COLD SPRING HARBOR
Molecular Case Studies

RESEARCH REPORT

Genetic and transcriptomic analyses in a rare case of human papillomavirus–related oropharyngeal squamous-cell carcinoma combined with small-cell carcinoma

Kuniaki Sato,[1] Kazuo Nishiyama,[1] Kenichi Taguchi,[2] Rina Jiromaru,[3,4] Hidetaka Yamamoto,[3] Akihide Matsunaga,[1] Ryozaburo Nagata,[1] Fumihide Rikimaru,[1] Satoshi Toh,[1] Yuichiro Higaki,[1] Shinya Oda,[5] Takashi Nakagawa,[4] and Muneyuki Masuda[1]

[1]Department of Head and Neck Surgery, [2]Department of Pathology, National Hospital Organization Kyushu Cancer Center 3-1-1 Minami-ku, Notame, Fukuoka, 874-0838, Japan; [3]Department of Anatomic Pathology, [4]Department of Otorhinolaryngology, Graduate School of Medical Sciences, Kyushu University, 3-1-1 Higashi-ku, Fukuoka, Fukuoka, 860-8556, Japan; [5]Clinical Research Institute, National Hospital Organization Kyushu Cancer Center 3-1-1 Minami-ku, Notame, Fukuoka, 874-0838, Japan

Abstract Human papillomavirus (HPV)-related oropharyngeal small-cell carcinoma (OPSmCC) is a rare malignancy with aggressive behavior, whereas HPV-related oropharyngeal squamous-cell carcinoma (OPSqCC) displays a favorable prognosis. Notably, these two malignancies occasionally arise in an identical tumor. In this case study, we explored the molecular characteristics that distinguishes these two carcinomas using a rare case of HPV-related oropharyngeal carcinoma (OPC) with the combined histology of SmCC and SqCC. Immunohistochemical analysis and HPV-RNA in situ hybridization (ISH) suggested that both SmCC and SqCC were HPV-related malignancies. Targeted exome sequencing revealed that SmCC and SqCC had no significant difference in mutations of known driver genes. In contrast, RNA sequencing followed by bioinformatic analyses suggested that aberrant transcriptional programs may be responsible for the neuroendocrine differentiation of HPV-related OPC. Compared to SqCC, genes up-regulated in SmCC were functionally enriched in inflammatory and immune responses (e.g., arachidonic acid metabolism). We then developed a SmCC-like gene module (top 10 up-regulated genes) and found that OPC patients with high module activity showed poor prognosis in The Cancer Genome Atlas (TCGA) and GSE65858 cohort. Gene set enrichment analysis of the SmCC-like gene module suggested its link to MYC proto-oncogene in the TCGA data set. Taken together, these findings suggest that the SmCC-like gene module may contribute to acquisition of aggressive phenotypes and tumor heterogeneity of HPV-related OPC. The present case study is the first report of genetic and transcriptomic aberrations in HPV-related OPSmCC combined with SqCC.

[Supplemental material is available for this article.]

Corresponding author:
mmuneyuki@icloud.com

Ontology term: neoplasia of the pharynx

Published by Cold Spring Harbor Laboratory Press

doi:10.1101/mcs.a006102

INTRODUCTION

The discovery and recent upsurge of human papillomavirus (HPV)-related oropharyngeal squamous-cell carcinoma (OPSqCC) had a substantial impact on the treatment of head

and neck cancer (Ang et al. 2010; Pytynia et al. 2014). This distinctive type of cancer demonstrates favorable responses to conventional chemo/radiation therapy and is thereby associated with significantly better survival than canonical tobacco- and alcohol-related HPV-negative OPSqCC. As a result, in 2017, the American Joint Committee on Cancer TNM classification specified HPV-related OPSqCC as a separate entity and set a lenient classification (i.e., downstaging) to this malignancy (O'Sullivan et al. 2016; Machczyński et al. 2020). Consequently, there has been a worldwide attempt to reduce the treatment intensity for HPV-related OPSqCC (Saito et al. 2020).

On the other hand, in accordance with the increase in HPV-related OPSqCC, the presence of HPV-related oropharyngeal small-cell carcinoma (OPSmCC) has been recognized recently (Bishop et al. 2012). Consistent with SmCC in lung and extrapulmonary regions (e.g., larynx) (Bean et al. 2019; Wakasaki et al. 2019), HPV-related OPSmCC demonstrates an aggressive phenotype and a highly dismal prognosis (Bishop and Westra 2011). Notably, HPV-related OPSqCC and OPSmCC often develop as a combined tumor (Bishop and Westra 2011; Nakano et al. 2017). Thus, it is of particular interest and of clinical importance to investigate the molecular mechanism that divides these two HPV-related carcinomas arising in the same organ. The occurrence of HPV-related oropharyngeal carcinoma (OPC) is explained as follows (Chung and Gillison 2009; Marur et al. 2010; Pytynia et al. 2014; Kang et al. 2015): (1) infection of high-risk HPV to the basal progenitor cells in the tonsil crypts; (2) HPV–DNA integration into the host genomic DNA; (3) production of viral E6 and E7 proteins and respective inhibition of p53 and pRB proteins, which lead to the increased expression of the p16 protein through a feedback loop; and (4) frequent mutations of *PIK3CA*. Interestingly, the genetic landscape of lung SmCC is characterized by essential inactivating mutations in two identical molecules inhibited in HPV-related OPSqCC and OPSmCC: *TP53* and *RB1* (George et al. 2015). In the case of lung SmCC, certain additional mutations (e.g., *NOTCH1*) were observed in 25% of the patients, suggesting that inactivation of notch signaling may affect SmCC development (George et al. 2015). Therefore, we speculated that a similar mechanism may endow HPV-infected OPC with an aggressive neuroendocrine phenotype.

In this context, we recently encountered a rare case of HPV-related OPC in which we detected the combined histology of SmCC and SqCC before treatment. This information enabled us to separately collect frozen samples from each component during surgery. We then analyzed the genetic and transcriptomic features of HPV-related SmCC and SqCC of the oropharynx using targeted exome sequencing and RNA sequencing (RNA-seq). Genetically, no significant difference in the mutations of known driver genes was observed between SmCC and SqCC. In contrast, RNA-seq suggested that epigenetic mechanisms may be involved in neuroendocrine differentiation of HPV-related OPC. By comparing the expression profiles, we developed a SmCC-like gene module. Interestingly, high activity of the SmCC-like gene module was significantly associated with the poor prognosis of OPC patients in The Cancer Genome Atlas (TCGA) and another cohort. Our results suggest that aberrant transcriptional programs may not only affect tumor heterogeneity and clonal evolution of HPV-related OPC, but also possibly contribute as a biomarker to uncover potential high-risk group in HPV-related OPC patients.

RESULTS

Clinical Presentation

A 53-yr-old Japanese female patient who presented with a sore throat and lymphadenopathy in her left neck visited a local clinic 3 wk prior to the first visit to our hospital. She had no history of smoking and drank socially. She was found to have an ulcerated mass in her left tonsil, which was initially diagnosed as SmCC by small-tissue biopsy. The patient was referred to our

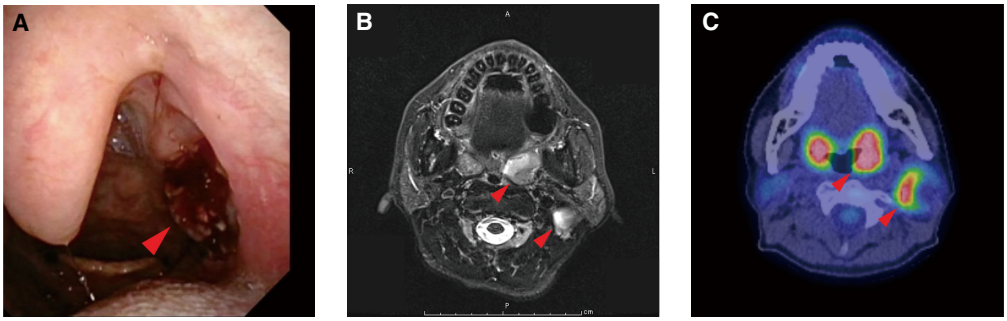

Figure 1. Clinical images of the patient with human papillomavirus (HPV)-related oropharyngeal carcinoma (OPC) in this study. (*A*) A fiberscopic image of the primary tumor in the left tonsil (red arrowhead). (*B*,*C*) Postcontrast T1-weighted magnetic resonance imaging (MRI; *B*) and positron emission tomography performed with 2-[fluorine-18]fluoro-2-deoxy-d-glucose (FDG-PET; *C*) revealed a left tonsillar mass and left cervical lymphadenopathy in the level II region (red arrowhead).

hospital for further examination. To examine the expression of p16, tissue biopsy from the tumor was retaken (Fig. 1A). Subsequent histopathological analysis indicated that this tumor contained SqCC, as it was immunohistochemically positive for p16 and neuroendocrine marker-positive necrotic cells. Taken together, the findings suggested that the tumor was a rare type of OPC showing the combined histology of SqCC and SmCC.

Contrast-enhanced magnetic resonance imaging (MRI) of the head and neck region showed a 35-mm well-defined, enhancing mass in the left tonsil with the involvement of multiple left lymph nodes (Fig. 1B). Whole-body positron emission tomography/computed tomography (PET/CT) indicated prominent fluorodeoxyglucose uptake within the left tonsil and level II lymph nodes in the left neck without distant metastasis (Fig. 1C). From these findings, the clinical stage of this oropharyngeal tumor was diagnosed as T2N1M0 and stage I disease according to the 8th edition of the Union for International Cancer Control-TNM (tumor, node, and metastasis; UICC-TNM) classification (Amin et al. 2017). Given the aggressive behavior of SmCC, she was treated with pull-through oropharyngectomy plus left modified neck dissection (level I-V), followed by reconstruction of the left lateral oropharyngeal wall with a right anterolateral thigh flap. Intraoperative histological examination of the tonsillar tumor specimen revealed the intratumor distribution of SmCC and SqCC, enabling separate collection of the SmCC and SqCC components from the tumor specimen. Then, these tumor samples and a noncancerous tonsillar tissue taken from the right tonsil were snap-frozen in liquid nitrogen for further molecular analyses. Adjuvant chemoradiation therapy (CRT) was performed with a total dose of 61.2-Gy radiation and 1 cycle of cisplatin (80 mg/m^2). At 26 months after adjuvant CRT, the patient is alive without locoregional recurrence or distant metastasis.

Pathological and Immunohistochemical Findings

The surgically resected primary tumor in the left tonsil was subjected to histological analysis. HE staining of the tumor showed the combined histology of SmCC and SqCC, in which small monotonous carcinoma cells and nonkeratinizing squamous carcinoma cells were synchronously observed (Fig. 2A–E). Immunohistochemically, the tumor cells of the SmCC component were positive for the neuroendocrine marker chromogranin A but negative for the squamous marker p40. In contrast, the tumor cells of the SqCC component were diffusely positive for p40 but negative for chromogranin A. Both components showed a positive staining pattern for p16 (Fig. 3A–D). In high-grade neuroendocrine carcinomas such as SmCC,

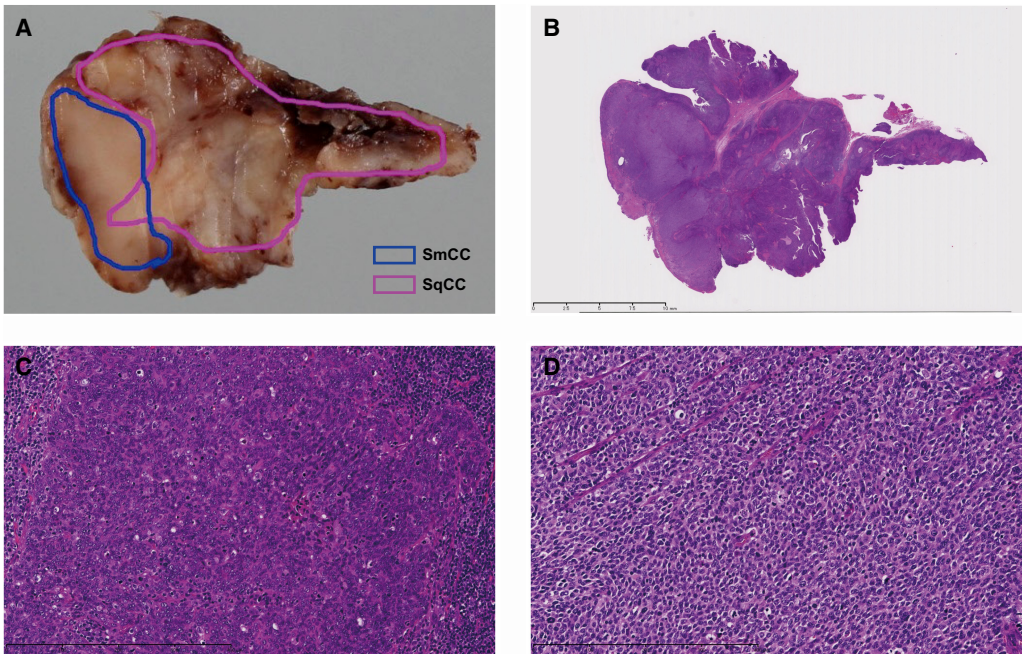

Figure 2. Representative histology of the primary OPC with combined histology of SmCC and SqCC. (*A*) A macroscopic image of the surgically resected primary tumor. The resected specimen showed small-cell carcinoma (SmCC; blue line) and squamous-cell carcinoma (SqCC; purple line). (*B–D*) Histological images of the primary tumor (*B*, hematoxylin and eosin staining) indicating the combined histology of SmCC (*C*) and SqCC (*D*) components.

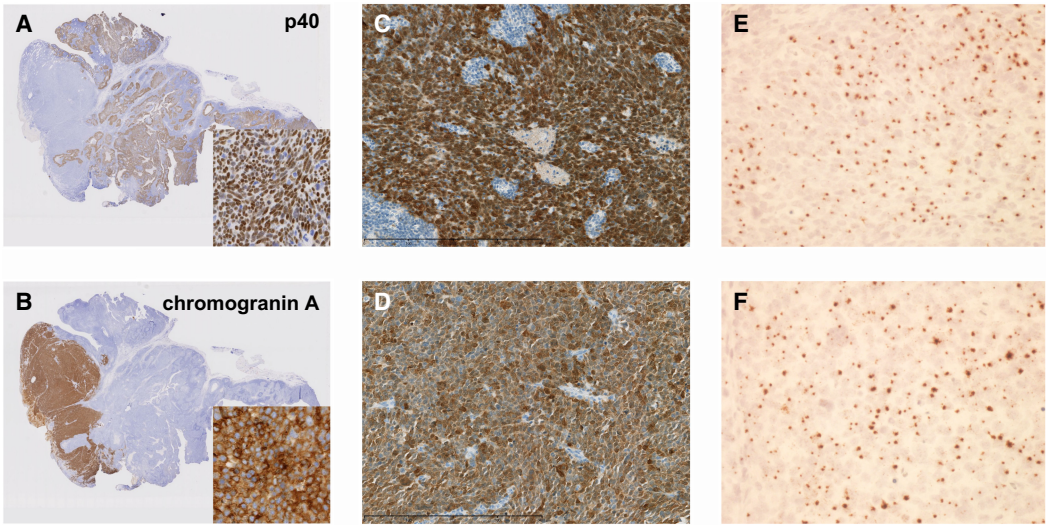

Figure 3. Immunohistochemistry (IHC) and HPV-RNA in situ hybridization (ISH) of the primary OPC. (*A*) IHC for p40 showed positive nuclear staining in SqCC but negative staining in SmCC. (*B*) In contrast, chromogranin A was immunohistochemically positive in SmCC but negative in SqCC. (*C*,*D*) IHC for p16 indicating positive staining in both SqCC (C) and SmCC (D). (*E*,*F*) High-risk HPV-RNA ISH showed positive nuclear staining in SqCC (*E*) and SmCC (*F*).

p16 is reportedly not suitable as a surrogate marker of HPV infection because p16 is frequently overexpressed via other mechanisms (Yuan et al. 1999; Bishop and Westra 2011). For this reason, we performed HPV-RNA in situ hybridization (ISH) on these tumor specimens. This analysis showed that both SmCC and SqCC were positive for high-risk HPV RNA, indicating that both of these tumor cells are HPV-related malignancies (Fig. 3E,F).

Genomic Analysis

To reveal the differences in somatic mutations of known driver genes between SmCC and SqCC, targeted exome sequencing was performed on the DNA extracted from tumor tissues and noncancerous tissues. In total, 44 and 42 somatic mutations were identified in SmCC and SqCC, respectively, and 42 mutations were shared between SmCC and SqCC. Of these, *PIK3CA* p.E542K, which is frequently mutated in head and neck cancers, including HPV-related OPSqCC, was identified as a known hotspot driver mutation listed in the COSMIC database (Table 1; Stransky et al. 2011; Nichols et al. 2013). Although two mutations, *CDKN2A* p.H66R and *ADGRB3* p.E436K, were found to be unique to SmCC, their functional and clinical significance has not yet been reported. Thus, we concluded that there was no significant difference in major driver mutations between SmCC and SqCC.

Transcriptomic Analyses with Publicly Available Data Sets

The results of the genomic analysis motivated us to investigate differences in the transcriptomes between SmCC and SqCC. For this purpose, we conducted RNA-seq using RNA extracted from SmCC, SqCC, and noncancerous tonsillar tissues. Unsupervised hierarchical clustering and principal component analysis (PCA) of the normalized read count data of SmCC, SqCC, and 69 OPSqCC tumors (including 46 HPV-positive and 23 HPV-negative cases) from the TCGA data set showed that both SmCC and SqCC belong to the HPV-positive cluster, suggesting that the overall transcriptional profiles of these two cancers are influenced by HPV infection (Fig. 4A; Supplemental Fig. S1A). Subsequent bioinformatics analyses revealed that 966 genes were up-regulated, whereas 993 genes were down-regulated significantly in SmCC compared with SqCC (Fig. 4B; Supplemental Fig. S1B). Gene Ontology (GO) analysis of the SmCC-specific up-regulated genes showed the enrichment of gene sets involved in arachidonic acid metabolism, inflammatory and immune responses and keratinocyte differentiation (Table 2). Among these up-regulated genes, we selected the top 10 up-regulated genes as the "SmCC-like gene module" (Fig. 4C). This module included genes reportedly involved in tumor progression, invasion and chemoresistance, such as Uroplakin 1B (*UPK1B*) and Kallikrein-related peptidases (*KLK7*, *8* and *12*) (Dong et al. 2010; Kryza et al. 2016; Wang et al. 2018; Gong et al. 2020). To explore oncogenic pathways correlated with the SmCC-like gene module activity (i.e., average of 10 module gene

Table 1. Somatic mutations observed in SmCC and SqCC

Gene	Chr	HGVS CDS	HGVS protein	Variant type	Predicted effect	dbSNP	Genotype
Hotspot mutation observed in both SmCC and SqCC							
PIK3CA	3	NM_006218.3: c.1624G > A	NP_006209.2: p.E542K	SNV	Missense	rs121913273	Heterozygous
Mutations unique to SmCC							
CDKN2A	9	NM_001195132.1: c.197A > G	NP_001182061.1: p.H66R	SNV	Missense	rs756750256	Heterozygous
ADGRB3	6	NM_001704.2: c.1306G > A	NP_001695.2: p.E436K	SNV	Missense	rs141698131	Heterozygous

(Chr) Chromosome, (HGVS CDS) Human Genome Variation Society coding sequence, (dbSNP) Single Nucleotide Polymorphism Database.

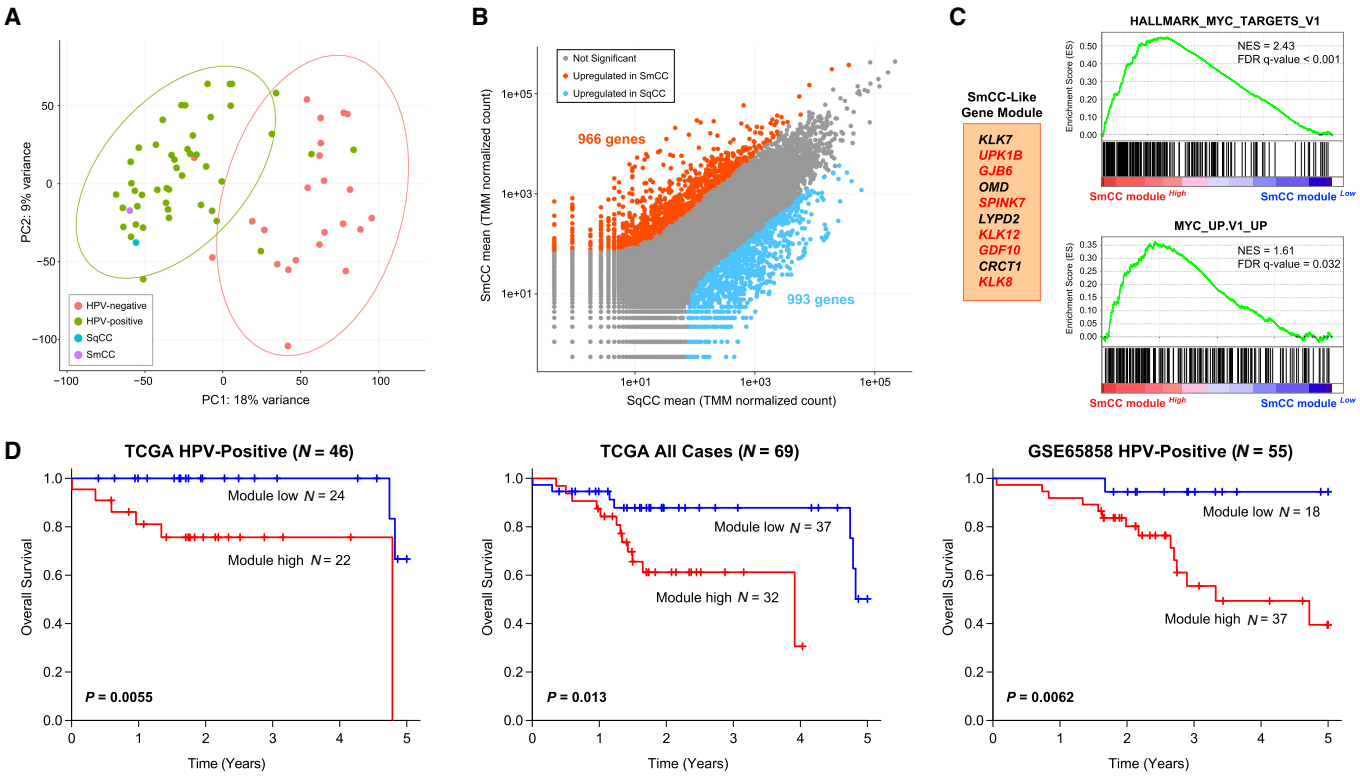

Figure 4. Clinical significance of SmCC-like gene module in OPC. (*A*) Principal component analysis of SmCC, SqCC, and 69 OPSqCC patients in the TCGA illustrating the clustering of samples. Both SqCC (blue dot) and SmCC (purple dot) were clustered in the HPV-positive group (green ellipse). (*B*) Differential expression analysis. Each dot represents the normalized read counts of genes. Red dots and blue dots indicate significantly up-regulated genes in SmCC and SqCC, respectively. (*C*) Gene Set Enrichment Analysis (GSEA) showing MYC target gene sets positively correlated with the SmCC-like gene module activity in the TCGA HPV-positive OPSqCC data set ($N = 46$). Predicted MYC target genes in the SmCC-like gene module are shown in red letters. (NES) Normalized enrichment score, (FDR) false discovery rate. (*D*) Clinical relevance of the SmCC-like gene module activity in OPSqCC patients. Kaplan–Meier survival curves for the overall survival of HPV-positive patients ($N = 46$, left) and all patients ($N = 69$, middle) in the TCGA data set. Kaplan–Meier survival curves for the overall survival of HPV-positive patients in GSE65858 ($N = 55$, right). *P*-values were calculated using the log-rank test.

expression levels; see Methods), we performed gene set enrichment analysis (GSEA) using the TCGA HPV-positive OPSqCC data set. Of note, SmCC-like gene module activity was positively correlated with the expression of target gene sets of MYC proto-oncogene. Consistent with these results, six out of 10 genes in SmCC-like gene module and 59% of up-regulated genes in SmCC were predicted as transcriptional targets of MYC (Fig. 4C).

The SmCC-Like Gene Module Predicts Poor Prognosis of OPC Patients

Finally, we explored clinical implications of the SmCC-like gene module activity in OPC patients. Notably, this activity was significantly correlated with poor prognosis in both HPV-positive ($N = 46$) and the overall cohort ($N = 69$) (log-rank $P = 0.0055$ and 0.013, respectively) (Fig. 4D). These results were further validated using GSE65858 microarray data set ($N = 55$) (Wichmann et al. 2015). The module activity was positively correlated with poor prognosis of HPV-positive OPC patients in this cohort (log-rank $P = 0.0062$) (Fig. 4D), suggesting that the SmCC-like gene expression program might be involved in the acquisition of the aggressive phenotypes of OPC.

Table 2. Enriched gene ontology (GO) categories and KEGG pathways of the up-regulated genes in small-cell carcinoma detected by DAVID (FDR < 0.01)

GO term (biological process)	GO ID	Gene count	FDR
Peptide cross-linking	GO:0018149	21	9.79×10^{-13}
Keratinocyte differentiation	GO:0030216	23	4.58×10^{-11}
Immune response	GO:0006955	52	1.49×10^{-10}
Epidermis development	GO:0008544	23	2.89×10^{-10}
Keratinization	GO:0031424	17	4.96×10^{-9}
Inflammatory response	GO:0006954	42	5.98×10^{-7}
Innate immune response	GO:0045087	38	1.01×10^{-3}
Cell adhesion	GO:0007155	37	9.13×10^{-3}
Adaptive immune response	GO:0002250	18	9.13×10^{-3}
Leukocyte chemotaxis	GO:0030595	6	9.13×10^{-3}
Humoral immune response	GO:0006959	11	9.13×10^{-3}
Response to lipopolysaccharide	GO:0032496	19	9.13×10^{-3}
Negative regulation of endopeptidase activity	GO:0010951	16	9.13×10^{-3}
KEGG pathway	**Pathway ID**	**Gene count**	**FDR**
Arachidonic acid metabolism	hsa00590	13	4.83×10^{-4}
B-cell receptor signaling pathway	hsa04662	13	9.29×10^{-4}
Hematopoietic cell lineage	hsa04640	14	1.46×10^{-3}
Staphylococcus aureus infection	hsa05150	11	1.46×10^{-3}
Intestinal immune network for IgA production	hsa04672	10	2.22×10^{-3}
Asthma	hsa05310	8	3.36×10^{-3}
Cytokine–cytokine receptor interaction	hsa04060	23	3.90×10^{-3}
Drug metabolism—cytochrome P450	hsa00982	11	5.02×10^{-3}
Primary immunodeficiency	hsa05340	8	5.02×10^{-3}
Complement and coagulation cascades	hsa04610	11	5.02×10^{-3}
Metabolism of xenobiotics by cytochrome P450	hsa00980	11	7.56×10^{-3}
Malaria	hsa05144	9	7.56×10^{-3}

(KEGG) Kyoto Ecyclopedia of Genes and Genomes, (FDR) false discovery rate.

DISCUSSION

As shown in the present study, HPV-related OPSmCC and SqCC arise in an identical organ under the influence of a common carcinogenic pathway (i.e., infection of high-risk HPV) (Bishop and Westra 2011; Nakano et al. 2017). However, these two types of cancer demonstrate opposing phenotypes: SmCC is recognized as one of the most aggressive forms of cancer (Bishop and Westra 2011), whereas SqCC is recognized as a curable cancer (Ang et al. 2010). Concerning tumor progression, SmCC appears to be a more advanced cancer that acquires neuroendocrine cell morphology and an aggressive phenotype compared to SqCC, although they develop on the commonly shared inhibition of p53 and pRB proteins. Our main concern was to clarify the molecular characteristics that distinguish these two cancers. In reference to the genetic landscape of pulmonary SmCC (George et al. 2015), which is described by essential mutations of *TP53* and *RB1* and additional mutations, including those of *TP73* and *NOTCH1*, we hypothesized that certain driver mutations may contribute to the development

of HPV-related SmCC. However, our results suggested that not genetic but rather transcriptional programs could be a main cause of neuroendocrine differentiation in the present case.

GO analysis indicated the enrichment of genes involved in inflammation and immune responses in the SmCC component. Given that the physiological role of the palatine tonsil is the barrier for pathogens, including viruses and bacteria, it is plausible that OPC is dependent on the transcriptional programs activated in response to organ-specific microenvironmental stressors. It is now a widely accepted concept that stress-triggered transcriptional programs are strong inducers of cancer hallmarks such as limitless self-renewal, anti-apoptosis, and dedifferentiation (Quail and Joyce 2013; Zanconato et al. 2019). This scenario is supported by our recent discovery that conditional activation of the stress sensor Yap1 can transcriptionally cause rapid carcinogenesis of oral SqCC in mice (Omori et al. 2020) and appears to account for the correlation between the prevailing activation of the SmCC-like gene module and the aggressive phenotype of OPC. Moreover, arachidonic acid metabolism, which is a representative inflammatory response, was the most prominently enriched pathway in the SmCC expression profile (Table 2). Analogous to OPC, colon adenocarcinomas arise from the progenitor cells in colonic crypts, where the end products of the arachidonic acid cascade, prostaglandins, play an important role in carcinogenesis (Terzić et al. 2010). A similar process may be involved in the carcinogenesis of OPC, particularly SmCC. GSEA in the TCGA data set showed positive correlations between SmCC-like gene module activity and the expression of target genes of MYC. A recent study reported that c-Myc and L-Myc determine molecular and histological subtypes of pulmonary SmCC (Patel et al. 2021). Similarly, activation of N-Myc transforms human prostate epithelial cells to neuroendocrine cancer cells (Lee et al. 2016). Interestingly, overexpressed MYC reportedly drives proliferation of HPV-positive cervical neuroendocrine carcinoma (Yuan et al. 2017). In the present case, we did not observe statistically significant differences in *MYC* expression levels between SmCC and SqCC (3.6-fold higher in SmCC compared with SqCC, data not shown). However, considering that more than half of the up-regulated genes in SmCC were predicted as transcriptional targets of MYC, the aberrant transcriptional program induced by MYC may play a key role in carcinogenesis and/or differentiation of OPSmCC. The precise interactions of the SmCC-like gene module with HPV infection, neuroendocrine differentiation as well as *MYC* proto-oncogene must be elucidated in future studies. It is also an intriguing subject to investigate whether the occurrence of SmCC is a de novo process from progenitor cells or a stepwise transformation process from SqCC.

Of note, activation of the SmCC-like gene module predicted poor survival of OPC patients. Our results illuminated the potential clinical risks associated with OPC diagnosis and treatment. It was coincidental that we have detected the combined histology of SmCC and SqCC in the respective small-biopsy specimens taken in the former clinic and our facility. Therefore, there was a high possibility that we might have misdiagnosed this case as HPV-related SqCC overlooking the SmCC component and then treated the patient with CRT alone, resulting in unfavorable outcomes. One potential pitfall of small-tissue biopsies is overlooking subpopulations with aggressive phenotype in the entire tumor, because these techniques provide only a limited number of cells from limited sites. Additionally, precise diagnosis on such small specimens acquired from oropharynx are still challenging (Palsgrove and Bishop 2020). Our data suggested that, as well as SmCC, not a small population of SqCC that retains squamous morphology may be malignantly reprogrammed by the SmCC expression profile in the OPC tumor microenvironment. However, comprehensive genomic/transcriptomic analysis combined with clinicopathological analysis in larger patient cohorts of HPV-related SmCC will be required to uncover the molecular basis and tumor heterogeneity of HPV-related OPSmCC as well as to develop strategies for detecting high-risk HPV-related OPC patients.

In conclusion, we described a comprehensive histologic, genomic, and transcriptomic analysis of a rare case of HPV-related OPC with the combined histology of SqCC and SmCC. Our findings provide biological and clinical insights into HPV-related OPSmCC.

METHODS

Immunohistochemical Analysis

Immunohistochemical staining was performed on 4-μm formalin-fixed, paraffin-embedded tissue sections. The primary antibodies used in the analyses were as follows: anti-p40 (BC28, 1:200; Abcam), anti-p16 (E6H4, prediluted; CIN Histology Kit; Roche), and anti-chromogranin A (prediluted; Nichirei Biosciences). For p16, immunohistochemistry (IHC) was performed with an ultraView universal DAB detection kit (Ventana Medical Systems) according to the manufacturer's protocol. An OptiView DAB detection kit (Ventana Medical Systems) was used for p40 and chromogranin A. All slides were counterstained with Mayer's hematoxylin (Sigma-Aldrich).

HPV RNA In Situ Hybridization (ISH)

ISH was performed to detect HPV-RNA using an E6/E7 mRNA probe set (Advanced Cell Diagnostics) for 18 types of high-risk HPV (types 16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 73, and 82) with an RNA scope 2.5HD detection kit (Advanced Cell Diagnostics) according to the manufacturer's protocols. Any nuclear or cytoplasmic staining was considered positive as previously described (Bishop et al. 2012; Jiromaru et al. 2020).

Sample Collection and Nucleic Acid Extraction

Intraoperative histological examination was performed by an experienced pathologist (K.T.) to evaluate the intratumor distribution of SmCC and SqCC in the surgically resected primary tumor specimen. Based on this evaluation, the SmCC and SqCC samples were obtained from each component by macrodissection. These tumor samples and a noncancerous sample obtained from the right tonsil were immediately frozen in liquid nitrogen and stored at −80°C until nucleic acid extraction. Genomic DNA was extracted from each sample by ethanol precipitation. Total RNA extraction was performed using ISOGEN-II (Nippon Gene) according to the manufacturer's protocols.

Targeted Sequencing for the Detection of Somatic Mutations

DNA extracted from the tumor samples and the noncancerous sample was subjected to targeted exome sequencing for 409 known cancer-related genes using an Oncomine tumor mutation load assay kit and the IonS5 System (Thermo Fisher) according to the manufacturer's protocols. The sequencing data were analyzed with Ion Reporter v5.1 and Torrent suite v5.8.0 software.

RNA Sequencing (RNA-seq) and Data Analysis

RNA extracted from the tumor samples and the noncancerous sample was sequenced on a DNBSEQ-G400 sequencer at Beijing Genomics Institutions (Shenzhen, China). The sequenced data were analyzed with the in-house pipeline Genomon2 v2.5.2 (https://genomon-project.github.io/GenomonPagesR) using the supercomputing system SHIROKANE (University of Tokyo). Briefly, sequencing reads were aligned to the human reference GRCh37/hg37 genome by STAR v2.5.2a using Gencode v22 annotations. Gene count tables were generated with htseq-count, a part of HTSeq v0.6.0 (Anders et al. 2015). Downstream analyses were carried out using R v3.6.3. (The R Foundation for

Statistical Computing). Normalization of the read count data and the detection of differentially expressed genes between SmCC and SqCC were carried out with NOIseq v2.28.0 (Tarazona et al. 2015). GO analysis was performed using the DAVID tool version 6.8 Beta (Huang et al. 2007).

Analysis of the Publicly Available Transcriptomic Data Set

The clinical information, RNA-seq read count data, and RSEM normalized expression data of 69 patients with primary OPSqCC in the TCGA were downloaded from the GDC data portal (https://portal.gdc.cancer.gov/). Microarray data with clinical information of OPSqCC patients in the GSE65858 data set were downloaded from NCBI Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo). The read count data of OPSqCC patients in the TCGA and the patient in this study were combined into one matrix. Subsequently, the merged data were normalized using DESeq2 v1.24.0 (Love et al. 2014). For sample clustering and principal component analysis, genes with zero counts across all samples were removed. SmCC-like gene module activity was calculated for each patient using the average of the mRNA expression levels of ten genes overexpressed in SmCC tissues (i.e., *KLK7*, *UPK1B*, *GJB6*, *OMD*, *SPINK7*, *LYPD2*, *KLK12*, *GDF10*, *CRCT1*, and *KLK8*) as previously described (Omori et al. 2019). GSEA was performed on RSEM-normalized RNA-seq data of the TCGA HPV-positive OPSqCC data set using GSEA v4.1.0 (Subramanian et al. 2005). Transcriptional target genes of MYC predicted by ChIP-seq experiments in ENCODE project were obtained from Harmonizome database (Rouillard et al. 2016). For survival analysis, Cutoff Finder (Budczies et al. 2012) was used to determine the cutoff values to stratify patients into two groups based on the SmCC-like gene module expression levels by calculating minimal Euclidean distances to the upper left corner on receiver operator characteristic (ROC) curves. Overall survival curves were plotted according to the Kaplan–Meier method and compared by the log-rank test using the R package survival. $P < 0.05$ was considered statistically significant.

ADDITIONAL INFORMATION

Database Deposition and Access

The RNA-seq data sets have been deposited into the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/) with accession number GSE166773.

Ethics Statement

The protocol of this study was reviewed and approved by the Institutional Review Board and Ethics Committee of National Kyushu Cancer Center (Protocol Number: 2015-43). Written informed consent was obtained from the patient. All experiments with human samples were conducted according to the principles expressed in the Declaration of Helsinki.

Acknowledgments

The authors thank all the technical staff of Clinical Research Institute and Department of Pathology of National Hospital Organization Kyushu Cancer Center for their assistance. This work used the supercomputing resource provided by the Human Genome Center, The Institute of Medical Science, The University of Tokyo (http://sc.hgc.jp/shirokane.html).

Competing Interest Statement

The authors have declared no competing interest.

Referees

Nobuhiko Oridate
Anonymous

Received April 27, 2021; accepted in revised form August 19, 2021.

Author Contributions

K.S. and M.M. designed the study. K.S., K.T., R.J., H.Y., and S.O. performed the experiments. K.S. performed the bioinformatics analyses. K.N., A.M., R.N., F.R., S.T., and Y.H. collected the surgically resected tumor and normal samples. K.S. wrote the original draft of the manuscript. K.S. and M.M. reviewed and edited the manuscript. T.N. and M.M. supervised the research.

Funding

This work was supported in part by the following grants and foundations: Japan Society for the Promotion of Science (Grant Number: 20K18300) and Takeda Science Foundation.

REFERENCES

Amin MB, Edge S, Greene F, Byrd DR, Brookland RK, Washington MK, Gershenwald JE, Compton CC, Hess KR, Sullivan DC, et al. 2017. *AJCC cancer staging manual*, 8th ed. Springer, New York.

Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31:** 166–169. doi:10.1093/bioinformatics/btu638

Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tân PF, Westra WH, Chung CH, Jordan RC, Lu C, et al. 2010. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med* **363:** 24–35. doi:10.1056/NEJMoa0912217

Bean MB, Liu Y, Jiang R, Steuer CE, Patel M, McDonald MW, Higgins KA, Beitler JJ, Shin DM, Saba NF. 2019. Small cell and squamous cell carcinomas of the head and neck: comparing incidence and survival trends based on surveillance, epidemiology, and end results (SEER) data. *Oncologist* **24:** 1562–1569. doi:10.1634/theoncologist.2018-0054

Bishop JA, Westra WH. 2011. Human papillomavirus–related small cell carcinoma of the oropharynx. *Am J Surg Pathol* **35:** 1679–1684. doi:10.1097/PAS.0b013e3182299cde

Bishop JA, Ma X-J, Wang H, Luo Y, Illei PB, Begum S, Taube JM, Koch WM, Westra WH. 2012. Detection of transcriptionally active high-risk HPV in patients with head and neck squamous cell carcinoma as visualized by a novel E6/E7 mRNA in situ hybridization method. *Am J Surg Pathol* **36:** 1874–1882. doi:10.1097/PAS.0b013e318265fb2b

Budczies J, Klauschen F, Sinn B, Győrffy B, Schmitt WD, Darb-Esfahani S, Denkert C. 2012. Cutoff finder: a comprehensive and straightforward web application enabling rapid biomarker cutoff optimization. *PLoS ONE* **7:** e51862. doi:10.1371/journal.pone.0051862

Chung CH, Gillison ML. 2009. Human papillomavirus in head and neck cancer: its role in pathogenesis and clinical implications: fig. 1. *Clin Cancer Res* **15:** 6758–6762. doi:10.1158/1078-0432.CCR-09-0784

Dong Y, Tan OL, Loessner D, Stephens C, Walpole C, Boyle GM, Parsons PG, Clements JA. 2010. Kallikrein-related peptidase 7 promotes multicellular aggregation via the α5β1 integrin pathway and paclitaxel chemoresistance in serous epithelial ovarian carcinoma. *Cancer Res* **70:** 2624–2633. doi:10.1158/0008-5472.CAN-09-3415

George J, Lim JS, Jang SJ, Cun Y, Ozretia L, Kong G, Leenders F, Lu X, Fernández-Cuesta L, Bosco G, et al. 2015. Comprehensive genomic profiles of small cell lung cancer. *Nature* **524:** 47–53. doi:10.1038/nature14664

Gong W, Liu Y, Preis S, Geng X, Petit-Courty A, Kiechle M, Muckenhuber A, Dreyer T, Dorn J, Courty Y, et al. 2020. Prognostic value of kallikrein-related peptidase 12 (KLK12) mRNA expression in triple-negative breast cancer patients. *Mol Med* **26:** 19. doi:10.1186/s10020-020-0145-7

Huang D, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaei J, Stephens R, Baseler MW, Lane HC, Lempicki RA. 2007. The DAVID gene functional classification tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol* **8:** R183. doi:10.1186/gb-2007-8-9-r183

Jiromaru R, Yamamoto H, Yasumatsu R, Hongo T, Nozaki Y, Hashimoto K, Taguchi K, Masuda M, Nakagawa T, Oda Y. 2020. HPV-related sinonasal carcinoma: clinicopathologic features, diagnostic utility of p16 and Rb immunohistochemistry, and *EGFR* copy number alteration. *Am J Surg Pathol* **44:** 305–315. doi:10.1097/PAS.0000000000001410

Kang SYC, Kannan N, Zhang L, Martinez V, Rosin MP, Eaves CJ. 2015. Characterization of epithelial progenitors in normal human palatine tonsils and their HPV16 E6/E7-induced perturbation. *Stem Cell Rep* **5:** 1210–1225. doi:10.1016/j.stemcr.2015.09.020

Kryza T, Silva ML, Loessner D, Heuzé-Vourc'h N, Clements JA. 2016. The kallikrein-related peptidase family: dysregulation and functions during cancer progression. *Biochimie* **122:** 283–299. doi:10.1016/j.biochi.2015.09.002

Lee JK, Phillips JW, Smith BA, Park JW, Stoyanova T, McCaffrey EF, Baertsch R, Sokolov A, Meyerowitz JG, Mathis C, et al. 2016. N-Myc drives neuroendocrine prostate cancer initiated from human prostate epithelial cells. *Cancer Cell* **29:** 536–547. doi:10.1016/j.ccell.2016.03.001

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15:** 550. doi:10.1186/s13059-014-0550-8

Machczyński P, Majchrzak E, Niewinski P, Marchlewska J, Golusiński W. 2020. A review of the 8th edition of the AJCC staging system for oropharyngeal cancer according to HPV status. *Eur Arch Otorhinolaryngol* **277:** 2407–2412. doi:10.1007/s00405-020-05979-9

Marur S, D'Souza G, Westra WH, Forastiere AA. 2010. HPV-associated head and neck cancer: a virus-related cancer epidemic. *Lancet Oncol* **11:** 781–789. doi:10.1016/S1470-2045(10)70017-6

Nakano T, Motoshita J, Tanaka R, Okabe M, Tamae A, Shiratsuchi H, Yasumatsu R, Nakashima T, Nakagawa T. 2017. Primary combined small cell carcinoma and squamous cell carcinoma of the oropharynx with special reference to EGFR status of small cell carcinoma component: case report and review of the literature. *Auris Nasus Larynx* **44:** 472–478. doi:10.1016/j.anl.2016.07.011

Nichols AC, Palma DA, Chow W, Tan S, Rajakumar C, Rizzo G, Fung K, Kwan K, Wehrli B, Winquist E, et al. 2013. High frequency of activating *PIK3CA* mutations in human papillomavirus–positive oropharyngeal cancer. *JAMA Otolaryngol Head Neck Surg* **139:** 617–622. doi:10.1001/jamaoto.2013.3210

Omori H, Sato K, Nakano T, Wakasaki T, Toh S, Taguchi K, Nakagawa T, Masuda M. 2019. Stress-triggered YAP1/SOX2 activation transcriptionally reprograms head and neck squamous cell carcinoma for the acquisition of stemness. *J Cancer Res Clin Oncol* **145:** 2433–2444. doi:10.1007/s00432-019-02995-z

Omori H, Nishio M, Masuda M, Miyachi Y, Ueda F, Nakano T, Sato K, Mimori K, Taguchi K, Hikasa H, et al. 2020. YAP1 is a potent driver of the onset and progression of oral squamous cell carcinoma. *Sci Adv* **6:** eaay3324. doi:10.1126/sciadv.aay3324

O'Sullivan B, Huang SH, Su J, Garden AS, Sturgis EM, Dahlstrom K, Lee N, Riaz N, Pei X, Koyfman SA, et al. 2016. Development and validation of a staging system for HPV-related oropharyngeal cancer by the International Collaboration on Oropharyngeal Cancer Network for Staging (ICON-S): a multicentre cohort study. *Lancet Oncol* **17:** 440–451. doi:10.1016/S1470-2045(15)00560-4

Palsgrove D, Bishop JA. 2020. HPV-related carcinoma of the oropharynx: challenges on small biopsy specimens. *J Am Soc Cytopathol* **9:** 359–368. doi:10.1016/j.jasc.2020.06.001

Patel AS, Yoo S, Kong R, Sato T, Sinha A, Karam S, Bao L, Fridrikh M, Emoto K, Nudelman G, et al. 2021. Prototypical oncogene family Myc defines unappreciated distinct lineage states of small cell lung cancer. *Sci Adv* **7:** eabc2578. doi:10.1126/sciadv.abc2578

Pytynia KB, Dahlstrom KR, Sturgis EM. 2014. Epidemiology of HPV-associated oropharyngeal cancer. *Oral Oncol* **50:** 380–386. doi:10.1016/j.oraloncology.2013.12.019

Quail DF, Joyce JA. 2013. Microenvironmental regulation of tumor progression and metastasis. *Nat Med* **19:** 1423–1437. doi:10.1038/nm.3394

Rouillard AD, Gundersen GW, Fernandez NF, Wang Z, Monteiro CD, McDermott MG, Ma'ayan A. 2016. The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database* **2016:** baw100. doi:10.1093/database/baw100

Saito Y, Hayashi R, Iida Y, Mizumachi T, Fujii T, Matsumoto F, Beppu T, Yoshida M, Shinomiya H, Kamiyama R, et al. 2020. Optimization of therapeutic strategy for p16 positive oropharyngeal squamous cell carcinoma: multi-institutional observational study based on the national Head and Neck Cancer Registry of Japan. *Cancer* **126:** 4177–4187. doi:10.1002/cncr.33062

Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, Kryukov G V, Lawrence MS, Sougnez C, McKenna A, et al. 2011. The mutational landscape of head and neck squamous cell carcinoma. *Science* **333:** 1157–1160. doi:10.1126/science.1208130

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci* **102:** 15545–15550. doi:10.1073/pnas.0506580102

Tarazona S, Furió-Tarí P, Turrà D, Di Pietro A, Nueda MJ, Ferrer A, Conesa A. 2015. Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. *Nucleic Acids Res* **43:** e140. doi:10.1093/nar/gkv711

Terzić J, Grivennikov S, Karin E, Karin M. 2010. Inflammation and colon cancer. *Gastroenterology* **138:** 2101–2114.e5. doi:10.1053/j.gastro.2010.01.058

Wakasaki T, Yasumatsu R, Masuda M, Matsuo M, Tamae A, Kubo K, Kogo R, Uchi R, Taura M, Nakagawa T. 2019. Small cell carcinoma in the head and neck. *Ann Otol Rhinol Laryngol* **8:** 000348941985360. doi:10.1177/0003489419853601

Wang F-H, Ma X-J, Xu D, Luo J. 2018. UPK1B promotes the invasion and metastasis of bladder cancer via regulating the Wnt/β-catenin pathway. *Eur Rev Med Pharmacol Sci* **22:** 5471–5480.

Wichmann G, Rosolowski M, Krohn K, Kreuz M, Boehm A, Reiche A, Scharrer U, Halama D, Bertolini J, Bauer U, et al. 2015. The role of HPV RNA transcription, immune response-related gene expression and disruptive *TP53* mutations in diagnostic and prognostic profiling of head and neck cancer. *Int J Cancer* **137:** 2846–2857. doi:10.1002/ijc.29649

Yuan H, Krawczyk E, Blancato J, Albanese C, Zhou D, Wang N, Paul S, Alkhilaiwi F, Palechor-Ceron N, Dakic A, et al. 2017. HPV positive neuroendocrine cervical cancer cells are dependent on Myc but not E6/E7 viral oncogenes. *Sci Rep* **7:** 45617. doi:10.1038/srep45617

Yuan J, Knorr J, Altmannsberger M, Goeckenjan G, Ahr A, Scharl A, Strebhardt K. 1999. Expression of p16 and lack of pRB in primary small cell lung cancer. *J Pathol* **189:** 358–362. doi:10.1002/(SICI)1096-9896(199911)189:3<358::AID-PATH452>3.0.CO;2-1

Zanconato F, Cordenonsi M, Piccolo S. 2019. YAP and TAZ: a signalling hub of the tumour microenvironment. *Nat Rev Cancer* **19:** 454–464. doi:10.1038/s41568-019-0168-y