


REVIEW

Open Access



# Exploring the roles of HPV16 variants in head and neck squamous cell carcinoma: current challenges and opportunities

Daniela Cochicho<sup>1,2†</sup>, Rui Gil da Costa<sup>3,4,5†</sup> and Ana Felix<sup>1,6\*†</sup> 

## Abstract

The incidence of squamous cell carcinomas of the head and neck (HNSCC) is consistently increasing, in association with human papillomavirus (HPV) infection, especially HPV16. HPV variants show heterogeneity in the pathogenicity of cervical cancer, but little has been established about their relevance on HNSCC. This review addresses the distribution of HPV16 variants in HNSCC and their potential contribution to clinical practice. A search was performed in PubMed using the keywords HNSCC HPV16 variants. Sixty articles were identified between 2000 and 2020 and 9 articles were selected for a systematic analysis. Clinical cohorts comprised 4 to 253 patients aged between 17 and 91 years with confirmed HPV16-positive HNSCC. Samples were collected from fresh biopsies of the tumour, oral rinse or formol fixed/paraffin embedded tissue, from the oral cavity, oropharynx, hypopharynx, larynx and Waldeyer's tonsillar ring. HPV16 variants were identified using Sanger sequencing techniques. Seven studies addressed the HPV16 E6 gene, one studied E6 and E7, another studied L1 and one focused on the long control region. European variants represent 25–95%, Asian-American 5–57% and African 2–4% of the total isolates, suggesting a marked predominance of European strains. No correlations could be drawn with patient prognosis, partly because many studies relied on small patient cohorts. Additional studies are needed, particularly those employing next generation sequencing techniques (NGS), which will allow faster and accurate analysis of large numbers of samples.

**Keywords:** HNSCC, HPV16, HPV16 variants, Next generation sequencing

## Background

Head and neck cancer is the sixth leading cancer by incidence worldwide [1] and comprises many different pathological entities. The diagnosis is often characterized by multifocal development and presentation at an advanced stage. [2] The age at diagnosis usually ranges between 50 and 70 years of age and men are significantly more likely to develop the disease than women (up to a 4:1 proportion, depending on geographical localization)

[3]. Currently, two major carcinogenic pathways leading to head and neck squamous cell carcinomas are recognized: the first is associated with risk factors like smoking and abusive alcohol consumption while the second is associated with human papillomavirus (HPV) infection [4, 5]. Over the past decade, several epidemiologic studies reported a 36.5% increase in the incidence of head and neck squamous cell carcinomas (HNSCC), particularly in high-income countries and among men < 60 years [4, 6–8]. Available data indicates that these changes specifically involve oropharyngeal cancers [9, 10]. Although tobacco consumption has decreased, the incidence of HPV-positive oropharyngeal cancers has increased [11], indicating that HPV infection is the underlying cause for the overall increase in HNSCC incidence [12–14].

\*Correspondence: ana.felix@nms.unl.pt

†Daniela Cochicho, Rui Gil da Costa and Ana Felix contributed equally to this work

<sup>1</sup> NOVA Medical School, NOVA University of Lisbon, CEDOC, Campo

Mártires da Pátria 130, 1169-056 Lisbon, Portugal

Full list of author information is available at the end of the article



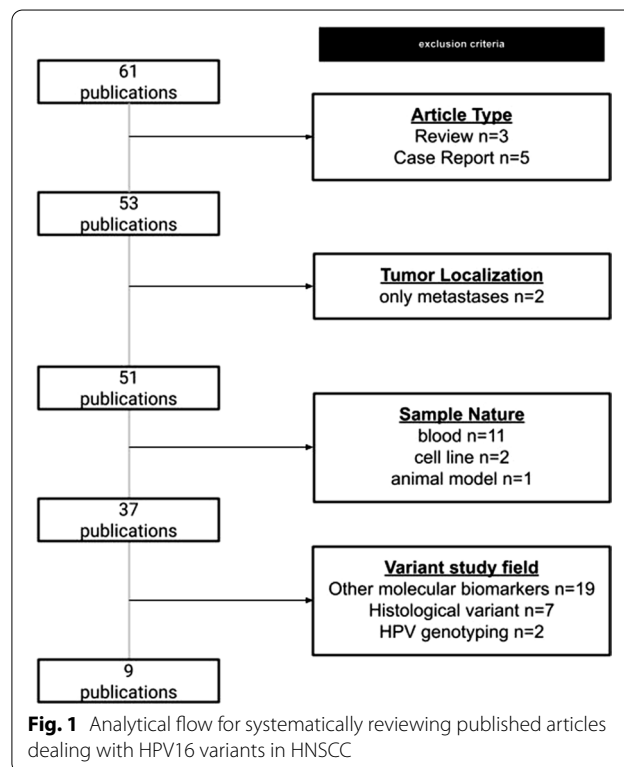
Patients showing oral HPV infection are 53 times more likely to develop HNSCC [15]. Presently, HPV is present in approximately 35% of HNSCC and most HPV-positive cases emerge in lingual and palatine tonsils [16]. The presence of HPV DNA in tumour cells defines a specific pathologic entity within HNSCC with specific epidemiology, molecular characteristics, and biological behaviour [17]. HPV interferes with key signaling pathways to promote carcinogenesis via its viral oncoproteins E6 and E7, which lead to the inactivation of tumour protein 53 (p53) and the retinoblastoma protein (pRB), respectively [18]. HPV is a group of small, double-stranded DNA viruses that infect the epidermis and keratinizing mucosae. More than 100 HPV types have been identified [19, 20]. Presently, 40 different HPV types are known to infect mucosal epithelia and are categorized into low-risk and high-risk HPV types according to their epidemiologic association with cervical cancer [21]. In HNSCC, the majority of HPV types belong to this group including HPV16, HPV18, HPV39 and HPV45 [22]. HPV16 is by far, the most detected type, accounting for 90% of all HPV-positive HNSCC cases [23], a significantly greater proportion than in cervical cancer where it accounts for little over 50% of cases [24]. Sequencing of the HPV genome also revealed intra-type variants with genetic differences ranging between 0.5 and 1% [25, 26]. HPV variants arise mainly from nucleotide substitutions in some restricted positions in the genome coding region or in the noncoding region [27, 28]. The prevalence of variants for each HPV type varies significantly in different geographical areas [27]. Regarding HPV16, whole-genome analysis allowed the characterization of five distinct phylogenetic clusters named according to their original geographical distribution: European (E), Asian (As), Asian-American (AA), African 1 (Af1) and African 2 (Af2) [29–31]. Subsequently, a new branch, North American 1 (NA1) was identified [32, 33]. Currently, emerging epidemiological, etiological, and molecular data suggest that intra-type HPV variants are biologically distinct and may be associated with different risk of cervical cancer progression [28]. *In vitro* studies using 3D organotypic epithelial cell cultures showed that HPV16 E6 variants differ in their ability to abolish keratinocyte differentiation and to induce p53 degradation [34–36]. Those experimental results are corroborated by clinical studies on cervical cancer. The T350G (L83V) HPV16 variant is the most frequently found among invasive cervical cancers [21, 37] and has been linked to a higher oncogenic potential than the prototype [38], possibly by facilitating persistent viral infection, a critical factor for cancer development [39–41]. Despite the growing number of studies concerning HPV16 variants on cervical cancer, there are currently very few reports describing their distribution

in HNSCC and nothing is established about their impact on the development, treatment response and impact on disease outcome. Nonetheless, HPV16 variants may also play an important role in head and neck carcinogenesis and studying their impact is needed due to the rising incidence of HPV-positive HNSCC [18, 23, 25, 42]. The purpose of this work is to systematically review the distribution of HPV16 variants in HNSCC and to assess the available knowledge concerning their potential contribution for HNSCC pathologic heterogeneity in published studies.

## Main text

### Review compilation data

A systematic review was performed on the PubMed electronic database (<https://pubmed.ncbi.nlm.nih.gov>) using the following keywords search criteria: HNSCC HPV16 variants. Observational studies reporting the distribution of HPV16 variants in HNSCC and published between 2000 and 2020 were included. Review articles, case reports and articles dealing with other types of cancer were excluded (Fig. 1). The selection search criteria were based on tumour localization, nature of sample and background of the variant study field. The studies were independently assessed to identify the prevalence of HPV16 variants in HNSCC and evaluate in respective cohorts for correlation with clinicopathological parameters. HNSCC



patients were categorized into different sites (oral cavity and oropharynx) and further into sub cohorts based on the demographic and clinical information provided. The published data were summarized using frequencies and percentages for HPV16 variants, stratified by tumour location, types of tumour samples, patient cohort size, sequencing methodology, and clinical variables such as age and gender (Fig. 2.)

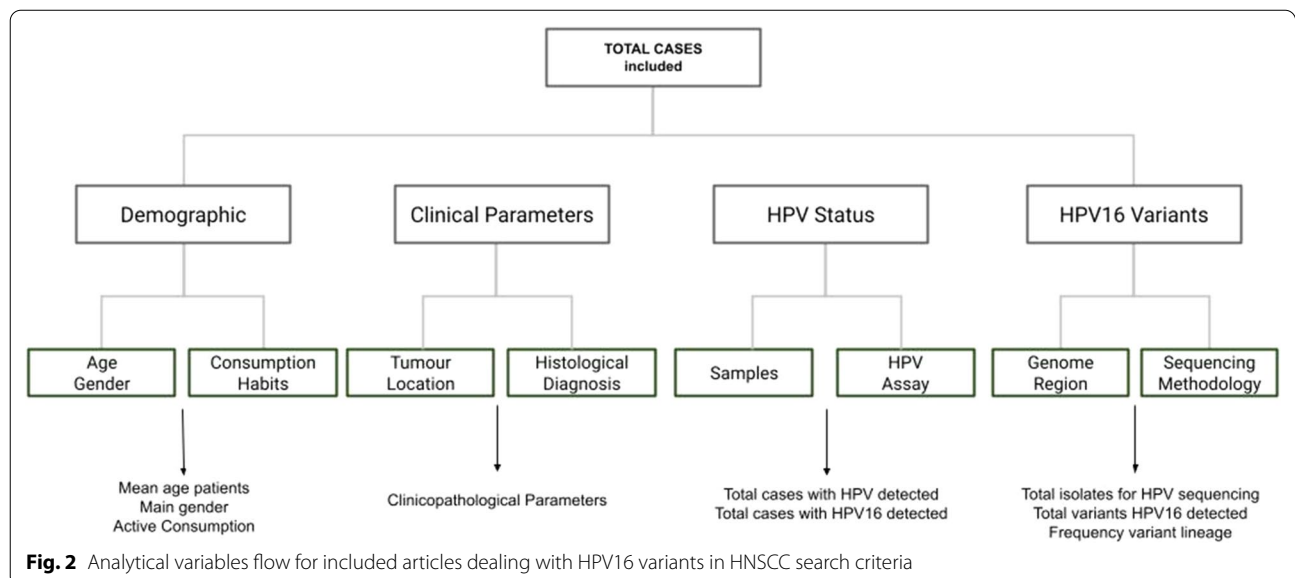
**General findings**

From 61 records found using the search criteria HNSCC HPV16 DNA cell carcinoma variants frequency patients’ cohorts, we selected 9 studies while 52 studies were excluded from further analysis, including review articles and papers which fit the exclusion criteria (Fig. 1). Selected articles report data from a total of 945 patients in the period between 2000 and 2020, distributed by 6 countries. (Table 1) Three studies were performed in Europe, three in the USA, two in Asia and Middle East (Japan and Iran) and one in South America (Brazil). HPV16 variants were evaluated in nine different clinical cohorts. All studies were different and did not contain the evaluation of the same cases. Overall, the study population consisted of patients whose age at diagnosis ranged between 17 and 91 years. The gender ratio was 1:3 or 1:4 (female to male), and all cases were histologically confirmed as HNSCC. The specimens used to evaluate HPV16 were primary tumours located in the oral cavity, oropharynx, hypopharynx, larynx and Waldeyer’s tonsillar ring. Regional lymph node metastases were also included in one study. The samples used to analyze were diverse and consisted of fresh tissue, oral rinse, or formalin fixed/paraffin embedded tissue.

The use of the p16INK4a marker is to clinically detect an oncogenically active HPV infection and is considered as a surrogate marker for HPV infection [43, 44]. From the 9 studies evaluated three described the expression pattern of p16INK4a among negative HNSCC HPV(−) and HNSCC HPV16(+) cases. For all studies there was a significant difference in p16INK4a expression between HNSCC HPV(−) and HNSCC HPV16(+). Pakdel et al. 2020 observed the expression of p16INK4a among HPV variants and showed that a strong p16INK4a expression was present in poorly differentiated tumour tissues infected with HPV16 sublineage A2 [45].

**HPV detection and genotyping**

HPV was detected using PCR-based techniques employing consensus degenerate primers: MY09/11 (4 studies) MY09/11 and/or GP05/06 (2 studies), PGMV (one study) and SPF1/2 (one study); all of which are complementary to the conserved L1 region. One study detected HPV using in situ hybridization. HPV genotyping was performed using different methods. Two studies using a commercial kit (INNO LIPA) which identifies 28 different HPV types (6, 11, 16, 18, 26, 31, 33, 35, 39, 40, 43, 44, 45, 51, 52, 53, 54, 56, 58, 59, 66, 68, 69, 70, 71, 73, 74, 82) by reverse blot hybridization. Two other studies used TaqMan PCR methods targeting the E6 or E6/LCR. One study used restriction fragment length polymorphism (RFLP) analysis. Two studies performed direct Sanger sequencing of the PCR products. The HPV frequency in the different cohorts ranged from 10 to 100% of the cases, and the HPV16 distribution ranged between 67 and 100% of all HPV positive cases (Table 1). Except the study done in 2020 in



**Table 1** Identification of HPV16 variants in HNSCC in the literature (PubMed 2000–2020)

Year publication	Reference	Country	Mean age patients (yrs)	Gender (% male)	Consumption (% active)	Samples	Total cases (n)	Total cases HPV detected % (n)	Total cases HPV16 detected % (n)	Total isolates for HPV sequencing	Genome region location	Amplicon size (pb)	Sequencing
2000	[48]	USA	63	no data available	87%	Fresh tissue	253	24% (62)	90% (56)	52	E6	455	Sanger
2004	[49]	Germany	no data available	no data available	no data available	FF/PE	24	100% (24)	100% (24)	21	E6 and E7	793	Sanger
2007	[50]	Italy	63	74%	67%	Fresh tissue FF/PE	115	18% (21)	67% (14/21)	13	L1	150	Sanger
2008	[51]	USA	57	77%	76%	Oral rinse FF/PE	135	32% (44)	100% (44/44)	19	E6	609	Sanger
2009	[52]	Japan	64	86%	79%	Fresh tissue	77	10% (8)	100% (8/8)	8	E6	323	Sanger
2013	[53]	Brazil	59	100%	50%	FF/PE	4	100% (4)	100% (4/4)	4	E6	no data available	Sanger
2015	[54]	Italy	65	88%	63%	FF/PE	24	100% (10)	100% (10/10)	10	E6	no data available	Sanger
2016	[55]	USA	60	86%	No data available	FF/PE	205	18% (36)	70% (25)	21	LCR	193	Sanger
2020	[56]	Iran	56	58%	No data available	FF/PE	108	23% (25)	16% (17)	13	E6	no data available	Sanger

E6 Early gene 6; E7 early gene 7; LCR long control region; L1 late gene 1; FF/PE Formalin fixed Paraffin embedded

Iran, where just 16% of the HPV detected were HPV16. The most common HPV in their series was HPV16 but also followed by HPV18 and HPV11 and the cases were almost located in Oral cavity and Larynx; only 6 tonsils in all 108 cases were evaluated, explaining the low HPV frequency overall.

#### Identification HPV16 variants

Of the 9 studies, seven studies identified HPV16 variants by sequencing the *E6* gene, one addressed variants in the *L1* gene and another the long control region. In one study the *E7* gene was also sequenced along with *E6* (Table 1). In all studies, the identification of HPV16 variants was performed by Sanger sequencing techniques and by comparison to the reference sequence (prototype). Next generation sequencing (NGS) was not used in any of the studies evaluated. However, the reliance on conventional Sanger sequencing is likely to limit the number of cases that can be analyzed and the reliability of results. [26, 46, 47] Sanger sequencing techniques show limitations concerning the size of amplicons, which may compromise the assay sensitivity. Additionally, to overcome polymerase errors and guarantee the integrity and robustness of the results, it is necessary to perform replicates for each case or vector molecular assays [48]. Such replicates add significantly to the workload and costs and constitute a limitation to analyze large patient cohorts. In fact, only two out of nine studies included over 200 HNSCC patients and none analyzed more than 52 HPV-positive patients for HPV16 variants. Future studies aiming to study larger patient cohorts are likely to benefit from

NGS techniques, which provide faster and reliable sequencing results of larger amplicons [25, 26, 46].

#### Prevalence of HPV16 variants

The frequency of each HPV16 variant (E, NA, AF, AS and AA) was estimated for each of the 9 studies included in the systematic review (Table 2). Gillison et al. [37] first reported data concerning the distribution of HPV16 *E6* variants in HNSCC. The authors evaluated 52 HPV16-positive patients from an overall HNSCC cohort comprising 259 patients. The age at diagnosis ranged between 17 and 91 years-old (median, 63 years-old) and the majority were smokers, with or without alcohol consumption (87%). Tumours were localized in the nasopharynx (n=2), oral cavity (n=84), oropharynx (n=60), hypopharynx (n=21), larynx (n=86). HPV-positive patients showed significantly improved disease outcomes compared with HPV-negative patients. All samples used were HNSCC fresh tumour specimens. The HPV16 variants were classified into the same phylogenetic group as the European prototype in 75% of cases, Asian in 17%, North American in 4.0% and African 1 in 4.0% of cases. Six novel variants not previously reported (E-G315T, E-G315G, E-C395G, E-A478T, E-A132T, Af1-C311, Af1-A389) were also identified. The authors remarked that the distribution of variants in HNSCC resembled that observed in cervical cancer. However, no conclusions were drawn concerning the potential contribution of specific HPV16 *E6* variants for increasing cancer risk or modifying tumour biopathology and prognosis. Four years later, Hoffmann et al. [38] identified HPV16 variants in 7 out of 21 tumour specimens of HNSCC. The authors analyzed the *E6* and *E7* ORFs. Altogether, the DNA samples carried HPV16 prototype European

**Table 2** Frequency of HPV16 variants

Reference	Total cases HPV16 detected (n)	Total isolates HPV sequencing (n)	HPV16 variant region location	European (E)		North American (NA) (n)	African (AF1 AF2) (n)	Asian (AS) (n)	Asia American (AA) (n)
				All lineages (n)	E-350-G (n)				
[48]	56	52	E6	39	6	2	2	9	0
[49]	24	21	E6 E7	15	8	0	0	0	0
[50]	14	13	L1	9	0	1	2	0	1
[51]	44	19	E6	18	4	0	0	1	0
[52]	8	8	E6	5	5	0	0	0	0
[53]	4	4	E6	2	1	0	0	0	2
[54]	10	10	E6	8	8	0	0	0	0
[55]	25	21	LCR	9	–	0	0	0	12
[56]	17	13	E6	11	–	0	0	0	2

(–) not evaluated; AA Asia-American HPV16 Variant; AF African HPV16 Variant; AS Asian HPV16 variant; E European HPV16 variant; E6 HPV Early protein 6; E7 HPV Early protein 7; LCR long control region; L1 late gene; NA North American HPV16 Variant

variant (29%), the European variant T350G (38%) and 33% Euro-German variant (A131G+C712A). Again, no conclusions could be drawn concerning the clinical-pathological relevance of the HPV16 variants identified, partly because of the cohort's small size. Badaracco et al. [49], published molecular analysis data on the HPV16 *L1* ORF, from an Italian cohort (total  $n=115$ , of which only 13 were tested for HPV variants), composed of 86 men and 29 women, with a mean age of 63.21 years old. A high percentage of patients were smokers (67%) and alcohol drinkers 43%. Tumours were localized mostly in the oral cavity ( $n=60$ ), followed by the larynx ( $n=30$ ), the oropharynx ( $n=10$ ), the tonsil area ( $n=8$ ), the hypopharynx ( $n=5$ ) and the sinus/nose ( $n=2$ ). In this study, the presence of HPV was not significantly associated with disease-free survival at 2 years. Sixty nine percent of the cases (13 cases) analyzed showed European-German variants (9 cases), 15% were African type 2 (2 cases), 8% Asian-American (1 case), and the remaining 8% (1 case) had an unclassified variant. The predominance of the European variant was unsurprising considering the Italian origin of the patients. No correlations were drawn between the presence of HPV16 variants and any epidemiological, pathological, or clinical data. Agrwal et al. [50], studied 19 HPV16 isolates from a universe of 135 HNSCC samples. The median age at diagnosis was 57 years-old, 77% of the patients were men, most patients had a history of smoking ( $n=62$  smokers versus 51 non-smokers) but were non-drinkers (73 non-drinkers versus 41 drinkers). All analyzed cases showed European variants and a single case carried an Asian variant. The most common European variant was E-350 T ( $n=6$ ), followed by E-350G ( $n=4$ ) and E-T131G ( $n=2$ ). Importantly, eight of the 19 isolates contained European variants with sequences unique to a single individual. Boscolo-Rizzo et al. [51] analyzed HPV E6 variants in a short ( $n=8$ ) case series. The authors showed the presence of the T350G mutation in 5 cases located in the oral cavity and oropharynx, while three tumours located in the larynx and hypopharynx contained HPV 16 prototype sequences. Joseph et al. [52] compared HPV16 variants present in four patients with bilateral tonsillar HNSCC. Two cases carried European variants while two others carried Asian-American variants. The results show that, in all 4 patients, the same HPV16 variant was present in the bilateral tumours, supporting the hypothesis that a single HPV infection, rather than independent infections with distinct agents, is responsible for those bilateral tumours. Hassani et al. [53] studied the HPV16 variants present in 10 cases of tonsillar HNSCC. The authors found that the E-350G-variant was present in 80% of cases while the European prototype was identified in the other 20%. Again, this short case series did not provide

data concerning the pathobiological relevance of HPV16 variants. In the following year, Betiol et al. [54] reported the distribution of HPV16 variants in a Brazilian cohort of 21 HNSCC patients. The authors analyzed the HPV16 LCR and found that 12 (57.1%) patients carried European and 9 carried Asian-American (42.9%) variants. The authors remarked that the slight predominance of European variants accompanied observation from their normal cervical samples, suggesting that the distribution of HPV16 variants reflects the overall frequency in each studied population. The most recent study was published by Pakdel et al. [45]. Thirteen HNSCC tissue specimens tested positive for HPV16 using overlapping PCR assays and were analyzed for the presence of *E6* variants. There was a marked predominance of European variants (84.6%) followed by Asian-American (15.4%) variants.

## Discussion

We reviewed the published literature on HPV16 variants in HNSCC, aiming to evaluate the biological meaning of those variants in this particular location. The clinical management of HNSCC improved greatly since the recognition of HPV-positive and HPV-negative lesions. Although tumour recurrence still occurs in 10–20% of HPV-positive HNSCC patients, the majority of these patients clearly benefit from therapeutic de-escalation [55]. The lack of adequate biomarkers to define more tailored approaches is still necessary in SCC HPV associated tumours [56]. The large majority of these tumours are associated with HPV16 and a deeper understanding of the bio pathological implications of distinct HPV16 variants would contribute to the molecular characterization of HPV-positive HNSCC and may help to define patient's subgroups that would benefit from specific therapeutic approaches. Data from cervical cancer patients showed an association between the presence of non-European HPV16 lineages, a longer viral persistence [57, 58] and an increased risk of developing high-grade cervical intraepithelial neoplasia [59]. Apparently, within individuals, HPV genomes harbour high levels of variability in HPV16 genome sequence upon normal to pre-cancer/cancer [60] revealing several fundamental discoveries and suggesting a paradigm shift from HPV16 as a single viral entity to theorize each HPV16 isolated to be a separate virus with distinct carcinogenic potential [61]. This would imply that within HPV16, the genetic variation partly predicts the risk of pre-cancer and cancer [62]. In particular, specific sublineages (A4, C, D2, and D3) have shown a significantly increased risk compared to the most common A1/A2 sublineages [26] and assured D2, for the strongest risk of cancer within glandular epithelium (adenocarcinomas) [26].

The major data from whole-sequences obtained from individual clinical specimens agrees that the genetic variation occurs more commonly on low-grade or benign HPV16 infections [61] and corrected explained In case–control analyses describing the highest amino acid changing variants HPV16 in the controls throughout the genome with cervix cancer. [62] Therefore, E7 oncogene lacks nonsynonymous (amino acid changing) variants in cervical cancers, suggesting the E7 conservation is admitted for carcinogenicity. [61] The specific conservation of the 98 amino acids of E7, that directly disrupts Rb function, was shown to be crucial for trigger carcinogenesis, owing to be commented as a highly specific target for etiologic and therapeutic research [26].

Presenting data from a cohort study that compares genomic characteristics of HPV associated with cervical versus oropharyngeal tumours using DNA sequence analysis showed no significant differences between distribution in HPV16 variants [24], both presenting major prevalence of the European variant. Instead, the HPV E6 gene amplified from oropharyngeal samples reported over more nonsynonymous mutations, but also for the E7 gene, no differences were found in mutation rates between the two anatomical locations [24, 62]. Notably, the E7 gene is conserved in both locations corroborating the recent findings on cervical cancer studies. The apparent restriction on E7 mutations seems to be present in the oropharynx, as well. Indeed, described all both share common biological features, but the important differences present in HPV-genome may explain their distinct pathophysiological mechanisms and susceptibility to treatment. Nevertheless, the invariability of E7 presents an attractive potential target for therapy at both locations. [24]

In all the studies evaluated in this analysis, the most prevalent variant was the European Variant, which belongs to the European lineage and within this lineage the E-350-G HPV16 variant was the most frequent in majority of the studies sequencing based E6 and/or E7 regions with 29% and 53% respectively, although the

cohorts had or not an European origin, (cohorts were from Brazil, United States of America, Japan, Italy and Germany (Table 3).

As it was established from molecular analysis studies in cervical cancer, that the European variant T350G, was the variant frequently found in cervical intraepithelial neoplasms and cancers, and has been associated with progression to cervical cancer particularly in North European women. The detection of the T350G variant in a large proportion of HNSCC patients, therefore, indicates that this variant might also play an important role in HN carcinogenesis. Hassani et al. in 2015 supported this hypothesis, elucidating that the HPV-16 E-350G variant has a polymorphism in residue 83, a leucine for valine (L83V), probably responsible for the increased cancer risk [53].

### Conclusions

In HNSCC different lineages of HPV16 variants can be identified and differ geographically. Although most reviews described only the distribution of HPV16 variants in HNSCC, LeConte in a more detailed analysis of those studies found important differences. They found the distribution of two HPV16 variant groups differ significantly in oropharyngeal cancer and cervical cancer. The European + South America (E + AS) variant groups showed a higher prevalence in the oropharyngeal samples, representing 90.2%, than in cervical carcinomas (71.4%). Moreover, the Asia-American (AA1 + AA2) variant groups were present in 22.5% of cervical cancers in contrast to 4.4% in the oropharyngeal cancers [24]. The number of cases studied does not allow us to explore the importance of those differences nor to understand the potential role determining the clinical behaviour and potential use to select treatment and prognosis. Hopefully, the new sequencing era will enrich the study of HPV and related cancers. The advances in HPV whole-genome sequencing [46] provided technically achievable large-scale longitudinal studies on HPV whole-genomic sequences and promotes an exhaustive understanding of

**Table 3** Frequency HPV16 Variants by region genome location

Reference	Total number of cases	Total isolates for HPV sequencing	HPV16 variant region location	European		North American (NA)	African (AF1 AF2)	Asian (AS)	Asia American (AA)
				All lineages	E-350-G				
[50]	115	115	L1	69%	Not Detected	-	15%	-	8%
[48, 51–54, 56]	621	106	E6	78%	29%	2%	2%	9%	4%
[49]	24	21	E6E7	71%	53%	-	-	-	-
[55]	205	21	LCR	42.8%	-	-	-	-	57.2%

(-) not evaluated; AA Asia-American HPV16 Variant; AF African HPV16 Variant; AS Asian HPV16 Variant; E European HPV16 variant; E6 HPV Early protein 6; E7 HPV Early protein 7; LCR long control region; L1 late gene; NA North American HPV16 variant

the viral genetic diversity within and between infected individuals and will make the link between variants and cancer risk, comprehensively [46, 61].

#### Abbreviations

AA: Asia-American HPV16 variant; AF: African HPV16 variant; AS: Asian HPV16 variant; E: European HPV16 variant; E6: HPV Early protein 6; E7: HPV Early protein 7; HNSCC: Head and neck squamous cell carcinoma; HPV: Human papillomavirus; NA: North American HPV16 VARIANT; NGS: Next generation sequencing.

#### Acknowledgements

Not applicable.

#### Authors' contributions

These authors contributed equally to this work. Study concepts: AF, DC, RGC. Study design: AF, DC, RGC. Data acquisition: DC, AF, RGC. Data analysis and interpretation: RGC, DC, AF. Manuscript preparation: DC. Manuscript editing: AF, DC, RGC. Manuscript review: AF, DC, RGC. All authors read and approved the final manuscript.

#### Funding

Not applicable.

#### Availability of data and materials

Not applicable.

#### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>NOVA Medical School, NOVA University of Lisbon, CEDOC, Campo Mártires da Pátria 130, 1169-056 Lisbon, Portugal. <sup>2</sup>Virology Laboratory IPOLFG, Rua Professor Lima Bastos, 1099-023 Lisbon, Portugal. <sup>3</sup>LEPABE, Laboratory for Process Engineering, Environment, Biotechnology and Energy, University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal. <sup>4</sup>Post-graduate Programme in Adult Health (PPGSAD), University Hospital (HUUFMA) and Morphology Department, Federal University of Maranhão, Av. dos Portugueses 1966 - Vila Bacanga, São Luís, MA 65080-805, Brazil. <sup>5</sup>Molecular Oncology and Viral Pathology Group, Research Center of IPO Porto (CI-IPOP) / RISE@CI-IPOP (Health Research Network), Portuguese Oncology Institute of Porto (IPO Porto) / Porto Comprehensive Cancer Center (Porto.CCC), Rua Dr. António Bernardino de Almeida, 4200-072 Porto, Portugal. <sup>6</sup>Pathology Department IPOLFG, Rua Professor Lima Bastos, 1099-023 Lisbon, Portugal.

Received: 27 July 2021 Accepted: 28 October 2021

Published online: 08 November 2021

#### References

1. Ferlay J, Soerjomataram I, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015. <https://doi.org/10.1002/ijc.29210>.
2. De Bree R, Leemans CR. Recent advances in surgery for head and neck cancer. *Curr Opin Oncol*. 2010. <https://doi.org/10.1097/CCO.0b013e3283380009>.
3. Bray F, Ferlay J, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018. <https://doi.org/10.3322/caac.21492>.
4. McDermott JD, Bowles DW. Epidemiology of head and neck squamous cell carcinomas: impact on staging and prevention strategies. *Curr Treat Options Oncol*. 2019. <https://doi.org/10.1007/s11864-019-0650-5>.
5. Gillison ML. Human papillomavirus-associated head and neck cancer is a distinct epidemiologic, clinical, and molecular entity. *Semin Oncol*. 2004. <https://doi.org/10.1053/j.seminoncol.2004.09.011>.
6. Chaturvedi AK, Anderson WF, et al. Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers. *J Clin Oncol*. 2013. <https://doi.org/10.1200/JCO.2013.50.3870>.
7. Hwang TZ, Hsiao JR, et al. Incidence trends of human papillomavirus related head and neck cancer in Taiwan. *Int J Can*. 2015. <https://doi.org/10.1002/ijc.29330>.
8. Mork J, Moller B, et al. Time trends in pharyngeal cancer incidence in Norway 1981–2005: a subsite analysis based on a re-abstraction and recoding of registered cases. *Can Causes Control*. 2010. <https://doi.org/10.1007/s10552-010-9567-9>.
9. Zumsteg ZS, Cook-Wiens G, et al. Incidence of oropharyngeal cancer among elderly patients in the United States. *JAMA Oncol*. 2016. <https://doi.org/10.1001/jamaoncol.2016.1804>.
10. Tota JE, Best AF, et al. Evolution of the oropharynx cancer epidemic in the United States: moderation of increasing incidence in younger individuals and shift in the burden to older individuals. *J Clin Oncol*. 2019. <https://doi.org/10.1200/JCO.19.00370>.
11. Faraji F, Eisele DW, et al. Emerging insights into recurrent and metastatic human papillomavirus-related oropharyngeal squamous cell carcinoma. *Laryngoscope Investig Otolaryngol*. 2017. <https://doi.org/10.1002/lio2.37>.
12. Chaturvedi AK, Engels EA, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. *J Clin Oncol*. 2011. <https://doi.org/10.1200/JCO.2011.36.4596>.
13. Hammarstedt L, Lindquist D, et al. Human papillomavirus as a risk factor for the increase in incidence of tonsillar cancer. *Int J Can*. 2006. <https://doi.org/10.1002/ijc.22177>.
14. Hocking JS, Stein A, et al. Head and neck cancer in Australia between 1982 and 2005 show increasing incidence of potentially HPV-associated oropharyngeal cancers. *Br J Can*. 2011. <https://doi.org/10.1038/sj.bjc.6606091>.
15. Gillison ML, D'Souza G, et al. Distinct risk factor profiles for human papillomavirus type 6-positive and human papillomavirus type 6-negative head and neck cancers. *J Natl Cancer Inst*. 2008. <https://doi.org/10.1093/jnci/djn025>.
16. El-Mofty SK. HPV-related squamous cell carcinoma variants in the head and neck. *Head Neck Pathol*. 2012;6:55. <https://doi.org/10.1007/s12105-012-0363-6>.
17. WHO, World Health Organization. International agency for research on cancer (IARC) monographs on the evaluation of carcinogenic risks to humans. 2007;90(Human Papilloma Virus):689. <https://monographs.iarc.who.int/wp-content/uploads/2018/06/mono90.pdf>
18. Polanska H, Raudenska M, et al. Clinical significance of head and neck squamous cell cancer biomarkers. *Oral Oncol*. 2014. <https://doi.org/10.1016/j.oraloncology.2013.12.008>.
19. De Villiers EM, Fauquet C, et al. Classification of papillomaviruses. In: *Virology*. Academic Press. 2004; <https://doi.org/10.1016/j.virol.2004.03.033>
20. De Villiers EM, Gunst K. Characterization of seven novel human papillomavirus types isolated from cutaneous tissue, but also present in mucosal lesions. *J Gen Virol*. 2009. <https://doi.org/10.1099/vir.0.011478-0>.
21. Muñoz N, Bosch FX, et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med*. 2003. <https://doi.org/10.1056/NEJMoa021641>.
22. Kreimer AR, Clifford GM, et al. Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systemic review. *Cancer Epidemiol Biomark Prev*. 2005. <https://doi.org/10.1158/1055-9965.EPI-04-0551>.
23. Huertas-Salgado A, Martín-Gómez DC, et al. E6 molecular variants of human papillomavirus HPV type 16: an updated and unified criterion for clustering and nomenclature. *Virology*. 2011. <https://doi.org/10.1016/j.virol.2010.10.039>.
24. LeConte BA, Szaniszló P, et al. Differences in the viral genome between HPV-positive cervical and oropharyngeal cancer. *PLOS ONE*. 2018. <https://doi.org/10.1371/journal.pone.0203403>.
25. Burk RD, Harari A, et al. Human papillomavirus genome variants. *Virology*. 2013. <https://doi.org/10.1016/j.virol.2013.07.018.Human>.



26. Mirabello L, Yeager M, et al. HPV16 sublineage associations with histology-specific cancer risk using HPV whole-genome sequences in 3200 women. *J Natl Cancer Inst*. 2016. <https://doi.org/10.1093/jnci/djw100>.
27. Sichero L, Simão Sobrinho J, et al. Oncogenic potential diverge among human papillomavirus type 16 natural variants. *Virology*. 2012. <https://doi.org/10.1016/j.virol.2012.06.011>(8).
28. Bernard HU, Calleja-Macias IE, et al. Genome variation of human papillomavirus types: phylogenetic and medical implications. *Int J Cancer*. 2006. <https://doi.org/10.1002/ijc.21655>.
29. Ho L, Chan SY, et al. Sequence variants of human papillomavirus type 16 in clinical samples permit verification and extension of epidemiological studies and construction of a phylogenetic tree. *J Clin Microbiol*. 1991;29:1765–72. <https://doi.org/10.1128/jcm.29.9.1765-1772.1991>.
30. Ho L, Chan SY, et al. The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *J Virol*. 1993;67:6413–23. <https://doi.org/10.1128/jvi.67.11.6413-6423.1993>.
31. Ho L, Tay SK, et al. Sequence variants of human papillomavirus type 16 from couples suggest sexual transmission with low infectivity and polyclonality in genital neoplasia. *J Infect Dis*. 1993. <https://doi.org/10.1093/infdis/168.4.803>.
32. Yamada T, Wheeler CM, et al. Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments. *J Virol*. 1995;69:7743–53.
33. Yamada T, Manos MM, et al. Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. *J Virol*. 1997;71:2463–72. <https://doi.org/10.1128/jvi.71.3.2463-2472.1997>.
34. Jackson R, Rosa BA, et al. Functional variants of human papillomavirus type 16 demonstrate host genome integration and transcriptional alterations corresponding to their unique cancer epidemiology. *BMC Genomics*. 2016. <https://doi.org/10.1186/s12864-016-3379-6>.
35. Nordfors C, Sobkowiak M, et al. Human papillomavirus (HPV) 16 E6 variants in tonsillar cancer in comparison to those in cervical cancer in Stockholm, Sweden. *PLOS ONE*. 2012. <https://doi.org/10.1371/journal.pone.0036239>.
36. Stöppler MC, Ching K, et al. Natural variants of the human papillomavirus type 16 E6 protein differ in their abilities to alter keratinocyte differentiation and to induce p53 degradation. *J Virol*. 1996. <https://doi.org/10.1128/jvi.70.10.6987-6993.1996>.
37. Gillison ML, Koch WM, et al. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *J Natl Cancer Inst*. 2000;92:709–20. <https://doi.org/10.1093/jnci/92.9.709>.
38. Hoffmann M, Lohrey C, et al. Human papillomavirus type 16 E6 and E7 genotypes in head-and-neck carcinomas. *Oral Oncol*. 2004. <https://doi.org/10.1016/j.oraloncology.2003.10.011>.
39. Xi LF, Demers GW, et al. Analysis of human papillomavirus type 16 variants indicates the establishment of persistent infection. *J Infect Dis*. 1995. [https://doi.org/10.1093/infdis/172.3.747\(3\)](https://doi.org/10.1093/infdis/172.3.747(3)).
40. Van Belkum A, Juffermans L, et al. Genotyping human papillomavirus type 16 isolates from persistently infected promiscuous individuals and cervical neoplasia patients. *J Clin Microbiol*. 1995. <https://doi.org/10.1128/jcm.33.11.2957-2962.1995>.
41. Zhang L, Liao H, et al. Variants of human papillomavirus type 16 predispose toward persistent infection. *Int J Clin Exp Pathol*. 2015;87:8453.
42. The Cancer Genome Atlas Network Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature*. 2015; doi: <https://doi.org/10.1038/nature14129>
43. Castellsagué X, Alemany L, et al. HPV involvement in head and neck cancers: comprehensive assessment of biomarkers in 3680 patients. *J Natl Cancer Inst*. 2016. <https://doi.org/10.1093/jnci/djv403>.
44. Dok R, Nuyts S. HPV positive head and neck cancers: Molecular pathogenesis and evolving treatment strategies. *Cancers (Basel)*. 2016. <https://doi.org/10.3390/cancers8040041>.
45. Pakdel F, Farhadi A, et al. The frequency of high-risk human papillomavirus types, HPV16 lineages, and their relationship with p16INK4a and NF-κB expression in head and neck squamous cell carcinomas in Southwestern Iran. *J Microbiol*. 2020. <https://doi.org/10.1007/s42770-020-00391-1>.
46. Cullen M, Boland JF, et al. Deep sequencing of HPV16 genomes: A new high-throughput tool for exploring the carcinogenicity and natural history of HPV16 infection. *Papillomavirus Res*. 2015. <https://doi.org/10.1016/j.pvr.2015.05.004>.
47. Lavezzo E, Masi G, et al. Characterization of intra-type variants of oncogenic human papillomaviruses by next-generation deep sequencing of the E6/E7 region. *Viruses*. 2016. <https://doi.org/10.3390/v8030079>.
48. van der Weele P, Meijer CJLM, et al. Whole-Genome Sequencing and Variant Analysis of Hpv16 Infections. *J Virol*. 2017. <https://doi.org/10.1128/JVI.00844-17>.
49. Badaracco G, Rizzo C, et al. Molecular analyses and prognostic relevance of HPV in head and neck tumours. *Oncol Rep*. 2007
50. Agrawal Y, Koch WM, et al. Oral HPV infection before and after treatment for HPV16-positive and negative head and neck squamous cell carcinoma. *Oral Oncol*. 2008. <https://doi.org/10.1158/1078-0432.CCR-08-0498.Oral>.
51. Boscolo-Rizzo P, Da Mosto MC, et al. HPV-16 E6 L83V variant in squamous cell carcinomas of the upper aerodigestive tract. *J Cancer Res Clin Oncol*. 2009. <https://doi.org/10.1007/s00432-008-0490-3>.
52. Joseph AW, Ogawa T, et al. Molecular etiology of second primary tumours in contralateral tonsils of human papillomavirus-associated index tonsillar carcinomas. *Oral Oncol*. 2013. <https://doi.org/10.1016/j.oraloncology.2012.09.009>.
53. Hassani S, Castillo A, Ohori J-I, Higashi M, Kurono Y, Akiba S, Koriyama C. Molecular pathogenesis of human papillomavirus type 16 in Tonsillar squamous cell carcinoma. *Anticancer Res*. 2015
54. Betiol JC, de Matos LL, et al. Prevalence of human papillomavirus types and variants and p16INK4a expression in head and neck squamous cells carcinomas in São Paulo, Brazil. *Infectious Agents Cancer*. 2016. <https://doi.org/10.1186/s13027-016-0067-8>.
55. Ang KK, Harris J, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med*. 2010. <https://doi.org/10.1056/nejmoa0912217>.
56. Beatty BT, Moon DH, et al. PIK3CA mutation in HPV-associated OPSCC patients receiving deintensified chemoradiation. *JNCI J Natl Cancer Inst*. 2019;112:855–8. <https://doi.org/10.1093/jnci/djz224>.
57. Villa LL, Sichero L, et al. Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. *J Gen Virol*. 2000. <https://doi.org/10.1099/0022-1317-81-12-2959>.
58. Schiffman M, Rodriguez AC, et al. A population-based prospective study of carcinogenic human papillomavirus (HPV) variant lineages, viral persistence, and cervical neoplasia. *NIH Public Access*. 2010. <https://doi.org/10.1158/0008-5472.CAN-09-4179.A>.
59. Sichero L, Ferreira S, et al. High grade cervical lesions are caused preferentially by non-European variants of HPVs 16 and 18. *Int J Cancer*. 2007. <https://doi.org/10.1002/ijc.22481>.
60. Hirose Y, Onuki M, et al. Within-host variations of human papillomavirus reveal APOBEC signature mutagenesis in the viral genome. *J Virol*. 2018. <https://doi.org/10.1128/jvi.00017-18>.
61. Mirabello L, Clarke MA, et al. The intersection of HPV epidemiology, genomics and mechanistic studies of HPV-mediated carcinogenesis. *Viruses*. 2018. <https://doi.org/10.3390/v10020080>.
62. Zhu B, Xiao Y, et al. Mutations in the HPV16 genome induced by APOBEC3 are associated with viral clearance. *Nat Commun*. 2020. <https://doi.org/10.1038/s41467-020-14730-1>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.