

## RESEARCH ARTICLE

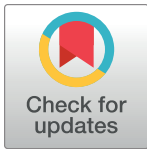
## Comparing serum protein levels can aid in differentiating HPV-negative and -positive oropharyngeal squamous cell carcinoma patients

Amy Dickinson<sup>1,2\*</sup>, Mayank Saraswat<sup>3,4</sup>, Stina Syrjänen<sup>5,6</sup>, Tiialotta Tohmola<sup>3</sup>, Robert Silén<sup>3</sup>, Reija Randén-Brady<sup>7</sup>, Timo Carpén<sup>1,7</sup>, Jaana Hagström<sup>5,7,8</sup>, Caj Haglund<sup>8,9</sup>, Petri Mattila<sup>1</sup>, Antti Mäkitie<sup>1,2,10</sup>, Sakari Joenväärä<sup>3,4</sup>, Suvi Silén<sup>1,11</sup>

**1** Department of Otorhinolaryngology—Head and Neck Surgery, University of Helsinki and Helsinki University Hospital, Helsinki, Finland, **2** Research Program in Systems Oncology, Faculty of Medicine, University of Helsinki, Helsinki, Finland, **3** Transplantation Laboratory, Haartman Institute, University of Helsinki, Helsinki, Finland, **4** HUSLAB, Helsinki University Hospital, Helsinki, Finland, **5** Department of Oral Pathology and Oral Radiology, University of Turku, Turku, Finland, **6** Department of Pathology, Turku University Hospital, Turku, Finland, **7** Department of Pathology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland, **8** Department of Surgery, University of Helsinki and Helsinki University Hospital, Helsinki, Finland, **9** Research Programs Unit, Translational Cancer Medicine, University of Helsinki, Helsinki, Finland, **10** Division of Ear, Nose and Throat Diseases, Department of Clinical Sciences, Intervention and Technology, Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden, **11** Department of Biosciences and Nutrition, Karolinska Institutet, Stockholm, Sweden

✉ Current address: Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, United States of America

\* [Amy.dickinson@helsinki.fi](mailto:Amy.dickinson@helsinki.fi)


 OPEN ACCESS

**Citation:** Dickinson A, Saraswat M, Syrjänen S, Tohmola T, Silén R, Randén-Brady R, et al. (2020) Comparing serum protein levels can aid in differentiating HPV-negative and -positive oropharyngeal squamous cell carcinoma patients. PLoS ONE 15(6): e0233974. <https://doi.org/10.1371/journal.pone.0233974>

**Editor:** Randall J. Kimple, University of Wisconsin, UNITED STATES

**Received:** February 21, 2020

**Accepted:** May 16, 2020

**Published:** June 15, 2020

**Copyright:** © 2020 Dickinson et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The dataset supporting the conclusions of this article is available in the [PrideXchange Consortium](https://proteomecentral.proteomexchange.org/cgi/GetDataset) via the PRIDE partner repository, Proteome Central, ProteomeXchange with the data set identifier PXD008445, available at (<http://proteomecentral.proteomexchange.org/cgi/GetDataset>).

**Funding:** This work was supported by the Helsinki University Hospital Research Fund to AM, University of Helsinki to S. Silén, Jane ja Aatos

## Abstract

## Background

The surrogate immunohistochemical marker, p16<sup>INK4a</sup>, is used in clinical practice to determine the high-risk human papillomavirus (HPV) status of oropharyngeal squamous cell carcinomas (OPSCC). With a specificity of 83%, this will misclassify some patients compared with direct HPV testing. Patients who are p16<sup>INK4a</sup>-positive but HPV DNA-negative, or RNA-negative, may be unsuitable for treatment de-escalation aimed at reducing treatment-related side effects. We aimed to identify cost-effective serum markers to improve decision making for patients at risk of misclassification by p16<sup>INK4a</sup> alone.

## Methods

Serum proteins from pre-treatment samples of 36 patients with OPSCC were identified and quantified using label-free mass spectrometry-based proteomics. HPV-status was determined using p16<sup>INK4a</sup>/HPV DNA and E6/E7 mRNA. Serum protein expressions were compared between groups of patients according to HPV status, using the unpaired t-test with a Benjamini-Hochberg correction. ROC curves (AUC) were calculated with SPSS (v25).

Erkon Säätiö to S. Silén, Korvatautien Tutkimussäätiö to S. Silén and AD, Emil Aaltonen Research Foundation to AD, and the Finnish-Norwegian Medical Foundation (Suomalais-Norjalainen Lääketieteen Säätiö) to AD. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Results

Of 174 serum proteins identified, complement component C7 (C7), apolipoprotein F (ApoF) and galectin-3-Binding Protein (LGALS3BP) significantly differed between HPV-positive and -negative tumors (AUC ranging from 0.84–0.87). ApoF levels were more than twice as high in the E6/E7 mRNA HPV-positive group than HPV-negative.

## Conclusions

Serum C7, ApoF and LGALS3BP levels discriminate between HPV-positive and HPV-negative OPSCC. Further studies are needed to validate these host immunity-related proteins as markers for HPV-associated OPSCC.

## Introduction

Human papilloma virus (HPV) is an established aetiological factor for oropharyngeal squamous cell carcinoma (OPSCC), with nearly 90% of HPV-positive OPSCCs being caused by the high-risk genotype HPV16 [1]. The incidence of OPSCC has been rising in Western countries, mainly due to the increase in HPV-positive disease. HPV prevalence in OPSCC is globally variable [1–5], and it also relates to the methods used for HPV detection as recently shown in a comprehensive assessment suggesting that only 25% of all OPSCCs are HPV positive worldwide [6]. Moreover, HPV-positive OPSCCs have a different pattern of histopathological features and clinical behavior compared with HPV-negative OPSCCs [7, 8]. They tend to occur in younger patients, and although HPV-positive tumors are more likely to present with nodal metastases, they generally have a favorable prognosis compared with alcohol- and smoking-related, non-HPV OPSCCs [9–11]. Furthermore, patients with HPV-related OPSCC may benefit from de-escalated treatment regimens, intended to minimize toxicity [12].

The difference in the tumor behavior is reflected in the latest WHO classification of head and neck tumors, in which HPV-positive and HPV-negative OPSCCs are classified as distinct clinical entities, using the overexpression of p16<sup>INK4a</sup> as a surrogate immunohistochemical marker for HPV status [13, 14]. According to a meta-analysis by Prigge et al., p16<sup>INK4a</sup> testing has a sensitivity of 94% [15]. CDKN2A, the gene encoding p16<sup>INK4a</sup>, is tumor suppressor gene through the inhibition of the catalytic activity of the cyclin D1/CDK4/6 complex. The cyclin D1/CDK4/6 complex phosphorylates the retinoblastoma protein (pRb), resulting in the release of the transcription factor E2F and the initiation of cell-cycle progression. Conversely, the pRb/E2F repressor complex inhibits transcription of several genes, including CDKN2A [16]. By inactivating pRb, HPV-E7 releases the CDKN2A gene from its transcriptional inhibition that results in p16<sup>INK4a</sup> overexpression, and it accumulates within the cells [17]. Despite its high sensitivity, p16<sup>INK4a</sup> has a specificity of 83% [15] as it is involved in other cellular processes such as cell senescence [18]. Frequent HPV-independent overexpression of p16<sup>Ink4a</sup> in head and neck carcinomas has been explained by mutation or amplification of CDKN2A or by mutations in RB1 or histone H3 lysine 36 methyltransferase genes [19].

The accepted gold standard method for the detection of HPV with an aetiological role in OPSCC is testing for E6/E7 mRNA, based on the assumption that E6 and E7 oncogenes are persistently transcribed in HPV-associated carcinomas [20]. However, E6/E7 mRNA detection is methodologically more challenging, thus limiting its use in clinical routine [15]. The meta-analysis by Prigge et al. compared p16<sup>INK4a</sup> and HPV DNA detection methods, alone and

combined, and suggested that p16<sup>INK4a</sup> immunohistochemistry (IHC) combined with HPV DNA PCR testing has an enhanced specificity (96%) while maintaining high sensitivity (93%) [15]. Mena et al. reported recently that the combination of HPV DNA and p16<sup>INK4a</sup> positivity also provided the strongest diagnostic accuracy as well as prognostic value in OPSCC [21]. Currently, the p16<sup>INK4a</sup> surrogate marker is used in the clinic to determine whether the OPSCC is HPV-positive or HPV-negative. However, there is a specific subgroup of OPSCCs, which are p16<sup>INK4a</sup>-positive but negative for HPV DNA, and these have been reported to have distinctive clinical and morphological features, and a worse prognosis and increased risk of distant metastases than those positive for both p16<sup>INK4a</sup> and HPV DNA [7, 22, 23].

At present, there is a consistent misclassification of some patients with OPSCC, as no method for determining the HPV status of an OPSCC is completely accurate. Therefore, we set about identifying serum proteins that could be simply and cost-effectively tested, to improve the detection of mis-classified patients.

To this end, we have used a label-free quantitative mass spectrometry methodology to identify the serum proteome in patients with OPSCC, comparing the protein expression levels between patients according to p16<sup>INK4a</sup>, DNA and E6/E7 mRNA statuses: E6/E7 mRNA -positive vs -negative, and p16<sup>INK4a</sup>/DNA -positive vs -negative. The aim of this was to identify serum proteins that could discriminate between the p16<sup>INK4a</sup>+ OPSCCs that are HPV-associated, and those that are p16<sup>INK4a</sup>+ but HPV-independent.

## Materials and methods

### Workflow and data availability

The overall workflow is shown in [Fig 1](#).

It is worth noting that the serum proteomics data is from an overarching serum proteomics mass spectrometry project and overlaps with another cohort that has previously been published by Tuhkuri et al. [24], Tuhkuri et al. compared serum proteins from early-stage OPSCC patients with controls and also separately compared early-stage p16+ and p16- OPSCCs with each other. In comparison with that study, we have used all stages of tumors, and have used additional methods to determine HPV-positivity. The mass spectrometry data for the samples used in both overlapping projects are available in the same overall project in the ProteomeX-change Consortium via the PRIDE partner repository with the data set identifier PXD008445 [25]. [S1 Table](#) explains which data files were used in this paper.

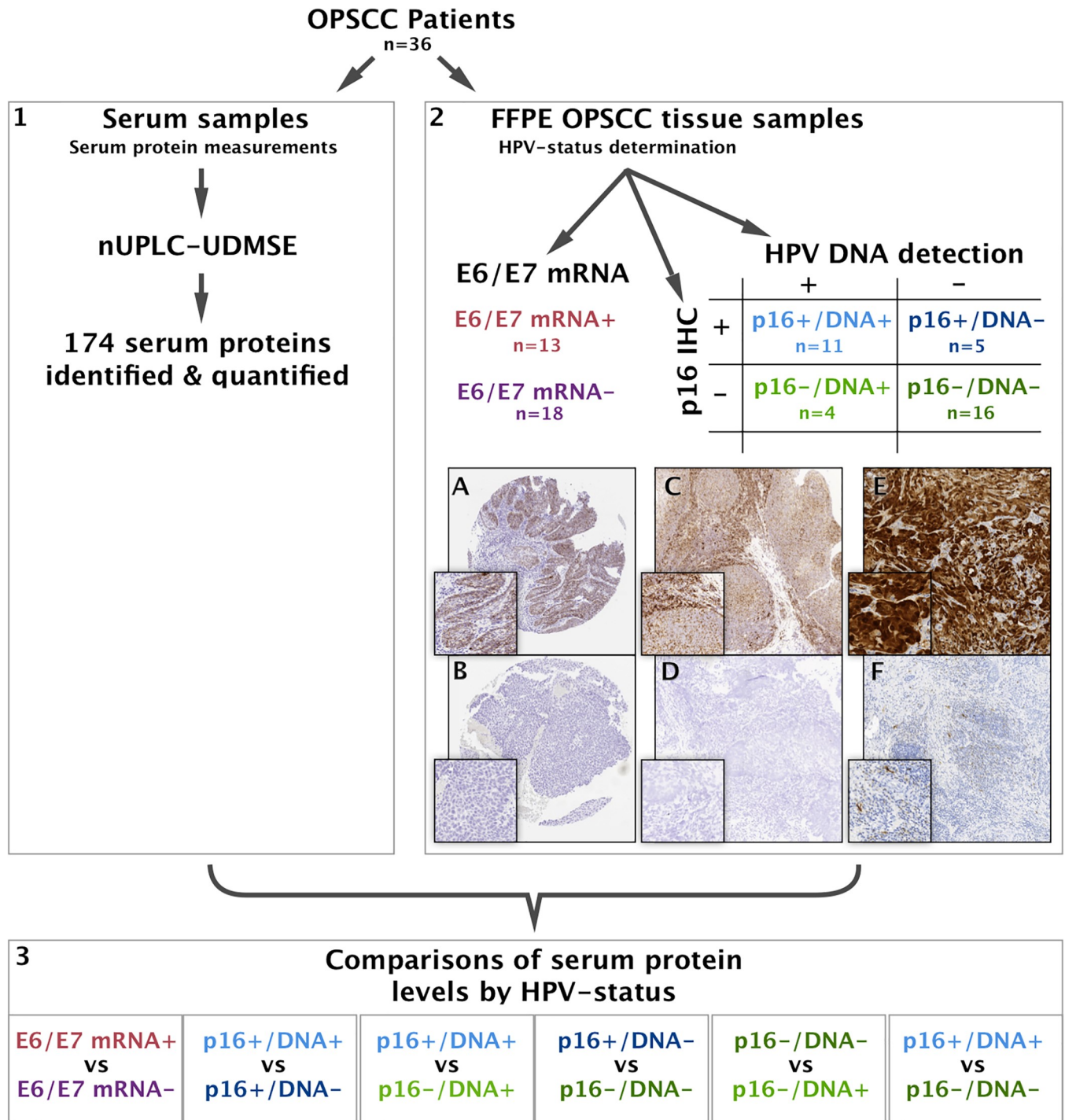
The data from HPV DNA detection, p16 and RNA scope on these patients has also been published elsewhere by Randen-Brady et al. [26] (the PM II cohort).

### Patients and serum samples

Pre-treatment serum samples, as well as formalin-fixed paraffin-embedded (FFPE) samples, were collected from 36 patients with both of these samples available, diagnosed with OPSCC between the years 2012 and 2016 at the Department of Otorhinolaryngology–Head and Neck Surgery, Helsinki University Hospital, Helsinki, Finland. Clinical data were collected from patient records. The UICC (Union for International Cancer Control) 8<sup>th</sup> edition TNM classification [14] was used to evaluate tumor stage.

After being allowed to clot at room temperature, blood samples were centrifuged at 4°C at 1000xg for 10 minutes in order to separate the serum. Sera were aliquoted and stored at -70°C, until they were all assayed at the same time.

The study plan was approved by the Helsinki University Hospital's institutional Research Ethics Board (DNr. 51/13/03/02/2013). Informed written consent was obtained from all



**Fig 1. Workflow.** Serum samples and FFPE samples were analysed from patients with OPSCC. 1. Serum proteins were identified and quantified using nano ultra-performance liquid chromatography—ultra definition mass spectrometry. 2. From formalin-fixed paraffin-embedded samples, HPV-status was determined using the following methods: E6/E7 mRNA was detected using RNAScope, to detect transcriptionally active HPV; p16 was detected using p16INK immunohistochemistry; HPV DNA was detected using PCR. 2A shows a representative TMA spot positive for E6/E7 mRNA; 2B shows a representative TMA spot negative for E6/E7 mRNA; 2C shows the positive PPIB probe for RNA analysis; 2D shows DAPB negative probe for the RNA analysis; 2E shows a representative p16 positive FFPE slide; 2F shows a representative p16 negative FFPE slide. Magnification is 100x and 200x. 3. Combining the serum proteomic data with the determined HPV-status, we compared the protein abundance between the groups.

<https://doi.org/10.1371/journal.pone.0233974.g001>

patients. The study was performed in accordance with the latest version of the Declaration of Helsinki.

### Tissue p16<sup>INK4a</sup> detection

p16<sup>INK4a</sup> was detected immunohistochemically on FFPE tissue samples on individual slides as previously described [26, 27]. Briefly, after deparaffinization and rehydration a pre-treatment-module (Lab Vision Corp., UK Ltd., UK) was used for treating the tissue slides in Tris-HCl buffer (pH 8.5). Endogenous peroxidase was blocked using 0.3% Dako REAL Peroxidase-Blocking Solution. The Dako REAL Envision detection system (Agilent Technologies Inc, CA, USA) was used according to manufacturer's instructions, and the primary antibody was the "ready-to-use" monoclonal mouse anti-human p16<sup>INK4a</sup> (9517 CINtec Histology Kit, MTM laboratories, Germany). If over 70% of tumor cells were strongly immunopositive for p16<sup>INK4a</sup>, the tumor was defined as p16<sup>INK4a</sup> positive [28]. As a positive control a known p16+ tonsil cancer was used, and negative controls were incubated without the primary antibody.

### Tissue HPV DNA detection

An HPV PCR assay was used to detect HPV DNA in the tissue samples. Firstly, DNA was extracted from 4 to 10  $\mu$ m diameter sections cut from formalin fixed paraffin embedded (FFPE) tissue sections using the high salt method as described earlier [29]. For HPV genotyping, Multiplex HPV Genotyping Kit® (DiaMEX GmbH, Germany) was used as described earlier [22, 26, 27]. Primer Set 1 contains all HPV primers: 9 biotinylated forward and 3 reverse primers for amplifying the HPV types under investigation. Primer Set 2 (DNA quality control primers) contains primers for the amplification of a  $\beta$ -globin gene fragment to verify the amount and the quality of human genomic sample DNA. Every eighth sample was a negative control without any genomic DNA to confirm the absence of contamination in the amplification reactions [30]. The labeled hybrids were analyzed with a Luminex LX-100 analyzer (Bio-Plex 200 System, Bio-Rad Laboratories, Hercules, USA). The Multiplex HPV Genotyping Kit detects 24 LR- and HR-HPV-genotypes as follows: LR-HPV6, 11, 42, 43, 44, and 70; and HR-HPV16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 73 and 82.

### In situ hybridization for high-risk HPV E6/E7 mRNA

In situ hybridization (ISH) for high risk HPV E6/E7 mRNA was performed manually using the RNAscope® 2.5 HD Reagent kit (Advanced Cell Diagnostics, Inc., Hayward, CA) according to the manufacturer's protocol as described also in detail in our recent papers [26, 31]. Five- $\mu$ m formalin-fixed and paraffin-embedded tissue microarray (TMA) sections were incubated for 1 hour at 58°C to fix the samples onto slides. After deparaffinization, the sections were pre-treated with hydrogen peroxidase (RNAscope® Hydrogen Peroxide) for 10 minutes at RT. Target retrieval was performed (RNAscope® Target Retrieval Reagents) for 15 minutes at 100°C. The sections underwent protease treatment (RNAscope® Protease Plus) for 30 minutes at 40°C in the hybridization oven followed by hybridization with the HR HPV 18 cocktail probe (RNAscope®) for genotypes 16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 73 and 82 for 2 hours at 40°C in the hybridization oven. Pre-amplifiers and amplifiers were hybridized consecutively, accompanied by chromogenic signal detection with 3,3'-diaminobenzidine (DAB). Finally, the slides were counterstained with hematoxylin. An endogenous housekeeping gene HS-PPIB (RNAscope®) probe was used as a positive control, and a bacterial gene DapB, diaminopimelate (RNAscope®) probe used as a negative control. The staining was examined using a qualitative scoring system: punctate brown, granular nuclear and



cytoplasmic dots were either observed or not in the individual TMA spots. 1–3 brown dots or more/tumor cell was recorded as a positive staining result.

### Serum treatment and protein digestion

These methods have been described previously in detail [32] and also by Tuhkuri et al. [24]. Briefly, the top 12 proteins were depleted from thawed serum samples, which increases the dynamic range of the lower abundance proteins. Total protein levels in each sample were measured using a bicinchoninic acid assay kit (Pierce, Thermo Scientific, Rockford, IL). The samples, amounting to 350 µg of total protein were dried in a speed vacuum then dissolved in a urea-containing digestion buffer. They were digested overnight using trypsin (Promega, Madison, WI), having been reduced and alkylated. Then they were purified using C18 columns (Thermo Scientific, Rockford, IL), and resuspended in a solution of 0.1% formic acid, which contained 12.5 gmol/µl of Hi3 E.coli peptides (Waters Corporation, MA). All described procedures were performed according to the manufacturer's instructions, where applicable.

### Liquid chromatography–ultra-definition mass spectrometry (LC-UDMS<sup>E</sup>)

This was performed as previously described in detail [24] and also by Tuhkuri et al. [24]. Four µL samples corresponding to approximately 1.4 µg of total protein, were injected into the nano ultra-performance liquid chromatography (UPLC) system (Waters Corporation, MA) [33]. The TRIZAIC nanoTile 85 µm x 100 mm HSS-T3u wTRAP was applied as a separating device before mass spectrometry (MS). Buffers were made up using UPLC-grade chemicals (Sigma-Aldrich, MO). Buffer A: 0.1% formic acid in water, Buffer B: 0.1% formic acid in acetonitrile. Samples were trapped and washed with a mixture of Buffer A and B, initially using 1% Buffer B, with a gradient of Buffer A/Buffer B occurring over 90 minutes for each sample, to elute specific peptides. Data were acquired with UDMSE with Synapt G2-S UDMS (Waters Corporation, MA) including ion mobility spectroscopy (IMS). The data range was 100–2000 m/z, scan time 1 s, IMS wave velocity 650 ms<sup>-1</sup> and collision energy ramped in trap between 20 and 60 V. Calibration was performed using Glu1-fibrinopeptide B MS2 fragments, and Glu1-fibrinopeptide B precursor ion was used during the acquisitions as a lock mass. 10% of the samples were acquired as triplicates to validate the results, and further analysis was done with Progenesis QI for Proteomics software (Nonlinear Dynamics, Newcastle, UK). The mass spectrometry proteomics data have been deposited into the ProteomeXchange Consortium via the PRIDE partner repository with the data set identifier PXD008445 [25]. [S1 Table](#) explains which data files were used.

### Mass spectrometry data analysis

The data analysis has been described previously in detail [34]. Briefly, Progenesis QI for proteomics software (v3, Nonlinear Dynamics) was used for raw data processing, using default parameters for peak picking and alignment. Peptide identification was run against Uniprot human FASTA sequences (UniprotKB Release 2015\_09, 20205 sequence entries), and label-free protein quantification performed using the Hi-N method (Protein Lynx Global Server) [35], whereby the samples were spiked with 12.5 fmol/µL of CLPB\_ECOLI (P63285, ClpB protein) peptides (Hi3 Escherichia Coli Standard, Waters). The peptide identification parameters were: fixed modification of cysteine (carbamidomethyl) and variable modification of methionine (oxidation). The peptide error tolerance was set to a maximum of 10 ppm, the false discovery rate limited to less than 2% and default values (in Progenesis QI for Proteomics) were used for the rest of the parameters.

## Comparisons

The following comparisons between the samples' protein abundances were performed:

- E6/E7 mRNA+ (n = 13) vs E6/E7 mRNA- (n = 18)  
 p16<sup>INK4a</sup>-/DNA- (n = 16) vs p16<sup>INK4a</sup>+ /DNA+ (n = 11)  
 p16<sup>INK4a</sup>+ /DNA- (n = 5) vs p16<sup>INK4a</sup>- /DNA- (n = 16)  
 p16<sup>INK4a</sup>- /DNA+ (n = 4) vs p16<sup>INK4a</sup>- /DNA- (n = 16)  
 p16<sup>INK4a</sup>- /DNA+ (n = 4) vs p16<sup>INK4a</sup>+ /DNA+ (n = 11)  
 p16<sup>INK4a</sup>+ /DNA- (n = 5) vs p16<sup>INK4a</sup>+ /DNA+ (n = 11)

## Statistical analysis

The data were imported into EZinfo 3.0.3.0 (Release date Dec 02, 2014, Umetrics, Sweden), a statistical package that works with Progenesis QI for proteomics. The quantified proteins in the comparisons were compared using an unpaired t-test on a protein-by-protein basis, with a p-value of 0.05 being the cut-off for statistical significance. The Benjamini-Hochberg multiple testing correction was performed, with an FDR cut-off of 0.1. ROC curves were calculated using SPSS (version 25).

## Results

### Patient demographics and tumor characteristics

Serum from 36 patients with newly diagnosed OPSCC was analyzed. Clinical parameters are shown in [S2 Table](#). All samples that were positive for HPV DNA contained HPV16 DNA, except for one sample that was HPV18 positive; this tumor was p16<sup>INK4a</sup>- but E6/E7 mRNA+.

Representative images of p16+ and p16-, E6/E7 mRNA+ and E6/E7 mRNA- are shown in [Fig 1](#) as well as in [S1 Fig](#). HPV DNA results are shown in [S3 Table](#).

### Proteomics

A total of 174 serum proteins with 2 or more unique peptides were identified and quantified by mass spectrometry. [Table 1](#) shows the significant proteins after multiple testing correction, between HPV-positive and HPV-negative as tested using E6/E7 mRNA and p16<sup>INK4a</sup>/DNA, respectively. The other comparisons, which included samples with discordant p16<sup>INK4a</sup>/DNA

**Table 1. Serum proteins differing between patients with a) E6/E7 mRNA -positive and -negative OPSCCs and b) p16<sup>INK4a</sup>/DNA -positive and -negative OPSCCs.**

E6/E7 mRNA+ vs mRNA- proteins					
Accession	Protein name	p-value	FDR	FC (mRNA-/mRNA+)	ROC
P10643	Complement component C7	0.000677	0.058859	1.37	0.87
Q13790	Apolipoprotein F	0.000542	0.058859	0.41	0.84
p16 <sup>INK4a</sup> +/DNA+ vs p16 <sup>INK4a</sup> -/DNA-					
Accession	Protein name	p-value	FDR	FC (p16 <sup>INK4a</sup> -DNA- / p16 <sup>INK4a</sup> + DNA+)	ROC
P10643	Complement component C7	0.000911	0.079284	1.42	0.86
Q08380	Galactin-3-Binding Protein	0.000879	0.079284	1.37	0.86

P-values were calculated using the unpaired t-test. These proteins remained significant following multiple-testing correction, with FDR set to 0.1. FC—fold change; ROC—receiver operating characteristic; abundance—average normalized abundance.

<https://doi.org/10.1371/journal.pone.0233974.t001>

yielded no significant proteins. [S4 Table](#) details all of the identified proteins for each comparison.

## Discussion

In this study, using nUPLC-UDMS<sup>E</sup>, we aimed to identify serum proteins that could be easily and economically implemented to make a differential diagnosis between HPV-associated and HPV-independent OPSCCs. This approach would improve the detection of patients whose OPSCCs are mis-classified with regards to HPV status by p16<sup>INK4a</sup> testing. For this, we compared the proteomic differences using different proxies for HPV: E6/E7 mRNA, and p16<sup>INK4a</sup> combined with HPV DNA. We identified three proteins (ApoF, Galactin-3-Binding protein (LGALS3BP) and complement component C7 (C7) that differentiate between HPV+ and HPV- tumors. C7 was the only protein that differentiated between HPV+ and HPV- tumors using either p16<sup>INK4a</sup>/DNA or E6/E7 mRNA as HPV detection methods.

ApoF is a sialoglycoprotein, also known as “lipid transfer inhibitor protein”, and is an inhibitor of cholesterol ester transferase that binds to LDL, and regulates cholesterol transport [36]. Lipid metabolic reprogramming is a known hallmark of cancer, and it has been shown that inhibition of cholesterol production with statins in head and neck cancers is associated with improved disease-specific survival [37]. In animal studies, ApoF has also been associated with immune and inflammation responses [38]. Transcription of interferon alpha (IFN $\alpha$ ) responsive genes was shown to be impaired in the ApoF knockout mice and attributable also to hypomorphic expression of Stat2 [39]. These genes and the type I IFN pathway are of major importance also in HPV infections as they modulate the transcription HPV E6 and E7 oncogenes and viral immune evasion [40]. Thus, the increased levels of ApoF in sera might be interpreted as being due to general metabolic reprogramming of malignant invasive cells, or due to reprogramming of the malignant cells and the surrounding stromal reaction by HPV *per se*.

We identified significantly different expression of Complement component C7 using either the E6/E7 mRNA proxy or the p16<sup>INK4a</sup>/DNA proxy ([Table 1](#)). C7 is a critical component of the terminal pathway of complement activation. C7, along with complement components 6, 8, and 9, is involved in membrane attack complex (MAC) formation [41]. Sublytic MACs, which form when the nucleated cells are not unequivocally identified as non-self can activate a few oncogenic pathways such as the mitogen-activated protein kinase (MAPK) family, extracellular regulated protein kinases (ERKs), p38 MAPK, the phosphatidylinositol 3-kinase (PI3K) pathway and Ras [42]. These pathways are of importance in HPV-infected cells in the progression toward malignancy [43].

Thus, similarly as with ApoF, these increased serum concentrations of C7 may result from local upregulation of complement genes in primary tumor sites, mainly by endothelial cells and cells participating in immune responses i.e. the integral component of both the premetastatic tumor niche the tumor microenvironment. Elevated serum levels of C7 have been linked to a better prognosis in ovarian and non-small cell lung cancer (NSCLC). Additionally, overexpression of C7 was shown to inhibit the colony formation of NSCLC cells, which indicates that C7 might be a potential tumor suppressor [44].

LGALS3BP is a secreted glycoprotein that was originally identified during investigations into different malignancies. Furthermore, the expression of LGALS3BP is upregulated in both chronic and acute viral infections, perhaps due to its induction by a variety of molecules that “either mimic or are characteristic for an ongoing inflammation and microbial infection, such as IFN- $\alpha$ , IFN- $\beta$ , IFN- $\gamma$ , TNF- $\alpha$ , poly(I:C), dsRNA, and dsDNA”, as thoroughly reviewed by Loimaranta et al. [45]. As one member of the scavenger receptor cysteine-rich (SRCR) domain-containing protein family, which are implicated in both self-nonsel self discrimination



and innate immunity-related proteins [45]. Thus, it was not unexpected to identify that LGALS3BP sera levels were associated with an HPV associated OPSCC. In the future, it would be important to study whether high sera levels of LGALS3BP could be associated with other viral associated malignancies such as EBV-associated nasopharyngeal carcinomas and lymphomas. Higher serum LGALS3BP levels have been linked to poor prognosis and progression in various cancers [46–49].

The main strength of the study is that, to our knowledge, this the first time that the serum proteomes have been assessed and compared in OPSCC patients with different p16<sup>INK4a</sup> and HPV statuses, to identify possible serum biomarkers that can improve classification of patients who would be misclassified when using p16<sup>INK4a</sup> or HPV DNA testing.

Importantly, all the identified three proteins could be linked also immunity and inflammatory responses also of importance for HPV *per se*. However, it remains to seen whether these proteins are specific enough for HPV or whether they are reflecting also other viral associated malignancies.

Limitations of this study include small group sizes. As a consequence of this, and insufficient smoking data, we were unable to compare the proteomic signatures of those with a <10 year smoking and >10 year smoking history, and different disease stages, which are known to affect the prognosis of these tumors [50, 51].

## Conclusions

We have identified three interesting serum proteins with the potential to improve stratification of OPSCC patients. Serum levels of C7, APOF and LGALS3BP differentiate between HPV-positive and HPV-negative OPSCC. Further research is warranted to confirm these findings on a larger cohort.

## Supporting information

**S1 Table. Details of files used from dataset PXD008445.**  
(XLSX)

**S2 Table. Clinical parameters of the patients.**  
(DOCX)

**S3 Table. HPV DNA PCR results.**  
(XLSX)

**S4 Table. All serum proteins with 2 or more significant proteins, with statistical information.**  
(XLSX)

**S1 Fig. Large versions of the p16-positive, p16-negative, E6/E7 mRNA+, E6/E7 mRNA- and RNAScope -positive and -negative probes, also shown as a small version in Fig 1.**  
(EPS)

## Acknowledgments

We are grateful to the patients for providing their samples for this research.

## Author Contributions

**Conceptualization:** Amy Dickinson, Suvi Silén.

**Data curation:** Timo Carpén.

**Formal analysis:** Amy Dickinson, Mayank Saraswat, Sakari Joenväärä.

**Funding acquisition:** Antti Mäkitie, Suvi Silén.

**Investigation:** Mayank Saraswat, Stina Syrjänen, Tiialotta Tohmola, Reija Randén-Brady, Timo Carpén, Jaana Hagström, Sakari Joenväärä.

**Methodology:** Mayank Saraswat, Sakari Joenväärä.

**Project administration:** Amy Dickinson.

**Resources:** Caj Haglund, Petri Mattila, Antti Mäkitie, Suvi Silén.

**Software:** Robert Silén.

**Supervision:** Caj Haglund, Antti Mäkitie, Suvi Silén.

**Validation:** Mayank Saraswat, Sakari Joenväärä.

**Writing – original draft:** Amy Dickinson, Stina Syrjänen.

**Writing – review & editing:** Amy Dickinson, Mayank Saraswat, Tiialotta Tohmola, Reija Randén-Brady, Timo Carpén, Jaana Hagström, Caj Haglund, Petri Mattila, Antti Mäkitie, Sakari Joenväärä, Suvi Silén.

## References

1. Ndiaye C, Mena M, Alemany L, Arbyn M, Castellsague X, Laporte L, et al. HPV DNA, E6/E7 mRNA, and p16INK4a detection in head and neck cancers: a systematic review and meta-analysis. *Lancet Oncol.* 2014; 15(12):1319–31. [https://doi.org/10.1016/S1470-2045\(14\)70471-1](https://doi.org/10.1016/S1470-2045(14)70471-1) PMID: 25439690
2. Haegglblom L, Ramqvist T, Tommasino M, Dalianis T, Nasman A. Time to change perspectives on HPV in oropharyngeal cancer. A systematic review of HPV prevalence per oropharyngeal sub-site the last 3 years. *Papillomavirus Res.* 2017; 4:1–11. <https://doi.org/10.1016/j.pvr.2017.05.002> PMID: 29179862
3. Carlander AF, Gronhoj Larsen C, Jensen DH, Garnaes E, Kiss K, Andersen L, et al. Continuing rise in oropharyngeal cancer in a high HPV prevalence area: A Danish population-based study from 2011 to 2014. *Eur J Cancer.* 2017; 70:75–82. <https://doi.org/10.1016/j.ejca.2016.10.015> PMID: 27888679
4. Jouhi L, Halme E, Ijala H, Saarihahti K, Koivunen P, Pukkila M, et al. Epidemiological and treatment-related factors contribute to improved outcome of oropharyngeal squamous cell carcinoma in Finland. *Acta Oncol.* 2018; 57(4):541–51. <https://doi.org/10.1080/0284186X.2017.1400688> PMID: 29145765
5. Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez BY, Xiao W, Kim E, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. *J Clin Oncol.* 2011; 29(32):4294–301. <https://doi.org/10.1200/JCO.2011.36.4596> PMID: 21969503
6. Castellsague X, Alemany L, Quer M, Halc G, Quiros B, Tous S, et al. HPV Involvement in Head and Neck Cancers: Comprehensive Assessment of Biomarkers in 3680 Patients. *J Natl Cancer Inst.* 2016; 108(6):djv403. <https://doi.org/10.1093/jnci/djv403> PMID: 26823521
7. Nauta IH, Rietbergen MM, van Bokhoven A, Bloemena E, Lissenberg-Witte BI, Heideman DAM, et al. Evaluation of the eighth TNM classification on p16-positive oropharyngeal squamous cell carcinomas in the Netherlands and the importance of additional HPV DNA testing. *Ann Oncol.* 2018; 29(5):1273–9. <https://doi.org/10.1093/annonc/mdy060> PMID: 29438466
8. Leemans CR, Braakhuis BJ, Brakenhoff RH. The molecular biology of head and neck cancer. *Nat Rev Cancer.* 2011; 11(1):9–22. <https://doi.org/10.1038/nrc2982> PMID: 21160525
9. Sood AJ, McIlwain W, O'Connell B, Nguyen S, Houlton JJ, Day T. The association between T-stage and clinical nodal metastasis in HPV-positive oropharyngeal cancer. *Am J Otolaryngol.* 2014; 35(4):463–8. <https://doi.org/10.1016/j.amjoto.2013.12.008> PMID: 24780201
10. Wang MB, Liu IY, Gornbein JA, Nguyen CT. HPV-Positive Oropharyngeal Carcinoma: A Systematic Review of Treatment and Prognosis. *Otolaryngol Head Neck Surg.* 2015; 153(5):758–69. <https://doi.org/10.1177/0194599815592157> PMID: 26124261
11. Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tan PF, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med.* 2010; 363(1):24–35. <https://doi.org/10.1056/NEJMoa0912217> PMID: 20530316

12. Mirghani H, Blanchard P. Treatment de-escalation for HPV-driven oropharyngeal cancer: Where do we stand? *Clin Transl Radiat Oncol*. 2018; 8:4–11. <https://doi.org/10.1016/j.ctro.2017.10.005> PMID: 29594236
13. Westra WH, Lewis JS Jr. Update from the 4th Edition of the World Health Organization Classification of Head and Neck Tumours: Oropharynx. *Head Neck Pathol*. 2017; 11(1):41–7. <https://doi.org/10.1007/s12105-017-0793-2> PMID: 28247229
14. UICC. (Union for International Cancer Control). TNM classification of malignant tumours. Brierley J GM, Wittekind C, editor: Wiley Blackwell; 2017.
15. Prigge ES, Arbyn M, von Knebel Doeberitz M, Reuschenbach M. Diagnostic accuracy of p16(INK4a) immunohistochemistry in oropharyngeal squamous cell carcinomas: A systematic review and meta-analysis. *Int J Cancer*. 2017; 140(5):1186–98. <https://doi.org/10.1002/ijc.30516> PMID: 27859245
16. Bremner R, Cohen BL, Sopta M, Hamel PA, Ingles CJ, Gallie BL, et al. Direct transcriptional repression by pRB and its reversal by specific cyclins. *Mol Cell Biol*. 1995; 15(6):3256–65. <https://doi.org/10.1128/mcb.15.6.3256> PMID: 7760821
17. Braakhuis BJ, Brakenhoff RH, Meijer CJ, Snijders PJ, Leemans CR. Human papilloma virus in head and neck cancer: the need for a standardised assay to assess the full clinical importance. *Eur J Cancer*. 2009; 45(17):2935–9. <https://doi.org/10.1016/j.ejca.2009.08.018> PMID: 19766476
18. Linge A, Schotz U, Lock S, Lohaus F, von Neubeck C, Gudziol V, et al. Comparison of detection methods for HPV status as a prognostic marker for loco-regional control after radiochemotherapy in patients with HNSCC. *Radiother Oncol*. 2018; 127(1):27–35. <https://doi.org/10.1016/j.radonc.2017.12.007> PMID: 29295747
19. Lechner M, Chakravarthy AR, Walter V, Masterson L, Feber A, Jay A, et al. Frequent HPV-independent p16/INK4A overexpression in head and neck cancer. *Oral Oncol*. 2018; 83:32–7. <https://doi.org/10.1016/j.oraloncology.2018.06.006> PMID: 30098776
20. Schache AG, Liloglou T, Risk JM, Jones TM, Ma XJ, Wang H, et al. Validation of a novel diagnostic standard in HPV-positive oropharyngeal squamous cell carcinoma. *Br J Cancer*. 2013; 108(6):1332–9. <https://doi.org/10.1038/bjc.2013.63> PMID: 23412100
21. Mena M, Taberna M, Tous S, Marquez S, Clavero O, Quiros B, et al. Double positivity for HPV-DNA/p16(ink4a) is the biomarker with strongest diagnostic accuracy and prognostic value for human papillomavirus related oropharyngeal cancer patients. *Oral Oncol*. 2018; 78:137–44. <https://doi.org/10.1016/j.oraloncology.2018.01.010> PMID: 29496041
22. Albers AE, Qian X, Kaufmann AM, Coords A. Meta analysis: HPV and p16 pattern determines survival in patients with HNSCC and identifies potential new biologic subtype. *Sci Rep*. 2017; 7(1):16715. <https://doi.org/10.1038/s41598-017-16918-w> PMID: 29196639
23. Rasmussen JH, Gronhoj C, Hakansson K, Friberg J, Andersen E, Lelkaitis G, et al. Risk profiling based on p16 and HPV DNA more accurately predicts location of disease relapse in patients with oropharyngeal squamous cell carcinoma. *Ann Oncol*. 2019; 30(4):629–36. <https://doi.org/10.1093/annonc/mdz010> PMID: 30657857
24. Tuhkuri A, Saraswat M, Makitie A, Mattila P, Silen R, Dickinson A, et al. Patients with early-stage oropharyngeal cancer can be identified with label-free serum proteomics. *Br J Cancer*. 2018; 119(2):200–12. <https://doi.org/10.1038/s41416-018-0162-2> PMID: 29961760
25. Vizcaino JA, Csordas A, Del-Toro N, Dianas JA, Griss J, Lavidas I, et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res*. 2016; 44(22):11033. <https://doi.org/10.1093/nar/gkw880> PMID: 27683222
26. Randen-Brady R, Carpen T, Jouhi L, Syrjanen S, Haglund C, Tarkkanen J, et al. In situ hybridization for high-risk HPV E6/E7 mRNA is a superior method for detecting transcriptionally active HPV in oropharyngeal cancer. *Hum Pathol*. 2019; 90:97–105. <https://doi.org/10.1016/j.humpath.2019.05.006> PMID: 31121191
27. Jouhi L, Datta N, Renkonen S, Atula T, Makitie A, Haglund C, et al. Expression of toll-like receptors in HPV-positive and HPV-negative oropharyngeal squamous cell carcinoma—an in vivo and in vitro study. *Tumour Biol*. 2015; 36(10):7755–64. <https://doi.org/10.1007/s13277-015-3494-z> PMID: 25941114
28. Fakhry C, Lacchetti C, Rooper LM, Jordan RC, Rischin D, Sturgis EM, et al. Human Papillomavirus Testing in Head and Neck Carcinomas: ASCO Clinical Practice Guideline Endorsement of the College of American Pathologists Guideline. *J Clin Oncol*. 2018; 36(31):3152–61. <https://doi.org/10.1200/JCO.18.00684> PMID: 30188786
29. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res*. 1988; 16(3):1215. <https://doi.org/10.1093/nar/16.3.1215> PMID: 3344216
30. Carpen T, Sjoblom A, Lundberg M, Haglund C, Markkola A, Syrjanen S, et al. Presenting symptoms and clinical findings in HPV-positive and HPV-negative oropharyngeal cancer patients. *Acta Otolaryngol*. 2018; 138(5):513–8. <https://doi.org/10.1080/00016489.2017.1405279> PMID: 29161981

31. Ruuskanen M, Irjala H, Minn H, Vahlberg T, Randen-Brady R, Hagstrom J, et al. Epstein-Barr virus and human papillomaviruses as favorable prognostic factors in nasopharyngeal carcinoma: A nationwide study in Finland. *Head Neck*. 2019; 41(2):349–57. <https://doi.org/10.1002/hed.25450> PMID: 30549170
32. Kuusela P, Saraswat M, Joenvaara S, Kaartinen J, Jarvinen A, Renkonen R. Changes in plasma protein levels as an early indication of a bloodstream infection. *PLoS One*. 2017; 12(2):e0172987. <https://doi.org/10.1371/journal.pone.0172987> PMID: 28235076
33. Distler U, Kuharev J, Navarro P, Levin Y, Schild H, Tenzer S. Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nat Methods*. 2014; 11(2):167–70. <https://doi.org/10.1038/nmeth.2767> PMID: 24336358
34. Saraswat M, Joenvaara S, Jain T, Tomar AK, Sinha A, Singh S, et al. Human Spermatozoa Quantitative Proteomic Signature Classifies Normo- and Asthenozoospermia. *Mol Cell Proteomics*. 2017; 16(1):57–72. <https://doi.org/10.1074/mcp.M116.061028> PMID: 27895139
35. Silva JC, Gorenstein MV, Li GZ, Vissers JP, Geromanos SJ. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Mol Cell Proteomics*. 2006; 5(1):144–56. <https://doi.org/10.1074/mcp.M500230-MCP200> PMID: 16219938
36. Wang X, Driscoll DM, Morton RE. Molecular cloning and expression of lipid transfer inhibitor protein reveals its identity with apolipoprotein F. *J Biol Chem*. 1999; 274(3):1814–20. <https://doi.org/10.1074/jbc.274.3.1814> PMID: 9880564
37. Gupta A, Stokes W, Eguchi M, Hararah M, Amini A, Mueller A, et al. Statin use associated with improved overall and cancer specific survival in patients with head and neck cancer. *Oral Oncol*. 2019; 90:54–66. <https://doi.org/10.1016/j.oraloncology.2019.01.019> PMID: 30846177
38. Yang Y, Fu Q, Zhou T, Li Y, Liu S, Zeng Q, et al. Analysis of apolipoprotein genes and their involvement in disease response of channel catfish after bacterial infection. *Dev Comp Immunol*. 2017; 67:464–70. <https://doi.org/10.1016/j.dci.2016.09.007> PMID: 27640368
39. Peloso GM, Demissie S, Collins D, Mirel DB, Gabriel SB, Cupples LA, et al. Common genetic variation in multiple metabolic pathways influences susceptibility to low HDL-cholesterol and coronary heart disease. *J Lipid Res*. 2010; 51(12):3524–32. <https://doi.org/10.1194/jlr.P008268> PMID: 20855565
40. Stanley MA, Pett MR, Coleman N. HPV: from infection to cancer. *Biochem Soc Trans*. 2007; 35(Pt 6):1456–60. <https://doi.org/10.1042/BST0351456> PMID: 18031245
41. Wurzner R. Modulation of complement membrane attack by local C7 synthesis. *Clin Exp Immunol*. 2000; 121(1):8–10. <https://doi.org/10.1046/j.1365-2249.2000.01263.x> PMID: 10886232
42. Bossi F, Rizzi L, Bulla R, Debeus A, Tripodo C, Picotti P, et al. C7 is expressed on endothelial cells as a trap for the assembling terminal complement complex and may exert anti-inflammatory function. *Blood*. 2009; 113(15):3640–8. <https://doi.org/10.1182/blood-2008-03-146472> PMID: 19179470
43. Ruutu M, Peitsaro P, Johansson B, Syrjanen S. Transcriptional profiling of a human papillomavirus 33-positive squamous epithelial cell line which acquired a selective growth advantage after viral integration. *Int J Cancer*. 2002; 100(3):318–26. <https://doi.org/10.1002/ijc.10455> PMID: 12115547
44. Ying L, Zhang F, Pan X, Chen K, Zhang N, Jin J, et al. Complement component 7 (C7), a potential tumor suppressor, is correlated with tumor progression and prognosis. *Oncotarget*. 2016; 7(52):86536–46. <https://doi.org/10.18632/oncotarget.13294> PMID: 27852032
45. Loimaranta V, Hepojoki J, Laaksoaho O, Pulliainen AT. Galectin-3-binding protein: A multitask glycoprotein with innate immunity functions in viral and bacterial infections. *J Leukoc Biol*. 2018; 104(4):777–86. <https://doi.org/10.1002/JLB.3VMR0118-036R> PMID: 29882603
46. Marchetti A, Tinari N, Buttitta F, Chella A, Angeletti CA, Sacco R, et al. Expression of 90K (Mac-2 BP) correlates with distant metastasis and predicts survival in stage I non-small cell lung cancer patients. *Cancer Res*. 2002; 62(9):2535–9. PMID: 11980646
47. Woodman N, Pinder SE, Tajadura V, Le Bourhis X, Gillett C, Delannoy P, et al. Two E-selectin ligands, BST-2 and LGALS3BP, predict metastasis and poor survival of ER-negative breast cancer. *Int J Oncol*. 2016; 49(1):265–75. <https://doi.org/10.3892/ijo.2016.3521> PMID: 27176937
48. Zeimet AG, Natoli C, Herold M, Fuchs D, Windbichler G, Daxenbichler G, et al. Circulating immunostimulatory protein 90K and soluble interleukin-2-receptor in human ovarian cancer. *Int J Cancer*. 1996; 68(1):34–8. [https://doi.org/10.1002/\(SICI\)1097-0215\(19960927\)68:1<34::AID-IJC7>3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1097-0215(19960927)68:1<34::AID-IJC7>3.0.CO;2-Y) PMID: 8895537
49. Wu KL, Chen HH, Pen CT, Yeh WL, Huang EY, Hsiao CC, et al. Circulating Galectin-1 and 90K/Mac-2BP Correlated with the Tumor Stages of Patients with Colorectal Cancer. *Biomed Res Int*. 2015; 2015:306964. <https://doi.org/10.1155/2015/306964> PMID: 26448934
50. Liskamp CP, Janssens GO, Bussink J, Melchers WJ, Kaanders JH, Verhoef CG. Adverse effect of smoking on prognosis in human papillomavirus-associated oropharyngeal carcinoma. *Head Neck*. 2016; 38(12):1780–7. <https://doi.org/10.1002/hed.24516> PMID: 27248701

51. Jouhi L, Hagstrom J, Atula T, Makitie A. Is p16 an adequate surrogate for human papillomavirus status determination? *Curr Opin Otolaryngol Head Neck Surg.* 2017; 25(2):108–12. <https://doi.org/10.1097/MOO.0000000000000341> PMID: 28141601