Molecular and evolutionary analysis of HPV16 E6 and E7 genes in Greek women

D. Tsakogiannis,1 A. Papadopoulou,1 G. Kontostathi,1 I. G. A. Ruether,1 Z. Kyriakopoulou,1 T. G. Dimitriou,1 G. Orfanoudakis2 and P. Markoulatos1

Correspondence
P. Markoulatos
markoulatos@bio.uth.gr

1University of Thessaly, School of Health Sciences, Department of Biochemistry & Biotechnology, Microbiology-Virology Laboratory, Larissa, Greece
2University of Strasbourg, Ecole Superieure de Biotechnologie de Strasbourg, CNRS FRE 3211-Oncoprotein Group, 67412 Illkirch, France

Human papillomavirus type 16 (HPV16) non-European variants have been associated with persistent infection and cervical cancer development, while the L83V variant of the E6 gene has been correlated with the progression of cervical malignancy. The present study investigated the presence of the HPV16 L83V variant in Greek women. Molecular evolutionary analysis of the HPV16 E6 and E7 oncogenes was conducted in order to estimate the evolution of the HPV16 genome in the Greek population. The E6 L83V variant was found in 78.2 % of high- and 64.28 % of low-grade specimens. Moreover, the prototype and E6 L83V variants were both prevalent in high- and low-grade malignancies in Greek women. Selective pressure analysis of the individual amino acid residues of HPV16 sequences from the Greek population indicates that codon 83 of the E6 protein, as well as codon 85 of the E7 protein, are undergoing positive selection. Novel sequence variations were recorded within the E6 and E7 genes in cervical samples, characterized as (T350G) European variants. However, no signal of intratypic recombination event was identified within the E6–E7 region. Molecular and evolutionary analyses of HPV16 genomes from distinct geographical locations might provide valuable information about viral evolution and oncogenicity.

INTRODUCTION

Human papillomavirus (HPV) is composed of a 7.9 kb double-stranded circular DNA encapsidated in an icosahedral capsid. HPV infects cutaneous and mucosal epithelial cells, leading to malignant transformation (Muñoz, 2000). A total of 40 HPV types have been related to anogenital tract malignancy and have been grouped as high-risk (HR) and low-risk (LR) genotypes (Bernard et al., 2010, 2006; zur Hausen, 1996). Persistent infection with HR HPV is a major cause of cervical cancer worldwide. HPV16 and HPV18 are the most common types identified in invasive cervical cancer cases (Li et al., 2011; de Sanjose et al., 2010).

The HPV genome is divided into three viral domains, consisting of the early region that encodes five early genes (E1, E2, E4, E6, E7), the late region that encodes two late genes (L1, L2) and the long control region (LCR). HPV types are defined according to the percentage of sequence identity at the L1 gene. Sequences from any two distinct HPV types differ by at least 10 % at the nucleotide level. Isolates derived from the same HPV type are referred as variants when less than 10 % sequence diversity in L1 gene is encountered (Bernard et al., 2006, 2010). The HPV genome has a low evolutionary rate, and is also characterized by the absence of recent recombination events (Chan et al., 1997; Halpern, 2000). Nevertheless, previous studies that have used alpha HPV sequences from the public database at Los Alamos National Laboratory identified significant recombination signals, located in E6, E7, L1 and L2 genes at different HPV types and namely within HPV16 (Carvajal-Rodriguez, 2008; Angulo et al., 2007).

The HPV16 genome has been extensively investigated for nucleotide polymorphisms (that occur via random mutations) and a number of HPV16 variants have been detected in distinct geographical locations and ethnic groups (Schiffman et al., 2010; Chen et al., 2005, 2009). In particular, HPV16 is divided into four distinct phylogenetic branches, the distribution of which varies geographically. These four distinct groups are designated as (i) European–Asian (including the sublineages European and Asian), (ii) African I, (iii) African II and (iv) North-
American/Asian American (Cornet et al., 2012; Yamada et al., 1995, 1997). Several analyses revealed a number of intratypic nucleotide polymorphisms within the E2, E6, E7, L1 and LCR regions (Wu et al., 2006; Swan et al., 2005; Eriksson et al., 1999; Wheeler et al., 1997; Yamada et al., 1995, 1997). Interestingly, intratypic nucleotide polymorphisms within the HPV16 E4 ORF are capable of correctly classifying an HPV16 genome in any of the four phylogenetic branches mentioned above (Tsakogiannnis et al., 2012). Different HPV16 variants might present distinct biological functions, thus resulting in different pathogenicity and immunogenicity. Epidemiological studies have suggested that the presence of HPV16 non-European variants constitute HR factors for development of cervical malignancy and invasive cancer (Sanchez et al., 2011; Quint et al., 2011; Cornet et al., 1997). Interestingly, intratypic nucleotide polymorphisms within the HPV16 E6 gene (rather than in other HPV16 genes) influence the progression of cervical malignancy, and that additional factors are likely to play a significant role as well (Lee et al., 2008).

This study considers the significant role of the E6 and E7 oncoproteins, and focuses on these two regions in order to investigate the nucleotide variability between different HPV16 variants that circulate in the Greek population. The most dominant HPV16 variants were detected in high- and low-grade cervical malignancies in Greek patients. Finally, molecular evolution analyses of the E6 and E7 genes were conducted in order to better understand viral evolution and pathogenicity.

**METHODS**

**Sample collection and DNA isolation.** A total of 53 HPV16 positive cervical specimens were collected from ThinPrep (46 specimens) and cervical-stained smears (seven specimens, A3, A4, A6, A13, A31, PNL7 and PNL16). In particular, 23 cervical samples were diagnosed as high-grade cervical intraepithelial neoplasia (CIN II, III), 28 as low-grade cervical intraepithelial neoplasia (CIN I) and two cervical specimens (PNL7, PNL16) were characterized as cervical cancers. The cervical samples were collected from patients attending for the annual PAP test, while smears were collected from patients with a history of abnormal smears. Patient ages ranged from 25 to 60 years.

DNA from ThinPrep samples was extracted using the chaotropic agent guanidine thiocyanate (GuSCN) (Casas et al., 1995). DNA from cervical-stained smears was extracted following the method described by Puranen et al. (1996). The quality of extracted DNA was evaluated by PCR amplification of a 498 bp fragment of human β-actin gene (Li et al., 2008). HPV typing was performed using the Nested Multiplex PCR described by Sotlar et al. (2004).

**E6–E7 amplification by PCR.** The cervical samples were subjected to PCR amplification for HPV16 E6–E7 genes using the primer set: HPV-16 41S-AGCCACCCGAAAGTACCA-3/HPV-16 75T-ATGGGGCACATCTCTTACCA-3. The forward primer hybridized 20 bp downstream from the transcriptional start site of the HPV16 E6 gene, while the reverse primer hybridized 10 bp upstream from the 3’ end of the HPV16 E7 gene.

PCR was performed in a final volume of 50 μl. PCR mixture contained 50 pmol of each primer, 5 x Green GoTaq Flexi buffer (Promega), 2 mM MgCl2, 1 mM dNTPs and 1.25 U of thermostable DNA Polymerase (GoTaq, Promega). The cycling conditions were as follows: 40 cycles of 40 s at 95 °C, 50 s at 55 °C and 1 min at 72 °C. The first cycle was proceeded by a 2 min denaturation step at 95 °C and the last cycle was followed by a 5 min elongation step at 72 °C.

**Cloning and sequencing of E6–E7 genes.** The 715 bp amplicons of the HPV16 E6–E7 genes were ligated into pGEM T-easy vector (Promega) and transformed into _Escherichia coli_ DH5a competent cells. Recombinant plasmid DNA was purified using the Nucleospin plasmid kit (Macherey-Nagel GmbH) and the plasmids were subjected to sequencing at Macrogen. In order to confirm nucleotide variations, two independent PCR assays were carried out for each individual cervical sample and amplicons from the two distinct PCR assays were subjected to cloning. Three clones from each individual amplicon were isolated. In addition, the two strands of all tested E6–E7 clones were used in sequencing reaction analysis.

In order to identify nucleotide polymorphism and amino acid substitutions into the E6 and E7 ORFs, multiple sequence alignments were performed with the MUSCLE algorithm, in MEGA v.5.5 software (Edgar, 2004; Tamura et al., 2007). The multiple sequence alignment was conducted between the cloned E6–E7 sequences and the reference sequence of HPV16 genome (HPV16R), available at the HPV16 Sequence Database (Los Alamos National Laboratory).

**Phylogenetic analysis.** Specific intratypic sequence polymorphisms within the E6 and E7 genes were used in order to classify the E6–E7 cloned sequences to the corresponding HPV16 variants. In order to examine the phylogenetic clustering of HPV16 E6–E7 cloned sequences, a multiple sequence alignment was carried out among the E6–E7 cloned sequences, the reference sequence of HPV16 (HPV16R) available at the HPV16 Sequence Database (Los Alamos National Laboratory), and those of representative sequences of HPV16 variants available in the GenBank sequence database under accession numbers for European variant AF536179 (European-German type), AY686580, AY686581, AF125673 and FJ006723 (North American), for East Asian variant AF534061, for African type I variant AF472508 and AF536180, for African type II variant AF472509 and for the Asian American variant, AF402678. The sequence alignment was performed using the MUSCLE algorithm (Edgar, 2004) of MEGA v.5.5 (Tamura et al., 2007).

**Selection and recombination analysis.** Maximum-likelihood (ML) and empirical Bayesian (EB) methods were implemented in order to identify the selective pressure acting upon the E6 and E7 genes. For the ML analysis, the fixed effects likelihood model (FEL) was implemented at the Datamonkey server (Kosakovsky Pond &11, 2007). Implementation of the EB analysis was through the Selection 2.4 program (Stern et al., 2007), using the highest precision level. In addition, a Bayesian analysis using Markov chain Monte Carlo methods using the program BEAST v.1.6.2 (Drummond & Rambaut, 2007) was also implemented. Bayesian analysis consisted of four chains with random starting trees. Convergence was assessed with effective sample size values, after a burning of 1 × 10⁶ steps. The median mutation rate in the first, second and third position of the codon in the E6 and E7 reading frames was estimated using Trace v.1.5 within the BEAST software package. The general time reversible
(GTR) plus Gamma substitution model was selected according to Akaike information criterion by the FindModel online software (www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html). In addition, the single recombination breakpoint (SBP) method in Datamonkey server (Kosakovský Pond et al., 2006) was used in order to identify any evidence of recombination signal.

**Phylogenetic tree.** An ML phylogenetic tree was computed (1000 bootstrap replicates) from the alignment of the reference HPV16 E6 sequence (HPV16R), the E6 cloned sequences derived from the present study (Tables 1 and 2) and the HPV16 E6 sequences that were presented in the study of Carvajal-Rodriquez (2008) (Fig. 1). Thus, it was possible to evaluate how the sequences in the present study classify with respect to codons 10, 14 and 83. Prior to phylogenetic estimation, the online software FindModel (www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html) was used to identify the best evolutionary model as mentioned above. The GTR plus Gamma nucleotide substitution model was used for the E6 phylogenetic tree. The nucleotide sequences used are available in GenBank under the following accession numbers: for European variant AF125673, AY686583, AY686580, AY686584 and AF536179 (European-German type), for East Asian variant AF534061, for African type I variant AF472508 and AF536180, for African type II variant AF472509 and for the Asian American variant AF402678, AY686579 and AY686582.

**RESULTS**

**E6–E7 sequence variations**

Sequence analysis of the HPV16 E6 gene shows that this viral region is polymorphic. The most frequently observed nucleotide changes the T350G that induces the L83V amino acid substitution. In high-grade cases, the nucleotide variation T350G was detected in 69.6% (16/23) of cervical specimens, while in low-grade dysplasia, the T350G polymorphism was identified in 64.3% (18/28) of cervical samples (Tables 1 and 2). Nevertheless, the nucleotide variations A184G, G201A, G219A, A280T, T302A, A336G, G514A, C523T and C539T were identified for the first time in the present study (Table 1). The nucleotide variations at positions 184, 201, 219, 302, 336, 514 and 539 induced the amino acid substitutions I27M, C33Y, R39Q, L76I, H78R, M137I and R146G, respectively. However, two silent mutations (A280T, C523T) were identified within the E6 ORF.

The nucleotide variations A184G, G201A, G219A, C523T and C539T were detected in high- and low-grade cervical samples (Table 1). In addition, the silent nucleotide variation A280T was detected in cloned sequences derived from the high-grade cervical samples AT5 and AT20 and the low-grade sample PNL2, while the nucleotide changes T302A, A336G and G514A were identified only in low-grade cervical intraepithelial neoplasias (CIN) (Table 1). The newly identified nucleotide variations were detected in HPV16 strains that harbour the T350G nucleotide variation. Nevertheless, G219A was identified in both European and African type II variants. In particular, G219A was found in low-grade sample C2279 and high-grade sample AT9 that were characterized as European variants. In addition, the G219A was identified in two cloned sequences derived from the low-grade cervical specimen A17 (A17 C1, A17 C2), characterized as African type II (Tables 1 and 2).

Sequence analysis of the HPV16 E7 gene revealed that this ORF appears to be more conserved among distinct HPV16 variants. However, seven nucleotide changes were identified for the first time in the present report. The nucleotide variations C565T, T396A, G685A, G813A, G647A and C818T were detected in cervical samples that harboured the nucleotide variation T350G. These variations induce the amino acid changes H2Y, M12K, A42T, M84I, G85S, G85D and T86I, respectively. The nucleotide variations T596A, G685A, G814A, G815A and C818T were detected in both high- and low-grade cervical intraepithelial neoplasias (CIN) (Table 1). In contrast, the nucleotide changes C565T, G813A were recorded only in low-grade cervical cases (Table 1). In particular, the C565T was detected in the low-grade cervical cases A24 and 1620, while the G813A was identified in two cloned sequences derived from the cervical sample A24 (A24 C2, A24 C3) and in the low-grade cervical sample 858 (Table 1).

**HPV16 variant determination**


According to nucleotide analysis of E6 and E7 ORFs, the HPV16 prototype sequence was identified in ten low-grade (2121, 2127, 2148, 2216, 796, ATT8, ATT10, ATT12, ATT18, ATT45) and in six high-grade cervical samples (PNL8, PNL10, PNL11, PNL12, ATT20, ATT24) (Tables 1 and 2). In addition, European sequences (T350G) were recorded in 18 low- and 15 high-grade CIN cases (Tables 1 and 2). Asian American sequences were reported in the high-grade cervical sample A3 and in the cancer case PNL7, while African type I sequence was identified in the high-grade cervical sample A4. The cancer case PNL16 was characterized as (T350G) European variant. In addition, two different HPV16 variants were identified in the low-grade cervical sample A17. In particular, one cloned sequence of the cervical specimen A17 (A17 C2) was characterized as a European variant, while in two other cloned sequences, A17 C1 and A17 C3, E6–E7 African type II sequences were detected (Tables 1 and 2). However, the possibility of an artefactual presence of both African type II
Table 1. Nucleotide variations in HPV16 E6 and E7 ORFs associated with both high- and low-grade cervical cases

- Absence of nucleotide variation or amino acid substitution; A, adenine; T, thymine; G, guanine; C, cytosine. A number of specific intratypic nucleotide polymorphisms were detected in the E6 and E7 genes that were used to describe the distinct HPV16 variants. The sequence variation T350G was identified in both high- and low-grade cervical specimens. The cervical samples PNL7 (invasive cancer) and A4 (CINIII) were characterized as Asian American and African type I variants, respectively. In addition, the low-grade cervical sample A17 presented sequence homology with African type II variant.

<table>
<thead>
<tr>
<th>Clone*</th>
<th>E6 region</th>
<th>E7 region</th>
<th>Variants</th>
<th>Accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td>A31C1</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td>-</td>
</tr>
<tr>
<td>A31C2</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>-</td>
</tr>
<tr>
<td>A13C1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td>A13C2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td>A13C3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td>A5C1</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td>-</td>
</tr>
<tr>
<td>PMLA1C1</td>
<td>-</td>
<td>-</td>
<td>T</td>
<td>-</td>
</tr>
<tr>
<td>PMLA2C1</td>
<td>-</td>
<td>-</td>
<td>T</td>
<td>-</td>
</tr>
<tr>
<td>ATT20C1</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td>-</td>
</tr>
<tr>
<td>ATT20C2</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td>-</td>
</tr>
<tr>
<td>European</td>
<td>G</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*The first letter or number corresponds to the code of the cervical sample. The last letter C1, C2, C3 corresponds to the number of the cloned sequence.
Table 2. Sequence analysis of HPV16 E6 and E7 ORFs in low- and high-grade cervical samples

<table>
<thead>
<tr>
<th>Clones</th>
<th>E6 region</th>
<th>E7 region</th>
<th>Variants</th>
<th>Accession number</th>
<th>Genbank</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIN-HG</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PNL1 C1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PNL2 C1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PNL11 C1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ATT22 C1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ATT10 C1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*The first letter or number corresponds to the code of the cervical sample. The last letter C1, C2, C3 corresponds to the number of the cloned sequence.
and European variants in cervical sample A17 cannot be excluded, due to sequencing only two PCR products.

**Selection and recombination analysis**

Molecular evolution analyses for E6 gene revealed that the relative substitution rates for the three codon position was CP1=0.826, CP2=1.015, CP3=1.159, whereas a corresponding analysis of the E7 ORF estimates CP1=0.728, CP2=1.2 and CP3=1.07. Only under the EB method (Stern et al., 2007) was the codon 83 detected as a positively selected site. Moreover, in the E7 gene, both the ML and EB methods detected codon 85 as a positively selected site, but only under the EB method was the finding statistically significant. The ML phylogenetic tree (Fig. 1) for the sequences in the present study and for codons 10, 14 and 83 (based on previous work by Carvajal-Rodrı́guez, 2008) revealed a sufficient classification of HPV16 genome in the corresponding phylogenetic branch. However, the SBP program (Kosakovsky Pond et al., 2006) revealed no evidence of recombination in the multiple sequence alignment between the E6–E7 cloned sequences, the reference sequence of HPV16 and those of representative sequences of HPV16 variants.

**DISCUSSION**

The main objective of this study was to understand the genomic evolution of the HPV16 genome, via the study of genomic diversity and evolution of the E6 and E7 oncogenes. According to the sequence analysis, specific nucleotide polymorphisms were detected and used to explore the intratypic heterogeneity of the HPV16 genome in the Greek population. The most frequently identified HPV16 variant was the European variant (T350G), following the HPV16 prototype strain (Tables 1 and 2) (Zuna et al., 2009; Andersson et al., 2000). In particular, the European variant (T350G) was identified in 65.2% of high-grade cervical cases and in 64.2% of low-grade specimens, while the prototype strain (350T) was detected in 26.1% of high-grade and in 35.7% of low-grade CIN samples (Tables 1 and 2). The nucleotide change T350G corresponds to amino acid substitution L83V, which is one of the changes detected as undergoing positive selection (Chen et al., 2005; DeFilippis et al., 2002; Carvajal-Rodrı́guez, 2008). Previous studies have revealed that HPV16 E6 L83V variant is prevalent in high-grade lesions and is associated with progression of cervical malignancy (Grodzi et al., 2006;
Andersson et al., 2000; Yamada et al., 1997). In the present study, the E6 L83V variant was detected in 69.5% of high-grade and 64.28% of low-grade cases. Moreover, the L83V variant was identified in the two cervical cancer cases (PNL7 and PNL16). The difference among high- and low-grade malignancies was not statistically significant. The prototype and the E6 L83V variant were both prevalent in high- and low-grade malignancies in Greek women.

Previous studies have proposed that specific intratypic variants might influence the progression of preinvasive cervical lesions to cancer. In particular, it was proposed that there is a strong relation between Asian American variant and cervical cancer development (Sanchez et al., 2011; Tornesello et al., 2011; Quint et al., 2010; Junes-Gill et al., 2008). According to sequence analyses of the E6–E7 cloned sequences and the representative sequences of HPV16 variants, non-European variants were identified in four cervical samples (PNL7, A3, A4, A17). The HPV16 Asian American variant was identified in cervical sample PNL7, diagnosed as invasive cervical cancer and in the high-grade cervical sample A3 (Table 1). However, two E6–E7 cloned sequences derived from the low-grade cervical sample A17 were characterized as African type II variants, while one E6–E7 cloned sequence derived from the cervical sample A17 was characterized as European variant (Tables 1 and 2). This was the only mixed infection that was recorded in the present study although an artefactual event could not be excluded due to sequencing only two PCR products. In addition, the cervical sample A4 was characterized as an African type I variant. Taking these data into account, we conclude that non-European variants are circulating in Greek population and are detected in high- and low-grade CINs (Tables 1 and 2).

Sequence analysis of the E6 gene revealed significant sequence heterogeneity, clustering the HPV16 variants in distinct phylogenetic branches (Pande et al., 2008; Swan et al., 2005; Wheeler et al., 1997; Yamada et al., 1995, 1997). Nevertheless, nine new nucleotide variations (A184G, G201A, G219A, A280T, T302A, A336G, G514A, C523T and C539T) were detected in the present study (Table 1) in cervical samples that were characterized as (T350G) European variants. In particular, the nucleotide variations A184G, G201A, G219A, T302A and A336G induced the amino acid changes I27M, C35Y, R39Q, L67I and H78R, respectively. These amino acid changes were located at the N-terminal domain of the E6 oncoprotein (Ghittoni et al., 2010; Boulet et al., 2007). Additionally, the nucleotide changes G514A and C539T caused the amino acid substitutions M137I and R146C, respectively. These changes were located at the C-terminal domain of the E6 protein into the Zn$^{2+}$ Finger portion (Ghittoni et al., 2010; Boulet et al., 2007). Moreover, the nucleotide variations A280T and C523T did not induce amino acid substitutions within the E6 protein. The nucleotide variation A280T was detected in both high- and low-grade cervical samples (Table 1).

Previous studies have proposed that the E7 gene is more conserved than E6 (de Boer et al., 2004; Wu et al., 2006; Pande et al., 2008). Nevertheless, seven new nucleotide variations that clustered to the same phylogenetic branch with the (T350G) European variant. The nucleotide changes C565T, T596A, G685A, G813A, G814A, G815A and C818T caused the amino acid substitutions H2Y, M12K, A42T, M84I, G85S, G85D and T86I, respectively. In particular, the amino acid substitutions H2Y and M12K were located at the CR1 domain of the E7 oncoprotein (Münger et al., 2004). In addition, the amino acid changes M84I, G85S, G85D and T86I were detected in the C-terminal domain of the E7 protein into the Zn$^{2+}$ Finger portion (Ghittoni et al., 2010) (Table 1). The new nucleotide and amino acid substitutions of E6 and E7 ORFs that were detected in European variants require further investigations, in order to discern their functional implications.

However, the ability of the E6 oncoprotein to interact with distinct p53 variants might induce strong selective pressure upon the E6 gene (DeFilippis et al., 2002). In particular, codons 10, 14 and 83 of the E6 gene are under positive selection and the L83V amino acid substitution is associated with cervical cancer development (Chen et al., 2005; DeFilippis et al., 2002; Carvajal-Rodrı´guez, 2008; Lee et al., 2008). In the present study, positive selection was found in codon 83 of the E6 gene. However, no signal of positive selection was found in codons 10 and 14, by the ML and EB methods. Considering the positive selection analysis and the ML phylogenetic tree (Fig. 1), we assume that codon 83, which is implicated in cervical cancer progression, has undergone positive selection within the European population. However, nucleotide and phylogenetic analyses revealed that codon 10 distinguishes African I and II, while codon 14 separates African type II from the non-African variants (Carvajal-Rodrı´guez, 2008). Absence of positive selection in codons 10 and 14 might be caused by dominance of HPV16 European variant in the Greek population (Fig. 1).

It was proposed that positive selection is acting upon codon 29 of the E7 protein (Sun et al., 2012). In the present analysis, only codon 85 was found to be under positive selection by EB analysis. Codon 85 is located within the C-terminal domain of E7 protein. This protein domain is implicated in protein interactions with the pRb and other cellular proteins, and may act as a dimerization domain (Münger et al., 2004). Amino acid substitution in codon 85 was detected only in cervical samples that were characterized as European variants. The role of positive selection in codon 85 in the Greek population should be further investigated in order to discern the interaction between the E7 oncoprotein and host cellular targets.

In conclusion, the HPV16 E6 and E7 genes are polymorphic regions of the viral genome and provide significant information about HPV16 intratypic heterogeneity and evolution. To the best of our knowledge, this is the first study describing nucleotide variations and
evolutionary pressure acting on E6 and E7 regions of HPV16 genomes from the Greek population. Molecular and evolutionary analyses of the HPV16 genome in distinct geographical locations might provide valuable information about the viral pathogenicity and genome evolution.

ACKNOWLEDGEMENTS

We are grateful to E. Panotopoulou, Papanicolaou Research Centre of Oncology and Experimental Surgery, Anticancer Oncology Hospital of Athens ‘St Savvas’, Athens, Greece, for providing the clinical samples and for valuable discussions. The work was supported by research grants of the Postgraduate Programme ‘Applications of Molecular Biology-Genetics. Diagnostic Biomarkers’ (3817) and ‘Biotechnology’ (3439), of the University of Thessaly, School of Health Sciences, Department of Biochemistry & Biotechnology. All authors declare that they have no conflicting or dual interests.

REFERENCES


