

Aus dem Institut für Pathologie  
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Proteogenomische Multiplexanalyse für die Tumordiagnostik

zur Erlangung des akademischen Grades  
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät  
Charité – Universitätsmedizin Berlin

von

Teresa Bockmayr  
aus München

Datum der Promotion: 26.06.2022

# Inhaltsverzeichnis

<b>Abkürzungsverzeichnis</b>	<b>IV</b>
<b>Zusammenfassung</b>	<b>V</b>
<b>Abstract</b>	<b>VI</b>
<b>1 Einleitung</b>	<b>1</b>
1.1 Tumorklassifikation und Grenzen der pathologischen Standarddiagnostik .	1
1.2 Molekulare Verfahren zur Tumordiagnostik und -klassifikation . . . . .	2
1.3 Bedeutung der Formalinfixierung zytologischer und histologischer Proben .	3
1.4 Zielsetzung der Arbeit . . . . .	4
<b>2 Methoden</b>	<b>5</b>
2.1 Next Generation Sequencing . . . . .	5
2.2 DNA-Methylierungs-Analyse . . . . .	6
2.3 Analyse der Proteinprofile mit der DigiWest-Methode . . . . .	7
2.4 Maschinelles Lernen . . . . .	8
<b>3 Ergebnisse</b>	<b>10</b>
3.1 Publikation 1: “Targeted deep sequencing of effusion cytology samples is feasible, informs spatiotemporal tumor evolution, and has clinical and diagnostic utility” . . . . .	10
3.2 Publikation 2: “Machine learning analysis of DNA methylation profiles distinguishes primary lung squamous cell carcinomas from head and neck metastases” . . . . .	12
3.3 Publikation 3: “Multiclass cancer classification in fresh frozen and formalin-fixed paraffin-embedded tissue by DigiWest multiplex protein analysis” . .	14
<b>4 Diskussion</b>	<b>16</b>
4.1 Nutzen der gezielten Tiefensequenzierung von Zellblöcken maligner Ergüsse	16
4.2 Bedeutung und Grenzen der Mutationsanalyse für die Tumorklassifikation .	16
4.3 Klinische Relevanz des DNA-Methylierungs-Tumorklassifikators . . . . .	17
4.4 Nutzen und Grenzen des Proteinprofil-basierten Tumorklassifikators . . . .	18
4.5 Vergleich der DNA-Methylierungs- und Proteinprofil-basierten Klassifikatoren	20
4.6 Fazit . . . . .	21
<b>Literaturverzeichnis</b>	<b>22</b>
<b>Eidesstattliche Erklärung</b>	<b>29</b>

Anteilserklärung	30
Publikationen	32
Publikation 1: “Targeted deep sequencing of effusion cytology samples is feasible, informs spatiotemporal tumor evolution, and has clinical and diagnostic utility”	33
Publikation 2: “Machine learning analysis of DNA methylation profiles distinguishes primary lung squamous cell carcinomas from head and neck metastases”	43
Publikation 3: “Multiclass cancer classification in fresh frozen and formalin-fixed paraffin-embedded tissue by DigiWest multiplex protein analysis”	53
Curriculum vitae	65
Komplette Publikationsliste	66
Danksagung	67

# Abkürzungsverzeichnis

BH:	Benjamini-Hochberg, Korrektur für multiples Testen
<i>BRAF</i> :	<i>B-Raf proto-oncogene, serine/threonine kinase</i>
COAD:	<i>Colorectal adenocarcinoma</i> ; kolorektales Adenokarzinom
CpG:	Cytosin-phosphatidyl-Guanin
CUP:	<i>Cancer of Unknown Primary</i> ; Krebserkrankung mit unbekanntem Primärtumor
dATP:	Desoxyadenosintriphosphat
dCTP:	Desoxycytidintriphosphat
dGTP:	Desoxyguanosintriphosphat
DNA:	<i>Deoxyribonucleic acid</i> ; Desoxyribonukleinsäure
dTTP:	Desoxythymidintriphosphat
<i>EGFR</i> :	<i>Epidermal growth factor receptor</i>
<i>ERBB2</i> :	<i>Erb-b2 receptor tyrosine kinase 2</i>
FFPE:	Formalin-fixiertes Paraffin-eingebettetes (Gewebe)
GEO:	<i>Gene Expression Omnibus</i>
HNSC:	<i>Head and neck squamous cell carcinoma</i> ; Plattenepithelkarzinom des Kopf-Halses
<i>KRAS</i> :	<i>KRAS proto-oncogene, GTPase</i>
LUAD:	<i>Lung adenocarcinoma</i> ; Adenokarzinom der Lunge
LUSC:	<i>Lung squamous cell carcinoma</i> ; Plattenepithelkarzinom der Lunge
MALDI:	<i>Matrix-Assisted Laser Desorption/Ionization</i> ; Matrix-unterstützte Laser-Desorption/Ionisierung
NGS:	<i>Next Generation Sequencing</i>
NN:	<i>Artificial neural network</i> ; künstliches neuronales Netzwerk
PAAD:	<i>Pancreatic adenocarcinoma</i> ; pankreatisches Adenokarzinom
PCR:	<i>Polymerase Chain Reaction</i> ; Polymerase-Kettenreaktion
PVDF:	Polyvinylidenfluorid
R:	Korrelationskoeffizient nach Pearson
RF:	<i>Random Forest</i>
RNA:	<i>Ribonucleic acid</i> ; Ribonukleinsäure
RPPA:	<i>Reverse Phase Protein Array</i>
SDS-PAGE:	<i>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</i> ; Natriumdodecylsulfat-Polyacrylamid-Gelelektrophorese
SVM:	<i>Support Vector Machine</i>
TCGA:	<i>The Cancer Genome Atlas</i>

## Zusammenfassung

Eine präzise Tumordiagnostik und -klassifikation bilden die Grundlage der onkologischen Therapieentscheidung. In der pathologischen Diagnostik werden dazu standardmäßig morphologische und immunhistochemische Verfahren eingesetzt. Diese stoßen bei der Identifikation des Organursprungs einer Metastase allerdings häufig an ihre Grenze und der maligne Tumor kann nicht eindeutig klassifiziert werden. Bei der Diagnostik therapierrelevanter Mutationen spielen darüber hinaus molekulargenetische Untersuchungen eine wichtige Rolle.

Im Rahmen dieser Arbeit wurden drei komplementäre molekulare Ansätze zur Unterstützung der histomorphologischen Tumordiagnostik untersucht. Zuerst wurde getestet, ob eine gezielte Tiefensequenzierung von Zellblöcken (Formalin-fixierten Paraffin-eingebetteten (FFPE) Zytologieproben) maligner Ergüsse möglich und diagnostisch relevant ist. In 15 von 20 Probenpaaren aus jeweils einer Tumorgewebeprobe und dem korrespondierenden Zellblock eines malignen Ergusses wurden übereinstimmende Genmutationen identifiziert. In vier dieser 15 Fälle konnten zusätzliche Mutationen entweder in der Gewebeprobe oder der Zytologie nachgewiesen werden, was auf eine räumlich-zeitliche Tumorentwicklung hinweist. In den übrigen fünf der 20 Probenpaare waren hingegen keine gemeinsamen Mutationen nachweisbar.

Anschließend wurden über die Analyse von DNA-Methylierungs- oder Proteinprofilen zwei unterschiedliche molekulare Methoden zur Identifikation des Organursprungs bei unklaren Metastasen untersucht. Die hohe Organspezifität der DNA-Methylierungs- und Proteinprofile ermöglichte die Klassifikation der Tumortypen mithilfe maschinellen Lernens.

Zunächst wurde ein DNA-Methylierungsprofil-basierter Tumorklassifikator etabliert, um pulmonale Metastasen eines Plattenepithelkarzinoms des Kopf-Halses (HNSC) von primären Plattenepithelkarzinomen der Lunge (LUSC) zu unterscheiden. Beide ähneln sich histomorphologisch und treten häufig bei Patienten mit HNSC auf. Zur Unterscheidung der DNA-Methylierungsprofile beider Tumortypen wurden drei maschinelle Lernverfahren eingesetzt. In einer unabhängigen Kohorte aus 279 Fällen (HNSC, LUSC, Lungennormalgewebe) erzielten die neuronalen Netzwerke (NN) eine Vorhersagegenauigkeit von 96,4 %, während *Support Vector Machines* (SVM) 95,7 % und *Random Forests* 87,8 % der Fälle korrekt klassifizierten. Außerdem wurden 50 von 51 Fällen (98 %) einer klinischen Kohorte mit den NN richtig als HNSC-Metastasen oder primäre LUSC eingeteilt.

Danach wurde ein Proteinprofil-basierter SVM-Klassifikator zur Unterscheidung von fünf Tumortypen (HNSC, LUSC, pulmonale, kolorektale und pankreatische Adenokarzinome) entwickelt. Hierfür wurde eine Western-Blot-basierte Methode zur Multiplex-Proteomik („DigiWest“) eingesetzt und zunächst deren Anwendbarkeit auf Routinediagnostik-FFPE-Gewebe evaluiert. 102 der getesteten 634 Antikörper zeigten sich dafür geeignet und wurden zur Tumorklassifikation verwendet. In 25 gepaarten Primärtumoren aus Frisch- und FFPE-Gewebe erzielte der SVM-Klassifikator eine Vorhersagegenauigkeit von 90,4 % in Frischgewebe und 77,6 % in den FFPE-Proben. Zudem wurden in einer unabhängigen Kohorte aus 25 FFPE-Tumortypen 88 % der Fälle richtig klassifiziert.

Zusammenfassend zeigen diese Ergebnisse, dass neben klassischen genomischen auch neuartige epigenetische und proteomische Verfahren einen wichtigen Beitrag zur Tumordiagnostik leisten können.

## Abstract

Precise cancer diagnostics and classification constitute the basis for oncological treatment selection. Histomorphology and immunohistochemistry are routinely used in pathological diagnostics but often reach their limits in identifying the tissue of origin of metastases, and cancer cannot be classified. Moreover, molecular genetic testing is crucial in the detection of therapeutically relevant mutations.

In this thesis, we explored three complementary molecular approaches to support histomorphological cancer diagnostics.

First, we tested the feasibility and diagnostic relevance of targeted deep sequencing of malignant effusion cell blocks (formalin-fixed paraffin-embedded (FFPE) cytology samples). 15 of 20 matched samples, each containing a tumor tissue sample and a malignant effusion cell block, showed concordant gene mutations. In four of these 15 cases, additional mutations were detected in either the tissue sample or cell block, indicating spatiotemporal tumor evolution. However, no concordant mutations were detected in the remaining five matched samples.

Next, two different methods based on the analysis of either DNA methylation or protein profiles were used to identify the tissue of origin of undetermined metastases. The high organ specificity of both methods enabled the classification of tumor samples by machine learning.

Firstly, a DNA methylation-based classifier was established to differentiate pulmonary metastases of head and neck squamous cell carcinomas (HNSC) from primary lung squamous cell carcinomas (LUSC). Both share similar histomorphology and frequently occur in patients with HNSC. Three machine learning methods were applied to distinguish both tumor types. Neural networks (NN) accurately classified 96.4% of 279 independent cases (HNSC, LUSC, normal lung tissue), while support vector machines (SVM) and random forests reached an accuracy of 95.7% and 87.8%, respectively. Furthermore, NN correctly distinguished primary LUSC from HNSC metastases in 98% of 51 clinical cases.

Secondly, a protein profile-based SVM classifier was developed to distinguish five tumor types (HNSC, LUSC, pulmonary, colorectal, and pancreatic adenocarcinomas). To this end, we used a western blot-based method for multiplex proteomics (“DigiWest”), which we first evaluated for its applicability to routine diagnostics FFPE tissue. 102 of 634 tested antibodies were suitable for this purpose and were applied for cancer classification. The SVM classifier yielded an accuracy of 90.4% in fresh frozen and 77.6% in FFPE tissue in a set of 25 paired fresh frozen and FFPE primary tumor samples. Moreover, 88% of 25 independent FFPE tumor samples were classified correctly.

In summary, these results demonstrate that besides classic genomic analysis, novel epigenetic and proteomic methods may valuably contribute to cancer diagnostics.

# 1 Einleitung

## 1.1 Tumorklassifikation und Grenzen der pathologischen Standarddiagnostik

Jedes Jahr wird bei ca. 500 000 Patienten in Deutschland eine Krebserkrankung neu diagnostiziert [1]. Eine genaue Diagnosestellung und korrekte Tumorklassifikation sind für diese Patienten von zentraler Bedeutung, da sich das weitere therapeutische Vorgehen danach richtet. Fehler in der Diagnosestellung können zur Auswahl einer ungeeigneten Therapie führen und die Prognose der Patienten somit erheblich verschlechtern. Maligne Tumore werden nach ihrem Organursprung und histologischen Typ (*Typing*), dem Differenzierungsgrad (*Grading*) und der anatomischen Ausbreitung (*Staging*) klassifiziert [2]. Neben klinischen, radiologischen und endoskopischen Untersuchungen spielt dabei vor allem die pathologische Diagnostik eine wesentliche Rolle. Zur pathologischen Standarddiagnostik gehören morphologische Untersuchungen von zytologischen oder histologischen Proben, am häufigsten an Hämatoxylin-Eosin-gefärbten Schnittpräparaten [2]. Ergänzend können immunhistochemische Färbungen angefertigt werden, um bestimmte Proteine durch Antikörpermarkierungen darzustellen [2]. Daneben erlangen auch molekularpathologische Analysen eine immer größere Bedeutung, insbesondere um therapierelevante genetische Aberrationen innerhalb eines Tumortyps nachzuweisen [2].

Durch diese Methoden gelingt es bei den meisten Patienten, den Organursprung und den histologischen Typ eines malignen Tumors zu bestimmen. Einige Fälle stellen allerdings eine diagnostische Herausforderung dar und der Tumortyp kann nicht eindeutig identifiziert werden. So lassen sich Primärtumore nicht immer zuverlässig von Metastasen unterscheiden, vor allem wenn es sich um Plattenepithelkarzinome oder schlecht differenzierte Tumore handelt. Plattenepithelkarzinome aus unterschiedlichen Lokalisationen weisen sowohl morphologisch als auch immunhistochemisch eine große Ähnlichkeit untereinander auf [3]. Dies betrifft insbesondere Patienten mit einem Plattenepithelkarzinom des Kopf-Halses (HNSC), bei denen eine weitere Raumforderung in der Lunge diagnostiziert wird. Einerseits treten ca. 70-80 % der Fernmetastasen eines HNSC in der Lunge auf, es handelt sich in diesem Fall um eine systemische Erkrankung, die mit einer schlechten Prognose einhergeht [4, 5]. Die Behandlung erfolgt dann meist durch eine palliative Chemotherapie und/oder Immuntherapie mit Checkpoint-Inhibitoren [5]. Andererseits haben Patienten mit einem HNSC auch ein erhöhtes Risiko einen zweiten Primärtumor in der Lunge zu entwickeln. Dort können sowohl synchrone als auch metachrone (bei ca. 7 % der Patienten mit HNSC) Plattenepithelkarzinome der Lunge (LUSC) auftreten [6]. Die Therapie eines lokal begrenzten LUSC erfolgt in der Regel in kurativer Intention und umfasst eine chirurgische Resektion des Tumors [7]. Die Unterscheidung zwischen einer pulmonalen Fernmetastase eines HNSC und einem zweiten Primärtumor der Lunge ist daher wesentlich, um eine adäquate Therapie einzuleiten und die Prognose dieser Patienten besser abschätzen zu können.

Des Weiteren kann es schwierig sein, den Organursprung einer Metastase zu identifizieren, wenn bei einem Patienten mehrere Primärtumore bekannt sind, oder es sich um eine primär metastasierte Erkrankung handelt, wie es insbesondere beim CUP-Syndrom (*Cancer of Unknown Primary* = Krebserkrankung mit unbekanntem Primärtumor) der Fall ist. Das CUP-Syndrom macht etwa 3-5 % aller Krebserkrankungen aus und umfasst eine heterogene Gruppe metastasierter Malignome, bei denen auch nach einer ausführlichen Diagnostik kein Primärtumor identifiziert werden kann [8]. Histologisch handelt es sich beim CUP-Syndrom am häufigsten um Adenokarzinome oder undifferenzierte Karzinome, seltener um Plattenepithelkarzinome oder neuroendokrine Tumore [9]. Wenn in einer Autopsie ein Primärtumor ermittelt werden konnte, befand er sich in etwa der Hälfte der Fälle entweder in der Lunge oder im Pankreas [10].

## 1.2 Molekulare Verfahren zur Tumordiagnostik und -klassifikation

Um die Zuordnung zu einer Tumorentität bei diagnostisch unklaren Fällen zu ermöglichen, wurden in den letzten Jahren verschiedene molekulare Methoden zur Tumordiagnostik und -klassifikation entwickelt. Diese wurden durch die raschen technischen Fortschritte der Messverfahren möglich und basieren auf der Analyse von Mutations-, mRNA-, microRNA-, DNA-Methylierungs- oder Proteinprofilen [11–15].

Somatische Mutationen sind häufig tumorentitätsübergreifend exprimiert [16], sie treten meist in verschiedenen Tumortypen mit unterschiedlichen Frequenzen auf. Deshalb ist vor allem eine vergleichende Analyse der Mutationen von mehreren Proben eines Patienten von diagnostischem Nutzen. So kann der Organursprung einer Metastase näher eingegrenzt werden, wenn es eine deutliche Übereinstimmung oder klare Unterschiede der Mutationsprofile zwischen dem Primärtumor und einer Metastase gibt [11]. Gleichzeitig können durch die Analyse der Mutationsprofile Treibermutationen diagnostiziert werden, die potenziell auf eine zielgerichtete Therapie ansprechen. Allerdings sind vergleichende Analysen unterschiedlicher Proben eines Patienten nicht immer möglich, da unter Umständen nicht genügend Material einer Probe verfügbar ist, oder im Falle des CUP-Syndroms kein Primärtumor identifiziert und so zum Vergleich herangezogen werden kann. Um ohne eine vergleichende Analyse auszukommen, wurden deshalb Klassifikatoren entwickelt, die tumortypspezifische molekulare Signaturen verwenden, um den Organursprung eines Tumors mithilfe von maschinellem Lernen zu bestimmen. Bisher beruhen diese Klassifikatoren am häufigsten auf der Analyse von Genexpressionsprofilen (mRNA und microRNA) [9, 12, 13]. Aber auch DNA-Methylierungs- und Proteinprofile sind spezifisch für einen Tumortyp und gewinnen daher zunehmend an Bedeutung [16, 17].

DNA-Methylierungen stellen einen wesentlichen Bestandteil der Epigenetik dar. Sie spielen eine wichtige Rolle in der Regulation der Genexpression, indem sie die Transkription der Gene hemmen [18]. Die Bindung der Methylgruppen erfolgt vorwiegend an Cytosinresten innerhalb von CpG-Dinukleotiden. In Tumorzellen treten sowohl verminderte als auch

vermehrte Methylierungen der CpG-Dinukleotide auf [17]. Einerseits geht eine globale Hypomethylierung der DNA mit einer genomischen Instabilität einher [18]. Andererseits führt eine DNA-Hypermethylierung in Promotoren von Tumorsuppressorgenen dazu, dass diese weniger exprimiert werden, was wiederum die Tumorentstehung fördern kann [19]. Da DNA-Methylierungsprofile charakteristisch für einen Tumortyp sind [17], wurden sie bereits erfolgreich genutzt, um Hirntumore, Sarkome oder CUP-Fälle zu klassifizieren [14, 20, 21].

Proteine sind zentrale Funktionsträger in der Zelle. Ihre Funktion kann durch posttranslationale Modifikationen verändert werden, dazu gehören unter anderem Phosphorylierungen, Glykosylierungen, Sulfatierungen oder Acetylierungen [22]. Durch posttranslationale Modifikationen und alternatives Spleißen ergibt sich eine deutlich größere Vielfalt an Proteinen in einer Zelle als an Genen. Neben der Massenspektrometrie stehen verschiedene Antikörperbasierte Verfahren zur Multiplexanalyse von Proteinen zur Verfügung, wie die *Reverse Phase Protein Arrays* (RPPA) oder die DigiWest-Methode, die einen klassischen Western Blot mit einer Bead-basierten Microarray-Plattform kombiniert [23, 24]. Die Analyse von Proteinprofilen in Gewebeproben ermöglichte bisher schon die Unterscheidung zwischen malignen und benignen Erkrankungen [25]. Da Proteinprofile zudem auch spezifisch für einen Tumortyp sind, können sie auch zur Tumorklassifikation verwendet werden [16]. Dies wurde bisher vor allem zur paarweisen Unterscheidung von Tumortypen genutzt, dagegen nur selten zur Klassifikation mehrerer Tumortypen mithilfe von Gewebeproben angewandt [15, 25–27].

### 1.3 Bedeutung der Formalinfixierung zytologischer und histologischer Proben

Damit die molekularen Analysen durchgeführt werden können, muss ausreichend Tumormaterial dafür zur Verfügung stehen. Frischgewebeproben eignen sich besonders gut für molekulare Untersuchungen, da bei der Kryokonservierung keine Formalin-induzierten Querverbindungen gebildet werden [28]. Allerdings ist die Gewinnung von Frischgewebeproben aufwändig, da die Proben möglichst zeitnah nach der Entnahme schockgefroren werden müssen, zudem ist ihre Lagerung teuer [28]. Dies führt dazu, dass Frischgewebeproben nur eingeschränkt verfügbar sind. Klinische Proben werden in der pathologischen Routinediagnostik standardmäßig in Formalin fixiert und anschließend in Paraffin eingebettet. Die Morphologie des Gewebes bleibt durch die Fixierung weitestgehend erhalten, dadurch sind Formalin-fixierte Paraffin-eingebettete (FFPE) Proben für histologische Untersuchungen gut geeignet [29]. Allerdings entstehen bei der Fixierung mit Formalin Querverbindungen (*cross-links*), die die Proteine untereinander oder mit Nukleinsäuren quervernetzen [29]. Die Extraktion der Makromoleküle aus FFPE-Gewebe kann dadurch erschwert werden [30]. Des Weiteren kann es durch die Fixierung zu einer Fragmentation der DNA und zu einer Verdickung oder Modifizierung von Epitopen der Proteine kommen, was deren Immunreaktivität beeinflussen kann [29, 31]. Es sollte daher für jede Messmethode zunächst geprüft werden,

inwieweit sie sich für eine Applikation auf FFPE-Gewebe eignet.

Neben histologischen Präparaten können auch zytologische Proben zur Tumordiagnostik herangezogen werden. Wenn zytologische Proben in Formalin fixiert und in Paraffin eingebettet werden, spricht man von Zellblöcken. Diese können unter anderem auch aus Körperflüssigkeiten wie malignen Ergüssen gewonnen werden. Zellblöcke sind lange haltbar und können zur Anfertigung zahlreicher Schnittpräparate für immunzytochemische oder auch molekulare Untersuchungen verwendet werden [32].

#### 1.4 Zielsetzung

Diese Arbeit beruht auf den Ergebnissen von insgesamt drei Publikationen, in denen verschiedene molekulare Methoden zur Unterstützung der histomorphologischen Tumordiagnostik entwickelt und getestet wurden. Im Folgenden wird die Zielsetzung der einzelnen Beiträge spezifiziert.

Im ersten Teil wird untersucht, ob eine *Next Generation Sequencing* (NGS)-basierte Multigen-Diagnostik von Zellblöcken (FFPE-Zytologieproben) maligner Ergüsse möglich ist und welchen Nutzen die daraus gewonnenen genetischen Informationen zur Komplementierung der Gewebediagnostik und in der Therapie haben. Dazu wurde eine gezielte Tiefensequenzierung mithilfe von vier Gen-Panels an 20 gepaarten Proben durchgeführt. Diese bestanden jeweils aus einer FFPE-Gewebeprobe des Primärtumors und dem korrespondierenden Zellblock eines malignen Ergusses [33].

Im zweiten Teil wird, basierend auf der Analyse von DNA-Methylierungsprofilen und Methoden des maschinellen Lernens, ein Klassifikator zur Unterscheidung von pulmonalen Metastasen eines Plattenepithelkarzinoms des Kopf-Halses (HNSC) und primären Plattenepithelkarzinomen der Lunge (LUSC) entwickelt [34].

Im dritten Teil wird getestet, ob die Analyse von Proteinprofilen mit der DigiWest-Methode es ermöglicht, Tumorproben entsprechend ihres Organursprungs zu klassifizieren. Da die Klassifikation sowohl in Frischgewebe als auch in FFPE-Proben angewandt werden soll, wurde zunächst geprüft, welche der untersuchten 634 Antikörper sich für eine Applikation auf FFPE-Gewebe mit der DigiWest-Methode eignen. Anschließend wurde analysiert, wie sich die Proteine in den folgenden Tumortypen unterscheiden: Plattenepithelkarzinome des Kopf-Halses (HNSC), Plattenepithelkarzinome der Lunge (LUSC), Adenokarzinome der Lunge (LUAD), kolorektale Adenokarzinome (COAD) und Adenokarzinome des Pankreas (PAAD). Schließlich wurden die Proteinprofile in Kombination mit einem Verfahren des maschinellen Lernens dazu genutzt, die Tumorproben nach ihrem Organursprung einzuteilen [35].

## 2 Methoden

In den drei Publikationen werden jeweils unterschiedliche Zielstrukturen (Genmutationen, DNA-Methylierungen, Proteine) analysiert. Daher werden in diesem Abschnitt zuerst die entsprechenden Messverfahren mit den dazugehörigen Kohorten beschrieben. Dabei handelt es sich um *Next Generation Sequencing*, DNA-Methylierungs-Arrays und die DigiWest-Methode. Anschließend wird ein Überblick über die Methoden des maschinellen Lernens gegeben, die zur Tumorklassifikation genutzt werden.

### 2.1 Next Generation Sequencing

Das *Next Generation Sequencing* (NGS) umfasst verschiedene Hochdurchsatzverfahren zur massiven parallelen Sequenzierung der DNA. Dabei können das gesamte Genom oder Exom sequenziert werden. Es ist aber auch möglich, bestimmte Gene auszuwählen und diese gezielt zu sequenzieren (Multigen-Panel-Sequenzierung). In dieser Arbeit wurden insgesamt vier Multigen-Panels (mit jeweils 50, 32, 25 und 161 Genen) mit einem Halbleiter-basierten NGS-Verfahren (Ion Torrent, Thermo Fisher Scientific, Waltham, USA) sequenziert, das im Folgenden kurz beschrieben wird [33].

Zuerst werden die Bibliotheken (*libraries*) erstellt. Dazu werden mindestens 10 ng DNA benötigt und mit dem Primer Pool gemischt, um die ausgewählten Gene über Polymerase-Kettenreaktion (PCR) zu amplifizieren [33]. Die Amplikons (DNA-Fragmente) werden dann partiell verdaut und an Adapter gebunden, bevor die Bibliothek quantifiziert wird [33]. Anschließend werden die DNA-Fragmente auf Beads fixiert und dort klonal mittels Emulsions-PCR amplifiziert, sodass jeder Bead zahlreiche Kopien eines bestimmten DNA-Fragments trägt. Dies ist notwendig, um bei der Sequenzierung ein ausreichend starkes Signal zu generieren. Jeder Bead kommt anschließend in eine Vertiefung (*well*) des Halbleiter-Chips, wo die Sequenzierung der DNA stattfindet [36]. Jedes Mal, wenn ein Nukleotid in die DNA eingebaut wird, kommt es zur Freisetzung eines Protons, was zu einem Absinken des pH-Werts in der jeweiligen Vertiefung führt. Die Veränderungen des pH-Werts werden über den Halbleiter-Chip gemessen und in ein digitales Signal umgewandelt, das proportional zu der Anzahl der eingebauten Nukleotide ist [36]. Da die vier Nukleotide (dATP, dCTP, dGTP, dTTP) nacheinander zugegeben werden und dazwischen jeweils ein Waschschrift erfolgt, kann so die Sequenz der DNA-Fragmente ermittelt werden [36, 37]. Ein Halbleiter-Chip hat Millionen an Vertiefungen, dadurch kann eine sehr große Anzahl an DNA-Fragmenten parallel sequenziert werden.

Die Multigen-Panel-Sequenzierung erfolgte an insgesamt 20 gepaarten Proben, die jeweils aus einer FFPE-Gewebeprobe des Tumors und dem korrespondierenden Zellblock eines malignen Ergusses bestanden [33]. Die Gewebeproben stammten aus insgesamt acht unterschiedlichen Tumorentitäten (mindestens 20 % Tumorzellgehalt), während die

Zytologieproben aus Pleura-, Aszites- oder Perikardpunktionen gewonnen wurden. In jedem Zellblock wurde der Tumorzellgehalt bestimmt. Er lag in neun der 20 Zellblöcke bei < 10 %, in diesen Zellblöcken waren auch keine Tumorzellcluster erkennbar [33].

## 2.2 DNA-Methylierungs -Analyse

Die Analyse der DNA-Methylierungen erfolgte mit dem Illumina Infinium Methylation EPIC BeadChip [34], der über 850 000 CpG-Dinukleotide abdeckt [38]. Dabei wird das Methylierungs-Level der einzelnen CpG-Dinukleotide untersucht, das Prinzip wird im Folgenden erläutert.

Die Methylierungen der DNA erfolgen weitestgehend an Cytosinen innerhalb von CpG-Dinukleotiden. Um methylierte von nicht-methylierten Cytosinen unterscheiden zu können, wird die DNA zuerst mit Natriumbisulfit behandelt [38]. Dies führt dazu, dass nicht-methylierte Cytosine in Uracile umgewandelt werden, während methylierte Cytosine unverändert bleiben [39]. Anschließend erfolgt die Amplifikation der Bisulfit-konvertierten DNA, wobei die Uracile (zuvor nicht-methylierte Cytosine) durch Thyminen ersetzt werden [39]. Die DNA wird dann fragmentiert und auf den BeadChip gegeben. Auf den Beads befinden sich Oligonukleotid-Sonden, die komplementär zu den untersuchten CpG-Dinukleotiden sind. Wenn die Bisulfit-konvertierte DNA daran hybridisiert, kann ein einzelnes fluoreszenz-markiertes Nukleotid in die Sonde eingebaut werden und so ein Signal generieren [38]. Je nach detektiertem Signal und verwendeter Sonde kann so das Level der DNA-Methylierung für ein bestimmtes CpG-Dinukleotid bestimmt werden.

In dem EPIC BeadChip werden Typ I und Typ II Sonden verwendet [38]. Bei den Typ I Sonden wird jedes untersuchte CpG-Dinukleotid durch zwei Sonden abgedeckt. Die eine Sonde ist spezifisch für die methylierte Form des CpG-Dinukleotids, während die andere Sonde komplementär zu der nicht-methylierten Form ist. Wenn die DNA mit einer der beiden Sonden hybridisiert, wird ein fluoreszenz-markiertes Nukleotid in die jeweilige Sonde eingebaut und so ein Signal generiert. Je nachdem in welcher Sonde das Signal detektiert wird, kann so bestimmt werden, inwieweit die DNA methyliert ist [38]. Bei den Typ II Sonden liegt nur eine Sonde pro CpG-Dinukleotid vor. Hier kann über den Fluoreszenzfarbstoff des eingebauten Nukleotids unterschieden werden, ob die Bisulfit-konvertierte DNA methyliert ist oder nicht. Ist die DNA methyliert, wird ein zu dem Cytosin komplementäres Guanin (z. B. grün markiert) eingebaut. Wenn die DNA hingegen nicht-methyliert ist, wird ein zu dem Thymin komplementäres Adenin (z. B. rot markiert) hinzugefügt. Durch die unterschiedlichen Fluoreszenzfarbstoffe können dann die DNA-Methylierungs-Level der untersuchten CpG-Dinukleotide bestimmt werden [38].

Die Analyse der DNA-Methylierungen wurde bei insgesamt 51 Patienten einer klinischen Kohorte durchgeführt. Dabei handelte es sich bei 31 Fällen um Metastasen eines HNSC und bei 20 Proben um primäre LUSC. Von dem Hersteller wurde empfohlen, 500 ng

DNA für die Untersuchung der DNA-Methylierungen zu verwenden [34]. Bei 15 Biopsien waren allerdings nur geringere Mengen an DNA verfügbar (mindestens 110 ng), eine erfolgreiche Durchführung der DNA-Methylierungs-Analyse war aber auch damit möglich [34]. Des Weiteren wurde eine statistische Kaplan-Meier-Analyse durchgeführt, um das krankheitsspezifische Überleben zu beurteilen.

### **2.3 Analyse der Proteinprofile mit der DigiWest-Methode**

Der DigiWest stellt ein Multiplexverfahren zur Proteinanalyse dar, in dem ein klassischer Western Blot mit einer Bead-basierten Microarray-Plattform kombiniert wird. Dadurch können bis zu 800 Proteine pro Probe gleichzeitig analysiert werden. Die DigiWest-Methode wurde gemäß Treindl et al. [24] durchgeführt und wird hier kurz zusammengefasst.

Zuerst werden die Proteine über eine Natriumdodecylsulfat-Polyacrylamid-Gelelektrophorese (SDS-PAGE) nach ihrem Molekulargewicht aufgetrennt. Anschließend werden sie auf eine Polyvinylidenfluorid (PVDF)-Membran geblottet und biotinyliert [24]. Jede Probenspur dieser Membran wird in 96 Streifen geschnitten, sodass jeder Streifen eine bestimmte Molekulargewichtsfraction abdeckt. Die 96 Streifen mit den biotinylierten Proteinen werden dann jeweils in eine Vertiefung einer Microarray-Plattform platziert. Dort werden die Proteine eluiert und mit farbkodierten Neutravidin-Beads inkubiert. Dadurch binden die biotinylierten Proteine an den farbkodierten Neutravidin-Beads, sodass jede Beadfarbe einer bestimmten Molekulargewichtsfraction entspricht [24]. Als Nächstes werden die Beads gemischt und bilden so wieder alle Molekulargewichtsfractionen der ursprünglichen Probenspur ab. Um die untersuchten Proteine nachzuweisen, wird jeweils ein kleines Aliquot der Bead-Mischung mit einem spezifischen Primärantikörper sowie einem fluoreszenz-markierten (Phycoerythrin) Sekundärantikörper inkubiert [24].

Die Proben werden mithilfe eines Durchflusszytometers ausgelesen, wobei gleichzeitig die Beadfarbe der jeweiligen Molekulargewichtsfraction und das Fluoreszenzsignal der Antikörperbindung detektiert werden. Für jeden Antikörper entstehen so 96 Signale pro Probe [24]. Diese werden in einem Diagramm visualisiert, in dem die Signalintensität in Abhängigkeit der Molekulargewichtsfraction abgebildet wird. Der Name „DigiWest“ leitet sich von dieser digitalen Repräsentation des ursprünglichen Blots ab. Zuletzt wird für jeden Antikörper das spezifische Signal bei dem erwarteten Molekulargewicht mithilfe eines Algorithmus identifiziert und quantifiziert [24].

Die Proteinanalyse mit der DigiWest-Methode erfolgte zuerst in einem Set aus 25 gepaarten Frischgewebe- und FFPE-Tumorproben. Dieses umfasste jeweils fünf Primärtumore aus den folgenden Tumortypen: Plattenepithelkarzinome des Kopf-Halses (HNSC), Plattenepithelkarzinome der Lunge (LUSC), Adenokarzinome der Lunge (LUAD), kolorektale Adenokarzinome (COAD) und Adenokarzinome des Pankreas (PAAD). Der Tumorzellgehalt wurde mittels Hämatoxylin-Eosin-gefärbter Schnittpräparate in jeder

Probe bestimmt. Er betrug mindestens 30 % im Frischgewebe und mindestens 40 % in den FFPE-Proben, mit durchschnittlich jeweils 70 % und 68 % [35].

Anschließend wurde die DigiWest-Methode in einer unabhängigen Kohorte durchgeführt. Diese bestand aus 25 FFPE-Primärtumoren mit jeweils fünf Proben pro Tumortyp (HNSC, LUSC, LUAD, COAD, PAAD). 18 dieser Proben waren mittelgradig und sieben schlecht differenziert. Der Tumorzellgehalt lag bei mindestens 40 % und betrug durchschnittlich 68 % [35].

Die benötigte Menge an Gesamtprotein für die DigiWest-Analyse variierte je nach Anzahl der gemessenen Antikörper. Zuerst erfolgte die Analyse von 634 Antikörpern in 16 FFPE-Proben des gepaarten Sets. Dafür wurden 2 x 20 µg Gesamtprotein verwendet. Anschließend wurden 303 Antikörper in den 25 gepaarten Proben aus Frisch- und FFPE-Gewebe untersucht, dazu wurde jeweils 20 µg Gesamtprotein benötigt [35]. Für diese Antikörper wurde die Korrelation zwischen den Signalen in Frisch- und FFPE-Gewebe mittels Pearson-Korrelationskoeffizienten berechnet. Das Signifikanzniveau wurde bei 0,05 festgelegt. Es erfolgte eine Korrektur für multiples Testen mit dem Benjamini-Hochberg Verfahren [40]. Bei dem Anti-Rabbit-Sekundärantikörper zeigte sich in den HNSC-Proben eine Kreuzreaktion bei ca. 50 kDa [35]. Daher wurden drei Primärantikörper von der Analyse ausgeschlossen, da eine Unterscheidung zwischen einer spezifischen Bindung des Primärantikörpers und einer Kreuzreaktion des Anti-Rabbit-Sekundärantikörpers nicht möglich war [35]. Zuletzt wurden in der unabhängigen Kohorte aus 25 FFPE-Proben 102 Antikörper gemessen, dafür waren 10 µg Gesamtprotein ausreichend [35]. Die Durchführung der DigiWest-Methode dauert mindestens zwei Wochen.

## 2.4 Maschinelles Lernen

Um die Tumorproben zu klassifizieren, wurden zwei unterschiedliche Herangehensweisen des maschinellen Lernens gewählt: ein Verfahren des unüberwachten Lernens (*unsupervised learning*) sowie mehrere Methoden des überwachten Lernens (*supervised learning*).

Einerseits erfolgte eine hierarchische Clusteranalyse, diese zählt zu den Methoden des unüberwachten Lernens. Die Klassenzugehörigkeit der Proben wird dabei nicht beachtet und das Verfahren dient vor allem dazu, Strukturen innerhalb eines Datensatzes zu erkennen. Proben, die sich ähneln, werden dabei in einem gemeinsamen Cluster zusammengefasst [41]. Um die Ähnlichkeit der Proben zu bestimmen, wird die Distanz zwischen zwei Proben bzw. Clustern berechnet (hier mittels Korrelation nach Pearson und *Average-Linkage*) [35, 41]. Eine geringe Distanz weist auf eine große Ähnlichkeit hin. Das Ergebnis der hierarchischen Clusteranalyse wird dann über ein Dendrogramm visualisiert [41].

Andererseits wurden insgesamt drei Methoden des überwachten Lernens verwendet, um die Proben entsprechend ihres Organursprungs zu klassifizieren. Dabei handelt es sich um *Support Vector Machines* (SVM), künstliche neuronale Netzwerke (NN) und *Random Forests* (RF). Bei den Methoden des überwachten Lernens lernt das Modell zuerst

anhand einer Trainingskohorte. In der Trainingskohorte ist die Klassenzugehörigkeit der Proben bekannt und das Modell wird darauf trainiert, die Klassen anhand bestimmter Muster zu unterscheiden [41]. Anschließend wird mit einer weiteren Kohorte getestet, wie hoch die Vorhersagegenauigkeit des Modells ist. Dazu werden Proben, bei denen die Klassenzugehörigkeit nicht vorgegeben ist, mithilfe dieses Modells klassifiziert. Die Vorhersagegenauigkeit gibt dann an, wie viele dieser Proben in die richtige Klasse eingeteilt werden [41]. Da zum Trainieren und Testen des Modells nicht die gleichen Proben verwendet werden dürfen, erfolgt die Bestimmung der Vorhersagegenauigkeit anhand einer unabhängigen Kohorte oder über verschachtelte Kreuzvalidierung (*nested cross-validation*). Vereinfacht betrachtet wird bei der verschachtelten Kreuzvalidierung eine Kohorte aufgeteilt, ein Teil der Proben wird zum Trainieren und Optimieren der Parameter des Modells verwendet und der andere zum Testen des Modells. Dieses Prozedere wird mehrmals mit unterschiedlichen Aufteilungen wiederholt [42, 43].

Für jedes Modell wird den Klassifikationsergebnissen jeweils eine bestimmte Wahrscheinlichkeit (Wahrscheinlichkeitswert zwischen 0 und 1) attribuiert. Durch den Wahrscheinlichkeitswert wird abgeschätzt, wie sicher das Modell in der Vorhersage einer Klasse ist. Die Einteilung des Modells erfolgt in die Klasse, für die die höchste Wahrscheinlichkeit vorliegt [34]. Ein ideales Modell würde die richtige Klasse mit einer Wahrscheinlichkeit von 1 vorhersagen und alle anderen Klassen mit einer Wahrscheinlichkeit von 0. Man geht davon aus, dass ein Modell höhere Vorhersagegenauigkeiten erzielen kann, wenn die Zuordnung zu einer Klasse mit hoher Wahrscheinlichkeit erfolgt. Um dies zu überprüfen, kann man einen bestimmten Wahrscheinlichkeitswert als Schwellenwert definieren und nur die Proben klassifizieren, deren Wahrscheinlichkeitswert über diesem Schwellenwert liegt. Alle anderen Proben, deren Wahrscheinlichkeitswert sich unterhalb dieses Schwellenwertes befindet, werden von der Klassifikation ausgeschlossen, da die Vorhersage des Klassifikators als zu unsicher betrachtet wird [34].

Im Folgenden wird ein kurzer Überblick über die drei verwendeten Klassifikatoren (SVM, NN, RF) gegeben. Diese drei Modelle wurden genutzt, um Metastasen eines HNSC von LUSC über die Analyse von DNA-Methylierungen zu unterscheiden [34]. Ihre Vorhersagegenauigkeiten wurden anschließend untereinander verglichen. Des Weiteren wurde das SVM-Modell verwendet, um fünf Tumortypen (HNSC, LUSC, LUAD, COAD, PAAD) basierend auf der Analyse von Proteinprofilen zu klassifizieren [35].

**Support Vector Machines:** Proben aus unterschiedlichen Klassen lassen sich nicht vollständig linear (durch eine Hyperebene im Stichprobenraum) voneinander abgrenzen, wenn es eine Überlappung ihrer Daten gibt. Das Prinzip der SVM besteht darin, die Daten mit einer Transformationsfunktion in einen höherdimensionalen Raum überzuführen, wo die verschiedenen Klassen dann linear voneinander getrennt werden können [41].

**Künstliche neuronale Netzwerke:** NN sind von dem Aufbau biologischer neuronaler Netzwerke inspiriert. NN bestehen aus Neuronen, die in Schichten angeordnet und über Kanten miteinander verbunden sind. Es gibt drei unterschiedliche Sorten von Neuronen: die Eingabeneuronen empfangen die Informationen von der Außenwelt, die verdeckten Neuronen verarbeiten diese schrittweise und die Ausgabeneuronen leiten das Ergebnis des Informationsflusses an die Außenwelt weiter. Die Kanten verbinden die Neuronen untereinander und sind gewichtet. Die Gewichte spiegeln den Einfluss eines Neurons auf das nächste wider, dieser kann exzitatorisch, inhibitorisch oder null sein. Vereinfacht betrachtet lernen die NN über Veränderungen der Gewichte. Wie diese Gewichte genau angepasst werden, wird mithilfe von Lernalgorithmen bestimmt [44].

**Random Forests:** RF bestehen aus einer Ansammlung vieler unterschiedlicher Entscheidungsbäume. Die Auswahl der Variablen für die Konstruktion der Entscheidungsbäume wird nach einem Zufallsprinzip bestimmt. Dadurch ergibt sich eine große Vielfalt unterschiedlicher Entscheidungsbäume. Um eine Probe zu klassifizieren, erfolgt durch jeden Entscheidungsbaum die Zuordnung zu einer Klasse. Als Klassifikationsergebnis der RF wird dann die Klasse angegeben, die am häufigsten durch die einzelnen Entscheidungsbäume gewählt wurde [41].

## 3 Ergebnisse

In diesem Abschnitt werden die wesentlichen Ergebnisse der drei Publikationen vorgestellt, die dieser Dissertation zugrunde liegen.

### 3.1 Publikation 1

*Leichsenring, J., Volckmar, A. L., Kirchner, M., Kazdal, D., Kriegsmann, M., Stögbauer, F., Bockmayr, T., Klauschen, F., Herth, F. J. F., Penzel, R., Warth, A., Schirmacher, P., Endris, V. & Stenzinger, A. Targeted deep sequencing of effusion cytology samples is feasible, informs spatiotemporal tumor evolution, and has clinical and diagnostic utility. *Genes Chromosomes Cancer*. 57, 70–79 (2018).*

In dieser ersten Arbeit wurde untersucht, ob die gezielte Tiefensequenzierung von Zellblöcken (FFPE-Zytologieproben) maligner Ergüsse möglich und von klinischer und diagnostischer Relevanz ist. Dazu erfolgte eine NGS-basierte Multigen-Diagnostik in 20 gepaarten FFPE-Proben mithilfe von insgesamt vier verschiedenen Gen-Panels. Die Paare bestanden aus jeweils einer FFPE-Gewebeprobe des Primärtumors und dem dazugehörigen Zellblock eines malignen Ergusses.

In allen FFPE-Gewebeproben des Primärtumors wurde mindestens eine *Missense*-, *Nonsense*- oder *Frameshift*-Mutation nachgewiesen. Bei allen Zellblöcken wurden die Bibliotheken erfolgreich generiert und sequenziert. Anschließend wurde analysiert, welche Gemeinsamkeiten oder Unterschiede der Mutationen zwischen den Gewebeproben und den dazugehörigen Zellblöcken bestanden. Einerseits zeigten 15 Patienten (75 %) übereinstimmende oder überlappende Mutationen zwischen beiden Proben. Davon waren in elf Zellblöcken die nachgewiesenen Mutationen identisch zu denen in den Gewebeproben. Bei vier Patienten gab es eine Überlappung von mindestens einer Mutation zwischen beiden Proben, wobei zusätzlich noch weitere Mutationen nachweisbar waren, entweder nur im Zellblock (zwei Patienten), nur in der Gewebeprobe (ein Patient) oder es gab zusätzliche unterschiedliche Mutationen in beiden (ein Patient). Andererseits bestanden bei fünf Patienten (25 %) keine Gemeinsamkeiten zwischen beiden gepaarten Proben. Davon zeigten zwei Patienten unterschiedliche Mutationen zwischen den Gewebeproben und Zellblöcken. Bei drei Patienten konnten jeweils nur Mutationen in der Gewebeprobe nachgewiesen werden, wobei im Zellblock keine Mutation (Wildtyp-Allele) identifiziert wurde. In diesen drei Zellblöcken lag der Tumorzellgehalt bei  $< 10\%$  und es waren keine Tumorzellcluster erkennbar.

Die gezielte Tiefensequenzierung der Zellblöcke der malignen Ergüsse hatte auch einen diagnostischen Nutzen. Bei einer Patientin wurden zwei verschiedene Primärtumore diagnostiziert. Einerseits war ein papilläres Schilddrüsenkarzinom mit einer *BRAF V600E*-Mutation bekannt, andererseits wurde in einem serösen Ovarialkarzinom eine *KRAS*-Mutation detektiert. Bei der Sequenzierung des Zellblocks der Aszites wurde die gleiche *KRAS*-Mutation wie in dem Ovarialkarzinom nachgewiesen, gleichzeitig lagen Wildtyp-Allele für *BRAF* vor. Somit konnte durch die gezielte Tiefensequenzierung die klinische Diagnose eines peritoneal metastasierten Ovarialkarzinoms bestätigt werden.

Des Weiteren wurden bei drei Patienten Mutationen mit potenziell klinischer Relevanz detektiert. Bei einer Patientin mit einem bekannten Adenokarzinom der Lunge wurde sowohl in der Gewebeprobe als auch im Zellblock des Pleuraergusses eine *EGFR*-Mutation identifiziert, die auf eine zielgerichtete Therapie ansprechen könnte. Bei einem weiteren Patienten mit einem Adenokarzinom der Lunge zeigte sich in beiden Proben eine aktivierende *ERBB2*-Mutation. Die Bedeutung einer zielgerichteten Therapie bei *ERBB2*-mutierten Adenokarzinomen der Lunge wird momentan in verschiedenen Studien evaluiert [45, 46]. Derzeit gehört sie allerdings nicht zur Standardtherapie. Bei dem dritten Patienten wurde ein kolorektales Karzinom mit einer *KRAS*-Mutation diagnostiziert, was mit einem schlechten Ansprechen auf eine EGFR-gerichtete Therapie einhergeht [47]. Allerdings konnte diese *KRAS*-Mutation nur in der Gewebeprobe identifiziert werden, während in dem Aszites-Zellblock Wildtyp-Allele vorlagen. In diesem Zellblock waren keine Tumorzellcluster erkennbar und der Tumorzellgehalt war gering ( $< 10\%$ ).

## 3.2 Publikation 2

*Jurmeister, P., Bockmayr, M., Seegerer, P., Bockmayr, T., Treue, D., Montavon, G., Vollbrecht, C., Arnold, A., Teichmann, D., Bressen, K., Schüller, U., von Laffert, M., Müller, K. R., Capper, D. & Klauschen, F. Machine learning analysis of DNA methylation profiles distinguishes primary lung squamous cell carcinomas from head and neck metastases. Sci Transl Med. 11 (2019).*

In dieser zweiten Arbeit wurde basierend auf der Analyse von DNA-Methylierungsprofilen in Tumorproben ein Klassifikator entwickelt, um pulmonale Metastasen eines Plattenepithelkarzinoms des Kopf-Halses (HNSC) von primären Plattenepithelkarzinomen der Lunge (LUSC) zu unterscheiden. Dieser wurde mithilfe von Methoden des maschinellen Lernens konstruiert.

Zuerst wurden die DNA-Methylierungen in Primärtumoren einer Referenzkohorte analysiert, um tumortypspezifische DNA-Methylierungsprofile zu erkennen, die die beiden Tumortypen voneinander unterscheiden können. Diese Referenzkohorte umfasste insgesamt 1071 Proben aus den Datenbanken *Gene Expression Omnibus* (GEO) und *The Cancer Genome Atlas* (TCGA), davon waren 528 primäre HNSC und 354 primäre LUSC. Zudem wurden 189 Normalgewebeproben der Lunge eingeschlossen, um auch Fälle erkennen zu können, die durch Lungenstroma oder angrenzendes normales Lungengewebe kontaminiert sind. Die 2000 variabelsten CpG-Stellen wurden identifiziert und dazu genutzt, drei unterschiedliche Methoden des maschinellen Lernens (künstliche neuronale Netzwerke (NN), *Support Vector Machines* (SVM), *Random Forests* (RF)) zu trainieren, um die drei Klassen (HNSC, LUSC, Lungennormalgewebe) voneinander zu unterscheiden.

Anschließend wurden diese drei Methoden des maschinellen Lernens in einer unabhängigen Kohorte auf ihre Vorhersagegenauigkeit getestet. Diese umfasste 279 Proben (110 HNSC-Primärtumore, 150 LUSC-Primärtumore, 19 Normalgewebeproben der Lunge) aus mehreren Studien (der Charité – Universitätsmedizin Berlin und aus sechs unterschiedlichen GEO-Datensätzen), die sich hinsichtlich des Probenmaterials (Frischgewebe- und FFPE-Proben), der Probenaufbereitung und der verwendeten DNA-Methylierungs-Arrays teilweise unterschieden. Die besten Ergebnisse konnten durch den NN-Klassifikator erzielt werden, der 96,4 % aller Proben richtig einteilte. Das SVM-Modell erzielte eine Vorhersagegenauigkeit von 95,7 %, während 87,8 % der Fälle durch den RF-Klassifikator richtig zugeordnet werden konnten.

Des Weiteren wurde untersucht, ob die Vorhersagegenauigkeit noch gesteigert werden kann, indem nur Proben berücksichtigt werden, für die die Zuordnung zu einer Klasse mit hoher Wahrscheinlichkeit angegeben wird. Der NN-Klassifikator konnte eine Vorhersagegenauigkeit von 98,1 % erreichen, wenn der Wahrscheinlichkeitswert für eine Klasse bei mindestens 0,8 lag, womit noch 269 der 279 Proben (96,4 %) eingeschlossen wurden. Bei

einem Wahrscheinlichkeitsschwellenwert von 0,95 wurde eine Vorhersagegenauigkeit von 99,2 % erzielt, wobei noch 257 der 279 Fälle (92,1 %) berücksichtigt wurden. Die Erhöhung der Vorhersagegenauigkeit geht somit mit einer geringeren Anzahl klassifizierbarer Proben einher. Die gleichen Vorhersagegenauigkeiten konnten zwar auch durch die SVM- und RF-Modelle erreicht werden, allerdings konnten dann deutlich weniger Fälle klassifiziert werden bzw. lagen niedrigere Wahrscheinlichkeitsschwellenwerte vor. Insgesamt weist der NN-Klassifikator somit auch hier die besten Ergebnisse auf und kann am meisten Proben mit sehr hoher Vorhersagegenauigkeit klassifizieren. Zudem zeigt der NN-Klassifikator beständig gute Ergebnisse, selbst bei Proben unterschiedlicher Herkunft und Verarbeitung. Insbesondere konnten auch FFPE-Proben richtig klassifiziert werden, obwohl für das Training des Klassifikators nur Frischgewebeproben herangezogen wurden.

Um die Validität des Klassifikators nicht nur für Primärtumore zu überprüfen, sondern um abzuschätzen, ob auch Metastasen eines HNSC von primären LUSC unterschieden werden können, erfolgte eine zusätzliche Validierung mithilfe einer klinischen Kohorte. Dazu wurden in einer retrospektiven Analyse an der Charité – Universitätsmedizin Berlin 408 Fälle identifiziert, bei denen sowohl ein HNSC als auch ein weiteres synchrones oder metachrones Plattenepithelkarzinom in der Lunge aufgetreten waren. Bei 64 der 408 Patienten (15,7 %) konnte in einer interdisziplinären Tumorkonferenz unter Berücksichtigung von klinischen und radiologischen sowie histopathologischen und molekularen Daten eine eindeutige Diagnose gestellt werden. Demnach handelte es sich bei 38 Fällen (59,4 %) um pulmonale Metastasen eines HNSC und bei 26 Tumoren (40,6 %) um primäre Plattenepithelkarzinome der Lunge. Diese Diagnosen dienten dann als Referenzdiagnose, um die Vorhersagegenauigkeit der Klassifikation basierend auf der Analyse von DNA-Methylierungsprofilen zu beurteilen. Bei den restlichen 344 der 408 Fälle (84,3 %) konnte keine präzise Diagnose gestellt werden, sie wurden daher ausgeschlossen. Da in 51 der 64 geeigneten Proben noch genügend Material vorhanden und die Qualität der DNA ausreichend war, erfolgte eine Analyse der DNA-Methylierungen in diesen Proben. Der NN-Klassifikator konnte 50 der 51 Proben richtig klassifizieren, nur eine HNSC-Metastase wurde fälschlicherweise als LUSC eingeteilt, was einer Vorhersagegenauigkeit von 98 % entspricht. Wenn nur die Proben mit einem Wahrscheinlichkeitswert über den vorher definierten Schwellenwerten von 0,8 und 0,95 eingeschlossen wurden, konnte die Vorhersagegenauigkeit des NN-Klassifikators bei beiden auf 100 % erhöht werden. Es wurden dann jeweils noch 48 (94 %) bzw. 46 (90 %) der 51 Fälle berücksichtigt. Bei der Kaplan-Meier-Analyse zeigten die Patienten, deren Tumor mit dem NN-Klassifikator als HNSC-Metastase diagnostiziert wurde, ein signifikant kürzeres krankheitsspezifisches Überleben ( $p < 0,05$ ) als diejenigen, deren Probe als LUSC klassifiziert wurde. Die SVM- und RF-Modelle waren dem NN-Klassifikator etwas unterlegen, sie teilten alle Proben mit einer Vorhersagegenauigkeit von jeweils 96,1 % und 84,3 % ein.

### 3.3 Publikation 3

*Bockmayr, T., Erdmann, G., Treue, D., Jurmeister, P., Schneider, J., Arndt, A., Heim, D., Bockmayr, M., Sachse, C. & Klauschen, F. Multiclass cancer classification in fresh frozen and formalin-fixed paraffin-embedded tissue by DigiWest multiplex protein analysis. Lab Invest. 100, 1288–1299 (2020).*

In dieser dritten Arbeit wurde geprüft, für welche Antikörper die DigiWest-Methode auf FFPE-Gewebe angewandt werden kann. Anschließend wurden diese Antikörper genutzt, um Proben von fünf Tumortypen zu klassifizieren und ihren Organursprung mithilfe eines SVM-Modells vorherzusagen. Dabei handelte es sich um Plattenepithelkarzinome des Kopf-Halses (HNSC), Plattenepithelkarzinome der Lunge (LUSC), Adenokarzinome der Lunge (LUAD), kolorektale Adenokarzinome (COAD) und Adenokarzinome des Pankreas (PAAD).

In einem ersten Schritt wurde getestet, welche der untersuchten 634 Antikörper sich für eine Anwendung auf FFPE-Gewebe mit der DigiWest-Methode eignen. Dazu wurden diese Antikörper zunächst in 16 FFPE-Tumorproben (3 HNSC, 4 LUSC, 3 LUAD, 3 COAD, 3 PAAD) gemessen. Bei 303 der 634 Antikörper wurde ein Signal in mehreren Proben detektiert, d. h. es konnte in mindestens vier Proben oder drei Proben desselben Tumortyps nachgewiesen werden. Bei diesen 303 Antikörpern wurde dann überprüft, ob es übereinstimmende Signale zwischen Frisch- und FFPE-Gewebe gibt. Dazu wurden sie in einem gepaarten Set aus 25 Frisch- und FFPE-Gewebeproben gemessen, das jeweils fünf Proben pro Tumortyp (HNSC, LUSC, LUAD, COAD, PAAD) umfasste. Für diese 303 Antikörper wurden dann die Korrelationskoeffizienten nach Pearson ( $R$ ) zwischen den Signalen in Frisch- und FFPE-Gewebe berechnet. Nach Korrektur für multiples Testen mit dem Benjamini-Hochberg ( $BH$ ) Verfahren zeigten 121 der 303 Antikörper (40 %) eine signifikante Korrelation ( $p-BH < 0,05$ ) zwischen Frischgewebe- und FFPE-Proben mit einem Korrelationskoeffizienten  $R \geq 0,47$ . Von diesen 121 Antikörpern wurden 19 ausgeschlossen, da die Signale nicht eindeutig erkennbar waren oder nicht beim erwarteten Molekulargewicht identifiziert wurden ( $> 20$  % verschoben). Daraus folgte, dass insgesamt 102 Antikörper für die Analyse in FFPE-Gewebe geeignet waren und zur weiteren Klassifikation der fünf Tumortypen verwendet werden konnten. In diesen 102 Antikörpern waren auch 14 Antikörper gegen Phosphoproteine enthalten. Bei 101 der 102 Antikörper lagen im Frischgewebe höhere mittlere Signalintensitäten ( $\log_2$ -Skala) als in den FFPE-Proben vor, zudem wurden die Signale durchschnittlich in einer größeren Anzahl an Proben detektiert (23 vs. 18 Proben).

Als Nächstes wurde untersucht, inwieweit sich die Expression der 102 Proteine zwischen den fünf Tumortypen unterscheidet und ob diese Proteine eine Klassifikation der Tumorproben nach ihrem Organursprung ermöglichen. Die DigiWest-Daten der 102 Antikörper in den

25 gepaarten Proben wurden jeweils für Frisch- und FFPE-Gewebe in einer hierarchischen Clusteranalyse zusammen mit einer Heatmap dargestellt. Dabei bilden die Tumorproben ein Cluster, die eine ähnliche Proteinexpression zeigen. Im Frischgewebe konnten relativ eindeutige Cluster für die einzelnen Tumortypen beobachtet werden. Nur ein LUSC zeigte eine größere Ähnlichkeit zu HNSC als zu den anderen LUSC, außerdem waren zwei LUAD näher bei den PAAD angeordnet als bei den anderen LUAD. Insgesamt bildeten im Frischgewebe zudem alle Plattenepithelkarzinome ein Cluster und alle Adenokarzinome waren in einem weiteren Cluster zusammengefasst. Im Gegensatz dazu ließen sich die verschiedenen Tumortypen in den FFPE-Proben nicht gut voneinander unterscheiden. Insbesondere die LUSC- und LUAD-Fälle zeigten häufig eine ähnliche Proteinexpression und waren nebeneinander angeordnet. Einzig die HNSC-Proben bildeten ein klar abgrenzbares Cluster, in dem viele Proteine stark exprimiert waren.

Anschließend wurde mithilfe der 102 Proteine ein *Support Vector Machine* (SVM)-Modell entwickelt, um die Proben entsprechend ihres Organursprungs zu klassifizieren. In dem gepaarten Set aus 25 Frischgewebe- und FFPE-Proben erfolgte die SVM-Klassifikation über eine wiederholte verschachtelte Kreuzvalidierung. Im Frischgewebe konnte eine Gesamtvorhersagegenauigkeit von 90,4 % (Standardabweichung von 5,4 % über 10 Wiederholungen) erreicht werden. Für alle Tumortypen wurden hohe individuelle Vorhersagegenauigkeiten erzielt. Am besten waren die Klassifikationsergebnisse für COAD und PAAD, deren Proben alle richtig eingeteilt wurden. Am geringsten war die Vorhersagegenauigkeit für LUAD, wo nur 78 % der Fälle richtig vorhergesagt wurden. Im FFPE-Gewebe wurde eine Gesamtvorhersagegenauigkeit von 77,6 % (Standardabweichung von 3,4 % über 10 Wiederholungen) erzielt, was etwas geringer ist als in Frischgewebe. Die meisten Fehler traten zwischen Plattenepithel- und Adenokarzinomen der Lunge auf. Es wurden nur 52 % der LUSC und 58 % der LUAD korrekt klassifiziert. Gleichzeitig wurden 40 % der LUSC als LUAD vorhergesagt und in reziproker Weise 18 % der LUAD als LUSC eingeteilt. Zusammen machte dies 11,6 % aller Fehler im FFPE-Gewebe aus. Andererseits wurden 100 % der PAAD, 98 % der HNSC und 80 % der COAD richtig klassifiziert.

Abschließend wurde der SVM-Klassifikator in einer unabhängigen Validierungskohorte von 25 FFPE-Proben angewandt. Dort konnten dieselben fünf Tumortypen mit einer Gesamtvorhersagegenauigkeit von 88 % klassifiziert werden. Einerseits wurden alle COAD-, HNSC- und LUAD-Proben richtig eingeteilt. Andererseits wurde ein LUSC fälschlicherweise als COAD vorhergesagt und jeweils ein PAAD als COAD und ein weiteres als LUAD fehlklassifiziert. Insgesamt wurden somit nur drei Proben falsch eingeteilt. Im Hinblick auf den Differenzierungsgrad wurden fünf der sieben schlecht differenzierten Proben und 17 der 18 mittelgradig differenzierten Fälle korrekt klassifiziert. Bei den schlecht differenzierten Tumorproben lagen im Durchschnitt etwas geringere Wahrscheinlichkeitswerte für die richtige Klasse vor als bei den mittelgradig differenzierten Fällen (durchschnittlich 0,45 vs. 0,61,  $p = 0,03$ ).

## 4 Diskussion

Insgesamt wurden im Rahmen dieser Arbeit drei unterschiedliche molekulare Verfahren zur Tumordiagnostik und -klassifikation untersucht, deren Relevanz und Grenzen im Folgenden diskutiert werden.

### 4.1 Nutzen der gezielten Tiefensequenzierung von Zellblöcken maligner Ergüsse

Molekulargenetische Untersuchungen haben einen wichtigen Stellenwert in der Tumordiagnostik, insbesondere um therapierelevante Mutationen zu identifizieren. In den untersuchten Zellblöcken der malignen Ergüsse konnte die NGS-basierte Multigen-Diagnostik erfolgreich durchgeführt werden. Bei 75 % der Fälle wurden übereinstimmende Mutationen zwischen dem Primärtumor und dem korrespondierenden Zellblock identifiziert [33]. Darunter befand sich auch eine *EGFR*-Mutation bei einer Patientin mit einem Adenokarzinom der Lunge, die auf das Ansprechen auf eine zielgerichtete Therapie hinweist. Da die Punktion maligner Ergüsse leicht durchführbar und wenig belastend für den Patienten ist, können die daraus gewonnenen Zellblöcke eine wichtige Quelle für eine gezielte Tiefensequenzierung darstellen, insbesondere wenn nicht ausreichend Tumorzellen aus histologischen Proben verfügbar sind, oder um (weitere) invasive Untersuchungen zu vermeiden.

Des Weiteren zeigten vier dieser Patienten neben gemeinsamen Mutationen noch zusätzliche Mutationen in dem Primärtumor und/oder dem Zellblock. Dies kann darauf hinweisen, dass sich Subklone des Tumors räumlich bzw. zeitlich unterschiedlich entwickelt haben. Die Analyse der Mutationen in Zellblöcken könnte nicht nur in der Primärdiagnostik nützlich sein, sondern auch um Resistenzmechanismen unter einer zielgerichteten Therapie zu erkennen.

Bei insgesamt drei Patienten wurden keine Mutationen in den Zellblöcken nachgewiesen, obwohl jeweils Mutationen in den korrespondierenden Gewebeproben identifiziert wurden. Bei diesen Zellblöcken war der Tumorzellgehalt gering ( $< 10\%$ ), zudem konnten keine Tumorzellcluster identifiziert werden. Möglicherweise war in diesen Zellblöcken der Anteil der tumorspezifischen DNA im Vergleich zur gesamten DNA zu gering und hat so zu falsch negativen Befunden geführt. Daher sollte der Tumorzellgehalt bei der gezielten Tiefensequenzierung von Zellblöcken beachtet werden und wenn möglich Zellblöcke mit einem hohen Tumorzellgehalt verwendet werden.

### 4.2 Bedeutung und Grenzen der Mutationsanalyse für die Tumorklassifikation

Die gleiche Treibermutation kann in verschiedenen Tumorentitäten auftreten, je nach Tumortyp aber unterschiedliche biologische und klinische Auswirkungen haben. Zum Beispiel

treten *BRAF V600E*-Mutationen unter anderem in Melanomen und kolorektalen Karzinomen auf [48]. Bei Melanomen ist eine Therapie mit BRAF-Inhibitoren häufig wirksam [49], selbst wenn es nach einiger Zeit zur Entwicklung von Resistenzen kommen kann [50]. Hingegen zeigt die gleiche Monotherapie mit BRAF-Inhibitoren keine Wirksamkeit bei kolorektalen Karzinomen [48]. Dies verdeutlicht, dass die Kenntnis des Organursprungs weiterhin entscheidend ist, um eine adäquate Therapie einzuleiten.

Um in diagnostisch unklaren Fällen die Bestimmung des Tumortyps zu ermöglichen, kann eine vergleichende Analyse der Mutationen von mehreren Proben eines Patienten nützlich sein. Zeigen bei einem Patienten zwei Proben ein ähnliches Mutationsprofil, ist es wahrscheinlich, dass es sich um den gleichen Ursprungstumor handelt. Hingegen weisen deutliche Unterschiede der Mutationsprofile entweder auf verschiedene Klone eines Tumors oder auf das Vorhandensein eines Zweittumors hin. Dies konnte auch bei einer Patientin in der gepaarten Analyse der Gewebeprobe und des korrespondierenden Zellblocks beobachtet werden [33]. Bei dieser Patientin entwickelte sich bei einem bekannten papillären Schilddrüsenkarzinom (*BRAF*-Mutation in der Gewebeprobe) im Verlauf Aszites, in dem dazugehörigen Zellblock wurde eine *KRAS*-Mutation nachgewiesen. Da sich die Mutationen zwischen beiden Proben unterschieden und die morphologischen und immunzytochemischen Untersuchungen des Zellblocks keine eindeutige Diagnose hinsichtlich des Tumortyps ermöglichten, wurde ein Zweittumor vermutet. Weitere Untersuchungen zeigten dann das Vorliegen eines peritoneal metastasierten Ovarialkarzinoms mit dergleichen *KRAS*-Mutation.

Allerdings ermöglicht eine vergleichende Analyse der Mutationen nicht in allen Fällen eine Identifikation des Tumortyps. Einerseits deckt das verwendete NGS-Panel nicht immer Mutationen ab, die es erlauben würden, zwei unterschiedliche Tumortypen voneinander abzugrenzen. Andererseits ist eine vergleichende Analyse dann nicht durchführbar, wenn nicht genügend Tumormaterial aus beiden Proben verfügbar ist. So kann eine Probe aufgebraucht sein oder extern gelagert werden. Im Falle des CUP-Syndroms kann kein Primärtumor identifiziert werden, der für eine vergleichende Analyse nötig wäre. Des Weiteren kann die Qualität der DNA durch Entkalkungen oder eine mehrjährige Lagerung beeinträchtigt sein.

Daher wurden in dieser Arbeit auch zwei Methoden zur Tumorklassifikation untersucht, die auf der Verwendung von tumortypspezifischen Signaturen beruhen. Über die Analyse von DNA-Methylierungs- oder Proteinprofilen kombiniert mit Methoden des maschinellen Lernens konnte so auf eine vergleichende Analyse mehrerer Tumorproben verzichtet werden, ohne dass die Aussagekraft der Ergebnisse eingeschränkt wurde.

### 4.3 Klinische Relevanz des DNA-Methylierungs-Tumorklassifikators

Der Klassifikator basierend auf der Analyse von DNA-Methylierungsprofilen wurde darauf ausgerichtet Metastasen eines Plattenepithelkarzinoms des Kopf-Halses (HNSC) von

Plattenepithelkarzinomen der Lunge (LUSC) zu unterscheiden. Dazu wurden drei Modelle des maschinellen Lernens miteinander verglichen. Die besten Ergebnisse erreichten die neuronalen Netzwerke (NN) mit einer Vorhersagegenauigkeit von insgesamt 96,4 % [34]. Aber auch die anderen beiden Modelle erzielten mit 95,7 % Vorhersagegenauigkeit für den SVM-Klassifikator und 87,8 % für das RF-Modell bessere Ergebnisse als vorherige Studien zur Unterscheidung von HNSC und LUSC. In einer früheren Arbeit von Lal et al. konnten beide Plattenepithelkarzinomtypen über die Untersuchung von Genexpressionsprofilen (RNA) mit einer Vorhersagegenauigkeit von 82,9 % zugeordnet werden [51]. In einer weiteren Studie wurde ein SVM-Klassifikator entwickelt, der basierend auf der Analyse von Proteinprofilen mit Massenspektrometrie LUSC von HNSC mit einer Vorhersagegenauigkeit von 86,8 % unterscheiden konnte [52].

Zudem konnte in der vorliegenden Arbeit die Vorhersagegenauigkeit von 96,4 % der NN-Modells noch weiter gesteigert werden, indem nur die Proben berücksichtigt wurden, die mit hoher Wahrscheinlichkeit einer Klasse zugeordnet werden konnten. Für den NN-Klassifikator lag in der Validierungskohorte bei 92,1 % der Proben ein Wahrscheinlichkeitswert von mindestens 0,95 vor. Wenn nur diese Proben eingeschlossen wurden, erreichte das NN-Modell eine Vorhersagegenauigkeit von 99 % [34]. Außerdem wurde der Klassifikator nicht nur mit einer großen Anzahl an Proben trainiert und in einer unabhängigen Kohorte aus Primärtumoren validiert, sondern zusätzlich auch bei Metastasen angewandt. In einer klinischen Kohorte konnten Metastasen eines HNSC zuverlässig von LUSC unterschieden werden. Zwischen diesen beiden Gruppen bestanden signifikante Unterschiede im krankheitsspezifischen Überleben, was die prognostische und klinische Relevanz der Klassifikationsergebnisse des NN-Modells verdeutlicht.

Neben den sehr präzisen Klassifikationsergebnissen hat die Analyse der DNA-Methylierungen den Vorteil, dass sie auf FFPE-Proben durchführbar und somit auch für klinische Proben geeignet ist. Des Weiteren kann die Analyse in einem für die Diagnostik angemessenen Zeitraum durchgeführt werden, da die Ergebnisse normalerweise innerhalb einer Woche vorliegen. Zugleich sind die Kosten mit ca. 300 EUR vergleichsweise günstig. Die Anwendbarkeit und klinische Relevanz der DNA-Methylierungsanalyse sollte in Zukunft noch weiter klinisch prospektiv validiert werden.

#### **4.4 Nutzen und Grenzen des Proteinprofil-basierten Tumorklassifikators**

Die Tumorklassifikation durch Analyse von Proteinprofilen mit der DigiWest-Methode wurde dazu entwickelt, Proben von fünf Tumortypen (HNSC, LUSC, LUAD, COAD, PAAD) entsprechend ihres Organursprungs zu klassifizieren. Da klinische Proben standardmäßig in Formalin fixiert und in Paraffin eingebettet werden, wurde diese Klassifikation auch für eine Anwendung auf FFPE-Gewebe getestet. Innerhalb der 634 untersuchten Antikörper wurden 102 Antikörper identifiziert, die sich für eine Applikation auf FFPE-Gewebe mit der DigiWest-Methode eignen. Diese 102 Antikörper wurden in 25 gepaarten Tumorproben

aus Frisch- und FFPE-Gewebe gemessen, dadurch konnten die Klassifikationsergebnisse zwischen beiden Probenmaterialien direkt miteinander verglichen werden.

Vorherige Studien, die mehrere Tumortypen mithilfe von Proteinprofilen aus Gewebeproben klassifizierten, wurden hingegen ausschließlich auf Frischgewebe durchgeführt. In zwei früheren Studien, die beide Vorhersagegenauigkeiten von 82 % erzielten, wurden jeweils sechs Arten von Adenokarzinomen mithilfe von Matrix-unterstützter Laser-Desorption/Ionisierung (MALDI)-Massenspektrometrie untersucht [15, 26]. In einer weiteren Arbeit konnten zehn verschiedene Tumortypen mithilfe von RPPA mit einer Vorhersagegenauigkeit von 93,6 % eingeteilt werden [27]. Insgesamt erzielten diese drei Studien somit ähnliche Ergebnisse wie die vorliegende Arbeit, in der eine Vorhersagegenauigkeit von 90,4 % in Frischgewebe mit dem SVM-Modell erreicht wurde [35]. Zusätzlich wurde die SVM-Klassifikation der fünf Tumortypen noch in den 25 korrespondierenden FFPE-Proben durchgeführt. Dort konnten insgesamt 77,6 % der Fälle richtig zugeordnet werden. Die besseren Klassifikationsergebnisse in Frischgewebe sind möglicherweise durch eine höhere Sensitivität der DigiWest-Methode in Frischgewebe im Vergleich zu den FFPE-Proben zu erklären. Im Frischgewebe lagen meist höhere mittlere Signalintensitäten vor, außerdem wurden die Signale eines Antikörpers im Durchschnitt in einer größeren Anzahl an Proben detektiert als in den FFPE-Proben. In der unüberwachten hierarchischen Clusteranalyse konnte auch eine bessere Abgrenzung der einzelnen Tumortypen in Frischgewebe im Vergleich zu den FFPE-Proben beobachtet werden.

Darüber hinaus wurde die SVM-Klassifikation in einer unabhängigen Kohorte aus 25 FFPE-Proben durchgeführt, in denen die gleichen 102 Antikörper analysiert wurden. Dort erzielte der SVM-Klassifikator eine Vorhersagegenauigkeit von 88 %. Diese Ergebnisse sind etwas besser als in den 25 FFPE-Proben des gepaarten Sets (77,6 %). Vermutlich lässt sich das darauf zurückführen, dass etwas mehr Proben zum Trainieren des SVM-Modells herangezogen wurden (25 Proben anstelle der 20 bei der verschachtelten Kreuzvalidierung des gepaarten Sets). Hinsichtlich des Differenzierungsgrads war es auch möglich schlecht differenzierte Proben korrekt zu klassifizieren. Obwohl dort durchschnittlich etwas niedrigere Wahrscheinlichkeitswerte für die korrekte Klasse vorlagen als in den mittelgradig differenzierten Fällen, konnten insgesamt fünf der sieben schlecht differenzierten Proben richtig eingeteilt werden [35].

Ein wichtiger Vorteil der DigiWest-Methode liegt darin, dass sie sich auch für die Untersuchung klinischer Proben eignet. Es werden nur geringe Mengen an Probenmaterial benötigt, um eine große Anzahl an Proteinen zu analysieren. So waren 10 µg Gesamtprotein ausreichend für die Analyse von 102 Antikörpern, bzw. 20 µg für 303 Antikörper. Des Weiteren werden die Proteine wie bei einem klassischen Western Blot nach ihrem Molekulargewicht aufgetrennt. Dadurch können unspezifische Signale besser erkannt werden als in RPPA [23]. Allerdings kann es wie bei jeder Antikörper-basierten Methode trotzdem zu Kreuzreaktionen kommen, wie es die Kreuzreaktion des Anti-Rabbit-Sekundärantikörpers

in den HNSC-Proben bei 50 kDa zeigt. Eine sorgfältige Auswahl und Auswertung der Antikörpersignale ist deshalb entscheidend. Insgesamt wurden 102 Antikörper identifiziert, die sich für eine Anwendung auf FFPE-Gewebe mit der DigiWest-Methode eignen, darunter auch 14 Antikörper gegen Phosphoproteine [35]. Neben der Verwendung zur Tumorklassifikation könnte die DigiWest-Methode in Zukunft auch auf FFPE-Proben angewandt werden, um einen besseren Einblick in Proteinaktivierungen und Signalkaskaden zu gewinnen. Möglicherweise können so auch neue Biomarker oder therapeutisch relevante Zielmoleküle identifiziert werden.

Insgesamt erzielte der Proteinprofil-basierte SVM-Klassifikator sowohl in Frischgewebe als auch in den FFPE-Proben gute Ergebnisse, um die Proben der fünf Tumortypen entsprechend ihres Organursprungs zu klassifizieren. Allerdings wurden nur 20 Proben in dem gepaarten Set bzw. 25 Proben in der unabhängigen Kohorte zum Training herangezogen. In Zukunft wäre es daher sinnvoll, die Anzahl der Proben zu erhöhen, um verschiedene Subtypen innerhalb eines Tumortyps umfassender abbilden zu können. Möglicherweise würde dies die Klassifikationsergebnisse verbessern. Da bisher nur Primärtumore eingeschlossen wurden, sollte die Klassifikation zudem auf Metastasen ausgeweitet werden. Um eine Klassifikation von CUP-Fällen durchzuführen, wäre ein Training des Klassifikators mit weiteren Tumortypen nützlich. Da es bei CUP-Fällen allerdings per Definition nicht möglich ist, den Primärtumor zu identifizieren [8], kann die Vorhersagegenauigkeit des Klassifikators dabei nicht überprüft werden. Es müsste dann eine prospektive Studie an einer größeren klinischen Kohorte durchgeführt werden, um zu überprüfen, inwieweit die Klassifikationsergebnisse und die darauf beruhenden Therapieentscheidungen einen klinischen Nutzen für die Patienten haben.

#### **4.5 Vergleich der DNA-Methylierungs- und Proteinprofil-basierten Klassifikatoren**

Sowohl über die Analyse von DNA-Methylierungs- als auch von Proteinprofilen können tumortypspezifische molekulare Signaturen identifiziert werden, die es ermöglichen den Organursprung einer Tumorprobe mithilfe von maschinellem Lernen vorherzusagen. Beide Klassifikatoren erzielten dabei gute Klassifikationsergebnisse und sind auf FFPE-Proben anwendbar, wodurch sie auch für eine klinische Anwendung nützlich sind.

Beide Modelle haben gemeinsam, dass sie HNSC- und LUSC-Proben klassifizieren. Obwohl in den FFPE-Proben des Proteinprofil-basierten SVM-Modells keine Fehlklassifikationen zwischen diesen beiden Tumortypen auftraten, ist der Klassifikator basierend auf der Analyse von DNA-Methylierungen für diese spezielle Fragestellung vorzuziehen. Er wurde gezielt darauf ausgerichtet, diese beiden Tumortypen zu unterscheiden und mit deutlich mehr Proben trainiert und validiert. Zudem wurde er auch für eine Anwendung bei Metastasen getestet und erzielte dabei sehr hohe Vorhersagegenauigkeiten.

Insgesamt lassen sich die Vorhersagegenauigkeiten der beiden Klassifikatoren allerdings

nur bedingt direkt miteinander vergleichen, da das Modell basierend auf der Analyse von Proteinprofilen fünf Tumortypen klassifiziert, während der Klassifikator basierend auf der Untersuchung von DNA-Methylierungsprofilen ausschließlich zwei Tumortypen einteilt. Die Klassifikation von mehreren Tumortypen über maschinelles Lernen ist in der Regel komplizierter als die Unterscheidung von nur zwei Tumortypen. Während die Analyse der DNA-Methylierungen zwar etwas schneller und günstiger durchgeführt werden kann als die DigiWest-Methode, hat diese den Vorteil, dass über die Analyse der Proteine gleichzeitig noch ein besseres Verständnis der Signalkaskaden, der Funktion der Zelle und damit zukünftig therapeutisch relevanter Zielstrukturen gewonnen werden kann.

#### **4.6 Fazit**

Zusammenfassend konnte in allen drei Studien ein Nutzen der verschiedenen molekularen Verfahren in der Tumordiagnostik bzw. -klassifikation gezeigt werden. Sind bei einem Patienten mehrere Tumore vorhanden, kann über eine vergleichende Analyse der Mutationsprofile festgestellt werden, ob es sich um den gleichen Ursprungstumor handelt. Zudem kann eine gezielte Tiefensequenzierung auch in Zellblöcken maligner Ergüsse durchgeführt werden, um therapierelevante Mutationen zu identifizieren. Des Weiteren kann der Organursprung eines Tumors sowohl durch die Analyse von DNA-Methylierungs- als auch von Proteinprofilen mithilfe von Methoden des maschinellen Lernens bestimmt werden, wobei beide Modelle gute Klassifikationsergebnisse erzielen. Die DigiWest-Methode könnte darüber hinaus in Zukunft auch auf FFPE-Proben angewandt werden, um Biomarker zu identifizieren oder bessere Einblicke in die Funktion der Zelle zu erlangen. Zuletzt eignet sich insbesondere der DNA-Methylierungs-basierte Klassifikator für eine Anwendung in der klinischen Diagnostik. Dieses Modell erzielt sehr hohe Vorhersagegenauigkeiten und ermöglicht somit eine präzise Unterscheidung zwischen Metastasen eines HNSC und Zweitumoren in der Lunge. Dies könnte bei den betroffenen Patienten die Auswahl einer adäquaten Therapie erleichtern. Während hier die genomischen, epigenetischen und proteomischen Verfahren voneinander unabhängig eingesetzt wurden, sollten zukünftige Studien untersuchen, ob Kombinationen der verschiedenen Datenmodalitäten zu noch besseren und robusteren Diagnosen führen können.

## Literatur

1. Krebs in Deutschland für 2015/2016. 12. Ausgabe. *Robert Koch-Institut und die Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (Hrsg)*. Berlin. (2019).
2. Böcker, W., Denk, H., Heitz, P., Moch, H., Höfler, G. & Kreipe, H. *Lehrbuch Pathologie* 6. Aufl. (Elsevier Health Sciences, München, 2019).
3. Pereira, T. C., Share, S. M., Magalhães, A. V. & Silverman, J. F. Can we tell the site of origin of metastatic squamous cell carcinoma? An immunohistochemical tissue microarray study of 194 cases. *Appl Immunohistochem Mol Morphol.* **19**, 10–14 (2011).
4. Duprez, F., Berwouts, D., De Neve, W., Bonte, K., Boterberg, T., Deron, P., Huvenne, W., Rottey, S. & Mareel, M. Distant metastases in head and neck cancer. *Head Neck.* **39**, 1733–1743 (2017).
5. Cohen, E. E. W., Bell, R. B., Bifulco, C. B., Burtness, B., Gillison, M. L., Harrington, K. J., Le, Q. T., Lee, N. Y., Leidner, R., Lewis, R. L., Licitra, L., Mehanna, H., Mell, L. K., Raben, A., Sikora, A. G., Uppaluri, R., Whitworth, F., Zandberg, D. P. & Ferris, R. L. The Society for Immunotherapy of Cancer consensus statement on immunotherapy for the treatment of squamous cell carcinoma of the head and neck (HNSCC). *J Immunother Cancer.* **7**, 184 (2019).
6. Erkal, H. S., Mendenhall, W. M., Amdur, R. J., Villaret, D. B. & Stringer, S. P. Synchronous and metachronous squamous cell carcinomas of the head and neck mucosal sites. *J Clin Oncol.* **19**, 1358–1362 (2001).
7. Postmus, P. E., Kerr, K. M., Oudkerk, M., Senan, S., Waller, D. A., Vansteenkiste, J., Escriu, C. & Peters, S. Early and locally advanced non-small-cell lung cancer (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* **28**, iv1–iv21 (2017).
8. Pavlidis, N. & Pentheroudakis, G. Cancer of unknown primary site. *Lancet.* **379**, 1428–1435 (2012).
9. Economopoulou, P., Mountzios, G., Pavlidis, N. & Pentheroudakis, G. Cancer of Unknown Primary origin in the genomic era: Elucidating the dark box of cancer. *Cancer Treat Rev.* **41**, 598–604 (2015).
10. Pentheroudakis, G., Golfinopoulos, V. & Pavlidis, N. Switching benchmarks in cancer of unknown primary: from autopsy to microarray. *Eur J Cancer.* **43**, 2026–2036 (2007).

11. Bochtler, T., Endris, V., Leichsenring, J., Reiling, A., Neumann, O., Volckmar, A. L., Kirchner, M., Allgäuer, M., Schirmacher, P., Krämer, A. & Stenzinger, A. Comparative genetic profiling aids diagnosis and clinical decision making in challenging cases of CUP syndrome. *Int J Cancer*. **145**, 2963–2973 (2019).
12. Erlander, M. G., Ma, X. J., Kesty, N. C., Bao, L., Salunga, R. & Schnabel, C. A. Performance and clinical evaluation of the 92-gene real-time PCR assay for tumor classification. *J Mol Diagn*. **13**, 493–503 (2011).
13. Rosenfeld, N., Aharonov, R., Meiri, E., Rosenwald, S., Spector, Y., Zepeniuk, M., Benjamin, H., Shabes, N., Tabak, S., Levy, A., Lebanony, D., Goren, Y., Silberschein, E., Targan, N., Ben-Ari, A., Gilad, S., Sion-Vardy, N., Tobar, A., Feinmesser, M., Kharenko, O., Nativ, O., Nass, D., Perelman, M., Yosepovich, A., Shalmon, B., Polak-Charcon, S., Fridman, E., Avniel, A., Bentwich, I., Bentwich, Z., Cohen, D., Chajut, A. & Barshack, I. MicroRNAs accurately identify cancer tissue origin. *Nat Biotechnol*. **26**, 462–469 (2008).
14. Capper, D., Jones, D. T. W., Sill, M., Hovestadt, V., Schrimpf, D., Sturm, D., Koelsche, C., Sahm, F., Chavez, L., Reuss, D. E., Kratz, A., Wefers, A. K., Huang, K., Pajtler, K. W., Schweizer, L., Stichel, D., Olar, A., Engel, N. W., Lindenberg, K., Harter, P. N., Braczynski, A. K., Plate, K. H., Dohmen, H., Garvalov, B. K., Coras, R., Hölsken, A., Hewer, E., Bewerunge-Hudler, M., Schick, M., Fischer, R., Beschorner, R., Schittenhelm, J., Staszewski, O., Wani, K., Varlet, P., Pages, M., Temming, P., Lohmann, D., Selt, F., Witt, H., Milde, T., Witt, O., Aronica, E., Giangaspero, F., Rushing, E., Scheurlen, W., Geisenberger, C., Rodriguez, F. J., Becker, A., Preusser, M., Haberler, C., Bjerkvig, R., Cryan, J., Farrell, M., Deckert, M., Hench, J., Frank, S., Serrano, J., Kannan, K., Tsirigos, A., Brück, W., Hofer, S., Brehmer, S., Seiz-Rosenhagen, M., Hänggi, D., Hans, V., Rozsnoki, S., Hansford, J. R., Kohlhof, P., Kristensen, B. W., Lechner, M., Lopes, B., Mawrin, C., Ketter, R., Kulozik, A., Khatib, Z., Heppner, F., Koch, A., Jouvet, A., Keohane, C., Mühleisen, H., Mueller, W., Pohl, U., Prinz, M., Benner, A., Zapatka, M., Gottardo, N. G., Driever, P. H., Kramm, C. M., Müller, H. L., Rutkowski, S., von Hoff, K., Frühwald, M. C., Gnekow, A., Fleischhack, G., Tippelt, S., Calaminus, G., Monoranu, C. M., Perry, A., Jones, C., Jacques, T. S., Radlwimmer, B., Gessi, M., Pietsch, T., Schramm, J., Schackert, G., Westphal, M., Reifenberger, G., Wesseling, P., Weller, M., Collins, V. P., Blümcke, I., Bendszus, M., Debus, J., Huang, A., Jabado, N., Northcott, P. A., Paulus, W., Gajjar, A., Robinson, G. W., Taylor, M. D., Jaunmuktane, Z., Ryzhova, M., Platten, M., Unterberg, A., Wick, W., Karajannis, M. A., Mittelbronn, M., Acker, T., Hartmann, C., Aldape, K., Schüller, U., Buslei, R., Lichter, P., Kool, M., Herold-Mende, C., Ellison, D. W., Hasselblatt, M., Snuderl, M., Brandner, S., Korshunov, A.,

- von Deimling, A. & Pfister, S. M. DNA methylation-based classification of central nervous system tumours. *Nature*. **555**, 469–474 (2018).
15. Bloom, G. C., Eschrich, S., Zhou, J. X., Coppola, D. & Yeatman, T. J. Elucidation of a protein signature discriminating six common types of adenocarcinoma. *Int J Cancer*. **120**, 769–775 (2007).
  16. Heim, D., Montavon, G., Hufnagl, P., Müller, K. R. & Klauschen, F. Computational analysis reveals histotype-dependent molecular profile and actionable mutation effects across cancers. *Genome Med*. **10**, 83 (2018).
  17. Fernandez, A. F., Assenov, Y., Martin-Subero, J. I., Balint, B., Siebert, R., Taniguchi, H., Yamamoto, H., Hidalgo, M., Tan, A. C., Galm, O., Ferrer, I., Sanchez-Cespedes, M., Villanueva, A., Carmona, J., Sanchez-Mut, J. V., Berdasco, M., Moreno, V., Capella, G., Monk, D., Ballestar, E., Ropero, S., Martinez, R., Sanchez-Carbayo, M., Prosper, F., Agirre, X., Fraga, M. F., Graña, O., Perez-Jurado, L., Mora, J., Puig, S., Prat, J., Badimon, L., Puca, A. A., Meltzer, S. J., Lengauer, T., Bridgewater, J., Bock, C. & Esteller, M. A DNA methylation fingerprint of 1628 human samples. *Genome Res*. **22**, 407–419 (2012).
  18. Kulis, M. & Esteller, M. DNA methylation and cancer. *Adv Genet*. **70**, 27–56 (2010).
  19. Pan, Y., Liu, G., Zhou, F., Su, B. & Li, Y. DNA methylation profiles in cancer diagnosis and therapeutics. *Clin Exp Med*. **18**, 1–14 (2018).
  20. Koelsche, C., Hartmann, W., Schrimpf, D., Stichel, D., Jabar, S., Ranft, A., Reuss, D. E., Sahm, F., Jones, D. T. W., Bewerunge-Hudler, M., Trautmann, M., Klingebiel, T., Vokuhl, C., Gessler, M., Wardelmann, E., Petersen, I., Baumhoer, D., Flucke, U., Antonescu, C., Esteller, M., Fröhling, S., Kool, M., Pfister, S. M., Mechttersheimer, G., Dirksen, U. & von Deimling, A. Array-based DNA-methylation profiling in sarcomas with small blue round cell histology provides valuable diagnostic information. *Mod Pathol*. **31**, 1246–1256 (2018).
  21. Moran, S., Martínez-Cardús, A., Sayols, S., Musulén, E., Balañá, C., Estival-Gonzalez, A., Moutinho, C., Heyn, H., Diaz-Lagares, A., de Moura, M. C., Stella, G. M., Comoglio, P. M., Ruiz-Miró, M., Matias-Guiu, X., Pazo-Cid, R., Antón, A., Lopez-Lopez, R., Soler, G., Longo, F., Guerra, I., Fernandez, S., Assenov, Y., Plass, C., Morales, R., Carles, J., Bowtell, D., Mileshkin, L., Sia, D., Tothill, R., Tabernero, J., Llovet, J. M. & Esteller, M. Epigenetic profiling to classify cancer of unknown primary: a multicentre, retrospective analysis. *Lancet Oncol*. **17**, 1386–1395 (2016).
  22. Pandey, A. & Mann, M. Proteomics to study genes and genomes. *Nature*. **405**, 837–846 (2000).

23. Akbani, R., Becker, K. F., Carragher, N., Goldstein, T., de Koning, L., Korf, U., Liotta, L., Mills, G. B., Nishizuka, S. S., Pawlak, M., Petricoin, E. F., Pollard, H. B., Serrels, B. & Zhu, J. Realizing the promise of reverse phase protein arrays for clinical, translational, and basic research: a workshop report: the RPPA (Reverse Phase Protein Array) society. *Mol Cell Proteomics*. **13**, 1625–1643 (2014).
24. Treindl, F., Ruprecht, B., Beiter, Y., Schultz, S., Döttinger, A., Staebler, A., Joos, T. O., Kling, S., Poetz, O., Fehm, T., Neubauer, H., Kuster, B. & Templin, M. F. A bead-based western for high-throughput cellular signal transduction analyses. *Nat Commun*. **7**, 12852 (2016).
25. Kriegsmann, J., Kriegsmann, M. & Casadonte, R. MALDI TOF imaging mass spectrometry in clinical pathology: a valuable tool for cancer diagnostics (review). *Int J Oncol*. **46**, 893–906 (2015).
26. Meding, S., Nitsche, U., Balluff, B., Elsner, M., Rauser, S., Schöne, C., Nipp, M., Maak, M., Feith, M., Ebert, M. P., Friess, H., Langer, R., Höfler, H., Zitzelsberger, H., Rosenberg, R. & Walch, A. Tumor classification of six common cancer types based on proteomic profiling by MALDI imaging. *J Proteome Res*. **11**, 1996–2003 (2012).
27. Zhang, P. W., Chen, L., Huang, T., Zhang, N., Kong, X. Y. & Cai, Y. D. Classifying ten types of major cancers based on reverse phase protein array profiles. *PLoS ONE*. **10**, e0123147 (2015).
28. Shabihkhani, M., Lucey, G. M., Wei, B., Mareninov, S., Lou, J. J., Vinters, H. V., Singer, E. J., Cloughesy, T. F. & Yong, W. H. The procurement, storage, and quality assurance of frozen blood and tissue biospecimens in pathology, biorepository, and biobank settings. *Clin Biochem*. **47**, 258–266 (2014).
29. Werner, M., Chott, A., Fabiano, A. & Battifora, H. Effect of formalin tissue fixation and processing on immunohistochemistry. *Am J Surg Pathol*. **24**, 1016–1019 (2000).
30. Broeckx, V., Boonen, K., Pringels, L., Sagaert, X., Prenen, H., Landuyt, B., Schoofs, L. & Maes, E. Comparison of multiple protein extraction buffers for GeLC-MS/MS proteomic analysis of liver and colon formalin-fixed, paraffin-embedded tissues. *Mol Biosyst*. **12**, 553–565 (2016).
31. Do, H. & Dobrovic, A. Sequence artifacts in DNA from formalin-fixed tissues: causes and strategies for minimization. *Clin Chem*. **61**, 64–71 (2015).
32. Nambirajan, A. & Jain, D. Cell blocks in cytopathology: An update. *Cytopathology*. **29**, 505–524 (2018).

33. Leichsenring, J., Volckmar, A. L., Kirchner, M., Kazdal, D., Kriegsmann, M., Stögbauer, F., Bockmayr, T., Klauschen, F., Herth, F. J. F., Penzel, R., Warth, A., Schirmacher, P., Endris, V. & Stenzinger, A. Targeted deep sequencing of effusion cytology samples is feasible, informs spatiotemporal tumor evolution, and has clinical and diagnostic utility. *Genes Chromosomes Cancer*. **57**, 70–79 (2018).
34. Jurmeister, P., Bockmayr, M., Seegerer, P., Bockmayr, T., Treue, D., Montavon, G., Vollbrecht, C., Arnold, A., Teichmann, D., Bressemer, K., Schüller, U., von Laffert, M., Müller, K. R., Capper, D. & Klauschen, F. Machine learning analysis of DNA methylation profiles distinguishes primary lung squamous cell carcinomas from head and neck metastases. *Sci Transl Med*. **11** (2019).
35. Bockmayr, T., Erdmann, G., Treue, D., Jurmeister, P., Schneider, J., Arndt, A., Heim, D., Bockmayr, M., Sachse, C. & Klauschen, F. Multiclass cancer classification in fresh frozen and formalin-fixed paraffin-embedded tissue by DigiWest multiplex protein analysis. *Lab Invest*. **100**, 1288–1299 (2020).
36. Rothberg, J. M., Hinz, W., Rearick, T. M., Schultz, J., Mileski, W., Davey, M., Leamon, J. H., Johnson, K., Milgrew, M. J., Edwards, M., Hoon, J., Simons, J. F., Marran, D., Myers, J. W., Davidson, J. F., Branting, A., Nobile, J. R., Puc, B. P., Light, D., Clark, T. A., Huber, M., Branciforte, J. T., Stoner, I. B., Cawley, S. E., Lyons, M., Fu, Y., Homer, N., Sedova, M., Miao, X., Reed, B., Sabina, J., Feierstein, E., Schorn, M., Alanjary, M., Dimalanta, E., Dressman, D., Kasinskas, R., Sokolsky, T., Fidanza, J. A., Namsaraev, E., McKernan, K. J., Williams, A., Roth, G. T. & Bustillo, J. An integrated semiconductor device enabling non-optical genome sequencing. *Nature*. **475**, 348–352 (2011).
37. Budczies, J., Bockmayr, M., Treue, D., Klauschen, F. & Denkert, C. Semiconductor sequencing: how many flows do you need? *Bioinformatics*. **31**, 1199–1203 (2015).
38. Pidsley, R., Zotenko, E., Peters, T. J., Lawrence, M. G., Risbridger, G. P., Molloy, P., Van Djik, S., Muhlhausler, B., Stirzaker, C. & Clark, S. J. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol*. **17**, 208 (2016).
39. Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., Molloy, P. L. & Paul, C. L. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci USA*. **89**, 1827–1831 (1992).
40. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Series B Stat Methodol*. **57**, 289–300 (1995).

41. James, G., Witten, D., Hastie, T. & Tibshirani, R. *An Introduction to Statistical Learning: With Applications in R* (Springer, New York, 2013).
42. Varma, S. & Simon, R. Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*. **7**, 91 (2006).
43. Krstajic, D., Buturovic, L. J., Leahy, D. E. & Thomas, S. Cross-validation pitfalls when selecting and assessing regression and classification models. *J Cheminform*. **6**, 10 (2014).
44. Aggarwal, C. C. *Neural Networks and Deep Learning - A Textbook* (Springer, Cham, 2018).
45. Li, B. T., Shen, R., Buonocore, D., Olah, Z. T., Ni, A., Ginsberg, M. S., Ulaner, G. A., Offin, M., Feldman, D., Hembrough, T., Cecchi, F., Schwartz, S., Pavlakakis, N., Clarke, S., Won, H. H., Brzostowski, E. B., Riely, G. J., Solit, D. B., Hyman, D. M., Drilon, A., Rudin, C. M., Berger, M. F., Baselga, J., Scaltriti, M., Arcila, M. E. & Kris, M. G. Ado-Trastuzumab Emtansine for Patients With HER2-Mutant Lung Cancers: Results From a Phase II Basket Trial. *J Clin Oncol*. **36**, 2532–2537 (2018).
46. Wang, Y., Jiang, T., Qin, Z., Jiang, J., Wang, Q., Yang, S., Rivard, C., Gao, G., Ng, T. L., Tu, M. M., Yu, H., Ji, H., Zhou, C., Ren, S., Zhang, J., Bunn, P., Doebele, R. C., Camidge, D. R. & Hirsch, F. R. HER2 exon 20 insertions in non-small-cell lung cancer are sensitive to the irreversible pan-HER receptor tyrosine kinase inhibitor pyrotinib. *Ann Oncol*. **30**, 447–455 (2019).
47. Karapetis, C. S., Khambata-Ford, S., Jonker, D. J., O’Callaghan, C. J., Tu, D., Tebbutt, N. C., Simes, R. J., Chalchal, H., Shapiro, J. D., Robitaille, S., Price, T. J., Shepherd, L., Au, H. J., Langer, C., Moore, M. J. & Zalcberg, J. R. K-ras mutations and benefit from cetuximab in advanced colorectal cancer. *N Engl J Med*. **359**, 1757–1765 (2008).
48. Hyman, D. M., Puzanov, I., Subbiah, V., Faris, J. E., Chau, I., Blay, J. Y., Wolf, J., Raje, N. S., Diamond, E. L., Hollebecque, A., Gervais, R., Elez-Fernandez, M. E., Italiano, A., Hofheinz, R. D., Hidalgo, M., Chan, E., Schuler, M., Lasserre, S. F., Makrutzki, M., Sirzen, F., Veronese, M. L., Tabernero, J. & Baselga, J. Vemurafenib in Multiple Nonmelanoma Cancers with BRAF V600 Mutations. *N Engl J Med*. **373**, 726–736 (2015).
49. Sosman, J. A., Kim, K. B., Schuchter, L., Gonzalez, R., Pavlick, A. C., Weber, J. S., McArthur, G. A., Hutson, T. E., Moschos, S. J., Flaherty, K. T., Hersey, P., Kefford, R., Lawrence, D., Puzanov, I., Lewis, K. D., Amaravadi, R. K., Chmielowski, B., Lawrence, H. J., Shyr, Y., Ye, F., Li, J., Nolop, K. B., Lee, R. J., Joe, A. K. & Ribas, A. Survival in BRAF V600-mutant advanced melanoma treated with vemurafenib. *N Engl J Med*. **366**, 707–714 (2012).

50. Kakadia, S., Yarlagadda, N., Awad, R., Kundranda, M., Niu, J., Naraev, B., Mina, L., Dragovich, T., Gimbel, M. & Mahmoud, F. Mechanisms of resistance to BRAF and MEK inhibitors and clinical update of US Food and Drug Administration-approved targeted therapy in advanced melanoma. *Onco Targets Ther.* **11**, 7095–7107 (2018).
51. Lal, A., Panos, R., Marjanovic, M., Walker, M., Fuentes, E., Kubicek, G. J., Henner, W. D., Buturovic, L. J. & Halks-Miller, M. A gene expression profile test to resolve head & neck squamous versus lung squamous cancers. *Diagn Pathol.* **8**, 44 (2013).
52. Bohnenberger, H., Kaderali, L., Ströbel, P., Yepes, D., Plessmann, U., Dharia, N. V., Yao, S., Heydt, C., Merkelbach-Bruse, S., Emmert, A., Hoffmann, J., Bodemeyer, J., Reuter-Jessen, K., Lois, A. M., Dröge, L. H., Baumeister, P., Walz, C., Biggemann, L., Walter, R., Häupl, B., Comoglio, F., Pan, K. T., Scheich, S., Lenz, C., Küffer, S., Bremmer, F., Kitz, J., Sitte, M., Beißbarth, T., Hinterthaler, M., Sebastian, M., Lotz, J., Schildhaus, H. U., Wolff, H., Danner, B. C., Brandts, C., Büttner, R., Canis, M., Stegmaier, K., Serve, H., Urlaub, H. & Oellerich, T. Comparative proteomics reveals a diagnostic signature for pulmonary head-and-neck cancer metastasis. *EMBO Mol Med.* **10** (2018).

# Eidesstattliche Erklärung

„Ich, Teresa Bockmayr, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema: „Proteogenomische Multiplexanalyse für die Tumordiagnostik“ selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.

Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren/innen beruhen, sind als solche in korrekter Zitierung kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen) werden von mir verantwortet.

Ich versichere ferner, dass ich die in Zusammenarbeit mit anderen Personen generierten Daten, Datenauswertungen und Schlussfolgerungen korrekt gekennzeichnet und meinen eigenen Beitrag sowie die Beiträge anderer Personen korrekt kenntlich gemacht habe (siehe Anteilserklärung). Texte oder Textteile, die gemeinsam mit anderen erstellt oder verwendet wurden, habe ich korrekt kenntlich gemacht.

Meine Anteile an etwaigen Publikationen zu dieser Dissertation entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem Erstbetreuer, angegeben sind. Für sämtliche im Rahmen der Dissertation entstandenen Publikationen wurden die Richtlinien des ICMJE (International Committee of Medical Journal Editors; [www.icmje.org](http://www.icmje.org)) zur Autorenschaft eingehalten. Ich erkläre ferner, dass ich mich zur Einhaltung der Satzung der Charité – Universitätsmedizin Berlin zur Sicherung Guter Wissenschaftlicher Praxis verpflichte.

Weiterhin versichere ich, dass ich diese Dissertation weder in gleicher noch in ähnlicher Form bereits an einer anderen Fakultät eingereicht habe.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§§156, 161 des Strafgesetzbuches) sind mir bekannt und bewusst.“

---

T. Bockmayr

# Anteilserklärung

Teresa Bockmayr hatte folgenden Anteil an den folgenden Publikationen:

## Publikation 1

Leichsenring, J., Volckmar, A. L., Kirchner, M., Kazdal, D., Kriegsmann, M., Stögbauer, F., Bockmayr, T., Klauschen, F., Herth, F. J. F., Penzel, R., Warth, A., Schirmacher, P., Endris, V. & Stenzinger, A. **Targeted deep sequencing of effusion cytology samples is feasible, informs spatiotemporal tumor evolution, and has clinical and diagnostic utility.** *Genes Chromosomes Cancer*. 57, 70–79 (2018). IF: 3.444

Beiträge im Einzelnen:

- Beteiligung an der Probengewinnung und Datenaufbereitung
- Beteiligung an der Revision des Manuskripts

## Publikation 2

Jurmeister, P.\*, Bockmayr, M.\*, Seegerer, P., Bockmayr, T., Treue, D., Montavon, G., Vollbrecht, C., Arnold, A., Teichmann, D., Bressemer, K., Schüller, U., von Laffert, M., Müller, K. R., Capper, D. & Klauschen, F. **Machine learning analysis of DNA methylation profiles distinguishes primary lung squamous cell carcinomas from head and neck metastases.** *Sci Transl Med*. 11 (2019). (\* equal contribution)  
IF: 16.304

Beiträge im Einzelnen:

- Beteiligung an der Probengewinnung, Datenaufbereitung und Auswertung
- Beteiligung an der Erstellung und Revision des Manuskripts

## Publikation 3

Bockmayr, T., Erdmann, G., Treue, D., Jurmeister, P., Schneider, J., Arndt, A., Heim, D., Bockmayr, M., Sachse, C. & Klauschen, F. **Multiclass cancer classification in fresh frozen and formalin-fixed paraffin-embedded tissue by DigiWest multiplex protein analysis.** *Lab Invest*. 100, 1288–1299 (2020). IF: 4.197

Beiträge im Einzelnen:

- Beteiligung an der Entwicklung der Fragestellung und des Studienkonzepts
- Identifikation geeigneter Tumorproben im pathologischen Archiv, histologische Evaluation und Qualitätskontrolle
- Probenaufbereitung und -prozessierung

- Auswertung der DigiWest-Daten, umfassende Qualitätskontrolle und Validierung der Ergebnisse der mit der DigiWest-Plattform *NMI TT Pharmaservices* durchgeführten Messungen
- Beteiligung an der statistischen/bioinformatischen Analyse der Daten (Training des Klassifikators und Durchführung der Klassifikation sowie insbesondere Prüfung der Ergebnisse im histopathologischen Kontext)
- Federführend bei der Visualisierung der Daten und Erstellung der Abbildungen (alle Abbildungen: 1, 2, 3 und 4)
- Federführend bei der Verfassung des gesamten Manuskripts (Einleitung, Methode, Ergebnisse, Diskussion)
- Federführung bei der Revision des Manuskripts

---

T. Bockmayr

---

Prof. Dr. F. Klauschen

## Publikationen

1) Leichsenring, J., Volckmar, A. L., Kirchner, M., Kazdal, D., Kriegsmann, M., Stögbauer, F., Bockmayr, T., Klauschen, F., Herth, F. J. F., Penzel, R., Warth, A., Schirmacher, P., Endris, V. & Stenzinger, A. **Targeted deep sequencing of effusion cytology samples is feasible, informs spatiotemporal tumor evolution, and has clinical and diagnostic utility.** *Genes Chromosomes Cancer*. 57, 70–79 (2018). <https://doi.org/10.1002/gcc.22509>

Seite 33–42

2) Jurmeister, P.\*, Bockmayr, M.\*, Seegerer, P., Bockmayr, T., Treue, D., Montavon, G., Vollbrecht, C., Arnold, A., Teichmann, D., Bressemer, K., Schüller, U., von Laffert, M., Müller, K. R., Capper, D. & Klauschen, F. **Machine learning analysis of DNA methylation profiles distinguishes primary lung squamous cell carcinomas from head and neck metastases.** *Sci Transl Med*. 11 (2019). <https://doi.org/10.1126/scitranslmed.aaw8513> (\* equal contribution)

Seite 43–52

3) Bockmayr, T., Erdmann, G., Treue, D., Jurmeister, P., Schneider, J., Arndt, A., Heim, D., Bockmayr, M., Sachse, C. & Klauschen, F. **Multiclass cancer classification in fresh frozen and formalin-fixed paraffin-embedded tissue by DigiWest multiplex protein analysis.** *Lab Invest*. 100, 1288–1299 (2020). <https://doi.org/10.1038/s41374-020-0455-y>

This is an open access article distributed under the terms of the Creative Commons CC BY license.

Seite 53–64

## Publikation 1

**Targeted deep sequencing of effusion cytology samples is feasible, informs spatiotemporal tumor evolution, and has clinical and diagnostic utility.**

Leichsenring, J., Volckmar, A. L., Kirchner, M., Kazdal, D., Kriegsmann, M., Stögbauer, F., Bockmayr, T., Klauschen, F., Herth, F. J. F., Penzel, R., Warth, A., Schirmacher, P., Endris, V. & Stenzinger, A.

*Genes, Chromosomes and Cancer.* 57, 70–79 (2018).

<https://doi.org/10.1002/gcc.22509>

Seite 33–42

## Publikation 2

**Machine learning analysis of DNA methylation profiles distinguishes primary lung squamous cell carcinomas from head and neck metastases.**

Jurmeister, P.\*, Bockmayr, M.\*, Seegerer, P., Bockmayr, T., Treue, D., Montavon, G., Vollbrecht, C., Arnold, A., Teichmann, D., Bressemer, K., Schüller, U., von Laffert, M., Müller, K. R., Capper, D. & Klauschen, F. (\* equal contribution)

*Science Translational Medicine*. 11 (2019).

<https://doi.org/10.1126/scitranslmed.aaw8513>

Seite 43–52



## Multiclass cancer classification in fresh frozen and formalin-fixed paraffin-embedded tissue by DigiWest multiplex protein analysis

Teresa Bockmayr<sup>1</sup> · Gerrit Erdmann<sup>2</sup> · Denise Treue<sup>1,3</sup> · Philipp Jurmeister<sup>1,4</sup> · Julia Schneider<sup>2</sup> · Anja Arndt<sup>2</sup> · Daniel Heim<sup>1</sup> · Michael Bockmayr<sup>1,5,6</sup> · Christoph Sachse<sup>2</sup> · Frederick Klauschen<sup>1,4</sup>

Received: 9 February 2020 / Revised: 2 June 2020 / Accepted: 7 June 2020 / Published online: 29 June 2020  
© The Author(s) 2020. This article is published with open access

### Abstract

Histomorphology and immunohistochemistry are the most common ways of cancer classification in routine cancer diagnostics, but often reach their limits in determining the organ origin in metastasis. These cancers of unknown primary, which are mostly adenocarcinomas or squamous cell carcinomas, therefore require more sophisticated methodologies of classification. Here, we report a multiplex protein profiling-based approach for the classification of fresh frozen and formalin-fixed paraffin-embedded (FFPE) cancer tissue samples using the digital western blot technique DigiWest. A DigiWest-compatible FFPE extraction protocol was developed, and a total of 634 antibodies were tested in an initial set of 16 FFPE samples covering tumors from different origins. Of the 303 detected antibodies, 102 yielded significant correlation of signals in 25 pairs of fresh frozen and FFPE primary tumor samples, including head and neck squamous cell carcinomas (HNSC), lung squamous cell carcinomas (LUSC), lung adenocarcinomas (LUAD), colorectal adenocarcinomas (COAD), and pancreatic adenocarcinomas (PAAD). For this signature of 102 analytes (covering 88 total proteins and 14 phosphoproteins), a support vector machine (SVM) algorithm was developed. This allowed for the classification of the tissue of origin for all five tumor types studied here with high overall accuracies in both fresh frozen (90.4%) and FFPE (77.6%) samples. In addition, the SVM classifier reached an overall accuracy of 88% in an independent validation cohort of 25 FFPE tumor samples. Our results indicate that DigiWest-based protein profiling represents a valuable method for cancer classification, yielding conclusive and decisive data not only from fresh frozen specimens but also FFPE samples, thus making this approach attractive for routine clinical applications.

### Introduction

Precise cancer diagnostics is crucial for the selection of the appropriate treatment and estimation of prognosis. Tumor classification includes the tissue of origin (histological type

and primary site), as well as the grade and stage of cancer. Besides clinical and radiographic data, the pathological examination is decisive in cancer diagnostics. Histomorphology complemented by immunohistochemistry allows for the determination of the tissue of origin in many cases. However, in certain instances, such as in squamous cell carcinomas, or when tumors have lost their specific features, this approach fails, and the tumor types cannot be determined by conventional methods. Moreover, in cancers of

**Supplementary information** The online version of this article (<https://doi.org/10.1038/s41374-020-0455-y>) contains supplementary material, which is available to authorized users.

✉ Christoph Sachse  
sachse@nmi-tt.de

✉ Frederick Klauschen  
frederick.klauschen@charite.de

<sup>1</sup> Institute of Pathology, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin and Berlin Institute of Health, Berlin, Germany

<sup>2</sup> NMI TT Pharmaservices, Berlin, Germany

<sup>3</sup> Central Biobank Charité (ZeBanC), Berlin, Germany

<sup>4</sup> German Cancer Consortium (DKTK), Partner Site Berlin, and German Cancer Research Center (DKFZ), Heidelberg, Germany

<sup>5</sup> Department of Pediatric Hematology and Oncology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

<sup>6</sup> Research Institute Children's Cancer Center Hamburg, Hamburg, Germany

unknown primary (CUP), which account for 3–5% of all malignant epithelial tumors, no primary tumor can be identified even after extensive diagnostic work-up [1]. CUP are often adenocarcinomas, and more than half of the primaries found in autopsy arose in the lung or the pancreas [2, 3]. Furthermore, patients sometimes develop more than one cancer, which makes it difficult to differentiate metastasis from a second primary tumor. For example, primary squamous cell carcinomas of the lung and metastases from head and neck squamous cell carcinomas (HNSC) often share a similar morphology and can, therefore, hardly be distinguished from each other [4].

Through the advances in molecular techniques, sophisticated diagnostic approaches have been developed and applied to cancer classification. Several studies describe the classification of tumors according to their tissue of origin using gene expression [5–7], microRNA [8, 9], and, more recently, DNA methylation [10, 11] profiling. Tissue-based protein profiling constitutes a further promising approach for cancer classification, as tumor types are characterized by specific protein profiles [12]. Proteins are the principal effector molecules in a cell, and their function can be influenced by posttranslational modifications such as phosphorylation, acetylation, glycosylation, or sulphation [13]. Proteomic approaches have been successfully applied to differentiate cancer from nonmalignant tissue [14–16] or for the pairwise distinction of tumor types or subtypes [17–19]. However, so far, only a few studies have been conducted to classify multiple tumor types based on protein profiles in tissue samples [20–22].

Mass spectrometry is a powerful method for comprehensive proteomics, but is associated with high initial and operating costs, requires specially trained staff and has not yet been implemented routinely in diagnostics [23]. Besides mass spectrometry, diverse targeted antibody-based techniques have been established for protein analysis. Immunohistochemistry is widely used for standard diagnostic purposes and enables protein identification and quantification in histological sections [23]. Western blots facilitate immunodetection according to molecular weight [24]. However, both methods are inconvenient for the analysis of a high number of proteins and require a relatively large amount of tissue. In contrast, reverse phase protein arrays (RPPA) and the DigiWest method are more suitable for the parallel measurement of multiple proteins [25, 26]. DigiWest relies on classical western blotting and combines it with bead-based multiplexing, which allows for the simultaneous measurement of 80–800 proteins in samples even with low amounts of material [26]. Furthermore, DigiWest shows a similar sensitivity, reproducibility, and signal linearity as a high-end western blot system [26]. It has been effectively used in the analysis of signaling pathways and for the verification of biomarker candidates [27–31].

Fresh frozen samples are often preferred for molecular analysis, as macromolecules are preserved without cross-links [32], but the availability of fresh frozen samples is often limited as their collection is laborious, expensive, and necessitates special logistics. Hence, in routine diagnostic pathology, tissue samples are fixed in formalin and embedded in paraffin for preservation. Formaldehyde induces cross-links among proteins or between proteins and nucleic acids, thereby preserving the tissue morphology adequately [33]. This ensures good quality for histological examination, but cross-links may impair the immunoreactivity of proteins by modifying their conformation and altering or masking the epitope [34]. It is also challenging to extract full-length proteins from formaldehyde-fixed tissue [35]. Different protocols have already been developed to analyze formalin-fixed paraffin-embedded (FFPE) samples by immunohistochemistry, western blotting, mass spectrometry, and RPPA [36–39].

In this study, we established FFPE sample extraction protocols suitable for DigiWest protein profiling, tested over 600 antibodies for their suitability on FFPE tissue, and identified antibodies yielding comparable results in FFPE and fresh frozen tissue. We showed that DigiWest multiplex protein profiles can be used to predict the tissue of origin of five different cancer types, including HNSC, lung squamous cell carcinomas (LUSC), lung adenocarcinomas (LUAD), colorectal adenocarcinomas (COAD), and pancreatic adenocarcinomas (PAAD) in both fresh frozen and FFPE tissue.

## Materials and methods

### Sample acquisition and preparation

A set of 25 paired fresh frozen and FFPE tumor samples, as well as an independent validation cohort of 25 FFPE tumor samples, were acquired from the archive of the Institute of Pathology of the Charité University Hospital Berlin, Germany. Informed consent was obtained from all patients in accordance with standard institutional guidelines. The samples were all primary tumors.

The set of 25 paired fresh frozen and FFPE tumor samples contained five HNSC, five LUSC, five LUAD, five COAD, and five PAAD. Tumors with different histological grades were included (1 well, 18 moderately, 2 moderately to poorly, and 4 poorly differentiated) to represent tumors that would occur in a realistic clinical setting. Tumor cell content was assessed by a board-certified pathologist based on hematoxylin-eosin-stained slides and was determined to be at least 30% in fresh frozen and at least 40% in FFPE samples. The average tumor cell content was 70% in fresh frozen and 68% in FFPE samples.

The independent validation cohort consisted of 25 FFPE tumor samples, also with five cases per tumor type (HNSC, LUSC, LUAD, COAD, and PAAD). Of these, 18 tumor samples were moderately and seven poorly differentiated. The tumor cell content was at least 40% and averaged 68%.

Fresh frozen samples used for DigiWest analysis were collected after surgical resection, snap-frozen in liquid nitrogen, and stored at  $-80^{\circ}\text{C}$ . The cold ischemic time was measured in 21 of 25 samples with a median of 14 min. They were cut in slices of  $15\ \mu\text{m}$  thickness at  $-20^{\circ}\text{C}$  using a cryostat (Leica Biosystems, Wetzlar, Germany). Tissue slices were lysed in CLB1 lysis buffer ( $10\ \mu\text{l}$  lysis buffer/mg tissue) containing PhosSTOP inhibitor cocktails (Roche Diagnostics GmbH) for 30 min in a thermomixer ( $4^{\circ}\text{C}$ , 1400 rpm). Samples were subsequently centrifuged for 5 min at  $4^{\circ}\text{C}$  and 18,200 g (Eppendorf, Hamburg, Germany). The supernatant was collected, divided into three aliquots, and stored at  $-80^{\circ}\text{C}$ . The total protein concentration was measured using Coomassie Plus (Bradford) Assay Kit (Thermo Scientific, Rockford, USA). All samples had a protein concentration of more than 1 mg/ml.

FFPE samples were cut in  $15\ \mu\text{m}$  thick curls. When needed, tumor-rich areas were manually macro-dissected to ensure a tumor cell content of at least 40%. For the extraction of proteins from FFPE curls, the Qproteome FFPE Tissue kit (Qiagen, Hilden, Germany) with its Heptan-based protocol was used according to the manufacturer's recommendations. The resulting protein lysates were further purified with the 2-D Clean-Up Kit (GE Healthcare, Chicago, USA) according to the vendor's protocol. The resulting protein pellets were re-suspended in LDS buffer containing 212 mM Tris HCL, 282 mM Tris base, 4% LDS (w/v), 1.01 mM EDTA and supplemented with 50 mM DTT (Invitrogen, Carlsbad, USA). Protein concentrations were determined using the 660 nm assay with IDCR (Invitrogen, Carlsbad, USA).

Protein concentrations of both fresh frozen and FFPE lysates were then adjusted by SDS PAGE, employing Coomassie Fluor Orange Protein Gel Stain (Invitrogen, Carlsbad, USA) according to the vendor's protocol, including an internal protein lysate standard to optimize sample loading for DigiWest.

### DigiWest multiplex protein analysis

DigiWest assays were performed as published (see [26] for details). In brief, for the initial analysis of 634 antibodies,  $2 \times 20\ \mu\text{g}$  of total protein per sample was loaded on an SDS-polyacrylamide gel ( $20\ \mu\text{g}/\text{lane}$ ) and size-separated via electrophoresis. For the subsequent analysis of 306 antibodies,  $1 \times 20\ \mu\text{g}$  of total protein was required, while only  $1 \times 10\ \mu\text{g}$  of total protein was used for the measurement of 102 antibodies in the independent validation cohort.

Size-separated proteins were blotted to a PVDF membrane and biotinylated. Every lane of the membrane was cut into 96 strips of 0.5 mm width, each corresponding to a certain molecular weight fraction. Each biotinylated protein strip was then placed in a specific well of a 96-well plate and elution buffer was added. The eluted proteins were incubated with magnetic color-coded beads (Luminex, Austin, USA) coated with neutravidin. The biotinylated proteins bind to the neutravidin beads such that each bead color represents proteins of one specific molecular weight fraction. The beads were mixed in pools of 96 bead identifies, thus resulting in a reconstitution of the original lane. For each protein measured, a small aliquot of the bead pool was incubated with a specific antibody and phycoerythrin-labeled secondary antibodies were added to generate signals.

Samples were read on a FlexMAP 3D flow cytometer (Luminex, Austin, USA), resulting in 96 values per antibody and sample, represented as graphs. Signal intensity was plotted against molecular weight and protein bands were visualized as peaks. While the molecular weight of each antibody was provided, an algorithm was used to identify adjacent peaks. The detected signals corresponded to the integral of the area of a peak, after subtraction of the local background. In a dedicated set-up, extraction of 4–16 FFPE samples plus DigiWest for up to 300 antibodies plus data analysis can be conducted within 10–12 days.

### Antibody selection

Antibody selection was performed on the set of 25 paired fresh frozen and FFPE tumor samples. From our collection of >1200 antibodies that had been pre-validated for DigiWest in fresh frozen materials, a selection of 634 antibodies was initially measured by DigiWest in 16 FFPE samples (3 HNSC, 4 LUSC, 3 LUAD, 3 COAD, and 3 PAAD). These 634 antibodies covered a broad range of molecular weights, targeting proteins and phosphoproteins in the cytoplasm and the nucleus. Among them, 306 antibodies were detectable in at least four samples or all samples of the same tumor type. These 306 antibodies were subsequently measured in the corresponding 16 fresh frozen samples and the additional nine pairs of fresh frozen and FFPE samples (2 HNSC, 1 LUSC, 2 LUAD, 2 COAD, and 2 PAAD).

A noticeable cross-reaction of the anti-rabbit secondary antibody (dk- $\alpha$ -rb-IgG (H + L)-RPE #711-116-152 Jackson, Westgrove, USA) was observed at 47–53 kDa in all HNSC samples, resulting in a stronger signal in fresh frozen than in FFPE samples. Therefore, we excluded a priori three antibodies (Cytokeratin 16, PPAR alpha-pS12, DAPK3 (ZIPK)-pT265) for which it was not possible to distinguish the specific signal from a cross-reaction. This resulted in a dataset of 303 antibodies.

Pearson correlations were computed for these 303 antibodies between fresh frozen and FFPE samples. Multiple testing correction for the significance of correlation scores was performed with the Benjamini–Hochberg (BH) method. Of the 303 antibodies, 121 showed a significant correlation ( $p\text{-BH} < 0.05$ ), corresponding to 128 of 407 detected signals. In the cases in which more than one signal was detected for a given antibody, only the signal with the highest correlation factor was retained for further analysis.

Among these 121 antibodies, 12 were excluded because the detected signal shifted over 20% of the expected molecular weight of the antibody, another five antibodies were discarded because the peaks were not clearly identifiable, and two antibodies were excluded due to limited availability or redundancy.

In total, 102 of 634 antibodies (Supplementary Table 1) resulted in clear peaks at the expected molecular weights ( $\pm 20\%$ ) and showed a significant correlation between the DigiWest signals detected in fresh frozen and FFPE samples. These 102 antibodies were subsequently used for cancer classification. Furthermore, DigiWest analysis in the independent validation cohort of 25 FFPE samples was also conducted with these 102 antibodies.

### Statistical analysis, classification, and data visualization

The analysis of processed DigiWest data was performed using the statistical programming language R [40] including the packages *gplots*, *kernelab*, *e1071*, and *caret* [41–44]. The data were transformed into log<sub>2</sub> scale, and Pearson correlation coefficient was applied for the analysis between log<sub>2</sub> expression values in fresh frozen and FFPE samples. The significance of correlation was assessed with the R-function *cor.test*. Heatmaps were generated based on the average-linkage method and Pearson correlation coefficient as similarity measure.

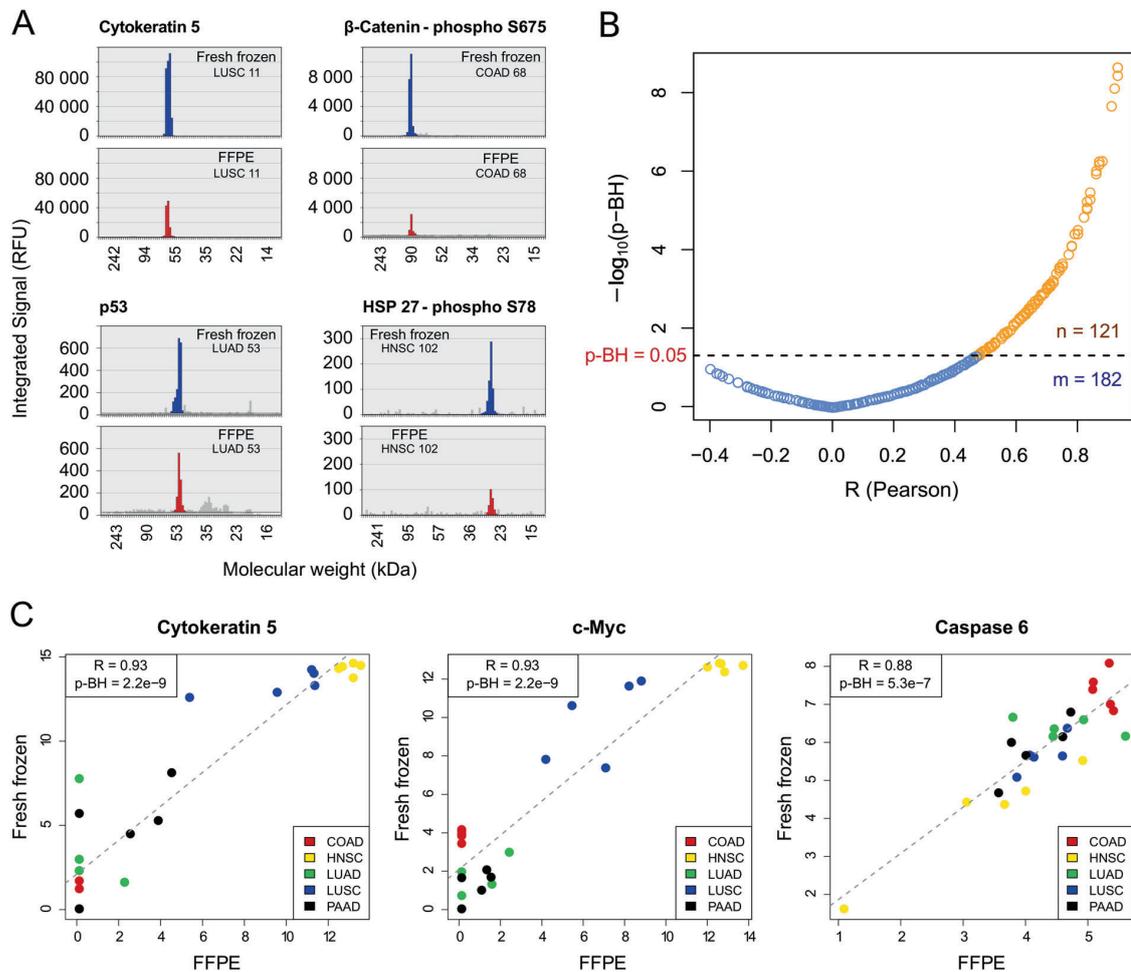
Radial basis function kernel support vector machines (SVM) were used as tumor classifiers. The model for the paired set was tuned and evaluated using nested cross-validation [45] with fivefold outer and fourfold inner cross-validation, repeated ten and five times, respectively. This ensures that no information from the validation samples was used for model selection at any point. The tuning parameters were chosen between  $C = 10^{(0.1, 2, 3)}$  and  $\sigma = 0.01 \times 10^{(-3, -2, \dots, 3)}$ . The classifier used on the independent validation set was trained on the 25 FFPE samples from the paired set using fivefold cross-validation (ten repeats) and the same tuning parameters. The optimal parameters were  $\sigma = 0.0001$  and  $C = 100$ . The SVM classifiers were compared with a random forest classifier, which yielded inferior classification accuracy for these prediction tasks.

The significance of differential expression between two groups was assessed with the *t*-test (R-function *pairwise.t.test* with default parameters). Multiple testing correction was performed with the BH method [46]. *p* values  $< 0.05$  were considered statistically significant.

## Results

After the establishment and optimization of a DigiWest-compatible extraction protocol for FFPE samples, we determined the performance of DigiWest in FFPE tissue for the initial 634 antibodies. From these, we selected 303 antibodies that were expressed in at least four samples or all samples of the same tumor type to compare their signals in fresh frozen and FFPE samples (see “Methods”). Those 303 antibodies were measured in the 25 pairs of fresh frozen and FFPE primary tumor samples, which included five samples for each of the following tumor types: HNSC, LUSC, LUAD, COAD, and PAAD.

Then we investigated the correlations of all detected signals between fresh frozen and FFPE samples via Pearson correlation coefficients (*R*) to identify those antibodies that performed similarly in both tissue types. Figure 1 presents selected proteins comparing fresh frozen and FFPE signal intensities (Fig. 1a). The detected signals correspond to the integrated area of a peak at a certain molecular weight. If more than one signal was detected for a given antibody, only the signal with the highest correlation coefficient was included. Without applying a multiple testing correction, 150 out of the 303 antibodies (50%) were found to be significantly correlated between fresh frozen and FFPE tissue. After multiple testing correction (BH method), 121 antibodies (40%) demonstrated a significant correlation ( $p\text{-BH} < 0.05$ ), with a correlation coefficient of  $R \geq 0.47$  (Fig. 1b). Nineteen of these antibodies were excluded because the peak was shifted ( $> 20\%$  of the expected molecular weight), or peaks were not clearly identifiable. Cytokeratin 5 and c-Myc showed the highest correlation coefficient among all proteins, with  $R = 0.93$  and  $p\text{-BH} = 2.2e - 9$ . The detected signals for these proteins were particularly high in HNSC and LUSC, for both fresh frozen and FFPE tissue (Fig. 1c). The remaining 182 antibodies (60%) were not significantly correlated ( $p\text{-BH} > 0.05$ ) and thus excluded. In total, 102 antibodies (34%) were used for further analysis. These 102 analytes comprised 14 out of 48 antibodies against phosphoproteins (29%) and 88 out of 255 against total proteins (35%). In fresh frozen samples, the mean signal intensities (log<sub>2</sub> scale) were generally higher, except for one of 102 antibodies, and the signals were detected more frequently than in FFPE tissue (23 vs. 18 tumor samples on average).



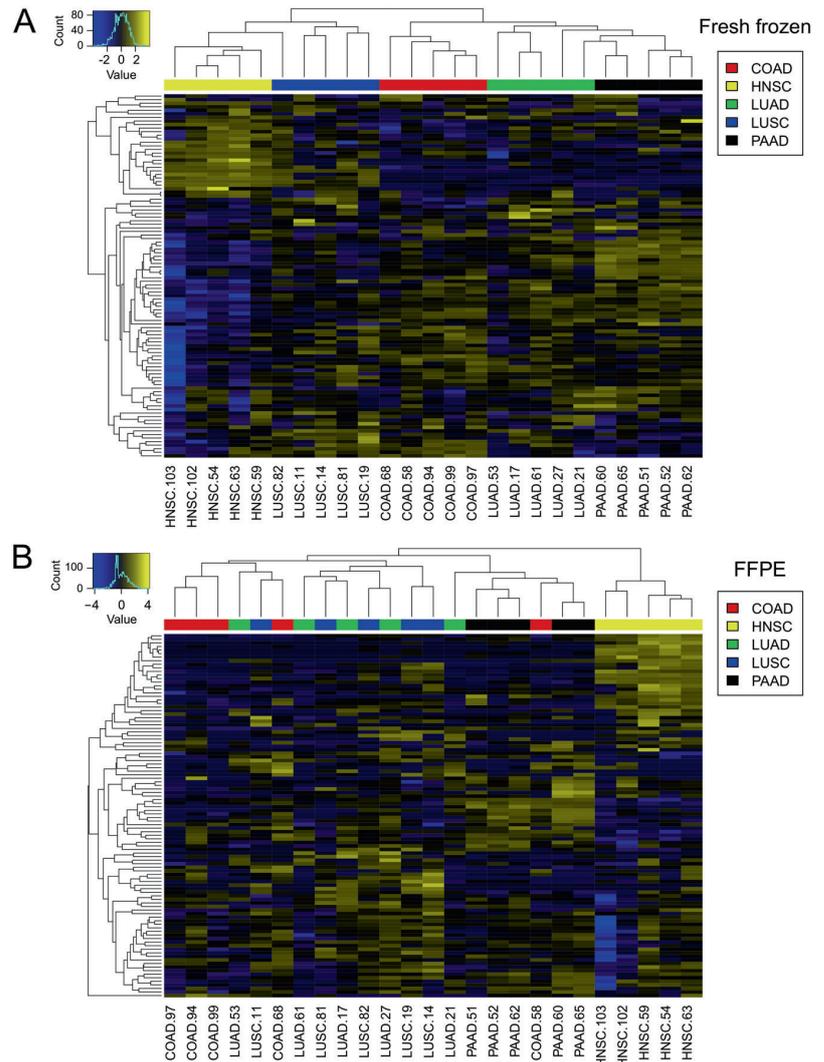
**Fig. 1** DigiWest protein profiling in fresh frozen and FFPE tissue. **a** DigiWest data displayed for four antibodies (Cytokeratin 5, p53,  $\beta$ -Catenin-phospho S675, HSP 27-phospho S78) in different tumor samples (fresh frozen in blue and FFPE tissue in red) with signal intensity (RFU: relative fluorescence units) plotted against molecular weight. **b** Pearson correlation coefficients ( $R$ ) were computed for each antibody between all signals detected in both fresh frozen and FFPE samples.  $p$  values were corrected for multiple testing with the Benjamini–Hochberg method ( $p$ -BH), using a significance level of 0.05. The corresponding volcano plot shows the  $p$  values ( $-\log_{10}$

transformed) plotted against correlation coefficients ( $R$ ) ( $n$ : number of antibodies with significant correlation ( $p$ -BH < 0.05), colored in orange;  $m$ : number of antibodies with  $p$ -BH > 0.05, colored in blue). **c** Relative signals ( $\log_2$ ) detected by DigiWest for three antibodies (Cytokeratin 5, c-Myc, Caspase 6) in 25 tumor samples, in both fresh frozen and FFPE samples. COAD colorectal adenocarcinomas, HNSC head and neck squamous cell carcinomas, LUAD lung adenocarcinomas, LUSC lung squamous cell carcinomas, PAAD pancreatic adenocarcinomas.

Furthermore, we explored to which extent the selected panel of 102 antibodies would qualify to classify cancers with respect to their histological type and organ origin in the paired set. We first applied an unsupervised learning approach, visualizing DigiWest data by heatmaps combined with hierarchical clustering in both fresh frozen and FFPE tissue (Fig. 2). The 25 tumor samples were grouped together based on the correlation between the antibody signals. For fresh frozen samples, the different tumor types formed relatively distinct clusters. Only one LUSC was more closely grouped with HNSC than with the other LUSC. In the

same way, two LUAD were more similar to PAAD than to the other LUAD. The two main clusters clearly separated adenocarcinomas from squamous cell carcinomas. In FFPE specimens, HNSC samples formed a distinct group, with high signals for a considerable number of proteins. All other tumor types were included in a second cluster, in which the tumor types were not well separated. Particularly, most of the LUAD and LUSC clustered together. Overall, the hierarchical clustering revealed clear differences between the investigated tumor types, especially in fresh frozen samples.

**Fig. 2 Heatmap and hierarchical clustering of DigiWest data.** Overall, 102 antibodies were analyzed in 25 tumor samples, including five tumor types (**a** fresh frozen tissue; **b** FFPE specimens) with columns = tumor samples and rows = antibody signals. The color gradient from blue to yellow corresponds to low or high antibody-specific signals among the 25 tumor samples. COAD colorectal adenocarcinomas, HNSC head and neck squamous cell carcinomas, LUAD lung adenocarcinomas, LUSC lung squamous cell carcinomas, PAAD pancreatic adenocarcinomas.



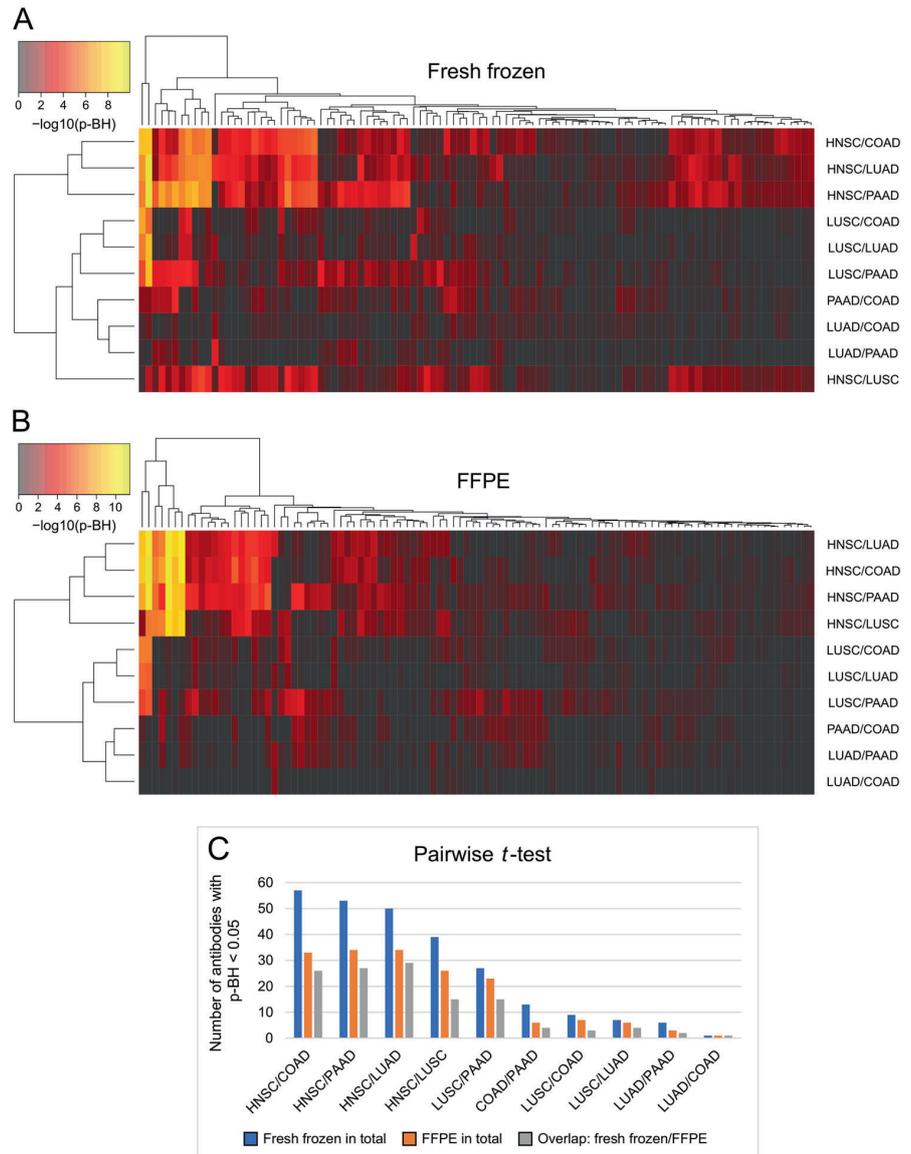
A *t*-test was carried out in fresh frozen and FFPE samples for each pair of tumor types to identify the proteins that were suitable for discrimination of the respective tumor types. The corresponding *p* values were visualized in a heatmap, displayed in Fig. 3. Generally, more proteins with a significant *p* value ( $p\text{-BH} < 0.05$ ) were found in fresh frozen than in FFPE tissue (Fig. 3a, b). Furthermore, a large proportion of the proteins that were significantly differentially expressed between the different tumor types in FFPE tissue were also significantly differentially expressed in fresh frozen samples (Fig. 3c). The expression of two proteins, Cytokeratin 5 and c-Myc, was significantly different between all squamous cell and adenocarcinomas, both in fresh frozen and in FFPE samples. Among all pairs of tumor types, those with the largest number of differentially expressed proteins all involved HNSC samples. For the

three pairs of HNSC and adenocarcinomas (COAD, PAAD, and LUAD), at least 50 proteins in fresh frozen tissue and 33 in FFPE samples were significantly different. Both types of squamous cell carcinomas (HNSC and LUSC) could also be discriminated by a considerable number of proteins (39 in fresh frozen and 26 in FFPE tissue). In contrast, only a few proteins were suitable for distinguishing LUSC vs. LUAD, as well as between two types of adenocarcinomas. For LUAD vs. COAD, only one protein with a significant *p* value was determined (Thyroid transcription factor-1 (TTF-1)).

To classify the tumor samples according to their tissue of origin, an SVM algorithm with repeated nested cross-validation (fourfold internal and fivefold external) was applied to the set of paired tumor samples. With the resulting classification, an overall accuracy of 90.4% was

**Fig. 3** Pairwise *t*-test performed for 102 antibodies for each pair of tumor types.

Heatmaps show *p* values ( $-\log_{10}$  transformed) of the pairwise *t*-test (a fresh frozen and b FFPE tissue) after Benjamini–Hochberg (BH) correction. Columns represent the tested antibodies, and rows indicate pairs of tumor types. c Bar chart of the number of antibodies with a significant *p* value ( $p\text{-BH} < 0.05$ ) in the pairwise *t*-test for each pair of tumors; comparison of fresh frozen samples, FFPE tissue, and overlap of both tissue types. COAD colorectal adenocarcinomas, HNSC head and neck squamous cell carcinomas, LUAD lung adenocarcinomas, LUSC lung squamous cell carcinomas, PAAD pancreatic adenocarcinomas.

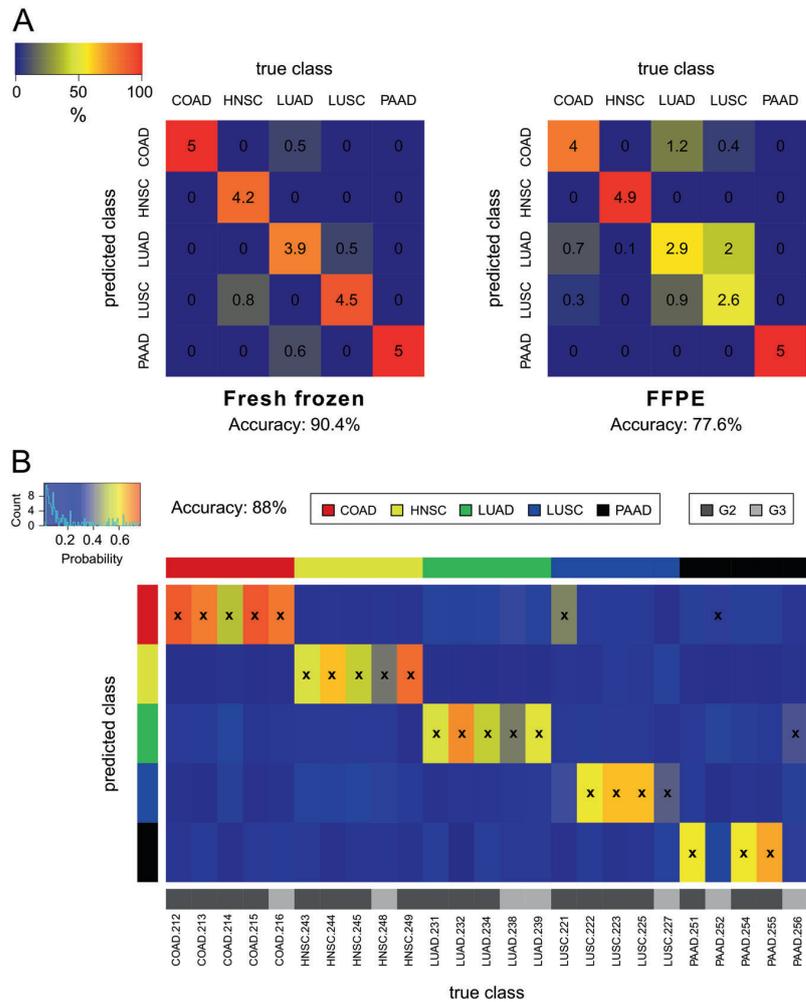


obtained in fresh frozen samples (standard deviation of 5.4% over ten repeats; Fig. 4a). All colorectal and PAAD were classified correctly. The lowest accuracy of 78% was attained for LUAD. For FFPE samples, the SVM classifier yielded an overall accuracy of 77.6% (standard deviation of 3.4% over ten repeats). The individual accuracies among the different tumor types varied more in FFPE than in fresh frozen specimens. On the one hand, all PAAD and almost all HNSC (98%) were correctly assigned to their tissue of origin. On the other hand, the classifier based on FFPE data was often not able to discriminate between LUSC and LUAD. Only 52% of LUSC were predicted correctly, while 40% of LUSC were misclassified as LUAD. Similar results

were obtained for LUAD, with 58% of the samples being correctly classified and 18% of LUAD categorized as LUSC. If squamous cell and adenocarcinomas of the lung were considered as only one tumor type since they arise in the same organ, the accuracy of this class increased to 84%. This resulted in an overall accuracy of 89.2%. Furthermore, some misclassification of FFPE tumor samples also occurred between LUAD and COAD.

In general, the classification model based on fresh frozen samples performed better than the classifier constructed with FFPE samples (overall accuracies of 90.4% vs. 77.6%). Both models performed best for PAAD and yielded there a prediction accuracy of 100%. For LUSC, LUAD,

**Fig. 4 Multiclass cancer classification by machine learning.** **a** Contingency matrices showing the classification accuracy of the SVM-based models obtained by repeated nested cross-validation with 25 tumor samples in fresh frozen and FFPE tissue. The numbers indicate how many of the five samples of each tumor type are classified on average in each class. The corresponding percentages are visualized by the color scheme (blue: low, red: high). **b** Classification results of the SVM algorithm in an independent validation cohort of 25 FFPE samples with five cases per tumor type. The predicted tumor type is marked with a cross; the color gradient indicates the confidence of the SVM classifier for each class (blue: low probability, red: high probability). Histological grades: G2 moderately and G3 poorly differentiated tumor samples. COAD colorectal adenocarcinomas, HNSC head and neck squamous cell carcinomas, LUAD lung adenocarcinomas, LUSC lung squamous cell carcinomas, PAAD pancreatic adenocarcinomas.



and COAD, the accuracy was at least 20% higher in fresh frozen than in FFPE samples. In contrast, the classification of HNSC yielded better results with FFPE samples.

Finally, we tested the performance of the SVM classifier in an independent validation cohort of 25 FFPE primary tumor samples, containing five cases per tumor type. This yielded an overall accuracy of 88%. The classification results and their probability scores are visualized in Fig. 4b. All COAD, HNSC, and LUAD cases were correctly assigned to their tissue of origin. In total, only three tumor samples were misclassified. One LUSC and one PAAD were categorized as COAD, while another PAAD was classified as LUAD. Concerning the histological grade, 17 out of 18 moderately (G2) and five out of seven poorly (G3) differentiated tumor samples were correctly predicted. The mean probability scores for the correct class in poorly differentiated cases were slightly lower than in moderately differentiated tumor samples (mean 0.45 vs. 0.61,  $p = 0.03$ ).

Overall, the SVM classifier performed better in the independent validation cohort than in the initial 25 FFPE samples from the paired set (88% vs. 77.6% overall accuracy).

## Discussion

Protein analysis in FFPE tissue is known to be challenging, as formaldehyde fixation induces cross-links, and proteins must be recovered before the analysis. In this study, we demonstrated that DigiWest multiplex protein analysis is feasible in FFPE samples and can be used for diagnostic cancer classification. By comparing the performance of 303 antibodies in fresh frozen and FFPE tissue, we identified 102 antibodies that yielded sound and comparable results in both. Importantly, this also included 14 antibodies against phosphoproteins, which allows getting better insights into protein activation and oncogenic signaling.

We identified proteins that were best suited for the pairwise distinction of tumor types. Some of these proteins are well known to be relevant in the corresponding tumor types. Exemplarily, we found that TTF-1 was highly expressed in many LUAD and was a useful protein to differentiate LUAD from every other tumor type both in fresh frozen and in FFPE tissue, consistent with its wide use in diagnostics. TTF-1 is usually used in immunohistochemistry to identify LUAD and to discern primary tumors of the lung from metastases [47]. According to our results, both Cytokeratin 5 and c-Myc were able to distinguish squamous cell carcinomas from adenocarcinomas, and their signal was particularly strong in LUSC and HNSC. Cytokeratin 5 is commonly used as an immunohistochemical marker for squamous cell carcinomas [48]. However, although c-Myc is often expressed in those tumor types, it is not known to be a specific marker for squamous cell carcinomas and has been detected in other tumor types before [49–52]. The expression of these proteins might not only be useful for diagnostic purposes but could further insight into the tumor's biology and indicate potential therapeutic targets.

Based on the 102 selected antibodies, we developed two approaches for cancer classification in the set of paired tumor samples. First, an unsupervised hierarchical clustering method and second a supervised SVM algorithm, which classified the tumors according to their tissue of origin. Both methods achieved a better distinction between the five tumor types in fresh frozen samples than in FFPE tissue. This is probably due to a higher sensitivity of DigiWest analysis in fresh frozen samples. Although the same 102 antibodies were used in both tissue types, the signals were generally detected more frequently in fresh frozen than in FFPE samples. Furthermore, the mean signal intensities were higher, and more proteins were suited for the pairwise distinction of tumor types in fresh frozen samples. Overall, the SVM algorithm reached an accuracy of 90.4% in fresh frozen and 77.6% in FFPE samples.

The reduced overall accuracy in FFPE samples from the paired set was mainly due to misclassification between squamous cell and adenocarcinomas of the lung, which accounted for 11.6% of all errors. These tumor types are closely related, arise in the same organ, and nonneoplastic cells may contribute to the signal. This could make them more difficult to distinguish, even though it does not fully explain the lower accuracy for lung cancer in FFPE compared with fresh frozen samples. If LUSC and LUAD were considered as only one cancer type, the overall accuracy of the FFPE classifier increased to 89.2%, which is very close to that in fresh frozen specimens.

Of note, in the independent validation cohort of 25 FFPE samples, the SVM classifier reached an overall accuracy of 88%, which is slightly higher than in the FFPE samples from the paired set (77.6% overall accuracy). This is likely

to be due to the larger number of training samples available to the classifier (25 instead of the 20 samples used in the nested cross-validation of the paired set). Furthermore, it was associated with a better distinction between LUSC and LUAD. The SVM classifier was also applicable to poorly differentiated tumor samples. Five of the seven poorly differentiated tumor samples were correctly assigned to their tissue of origin. However, as expected from tumor biology, the probability scores for the correct class of poorly differentiated tumor samples were, on average, slightly lower than those of moderately differentiated cases.

Our classifier was able to distinguish with high accuracy between squamous cell carcinomas of different origins. In a previous study, Bohnenberger et al. [17] developed a classifier based on quantitative mass spectrometry data in FFPE samples, which differentiated between HNSC and LUSC with an accuracy of 86.8% in an independent test set [17]. However, more than 1100 proteins were necessary to achieve these results, and the accuracy decreased to 76.8% when only 100 proteins were included [17]. We demonstrated, based on the data of 102 antibodies, that it is possible to generate a classifier with comparable accuracy for discriminating five tumor types (77.6% accuracy in the paired set and 88% in the independent validation cohort) with DigiWest multiplex protein profiling. In our study, no misclassifications occurred between LUSC and HNSC in FFPE samples. As patients with a primary HNSC often develop distant metastasis in the lungs, but at the same time have an increased risk for the occurrence of a second primary tumor of the lungs, the classifier might be used to complement current diagnostic methods.

The overall accuracy of our classifier is comparable with those of previous studies establishing a multiclass cancer classifier based on protein profiles, even though those had only been carried out on fresh frozen samples. On the one hand, our classifier performed slightly better in fresh frozen samples (90.4% accuracy) than two studies that each classified six types of adenocarcinomas using MALDI mass spectrometry with an average accuracy of 82% [20, 21]. On the other hand, the accuracy of our classifier was slightly lower than that of Zhang et al., which reached an accuracy of 93.6% [22]. However, they used a considerably larger number of samples for training, originating from RPPA data of ten tumor types from The Cancer Proteome Atlas [22]. While this accuracy is relatively high, RPPA is not applicable in the typical clinical setting, as outlined below.

Altogether, our classifier achieved a high accuracy using only a modest number of samples for training. A greater number of samples might better represent various differentiation statuses or molecular subtypes within each tumor type, possibly leading to increased accuracy.

A further advantage of DigiWest analysis is that only a low amount of material is required, as 20 µg of protein was sufficient for the measurement of 303 antibodies (resp. 10 µg for 102 antibodies). This makes DigiWest particularly useful for clinical samples, which are often limited in their amounts and routinely collected as FFPE tissue. RPPA represents another approach for the multiplex analysis of proteins, with the advantage that several hundreds of samples can be measured in parallel [25]. However, RPPA is less appropriate for the analysis of single or only few samples as is the case in routine diagnostics where samples have to be processed as they accrue for timely diagnoses. In contrast to RPPA, the proteins in DigiWest analysis are separated according to their molecular weight, which enables direct quality control of the raw data for each antibody in each sample to rule out unspecific signals. This step of identification and interpretation of the specific signals remains critical, as seen in the above-mentioned case of anti-rabbit secondary antibody cross-reaction at 50 kDa in HNSC samples. This emphasizes the need for a careful selection of antibodies.

In summary, our study demonstrates that DigiWest multiplex protein profiling can be performed on FFPE tissue specimens. We identified 102 antibodies against a variety of proteins and phosphoproteins that showed expression in both fresh frozen and FFPE samples, as well as correlating signals between the two. These 102 analytes were used to develop an SVM algorithm capable of classifying samples of five tumor types according to their tissue of origin with high accuracy in both fresh frozen and FFPE samples. Thus, DigiWest analysis constitutes a promising approach for analyzing the protein expression of FFPE samples, which can be used for diagnostic cancer classification and might also help to identify novel biomarkers or potential therapeutic targets in the future.

**Acknowledgements** The authors gratefully acknowledge financial support of this study through ProFIT grants 10163016 and 10163019 of IBB Berlin. The authors wish to thank Drs Markus Templin and Michael Pawlak for helpful input and discussions on the establishment of DigiWest in FFPE samples. We gratefully acknowledge the excellent technical assistance of Przemyslaw Dudys, Barbara Meyer-Bartell, and Ines Koch. Open Access Funding provided by Projekt DEAL.

**Author contributions** Conceptualization: TB, CS, and FK; Methodology: TB, GE, DT, DH, MB, CS, and FK; Formal analysis: TB, GE, DH, MB, and FK; Investigation: All authors; Resources: CS and FK; Data curation: TB, GE, DH, and MB; Writing—Original draft: TB; Writing—Review and editing: All authors; Visualization: TB and MB; and Supervision: CS and FK.

### Compliance with ethical standards

**Conflict of interest** GE, JS, AA, and CS are employees of NMI TT Pharmaservices, a company offering DigiWest service studies. All other authors declare no conflicts of interest.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

### References

- Pavlidis N, Pentheroudakis G. Cancer of unknown primary site. *Lancet*. 2012;379:1428–35.
- Economopoulou P, Mountzios G, Pavlidis N, Pentheroudakis G. Cancer of unknown primary origin in the genomic era: elucidating the dark box of cancer. *Cancer Treat Rev*. 2015;41:598–604.
- Pentheroudakis G, Golfopoulos V, Pavlidis N. Switching benchmarks in cancer of unknown primary: from autopsy to microarray. *Eur J Cancer*. 2007;43:2026–36.
- Pereira TC, Share SM, Magalhães AV, Silverman JF. Can we tell the site of origin of metastatic squamous cell carcinoma? An immunohistochemical tissue microarray study of 194 cases. *Appl Immunohistochem Mol Morphol*. 2011;19:10–14.
- Ramaswamy S, Tamayo P, Rifkin R, Mukherjee S, Yeang C-H, Angelo M, et al. Multiclass cancer diagnosis using tumor gene expression signatures. *Proc Natl Acad Sci USA*. 2001;98:15149–54.
- Erlander MG, Ma X-J, Kesty NC, Bao L, Salunga R, Schnabel CA. Performance and clinical evaluation of the 92-gene real-time PCR assay for tumor classification. *J Mol Diagn*. 2011;13:493–503.
- Pillai R, Deeter R, Rigl CT, Nystrom JS, Miller MH, Buturovic L, et al. Validation and reproducibility of a microarray-based gene expression test for tumor identification in formalin-fixed, paraffin-embedded specimens. *J Mol Diagn*. 2011;13:48–56.
- Rosenfeld N, Aharonov R, Meiri E, Rosenwald S, Spector Y, Zepeniuk M, et al. MicroRNAs accurately identify cancer tissue origin. *Nat Biotechnol*. 2008;26:462–9.
- Meiri E, Mueller WC, Rosenwald S, Zepeniuk M, Klinke E, Edmonston TB, et al. A second-generation microRNA-based assay for diagnosing tumor tissue origin. *Oncologist*. 2012;17:801–12.
- Moran S, Martínez-Cardús A, Sayols S, Musulén E, Balañá C, Estival-Gonzalez A, et al. Epigenetic profiling to classify cancer of unknown primary: a multicentre, retrospective analysis. *Lancet Oncol*. 2016;17:1386–95.
- Jurmeister P, Bockmayr M, Seegerer P, Bockmayr T, Treue D, Montavon G, et al. Machine learning analysis of DNA methylation profiles distinguishes primary lung squamous cell carcinomas from head and neck metastases. *Sci Transl Med*. 2019;11. <https://doi.org/10.1126/scitranslmed.aaw8513>.
- Heim D, Montavon G, Hufnagl P, Müller K-R, Klauschen F. Computational analysis reveals histotype-dependent molecular profile and actionable mutation effects across cancers. *Genome Med*. 2018;10:83.
- Pandey A, Mann M. Proteomics to study genes and genomes. *Nature*. 2000;405:837–46.

14. Schwamborn K, Krieg RC, Reska M, Jakse G, Knuechel R, Wellmann A. Identifying prostate carcinoma by MALDI-Imaging. *Int J Mol Med*. 2007;20:155–9.
15. Le Faouder J, Laouirem S, Chapelle M, Albuquerque M, Belghiti J, Degos F, et al. Imaging mass spectrometry provides fingerprints for distinguishing hepatocellular carcinoma from cirrhosis. *J Proteome Res*. 2011;10:3755–65.
16. Kim HK, Reyzer ML, Choi IJ, Kim CG, Kim HS, Oshima A, et al. Gastric cancer-specific protein profile identified using endoscopic biopsy samples via MALDI mass spectrometry. *J Proteome Res*. 2010;9:4123–30.
17. Bohnenberger H, Kaderali L, Ströbel P, Yepes D, Plessmann U, Dharia NV, et al. Comparative proteomics reveals a diagnostic signature for pulmonary head-and-neck cancer metastasis. *EMBO Mol Med*. 2018;10:e8428.
18. Kriegsmann M, Casadonte R, Kriegsmann J, Dienemann H, Schirmacher P, Hendrik Kobarg J, et al. Reliable entity subtyping in non-small cell lung cancer by matrix-assisted laser desorption/ionization imaging mass spectrometry on formalin-fixed paraffin-embedded tissue specimens. *Mol Cell Proteomics*. 2016;15:3081–9.
19. Casadonte R, Kriegsmann M, Zweynert F, Friedrich K, Baretton G, Bretton G, et al. Imaging mass spectrometry to discriminate breast from pancreatic cancer metastasis in formalin-fixed paraffin-embedded tissues. *Proteomics*. 2014;14:956–64.
20. Bloom GC, Eschrich S, Zhou JX, Coppola D, Yeatman TJ. Elucidation of a protein signature discriminating six common types of adenocarcinoma: protein signature discriminating six adenocarcinomas. *Int J Cancer*. 2007;120:769–75.
21. Meding S, Nitsche U, Balluff B, Elsner M, Rauser S, Schöne C, et al. Tumor classification of six common cancer types based on proteomic profiling by MALDI imaging. *J Proteome Res*. 2012;11:1996–2003.
22. Zhang P-W, Chen L, Huang T, Zhang N, Kong X-Y, Cai Y-D. Classifying ten types of major cancers based on reverse phase protein array profiles. *PLoS ONE*. 2015;10:e0123147.
23. Longuespée R, Casadonte R, Schwamborn K, Reuss D, Kazdal D, Kriegsmann K, et al. Proteomics in pathology. *Proteomics*. 2018;18:1700361.
24. Spurrier B, Ramalingam S, Nishizuka S. Reverse-phase protein lysate microarrays for cell signaling analysis. *Nat Protoc*. 2008;3:1796–808.
25. Akbani R, Becker K-F, Carragher N, Goldstein T, König L, de Korf U, et al. Realizing the promise of reverse phase protein arrays for clinical, translational, and basic research: a workshop report: the RPPA (Reverse Phase Protein Array) Society. *Mol Cell Proteomics*. 2014;13:1625–43.
26. Treindl F, Ruprecht B, Beiter Y, Schultz S, Döttinger A, Staebler A, et al. A bead-based western for high-throughput cellular signal transduction analyses. *Nat Commun*. 2016;7. <https://doi.org/10.1038/ncomms12852>.
27. Lomnytska M, Pinto R, Becker S, Engström U, Gustafsson S, Björklund C, et al. Platelet protein biomarker panel for ovarian cancer diagnosis. *Biomark Res*. 2018;6:2.
28. Schumacher D, Andrieux G, Boehnke K, Keil M, Silvestri A, Silvestrov M, et al. Heterogeneous pathway activation and drug response modelled in colorectal-tumor-derived 3D cultures. *PLoS Genet*. 2019;15:e1008076.
29. Naskou J, Beiter Y, van Rensburg R, Honisch E, Rudelius M, Schlenz M, et al. EZH2 loss drives resistance to carboplatin and paclitaxel in serous ovarian cancers expressing ATM. *Mol Cancer Res*. 2019. <https://doi.org/10.1158/1541-7786.MCR-19-0141>.
30. Inder S, Bates M, Ni Labhrai N, McDermott N, Schneider J, Erdmann G, et al. Multiplex profiling identifies clinically relevant signalling proteins in an isogenic prostate cancer model of radioresistance. *Sci Rep*. 2019;9. <https://doi.org/10.1038/s41598-019-53799-7>.
31. Laner-Plamberger S, Oeller M, Poupardin R, Krisch L, Hochmann S, Kalathur R, et al. Heparin differentially impacts gene expression of stromal cells from various tissues. *Sci Rep*. 2019;9:1–15.
32. Shabihkhani M, Lucey GM, Wei B, Mareninov S, Lou JJ, Vinters HV, et al. The procurement, storage, and quality assurance of frozen blood and tissue biospecimens in pathology, biorepository, and biobank settings. *Clin Biochem*. 2014;47:258–66.
33. Fox CH, Johnson FB, Whiting J, Roller PP. Formaldehyde fixation. *J Histochem Cytochem*. 1985;33:845–53.
34. Werner M, Chott A, Fabiano A, Battifora H. Effect of formalin tissue fixation and processing on immunohistochemistry. *Am J Surg Pathol*. 2000;24:1016.
35. Broeckx V, Boonen K, Pringels L, Sagaert X, Prenen H, Landuyt B, et al. Comparison of multiple protein extraction buffers for GeLC-MS/MS proteomic analysis of liver and colon formalin-fixed, paraffin-embedded tissues. *Mol Biosyst*. 2016;12:553–65.
36. Shi SR, Key ME, Kalra KL. Antigen retrieval in formalin-fixed, paraffin-embedded tissues: an enhancement method for immunohistochemical staining based on microwave oven heating of tissue sections. *J Histochem Cytochem*. 1991;39:741–8.
37. Ikeda K, Monden T, Kanoh T, Tsujie M, Izawa H, Haba A, et al. Extraction and analysis of diagnostically useful proteins from formalin-fixed, paraffin-embedded tissue sections. *J Histochem Cytochem*. 1998;46:397–403.
38. Casadonte R, Caprioli RM. Proteomic analysis of formalin-fixed paraffin-embedded tissue by MALDI imaging mass spectrometry. *Nat Protoc*. 2011;6:1695–709.
39. Bader S, Zajac M, Friess T, Ruge E, Rieder N, Gierke B, et al. Evaluation of protein profiles from treated xenograft tumor models identifies an antibody panel for formalin-fixed and paraffin-embedded (FFPE) tissue analysis by reverse phase protein arrays (RPPA). *Mol Cell Proteomics*. 2015;14:2775–85.
40. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.r-project.org/> (2019). Accessed 10 Dec 2019.
41. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Liaw WHA, Lumley T, et al. gplots: various R Programming Tools for Plotting Data. <https://CRAN.R-project.org/package=gplots> (2019). Accessed 10 Dec 2019.
42. Karatzoglou A, Smola A, Hornik K, Australia (NICTA) NI, Maniscalco MA, Teo CH. kernlab: Kernel-Based Machine Learning Lab. <https://CRAN.R-project.org/package=kernlab> (2019). Accessed 10 Dec 2019.
43. Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F. C++-code C-CC (libsvm et al. e1071): Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. <https://CRAN.R-project.org/package=e1071> (2019). Accessed 18 May 2020.
44. Jed Wing MKC, Weston S, Williams A, Keefer C, Engelhardt A, Cooper T, et al. caret: Classification and Regression Training. <https://CRAN.R-project.org/package=caret> (2019). Accessed 10 Dec 2019.
45. Raschka S Model evaluation, model selection, and algorithm selection in machine learning. arXiv:1811.12808 [cs, stat]. <http://arxiv.org/abs/1811.12808> (2018). Accessed 10 Dec 2019.
46. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol*. 1995;57:289–300.
47. Moldvay J, Jackel M, Bogos K, Soltész I, Agócs L, Kovács G, et al. The role of TTF-1 in differentiating primary and metastatic lung adenocarcinomas. *Pathol Oncol Res*. 2004;10:85–88.
48. Kaufmann O, Fietze E, Dietel M, Mengers J. Value of p63 and cytokeratin 5/6 as immunohistochemical markers for the differential diagnosis of poorly differentiated and undifferentiated carcinomas. *Am J Clin Pathol*. 2001;116:823–30.

49. Pai RB, Pai SB, Lalitha RM, Kumaraswamy SV, Lalitha N, Johnston R, et al. Over-expression of c-Myc oncoprotein in oral squamous cell carcinoma in the south Indian population. 2009. <https://doi.org/10.3332/ecancer.2009.128>.
50. Volm M, Rittgen W, Drings P. Prognostic value of ERBB-1, VEGF, cyclin A, FOS, JUN and MYC in patients with squamous cell lung carcinomas. *Br J Cancer*. 1998;77:663–9.
51. Toon CW, Chou A, Clarkson A, DeSilva K, Houang M, Chan JCY, et al. Immunohistochemistry for Myc predicts survival in colorectal cancer. *PLoS ONE*. 2014;9:e87456.
52. Schleger C, Verbeke C, Hildenbrand R, Zentgraf H, Bleyl U. c-MYC activation in primary and metastatic ductal adenocarcinoma of the pancreas: incidence, mechanisms, and clinical significance. *Mod Pathol*. 2002;15:462–9.

## **Curriculum vitae**

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

## Komplette Publikationsliste

1) Leichsenring, J., Volckmar, A. L., Kirchner, M., Kazdal, D., Kriegsmann, M., Stögbauer, F., Bockmayr, T., Klauschen, F., Herth, F. J. F., Penzel, R., Warth, A., Schirmacher, P., Endris, V. & Stenzinger, A. **Targeted deep sequencing of effusion cytology samples is feasible, informs spatiotemporal tumor evolution, and has clinical and diagnostic utility.** *Genes Chromosomes Cancer*. 57, 70–79 (2018).

Impact Factor: 3.444

2) Jurmeister, P.\*, Bockmayr, M.\*, Seegerer, P., Bockmayr, T., Treue, D., Montavon, G., Vollbrecht, C., Arnold, A., Teichmann, D., Bressen, K., Schüller, U., von Laffert, M., Müller, K. R., Capper, D. & Klauschen, F. **Machine learning analysis of DNA methylation profiles distinguishes primary lung squamous cell carcinomas from head and neck metastases.** *Sci Transl Med*. 11 (2019). (\* equal contribution)

Impact Factor: 16.304

3) Bockmayr, T., Erdmann, G., Treue, D., Jurmeister, P., Schneider, J., Arndt, A., Heim, D., Bockmayr, M., Sachse, C. & Klauschen, F. **Multiclass cancer classification in fresh frozen and formalin-fixed paraffin-embedded tissue by DigiWest multiplex protein analysis.** *Lab Invest*. 100, 1288–1299 (2020).

Impact Factor: 4.197

## Danksagung

An erster Stelle möchte ich mich bei meinem Doktorvater Herrn Prof. Dr. Frederick Klauschen für das spannende Thema und die intensive Betreuung dieser Arbeit herzlich bedanken. Er hat mich in den letzten Jahren stets unterstützt, motiviert und sich immer Zeit für meine Fragen genommen. Für seine Hilfe und Geduld, sowie die zahlreichen Anregungen und unverzichtbaren Ratschläge bin ich sehr dankbar.

Besonders möchte ich auch Frau Dr. Denise Treue danken. Sie stand mir mit Rat und Tat zur Seite und hat mich bei der Durchführung der Arbeit sehr unterstützt. Vielen Dank für die ausführlichen Erklärungen und praktischen Tipps!

Mein Dank gilt außerdem Herrn Daniel Heim, Herrn Dr. Philipp Jurmeister, Frau Ines Koch und Frau Barbara Meyer-Bartell, die mir bei meinen Fragen immer weitergeholfen und mich unterstützt haben. Des Weiteren danke ich Herrn Dr. Christoph Sachse, Herrn Dr. Gerrit Erdmann, Frau Anja Arndt und Frau Julia Schneider für die fruchtbare Kooperation im Rahmen dieser Arbeit.

Herzlich möchte ich mich auch bei meiner Familie und Freunden bedanken, die durch ihre Unterstützung und Ermutigungen wesentlich zum Gelingen dieser Arbeit beigetragen haben. Großer Dank gilt meinen Eltern, die mir den Weg bis hierher ermöglicht haben und immer an mich geglaubt haben. Für den liebevollen Zuspruch danke ich meiner Oma, meiner Schwester Sophia, Camillo, Agnes und Theresa. Darüber hinaus bedanke ich mich besonders bei meinem Bruder Michael für die konstruktive Kritik und die vielen motivierenden Worte. Kimberley und Bernadette danke ich für die hilfreichen Anmerkungen und das Korrekturlesen des Manuskripts. Zuletzt möchte ich meinem Freund Michael für die liebevolle Unterstützung und stets aufheiternden Worte von ganzem Herzen danken.