Aus dem Institut für Pathologie
(Prof. Dr. med. P. Ströbel)
der Medizinischen Fakultät der Universität Göttingen

Proteomische Analyse von Plattenepithelkarzinomen des Kopf-Hals-Bereichs und Plattenepithelkarzinomen der Lunge

INAUGURAL-DISSERTATION

zur Erlangung des Doktorgrades
der Medizinischen Fakultät der
Georg-August-Universität zu Göttingen

vorgelegt von

Jonatan Felix Hoffmann

aus

Augsburg

Göttingen 2021

Dekan:	Prof. Dr. W. Brück

Betreuungsausschuss

Betreuer:	Prof. Dr. P. Ströbel
Ko-Betreuer:	Prof. Dr. M. Canis

Prüfungskommission

Referent/in	Prof. Dr. P. Ströbel
Ko-Referent/in:	Prof. Dr. D. Beutner
Drittreferent/in:	Prof. Dr. T. Meyer

Datum der mündlichen Prüfung: 18.08.2022

Hiermit erkläre ich, die Dissertation mit dem Titel "Proteomische Analyse von Plattenepithelkarzinomen des Kopf-Hals-Bereichs und Plattenepithelkarzinomen der Lunge" eigenständig angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Göttingen, den …… ……

(Unterschrift)

Die Daten, auf denen die vorliegende Arbeit basiert, wurden teilweise publiziert:

Bohnenberger H, Kaderali L, Ströbel P, Yepes D, Plessmann U, Dharia NV, Yao S, Heydt C, Merkelbach-Bruse S, Emmert A, **Hoffmann J** et al. (2018): Comparative proteomics reveals a diagnostic signature for pulmonary head-and-neck cancer metastasis. EMBO Mol Med 10, e8428

Inhaltsverzeichnis

Abbildungsverzeichnis ..... **III**

Tabellenverzeichnis ..... **IV**

Abkürzungsverzeichnis ..... **V**

1 Einleitung **1**

1.1 Maligne Neoplasien des Kopf-Hals-Bereichs 1

1.1.1 Diagnostische Unterscheidung von Lungenmetastasen und metachronen Primärkarzinomen der Lunge 2

1.1.2 Therapeutische Angriffspunkte in der HNSCC-Therapie 3

1.2 Massenspektrometrische Proteomik 4

1.2.1 Super-SILAC als Proteinquantifizierungsstandard 5

1.2.2 Proteinisolation mithilfe der filtergestützten Probenaufbereitung 6

1.3 Ziel der Studie 8

2 Material und Methoden **9**

2.1 Ethik 13

2.2 Molekularbiologische Methoden 13

2.2.1 Isolation des Tumorgewebes aus Paraffinblöcken 13

2.2.2 Proteinisolation aus FFPE-Gewebe 14

2.2.3 Filter-gestützte Probenaufbereitung (FASP) 14

2.2.4 Messung der Proteinkonzentration 15

2.2.5 Messung der Peptidmenge 16

2.3 Zellbiologische Methoden 16

2.3.1 Generelle Kulturbedingungen 16

2.3.2 Auftauen und Einfrieren von Zellen 16

2.3.3 Zelllyse 17

2.3.4 Herstellung des Super-SILAC-Standards 17

2.4 Massenspektrometrische Methoden 18

2.4.1 Massenspektrometrie 18

2.4.2 MS-Datenverarbeitung 18

2.5 Statistische Datenanalyse 19

2.6 Datenverfügbarkeit 20

3 Ergebnisse **21**

3.1 Zusammenhang zwischen Gewebevolumen und isolierter Protein- bzw. Peptidmenge 21

3.2 Evaluation des Proteinquantifizierungsstandards 23

3.3 Zusammensetzung des Patient*innenkollektivs 24

3.4 Anzahl quantifizierter Proteine und technische Reproduzierbarkeit der massenspektrometrischen Analyse 27

3.5 Proteomische Unterscheidung von SQCLC und HNSCC ... 29
3.6 Unterschiedlich exprimierte Proteine zwischen HNSCC und SQCLC ... 30

4 Diskussion ... 34

4.1 Geringe Gewebemengen sind ausreichend für eine massenspektrometrische Proteomanalyse ... 34
4.2 Komplexe proteomische Signaturen ermöglichen die Unterscheidung von HNSCC und SQCLC ... 35
4.3 Identifikation und Quantifizierung bereits bekannter onkogener Proteine ... 37
4.4 Immunhistochemische Validierung und Validierung der Proteinsignatur ... 38
4.5 Einsatz der Proteinsignatur zur klinischen Risikostratifizierung ... 38
4.6 Limitationen, Chancen und Ausblick auf zukünftige Entwicklungen ... 39

5 Zusammenfassung ... 43

6 Literaturverzeichnis ... 44

Danksagung ... 52

Abbildungsverzeichnis

Abbildung 1: Zusammenhang zwischen Gewebevolumen und Proteinmenge. 22
Abbildung 2: Aus 100 µg Protein gewonnene Peptidmenge. 23
Abbildung 3: Evaluation des Proteinquantifizierungsstandards. 24
Abbildung 4: Immunhistochemische Färbung von CK5/6 und p63. 25
Abbildung 5: Anzahl quantifizierter Proteine. 27
Abbildung 6: Technische Reproduzierbarkeit der massenspektrometrischen Analyse. 28
Abbildung 7: Gemessene Proteinintensitäten. 28
Abbildung 8: Hauptkomponentenanalyse der Proteinexpression von SQCLC und HNSCC. 29
Abbildung 9: In HNSCC bzw. SQCLC signifikant stärker exprimierte Proteine. 31
Abbildung 10: Top 10 der hochregulierten Proteine. 32

Tabellenverzeichnis

Tabelle 1: Geräte. 9
Tabelle 2: Materialien. 10
Tabelle 3: Chemikalien und Reagenzien. 11
Tabelle 4: Puffer. 12
Tabelle 5: Kits. 12
Tabelle 6: Computersoftware und Internetadressen. 12
Tabelle 7: Zelllinien. 13
Tabelle 8: Zusammensetzung des Patient*innenkollektivs. 26
Tabelle 9: Top 10 der in HNSCC hochregulierten Proteine. 33
Tabelle 10: Top 10 der in SQCLC hochregulierten Proteine. 33

Abkürzungsverzeichnis

ABC	Ammoniumhydrogencarbonat
AGC	*automatic gain control*
ALK1	Anaplastische Lymphomkinase 1
AN	Acetonitril
ATCC	*American Type Culture Collection*
BSA	Bovines Serumalbumin
C-	Carboxy-
CK	Cytokeratin
ddH_2O	doppelt deionisiertes Wasser
DCA	Desoxycholsäure
DE	*dynamic exclusion*
DGHO	Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure
EBV	Epstein-Barr-Virus
EDTA	Ethylendiamintetraacetat
EGFR	*epidermal growth factor receptor*
FASP	*filter-aided sample preparation*
FC	*fold change*
FCS	*fetal calf serum*
FDR	*false detection rate*
FFPE	*formalin-fixed and paraffin-embedded*
HCD	*higher-energy collisional dissociation*
HE	Hämatoxylin-Eosin
HMGCS	Hydroxymethylglutaryl-CoA-Synthase
HNSCC	*head and neck squamous cell carcinoma*
HPLC	*high performance liquid chromatography*
HPV	Humane Papillomaviren
IAA	Iodacetamid
IT	*ion injection time*
LOH	*loss of heterozygosity*
NSCLC	Nicht-kleinzelliges Bronchialkarzinom
MA	Massachusetts
MS	Massenspektrometrie
MS/MS	Tandem-Massenspektrometrie
MW	Mittelwert
NCE	*normalized collision energy*
PBS	*phosphate buffered saline*
PC	*principal component*

PD-L1	*programmed cell death protein ligand 1*
PS	Penicillin/Streptomycin
RPMI	*Roswell Park Memorial Institute*
SCLC	Kleinzelliges Bronchialkarzinom
SDS	*sodium dodecyl sulfate*
SILAC	*stable isotope labeling by/with amino acids in cell culture*
SQCLC	*squamous cell lung carcinoma*
SVM	*support vector machine*
TCEP	Tris(2-carboxyethyl)phosphin
TFA	Trifluoressigsäure
Th	Thomson (Einheit)

1 Einleitung

Inhalt dieser Studie ist die vergleichende proteomische Analyse von Plattenepithelkarzinomen des Kopf-Hals-Bereichs (*head and neck squamous cell carcinoma*, HNSCC) und Plattenepithelkarzinomen der Lunge (*squamous cell lung carcinoma*, SQCLC) mithilfe von quantitativer Massenspektrometrie mit dem Ziel, biologische Unterschiede zwischen diesen beiden Tumorarten, mögliche therapeutische Angriffspunkte sowie potentielle Biomarker bzw. proteomische Signaturen, welche im Rahmen der klinischen Entscheidungsfindung zur Unterscheidung von HNSCC und SQCLC benötigt werden, zu identifizieren.

1.1 Maligne Neoplasien des Kopf-Hals-Bereichs

Maligne Neoplasien des Kopf-Hals-Bereichs machen weltweit die sechst häufigste Tumorerkrankung aus (Leemans et al. 2011). Mit etwa 890.000 Neuerkrankungen und 450.000 Todesfällen jährlich sind sie für 4,9% aller weltweiten Tumorneuerkrankungen und 4,7% der tumorassoziierten Todesfälle verantwortlich (Bray et al. 2018). Das kumulative Risiko bis zu einem Alter von 74 Jahren an einer malignen Neoplasie des Kopf-Hals-Bereichs zu erkranken bzw. zu versterben beträgt für Männer 1,8% bzw. 0,9%, für Frauen 0,5% bzw. 0,3% (Bray et al. 2018). Zum Kopf-Hals-Bereich werden in diesem Zusammenhang maligne Neoplasien der Mundhöhle, des Pharynx (Naso-, Oro-, Hypopharynx), des Larynx, der Nasen- und Nasennebenhöhlen sowie der Speicheldrüsen gezählt (Lydiatt et al. 2017). Bei über 95% der malignen Neoplasien des Kopf-Hals-Bereichs handelt es sich um Plattenepithelkarzinome, weshalb im englischen Sprachraum die Bezeichnung *head and neck squamous cell carcinoma* (HNSCC) gebräuchlich ist (Leemans et al. 2011; Jou und Hess 2017). Die 5-Jahres-Überlebesrate beträgt ca. 40-50% (Leemans et al. 2011).

Zu den Hauptrisikofaktoren für die Entwicklung von HNSCC zählen der Kontakt zu Karzinogenen, allen voran durch Tabak- und Alkoholkonsum, die Infektion mit Hochrisiko-HPV(humane Papillomaviren)-Typen und EBV (Epstein-Barr-Virus) sowie genetische Prädisposition (z. B. Fanconi-Anämie) (Sankaranarayanan et al. 1998; Kutler et al. 2003; Hopkins et al. 2008; Leemans et al. 2011). Die Kombination von Tabak- und Alkoholkonsum hat dabei einen synergistischen, multiplikativen Effekt auf die Entwicklung (Talamini et al. 2002; Hashibe et al. 2009; Leemans et al. 2011).

Häufig kommt es bei betroffenen Patient*innen zum Auftreten von syn- und/oder metachronen Zweitkarzinomen (meist des Kopf-Hals-Bereichs oder der Lunge), lokoregionären Rezidiven, oder Fernmetastasen (Atabek et al. 1987; Jones et al. 1995; Leemans et al. 2011;

Jou und Hess 2017). Dies wird mit dem Konzept der Feldkanzerisierung in Verbindung gebracht, wirkt sich negativ auf das Überleben aus und trägt dazu bei, dass sich die Prognose in den vergangenen Jahren kaum verbessert hat (Slaughter et al. 1953; León et al. 1999; de Bree et al. 2000; Leemans et al. 2011; Takes et al. 2012).

Obwohl Fernmetastasen im Vergleich zu anderen Malignomen seltener auftreten, haben sie entscheidende prognostische und therapeutische Konsequenzen und treten mit rund 66-80% der Fälle bevorzugt in der Lunge auf (Calhoun et al. 1994; Ferlito et al. 2001; Takes et al. 2012). Als Risikofaktoren für die Fernmetastasierung gelten unter anderem der Ort des Primärtumors (v. a. Hypopharynx), fortgeschrittene T- und N-Stadien, unzureichende lokoregionäre Kontrolle, Zweitkarzinome und das histologische Grading (de Bree et al. 2000; Liao et al. 2007; Leitlinienprogramm Onkologie 2012; Takes et al. 2012). Vor allem bei HNSCC-Patient*innen mit einem hohen Risiko für Fernmetastasierung sollte daher ein Screening in Bezug auf Lungentumore durchgeführt werden, wofür die CT-Thorax ein geeignetes Verfahren darstellt (Atabek et al. 1987; de Bree et al. 2000; Takes et al. 2012; Leitlinienprogramm Onkologie 2012; 2019).

Da Tabakkonsum jedoch nicht nur einen wesentlichen Risikofaktor für Plattenepithelkarzinome des Kopf-Hals-Bereichs, sondern auch für Plattenepithelkarzinome des Bronchialsystems und der Lunge darstellt, müssen beim Auftreten von Lungentumoren neben Metastasen von HNSCC (metHNSCC) auch metachrone Primärkarzinome der Lunge (SQCLC) in Betracht gezogen werden (Atabek et al. 1987; Jones et al. 1995; León et al. 1999; Ferlito et al. 2001; DGHO 2018). Es wird angenommen, dass es bei ca. 8% der HNSCC-Patient*innen zu einem metachronen Primärkarzinom der Lunge kommt und dieses im Durchschnitt 6,7 Jahre später diagnostiziert wird (Crippen et al. 2019).

1.1.1 Diagnostische Unterscheidung von Lungenmetastasen und metachronen Primärkarzinomen der Lunge

Die Unterscheidung zwischen metHNSCC und SQCLC ist für das therapeutische Vorgehen von entscheidender Bedeutung: Während auf der einen Seite für metHNSCC vor allem eine palliative, systemische Therapie und nur in ausgewählten Fällen die Resektion der Metastasen in Frage kommt (Rusch 1995; 2010; Younes et al. 1997; Young et al. 2015; Leitlinienprogramm Onkologie 2012; 2019), ist der Therapieansatz bei SQCLC in erster Linie kurativ und die Prognose deutlich besser (Kuriakose et al. 2002; Henschke et al. 2003; Wisnivesky et al. 2004; Geurts et al. 2005; DGHO 2018). Bei der kurativen Therapie von SQCLC steht häufig die operative Resektion in Kombination mit einer Chemotherapie im

Vordergrund, bei kleinen Tumoren kann jedoch auch eine rein chirurgische Therapie erfolgen (DGHO 2018).

Da metHNSCC und SQCLC sich histomorphologisch nicht unterscheiden, stellt die Differentialdiagnose eine große Herausforderung und ein diagnostisches Dilemma dar. Aktuell kommen verschiedene klinische und bildgebende Unterscheidungskriterien (z. B. Histologie, Grading, TNM-Stadium, Zeitintervall zwischen HNSCC und Auftreten des Lungentumors, Anzahl/Ort der Lungentumore) zum Einsatz, die jedoch wenig validiert und nicht einheitlich festgelegt sind und so keine abschließend sichere Einschätzung erlauben, ob es sich um eine Lungenmetastase oder ein metachrones Primärkarzinom der Lunge handelt (Kuriakose et al. 2002; Geurts et al. 2005; Ichinose et al. 2016a). In den letzten Jahren waren die Genome und Transkriptome von HNSCC und SQCLC Gegenstand zahlreicher Untersuchungen. Zusätzlich zu ihrer gemeinsamen Histomorphologie zeigten sich dabei ähnliche Genexpressionsprofile und *copy number alterations*, jedoch auch Hinweise, dass Genexpressionsprofile oder Unterschiede im *loss of heterozygosity* (LOH) hilfreich bei der Unterscheidung sein könnten (van Oijen et al. 2000; Tabor et al. 2002; Geurts et al. 2005; 2009; Talbot et al. 2005; Vachani et al. 2007; Cancer Genome Atlas Research Network 2012; Cancer Genome Atlas Network 2015; Ichinose et al. 2016b). Dennoch fehlen nach wie vor klare, validierte Unterscheidungsmerkmale, insbesondere Biomarker, die klinisch dringend benötigt werden. Eine tiefergehende Untersuchung dieser Tumoren, vor allem auch auf proteomischer Ebene, ist daher unabdingbar.

1.1.2 Therapeutische Angriffspunkte in der HNSCC-Therapie

In der Therapie maligner Neoplasien kommen vermehrt sogenannte *targeted therapies* zum Einsatz, auch spielen personalisierte Behandlungsansätze eine zunehmende Rolle. Ein Beispiel hierfür ist das nicht-kleinzellige Lungenkarzinom, bei dem unter anderem Tyrosinkinase- (z. B. ALK1(Anaplastische Lymphomkinase 1)) oder PD-L1(*programmed cell death protein-ligand 1*)-Inhibitoren zur Anwendung kommen (DGHO 2018). Die Therapie der HNSCC richtet sich aktuell jedoch vor allem nach der anatomischen Lokalisation und nicht nach biologischen Merkmalen der jeweiligen Neoplasie, wobei unter anderem chirurgische Verfahren, Strahlentherapie und die Platin-basierte Chemotherapie zum Einsatz kommen (Jou und Hess 2017). Der gegen *epidermal growth factor receptor* (EGFR) gerichtete Antikörper Cetuximab ist die bisher erste und einzige erfolgreiche spezifische Therapie und wird meist in der Kombination mit einer Strahlentherapie eingesetzt (Bonner et al. 2006; Leemans et al. 2011). Das Beispiel Cetuximab zeigt, dass sich mit verbessertem Verständnis der zugrunde liegenden molekularen Mechanismen neue, vielversprechende Therapieoptionen entwickeln können

und sich die Therapie weiter in die Richtung personalisierter Therapiekonzepte entwickeln kann (Leemans et al. 2011). Zur Identifikation entsprechender therapeutischer Angriffspunkte ist die Untersuchung biochemischer Grundlagen der entsprechenden malignen Neoplasien notwendig.

1.2 Massenspektrometrische Proteomik

Da Proteine durch ihre Aktivität und Funktion eine entscheidende Rolle in allen biologischen Prozessen und insbesondere bei der Entstehung von Krankheiten und malignen Neoplasien spielen, ist eine genauere Untersuchung ihrer Expression von elementarer Bedeutung für das biomedizinische Verständnis. Hierzu wurden in den letzten Jahren systematische, proteomische Untersuchungsmethoden auf der Basis massenspektrometrischer (MS) Analysen entwickelt, die nun eine wichtige und vielversprechende Rolle in der biomedizinischen Forschung spielen (Aebersold und Mann 2003; 2016). Als Proteom wird dabei die Gesamtheit der von einem Lebewesen, einem Gewebe, einer Zelle oder einem Zellkompartiment zu einem bestimmten Zeitpunkt exprimierten Proteine bezeichnet. Unter Proteomik versteht man die Erforschung des Proteoms. Durch die technischen Fortschritte der vergangenen Jahrzehnte ist es heutzutage möglich, die absolute oder relative Menge mehrerer tausend Proteine in einem einzigen massenspektrometrischen Durchlauf mit hoher Genauigkeit und Sensitivität zu messen (Aebersold und Mann 2003; 2016).

Eine vielfach angewandte und typische Herangehensweise ist dabei die sogenannte *shotgun* oder *bottom-up* Proteomik. Sie zielt darauf ab, das Proteom möglichst unvoreingenommen und vollständig zu identifizieren und ist geeignet zur Identifikation von Biomarkern bzw. proteomischen Signaturen, insbesondere zur Tumorcharakterisierung (Washburn et al. 2001; Aebersold und Mann 2003; 2016). So konnten beispielsweise die beiden häufigsten Subtypen des nicht-kleinzelligen Lungenkarzinoms, das Adeno- und das Plattenepithelkarzinom, in einem Xenograft-Modell anhand proteomischer Signaturen unterschieden werden (Zhang et al. 2014). In anderen Studien gelang die Differenzierung zweier Subtypen des diffus-großzelligen B-Zell-Lymphoms (Deeb et al. 2012; 2015), die histologisch nicht zu unterscheiden sind, sowie die Klassifikation verschiedener Subtypen des Mammakarzinoms (Tyanova et al. 2016a) anhand der massenspektrometrisch ermittelten Proteinexpressionsprofile.

Im Rahmen der *bottom-up* Proteomik werden zunächst die zu untersuchenden Proteine aus Zellen oder Gewebe isoliert und durch Proteasen wie Trypsin zu Peptiden gespalten, da die Massenspektrometrie von Peptiden sensitiver ist als die von Proteinen (Aebersold und Mann 2003; 2016). Die Peptide werden anschließend mittels Hochleistungsflüssigkeitschromato-

graphie (*high performance liquid chromatography*, HPLC) aufgetrennt und mittels Elektrospray-Ionisation (Fenn et al. 1989) ionisiert (Aebersold und Mann 2003). Daraufhin werden die Peptide massenspektrometrisch analysiert und das Verhältnis von Masse zu Ladung (m/z) und die Intensität jedes Peptids bestimmt und in einem MS-Spektrum mit dem m/z-Verhältnis auf der Abszisse und der Intensität auf der Ordinate dargestellt (Steen und Mann 2004). Da diese sogenannte MS-Analyse bzw. die Masse eines Peptids allein nicht ausreicht um dieses Peptid zu identifizieren, werden die Peptide fragmentiert und die Fragmente erneut massenspektrometrisch analysiert (Aebersold und Mann 2003). Dies wird als tandem MS bzw. MS/MS-Analyse bezeichnet und erlaubt die Bestimmung eines Teils der Aminosäuresequenz des Peptids (Steen und Mann 2004). Die MS- und MS/MS-Daten sowie die Daten zur Aminosäuresequenz eines Peptids werden mit einer Datenbank abgeglichen und so das zugehörige Protein identifiziert (Aebersold und Mann 2003). Mithilfe von Software wie MaxQuant (Cox und Mann 2008) und Andromeda (Cox et al. 2011) gelingt insbesondere die quantitative Datenauswertung.

1.2.1 Super-SILAC als Proteinquantifizierungsstandard

Oft ist nicht nur die reine Anwesenheit eines Proteins, sondern vielmehr die quantitative Veränderung zwischen zwei unterschiedlichen Situationen, wie beispielsweise verschiedenen malignen Neoplasien oder experimentell unterschiedlich behandelten Zellen, von Interesse (Walther und Mann 2010). Eine gut etablierte Methode zur massenspektrometrischen Proteinquantifikation stellt das *stable isotope labeling by/with amino acids in cell culture* (SILAC) dar, wobei schwere Aminosäuren in die Proteome verschiedener Zellen eingebaut und anschließend zur Quantifizierung genutzt werden (Jiang und English 2002; Ong et al. 2002; Zhu et al. 2002). Da die klassische SILAC-Methode auf den Einsatz in der Zellkultur beschränkt ist, wurde der sogenannte Super-SILAC-Standard als Proteinquantifizierungsstandard entwickelt (Geiger et al. 2010). Dabei handelt es sich um eine exakte und gut etablierte Methode, um Proteine in menschlichem Gewebe massenspektrometrisch zu quantifizieren (Shenoy und Geiger 2015).

Zur Herstellung des Proteinquantifizierungsstandards werden in den Proteomen von vier bis sechs Zelllinien die natürlichen, essentiellen Aminosäuren Arginin (Scott et al. 2000) und Lysin durch Analoga mit einer höheren Masse ersetzt (Shenoy und Geiger 2015). Dieser Massenunterschied entsteht durch die Markierung der Aminosäuren mit stabilen, nicht radioaktiven Isotopen, beispielsweise ^{13}C anstatt ^{12}C (Ong et al. 2002; Geiger et al. 2010). Durch den Einsatz eines speziellen SILAC-Zellkulturmediums, in dem die natürlichen Varianten durch die markierten Analoga ersetzt wurden, werden die markierten Aminosäuren

vollständig in die Proteome der Zelllinien integriert (Geiger et al. 2010). Die Kombination von Arginin und Lysin bietet sich besonders mit dem Einsatz der Protease Trypsin an. Da Trypsin Proteine ausschließlich an der C-terminalen Seite von Arginin und Lysin spaltet (Olsen et al. 2004), verfügt somit jedes Peptid, mit Ausnahme des C-terminalen Peptids eines Proteins, über eine markierte Aminosäure und kann zur Quantifizierung genutzt werden (Geiger et al. 2011; Shenoy und Geiger 2015). Außerdem erzeugt Trypsin Peptide in einer idealen Größe für die Massenspektrometrie (Ong et al. 2003).

Nach der vollständigen Integration der markierten Aminosäuren in die Proteome der Zelllinien werden diese lysiert, die Proteine isoliert und zum Proteinquantifizierungsstandard gemischt (Geiger et al. 2010; 2011). Dieser Standard wird den Proteinen des zu untersuchenden Gewebes im Verhältnis 1:1 hinzugegeben, beides gemeinsam durch Trypsin proteolytisch verdaut und anschließend massenspektrometrisch analysiert (Geiger et al. 2010; 2011). Durch den bekannten Massenunterschied, der durch die isotopisch markierten Aminosäuren erzeugt wurde, können Peptide des Standards von Peptiden des Gewebes unterschieden und das Verhältnis der Intensität eines Peptid im zu untersuchenden Gewebe zur Intensität dieses Peptids im Standard berechnet werden (Peptidintensität in der Gewebeprobe/Peptidintensität im Proteinquantifizierungsstandard). Setzt man dieses Verhältnis nun ins Verhältnis zu einer anderen Gewebeprobe (Verhältnis der Verhältnisse) lässt sich der Proteinquantifizierungsstandard kürzen und die Peptide/Proteine zweier Proben können relativ zueinander quantifiziert werden (Geiger et al. 2010; 2011).

Die Auswahl der für den Proteinquantifizierungsstandard eingesetzten Zelllinien richtet sich nach dem zu untersuchenden Gewebe und sollte dieses möglichst gut repräsentieren. Die Zelllinien sollten daher eine gewisse Ähnlichkeit zu diesem Gewebe haben, sich aber gleichzeitig voneinander unterscheiden (Geiger et al. 2011).

Vorteile des Super-SILAC sind, dass neben der genauen Quantifizierung an menschlichem Gewebe auch die Lagerung des Standards möglich, die Herstellung günstig und somit auch eine Anwendung in großem Stil möglich ist (Shenoy und Geiger 2015). Ein Nachteil stellt die Beschränkung auf biologische Proben dar, für die repräsentative Zelllinien existieren (Shenoy und Geiger 2015).

1.2.2 Proteinisolation mithilfe der filtergestützten Probenaufbereitung

Neben der Proteinquantifizierung ist auch die Proteinisolation ein entscheidender Bestandteil der massenspektrometrischen Proteomanalyse. Insbesondere die Isolation und Analyse von Proteinen aus formalinfixiertem und in Paraffin eingebettetem Gewebe (*formalin-fixed and*

paraffin-embedded, FFPE) ist hierbei sehr interessant, da im Rahmen der Diagnostik verschiedenster Erkrankungen menschliche Gewebe entnommen und zur histopathologischen Untersuchung in Formalin fixiert und in Paraffin eingebettet werden. Diese FFPE-Gewebe werden in Pathologischen Instituten und Biobanken archiviert und stellen wertvolle Gewebedatenbanken, insbesondere zur retrospektiven Proteomanalyse und Biomarkeridentifikation, dar.

Zur Isolation von Proteinen aus FFPE-Gewebe wurden mehrere Protokolle entwickelt, bei denen unter anderem Natriumdodecylsulfat (*sodium dodecyl sulfate*, SDS) zum Einsatz kommt (Shi et al. 2006; Guo et al. 2007; Jiang et al. 2007). SDS ist jedoch für die massenspektrometrische Analyse ungeeignet und muss daher wieder entfernt werden (Nagaraj et al. 2008), wofür die filtergestützte Probenaufbereitung (*filter-aided sample preparation*, FASP) ein geeignetes Verfahren darstellt (Wiśniewski et al. 2009b). Dabei wird SDS in einem Zentrifugalfilter durch Urea ausgetauscht, die Proteine werden proteolytisch verdaut und die dabei entstehenden Peptide isoliert, welche dann der massenspektrometrischen Analyse zugeführt werden können (Wiśniewski et al. 2009b). In diversen Untersuchungen konnte gezeigt werden, dass diese Methode zur Proteomanalyse geeignet ist und mehrere Tausend Proteine in einem einzigen MS/MS-Durchlauf identifiziert und quantifiziert werden können (Wiśniewski et al. 2009a; 2010; 2011; Ostasiewicz et al. 2010; Wiśniewski 2013; Erde et al. 2014; Bohnenberger et al. 2015). Der Einsatz von FASP, insbesondere in der Kombination mit einem Proteinquantifizierungsstandard wie auch in dieser Studie angewendet, stellt einen leistungsfähigen Ansatz zur Identifikation von Biomarkern und proteomischen Signaturen zur Charakterisierung maligner Neoplasien dar (Wiśniewski et al. 2011; Deeb et al. 2012; 2015; Zhang et al. 2014; Bohnenberger et al. 2015) und ist damit gut geeignet, die klinisch dringend benötigten Biomarker zur Unterscheidung von metHNSCC und SQCLC zu identifizieren.

1.3 Ziel der Studie

Plattenepithelkarzinomen des Kopf-Hals-Bereichs (*head and neck squamous cell carcinoma*, HNSCC) sind weltweit häufig auftretende maligne Neoplasien. Da Tabakkonsum nicht nur ein Risikofaktor für die Entstehung von HNSCC sondern auch von Plattenepithelkarzinomen der Lunge (*squamous cell lung carcinoma*, SQCLC) ist, können bei Patient*innen mit HNSCC im Verlauf der Erkrankung sowohl Lungenmetastasen des HNSCC als auch metachrone SQCLC entstehen. Die Unterscheidung zwischen diesen beiden malignen Neoplasien basiert, obwohl therapeutisch hoch relevant, auf wenig validierten Entscheidungskriterien. Trotz ausführlicher Untersuchung auf genetischer und transkriptomischer Ebene fehlen nach wie vor hilfreiche Biomarker und klare Unterscheidungsmerkmale.

Die massenspektrometrische Proteomanalyse hat sich in den letzten Jahren zu einer wichtigen Technologie in der biomedizinischen Forschung entwickelt. In einem einzelnen massenspektrometrischen Durchlauf können mehrere Tausend Proteine identifiziert und quantifiziert werden. Dies ist geeignet zur Identifikation von Biomarkern bzw. proteomischen Signaturen, insbesondere zur Tumorcharakterisierung.

In dieser Studie sollen daher die Tumorgewebe von insgesamt 30 HNSCC, bei denen es im Verlauf der Erkrankung zum Auftreten von Lungentumoren mit Plattenepithelkarzinom-Histologie kam, sowie insgesamt 44 SQCLC massenspektrometrisch analysiert und quantitativ miteinander verglichen werden. Die Proteine werden dazu mittels der filtergestützten Probenaufbereitung (*filter-aided sample preparation*, FASP) aus formalinfixiertem und in Paraffin eingebettetem Gewebe (*formalin-fixed and paraffin-embedded*, FFPE) isoliert und in Kombination mit einem Proteinquantifizierungsstandard massenspektrometrisch analysiert.

2 Material und Methoden

Tabelle 1: Geräte.

Produkt	Hersteller
Akkupipettierhilfe accu-jet® pro	Brand, Wertheim
Autoklav VX-100	Systec, Linden
Feinwaage R160P	Sartorius, Göttingen
Massenspektrometer Orbitrap Fusion™	Thermo Fisher Scientific, Waltham, MA
Mikrotom SM2000R	Leica, Wetzlar
Photometer Infinite® M200 Pro	Tecan Group, Männedorf, Schweiz
Sicherheitswerkbank HeraSafe HS15	Heraeus, Hanau
Thermomixer Comfort	Eppendorf, Hamburg
Ultraschallgerät Sonopuls HD 70	Bandelin electronic, Berlin
Vakuumzentrifuge Concentrator 5301	Eppendorf, Hamburg
Vortex Genie 2	Schütt Labortechnik, Göttingen
Water Purification System Milli-Q	Merck, Darmstadt
Zellbrutschrank	Binder, Tuttlingen
Zellmikroskop Axiovert 25	Carl Zeiss, Jena
Zentrifugen 5430 R, 5417 R und 5804	Eppendorf, Hamburg

Tabelle 2: Materialien.

Produkt	Hersteller
1,5 ml und 2 ml Verschlussröhrchen	Eppendorf, Hamburg
1,5 ml Verschlussröhrchen Protein LoBind® *Tube*	Eppendorf, Hamburg
2,0 ml Verschlussröhrchen *Low Protein Binding Tube*	Thermo Fisher Scientific, Waltham, MA
1,5ml Verschlussröhrchen lichtgeschützt	Sarstedt, Nümbrecht
15ml und 50ml Polypropylen Röhrchen	Greiner Bio-One, Frickenhausen
Disperser	Eppendorf, Hamburg
HPLC C18 Säulen	Dr. Maisch, Ammerbuch-Entringen
HPLC-System UltiMate 3000 RSLCnano	Thermo Fisher Scientific, Waltham, MA
Kühlbehälter Mr. Frosty™	Thermo Fisher Scientific, Waltham, MA
Kryo-Röhrchen CryoPure 1,8ml	Sarstedt, Nümbrecht
Mikrotestplatte 384 Well, PS, F-Boden, Fluotrac.	Greiner Bio-One, Frickenhausen
Mikrotestplatte 96 Well	Sarstedt, Nümbrecht
Parafilm „M“ Laboratory Films	Bemis Company, Neenah, Wisconsin
Pipetten (2,5; 20; 100; 1.000 µl)	Eppendorf, Hamburg; Sartorius, Göttingen
Pipettenspitzen Biosphere® 20µl, 200µl und 1000µl	Sarstedt, Nümbrecht
Serologische Pipette 5ml	Sarstedt, Nümbrecht
Umkehrphasen-C18-Vorsäule (5 mm lang, 0,3 mm Innendurchmesser)	Thermo Fisher Scientific, Waltham, MA
Umkehrphasenkieselgel ReproSil-Pur C18-AQ 1,9 µm	Dr. Maisch, Ammerbuch-Entringen
Zellkultur TC-Flasche T25, T75 und T175, Stand., Bel. K	Sarstedt, Nümbrecht
Zellschaber 25cm	Sarstedt, Nümbrecht
Zentrifugalfilter Microcon® 30kDa	Merck Millipore, Tullagreen, Irland
ZipTip® C18 Pipettenspitzen	Merck Millipore, Tullagreen, Irland

Tabelle 3: Chemikalien und Reagenzien.

Produkt	Hersteller
Acetonitril (AN)	AppliChem, Darmstadt
Ameisensäure	Sigma, St. Louis, Missouri
Ammoniumhydrogencarbonat (ABC)	Th. Geyer, Renningen
L-Arginin +6 ($^{13}C_6$-Arginin)	Cambridge Isotope Laboratories, Tewksbury, MA
Desoxycholsäure (DCA)	Sigma, St. Louis, Missouri
Dimethylsulfoxid (DMSO)	Honeywell, Muskegon, Michigan
Dinatriumhydrogenphosphat (Na_2HPO_4)	Th. Geyer, Renningen
Ethanol	Th. Geyer, Renningen
Ethylacetat	Th. Geyer, Renningen
Dialysiertes fetales Kälberserum	Gibco, Darmstadt
L-Glutamin	Gibco, Darmstadt
Iodoazetamid (IAA)	Merck, Darmstadt
Kaliumchlorid (KCl)	Th. Geyer, Renningen
Kaliumdihydrogenphosphat (KH_2PO_4)	Th. Geyer, Renningen
L-Lysin +4 (D_4-Lysin)	Cambridge Isotope Laboratories, Tewksbury, MA
Methanol	Th. Geyer, Renningen
Natriumchlorid (NaCl)	Th. Geyer, Renningen
Natriumdodecylsulfat (SDS)	Th. Geyer, Renningen
Penicillin/Streptomycin (PS)	Gibco, Darmstadt
Pierce™ Water, LC-MS Grade	Thermo Fisher Scientific, Waltham, MA
RPMI-1640-Medium	Gibco, Darmstadt
SILAC-RPMI-1640-Medium	Gibco, Darmstadt
Trifluoressigsäure (TFA)	Sigma, St. Louis, Missouri
Tris(2-carboxyethyl)phosphinhydrochlorid (TCEP)	Sigma, St. Louis, Missouri
Trypsin-EDTA	Gibco, Darmstadt
Trypsin, *Sequencing Grade Modified*	Promega, Madison, Wisconsin
L-Tryptophan	Sigma, St. Louis, Missouri
Urea	AppliChem, Darmstadt
Xylol	Th. Geyer, Renningen

Tabelle 4: Puffer.

Puffer	Zusammensetzung
Lysepuffer	4% SDS, 0,2% DCA, 50 mM TCEP, 100 mM ABC, pH 8,0
Austauschpuffer	8 M Urea, 0,2% DCA, 100 mM ABC, pH 8,0
Alkylierungspuffer	8 M Urea, 50 mM IAA, 100 mM ABC, pH 8,0
Aufschlusspuffer	0,2% DCA, 50 mM ABC, pH 8,0
Peptidwiederherstellungspuffer	50 mM ABC, pH 8,0
Probenbeladungspuffer	2% AN, 0,05% TFA
PBS	137,5 mM NaCl, 2,7 mM KCl, 1,5 mM KH_2PO_4, 8,6 mM Na_2HPO_4

Sofern nicht anders angegeben, wurden alle Puffer in doppelt deionisiertem Wasser (ddH_2O) angesetzt.

Tabelle 5: Kits.

Produkt	Hersteller
Ionic Detergent Compatibility Reagent (IDCR)	Thermo Fisher Scientific, Waltham, MA
Pre-Diluted Protein Assay Standards: Bovine Serum Albumin Set	Thermo Fisher Scientific, Waltham, MA
Pierce™ *660nm Protein Assay Reagent*	Thermo Fisher Scientific, Waltham, MA

Tabelle 6: Computersoftware und Internetadressen.

Computersoftware bzw. Internetadresse	Verwendung
Bioconductor 3.2 (http://www.bioconductor.org)	Statistische Datenanalysen
Excel 16.26, Microsoft, Redmond, Washington	Statistik, Erstellung von Diagrammen
https://www.ncbi.nlm.nih.gov	Literaturrecherche
https://www.uniprot.org (UniProtKB und Swiss-Prot)	Protein-Datenbank
i-control 1.10, Tecan Group, Männedorf, Schweiz	Auswertung photometrischer Messungen
MaxQuant 1.0.12.31, MPI, München	Auswertung MS-Daten
Photoshop CS6, Adobe, San José, Kalifornien	Bildbearbeitung
R 3.4.1, The R Foundation (https://www.r-project.org)	Statistische Datenanalysen

Tabelle 7: Zelllinien.

Zelllinie	Beschreibung
NCI-H2228 (ATCC® CRL-5935™)	Diese humane, adhärente Zelllinie wurde im Jahr 1989 aus einem Adenokarzinom der Lunge einer nichtrauchenden Patientin etabliert. Weitere Daten zur Patientin oder zum Stadium der Erkrankung sind nicht bekannt (Phelps et al. 1996).
NCI-H2228 (ATCC® CRL-5935™)	Diese humane, adhärente Zelllinie wurde im Jahr 1986 aus dem Pleuraerguss einer 49-jährigen Patientin mit kleinzelligem Bronchialkarzinom (Stadium: *extensive disease*) etabliert. Die Patientin war Raucherin mit 60 *pack-years* und erhielt vor der Entnahme der Zellen keinerlei Therapie (Phelps et al. 1996; Virmani et al. 1998).
HCC-15	Diese humane, adhärente Zelllinie wurde aus einem Plattenepithelkarzinom der Lunge eines 47-jährigen Patienten etabliert (Virmani et al. 1998; DSMZ 2019a).
HCC-44	Diese humane, adhärente Zelllinie wurde aus einem Adenokarzinom der Lunge einer 54-jährigen Patienten etabliert (Virmani et al. 1998; DSMZ 2019b).

2.1 Ethik

Das Einverständnis für die Nutzung menschlichen Patient*innenmaterials wurde durch die Ethikkommission der Universitätsmedizin Göttingen (#1-2-08, 9-12-10, DOK_99_2016) eingeholt. Eine informierte Einwilligung wurde von allen Patient*innen eingeholt. Alle Maßnahmen wurden in Einklang mit der Deklaration von Helsinki sowie institutionellen, landes- und bundesweiten Richtlinien bzw. Gesetzen durchgeführt.

2.2 Molekularbiologische Methoden

2.2.1 Isolation des Tumorgewebes aus Paraffinblöcken

Die untersuchten Tumorgewebeproben waren durch die Klinik für Thorax-, Herz- und Gefäßchirurgie und die Klinik für Hals-, Nasen- und Ohrenheilkunde der Universitätsmedizin Göttingen im Rahmen der chirurgischen Therapie entnommen und durch das Institut für Pathologie zur Diagnostik in Formalin fixiert und in Paraffin eingebettet worden. Zunächst wurde durch Mikroskopieren der HE (Hämatoxylin-Eosin)-gefärbten Schnitte der entsprechenden Tumorproben ein Areal mit einer Tumorzelldichte von >80% identifiziert. Anschließend wurden am Mikrotom jeweils fünf Schnitte á 10 µm Dicke dieses Areals aus dem korrespondierenden Paraffinblock entnommen und bis zur weiteren Verarbeitung in einem 1,5 ml Verschlussröhrchen bei Raumtemperatur lichtgeschützt aufbewahrt. Dieser als

Makrodissektion bezeichnete Schritt wurde bereits zuvor als mögliches Verfahren zur Tumorisolation aus FFPE(*formalin-fixed and paraffin-embedded*)-Gewebe beschrieben (Cree et al. 2014).

2.2.2 Proteinisolation aus FFPE-Gewebe

Zur Deparaffinisierung wurde das isolierte FFPE-Gewebe (s. 2.2.1) je zwei Mal für je 5 min bei Raumtemperatur in Xylol und 100% Ethanol inkubiert. Noch verbleibendes Ethanol wurde durch die Zentrifugation in einer Vakuumzentrifuge für 10 min bei 45°C verflüchtigt. Nach der anschließenden Zugabe von 500 µl Lysepuffer wurde das Gewebe mithilfe eines Dispersers mechanisch zerkleinert, für 60 min bei 90°C und 600 rpm im Thermomixer inkubiert, für 3 min (Power 40%, Cycle 20%) mithilfe des Ultraschallgeräts Sonopuls HD 70 (Bandelin electronic, Berlin) sonifiziert und für 15 min bei 20.000 g bei Raumtemperatur zentrifugiert. Nun lagen die Proteine in Lysepuffer gelöst vor (im Folgenden als FFPE-Lysat bezeichnet).

2.2.3 Filter-gestützte Probenaufbereitung (FASP)

Die Grundlage für die weitere Aufarbeitung der FFPE-Lysate (s. 2.2.2) bildete die bereits zuvor beschriebene Filter-gestützte Probenaufbereitung (*filter-aided sample preparation*, FASP) (Wiśniewski et al. 2009b; Wiśniewski 2013; Bohnenberger et al. 2015). Diese diente vor allem dazu, Chemikalien, welche für die Lyse und Solubilisierung verwendet worden waren, jedoch ungeeignet für die Massenspektrometrie sind, beispielsweise SDS, wieder aus den Proben zu entfernen.

Direkt im Anschluss an die Messung der jeweiligen Proteinkonzentrationen (s. 2.2.4) wurden je 25 µg Protein des FFPE-Lysats und des Super-SILAC-Standards im Verhältnis 1:1 (m/m) gemischt und in einen 30 kDa Microcon® Zentrifugalfilter (Merck Millipore, Tullagreen, Irland) überführt, welcher auf ein dafür vorgesehenes 1,5 ml Verschlussröhrchen als Auffangbehälter aufgesetzt wurde und nur für Moleküle mit einer Größe <30kDa durchlässig war. Insgesamt vier Mal wurden der Filtereinheit 200 µl Austauschpuffer hinzugegeben und anschließend zentrifugiert, um SDS zu entfernen. Alle in diesem Abschnitt beschriebenen Zentrifugationsschritte wurden, sofern nicht anders angegeben, für 10 min bei 14.000 g bei 20°C durchgeführt und anschließend das Filtrat aus dem Auffangbehälter entfernt. Danach wurden 100 µl Alkylierungspuffer hinzugefügt, die Probe für 60 min bei 37°C bei 300 rpm lichtgeschützt im Thermomixer inkubiert und anschließend zentrifugiert. Daraufhin wurden insgesamt drei Mal 200 µl Aufschlusspuffer hinzugefügt und zentrifugiert, um Urea zu

entfernen. Im Anschluss wurde die Filtereinheit auf ein neues Verschlussröhrchen übertragen und 90 µl Aufschlusspuffer sowie 1 µg *sequencing grade modified* Trypsin (Promega, Madison, Wisconsin) hinzugefügt. Somit lag das Verhältnis von Trypsin:Protein bei 1:50 (m/m). Danach wurde die Filtereinheit mithilfe von Parafilm verschlossen und über Nacht (mind. 12 h) bei 37°C und 300 rpm im Thermomixer inkubiert. Daraufhin wurde zentrifugiert, ohne das Filtrat zu verwerfen. Insgesamt zwei Mal wurden 50 µl Peptidwiederherstellungspuffer hinzugefügt und zentrifugiert, ohne das Filtrat zu verwerfen, um die Peptide im Auffangbehältnis zu sammeln.

Um Desoxycholsäure (DCA) zu entfernen, wurde das die Peptide beinhaltende Filtrat anschließend mit 200 µl Ethylacetat und 2,5 µl TFA (Trifluoressigsäure) gemischt und in ein 1,5 ml LoBind® Verschlussröhrchen (Eppendorf, Hamburg) transferiert. Insgesamt drei Mal wurden nun 200 µl Ethylacetat hinzugefügt, für 10 s sonifiziert (Power 40%, Cycle 20%), für 10 min bei 16.000 g bei Raumtemperatur zentrifugiert und die anschließend oben liegende organische Schicht entfernt. Nach der letzten Wiederholung wurde die Probe für 5 min bei 60°C inkubiert, um verbliebenes Ethylacetat zu verflüchtigen. Anschließend wurde die Probe in einer Vakuumzentrifuge bei 45°C getrocknet. Das daraufhin vorliegende Pellet wurde insgesamt drei Mal mit 50% Methanol resuspendiert und erneut bei 45°C in der Vakuumzentrifuge getrocknet. Vor dem letzten Trocknungsschritt wurde durch die Messung der Peptidmenge (s. 2.2.5) sichergestellt, dass ausreichend Peptid für die massenspektrometrische Analyse vorhanden war. Die Peptide wurden mit ZipTip® C18 Pipettenspitzen (Merck Millipore, Tullagreen, Irland) entsprechend der Herstellerangaben entsalzt und erneut in der Vakuumzentrifuge bei 45°C getrocknet. Die getrockneten Proben wurden bis zur massenspektrometrischen Analyse bei -80°C aufbewahrt.

2.2.4 Messung der Proteinkonzentration

Die Messung der Proteinkonzentration erfolgte für FFPE-Lysate, Zelllysate aus der Zellkultur und den Super-SILAC-Standard in gleicher Weise: Unter Verwendung des Pierce™ 660 nm *Protein Assay Reagent* (Thermo Fisher Scientific, Waltham, MA) in Kombination mit *Ionic Detergent Compatibility Reagent* (IDCR; Thermo Fisher Scientific, Waltham, MA) entsprechend der Herstellerangaben wurde die Absorption der zu messenden Proben sowie einer standardisierten Verdünnungsreihe aus bovinem Serum-Albumin (*Pre-Diluted Protein Assay Standards: Bovine Serum Albumin Set;* Thermo Fisher Scientific, Waltham, MA) im Photometer Infinite® M200 Pro (Tecan Group, Männedorf, Schweiz) bei 660 nm gemessen. Die Proteinkonzentration wurde anschließend anhand des Vergleichs mit der Verdünnungsreihe in Microsoft Excel berechnet.

2.2.5 Messung der Peptidmenge

Tryptophan kommt in der Natur mit einer Häufigkeit von 1,1% der Aminosäuren in Proteinen vor. Zusätzlich wird Tryptophan unter UV-Licht zur Fluoreszenz angeregt. Über die photometrische Messung der Fluoreszenzemission bei 360 nm wurde im Vergleich mit einer standardisierten Verdünnungsreihe auf die Tryptophan- und damit auch auf die Peptidmenge der Probe geschlossen. Zur Fluoreszenzanregung wurde UV-Licht der Wellenlänge 280 nm verwendet. Die Verdünnungsreihe wurde unmittelbar vor der Messung in lichtgeschützten 1,5 ml Reagenzgefäßen mit 50% Methanol als Trägerlösung hergestellt. Es kam das Photometer Infinite® M200 Pro (Tecan Group, Männedorf, Schweiz) in Verbindung mit einer 384-Well Mikroplatte (Greiner Bio-One, Frickenhausen) zum Einsatz.

2.3 Zellbiologische Methoden

2.3.1 Generelle Kulturbedingungen

Bei allen eingesetzten Zelllinien handelte es sich um adhärente Zellen. Die Zellen wurden bei 37°C, 5% CO_2 und 90% Luftfeuchtigkeit im Medium (s. 2.3.4) kultiviert. Je nach Teilungsrate und Dichte wurde alle drei bis fünf Tage ein Teil der Zellen verworfen oder die Zellen auf weitere Kulturflaschen expandiert. Hierzu wurde das Medium verworfen, die Kulturflasche einmal mit PBS (*phosphate buffered saline*) gewaschen, 2 ml Trypsin-EDTA-Lösung hinzugegeben und anschließend für 2 min bei 37°C inkubiert, sodass sich die adhärenten Zellen lösten. Das Trypsin wurde durch RPMI-1640-Medium inaktiviert, die Zellen für 4 min bei 400 g bei Raumtemperatur zentrifugiert, anschließend in neuem Medium aufgenommen und entweder auf eine Kulturflasche verdünnt aufgetragen oder auf mehrere Kulturflaschen expandiert.

2.3.2 Auftauen und Einfrieren von Zellen

Zur längeren Einlagerung wurden die adhärenten Zellen wie bereits beschrieben (s. 2.3.1) mithilfe von Trypsin-EDTA gelöst, für 4 min bei 400 g bei Raumtemperatur zentrifugiert, in 1 ml FCS mit 10% DMSO (Dimethylsulfoxid) resuspendiert und in Kryo-Röhrchen überführt. Die Zellen wurden danach mithilfe eines Gefrierbehälters um -1°C/min auf -80°C abgekühlt und anschließend in flüssigem Stickstoff gelagert.

Zur erneuten Verwendung wurden die Zellen in einem 37°C warmen Wasserbad aufgetaut, in Medium aufgenommen und für 4 min bei 400 g bei Raumtemperatur zentrifugiert.

Danach wurden die Zellen in frischem Medium aufgenommen und eine Kulturflasche überführt.

2.3.3 Zelllyse

Die Kulturflasche der zu lysierenden Zellen wurde zunächst zwei Mal mit je 5 ml PBS gewaschen und die Zellen in 5 ml PBS mithilfe eines Zellschabers abgelöst. Anschließend wurden die Zellen für 4 min bei 400 g bei Raumtemperatur zentrifugiert, der Überstand verworfen, die Zellen erneut in 1 ml PBS aufgenommen und in ein 1,5 ml Verschlussröhrchen überführt. Die Zellen wurden für 4 min bei 400 g bei 4°C zentrifugiert und der Überstand verworfen. Danach wurden die Zellen in 1 ml PBS resuspendiert und erneut für 4 min bei 400 g bei 4°C zentrifugiert, der Überstand verworfen und 200 µl Lysepuffer hinzugefügt. Hierbei kam der gleiche Puffer wie bei der Lyse des FFPE-Gewebes zum Einsatz. Es folgte die Inkubation im Thermomixer bei 90°C und 600 rpm für 15 min und die Sonifikation für 30 s (Power 40%, Cycle 20%). Das Zelllysat wurde für 15 min bei 20.000 g bei Raumtemperatur zentrifugiert, der Überstand in ein neues 1,5 ml Verschlussröhrchen überführt und bis zur weiteren Verwendung bei -80°C gelagert.

2.3.4 Herstellung des Super-SILAC-Standards

Zur Herstellung des Super-SILAC-Standards kamen die vier Bronchialkarzinomzelllinien NCI-H2228, NCI-H1339, HCC-15 und HCC-44, welche bei der *American Type Culture Collection* erworben worden waren, zum Einsatz. Als Medium wurde SILAC-RPMI-1640-Medium mit 10% dialysiertem FCS, 2 mM L-Glutamin, 50 U/ml Penicillin, 50 µg/ml Streptomycin, 2 mg/ml $^{13}C_6$-Arginin (+6U) (Cambridge Isotope Laboratories, Tewksbury, MA) und 4 mg/ml D_4-Lysin (+4U) (Cambridge Isotope Laboratories, Tewksbury, MA) verwendet. Die weiteren Parameter unterschieden sich nicht von den generellen Zellkulturbedingungen (s. 2.3.1). Um das Proteom der Zellen vollständig mit den schwereren Aminosäuren zu markieren, wurden die Zellen für mindestens zehn Zellzyklen in SILAC-Medium kultiviert. Anschließend wurden die Zellen lysiert (s. 2.3.3), die Proteinkonzentration der Lysate bestimmt (s. 2.2.4) und die Lysate der vier Zelllinien im Mengenverhältnis 1:1:1:1 zum Super-SILAC-Standard gemischt. Dieser wurde aliquotiert und bis zur weiteren Verwendung bei -80°C gelagert.

2.4 Massenspektrometrische Methoden

Die in diesem Abschnitt beschriebenen Methoden wurden von der Forschungsgruppe „Bioanalytische Massenspektrometrie“ am Max-Planck-Institut für biophysikalische Chemie Göttingen (Prof. Dr. Henning Urlaub) und von der Medizinischen Klinik II, Universitätsklinik Frankfurt am Main (Prof. Dr. Thomas Oellerich) durchgeführt.

2.4.1 Massenspektrometrie

Die Peptide wurden in Probenbeladungspuffer resuspendiert und durch ein UltiMate 3000 RSLCnano HPLC(*high performance liquid chromatography*)-System (Thermo Fisher Scientific, Waltham, MA) getrennt, das mit einem Orbitrap Fusion (Thermo Fisher Scientific, Waltham, MA) gekoppelt war. Zuerst wurden die Peptide auf einer Umkehrphasen-C18-Vorsäule (5 mm lang, Innendurchmesser 0,3 mm) für 3 min entsalzt. Nach 3 min wurde die Vorsäule durch die mit ReproSil-Pur C18-AQ 1,9 µm (Dr. Maisch, Ammerbuch-Entringen) gepackte Analysesäule (30 cm lang, Innendurchmesser 75 µm) ausgewechselt. Lösungsmittel A bestand aus 0,1% Ameisensäure in Wasser und Lösungsmittel B bestand aus 80% Acetonitril und 0,1% Ameisensäure in Wasser. Die Peptide wurden mit einer Flussrate von 300 nl/min über 166 min (5 – 42% Gradient) eluiert. Die Temperatur der Vorsäule und der Säule wurde während der Chromatographie auf 50°C eingestellt. MS-Daten wurden auf dem Orbitrap Fusion im datenabhängigen Top-Speed-Modus erfasst, wobei zahlreiche Vorläufer im m/z-Bereich von 350 – 1.500 Th und mit einem Ladezustand von 2 – 6 aus dem MS1-Scan für die MS2-Fragmentierung, mit einem Isolationsfenster von 1,6 Th und einer maximalen Zykluszeit von 3 s, ausgewählt wurden. MS1-Scans wurden mit einer Auflösung von 120.000 und einem AGC(*automatic gain control*)-Ziel von 4E5 aufgenommen. Ausgewählte Vorläufer wurden einer HCD(*higher-energy collisional dissociation*)-Fragmentierung mit einer normalisierten Kollisionsenergie (*normalized collision energy*, NCE) von 35 unterzogen. MS2-Scans wurden in der Ionenfalle im Rapid-Scan-Modus mit einer maximalen Ioneninjektionszeit (*ion injection time*, IT) von 250 ms und einem AGC-Ziel von 3E3 aufgenommen. Alle Scan-Ereignisse wurden automatisch parallelisiert. Der dynamische Ausschluss (*dynamic exclusion*, DE) wurde auf 60 s eingestellt.

2.4.2 MS-Datenverarbeitung

MS/MS-Spektren wurden mit einer humanen UniProtKB/Swiss-Prot-Datenbank (https://www.uniprot.org) abgeglichen, die 134.921 Proteineinträge (heruntergeladen im Juli 2014) enthielt, ergänzt mit 245 häufig beobachteten Kontaminanten, die mit der

Andromeda-Suchmaschine gesammelt worden waren (Cox et al. 2011). Die Massentoleranzen für Vorläufer- und Fragmentionen wurden auf 6 bzw. 20 ppm eingestellt. Protein-N-terminale Acetylierung und Methioninoxidation wurden als variable Modifikationen zugelassen. Die Carbamidomethylierung von Cystein wurde als feste Modifikation definiert. Die minimale Peptidlänge wurde auf sieben Aminosäuren mit einem Maximum an zwei Fehlspaltungen eingestellt. Die Falschentdeckungsrate (*false detection rate*, FDR) wurde sowohl auf Peptid- als auch auf Proteinebene unter Verwendung einer *forward-and-reverse concatenated decoy database* auf 1% festgelegt. Für die SILAC-Quantifizierung wurde die Multiplizität für die doppelte SILAC-Markierung (Lys +0 / Arg +0, Lys +4 / Arg +6) auf zwei gesetzt. Für die Peptidquantifizierung waren mindestens zwei Erfassungen des Verhältnisses erforderlich. In MaxQuant wurde die Option "*match between runs*" deaktiviert und die Option "*re-quantify*" aktiviert. Die weitere Datenauswertung erfolgte mit der MaxQuant Software (Cox und Mann 2008; Tyanova et al. 2016b).

2.5 Statistische Datenanalyse

Die im folgenden Abschnitt beschriebenen statistischen Datenanalysen wurden von Prof. Dr. Lars Kaderali, Institut für Bioinformatik, Universitätsmedizin Greifswald durchgeführt.

Die Datenanalyse wurde mit R Version 3.4.1 (https://www.r-project.org) und Bioconductor Version 3.2 (Huber et al. 2015) durchgeführt. Alle verfügbaren Proben wurden in die Analyse einbezogen. Proteomdaten wurden wie folgt verarbeitet: Proteine mit mehr als 2/3 fehlenden Werten wurden von der weiteren Analyse ausgeschlossen. Für die verbleibenden 2.586 Proteine wurden fehlende Werte mittels k-*nearest-neighbour*-Imputation mit k = 10 imputiert. Eine univariate Analyse nach dem empirischen Bayes-Ansatz wurde mit dem Bioconductor-Paket *limma* (Ritchie et al. 2015) unter Verwendung der Funktionen lmFit und eBayes durchgeführt. P-Werte wurden unter Verwendung der Falschentdeckungsrate korrigiert. Für Tests mit Vergleich der Mittelwerte wurden die Daten auf Normalität und Ähnlichkeit der Varianzen zwischen den Gruppen bewertet und ein t-Test (gegebenenfalls mit Korrektur für ungleiche Varianzen) oder ein Wilcoxon-Test entsprechend ausgewählt.

Das Training der *support vector machine* (SVM) wurde unter Verwendung von Bootstrapping mit 25 Wiederholungen mit dem Parameter C = 1 in der SVM und einem linearen Kernel durchgeführt. Die Merkmalsauswahl wurde unter Verwendung der *recursive feature elimination* durchgeführt.

2.6 Datenverfügbarkeit

Alle Rohdaten und MaxQuant-Suchergebnisse dieser Studie wurden gemeinsam mit den Ergebnissen weiterführender Analysen (Bohnenberger et al. 2018) im ProteomeXchange-Konsortium (http://proteomecentral.proteomexchange.org/cgi/GetDataset) über das PRIDE-Partner-Repository (Vizcaíno et al. 2014) mit der Datensatz-ID PXD007705 hinterlegt.

3 Ergebnisse

Ziel dieser Studie war es, durch die vergleichende proteomische Analyse von Plattenepithelkarzinomen des Kopf-Hals-Bereichs (*head and neck squamous cell carcinoma*, HNSCC) einerseits und der Lunge (*squamous cell lung carcinoma*, SQCLC) andererseits biologische Unterschiede, mögliche therapeutische Angriffspunkte und Biomarker bzw. proteomische Signaturen zur Unterscheidung zwischen diesen beiden malignen Neoplasien zu identifizieren. Hierzu wurden Proteine aus in Formalin-fixiertem und in Paraffin eingebettetem (*formalin-fixed and paraffin-embedded*, FFPE) Tumorgewebe isoliert und massenspektrometrisch analysiert.

Die in diesem Abschnitt beschriebenen massenspektrometrischen Methoden wurden von der Forschungsgruppe „Bioanalytische Massenspektrometrie" am Max-Planck-Institut für biophysikalische Chemie Göttingen (Prof. Dr. Henning Urlaub) und von der Medizinischen Klinik II, Universitätsklinik Frankfurt am Main (Prof. Dr. Thomas Oellerich) durchgeführt.

Die in diesem Abschnitt beschriebenen statistischen Datenanalysen der massenspektrometrischen Analysen wurden von Prof. Dr. Lars Kaderali, Institut für Bioinformatik, Universitätsmedizin Greifswald durchgeführt.

3.1 Zusammenhang zwischen Gewebevolumen und isolierter Protein- bzw. Peptidmenge

Eine grundlegende Voraussetzung für die proteomische Analyse von HNSCC und SQCLC war die Isolation von Proteinen aus Tumorgewebe. Daher wurde in einem ersten Schritt der Zusammenhang zwischen dem entnommenen Gewebevolumen und der daraus isolierten Proteinmenge untersucht. Hierzu wurde analog zum späteren Vorgehen Tumorgewebe mittels Makrodissektion aus Paraffinblöcken herausgeschnitten (s. 2.2.1). Durch die Zugabe von Lysepuffer, mechanischer Zerkleinerung und Sonifikation wurde das Gewebe aufgeschlossen, die Zellen lysiert und die Proteine isoliert (s. 2.2.2 und 2.2.3). Das entnommene Gewebevolumen wurde durch Messung der Schnittfläche und Multiplikation mit der Schnittdicke berechnet. Die Proteinkonzentrationen der Lysate wurden photometrisch bestimmt (s. 2.2.4) und daraus die in der jeweiligen Probe enthaltene Proteinmenge berechnet. Insgesamt wurden 36 Proben aus HNSCC analysiert, die jedoch nicht Teil des im späteren Verlauf untersuchten Patient*innenkollektivs waren. Drei dieser 36 Proben waren Leervergleiche ohne Gewebe. In den übrigen 33 Proben lag die Schnittfläche im Bereich 8 – 20,25 mm^2, die Schnittdicke im Bereich 10 – 160 µm und das Gewebevolumen im Bereich 0,08 – 3,12 µl.

Die Gesamtmenge an isoliertem Protein aller Proben lag im Bereich 35,0 – 347,6 µg. Aus 1 µl Gewebevolumen konnten durchschnittlich 109 µg Protein gewonnen werden.

Abbildung 1 stellt den Zusammenhang zwischen Gewebevolumen in µl und Proteinmenge in µg grafisch dar. Es zeigte sich ein linearer Zusammenhang zwischen entnommenem Gewebevolumen und daraus isolierter Proteinmenge.

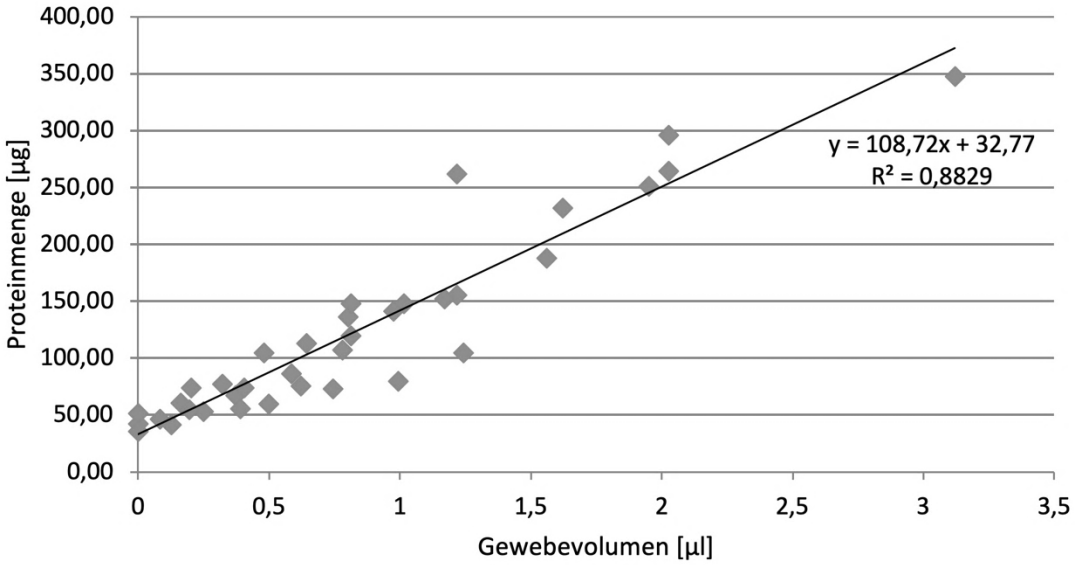

Abbildung 1: Zusammenhang zwischen Gewebevolumen und Proteinmenge. Proteinmenge nach Lyse der Zellen in Abhängigkeit vom Gewebevolumen, das den Paraffinblöcken entnommen worden war.

Des Weiteren wurde untersucht, welche Menge an Peptid nach der weiteren Aufarbeitung der Proteine mittels der filtergestützten Probenaufbereitung (*filter-aided sample preparation*, FASP) noch vorhanden war. Dazu wurden insgesamt 70 Proben mit jeweils 100 µg Protein entsprechend des FASP-Protokolls (s. 2.2.3) aufbereitet und anschließend die Peptidmenge bestimmt (s. 2.2.5). Auffällig war dabei eine starke Streuung der Peptidmenge, die in Abbildung 2 veranschaulicht wird. Das Minimum betrug 1,7 µg, das Maximum 62,6 µg. Der Median lag bei 28,4 µg, das 25%-Quartil bei 18,7 µg, das 75%-Quartil bei 40,1 µg. In 9 (12,9%) der 70 Proben betrug die Peptidmenge <10 µg und damit weniger als 10% der ursprünglichen Proteinmenge. Es erschien daher sinnvoll, die Peptidmenge nach Durchlaufen des FASP-Protokolls zu bestimmen, um sicherzustellen, dass ausreichend Peptid für die anschließende massenspektrometrische Analyse vorhanden war.

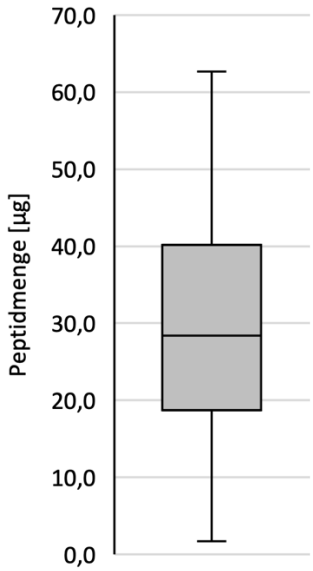

Abbildung 2: Aus 100 µg Protein gewonnene Peptidmenge. Der Boxplot veranschaulicht die Peptidmenge, die in 70 Proben aus 100 µg Protein gewonnen werden konnte. Die mittlere Linie innerhalb der Boxen entspricht dem Median. Die obere bzw. untere Begrenzung der Box repräsentiert das 75%- bzw. 25%-Quartil. Die Whisker zeigen Minimum und Maximum.

3.2 Evaluation des Proteinquantifizierungsstandards

Es existieren mehrere Techniken, um die Proteine massenspektrometrisch zu quantifizieren, wobei das *stable isotope labeling by/with amino acids in cell culture* (SILAC) eine weit verbreitete und gut etablierte Methodik darstellt (s. 1.2.1). Im Rahmen dieser Arbeit wurde ein sogenannter Super- bzw. SpikeIn-SILAC-Proteinquantifizierungsstandard, eine Weiterentwicklung des klassischen SILAC-Verfahrens, eingesetzt. Hierbei handelt es sich um ein mit nicht radioaktiven Isotopen markiertes, in Zellkultur gewonnenes Proteingemisch, das den zu untersuchenden Tumorproteinen als interner Standard hinzugegeben wird und somit die relative Quantifizierung der Proben untereinander ermöglicht.

Zur Herstellung dieses Standards (s. 2.3.4) wurden zunächst die natürlichen, essentiellen Aminosäuren Arginin und Lysin in den Proteomen der vier Lungenkarzinomzelllinien NCI-H2228, NCI-H1339, HCC-15 und HCC-44 durch die Isotop-markierten Aminosäuren $^{13}C_6$-Arginin und D_4-Lysin ersetzt und somit ein Massenunterschied von +6 U (Arginin) bzw. +4 U (Lysin) im Vergleich zu den natürlichen Varianten erzeugt. Im Anschluss daran wurden die Zellen dieser vier Zelllinien lysiert, die jeweiligen Proteinkonzentrationen bestimmt und im Mengenverhältnis 1:1:1:1 zum Proteinquantifizierungsstandard gemischt.

Um eine exakte Quantifizierung zu ermöglichen, sollte der Proteinquantifizierungsstandard die Proteome der zu untersuchenden Gewebe möglichst gut repräsentieren. Um dies zu evaluieren wurde das Tumorgewebe je einer HNSCC- bzw. SQCLC-Probe

massenspektrometrisch analysiert. Hierfür wurde zunächst unter mikroskopischer Kontrolle ein Bereich mit einer Tumorzelldichte von >80% markiert und mittels Makrodissektion das entsprechende FFPE-Gewebe entnommen. Hieraus wurden die Proteine isoliert und im Mengenverhältnis von 1:1 mit dem super-SILAC-Standard gemischt. Dieses Proteingemisch wurde mithilfe der Filter-unterstützen Probenaufbereitung (*filter-aided sample preparation*, FASP, s. 2.2.3) aufgereinigt, durch die Protease Trypsin zu Peptiden gespalten und massenspektrometrisch analysiert.

Abbildung 3 zeigt die Histogramme der Verteilung der SILAC-Ratios der beiden analysierten Proben. Die SILAC-Ratio stellt dabei das Verhältnis der massenspektrometrisch gemessenen Proteinintensität in der Probe im Vergleich zur Intensität desselben Proteins im Proteinquantifizierungsstandard dar (SILAC-Ratio = $\log_2$ (Proteinintensität in der Probe/Proteinintensität im Proteinquantifizierungsstandard)). Die Unterscheidung zwischen den Peptiden/Proteinen der Probe einerseits und des Proteinquantifizierungsstandards andererseits gelingt dabei durch den definierten Massenunterschied, welcher durch die isotopische Markierung im Proteinquantifizierungsstandard erzeugt wurde.

Da bei beiden Proben >90% der gemessenen Proteine eine $\log_2$ SILAC-Ratio zwischen -2 und +2 aufwiesen, konnte von einer guten Repräsentation der Proteome der zu untersuchenden Gewebe durch den Proteinquantifizierungsstandard ausgegangen werden.

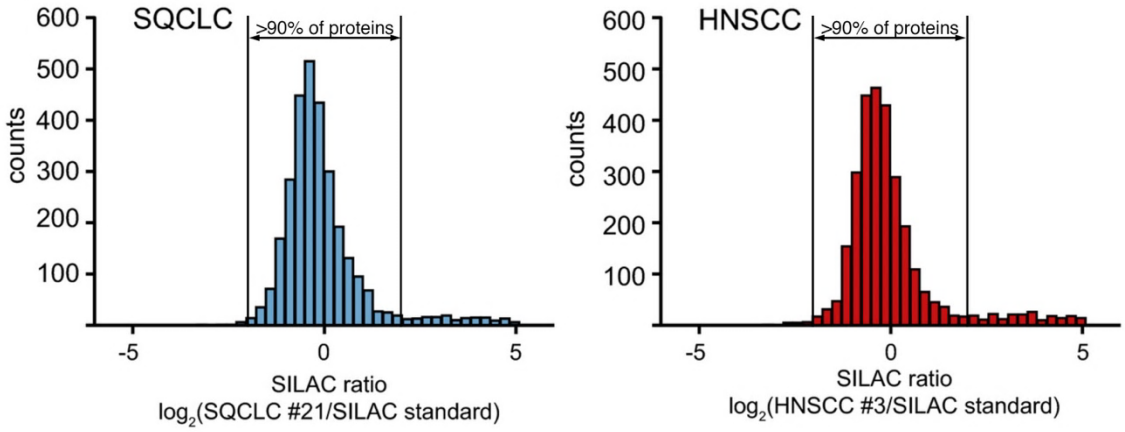

Abbildung 3: Evaluation des Proteinquantifizierungsstandards. Die Histogramme zeigen die Verteilung der SILAC-Ratios der gemessenen Proteine in einer SQCLC- (links) bzw. HNSCC-Probe (rechts). Modifiziert nach Bohnenberger et al. (2018), CC BY 4.0

3.3 Zusammensetzung des Patient*innenkollektivs

Nachdem die technischen Voraussetzungen für die vergleichende proteomische Analyse vorlagen, wurde ein Kollektiv bestehend aus insgesamt 74 Patient*innen identifiziert. Davon waren 44 Patient*innen an einem primären Plattenepithelkarzinom der Lunge (SQCLC) und

30 Patient*innen an einem primären Plattenepithelkarzinom des Kopf-Hals-Bereichs (HNSCC) erkrankt. Bei den 30 HNSCC-Patient*innen kam es im Verlauf der Erkrankung zusätzlich zu einem Lungentumor mit Plattenepithelkarzinom-Histologie. Alle Patient*innen erhielten eine chirurgische Therapie und 40 zusätzlich eine adjuvante (Chemo)Radiotherapie. Keine*r der Patient*innen erhielt eine neoadjuvante Therapie oder primäre Chemotherapie. Drei der HNSCC-Proben waren positiv für HPV-16. Durch die mikroskopische Einschätzung eines erfahrenen Pathologen sowie die immunhistochemische Färbung (vgl. exemplarisch Abbildung 4) von p63 und CK5/6 (Cytokeratin) wurde bestätigt, dass es sich bei allen 74 Tumoren histologisch um Plattenepithelkarzinome handelte. Tabelle 8 gibt einen Überblick über die Zusammensetzung des Patient*innenkollektivs.

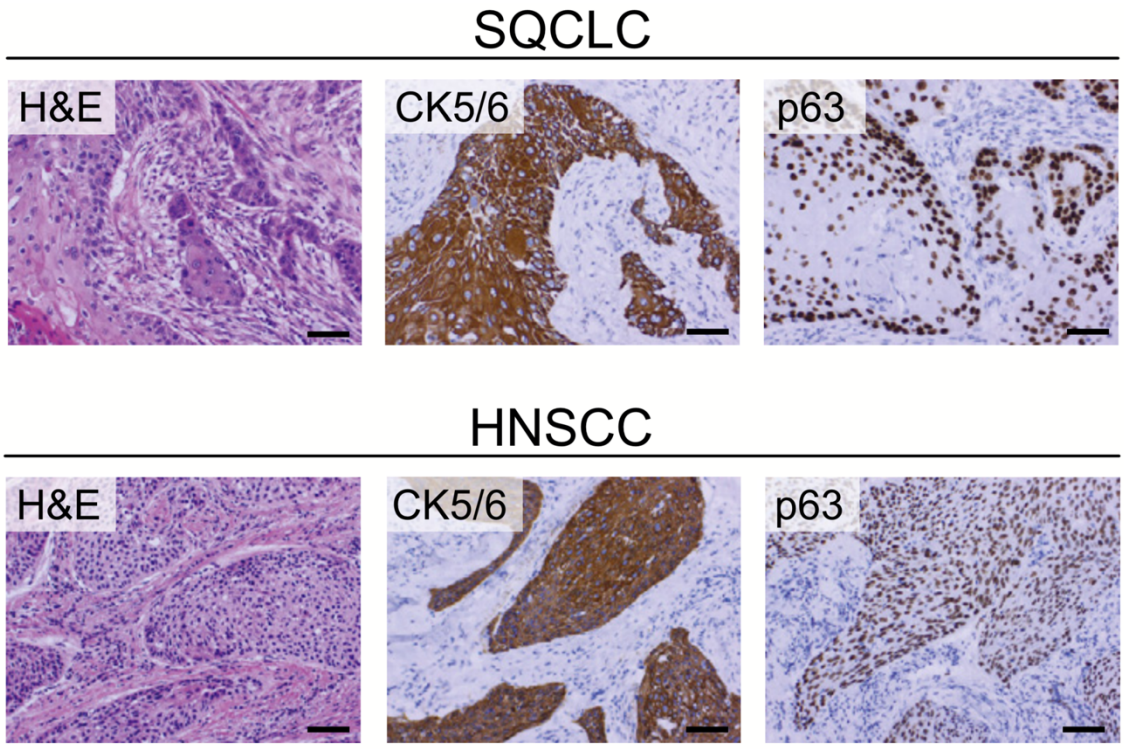

Abbildung 4: Immunhistochemische Färbung von CK5/6 und p63. Die Abbildung zeigt exemplarisch an je einer SQCLC- bzw. HNSCC-Probe die HE(Hämatoxylin-Eosin)-Färbung sowie die immunhistochemische Anfärbung von CK5/6 und p63, welche zur Bestätigung der Plattenepithelkarzinomhistologie genutzt wurden. Modifiziert nach Bohnenberger et al. (2018), CC BY 4.0

Alle Patient*innen waren an der Universitätsmedizin Göttingen in Behandlung. Das untersuchte Tumorgewebe entstammte den FFPE-Geweben, welche im Anschluss an die chirurgische Therapie durch das Pathologische Institut im Rahmen der normalen klinischen Routine zur histologischen Diagnostik eingebettet worden waren.

Tabelle 8: Zusammensetzung des Patient*innenkollektivs.

Merkmal	**SQCLC (n = 44)**	**HNSCC (n = 30)**
Alter: Median (Bereich) [Jahre]	66 (49-81)	55 (31-76)
Geschlecht männlich [Anzahl (%)]	32 (72,7)	26 (86,7)
Rauchen bzw. Tabakkonsum [Anzahl (%)]		
Aktuell oder früher	22 (50,0)	28 (93,3)
Nie	0 (0,0)	2 (6,7)
Unbekannt	22 (50,0)	0 (0,0)
Lokalisation des Primärtumors [Anzahl (%)]		
Lunge	44 (100,0)	-
Mundhöhle	-	5 (16,7)
Pharynx	-	13 (43,3)
Larynx	-	12 (40,0)
Staging UICC (7te Edition) [Anzahl (%)]		
I	14 (31,8)	1 (3,3)
II	11 (25,0)	5 (16,7)
III	19 (43,2)	7 (23,3)
IV	0 (0,0)	17 (56,7)
Systemisches Therapieregime [Anzahl (%)]		
Neoadjuvant	0 (0,0)	0 (0,0)
Adjuvant	22 (50,0)	18 (60,0)
Generelles Überleben: Mediane *Follow-up*-Zeit (Bereich) [Monate]	19 (1-38)	57 (7-169)
Todesfälle [Anzahl (%)]	17 (31,8)	20 (66,7)
Lokalrezidiv der HNSCC [Anzahl (%)]	-	5 (16,7)
Zeit von HNSCC bis Lungentumor: Median (Bereich) [Monate]	-	20 (1-121)
TP53 Mutationstatus [Anzahl (%)]		
Wildtyp	12 (27,3)	12 (40,0)
Mutiert	31 (70,5)	12 (40,0)
Unbekannt	1 (2,2)	6 (20,0)
HPV-Status [Anzahl (%)]		
Negativ	-	27 (90,0)
Positiv	-	3 (10,0)

3.4 Anzahl quantifizierter Proteine und technische Reproduzierbarkeit der massenspektrometrischen Analyse

Analog zum Vorgehen bei der Evaluation des Proteinquantifizierungsstandards wurden die insgesamt 44 SQCLC- und 30 HNSCC-Proben aufgearbeitet und die Proteine massenspektrometrisch analysiert. Es konnten insgesamt 6.214 Proteine quantifiziert werden. Der Mittelwert der pro Probe quantifizierten Proteine betrug 2.251. Das Minimum an quantifizierten Proteinen lag bei 1.317, das Maximum bei 3.314 (vgl. Abbildung 5).

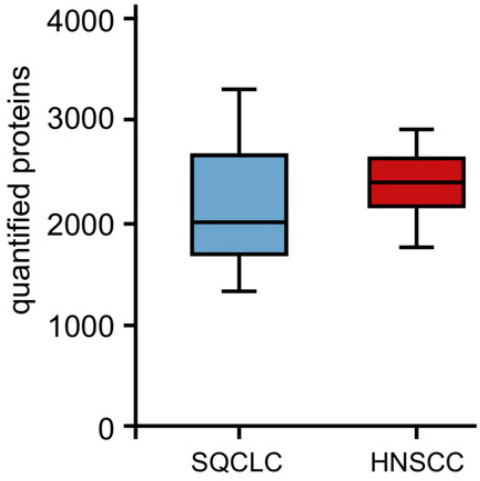

Abbildung 5: Anzahl quantifizierter Proteine. Die Boxplots veranschaulichen die Zahl der quantifizierten Proteine der 44 SQCLC- (links) und 30 HNSCC-Proben (rechts). Die mittlere Linie innerhalb der Boxen entspricht dem Median. Die obere bzw. untere Begrenzung der Box repräsentiert das 75%- bzw. 25%-Quartil. Die Whisker zeigen Minimum und Maximum. Modifiziert nach Bohnenberger et al. (2018), CC BY 4.0

Zudem wurde die Reproduzierbarkeit der massenspektrometrischen Analysen untersucht. Jede Probe wurde in zwei Replikaten massenspektrometrisch analysiert. Für jede Probe wurde eine Pearson-Korrelationsanalyse zwischen den SILAC-Ratios aller quantifizierten Proteine der beiden Replikate durchgeführt (exemplarisch in Abbildung 6 dargestellt). Die Korrelationskoeffizienten lagen dabei in einem Bereich von 0,802 – 0,975 (MW: 0,949 ± 0,0266).

Die massenspektrometrisch gemessenen Proteinintensitäten aller 74 Proben umfassten eine Größenordnung von insgesamt etwa sechs Zehnerpotenzen, was die Identifikation von Proteinen mit starker bzw. schwacher Expression widerspiegelt. Die Intensitäten von 90% der Proteine lag in einem Bereich von etwa vier Zehnerpotenzen (vgl. Abbildung 7).

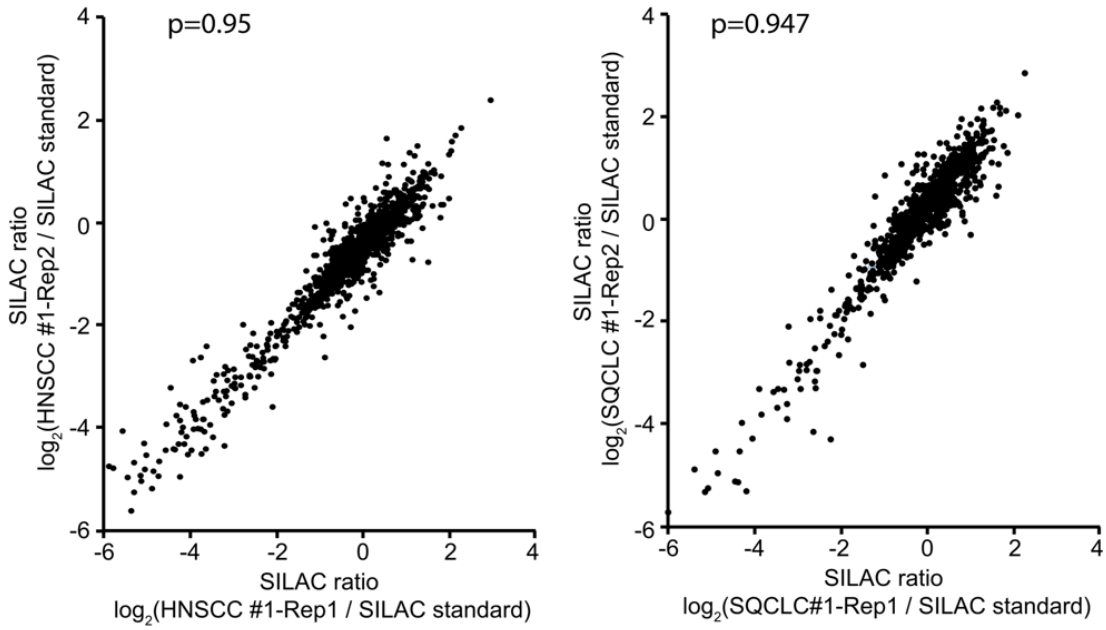

Abbildung 6: Technische Reproduzierbarkeit der massenspektrometrischen Analyse. Die x- bzw. y-Achse zeigt die SILAC-Ratio aller quantifizierten Proteine zweier Replikate (Rep1 bzw. Rep2) je einer HNSCC- (links) und SQCLC-Probe (rechts) mit Angabe des Pearsonschen Korrelationskoeffizienten (p). Modifiziert nach Bohnenberger et al. (2018), CC BY 4.0

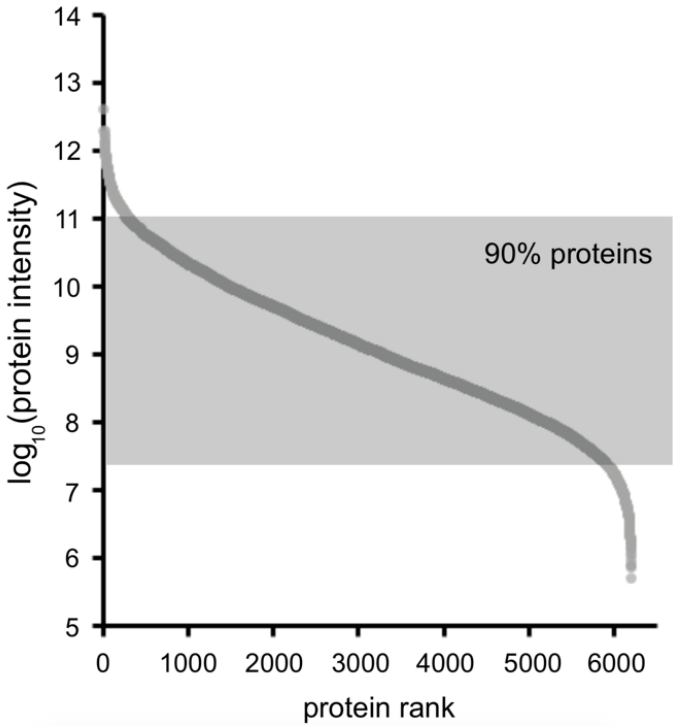

Abbildung 7: Gemessene Proteinintensitäten. Die Abbildung zeigt die Verteilung der massenspektrometrisch gemessenen Proteinintensitäten aller 74 Proben. Modifiziert nach Bohnenberger et al. (2018), CC BY 4.0

3.5 Proteomische Unterscheidung von SQCLC und HNSCC

Ein zentrales Ziel dieser Arbeit war die bessere Unterscheidung von SQCLC und HNSCC. Auf Basis des gesamten durch die massenspektrometrische Analyse der 74 Proben gewonnenen proteomischen Datensatzes wurde eine Hauptkomponentenanalyse (vgl. Abbildung 8) der Proteinexpression in SQCLC bzw. HNSCC durchgeführt. Diese zeigte signifikante Unterschiede der Proteinexpression, indem sie die Proben entsprechend ihres jeweiligen Ursprungsgewebes in zwei verschiedene, sich partiell überlappende Gruppen trennte.

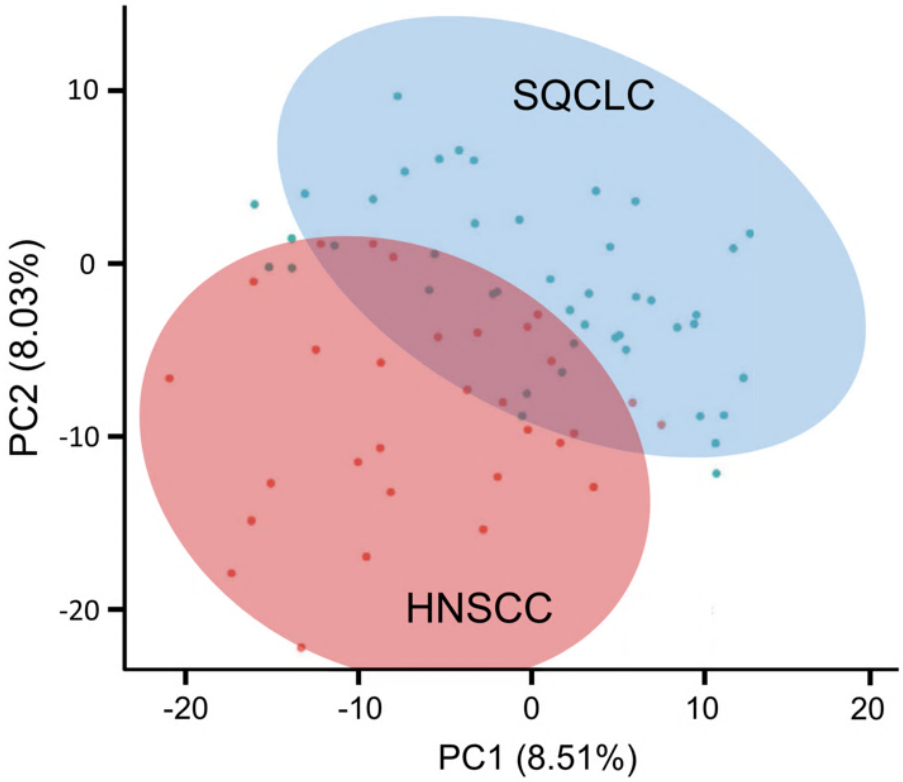

Abbildung 8: Hauptkomponentenanalyse der Proteinexpression von SQCLC und HNSCC. Es wurden die ersten beiden Hauptkomponenten (PC = *principal component*) aufgetragen, die für 8,51% bzw. 8,03% der gesamten Varianz verantwortlich sind. Modifiziert nach Bohnenberger et al. (2018), CC BY 4.0

Um eine diagnostische Signatur mit hoher Klassifizierungsgenauigkeit zur Unterscheidung von SQCLC und HNSCC zu erhalten, wurde als nächstes eine lineare *support vector machine* (SVM) trainiert. Eine SVM ist eine Methode des *Supervised Learnings*, einer Unterform des maschinellen Lernens. Dabei wird ein Trainingsdatensatz mit bekannten Merkmalen und bekannter Gruppenzugehörigkeit (hier die Proteinexpression der 44 SQCLC und 30 HNSCC) genutzt, um Regeln aus diesen Daten zu extrahieren, welche die Gruppenzugehörigkeit (hier SQCLC oder HNSCC) bestmöglich unterscheiden bzw. vorhersagen können. Das Training wurde unter Verwendung eines Bootstrapping-Verfahrens durchgeführt, um die Robustheit der Vorhersagen zu bewerten und vor Überanpassung zu schützen. Unter Verwendung des

vollständigen Satzes an Proteinen als Merkmale in einem Kreuzvalidierungsansatz erreichte der Prädiktor eine Bootstrap-Genauigkeit von 90,5% ± 0,05, eine Sensitivität von 85,7% und eine Spezifität von 93,7%. Anschließend wurde untersucht, inwieweit sich eine Verringerung der Größe des Signatursatzes auf die Klassifizierungsleistung auswirken würde. Unter Verwendung der *recursive feature elimination* konnte der Satz auf etwa 1.100 Proteine reduziert werden, ohne die Klassifizierungsleistung zu beeinträchtigen. Danach begann die Genauigkeit der Klassifizierung allmählich abzunehmen, mit einer Bootstrap-Genauigkeit von 76,8% unter Verwendung von 100 Proteinen, 65,0% unter Verwendung von 25 Proteinen und nur 57,2% unter Verwendung eines einzelnen Proteins.

3.6 Unterschiedlich exprimierte Proteine zwischen HNSCC und SQCLC

In allen 74 Proben konnten insgesamt 518 Proteine identifiziert werden, die eine signifikant unterschiedliche Expression zwischen HNSCC und SQCLC aufwiesen. Dazu wurde eine univariate Datenanalyse unter Verwendung einer empirischen Bayes-Methode mit linearem Regressionsmodell verwendet. Als Signifikanzniveau wurde $p<0,05$ festgelegt.

Abbildung 9 veranschaulicht die unterschiedliche Proteinexpression anhand eines *volcano plots*, wobei jeder Punkt ein Protein repräsentiert. Proteine, die im Vergleich zur jeweils anderen Gruppe (HNSCC bzw. SQCLC) signifikant und mindestens um den Faktor $\log_2(0,5)$ stärker exprimiert waren, sind farblich hervorgehoben. Rot markierte Proteine waren in HNSCC stärker exprimiert als in SQCLC, blau markierte waren in SQCLC stärker exprimiert als in HNSCC. Die jeweils 10 Proteine, welche mit der höchsten Signifikanz in HNSCC bzw. SQCLC hochreguliert waren, werden in Tabelle 9, Tabelle 10 und Abbildung 10 dargestellt.

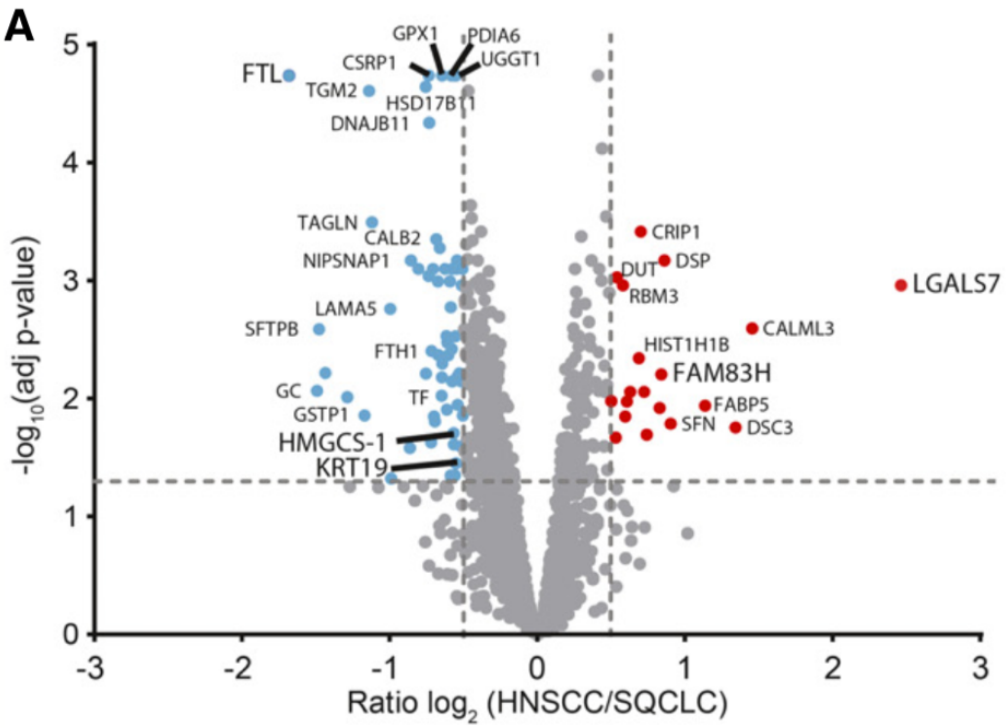

Abbildung 9: In HNSCC bzw. SQCLC signifikant stärker exprimierte Proteine. Jeder Punkt repräsentiert ein Protein. Die x-Achse zeigt die $\log_2$ SILAC-Ratio zwischen HNSCC und SQCLC der beiden gemessenen Replikate. Die y-Achse zeigt die $-\log_{10}$ adjustierten p-Werte der unterschiedlichen Proteinexpression. Die gestrichelte, horizontale Linie markiert das Signifikanzniveau von $p = 0{,}05$ ($-\log_{10} \approx 1{,}3$). Blau (oben links) markierte Proteine waren in SQCLC signifikant und mindestens 1,41-fach ($\log_2 \approx -0{,}5$ bzw. $+0{,}5$; vertikale gestrichelte Linien) stärker exprimiert als in HNSCC, rot (oben rechts) markierte in HNSCC stärker als in SQCLC. Modifiziert nach Bohnenberger et al. (2018), CC BY 4.0

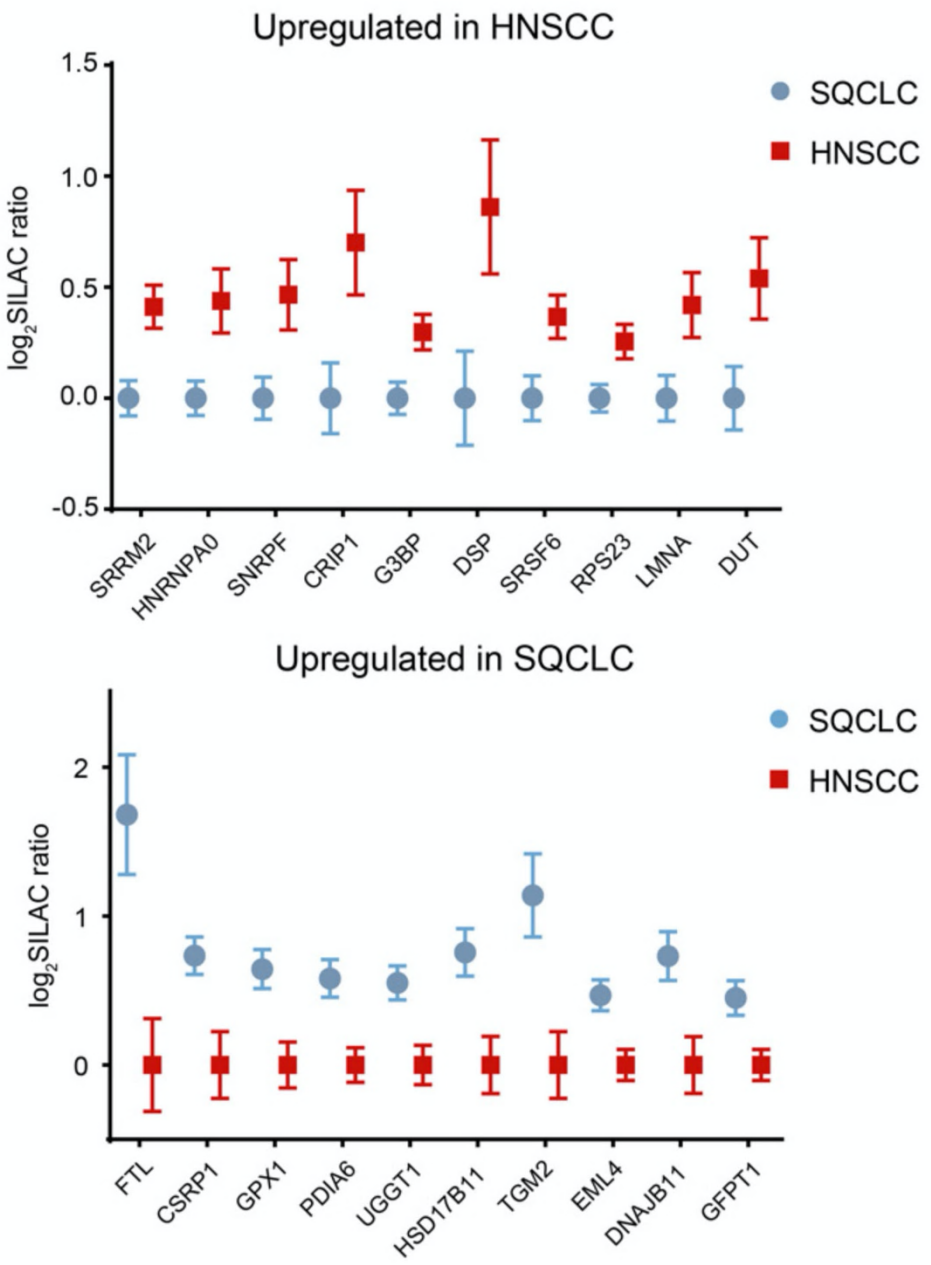

Abbildung 10: Top 10 der hochregulierten Proteine. Die Abbildung zeigt die 10 Proteine, die mit der höchsten Signifikanz in HNSCC (oben) bzw. SQCLC (unten) hochreguliert waren. Die Kreise/Quadrate zeigen den Mittelwert, die Whisker das 95%-Konfidenzintervall. Modifiziert nach Bohnenberger et al. (2018), CC BY 4.0

Tabelle 9: Top 10 der in HNSCC hochregulierten Proteine.

Genname	Proteinname	UniprotID	MW log_2 FC HNSCC/SQCLC	p-Wert
SRRM2	*Serine/arginine repetitive matrix protein 2*	Q9UQ35	0,41	0,00002
HNRNPA0	*Heterogeneous nuclear ribonucleoprotein A0*	Q13151	0,44	0,00008
SNRPF	*Small nuclear ribonucleoprotein F*	P62306	0,47	0,0003
CRIP1	*Cysteine-rich protein 1*	P50238	0,70	0,0004
G3BP	*Ras GTPase-activating protein-binding protein 1*	Q5U0Q1	0,30	0,0004
DSP	*Desmoplakin*	P15924	0,86	0,0007
SRSF6	*Serine/arginine-rich splicing factor 6*	Q59GY3	0,37	0,0007
RPS23	*40S ribosomal protein S23*	A8K517	0,26	0,0008
LMNA	*Prelamin-A/C*	P02545-2	0,42	0,0008
DUT	*Deoxyuridine 5-triphosphate nucleotidohydrolase, mitochondrial*	P33316	0,54	0,0009

MW = Mittelwert, FC = *fold change*

Tabelle 10: Top 10 der in SQCLC hochregulierten Proteine.

Genname	Proteinname	UniprotID	MW log_2 FC HNSCC/SQCLC	p-Wert
FTL	*Ferritin light chain*	P02792	-1,68	0,00002
CSRP1	*Cysteine and glycine-rich protein 1*	B4DY28	-0,73	0,00002
GPX1	*Glutathione peroxidase*	Q7L4Q3	-0,64	0,00002
PDIA6	*Protein disulfide-isomerase A6*	B5MCQ5	-0,58	0,00002
UGGT1	*UDP-glucose:glycoprotein glucosyltransferase 1*	Q9NYU2-2	-0,55	0,00002
HSD17B11	*Estradiol 17-beta-dehydrogenase 11*	Q8NBQ5	-0,76	0,00002
TGM2	*Protein-glutamine gamma-glutamyltransferase 2*	P21980	-1,14	0,00002
EML4	*Echinoderm microtubule-associated protein-like 4*	Q9HC35	-0,47	0,00002
DNAJB11	*DnaJ homolog subfamily B member 11*	Q9UBS4	-0,73	0,00005
GFPT1	*Glutamine--fructose-6-phosphate aminotransferase [isomerizing] 1*	Q06210-2	-0,45	0,0002

MW = Mittelwert, FC = *fold change*

4 Diskussion

Bei Patient*innen, die an einem Plattenepithelkarzinom des Kopf-Hals-Bereichs (*head and neck squamous cell carcinoma*, HNSCC) erkrankt sind, kann im Verlauf sowohl eine Lungenmetastase dieses Plattenepithelkarzinoms (metHNSCC) als auch ein metachrones Primärkarzinom der Lunge (*squamous cell lung carcinoma*, SQCLC) auftreten. Die Unterscheidung zwischen diesen beiden Entitäten ist klinisch von Bedeutung für die Wahl der entsprechenden Therapie, ob in palliativer oder kurativer Intention behandelt wird. Aufgrund großer morphologischer und genomischer Ähnlichkeiten zwischen diesen beiden Tumorentitäten kommen aktuell verschiedene klinische und bildgebende Unterscheidungskriterien (z. B. Histologie, Grading, TNM-Stadium, Zeitintervall zwischen HNSCC und Auftreten des Lungentumors, Anzahl/Ort der Lungentumore) zum Einsatz, die jedoch wenig validiert und nicht einheitlich festgelegt sind und so keine abschließend sichere Einschätzung erlauben, ob es sich um eine Lungenmetastase oder ein metachrones Primärkarzinom der Lunge handelt (Kuriakose et al. 2002; Geurts et al. 2005; Ichinose et al. 2016a).

In dieser Studie wurde daher eine vergleichende proteomische Analyse der Tumorgewebe von insgesamt 30 HNSCC-Patient*innen, bei denen im Verlauf ein Lungentumor mit Plattenepithelkarzinomhistologie auftrat, und 44 SQCLC-Patient*innen durchgeführt. Sie ist damit zum aktuellen Zeitpunkt die erste und einzige vergleichende massenspektrometrische Proteomanalyse von HNSCC und SQCLC an menschlichem Tumorgewebe. Diese Arbeit konnte eine proteomische Signatur identifizieren, welche zur zuverlässigen Unterscheidung von HNSCC und SQCLC geeignet ist und stellt außerdem einen umfassenden proteomischen Datensatz zum Vergleich dieser Tumorentitäten als Basis für weitere Untersuchungen zur Verfügung.

4.1 Geringe Gewebemengen sind ausreichend für eine massenspektrometrische Proteomanalyse

Vor der Analyse der Tumorgewebeproben wurde der Zusammenhang zwischen dem aus den FFPE-Blöcken entnommenen Gewebevolumen und der daraus isolierten Proteinmenge untersucht. Die isolierte Proteinmenge stieg dabei wie erwartet linear mit dem Gewebevolumen an. Aus 1 µl Gewebevolumen konnten durchschnittlich 109 µg Protein isoliert werden. Da pro 100 µg Protein nach der filter-gestützten Probenaufbereitung (*filter-aided sample preparation*, FASP) durchschnittlich noch 28 µg Peptid vorhanden waren und für die

massenspektrometrische Analyse nur wenige µg Peptid notwendig sind, kann bereits 1 µl FFPE-Gewebevolumen als ausreichend für eine proteomische Analyse erachtet werden.

In den 74 analysierten Tumorgewebeproben konnten insgesamt 6.214 Proteine quantifiziert werden. Pro Probe wurden im Mittel 2.251, mindestens 1.317 und maximal 3.314 Proteine quantifiziert. Diese Werte sind damit vergleichbar mit den Ergebnissen ähnlich konzipierter Studien an anderen humanen Geweben, bei denen in einem einzelnen massenspektrometrischen Durchlauf ebenfalls mehrere Tausend Proteine identifiziert und quantifiziert werden konnten (Wiśniewski et al. 2011; Deeb et al. 2012; Bohnenberger et al. 2015).

4.2 Komplexe proteomische Signaturen ermöglichen die Unterscheidung von HNSCC und SQCLC

Mithilfe einer Hauptkomponentenanalyse des gesamten proteomischen Datensatzes konnten HNSCC und SQCLC entsprechend ihres jeweiligen Ursprungs in zwei verschiedene, sich teilweise überlappende Gruppen aufgetrennt werden. Die proteomische Unterscheidung von HNSCC und SQCLC ist somit prinzipiell möglich. Um eine diagnostische Signatur mit hoher Klassifizierungsgenauigkeit zu erhalten, wurde auf Basis des gewonnen proteomischen Datensatzes eine lineare *support vector machine* (SVM) auf die Unterscheidung zwischen HNSCC und SQCLC trainiert. Dabei konnte in einem Kreuzvalidierungsverfahren eine Vorhersagegenauigkeit von 90,5%, eine Sensitivität von 85,7% und eine Spezifität von 93,7% erzielt werden. Um diese Genauigkeit zu erreichen, war eine Signaturgröße von etwa 1.100 Proteinen notwendig. Mit abnehmender Anzahl der berücksichtigten Proteine, nahm auch die Vorhersagegenauigkeit ab und lag bei 76,8% unter Verwendung von 100 Proteinen, bei 65,0% unter Verwendung von 25 Proteinen und nur bei 57,2% unter Verwendung eines einzelnen Proteins. Die Unterscheidung von HNSCC und SQCLC ist mithilfe von Proteinsignaturen also zuverlässig möglich. Allerdings müssen hierfür komplexe Proteinsignaturen herangezogen werden, während einzelne Proteine bzw. Biomarker nicht ausreichend geeignet sind.

Vergleicht man diese Arbeit mit der bisherigen Literatur, so muss festgestellt werden, dass bislang vor allem genetische und transkriptomische Analysen zur Unterscheidung von HNSCC und SQCLC durchgeführt wurden, während nur eine Studie zu Unterschieden in der Proteinexpression existiert. So wurde unter anderem die Differenzierung dieser beiden Tumorentitäten anhand des *loss of heterozygosity*, *TP53*-Mutationstatus, *copy number alterations*, DNA-Methylierung oder der Genexpression untersucht (Geurts et al. 2005; 2009; Talbot et al. 2005; Vachani et al. 2007; Campbell et al. 2018; Jurmeister et al. 2019). Auch in diesen

Studien bestätigte sich die hier gewonnene Erkenntnis, dass einzelne Biomarker zur Unterscheidung nicht ausreichend sind, sondern komplexere Signaturen benötigt werden.

In einer von Talbot et al. (2005) durchgeführten Genexpressionsanalyse an 18 SQCLC und 31 HNSCC konnte mithilfe einer SVM eine 500 Gene umfassende Signatur identifiziert werden, die in der Lage war, HNSCC und SQCLC der Trainingskohorte zuverlässig zu unterscheiden. Wurden Signaturen mit weniger als 100 Genen genutzt, nahm die Vorhersagegenauigkeit ab (Talbot et al. 2005). Zur Unterscheidung von metHNSCC und SQCLC wurde die Signatur zudem an zwölf Lungentumoren unbekannten Ursprungs bei Patient*innen mit HNSCC in der Vorgeschichte getestet. Hierbei gelang es, alle Proben zuverlässig in eine der beiden Gruppen einzuordnen (Talbot et al. 2005). Allerdings kam hier nur die geringe Zahl von zwölf Lungentumoren zum Einsatz, wovon es sich nur in einem Fall um ein metHNSCC handelte. Ähnlich ist eine von Vachani et al. (2007) durchgeführte Genexpressionsanalyse an zehn SQCLC und 18 HNSCC zu werten, bei der eine zehn Gene umfassende Signatur erfolgreich identifiziert werden konnte. Auch hier zeigte sich eine deutliche Abnahme der Vorhersagegenauigkeit, wenn weniger als zehn Gene genutzt wurden (Vachani et al. 2007). Die Signatur konnte ebenfalls erfolgreich an Lungentumoren unbekannten Ursprungs validiert werden, allerdings handelte es sich hierbei um dieselben zwölf Tumore wie in der Studie von (Talbot et al. 2005).

Zu Unterschieden in der Proteinexpression existiert lediglich eine Studie von Ichinose et al. (2016b). Dabei wurde ein diagnostischer Algorithmus zur Unterscheidung von metHNSCC und SQCLC auf der Basis der immunhistochemischen Analyse von CK19, MMP3 und PI3 an 48 HNSCC und 39 SQCLC entwickelt (Ichinose et al. 2016b). Waren die Tumore CK19 positiv, MMP3 negativ und PI3 negativ wurden sie als SQCLC gewertet, andernfalls als HNSCC (Ichinose et al. 2016b). Allerdings mussten bestimmte Tumore ausgeschlossen werden, um eine Genauigkeit von 85% in der Trainings- und 82% in der Validierungskohorte (bestehend aus 23 HNSCC und 32 SQCLC) zu erzielen (Ichinose et al. 2016b). Der Algorithmus konnte auch an 28 Lungentumoren unbekannten Ursprungs erfolgreich validiert werden und dabei das Rezidivrisiko besser vorhersagen als gängige klinisch-histopathologische Merkmale (Ichinose et al. 2016b). Analog dazu war CK19 auch in der SQCLC-Kohorte der hier vorliegenden Arbeit deutlich hochreguliert ($\log_2$ FC HNSCC/SQCLC = -0,99; p = 0,048). MMP3 und PI3 konnten nicht bzw. nicht in ausreichend vielen Tumorgewebeproben identifiziert werden, um eine Quantifizierung zu ermöglichen.

Auch wenn die hier vorliegende Arbeit die bislang erste und einzige ist, in der die massenspektrometrische Proteomanalyse an HNSCC und SQCLC angewendet wurde, konnte diese

Methode bereits erfolgreich zur Klassifizierung von DLBCL-Subtypen eingesetzt werden, die histomorphologisch nicht zu differenzieren sind (Deeb et al. 2012). Mithilfe einer SVM konnte auf Basis der proteomischen Analyse von 20 menschlichen Gewebeproben eine aus vier Proteinen bestehende Signatur identifiziert werden, welche die Unterscheidung der DLBCL-Subtypen ermöglichte (Deeb et al. 2015). Immunhistochemische Analysen von lediglich drei Proteinen waren für die Differenzierung jedoch nicht ausreichend (Benesova et al. 2013). Bei genetischen Untersuchungen waren die Expressionsprofile mehrerer hundert Gene für die Unterscheidung notwendig (Alizadeh et al. 2000; Rosenwald et al. 2002).

Zur Differenzierung zweier Tumorentitäten mit hoher Ähnlichkeit sind nicht nur komplexe genomische oder proteomische Signaturen, sondern auch Methoden des maschinellen Lernens notwendig, wie sie auch in der hier vorliegenden Arbeit und einigen der erwähnten Studien über die SVM Anwendung fanden (Talbot et al. 2005; Vachani et al. 2007; Deeb et al. 2015). Unterstützt wird dies durch eine Studie von Jurmeister et al. (2019), in der ein künstliches neuronales Netzwerk auf der Basis der Analyse von DNA-Methylierung zur Unterscheidung von HNSCC und SQCLC trainiert wurde. Dieses neuronale Netzwerk war in einer großen Validierungskohorte bestehend aus insgesamt 279 Proben (110 HNSCC, 150 SQCLC, 19 normales Lungengewebe) in der Lage, die Gewebeproben mit einer hohen Genauigkeit von 96,4% korrekt zu klassifizieren (Jurmeister et al. 2019). Diese hohe Genauigkeit konnte durch das künstliche neuronale Netzwerk und eine SVM als Methoden des maschinellen Lernens erzielt werden, während einfachere Methoden, wie die hierarchische Clusteranalyse nicht in der Lage waren, HNSCC und SQCLC zu unterscheiden (Jurmeister et al. 2019).

4.3 Identifikation und Quantifizierung bereits bekannter onkogener Proteine

Im Rahmen der massenspektrometrischen Analyse konnten die HMGCS-1 (Hydroxymethylglutaryl-CoA-Synthase) und Proteine der FAM83-Familie (FAM83A-H) identifiziert und quantifiziert werden, welche bereits früher als onkogene Faktoren beschrieben worden waren (Cipriano et al. 2012; Nalla et al. 2016; Zhao et al. 2017). Die FAM83-Proteine spielen dabei eine Rolle in Verbindung mit EGFR, RAS und MAPK (Cipriano et al. 2012). Bei HNSCC wurde ein Zusammenhang zwischen einer erhöhten FAM83H-Expression und einer erhöhten Sensitivität für eine Therapie mit EGFR-Inhibitoren beobachtet (Lepikhova et al. 2018). Dies könnte möglicherweise eine Erklärung dafür sein, dass in einer weiteren Studie die verminderte Expression von FAM83H in Neoplasien des Kopf-Hals-Bereichs mit einem

schlechteren Überleben assoziiert war (Snijders et al. 2017). Im hier gewonnenen Datensatz war FAM83H in HNSCC deutlich hochreguliert ($\log_2$ FC HNSCC/SQCLC = 0,84; p = 0,006).

Die Aktivität der HMGCS-1 scheint in Zusammenhang mit mutiertem B-Raf ein wichtiger onkogener Faktor in diversen malignen Neoplasien zu sein (Zhao et al. 2017) und war im hier vorliegenden Datensatz in SQCLC hochreguliert ($\log_2$ FC HNSCC/SQCLC = -0,57; p = 0,02). Dies zeigt, dass die massenspektrometrische Proteomanalyse einen wichtigen Beitrag zur Identifikation und weiteren Charakterisierung von Proteinen leisten kann, die in der Tumorentstehung und -therapie von Bedeutung sind.

4.4 Immunhistochemische Validierung und Validierung der Proteinsignatur[1]

Zur immunhistochemischen Validierung der proteomischen Daten wurden HMGCS-1, FAM83H, LGALS7 und FTL (*ferritin light chain*) an einer unabhängigen Kohorte von 212 SQCLC und 343 HNSCC immunhistochemisch angefärbt (Bohnenberger et al. 2018). Dabei zeigte sich, wie bereits in den in dieser Studie gewonnenen proteomischen Daten, dass HMGCS-1 und FTL in SQCLC stärker exprimiert waren, FAM83H und LGALS7 hingegen stärker in HNSCC (Bohnenberger et al. 2018). Dies zeigt, dass sich die Ergebnisse der massenspektrometrischen Proteomanalyse mit klinischen Routineverfahren wie der Immunhistochemie validieren lassen. Allerdings war aufgrund der großen Ähnlichkeit zwischen HNSCC und SQCLC keiner der immunhistochemischen Marker geeignet, um zwischen diesen beiden Entitäten zu differenzieren (Bohnenberger et al. 2018), was erneut die Erkenntnis unterstreicht, dass dafür komplexe proteomische Signaturen notwendig sind.

4.5 Einsatz der Proteinsignatur zur klinischen Risikostratifizierung[1]

Die aus etwa 1.100 Proteinen bestehende Signatur, welche in dieser Arbeit identifiziert worden war, konnte an einem unabhängigen Kollektiv aus insgesamt 38 Tumorgewebeproben (19 HNSCC und 19 SQCLC) mit einer diagnostischen Genauigkeit von 86,8%, einer Sensitivität von 85,0% und einer Spezifität von 88,9% erfolgreich validiert werden (Bohnenberger et al. 2018). In einem weiteren Schritt wurden insgesamt 51 Plattenepithelkarzinome der Lunge unbekannter Herkunft, bei denen zusätzlich ein HNSCC in der Vorgeschichte

[1] Die im folgenden Abschnitt beschriebenen Untersuchungen wurden nicht mehr im Rahmen dieser Arbeit, sondern auf dem hier gewonnenen proteomischen Datensatz aufbauend von Bohnenberger et al. (2018) durchgeführt.

aufgetreten war, massenspektrometrisch analysiert und anhand der proteomischen Signatur mithilfe der zuvor trainierten SVM entweder als Metastase des HNSCC (metHNSCC) oder metachrones Primärkarzinom der Lunge (SQCLC) klassifiziert (Bohnenberger et al. 2018). Zusätzlich wurde die Einteilung auch nach einer derzeit üblichen klinischen Klassifikation vorgenommen (Ichinose et al. 2016a), bei der klinische Merkmale wie Lokalrezidiv des HNSCC, die Zeit zwischen HNSCC und Auftreten des Lungentumors sowie die Anzahl an Tumorherden in der Lunge einbezogen wurden (Bohnenberger et al. 2018). In beiden Fällen wurde die getroffene Einteilung mit dem Gesamtüberleben korreliert. Während durch die klinische Klassifikation keine signifikanten Unterschiede im Gesamtüberleben identifiziert werden konnten, zeigte sich bei der Einteilung nach der proteomischen Signatur ein signifikanter Unterschied mit einem medianen Überleben von acht Monaten in der als metHNSCC vorhergesagten Gruppe und demgegenüber einem medianen Überleben von 55 Monaten in der als SQCLC vorhergesagten Gruppe, was dem erwarteten Unterschied zwischen diesen beiden Krankheitssituationen entspricht (Bohnenberger et al. 2018). Kritisch anzumerken ist hierbei jedoch, dass die HNSCC von 23 der untersuchten Patient*innen bereits Teil der initialen, in dieser Arbeit beschriebenen, Trainingskohorte gewesen waren. Auch unter Ausschluss dieser Patient*innen zeigte sich jedoch ein signifikanter Unterschied im Überleben lediglich für die Einteilung anhand der proteomischen Signatur und nicht für die klinische Klassifikation (Bohnenberger et al. 2018).

Dies spricht dafür, dass Lungentumore unklarer Herkunft mithilfe komplexer proteomischer Signaturen in verschiedene Risikogruppen mit unterschiedlicher Prognose stratifiziert werden können (Bohnenberger et al. 2018) und die Unterscheidung von metHNSCC und SQCLC mithilfe komplexer proteomischer Signaturen besser gelingt als auf der Basis gängiger klinischer Merkmale.

4.6 Limitationen, Chancen und Ausblick auf zukünftige Entwicklungen

Einschränkend müssen sowohl in Bezug auf diese Arbeit als auch auf die erwähnten vorhergehenden Studien zur Unterscheidung von metHNSCC und SQCLC einige Punkte berücksichtigt werden: Bei allen bisherigen Studien handelte es sich um retrospektive, unizentrische Studien mit vergleichsweise geringen Fallzahlen. Sie liefern jeweils Hinweise, dass verschiedene biologische Aspekte (z. B. *loss of heterozygosity*, DNA-Methylierung, Genexpression, Proteinexpression) bei der Unterscheidung zwischen metHNSCC und SQCLC hilfreich sein können. Es gibt bislang allerdings keine Studie, welche die Kombination dieser Aspekte

untersucht bzw. sie miteinander vergleicht. Hierzu werden prospektive klinische Studien mit höheren Fallzahlen dringend benötigt.

Bei der Identifikation von genetischen oder proteomischen Signaturen und vor allem bei deren Validierung an Lungentumoren unbekannten Ursprungs ergeben sich zwei grundlegende Probleme: Zum einen existiert wie bereits erwähnt kein etablierter klinischer Standard zur sicheren Unterscheidung von metHNSCC und SQCLC. Um diese Schwierigkeit zu adressieren, könnte eine Betrachtung des Gesamtüberlebens bzw. Rezidivrisikos zwischen den beiden Gruppen hilfreich sein, da im Fall eines metastasierten HNSCC von einem geringeren Gesamtüberleben bzw. höherem Rezidivrisiko als bei metachronen Primärkarzinomen der Lunge auszugehen ist. Erfolgreich konnte dies bereits von Bohnenberger et al. (2018) bzw. Ichinose et al. (2016b) eingesetzt werden. Auch bei der Differenzierung von DLBCL-Subtypen war die Korrelation mit dem Gesamtüberleben hilfreich (Alizadeh et al. 2000). Zum anderen besteht ein gewisser Bias, da Lungentumore bevorzugt dann operativ reseziert werden, wenn klinisch der Verdacht auf ein metachrones Primärkarzinom der Lunge und nicht auf eine Lungenmetastase des HNSCC besteht. Um das Risiko für einen derartigen Bias zu reduzieren, wären prospektive, randomisierte klinische Studien notwendig.

In dieser Studie wurden HNSCC unabhängig von ihrer Lokalisation (5x Mundhöhle, 13x Pharynx, 12x Larynx) und karzinogenen Faktoren (3x HPV positiv, 27x HPV negativ) eingeschlossen, obwohl es sich hierbei um eine sehr heterogene Gruppe an malignen Neoplasien handelt. So ist beispielsweise das Risiko für Fernmetastasierung bei HNSCC des Hypopharynx deutlich höher als bei anderen Lokalisationen (Takes et al. 2012). HPV positive oropharyngeale Karzinome haben eine bessere Prognose als HPV negative (Fakhry et al. 2014). Weitere Studien sind notwendig, um diagnostische Kriterien/Signaturen entsprechend der jeweiligen Lokalisation bzw. dem Vorhandensein karzinogener Faktoren zu entwickeln.

Auch in Bezug auf die hier angewandte Methode der massenspektrometrischen Proteomanalyse gilt es einige Aspekte zu berücksichtigen. So ist zunächst lediglich die relative Quantifizierung verschiedener Proben untereinander möglich. Es konnte beispielsweise gezeigt werden, dass die onkogene HMGCS-1 in SQCLC im Vergleich zu HNSCC stärker exprimiert ist, über die absolute Expression in HNSCC bzw. SQCLC kann jedoch keine direkte Aussage getroffen werden. Dies ließe sich allerdings durch methodische Anpassungen realisieren. So könnte ein Biomarker, dessen absolute Expression ermittelt werden soll, in einer definierten, bekannten Menge relativ zum Proteinquantifizierungsstandard quantifiziert werden. Im Vergleich mit der Expression in HNSCC oder SQCLC ließe sich der Standard als

verbindendes Element kürzen und so die absolute Expression des Biomarkers in HNSCC bzw. SQCLC bestimmen (Shenoy und Geiger 2015).

Ein Vorteil der massenspektrometrischen Proteomanalyse ist, dass auch noch Monate oder Jahre später weitere Gewebeproben analysiert und unmittelbar mit den ursprünglichen Ergebnissen verglichen werden können, da der Proteinquantifizierungsstandard günstig in großem Stil produziert und langfristig gelagert werden kann (Shenoy und Geiger 2015). So könnte beispielsweise normales Plattenepithel aus dem Kopf-Hals-Bereich und der Lunge analysiert und die Proteinexpression quantitativ mit dem bereits untersuchten Karzinomgewebe verglichen werden. Auf diese Weise können proteomische Datensätze kontinuierlich ergänzt bzw. verfeinert werden und so Erkenntnisse über die biologischen Grundlagen maligner Neoplasien im Allgemeinen gewonnen werden. Ein weiterer Vorteil des Proteinquantifizierungsstandards ist die Tatsache, dass die proteomische Analyse nicht von bestimmten Laboratorien oder Massenspektrometern abhängig ist. Somit sind auch große multizentrische Studien, bei denen die Proteomanalyse zum Einsatz kommt, denkbar.

Da sich die in dieser Arbeit beschriebene Methode, von der Entnahme des Gewebes bis zur massenspektrometrischen Analyse, innerhalb einer Woche durchführen lässt, was in etwa der Dauer anderer molekularbiologischer Untersuchungsmethoden wie z. B. dem Next-Generation-Sequencing entspricht, ist die Integration in die klinische Routine prinzipiell möglich. In der klinischen Diagnostik könnten die proteomischen Daten so in Verbindung mit weiteren molekularbiologischen und klinischen Merkmalen zur Entscheidungsfindung herangezogen werden. Allerdings sind hierfür ebenfalls zunächst prospektive klinische Studien notwendig.

Mit der hier erfolgreich angewandten Methode der Proteinisolation aus FFPE-Gewebe stehen darüber hinaus die FFPE-Gewebe, die weltweit in den Archiven Pathologischer Institute oder Biobanken lagern, der proteomischen Analyse mittels Massenspektrometrie zur Verfügung. Die für die Analyse notwendige Gewebemenge von weniger als 1 µl sollte dabei in der Regel gut verfügbar sein. Der Proteinquantifizierungsstandard stellt dafür eine exakte Quantifizierungsmethode dar, mit der proteomische Analysen günstig, in großem Stil und über einen langen Zeitraum möglich sind (Shenoy und Geiger 2015). Die untersuchten Fragestellungen müssen dabei nicht auf maligne Neoplasien beschränkt sein.

Es ist anzunehmen, dass es nach den technischen Entwicklungen im Bereich der Massenspektrometrie in den letzten Jahren (Aebersold und Mann 2003; 2016) auch zukünftig zu Weiterentwicklungen kommen wird und massenspektrometrische Analysen dadurch noch präziser werden. Durch die Fortschritte im Bereich der Bioinformatik besteht einerseits die Möglichkeit, die großen Datenmengen der Massenspektrometrie auszuwerten. Andererseits

können auch verschiedene sogenannte *-omics* Daten (*Genomics, Transcriptomics, Metabolomics, Proteomics*) zusammengeführt und als Ganzes betrachtet werden (Manzoni et al. 2018). Vielversprechend sind auch die Entwicklungen im Bereich des maschinellen Lernens, insbesondere im Bereich künstlicher neuronaler Netzwerke (Jurmeister et al. 2019). Diese Aspekte können zum einen hilfreich sein, um die bereits durchgeführten Studien auf genomischer, transkriptomischer und proteomischer Ebene zur Unterscheidung von HNSCC und SQCLC gemeinsam zu analysieren. Zum anderen kann dies zukünftig einen wichtigen Beitrag zu einem besseren Gesamtbild molekularbiologischer Prozesse im Allgemeinen und insbesondere auch bei Erkrankungen wie malignen Neoplasien leisten.

5 Zusammenfassung

Bei Patient*innen, die an einem Plattenepithelkarzinom des Kopf-Hals-Bereichs erkrankt sind, kann es aufgrund gemeinsamer Risikofaktoren wie z. B. Tabakrauch im Verlauf sowohl zu einer Lungenmetastase dieses Karzinoms als auch zu einem metachronen Plattenepithelkarzinom der Lunge kommen. Die Unterscheidung dieser beiden Tumorentitäten ist essentiell, da entweder palliativ oder kurativ therapiert wird, gestaltet sich aufgrund großer genomischer und histomorphologischer Ähnlichkeiten in der klinischen Routine jedoch als große Herausforderung. In dieser Studie wurde eine vergleichende proteomische Analyse dieser Neoplasien durchgeführt mit dem Ziel, sie anhand von Proteinsignaturen voneinander zu unterscheiden. Dazu wurden Proteine aus insgesamt 74 Tumorgewebeproben (44 Plattenepithelkarzinome der Lunge und 30 Plattenepithelkarzinome des Kopf-Hals-Bereichs) isoliert und mittels eines Proteinquantifizierungsstandards massenspektrometrisch analysiert.

Dabei konnten pro Probe durchschnittlich 2.251 und insgesamt 6.214 Proteine quantifiziert werden. Bei 518 Proteinen zeigte sich eine signifikant unterschiedliche Expression zwischen Plattenepithelkarzinomen der Lunge und des Kopf-Hals-Bereichs. Darunter waren auch Proteine, wie beispielsweise HMGCS-1 oder FAM83H, die bereits zuvor als onkogene Faktoren beschrieben worden waren. Mittels einer aus ca. 1.100 Proteinen bestehenden Signatur konnten Plattenepithelkarzinomen der Lunge und des Kopf-Hals-Bereichs erfolgreich differenziert werden. Einzelne Biomarker waren hingegen nicht geeignet.

Diese Arbeit stellt damit einen umfassenden proteomischen Datensatz zur Verfügung, der als Grundlage weiterer Untersuchungen dienen kann. Sie zeigt außerdem, dass die Unterscheidung von malignen Neoplasien mit hoher Ähnlichkeit durch komplexe proteomische Signaturen möglich ist. Die massenspektrometrische Proteomanalyse könnte zukünftig eine Rolle in der Diagnostik von Lungentumoren bei Patient*innen mit einem Plattenepithelkarzinom des Kopf-Hals-Bereichs einnehmen, jedoch sind hierfür insbesondere noch prospektive klinische Studien notwendig. Die in dieser Studie gewählte Methode bietet außerdem die Möglichkeit, formalinfixiertes und in Paraffin eingebettetes Gewebe, das weltweit in den Archiven Pathologischer Institute und Biobanken lagert, massenspektrometrisch zu analysieren und so zu einem besseren Verständnis maligner Neoplasien und der Beantwortung klinisch relevanter Fragestellungen beizutragen.

6 Literaturverzeichnis

Aebersold R, Mann M (2003): Mass spectrometry-based proteomics. Nature 422, 198–207

Aebersold R, Mann M (2016): Mass-spectrometric exploration of proteome structure and function. Nature 537, 347–355

Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X et al. (2000): Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 403, 503–511

Atabek U, Mohit-Tabatabai MA, Raina S, Rush BF, Dasmahapatra KS (1987): Lung cancer in patients with head and neck cancer. Incidence and long-term survival. Am J Surg 154, 434–438

Benesova K, Forsterova K, Votavova H, Campr V, Stritesky J, Velenska Z, Prochazka B, Pytlik R, Trneny M (2013): The Hans algorithm failed to predict outcome in patients with diffuse large B-cell lymphoma treated with rituximab. Neoplasma 60, 68–73

Bohnenberger H, Ströbel P, Mohr S, Corso J, Berg T, Urlaub H, Lenz C, Serve H, Oellerich T (2015): Quantitative mass spectrometric profiling of cancer-cell proteomes derived from liquid and solid tumors. J Vis Exp 96, e52435

Bohnenberger H, Kaderali L, Ströbel P, Yepes D, Plessmann U, Dharia NV, Yao S, Heydt C, Merkelbach-Bruse S, Emmert A et al. (2018): Comparative proteomics reveals a diagnostic signature for pulmonary head-and-neck cancer metastasis. EMBO Mol Med 10, e8428

Bonner JA, Harari PM, Giralt J, Azarnia N, Shin DM, Cohen RB, Jones CU, Sur R, Raben D, Jassem J et al. (2006): Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck. N Engl J Med 354, 567–578

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018): Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68, 394–424

Calhoun KH, Fulmer P, Weiss R, Hokanson JA (1994): Distant metastases from head and neck squamous cell carcinomas. Laryngoscope 104, 1199–1205

Campbell JD, Yau C, Bowlby R, Liu Y, Brennan K, Fan H, Taylor AM, Wang C, Walter V, Akbani R et al. (2018): Genomic, pathway network, and immunologic features distinguishing squamous carcinomas. Cell Rep 23, 194–212

Cancer Genome Atlas Network (2015): Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature 517, 576–582

Cancer Genome Atlas Research Network (2012): Comprehensive genomic characterization of squamous cell lung cancers. Nature 489, 519–525

Cipriano R, Graham J, Miskimen KLS, Bryson BL, Bruntz RC, Scott SA, Brown HA, Stark GR, Jackson MW (2012): FAM83B mediates EGFR- and RAS-driven oncogenic transformation. J Clin Invest 122, 3197–3210

Cox J, Mann M (2008): MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26, 1367–1372

Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M (2011): Andromeda: A Peptide search engine integrated into the MaxQuant environment. J Proteome Res 10, 1794–1805

Cree IA, Deans Z, Ligtenberg MJL, Normanno N, Edsjö A, Rouleau E, Solé F, Thunnissen E, Timens W, Schuuring E et al. (2014): Guidance for laboratories performing molecular pathology for cancer patients. J Clin Pathol 67, 923–931

Crippen MM, Brady JS, Burke LA, Eloy JA, Baredes S, Park RCW (2019): Second primary lung malignancy following head and neck squamous cell carcinoma. Laryngoscope 129, 903–909

de Bree R, Deurloo EE, Snow GB, Leemans CR (2000): Screening for distant metastases in patients with head and neck cancer. Laryngoscope 110, 397–401

Deeb SJ, D'Souza RCJ, Cox J, Schmidt-Supprian M, Mann M (2012): Super-SILAC allows classification of diffuse large B-cell lymphoma subtypes by their protein expression profiles. Mol Cell Proteomics 11, 77–89

Deeb SJ, Tyanova S, Hummel M, Schmidt-Supprian M, Cox J, Mann M (2015): Machine learning-based classification of diffuse large B-cell lymphoma patients by their protein expression profiles. Mol Cell Proteomics 14, 2947–2960

DGHO (2018): Lungenkarzinom, nicht-kleinzellig (NSCLC). Leitlinie. Hrsg. v. der Deutschen Gesellschaft für Hämatologie und Medizinische Onkologie e.V. Stand: November 2018. https://www.onkopedia.com/de/onkopedia/guidelines/lungenkarzinom-nicht-kleinzellig-nsclc/@@guideline/html/index.html; abgerufen am 18.07.2019.

DSMZ (2019a): HCC-15 (ACC 496). https://www.dsmz.de/collection/catalogue/details/culture/ACC-496; Zugriff am 17.07.2019.

DSMZ (2019b): HCC-44 (ACC 534). https://www.dsmz.de/collection/catalogue/details/culture/ACC-534; Zugriff am 17.07.2019.

Erde J, Loo RRO, Loo JA (2014): Enhanced FASP (eFASP) to increase proteome coverage and sample recovery for quantitative proteomic experiments. J Proteome Res 13, 1885–1895

Fakhry C, Zhang Q, Nguyen-Tan PF, Rosenthal D, El-Naggar A, Garden AS, Soulieres D, Trotti A, Avizonis V, Ridge JA et al. (2014): Human papillomavirus and overall survival after progression of oropharyngeal squamous cell carcinoma. J Clin Oncol 32, 3365–3373

Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM (1989): Electrospray ionization for mass spectrometry of large biomolecules. Science 246, 64–71

Ferlito A, Shaha AR, Silver CE, Rinaldo A, Mondin V (2001): Incidence and sites of distant metastases from head and neck cancer. ORL J Otorhinolaryngol Relat Spec 63, 202–207

Geiger T, Cox J, Ostasiewicz P, Wisniewski JR, Mann M (2010): Super-SILAC mix for quantitative proteomics of human tumor tissue. Nat Methods 7, 383–385

Geiger T, Wisniewski JR, Cox J, Zanivan S, Kruger M, Ishihama Y, Mann M (2011): Use of stable isotope labeling by amino acids in cell culture as a spike-in standard in quantitative proteomics. Nat Protoc 6, 147–157

Geurts TW, Nederlof PM, van den Brekel MWM, van't Veer LJ, de Jong D, Hart AAM, van Zandwijk N, Klomp H, Balm AJM, van Velthuysen MLF (2005): Pulmonary squamous cell carcinoma following head and neck squamous cell carcinoma: metastasis or second primary? Clin Cancer Res 11, 6608–6614

Geurts TW, van Velthuysen MLF, Broekman F, van Huysduynen TH, van den Brekel MWM, van Zandwijk N, van Tinteren H, Nederlof P, Balm AJM, Brakenhoff RH (2009): Differential diagnosis of pulmonary carcinoma following head and neck cancer by genetic analysis. Clin Cancer Res 15, 980–985

Guo T, Wang W, Rudnick PA, Song T, Li J, Zhuang Z, Weil RJ, DeVoe DL, Lee CS, Balgley BM (2007): Proteome analysis of microdissected formalin-fixed and paraffin-embedded tissue specimens. J Histochem Cytochem 55, 763–772

Hashibe M, Brennan P, Chuang SC, Boccia S, Castellsague X, Chen C, Curado MP, Dal Maso L, Daudt AW, Fabianova E et al. (2009): Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. Cancer Epidemiol Biomarkers Prev 18, 541–550

Henschke CI, Wisnivesky JP, Yankelevitz DF, Miettinen OS (2003): Small stage I cancers of the lung: genuineness and curability. Lung Cancer 39, 327–330

Hopkins J, Cescon DW, Tse D, Bradbury P, Xu W, Ma C, Wheatley-Price P, Waldron J, Goldstein D, Meyer F et al. (2008): Genetic polymorphisms and head and neck cancer outcomes: a review. Cancer Epidemiol Biomarkers Prev 17, 490–499

Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, Bravo HC, Davis S, Gatto L, Girke T et al. (2015): Orchestrating high-throughput genomic analysis with Bioconductor. Nat Methods 12, 115–121

Ichinose J, Shinozaki-Ushiku A, Takai D, Fukayama M, Nakajima J (2016a): Differential diagnosis between primary lung squamous cell carcinoma and pulmonary metastasis of head and neck squamous cell carcinoma. Expert Rev Anticancer Ther 16, 403–410

Ichinose J, Shinozaki-Ushiku A, Nagayama K, Nitadori JI, Anraku M, Fukayama M, Nakajima J, Takai D (2016b): Immunohistochemical pattern analysis of squamous cell carcinoma: Lung primary and metastatic tumors of head and neck. Lung Cancer 100, 96–101

Jiang H, English AM (2002): Quantitative analysis of the yeast proteome by incorporation of isotopically labeled leucine. J Proteome Res 1, 345–350

Jiang X, Jiang X, Feng S, Tian R, Ye M, Zou H (2007): Development of efficient protein extraction methods for shotgun proteome analysis of formalin-fixed tissues. J Proteome Res 6, 1038–1047

Jones AS, Morar P, Phillips DE, Field JK, Husband D, Helliwell TR (1995): Second primary tumors in patients with head and neck squamous cell carcinoma. Cancer 75, 1343–1353

Jou A, Hess J (2017): Epidemiology and molecular biology of head and neck cancer. Oncol Res Treat 40, 328–332

Jurmeister P, Bockmayr M, Seegerer P, Bockmayr T, Treue D, Montavon G, Vollbrecht C, Arnold A, Teichmann D, Bressem K et al. (2019): Machine learning analysis of DNA methylation profiles distinguishes primary lung squamous cell carcinomas from head and neck metastases. Sci Transl Med 11, eaaw8513

Kuriakose MA, Loree TR, Rubenfeld A, Anderson TM, Datta RV, Hill H, Rigual NR, Orner J, Singh A, Hicks WL (2002): Simultaneously presenting head and neck and lung cancer: a diagnostic and treatment dilemma. Laryngoscope 112, 120–123

Kutler DI, Auerbach AD, Satagopan J, Giampietro PF, Batish SD, Huvos AG, Goberdhan A, Shah JP, Singh B (2003): High incidence of head and neck squamous cell carcinoma in patients with Fanconi anemia. Arch Otolaryngol Head Neck Surg 129, 106–112

Leemans CR, Braakhuis BJM, Brakenhoff RH (2011): The molecular biology of head and neck cancer. Nat Rev Cancer 11, 9–22

Leitlinienprogramm Onkologie (2012): Diagnostik und Therapie des Mundhöhlenkarzinoms. S3-Leitlinie Langversion. Hrsg. v. Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF). AWMF-Registernummer: 007/100OL. Stand: November 2012. https://www.awmf.org/leitlinien/detail/ll/007-100OL.html; abgerufen am 18.07.2019.

Leitlinienprogramm Onkologie (2019): Diagnostik, Therapie und Nachsorge des Larynxkarzinoms. S3-Leitlinie Langversion 1.0. Hrsg. v. Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF). AWMF-Registernummer: 017/076OL. Stand: Januar 2019. http://www.leitlinienprogramm- onkologie.de/leitlinien/larynxkarzinom/; abgerufen am 18.07.2019.

León X, Quer M, Diez S, Orús C, López-Pousa A, Burgués J (1999): Second neoplasm in patients with head and neck cancer. Head Neck 21, 204–210

Lepikhova T, Karhemo PR, Louhimo R, Yadav B, Murumägi A, Kulesskiy E, Kivento M, Sihto H, Grénman R, Syrjänen SM et al. (2018): Drug-Sensitivity screening and genomic characterization of 45 HPV-Negative head and neck carcinoma cell lines for novel biomarkers of drug efficacy. Mol Cancer Ther 17, 2060–2071

Liao CT, Wang HM, Chang JTC, Ng SH, Hsueh C, Lee LY, Lin CH, Chen IH, Huang SF, Yen TC (2007): Analysis of risk factors for distant metastases in squamous cell carcinoma of the oral cavity. Cancer 110, 1501–1508

Lydiatt WM, Patel SG, Ridge JA, O'Sullivan B, Shah JP: Part II Head and Neck. In: Amin MB, American Joint Committee on Cancer, American Cancer Society (Hrsg.): AJCC cancer staging manual. Eight edition; American Joint Committee on Cancer, Springer, Chicago IL 2017, 55–181

Manzoni C, Kia DA, Vandrovcova J, Hardy J, Wood NW, Lewis PA, Ferrari R (2018): Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences. Brief Bioinform 19, 286–302

Nagaraj N, Lu A, Mann M, Wiśniewski JR (2008): Detergent-based but gel-free method allows identification of several hundred membrane proteins in single LC-MS runs. J Proteome Res 7, 5028–5032

Nalla AK, Williams TF, Collins CP, Rae DT, Trobridge GD (2016): Lentiviral vector-mediated insertional mutagenesis screen identifies genes that influence androgen independent prostate cancer progression and predict clinical outcome. Mol Carcinog 55, 1761–1771

Olsen JV, Ong SE, Mann M (2004): Trypsin cleaves exclusively C-terminal to arginine and lysine residues. Mol Cell Proteomics 3, 608–614

Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (2002): Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1, 376–386

Ong SE, Kratchmarova I, Mann M (2003): Properties of 13C-Substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). J Proteome Res 2, 173–181

Ostasiewicz P, Zielinska DF, Mann M, Wiśniewski JR (2010): Proteome, phosphoproteome, and N-glycoproteome are quantitatively preserved in formalin-fixed paraffin-embedded tissue and analyzable by high-resolution mass spectrometry. J Proteome Res 9, 3688–3700

Phelps RM, Johnson BE, Ihde DC, Gazdar AF, Carbone DP, McClintock PR, Linnoila RI, Matthews MJ, Bunn PA, Carney D et al. (1996): NCI-navy medical oncology branch cell line data base. J Cell Biochem Suppl 63, 32–91

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015): limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43, e47

Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, Fisher RI, Gascoyne RD, Muller-Hermelink HK, Smeland EB, Giltnane JM et al. (2002): The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. N Engl J Med 346, 1937–1947

Rusch VW (1995): Pulmonary metastasectomy. Current indications. Chest 107, 322S-331S

Rusch VW (2010): Pulmonary metastasectomy: a moving target. J Thorac Oncol 5, S130-131

Sankaranarayanan R, Masuyer E, Swaminathan R, Ferlay J, Whelan S (1998): Head and neck cancer: a global perspective on epidemiology and prognosis. Anticancer Res 18, 4779–4786

Scott L, Lamb J, Smith S, Wheatley DN (2000): Single amino acid (arginine) deprivation: rapid and selective death of cultured transformed and malignant cells. Br J Cancer 83, 800–810

Shenoy A, Geiger T (2015): Super-SILAC: current trends and future perspectives. Expert Rev Proteomics 12, 13–19

Shi SR, Liu C, Balgley BM, Lee C, Taylor CR (2006): Protein extraction from formalin-fixed, paraffin-embedded tissue sections: quality evaluation by mass spectrometry. J Histochem Cytochem 54, 739–743

Slaughter DP, Southwick HW, Smejkal W (1953): Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. Cancer 6, 963–968

Snijders AM, Lee SY, Hang B, Hao W, Bissell MJ, Mao JH (2017): FAM83 family oncogenes are broadly involved in human cancers: an integrative multi-omics approach. Mol Oncol 11, 167–179

Steen H, Mann M (2004): The abc's (and xyz's) of peptide sequencing. Nat Rev Mol Cell Biol 5, 699–711

Tabor MP, van Houten VMM, Kummer JA, Vosjan MJWD, Vlasblom R, Snow GB, Leemans CR, Braakhuis BJM, Brakenhoff RH (2002): Discordance of genetic alterations between primary head and neck tumors and corresponding metastases associated with mutational status of the TP53 gene. Genes Chromosomes Cancer 33, 168–177

Takes RP, Rinaldo A, Silver CE, Haigentz M, Woolgar JA, Triantafyllou A, Mondin V, Paccagnella D, de Bree R, Shaha AR et al. (2012): Distant metastases from head and neck squamous cell carcinoma. Part I. Basic aspects. Oral Oncol 48, 775–779

Talamini R, Bosetti C, La Vecchia C, Dal Maso L, Levi F, Bidoli E, Negri E, Pasche C, Vaccarella S, Barzan L, Franceschi S (2002): Combined effect of tobacco and alcohol on laryngeal cancer risk: a case-control study. Cancer Causes Control 13, 957–964

Talbot SG, Estilo C, Maghami E, Sarkaria IS, Pham DK, O-charoenrat P, Socci ND, Ngai I, Carlson D, Ghossein R et al. (2005): Gene expression profiling allows distinction between primary and metastatic squamous cell carcinomas in the lung. Cancer Res 65, 3063–3071

Tyanova S, Albrechtsen R, Kronqvist P, Cox J, Mann M, Geiger T (2016a): Proteomic maps of breast cancer subtypes. Nat Commun 7, 10259

Tyanova S, Temu T, Cox J (2016b): The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat Protoc 11, 2301–2319

Vachani A, Nebozhyn M, Singhal S, Alila L, Wakeam E, Muschel R, Powell CA, Gaffney P, Singh B, Brose MS et al. (2007): A 10-gene classifier for distinguishing head and neck squamous cell carcinoma and lung squamous cell carcinoma. Clin Cancer Res 13, 2905–2915

van Oijen MG, Leppers Vd Straat FG, Tilanus MG, Slootweg PJ (2000): The origins of multiple squamous cell carcinomas in the aerodigestive tract. Cancer 88, 884–893

Virmani AK, Fong KM, Kodagoda D, McIntire D, Hung J, Tonk V, Minna JD, Gazdar AF (1998): Allelotyping demonstrates common and distinct patterns of chromosomal loss in human lung cancer types. Genes Chromosomes Cancer 21, 308–319

Vizcaíno JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Ríos D, Dianes JA, Sun Z, Farrah T, Bandeira N et al. (2014): ProteomeXchange provides globally coordinated proteomics data submission and dissemination. Nat Biotechnol 32, 223–226

Walther TC, Mann M (2010): Mass spectrometry-based proteomics in cell biology. J Cell Biol 190, 491–500

Washburn MP, Wolters D, Yates JR (2001): Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 19, 242–247

Wiśniewski JR (2013): Proteomic sample preparation from formalin fixed and paraffin embedded tissue. J Vis Exp 79, e50589

Wiśniewski JR, Zougman A, Mann M (2009a): Combination of FASP and StageTip-based fractionation allows in-depth analysis of the hippocampal membrane proteome. J Proteome Res 8, 5674–5678

Wiśniewski JR, Zougman A, Nagaraj N, Mann M (2009b): Universal sample preparation method for proteome analysis. Nat Methods 6, 359–362

Wiśniewski JR, Nagaraj N, Zougman A, Gnad F, Mann M (2010): Brain phosphoproteome obtained by a FASP-based method reveals plasma membrane protein topology. J Proteome Res 9, 3280–3289

Wiśniewski JR, Ostasiewicz P, Mann M (2011): High recovery FASP applied to the proteomic analysis of microdissected formalin fixed paraffin embedded cancer tissues retrieves known colon cancer markers. J Proteome Res 10, 3040–3049

Wisnivesky JP, Yankelevitz D, Henschke CI (2004): The effect of tumor size on curability of stage I non-small cell lung cancers. Chest 126, 761–765

Younes RN, Gross JL, Silva JF, Fernandez JA, Kowalski LP (1997): Surgical treatment of lung metastases of head and neck tumors. Am J Surg 174, 499–502

Young ER, Diakos E, Khalid-Raja M, Mehanna H (2015): Resection of subsequent pulmonary metastases from treated head and neck squamous cell carcinoma: systematic review and meta-analysis. Clin Otolaryngol 40, 208–218

Zhang W, Wei Y, Ignatchenko V, Li L, Sakashita S, Pham NA, Taylor P, Tsao MS, Kislinger T, Moran MF (2014): Proteomic profiles of human lung adeno and squamous cell carcinoma using super-SILAC and label-free quantification approaches. Proteomics 14, 795–803

Zhao L, Fan J, Xia S, Pan Y, Liu S, Qian G, Qian Z, Kang HB, Arbiser JL, Pollack BP et al. (2017): HMG-CoA synthase 1 is a synthetic lethal partner of BRAFV600E in human cancers. J Biol Chem 292, 10142–10152

Zhu H, Pan S, Gu S, Bradbury EM, Chen X (2002): Amino acid residue specific stable isotope labeling for quantitative proteomics. Rapid Commun Mass Spectrom 16, 2115–2123

Danksagung

Mein besonderer Dank gilt Dr. Hanibal Bohnenberger, der mich als direkter Betreuer entscheidend bei der Planung und Durchführung dieser Arbeit unterstützt hat und mir bei technischen wie wissenschaftlichen Fragen und Unklarheiten jederzeit exzellent mit Rat und Tat zur Seite stand.

Bedanken möchte ich mich außerdem bei meinem Doktorvater Prof. Dr. Philipp Ströbel, der mir diese Arbeit mit seiner Unterstützung und dem mir entgegengebrachten Vertrauen ermöglicht hat.

Für die gelungene Zusammenarbeit und die Ermöglichung der massenspektrometrischen Messungen möchte ich Prof. Dr. Henning Urlaub (Max-Planck-Institut für biophysikalische Chemie Göttingen) und Prof. Dr. Thomas Oellerich (Medizinische Klinik II, Universitätsklinik Frankfurt am Main) danken.

Für die statistische Auswertung der massenspektrometrischen Daten möchte ich mich bei Prof. Dr. Lars Kaderali bedanken.

Dr. Stefan Küffer, Monique Küffer, Stefanie Schwager und Jennifer Appelhans möchte ich für die technische Unterstützung während der experimentellen Laborarbeit danken.

Für die konstruktiven und ausführlichen Korrekturen dieser Arbeit möchte ich mich bei Johanna Leyhe und Ole Lehmbäcker bedanken.