Quantitative proteomics identifies biomarkers to distinguish pulmonary from head and neck squamous cell carcinomas by immunohistochemistry

Annika Richter1†*, Alexander Fichtner1†, Jasmin Joost1†, Philipp Brockmeyer2, Philipp Kauffmann2, Henning Schliephake2, Alexander Hammerstein-Equord3, Stefan Kueffer1, Henning Urlaub4,5, Thomas Oellerich6,7, Philipp Ströbel1, Hanibal Bohnenberger1† and Felix Bremmer1

1Institute of Pathology, University Medical Centre Göttingen, Göttingen, Germany
2Department of Oral and Maxillofacial Surgery, University Medical Centre Göttingen, Göttingen, Germany
3Department of Thoracic and Cardiovascular Surgery, University Medical Centre Göttingen, Göttingen, Germany
4Bioanalytical Mass Spectrometry Group, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany
5Bioanalytics, Institute for Clinical Chemistry, University Medical Centre Göttingen, Göttingen, Germany
6Department of Medicine II, Haematology/Oncology, Goethe University, Frankfurt, Germany
7German Cancer Research Centre and German Cancer Consortium, Heidelberg, Germany

*Correspondence to: Annika Richter, Institute of Pathology, University Medical Centre, Robert-Koch-Str. 40, 37075 Göttingen, Germany. E-mail: annika.richter1@med.uni-goettingen.de
†These authors contributed equally to this study.

Abstract

The differentiation between a pulmonary metastasis and a newly developed squamous cell carcinoma of the lung in patients with prior head and neck squamous cell carcinoma (HNSCC) is difficult due to a lack of biomarkers but is crucially important for the prognosis and therapy of the affected patient. By using high-resolution mass spectrometry in combination with stable isotope labelling by amino acids in cell culture, we identified 379 proteins that are differentially expressed in squamous cell carcinomas of the lung and the head and neck. Of those, CAV1, CAV2, LGALS1, LGALS7, CK19, and UGDH were tested by immunohistochemistry on 194 tissue samples (98 lung and 96 HNSCCs). The combination of CAV1 and LGALS7 was able to distinguish the origin of the squamous cell carcinoma with high accuracy (area under the curve 0.876). This biomarker panel was tested on a cohort of 12 clinically classified lung tumours of unknown origin after HNSCC. Nine of those tumours were immunohistochemically classifiable.

Keywords: SILAC-mass spectrometry; squamous cell carcinoma; head and neck; lung; metastasis; cell culture

Received 23 April 2021; Revised 23 July 2021; Accepted 1 September 2021

Introduction

Head and neck squamous cell carcinomas (HNSCCs) comprise a group of neoplasms developing in the nose, oral cavity, pharynx, and larynx [1] and make up more than 90% of neoplasms of the head and neck [2]. The long-term survival of patients with HNSCC treated with curative intent has not improved significantly over recent decades [2] due to lymph node metastasis, local tumour recurrence, and mainly because of distant metastasis (met-HNSCC), for example to the lung [3]. The differentiation between a pulmonary metastasis of a primary known HNSCC and a secondary primary tumour (SPT) of the lung after a prior HNSCC is difficult. More than two-thirds of SPTs are squamous cell carcinomas [4,5] and both tumours have similar patterns of genetic mutations and pathological molecular alterations [5–8]. No biomarker has yet been established for their differential diagnosis. Therefore, the differentiation is mainly based on radiological and clinical criteria [5,9,10].

However, the distinction between met-HNSCC and SPT is crucially important for further prognosis and the decision between a curative and palliative treatment...
We therefore aimed to identify, by high-resolution mass spectrometry (MS)-based proteomics, differentially expressed proteins that can be used as immunohistochemical markers to distinguish between squamous cell lung carcinoma (SQCLC) and HNSCC.

Methods
Proteomic analysis of squamous cell carcinoma cell lines

The human tumour cell line NCI-H2170 (CRL-5928; American Type Culture Collection, Manassas, VA, USA) originates from a squamous cell carcinoma of the lung. The PCI-13.1 cell line (Department of Pathology, UPMC, PA, USA) derives from HPV-negative squamous cell carcinoma of the oropharynx. Labelling with stable isotopes in cell culture and proteomic comparison were conducted as described previously [8,11,12]. In brief, NCI-H2170 and PCI3.1 cells were cultured in RPMI 1640 medium supplemented with 10% dialysed foetal calf serum (Invitrogen, Waltham, MA, USA), 4 mM glutamine, antibiotics, 0.115 mM L-arginine-13C6 and 0.275 mM L-lysine-4,4,5,5-D4 (Eurisotop, St-Aubin, France), or equimolar levels of the corresponding non-labelled (light) amino acids (Merck Millipore, Burlington, MA, USA) for at least 10 cell cycles. Labelled cells were lysed in 0.5% Nonidet P-40 buffer containing 50 mM Tris/HCl, pH 7.8, 150 mM NaCl, 1 mM Na3VO4, 1 mM NaF, 0.2% lauryl maltoside, and protease inhibitors (Complete, Roche, Basel, Switzerland). Equal amounts of proteins of light-labelled NCI-H2170 were mixed with heavy-labelled PCI-13.1 and vice versa to obtain two biological replicates with different stable isotope labelling by amino acids in cell culture (SILAC). Afterwards, they were separated by 1D-PAGE (4–12% NuPAGE Bis-Tris Gel; Invitrogen) and stained with Coomassie brilliant blue. Next, the stained gel was separated in 23 slices and each one was reduced with 10 mM DTT for 55 min at 56°C, alkylated with 55 mM iodoacetamide (IAA) for 20 min at 26°C, and digested with modified trypsin (Promega, Madison, WI, USA) overnight at 37°C. Resulting peptides were separated by a C18 precolumn (2.5 cm, 360 µm o.d., 100 µm i.d., Reprosil-Pur 120 Å, 5 µm, C18-AQ; Dr. Maisch GmbH, Ammerbuch, Germany) at a flow rate of 10 µl/min and a C18 capillary column (20 cm, 360 µm o.d., 75 µm i.d., Reprosil-Pur 120 Å, 3 µm, C18-AQ; Dr. Maisch GmbH) at a flow rate of 300 nl/min, with a gradient of acetonitrile ranging from 5 to 35% in 0.1% formic acid for 90 min using an Proxeon nano LC coupled to an Q Exactive mass spectrometer (Thermo Fisher Scientific).
Electron, Waltham, MA, USA). MS conditions were as follows: spray voltage, 1.8 kV; heated capillary temperature, 270 °C; and normalised collision energy, 28. The mass spectrometer automatically switched between MS and MS/MS acquisitions (data-dependent mode). Survey MS spectra were acquired

Table 2. Results of the marker candidates in the HPA and cohort 1.

<table>
<thead>
<tr>
<th>Marker</th>
<th>SILAC-MS (significantly stronger in)</th>
<th>IHC (HPA) (n) positive</th>
<th>IHC (cohort 1) (n) positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAV1</td>
<td>PCI-13.1 (HNSCC)</td>
<td>3/3 HNSCC</td>
<td>6/6 HNSCC</td>
</tr>
<tr>
<td>CAV2</td>
<td>PCI-13.1 (HNSCC)</td>
<td>3/3 HNSCC</td>
<td>6/6 HNSCC</td>
</tr>
<tr>
<td>LGALS1</td>
<td>PCI-13.1 (HNSCC)</td>
<td>3/4 HNSCC</td>
<td>5/6 HNSCC</td>
</tr>
<tr>
<td>LGALS7</td>
<td>PCI-13.1 (HNSCC)</td>
<td>3/3 HNSCC</td>
<td>6/6 HNSCC</td>
</tr>
<tr>
<td>CK19</td>
<td>H2170 (SQCLC)</td>
<td>1/4 SQCLC</td>
<td>2/6 HNSCC</td>
</tr>
<tr>
<td>UGDH</td>
<td>H2170 (SQCLC)</td>
<td>1/3 SQCLC</td>
<td>2/6 HNSCC</td>
</tr>
</tbody>
</table>

The table shows the results of the SILAC-MS as well as those of the IHC in the HPA and cohort 1. The absolute number of positive cases and the total number of cases are given for each case. The proteins that were upregulated in HNSCC are highlighted in blue and the proteins that were downregulated are highlighted in grey.

HPA, Human Protein Atlas; IHC, immunohistochemistry; SILAC-MS, SILAC-based mass spectrometry.

Table 3. Clinical and pathological data.

<table>
<thead>
<tr>
<th></th>
<th>SQCLC (n = 98)</th>
<th>HNSCC (n = 98)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>65.5 ± 8.4</td>
<td>62.1 ± 10.4</td>
</tr>
<tr>
<td>Age range</td>
<td>43–81</td>
<td>24–83</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>86 (87.8%)</td>
<td>71 (74.0%)</td>
</tr>
<tr>
<td>Female</td>
<td>12 (12.2%)</td>
<td>25 (26.0%)</td>
</tr>
<tr>
<td><strong>Localisation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oral cavity</td>
<td>–</td>
<td>48 (50.0%)</td>
</tr>
<tr>
<td>Pharynx</td>
<td>–</td>
<td>25 (26.0%)</td>
</tr>
<tr>
<td>Larynx</td>
<td>–</td>
<td>23 (24.0%)</td>
</tr>
<tr>
<td>Lung</td>
<td>98 (100%)</td>
<td>–</td>
</tr>
<tr>
<td><strong>pT stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pT1</td>
<td>19 (19.4%)</td>
<td>22 (22.9%)</td>
</tr>
<tr>
<td>pT2</td>
<td>63 (64.3%)</td>
<td>41 (42.7%)</td>
</tr>
<tr>
<td>pT3</td>
<td>12 (12.2%)</td>
<td>19 (19.8%)</td>
</tr>
<tr>
<td>pT4</td>
<td>4 (4.1%)</td>
<td>14 (14.6%)</td>
</tr>
<tr>
<td><strong>pN stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pN0</td>
<td>59 (60.2%)</td>
<td>52 (54.2%)</td>
</tr>
<tr>
<td>pN1</td>
<td>22 (22.4%)</td>
<td>23 (24.0%)</td>
</tr>
<tr>
<td>pN2</td>
<td>17 (17.3%)</td>
<td>21 (21.9%)</td>
</tr>
<tr>
<td><strong>pM stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pM0</td>
<td>98 (100%)</td>
<td>95 (99.0%)</td>
</tr>
<tr>
<td>Uncertain pM1</td>
<td>0 (0%)</td>
<td>1 (1.0%)</td>
</tr>
<tr>
<td><strong>UICC stage (eighth edition)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>33 (33.7%)</td>
<td>16 (16.7%)</td>
</tr>
<tr>
<td>II</td>
<td>43 (43.9%)</td>
<td>26 (27.1%)</td>
</tr>
<tr>
<td>III</td>
<td>22 (22.4%)</td>
<td>25 (26.0%)</td>
</tr>
<tr>
<td>IV</td>
<td>0 (0%)</td>
<td>29 (30.2%)</td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>G2</td>
<td>70 (70.7%)</td>
<td>87 (90.6%)</td>
</tr>
<tr>
<td>G3</td>
<td>28 (29.3%)</td>
<td>9 (9.4%)</td>
</tr>
<tr>
<td><strong>p16 positive</strong></td>
<td>–</td>
<td>12 (12.5%)</td>
</tr>
</tbody>
</table>

Figure 2. Legend on next page.
in the Orbitrap (m/z 350–1,600) with the resolution set to 70,000 at m/z 200. The 15 most intense ions were sequentially isolated for higher-energy collisional dissociation (HCD) MS/MS fragmentation and detection. Raw data were analysed with MaxQuant (version 1.3.0.5) using Uniprot human as a sequence database. Up to two missed cleavages of trypsin were allowed. Oxidised methionine was searched as variable modification and cysteine carbamidomethylation as fixed modification. The modifications corresponding to arginine and lysine labelled with heavy stable isotopes were handled as fixed modifications. The false positive rate was set to 1% at the peptide level, the false discovery rate was set to 1% at the protein level, and the minimum required peptide length was set to six amino acids.

Resulting data from MaxQuant analysis of the raw data were further analysed by Perseus (version 1.5.2.6; Max Planck Institute for Biochemistry, Martinsried, Germany). Ratios of intensity of heavy- and light-labelled proteins corresponding to either of the two biological replicates were calculated. Intensity values were log\textsubscript{10}-transformed. Next, an outlier significance score for SILAC ratios depending on intensity values (significance B in Perseus, see citation for more details [13]) for every protein was calculated and fold discovery rate was corrected by Benjamini–Hochberg correction. The complete results of the statistical analysis with Perseus including the number of unique peptides can be found in supplementary material, Table S1.

Tissue samples

The patient samples were collected at the University Medical Centre in Göttingen, Germany. In total, 98 SQCLC, 96 HNSCC, and 12 lung tumours with squamous cell carcinoma histology of unknown origin after primary HNSCC were included in this study. The samples of the main cohort (98 SQCLC, 96 HNSCC) derived from oncological resections. Approval for using the human patient material in this study was obtained from the Ethics Committee of the University Medical Centre Göttingen (vote no. 07/06/09, updated in April 2018). All procedures were conducted in accordance with the Declaration of Helsinki and institutional, state, and federal guidelines.

Immunohistochemistry

Immunohistochemical reactions were performed on 2-μm formalin-fixed and paraffin-embedded tissue sections, as described previously [14]. Antigen retrieval was carried out at 97 °C in citrate buffer (pH 6) or EDTA buffer (pH 9). The antibodies and dilutions used are listed in supplementary material, Table S2. The sections were incubated with a ready-to-use horseradish peroxidase-labelled secondary antibody at room temperature for 25 min (anti-rabbit/mouse, produced in goat; REAL EnVision Detection System; Dako, Agilent Technologies, Waldbronn, Germany). The substrate DAB + Chromogen system produces a brown end product, and is applied to visualise the target antigen (REAL DAB + Chromogen; Dako, Agilent Technologies). Tissue samples were counterstained with Mayer’s haematoxylin (Dako, Agilent Technologies) for 8 min and analysed using light microscopy.

Two independent investigators evaluated all stained tissue sections by using an immunoreactivity staining score (IRS) as described previously [12]. The percentage of positively stained cells was first classified using a 0–4 scoring system: score 0 = 0% positive cells, score 1 = less than 10% positive cells, score 2 = 10–50% positive cells, score 3 = 51–80% positive cells, and score 4 = >80% positive cells. The intensity of staining was evaluated on a four-tiered scale (0 = negative, 1 = weak, 2 = intermediate, and 3 = strong). Afterwards, the scores of intensity and staining were multiplied and the mean value per patient was calculated, where 0–1 point was interpreted as negative, 2–3 as weakly positive, 4–6 as moderately positive, and 8–12 points as strongly positive. Therefore, ‘immunohistochemically positive’ tumours

Figure 2. Staining of CAV1 and CAV2 in SQCLC and HNSCC. (A) The SQCLC tumour cells show no or only a slight expression of CAV1 (total magnification ×100). (B) The tumour cells of HNSCC show cytoplasmic and strong membranous expression of CAV1 (total magnification ×50). (C) Box plot for CAV1 expression in SQCLC and HNSCC; the horizontal lines within the boxes represent the median IRS values, ***p < 0.001. (D) CAV1 had an AUC value of 0.781, a sensitivity of 89.9%, and a specificity of 66.3%. (E) The SQCLC tumour cells show no or only slight expression of CAV2 (total magnification ×100). (F) Box plot for CAV2 expression in SQCLC and HNSCC; the horizontal lines within the boxes represent the median IRS values, ***p < 0.001. (G) The tumour cells of HNSCC show cytoplasmic and strong membranous expression of CAV2 (total magnification ×50). (H) CAV2 had an AUC value of 0.613, a sensitivity of 96.8%, and a specificity of 25.8%.

© 2021 The Authors. The Journal of Pathology: Clinical Research published by The Pathological Society of Great Britain and Ireland and John Wiley & Sons, Ltd.

Figure 3. Legend on next page.
have an IRS ≥ 2. Examples of the various IRSs are shown in supplementary material, Figure S1.

Statistical analysis
Differences of the obtained IRS between the different subtypes of squamous cell carcinomas were statistically evaluated using the Mann–Whitney U-test (GraphPad Software, San Diego, CA, USA). A P value of <0.05 was considered significant. Using receiver operating characteristic (ROC) curve analysis, we were able to determine the area under the curve (AUC) value and sensitivity and specificity of the antibodies. The cut-off of the ROC analyses was set at IRS ≥ 2 as this value was considered immunohistochemically positive.

Results
Quantitative proteomic comparison of pulmonary and HNSCC cell lines
The differential diagnosis of metachronous primary and metastatic squamous cell carcinoma in patients with HNSCC is crucial for clinical treatment decisions. However, diagnostic biomarkers are still not established in routine diagnostic pathology, partially due to overlapping genetic and morphological features. We therefore aimed to identify new proteomic biomarkers that can be translated to immunohistochemical support for the pathological diagnosis. We therefore compared the protein expression profile of the two human cell lines NCI-H2170 and PCI-13.1. NCI-H2170 originates from a squamous cell carcinoma of the lung and PCI-13.1 from an HPV-negative squamous cell carcinoma from the oropharynx. By using high-resolution MS in combination with SILAC, we were able to quantify 4,606 proteins and 379 of those were significantly differently expressed (Figure 1 and supplementary material, Tables S1 and S3).

We compared those proteins with the expression data of the Human Protein Atlas and chose six markers (two upregulated in the lung cancer cell line H2170 and four upregulated in the HNSCC cell line PCI-13.1) for immunohistochemical validation (Table 1). Staining for these six markers was examined in a small training cohort of HNSCC (n = 6) and SQCLC (n = 6) (Table 2).

Patient collection for immunohistochemical validation
The clinical and pathological data of 98 patients with SQCLC and 96 patients with HNSCC were evaluated. Of the 96 HNSCC examined, 12 tumours were p16 positive (12.5%); of these, six tumours were located in the oropharynx and six in the pharynx or larynx. In contrast, 84 tumours were p16 negative (87.5%). The SQCLCs were UICC stage I–III and the HNSCCs ranged from stage I to IV (Table 3). All patients were treated by surgery and none of the patients had received neoadjuvant therapy or primary chemoradiotherapy. Squamous cell histology of all samples was confirmed by expert pathological review (FB, PS, and HB).

Distinction of SQCLC and HNSCC by immunohistochemical biomarkers
We immunohistochemically stained the described collection of squamous cell carcinomas with antibodies specific for CAV1, CAV2, LGALS1, and LGALS7 (upregulated in HNSCC) as well as CK19 and UDP-glucose-6-dehydrogenase (UGDH) (upregulated in SQCLC) in order to test their ability to distinguish SQCLC and HNSCC.

Analysis of proteins upregulated in HNSCC
Staining of CAV1 in SQCLC and HNSCC
Tumour cells assessed as positive showed cytoplasmic and often membranous staining. For SQCLC, 33.7% (31/92) of tumours were positive (Figure 2A), HNSCC presented significantly higher expression of CAV1 (p < 0.001), with 89.9% (80/89) of the tumours assessed as immunohistochemically positive (Figure 2B). Median IRSs of 0 for SQCLC and 6 for HNSCC were observed.
Figure 4. Legend on next page.
(Figure 2C). In our analysis, CAV1 showed an AUC value of 0.781, a sensitivity of 89.9%, a specificity of 66.3%, a positive predictive value of 72.1%, and a negative predictive value of 87.1% (Figure 2D). The immunohistochemical data of CAV1 depending on the subsites of HNSCC are listed in supplementary material, File S1.

Staining of CAV2 in SQCLC and HNSCC
Tumour cells assessed as positive showed cytoplasmic and often membranous staining. For SQCLC, 74.2% (72/97) of tumours were positive (Figure 2E). HNSCC presented significantly higher expression of CAV2 (p < 0.001), with 96.8% (90/93) of the tumours assessed as immunohistochemically positive (Figure 2F). Median IRSs of 3 for SQCLC and 6 for HNSCC were observed (Figure 2G). CAV2 showed an AUC value of 0.613, a sensitivity of 96.8%, a specificity of 25.8%, a positive predictive value of 55.6%, and a negative predictive value of 89.3% (Figure 2H).

The immunohistochemical data of CAV2 depending on the subsites of HNSCC are listed in supplementary material, File S1.

Expression of LGALS1 in SQCLC and HNSCC
Tumour cells assessed as positive showed cytoplasmic and often nuclear staining. For SQCLC, 64.6% (62/96) of tumours were positive (Figure 3A). HNSCC presented significantly higher expression of LGALS1 (p < 0.001), with 92.1% (82/89) of the tumours assessed as immunohistochemically positive (Figure 3B). Median IRSs of 2 for SQCLC and 6 for HNSCC were observed (Figure 3C). LGALS1 showed an AUC value of 0.638, a sensitivity of 92.1%, a specificity of 35.4%, a positive predictive value of 56.9%, and a negative predictive value of 86.7% (Figure 3D).

The immunohistochemical data of LGALS1 depending on the subsites of HNSCC are listed in supplementary material, File S1.

Expression of LGALS7 in SQCLC and HNSCC
Tumour cells assessed as positive showed mainly cytoplasmic and focally nuclear staining. For SQCLC, 34% (33/97) of tumours were positive (Figure 3E). HNSCC presented significantly higher expression of LGALS7 (p < 0.001), with 94.5% (86/91) of the tumours assessed as immunohistochemically positive (Figure 3F). Median IRSs of 0 for SQCLC and 6 for HNSCC were observed (Figure 3G). LGALS7 showed an AUC value of 0.802, a sensitivity of 94.5%, a specificity of 66.0%, a positive predictive value of 72.3%, and a negative predictive value of 92.3% (Figure 3H).

The immunohistochemical data of LGALS7 depending on the subsites of HNSCC are listed in supplementary material, File S1.

Analysis of proteins upregulated in SQCLC
Staining of CK19 in SQCLC and HNSCC
Tumour cells assessed as positive showed mainly cytoplasmic and membranous staining. For SQCLC, 97.9% (96/98) of tumours were positive (Figure 4A). HNSCC presented significantly lower expression of CK19 (p < 0.001), with 69.5% (64/92) of the tumours assessed as immunohistochemically positive (Figure 4B). Median IRSs of 8 for SQCLC and 3 for HNSCC were observed (Figure 4C). CK19 showed an AUC value of 0.624, a sensitivity of 98.0%, a specificity of 30.4%, a positive predictive value of 60.0%, and a negative predictive value of 93.3% (Figure 4D).

Staining of UGDH in SQCLC and HNSCC
Tumour cells assessed as positive showed mainly cytoplasmic and nuclear staining. For SQCLC, 78.1% (75/96) of tumours were positive (Figure 4E). HNSCC presented a significant difference in expression of UGDH (p = 0.0131), while showing a nearly similar immunohistochemical positivity of 76.8% (73/95) (Figure 4F). This can be explained by a median IRS of 6 for SQCLC and 3 for HNSCC (Figure 4G). UGDH showed an AUC value of 0.506, a sensitivity of 78.1%, a specificity of 23.2%, a positive predictive value of 50.7%, and a negative predictive value of 51.2% (Figure 4H).

Expression of combination of CAV1 and LGALS7 in SQCLC and HNSCC
With an AUC value between 0.613 and 0.802 (Figures 4D,H and 3D,H), we observed that single
Figure 5. Legend on next page.
immunohistochemical marker was able to distinguish between HNSCC and SQCLC, but the specificity in particular was too low for routine diagnostic use. Therefore, we tested whether a combination of the two most promising markers, CAV1 and LGALS7, was able to further increase the diagnostic accuracy. If only one antibody had to be positive to diagnose an HNSCC, the sensitivity increased to 98.7%. The specificity decreased to 47.3%. The positive predictive value was 61.0% and the negative predictive value 97.7%. The new AUC value was 0.730 (Figure 5A). If both antibodies showed a positive reaction, the sensitivity was 87.2%, specificity improved to 87.9%, the positive predictive value was 86.1%, the negative predictive value 88.9%, and the AUC value was 0.876 (Figure 5B).

Immunohistochemical examination of lung tumours of unknown origin

Based on our findings mentioned above, we used the antibody panel consisting of CAV1 and LGALS7 to examine 12 lung tumours from patients diagnosed previously with an HNSCC (Table 4). Both markers had to be assessed positive (IRS ≥ 2) in order to diagnose an HNSCC. Of the 12 carcinomas, 7 showed expression of both markers and were therefore classified as met-HNSCC (Figure 5C,D). Two tumours were negative and were classified as SQCLC (Figure 5E,F). Three cases (Tu-2, Tu-5, and Tu-10) were positive for CAV1 and negative for LGALS7 (Figure 5G,H). These were assessed as tumours of uncertain origin. Further clinical information can be found in Supplementary Table S4.

Discussion

Patients with HNSCC can develop both pulmonary metastases and metachronous primary squamous cell carcinoma of the lung in the course of their disease due to the similar aetiology and risk factors. The distinction between these two entities is a major diagnostic challenge, because the morphology and the genetic alterations are very similar and there are no reliable biomarkers known yet. However, the correct diagnosis is important for the prognosis and therapy of the patient. Until now, the distinction of those cases was mainly based on clinical and radiological criteria [6,7,10,15]. Therefore, it would be a great diagnostic gain to discover biomarkers that can differentiate between both entities with a high sensitivity and specificity.

To detect new markers, we analysed cell cultures of HNSCC and SQCLC with SILAC-based MS and tested the chosen proteins immunohistochemically on a cohort of HNSCC and SQCLC. These tests were performed using primary tumours, as the differentiation between met-HNSCC and SQCLC is still a problem and we have no gold standard or ground truth. A limitation of this method is tumour evolution between primary and metastasis and therefore a risk of antigen loss. This has to be further analysed by staining more lung tumours in future studies.

Caveolin-1 (CAV1) belongs to a group of caveolins and is an integral membrane protein. As part of the caveolin scaffolding domain, which are cytoskeletal associated proteins, it links cell adhesion molecules and signalling molecules. Therefore, it participates in multiple processes of malignant tumour cells including signal transduction, cell transformation, cell migration, and metastasis [16–19]. Different studies were able to show high expression of CAV1 in HNSCC, squamous cell carcinoma of the oesophagus and the uterine cervix as well as adenocarcinoma of the prostate. In contrast, carcinomas of the lung, pancreas, ovaries, and breast showed low expression [20–25]. With SILAC-based MS as well as immunohistochemistry, significantly higher expression of CAV1 was identified in HNSCC than in SQCLC. The study of Vachani et al., who compared the gene expression of 18 HNSCC and 10 SQCLC, showed similar results. The dominant immunohistochemical expression pattern in our series showed strong staining in the periphery and absent or weak staining in the middle of the tumour. This
Table 4. Immunohistochemistry of lung tumours of unknown origin in patients with prior HNSCC.

<table>
<thead>
<tr>
<th>Case</th>
<th>Age (years)</th>
<th>Gender</th>
<th>Location</th>
<th>Recurrence</th>
<th>Time interval (years)</th>
<th>Pulmonary foci (n)</th>
<th>Clinical classification</th>
<th>IHC classification</th>
<th>LGALS7</th>
<th>CAV1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tu-01</td>
<td>61</td>
<td>M</td>
<td>Larynx</td>
<td>Yes</td>
<td>2</td>
<td>1</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>Tu-02</td>
<td>76</td>
<td>M</td>
<td>Larynx</td>
<td>Yes</td>
<td>1</td>
<td>4</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>Tu-03</td>
<td>62</td>
<td>M</td>
<td>Oral cavity</td>
<td>Yes</td>
<td>2</td>
<td>3</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>Tu-04</td>
<td>73</td>
<td>M</td>
<td>Hypopharynx</td>
<td>No</td>
<td>2</td>
<td>3</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>Tu-05</td>
<td>71</td>
<td>M</td>
<td>Larynx</td>
<td>No</td>
<td>1</td>
<td>4</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>Tu-06</td>
<td>66</td>
<td>M</td>
<td>Hypopharynx</td>
<td>No</td>
<td>1</td>
<td>1</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>Tu-07</td>
<td>79</td>
<td>M</td>
<td>Larynx</td>
<td>No</td>
<td>1</td>
<td>1</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>Tu-08</td>
<td>57</td>
<td>M</td>
<td>Hypopharynx</td>
<td>No</td>
<td>1</td>
<td>0</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>Tu-09</td>
<td>49</td>
<td>M</td>
<td>Larynx</td>
<td>No</td>
<td>1</td>
<td>1</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>Tu-10</td>
<td>23</td>
<td>F</td>
<td>Oral cavity</td>
<td>No</td>
<td>1</td>
<td>1</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>Tu-11</td>
<td>39</td>
<td>M</td>
<td>Oropharynx</td>
<td>No</td>
<td>1</td>
<td>1</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>Tu-12</td>
<td>62</td>
<td>M</td>
<td>Larynx</td>
<td>No</td>
<td>1</td>
<td>0</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
</tr>
</tbody>
</table>

Data for the 12 lung tumours (Tu-01–Tu-12) of unknown origin, those of the previously known HNSCC localisation, relapse, pT, pN, grade as well as the results of the IHC staining of the lung tumours are presented. The negative IHC results and the classification as SQCLC are marked in bold and the positive results and the classification as met-HNSCC in italics.

M, male; F, female; IHC, immunohistochemical.

The two proteins CK19 and UGDH, which were upregulated in SQCLC, are also of interest. UGDH was not suitable to distinguish SQCLC from HNSCC. In
contrast, CK19 is positive in 97.9% of the SQCLC (AUC 0.624, sensitivity 98.0%, and specificity 30.4%). This is in line with the results of Ichinose et al who showed CK19 in combination with MMP3, ZNF830, and PI3 as a protein to differentiate between SQCLC and HNSCC [56]. As a combination of antibodies in a biomarker panel can increase the sensitivity and specificity regarding the differential diagnosis, we combined the two most promising antibodies anti-CAV1 and anti-LGALS7. The quality criteria were improved when both markers had to be positive to diagnose an HNSCC.

We tested our biomarker panel on a cohort of 12 lung tumours of unknown origin after HNSCC. The tumours were clinically classified using a classification score according to Ichinose et al [56]. This score includes clinical criteria such as local relapse of the HNSCC, the number of pulmonary tumours, and the time between the HNSCC and the lung tumour. Bohnenberger et al showed in their study that the clinical classification (here especially by Ichinose et al) did not show a difference in survival [8]. They presented differentiation into met-HNSCC and SQCLC based on proteomic analyses. The newly established groups showed a significant difference in survival. Thereby, they were able to demonstrate that the clinical classification is unreliable and therefore not ground truth. In our cohort, 3 of the 12 tumours (25%) showed concordance and 6 discordance (50%) between the immunohistochemical and clinical classification. Tu-04 and Tu-06 were immunohistochemically classified as SQCLC in contrast to the clinical classification. The pathological parameters (solitary pulmonary nodule, no relapse, pN0) of Tu-04 supported the immunohistochemical diagnosis of SQCLC. The new classification of Tu-06 is supported by the clinical information that there was no local relapse. The pN2b and death due to tumour are in favour of the diagnosis of met-HNSCC. This is a good example of difficulties trying to differentiate SQCLC and met-HNSCC clinically. Tu-07, Tu-08, Tu-09, and Tu-12 were immunohistochemically categorised as met-HNSCC in contrast to the clinical classification. The clinical and pathological data of Tu-07, Tu-08, and Tu-12 (pN2b/pN2c, death) supported the immunohistochemical classification, whereas the data for Tu-09 (pN0, no local recurrence, alive) are in favour of SQCLC.

In summary, we were able to establish an immunohistochemical marker panel that can be used to differentiate between squamous cell carcinomas of the lung (SQCLC) and the head and neck (HNSCC). We conclude that CAV1 and LGALS7, especially in combination as an antibody panel, represent promising marker candidates to differentiate between primary SQCLC and pulmonary metastases (met-HNSCC) of previously known HNSCC.

Author contributions statement

FB, HB and AR conceived the study concepts and together with JJ and PS the study design. JJ, AR, AF, PB, PK, HS and AH-E acquired data and HB, HU and TO analysed and interpreted them. All authors were involved in writing the paper and had final approval of the submitted and published versions.

References

real-time polymerase chain reaction. Histopathology 2021; 78: 593–606.


**SUPPLEMENTARY MATERIAL ONLINE**

File S1. Expression of immunohistochemical markers according to tumour subsite

Figure S1. Examples of IRS

Table S1. Complete proteomic data

Table S2. Antibodies used in this study

Table S3. Differential protein expression analysis

Table S4. Immunohistochemistry and complete clinical data of lung tumours of unknown origin in patients with prior HNSCC