasia. J Natl Cancer Inst 89:796–802.

Yamada T, Wheeler CM, Halpern AL, Stewart AC, Hildesheim A, Jenison SA. 1995. Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments. J Virol 69: 7743–7753.

Yamada T, Manos MM, Peto J, Greer CE, Munoz N, Bosch FX, Wheeler CM. 1997. Human papillomavirus type 16 sequence variation in cervical cancers: A worldwide perspective. J Virol 71:2463–2472.

Zehbe I, Voglino G, Delius H, Wilander E, Tommasino M. 1998a. Risk of cervical cancer and geographical variations of human papillomavirus 16 E6 polymorphisms. Lancet 352:1441–1442.

Zehbe I, Wilander E, Delius H, Tommasino M. 1998b. Human papillomavirus 16 E6 variants are more prevalent in invasive cervical carcinoma than the prototype. Cancer Res 58:829–833.

Zur Hausen H. 1991. Human papillomaviruses in the pathogenesis of anogenital cancer. Virology 184:9–13.