

Sequence variation analysis of the E1 and E2 genes of human papillomavirus type 16 in cervical lesions in women from the south of Poland

Katarzyna Sitarz^{1,2}, Jolanta Kopec¹, Barbara Zawilinska¹, Malgorzata Klimek³ and Slawa Szostek¹✉

¹Jagiellonian University Medical College, Faculty of Medicine, Chair of Microbiology, Department of Molecular Medical Microbiology, Kraków, Poland; ²Jagiellonian University, Faculty of Chemistry, Chiraloptical Spectroscopy Group, Kraków, Poland; ³National Research Institute of Oncology, Krakow Branch, Clinic of Radiotherapy, Kraków, Poland

The E1 and E2 genes of the human papillomavirus encode the so-called early proteins, their sequences are conserved, and regulatory functions are associated with the viral oncoproteins. The purpose of this study is to determine the HPV16 E1 and E2 mutations appearing in the female population of southern Poland, depending on the severity of cervical pathological changes. We also take into account the number of E1 and E2 mutations detected in the E6 gene variant (350G or 350T). This publication is one of the first in the Central and Eastern Europe to deal with this topic. We identified 4 mutations in the E1 gene and 24 mutations in the E2 gene that have not been described so far. In three cases of squamous cell carcinoma a C3409T mutation occurred, which is widely described as oncogenic. This mutation lies in the 3243-3539 area of the E2 hinge region. Statistical analyses show a possible relationship of mutations in this area with oncogenesis. The discovered dependencies may be important in the context of oncogenesis, however, a study with a larger group of patients is needed in order to confirm this view.

Key words: human papillomavirus, E1, E2, polymorphism, cervical cancer, LSIL

Received: 23 January, 2021; revised: 19 February, 2021; accepted: 06 March, 2021; available on-line: 28 April, 2021

✉ e-mail: slawa.szostek@uj.edu.pl

Acknowledgements of Financial Support: Katarzyna Sitarz acknowledges fellowship with the project no. POWR.03.02.00-00-I013/16. Katarzyna Sitarz acknowledges support of the Inter-DokMed project no. POWR.03.02.00-00-I013/16. Research partially financed with subsidies for the maintenance and development of the research potential of the Faculty of Medicine of the Jagiellonian University Medical College nr N41/DBS/000172.

Abbreviations: ATP, adenosine triphosphate; BAX, Bcl-2-associated X protein; DNA, deoxyribonucleic acid; dNTP, deoxynucleotide triphosphate; dsDNA, double-stranded deoxyribonucleic acid; FIGO, The International Federation of Gynecology and Obstetrics; HPV, human papillomavirus; HPV16, human papillomavirus type 16; HPV18, human papillomavirus type 18; LSIL, low-grade squamous intraepithelial lesions; mRNA, messenger ribonucleic acid; ORF, open reading frame; SCC, squamous cervical carcinoma; SF3 helicase, superfamily helicase

INTRODUCTION

According to the current data, cervical cancer ranks fourth in the world in terms of incidence and mortality among all cancer types (Bray *et al.*, 2018). In Poland, cervical cancer in women has been consistently the seventh most prevalent type of cancer and it has been listed as

the ninth most frequent cause of death (among women between ages 15–44, it has ranked as the fourth and the third, respectively) (Bruni *et al.*, 2019). It is estimated that 99.7% of cervical cancer cases correspond to an oncogenic HPV type infection (Aref-Adib & Freeman-Wang, 2016). HPV16 is the most common type of HPV found in women with cervical cancer (present in 63.3% of cases) (Zampronha *et al.*, 2013). It should be noted, however, that in 26.9% of cases, coinfection with the second most common HPV virus – HPV18 – has been reported (Zampronha *et al.*, 2013).

HPV16 belongs to the Alphapapillomavirus genus of the Papillomaviridae family. It is the most common type of HPV detected not only in cervical cancer, but also in the head and neck cancers, as well as in the vulvar, penile and anal cancers (Alemany *et al.*, 2015; Anic & Giuliano, 2011; Celebi *et al.*, 2018; Pils *et al.*, 2017). The HPV16 genome is composed of approximately 8 kb-long dsDNA consisting of six early genes (E1, E2, E4-E7), two late genes (L1, L2) and the LCR region that controls transcription and replication of the viral DNA (de Sanjosé *et al.*, 2018). The E1 and E2 proteins maintain expression of the oncogenic E6 and E7 proteins at a low level (Doorbar, 2006). A low concentration of oncogenic proteins is sufficient to realize the life cycle of the virus while not causing disturbances in the process of epithelial cell differentiation (Doorbar, 2006). However, when the virus integrates with the host genome, the E1 or E2 genes are interrupted or deleted, resulting in cell cycle dysregulation and increased expression of viral oncoproteins (Cricca *et al.*, 2009; Schmidt *et al.*, 2005). In case of HPV16, however, this mechanism may work differently as it has been recently shown that for this type of virus the E1 mRNA levels increase along progression of pathological changes (Baedyananda *et al.*, 2018).

According to the UniProt database (www.uniprot.org), the HPV16 E1 protein has four essential domains: the nuclear localization signal, nuclear export signal, SF3 helicase and DNA-binding region. In turn, the HPV16 E2 protein has three principal essential domains: the trans-activation domain, DNA-binding domain and the hinge region.

E6 is an oncogenic protein characterized by two zinc finger motifs (Liu *et al.*, 2009). Its oncogenic potential is generally caused by inactivation of the p53 tumor transformation suppressor and the BAX and Bcl proapoptotic proteins (Thomas & Banks, 1999; Thomas *et al.*, 1999; Vogt *et al.*, 2006).

Presence of specific mutations in genes encoding E1, E2 and E6 proteins is also associated with an increased oncogenic potential of the virus (Bae *et al.*, 2009; Hu *et al.*, 2001; Kahla *et al.*, 2014; Szostek *et al.*, 2017; Tsakogiannis *et al.*, 2014; Yao *et al.*, 2019). Many publications show that the distribution of different variants of HPV16 mutations in the world population is not identical, some variants are characteristic of the European population, others are typical for Asian population and yet they are different from the African one (Burk *et al.*, 2009).

Based on single nucleotide polymorphisms, 4 main HPV16 lineages were distinguished: European Asian, African 1, African 2 and Asian American (Burk *et al.*, 2009). In the European Asian lineage there are European and Asian sublineages (Cornet *et al.*, 2012). In turn, the European subtype is divided into two variants: EUR-p (EUR-350T) and EUR-350G (Cornet *et al.*, 2012). This division is based on the replacement of the T>G nucleotide at position 350 of the E6 gene, which results in an amino acid change at position 83, where leucine replaces valine (L83V) (Zehbe *et al.*, 2001). The HPV16 E6 350G genome variant is more common in women with persistent infections and high-grade cervical diseases than the HPV16 E6 350T prototype (Gheit *et al.*, 2011), but an increase in the chance of cancer progression probably does not occur only because of this additional nucleotide change (Szostek *et al.*, 2017).

The purpose of this study is to determine the HPV16 E1 and E2 mutations in females living in southern Poland and their relation to the severity of cervical pathological changes, as no similar report in the Central and Eastern Europe has been produced yet (Sabol *et al.*, 2012). Our aim was to search for unknown mutations of these genes and assess whether the E1 and E2 sequence variations differ depending on the HPV 16 350T or 350G E6 gene variant.

MATERIALS AND METHODS

The study was conducted on 22 women from southern Poland, aged 28–62 years (mean 43±12) with HPV16 infection confirmed by PCR and reverse hybridization (INNO-LiPA, Innogenetics, Belgium). Clinical material (cervical smears) was taken from women with two different diagnoses: low-grade squamous intraepithelial lesions – LSIL (n=11) and squamous cervical carcinoma – SCC (n=11), FIGO stage I-III. All samples were taken at the National Research Institute of Oncology, Krakow Branch, Poland. The study had been approved by the Ethics Committee of the Jagiellonian University (identification code: 1072.6120.29.2018).

All samples were collected from participating patients before the beginning of treatment. The smears were taken into 2 ml of the 0.9% NaCl solution and stored at –70°C before further processing.

Genomic DNA was isolated from cervical smears using the Genomic DNA Prep Plus kit (A&A Biotechnology, Poland).

Four overlapping fragments of the E1 gene sequence (645 bp, 488 bp, 533 bp, and 405 bp in length) were amplified using 4 overlapping primer sets (Tsakogiannis *et al.*, 2014). The reactions were carried out in the S1000 Thermal Cycler (BIO-RAD Laboratories, Germany), in 25 µl of a mixture containing: 50 ng of isolated DNA, 1× PCR Buffer II, 2 mM MgCl₂, 1 mM dNTPs, 50 pmol of each primer set, 1.25 U of AmpliTaq Gold DNA polymerase (Applied Biosystems, USA). The tem-

perature profile was as follows: initial denaturation of 2 min/95°C, and then 50 cycles: denaturation 40 s/95°C, annealing 30 s/50°C, elongation 72°C for 30 s, final extension at 72°C/5 min.

Two overlapping fragments of the E2 gene sequence (632 bp and 598 bp in length) and E2 complete ORF (1176 bp in length) were amplified using three primer sets (Tsakogiannis *et al.*, 2012). The reaction mixture (25 µl), contained the following: 50 ng of isolated DNA, 1× PCR Buffer II, 2 mM MgCl₂, 0.25 mM dNTPs, 50 pmol of each primer, 2.5 U of AmpliTaq Gold DNA polymerase (Applied Biosystems, USA). The temperature profile was as follows: initial denaturation of 2 min/95°C, and then 40 cycles: denaturation 40 s/95°C, annealing 30 s/56°C, elongation 72°C for 60 s, final extension at 72°C/5 min.

The positive (SiHa cell line containing HPV16 insert) and negative controls (H₂O) were added to each reaction.

The obtained PCR products were enzymatically purified and sequenced at the Genomed company (Poland). The same primers were used for amplification and sequencing of both strands (sense and antisense). Tested sequences were compared with the prototype HPV16 sequence (GenBank K02718) using the BLAST 2.0 database and the ChromasPro 1.5 program.

Identification of the HPV16 E6 gene variant was carried out in the framework of the previously described studies (Szostek *et al.*, 2017).

In order to determine the variants of HPV16 genomes in the tested samples, their sequences were compared with the reference sequences for European (K02718, AF536179, HQ644236), Asian (AF534061), African (AF536180, HQ644298, AF472509), North American (HQ644257) and Asian-American (AY686579, AF402678) variants.

The complete E1 and E2 ORF nucleotide sequences from 22 patients were compared with the HPV16 prototype sequence (GenBank K0.2718). The analysis was carried out according to the advancement of changes and sequence affiliation to the prototype (n=10) and the 350G variant (n=12) of the HPV16 E6 gene.

A statistical analysis was done using the STATISTICA 13.3 software package. We used the Wald-Wolfowitz test to calculate the statistical significance included in this publication. *p* Value≤0.05 was considered as statistically significant.

RESULTS AND DISCUSSION

The E1 gene sequence polymorphisms were found in six samples and in the E2 gene in seven samples.

The analysis of the E1 gene polymorphism is presented in Table 1. Any changes in the gene sequence when compared to the prototype sequence were found in 6/22 samples (27%) including 4/11 LSIL and 2/11 SCC (36% and 18%, respectively, *p*=0.016). Three mutations were detected in one SCC sample, two mutations were found in one LSIL sample, and a single mutation was found in the remaining four samples (3 LSIL, 1 SCC). The sequence analysis showed substitutions at 7 nucleotides located between positions 1053 and 2456 (A1053C, G1189T, G1222A, C1225T, G1345C, A1892C, G2034T) leading to an amino acid change (E63D, R109I, R120K, E121D, S161T, Q343H, A391S), while another nucleotide substitution (C2456T) was silent. Furthermore, for the first time, 4 previously unknown missense mutations (G1189T, G1345C, A1892C, and G2034T) were ob-

Table 1. HPV16 E1 variants identified in 6/22 patients with LSIL and SCC.

Capital letters indicate variants with an amino acid change, while small letters indicate the silent mutation; Symbol “*” means a mutation not previously described. LSIL – low-grade squamous intraepithelial lesions; SCC – squamous cervical cancer; aa – amino acid; NES – nuclear export signal; DBR – DNA binding region; SF3 H – SF3 helicase. Eur-p means standard nucleotide (thymine) at position 350 of the E6 sequence; Eur-350G means guanine at this position. Designations in the “Class/subclass” section indicate nucleotides at the given positions of the E6 protein. PCR was performed 3 times in each sample, and sequencing was done in two directions by a reputable company.

Patient's number (initials)	1 (T.P.)	4 (E.K.)	8 (A.L.)	13 (T.K.)	16 (B.K.)	21 (K.B.)	No. of mutations (study group)	Codon	Domain	Reference ^{aa}	Mutated ^{aa}		
LSIL/SCC	LSIL	SCC	LSIL	SCC	LSIL	LSIL							
Sublineage	Eur-p		Eur-350G										
Class/subclass	350T			350G		350G/109C							
Nucleotide number	1053	A	-	-	-	C	C	2 (LSIL)	63	-	E	D	
	1189	G	-	-	-	T*	-	1 (SCC)	109	NES	R	I	
	1222	G	-	-	-	A	-	1 (SCC)	120	-	R	K	
	1225	C	-	-	-	-	-	T	1 (LSIL)	121	-	E	D
	1345	G	-	-	-	C*	-	-	1 (SCC)	161	-	S	T
	1892	A	C*	-	-	-	-	-	1 (LSIL)	343	DBR	Q	H
	2034	G	-	-	T*	-	-	-	1 (LSIL)	391	-	A	S
	2456	C	-	t	-	-	-	-	1 (SCC)	531	SF3 H	N	-
No. of mutations	1	1	1	3	1	2							

served. The A1053C mutation was found in two LSIL samples, the C1225T, A1892C and G2034T mutations in single LSIL samples, while the G1189T, G1222A and G1345C were found in SCC samples. Only in two cases amino acid substitutions were non-conservative (R109I, Q343H), while other substitutions were conservative. Using the UniProt database, it was found that two of the above missense mutations were related to specific E1 protein domains: G1189T – the nuclear export signal, A1892C – the DNA binding region.

Analysis of the E2 gene polymorphism is presented in Table 2. Changes in the gene sequence (in comparison to the prototype sequence) were found in 7/22 samples (32%) – in 3/11 LSIL and 4/11 SCC (27% and 36%, respectively, $p=0.13$). The highest number of mutations was detected in three patients with LSIL – 13, 12 and 9 mutations, respectively. Four mutations were detected in one SCC sample and identical single mutations were detected in the remaining three samples. Furthermore, for the first time, 18 previously unknown missense mutations (A2815G, C2833T, T2835G, C2839A, T2841A, C2876G, C2876T, T2877A, A2885T, A2910C, A2918T, G3187T, C3401A, A3509T, T3523G, G3533T, A3534C, A3691G) and 2 nonsense mutations (C2948G, A2950T) were observed. The sequence analysis showed 27 substitutions at 25 nucleotides located between positions 2815 and 3766 (A2815G, T2817G, C2833T, T2835G, C2839A, T2841A, C2876G, C2876T, T2877A, T2881G, T2881A, A2885T, A2910C, A2918T, G3187T, C3401A, C3409T, C3461T, C3464T, A3509T, T3523G, G3533T, A3534C, C3535T, C3573A, A3691G, G3766A) leading to an amino acid change (N21E, R27W, H29K, A41G, A41V, Y43D, Y43N, Y44F, K52N, N55I, V145L, A216E, P219S, T236M, T237I, H252L, L257V, D260S, S261L, S273R, T313A, D338N), while other nucleotide substitutions (C2892G, C2922T, A2937G, A3024G, A3186T, C3408T, C3438A, A3454T, C3504T, T3560G) were silent or encoded the STOP codon (C2948G, A2950T). The A2937G mutation was found twice in the LSIL samples, while the C3409T mutation was found three times in the SCC samples. This mutation was the only mutation

found in these samples. In six cases, amino acid substitutions were non-conservative (R27W, Y43D, Y43N, K52N, L257V, D260S, D338N) while other substitutions were conservative. Using the UniProt database, it was found that all of the above missense mutations were related to the specific E2 protein domains and regions: A2815G, T2817G, C2833T, T2835G, C2839A, T2841A, C2876G, C2876T, T2877A, T2881G, T2881A, A2885T, A2910C, A2918T, G3187T – the transactivation domain; C3401A, C3409T, C3461T, C3464T, A3509T, T3523G, G3533T, A3534C, C3535T, C3573A – the hinge domain; and A3691G and G3766A – the DNA binding domain.

It was also found that the difference in the number of nucleotide changes in the E2 gene between the E6 350T (13) and 350G (29) variants is statistically significant ($p=0.03$).

There were no mutations in the transactivation domain in the Eur-p variant, while in the Eur-350G variant they were dominant. In 3/4 of the samples from SCC, the nucleotide changes were only in the hinge region (Table 2).

HPV16 is the most frequently detected oncogenic type of HPV among women in Poland and worldwide. In this study, we examined the polymorphisms of the E1 and E2 genes of this virus found in women living in southern Poland. The E1 gene sequences were analyzed first. It is one of the key proteins in the process of replication of a viral genome – it has an ATP helicase activity (McBride & Warburton, 2017). The influence of the E1 protein mutations on oncogenesis has been studied by several authors and their results clearly show a correlation of some mutations of this protein with the occurrence of HSIL and cervical cancers (Sabol *et al.*, 2012; Tsakogiannis *et al.*, 2014; Yao *et al.*, 2019). The relationship with the process of cervical oncogenesis was found in the case of following mutations: T933A, T1014G, A1668G, G2073A, G2160A, T2169C, T2189C, T2232C, G2337A, A2453T, C2454T, A2547G, A2587T and G2650A. None of the above mutations were found in samples tested in this study. However, in two LSIL samples the A1053C mutation was found, previously de-

Table 2. HPV16 E2 variants identified in 7/22 patients with LSIL and SCC.

Capital letters indicate variants with an amino acid change, while small letters indicate the silent mutation; Symbol "*" means a mutation not previously described. LSIL – low-grade squamous intraepithelial lesions; SCC – squamous cervical cancer; aa – amino acid; T – trans-activation; H – hinge; DB – DNA binding. Eur-p means standard nucleotide (thymine) at position 350 of the E6 sequence; Eur-350G means guanine at this position. Designations in the "Class/subclass" section indicate nucleotides at the given positions of the E6 protein. PCR was performed 3 times in each sample, and sequencing was done in two directions by a reputable company.

Patient's number (initials)	3 (B.S.)	6 (M.K.)	7 (M.J.M.)	10 (I.S.)	13 (T.K.)	18 (S.Z.)	20 (K.M.)						
LSIL/SCC	LSIL	SCC	LSIL	LSIL	SCC	SCC	SCC	No. of mutations	Codon	Domain	Referen- ce ^{aa}	Muta- ted ^{aa}	
Sublineage	Eur- p	Eur-350G											
Class/subclass	350T		350G										
2815	A	-	G*	-	-	-	-	1 (SCC)	21	T	N	E	
2817	T	-	G	-	-	-	-	1 (SCC)					
2833	C	-	-	-	T*	-	-	1 (LSIL)	27	T	R	W	
2835	T	-	-	-	G*	-	-	1 (LSIL)					
2839	C	-	-	-	A*	-	-	1 (LSIL)	29	T	H	K	
2841	T	-	-	-	A*	-	-	1 (LSIL)					
2876	C	-	-	G*	T*	-	-	2 (LSIL)	41	T	A	G/V	
2877	T	-	-	A*	-	-	-	1 (LSIL)	41				
2881	T	-	-	G	A	-	-	2 (LSIL)	43	T	Y	D/N	
2885	A	-	-	-	T*	-	-	1 (LSIL)	44	T	Y	F	
2892	C	-	-	-	g*	-	-	1 (LSIL)	46	T	A	-	
2910	A	-	-	-	C*	-	-	1 (LSIL)	52	T	K	N	
2918	A	-	-	-	T*	-	-	1 (LSIL)	55	T	N	I	
2922	C	-	-	-	t*	-	-	1 (LSIL)	56	T	H	-	
2937	A	-	-	g	g	-	-	2 (LSIL)	61	T	T	-	
2948	C	-	-	G*	-	-	-	1 (LSIL)	65	T	S	STOP	
2950	A	-	-	T*	-	-	-	1 (LSIL)	66	T	K	STOP	
3024	A	-	-	g*	-	-	-	1 (LSIL)	90	T	E	-	
3186	A	-	-	t	-	-	-	1 (LSIL)	144	T	S	-	
3187	G	-	-	T*	-	-	-	1 (LSIL)	145	T	V	L	
3401	C	-	A*	-	-	-	-	1 (SCC)	216	H	A	E	
3408	C	-	t	-	-	-	-	1 (SCC)	218	H	H	-	
3409	C	-	-	-	-	T	T	3 (SCC)	219	H	P	S	
3438	C	A	-	-	-	-	-	1 (LSIL)	228	H	A	-	
3454	A	t*	-	-	-	-	-	1 (LSIL)	234	H	T	-	
3461	C	T	-	-	-	-	-	1 (LSIL)	236	H	T	M	
3464	C	T	-	-	-	-	-	1 (LSIL)	237	H	T	I	
3504	C	t*	-	-	-	-	-	1 (LSIL)	250	H	P	-	
3509	A	T*	-	-	-	-	-	1 (LSIL)	252	H	H	L	
3523	T	G*	-	-	-	-	-	1 (LSIL)	257	H	L	V	
3533	G	T*	-	-	-	-	-	1 (LSIL)	260	H	D	S	
3534	A	C*	-	-	-	-	-	1 (LSIL)					
3535	C	T	-	-	-	-	-	1 (LSIL)	261	H	S	L	
3560	T	g*	-	-	-	-	-	1 (LSIL)	269	H	T	-	
3573	C	A	-	-	-	-	-	1 (LSIL)	273	H	S	R	
3691	A	G*	-	-	-	-	-	1 (LSIL)	313	DB	T	A	
3766	G	-	A	-	-	-	-	1 (LSIL)	338	DB	D	N	
No. of mutations	13	5	9	12	1	1	1						

scribed in the literature as frequent (Sabol *et al.*, 2012) but with no information about its relationship to oncogenesis. In one of the SCC samples, 3 nucleotide changes were observed, which resulted in a change in amino acid sequence in the protein, including one of which was non-conservative and considered important for the protein function.

Next, we analyzed the E2 gene. Its most known function is to control the level of E6 and E7 oncoproteins (Nishimura *et al.*, 2000), and moreover the E2 gene also participates in the process of viral genome replication and post-transcriptional processing (Schwartz *et al.*, 2013). For the E2 gene, mutations have been described that are directly associated with a pro-oncogenic change in the protein function, such as the lack of control over E6 and E7 levels (Kahla *et al.*, 2014); mutations that have an influence on the negative prognosis when using radiation therapy are also described (Kahla *et al.*, 2018). For the E2 gene, we found the C3409T mutation in three SCC samples. Importantly, this mutation was not found in any of the LSIL samples. This mutation was also described by other authors (Casas *et al.*, 1999; Swan *et al.*, 2005; Tsakogiannis *et al.*, 2012). It is a nucleotide change associated with the E6 European variant 350G, as well as with other types outside Europe (Casas *et al.*, 1999; Swan *et al.*, 2005). There are also reports that this polymorphism is associated with a higher risk of developing high-grade dysplasia or neoplasia (Graham & Herrington, 2000). Importantly, the C3409T mutation is located within the hinge region. The literature reports that deletions in this region, particularly at nucleotides 3243 to 3539, are associated with cervical cancer (Arias-Pulido *et al.*, 2006). In our samples, mutations in this range were found in four SCC samples and one LSIL sample ($p=0.049$). We therefore conclude that both, deletions and mutations in this region may affect transformation of the cervical epithelial cells. The nucleotide changes in the E2 gene that herein are described for the first time, were detected in individual samples. It is a good starting point for a search of identified mutations in a subsequent study that would lead to establishing their relationship with oncogenesis.

REFERENCES

- Alemay L, Saunier M, Alvarado-Cabrero I, Quirós B, Salmeron J, Shin HR, Pirog E, Guimerà N, Hernández GA, Felix A, Clavero O, Lloveras B, Kasamatsu E, Goodman MT, Hernandez BY, J Laco, Tinoco L, Geraets DT, Lynch CF, Mandys V, Poljak M, Jach R, Verge J, Clavel C, Ndiaye C, Klaustermeier J, Cubilla A, Castell-sagué X, Bravo IG, Pawlita M, Quint W, Muñoz N, Bosch FX, Sanjosé S (2015) HPV DNA Prevalence and Type Distribution in Anal Carcinomas Worldwide. *Int J Cancer* **136**: 98–107. <https://doi.org/10.1002/ijc.28963>
- Anic GM, Giuliano AR (2011) Genital HPV Infection and Related Lesions in Men. *Prev Med* **53**: 1–14. <https://doi.org/10.1016/j.ypmed.2011.08.002>
- Aref-Adib M, Freeman-Wang T (2016) Cervical Cancer Prevention and Screening: The Role of Human Papillomavirus Testing. *Obstet Gynaecol* **18**: 251–263. <https://doi.org/10.1111/tog.12279>
- Arias-Pulido H, Peyton CL, Joste NJ, Vargas H, Wheeler CM (2006) Human papillomavirus type 16 integration in cervical carcinoma *in situ* and in invasive cervical cancer. *J Clin Microbiol* **44**: 1755–1762. <https://doi.org/10.1128/JCM.44.5.1755-1762.2006>
- Bae JH, Chung JL, Lee SJ, Luk AC, Tong SY, Chan PK, Park JS (2009) Distribution of human papillomavirus type 58 variants in progression of cervical dysplasia in Korean women. *J Microbiol Biotechnol* **19**: 1051–1054. <https://doi.org/10.4014/jmb.0812.679>
- Baediananda F, Chaiwongkot A, Bhattarakosol P (2018) Elevated HPV16 E1 expression is associated with cervical cancer progression. *Intervirology* **60**: 171–180. <https://doi.org/10.1159/000487048>
- Bray F, Ferlay J, Soerjomataram J, Siegel RL, Torre LA, Jemal A (2018) Global Cancer Statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin* **68**: 394–424. <https://doi.org/10.3322/caac.21492>
- Bruni L, Albero G, Serrano B, Mena M, Gómez D, Muñoz J, Bosch FX, de Sanjosé S (2019) Human papillomavirus and related diseases report. ICO/IARC Information Centre on HPV and Cancer (HPV Information Centre): **307**
- Burk RD, Zigi C, Koenraad VD (2009) Human Papillomaviruses: Genetic Basis of Carcinogenicity. *Public Health Genom* **12**: 281–290. <https://doi.org/10.1159/000214919>
- Casas L, Galvan SC, Ordóñez RM, López N, Guido M, Berumen J (1999) Asian-American variants of human papillomavirus type 16 have extensive mutations in the E2 gene and are highly amplified in cervical carcinomas. *Int J Cancer* **83**: 449–455. [https://doi.org/10.1002/\(sici\)1097-0215\(19991112\)83:4<449::aid-ijc3>3.0.co;2-0](https://doi.org/10.1002/(sici)1097-0215(19991112)83:4<449::aid-ijc3>3.0.co;2-0)
- Celebi OO, Sener E, Hosal S, Cengiz M, Gullu I, Tezel GG (2018) Human papillomavirus infection in patients with laryngeal carcinoma. *BMC Cancer* **18**: 1–5. <https://doi.org/10.1186/s12885-018-4890-8>. PMID: 30342481
- Cornet I, Gheit T, Franceschi S, Vignat J, Burk RD, Sylla BS, Tommasino M, Clifford GM; IARC HPV Variant Study Group (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>
- Cricca M, Venturoli S, Leo E, Costa S, Musiani M, Zerbinì M (2009) Disruption of HPV 16 E1 and E2 genes in precancerous cervical lesions. *J Virol Methods* **158**: 180–183. <https://doi.org/10.1016/j.jviromet.2009.01.005>
- Doorbar J (2006) Molecular biology of human papillomavirus infection and cervical cancer. *Clin Sci (London)* **110**: 525–541. <https://doi.org/10.1042/CS20050369>
- Gheit T, Cornet I, Clifford GM, Iftner T, Munk C, Tommasino M, Kjaer SK (2011) Risks for persistence and progression by human papillomavirus type 16 variant lineages among a population-based sample of Danish women. *Cancer Epidemiol Biomarkers Prev* **20**: 1315–1321. <https://doi.org/10.1158/1055-9965.EPI-10-1187>
- Graham DA, Herrington C (2000) HPV-16 E2 Gene disruption and sequence variation in CIN 3 lesions and invasive squamous cell carcinomas of the cervix: relation to numerical chromosome abnormalities. *Mol Pathol* **53**: 201–206. <https://doi.org/10.1136/mp.53.4.201>
- Hu X, Pang T, Guo Z, Mazurenko N, Kissel'jov F, Ponten J, Nister M (2001) HPV16 E6 Gene variations in invasive cervical squamous cell carcinoma and cancer *in situ* from Russian patients. *Br J Cancer* **84**: 791–795. <https://doi.org/10.1054/bjoc.2000.1619>
- Kahla S, Kochbati L, Hammami S, Chanoufi MB, Maalel M, Oueslati R (2014) Sequence Variation in the E2-binding domain of HPV16 and biological function evaluation in Tunisian cervical cancers. *Biomed Res Inter* **2014**: 639321. <https://doi.org/10.1155/2014/639321>
- Kahla S, Hammami S, Kochbati L, Chanoufi MB, Oueslati R (2018) HPV16 E2 variants correlated with radiotherapy treatment and biological significance in cervical cell carcinoma. *Infect Genet Evol* **65**: 238–243. <https://doi.org/10.1016/j.meegid.2018.08.001>
- Liu Y, Cherry L J, Dineen JV, Androphy EJ, Baleja JD (2009) Determinants of stability for the E6 protein of papillomavirus type 16. *J Mol Biol* **386**: 1123–1137. <https://doi.org/10.1016/j.jmb.2009.01.018>
- McBride AA, Warburton A (2017) The role of integration in oncogenic progression of HPV-associated cancers. *PLoS Pathog* **13**: 1–7. <https://doi.org/10.1371/journal.ppat.1006211>
- Nishimura A, Takeshi O, Ishimoto T, Dowhanick JJ, Frizzell MA, Howley PM, Sakai H (2000) Mechanisms of human papillomavirus E2-mediated repression of viral oncogene expression and cervical cancer cell growth inhibition. *J Virol* **74**: 3752–3760. <https://doi.org/10.1128/jvi.74.8.3752-3760.2000>
- Pils S, Gensthaler L, Alemany L, Horvat R, de Sanjosé S, Joura EA (2017) HPV Prevalence in vulvar cancer in Austria. *Wien Klin Wochenschr* **129**: 805–809. <https://doi.org/10.1007/s00508-017-1255-2>
- Sabol I, Matovina M, Si-Mohamed A, Grce M (2012) Characterization and whole genome analysis of human papillomavirus type 16 E1-1374A63nt variants. *PLoS ONE* **7**: e41045. <https://doi.org/10.1371/journal.pone.0041045>
- de Sanjosé S, Brotons M, Pavón MA (2018) The natural history of human papillomavirus infection. *Best Pract Res Clin Obstet Gynaecol* **47**: 2–13. <https://doi.org/10.1371/journal.pone.0041045>
- Schmidt MT, Olejnik AK, Gozdzička-Jozefiak A (2005) The HPV16 E2 transcriptional regulator mode of action depends on the physical state of the viral genome. *Acta Biochim Pol* **55**: 687–692. https://doi.org/10.18388/abp.2009_2466
- Schwartz S (2013) Papillomavirus transcripts and posttranscriptional regulation. *Virology* **445**: 187–95. <https://doi.org/10.1016/j.virol.2013.04.034>
- Swan DC, Rajeevan M, Tortolero-Luna G, Follen M, Tucker RA, Unger ER (2005) Human papillomavirus type 16 E2 and E6/E7 variants. *Gynecol Oncol* **96**: 695–700. <https://doi.org/10.1016/j.ygyno.2004.11.045>

- Szostek S, Zawilinska B, Klimek M, Kosz-Vnenchak M (2017) HPV16 E6 polymorphism and physical state of viral genome in relation to the risk of cervical cancer in women from the south of Poland. *Acta Biochim Pol* **64**: 143–149. https://doi.org/10.18388/abp.2016_1364
- Thomas M, Banks L (1999) Human Papillomavirus (HPV) E6 Interactions with bak are conserved amongst E6 proteins from high and low risk HPV types. *J Gen Virol* **80**: 1513–1517. <https://doi.org/10.1099/0022-1317-80-6-1513>
- Thomas M, Pim D, Banks L (1999) The role of the E6-P53 interaction in the molecular pathogenesis of HPV. *Oncogene* **18**: 7690–7700. <https://doi.org/10.1038/sj.onc.1202953>
- Tsakogiannis D, Ruether IGA, Kyriakopoulou Z, Pliaka V, Theoharopoulou A, Skordas V, Panotopoulou E, Nepka C, Markoulatos P (2012) Sequence variation analysis of the E2 gene of human papilloma virus type 16 in cervical lesions from women in Greece. *Arch Virol* **157**: 825–832. <https://doi.org/10.1007/s00705-012-1236-8>
- Tsakogiannis D, Darmis F, Gortsilas P (2014) Nucleotide polymorphisms of the human papillomavirus 16 E1 gene. *Arch Virol* **159**: 51–63. <https://doi.org/10.1007/s00705-013-1790-8>
- Vogt M, Butz K, Dymalla S, Hoppe-Seyler F (2006) Inhibition of bax activity is crucial for the antiapoptotic function of the human papillomavirus E6 oncoprotein. *Oncogene* **25**: 4009–4015. <https://doi.org/10.1038/sj.onc.1209429>
- Yao Y, Yan Z, Dai S, Li C, Yang L, Liu S, Zhang X, Shi L, Yao Y (2019) Human papillomavirus type 16 E1 mutations associated with cervical cancer in a Han Chinese population. *Int J Med Sci* **16**: 1042–1049. <https://doi.org/10.7150/ijms.34279>
- Zampronha RAC, Freitas-Junior R, Murta EF, Michelin MA, Barbaresco AA, Adad SJ, de Oliveira AM, Rassi AB, Oton GJB (2013) Human papillomavirus types 16 and 18 and the prognosis of patients with stage 1 cervical cancer. *Clinics (Sao Paulo)* **68**: 809–814. [https://doi.org/10.6061/clinics/2013\(06\)14](https://doi.org/10.6061/clinics/2013(06)14)
- Zehbe I, Tachezy R, Mytilineos J, Voglino G, Mikyskova I, Delius H, Marongiu A, Gissmann L, Wilander E, Tommasino M (2001) Human papillomavirus 16 E6 polymorphisms in cervical lesions from different European populations and their correlation with human leukocyte antigen class II haplotypes. *Int J Cancer* **94**: 711–716. <https://doi.org/10.1002/ijc.1520>