Infection, Genetics and Evolution 68 (2019) 185–193

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

Identification of human papillomavirus type 16 variants circulating in the Calabria region by sequencing and phylogenetic analysis of HPV16 from cervical smears

Luisa Galati[a,c], Michele Equestre[b], Roberto Bruni[a], Luisa Accardi[a], Carlo Torti[d], Maria Teresa Fiorillo[e], Giovanni Surace[e], Giorgio Settimo Barreca[c], Maria Carla Liberto[c], Alfredo Focà[c], Anna Rita Ciccaglione[a], Paola Di Bonito[a,*]

[a] Department Infectious Diseases, EVOR unit, Istituto Superiore di Sanità, Rome, Italy
[b] Department Cell Biology and Neurosciences, Istituto Superiore di Sanità, Rome, Italy
[c] Institute of Clinical Microbiology, Department of Health Sciences, "Magna Graecia" University, Catanzaro, Italy
[d] Department of Medical and Surgical Sciences, Unit of Infectious and Tropical Diseases, School of Medicine, University of "Magna Graecia", Catanzaro, Italy
[e] Unit of Microbiology and Virology, Polo Sanitario Nord ASP 5, Reggio Calabria, Italy

ARTICLE INFO

Keywords:
Human Papillomavirus
HPV16
Variant
Lineage
Sublineage

ABSTRACT

Sequence analysis of HPV16 isolates reveals the presence of genome variants with characteristic mutations. The HPV16 variants have different geographical distribution and diverge into four phylogenetic lineages (A, B, C and D) and 16 sub-lineages: A1, A2, A3 (previously known as European variants), A4 (Asian variant), B1, B2, B3, B4, C1, C2, C3, and C4 (African variants), D1 (North-American variant), D2, D3 (Asian-American variants) and D4. Population studies showed that infections with viruses belonging to specific HPV16 sublineages confer different risks of viral persistence and cancer. In this study, 39 HPV16-positive cervical smears from European women living in Calabria (Italy) were analyzed for the presence of HPV16 variants. Cervical DNA extracts were processed by PCR to amplify L1, the Long Control Region (LCR), E6 and E7, which were sequenced. The sequences were concatenated and the 3169 nucleotides long fragments were characterized by BLAST and phylogenetic analysis. A total of 96 Single Nucleotide Polymorphism (SNPs) were detected, 29 of which mapping in the L1, 45 in the LCR, 15 in the E6 and 7 in the E7. The most common SNP was the T350G (29/39 samples, 74.4%), causing the L83 V amino acid change in the E6. Most of the HPV16 isolates (89.7%) had 99% of nucleotide (nt) identity to members of the A1 and A2 sublineages, while 4 isolates had 99% nt identity to members of the B2, B4, C1 and D4 sublineages. In conclusion, viruses belonging to the A1, A2, B2, B4, C1 and D4 HPV16 sublineages were found to circulate in the Calabria region.

1. Introduction

Infections with Human Papillomaviruses (HPVs) cause benign and malignant lesions of the oral and genital tract. HPV16 is responsible for 61% of the cervical cancer (CC) cases (de Sanjose et al., 2010), is the prevalent genotype in anal, vaginal, penile and vulvar tumors, and is also associated with a subset of oral and oropharyngeal cancers (Schiffman et al., 2016). It has been well established that CC is caused by persistent infections of oncogenic HPV types, epidemiologically classified as 'high-risk" (HR-HPV) and categorized in either class 1 or class 2A/2B, including possibly/probably human carcinogens, by the International Agency for Research on Cancer (IARC) (Bouvard et al., 2009). Despite the availability of vaccines against most HR-HPVs and the implementation of a screening campaign for CC prevention in many countries, CC remains a public health problem at a global level. The most recent estimates indicate that, every year, almost 500,000 women are diagnosed with CC worldwide, with 236,000 deaths (De Martel et al., 2017). In Italy, it is estimated that 1515 new cases of CC occur every year, with 697 deaths (Rossi et al., 2013). HPV16 infections are very common in the sexually active population. However, only a small fraction of these infections progress to cancer, where host and viral factors are the leading causes for the development of cancerous lesions (Burk et al., 2013; Burk et al., 2017). In immunocompromised and HIV-positive hosts, HPV infections are more frequent and severe (Reusser

* Corresponding author at: Department of Infectious Diseases, Viral Hepatitis, Oncovirus and Retroviruses (EVOR) Unit, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Roma, Italy.
E-mail address: paola.dibonito@iss.it (P. Di Bonito).

https://doi.org/10.1016/j.meegid.2018.12.024
Received 26 September 2018; Received in revised form 17 December 2018; Accepted 18 December 2018
Available online 20 December 2018

et al., 2015).

HPV16 has a circular double-stranded (ds) DNA genome approximately 7900 bp long, divided into three functional regions: the Long Control Region (LCR), also known as Upstream Regulatory Region (URR), containing regulatory elements for viral replication and transcription; an Early region encoding the E6, E7, E1, E2, E4, E5 proteins, and a Late region encoding the L2 and L1 capsid proteins (Egawa et al., 2015). The E6 and E7 proteins are the most important viral oncoproteins as they subvert the cell cycle control mainly through the impairment of p53 and pRb tumor suppressor activities. L1 and L2 are respectively the major and minor capsid protein and they are responsible for the viral DNA packaging and assembly in an icosahedral structure.

Sequence analysis of viral isolates highlighted the existence of genome variants with characteristic mutations. The L1 is the most conserved gene within the HPV genome and is used for PV classification. L1 nucleotide variability higher than 10% identifies a new HPV. The analysis of complete genomes allows the HPV16 classification in lineages and sublineages. A distinct HPV16 variant lineage has nt variability between 1.0 and 10.0% compared to the prototype genome, while a sub-lineage has nt variability between 0.5 and 1.0% (Burk et al., 2013; Chen et al., 2015). The host-pathogen coevolution has been the main evolutionary driving force of Papillomaviruses (Bravo and Felez-Sanchez, 2015). DsDNA viruses have a lower rate of mutation, if compared to RNA viruses, and this rate should be even slower in the case of HPVs, because these viruses do not code for their own DNA polymerase, but use the high-fidelity cellular enzymes. However, the presence of > 500 HPV16 genome isolates in GenBank suggests that further selective driving forces exist. Sequence and phylogenetic analysis have shown the presence of numerous silent mutations in the HPV16 genomes, not all ascribable to sequencing errors. Several lines of evidence suggest that cytosine deaminases of the APOBEC family act as a mutator by targeting the HPV genome during the infection. The high number of SNPs observed in HPVs could be ascribed to these enzymes, which cause the C to T or C to G nt change (Warren et al., 2017; Hirose et al., 2018). Furthermore, changes in the coding and non-coding regions of HPV genomes progress at differential rates (Pimenoff et al., 2016). The most informative regions for phylogenetic inference are the E4 and E5 genes and the Long Control Region (Nicolás-Párraga et al., 2017), suggesting that intracellular mechanisms along with the genetic drift are the forces driving the HPV16 evolution.

The HPV16 variants show different geographical distribution and diverge into the A, B, C and D phylogenetic lineages and the following sublineages: A1, A2, A3 (previously named European variants), A4 (Asian variant), B1, B2, B3, B4, C1, C2, C3, and C4 (African variants), D1 (North American variant), D2, D3 (Asian-American variants) and D4 (Burk et al., 2013; Mirabello et al., 2018). The HPV16 A1/A2/A3 sublineages account for the largest proportion of HPV16 infections worldwide (Cornet et al., 2012; Pimenoff et al., 2016). Importantly, the frequency of B, C, and D (non-European) variants was reported to increase with the severity of cervical lesions (Tornesello et al., 2004; Freitas et al., 2014) and with progression to cancer (Mirabello et al., 2016).

In Italy, the presence of HPV16 variants in cervical samples has been previously studied in Caucasian (Cento et al., 2009; Tornesello et al., 2004) and non-Caucasian women (Tornesello et al., 2007), in men and in HIV-positive or -negative cohorts of few Italian Regions (Garbuglia et al., 2007; Tornesello et al., 2008a; Tanzi et al., 2009; Tornesello et al., 2008b). Overall, the results showed that A is the main HPV16 lineage circulating in Italy, while the B and D lineages are often associated with high-grade cervical lesions and cancer (Cento et al., 2009; Tornesello et al., 2011; Tornesello et al., 2004). The data about the circulation of high-risk HPV variants in a specific geographical region could be of great importance for planning cervical cancer prevention interventions. Recently, we reported that HPV16 is the high-risk type most frequently detected in Calabrian women (Galati et al., 2017); in the present study, we analyzed by sequencing the variants of HPV16 circulating in women of the same geographical region. A total of 39 samples of exfoliated cervical cells were collected in women with different conditions of the cervix. The L1, LCR, E6 and E7 viral regions were amplified from cervical cell DNA and sequenced by the Sanger method. The L1-LCR-E6-E7 concatenated sequences were used as input in BLAST search to detect the most related HPV16 genomes. A multiple sequences alignment of the 39 L1-LCR-E6-E7 concatenated sequences, their most related HPV16 genomes and the 16 HPV16 sublineage prototypes, previously identified (Burk et al., 2013; Pimenoff et al., 2016; Mirabello et al., 2018), was performed using Clustal W. Then, phylogenetic trees were inferred to show the relationship among the sequences. The analyzed viral region encodes proteins with relevant functions in the virus life cycle and in the host immune response; in addiction, it contains useful targets for HPV16 diagnostic kits and vaccines and sufficient diagnostic SNPs to distinguish the variants.

2. Materials and methods

2.1. Sample collection and HPV detection

Thirty-nine HPV16-positive cervical specimens were selected among 1298 samples subjected to HPV-DNA diagnosis at Polo Sanitario Nord, Azienda Sanitaria Provinciale 5 of Reggio Calabria, from February 2016 to June 2017. Cervical scrapes were collected with a cytobrush in PCR Cell Collection Media (Roche) during clinical examination. The Linear Array HPV Genotyping® test (Roche, USA) was used for HPV DNA testing, while E6 and E7 mRNAs from high-risk HPV types (16, 18, 31, 33 and 45) were detected by a commercial Nucleic Acid Sequence Based Amplification (NASBA) assay (NucliSENS EasyQ HPV; BioMérieux, France). The women with a HPV-DNA test positive for HPV16 were invited to participate in the study and the cervical specimens were further analyzed, after an informed consent was signed. Valid Pap smear results and some demographic information were collected from the participating women. This study was approved by the Regional Ethical Committee in compliance with the Declaration of Helsinki.

2.2. PCR reactions, DNA sequencing, and analysis

DNA was extracted from endocervical cells collected in PCR Cell Collection Media (Roche) by the automated system NucliSENS EasyMag (Biomérieux). A cell suspension was processed by "specific B" protocol of DNA extraction according to the manufacturer's directions. DNA was quantified at 260 nm using a NanoDrop spectrophotometer (ThermoFisher scientific). Cervical DNA extracts were subjected to PCR amplification using oligonucleotides designed to amplify L1, LCR, E6, and E7 of all HPV16 variants (Supplementary Table S1). The sequences of the L1 gene (1595 bp) were obtained by sequencing three overlapping PCR-products; the sequences of the E6-E7 (867 bp) genes and LCR (888 bp) were obtained by sequencing a single PCR-product. PCR reaction contained 100 ng of DNA, AmpliTaq Gold® 360 Master mix (ThermoFisher Scientific), reaction buffer 10% v/v, 3 mM $MgCl_2$, dNTPs 0.8 mM each, oligonucleotide primers 0.2 μM each, and autoclaved distilled water, in a final volume of 50 μl. The PCR conditions are in Supplementary Table S1 (column 5). PCR products were visualized by 1.2% agarose gel electrophoresis and purified by the PCR reaction or agarose gel using Illustra GFX PCR DNA and gel band purification kit (GE Healthcare). Fragments were sequenced by an automated DNA Sanger sequencer based on fluorescent dye terminator (genetic analyzer Beckman Coulter Inc., Fullerton, CA). Double-stranded DNA sequencing was performed using the same PCR primers. For the E6-E7 region, an additional sequencing reaction was carried out using an internal primer (Supplementary Table S1). Nucleotide sequences were manually edited and assembled into the L1-LCR-E6-E7 whole region (3169 bp) using Codon Code Aligner software (CodonCode Corporation, MA, US). The L1-LCR-E6-E7 concatenated sequences were aligned to those of the

NCBI GeneBank database (http://blast.ncbi.nlm.nih.gov/Blast.cgi), using the Basic Local Alignment Search Tool (BLAST), and the most related HPV16 genomes were retrieved from GeneBank.

The SNPs were visually interpreted using the Molecular Evolutionary Genetics Analysis, MEGA software V 7.0. Nucleotide position of each SNP was given in the test, by convention, following the numbering of K02718 HPV16 genome, taking into account the updates contained in NC_001526.4. To identify new and already known SNPs in the ISS-UMG samples, the sequences deposited either in GeneBank or described in the inherent literature were considered. The ISS-UMG sequences of each LCR, L1, E6, and E7 fragment were aligned by Clustal W, and the SNPs recorded.

HPV16 variant sequences previously isolated in Italy were searched in the NCBI database. Records for 63 LCR, 57 E6, 56 E7, and 85 L1 partial sequences were found, downloaded by NCBI Batch-Entrez and aligned by Clustal W with the ISS-UMG sequences to search conserved SNPs. These partial HPV16 sequences were previously described (Tanzi et al., 2009; Cento et al., 2009; Garbuglia et al., 2007). Since not all the HPV16 sequences have been deposited in public databases, we cannot rule out any error in the list of SNPs identified as new in this study.

Nucleotide sequences were translated into amino acid sequences by MEGA software V 7.0 to record any amino acids (aa) change. The L1 gene was considered to start from the MSLW motif. To map the position of the L1 amino acid changes in the secondary structure of the protein (Table 2, last row) the sequence reported in Bishop and co-authors was used (Bishop et al., 2007).

2.3. Phylogenetic analysis

The overlapping sequences of L1, LCR, E6 and E7 were assembled by CodonCode Aligner into a unique region 3169 nt long, which was then used to performe a Phylogenetic analysis by MEGA, software version 7.0. The ISS-UMG sequences, their most related genomes as found by BLAST analysis, the HPV16 reference genomes of A1, A2, A3, A4, B1, B2, B3, B4, C1, C2, C3, C4, D1, D2, D3, and D4 variant sublineages present in either PAVE (Papillomavirus Episteme database https://pave.niaid.nih.gov/) or in Mirabello et al., 2018, were aligned using the ClustalW algorithm. The best substitution model to apply in the phylogenetic inference was assessed by the “Models” tool in MEGA: the results showed T92 + G to be the best fitting model for the dataset analyzed, that included reference sequences. Phylogenetic tree reconstruction was performed using both the Maximum Likelihood and the Neighbor-Joining methods.

To avoid any ambiguity, we followed the HPV16 variant terminology and variants classification previously reported by Burk and Mirabello (Burk et al., 2013; Mirabello et al., 2018).

The ISS-UMG ID 1-39 sequences were submitted to NCBI database and their accession numbers in GenBank are from MH937376 to MH937414.

The Accession Numbers of the HPV16 genomes, previously classified in sublineages, used as references for the phylogenetic analysis are: KU053891 A1, KU053889 A1, KU053883 A1, HQ644268 A1,HQ644271 A1, KU053853 A1, K02718 A1, HQ644284 A1, AF536179 A2, KU053896 A2, KU053903 A2, KU053895 A2, HQ644236 A3, AF534061 A4, JQ004096 A4, KU053904 A4, HQ644234 A4, HQ644293 B1, KU053907 B1, KU053907 B1, KU053906 B1, KU053908 B1, KU053909 B2, KU053910 B2, HQ644298 B3, KU053915 B3, KU053911 B4, KU053912 B4, KU053913 B4, KU053914 B4, KU053919 C1, KU053917 C1, KU053918 C1, AF472509 C1, KU053916 C1, HQ644244 C2, HQ644243 C2, KU053921 C3, KU053925 C4, KU053922 C4, KU053923 C4, KU053924 C4, KU053929 D1, KU053930 D1, KU053928 D1, KU053926 D1, KU053927 D1, HQ644257 D1, AY686579 D2, KU053935 D2, KU053936 D2, KU053943 D3, KU053944 D3, AF402678 D3, KU053942 D3, KU053941 D3, KU053940 D3, KU053933 D4, KU053934 D4, KU053931 D4, KU053932 D4.

3. Results

3.1. Sequence analysis and phylogeny

Overall, 1296 women were subjected to HPV-DNA diagnosis in a regional public surgery, during a period of 16 months. Of these, 461 (35.5%) women were HPV-positive for one (n = 266, 57.7%) or more (n = 195, 42.3%) HPV genotypes. The most prevalent genotype was HPV16 (n = 75, 16.3%) followed by HPV53 (n = 50, 10.8%), HPV31 (n = 49, 10.6%) and HPV66 (n = 46, 10%). The HPV16-positive women were invited to participate in the present study, and a total of 39 women were enrolled. All women were European, 34 of which born in Italy, 4 in Romania and 1 in Ukraine (Supplementary Table S2); they were in an age range 20–50 years (median age = 33). Thirty-one women showed a single HPV16 infection while 8 showed multiple infections (Supplementary Table S2). The women had cervical lesions of different grade. In 11 women, the Pap-test did not show any atypical cells (Normal), 7 women had cell smears with atypical squamous cells of undetermined significance (ASC-US), 12 had cell smears compatible with a low-grade squamous intraepithelial lesion (L-SIL) and 9 had cell smears compatible with high-grade squamous intraepithelial lesions (H-SIL). The women with high and low SIL lesions followed the triage established by the National guidelines for cervical cancer prevention.

The cervical DNA samples were subjected to PCR amplification to obtain 5 overlapping fragments, of which 3 for the L1 gene, 1 for the E6-E7 genes and 1 for the LCR region. These fragments were sequenced and concatenated into 39 full-length sequences of 3169 bp (ISS-UMG ID), which were analyzed by BLAST to find the best related sequences in GeneBank. The Blast results are reported in Supplementary Table S2 (column 8), with the accession numbers of the best related HPV16 genomes identified (van der Weele et al., 2017; Makowsky et al., 2016; Burk et al., 2013; Chen et al., 2015; Smith et al., 2011). Among the ISS-UMG sequences, 17 sequences showed 100% of nt-identity and 22 sequences showed 99% of nt-identity with the HPV16 genomes in GenBank (column 8). The relationship among the sequences was visualized by a phylogenetic tree reconstruction as described in Materials and Methods. Fig. 1 shows the phylogenetic tree obtained by using the Maximum Likelihood method, with the best substitution model (T92 + G). Only bootstrap values above 70 are reported in the figure. Analysis of the same dataset by the Neighbor-Joining phylogenetic approach, based on genetic distances corrected according to the T92 + G model, produced substantially similar results (see Supplementary Fig. S1). The majority of women (35/39, corresponding to 89.7%) were infected with HPV16 isolates clustering with the A1/A2 sublineages. However, only some of them form a well-supported group with the A1 or A2 prototypes. In fact, ISS-UMG 1, 15, 18 and 28 form a well-supported clade (bootstrap value: 85) with the A1 prototype K02718, while ISS-UMG 23 and 33 cluster with the A2 prototype KU053903 (bootstrap value: 80). No sequences clustering with the A3 and A4 sublineages were found in ISS-UMG dataset. Four samples cluster with HPV16 members of B, C and D lineages. Specifically, ISS-UMG 21 sequence clusters with viruses belonging to the C1 sublineage (bootstrap value: 86), in agreement with BLAST results that showed it was identical to KY549209 sequence previously classified as C1 (van der Weele et al., 2017). ISS-UMG 2 sequence clusters with viruses classified in the D4 sublineage (bootstrap value: 96) and shows 99% of nt identity with the KU053934 sequence (Mirabello et al., 2018). ISS-UMG14 sequence clusters with viruses belonging to the B4 sublineage (bootstrap value: 97) and shows 99% of nt identity with KU053911 (Mirabello et al., 2018). ISS-UMG 32 sequence shows 99% nt identity with KU053910 classified in the B2 sublineage (Mirabello et al., 2018). The phylogenetic tree reconstructed using the ML method groups ISS-UMG 32 with KU053909 and KU053910 belonging to B2, but with a bootstrap value < 70. Some ISS-UMG sequences belonging to the A lineage were identical to each other: ISS-UMG 18 and 28; ISS-UMG 7 and 39; ISS-UMG 3, 10, 11, 12, 13 and 16. Out of the 39 ISS-UMG

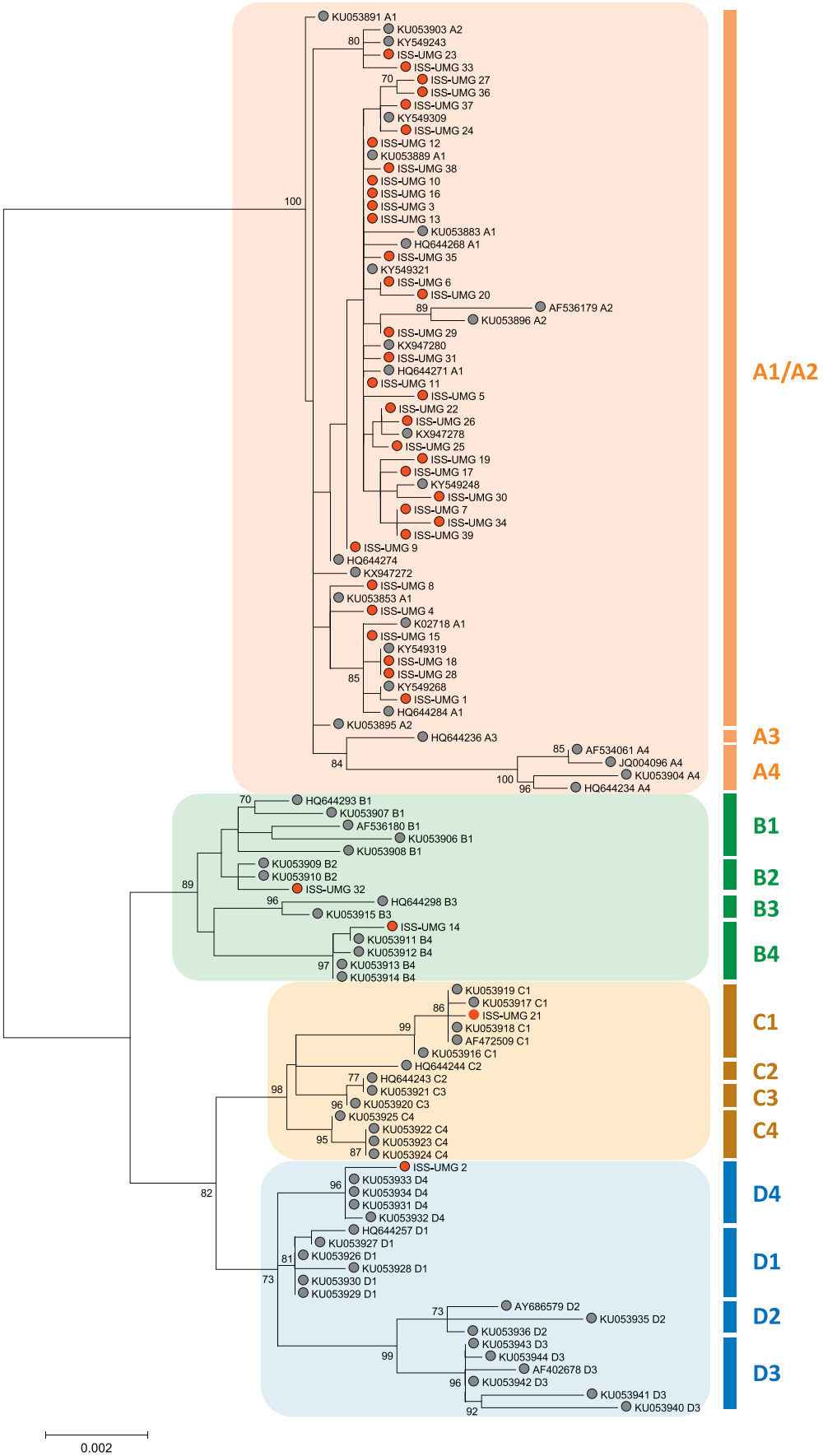

Fig. 1. Phylogenetic analysis of HPV16 sequences (n = 39) from women of Calabria region. The tree was built using the Maximum Likelihood method based on T92 + G model in MEGA 7 software. The sequence dataset used for the analysis included: 1) 39 ISS-UMG sequences; 2) 61 HPV16 reference sequences for lineages and sublineages classification (Mirabello et al., 2018; Burk et al., 2013); 3) 15 best related HPV16 genomes with the ISS-UMG sequences found by BLAST search. In the tree, the ISS-UMG sequences are indicated by a red circle; reference sequences by a grey circle, accession number, lineage or sublineage; the best related genomes by a grey circle and accession number. The accession numbers of the reference sequences are reported in Material and Method. The ISS-UMG 1-39 accession numbers in GenBank are from MH937376 to MH937414.

Table 1
The SNPs detected in the L1 gene of ISS-UMG samples are shown. The rows report from the top: 1) nucleotides of the K02718 HPV16 reference genome; 2) nucleotide position in the K02718 sequence; 3) nucleotide mutations found in ISS-UMG samples; 4) number of ISS-UMG samples carrying the above mutations; 5) amino acid changes and positions; 6) secondary structure of the L1 protein with amino acid changes (Bishop et al., 2007); 7) function of the amino acids changed: T and B epitopes for humoral and cell-mediated immune responses (Gurgel et al., 2015); AS = Assembly indicates the amino acid involved in the L1-assembly for virus-like-particle formation (Pillai et al., 2009).

L 1																													
KO2718	T	A	C	T	T	T	A	C	A	C	T	A	A	A	A	T	C	T	T	A	G	C	C	A	C	G	A	G	A
POSITION	5657	5696	5862	5909	5960	5999	6059	6163	6178	6240	6245	6314	6389	6423	6432	6480	6557	6566	6609	6693	6719	6852	6863	6962	6968	6992	7049	7058	7070
NT CHANGE	C	G	T	C	C	G	G	A	C	G	C	G	G	C	G	C	T	A	C	C	A	T	T	G	T	A	G	T	C
N (%) ISS-UMG SAMPLES	2 (5.1)	4 (10.2)	4 (10.2)	4 (10.2)	4 (10.2)	1 (2.5)	2 (5.1)	4 (10.2)	1 (2.5)	39 (100)	4 (10.2)	4 (10.2)	2 (5.1)	2 (5.1)	32 (82)	1 (2.5)	4 10.2)	1 (2.5)	1 (2.5)	2 (2.5)	3 (7.7)	4 (10.2)	2 (5.1)	1 (2.5)	4 (10.2)	3 (7.7)	2 (5.1)	4 (10.2)	1 (2.5)
AA CHANGE	—	—	H76Y	—	—	—	—	T176N	N181T	H202D	—	—	—	—	T266A	S282P	—	—	W325R	T353P	—	—	—	—	—	—	—	L474F	—
SECONDARY STRUCTURE			B-C					EF-LOOP	EF-LOOP	EF-LOOP					FG-LOOP	FG-LOOP			β-G2	HI-LOOP								α-5	
FUNCTION			T/B					T/B AS	B	AS					B AS					B									

SNPs typical of BCD
not-lineage dependent SNPs
new SNPs

Table 2
The SNPs detected in the Long Control Region of ISS-UMG samples are shown. The rows report from the top: 1) nucleotides of the K02718 HPV16 reference genome; 2) nucleotide position in the K02718 sequence; 3) nucleotide mutations in ISS-UMG samples; 4) number of the samples carrying the mutation; 5) percentage of the samples with the same mutation; 6) transcription factors binding to that position (Watts et al., 2014; Gurgel et al., 2015). The bottom row indicates the known functions of the region: the Transcription Terminator signal of the late transcripts, the Enhancer and silencer elements and the E6 promoter.

LCR

KO2718	G	A	G	A	A	A	A	A	A	G	G	G	C	T	G	G	T	T	A	A	G	T	A	G	A	C	C	G	C	T	A	C	C	C	G	G	A	A	G	C	C	C	C	C	C
POSITION	7173	7175	7193	7227	7232	7233	7303	7316	7348	7359	7360	7387	7394	7401	7429	7435	7441	7450	7458	7485	7489	7496	7507	7521	7576	7595	7669	7677	7689	7714	7729	7764	7784	7786	7826	7834	7837	7839	7869	7886	13	24	25	31	43
NT CHANGE	A/C	C	T	C	G/C	C	C	C	C	A	C/A	C	T	G	A	A	G	C	T	C	A	C	C	A	C	T	T	A	A	G	C	T	T	T	A	T	C	G	A	A	G	G	T	T	T
ISS-UMG SAMPLES	3	1	33	1	2	2	1	2	1	1	2	1	1	1	1	1	1	6	1	2	4	2	1	34	1	1	2	1	3	1	1	4	1	4	1	4	1	1	3	2	1	2	1	2	1
%	7.7	2.5	84.6	2.5	5.1	5.1	2.5	5.1	2.5	2.5	5.1	2.5	2.5	2.5	2.5	2.5	2.5	15.4	2.5	5.1	10.2	5.1	2.5	87.2	2.5	2.5	5.1	2.5	7.7	2.5	2.5	10.2	2.5	10.2	2.5	10.2	2.5	2.5	78.7	5.1	2.5	5.1	2.5	5.1	2.5
FUNCTION		C/EBP	TEF-1										ETS-1			GRE/I	YY-1			GRE,YY-1	GRE,YY-1			YY-1		NF1	OCT-1*	OCT-1	TEF-1	NF-1	NF-1, YY-1	YY-1		YY-1	YY-1 TEF-1	NF-1*		OCT-1*	OCT-1*	E2			Tatabox		AP-1	E2

5' LCR
TRANSCRIPTION TERMINATOR SIGNAL OF THE LATE TRANSCRIPTS

CENTRAL LCR
ENHANCER

3' LCR
REPLICATION ORIGIN AND E6 PROMOTER

Yellow: Diagnostic SNPs of BCD lineages
Orange: not-lineage dependent SNPs
Green: new SNPs
White: SNPs already reported

sequences, 28 showed at least one different SNP.

3.2. Genetic variability and SNPs characteristic of samples from Calabria

DNA sequence analysis of the 39 ISS-UMG samples compared to the K02718 reference sequence revealed a total of 96 SNPs, 29 of which mapping in the L1 (Table 1), 45 in the LCR (Table 2), 15 in the E6 and 7 in the E7 (Table 3). In all the Tables, the typical SNPs mutated in the BCD variants are shown in the yellow columns, those not-lineage dependent in the orange columns, the SNPs already detected in other studies in the white columns and the SNPs detected for the first time in this study are in the green columns.

Table 1 reports the 29 SNPs found in the L1 ISS-UMG genes. All the sequences had an ATC insertion (aa S) at nt 6901, while the GAT triplet (aa D) was deleted at nt 6950. All samples had a C6240G SNP causing the H202D amino acid change. At the A6432 polymorphic site, 32 samples had G resulting in aa T instead of A. In the ISS-UMG 5, the T6609C SNP causing the W325R amino acid change was found in this study for the first time. Some SNPs typical of BCD lineages, causing six amino acid changes in L1 protein, were found (Table 1, row AA-Change). Premature stop codons were not detected in any samples. The last two rows of Table 1 report the known L1 functions: the secondary structure (Bishop et al., 2007), the B and T epitopes critical for the immune response (Gurgel et al., 2015) and the sites for capsid assembly (As) (Pillai et al., 2009).

Table 2 reports the 45 SNPs found in the LCR ISS-UMG samples compared to the K02718 reference sequence. The specific combination of SNPs in the LCR determines the HPV16 clustering in lineages and sublineages (Cornet et al., 2012). The most commonly detected variations were the C7433 insertion and the A7863 deletion that were found in 100% of the samples. Moreover, the G7193 T variation was found in 33/39 (84.6%) samples, followed by G7521A in 34/39 (87.2%) and T7450C in 6/39 (15.4%) cases. Twenty-five nt changes were detected at the putative binding sites of the TEF-1, GRE-YY1, OCT1 and E2 transcription factors (Table 2, row Function). The last row indicates the functions corresponding to known regulatory elements of the URR map.

Regarding the E6-E7 genes, a total of 22 SNPs, 15 in E6 and 7 in E7, were detected in the HPV16 Calabrian dataset compared to the K02718 reference (Table 3). Eight SNPs in E6 and 3 in E7 cause an amino acid change. The most studied E6 SNPs for virus pathogenesis, namely T350G causing the L83 V and A131G causing the R10G amino acid change, were detected in 29 (74.4%) and 3 (7.7%) ISS-UMG samples, respectively. In E7, the A647G and C757A SNPs, causing the N29S and R66W amino acid change respectively, have already been described; conversely, the C794T SNP causing the T78I amino acid change was found for the first time in this study, to the best of our knowledge.

The ISS-UMG sequences were compared to the partial HPV16 sequences of the Italian isolates deposited in NCBI GenBank bank. These sequences were from samples collected in Northern Italy (Milan; Tanzi et al., 2009), Central Italy (Rome; Garbuglia et al., 2007; Cento et al., 2009) and in Sardinia (Cagliari; Montaldo et al., 2007). In BLAST analysis, the partial EU650438 URR sequence (Tanzi et al., 2009) showed 99% of identity with the ISS-UMG 7, 17, 19, 34 sequences of the A lineage; the partial EU650455–451-454 URRs (Tanzi et al., 2009) showed 100% of identity with ISS-UMG18 and ISS-UMG 28. The T795G and T789C E7 SNPs, present in the ISS-UMG 2, 14, 21, 32 E7-sequences, were also found in the EF422131 (Montaldo et al., 2007; Garbuglia et al., 2007) and FJ644988 sequences (Cento et al., 2009). ISS-UMG 2, 14, 21, 32 share several SNPs with EU650474 and EU650448 URR, sequences previously classified as African-2 and now recognized as belonging to the C1 sublineage.

4. Discussion

It has been well established that HPV16 is the most prevalent genotype detected in cervical cancer,worldwide. However, the prevalence of the A, B, C and D HPV16 variants is variable in the different continents (Pimenoff et al., 2016). Population studies demonstrated the

Table 3
The SNPs detected in E6-E7 genes of ISS-UMG samples are shown. The rows report from the top: 1) nucleotides of the K02718 HPV16 reference genome; 2) nucleotide position in the K02718 sequence; 3) nucleotide mutations in ISS-UMG samples; 4) number of samples carrying the mutations above; 5) position of the amino acid changes in the E6 or E7 proteins; 6) known functions of the amino acid: B and T epitopes involved in the immune response, p53 binding, NLS = Nuclear Localization Signal and NES = Nuclear Export Signal function, formation of Zinc-finger domain (Pillai et al., 2009).

	E 6														E 7							
KO2718	A	A	T	A	G	C	G	A	C	T	A	T	C	T	A	A	C	T	C	T	A	T
POSITION	93	103	109	131	132	*143*	*145*	189	256	*286*	*289*	295	*335*	*350*	403	*647*	757	789	794	*795*	822	846
NT CHANGE	T	G	C	G	T	G	T	G	T	A	G	G	T	G	G	G	A	C	T	G	G	A
N (%) ISS-UMG SAMPLES	1 (2.6)	1 (2.6)	4 (10.2)	3 (7.7)	1 (2.6)	3 (7.7)	4 (10.2)	1 (2.6)	1 (2.6)	4 (10.2)	4 (10.2)	2 (5.1)	5 (12.8)	29 (74.3)	1 (2.6)	1 (2.6)	1 (2.6)	4 (10.2)	1 (2.6)	4 (10.2)	2 (5.1)	1 (2.6)
AA CHANGE	—	—	—	R10G	R10I	Q14D	Q14H	E29G	—	—	—	D64E	H78Y	L83V	—	N29S	R66W	—	T78I	—	—	—
FUNCTION				T/B p53			T/B p53	T					T p53	T p53	T/B	pRB NLS			NES Zn finger domain			

Yellow: Diagnostic SNPs of BCD lineages
Orange: non-lineage-specific-SNPs
Green: new SNPs
White: SNPs already reported

association between specific sublineages and an increased risk to develop: 1) high-grade cervical lesions (C, D2 and D3); 2) squamous cell carcinoma (A1 and A2) or adenocarcinoma (A4) of the cervix. In addition, also the genetic background of populations seems to be involved in the harmfulness of some variants (Mirabello et al., 2016; Mirabello et al., 2018).

In Italy, HPV16 is the prevalent genotype in cervical infections (Giambi et al., 2013) and this is true also for the Calabria region (Galati et al., 2017). Studies on the HPV prevalence in this Region exist, but data on the circulating HPV16 variants are still lacking. Indeed, little is known about the different oncogenic potential of HPV16 variants in Italian women. In Italy, HPV DNA-test will become the primary test of cervical cancer screening for women over 30 by the end of 2018. Since HPV16 mutants could directly affect the proficiency of the genetic diagnostic kits, the monitoring of HPV variants is becoming a priority.

In the present study, to possibly shed light on the HPV16 genomes currently circulating in Calabria, 39 cervical DNA samples HPV16-positive were sequenced in the 3169 nt viral region covering the L1, LCR, E6 and E7. The obtained sequences were compared to those belonging to previously identified HPV16 variant prototypes (Mirabello et al., 2018). The sequence and phylogenetic analysis showed that the ISS-UMG samples have 99–100% of nt identity with HPV16 variants belonging to the A, B, C, and D lineages. Most of the HPV16 isolates belong to the A lineages (89.7%), in agreement with previous data on the Italian population (Tornesello et al., 2004; Cento et al., 2009). Four HPV16 variants belonging to lineages B (ISS-IMG32, B2, ASC-US; ISS-UMG14, B4, H-SIL), C (ISS-UMG21, C1, ASC-US), and D (ISS-UMG2, D4, H-SIL) were also detected. The C lineage was previously identified in Italy in HIV seropositive subjects and in high-risk immigrant women from Africa, but not in the Italian general population (Tornesello et al., 2008a, 2008b; Tornesello et al., 2007; Tanzi et al., 2009). Moreover, B and D lineages were reported in Italian women with CIN lesions and cancer, as well as in patients with penile carcinoma (Tornesello et al., 2004; Tornesello et al., 2011; Tornesello et al., 2008b). The two ISS-UMG14 and ISS-UMG2 sequences show 99% of nt identity with members of B4 and D4, respectively, suggesting the presence in Italy of these recently discovered HPV16 sublineages (Mirabello et al., 2016; Mirabello et al., 2018), that were unknown at the time of the previous Italian studies.

A total of 96 SNPs, of which 29 mapping in L1, 45 in the LCR, 15 in E6 and 7 in E7, were found in the ISS-UMG sequence dataset. The most common SNP was the E6 T350G leading to the L83 V amino acid change (29/39 samples; 74.4%), which was found in all the spectrum of cytological pap-test results. This E6 mutation has been studied for a long time, since the late 90s. A recent *in vitro* study demonstrated that the E6 L83 V, in the absence of E7, can immortalize but not transform keratinocytes and has the unique ability to down-regulate E-cadherin, protein involved in cell-cell adhesion of epithelial cells and in their transition to invasive cells (Togtema et al., 2015). Over the years, epidemiological studies have given controversial results about the association of this mutation with invasive CC and/or cervical lesion progression to cancer. Such studies have clearly demonstrated that populations of different geographic origin have a different prevalence of the L83 V polymorphism (Zehbe et al., 1998a; Zehbe et al., 1998b; Villa et al., 2000; Nindl et al., 1999; Zehbe et al., 2001a, 2001b; Tornesello et al., 2004; Grodzki, 2006). This diversity suggests that the risk conferred by the E6 L83 V mutation could be population-dependent (Cornet et al., 2013; Nicolás-Párraga et al., 2016). In our study, the prevalence of the T350G polymorphism is high (74.4%) for a European population, where it usually ranges from 47% to 59% (Nicolás-Párraga et al., 2016). The high prevalence found is similar to that reported for a Greek population and might be representative of a higher prevalence of this mutation in the southern European population (Tsakogiannis et al., 2018). However, we cannot exclude a confounding effect due to the sample size.

Regarding the LCR sequence characterization, G7521A was the most frequent SNP, observed in 85.7% of the cases. As previously reported (Kurvinen et al., 2000), this SNP maps in the transcription factor YY1 binding sites and it is the most frequently detected worldwide, especially in the European population (Ramas et al., 2018).

This is the first study reporting the circulation of HPV16 variants in the Calabria region. Because of its geographical position in the

Mediterranean area, Calabria is a crossroad for European, Asian and African populations and can represent a privileged observation point to study the geographical HPV16 variants circulation. Hence, larger size studies on the circulation of HPV variants in the Calabria region would be useful for a better characterization of the sub-lineages circulating in this region.

5. Conclusion

HPV16 variants belonging to the four currently recognized A, B, C, D lineages were found to circulate in Calabria, Italy. Although most of the analyzed sequences were 99–100% identical to sequences previously described, 8 new SNPs, causing 2 amino acid substitutions, were reported for the first time in this study, to the best of our knowledge. Three SNPs were found in L1: T6609C, causing the W325R amino acid change located in a β-sheet region of the protein, and A6059G and A7070G, causing no amino acid change. In the LCR, three new SNPs were identified: T7401G in the Transcription Terminator signal of the late transcript, C7595T in the Enhancer region and C13G in the E6 early promoter. Two new SNPs were found in the E6-E7 region: A93T and C794T, the latter causing the Y/I substitution in the nuclear export signal of the E7 zinc-finger domain. The study provides important information for guiding future investigations in the context of HPV16 variant diversity and cervical cancer prevention in Calabria.

Conflicts of interest

All the authors declare that they have no conflict of interest.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2018.12.024.

Acknowledgment

LG carried out this work at the Istituto Superiore di Sanità in Rome during her PhD course “Life Science”, University of “Magna Graecia”, Catanzaro, Italy. The authors did not receive any specific grant from funding agencies in the public, commercial or not-for profit sectors to perform this study.

References

Bishop, B., Dasgupta, J., Klein, M., Garcea, R.L., Christensen, N.D., Zhao, R., Chen, X.S., 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. https://doi.org/10.1074/jbc.M706380200.

Bouvard, V., Baan, R., Straif, K., Grosse, Y., Secretan, B., El Ghissassi, F., Benbrahim-Tallaa, L., Guha, N., Freeman, C., Galichet, L., Cogliano, V., 2009. A review of human carcinogens—Part B: biological agents. Lancet Oncol. 10, 321–322. https://doi.org/10.1016/S1470-2045(09)70096-8.

Bravo, I.G., Felez-Sanchez, M., 2015. Papillomaviruses: viral evolution, cancer and evolutionary medicine. Evol. Med. Public Heal. 2015, 32–51. https://doi.org/10.1093/emph/eov003.

Burk, R.D., Harari, A., Chen, Z., 2013. Human papillomavirus genome variants. Virology 445, 232–243. https://doi.org/10.1016/j.virol.2013.07.018.

Burk, R.D., Chen, Z., Saller, C., Tarvin, K., Carvalho, A.L., Scapulatempo-Neto, C., Silveira, H.C., Fregnani, J.H., Creighton, C.J., Anderson, M.L., Castro, P., Wang, S.S., Yau, C., Benz, C., Gordon Robertson, A., Mungall, K., Lim, L., Bowlby, R., Sadeghi, S., Brooks, D., Sipahimalani, P., Mar, R., Ally, A., Clarke, A., Mungall, A.J., Tam, A., Lee, D., Chuah, E., Schein, J.E., Tse, K., Kasaian, K., Ma, Y., Marra, M.A., Mayo, M., Balasundaram, M., Thiessen, N., Dhalla, N., Carlsen, R., Moore, R.A., Holt, R.A., Jones, S.J.M., Wong, T., Pantazi, A., Parfenov, M., Kucherlapati, R., Hadjipanayis, A., Seidman, J., Kucherlapati, M., Ren, X., Xu, A.W., Yang, L., Park, P.J., Lee, S., Rabeno, B., Huelsenbeck-Dill, L., Borowsky, M., Cadungog, M., Iacocca, M., Petrelli, N., Swanson, P., Ojesina, A.I., Ojesina, A.I., Ojesina, A.I., Le, X., Sandusky, G., Adebamowo, S.N., Akeredolu, T., Adebamowo, C., Reynolds, S.M., Shmulevich, I., Shelton, C., Crain, D., Mallery, D., Curley, E., Gardner, J., Penny, R., Morris, S., Shelton, T., Liu, J., Lolla, L., Chudamani, S., Wu, Y., Birrer, M., McLellan, M.D., Bailey, M.H., Miller, C.A., Wyczalkowski, M.A., Fulton, R.S., Fronick, C.C., Lu, C., Mardis, E.R., Appelbaum, E.L., Schmidt, H.K., Fulton, L.A., Cordes, M.G., Li, T., Ding, L., Wilson, R.K., Rader, J.S., Behmaram, B., Uyar, D., Bradley, W., Wrangle, J., Pastore, A., Levine, D.A., Dao, F., Gao, J., Schultz, N., Sander, C., Ladanyi, M., Einstein, M., Teeter, R., Benz, S., Wentzensen, N., Felau, I., Zenklusen, J.C., Bodelon, C., Demchok, J.A., Yang, L., Sheth, M., Ferguson, M.L., Tarnuzzer, R., Yang, H., Schiffman, M., Zhang, J., Wang, Z., Davidsen, T., Olaniyan, O., Hutter, C.M., Sofia, H.J., Gordenin, D.A., Chan, K., Roberts, S.A., Klimczak, L.J., Van Waes, C., Chen, Z., Saleh, A.D., Cheng, H., Parfitt, J., Bartlett, J., Albert, M., Arnaout, A., Sekhon, H., Gilbert, S., Peto, M., Myers, J., Harr, J., Eckman, J., Bergsten, J., Tucker, K., Anne Zach, L., Karlan, B.Y., Lester, J., Orsulic, S., Sun, Q., Naresh, R., Pihl, T., Wan, Y., Zaren, H., Sapp, J., Miller, J., Drwiega, P., Murray, B.A., Zhang, H., Cherniack, A.D., Sougnez, C., Sekhar Pedamallu, C., Lichtenstein, L., Meyerson, M., Noble, M.S., Heiman, D.I., Voet, D., Getz, G., Saksena, G., Kim, J., Shih, J., Cho, J., Lawrence, M.S., Gehlenborg, N., Lin, P., Beroukhim, R., Frazer, S., Gabriel, S.B., Schumacher, S.E., Leraas, K.M., Lichtenberg, T.M., Zmuda, E., Bowen, J., Frick, J., Gastier-Foster, J.M., Wise, L., Gerken, M., Ramirez, N.C., Danilova, L., Cope, L., Baylin, S.B., Salvesen, H.B., Vellano, C.P., Ju, Z., Diao, L., Zhao, H., Chong, Z., Ryan, M.C., Martinez-Ledesma, E., Verhaak, R.G., Averett Byers, L., Yuan, Y., Chen, K., Ling, S., Mills, G.B., Lu, Y., Akbani, R., Seth, S., Liang, H., Wang, J., Han, L., Weinstein, J.N., Bristow, C.A., Zhang, W., Mahadeshwar, H.S., Sun, H., Tang, J., Zhang, J., Song, X., Protopopov, A., Mills Shaw, K.R., Chin, L., Olabode, O., Disaia, P., Radenbaugh, A., Haussler, D., Zhu, J., Stuart, J., Chalise, P., Koestler, D., Fridley, B.L., Godwin, A.K., Madan, R., Ciriello, G., Martinez, C., Higgins, K., Bocklage, T., Todd Auman, J., Perou, C.M., Tan, D., Parker, J.S., Hoadley, K.A., Wilkerson, M.D., Mieczkowski, P.A., Skelly, T., Veluvolu, U., Neil Hayes, D., Kimryn Rathmell, W., Hoyle, A.P., Simons, J.V., Wu, J., Mose, L.E., Soloway, M.G., Balu, S., Meng, S., Jefferys, S.R., Bodenheimer, T., Shi, Y., Roach, J., Thorne, L.B., Boice, L., Huang, M., Jones, C.D., Zuna, R., Walker, J., Gunderson, C., Snowbarger, C., Brown, D., Moxley, K., Moore, K., Andrade, K., Landrum, L., Mannel, R., McMeekin, S., Johnson, S., Nelson, T., Elishaev, E., Dhir, R., Edwards, R., Bhargava, R., Tiezzi, D.G., Andrade, J.M., Noushmehr, H., Gilberto Carlotti, C., da Cunha Tirapelli, D.P., Weisenberger, D.J., Van Den Berg, D.J., Maglinte, D.T., Bootwalla, M.S., Lai, P.H., Triche, T., Swisher, E.M., Agnew, K.J., Simon Shelley, C., Laird, P.W., Schwarz, J., Grigsby, P., Mutch, D., 2017. Integrated genomic and molecular characterization of cervical cancer. Nature 543, 378–384. https://doi.org/10.1038/nature21386.

Cento, V., Ciccozzi, M., Ronga, L., Perno, C.F., 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, 1627–1634. https://doi.org/10.1002/jmv.21552.

Chen, Z., de Freitas, L.B., Burk, R.D., 2015. Evolution and classification of oncogenic human papillomavirus types and variants associated with cervical cancer. Methods Mol. Biol. 1249, 3–26. https://doi.org/10.1007/978-1-4939-2013-6_1.

Cornet, I., Gheit, T., Franceschi, S., Vignat, J., Burk, R.D., Sylla, B.S., Tommasino, M., Clifford, G.M., 2012. Human Papillomavirus type 16 genetic variants: phylogeny and classification based on E6 and LCR. J. Virol. 86, 6855–6861. https://doi.org/10.1128/JVI.00483-12.

Cornet, I., Gheit, T., Clifford, G.M., Combes, J.-D., Dalstein, V., Franceschi, S., Tommasino, M., Clavel, C., 2013. Human papillomavirus type 16 E6 variants in France and risk of viral persistence. Infect. Agent. Cancer 8, 4. https://doi.org/10.1186/1750-9378-8-4.

De Martel, C., Plummer, M., Vignat, J., Franceschi, S., 2017. Worldwide Burden of Cancer Attributable to HPV by Site, Country and HPV Type 33. pp. 664–670. https://doi.org/10.1002/ijc.30716.

de Sanjose, S., Quint, W.G.V., Alemany, L., Geraets, D.T., Klaustermeier, J.E., Lloveras, B., Tous, S., Felix, A., Bravo, L.E., Shin, H.R., Vallejos, C.S., de Ruiz, P.A., Lima, M.A., Guimera, N., Clavero, O., Alejo, M., Llombart-Bosch, A., Cheng-Yang, C., Tatti, S.A., Kasamatsu, E., Iljazovic, E., Odida, M., Prado, R., Seoud, M., Grce, M., Usubutun, A., Jain, A., Suarez, G.A.H., Lombardi, L.E., Banjo, A., Menéndez, C., Domingo, E.J., Velasco, J., Nessa, A., Chichareon, S.C.B., Qiao, Y.L., Lerma, E., Garland, S.M., Sasagawa, T., Ferrera, A., Hammouda, D., Mariani, L., Pelayo, A., Steiner, I., Oliva, E., Meijer, C.J.L.M., Al-Jassar, W.F., Cruz, E., Wright, T.C., Puras, A., Llave, C.L., Tzardi, M., Agorastos, T., Garcia-Barriola, V., Clavel, C., Ordi, J., Andújar, M., Castellsagué, X., Sánchez, G.I., Nowakowski, A.M., Bornstein, J., Muñoz, N., Bosch, F.X., 2010. Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. Lancet Oncol. 11, 1048–1056. https://doi.org/10.1016/S1470-2045(10)70230-8.

Egawa, N., Egawa, K., Griffin, H., Doorbar, J., 2015. Human papillomaviruses; Epithelial tropisms, and the development of neoplasia. Viruses 7, 3863–3890. https://doi.org/10.3390/v7072802.

Freitas, L.B., Chen, Z., Muqui, E.F., Boldrini, N.A.T., Miranda, A.E., Spano, L.C., Burk, R.D., 2014. Human papillomavirus 16 non-european variants are preferentially associated with high-grade cervical lesions. PLoS One 9, 1–7. https://doi.org/10.1371/journal.pone.0100746.

Galati, L., Peronace, C., Fiorillo, M.T., Masciari, R., Giraldi, C., Nisticò, S., Minchella, P., Maiolo, V., Barreca, G.S., Marascio, N., Lamberti, A.G., Giancotti, A., Lepore, M.G., Greco, F., Mauro, M.V., Borelli, A., Bocchiaro, G. Lo, Surace, G., Liberto, M.C., Focà, A., 2017. Six years genotype distribution of Human Papillomavirus in Calabria Region, Southern Italy: a retrospective study. Infect. Agent. Cancer 12. https://doi.org/10.1186/s13027-017-0154-5.

Garbuglia, A.R., Carletti, F., Minosse, C., Piselli, P., Zaniratti, M.S., Serraino, D., Capobianchi, M.R., 2007. Genetic variability in E6 and E7 genes of human papillomavirus −16, −18, −31 and −33 from HIV-1-positive women in Italy. New Microbiol. 30, 377–382.

Giambi, C., Donati, S., Carozzi, F., Salmaso, S., Declich, S., Atti, M.L.C. degli, Ronco, G., Alibrandi, M.P., Brezzi, S., Collina, N., Franchi, D., Lattanzi, A., Minna, M.C., Nannini, R., Barretta, E., Burroni, E., Gillio-Tos, A., Macallini, V., Pierotti, P., Bella, A., 2013. A cross-sectional study to estimate high-risk human papillomavirus prevalence and type distribution in Italian women aged 18-26 years. BMC Infect. Dis. 13, 1. https://doi.org/10.1186/1471-2334-13-74.

Grodzki, M., 2006. Increased risk for cervical disease Progression of French women infected with the human Papillomavirus type 16 E6-350G variant. Cancer Epidemiol.

Biomark. Prev. 15, 820–822. https://doi.org/10.1158/1055-9965.EPI-05-0864.
Gurgel, A.P.A.D., Chagas, B.S., Do Amaral, C.M., Nascimento, K.C.G., Leal, L.R.S., Silva Neto, J.D.C., Cartaxo Muniz, M.T., De Freitas, A.C., 2015. Prevalence of human papillomavirus variants and genetic diversity in the L1 gene and long control region of HPV16, HPV31, and HPV58 found in north-East Brazil. Biomed. Res. Int. 2015. https://doi.org/10.1155/2015/130828.
Hirose, Y., Onuki, M., Tenjimbayashi, Y., Mori, S., Ishii, Y., Takeuchi, T., Tasaka, N., Satoh, T., Morisada, T., Iwata, T., Miyamoto, S., Matsumoto, K., Sekizawa, A., Kukimoto, I., 2018. Within-host variations of human Papillomavirus reveal APOBEC-signature mutagenesis in the viral genome. J. Virol. JVI. https://doi.org/10.1128/JVI.00017-18. 00017–18.
Kurvinen, K., Yliskoski, M., Saarikoski, S., Syrjänen, K., Syrjänen, S., 2000. Variants of the long control region of human papillomavirus type 16. Eur. J. Cancer 36, 1402–1410.
Makowsky, R., Lhaki, P., Wiener, H.W., Bhatta, M.P., Cullen, M., Johnson, D.C., Perry, R.T., Lama, M., Boland, J.F., Yeager, M., Ghimire, S., Broker, T.R., Shrestha, S., 2016. Genomic diversity and phylogenetic relationships of human papillomavirus 16 (HPV16) in Nepal. Infect. Genet. Evol. 46, 7–11. https://doi.org/10.1016/j.meegid.2016.10.004.
Mirabello, L., Yeager, M., Cullen, M., Boland, J.F., Chen, Z., Wentzensen, N., Zhang, X., Yu, K., Yang, Q., Mitchell, J., Roberson, D., Bass, S., Xiao, Y., Burdett, L., Raine-Bennett, T., Lorey, T., Castle, P.E., Burk, R.D., Schiffman, M., 2016. HPV16 sublineage associations with histology-specific cancer risk using HPV whole-genome sequences in 3200 women. J. Natl. Cancer Inst. https://doi.org/10.1093/jnci/djw100.
Mirabello, L., Clarke, M., Nelson, C., Dean, M., Wentzensen, N., Yeager, M., Cullen, M., Boland, J., Schiffman, M., Burk, R., 2018. The intersection of HPV epidemiology, genomics and mechanistic studies of HPV-mediated carcinogenesis. Viruses 10, 80. https://doi.org/10.3390/v10020080.
Montaldo, C., Mastinu, A., Quartuccio, M., Piras, V., Denotti, G., Pisano, E., Orrù, G., 2007. Detection and genotyping of human papillomavirus DNA in samples from healthy Sardinian patients: a preliminary study. J. oral Pathol. Med. 36, 482–487. https://doi.org/10.1111/j.1600-0714.2007.00556.x.
Nicolás-Párraga, S., Gandini, C., Pimenoff, V.N., Alemany, L., de Sanjosé, S., Xavier Bosch, F., Bravo, I.G., 2016. HPV16 variants distribution in invasive cancers of the cervix, vulva, vagina, penis, and anus. Cancer Med. 5, 2909–2919. https://doi.org/10.1002/cam4.870.
Nicolás-Párraga, S., Alemany, L., de Sanjosé, S., Bosch, F.X., Bravo, I.G., 2017. Differential HPV16 variant distribution in squamous cell carcinoma, adenocarcinoma and adenosquamous cell carcinoma. Int. J. Cancer 140, 2092–2100. https://doi.org/10.1002/ijc.30636.
Nindl, I., Rindfleisch, K., Lotz, B., Schneider, A., Dürst, 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 (https://doi.org/10.1002/(SICI)1097-0215(19990719)82:2 < 203::AID-IJC9 > 3.0.CO;2-9).
Pillai, M.R., Hariharan, R., Babu, J.M., Lakshmi, S., Chiplunkar, S.V., Patkar, M., Tongaonkar, H., Dinshaw, K., Jayshree, R.S., Reddy, B.K., Siddiqui, M., Roychoudury, S., Saha, B., Abraham, P., Gnanamony, M., Peedicayil, A., Subhashini, J., Ram, T.S., Dey, B., Sharma, C., Jain, S.K., Singh, N., 2009. Molecular variants of HPV-16 associated with cervical cancer in Indian population. Int. J. Cancer 125, 91–103. https://doi.org/10.1002/ijc.24322.
Pimenoff, V.N., De Oliveira, C.M., Bravo, I.G., 2016. Transmission between Archaic and Modern Human Ancestors during the Evolution of the Oncogenic Human Papillomavirus 16 Article Fast Track. vol. 34. pp. 4–19. https://doi.org/10.1093/molbev/msw214.
Ramas, V., Mirazo, S., Bonilla, S., Ruchansky, D., Arbiza, J., 2018. Analysis of human papillomavirus 16 E6, E7 genes and Long Control Region in cervical samples from Uruguayan women. Gene 654, 103–109. https://doi.org/10.1016/j.gene.2018.02.023.
Reusser, N., Downing, C., Guidry, J., Tyring, S., 2015. HPV carcinomas in immunocompromised patients. J. Clin. Med. 4, 260–281. https://doi.org/10.3390/jcm4020260.
Rossi, S., Crocetti, E., Capocaccia, R., Gatta, G., 2013. Estimates of cancer burden in Italy. Tumori 99, 416–424. https://doi.org/10.1700/1334.14803.
Schiffman, M., Doorbar, J., Wentzensen, N., De Sanjosé, S., Fakhry, C., Monk, B.J., Stanley, M.A., Franceschi, S., 2016. Carcinogenic human papillomavirus infection. Nat. Rev. Dis. Prim. 2. https://doi.org/10.1038/nrdp.2016.86.
Smith, B., Chen, Z., Reimers, L., Van Doorslaer, K., Schiffman, M., Desalle, R., Herrero, R., Yu, K., Wacholder, S., Wang, T., Burk, R.D., 2011. Sequence Imputation of HPV16 Genomes for Genetic Association Studies. vol. 6https://doi.org/10.1371/journal.pone.0021375.
Tanzi, E., Amendola, A., Bianchi, S., Fasolo, M.M., Beretta, R., Pariani, E., Zappa, A., Frati, E., Orlando, G., 2009. Human papillomavirus genotypes and phylogenetic analysis of HPV-16 variants in HIV-1 infected subjects in Italy. Vaccine. https://doi.org/10.1016/j.vaccine.2008.10.091.
Togtema, M., Jackson, R., Richard, C., Niccoli, S., Zehbe, I., 2015. The human papillomavirus 16 European-T350G E6 variant can immortalize but not transform keratinocytes in the absence of E7. Virology 485, 274–282. https://doi.org/10.1016/j.virol.2015.07.025.
Tornesello, M.L., Duraturo, M.L., Salatiello, I., Buonaguro, L., Losito, S., Botti, G., Stellato, G., Greggi, S., Piccoli, R., Pilotti, S., Stefanon, B., De Palo, G., Franceschi, S., Buonaguro, F.M., 2004. Analysis of human papillomavirus type-16 variants in Italian women with cervical intraepithelial neoplasia and cervical cancer. J. Med. Virol. 74, 117–126. https://doi.org/10.1002/jmv.20154.
Tornesello, M.L., Duraturo, M.L., Buonaguro, L., Vallefuoco, G., Piccoli, R., Palmieri, S., Buonaguro, F.M., 2007. Prevalence of human papillomavirus genotypes and their variants in high risk West Africa women immigrants in South Italy. Infect. Agent. Cancer 2, 1–9. https://doi.org/10.1186/1750-9378-2-1.
Tornesello, M.L., Duraturo, M.L., Giorgi-Rossi, P., Sansone, M., Piccoli, R., Buonaguro, L., Buonaguro, F.M., 2008a. Human papillomavirus (HPV) genotypes and HPV16 variants in human immunodeficiency virus-positive Italian women. J. Gen. Virol. 89, 1380–1389. https://doi.org/10.1099/vir.0.83553-0.
Tornesello, M.L., Duraturo, M.L., Losito, S., Botti, G., Pilotti, S., Stefanon, B., De Palo, G., Gallo, A., Buonaguro, L., Buonaguro, F.M., 2008b. Human papillomavirus genotypes and HPV16 variants in penile carcinoma. Int. J. Cancer 122, 132–137. https://doi.org/10.1002/ijc.23062.
Tornesello, M.L., Losito, S., Benincasa, G., Fulciniti, F., Botti, G., Greggi, S., Buonaguro, L., Buonaguro, F.M., 2011. Human papillomavirus (HPV) genotypes and HPV16 variants and risk of adenocarcinoma and squamous cell carcinoma of the cervix. Gynecol. Oncol. 121, 32–42. https://doi.org/10.1016/j.ygyno.2010.12.005.
Tsakogiannis, D., Papadopoulou, A., Kontostathi, G., Ruether, I.G.A., Kyriakopoulou, Z., Dimitriou, T.G., Orfanoudakis, G., Markoulatos, P., 2018. Molecular and Evolutionary Analysis of HPV16 E6 and E7 Genes in Greek Women 1688–1696. https://doi.org/10.1099/jmm.0.055491-0.
van der Weele, P., Meijer, C.J.L.M., King, A.J., 2017. Whole-genome sequencing and variant analysis of human papillomavirus 16 infections. J. Virol. JVI. https://doi.org/10.1128/JVI.00844-17. 00844–17.
Villa, L.L., Caballero, O., Ferenczy, A., Sichero, L., Rohan, T., Franco, E.L., Rahal, P., 2000. Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. J. Gen. Virol. 81, 2959–2968. https://doi.org/10.1099/0022-1317-81-12-2959.
Warren, C., Westrich, J., Doorslaer, K., Pyeon, D., 2017. Roles of APOBEC3A and APOBEC3B in human Papillomavirus infection and disease progression. Viruses 9, 233. https://doi.org/10.3390/v9080233.
Watts, K.J., Thompson, C.H., Cossart, Y.E., Rose, B.R., 2014. Variable Oncogene Promoter Activity of Human Papillomavirus Type 16 Cervical Cancer Isolates from Australia 39, 2009–2014. https://doi.org/10.1128/JCM.39.5.2009.
Zehbe, I., Voglino, G., Delius, H., Wilander, E., Tommasino, M., Duffner, U., Bergstraesser, E., Sauter, S., 1998a. Risk of cervical cancer and geographical variations of human 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.
Zehbe, I., Tachezy, R., Mytilineos, J., Voglino, G., Mikyoskova, I., Delius, H., Marongiu, A., Gissman, L., Wilander, E., Tommasino, M., 2001a. Risk of cervical cancer and geographical variations of human papillomavirus 16 E6 polymorphisms. Lancet 80, 793–794. https://doi.org/10.1002/ana.24784.
Zehbe, I., Tachezy, R., Mytilineos, J., Voglino, G., Mikyskova, I., Delius, H., Marongiu, A., Gissmann, L., Wilander, E., Tommasino, M., 2001b. 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, 711–716. https://doi.org/10.1002/ijc.1520.