


# Differences in gene expression between high-grade dysplasia and invasive HPV<sup>+</sup> and HPV<sup>-</sup> tonsillar and base of tongue cancer

Linnea Haeggblom<sup>1</sup> | Andreas Ährlund-Richter<sup>1</sup> | Leila Mirzaie<sup>1</sup> |  
Pedro Farrajota Neves da Silva<sup>1,2</sup> | Ramona G. Ursu<sup>3</sup> | Torbjörn Ramqvist<sup>1</sup>  |  
Anders Näsman<sup>1,2</sup> 

<sup>1</sup>Department of Oncology-Pathology, Karolinska Institute, Stockholm, Sweden

<sup>2</sup>Department of Clinical Pathology, Karolinska University Hospital, Stockholm, Sweden

<sup>3</sup>Department of Microbiology, University of Medicine and Pharmacy Grigore T. Popa Iasi, Iasi, Romania

## Correspondence

Anders Näsman, Department of Oncology-Pathology, Karolinska Institute, Bioclinicum J6:20, Karolinska University Hospital, 171 64 Stockholm, Sweden.  
Email: anders.nasman@ki.se

## Funding information

This work was supported by Cancer och Allergifonden, Åke Wibergs stiftelse, Stiftelsen Clas Groschinsky's minnesfond, Stiftelsen Tornspiran, Svenska läkarsällskapet (SLS), Stiftelsen Sigurd och Elsa Goljes Minne, Magnus Bergvalls stiftelse, Eva och Oscar Ahréns stiftelse, Tore Nilssons Stiftelse för medicinsk forskning, Svenska Sällskapet för Medicinsk Forskning (SSMF), the Stockholm County Council (SLL) and Karolinska Institutet. RGU was funded by the University of Medicine and Pharmacy Grigore T Popa Iasi, based on contract no. 30336/28.12.2017. The funding bodies had no role in the study design, data collection, analysis, interpretation of data, or in writing the manuscript.

## Abstract

**Background:** Human papillomavirus (HPV) is a causative agent for tonsillar and base of tongue squamous cell carcinoma (TSCC/BOTSCC), as well as for cervical cancer. Premalignant stages in cervical cancer have been studied extensively, while little is known about premalignant stages in TSCC/BOTSCC and the role of HPV. Here we analyzed differences in gene and protein expression between high-grade dysplasia and invasive cancer in both HPV-positive (HPV<sup>+</sup>) and HPV-negative (HPV<sup>-</sup>) TSCC/BOTSCC.

**Methods:** High-grade dysplasia and invasive carcinoma were laser microdissected from HPV<sup>+</sup> and HPV<sup>-</sup> TSCC/BOTSCC tumor sections. Differential gene expression was studied utilizing nanoString RNA-panels and genes of interest were validated on the protein level by immunohistochemistry.

**Results:** Forty genes in the HPV<sup>+</sup> tumors showed significantly different expression between high-grade dysplasia and invasive cancer and 33 genes in the HPV<sup>-</sup> tumors. Five out of the nine most significant pathways showed similar increased activity in invasive cancer as compared to high-grade dysplasia in both HPV<sup>+</sup> and HPV<sup>-</sup> tumors. Lastly, significant differences in protein expression was confirmed for SPARC, psoriasin, type I collagen and galectin-1 in both HPV<sup>+</sup> and HPV<sup>-</sup> tumors.

**Conclusions:** This is to our knowledge the first study disclosing differences and similarities in gene expression between dysplastic and invasive HPV<sup>+</sup> and HPV<sup>-</sup> TSCC/BOTSCC.

## KEYWORDS

base of tongue cancer, cancer in situ, cancer progression, high-grade dysplasia, human papillomavirus, invasive carcinoma, oropharyngeal cancer, premalignant stages, tonsillar cancer

**Abbreviations:** AJCC, American joint committee on cancer; BOTSCC, base of tongue squamous cell carcinoma; CIN, cervical intraepithelial neoplasia; DCIS, ductal carcinoma in situ; ECM, extracellular matrix; FFPE, formalin fixed paraffin embedded; HNSCC, head and neck squamous cell carcinoma; HPV, human papillomavirus; HPV<sup>-</sup>, human papillomavirus negative; HPV<sup>+</sup>, human papillomavirus positive; HSIL, high-grade squamous intraepithelial lesions; IHC, immunohistochemistry; LCM, laser capture microdissection; LSIL, low-grade squamous intraepithelial lesions; OPSCC, oropharyngeal squamous cell carcinoma; p16, p16<sup>INK4a</sup> tumor suppressor protein; SPARC, secreted protein acidic and rich in cysteine; TIL, tumor infiltrating lymphocyte; TSCC, tonsillar squamous cell carcinoma.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Cancer Medicine* published by John Wiley & Sons Ltd.

## 1 | BACKGROUND

Previous studies by us and others have demonstrated an epidemic increase in the incidence of tonsillar and base of tongue squamous cell carcinoma (TSCC and BOTSCC), and this has been attributed to an increase in human papillomavirus (HPV) infection.<sup>1–8</sup> Moreover, patients with HPV positive (HPV<sup>+</sup>) TSCC and BOTSCC are generally younger and do not usually present with the classical head and neck squamous cell carcinoma (HNSCC) risk factors, for example, smoking and alcohol abuse.<sup>9,10</sup> However, despite an epidemic increase in a younger patient population, no screening methods, in line with those for cervical cancer, have been developed.

The absolute majority of all carcinomas develop in a similar manner, starting with dysplasia followed by high-grade dysplasia/cancer in situ, which eventually progresses into an invasive cancer and thereafter becomes metastatic. High-grade dysplasia/cancer in situ refers to the earliest stage of cancer when the cancerous growth is still contained by the basal membrane and has not invaded the surrounding tissues or spread to other organs in the body.<sup>11</sup> In cervical carcinoma, the premalignant and malignant stages are very well characterized and divided into low-grade squamous intraepithelial lesions (LSIL) and high-grade squamous intraepithelial lesions (HSIL) (former cervical intraepithelial neoplasia [CIN] 1–3 grades) and invasive cancer. Analogous stages are also found in other HPV-associated anogenital carcinomas.<sup>12</sup> A similar, yet less studied, scenario is also observed in the epithelium coating the tonsils and the base of tongue, where occasionally premalignant lesions are found close by the invasive tumor. However, it has been debated if pure premalignant lesions occur in HPV<sup>+</sup> TSCC or BOTSCC.<sup>13–16</sup>

Increased knowledge of genes involved in tumor invasion mechanisms and identification of new progression markers could potentially in the future improve and personalize treatment and/or facilitate development of new strategies for early detection and preventative measures.

Due to its rarity, no studies have, to our knowledge, been performed on differences between premalignant and invasive lesions in HPV<sup>+</sup> TSCC and BOTSCC. Instead, a handful of studies have been performed on premalignant changes in HPV<sup>+</sup> cervical carcinomas. In two independent studies by Gius et al and den Boon et al, differences in gene expression in the different stages of premalignant lesions (CIN1–3) leading up to cervical cancer were examined, and both studies show specific changes displayed from stage-to-stage.<sup>17,18</sup> However, no comparisons between HPV<sup>+</sup> and HPV negative (HPV<sup>-</sup>) precancerous and invasive tumors were possible in these studies, since >99% of all cervical carcinomas are HPV<sup>+</sup>.<sup>19,20</sup>

Moreover, it is imperative for the patient treatment and prognosis to distinguish dysplasia/cancer in situ from an invasive carcinoma. Whereas an invasive tumour diagnosis

most often will result in heavy oncological treatment and/or radical surgery, a diagnosis of cancer in situ/severe dysplasia diagnosis may lead to surgical extirpation, but not with extensive margins, and only clinical follow-up. Normally, the distinction between premalignant and invasive disease is unproblematic, but there are cases, especially in head neck cancer pathology when the diagnosis is vague. Therefore, some invasion markers, for example, Laminin-5,<sup>21</sup> have been proposed, but these markers have all very poor sensitivity and specificity in the head and neck area. Therefore, there is an urgent need for new clinical invasion markers in head and neck pathology, for better and safer diagnosis of invasive disease.

In this study we aimed, for the first time, to examine and better understand differences between HPV<sup>+</sup> and HPV<sup>-</sup> pre-invasive TSCC and BOTSCC lesions and invasive tumors by laser capture microdissection (LCM) of the malignant epithelial cells followed by gene analysis utilizing the PanCancer Progression Panel (NanoString). In addition, we also aspired to identify usable clinical invasion markers by visual verification of obtained RNA data on protein level by immunohistochemistry (IHC).

## 2 | METHODS

### 2.1 | Patients and tumor samples

In this study 24 TSCC and BOTSCC (ICD-10: C09.0, C09.1, C09.8, C09.9, C01.9) biopsies were selected from a larger consecutive cohort diagnosed at Karolinska university hospital 2007–2015.<sup>5,22</sup> All biopsies, previously tested for both presence of HPV DNA and p16<sup>INK4a</sup> (p16) expression,<sup>5,22</sup> were histologically evaluated and selected for containing both high-grade dysplasia/cancer in situ and invasive carcinoma within the same formalin fixed paraffin embedded (FFPE) section. Thirteen HPV driven tumors, being both HPV DNA positive and p16 positive (HPV<sup>+</sup>p16<sup>+</sup>) were identified from the cohort and 11 HPV DNA negative and p16 negative (HPV<sup>-</sup>p16<sup>-</sup>) tumors, thus driven by other factors, were selected. From these 24 samples, the six HPV<sup>+</sup> and HPV<sup>-</sup> patient samples with highest quality and satisfactory amount of RNA were used in the RNA expression assay (please see below). Differentially expressed genes were then validated in all 24 samples by IHC. For more details about the study subjects see Supplementary Tables S1 and S2. The study was performed according to approval 2009/1278-31/4 from the Regional Ethics Committee, Karolinska Institutet.

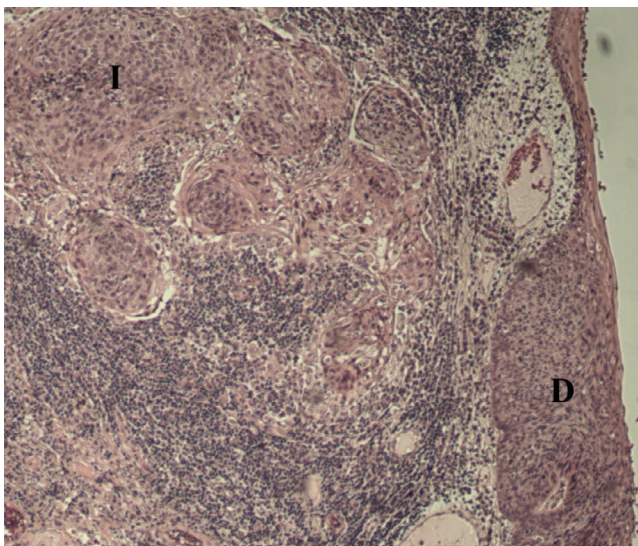
### 2.2 | Tumor selection and microdissection

Hematoxylin/eosin stained FFPE tumor sections were histologically examined by three independent researchers

(LH, PFNS, and AN) and selected for containing sufficient amounts of high-grade dysplasia/cancer in situ as well as invasive cancer (Figure 1). FFPE biopsies with distinct differences for dysplasia and invasive carcinoma were sectioned (six 5  $\mu\text{m}$ -sections per sample) by microtome with RNase-free water and mounted on Membrane Slides PEN-Membrane 2.0  $\mu\text{m}$  (No.11505158) (Leica Microsystems AB). Tissue sections were deparaffinized for  $2 \times 15$  minutes in fresh xylene followed by rehydration in decreasing percentages of fresh ethanol (5 minutes each in 100%, 95%, 70%, 50%, 0% ethanol) and thereafter stained with fresh hematoxylin for 30 seconds. On a Leica LMD 7000 microscope (Leica Microsystems AB), using the Laser Microdissection System (version 7.6.5684) the dysplastic and invasive carcinoma areas of the six replicate tumor sections were laser micro dissected and collected separately in microcentrifuge tubes containing PKD buffer from the RNeasy FFPE kit (Qiagen).

### 2.3 | RNA extraction and multiplex gene expression analysis

Samples were kept on ice whenever possible and RNA was extracted immediately upon laser microdissection to ensure best RNA quality, using the RNeasy FFPE kit (Qiagen) according to manufacturer's instructions. RNA concentration was measured on Qubit 4 Fluorometer (Thermo Fisher Scientific) using the Qubit<sup>TM</sup> RNA HS Assay Kit (Thermo Fisher Scientific). Twelve samples (six HPV<sup>+</sup>p16<sup>+</sup> tumors and six HPV<sup>-</sup>p16<sup>-</sup> tumors) with sufficient RNA concentrations ( $>2$  ng/ $\mu\text{L}$ ) were selected for the multiplex gene expression assay. Due to low RNA concentrations from FFPE material, the nCounter Low RNA Input Amplification Kit (NanoString Technologies) was



**FIGURE 1** Histological representation of contained high-grade dysplasia/cancer in situ (D) and invasive carcinoma (I)

used for the cDNA conversion and multiplex target enrichment steps. Thereafter the hybridization reactions were set up using provided master kit and progression primers needed for the nCounter PanCancer Progression Panel (NanoString Technologies). The gene expression assay was performed according to manufacturer's instructions utilizing the nCounter<sup>®</sup> Sample Prep Station with FLEX configuration (NanoString Technologies) as well as the nCounter<sup>®</sup> Digital Analyzer 5s (NanoString Technologies) for reading the samples. For data analysis the nSolver Analysis Software (version 4.0) was used, including the nCounter Advanced Analysis add-on software (version 2.0.115). More specifically, fold changes and *P*-values were calculated using nCounter default settings. Pathway scores were calculated using measurements of genes included in 34 different pathways, using the nCounter Advanced Analysis Software.<sup>23</sup>

### 2.4 | Immunohistochemistry evaluation of differentially expressed genes

Six of the top differentially expressed genes between high-grade dysplasia/cancer in situ and invasive carcinoma, identified from the multiplex gene expression analysis, were further evaluated for protein expression by IHC. FFPE sections from the 24 TSCC/BOTSCC biopsies were stained by IHC with six different antibodies as per standard protocol. Further details about antibodies and methodologies are presented in Supplementary Table S3. Stained tissue sections were histologically evaluated by light microscopy by three trained scientists (LH, AN, LM) as described separately for each antibody in Supplementary Table S3. Evaluation scores are presented in Supplementary Table S4.

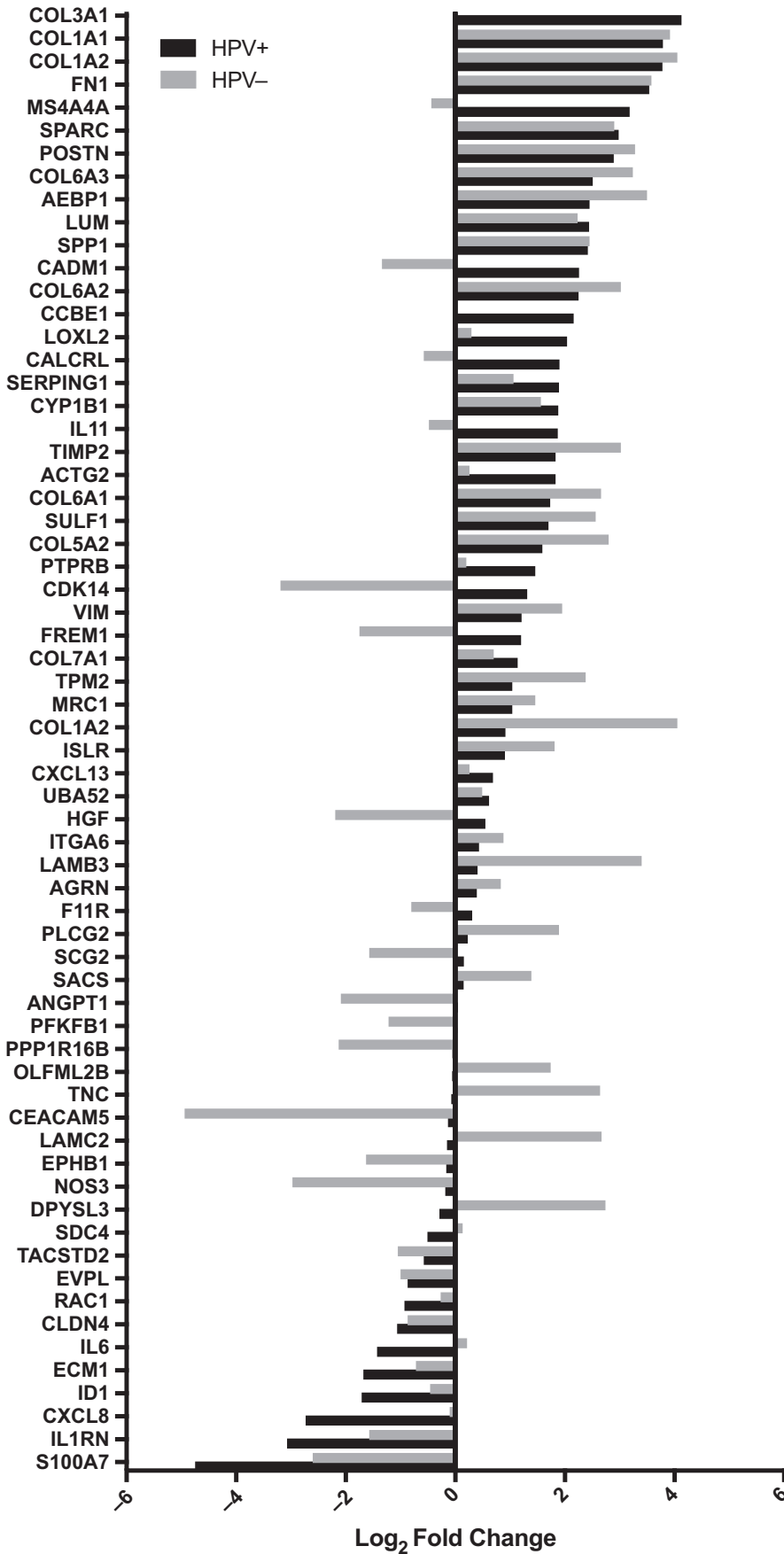
### 2.5 | Statistical analysis

Nanostring data and pathway analysis are described above. Differences in protein expression were assessed with the Wilcoxon signed rank-test. Two sided *P*-values were reported.

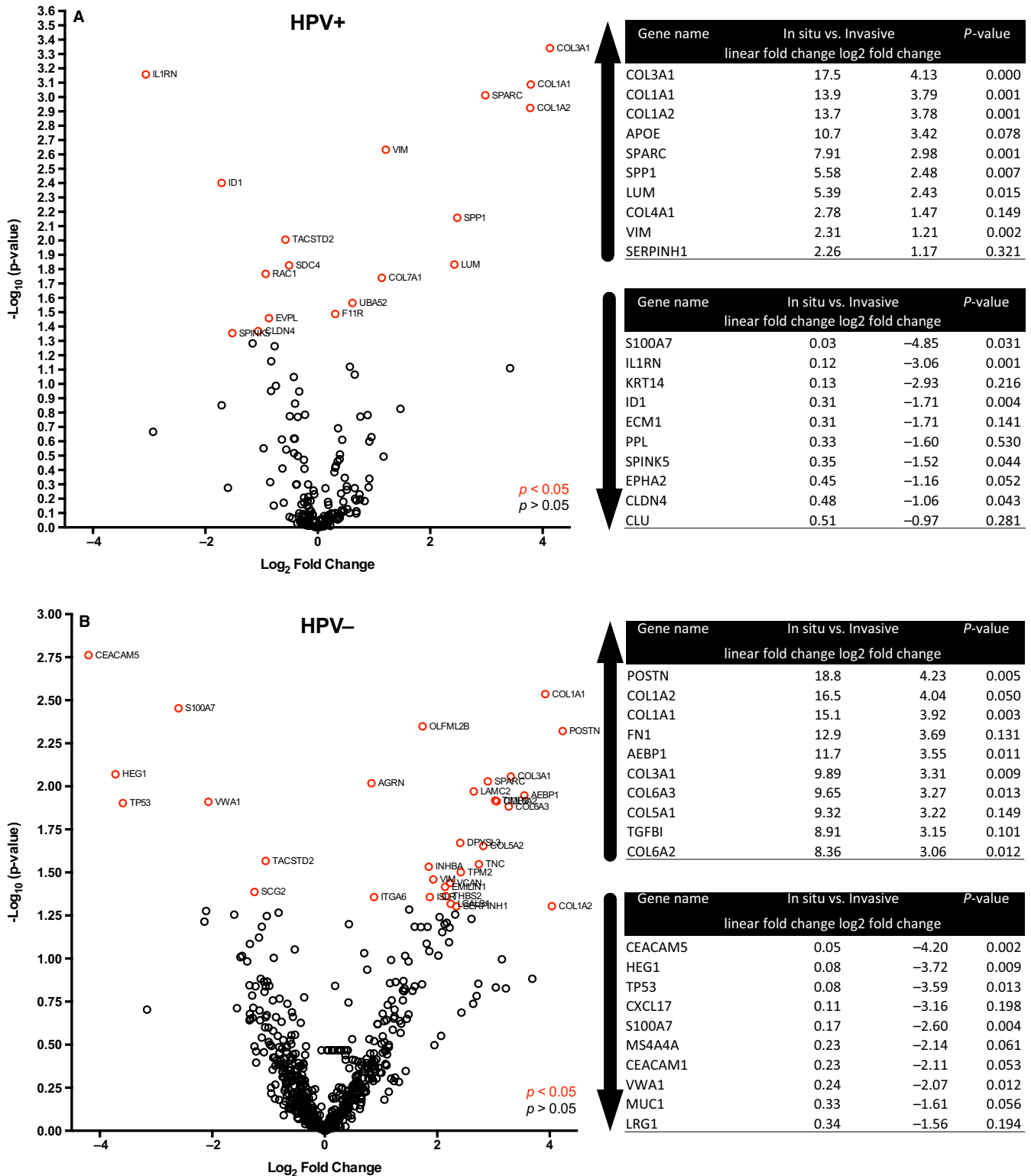
## 3 | RESULTS

### 3.1 | Gene expression and pathway analysis of high-grade dysplasia and invasive TSCC and BOTSCC

Gene expression was measured in six HPV<sup>+</sup> and six HPV<sup>-</sup> TSCC and BOTSCC samples from a panel of 770 cancer progression related genes (PanCancer Progression Panel, NanoString). In total, 40 genes in the HPV<sup>+</sup> tumors and 33 genes in the HPV<sup>-</sup> tumors showed differential expression ( $P < .05$ ), comparing high-grade dysplastic epithelium and invasive cancer. Significant genes are presented in Figure



**FIGURE 2** Log<sub>2</sub> fold change difference between high-grade dysplasia and invasive carcinoma comparing statistically significant differences ( $P < .05$ ) in mRNA expression in the HPV<sup>+</sup> and HPV<sup>-</sup> samples separately. All genes showing a positive fold change value have a higher expression in invasive carcinoma compared to high-grade dysplasia, and genes showing a negative fold change value have a lower expression in invasive carcinoma compared to high-grade dysplasia



**FIGURE 3** Gene expression in high-grade dysplasia and invasive carcinoma in HPV<sup>+</sup> and HPV<sup>-</sup> tonsillar and base of tongue squamous cell carcinoma separately. (A) Volcano plot showing differentially expressed genes between high-grade dysplasia and invasive carcinoma by log<sub>2</sub> fold change (x-axis) and -log<sub>10</sub> of P-value (y-axis) in HPV<sup>+</sup> tonsillar and base of tongue squamous cell carcinoma cases. The top 10 genes with increased and decreased linear fold change as well as log<sub>2</sub> fold change expression between high-grade dysplasia and invasive carcinoma are together with P-values presented in the tables to the right. (B) Same as in (A), here however with HPV<sup>-</sup> tonsillar and base of tongue squamous cell carcinoma cases

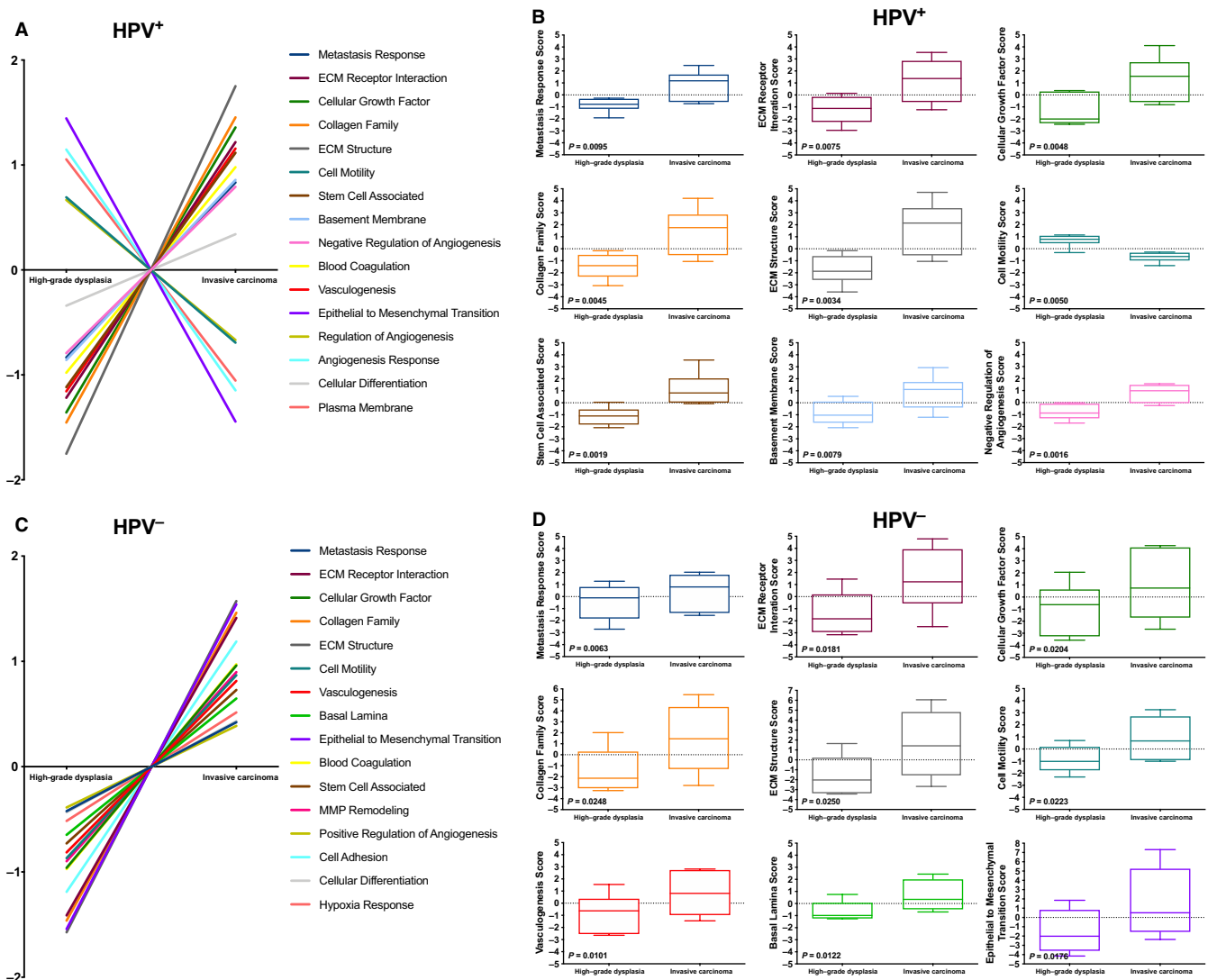
2 for both HPV<sup>+</sup> and HPV<sup>-</sup> tumors. Of these genes, 10 were common for both HPV<sup>+</sup> and HPV<sup>-</sup> tumors. More specifically, the genes *COL3A1*, *COL1A1*, *COL1A2*, *SPARC*, *COL6A3*, *AEBP1*, *COL6A2*, and *VIM* were significantly up-regulated, whereas *SI00A7* and *TACSTD2* were significantly down-regulated, in invasive cancer as compared to the dysplastic epithelium (Figure 2).

Consequently, 30 genes differed between high-grade dysplasia vs. invasive cancer uniquely for HPV<sup>+</sup> samples. The following five genes, *FNI*, *MS4A4A*, *POSTN*, *LUM*, and *SPPI*, were the most significantly up-regulated, in invasive cancer as compared to the dysplastic epithelium. Whereas *IL1RN*, *CXCL8*, *IDI*, *ECM1*, and *IL6*, were the five most significantly down-regulated genes in invasive cancer as compared to the dysplastic epithelium. For more detailed data see Figure 2.

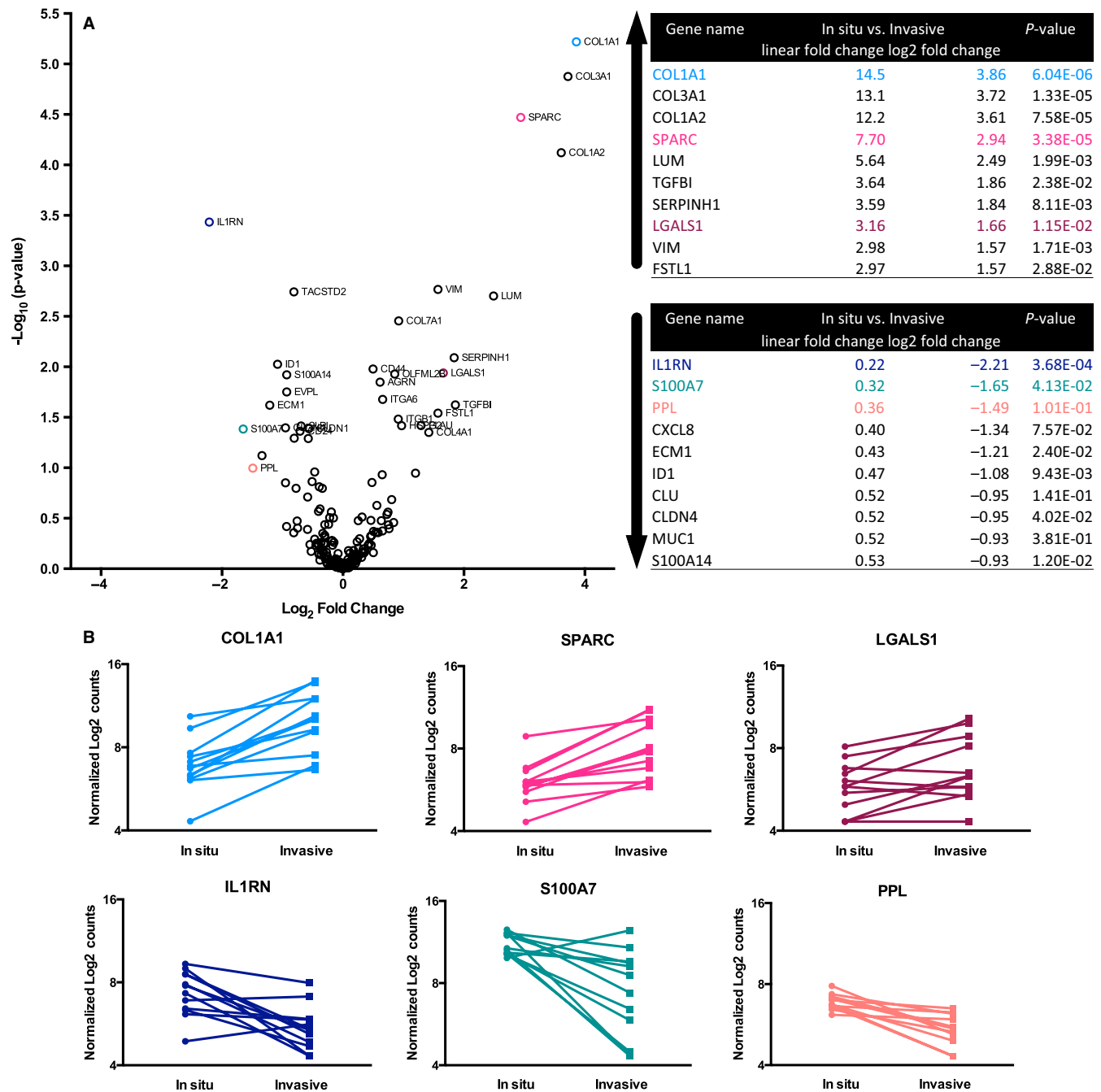
Exclusively for HPV<sup>-</sup> cases 23 genes differed between high-grade dysplasia and invasive cancer. The following five genes, *LAMB3*, *TIMP2*, *COL5A2*, *DPYSL3*, and *LAMC2*, were the most significantly up-regulated, in invasive cancer as compared to the dysplastic epithelium. Whereas *CEACAM5*, *CDK14*, *NOS3*, *HGF*, and *PPP1R16B*, were the five most significantly down-regulated genes in invasive cancer as compared to the dysplastic epithelium (Figure 2).

Upon further evaluation of the above mentioned genes, a more stringent cut-off was applied in order to exclude samples with a low count value, resulting in that 19 genes in the HPV<sup>+</sup> samples (Figure 3A) and 32 in the HPV<sup>-</sup> samples (Figure 3B) presented with differentially expressed genes.

To characterize the effect of altered gene expression on various cancer progression pathways, the pathway analysis



**FIGURE 4** Pathway differences between high-grade dysplasia and invasive carcinoma. Individual pathway scores for high-grade dysplasia and invasive carcinoma are shown for (A) and (C) respectively in HPV<sup>+</sup> and HPV<sup>-</sup> tonsillar and base of tongue squamous cell carcinoma respectively. Boxplots for individual pathway scores for the nine most statistically significant different pathways are illustrated in (B) and (D) respectively for HPV<sup>+</sup> and HPV<sup>-</sup> tonsillar and base of tongue squamous cell carcinoma

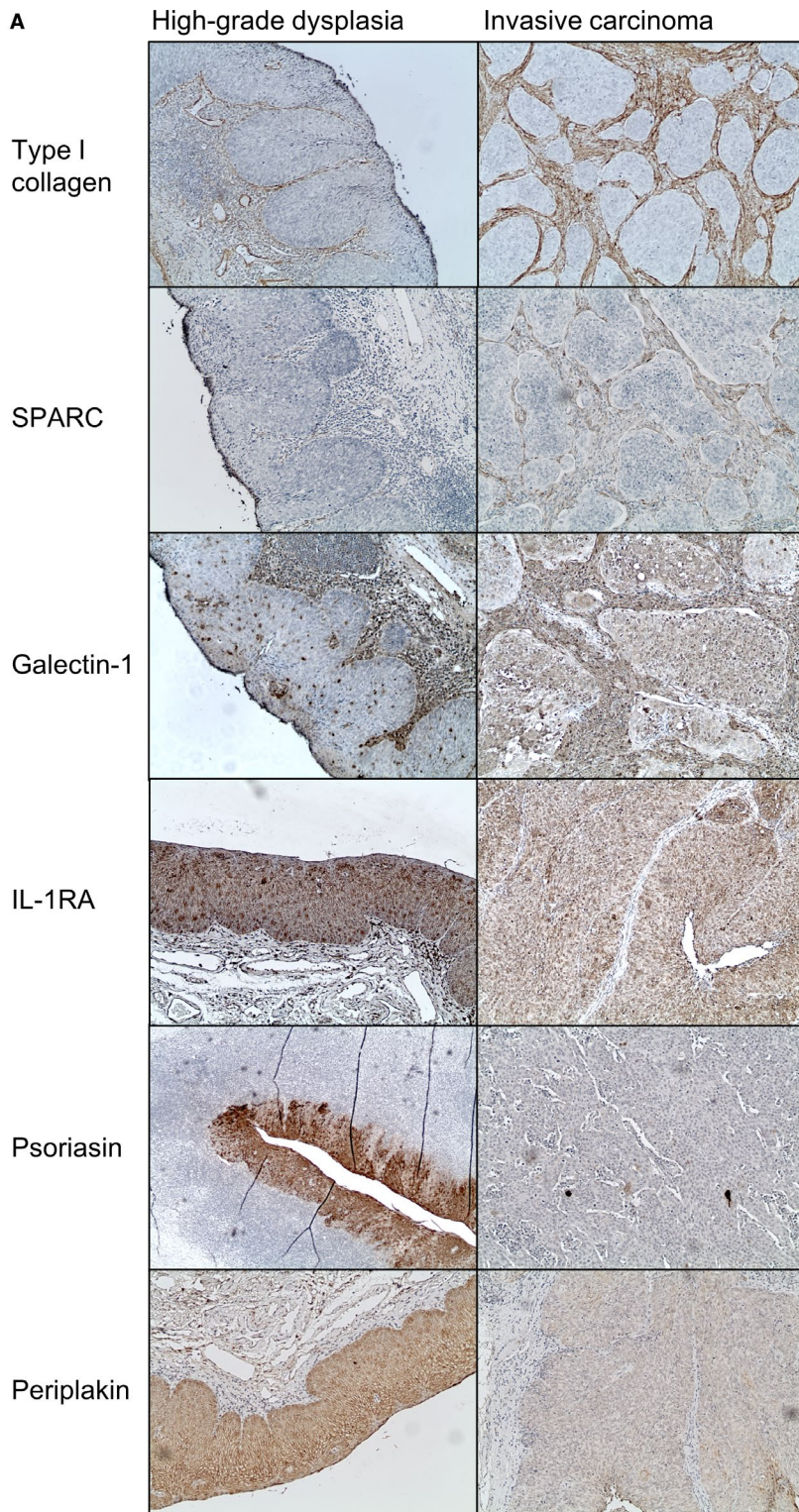


**FIGURE 5** Gene expression in high-grade dysplasia and invasive carcinoma of both HPV<sup>+</sup> and HPV<sup>-</sup> tonsillar and base of tongue squamous cell carcinoma samples combined, adjusted for HPV<sup>+</sup> status. (A) Volcano plot showing differentially expressed genes between high-grade dysplasia and invasive carcinoma by log<sub>2</sub> fold change (x-axis) and -log<sub>10</sub> of P-value (y-axis) in HPV<sup>+</sup> and HPV<sup>-</sup> tonsillar and base of tongue squamous cell carcinoma cases. The top 10 genes with increased and decreased linear as well as log<sub>2</sub> fold change expression between high-grade dysplasia and invasive carcinoma are together with P-values presented in the tables to the right. Genes marked with a color were selected for immunohistochemical protein analysis. (B) Graphs of selected genes showing normalized log<sub>2</sub> counts for tonsillar and base of tongue dysplastic lesions and invasive carcinoma (HPV<sup>+</sup> and HPV<sup>-</sup>) for each individual sample

tool provided in the nSolver Advanced Analysis Software was utilized. This tool condenses each sample's gene expression profile to calculate a pathway score using a first principal component analysis.<sup>23</sup> Notably, among the nine most significant pathways, there was a similar increased activity in invasive cancer as compared to high-grade dysplasia in

both HPV<sup>+</sup> and HPV<sup>-</sup> tumors in five pathways: metastasis response, extracellular matrix (ECM) receptor interaction, cellular growth factor, collagen family, and ECM structure (Figure 4).

In the search of new clinical invasion biomarkers in head and neck cancer, and because of low sample number,



**FIGURE 6** (A) Examples of immunohistochemistry stainings of type I collagen, SPARC, galectin-1, IL-1RA, psoriasin, and periplakin, in high-grade dysplasia/carcinoma in situ and in invasive carcinoma. Expression of type I collagen and SPARC is higher in the tissue surrounding invasive carcinoma, and galectin-1 has a higher expression in invasive carcinoma, whereas IL-1RA, psoriasin, and periplakin show a higher expression in high-grade dysplasia compared to invasive carcinoma. (B) Summary of immunohistochemical protein evaluation scores and *P*-values for selected proteins, HPV<sup>+</sup>, and HPV<sup>-</sup> tumors combined

**B**

	Number of samples with higher expression in high-grade dysplasia	Number of samples with higher expression invasively	Number of samples with no difference between high-grade dysplasia and invasive carcinoma	<i>P</i> -value
SPARC	0	17	6	<0.0001
Psoriasin	14	0	10	0.0001
Type I Collagen	0	13	10	0.0002
Galectin-1	1	9	13	0.0215
IL1-RA	5	0	19	0.0625
Galectin-1 TILs	1	6	16	0.1250
Periplakin	8	3	12	0.2266



the HPV<sup>+</sup> and HPV<sup>-</sup> groups were combined to investigate gene expression, adjusted for HPV status (Figure 5A). In invasive cancer vs. high-grade dysplasia, *COL1A1*, *COL3A1*, *COL1A2*, and *SPARC* were the most up-regulated genes, while *IL1RN*, *S100A7*, and *PPL* were the most down-regulated genes (Figure 5A). After analyzing the fold-change values for mRNA expression, *P*-values and published gene functions, six genes: *COL1A1*, *SPARC*, *LGALS1*, *IL1RN*, *S100A7*, and *PPL* (Figure 5B), were chosen as potential pathological biomarker candidates for further evaluation by IHC for protein expression.

### 3.2 | Protein expression in high-grade dysplasia and invasive TSCC and BOTSCC

Based on the results from the gene expression assay for the combined cohort of HPV<sup>+</sup> and HPV<sup>-</sup> TSCC, six proteins (type I collagen (*COL1A1*), SPARC (*SPARC*), galectin-1 (*LGALS1*), IL1-RA (*IL1RN*), psoriasin (*S100A7*) and periplakin (*PPL*)) were selected for evaluation by IHC as potential invasion marker candidates. The transcripts of these six proteins showed with few exceptions either an increase (type I collagen, SPARC and galectin-1) or decrease (IL1-RA, psoriasin and periplakin) between high-grade dysplasia and invasive cancer (Figure 5B).

Protein expression of the above genes was therefore evaluated on the extended cohort with 24 different FFPE tumor sections, but due to lack of sufficient tumor material one sample was excluded for type I collagen, SPARC, galectin-1, and periplakin. Details of the staining evaluation and scores are shown in Figure 6 and Supplementary Tables S3 and S4 respectively, and examples of the stainings are presented in Figure 6.

Neither type I collagen nor SPARC were found to be expressed at all in dysplastic epithelium or in invasive carcinomas. However, type I collagen was often expressed in the immediate ECM surrounding the cancer, but mostly to a lesser extent around high-grade dysplasia as compared to invasive tumor areas ( $P = .0002$ ) (Figure 6). Similarly, SPARC was often expressed in the peritumoral stroma in peritumoral fibroblasts, mostly with absent or little expression around dysplasia, whereas higher expression was found around 17/23 invasive cancer areas, while 6/23 samples showed no difference in expression ( $P < .0001$ ) (Figure 6).

Galectin-1 was mainly highly expressed in most tumors, with similar expression in high-grade dysplastic epithelium and invasive tumor areas in 13/23 cases, while 9/23 cases showed higher expression in the invasive component, and 1/23 showed lower expression ( $P = .02$ ) (Figure 6). Since galectin-1 is variably expressed on immune cells,<sup>24</sup> we also assessed tumor infiltrating lymphocytes (TILs) in dysplasia and invasive components, but no differences were observed in galectin-1 positive TILs in the different components (Figure 6).

Notably, 14/24 tumors showed that psoriasin was highly expressed in epithelial dysplasia, while the remaining samples showed no difference in expression ( $P = 0001$ ) (Figure 6). Periplakin was expressed very heterogeneously in both dysplasia and invasive tumor areas, often with no clear differences between dysplasia and invasive tumor, with a tendency toward having higher expression in invasive carcinoma. Likewise, IL1-RA expression was mostly strongly expressed in cancer cells, however, a significant difference between high-grade dysplasia and invasive cancer was not observed. (Figure 6).

## 4 | DISCUSSION

In this study we report, for the first time, differences in gene expression between high-grade dysplasia/cancer in situ and invasive HPV<sup>+</sup> and HPV<sup>-</sup> TSCC/BOTSCC. Some of the most significant differential mRNA expressions were found for *COL1A1*, *SPARC*, and *LGALS1* that were stronger expressed in invasive carcinoma compared to high-grade dysplasia, and the opposite was shown for *S100A7*, which was stronger expressed in high-grade dysplasia. These findings were also confirmed on a protein level by IHC. Moreover, we show that there are many similarities in gene expression and signal pathway activities between HPV<sup>+</sup> and HPV<sup>-</sup> tumors—especially in genes related to ECM, highlighting the importance of the tumor micro milieu.

As noted in the introduction, malignant invasive epithelial tumors develop from a dysplastic epithelium and detection of dysplasia has been utilized in different screening programs, for example, in cervical cancer. However, dysplasia in HPV<sup>+</sup> TSCC/BOTSCC is rarely detected and pure premalignant stages have not been found, but common in HPV<sup>-</sup> TSCC/BOTSCC. Some authors have even postulated that HPV<sup>+</sup> premalignant phases do not exist in TSCC/BOTSCC,<sup>16</sup> but have been described by others.<sup>14,25,26</sup> Subsequently, only two studies have to our knowledge investigated HPV<sup>+</sup> TSCC/BOTSCC dysplasia on a molecular level. Mooren et al, studied presence of p16INK4a expression in tonsillar dysplasia and found that p16INK4a positivity correlated to HPV status.<sup>25</sup> Additionally, another study by Masterson et al confirmed increased SYPC2 expression in premalignant carcinoma in situ and invasive carcinoma vs. normal epithelium by LCM and qPCR.<sup>26</sup> However, in this study we aimed to examine a wider range of genes and their differential expression between dysplastic and invasive HPV<sup>+</sup> and HPV<sup>-</sup> TSCC/BOTSCC. Notably, numerous similarities in differential expression were here observed between HPV<sup>+</sup> and HPV<sup>-</sup> TSCC/BOTSCC dysplasia and invasive cancer, suggesting that HPV<sup>+</sup> dysplasia in tonsils/base of tongue is a distinguishable entity from invasive HPV<sup>+</sup> TSCC/BOTSCC. Therefore, it is very likely that premalignant stages of HPV<sup>+</sup> TSCC/BOTSCC exist. Why such stages have not been described so far,<sup>14,15</sup> may be due to that tonsillar/

base of tongue neoplasias are uncommon or possibly due to a more aggressive behavior of HPV<sup>+</sup> dysplasia, an assumption that is not quite supported by data presented above. Hence, larger cohorts may be needed to identify premalignant phases in tonsils/base of tongue—possibly also by utilizing markers described here, such as psoriasin.

On the signaling pathway level, it is not surprising that pathways involved in tumor progression, such as metastasis response, ECM receptor interaction, ECM structure, collagen family, and cellular growth factor pathways are up-regulated both in HPV<sup>+</sup> and HPV<sup>-</sup> invasive carcinoma. Interestingly, we also observed a higher stem cell associated score in invasive cancer compared to dysplasia, which was especially observed in the HPV<sup>+</sup> samples while present to a lesser extent in the HPV<sup>-</sup> samples. Notably, a study by Zhang et al showed that the cancer stem cell pool was much higher in HPV<sup>+</sup> OPSCC as compared to HPV<sup>-</sup> cancer.<sup>27</sup> However, high expression of stem cell markers have been shown to be associated with a poor prognosis in HNSCC,<sup>28-30</sup> yet most patients with HPV<sup>+</sup> tumors in general have a very good prognosis.

On the transcript level, the collagen genes *COL1A1*, *COL1A2*, *COL3A1*, where in our study found to be the most significantly increased ones in invasive cancer as compared to dysplasia in both HPV<sup>+</sup> and HPV<sup>-</sup> tumors. Moreover, the proteins encoded by these genes belong to a family of proteins that strengthen and support many tissues in the body, being the most abundant part of the ECM.<sup>31</sup> Their expression has however also been linked to cell proliferation and tumor progression and the correlation to prognosis in various cancer types.<sup>32</sup> Thus, collagen has been suggested as a double-edged sword in tumor progression, in one way acting as a passive barrier to resist tumor cells, in another way sending out biochemical and biophysical signals that affect, for example, cell adhesion, migration, angiogenesis, and repair systems that in cancer typically can be deregulated causing tumor progression.<sup>33</sup> Another matricellular protein, SPARC (secreted protein acidic and rich in cysteine), which regulates cell-matrix interactions and signaling pathways in cells, was also observed to increase in invasive cancer as compared to dysplasia. SPARC, which is considered to be solely produced by cancer associated fibroblasts in the tumor stroma, is suggested to regulate tumor cell growth and metastasis.<sup>34,35</sup> In a study by Witkiewicz et al, SPARC expression was correlated to tumor recurrence in ductal carcinoma in situ (DCIS) of the breast.<sup>36</sup> Moreover, Chang et al show that in HNSCC a higher SPARC expression correlated to a higher tumor grade,<sup>37</sup> but the prognostic role of SPARC expression is however still inconsistent within and between cancer types.<sup>35</sup>

*S100A7* encodes the protein psoriasin part of the S100 family containing calcium-binding motifs and is an important cell mediator for, for example, cell survival and maturation and has been associated to tumor progression and survival.<sup>38</sup> In a study by Liu et al, psoriasin was shown to promote

invasion and survival of pancreatic cancer cells.<sup>39</sup> Other studies have shown that psoriasin is more highly expressed in, for example, squamous skin cancer in situ compared to invasive squamous skin cancer,<sup>40</sup> and in DCIS compared to invasive breast cancer,<sup>41,42</sup> which is in line with our findings. Tripathi et al have shown that psoriasin is significantly overexpressed in oral leukoplakia lesions with squamous cell hyperplasia or dysplasia as well as in HNSCC, compared to normal tissue.<sup>43</sup> Moreover, they show that patients with a nuclear accumulation of psoriasin in HNSCC have a reduced disease-free survival.<sup>43</sup>

*IL1RN* encodes the interleukin-1 receptor antagonist (IL-1RA) and is a natural inhibitor of the IL-1 receptor. Expression of IL-1RA protein has been suggested to be able to decrease tumor growth, metastasis, and angiogenesis in mouse models.<sup>44</sup> In this study, we show that *IL1RN* gene expression is lower in invasive cancer compared to dysplastic epithelium, especially in HPV<sup>+</sup> TSCC/BOTSCC. Interestingly, similar results have previously also been published in CIN3 vs. invasive cervical carcinoma.<sup>18</sup>

Lastly, in order to validate our data and in search for new pathological invasion biomarkers, we assessed six gene transcripts for protein expression and tissue distribution by IHC and were able to validate four out of six transcripts at the protein level. However, while the relative expression in dysplasia and invasive carcinoma often corresponded to the trend of transcriptional data, the absolute expression varied considerably between samples. Therefore, using any of these markers as markers of invasion in a clinical setting would probably need further tuning of staining interpretation.

There are some limitations in this study. First, tumor areas containing high-grade dysplasia/carcinoma in situ were often scanty and decreased upon serial sectioning in some cases. Hence, the RNA concentration was very low after extraction. Due to low RNA concentrations and fragmented FFPE material, the NanoString platform using tumor progression RNA panel was utilized. However, using a more “global” method, such as RNAseq, would have yielded more data and a better picture of the transition between dysplasia and invasive tumor. Secondly, few patient samples were included in this study, reflecting that dysplasia is uncommonly detected in HPV<sup>+</sup> TSCC/BOTSCC.

In conclusion, this is the first study to our knowledge, examining and disclosing differences in gene expression between dysplastic and invasive HPV<sup>+</sup> and HPV<sup>-</sup> TSCC/BOTSCC. However, larger studies are needed to distinguish pure premalignant from invasive tonsillar/base of tongue neoplasias.

## ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Ethical approval and consent to participate was conducted according to ethical permission 2005/431-31/4 from the Ethics Committee at Karolinska Institute, Stockholm, Sweden. The

study was performed in accordance with the Declaration of Helsinki.

## ACKNOWLEDGMENTS

The authors thank professor Tina Dalianis for providing tumor material, the Core Facility for Tissue Based Diagnostics and Research at Karolinska Institute for assistance with the micro dissector and the Ki Gene core facility at Karolinska Institutet for the use of their Nanostring technologies as well as their expertise and support during the gene expression assays.

## CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

## AUTHOR CONTRIBUTIONS

L.H., A.N., and T.R. formulated the research question and came up with the study design. Sample selection and collection was done by A.N., P.F., and L.H. L.H. and A.N. performed the tissue microdissection, RNA extraction, and gene expression assay. A.Ä.R., L.M., and R.G.U. performed the IHC stainings. L.H., A.N., and L.M. evaluated the IHC stainings. L.H. and A.N. analyzed, summarized, and interpreted the data and wrote the manuscript which was revised and approved by all co-authors.

## ORCID

Torbjörn Ramqvist  <https://orcid.org/0000-0003-3891-496X>  
Anders Näsman  <https://orcid.org/0000-0003-4602-4297>

## REFERENCES

- International Agency for Research on Cancer. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Volume 90, Human Papillomaviruses*. Lyon, France: IARC; 2007.
- Gillison ML. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *J Natl Cancer Inst*. 2000;92:709-720.
- Lindquist D, Romanitan M, Hammarstedt L, et al. Human papillomavirus is a favourable prognostic factor in tonsillar cancer and its oncogenic role is supported by the expression of E6 and E7. *Mol Oncol*. 2007;1:350-355.
- Attner P, Du J, Näsman A, et al. Human papillomavirus and survival in patients with base of tongue cancer. *Int J Cancer*. 2011;128:2892-2897.
- Näsman A, Nordfors C, Holzhauser S, et al. Incidence of human papillomavirus positive tonsillar and base of tongue carcinoma: a stabilisation of an epidemic of viral induced carcinoma? *Eur J Cancer*. 2015;51:55-61.
- Näsman A, Attner P, Hammarstedt L, et al. Incidence of human papillomavirus (HPV) positive tonsillar carcinoma in Stockholm, Sweden: an epidemic of viral-induced carcinoma? *Int J Cancer*. 2009;125:362-366.
- Hocking JS, Stein A, Conway EL, et al. Head and neck cancer in Australia between 1982 and 2005 show increasing incidence of potentially HPV-associated oropharyngeal cancers. *Br J Cancer*. 2011;104(5):886-891.
- Chaturvedi AK, Anderson WF, Lortet-Tieulent J, et al. Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers. *J Clin Oncol*. 2013;31(36):4550-4559.
- Grisar K, Dok R, Schoenaers J, et al. Differences in human papillomavirus-positive and -negative head and neck cancers in Belgium: an 8-year retrospective, comparative study. *Oral Surg Oral Med Oral Pathol Oral Radiol*. 2016;121:456-460.
- Blitzer GC, Smith MA, Harris SL, Kimple RJ. A review of the clinical and biologic aspects of HPV-positive HNC. *Int J Radiat Oncol Biol Phys*. 2014;15:761-770.
- Weinberg RA. *The Biology of Cancer*. 1st edn. New York: Garland Science, Taylor & Francis Group; 2007.
- Waxman AG, Chelmsow D, Darragh TM, Lawson H, Moscicki AB. Revised terminology for cervical histopathology and its implications for management of high-grade squamous intraepithelial lesions of the cervix. *Obstet Gynecol*. 2012;120(6):1465-1471.
- Fossum CC, Chintakuntlawar AV, Price DL, Garcia JJ. Characterization of the oropharynx: anatomy, histology, immunology, squamous cell carcinoma and surgical resection. *Histopathology*. 2017;70:1021-1029.
- Begum S, Cao D, Gillison M, Zahurak M, Westra WH. Tissue distribution of human papillomavirus 16 DNA integration in patients with tonsillar carcinoma. *Clin Cancer Res*. 2005;11:5694-5699.
- Palmer E, Newcombe RG, Green AC, et al. Human papillomavirus infection is rare in nonmalignant tonsil tissue in the UK: implications for tonsil cancer precursor lesions. *Int J Cancer*. 2014;135:2437-2443.
- Holmes BJ, Wenig BM. Virus-associated carcinomas of the head & neck: Update from the 2017 WHO classification. *Ann Diagn Pathol*. 2019;38:29-42. <https://doi.org/10.1016/j.anndiagpath.2018.10.008>.
- Gius D, Funk MC, Chuang EY, et al. Profiling microdissected epithelium and stroma to model genomic signatures for cervical carcinogenesis accommodating for covariates. *Cancer Res*. 2007;67:7113-7123.
- den Boon JA, Pyeon D, Wang SS, et al. Molecular transitions from papillomavirus infection to cervical precancer and cancer: role of stromal estrogen receptor signaling. *Proc Natl Acad Sci*. 2015;112:E3255-64. <https://doi.org/10.1073/pnas.1509322112>.
- Herrington CS. Do HPV-negative cervical carcinomas exist?—revisited. *J Pathol*. 1999;189:1-3.
- Burd EM. Human papillomavirus and cervical cancer. *Clin Microbiol Rev*. 2003;16:1-17.
- Skyldberg B, Salo S, Eriksson E, et al. Laminin-5 as a marker of invasiveness in cervical lesions. *J Natl Cancer Inst*. 1999;91(21):1882-1887.
- Haeggbloom L, Attoff T, Yu Y, et al. Changes in incidence and prevalence of human papillomavirus in tonsillar and base of tongue cancer during 2000–2016 in the Stockholm region and Sweden. *Head Neck*. 2018;41(6):1583-1590.
- Tomfohr J, Lu J, Kepler TB. Pathway level analysis of gene expression using singular value decomposition. *BMC Bioinformatics*. 2005;6:225.

24. Cedeno-Laurent F, Dimitroff CJ. Galectin-1 research in T cell immunity: past, present and future. *Clin Immunol.* 2012;142:107-116.
25. Mooren JJ, Gültekin SE, Straetmans J, et al. P16INK4A immunostaining is a strong indicator for high-risk-HPV-associated oropharyngeal carcinomas and dysplasias, but is unreliable to predict low-risk-HPV-infection in head and neck papillomas and laryngeal dysplasias. *Int J Cancer.* 2014;134:2108-2117. <https://doi.org/10.1002/ijc.28534>.
26. Masterson L, Sorgeloos F, Winder D, et al. Deregulation of SYCP2 predicts early stage human papillomavirus-positive oropharyngeal carcinoma: a prospective whole transcriptome analysis. *Cancer Sci.* 2015;106:1568-1575.
27. Zhang M, Kumar B, Piao L, et al. Elevated intrinsic cancer stem cell population in human papillomavirus-associated head and neck squamous cell carcinoma. *Cancer.* 2014;120(7):992-1001.
28. Fan Z, Li M, Chen X, et al. Prognostic value of cancer stem cell markers in head and neck squamous cell carcinoma: a meta-analysis. *Sci Rep.* 2017;7:43008. <https://doi.org/10.1038/srep43008>.
29. Joshua B, Kaplan MJ, Doweck I, et al. Frequency of cells expressing CD44, a Head and Neck cancer stem cell marker: Correlation with tumor aggressiveness. *Head Neck.* 2012;34:42-49.
30. Näsman A, Nordfors C, Grün N, et al. Absent/weak CD44 intensity and positive human papillomavirus (HPV) status in oropharyngeal squamous cell carcinoma indicates a very high survival. *Cancer Med.* 2013;2:507-518.
31. Ricard-Blum S. The collagen family. *Cold Spring Harb Perspect Biol.* 2011;3:1-19.
32. Kirkland SC. Type I collagen inhibits differentiation and promotes a stem cell-like phenotype in human colorectal carcinoma cells. *Br J Cancer.* 2009;101:320-326.
33. Fang M, Yuan J, Peng C, Li Y. Collagen as a double-edged sword in tumor progression. *Tumor Biol.* 2014;35:2871-2882.
34. Framson PE, Sage EH. SPARC and tumor growth: where the seed meets the soil? *J Cell Biochem.* 2004;92:679-690.
35. Paulsson J, Mücke P. Prognostic relevance of cancer-associated fibroblasts in human cancer. *Semin Cancer Biol.* 2014;25:61-68.
36. Witkiewicz AK, Freydy B, Chervoneva I, et al. Stromal CD10 and SPARC expression in ductal carcinoma in situ (DCIS) patients predicts disease recurrence. *Cancer Biol Ther.* 2010;10:391-396.
37. Chang C-H, Yen M-C, Liao S-H, et al. Secreted protein acidic and rich in cysteine (SPARC) enhances cell proliferation, migration, and epithelial mesenchymal transition, and SPARC expression is associated with tumor grade in head and neck cancer. *Int J Mol Sci.* 2017;18:1556.
38. Hattinger E, Zwicker S, Ruzicka T, Yuspa SH, Wolf R. Opposing functions of psoriasis (S100A7) and koebnerisin (S100A15) in epithelial carcinogenesis. *Curr Opin Pharmacol.* 2013;13:588-594.
39. Liu Y, Bunston C, Hodson N, et al. Psoriasis promotes invasion, aggregation and survival of pancreatic cancer cells; association with disease progression. *Int J Oncol.* 2017;50:1491-1500.
40. Alowami S, Qing G, Emberley E, Snell L, Watson PH. Psoriasis (S100A7) expression is altered during skin tumorigenesis. *BMC Dermatol.* 2003;3:1. <https://doi.org/10.1186/1471-5945-3-1>.
41. Petersson S, Bylander A, Yhr M, Enerbäck C. S100A7 (Psoriasis), highly expressed in Ductal Carcinoma In Situ (DCIS), is regulated by IFN-gamma in mammary epithelial cells. *BMC Cancer.* 2007;7:205. <https://doi.org/10.1186/1471-2407-7-205>.
42. Emberley ED, Murphy LC, Watson PH. S100A7 and the progression of breast cancer. *Breast Cancer Res.* 2004;6:153-159. <https://doi.org/10.1186/bcr816>.
43. Tripathi SC, Matta A, Kaur J, et al. Nuclear S100A7 is associated with poor prognosis in head and neck cancer. *PLoS ONE.* 2010;5:e11939. <https://doi.org/10.1371/journal.pone.0011939>.
44. Lewis AM, Varghese S, Xu H, Alexander HR. Interleukin-1 and cancer progression: the emerging role of interleukin-1 receptor antagonist as a novel therapeutic agent in cancer treatment. *J Transl Med.* 2006;4:48. <https://doi.org/10.1186/1479-5876-4-48>.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Haegglom L, Ährlund-Richter A, Mirzaie L, et al. Differences in gene expression between high-grade dysplasia and invasive HPV<sup>+</sup> and HPV<sup>-</sup> tonsillar and base of tongue cancer. *Cancer Med.* 2019;8:6221-6232. <https://doi.org/10.1002/cam4.2450>