Journal of Translational Medicine

RESEARCH

Open Access

Intratumoral spatiotemporal heterogeneity of HPV DNA status in HPV-associated oropharyngeal squamous cell carcinoma with clinical record

Bokyung Ahn[1†], Chae Won Park[1,2†], Joon Seon Song[1], Hee-Jin Lee[1], Wonkyung Kim[1,2], Yoon Se Lee[3], Young Ho Jung[3], Seung-Ho Choi[3], Soon Yuhl Nam[3], Sang Yoon Kim[3], Ji-Hye Oh[4], Young Gwang Kang[1,2], Da Eun Oh[4], Chang Ohk Sung[1,2,4*] and Kyung-Ja Cho[1*]

Abstract

Background Recently, the heterogeneity of human papillomavirus (HPV) status within HPV-associated oropharyngeal squamous cell carcinoma (HPV-OPSCC) has garnered substantial attention. However, the origin and relevance of a subset of cancer cells with HPV-loss remain largely unknown.

Methods We encountered seven HPV-OPSCCs with spatial heterogeneity identified by HPV-DNA in situ hybridization among 83 OPSCCs, including 48 HPV-OPSCCs. We analyzed the characteristics of HPV-loss and HPV-positive cancer cells using histomorphology and immunohistochemistry. Spatial transcriptomic profiling was performed to investigate the evolutionary dynamics and pathway alterations associated with the emergence of HPV-loss tumor cells, as well as the corresponding changes in the tumor microenvironment. Additionally, we evaluated the clinical prognosis of OPSCCs exhibiting intratumoral HPV-DNA heterogeneity.

Results We demonstrated that HPV-loss subsets originate from HPV-positive cancer cells and diverge into distinct lineages during spatiotemporal evolution. This transition was associated with hypoxia and PI3K/AKT/mTOR signaling. Although HPV-loss clones showed more aggressive features such as poor differentiation and strong p16 expression, the overall prognosis of patients with HPV-DNA spatial heterogeneity was comparable to those with homogeneous HPV status. This may be due to increased interferon signaling and elevated cytotoxic lymphocyte infiltration in HPV-loss clones, suggesting a loss of immune evasion. In contrast, HPV-positive clones retained immune evasion mechanisms, potentially mediated by Fibroblast Growth Factor Receptor 3 signaling.

†Bokyung Ahn and Chae Won Park contributed equally to this work.

*Correspondence:
Chang Ohk Sung
co.sung@amc.seoul.kr
Kyung-Ja Cho
kjc@amc.seoul.kr

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

Conclusions We identified the emergence of HPV-loss clones in HPV-OPSCC and uncovered their spatiotemporal evolution characteristics, as well as their impact on the tumor microenvironment.

Keywords Oropharyngeal squamous cell carcinoma, HPV, Spatial heterogeneity, Spatial transcriptomics, Microenvironment

Background

Human papillomavirus (HPV) has become a major etiological factor in oropharyngeal squamous cell carcinoma (OPSCC) since its identification by Gillison et al. [1]. HPV-associated OPSCC (HPV-OPSCC) has better sensitivity to radiation therapy and is clinically more favorable than HPV-negative OPSCC [2–6]. However, recent multi-center clinical trial results have shown that some patients experience poor outcomes despite aggressive treatment, suggesting that HPV-OPSCC comprises clinically and biologically heterogeneous subtypes [7].

HPV-OPSCC is characterized by a two-fold lower mutational load [8], higher proliferative index with upregulation of cell cycle genes [9], *p53* integrity, and more frequent alteration of the *PTEN/PIK3CA* pathway [10] than the HPV-negative OPSCC. Consequently, HPV-OPSCC is classified separately by the World Health Organization and the American Joint Committee on Cancer (AJCC) based on its distinct epidemiological, clinicopathological, and molecular characteristics [11].

Given that HPV status is crucial for patient management, several HPV detection methods have been developed at DNA, mRNA, and protein levels, and using more than one testing method may further increase the detection sensitivity [12]. According to the National Comprehensive Cancer Network, p16 immunohistochemistry (IHC) and polymerase chain reaction (PCR)-based assays are highly reliable in terms of sensitivity, while HPV-DNA in situ hybridization (ISH) elicits the highest specificity [13]. Compared with PCR-based assays, ISH can localize HPV to the neoplastic area and further distinguish episomal (diffuse staining) from integrated HPV-DNA (dot-like or punctuate positivity) [14]. Owing to the moderate specificity of p16 IHC, HPV is reportedly absent in 8–20% of patients with p16-positive (+) OPSCC, indicating that the presence of HPV should be confirmed using an alternate method [15, 16]. Therefore, combining a sensitive test (e.g., p16 IHC) with a specific test (e.g., ISH) was observed to have the best potential in detecting HPV status [14].

Unlike the well-known clinical implication of discrepant p16 IHC results and HPV status [17], intratumoral heterogeneity of p16 IHC staining or HPV status within the HPV-OPSCC has been under-recognized and poorly understood until recently. Puram et al. discovered a subset of cells within HPV-OPSCC with lost or repressed HPV expression (called *HPVoff* cells) [18].

Considering that HPV-DNA ISH is the gold standard for determining HPV positivity rather than RNA ISH, we confirmed HPV positivity using HPV-DNA ISH. Subsequently, we evaluated the peculiar differential findings between spatially segregated HPV-DNA positive (HPV-positive) and HPV-DNA loss (HPV-loss) components within HPV-OPSCC using HPV-DNA ISH and assessed their clinicopathologic significance. We performed spatial transcriptomics to investigate the evolutionary history of HPV-loss clones, confirming that spatial heterogeneity of cytotoxic lymphocytes interfaced with cancer cells is associated with HPV status.

Methods

Patient selection

All individuals diagnosed with OPSCC via biopsy or resection were selected from October 2020, when Asan Medical Center, Seoul, Korea, started performing high-risk HPV-DNA ISH, until August 2021. Patients lacking formalin-fixed paraffin-embedded (FFPE) blocks were excluded from this study. Finally, 83 individuals with OPSCC were included. Clinical data, including age, sex, procedure, distant metastatic status by pre-operative images, and survival outcomes, were obtained from electronic medical records. Clinical tumor staging was applied for individuals who underwent neo-adjuvant therapy without resection; pathologic tumor staging was applied for individuals who underwent surgical resection. HPV-associated and HPV-negative OPSCC were staged using the staging criteria of the AJCC [11]. Overall survival was evaluated using only disease-specific death, and recurrence was defined as newly appeared, local recurrence, or distant metastasis after the initial diagnosis. This study was reviewed and approved by the Institutional Review Board of Asan Medical Center (Approval number: 2021 – 1235).

Histological evaluation

Two pathologists (B.A. and K.C., head and neck specialists) carefully re-evaluated all hematoxylin and eosin (H&E)-stained slides of biopsy and resection samples of the primary and/or metastatic sites of all individuals. Tumor differentiation, tumor size, presence of lymphovascular and perineural invasions, margin status, and lymph node metastatic status were evaluated.

HPV-DNA ISH

HPV-DNA status was evaluated using the Ventana HPV III Family 16 Probe B, a cocktail recognizing high-risk HPV types 16, 18, 31, 33, 35, 45, 52, 56, 58, and 66. The test was interpreted following the manufacturer's guidelines. Determining a true positive required either the episomal (large, homogeneous, globular navy-blue precipitates within cell nuclei) or integrative pattern (discrete, stippled navy-blue nuclear pattern). Samples with no signal within carcinoma cells were deemed negative.

HPV genotyping analysis using semi-quantitative real-time PCR (sqPCR)

Samples that exhibited intratumoral spatial heterogeneity of high-risk HPV-DNA status were subjected to sqPCR analysis, which was performed to confirm the specific virus genotype and exclude the presence of any virus possibly not covered by the ISH probe. One sample was excluded because of insufficient tissue for the genotyping analysis. Nucleic acids were extracted from 10-μm (× 5) paraffin tissue sections, and the CFX96TM RT-PCR system (Bio-Rad Laboratories, Hercules, CA, USA) and Anyplex II HPV28 Panel A detection system (31744024; Seegene, Seoul, Korea) were used. Anyplex II HPV28 detection (A) can detect 19 high-risk HPV types (16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 69, 73, 82) and 9 low-risk HPV types (6, 11, 40, 42, 43, 44, 54, 61, 70).

IHC

Representative sections of both primary tumor and metastatic lymph nodes, if present, were subjected to immunohistochemical staining. In brief, 4-μm thick tissue sections were deparaffinized and hydrated in xylene and exposed to serial dilutions of ethanol. Endogenous peroxidase was blocked via incubation in 3% H_2O_2 for 10 min. Thereafter, heat-induced antigen retrieval was performed. Primary antibodies were used with a Benchmark autostainer (Ventana Medical Systems, Tusan, AZ, USA) following the manufacturer's protocol. Sections were incubated at room temperature for 32 min with primary antibodies for p16 (Clone E6H4; 1:6; VENTANA, USA), p53 (clone DO-7; 1:1000; DAKO, Denmark, Glostrup), RB Transcriptional Corepressor 1 (RB1, clone 3C8; 1:10000; QED Bioscience, California, USA), Ki-67 (clone MIB1; 1:200; DAKO), and CD8 (clone C8/144B; 1:400, Mouse monoclonal, catalog no. NCL-L-CD8-4B11, Novo, CA, USA). The sections were then labeled with an automated immunostaining system and an I-View detection kit (Benchmark XT; Ventana Medical Systems). The immunolabeled sections were lightly counterstained with hematoxylin, dehydrated in ethanol, and cleared in xylene. Positivity, intensity, and spatial distribution were all assessed by light microscopic examination and independently evaluated by two pathologists (B.A. and K.C.) who were blinded to the clinicopathologic information.

Image analysis for CD8 positive T cells

IHC images were analyzed using ImageJ version 1.54 g [19]. Total cell numbers were quantified based on hematoxylin-stained nuclei, and CD8 + T cells were identified based on positive (brown color) staining. Image analysis included color deconvolution to separate hematoxylin and IHC signals, threshold adjustment (including color thresholding for CD8 IHC-positive cells), and watershed segmentation for cell separation. The ratio of CD8 + T cells was calculated as the proportion of CD8 IHC-positive cells among total hematoxylin-stained nuclei.

Spatial transcriptomic profiling (10× genomics visium)

For spatial transcriptomics analyses, we used FFPE tissues of three available OPSCCs with spatial heterogeneity of HPV status with spatially segregated HPV-positive and HPV-loss components by HPV-DNA ISH. Selection of the 6.5 mm × 6.5 mm capture areas in H&E stained slides for spatial transcriptomics was based on the consensus of three expert pathologists (C.O.S., B.A., and K.C.). Tissue sections meeting quality control standards (DV200, the percentage of RNA fragments with ≥ 200 nucleotides, ≥ 30%) underwent spatial transcriptomics using the Visium CytAssist Spatial Gene Expression for FFPE platform from 10x Genomics Inc. Initial steps included deparaffinization, H&E staining, imaging, and subsequent decrosslinking. Subsequent decrosslinking, tissue permeabilization, probe hybridization, and library preparation were performed meticulously according to the manufacturer's protocol. Sample libraries were sequenced on an Illumina NovaSeq 6000 platform, using specified read lengths as recommended by the manufacturer.

Preprocessing of spatial transcriptomics data

Sequences and histology images were processed using a 10x Genomics Space Ranger v2.0.1. Illumina basecall files generated by the sequencing instrument were converted to FASTQ format for each sample using the 'mkfastq' command. Visium spatial expression libraries were analyzed with the count command. Image processing proceeded in two stages: first, the fiducial alignment grid of the tissue image was used to determine the orientation and position of the input image; subsequently, the region of the tissue covered on the slide was identified. Sequencing reads were aligned to the Visium Human Transcriptome Probe Set v2.0 GRCh38-2020-A reference using the STAR (v2.7.2a) aligner.

Manual histologic annotation at the individual spot level by pathologists

All spots were reviewed and annotated by two pathologists (B.A. and C.O.S.) based on the H&E staining and HPV-DNA ISH. Spots were categorized into four groups: HPV-loss cancer cells, HPV-positive cancer cells, normal epithelium, and stroma. Pathologist annotations were exported from a 10× Genomics Loupe Browser v6.4.1 and imported into Seurat for further analysis.

Processing and visualization of spatial transcriptomics data

For the spatial transcriptomics analysis, three patient samples (individual #1, individual #2, and individual #3) were processed using Seurat (v4.4.0) [20] in R (v4.1.2) to identify and filter out low-quality spots. The visium data from individual #3 exhibited lower nCount and nFeature values in the six rows at the bottom, indicating potential technical artifacts. Based on these findings, the six rows at the bottom were excluded from the analysis. Consequently, individual #1 and individual #2 retained 4,925 and 4,962 spots, respectively, and individual #3 retained 4,579 spots for downstream analysis. Normalization and scaling of the data were conducted using the NormalizeData and ScaleData functions from Seurat (v4.4.0). The selection of the principal component (PC) followed guidelines from the Harvard Chan Bioinformatics Core's single cell RNA sequencing (scRNA-seq) protocols. Elbow plot analysis was utilized to quantitatively determine the number of PCs, leading to the selection of 14 PCs for further analysis. To remove batch effects, harmony (v.1.2.1) [21] was employed, and clustering was subsequently performed using the FindClusters function with the Leiden algorithm at a resolution of 1.0. To prevent over-clustering, the single-cell significant hierarchical clustering (scSHC) package (v.0.1.0) was used [22]. This integrated approach yielded seven large unsupervised clusters from a total of 14,466 spots across the three samples. Then, the clusters were annotated using histology annotation information. Spots corresponding histologically to stroma or normal epithelium were categorized into stroma or normal clusters, respectively. Finally, we annotated clusters for cancer spots as HPV_*pos* or HPV_*loss*_C# (e.g., HPV_pos_C1), incorporating both transcriptomic and histological features.

Pathway analysis

Differentially expressed genes (DEGs) between HPV-positive and HPV-loss tumor spots were identified, and Gene Set Enrichment Analysis (GSEA) with Reactome pathway data from the Molecular Signature Database (MSigDB) was performed using the ClusterProfiler (v4.8.3) R package. GSEA results were visualized using the enrichplot (v1.20.3) R package. ClusterProfiler was used to perform Reactome enrichment analyses for the DEG set. The permutation or hypergeometric test results were adjusted for *P* values using a false discovery rate. For Hallmark and Gene Ontology Biological Process (GOBP) pathway analyses, differential expression analysis was conducted using the FindMarkers function of the Seurat package (v4.4.0), with the DESeq2 method to identify genes showing significant differences between the HPV-positive and HPV-loss clusters for each individual. To explore functional enrichment, Hallmark gene sets from MSigDB ("H") and GOBP gene sets from MSigDB ("C5") were utilized. Enrichment analyses were performed separately for each HPV group using a gene list aggregated by the mean, applying the enricher function of the clusterProfiler package (v4.2.2), with the corresponding expressed gene lists as the background universe. This integrated approach combines differential expression analysis with functional enrichment to provide insights into the biological implications of HPV-DNA status.

Copy number alteration (CNA) analysis

Copy number alterations (CNAs) were inferred using the inferCNV tool (v1.21.0) [23] for each subset of samples derived from the merged Seurat object. The stroma and normal epithelium clusters were used as the normal reference group for inferCNV. Copy number variant (CNV) prediction was performed using a Hidden Markov Model (HMM) approach, and hierarchical clustering was executed using the 'ward.D2' method. Detailed parameter settings included: cutoff = 0.1, cluster by groups = FALSE, analysis mode = 'subclusters,' tumor subcluster partition method = 'random_trees,' scale data = TRUE, denoise = TRUE, noise filter = 0.12, and HMM type = 'i6.' Gene expression values for each cell were constrained between 0 and 2, centered at 1. The CNV score for each cell was calculated as the quadratic sum of CNVregion – 1 [24]. Evolutionary tree reconstruction was conducted using Uphyloplot2 (v2.3) [25], based on the inferCNV output file. Copy number gain and loss were visualized with GenVisR (v1.35.3) [26].

Pseudotime trajectory analysis

To infer pseudotime trajectories, the dyno package (v0.1.2) [27] was utilized, adhering to the provided pipelines and employing the ti_paga function, which implements the Partition-based graph abstraction (PAGA) method. The wrap_expression function was used to structure the gene expression data as input, with parameters set according to reproducibility guidelines: n_cells = 7,193 (excluding stroma), n_features = 18,085, and prior_information = c('start_id' and 'groups_id'). These settings were standardized using the guidelines_shiny function, which incorporates an interactive Shiny application to optimize method selection. The 'start_id' parameter was set to the 'Normal_epithelium' start cell list. Because the

input object included merged samples, the parameter 'multiple_disconnected' was set to TRUE to accommodate the possibility of distinct origins. In addition, spatial trajectories were examined using stLearn (v0.4.12) [28] in accordance with the downstream analysis guidelines specific to spatial trajectory inference. Individual samples were subsetted from the merged Seurat object, and the 'iroot' parameter of the trajectory.pseudotime function was specified for the 'Normal_epithelial,' 'HPV_loss_C4,' and 'HPV_pos_C7' clusters. These clusters were ordered based on each individual's PAGA pseudotime values. Given the potential for status transitions, such as HPV-positive to HPV-loss or vice versa, max_nodes was set to 5 for individual #1 and 4 for the other samples. This differentiation was due to individual #1 uniquely containing a 'Normal_epithelial' group, allowing the examination of the trajectory between HPV-positive and HPV-loss clones with a higher level of detail.

Tumor-immune type deconvolution

Immune cell types in tumors were analyzed using MCPcounter (v1.2.0) [29] in spatial transcriptomics, as described previously [30]. The output consisted of tissue-infiltrating immune and stromal cell population scores, which were subsequently used for visualizing spatial clustering. Given that MCPcounter is a reference-free cell-type deconvolution approach, we used a reference-based cell-type deconvolution approach, such as robust cell-type decomposition (RCTD) [31]. For the RCTD analysis, we utilized a public scRNA-seq dataset of oropharyngeal carcinoma [18] as the reference to obtain cell distribution in the spatial regions. To map the cell types detected in the reference dataset to spatial data, we performed RCTD using the spacexr R package (v2.2.1), with the parameter doublet_mode = 'full.'

Identifying cell-cell interactions using stLearn

Cell-cell interactions were examined using stLearn (v0.4.12) [28], a comprehensive toolkit designed for analyzing spatial transcriptomic data in Python (v3.8.12). To avoid the selection of non-meaningful nodes, the Cell-Cell Interaction Analysis pipeline was applied to each sample subset, excluding the 'stroma' using transcriptome annotation. The tl.cci.run function was employed with parameters set to min_spots = 20, distance = None, and n_pairs = 10,000, as per the recommended guidelines. The ligand-receptor dataset used for this analysis was 'connectomeDB2020_lit' for human samples. The significance scores of ligand-receptor interactions, as produced by the tl.cci.run function, were extracted to evaluate ligand-receptor significance within each spot.

Signature scoring

Signature scores were calculated using the singscore package (v1.19.1) [32] based on gene lists derived from the signature gene sets, including interferon, partial epithelial-mesenchymal transition (pEMT), HPV*off*, and HPV*on* signature, reported by Puram et al. [18]. Considering our dataset, the Hallmark gene sets from the msigdbr package (v7.5.1) were applied.

Public data preprocessing and survival analysis using signature scoring

We analyzed six public datasets for head and neck cancer containing RNA sequencing data, survival data, and HPV positivity [33–38]. The GSE65858, GSE208576, and GSE117973 datasets underwent preprocessing, including log2 transformation to normalize distribution, followed by robust scaling normalization to mitigate the impact of outliers. Given the limitations in determining HPV heterogeneity in public datasets, we utilized the HPV*off* signature score within HPV-associated tumor groups. The upper quartile (0.75) was applied as a cutoff to classify HPV-heterogeneous and HPV-homogeneous groups among HPV-associated tumors, considering the actual proportion of HPV heterogeneity in our study. To validate the HPV*on* signature, we compared HPV*on* signature scores between HPV-associated and HPV-negative tumors.

Statistical analysis

Continuous variables were compared using the independent Student's *t*-test or Wilcoxon rank sum test. Categorical variables were examined using the χ^2 or Fisher exact tests. The correlation between the two groups was examined using either the Pearson correlation or Spearman correlation test. Survival curves were estimated using the Kaplan–Meier method, and differences between curves were compared using the log-rank test. Multivariate Cox proportional hazards regression analyses were performed to evaluate the prognostic potential of the factors. The relationship between multiple variables was analyzed using multivariate linear regression analysis. All statistical evaluations were performed using R (version 4.0.3; The R Foundation for Statistical Computing, Vienna, Austria). Statistical significance was set at $P < 0.05$.

Results

Study cohort and design

HPV-DNA ISH was performed on whole-slide images of 83 OPSCC patients diagnosed over approximately one year (2020–2021) at Asan Medical Center, Seoul, Korea. Among them, 48 individuals were confirmed to be HPV-DNA ISH positive, named HPV-OPSCC. Of these 48 individuals, seven exhibited heterogeneities in HPV-DNA ISH results (HPV-heterogeneous OPSCC)

(Fig. 1A). We compared the clinical and pathological characteristics of these individuals. For three of the seven patients with HPV-heterogeneous OPSCC who underwent surgical resection, spatial transcriptomics was conducted to compare and analyze molecular and tumor microenvironmental characteristics of HPV-positive and HPV-loss cancer cell subsets (Fig. 1B). HPV positivity was determined based on HPV-DNA ISH.

Clinicopathologic characteristics of HPV-OPSCC with HPV-DNA intratumoral spatial heterogeneity

The clinicopathologic features of the seven patients with HPV-OPSCC and spatial heterogeneity determined using HPV-DNA ISH are summarized in Fig. 2A. All seven individuals were male, primarily aged between 60 and 70 years, with OPSCC located in the tonsil. Six individuals underwent HPV genotyping, revealing a single HPV type of 16, 18, and 59 in 4, 1, and 1 individual, respectively (Supplementary Fig. S1). Of the seven individuals, four exhibited differential histomorphologic characteristics between the HPV-positive and HPV-loss components; the other three individuals had markedly limited biopsy tissue specimens to perform a comparison and did not show frank histomorphologic distinction between components.

Tumor tissue specimens derived from three individuals who underwent surgical resection were subjected to spatial transcriptomics analysis, yielding expression values for 14,466 genes from an average of 4,946 spots per individual (Fig. 2B). The spots of HPV-positive and HPV-loss tumor areas were annotated by two pathologists (B.A. & C.O.S.) based on HPV-DNA ISH results and histological images. This result correlated well with the transcriptome-based unsupervised clustering results (Supplementary Fig. S2). HPV-DNA sqPCR was performed on microdissected HPV-positive and HPV-loss areas from tumor tissues of three patients. HPV-loss areas exhibited either HPV negativity ($n=1$) or a reduced HPV signal ($n=2$) compared with HPV-positive areas (Supplementary Fig. S3).

Comparing the expression patterns of p16 (*CDKN2A*), Ki-67, and p53 genes, known to differentiate HPV-associated and HPV-negative OPSCC, their expression levels differed between HPV-positive and HPV-loss cancer cells in HPV-heterogeneous OPSCC (Fig. 2C). For example, in individual #3, the HPV-positive component was located in the deep portion of the tonsil; the HPV-positive component showed marked pleomorphism, while the HPV-loss component showed a relatively uniform morphology. Although both HPV-positive and -loss components showed p16 positivity, the HPV-positive component displayed less intensity than the HPV-loss component. While the HPV-positive component had a Ki-67 labeling index of 60–70% and a wild-type expression pattern in p53 IHC, the HPV-loss component had a Ki-67 labeling index of ˃ 95% and a null mutation pattern in p53 IHC. Finally, RB1 IHC showed a partial loss in both components. These features were similar in the other individuals with differential histomorphologic features between the HPV-positive and -loss components, although the

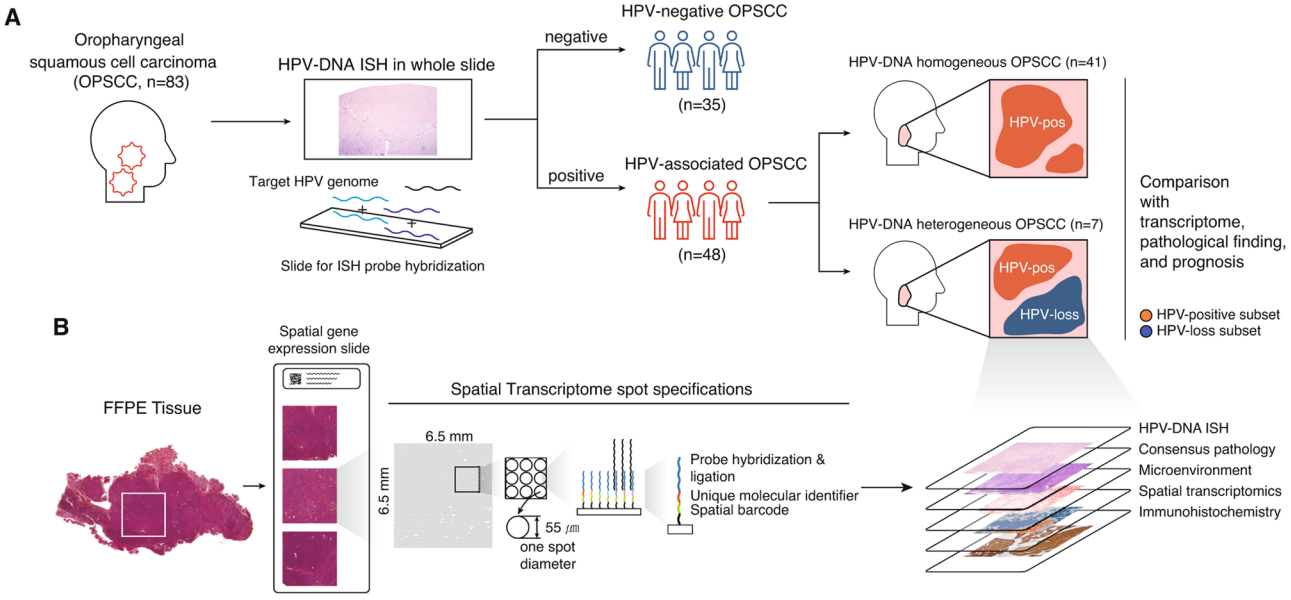

Fig. 1 Study cohort and design. **(A)** Tumors of 83 individuals diagnosed with OPSCC were classified using HPV-DNA ISH, and tumors exhibiting heterogeneity in HPV-positivity within the tumor tissue were further identified and categorized. **(B)** Spatial transcriptomics using the Visium platform was performed on OPSCCs exhibiting HPV heterogeneity, and the results were compared and analyzed alongside histology, immunohistochemistry, and HPV-DNA ISH findings. HPV, human papillomavirus; OPSCC, oropharyngeal squamous cell carcinoma; ISH, in situ hybridization

A

Individual	Sex/Age	Procedure	Location	HPV type by RT-PCR	T stage	N stage and intranodal HPV status	Histologic features		p16 IHC		Ki-67 IHC		p53 IHC		Rb1 IHC
							HPV- pos	HPV- abs	HPV- pos	HPV- abs	HPV- pos	HPV- abs	HPV- pos	HPV- abs	
1	M/63	Resection	Right tonsil	16 (++)	pT1	pN1, HPV+, heterogeneous	More pleomorphic	Less pleomorphic	Weak	Strong	Lower	Higher	Wild type pattern	Null-mutation pattern	Partial loss
2	M/61	Resection	Right tonsil	18 (++)	pT2	pN0									
3	M/73	Resection	Right tonsil	16 (++)	pT2	pN0									
4	M/54	Biopsy	Right tonsil	16 (++)	cT2	cN1, HPV+, homogeneous									
5	M/73	Biopsy	Right tonsil	59 (++)	ypT1	ypN2, HPV+, heterogeneous	Not significantly different between the heterogeneous components		Strong		High		Heterogeneous		
6	M/63	Biopsy	Left tonsil	16 (+++)	cT3	cN2, No Bx									
7	M/63	Biopsy	Right tonsil	NA	cT2	cN3, No Bx									

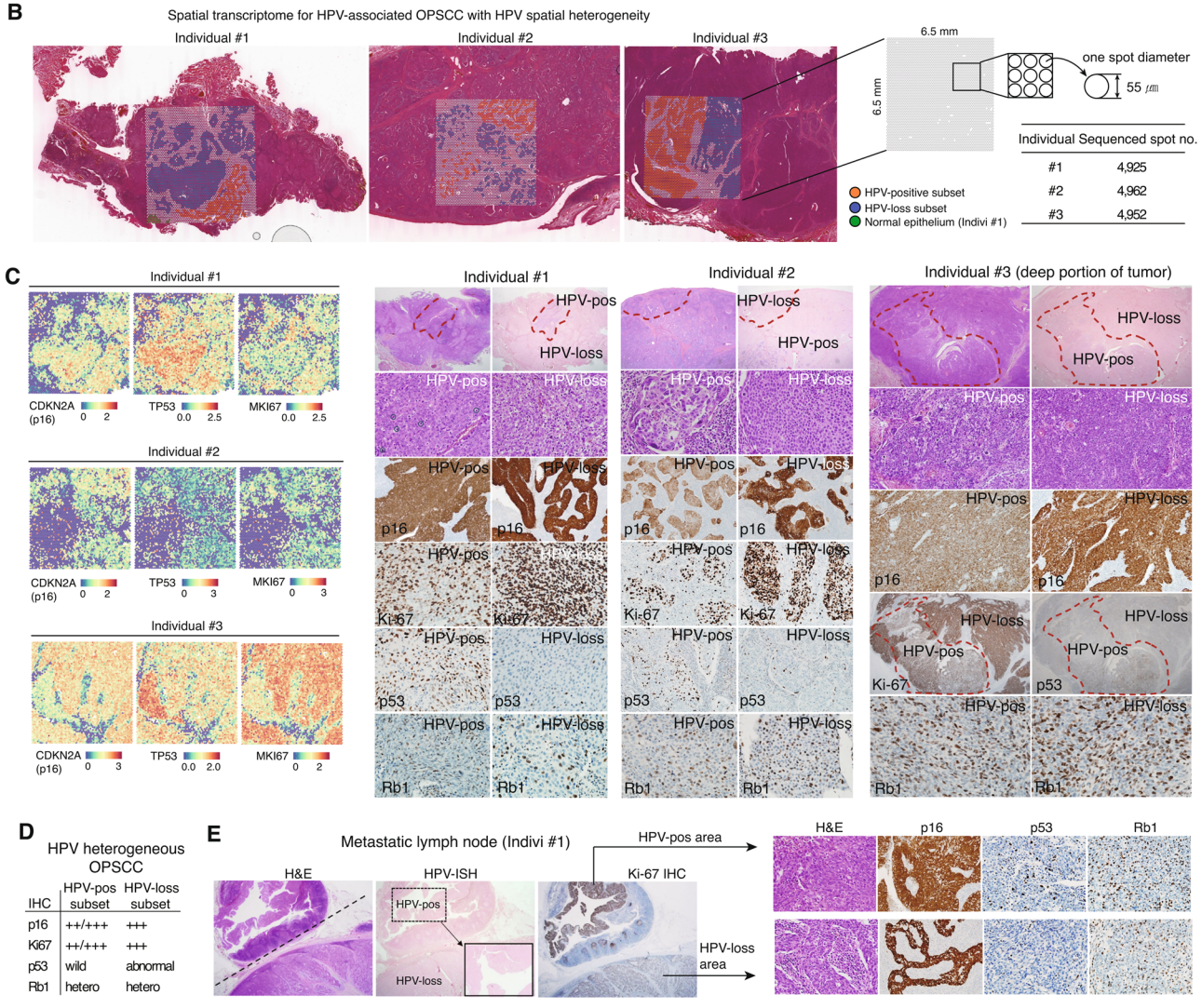

Fig. 2 (See legend on next page.)

location of HPV-positive components was relatively superficial in the other two resected patients. Altogether, HPV-loss tumor cells showed a tendency of higher p16 and Ki67 expression, along with an abnormal (null) p53 protein expression pattern (Fig. 2D). Two patients (individuals #1 and #5) underwent dissection of metastatic lymph nodes and showed spatial heterogeneity of HPV-DNA ISH, similar to the primary lesion. Two metastatic

(See figure on previous page.)

Fig. 2 Characteristics of HPV-associated OPSCC with HPV spatial heterogeneity. **(A)** Clinicopathological characteristics of seven individuals exhibiting HPV spatial heterogeneity. (**B)** Spatial transcriptomics was performed on three individuals with HPV-associated OPSCC who underwent surgical resection. HPV-positive and HPV-loss cancer cells form spatially distinct subsets. **(C)** Histological and gene expression characteristics of tumor tissues from three individuals who underwent tumor resection. The cancer morphology and immunoprofile of p16, Ki-67, p53, and RB1 differ between HPV-DNA ISH-positive (HPV-positive) and HPV-DNA ISH-negative (HPV-loss) subsets. **(D)** Summary of immunoprofiles of HPV-positive and HPV-loss cancer cells. **(E)** Histologic features and immunohistochemical results of metastatic lymph nodes. Low power view of HE-stained and HPV-DNA ISH of two metastatic lymph nodes shows one lymph node with HPV-positive (upper left) and the other with HPV-loss (lower right) components, respectively. The HPV-positive component shows a solid sheet-like appearance, while the HPV-loss component shows trabecular architecture with entrapped inflammatory cells. p16 IHC shows block positivity for both components but stronger intensity in the HPV-loss component. p53 IHC shows a heterogeneous pattern in the HPV-positive component and a null mutation pattern in the HPV-loss component. RB1 IHC shows partial loss in both HPV-positive and HPV-loss components. In contrast to the primary lesion, the Ki-67 labeling index value is markedly higher in the HPV-positive component. HPV, human papillomavirus; OPSCC, oropharyngeal squamous cell carcinoma; RB1, RB transcriptional corepressor 1; ISH, in situ hybridization; IHC, immunohistochemistry; HE, hematoxylin and eosin

lymph nodes of individual #1 were each composed of either HPV-positive or HPV-loss cancer cells (Fig. 2E); the HPV-positive component appeared solid with oval to spindle-shaped tumor cells, while the HPV-loss component showed round-shaped tumor cells with intercellular bridge and peripheral palisading. Similar to the primary site, p53 expression showed a wild expression pattern in the HPV-positive component and a null mutation pattern in the HPV-loss component. In contrast to the primary site, the HPV-positive component had a higher Ki-67 labeling index than the HPV-loss component. Overall, these findings suggested that the HPV-loss subset in HPV-heterogeneous OPSCC is a unique tumor subset with distinct characteristics. The metastatic lymph node of individual #5, who underwent neo-adjuvant therapy, showed multifocally scattered islands of HPV-positive and HPV-loss components (Supplementary Fig. S4). The HPV-positive component showed more pleomorphism with basaloid appearance, while the HPV-loss component showed vague keratinization with a comparatively homogeneous appearance with a lower nuclear/cytoplasm ratio. p16 IHC was almost negative in the HPV-positive components while showing block positivity in the HPV-loss components. Limited tumor tissue availability restricted the performance of p53, RB1, and Ki-67 IHC. Both HPV-positive and HPV-loss tumor cells were non-dominant in metastatic lymph nodes.

Increased infiltration of intratumoral cytotoxic lymphocytes in HPV-loss tumor spots

Upon pooling data from the three individuals for GSEA using whole gene expression levels, Major histocompatibility complex (MHC) class I-mediated antigen processing signaling and interferon signaling were enriched in HPV-loss spots (Fig. 3A). In pathway analysis based on DEGs between HPV-positive and HPV-loss spots from each individual, inflammatory signaling, including interferon signaling, was increased in HPV-loss spots compared with that in HPV-positive spots (Fig. 3B and Supplementary Fig. S5A-B). This finding suggested that the mechanism of interferon response evasion by the virus is enriched in HPV-positive cells, and as they transition to HPV-loss cells, this immune evasion mechanism diminishes, leading to an increase in interferon response and infiltration of cytotoxic lymphocytes.

Given that the diameter of a single spot was 55 μm and comprised several cells, we performed deconvolution analysis at the spot level to characterize immune cells interfacing with cancer cells in the tumor area (Fig. 3C). Among the various immune cell types, infiltration of cytotoxic lymphocytes was higher in HPV-loss spots than that in HPV-positive spots (Fig. 3D and Supplementary Fig. S5C). Intratumoral infiltration of CD8 + lymphocytes on whole-slide analysis showed an inverse pattern relative to HPV-DNA ISH positivity (Fig. 4A and Supplementary Fig. S6A-C).

Calculating the HPV*off* signature score using the HPV*off* signature gene set identified by Puram et al. [18], we detected a significant correlation between the HPV*off* signature and HPV-loss spots (Fig. 4B). Upon investigating the pathways used by Puram et al. [18], the interferon signature differed most significantly between HPV-loss and HPV-positive spots (Fig. 4C). Collectively, HPV-loss status in HPV-OPSCC appears to be associated with the loss of HPV expression similar to the HPV*off* state, which is related to increased infiltration of intratumoral cytotoxic lymphocytes (Fig. 4D).

Spatiotemporal evolution of tumor clusters between HPV-positive and HPV-loss subsets

Using unsupervised clustering of the transcriptome, we identified subclones (clusters) based on HPV-DNA status (Supplementary Fig. S7). Subsequently, pseudotime trajectory analysis was performed to determine the relative differentiation order of each cluster (Fig. 5A and Supplementary Fig. S8). Based on pseudotime values, we mapped the spatiotemporal evolutionary path among tumor clusters; as clusters progressed, the infiltration of cytotoxic lymphocytes was dynamically altered (Fig. 5A). Upon jointly analyzing tumor tissues from the three individuals, the spatial transcriptomic tumors exhibited distinct temporal order (Fig. 5B). The tumor tissue from individual #1, which included normal squamous epithelium and had a superficially located tumor,

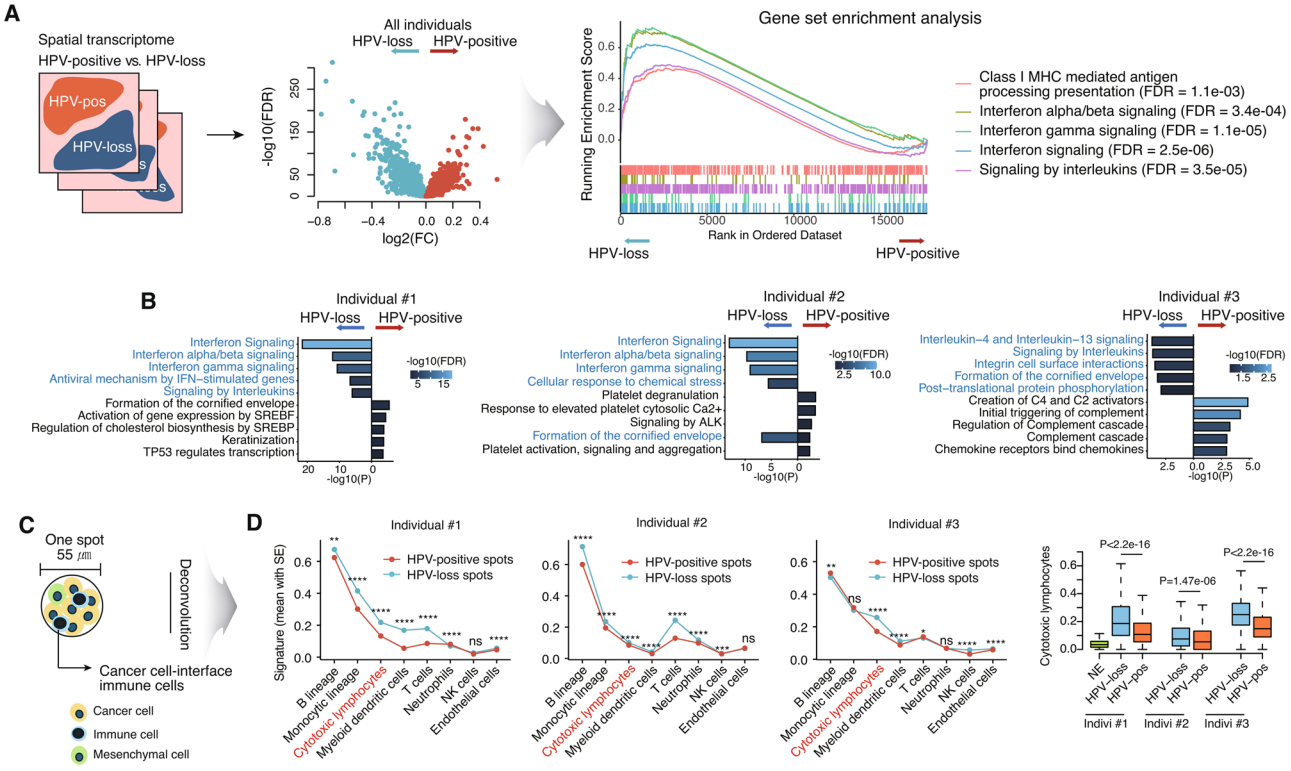

Fig. 3 Enriched CD8 + cytotoxic lymphocyte infiltration in HPV-loss spots within HPV heterogeneous OPSCC. **(A)** Data pooled from the three individuals show enriched MHC class I signaling and interferon signaling in the HPV-loss clone (GSEA). **(B)** Pathway analysis using differentially expressed genes between HPV-positive and HPV-loss spots shows that interferon and interleukin signaling are enriched in the HPV-loss cluster. **(C)** Each sequenced spot, with a diameter of 55 μm, consists of several cells. Therefore, deconvolution was performed at the spot level to profile immune cells closely associated with cancer cells. **(D)** In all three individuals, cytotoxic lymphocyte infiltration is higher in HPV-loss clones (Wilcoxon-rank sum test, two-sided). HPV, human papillomavirus; OPSCC, oropharyngeal squamous cell carcinoma; GSEA, gene set enrichment analysis

was the earliest in the temporal sequence. In the tumor area of individual #3, the temporal sequence was located in the deep portion of the tumor histologically, aligning well with the most progressed state among tissues from the three individuals in pseudotime. Based on this integrated tree, albeit with limitations, it appears that in HPV-OPSCC, the HPV-positive clone first emerged from normal epithelium infected with HPV, followed by the development of clones transitioning into HPV-loss clones (Fig. 5C). Although reversion from HPV-loss to HPV-positive appeared possible, this may be due to differences in proliferation rates between HPV-positive and HPV-loss clones, leading to a relative time discrepancy in the evolutionary process, potentially limiting interpretation based on the pseudotime analysis results. However, for individual #3, the HPV-positive clone was detected deep within the tumor tissue, suggesting the possibility of HPV reactivation. Despite considerable cytotoxic lymphocyte infiltration in the HPV-loss component, it appears to decrease with tumor progression (Fig. 5D). Accordingly, the reduction in the lymphocyte infiltration gap between HPV-positive and HPV-loss components in individuals #2 and #3, compared with individual #1, may be related.

Evolutionary trees based on CNAs suggested that HPV-positive and HPV-loss clones represent distinctive subgroups with different differentiation lineages (Fig. 5E, Supplementary Fig. S9-S10). These findings indicated that, after the initial development of HPV-OPSCC from HPV-infected normal epithelium, HPV-loss clones emerge and expand, eventually progressing into a unique subgroup evolutionary lineage. To define cancer cells and observe their progression patterns, we analyzed inferCNV, PAGA pseudotime, and stLearn pseudotime, with the results revealing their correlation (Supplementary Fig. S11). As cancer cells progressed over time, CNA gradually increased, resulting in a decreasing trend in cytotoxic lymphocyte infiltration (Fig. 5F). Given that CNA and individual-specific characteristics were associated with lymphocyte infiltration within the tumor, HPV-DNA status remained independently associated with regional intratumoral cytotoxic lymphocyte infiltration even after adjusting for these two factors (Fig. 5G).

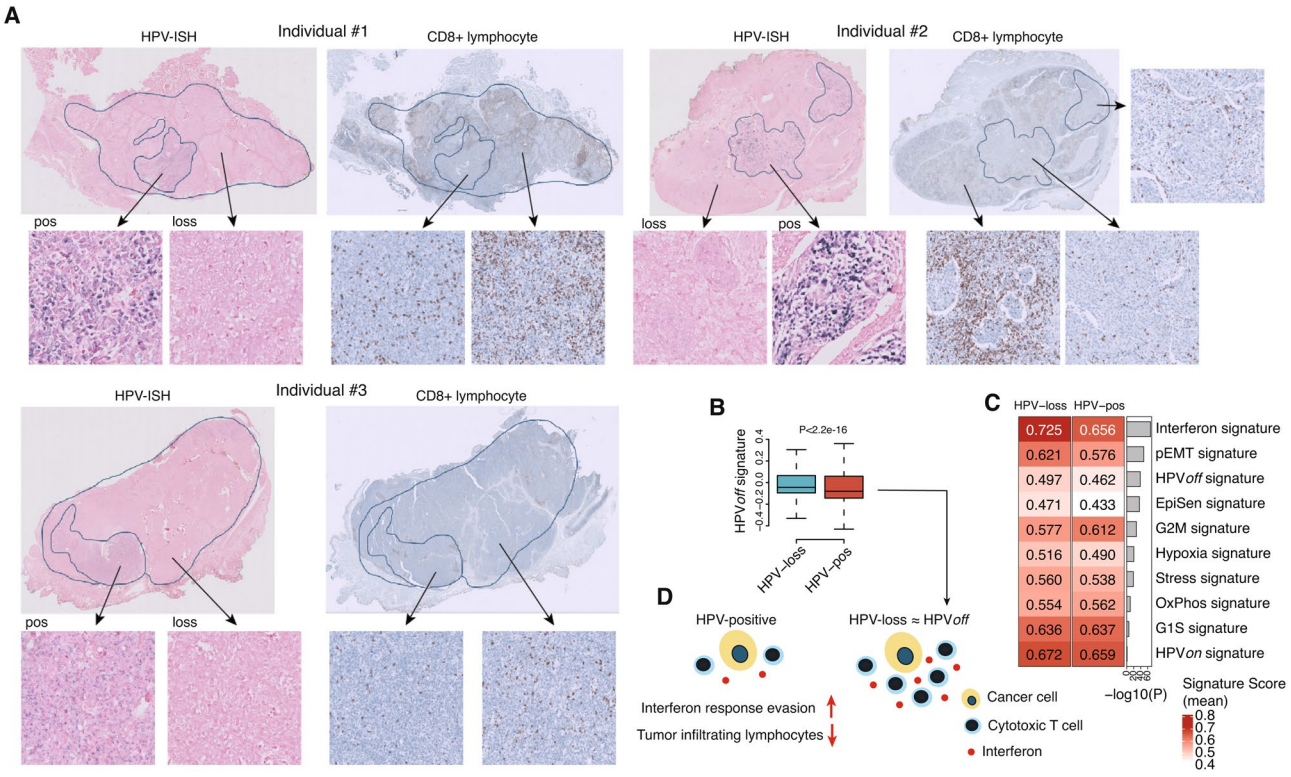

Fig. 4 Immune microenvironment and interferon response associated with HPV-loss in OPSCC. **(A)** Comparison of HPV-DNA ISH and CD8 IHC results on the original whole slide used for Visium spatial transcriptome analysis. Increased intratumoral infiltration of CD8 + lymphocytes can be observed in HPV-DNA ISH-positive areas. **(B)** HPV-loss subset in this study is associated with the HPV*off* signature (Wilcoxon-rank sum test, two-sided). **(C)** Among the signatures examined in HPV-associated OPSCC, the interferon signature shows the greatest difference between HPV-positive and HPV-loss spots (Wilcoxon-rank sum test, two-sided). **(D)** HPV-loss subset exhibits decreased evasion of the antiviral interferon response and increased infiltration of cytotoxic lymphocytes. HPV, human papillomavirus; OPSCC, oropharyngeal squamous cell carcinoma; ISH, in situ hybridization; IHC, immunohistochemistry

Pathway analysis between HPV-positive and HPV-loss clusters

The ligand-receptor interaction patterns in interfaced cells also differed between HPV-positive and HPV-loss clusters (Fig. 6A). Fibroblast Growth Factor Receptor 3 (FGFR3) signaling was frequently detected in HPV-positive clusters; FGFR3 was the most prevalent among all ligand-receptor interactions and FGFR3 ligands included FGF1, FGF2, FGF3, and FGF7 (Fig. 6B). This finding suggested that FGFR3-mediated signaling may be associated with the reduced infiltration of cytotoxic lymphocytes in HPV-positive clones. The FGFR3/CD8A expression ratio was higher in HPV-positive clones, and as FGFR3 expression increased, the infiltration of cytotoxic lymphocytes decreased (Fig. 6C). Considering individual #2, exhibiting the simplest cluster composition, pathway analysis according to HPV positivity showed an increase in interferon signaling, epithelial-mesenchymal transition (EMT), and response to virus-related pathways in the HPV-loss clusters, whereas the HPV-positive clusters exhibited an upregulation of cell cycle and keratin differentiation pathways (Fig. 6D). Upon examining pathway changes along the pseudotime differentiation of these clusters associated with HPV-DNA status change, pEMT, interferon, hypoxia, and phosphatidylinositol 3-kinase (PI3K)/AKT/mTOR signaling pathway were significantly altered (Fig. 6E). The hypoxia and PI3K/AKT/mTOR signaling pathways showed a significant correlation (Fig. 6F). Collectively, the transition from HPV-positive to HPV-loss status was associated with hypoxia and the PI3K/AKT/mTOR signaling pathway (Fig. 6G). After transitioning to the HPV-loss state, pEMT signaling increased sharply (Fig. 6G). In the combined analysis with the remaining two individuals, hypoxia signaling was higher in the HPV-loss clusters (Fig. 6H).

Clinical significance of HPV-OPSCC with HPV spatiotemporal heterogeneity

Detailed clinicopathologic features of the 83 individuals with OPSCC are summarized in Supplementary Table 1. The mean patient age was 58.5 (median, 63; range, 37–89) years (Fig. 7A). Of the 83 individuals, 76 (91.6%) were males and 7 (8.4%) were females. The mean tumor size was 2.6 ± 0.7 cm. Of these, 16 (19.3%) were pT1, 37 (44.6%) were pT2, 12 (14.5%) were pT3, and 17 (20.6%) were patients with pT4. Of the 83 individuals, 62 (74.7%)

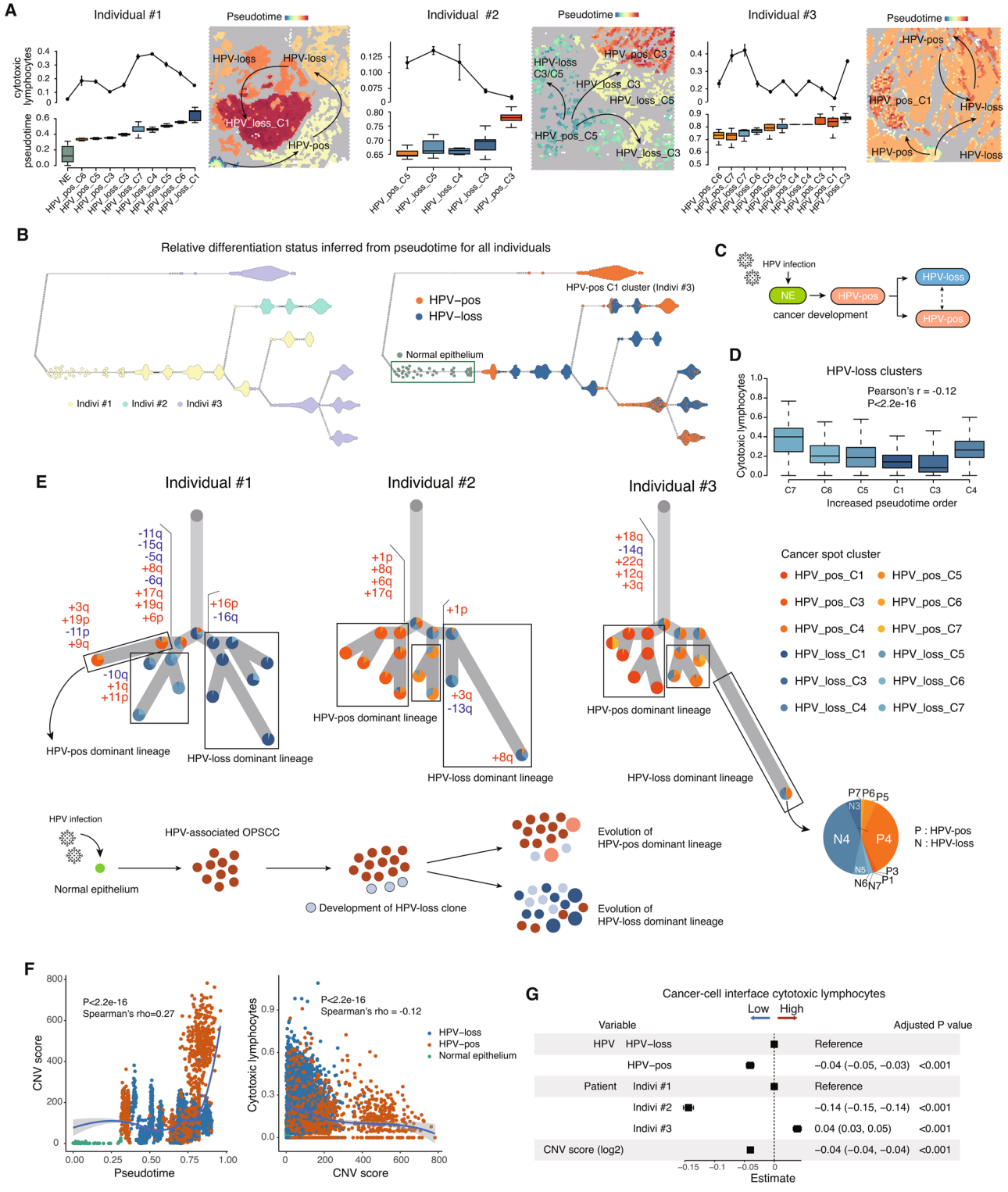

Fig. 5 (See legend on next page.)

and 2 (2.4%) presented with lymph node metastasis and distant metastasis at the time of diagnosis, respectively. Fifty-one (61.4%) individuals underwent surgical resection, and 13 (15.7%) received neo-adjuvant treatment before surgical resection. Thirty-two (38.6%) individuals were biopsied for diagnosis only, without further surgical resection, and of those, 26 (31.3%) underwent neo-adjuvant treatment with or without adjuvant therapy. The

(See figure on previous page.)

Fig. 5 Spatiotemporal evolution of tumor clusters. **(A)** Evolutionary history of subclusters in space through pseudotime differentiation shows dynamic changes in the infiltration of intratumoral cytotoxic lymphocytes as the tumor subclusters progress. **(B)** Combined evolutionary tree of tumor tissue samples from the three individuals shows that the tumor in individual #3 is relatively the most progressed, while the tissue from individual #1, which is a superficially located tumor, including adjacent normal epithelium, represents a relatively early stage of the cancer progression. **(C)** Changes in HPV-DNA status in the combined evolutionary tree suggest that the HPV-positive cluster may initially arise from the normal epithelium and transition to the HPV-loss cluster. **(D)** In HPV-loss clusters, as cancer cells progress, there is a tendency for cytotoxic lymphocyte infiltration to decrease gradually (Pearson correlation test, two-sided). **(E)** Upon examining evolution based on CNAs, HPV-positive and HPV-loss clusters follow separate evolutionary trees. **(F)** With advancing pseudotime progression, CNAs in cancer cells increase while cytotoxic lymphocyte infiltration decreases (Spearman correlation test, two-sided). **(G)** Even after adjusting for factors that may influence cytotoxic lymphocyte infiltration in tumor tissue, such as individual variations and CNAs, HPV remains independently associated with cytotoxic lymphocyte infiltration in the tumor (Multivariate linear regression analysis). HPV, human papillomavirus; CNAs, copy number alterations

median follow-up period was 20 (range, 1–36) months. Fourteen individuals had disease-specific deaths, and six experienced disease recurrence. p16 IHC was positive in 65 individuals (78.3%), and this was associated with HPV-DNA ISH positivity (Fig. 7B). HPV-OPSCC was detected at an earlier stage than HPV-negative OPSCC (Fig. 7B). Although HPV-OPSCC was detected at a younger age, there was no difference in tumor size (Fig. 7C). Patients with HPV-OPSCC had a better prognosis than those with HPV-negative OPSCC (Fig. 7D), and even after adjusting for factors associated with HPV positivity (age and clinical stage), HPV positivity remained independently associated with favorable survival (Fig. 7E). However, there was no significant difference in survival rates between HPV-OPSCC with HPV spatial heterogeneity and homogeneity (Fig. 7F). Additionally, using the HPV*off* gene set identified by Puram et al. [18], we analyzed the impact of the HPV*off* signature on the prognosis of HPV-associated head and neck cancer in 633 individuals across six public datasets. No statistically significant differences were detected (Supplementary Fig. S12).

Summary of spatial heterogeneity in the tumor microenvironment by HPV heterogeneity

HPV infection of the oropharyngeal normal epithelium, resulting in the development of HPV-OPSCC, may result in the subsequent emergence of HPV-loss clones that have undergone HPV loss. These clones undergo increased infiltration of cytotoxic lymphocytes owing to the loss of antiviral immune evasion mechanisms, including interferon signaling evasion. These changes in HPV phenotype status and the associated differences in regional immune cell infiltration contribute to increased spatial heterogeneity within the tumor (Fig. 8).

Discussion

In our cohort of 83 individuals with OPSCC, more than three-fourths were p16-positive, and more than 50% were positive for HPV-DNA ISH. We primarily focused on HPV-OPSCCs with spatially distinct areas of HPV-positive and HPV-loss components. The HPV-positive and HPV-loss components showed distinct histologic and molecular profiles.

p16 immunostaining is used as a surrogate marker for HPV-associated cancer. Herein, positive p16 immunostaining was significantly associated with HPV-DNA ISH positivity; however, a few individuals tested positive for p16 despite being negative for HPV-DNA ISH. This suggests that p16 IHC positivity may occur in a few HPV-negative OPSCCs. D'Souza et al. [39] reported individuals tested HPV-negative among individuals with p16-positive OPSCC, with rates of 29% in White/Hispanic individuals, 14% in Asians, and 60% in African American individuals. However, our findings indicate the possibility of initial HPV-OPSCC transitioning to an HPV-loss status. This is supported by instances where p16 IHC was positive with even stronger expression in HPV-loss cancer cells in HPV-OPSCC. Puram et al. demonstrated that p16 expression is maintained in HPV*off* cells, suggesting that p16 reflects the robust early impact of HPV infection [18]. Sustained p16 expression despite HPV loss or reduction indicates the presence of an HPV-negative, unknown mechanism maintaining p16 activation [18]. Therefore, detecting p16 in HPV-loss cancer cells in patients with HPV-OPSCC further implies that these cells likely had initial HPV expression, which was subsequently lost. In our study, HPV-loss clones appeared relatively advanced, suggesting that p16 expression may increase with tumor progression.

Puram et al. [18] performed a single-cell RNA sequencing analysis and reported that HPV*off* cancer cells lost the cell cycle signature observed in HPV-positive cells. However, in the current study, the Ki67 proliferation index, assessed using IHC, tended to be higher in HPV-loss cells, with enrichment of the cell cycle pathway observed in HPV-positive clusters. This finding could be attributed to the intrinsic heterogeneity of HPV-loss cancer cells and their relatively advanced cancer cell state. Collectively, this evidence may provide a deeper understanding of tumor heterogeneity and the dynamic role of HPV-loss cells in cancer progression.

The tumor microenvironment is known to differ between HPV-associated and HPV-negative head and neck squamous cell carcinomas [40, 41]. In this study, we observed that HPV-loss cells are surrounded by increased cytotoxic T-cell infiltration and enhanced

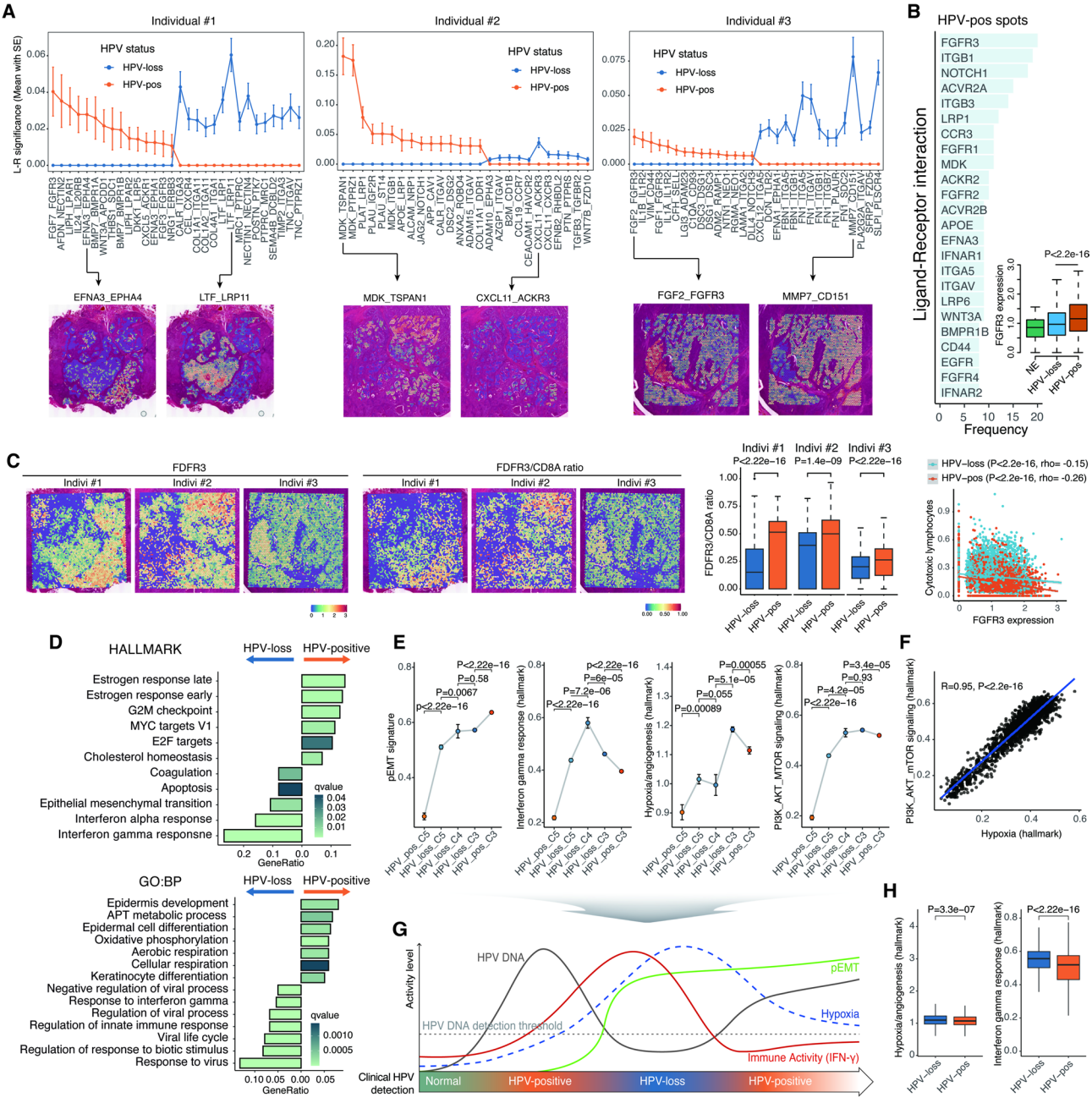

Fig. 6 Ligand-receptor interaction signaling between HPV-positive and HPV-loss clusters. **(A)** Examining ligand-receptor interactions among cells within each spot reveals significant differences between HPV-positive and HPV-loss clusters. **(B)** Among the ligand-receptor interactions, FGFR3 signaling is most frequent in the HPV-positive cluster. FGFR3 expression is elevated in the HPV-positive cluster (Wilcoxon-rank sum test, two-sided). **(C)** FGFR3/CD8A ratio is significantly higher in the HPV-positive subset (Wilcoxon-rank sum test, two-sided), and FGFR expression inversely relates to the infiltration of cytotoxic lymphocytes (Spearman correlation test, two-sided). **(D)** Gene set enrichment analysis was conducted using DEGs identified from cluster-matched comparisons across individual samples and the combined dataset. **(E, F)** Partial EMT (pEMT) signature score and three Hallmark signature scores are expressed as the median with standard error (SE) for individual #2 (t-test, two-sided), along with the correlation between the hypoxia and PI3K/AKT/mTOR signature scores (Pearson correlation test, two-sided). **(G)** Temporal dynamics of the hypoxia signature, pEMT signature, and immune activity (interferon-γ) are depicted over pseudotime. **(H)** Hypoxia and interferon signatures according to HPV status clusters in all three individuals (t-test, two-sided). HPV, human papillomavirus; FGFR3, Fibroblast Growth Factor Receptor 3; PI3K, phosphatidylinositol 3-kinase; DEGs, Differentially expressed genes; EMT, epithelial-mesenchymal transition

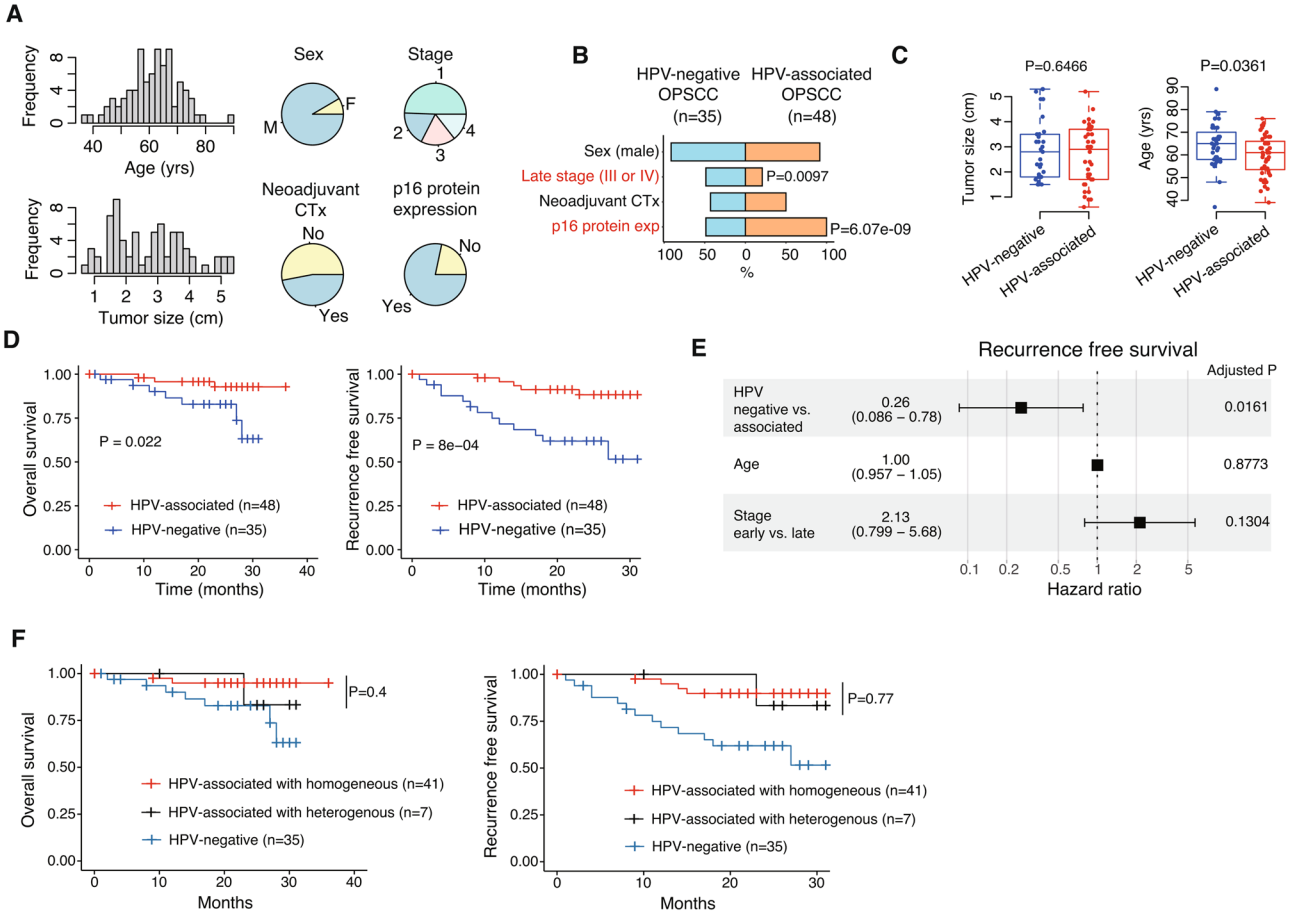

Fig. 7 Clinical significance of HPV spatial heterogeneity. **(A)** Summary of the clinical characteristics of individuals with OPSCC. **(B)** HPV-associated OPSCC appears to be related to p16 protein expression and is associated with an earlier stage (Fisher exact test, two-sided). **(C)** HPV-associated OPSCC occurs in younger individuals and is not associated with tumor size (Wilcoxon-rank sum test, two-sided). **(D)** HPV-associated OPSCC has a better prognosis than HPV-negative OPSCC (Log-rank test, two-sided). **(E)** Even after adjusting for tumor size and stage, positive HPV status is independently associated with a favorable prognosis (Multivariate Cox regression analysis with hazard ratio and 95% confidence interval). **(F)** Among individuals with HPV-associated OPSCC, no difference in survival rates can be observed between individuals with HPV spatial heterogeneity and those with HPV spatial homogeneity (log-rank test, two-sided). HPV, human papillomavirus; OPSCC, oropharyngeal squamous cell carcinoma

interferon signaling, while HPV-positive clusters exhibited reduced cytotoxic T-cell infiltration and elevated FGFR3 signaling. Recently, FGFR signaling contributed to immune evasion by tumor cells, and FGFR inhibition led to increased interferon-gamma signaling [42]. These findings support our observation that FGFR3 signaling is elevated in the HPV-DNA positive clusters, where T-cell immune infiltration is reduced, whereas FGFR3 signaling is relatively reduced in the HPV-loss clusters, accompanied by enhanced T-cell infiltration and interferon signaling. However, the inability to experimentally validate these observations has been noted as a limitation of this study.

Herein, we observed differences in cancer cell histomorphology between HPV-positive and HPV-loss cancer cells. HPV-positive cancer cells exhibited more pleomorphism, whereas HPV-loss cancer cells demonstrated a more uniform histology with less keratinization. These histologic differences could be attributed to HPV-induced effects. However, cancer cell destruction mediated by infiltrating inflammatory cells may be responsible for the selective survival of immune-evasive cancer cells, potentially contributing to altered morphology over time.

We focused on HPV-DNA because HPV-DNA, not HPV-RNA, testing is the gold standard test. Upon sqPCR analysis of HPV-DNA, we observed that HPV-ISH negative areas exhibited a loss or reduced level of HPV-DNA compared with HPV-ISH positive areas. Although this is partly attributable to the higher sensitivity of qPCR, it indicates that HPV-DNA is not completely lost. This suggests that during tumor evolution, the chromatin accessibility of the HPV genome in specific cancer cells may be altered during different cell cycle phases, limiting HPV detection. Additionally, it is possible that HPV was integrated into the cancer cell genome, leaving only remnants of the HPV genome; alternatively, HPV-DNA may

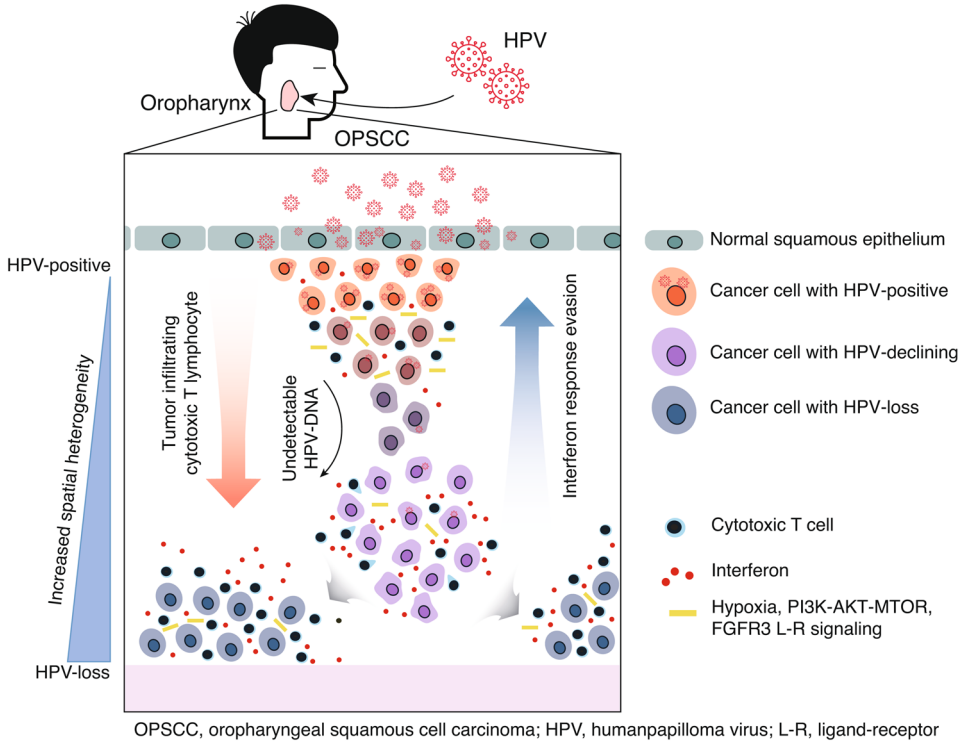

Fig. 8 Summary of results. Combining overall findings, HPV-associated OPSCC is initiated by HPV infection of the oropharyngeal mucosa, and with tumor progression, negative conversion of HPV within cancer cells creates new tumor clusters, increasing tumor heterogeneity. The tumor clusters that underwent HPV-loss conversion exhibit reduced evasion of the virus-induced interferon response, resulting in an increased infiltration of cytotoxic lymphocytes, which contributes to the spatial heterogeneity of immune cell infiltration. HPV-positive and HPV-loss clones differ in molecular features, such as p16, p53, and cell proliferation rate, forming distinct molecular and microenvironmental characteristics. HPV, human papillomavirus; OPSCC, oropharyngeal squamous cell carcinoma

be gradually eliminated in certain cancer cells by immune cell attack. Puram et al. [18] suggested that cancer cells may exist in an intermediate state between HPV*on* and HPV*off*. The findings of our study differ from those of Puram et al., given that we focused on HPV-DNA rather than HPV gene expression. Therefore, our DNA-based measurements may not reflect differences in HPV gene expression. However, quantitative analysis using sqPCR suggests that classifying certain areas where HPV can be quantitatively measured in a binary manner may not be feasible.

We detected the presence of HPV heterogeneity even in metastatic lymph nodes. Although a definitive explanation could not be established, we speculate that instead of two subtypes metastasizing independently, HPV-positive cancer cells may have initially metastasized, followed by the emergence of the HPV-loss subtype. This suggests that HPV-loss conversion can occur independently in multiple cancer cells and may arise not only in the primary site but also in metastatic cancer cells.

Under specific conditions, HPV-loss clones may transition back to an HPV-positive status due to alterations in the tumor microenvironment and selective pressures. Based on the pseudotime analysis, individual #2 exhibited transitions from HPV-positive to HPV-loss and subsequently back to HPV-positive states. This pattern resembles "apparent clearance," wherein HPV-DNA persists at subclinical levels due to immune surveillance [43]. Additionally, HPV-loss clones exhibited markedly higher hypoxia signature scores, suggesting that these clones potentially entered a hypoxia-induced dormant state, which could contribute to reduced viral gene expression and detection thresholds [44]. We observed that as dormant HPV-positive clones transitioned back to an HPV-DNA detectable state, the hypoxia/angiogenesis score ratio decreased, suggesting that alleviation of hypoxia facilitates this process. This is consistent with previous findings, showing that reoxygenation promotes HPV-DNA re-emergence and cellular proliferation, potentially driving tumor progression [45]. Our data further support this association, given that we observed a positive correlation between the PI3K/AKT/mTOR signature score and hypoxia signaling, corroborating reports that hypoxia-induced PI3K/AKT/mTOR pathway activation mediates E6/E7 repression and dormancy in HPV-positive cancer cells [46]. This reinforces the role of hypoxia-driven oncogene repression in HPV-associated tumor dormancy and progression, indicating that hypoxic conditions may

create a tumor microenvironment conducive to HPV-DNA persistence and expansion.

Although the mechanisms underlying the reactivation of HPV expression following an HPV-loss transition remain incompletely understood, several possibilities have been proposed in addition to such hypoxia and its associated molecular pathways, which are also supported by our findings. According to studies by Puram et al. [18], HPV silencing appears to be regulated by epigenetic mechanisms such as DNA methylation and histone modifications. Consequently, the reversibility of these epigenetic states suggests that transitions between HPV*on* and HPV*off* phenotypes may be bidirectional. Furthermore, given the well-established interplay between hypoxia and epigenetic regulation [47], the reactivation of HPV expression may be influenced not only by intrinsic signaling pathways of tumor cells but also by dynamic changes in the tumor microenvironment.

Clinically, the most relevant question is whether prognostic significance and metastatic potential differ between HPV-positive and HPV-loss components. Survival data of our cohort is limited owing to the retrospective nature of the cohort, along with a short follow-up period. However, individuals with intratumoral spatial heterogeneity of HPV-DNA status showed no significant prognostic differences compared with HPV-associated OPSCC without HPV heterogeneity. The HPV-loss clones do not significantly impact clinical outcomes. This finding was supported by an analysis of large-scale public data, which revealed no significant difference in survival outcomes. Their favorable prognosis may be related to the increased infiltration of cytotoxic lymphocytes within HPV-loss clones despite the increased heterogeneity of cancer cells by HPV status change. However, the clinical impact of HPV-loss warrants further validation using larger datasets and accurate methods for identifying HPV-loss cells.

The main limitation of our study is that our conclusions are based on spatial transcriptomic data from only three patients. Although the large number of spots included in the analysis showed high statistical significance, a sample size of three patients limits the generalization of our findings. Therefore, further studies using a larger dataset are needed to validate the precise impact of HPV heterogeneity on the tumor microenvironment and clinical prognosis.

Conclusions

Our study focused on the intratumoral spatiotemporal heterogeneity of HPV-DNA status and its clinicopathologic characteristics in individuals with OPSCC. HPV-loss clones may have differentiated from HPV-positive clones, displaying unique molecular characteristics. The HPV-loss clone is related to regionally distinct cytotoxic lymphocyte infiltration, providing a pivotal basis for the spatiotemporal heterogeneity of cancer cells and the tumor microenvironment.

Abbreviations

HPV	Human papillomavirus
OPSCC	Oropharyngeal squamous cell carcinoma
ISH	In situ hybridization
DEG	Differentially expressed gene
GSEA	Gene Set Enrichment Analysis
CAN	Copy number alteration
sqPCR	Semi-quantitative real-time PCR
IHC	Immunohistochemistry
HE	Hematoxylin and eosin
pEMT	Partial epithelial-mesenchymal transition
scRNA-seq	Single cell RNA sequencing
PCR	Polymerase chain reaction

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-026-08082-5.

Supplementary Material 1

Acknowledgements

We thank the Bioinformatics core facility at the ConveRgence mEDIcine research cenTer (CREDIT), Asan Medical Center, for their support.

Author contributions

K-J.C. conceptualized the study. K-J.C. and C.O.S. designed the study. C.W.P. and C.O.S. analyzed sequencing data and generated figures. W.K., J.H.O., and Y.K.K. contributed to the sequencing data analysis. B.A., J.S.S., H-J.L., and K-J.C. performed the experiments, analyzed, and interpreted the data. B.A., C.O.S., and K-J.C. performed the immunohistochemical staining and pathological assessment. Y.S.L., Y.H.J., S-H.C., and S.Y.N. collected individual samples and data. B.A., C.W.P., C.O.S., and K-J.C. wrote the manuscript. J.S.S. and H-J.L. reviewed and edited the manuscript. All authors have read and approved the final manuscript.

Funding

This study was supported by grants from the Basic Science Research Program through the National Research Foundation of Korea (grant number: NRF-RS-2024-00338555, RS-2023-00275813, RS-2025-00557349); a grant from the Asan Institute for Life Sciences of Asan Medical Center (grant numbers: 2022IP0046, 2023IP0083-2 and 2024IP0082-1); and a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HR21C0198).

Data availability

Visium data generated in this study have been deposited in the NCBI Gene Expression Omnibus (GEO) database (GEO accession no: GSE289363).

Declarations

Ethics approval and consent to participate

This study was approved by the Institutional Review Board of the Asan Medical Center (IRB no. 2021 – 1235).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details
[1]Department of Pathology, Asan Medical Center, University of Ulsan College of Medicine, 88, Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Republic of Korea
[2]Department of Medical Science, Asan Medical Center, University of Ulsan College of Medicine, Brain Korea 21 project, Seoul, Republic of Korea
[3]Department of Otolaryngology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
[4]Bioinformatics Core Laboratory, Convergence Medicine Research Center, Asan Institute for Life Sciences, Asan Medical Center, Seoul, Republic of Korea

Received: 18 March 2025 / Accepted: 27 May 2025
Published online: 06 April 2026

References

1. 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.
2. Lewis JS. Morphologic diversity in human papillomavirus-related oropharyngeal squamous cell carcinoma. Catch Me If You Can! Mod Pathol. 2017;30:S44–53.
3. Perry ME. The specialised structure of crypt epithelium in the human palatine tonsil and its functional significance. J Anat. 1994;185(Pt 1):111–27.
4. Lassen P, Lacas B, Pignon JP, Trotti A, Zackrisson B, Zhang Q, Overgaard J, Blanchard P. Prognostic impact of HPV-associated p16-expression and smoking status on outcomes following radiotherapy for oropharyngeal cancer: The MARCH-HPV project. Radiother Oncol. 2018;126:107–15.
5. Fakhry C, Waterboer T, Westra WH, Rooper LM, Windon M, Troy T, Koch W, Gourin CG, Bender N, Yavvari S, Kiess AP, Miles BA, Ryan WR, Ha PK, Eisele DW, D'Souza G. Distinct biomarker and behavioral profiles of human papillomavirus-related oropharynx cancer patients by age. Oral Oncol. 2020;101:104522–104522.
6. Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tân PF, Westra WH, Chung CH, Jordan RC, Lu C, Kim H, Axelrod R, Silverman CC, Redmond KP, Gillison ML. Human papillomavirus and survival of patients with oropharyngeal cancer. N Engl J Med. 2010;363:24–35.
7. Lim YX, Mierzwa ML, Sartor MA, D'Silva NJ. Clinical, morphologic and molecular heterogeneity of HPV-associated oropharyngeal cancer. Oncogene. 2023;42:2939–55.
8. Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, Kryukov GV, Lawrence MS, Sougnez C, McKenna A, Shefler E, Ramos AH, Stojanov P, Carter SL, Voet D, Cortés ML, Auclair D, Berger MF, Saksena G, Guiducci C, Onofrio RC, Parkin M, Romkes M, Weissfeld JL, Seethala RR, Wang L, Rangel-Escareño C, Fernandez-Lopez JC, Hidalgo-Miranda A, Melendez-Zajgla J, Winckler W, Ardlie K, Gabriel SB, Meyerson M, Lander ES, Getz G, Golub TR, Garraway LA, Grandis JR. The mutational landscape of head and neck squamous cell carcinoma. Science. 2011;333:1157–60.
9. Rosty C, Sheffer M, Tsafrir D, Stransky N, Tsafrir I, Peter M, de Crémoux P, de La Rochefordière A, Salmon R, Dorval T, Thiery JP, Couturier J, Radvanyi F, Domany E, Sastre-Garau X. 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.
10. Sastre-Garau X, Harlé A. Pathology of HPV-Associated Head and Neck Carcinomas: Recent Data and Perspectives for the Development of Specific Tumor Markers. Front Oncol. 2020;10.
11. Amin MB, Greene ES FL, editors. AJCC cancer staging manual. 8th ed. New York: Springer; 2017.
12. Fauzi FH, Hamzan NI, Rahman NA, Suraiya S, Mohamad S. Detection of human papillomavirus in oropharyngeal squamous cell carcinoma. J Zhejiang Univ Sci B. 2020;21:961–76.
13. Caudell JJ, Gillison ML, Maghami E, Spencer S, Pfister DG, Adkins D, Birkeland AC, Brizel DM, Busse PM, Cmelak AJ, Colevas AD, Eisele DW, Galloway T, Geiger JL, Haddad RI, Hicks WL, Hitchcock YJ, Jimeno A, Leizman D, Mell LK, Mittal BB, Pinto HA, Rocco JW, Rodriguez CP, Savvides PS, Schwartz D, Shah JP, Sher D, St John M, Weber RS, Weinstein G, Worden F, Yang Bruce J, Yom SS, Zhen W, Burns JL, Darlow SD. NCCN Guidelines® Insights: Head and Neck Cancers, Version 1.2022. J Natl Compr Canc Netw. 2022;20:224–34.
14. Cantley RL, Gabrielli E, Montebelli F, Cimbaluk D, Gattuso P, Petruzzelli G. Ancillary Studies in Determining Human Papillomavirus Status of Squamous Cell Carcinoma of the Oropharynx: A Review. Pathol Res Int. 2011;2011:138469.
15. Mirghani H, Casiraghi O, Amen F, He M, Ma XJ, Saulnier P, Lacroix L, Drusch F, Ben Lakdhar A, Saint Guily JL, Badoual C, Scoazec JY, Vielh P. Diagnosis of HPV-driven head and neck cancer with a single test in routine clinical practice. Mod Pathol. 2015;28:1518–27.
16. Rietbergen MM, Snijders PJ, Beekzada D, Braakhuis BJ, Brink A, Heideman DA, Hesselink AT, Witte BI, Bloemena E, Baatenburg-De Jong RJ, Leemans CR, Brakenhoff RH. Molecular characterization of p16-immunopositive but HPV DNA-negative oropharyngeal carcinomas. Int J Cancer. 2014;134:2366–72.
17. Lewis JS Jr., Thorstad WL, Chernock RD, Haughey BH, Yip JH, Zhang Q, El-Mofty SK. p16 positive oropharyngeal squamous cell carcinoma:an entity with a favorable prognosis regardless of tumor HPV status. Am J Surg Pathol. 2010;34:1088–96.
18. Puram SV, Mints M, Pal A, Qi Z, Reeb A, Gelev K, Barrett TF, Gerndt S, Liu P, Parikh AS, Ramadan S, Law T, Mroz EA, Rocco JW, Adkins D, Thorstad WL, Gay HA, Ding L, Paniello RC, Pipkorn P, Jackson RS, Wang X, Mazul A, Chernock R, Zevallos JP, Silva-Fisher J, Tirosh I. Cellular states are coupled to genomic and viral heterogeneity in HPV-related oropharyngeal carcinoma. Nat Genet. 2023;55:640–50.
19. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9:671–5.
20. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, Hoffman P, Stoeckius M, Papalexi E, Mimitou EP, Jain J, Srivastava A, Stuart T, Fleming LM, Yeung B, Rogers AJ, McElrath JM, Blish CA, Gottardo R, Smibert P, Satija R. Integrated analysis of multimodal single-cell data. Cell. 2021;184:3573–e35873529.
21. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, Brenner M, Loh PR, Raychaudhuri S. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods. 2019;16:1289–96.
22. Grabski IN, Street K, Irizarry RA. Significance analysis for clustering with single-cell RNA-sequencing data. Nat Methods. 2023;20:1196–202.
23. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, Cahill DP, Nahed BV, Curry WT, Martuza RL, Louis DN, Rozenblatt-Rosen O, Suvà ML, Regev A, Bernstein BE. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science. 2014;344:1396–401.
24. Peng J, Sun BF, Chen CY, Zhou JY, Chen YS, Chen H, Liu L, Huang D, Jiang J, Cui GS, Yang Y, Wang W, Guo D, Dai M, Guo J, Zhang T, Liao Q, Liu Y, Zhao YL, Han DL, Zhao Y, Yang YG, Wu W. Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. Cell Res. 2019;29:725–38.
25. Kurtenbach S, Cruz AM, Rodriguez DA, Durante MA, Harbour JW. Uphyloplot2: visualizing phylogenetic trees from single-cell RNA-seq data. BMC Genomics. 2021;22:419.
26. Skidmore ZL, Wagner AH, Lesurf R, Campbell KM, Kunisaki J, Griffith OL, Griffith M. GenVisR: Genomic Visualizations in R. Bioinformatics. 2016;32:3012–4.
27. Wolf FA, Hamey FK, Plass M, Solana J, Dahlin JS, Göttgens B, Rajewsky N, Simon L, Theis FJ. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. Genome Biol. 2019;20:59.
28. Pham D, Tan X, Balderson B, Xu J, Grice LF, Yoon S, Willis EF, Tran M, Lam PY, Raghubar A, Kalita-de Croft P, Lakhani S, Vukovic J, Ruitenberg MJ, Nguyen QH. Robust mapping of spatiotemporal trajectories and cell-cell interactions in healthy and diseased tissues. Nat Commun. 2023;14:7739.
29. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautès-Fridman C, Fridman WH, de Reyniès A. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016;17:218.
30. Meylan M, Petitprez F, Becht E, Bougoüin A, Pupier G, Calvez A, Giglioli I, Verkarre V, Lacroix G, Verneau J, Sun CM, Laurent-Puig P, Vano YA, Elaïdi R, Méjean A, Sanchez-Salas R, Barret E, Cathelineau X, Oudard S, Reynaud CA, de Reyniès A, Sautès-Fridman C, Fridman WH. Tertiary lymphoid structures generate and propagate anti-tumor antibody-producing plasma cells in renal cell cancer. Immunity. 2022;55:527–e541525.
31. Cable DM, Murray E, Zou LS, Goeva A, Macosko EZ, Chen F, Irizarry RA. Robust decomposition of cell type mixtures in spatial transcriptomics. Nat Biotechnol. 2022;40:517–26.
32. Foroutan M, Bhuva DD, Lyu R, Horan K, Cursons J, Davis MJ. Single sample scoring of molecular phenotypes. BMC Bioinformatics. 2018;19:404.

33. Network CGA. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature. 2015;517:576–82.
34. Metzger K, Moratin J, Freier K, Hoffmann J, Zaoui K, Plath M, Stögbauer F, Freudlsperger C, Hess J, Horn D. A six-gene expression signature related to angiolymphatic invasion is associated with poor survival in laryngeal squamous cell carcinoma. Eur Arch Otorhinolaryngol. 2021;278:1199–207.
35. Nan Z, Dou Y, Chen A, Wang K, Sun J, Meng Z, Neckenig M, Ai D, Liu S, Dong Z, Ma C, Cheng Y, Qu X. Identification and validation of a prognostic signature of autophagy, apoptosis and pyroptosis-related genes for head and neck squamous cell carcinoma: to imply therapeutic choices of HPV negative patients. Front Immunol. 2022;13:1100417.
36. Novotný J, Bandúrová V, Strnad H, Chovanec M, Hradilová M, Šáchová J, Šteffl M, Grušanović J, Kodet R, Pačes V, Lacina L, Smetana K Jr., Plzák J, Kolář M, Vomastek T. Analysis of HPV-Positive and HPV-Negative Head and Neck Squamous Cell Carcinomas and Paired Normal Mucosae Reveals Cyclin D1 Deregulation and Compensatory Effect of Cyclin D2. Cancers (Basel). 2020;12.
37. Walter V, Yin X, Wilkerson MD, Cabanski CR, Zhao N, Du Y, Ang MK, Hayward MC, Salazar AH, Hoadley KA, Fritchie K, Sailey CJ, Weissler MC, Shockley WW, Zanation AM, Hackman T, Thorne LB, Funkhouser WD, Muldrew KL, Olshan AF, Randell SH, Wright FA, Shores CG, Hayes DN. Molecular subtypes in head and neck cancer exhibit distinct patterns of chromosomal gain and loss of canonical cancer genes. PLoS ONE. 2013;8:e56823.
38. Wichmann G, Rosolowski M, Krohn K, Kreuz M, Boehm A, Reiche A, Scharrer U, Halama D, Bertolini J, Bauer U, Holzinger D, Pawlita M, Hess J, Engel C, Hasenclever D, Scholz M, Ahnert P, Kirsten H, Hemprich A, Wittekind C, Herbarth O, Horn F, Dietz A, Loeffler M. 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. 2015;137:2846–57.
39. D'Souza G, Westra WH, Wang SJ, van Zante A, Wentz A, Kluz N, Rettig E, Ryan WR, Ha PK, Kang H, Bishop J, Quon H, Kiess AP, Richmon JD, Eisele DW, Fakhry C. Differences in the Prevalence of Human Papillomavirus (HPV) in Head and Neck Squamous Cell Cancers by Sex, Race, Anatomic Tumor Site, and HPV Detection Method. JAMA Oncol. 2017;3:169–77.
40. Lee H, Park S, Yun JH, Seo C, Ahn JM, Cha HY, Shin YS, Park HR, Lee D, Roh J, Heo HJ, Baek SE, Kim EK, Lee HS, Kim CH, Kim YH, Jang JY. Deciphering head and neck cancer microenvironment: Single-cell and spatial transcriptomics reveals human papillomavirus-associated differences. J Med Virol. 2024;96:e29386.
41. Hong MH, Park S, Vo T, Cho J, Jung HA, Lee SH, Kim SH, Zhou H, Chirovsky D, Koh YW, Yoon SO, Webber AL, Gumuscu B, Cho BC, Ahn MJ. Expression of PD-L1, PD-L2, and inflammatory gene expression profile in locally advanced head and neck squamous cell carcinoma. ESMO Open. 2024;9:103961.
42. Ruan R, Li L, Li X, Huang C, Zhang Z, Zhong H, Zeng S, Shi Q, Xia Y, Zeng Q, Wen Q, Chen J, Dai X, Xiong J, Xiang X, Lei W, Deng J. Unleashing the potential of combining FGFR inhibitor and immune checkpoint blockade for FGF/FGFR signaling in tumor microenvironment. Mol Cancer. 2023;22:60.
43. Doorbar J. The human Papillomavirus twilight zone - Latency, immune control and subclinical infection. Tumour Virus Res. 2023;16:200268.
44. Porter VL, Marra MA. The Drivers, Mechanisms, and Consequences of Genome Instability in HPV-Driven Cancers. Cancers (Basel). 2022;14.
45. Hoppe-Seyler K, Bossler F, Lohrey C, Bulkescher J, Rösl F, Jansen L, Mayer A, Vaupel P, Dürst M, Hoppe-Seyler F. Induction of dormancy in hypoxic human papillomavirus-positive cancer cells. Proc Natl Acad Sci U S A. 2017;114:E990–8.
46. Bossler F, Kuhn BJ, Günther T, Kraemer SJ, Khalkar P, Adrian S, Lohrey C, Holzer A, Shimobayashi M, Dürst M, Mayer A, Rösl F, Grundhoff A, Krijgsveld J, Hoppe-Seyler K, Hoppe-Seyler F. Repression of Human Papillomavirus Oncogene Expression under Hypoxia Is Mediated by PI3K/mTORC2/AKT Signaling. mBio. 2019;10.
47. Verdikt R, Thienpont B. Epigenetic remodelling under hypoxia. Semin Cancer Biol. 2024;98:1–10.

Publisher's note