

Comparative analysis of HPV16 gene expression profiles in cervical and in oropharyngeal squamous cell carcinoma

Andrea Cerasuolo¹, Clorinda Annunziata¹, Marianna Tortora¹, Noemy Starita¹, Giovanni Stellato², Stefano Greggi², Maria Grazia Maglione³, Franco Ionna³, Simona Losito⁴, Gerardo Botti⁴, Luigi Buonaguro¹, Franco M. Buonaguro¹ and Maria Lina Tornesello¹

¹ Molecular Biology and Viral Oncology Unit, Istituto Nazionale Tumori IRCCS "Fondazione G. Pascale", Napoli, Italy

² Gynecology Oncology Division, Istituto Nazionale Tumori IRCCS "Fondazione G. Pascale", Napoli, Italy

³ Department of Maxillofacial and Ear Nose and Throat Surgery, Istituto Nazionale Tumori IRCCS "Fondazione G. Pascale", Napoli, Italy

⁴ Department of Pathology, Istituto Nazionale Tumori IRCCS "Fondazione G. Pascale", Napoli, Italy

Correspondence to: Maria Lina Tornesello, **email:** m.tornesello@istitutotumori.na.it

Keywords: HPV16; E6 gene; E7 gene; cervical carcinoma; cervical intraepithelial neoplasia; Pathology Section

Received: February 07, 2016

Accepted: February 28, 2017

Published: March 07, 2017

Copyright: Cerasuolo et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC-BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Human papillomavirus type 16 (HPV16) is the major cause of cervical cancer and of a fraction of oropharyngeal carcinoma. Few studies compared the viral expression profiles in the two types of tumor. We analyzed HPV genotypes and viral load as well as early (E2/E4, E5, E6, E6*I, E6*II, E7) and late (L1 and L2) gene expression of HPV16 in cervical and oropharyngeal cancer biopsies. The study included 28 cervical squamous cell carcinoma (SCC) and ten oropharyngeal SCC, along with pair-matched non-tumor tissues, as well as four oropharynx dysplastic tissues and 112 cervical intraepithelial neoplasia biopsies. Viral load was found higher in cervical SCC (<1 to 694 copies/cell) and CIN (<1 to 43 copies/cell) compared to oropharyngeal SCC (<1 to 4 copies/cell). HPV16 E2/E4 and E5 as well as L1 and L2 mRNA levels were low in cervical SCC and CIN and undetectable in oropharynx cases. The HPV16 E6 and E7 mRNAs were consistently high in cervical SCC and low in oropharyngeal SCC. The analysis of HPV16 E6 mRNA expression pattern showed statistically significant higher levels of E6*I versus E6*II isoform in cervical SCC ($p = 0.002$) and a slightly higher expression of E6*I versus E6*II in oropharyngeal cases. In conclusion, the HPV16 E5, E6, E6*I, E6*II and E7 mRNA levels were more abundant in cervical SCC compared to oropharyngeal SCC suggesting different carcinogenic mechanisms in the two types of HPV-related cancers.

INTRODUCTION

Cancers of the cervix and of the head and neck region are among the most common tumors in the world accounting for approximately 528,000 and 686,000 new cases in 2012, respectively [1]. Although cervical cancer incidence has decreased over the last decades in many high-resource countries, due to the introduction of cervical screening programs, a stable or increasing incidence has been reported in low-income countries [2-4]. In parallel, the incidence of head and neck tumors, including larynx,

hypopharynx and oral cavity carcinoma, has decreased as a consequence of smoking and alcohol use decline [5, 6]. However, during the same time period an increased incidence of oropharyngeal cancer has been observed in European countries, in Australia, Canada, and in the United States suggesting a new and emerging risk factor particularly among young men [7, 8].

Twelve high risk human papillomaviruses (HPV) have been recognized as the necessary cause of cervical cancer and of a subgroup of head and neck tumors. Indeed, the viral DNA has been identified in more than

99% of cervical carcinoma and in approximately 30% of head and neck tumors with a greater prevalence in the oropharyngeal SCC (45.8%) [9-12]. The HPV16 is the predominant cause of infection in both cervical and oropharyngeal SCC accounting for 65-70% and 82-87% of all HPV-positive cases, respectively, across the world [12-14].

The pathogenesis of oncogenic HPVs has been well studied in cervical neoplasia. The HPV infection initiates at the basal layer of the epithelium with the expression of early proteins E5, E6 and E7 which stimulate cell growth and viral DNA replication [15-19]. During the productive phase of HPV infection the viral genomes markedly replicate in the spinous layers and are encapsidated in the upper terminally differentiated epithelia where the late viral genes L1 and L2 are highly expressed [15, 20]. In cervical intraepithelial neoplasia grade 3 (CIN 3) and cervical carcinoma the virus undergoes abortive infection in which early genes E6 and E7 are expressed in all epithelial layers and cause the alteration of a number of cellular pathways involved in cell cycle regulation and apoptosis [15, 20, 21]. Specifically, the E6 protein binds to p53 oncosuppressor causing its ubiquitination and proteasomal degradation [22-24], and E7 oncoprotein abrogates pRB activity [25-27]. The HPV16 E5 protein induces dimerization of the epidermal growth factor receptor (EGFR) on the cell membrane concurring to the activation of mitogenic signals [18, 28].

The abortive infections are often accompanied by the integration of viral genome into human chromosomes determining the over expression of E6 and E7 oncoproteins either by the interruption of E2-mediated transcriptional repression or by functional alterations in the long control region [29-32].

Transcription of HPV16 E6 and E7 as well as E1, E2, E4 and E5 early viral genes, driven by enhancer elements located in the long control region and early promoter P97, produces multiple polyadenylated mRNAs containing exonic and intronic sequences which are alternatively spliced [33]. The alternative splicing of full-length polycistronic E6/E7 mRNAs, through the differential use of the splice donor at nucleotide 226 and splice acceptors at nucleotides 409 and 526, produces the E6*I and E6*II isoforms, respectively [33-35]. The E6*I and E6*II transcripts have shown to be more abundant in high grade neoplasia and invasive cancers than low grade cervical lesions [36]. The post-transcription regulation of HPV16 genes is controlled by several cellular splicing factors, including serine-arginine-rich (SR) proteins, heterogeneous nuclear ribonucleoproteins (hnRNPs), cleavage stimulation factor 64 kDa subunit (CSTF64) and CUG triplet repeat RNA-binding protein1 (CUGBP1), which are highly expressed in basal and middle layers cells of the cervical epithelium [37, 38]. No study has compared the HPV16 gene expression profiles in cervical and oropharyngeal SCC.

The aim of our study was to analyze the viral load and expression levels of HPV16 early genes, particularly E5, E6, E6*I, E6*II and E7 genes as well as L1 and L2 late genes in cervical carcinoma and in oropharyngeal cancer. We sought to identify similarities and differences in viral related transformation mechanisms between the two types of tumor.

RESULTS

HPV genotype distribution and viral load analysis

The study included women diagnosed with cervical SCC (n=28), borderline to mild dyscaryosis cytology (BMD) (n=40), CIN1 (n=66), CIN2 (n=4) and CIN3 (n=2) with a mean age of 52.6 (\pm 12.7), 40.2 (\pm 10.9), 38.2 (\pm 10.2), 32.8 (\pm 5.3), 22.5 (\pm 4.9) years at diagnosis, respectively. Six men (mean age of 52.3 [\pm 15.2] years) and four women (mean age of 58.5 [\pm 15.9] years) affected by oropharyngeal SCC, as well as three men (mean age of 45.7 [\pm 11.7] years) and one woman (41 years) with oropharyngeal dysplasia, were also included in the study. HPV DNA sequences were identified in 92.8 % (26 out of 28) of cervical SCC, 70% (7 out of 10) of oropharyngeal SCC, 65% (26 out of 40) of BMD, 65.3% (47 out of 72) of CIN, and 25% (1 out of 4) of oropharyngeal dysplastic tissues. Among the HPV positive samples the type 16 was the most prevalent representing 65.4% (17 out of 26) and 19% (9 out of 47) of all infections in cervical SCC and CIN, respectively, as well as 100% of all infections in oropharyngeal SCC and dysplastic tissues (Table 1). Other frequent genotypes were HPV33, 45, and 58 identified in 11.5% (3 out of 26), 7.7% (2 out of 26) and 7.7% (2 out of 26) of cervical SCC as well as HPV42 identified in 6.4% (3 out of 47) of CIN. The HPV sequences, detected in 57% (4 out of 7) and 44.4% (4 out of 9) of cervical and oropharyngeal pair matched non-tumor tissues, were always concordant with the genotype identified in the corresponding tumor. Samples positive for HPV16 were selected for further analyses.

HPV16 viral load ranged from <1 to 694 copies per cell in cervical SCC (median value 4.5 copies/cell) and from <1 to 43 copies per cell in CIN (median 0.1 copy/cell), while there was less than one copy per cell in paired non-tumor tissues (median 10^{-6} copy/cell). The viral load ranged from <1 to 4 copies per cell in oropharyngeal SCC (median 5×10^{-5} copy/cell) and from <1 to 30 (median 7×10^{-4} copy/cell) in paired non-tumor tissues. Less than one copy per cell was found in dysplasia and paired non-dysplastic tissues (Figure 1). The viral copy number was statistically significant higher in cervical SCC compared to paired non-tumor tissues ($p = 0.007$), but not compared to CIN. No significant differences were observed between viral load in oropharyngeal SCCs and respective paired non-tumor tissues (Figure 1).

HPV16 gene expression analysis

Viral gene expression profiles were analyzed in all HPV16-positive samples by real time PCR using SiHa cell line cDNA as a positive control. The *GAPDH* cDNA was amplified in all samples to normalize the viral gene expression levels. The RNA quality was suitable for the analysis in 10 cervical SCC, six CIN, seven oropharyngeal SCC and one oropharyngeal dysplasia, along with three cervical and five oropharyngeal paired non-tumor tissues and one paired non-dysplastic tissue. Oligonucleotide pairs encompassing the untranslated region upstream the HPV16 promoter P97 were used to exclude the presence of HPV16 DNA in the cDNA samples (Figure 2).

The HPV16 E6 full length mRNA was detected in all cervical SCC, in 14.3% (1 out of 7) of oropharyngeal

SCC and 17% (1 out of 6) of CIN. The full E6 transcripts were also detected in 33.3% (1 out of 3) and in 20% (1 out of 5) of cervical and oropharyngeal paired non-tumor tissues, respectively (Figure 3). The E6*I and E6*II isoforms were both expressed in all cervical SCC, in 50% (3 out of 6) of CIN, in 14.3% (1 out of 7) of oropharyngeal SCC, in 33.3% (1 out of 3) and 20% (1 out of 5) of cervical and oropharyngeal paired non-tumor tissues, respectively (Figure 3). The analysis of E6 mRNA patterns showed that E6*I was more expressed than full-length E6 and E6*II in all samples, but the difference between E6*I and E6*II levels was statistically significant only in cervical SCC ($p = 0.002$). The oncogene E7 was highly expressed in all cervical SCC and paired non-tumor tissues, in 83% (5 out of 6) of CIN, in 14.3% (1 out of 7) of oropharyngeal SCC and in 40% (2 out of 5) of paired

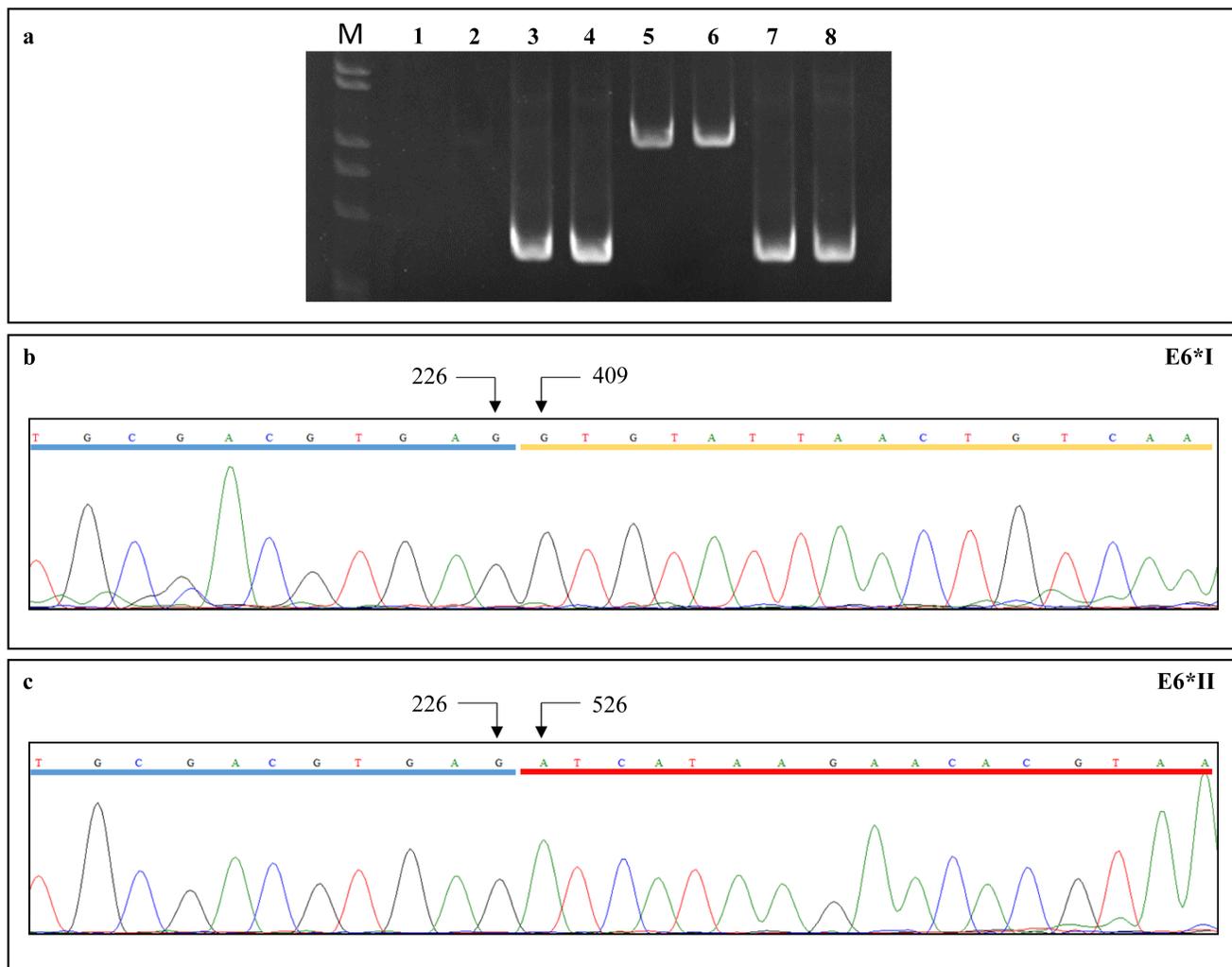


Figure 2: (a) PCR reactions using primer pairs targeting the full-length E6 cDNA (lanes 3-4) and genome viral DNA (lanes 7-8). PCR amplifications with primer pairs encompassing the untranslated region upstream the HPV16 P97 promoter gave negative results on cDNA (lanes 1-2) and positive results on viral DNA (lanes 5-6), suggesting the absence of DNA contamination in cDNA samples. (b) E6*I and (c) E6*II cDNA sequence electropherograms showing donor (nt 226) and acceptor (nt 409 for E6*I and nt 526 for E6*II) splicing nucleotides; exons are underlined.

non-tumor tissues (Figure 3). The E6, E6*I, E6*II and E7 levels were significantly higher in cervical SCC than in paired non-tumor tissues, CIN and oropharyngeal SCC ($p < 0.02$). A positive linear correlation was observed in cervical SCC between E6*I and E7 expression levels ($R = 0.68$, $p = 0.04$) and between E6*II and E7 ($R = 0.78$, $p = 0.01$), (Figure 4). In oropharyngeal dysplasia only E6*I and E7 were found expressed, while none of viral mRNAs were detected in paired non-dysplastic tissues.

The HPV16 E2/E4 transcripts were detected at low levels in 80% (8 out of 10) of cervical SCC and in 17% (1 out of 6) of CIN, but not detected in oropharyngeal SCC. In one CIN only E4 mRNA and not E2 was revealed. The levels of E2/E4 and E5 mRNAs were higher in cervical SCC than in CIN, but the difference was statistical significant only for E2/E4 levels ($p < 0.01$). The E5 mRNA was identified in 80% (8 out of 10) of cervical SCC, in 33.3% (2 out of 6) of CIN, and in none of oropharyngeal SCC and paired non-tumor tissues (Figure 3).

L1 and L2 bicistronic transcripts were expressed at low levels in all cervical SCC and in 83% (5 out of 6) of CIN, however no expression was observed in oropharyngeal SCC and in paired non-tumor tissues.

There was no statistically significant correlation between viral load and oncogene expression in all analyzed samples.

DISCUSSION

Persistent infection with high risk HPVs, mainly HPV16, is strictly associated with cervical SCC [39] and is becoming an emerging etiological factor for oropharyngeal SCC [40]. In this study the viral load and the expression pattern of HPV16 early (E2, E4, E5, E6, E6*I, E6*II, E7) and late (L1 and L2) genes were analyzed in cervical SCC and paired tissues, in CIN, in oropharyngeal SCC and paired non-tumor tissues, as well as in oropharyngeal dysplasia to uncover differences in HPV-related transformation mechanisms in different types of HPV-related tumors. In agreement with previous studies, we found that HPV16 was the most frequent genotype in cervical and oropharyngeal cancer [12, 14, 41].

Studies on the positive correlation between HPV16 viral load and cervical or oropharyngeal SCC are controversial [42-45], with some reporting that viral

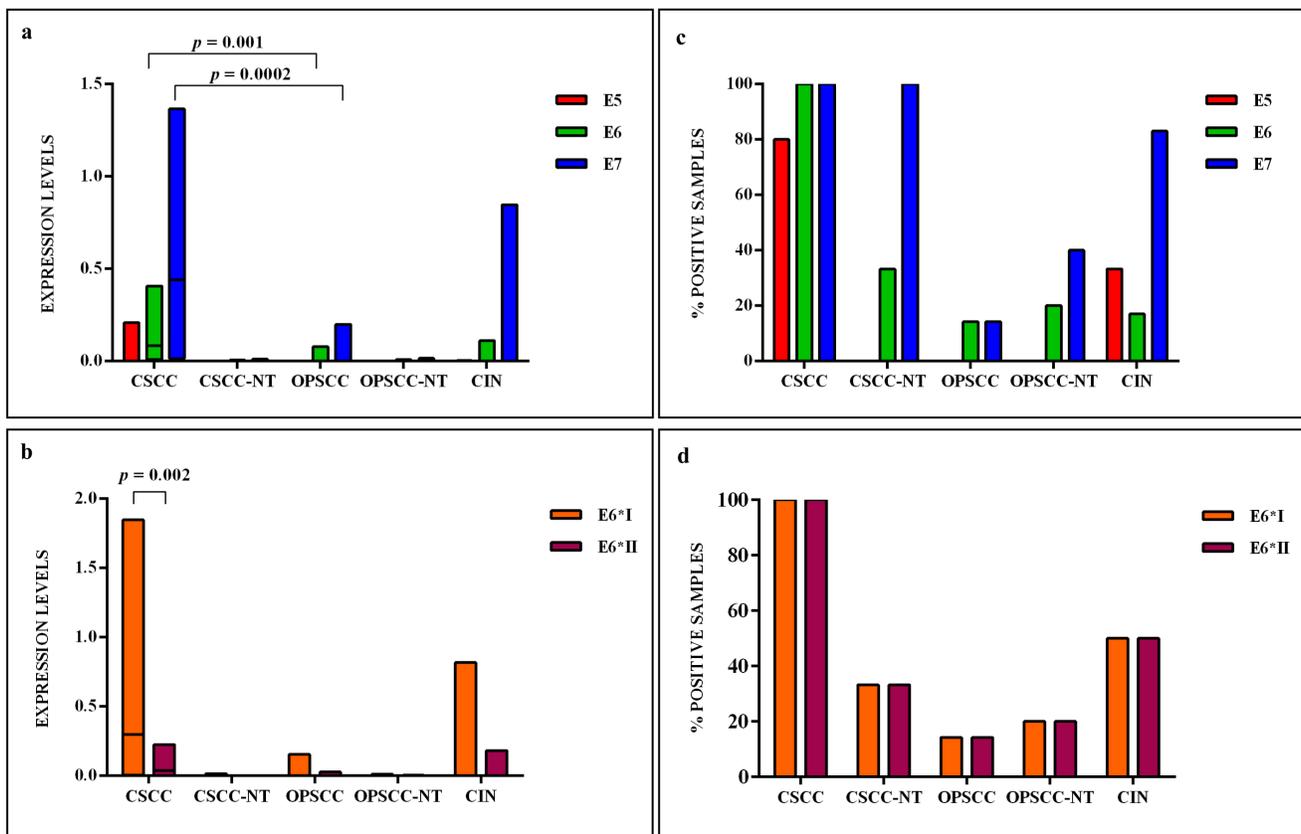


Figure 3: (a) The relative expression values, adjusted to GAPDH expression, of HPV16 E5, E6, E7 and (b) E6 isoforms, in cervical SCC (CSCC), paired cervical non-tumor tissues (CSCC-NT), oropharyngeal SCC (OPSCC), paired oropharyngeal non-tumor tissues (OPSCC-NT) and cervical intraepithelial neoplasia (CIN). The horizontal line in the bars indicates the median. (c) The percentage of cases expressing HPV16 E5, E6, E7 and (d) E6 isoforms in each histological group.

copy number increases with CIN severity and is highest in cervical SCC [46]. We found that HPV16 viral load was broadly variable especially in cervical SCC with a viral copy number significantly higher than in cervical paired non-tumor tissues. In oropharynx we found a slightly higher viral load in non-tumor tissues than in carcinoma. Until now no studies have compared the viral load in oropharyngeal SCC versus autologous non-tumor tissues. A recent study by Dang and Feng (2016) showed that oral and oropharyngeal cancer samples contained a higher HPV16 DNA load than normal tissues from non-cancer patients [47], however these data are not informative regarding the difference of viral copy number between non-tumor tissues and oropharyngeal carcinoma in the same patient.

Several studies showed that HPV gene expression pattern could be useful as viral molecular marker of tumor progression [48-50]. Analysis of E6 and E7 mRNA levels, together with other clinical data, seems useful to assess the risk of progression of cervical SCC and oropharyngeal SCC cases [43, 51, 52]. In our study, the E6 and E7 expression in oropharyngeal SCC and paired non-tumor tissues was much lower than in cervical SCC. These results are consistent with previous studies reporting that while HPV is always transcriptionally active in cervical SCC the viral oncogenes may not be expressed in HPV16 DNA positive oropharyngeal cancer suggesting that the virus in some HPV-positive head and neck cancers is not the main carcinogenic factor [53, 54]. Oropharyngeal SCC patients without biologically active HPV are considered at risk for cancer progression similarly to HPV-negative smoker patients [55-57].

Holzinger et al (2012) reported that both high viral load and increased expression of E6 and E7 mRNAs could define oropharyngeal SCC with active HPV16 involvement. However, we observed that there was no significant correlation between viral copy number and E6 and E7 viral mRNA levels in both cervical and oropharyngeal SCC [50]. Our observation is in agreement with previous studies reporting no correlation between DNA copy number and the E6 and E7 mRNA expression in cervical cancer, implying that not all viral genomes are transcriptionally active in tumors and cancer derived cell lines [58]. Paradigmatic is the highly divergent HPV16 copy number in CaSki (above 400 viral copies per cell) and SiHa cells (1 copy per cell) and the comparable levels of E7 transcripts (41.6 ± 1.5 and 14.8 ± 2.6 transcript copies per cell in CaSki and SiHa cells, respectively) in the two cervical cancer derived cell lines [59].

Recent in vitro molecular studies showed that E6* isoforms of high risk HPVs cooperate with E6 and are able to target the Discs Large MAGUK Scaffold Protein 1 (DLG1), the Pals1-Associated Tight Junction Protein (PATJ), the Membrane-Associated Guanylate Kinase Inverted 1 (MAGI-1), and to a lesser extent the Scribbled Planar Cell Polarity Protein (SCRIB), for proteasomal

degradation, causing invasive growth of epithelial cells and disrupting epithelial cells junctions [60-62]. Several other E6 mRNA species, besides E6*I and E6*II, derived from alternative splicing have been identified, but their role in tumor progression has been poorly characterized [63-67].

The over expression of HPV16 E6*I and E6*II isoforms has been also shown to modify the levels of many proteins involved in mitochondrial dysfunction and oxidative phosphorylation in C33A cervical cells, and the β -integrin signaling pathway in HPV16-positive SiHa cells [66]. In high grade cervical intraepithelial neoplasia and in cervical cancer the E6*I is more expressed than E6*II [36, 68] and E6*I/E6*II ratio seems to be a predictive marker of clinical outcome in HPV-related oropharyngeal cancer [69].

In the present study the expression levels of HPV16 E6*I and E6*II were found high in all cervical SCC but not in oropharyngeal SCC. In agreement with other studies, we observed a linear correlation between E6*I and E7 mRNA as well as between E6*II and E7 mRNA [33, 70].

Interestingly, in the analyzed oropharyngeal dysplasia the E7 oncogene and the E6*I isoform were both detected leading to the hypothesis that HPV oncogenic activity may be important for the early phases of oropharyngeal neoplasia [53, 71]. It has been suggested that during the early stages of oropharyngeal carcinogenesis the viral oncoproteins may synergize with other carcinogens, such as smoke and alcohol, and may increase the risk of tumor progression [53, 71]. In our study the HPV16 E7 oncogene was much more expressed in cervical SCC than oropharyngeal SCC, suggesting that the expression of E7 is the main driver of tumor cell proliferation in cervical cancer but not in oropharyngeal SCC [72].

Several studies showed that E5 expression synergizes with E6 and E7 causing a more severe cancer phenotype [17, 73]. In our study the E5 expression was variable in cervical SCC, low or absent in CIN (except one sample showing also high E6 and E7 mRNA levels) and in cervical paired non-tumor tissues. Variable levels of E5 in cervical SCC could be due to the presence of both episomal and integrated forms of HPV16, as reported by Das et al. (2015) [74]. In fact, integration of HPV16 in the host genome could disrupt E5 ORF together with E2 ORF [75].

In oropharyngeal SCC and respective paired non-tumor tissues the E5 mRNA was found not expressed, in accordance to previous studies reporting variable expressions of E5 in head and neck cancer [73, 76]. The E2/E4 as well as L1 and L2 gene expression was generally low in cervical SCC and CIN and undetectable in oropharyngeal SCC [58, 77].

The main limitations of our study are the limited number of patients and the inability to evaluate long term

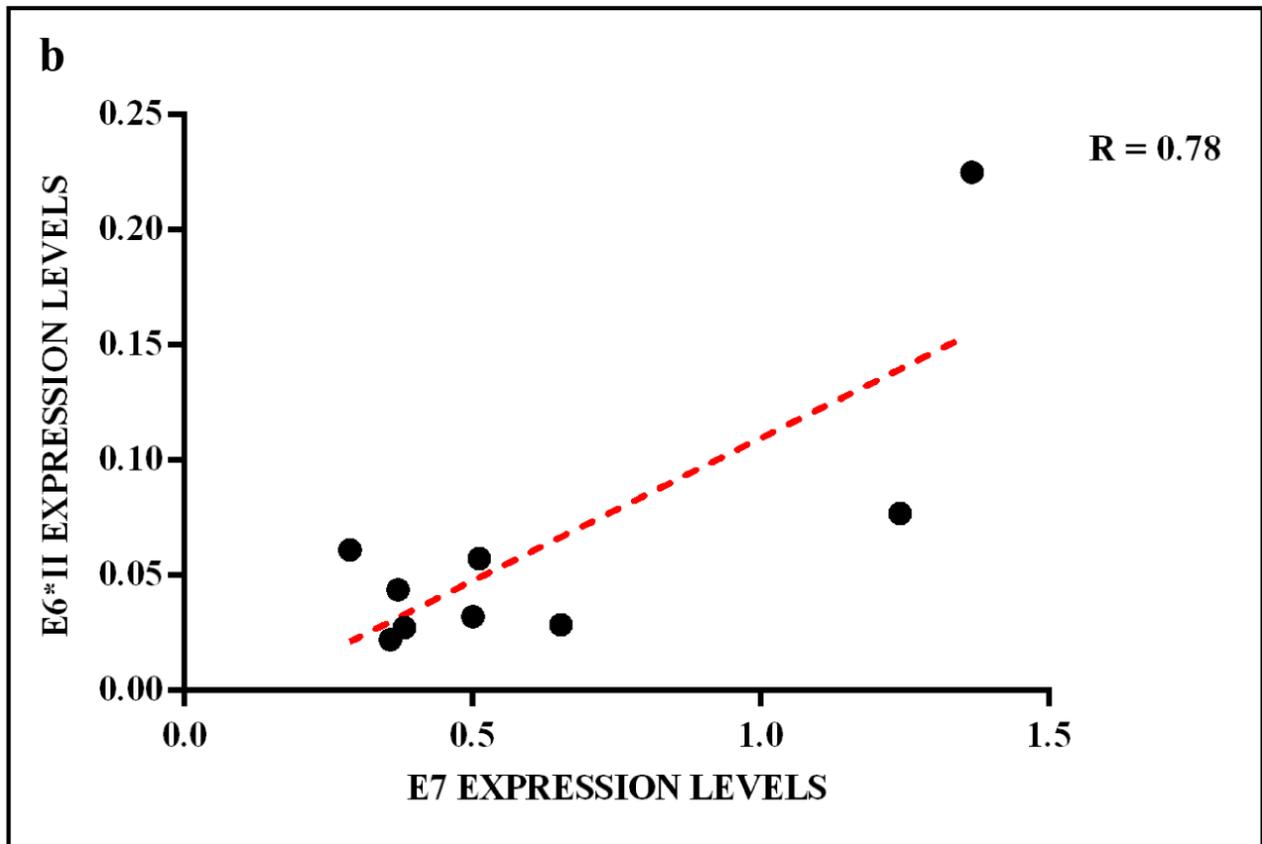
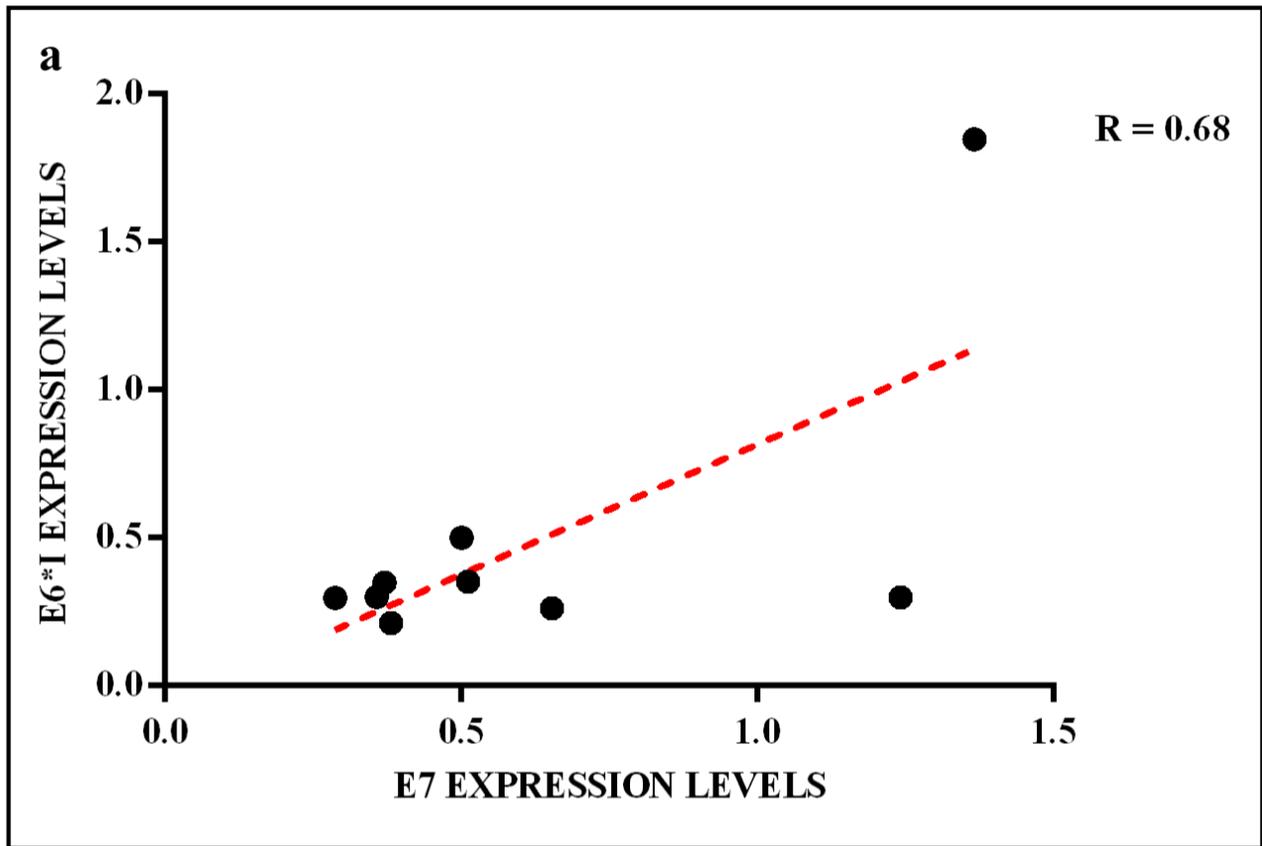


Figure 4: Linear correlation between (a) E6*I/E7 and (b) E6*II/E7 expression levels in cervical SCC.

disease outcome in relation to viral expression profiles. However, this is the first study comparing the HPV16 traits between the cervical and oropharyngeal SCC along with autologous non-tumor tissues.

In conclusion, this study confirmed that HPV16 is highly prevalent in cervical and oropharyngeal SCC. However the viral load is very low in oropharyngeal SCC compared to cervical SCC and importantly the viral oncogene mRNA levels and expression profiles are very different between cervical SCC and oropharyngeal SCC. Indeed, E5, E6, E6*I, E6*II and E7 mRNA were significantly more abundant in cervical SCC than in oropharyngeal SCC suggesting the presence of different carcinogenic mechanisms in the two different virus-related tumors.

MATERIALS AND METHODS

Patients and samples

Twenty-eight cervical SCC biopsies along with paired non-tumor tissues and 112 cervical biopsies, comprising 40 cases of borderline to mild dyscaryosis (BMD) cytology and normal histology, 66 cervical intraepithelial neoplasia (CIN) grade 1, four CIN2 and two CIN3 biopsies, obtained from patients attending the Gynecology Unit of Istituto Nazionale Tumori “Fond Pascale” from November 2013 to December 2015. Ten oropharyngeal SCC and paired non-tumor tissues, as well as four oropharyngeal dysplastic biopsies with paired non-dysplastic tissues, were obtained from patients referred to the Head and Neck Surgery Unit of the Istituto Nazionale Tumori “Fond Pascale” from January 2012 to December 2015 (Table 1). Each biopsy was divided in two sections: the first section was stored in RNA Later (Ambion, Austin, Texas) at -80°C, the second was subjected to histopathologic examination. Similarly, paired non-tumor biopsies were divided in two sections and processed for molecular analysis and histopathologic examination. This study was approved by the Institutional Scientific Board and by the Ethical Committee of the Istituto Nazionale Tumori “Fond Pascale”, and it is in accordance with the principles of the Declaration of Helsinki.

DNA extraction, HPV genotyping and HPV16 viral load

Genomic DNA was extracted according to published procedures [78]. Specifically, 10 mg tissue samples were digested with proteinase K (150 µg per ml at 60°C for 30 min) in 100 µl of lysis buffer (10 mM Tris-HCl pH 7.6, 5 mM EDTA, 150 mM NaCl, 1% SDS), followed by DNA purification with phenol and phenol-chloroform-isoamyl alcohol (25:24:1) extraction and ethanol precipitation

in 0.3 M sodium acetate (pH 4.6). The quantity of isolated DNA was spectrophotometrically assessed with Nanodrop 2000c (Thermo Fisher Scientific, Waltham, Massachusetts).

HPV detection was carried out by nested PCR amplifying 300 ng of genomic DNA with MY09/MY11 primer pairs [79] for the outer reaction and MGP primer system for the inner reaction in 50 µl reaction mixture containing 5 µl of outer reaction, as previously described [80]. HPV genotypes were identified by direct automated DNA sequencing analysis of MGP amplified products using the primer GP5+ [81] at Eurofins Laboratories (Milan, Italy). HPV type identification was performed by alignments of HPV sequences with those present in the GenBank database using the BLASTn software (<http://www.ncbi.nlm.nih.gov/blast/html>). HPV16 positive samples were selected for gene expression and viral load analysis.

HPV16 viral load quantization was performed in the Bio-Rad CFX96 Real-time PCR Detection System using 300 ng of template DNA, 12.5 µl of 1x iQ™ SYBR® Green supermix (Bio-Rad, Hercules, California) and 10 pmol each of E7 forward and reverse primers (Supplementary table) in a final volume of 25 µL. Thermal cycling consisted of a denaturation step at 95°C for 3 min, followed by 40 cycles of annealing at 54.3°C for 30 s, extension at 72°C for 30 s and denaturation at 95°C for 30 s. Exon 7 of *TP53* human gene was also amplified with primers targeting the exon 7 to normalize the viral load in each sample, as previously described [82]. Two replicates were performed for each sample and real-time PCR data were analyzed using Bio-Rad CFX manager software. Two standard curves were constructed to calculate absolute numbers of HPV16 E7 and *TP53* copies, respectively, by amplifying serial dilutions (10⁵ to 1 cell) of SiHa cell genomic DNA, containing 1 copy per cell of integrated HPV16 genome. The viral load per cell in each sample was calculated by normalizing the E7 copy number against the amount of cellular DNA (*TP53*) according to the formula: HPV copies/cell = Number of E7 copies/(number of *TP53* copies/2).

HPV16 gene expression analysis

Total RNA was extracted from all samples using RNeasy MiniKit (Qiagen, Hilden, Germany) according to manufacturer procedure. The quality and quantity of isolated RNA was determined using the Nanodrop 2000c and calculating the ratio of absorbance at 260 nm and 280 nm. All RNA samples with a ratio in the range of 1.8-2.0 were included in further analyses. For each sample 250 ng of total RNA were reverse transcribed in 20 µL volume containing 4 µL of iScript reaction mix (Bio-Rad), 1 µL of iScript reverse transcriptase (Bio-Rad) and nuclease-free water. The reaction was incubated at 25°C for 5 min and at 42°C for 30 min, finally the enzyme was inactivated

at 85°C for 5 min in the Gene Amp PCR System 2400 (Applied Biosystems, Foster City, California).

The cDNA samples were amplified for HPV16 early (E2/E4, E5, E6, E6*I, E6*II, E7) and late (L1 and L2) transcripts by real time PCR using specific primer pairs described in Supplementary table. The reverse primers specific to each E6 isoform encompassed the respective splicing acceptor nucleotides ensuring the specific amplification of each spliced isoform, as confirmed by nucleotide sequence analysis of PCR amplicons (Figure 2). The reaction mixture included 12.5 µL of 1x iQ™ SYBR® Green supermix (Bio-Rad), 10 pmol of each primer, 1 µL of cDNA and nuclease-free water in a final volume of 25 µL. All reactions were performed in duplicate. The amplifications were carried out on the Bio-Rad CFX96 real time PCR Detection System following the protocols described in Supplementary table. The expression of each viral gene was analyzed with the 2^{-ΔCt} method using *GAPDH* as a reference gene. The ΔCt values for each amplified transcript were calculated by subtracting the respective Ct value from the corresponding *GAPDH* Ct ($\Delta C_t = C_{t_x} - C_{t_{GAPDH}}$). The Ct values were corrected for primer pairs efficiency to compare the expression levels of the different genes. Primer efficiency was calculated generating standard curves of SiHa cDNA serial dilutions for each analyzed gene.

Statistical analysis

Statistical analysis was performed with GraphPad version 6 (Prism). Spearman's rank correlation coefficient (r) was calculated to evaluate correlation between viral load and oncogenes expression levels, Pearson's coefficient (R) was calculated to evaluate linear correlation between gene expression levels, while ANOVA Kruskal–Wallis test and U Mann–Whitney test were used to evaluate differences in gene expression levels and viral load. All variables with $p < 0.05$ were considered statistically significant.

Abbreviations

Human Papillomavirus, HPV; squamous cell carcinoma, SCC; cervical intraepithelial neoplasia, CIN.

ACKNOWLEDGMENTS

This study was funded by the Ministero della Salute (Ricerca Corrente 2013-2017). We thank Serena Salerno for English editing.

CONFLICTS OF INTEREST

The authors declare no competing financial interests.

Editorial note

This paper has been accepted based in part on peer-review conducted by another journal and the authors' response and revisions as well as expedited peer-review in Oncotarget.

REFERENCES

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11. In: International Agency for Research on Cancer, editor. Lyon, France: 2013.
2. Jemal A, Bray F, Forman D, O'Brien M, Ferlay J, Center M, Parkin DM. Cancer burden in Africa and opportunities for prevention. *Cancer*. 2012; 118: 4372-84.
3. Vaccarella S, Lortet-Tieulent J, Plummer M, Franceschi S, Bray F. Worldwide trends in cervical cancer incidence: Impact of screening against changes in disease risk factors. *Eur J Cancer*. 2013.
4. Arbyn M, Castellsague X, de SS, Bruni L, Saraiya M, Bray F, Ferlay J. Worldwide burden of cervical cancer in 2008. *Ann Oncol*. 2011; 22: 2675-86.
5. Hashibe M, Brennan P, Benhamou S, Castellsague X, Chen C, Curado MP, Dal ML, Daudt AW, Fabianova E, Fernandez L, Wunsch-Filho V, Franceschi S, Hayes RB, et al. Alcohol drinking in never users of tobacco, cigarette smoking in never drinkers, and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. *J Natl Cancer Inst*. 2007; 99: 777-89.
6. Garavello W, Bertuccio P, Levi F, Lucchini F, Bosetti C, Malvezzi M, Negri E, La VC. The oral cancer epidemic in central and eastern Europe. *Int J Cancer*. 2010; 127: 160-71.
7. Carvalho AL, Nishimoto IN, Califano JA, Kowalski LP. Trends in incidence and prognosis for head and neck cancer in the United States: a site-specific analysis of the SEER database. *Int J Cancer*. 2005; 114: 806-16.
8. Chaturvedi AK, Anderson WF, Lortet-Tieulent J, Curado MP, Ferlay J, Franceschi S, Rosenberg PS, Bray F, Gillison ML. Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers. *J Clin Oncol*. 2013; 31: 4550-9.
9. Guan P, Howell-Jones R, Li N, Bruni L, de SS, Franceschi S, Clifford GM. Human papillomavirus types in 115,789 HPV-positive women: a meta-analysis from cervical infection to cancer. *Int J Cancer*. 2012; 131: 2349-59.
10. Kreimer AR, Clifford GM, Boyle P, Franceschi S. Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. *Cancer Epidemiol Biomarkers Prev*. 2005; 14: 467-75.
11. Gillison ML, D'Souza G, Westra W, Sugar E, Xiao W, Begum S, Viscidi R. Distinct risk factor profiles for human

- papillomavirus type 16-positive and human papillomavirus type 16-negative head and neck cancers. *J Natl Cancer Inst.* 2008; 100: 407-20.
12. Li N, Franceschi S, Howell-Jones R, Snijders PJ, Clifford GM. Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: Variation by geographical region, histological type and year of publication. *Int J Cancer.* 2011; 128: 927-35.
 13. IARC. IARC monographs on the evaluation of carcinogenic risks to humans. A Review of Human Carcinogens Part B: Biological Agents. Lyon, France: 2011.
 14. Agalliu I, Gapstur S, Chen Z, Wang T, Anderson RL, Teras L, Kreimer AR, Hayes RB, Freedman ND, Burk RD. Associations of Oral alpha-, beta-, and gamma-Human Papillomavirus Types With Risk of Incident Head and Neck Cancer. *JAMA Oncol.* 2016.
 15. Doorbar J. Molecular biology of human papillomavirus infection and cervical cancer. *Clin Sci (Lond).* 2006; 110: 525-41.
 16. Pyeon D, Pearce SM, Lank SM, Ahlquist P, Lambert PF. Establishment of human papillomavirus infection requires cell cycle progression. *PLoS Pathog.* 2009; 5: e1000318.
 17. DiMaio D, Mattoon D. Mechanisms of cell transformation by papillomavirus E5 proteins. *Oncogene.* 2001; 20: 7866-73.
 18. Venuti A, Paolini F, Nasir L, Corteggio A, Roperto S, Campo MS, Borzacchiello G. Papillomavirus E5: the smallest oncoprotein with many functions. *Mol Cancer.* 2011; 10:140.
 19. Chow LT, Broker TR, Steinberg BM. The natural history of human papillomavirus infections of the mucosal epithelia. *APMIS.* 2010; 118: 422-49.
 20. Moody CA, Laimins LA. Human papillomavirus oncoproteins: pathways to transformation. *Nat Rev Cancer.* 2010; 10: 550-60.
 21. Tornesello ML, Buonaguro L, Buonaguro FM. An overview of new biomolecular pathways in pathogen-related cancers. *Future Oncol.* 2015; 11: 1625-39.
 22. Scheffner M, Huibregtse JM, Vierstra RD, Howley PM. The HPV-16E6 and E6AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell.* 1993; 75: 495-505.
 23. Wallace NA, Galloway DA. Novel Functions of the Human Papillomavirus E6 Oncoproteins. *Annu Rev Virol.* 2015; 2: 403-23.
 24. Longworth MS, Laimins LA. Pathogenesis of human papillomaviruses in differentiating epithelia. *Microbiol Mol Biol Rev.* 2004; 68: 362-72.
 25. Munger K, Yee CL, Phelps WC, Pietenpol JA, Moses HL, Howley PM. Biochemical and biological differences between E7 oncoproteins of the High- and low-Risk human papillomavirus types are determined by amino-terminal sequences. *J Virol.* 1991; 65: 3943-8.
 26. Jansma AL, Martinez-Yamout MA, Liao R, Sun P, Dyson HJ, Wright PE. The high-risk HPV16 E7 oncoprotein mediates interaction between the transcriptional coactivator CBP and the retinoblastoma protein pRb. *J Mol Biol.* 2014; 426: 4030-48.
 27. Liu X, Clements A, Zhao K, Marmorstein R. Structure of the human Papillomavirus E7 oncoprotein and its mechanism for inactivation of the retinoblastoma tumor suppressor. *J Biol Chem.* 2006; 281: 578-86.
 28. Pim D, Collins M, Banks L. Human papillomavirus type 16 E5 gene stimulates the transforming activity of the epidermal growth factor receptor. *Oncogene.* 1992; 7: 27-32.
 29. Jeon S, Allen-Hoffmann L, Lambert PF. Integration of human papillomavirus type 16 into the human genome correlate with a selective growth advantage of cells. *J Virol.* 1995; 69: 2989-97.
 30. Tornesello ML, Buonaguro FM, Buonaguro L, Salatiello I, Beth-Giraldo E, Giraldo G. Identification and functional analysis of sequence rearrangements in the long control region of human papillomavirus type 16 Af-1 variants isolated from Ugandan penile carcinomas. *J Gen Virol.* 2000; 81: 2969-82.
 31. Pett MR, Alazawi WO, Roberts I, Downen S, Smith DI, Stanley MA, Coleman N. Acquisition of high-level chromosomal instability is associated with integration of human papillomavirus type 16 in cervical keratinocytes. *Cancer Res.* 2004; 64: 1359-68.
 32. Tornesello ML, Buonaguro FM, Meglio A, Buonaguro L, Beth-Giraldo E, Giraldo G. Sequence variations and viral genomic state of human papillomavirus type 16 in penile carcinomas from Ugandan patients. *J Gen Virol.* 1997; 78: 2199-208.
 33. Zheng ZM, Baker CC. Papillomavirus genome structure, expression, and post-transcriptional regulation. *Front Biosci.* 2006; 11: 2286-302.
 34. Filippova M, Johnson MM, Bautista M, Filippov V, Fodor N, Tungteakkhun SS, Williams K, Duerksen-Hughes PJ. The large and small isoforms of human papillomavirus type 16 E6 bind to and differentially affect procaspase 8 stability and activity. *J Virol.* 2007; 81: 4116-29.
 35. Filippova M, Evans W, Aragon R, Filippov V, Williams VM, Hong L, Reeves ME, Duerksen-Hughes P. The small splice variant of HPV16 E6, E6, reduces tumor formation in cervical carcinoma xenografts. *Virology.* 2014; 450-451: 153-64.
 36. Pastuszek-Lewandoska D, Bartosinska-Dyc A, Migdalska-Sek M, Czarnecka KH, Nawrot E, Domanska D, Szylllo K, Brzezianska E. HPV16 E6*II gene expression in intraepithelial cervical lesions as an indicator of neoplastic grade: a pilot study. *Med Oncol.* 2014; 31: 842.
 37. Johansson C, Schwartz S. Regulation of human papillomavirus gene expression by splicing and polyadenylation. *Nat Rev Microbiol.* 2013; 11: 239-51.
 38. McFarlane M, MacDonald AI, Stevenson A, Graham SV.

- Human Papillomavirus 16 Oncoprotein Expression Is Controlled by the Cellular Splicing Factor SRSF2 (SC35). *J Virol.* 2015; 89: 5276-87.
39. Bouvard V, Baan R, Straif K, Grosse Y, Secretan B, El GF, Benbrahim-Tallaa L, Guha N, Freeman C, Galichet L, Cogliano V. A review of human carcinogens—Part B: biological agents. *Lancet Oncol.* 2009; 10: 321-2.
 40. Stein AP, Saha S, Kraninger JL, Swick AD, Yu M, Lambert PF, Kimple RJ. Prevalence of Human Papillomavirus in Oropharyngeal Cancer: A Systematic Review. *Cancer J.* 2015; 21: 138-46.
 41. Carozzi FM, Tornesello ML, Burrioni E, Loquercio G, Carillo G, Angeloni C, Scalisi A, Macis R, Chini F, Buonaguro FM, Giorgi RP. Prevalence of Human Papillomavirus Types in High-Grade Cervical Intraepithelial Neoplasia and Cancer in Italy. *Cancer Epidemiol Biomarkers Prev.* 2010; 19: 2389-400.
 42. Beachler DC, Sugar EA, Margolick JB, Weber KM, Strickler HD, Wiley DJ, Cranston RD, Burk RD, Minkoff H, Reddy S, Xiao W, Guo Y, Gillison ML, et al. Risk factors for acquisition and clearance of oral human papillomavirus infection among HIV-infected and HIV-uninfected adults. *Am J Epidemiol.* 2015; 181: 40-53.
 43. Woodman CB, Collins SI, Young LS. The natural history of cervical HPV infection: unresolved issues. *Nat Rev Cancer.* 2007; 7: 11-22.
 44. Wanram S, Limpaboon T, Leelayuwat C, Yuenyao P, Guiney DG, Lulitanond V, Jearanaikoon P. The use of viral load as a surrogate marker in predicting disease progression for patients with early invasive cervical cancer with integrated human papillomavirus type 16. *Am J Obstet Gynecol.* 2009; 201: 79-7.
 45. Andersson S, Safari H, Mints M, Lewensohn-Fuchs I, Gyllensten U, Johansson B. Type distribution, viral load and integration status of high-risk human papillomaviruses in pre-stages of cervical cancer (CIN). *Br J Cancer.* 2005; 92: 2195-200.
 46. Wu Z, Qin Y, Yu L, Lin C, Wang H, Cui J, Liu B, Liao Y, Warren D, Zhang X, Chen W. The association between human papillomavirus (HPV) 16, HPV18, and other HR-HPV viral load and the histological classification of cervical lesions: Results from a large-scale cross-sectional study. *J Med Virol.* 2016.
 47. Dang J, Feng Q. HPV16 infection in oral cavity and oropharyngeal cancer patients. *J Oral Sci.* 2016; 58: 265-9.
 48. Hoffer D, Bohmer G, von WR, Neumann H, Halec G, Holzinger D, Dondog B, Gissmann L, Pawlita M, Schmitt M. HPV16 RNA patterns defined by novel high-throughput RT-qPCR as triage marker in HPV-based cervical cancer precursor screening. *Gynecol Oncol.* 2015; 138: 676-82.
 49. Schmitt M, Dalstein V, Waterboer T, Clavel C, Gissmann L, Pawlita M. Diagnosing cervical cancer and high-grade precursors by HPV16 transcription patterns. *Cancer Res.* 2010; 70: 249-56.
 50. Holzinger D, Schmitt M, Dyckhoff G, Benner A, Pawlita M, Bosch FX. Viral RNA patterns and high viral load reliably define oropharynx carcinomas with active HPV16 involvement. *Cancer Res.* 2012; 72: 4993-5003.
 51. Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez BY, Xiao W, Kim E, Jiang B, Goodman MT, Sibug-Saber M, Cozen W, Liu L, Lynch CF, Wentzensen N, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. *J Clin Oncol.* 2011; 29: 4294-301.
 52. Tornesello ML, Buonaguro L, Giorgi-Rossi P, Buonaguro FM. Viral and cellular biomarkers in the diagnosis of cervical intraepithelial neoplasia and cancer. *Biomed Res Int.* 2013; 2013: 519619.
 53. Braakhuis BJ, Snijders PJ, Keune WJ, Meijer CJ, Ruijter-Schippers HJ, Leemans CR, Brakenhoff RH. Genetic patterns in head and neck cancers that contain or lack transcriptionally active human papillomavirus. *J Natl Cancer Inst.* 2004; 96: 998-1006.
 54. Ndiaye C, Mena M, Alemany L, Arbyn M, Castellsague X, Laporte L, Bosch FX, de Sanjosé S, Trottier H. HPV DNA, E6/E7 mRNA, and p16INK4a detection in head and neck cancers: a systematic review and meta-analysis. *Lancet Oncol.* 2014; 15: 1319-31.
 55. Snijders PJ, van den Brule AJ, Meijer CJ. The clinical relevance of human papillomavirus testing: relationship between analytical and clinical sensitivity. *J Pathol.* 2003; 201: 1-6.
 56. Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tan PF, Westra WH, Chung CH, Jordan RC, Lu C, Kim H, Axelrod R, Silverman CC, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med.* 2010; 363: 24-35.
 57. Tornesello ML, Perri F, Buonaguro L, Ionna F, Buonaguro FM, Caponigro F. HPV-related oropharyngeal cancers: From pathogenesis to new therapeutic approaches. *Cancer Lett.* 2014.
 58. de Boer MA, Jordanova ES, Kenter GG, Peters AA, Corver WE, Trimpos JB, Fleuren GJ. High human papillomavirus oncogene mRNA expression and not viral DNA load is associated with poor prognosis in cervical cancer patients. *Clin Cancer Res.* 2007; 13: 132-8.
 59. Rajeevan MS, Swan DC, Nisenbaum R, Lee DR, Vernon SD, Ruffin MT, Horowitz IR, Flowers LC, Kmak D, Tadros T, Birdsong G, Husain M, Srivastava S, et al. Epidemiologic and viral factors associated with cervical neoplasia in HPV-16-positive women. *Int J Cancer.* 2005; 115: 114-20.
 60. Pim D, Tomaic V, Banks L. The human papillomavirus (HPV) E6* proteins from high-risk, mucosal HPVs can direct degradation of cellular proteins in the absence of full-length E6 protein. *J Virol.* 2009; 83: 9863-74.
 61. Storrs CH, Silverstein SJ. PATJ, a tight junction-associated PDZ protein, is a novel degradation target of high-risk human papillomavirus E6 and the alternatively spliced isoform 18 E6. *J Virol.* 2007; 81: 4080-90.

62. Mantovani F, Banks L. The human papillomavirus E6 protein and its contribution to malignant progression. *Oncogene*. 2001; 20: 7874-87. doi: 10.1038/sj.onc.1204869.
63. Zheng ZM, Tao M, Yamanegi K, Bodaghi S, Xiao W. Splicing of a cap-proximal human Papillomavirus 16 E6E7 intron promotes E7 expression, but can be restrained by distance of the intron from its RNA 5' cap. *J Mol Biol*. 2004; 337: 1091-108.
64. Ajiro M, Zheng ZM. E6^{E7}, a novel splice isoform protein of human papillomavirus 16, stabilizes viral E6 and E7 oncoproteins via HSP90 and GRP78. *MBio*. 2015; 6: e02068-14.
65. Lin K, Lu X, Chen J, Zou R, Zhang L, Xue X. E6-associated transcription patterns in human papilloma virus 16-positive cervical tissues. *Oncol Lett*. 2015; 9: 478-82.
66. Evans W, Philippova M, Filippov V, Bashkirova S, Zhang G, Reeves ME, Duerksen-Hughes P. Overexpression of HPV16 E6* Alters beta-Integrin and Mitochondrial Dysfunction Pathways in Cervical Cancer Cells. *Cancer Genomics Proteomics*. 2016; 13: 259-73.
67. Ajiro M, Jia R, Zhang L, Liu X, Zheng ZM. Intron definition and a branch site adenosine at nt 385 control RNA splicing of HPV16 E6*I and E7 expression. *PLoS One*. 2012; 7: e46412.
68. Kosel S, Burggraf S, Engelhardt W, Olgemoller B. Increased levels of HPV16 E6*I transcripts in high-grade cervical cytology and histology (CIN II+) detected by rapid real-time RT-PCR amplification. *Cytopathology*. 2007; 18: 290-9.
69. Hong A, Zhang X, Jones D, Zhang M, Lee CS, Lyons JG, Veillard AS, Rose B. E6 viral protein ratio correlates with outcomes in human papillomavirus related oropharyngeal cancer. *Cancer Biol Ther*. 2016; 17: 181-7.
70. Chung CH, Gillison ML. Human papillomavirus in head and neck cancer: its role in pathogenesis and clinical implications. *Clin Cancer Res*. 2009; 15: 6758-62.
71. Gillison ML, Koch WM, Capone RB, Spafford M, Westra WH, Wu L, Zahurak ML, Daniel RW, Viglione M, Symer DE, Shah KV, Sidransky D. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *J Natl Cancer Inst*. 2000; 92: 709-20.
72. Rosty C, Sheffer M, Tsafirir D, Stransky N, Tsafirir I, Peter M, de CP, de La RA, Salmon R, Dorval T, Thierry JP, Couturier J, Radvanyi F, et al. Identification of a proliferation gene cluster associated with HPV E6/E7 expression level and viral DNA load in invasive cervical carcinoma. *Oncogene*. 2005; 24: 7094-104.
73. Um SH, Mundi N, Yoo J, Palma DA, Fung K, MacNeil D, Wehrli B, Mymryk JS, Barrett JW, Nichols AC. Variable expression of the forgotten oncogene E5 in HPV-positive oropharyngeal cancer. *J Clin Virol*. 2014; 61: 94-100.
74. Das P, Thomas A, Kannan S, Deodhar K, Shrivastava SK, Mahantshetty U, Mulherkar R. Human papillomavirus (HPV) genome status & cervical cancer outcome—A retrospective study. *Indian J Med Res*. 2015; 142: 525-32.
75. zur Hausen H. Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer*. 2002; 2: 342-50.
76. Parfenov M, Pedamallu CS, Gehlenborg N, Freeman SS, Danilova L, Bristow CA, Lee S, Hadjipanayis AG, Ivanova EV, Wilkerson MD, Protopopov A, Yang L, Seth S, et al. Characterization of HPV and host genome interactions in primary head and neck cancers. *Proc Natl Acad Sci U S A*. 2014; 111: 15544-9.
77. Olthof NC, Huebbers CU, Kolligs J, Henfling M, Ramaekers FC, Cornet I, van Lent-Albrechts JA, Stegmann AP, Silling S, Wieland U, Carey TE, Walline HM, Gollin SM, et al. Viral load, gene expression and mapping of viral integration sites in HPV16-associated HNSCC cell lines. *Int J Cancer*. 2015; 136: E207-E218.
78. Annunziata C, Buonaguro L, Losito S, Buonaguro FM, Tornesello ML. Somatic mutations of STK11 gene in human papillomavirus positive and negative penile cancer. *Infect Agent Cancer*. 2013; 8: 2.
79. Resnick RM, Cornelissen MT, Wright DK, Eichinger GH, Fox HS, ter Schegget J, Manos MM. Detection and typing of human papillomavirus in archival cervical cancer specimens by DNA amplification with consensus primers. *J Natl Cancer Inst*. 1990; 82: 1477-84.
80. Tortora M, Annunziata C, Liguori G, Losito S, Botti G, Greggi S, Buonaguro L, Buonaguro FM, Tornesello ML. Detection of human papillomavirus DNA in peri-tumor tissues and pelvic lymph nodes as potential molecular marker of micrometastasis in cervical cancer. *Infect Agent Cancer*. 2016; 11: 22.
81. de Roda Husman AM, Walboomers JM, Meijer CJ, Risse EK, Schipper ME, Helmerhorst TM, Bleker OP, Delius H, van den Brule AJ, Snijders PJ. Analysis of cytologically abnormal cervical scrapes for the presence of 27 mucosotropic human papillomavirus genotypes, using polymerase chain reaction. *Int J Cancer*. 1999; 56: 802-6.
82. Tornesello ML, Duraturo ML, Losito S, Botti G, Pilotti S, Stefanon B, De PG, Gallo A, Buonaguro L, Buonaguro FM. Human papillomavirus genotypes and HPV16 variants in penile carcinoma. *Int J Cancer*. 2008; 122: 132-7.