



JOHANNES GUTENBERG
UNIVERSITÄT MAINZ

Relevanz von Protease- Netzwerken in Krebs und Entwicklung

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

am Fachbereich Biologie der
Johannes Gutenberg-Universität Mainz



vorgelegt von Dorothee Gößwein
geboren am 01.10.1979 in Stuttgart

Mainz, 2018

Was du liebst, lass frei. Kommt es zurück, gehört es dir - für immer.

- Konfuzius

INHALTSVERZEICHNIS

1	ZUSAMMENFASSUNG	1
	Abstract.....	4
2	EINLEITUNG	7
2.1	Erbinformation und Genexpression	8
2.2	Epigenetik und Entwicklungsprozesse.....	12
2.3	Tumorgenese	17
2.3.1	Tumor-Mikroumgebung	20
2.3.2	Extrazelluläre Matrix (ECM).....	21
2.4	Proteasen.....	22
2.4.1	Proteolytische Netzwerke	24
2.4.2	Threonin Aspartase 1	28
2.5	Personalisierte Therapie.....	32
2.6	Zielsetzung.....	34
3	ERGEBNISSE	36
3.1	Expressionsprofiling zur Identifikation neuer krankheits-relevanter Signalwege und potenzieller Biomarker	36
3.1.1	“Identification of novel disease-relevant signaling-pathways in head and neck cancer by genome-wide transcriptional profiling”	99
3.1.2	“Expressional analysis of disease-relevant signalling pathways in primary tumours and metastasis of head and neck cancers”	101
3.2	Proteasen und Steuerungsmechanismen in Karzinogenese und Entwicklung	117

3.2.1	„Protein Translocation Assays to probe protease function and screen for inhibitors“	117
3.2.2	„Overexpression of the catalytically impaired Taspase1 T234V or Taspase1 D233A variants does not have a dominant negative effect in T(4;11) leukemia cells“	119
3.2.3	„Fly versus man: evolutionary impairment of nucleolar targeting affects the degradome of Drosophila's Taspase1“	133
3.2.4	„TFIIA transcriptional activity is controlled by a ‚cleave-and-run‘ Exportin1 / Taspase1-switch“	148
3.3	Systematische, wissenschaftliche und epidemiologische Erfassung von Kopf-Hals-Tumordaten	150
3.3.1	Struktureller Aufbau der Datenbank „TUBA“	150
3.3.2	Epidemiologische und wissenschaftliche Daten	152
4	DISKUSSION UND AUSBLICK	160
5	MATERIAL UND METHODEN	173
5.1	Material	173
5.1.1	Allgemeine Verbrauchsmaterialien	173
5.1.2	Chemikalien und Reagenzien	173
5.1.3	Puffer und Lösungen	175
5.1.4	Geräte	178
5.1.5	Antikörper	179
5.1.6	Oligonukleotide	181
5.1.7	Vektoren	185
5.1.8	Zelllinien	188
5.1.9	Patienten-Charakteristika und -Gewebe	189
5.1.10	Software	190
5.2	Methoden	192
5.2.1	DNA- und RNA-Techniken	192
5.2.2	Zellkulturtechniken	200
5.2.3	Proteintechniken	203

5.2.4	Bioinformatik.....	206
6	REFERENZEN	209
7	PUBLIKATIONEN.....	227
7.1	Artikel.....	227
7.2	Vorträge.....	228
7.3	Posterbeiträge	228
8	APPENDIX.....	230
8.1	Abkürzungen.....	230
8.2	Zusätzliche Tabellen und Listen.....	234
8.3	Lebenslauf.....	244
8.4	Eidesstattliche Versicherung	245
8.5	Danksagung	246

ABBILDUNGSVERZEICHNIS

Abb. 1: Transkription und Translation.	9
Abb. 2: Schematische Darstellung des Prä-Initiationskomplex (PIC).	10
Abb. 3: Schematische Darstellung epigenetischer Modifizierungen an Chromatin.	13
Abb. 4: Schematische Darstellung der Domänen des Mixed Lineage Leukemia-Proteins MLL.	16
Abb. 5: "Hallmarks of Cancer".	18
Abb. 6: Übersicht Krebs-assoziiertes Signalwege.	19
Abb. 7: Übersicht über das Degradom und die Klassifizierung humaner Proteasen.	24
Abb. 8: Proteolytisches Interaktionsnetzwerk um Cathepsin B, MMP2 und MMP9.	26
Abb. 9: Schematische Darstellung des Taspase1-Transportmechanismus über Importin- α und NPM1/Crm1.	31
Abb. 10: Validierung der Microarray-Genexpressionsdaten mittels RT-PCR.	38
Abb. 11: Vergleich der relativen Expression des Gens FN1 in PTvs.N.	39
Abb. 12: Vergleich der relativen Expression für das Gen KRT24 in PTvs.N.	40
Abb. 13: Vergleich der relativen Expression für das Gen OLR1 in PTvs.N.	41
Abb. 14: Vergleich der relativen Expression für das Gen PRR4 in PTvs.N.	42
Abb. 15: Vergleich der relativen Expression für das Gen ARHGAP25 in Mvs.PT.	43
Abb. 16: Vergleich der relativen Expression für das Gen SERPINB4 in Mvs.PT.	44
Abb. 17: Grafische Darstellung der RMA-Normalisierung der rohen Microarray-Intensitätswerte.	45
Abb. 18: Grafische Darstellung der Quantilnormalisierung der RMA-normalisierten Arraydaten.	46
Abb. 19: Hierarchisches Clustering aller Affy-IDs (alle Gene) über alle Gewebeproben.	48
Abb. 20: Hierarchisches Clustering aller Affy-IDs (alle Gene) über die Gewebeproben aus Primärtumoren (PT) und Normalgewebe (N).	49
Abb. 21: Hierarchisches Clustering aller Affy-IDs (alle Gene) über die Gewebeproben aus Metastasen (M) und Primärtumoren (PT).	50
Abb. 22: Hierarchisches Clustering gefilterter Affy-IDs über die Gewebeproben aus Primärtumoren (PT) und Normalgewebe (N).	52
Abb. 23: Hierarchisches Clustering gefilterter Affy-IDs über die Gewebeproben aus Metastasen (M) und Primärtumoren (PT).	53
Abb. 24: Hierarchisches Clustering der Gensignatur von Primärtumoren (PT) und Normalgewebe (N).	54
Abb. 25: Hierarchisches Clustering der Gensignatur von Metastasen (M) und Primärtumoren (PT).	55
Abb. 26: Hierarchisches Clustering der Gensignatur von Metastasen (M) und Normalgewebe (N).	56
Abb. 27: Top 01-Netzwerk (<i>Ingenuity Pathway Analysis</i> [®]) differenziell exprimierter Gene in PTvs.N.	69
Abb. 28: Top 02-Netzwerk (<i>Ingenuity Pathway Analysis</i> [®]) differenziell exprimierter Gene in PTvs.N.	70
Abb. 29: Top 03-Netzwerk (<i>Ingenuity Pathway Analysis</i> [®]) differenziell exprimierter Gene in PTvs.N.	71
Abb. 30: Top 04-Netzwerk (<i>Ingenuity Pathway Analysis</i> [®]) differenziell exprimierter Gene in PTvs.N.	72
Abb. 31: Top 05-Netzwerk (<i>Ingenuity Pathway Analysis</i> [®]) differenziell exprimierter Gene in PTvs.N.	73
Abb. 32: Top 06-Netzwerk (<i>Ingenuity Pathway Analysis</i> [®]) differenziell exprimierter Gene in PTvs.N.	74
Abb. 33: Top 07-Netzwerk (<i>Ingenuity Pathway Analysis</i> [®]) differenziell exprimierter Gene in PTvs.N.	75
Abb. 34: Top 08-Netzwerk (<i>Ingenuity Pathway Analysis</i> [®]) differenziell exprimierter Gene in PTvs.N.	76

Abb. 35: Top 09-Netzwerk (<i>Ingenuity Pathway Analysis</i> [®]) differenziell exprimierter Gene in PTvs.N.....	77
Abb. 36: Top 01-Netzwerk (<i>Ingenuity Pathway Analysis</i> [®]) differenziell exprimierter Gene in Mvs.PT.....	78
Abb. 37: Top 02-Netzwerk (<i>Ingenuity Pathway Analysis</i> [®]) differenziell exprimierter Gene in Mvs.PT.....	79
Abb. 38: Top 03-Netzwerk (<i>Ingenuity Pathway Analysis</i> [®]) differenziell exprimierter Gene in Mvs.PT.....	80
Abb. 39: Top 04-Netzwerk (<i>Ingenuity Pathway Analysis</i> [®]) differenziell exprimierter Gene in Mvs.PT.....	81
Abb. 40: Top 05-Netzwerk (<i>Ingenuity Pathway Analysis</i> [®]) differenziell exprimierter Gene in Mvs.PT.....	82
Abb. 41: Top 06-Netzwerk (<i>Ingenuity Pathway Analysis</i> [®]) differenziell exprimierter Gene in Mvs.PT.....	83
Abb. 42: Top 07-Netzwerk (<i>Ingenuity Pathway Analysis</i> [®]) differenziell exprimierter Gene in Mvs.PT.....	84
Abb. 43: Der Upstream-Regulator CD44 in PTvs.N (<i>Ingenuity Pathway Analysis</i> [®]).....	87
Abb. 44: Der Wachstumsfaktor CTGF in PTvs.N (<i>Ingenuity Pathway Analysis</i> [®]).	88
Abb. 45: Der Upstream-Regulator FN1 in PTvs.N (<i>Ingenuity Pathway Analysis</i> [®]).....	89
Abb. 46: Der Upstream-Regulator MYC in PTvs.N (<i>Ingenuity Pathway Analysis</i> [®]).....	89
Abb. 47: Der Wachstumsfaktor FGF2 in PTvs.N (<i>Ingenuity Pathway Analysis</i> [®]).	90
Abb. 48: Der Wachstumsfaktor TGFB3 in PTvs.N (<i>Ingenuity Pathway Analysis</i> [®]).	91
Abb. 49: Der Rezeptor ERBB3 in PTvs.N (<i>Ingenuity Pathway Analysis</i> [®]).	92
Abb. 50: Das Zytokin Interferon Gamma (IFNG) in Mvs.PT (<i>Ingenuity Pathway Analysis</i> [®]).	93
Abb. 51: Die PI3K-Familie in Mvs.PT (<i>Ingenuity Pathway Analysis</i> [®]).....	94
Abb. 52: Der NFκB-Komplex in Mvs.PT (<i>Ingenuity Pathway Analysis</i> [®]).....	95
Abb. 53: Vergleich von NGS-Amplifikationsdaten ¹⁶⁷ mit Microarray-Expressionsdaten der vorliegenden Studie.	98
Abb. 54: Stark vereinfachte, beispielhaft dargestellte Datenbankstruktur.	151
Abb. 55: Kaplan-Meier Überlebenskurven für die UICC Stadien I, II, III, IVab (IV) und IVc. Die einzelnen Stadien sind, wie in der Legende angegeben, farblich unterschieden.....	155
Abb. 56: Kaplan-Meier-Überlebenskurven nach Alter, Geschlecht, Lokalisation des Primärtumors und Therapie.	158
Abb. 57: Schematische Darstellung der Zusammenhänge aus Hypoxie, epithelialer-zu-mesenchymaler Transition (EMT) und Remodellierung der extrazellulären Matrix (ECM).....	161
Abb. 58: Schematische Darstellung einer Splice Overlap Extension PCR.	196

TABELLENVERZEICHNIS

Tab. 1: Klinische Charakteristika der HNSCC Patienten.	37
Tab. 2: Gensignatur des Vergleichs PTvs.N (Filter: $\log_2FC \geq 1,5$ und p-Wert $< 0,0001$)	58
Tab. 3: Gensignatur des Vergleichs Mvs.PT (Filter: $\log_2FC \geq 1,5$ und p-Wert $< 0,0001$)	61
Tab. 4: Top Netzwerke PTvs.N (Ingenuity® Pathway Analysis)	67
Tab. 5: Top Netzwerke Mvs.PT (Ingenuity® Pathway Analysis)	68
Tab. 6: Aktivierte/Inhibierte Upstream Regulatoren PTvs.N (<i>Ingenuity Pathway Analysis</i> ®)	85
Tab. 7: Aktivierte/Inhibierte Upstream Regulatoren Mvs.PT (<i>Ingenuity Pathway Analysis</i> ®)	86
Tab. 8: Vergleich der Expressionsdaten mit Amplifizierungsdaten aus Next Generation Sequencing.	96
Tab. 9: Übersicht über alle in der Datenbank aufgezeichneten Tumore, aufgeteilt nach Lokalisation.	153
Tab. 10: Therapieerfolge in Abhängigkeit der UICC-Stadien.....	156
Tab. 11: Therapieerfolge in Abhängigkeit der Behandlung nach Primärtumoren (Alle).....	156
Tab. 12: Verwendete Chemikalien und Reagenzien	173
Tab. 13: Verwendete Puffer und Lösungen	175
Tab. 14: Verwendete Geräte.....	178
Tab. 15: Verwendete Antikörper	179
Tab. 16: Verwendete Sekundärantikörper.....	181
Tab. 17: Primer zur Sequenzanalyse	181
Tab. 18: Primer für Oligoannealing	182
Tab. 19: Primer zur quantitativen Analyse via RT-PCR.....	183
Tab. 20: Primer zur PCR-Synthese.....	184
Tab. 21: Ausgangsvektoren	185
Tab. 22: Humane Expressionsvektoren.....	186
Tab. 23: Drosophila-Expressionsvektoren	186
Tab. 24: Zelllinien und deren Ursprung.	188
Tab. 25: Medien zur Kultivierung von Zellen.	189
Tab. 26: Verwendete Software	190
Tab. 27: PCR Programme	193
Tab. 28: Pipettierschema für einen 50µl PCR-Ansatz.	193
Tab. 29: qPCR Programm	194
Tab. 30: Pipettierschema für einen 10µl Ansatz zur Phosphorylierung der 5'-Enden.	196
Tab. 31: Pipettierschema DNA-Restriktionsverdau: Kontrollansatz (10µl) und präparativer Ansatz (50µl).	197
Tab. 32: Zusammensetzung der SDS Acrylamidgelle	205
Tab. 33: Abkürzungsverzeichnis.....	230
Tab. 34: Durchschnittliche Microarray-Daten validierter Gene (PTvs.N)	234
Tab. 35: Durchschnittliche Microarray Daten validierter Gene (Mvs.PT)	234

Tab. 36: TOP100 Liste differenziell exprimierter Gene in PTVs.N, sortiert nach der FDR.	235
Tab. 37: TOP100 Liste differenziell exprimierter Gene in Mvs.PT, sortiert nach der FDR.	238
Tab. 38: TOP100 Liste differenziell exprimierter Gene in Mvs.N, sortiert nach der FDR.	241

1 ZUSAMMENFASSUNG

Immer mehr Menschen erkranken und sterben jedes Jahr an Krebs. Trotz einer sich ständig erweiternden Bandbreite an technischen und pharmakologischen Möglichkeiten haben die meisten Krebspatienten oftmals eine schlechte Prognose. Der Erfolg einer Behandlung könnte deutlich verbessert werden, wenn die durch evidenzbasierte Medizin evaluierte Therapie auf einem soliden Verständnis molekularbiologischer Grundlagen und Zusammenhänge beruht. Dieser translationale Ansatz (*from bench to bedside*) wird durch neueste 'Omics'-Technologien wie Genomsequenzierung, Expressionsprofiling und Proteinanalysen vorangetrieben sowie idealerweise durch Biobanken (für Gewebe, Zelllinien und Flüssigbiopsien) mit klinischen Daten in umfassenden Datenbanken verknüpft. Die an Bedeutung gewinnende Bioinformatik soll es dabei ermöglichen die expandierenden Datenmengen sinnvoll zu verknüpfen, bereinigend zu filtern, auszuwerten und für die weiterführende Interpretation zum Wohl des Patienten verständlich darzustellen.

Diese Arbeit konzentriert sich auf die Erforschung krebsrelevanter Signalwege, einschließlich Proteasenetze und Mechanismen, die zur Tumorentstehung und Metastasierung beitragen könnten. Zusätzlich scheinen diese Netzwerke zur Entwicklung von Therapieresistenzen beizutragen. Somit ist diese Wissensgrundlage essentiell zur Definition krebsrelevanter Zielstrukturen und Behandlungsstrategien.

Proteasen sind bereits klinisch anerkannte Zielstrukturen für eine Vielzahl von Krebserkrankungen. Die im Rahmen der Arbeit untersuchte Threonin-Aspartase 1 (Taspase1) wurde als proteolytischer Regulator der für die Entwicklung essentiellen Hox-Genexpression durch die Spaltung von Mixed Lineage Leukemia (MLL) beschrieben. Da das Taspase1-Proenzym sich autoproteolytisch zur Eigenaktivierung in eine α - und eine β -Untereinheit spaltet und postuliert wurde, dass sich je vier Untereinheiten zu einem $\alpha\beta\alpha$ -Heterodimer zusammenlagern, wurden die katalytisch inaktive Taspase1^{T234V} und die proteolytisch stark eingeschränkte Taspase1^{D233A} Variante in Zellen überexprimiert und mit Hilfe eines zellbasierten Translokations-Biosensor-Assays auf ihre Aktivität überprüft. Dabei konnte selbst eine 9-fache Überexpression der jeweiligen Mutante weder die *trans*- noch die *cis*-Aktivität einschränken und eine Prozessierung des AF4-MLL-Onkoproteins *in vivo* verhindern. Eine Überexpression der einzelnen Untereinheiten zeigte keinen *trans*-dominanten

ZUSAMMENFASSUNG

Effekt an der ko-exprimierten oder endogenen Protease. Eine weitere Überprüfung durch Ko-Expression der individuellen α - und β -Untereinheiten der Taspase1 zeigte keine Zusammenlagerung der Untereinheiten zu einem aktiven Enzymkomplex.

Die molekularbiologischen Details der autoproteolytischen Spaltung, bei dem das nukleophile Threonin²³⁴ der β -Untereinheit N-terminal freigelegt wird und sich das katalytische Zentrum konformativ ausbildet, sind bisher noch unklar, scheinen aber essentiell für die Aktivität der Taspase1 zu sein und sind evolutiv enorm konserviert. In *Drosophila melanogaster* konnte in der Arbeit das katalytisch aktive Nukleophil der *Drosophila* Taspase1 (dTaspase1) an der Stelle Threonin¹⁹⁵ identifiziert werden. Ein Vergleich der Spaltungsmotive der humanen Taspase1 (Q³[F,I,L,V]²D¹↓G^{1'}X^{2'}D^{3'}D^{4'}) gegenüber der dTaspase1 (Q³[F,I,L,M]²D¹↓G^{1'}X^{2'}X^{3'}) veranschaulichte zudem deutliche Unterschiede. Während für die humane Taspase1 ein Degradom von 27 Proteinen *in silico* vorgesagt werden konnte, ist das Degradom der dTaspase1 unter anderem auf Grund des weniger spezifischen Spaltungsmotifs auf potentiell 70 Proteine erweitert. Die humane Taspase1 wird über die Interaktion mit Nukleophosmin1 (NPM1) über Importin- α in den Kern transportiert und lokalisiert dort in den Nukleoli, wogegen der dTaspase1 die Interaktionsdomäne mit NPM1 (NoLS) fehlt und sie nicht an den Nukleoli akkumuliert. Wurde eine NoLS künstlich hinzugefügt, konnte eine Interaktion mit NPM1 nachgewiesen werden, die dTaspase1 sammelte sich an den Nukleoli und spaltete auch nukleäre Proteine. Die Translokation der Protease scheint als evolutiver Mechanismus der Feinregulation von Degradom und proteolytischer Aktivität zu fungieren. Nicht nur die Protease selbst, sondern auch ihre Ziele können durch die Veränderung ihrer Lokalisation reguliert werden. Das Taspase1-Substrat TFIIA, ein basaler Transkriptionsfaktor, wird nach der Produktion durch ein nukleäres Exportsignal (NES) gekoppelt an den Exportin1/Crm1 Exportrezeptor im Zytoplasma angereichert. Wird die Exportaktivität inhibiert oder wird der Transkriptionsfaktor durch Taspase1 proteolytisch prozessiert, verbleibt TFIIA im Kern und seine transkriptionelle Aktivität verändert. Die Bindung des TATA Box bindenden Proteins (TBP) an TFIIA wird verringert. Wird im Gegenzug die Proteolyse verhindert und somit der Export ermöglicht, wird der Effekt umgekehrt und führt zu einer Steigerung der Transkription von p16^{INK4a}. Eine Erhöhung der Expression der Taspase1 führt also dazu, dass ein Zellzyklusinhibitor und Tumorsuppressor niedriger exprimiert wird, was ihre Funktion als *non-oncogene addiction protease* bestätigt.

Solche transkriptionellen Veränderungen sind maßgeblich für die weiterführende Malignität hin zu einer metastasierenden Erkrankung bei Krebs verantwortlich. Um solche Verschiebungen

ZUSAMMENFASSUNG

weiterführend sichtbar zu machen, wurde der Expressionsstatus von Primärtumoren aus 15 Kopf-Hals-Tumor-Patienten mit ihren korrespondierenden Lymphknotenmetastasen sowie Normalgewebe als jeweilige Kontrolle mit Hilfe von genomweiten Microarrays verglichen und bioinformatisch ausgewertet. Unterschiedliche Signalwege waren dereguliert, beispielsweise involviert in die Remodellierung der extrazellulären Matrix, die epithelial-mesenchymale Transition, den Metabolismus, die Immunreaktivität wie auch in Protease-steuernde Netzwerke. Kürzlich veröffentlichte Hochdurchsatzsequenzierungsdaten des *Cancer Genome Atlas*-Netzwerks wurden mit den Expressionsdaten der Microarray-Analyse verglichen, und es konnten in der Tat wesentliche Unterschiede herausgestellt werden. Durch die Anwendung stringenter Filterkriterien konnten die wesentlichen differenziell exprimierten Gene herausgefiltert werden, und Zytokeratin24 (KRT24) als bisher unbekannter, potentiell prognostischer Biomarker identifiziert werden. Die reduzierte Expression von KRT24 in Tumoren im Vergleich zu gesunder Schleimhaut konnte sowohl auf mRNA- als auch auf Protein-Ebene nachgewiesen werden. Die transiente Expression von KRT24 beeinflusste nicht nur das Wachstum von Kopf-Hals-Tumorzelllinien, sondern reduzierte auch das Wachstum im murinen Xenograft-Modell, was die Wirkung von KRT24 als möglicher Tumorsuppressor und prognostischer Biomarker unterstreicht. Zur weiteren Erforschung potentieller neuer Biomarker wurde eine Tumordatenbank erstellt, in welcher klinische mit histopathologischen Daten zusammengeführt werden. Immunhistochemische Färbungen vielversprechender Kandidatenproteine können so neue Hinweise auf differentielle Proteinexpressions- und Lokalisationsmuster geben und im Zusammenhang mit korrespondierenden Patientendaten wie dem Krankheitsverlauf ausgewertet werden.

Zusammenfassend konnten im Rahmen dieser Arbeit zum einen neue Erkenntnisse der Protease-Regulationsmechanismen durch Transportereignisse gewonnen werden. Darüber hinaus erlauben die hier durchgeführten Studien neue Einblicke in die Erforschung tumor-assoziiertes Signalwege sowie die Identifizierung neuer Biomarker durch ein vertieftes molekulares Verständnis pathobiologischer Zusammenhänge, insbesondere bei Tumorerkrankungen des Kopf-Hals-Bereichs.

Abstract

More and more people are dying of cancer every year. Despite a constantly expanding range of technical and pharmacological possibilities, most cancer patients often have a poor prognosis. The success of treatment could be significantly improved if the therapy evaluated by evidence-based medicine is based on a solid understanding of biomolecular background and context. This translational approach (“from bench to bedside”) is driven by the latest 'omics' technologies such as genome sequencing, expression profiling and protein analysis and ideally linked by biobanks (for tissues, cell lines and liquid biopsies) to clinical data in comprehensive databases. Bioinformatics, which is gaining in importance, should make it possible to meaningfully combine the expanding data volumes, to clean them up by filtering, to evaluate them and to make them comprehensible for further interpretation for the benefit of the patient.

This thesis focuses on the study of cancer-related signalling pathways, including protease networks and mechanisms that could contribute to tumorigenesis and metastasis. In addition, these networks appear to contribute to the development of therapy resistance. Thus, this knowledge base is essential for the definition of cancer-relevant target structures and treatment strategies.

Proteases are already clinically recognized targets for a variety of cancers. Threonine aspartase 1 (Taspase1), which was investigated in this study, was described as a proteolytic regulator of developmentally essential Hox gene expression through the cleavage of mixed lineage leukemia (MLL). Since the Taspase1 proenzyme autoproteolytically cleaves into an α - and a β -subunit for self-activation and has been postulated to assemble four subunits into one $\alpha\beta\beta\alpha$ -heterodimer, the catalytically inactive Taspase1^{T234V} and the proteolytically restricted Taspase1^{D233A} variant were overexpressed in cells and checked for activity using a cell-based translocation biosensor assay. Even a 9-fold overexpression of the respective mutant could neither restrict the *trans* nor the *cis* activity and prevent processing of the AF4 · MLL oncoprotein *in vivo*. Overexpression of the individual subunits showed no *trans*-dominant effect on the co-expressed or endogenous protease. Further review by co-expression of the individual α - and β -subunits of Taspase 1 did not reveal any assembly of the subunits into an active enzyme complex.

ZUSAMMENFASSUNG

The biomolecular details of autoproteolytic cleavage, in which the nucleophilic threonine234 of the β -subunit is exposed N-terminally and the catalytic centre is conformationally formed, are still unclear, but seem to be essential for the activity of Taspase1 and are highly conserved. In *Drosophila melanogaster*, the catalytically active nucleophile of Drosophila Taspase1 (dTaspase1) at the site threonine195 could be identified. A comparison of the cleavage motifs of human Taspase1 ($Q^3[F,I,L,V]^2D^1\downarrow G^1X^{2'}D^{3'}D^{4'}$) with dTaspase1 ($Q^3[F,I,L,M]^2D^1\downarrow G^1X^{2'}X^{3'}$) also illustrated significant differences. While the human Taspase1 could be predicted to degrade 27 proteins *in silico*, the degradation of dTaspase1 is extended to potentially 70 proteins due to the less specific cleavage motif. The human Taspase1 is transported via the interaction with nucleophosmin1 (NPM1) by importin- α into the nucleus and localized in the nucleoli, whereas the dTaspase1 lacks the interaction domain with NPM1 (NoLS) and does not accumulate at the nucleoli. When a NoLS was added artificially, interaction with NPM1 could be detected; dTaspase1 accumulated on the nucleoli and also cleaved nuclear proteins. The translocation of the protease appears to function as an evolutionary mechanism to finely regulate degradation and proteolytic activity. Not only the protease itself but also its targets can be regulated by changing its localization. The Taspase1 substrate TFIIA, a basal transcription factor, is accumulated in the cytoplasm after production by a nuclear export signal (NES) coupled to the Exportin1/Crm1 export receptor. If the export activity is inhibited or if Taspase1 proteolytically processes the transcription factor, TFIIA remains in the nucleus and its transcriptional activity is altered. The binding of TATA box binding protein (TBP) to TFIIA is reduced. If, in turn, proteolysis is prevented and export is possible, the effect is reversed and leads to an increase in the transcription of p16^{INK4a}. An increase in the expression of Taspase1 thus leads to a lower expression of a cell cycle inhibitor and tumour suppressor, which confirms its function as a *non-oncogene addiction* protease.

Such transcriptional changes are critically responsible for the further malignancy towards metastatic cancer disease. In order to further visualize such shifts, the expression status of primary tumours from 15 head and neck tumour patients with their corresponding lymph node metastases as well as normal tissue as respective control was compared with the aid of genome-wide microarrays and evaluated bioinformatics. Different signalling pathways have been deregulated, for example, involved in extracellular matrix remodelling, epithelial-mesenchymal transition, metabolism, immunoreactivity, as well as in protease-controlling networks. Recently published high-throughput sequencing data from the Cancer Genome Atlas network was compared to expression data from

ZUSAMMENFASSUNG

microarray analysis, and significant differences could indeed be highlighted. By applying stringent filter criteria, the essential differentially expressed genes could be filtered out, and cytokeratin24 (KRT24) could be identified as a previously unknown, potentially prognostic biomarker. The reduced expression of KRT24 in tumours compared to healthy mucosa could be detected at both mRNA and protein levels. The transient expression of KRT24 not only affected the growth of head and neck tumour cell lines, but also reduced growth in the murine xenograft model, highlighting the effect of KRT24 as a potential tumour suppressor and prognostic biomarker. To further explore potential new biomarkers, a tumour database has been created in which clinical data are combined with histopathological data. Immunohistochemical staining of promising candidate proteins can thus provide new information on differential protein expression and localization patterns which can be evaluated in conjunction with corresponding patient data such as the course of the disease.

In summary, new insights into transport mechanisms to regulate protease activity could be gained in the context of this thesis. In addition, the studies conducted here allow new insights into the study of tumour-associated signalling pathways as well as the identification of new biomarkers through an in-depth molecular understanding of pathobiological relationships, especially in tumours of the head and neck.

2 EINLEITUNG

Krebs ist der Oberbegriff für eine Vielzahl maligner, neoplastischer Erkrankungen, welche nach Herz-Kreislaufkrankungen zu den Haupttodesursachen weltweit zählen ^{1,2}. All diese Neoplasmen zeichnen sich durch eine erhöhte Proliferation von Zellen aus, welche über natürliche Gewebegrenzen hinweg invasiv wachsen und schließlich im Endstadium in andere Organe und Teile des Körpers metastasieren können ³. Durch die Verdrängung gesunden Gewebes verlieren die Organe nach und nach ihre Funktion, was die Haupttodesursache bei Krebserkrankungen darstellt ². Maligne Neoplasmen waren im Jahr 2015 weltweit für 8,7 Mio. respektive 15,5% der Todesfälle verantwortlich, mit deutlich steigenden Zahlen seit dem Jahr 2000 ¹. Lungen-, Leber-, Magen- und Dickdarmkrebs gehören dabei zu den häufigsten tödlich endenden Tumorarten ^{1,2}.

Krebserkrankungen sind immer genetische Erkrankungen. Während in den gesunden Zellen des Körpers (mit Ausnahmen einiger Zellen des Immunsystems und der Reproduktion) die genetische Information überall identisch ist, ist sie in Krebszellen maßgeblich verändert. Die zellulären Programme, welche für die reguläre Funktion der Zellen sorgen, sind dadurch entweder ausgeschaltet oder derart modifiziert, dass sie dem Erhalt sowie der Vergrößerung und Verbreitung des Tumors dienen und nicht mehr dem Organismus als Ganzes. Es kann dem Verlauf der malignen Erkrankung nutzen, wenn der Tumor in der Lage ist, Mechanismen zu aktivieren, welche ursprünglich der Regeneration (z.B. Wundheilung) oder dem Aufbau von Gewebe und Organen in der Entwicklung dienen. Diese sind normalerweise stringent reguliert und erfordern ein Zusammenspiel mehrerer komplexer Prozesse, welche im Tumor durch die genetischen Veränderungen ausgehebelt werden oder teilweise sogar ins Gegenteil verkehrt werden. Genetische „Unfälle“ passieren dabei meist im „Alltag“: Durch Fehler in der Replikation oder verschiedene Umwelteinflüsse wie Karzinogene oder UV-Strahlung kann die DNA geschädigt werden - normalerweise werden solche Schäden durch DNA Reparaturmechanismen sofort korrigiert. Ist die ansonsten gesunde Zelle jedoch nicht in der Lage, die erforderlichen Reparaturen erfolgreich durchzuführen, leitet sie den programmierten Zelltod (Apoptose) ein. Damit wird verhindert, dass die geschädigte Zelle dem Organismus als Ganzes schaden kann. In Krebszellen ist auch dieser natürliche Mechanismus deaktiviert, so dass die Zelle trotz teils massiver DNA-Schäden weiterleben und sich teilen kann, was wiederum zur Anreicherung

weiterer Mutationen führt. Die Ursachen und das genetische Material der gleichen Krebsart ist in verschiedenen Tumoren desselben Typs daher niemals genau gleich und kann selbst innerhalb des Tumors variieren. Diese Heterogenität erschwert es enorm, erfolgreiche Therapien für diese Erkrankungen zu entwickeln. Um Krebserkrankungen zu verstehen, ist es demnach essentiell, die zu Grunde liegenden zellbiologischen und systemischen Mechanismen im Detail zu erforschen und Gemeinsamkeiten zu finden, welche es schließlich ermöglichen, sinnvolle Therapieansätze zu entdecken (frei zusammengefasst nach ³ und ⁴).

2.1 Erbinformation und Genexpression

Die Erbinformation einer Zelle wird in der DNA gespeichert und ist – abgesehen von hochspezialisierten Zellen wie einigen Immunzellen und reproduktiven Zellen sowie durch externe Einflüsse modifizierte Zellen – in allen Zellen eines Organismus gleich. Die Ausprägung einer Zelle, welche die Zellart bestimmt, wird durch die Auswahl der Gene definiert, welche wiederum von der DNA abgelesen und in regulatorische sowie strukturelle Komponenten oder Proteine umgesetzt werden (**Abb. 1**) ⁴.

Die Entscheidung in der Entwicklung einer Zelle, welche spezifische Kombination an Genen exprimiert wird, wird als Differenzierung bezeichnet und ist einer der wesentlichen Schritte zur Ausbildung von Geweben und essentiell für die Entwicklung des Organismus. Auch im späteren Verlauf des Lebenszyklus einer Zelle können Gene an- oder abgeschaltet werden, meist in Reaktion auf äußere Einflüsse. Hierzu wird vor allem die Transkription eines Gens oder einer bestimmten Gruppe von Genen gezielt erhöht oder erniedrigt, so dass mehr oder weniger Kopien des gleichen Gens hergestellt werden ⁴.

Ein Gen kann dabei nicht nur mehrere Proteine kodieren, sondern auch regulatorische RNA, Komponenten für Ribosomen (rRNA) sowie Spleißosomen (snRNA) oder für die Translation essentielle Transfer-RNA (tRNA) ⁵. Da die proteinkodierenden Gene nur etwa 2-3% des menschlichen Genoms ausmachen, wird angenommen, dass die Anzahl der nicht-kodierenden RNAs deutlich größer ist als die der mRNAs ⁶. Die Transkription wird in Eukaryoten von drei Enzymen durchgeführt: RNA Polymerase I (Pol I), Pol II und Pol III. Während Pol I und Pol III vor allem nicht-kodierende RNA herstellen, produziert die RNA Polymerase II mRNA, welche in Proteine umgeschrieben wird ⁵. Die

Regulation der Pol II ist abhängig von der Zelldifferenzierung, dient dem Erhalt der Zellidentität und der Reaktion auf Umwelteinflüsse. Generell kann der Prozess der Transkription an unterschiedlichen Stellen reguliert werden, am entscheidendsten ist jedoch die Phase der Transkriptionsinitiation, bei welcher der Prä-initiationskomplex (PIC) im Bereich des Promotors gebildet wird. Dabei assoziiert die Pol II mit den Transkriptionsfaktoren TFIIA, TFIIB, TFIID, TFIIIE, TFIIF und TFIIFH (**Abb. 2**)^{5,7}.

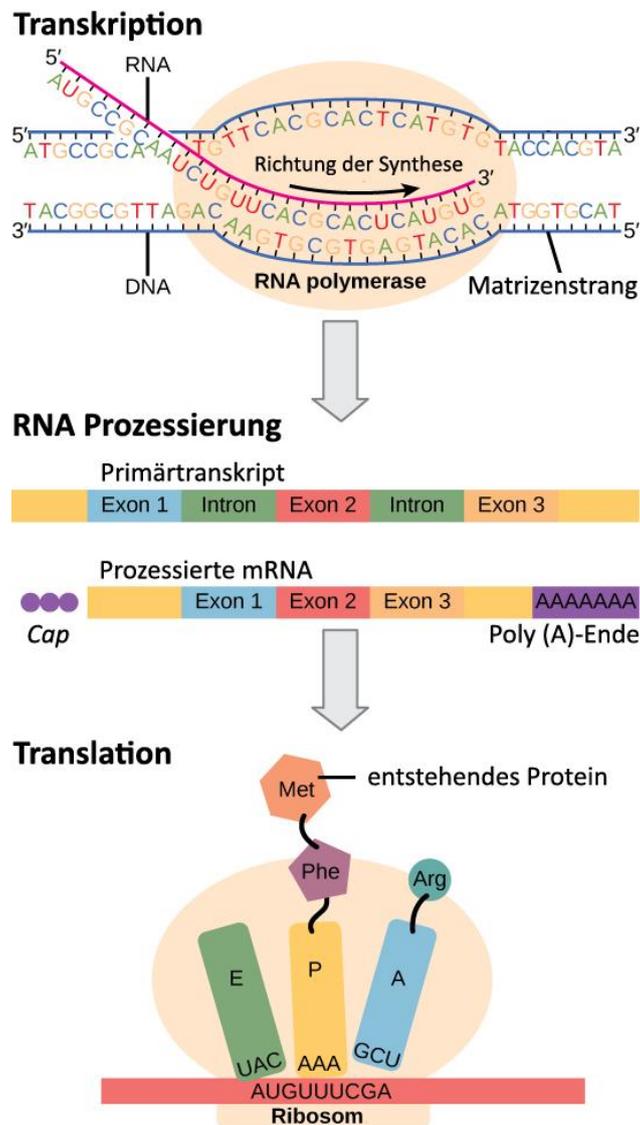
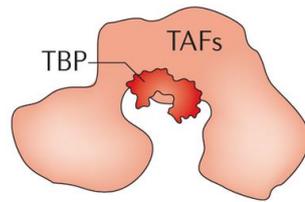


Abb. 1: Transkription und Translation.

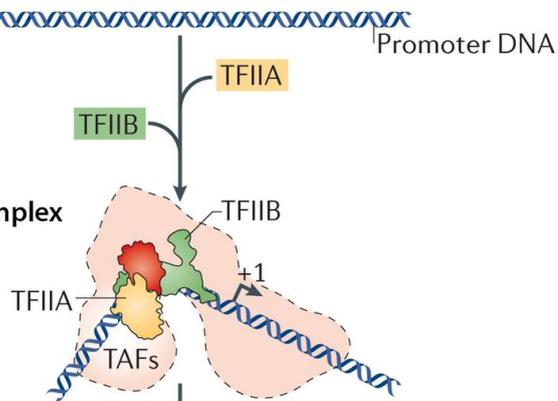
Die RNA (rot) wird durch das Enzym RNA Polymerase (rosa) komplementär zum DNA-Matrizenstrang (blau) als Primärtranskript im Zellkern synthetisiert. Anschließend wird die RNA im abgebildeten schematischen Beispiel durch das Herausschneiden von Introns (grün) gespleißt, dem Hinzufügen einer Kappe (engl. *Cap*, violette Punkte) und dem Poly(A)-Ende (violett) weiter prozessiert und als fertige mRNA aus dem Kern heraus zu den Ribosomen (rosa) transportiert. Dort wird die mRNA abgelesen und in eine Proteinsequenz translatiert. Adaptiert nach ⁸.

EINLEITUNG

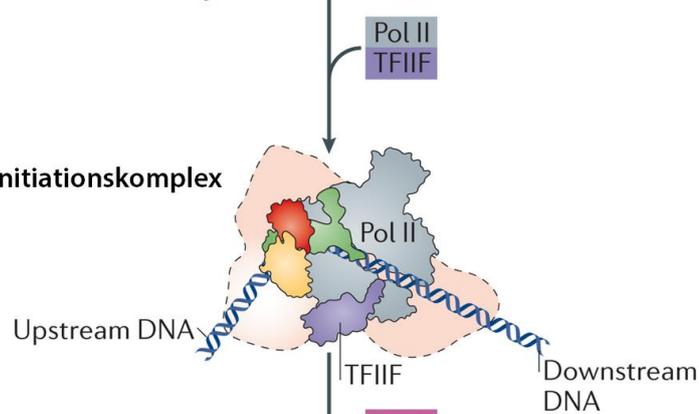
Ungebundener Promotor



Upstream Promotor Komplex



„Core“ Pre-Initiationskomplex (PIC)



Geschlossener PIC

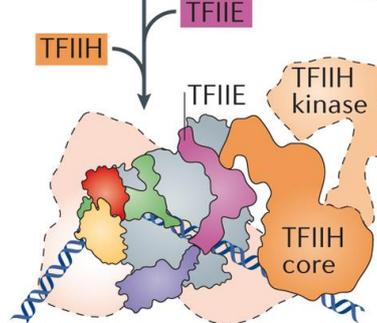


Abb. 2: Schematische Darstellung des Prä-Initiationskomplex (PIC).

Dargestellt ist das kanonische Modell der schrittweisen PIC-Zusammensetzung aus verschiedenen generellen Transkriptionsfaktoren und der Pol II an der Promotor-DNA (blau). TBP (rot) bindet direkt an die DNA, induziert damit eine 90° Beugung der DNA und rekrutiert TFIIB (grün) und TFIIA (gelb), die TBP flankieren und stabilisieren. Der resultierende Upstream Promotor-Komplex wird durch die Pol II (grau) und TFIIF (violett) zum Kern des PIC vervollständigt, der sog. „Core“ PIC. Unter Verwendung von ATP kann die DNA zur Transkriptionsblase geöffnet werden, nachdem TFIIE (orange) und TFIIE (magenta) gebunden haben, und die RNA Synthese wird initiiert. Verändert nach ⁷.

EINLEITUNG

Die generellen Transkriptionsfaktoren binden und öffnen die DNA im Bereich des Promotors und ermöglichen so den Zugang der Pol II zur DNA. Zudem stimulieren sie das Weiterlaufen der Pol II stromabwärts zur Synthese der RNA. Der Kernpromotor, um den sich der PIC bildet, besteht oft aus einer Kombination einer TATA-Box, dem Initiator-Motiv (Inr) und einem weiteren Downstream Promotor Element (DPE) und/oder dem Motif Ten Element (MTE) ^{5,7,9}.

Die TATA-Box ist etwa 30 Basenpaare (bp) stromaufwärts des Transkriptionsstarts lokalisiert und besitzt die Konsensussequenz TATAAWR (mit W=A/T und R=A/G) ¹⁰. Das TATA-Box bindende Protein (TBP) bindet universell an jeden Promotor als ersten Schritt der Bildung des PIC ⁷. Obwohl nicht alle Promotoren eine kanonische TATA-Box besitzen, konnte TBP an genomweiten Promotoren im Modellorganismus Hefe detektiert werden ¹¹.

TBP bindet in der kleinen Furche der DNA mit nanomolarer Affinität und knickt dabei den DNA-Strang um etwa 90° ab ¹². Mehrere TBP-assoziierte Faktoren (TAFs), die je nach Promotor und Gen variieren können, werden zu TBP rekrutiert und bilden so den Promotor-spezifischen generellen Transkriptionsfaktor TFIID ^{7,13}. TFIID ist in seiner konformativen Struktur hochkomplex. Etwa ein Drittel der Struktur ist enorm flexibel und kann seine Position um über 100 Å verändern. Diese konformative Plastizität steht dabei im direkten Zusammenhang mit der Bindekapazität des TFIID-Komplexes zur DNA, und damit zur Feinregulation der Promotor-Erkennung und schlussendlich der Fähigkeit zur Transkriptionsinitiation ⁹.

Der Transkriptionsfaktor TFIIA wird zur basalen und aktiven Transkription einer Zelle rekrutiert ¹⁴, ist aber nicht zwingend für die Transkriptionsinitiation benötigt. Unter anderem deshalb ist es derzeit in der Fachwelt strittig, ob TFIIA zu den generellen Transkriptionsfaktoren gezählt oder als Hilfs-Transkriptionsfaktor angesehen werden muss. Humanes TFIIA besteht aus den drei Untereinheiten α (35 kDa), β (19 kDa) und γ (12 kDa), welche von den Genen GTF2A1 ($\alpha\beta$) und GTF2A2 (γ) kodiert werden ¹⁵. TFIIA bindet zwar nicht direkt an die DNA, ist aber in der Lage, den TBP-DNA Komplex zu stabilisieren ¹⁶ und ist damit Pol II-spezifisch ^{7,17}. Die Stabilisierung des TBP-DNA Komplexes erfolgt über eine kompetitive Verdrängung der inhibitorischen TAND2-Domäne von TAF1 ¹⁸. Damit sind Konformationsänderungen innerhalb von TFIID verbunden, welche die zeitliche Abfolge und die Positionierung innerhalb des PIC Komplexes leicht verändern. TFIIA ist essentiell für die Zell-Viabilität und nimmt eine vielfältigere Rolle ein als die des alleinigen Transkriptionsfaktors ^{14,15,19}. Die zusätzlich assoziierte Funktion von TFIIA als „Anti-Repressor“ wurde schon früh verstanden und bereits

biochemisch bestätigt, indem der inhibitorische Effekt verschiedener negativer Co-Faktoren aufgehoben wurde²⁰⁻²². TFIIA scheint zudem die Transkriptionsregulation durch TBP-related Factors (TRFs) zu beeinflussen^{23,24} sowie durch das TBP-Like Protein (TLP) mit hoher Affinität gebunden zu werden^{25,26}, welches somit wiederum vor dem proteasomalen Abbau geschützt wird²⁷. TFIIA fungiert darüber hinaus als Co-Aktivatoren für mehrere Aktivatoren und verstärkt den Effekt anderer Co-Aktivatoren¹⁵.

2.2 Epigenetik und Entwicklungsprozesse

Zelluläre Programme zur Genexpression werden nicht nur von sequenzspezifischen Transkriptionsfaktoren beeinflusst, sondern auch maßgeblich durch Chromatin-Modifikationen reguliert. Das Zusammenspiel von DNA-Methylierung, Histonmodifikationen, der Packdichte der Nukleosomen und anderen Faktoren wie kleiner RNAs trägt zum sogenannten Epigenom bei. Durch dieses wird die Genexpression reguliert und die Zellen dazu befähigt, ihre Identität auch über Teilungen hinweg zu erhalten (**Abb. 3**)²⁸. Histonproteine dienen der Verpackung der DNA und können durch Methylierung, Acetylierung, Phosphorylierung, Ubiquitylierung und ADP-Ribosylierung post-translational modifiziert werden, um so die Verpackungsdichte vor allem durch Ladungsveränderungen zu steuern²⁹. Heterochromatische Regionen zeichnen sich durch Di- und Trimethylierungen am Lysin 9 des Histons H3 (H3K9me2 und H3K9me3) aus, welche die Voraussetzung für die darauffolgende Bindung des Heterochromatin Protein 1 (HP1) darstellen³⁰. Kleine funktionelle RNAs tragen zum Erhalt des Heterochromatins bei. DNA-Methylierung ist ubiquitär im Genom implementiert, mit Ausnahme der CpG Inseln (reich an Cytosin und Guanin)³¹, welche als Promotoren und Enhancer fungieren können und so häufig mit bestimmten Genen assoziiert sind³². Werden diese Regionen methyliert, führt dies zur Stilllegung des jeweiligen Gens³³. Die Trimethylierung H3K27me3 inaktiviert Gene in der entsprechenden DNA-Region, während die Monomethylierung von H3K4, H3K9, H3K27, H3K20 und H2BK5 und Histonacetylierung aktiv transkribierte Regionen markiert, vor allem am 5'-Ende der Gene, H3K36me am 3'-Ende^{34,35}. Am Transkriptionsstart aktiver Gene finden sich häufig die Modifikationen H3K4me3, H3K4me2, H3K4me1, sowie Histonacetylierungen und die Histonvariante H2A.Z²⁸.

EINLEITUNG

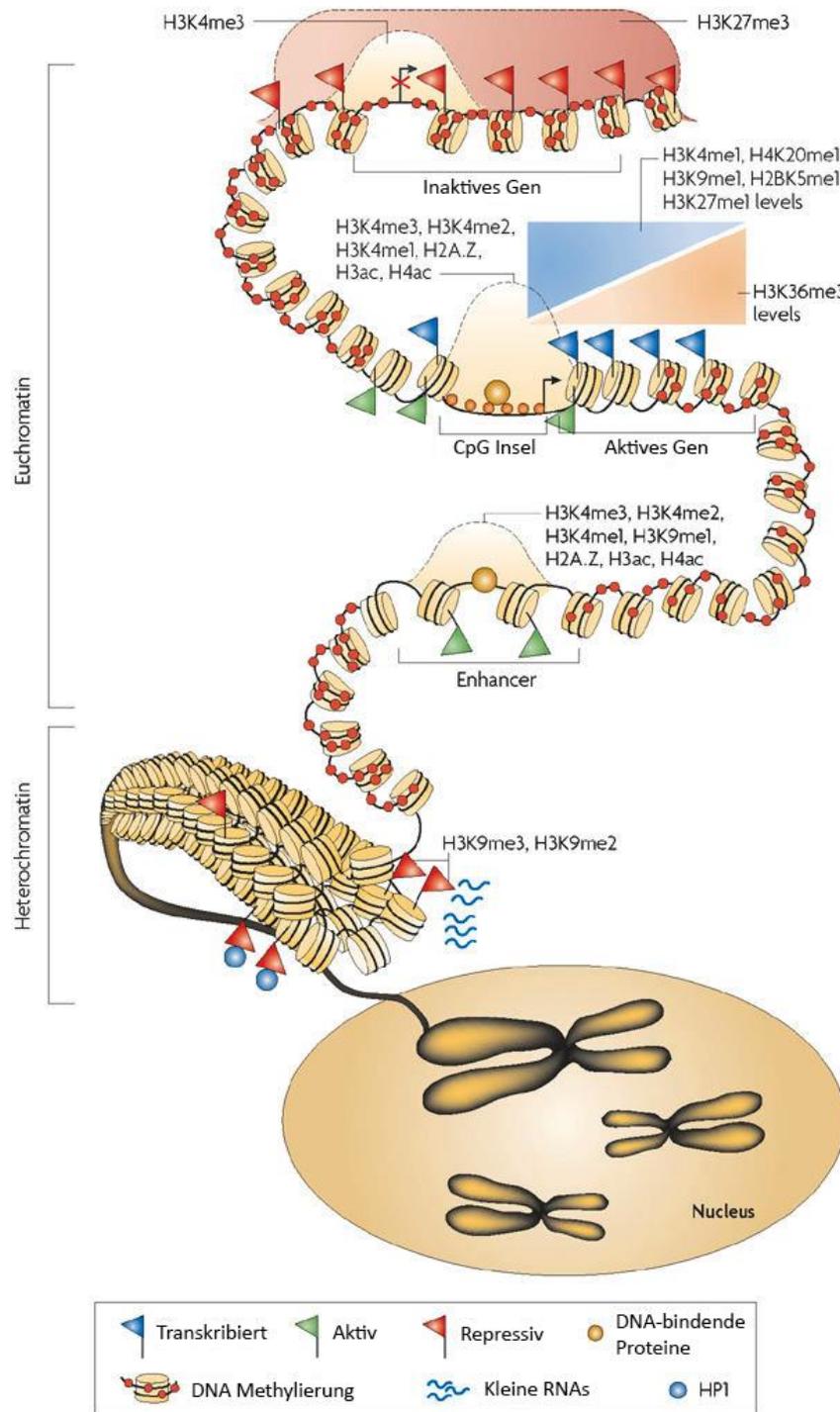


Abb. 3: Schematische Darstellung epigenetischer Modifizierungen an Chromatin.

Chromosomen (~1.400nm Durchmesser, dunkelgelb) sind in dichter gepackte, unzugängliche Regionen (Heterochromatin, ~30-300nm) und offenere, ablesbare Regionen (Euchromatin, ~2-30nm) aufgeteilt. Die DNA ist in Nukleosomen um Histone (gelb) gewickelt, welche durch unterschiedliche Modifikationen (rot) entweder einen leichteren (aktives Gen) oder schwereren Zugang (inaktives Gen) für den gesamten Transkriptionskomplex ermöglichen (modifiziert nach ²⁸).

EINLEITUNG

Das Epigenom kann sich je nach Gewebetyp und Entwicklungszustand der Zelle innerhalb eines Organismus unterscheiden. Für die Ausprägung einer Zelle und die Initiation eines Gewebes sind in der Entwicklung im Wesentlichen die Hox-Gene, eine Subgruppe der Homeobox Genfamilie, verantwortlich. Hox-Gene kodieren für Transkriptionsfaktoren, welche als Monomere oder Dimere direkt die Transkription steuern oder als Heterodimere und Heterotrimere mit TALE-Kofaktoren, einer weiteren Subgruppe der Homeobox Genfamilie, abhängig vom Zelltyp interagieren. Sie sind auf verschiedene Chromosomen verteilt und werden nach drei Grundprinzipien abgelesen: Räumliche Kollinearität, posteriore Überlegenheit und temporale Kollinearität. Mit räumlicher Kollinearität ist die Position der Gene auf den Chromosomen gemeint, die von 3' nach 5' „abgearbeitet“ werden. Gene, die also weiter in 3'-Richtung liegen, sind mit cranialen und anterioren Strukturen verbunden, während 5'-gelagerte Gene für kaudale und distale Strukturen zuständig sind. Gene in 5'-Richtung prägen in der Regel einen dominanteren Phänotyp aus und eine höhere Komplexität als Gene in 3'-Richtung, was als posteriore Überlegenheit bezeichnet wird. Die Hox Gene sind nicht nur räumlich zugeordnet, sondern werden korrespondierend zeitlich nacheinander abgelesen, was unter temporaler Kollinearität verstanden wird. (aus ³⁶)

Die Transkription der Hox-Gene selbst wird über Trithorax (TrxG)-Proteine hoch- und über Polycomb (PcG)-Proteine herunterreguliert, welche beide ursprünglich in der umgangssprachlich als Fruchtfliege bezeichneten Taufliege *Drosophila melanogaster*, einem Modellorganismus für Entwicklungsstudien, identifiziert wurden. Die Grundlage der Funktion der TrxG- und PcG-Proteine ist die gezielte Chromatinmodifikation in bestimmten DNA-Regionen, welche die Aktivierung oder Stilllegung spezifischer Gene erlaubt. Abhängig von ihrer Zugehörigkeit zu den Polycomb-Komplexen PRC1 und PRC2 sind PcG-Proteine grob in zwei Gruppen eingeteilt. Der PRC2-Komplex beinhaltet die Komponenten EZH2, EED, RbAp46/48 und SUZ12. Die katalytische Aktivität von EZH2 kann Lysine an Position 27 des Histons 3 methylieren (H3K27me). Zum PRC1-Komplex gehören neben weiteren Faktoren HPC13, HPH13, BMI1, RING1A. HPC13 bindet über seine Chromodomäne an H3K27me, welches durch den PRC2 Komplex modifiziert wurde, und verhindert so eine erneute Aktivierung des entsprechenden DNA-Abschnitts (zusammengefasst in ³⁷). Die TrxG-Proteine finden sich ebenfalls in Multiprotein-Komplexen wieder, können jedoch nicht so klar eingeteilt werden wie die PcG-Proteine. Einer der wichtigsten TrxG-Komplexe ist der Mixed Lineage Leukemia (MLL)-Komplex, welcher aus MLL1, MLL2, MLL3, WDR5, ASH2L, RbBP4 und CFP1 zusammengesetzt ist, und dessen ausschlaggebende Eigenschaft die Methylierung des Histons 3 am Lysin 4 (H3K4me) ist (aus ^{38,39}).

EINLEITUNG

Eine abnormale Regulation durch TrxG- und PcG-Proteine führt zu einer veränderten Genexpression, welche zur Tumorentstehung beitragen kann³⁷. Insbesondere das MLL-Gen ist an einer auf dem langen Arm von Chromosom 11 (Position 11q23) lokalisiert, einem strukturell instabilen Bereich, in welchem es häufig zu chromosomalen Brüchen kommt. Damit stellt das MLL-Gen ein häufiges Ziel von chromosomalen Translokationen dar, welche mit der Entstehung humaner Leukämien wie der „Acute Myeloid Leukemia“ (AML), der „Acute Lymphoblastic Leukemia“ (ALL) und der biphänotypischen „Mixed Lineage Leukemia“ (wie das Gen ebenfalls mit MLL abgekürzt) assoziiert sind³⁸. Insbesondere in Kleinkind-Leukämien tritt eine derartige Translokation im MLL Gen gehäuft (>70%) auf⁴⁰, während MLL-Translokationen insgesamt für etwa 10% aller humanen Leukämien verantwortlich gemacht werden⁴¹.

MLL (Gensymbol: KMT2A) kodiert mit insgesamt 37 Exons und 3969 Aminosäuren (AS) für eine große Histon-Methyltransferase, welche ubiquitär in allen Geweben sowie hämatopoetischen Zellen, darunter auch Stamm- und Progenitorzellen, exprimiert wird⁴²(**Abb. 4**). In *ChIP-on-chip*-Studien (Kombination aus Chromatin-Immunpräzipitation und DNA-Microarray-Hybridisierung) wurde MLL an 90% der 5.428 Gene, welche transkriptionell aktiv waren und von Pol II abgelesen wurden, detektiert und assoziiert mit mehreren Tausend Promotoren^{43,44}. Das 500 kDa-umfassende Progenitor-Protein wird von der Protease Taspase1 proteolytisch nach der AS 2.666 (Cleavage Site 1, CS1) und der AS 2.718 (Cleavage Site 2, CS2) in die nicht kovalent assoziierten Fragmente MLL^N (300 kDa) und MLL^C (180 kDa) gespalten⁴⁵.

MLL^N enthält drei AT-Hook Domänen (ATH1-3), zwei „gefleckte“ nukleäre Lokalisationssignale (SNL1/2), eine transkriptionelle Repressionsdomäne (TRD) und eine Pflanzenhomologiedomäne (PHD). In MLL^C befindet sich eine transkriptionelle Aktivierungsdomäne (TA) sowie die (Su(var)3-9, enhancer-of-zeste, trithorax)-Methyltransferasedomäne (SET). Über die AT1-3 und die TRD kann MLL die DNA unspezifisch direkt binden, interagiert aber auch mit anderen Proteinen wie Menin (MEN1) und gewinnt dadurch an Sequenzspezifität⁴⁶⁻⁴⁸. Die TA rekrutiert das CREB-binding Protein (CBP), das als Histon-Acetyltransferase die Genexpression aktivieren kann⁴⁹. Die SET-Domäne kann H3K4 methylieren und ist an der Komplexformation, Substraterkennung und -bindung beteiligt⁵⁰. C-Terminal der PHD bestehend aus den AS 1991-2104 befindet sich die FYRN-Domäne auf MLL^N, die mit der FYRC-Domäne, N-terminal der SET-Domäne an den AS 3651-3752 auf MLL^C, nach der proteolytischen Spaltung durch Taspase1 interagiert⁵¹.

EINLEITUNG

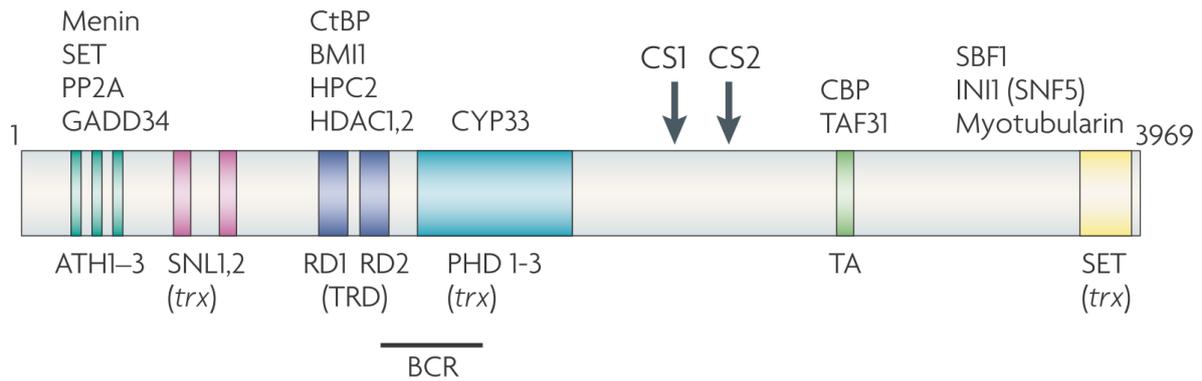


Abb. 4: Schematische Darstellung der Domänen des Mixed Lineage Leukemia-Proteins MLL.

Unterhalb der abstrahierten Proteinstruktur sind die Domänen des MLL Proteins benannt, oberhalb die potentiellen Bindepartner der jeweiligen Domänen und die zwei Spaltstellen CS1 und CS2 (Pfeile). ATH: AT-Hook-Domäne (dunkelgrün), SNL: „geflecktes“ nukleäres Lokalisationssignal (magenta), (T)RD: (Transkriptionelle) Repressionsdomäne (blau), PHD: Pflanzenhomologiedomäne (cyan), TA: Transkriptionelle Aktivationsdomäne (hellgrün), SET: (Su(var)3-9, enhancer-of-zeste, trithorax)-Methyltransferasedomäne (gelb), BCR: Bruchstellencluster (Exons 8-13) (modifiziert nach ³⁸).

Die häufigsten Translokationspartner in MLL-vermittelten Leukämien sind die Serin- und Prolinreichen nukleären Proteine AF4, AF9, ENL und AF10, respektive t(4;11)(q21;q23), t(9;11)(p22;q23), t(11;19)(q23;p13.3), und t(10;11)(p12;q23) ^{41,52}. Aus der Translokation t(4;11)(q21;q23) entstehen die beiden Fusionsproteine MLL·AF4 und AF4·MLL. Das Translokationsprodukt MLL·AF4 wurde als das für die Ausbildung der Leukämie verantwortliche Protein identifiziert und erhält sein onkogenes Potential meist gemeinsam mit nur wenigen weiteren Mutationen ^{53,54} durch die epigenetische Aktivierung der Transkription multipler Genloci ^{55,56}. MLL·AF4 bindet zudem an Enhancer der antiapoptotischen Bcl-2 Familie, moduliert die H3K27-Acetylierung und aktiviert so die Transkription von BCL-2 und MCL-1 ⁵⁷. Die Rolle des AF4·MLL Produkts wird kontrovers diskutiert, da AF4·MLL zwar im murinen Modell ALL ohne MLL·AF4 auslösen kann, aber nur in etwa 50%-80% der Patienten detektiert wird ^{52,54,58,59}. Das AF4·MLL Fusionsprodukt verändert ebenfalls die epigenetische Signatur, wird aber im unprozessierten Zustand von den E3-Ubiquitin-Ligasen SIAH1 und SIAH2 zum proteasomalen Abbau markiert ^{51,60}, was die Detektion und experimentelle Analyse des AF4·MLL Produkts deutlich erschwert. Wird AF4·MLL jedoch durch Taspase1 proteolytisch gespalten, lagert sich das AF4·MLL^N Spaltprodukt durch die FYRN-Domäne an die FYRC-Domäne des MLL^C Spaltprodukts an, was eine Ubiquitylierung durch SIAH1 und 2 verhindert ⁵¹.

2.3 Tumorgenese

Die Vorgänge, welche zur Krebsentstehung führen, gestalten sich nicht nur für Leukämien hochkomplex. Während sich in letzteren häufig Translokationen ursächlich für das onkogene Potential in hämatopoetischen Zellen erweisen, verändert in soliden Tumoren oft eine Ansammlung ungünstiger Mutationen maßgeblich die Eigenschaften der verantwortlichen Zellen.

Es gibt mehrere Modelle, welche die Gemeinsamkeiten und allgemeinen Zusammenhänge darzustellen versuchen. Das sicherlich bekannteste sind die *Hallmarks of Cancer*, die von Hanahan und Weinberg erstmals 2000 beschrieben wurden und 10 Jahre später mit neuen Erkenntnissen stark überarbeitet und ausgeweitet wurden (**Abb. 5**)^{3,61}. Hanahan und Weinberg sehen die Tumorentstehung als einen schrittweisen Prozess, in dem nach und nach zehn Hauptcharakteristika eines Tumors ausgebildet werden. Die ursprünglichen sechs Eigenschaften beinhalteten eine fortlaufende Stimulation zur Proliferation, die Verhinderung des Zelltods durch Apoptose, die Vermeidung von Wachstumssuppressoren, die Fähigkeit zur replikativen Unsterblichkeit, die Induzierung der Angiogenese und die Aktivierung von Invasion und Metastasierung. Hinzu kamen später die Deregulierung des Zellmetabolismus, genomische Instabilität, tumorfördernde Entzündung und Vermeidung der Zerstörung durch das Immunsystem.

In gesundem Gewebe ermöglichen fein abgestimmte Signale, welche zwischen den Zellen ausgetauscht werden, dass Lücken im Zellverband sehr präzise geschlossen werden und nicht zu wenige oder zu viele Zellteilungen stattfinden. Die Proliferationssignale, die die Zelle erhält, werden über sich amplifizierende Signalwege in den Zellkern weitergereicht und verändern dort die Transkription, so dass die Zelle ihre Masse erhöht, die DNA repliziert wird und es schließlich zur Teilung kommt. An jedem Punkt dieses mehrstufigen Prozesses kann es in Zellen zu malignen Veränderungen kommen, welche meist durch aktivierende Mutationen sogenannter Onkogene (z.B. RAS) oder inaktivierende Mutationen in Tumorsuppressorgenen (z.B. p53) ausgelöst werden. RAS ist ein Teil der Signalkaskade, die zur Proliferation der Zelle führt, zum Beispiel als Mechanismus der Wundheilung. Wird RAS übermäßig produziert oder durch Mutation dauerhaft aktiviert, wird sich die Zelle unkontrolliert teilen, ihren Energiemetabolismus umstellen und diese Fähigkeit wahrscheinlich an die nächste Zell-Generation weiterreichen. p53 dient in gesunden Zellen dazu, im Falle von DNA-Schäden die Zellteilung zu verhindern und bei Irreparabilität die Apoptose einzuleiten. Wird p53 also

ausgeschaltet, werden gleich mehrere „Hallmarks“ abgedeckt: genomische Instabilität und Vermeidung von Apoptose (Abb. 6). Die genomische Instabilität ermöglicht es dem entstehenden Tumor, durch weitere Mutationen alle „Hallmarks of Cancer“ schrittweise über Mechanismen der Klon-Selektion zu aktivieren, solange einzelne Zellen überleben (frei zusammengefasst nach ³).

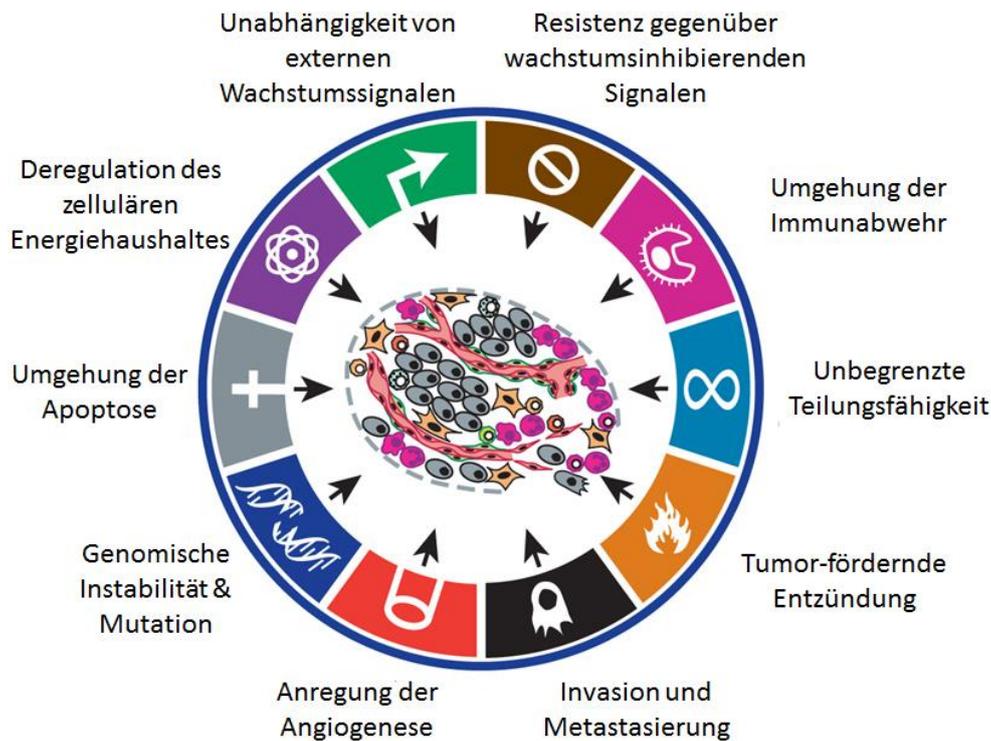


Abb. 5: "Hallmarks of Cancer".
Details siehe Text (verändert nach ³).

Basierend auf dem Ansatz einzelner, treibender Initial-Mutationen und der Möglichkeit, die Ursache von Krebs auf wenige molekulare Mechanismen vereinfachen zu können, ging man vor der Entwicklung von Hochdurchsatzsequenzierungstechniken (NGS) davon aus, dass man nur wenige, immer gleiche oder zumindest ähnliche Mutationen finden würde. Die Signalweg-Theorie, aus der sich die *Hallmarks of Cancer* entwickelten, bietet eine mögliche Erklärung für die schier unerschöpfliche Zahl heterogener Mutationen, welche mittels NGS entdeckt wurden ⁶¹.

EINLEITUNG

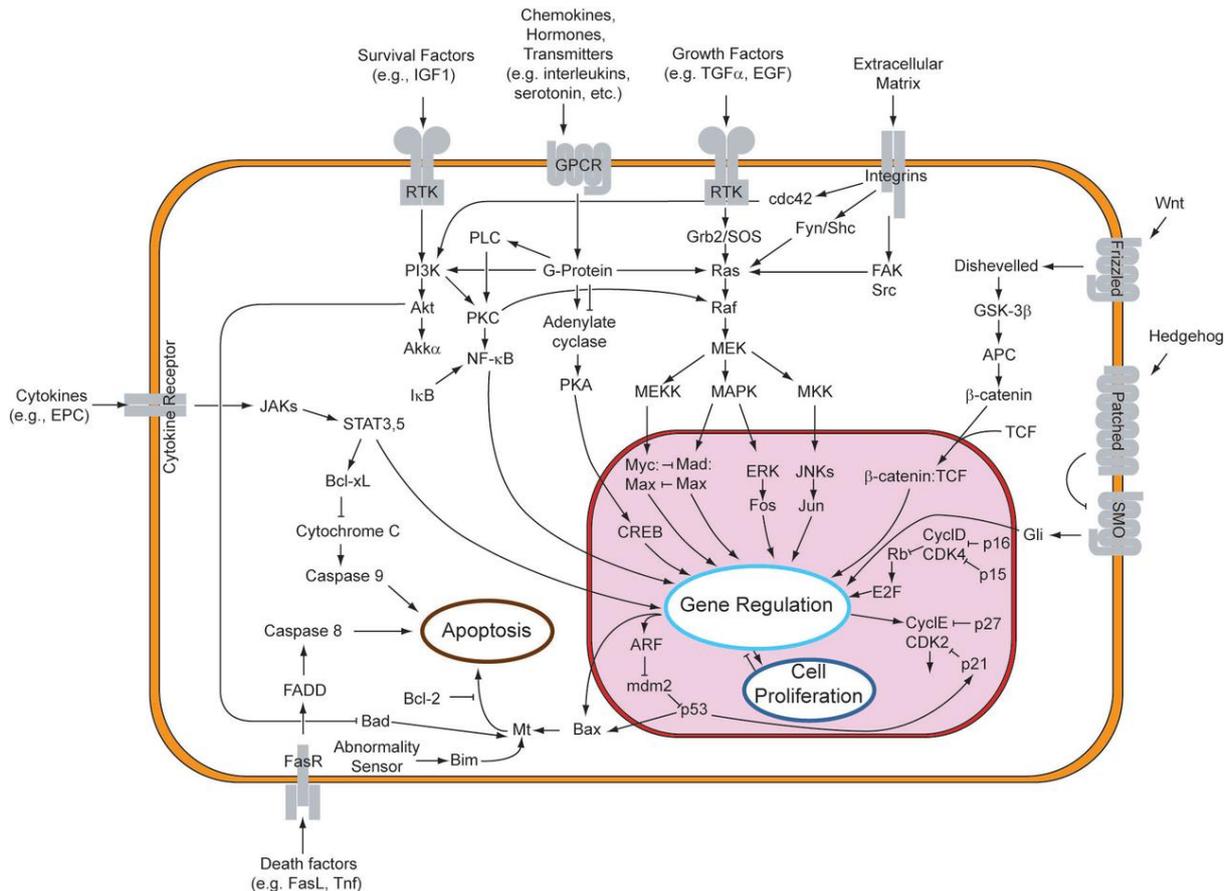


Abb. 6: Übersicht Krebs-assoziiertes Signalwege.

Diese stark gekürzte schematische Darstellung bietet eine Übersicht über die Zusammenhänge und Vernetzung der wichtigsten mit Krebs in Verbindung gebrachten Signalwege. Zu den wichtigsten Knotenpunkten der in dieser Arbeit näher betrachteten Signalwege gehören RAS, PI3K, G-Proteine, GSK-3β, die JAK/STAT-Achse, FADD und p53. Neben der schematisierten Zellmembran (orange) und dem Zellkern (rot) sind auch einige der assoziierten Zellprozesse wie die Apoptose (braun), Genregulation (cyan) und Proliferation (blau) farblich hervorgehoben (Quelle: Wikipedia Commons, freie Lizenz, abgerufen am 08.08.2017.).

Die Vielfalt der Gensignaturen von Krebs, die seither eine Klassifizierung erschweren, hat jedoch ebenfalls eine neue Theorie der Genom-Evolution hervorgebracht ⁶². Basis dieser sogenannten Genom-Evolutionstheorie ist die kausale Komplexität einer Tumorerkrankung und deren ständige Weiterentwicklung und Veränderung innerhalb der Erkrankung selbst. Die Theorie geht davon aus, dass die Hallmarks nicht unabhängig voneinander sind, sich teilweise selbst widersprechen und damit keine Basis liefern, um die Kanzerogenese als Gesamt-Theorie zu erklären. Innerhalb der Genom-Evolutionstheorie wird das Genom an sich umdefiniert und nicht mehr als Gesamtheit gleicher Gene eines Organismus gesehen, sondern als sich veränderndes System evolutionärer Adaptation auch auf der somatischen Ebene. Krebs wäre damit als genomische Systemerkrankung zu sehen, welche alle

Charakteristika von Krebszellen miteinander als „unterschiedlich-zu-Normalgewebe“ vereint und sich beständig nach evolutionären Mechanismen heterogen verändert. Diese Mechanismen funktionieren im Allgemeinen nach den drei Prinzipien Systemstress, Populationsdiversität und genomgesteuerte makrozelluläre Evolution (frei zusammengefasst nach ⁶²).

2.3.1 Tumor-Mikroumgebung

Unabhängig davon, welche Theorie zu Grunde gelegt wird, kristallisiert sich in den letzten Jahren immer mehr heraus, dass die Gesamtheit des Organismus und die Umgebung, in welcher der Tumor entsteht, mit betrachtet werden muss. Während Tumorzellen sich durch genetische Instabilität auszeichnen, sind Stromazellen, die den Tumor umgeben und durchziehen, in der Regel genetisch stabil und fungieren als konstantes Unterstützungssystem für den Tumor. Die Fähigkeit zur bidirektionalen Kommunikation zwischen Zellen und ihrer Mikroumgebung besitzen auch normale Gewebe, und sie ist für den Erhalt der zellulären wie organismischen Homöostase unerlässlich. Die Interaktionen zwischen Tumorzellen und dem Stroma unterliegen jedoch einer eigenen Dynamik, die den Verlauf der Erkrankung maßgeblich beeinflussen kann ⁶³. Die Heterogenität der Tumore betrifft also nicht nur die Tumorzellen selbst, sondern auch die Zusammensetzung der Art der Stromazellen und ihren Aktivierungsstatus ⁶⁴. Zonen chronischer Entzündung bieten häufig günstige Entstehungsbedingungen für Tumore und korrelieren mit der Tumorzinzidenz ⁶⁵. Eine Beeinträchtigung des Immunsystems hingegen, z.B. ausgelöst durch AIDS oder Immunsuppressiva nach Organtransplantationen, begünstigen ebenfalls direkt die Entstehung von Tumoren, während ein gesundes Immunsystem schützend wirkt ^{66,67}. Die Komplexität der Situation wird durch die Tatsache weiter erhöht, dass einige tumor-assoziierte Makrophagen (TAMs) gerade unter Einfluss bestimmter Zytokine wie Interleukin-4 (IL-4) und Transforming Growth Factor- β (TGF- β) tumorfördernde Eigenschaften besitzen, aber mit einigen pharmakologischen Substanzen zu einem tumorbekämpfenden Verhalten umprogrammiert werden können ^{68,69}. Diese TAMs werden durch Sauerstoffarmut und dadurch ausgeschüttete Rekrutierungsstoffe wie *Endothelin-2* und dem *Vascular Endothelial Growth Factor* (VEGF) angezogen und akkumulieren so im hypoxischen Bereich eines schnell wachsenden Tumors, was mit der tumorfördernden Polarisierung der Makrophagen korreliert und möglicherweise als molekularer Schalter fungiert ^{70,71}. VEGF wird dabei vor allem von Krebs-Assoziierten-Fibroblasten (*cancer-associated fibroblasts*, CAFs) sekretiert, die so die vaskuläre

Permeabilität ermöglichen, um die dauerhafte Blutversorgung des wachsenden Tumors mittels Angiogenese zu gewährleisten⁷². Diese Fibroblasten sind zudem in der Lage, die extrazelluläre Matrix (ECM), die dem Gewebe Struktur gibt und mit speziellen Membranen Gewebegrenzen festlegt, durch die Rekrutierung spezifischer Proteasen umzumodellieren und abzubauen. Durch die Ausschüttung von *A Disintegrin And Metalloprotease (ADAM)-12* und ADAM-17 wird in CAFs und umgebenden Tumorzellen die Sekretion der Matrix-Metalloproteinasen (MMP)-2 und MMP-9 angeregt, die in der Lage sind, die ECM maßgeblich umzugestalten und Gewebegrenzen durchlässig zu machen⁷³. CAFs schütten des Weiteren das Chemokin C-X-C Motif Ligand 12 (CXCL12) aus, das nach Bindung an dem Chemokin C-X-C Motif Receptor Type 4 (CXCR4) die Sekretion von MMP-9 aus den Tumorzellen in deren Mikroumgebung stimuliert^{74,75}.

2.3.2 Extrazelluläre Matrix (ECM)

Grundsätzlich moduliert die extrazelluläre Matrix die Kommunikation zwischen den Zellen innerhalb von Geweben und ist für den Erhalt der Gewebe-Homöostase unerlässlich⁷⁶. Die ECM besteht aus einem komplexen Geflecht von Makromolekülen, die in dreidimensionalen Strukturen geordnet sind. Spezielle biomechanische und biochemische Eigenschaften ermöglichen die Regulation von Zellwachstum, Überleben, Beweglichkeit und Differenzierung durch die ligandenspezifische Bindung von Rezeptoren wie Integrin, Syndecan und Discoidin Rezeptoren und bilden das Fundament für die mechanische Integrität eines Gewebes, das flexibel auf Umweltbedingungen reagieren und dementsprechend angepasst werden kann^{77,78}. Die Hauptbestandteile der extrazellulären Matrix sind Faserproteine wie Kollagene, Elastine, Fibronektine und Laminine, Proteoglykane wie Chondroitin-, Heparan- und Keratan-Sulfat sowie Hyaluronsäure, die von Bindegewebszellen wie Fibroblasten sekretiert werden⁷⁹. Erst der gezielte Umbau der ECM ermöglicht es Tumorzellen, in umliegende Gewebe invasiv einzuwachsen zu können, sich aus dem Zellverband zu lösen und in die Blutbahn zu gelangen⁷⁷. Mit dem Durchbrechen der physikalischen Barrieren werden Rho und Rac GTPasen ausgeschüttet, welche die Mobilität der Zellen stimulieren und deren epitheliale-zu-mesenchymale-Transition (EMT) begünstigen⁸⁰. Die EMT wird weiter durch TGF- β gefördert, das durch infiltrierende Immunzellen ausgeschüttet wird und den Phosphatidylinositol 3-Kinase (PI3K)-Signalweg aktivieren kann^{81,82}. Dieser Signalweg erhöht zudem die Expression der Glucose-Rezeptoren GLUT1 und GLUT4, was die zelluläre Aufnahme von Glucose deutlich erhöht und den

Metabolismus zu Gunsten des Warburg-Effektes verschiebt ⁸³. PI3K, AKT und *Mammalian Target of Rapamycin* (mTOR) kontrollieren vielfältige zelluläre Funktionen wie den Zellzyklus, Überleben, Metabolismus, Mobilität und genomische Instabilität und decken damit viele der oben beschriebenen *Hallmarks of Cancer* ab ³. Das PI3KCA-Onkogen, welches für die katalytische PI3K-Isoform p110 α kodiert, und das Tumorsuppressorgen *Phosphatase And Tensin Homolog* (PTEN), das für die wichtigste PtdIns(3,4,5)P₃-Phosphatase kodiert, gehören zu den am häufigsten mutierten Genen ⁸⁴. Selbst wenn diese zwei Gene nicht mutiert sind, ist die PI3K-AKT-mTOR Achse in fast allen Fällen auf unterschiedlichen Wegen aktiviert und stellt einen wesentlichen Marker für die Malignität eines Tumors dar ⁸⁵.

Zu welchem Zeitpunkt ein Tumor die Fähigkeit zu Invasion und Metastasierung erlangt, ist sehr unterschiedlich und kann entgegen früherer Annahmen bereits in sehr frühen Stadien während der Tumorentwicklung geschehen ⁸⁶. Metastasen sind unbestritten die Haupttodesursache bei Krebs, da durch die unkontrollierte Verdrängung von gesundem Gewebe in lebenswichtigen Organen nach und nach die Funktion des Organismus verloren geht ². Sogar kleine Tumore können Millionen zirkulierender Krebszellen (*circulating tumour cells*, CTCs) in die Blutbahn abgeben, was jedoch nicht zwangsläufig zur Ausbildung von Fernmetastasen führt. Die Wahrscheinlichkeit, dass sich diese Zellen in anderen Geweben dauerhaft einnisten können, ist sehr gering, und selbst niedergelassene, sogenannte disseminierte Tumorzellen (*disseminated tumour cells*, DTCs) entwickeln sich nur zu etwa 50% zu Metastasen ^{87,88}. Der Erfolg einer solchen Kolonisierung hängt maßgeblich davon ab, ob im Zielgewebe eine entsprechend unterstützende Mikroumgebung vorhanden ist, z.B. durch im Gewebe prä-existierende CXCR4/CXCL12-Signale, TGF- β -Sekretion aus den eindringenden Tumorzellen und darauffolgende Periostin (POSTN)-Aktivierung oder systemisch über LOX oder VEGF vermittelte Signale ⁸⁶.

2.4 Proteasen

Proteasen spielen nicht nur beim Abbau und der Remodellierung der extrazellulären Matrix eine wichtige Rolle, sondern auch bei vielen anderen zellulären Prozessen. Während Proteasen ursprünglich als unspezifisch agierende Abbau-Enzyme gesehen wurden, erlauben sie über die spezifische Hydrolyse von Peptidverbindungen tatsächlich eine äußerst präzise Kontrolle biologischer

Vorgänge auf zellulärer Ebene. Durch die irreversible Spaltung eines Substrats können dessen biologische Funktion sowie seine intra- oder extrazelluläre Lokalisation verändert und reguliert, Rezeptoren modifiziert, Zelloberflächen gestaltet, andere Enzyme, Zytokine, Hormone oder Wachstumsfaktoren aktiviert oder deaktiviert werden. Folglich können Proteasen bei unterschiedlichsten Zellfunktionen regulierend und modulierend eingreifen, wie z.B. bei der DNA-Replikation, dem Zellzyklus, der Proliferation, der Differenzierung und der Migration, Morphogenese und Gewebeausbildung, der Wundheilung, Angiogenese und Apoptose (frei zusammengefasst nach ⁸⁹).

Es gibt derzeit etwa 570 bekannte humane Proteasen, die in fünf katalytische Klassen eingeteilt werden: Metalloproteinasen, Serin-, Threonin-, Cystein- und Aspartat-Proteasen (**Abb. 7**). Extrazelluläre Proteasen sind vor allem in den Abbau und die Remodellierung der extrazellulären Matrix involviert, wie die oben bereits beschriebenen MMPs und ADAMs. Die weitaus größere Zahl intrazellulärer Proteasen haben für gewöhnlich die Funktion, unerwünschte oder beschädigte Proteine zu entfernen. Lysosomale Cysteine und Aspartyl-Cathepsine sorgen für den Abbau endozytotisch aufgenommener Proteine, während die Cystein-Proteasen der Caspase-Familie für die proteolytische Abbaukaskade der Apoptose zuständig sind ^{89,90}. Die bekannten Schnittstellen zwischen Proteasen und Wachstumsfaktoren, Zytokinen und Signalmolekülen werden dabei immer zahlreicher. So kann die Aktivierung von Signalmolekülen durch Proteasen zu einer Rückkopplungsschleife führen, über die Chemokine, Zytokine und Wachstumsfaktoren die Protease-Expression hochregulieren ⁹¹, wie im vorangegangenen Kapitel beschrieben. Weitere Beispiele sind Interleukin IL-8, welches in der Lage ist, die Matrix-Metalloproteinasen MMP2 und MMP9 in Blasenkrebs hochzuregulieren ⁹², sowie Interleukin IL-4, welches die Cathepsin-Aktivität in tumor-assoziierten Makrophagen induziert ⁹³.

Die strikte Regulation von Proteasen stellt einen äußerst kritischen Prozess dar, da die von ihnen katalysierte Hydrolyse irreversibel ist. Jegliche Fehlfunktion innerhalb dieses komplexen Enzymnetzwerkes kann zu gravierenden neurodegenerativen und kardiovaskulären Krankheiten sowie zur Entstehung von Krebs führen ^{89,91}. Das Forschungsgebiet *Degradomics* betrachtet die Substrate von Proteasen für Therapieansätze und versucht zu verstehen, welche Funktion die Protease in welchem zellulären Kontext erfüllt. Die für die Analyse notwendige Identifizierung potentieller Substrate und Inhibitoren wird durch genetische und bioinformatische Analysen von Gen- und Proteinsequenzen erleichtert.

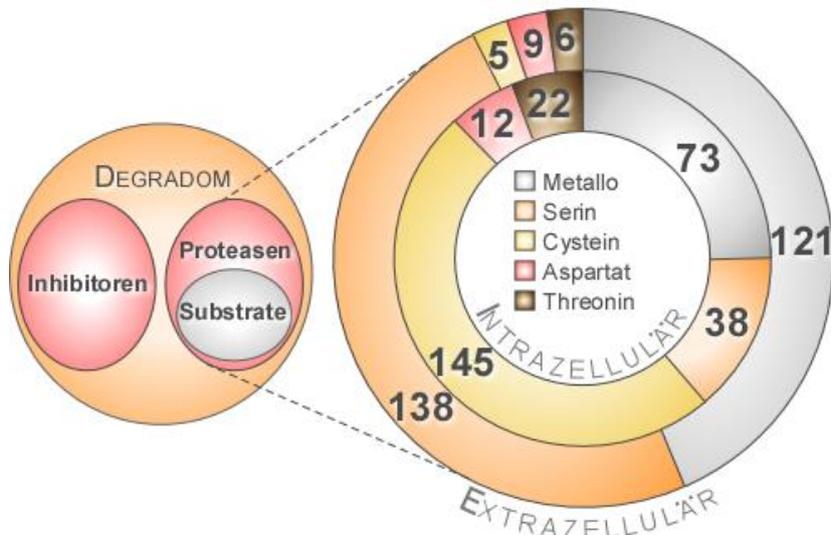


Abb. 7: Übersicht über das Degradom und die Klassifizierung humaner Proteasen.

Proteasen, ihre Substrate und Inhibitoren bilden das Degradom (links). In der detaillierten Auflistung der unterschiedlichen Arten von Proteasen (rechts) ist im inneren Ring die jeweilige Anzahl der intrazellulären Vertreter, im äußeren Ring die der extrazellulären Vertreter der jeweiligen Klasse aufgetragen (verändert nach ⁹⁴).

Protease-Datenbanken wie MEROPS (<http://merops.sanger.ac.uk/>) erlauben die taxonomische Einordnung von Proteasen sowie den Vergleich ihres Degradoms und beinhalten zudem Werkzeuge zur Zuordnung von Spaltsequenzen und Homologiedomänen ⁸⁹. Aus dem mechanistischen Verständnis der Spaltaktivität einer Protease kann man so auf seine Position innerhalb eines mittlerweile großen proteolytischen Netzwerkes schließen, das über die Regulation Proteasenspezifischer Inhibitoren Signalwege An- oder Abschalten kann, darunter die Modulation unterschiedlichster Prozesse wie Angiogenese, Invasion, ECM-Komposition, aber auch die Zusammensetzung der Mikroumgebung ^{91,95}.

2.4.1 Proteolytische Netzwerke

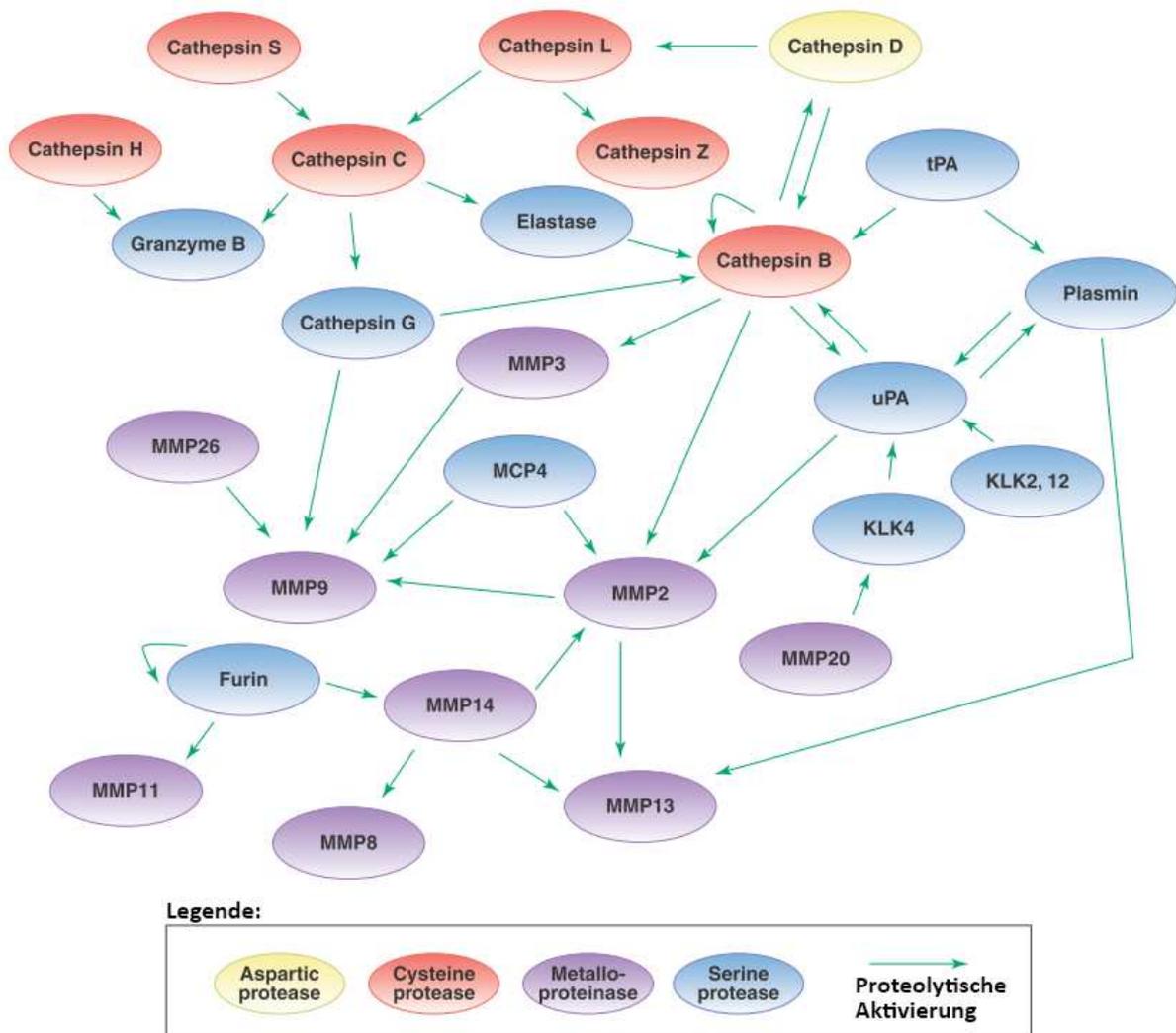
Die meisten Proteasen werden als inaktive oder wenig aktive Zymogene synthetisiert und müssen erst aktiviert werden, meist durch die Spaltung mittels anderer Proteasen oder Auto-Hydrolyse ⁹⁵. Die am besten untersuchte Aktivierungs-Kaskade dieser Art ist die Caspase-Kaskade, die zum apoptotischen Zelltod führt. Initiator-Caspasen wie Caspase 2, 8, 9 oder 10 können direkt oder

indirekt die Effektor-Caspasen 3, 6 und 7 durch Abspaltung ihrer Prodomänen aktivieren. Werden diese Effektor-Caspasen aktiviert, ist der Zelltod unausweichlich ⁹⁰. Die Wichtigkeit der Initiator-Caspasen für die Entwicklung und das Überleben des Organismus als Ganzes wird in Maus-Experimenten deutlich, in welchen der Knockout von Caspase 8 schon vor der Geburt letal ist ⁹⁶. Caspasen können durch mehrere endogene Protease-Inhibitoren wie XIAP gehemmt werden, welche wiederum von anderen Proteasen wie Cystein-Cathepsinen inhibiert werden können ⁹⁷.

Im Gegensatz zur linearen Caspase-Kaskade ist Cathepsin B in ein komplexes proteolytisches Netzwerk eingebunden (**Abb. 8**). Cathepsin B ist der Mikroumgebung vieler unterschiedlicher hochreguliert, und kann sich unter besonderen Bedingungen durch Auto-Proteolyse selbst aktivieren ⁹⁸. Es wird durch Cathepsin D aktiviert, welches in der Folge Cathepsin L und dieses wiederum das ECM-abbauende Enzym Heparanase aktiviert ⁹⁹. Auch Cathepsin G, *Urokinase-Type Plasminogen Activator* (uPA), *Tissue-Type Plasminogen Activator* (tPA) und Elastase können Cathepsin B aktivieren ¹⁰⁰. Dabei gibt es auch einige reziproke Interaktionen, denn uPA kann beispielsweise nicht nur Cathepsin B aktivieren, sondern auch selbst von Cathepsin B aktiviert werden ¹⁰¹.

In den meisten Fällen wird uPA bei der Aktivierung über einen Rückkopplungsmechanismus durch die Bindung von Pro-uPA über Plasmin zu uPAR transportiert ¹⁰⁰. Auch hier gibt es Alternativen, z.B. vermittelt über die Kallikreine (KLK) 2,4 und 12 ¹⁰². Die über CAFs stimulierte uPAR-uPA-Kaskade ist hinsichtlich des metastatischen Potentials von Krebszellen maßgeblich an der durch MMP2 vermittelten Aktivierung von MMP9 beteiligt ¹⁰³.

Die 24 verschiedenen Matrix-Metalloproteinasen gehören zu den in Tumoren am häufigsten überexprimierten Proteasen und sind essentiell für die Degradation und Remodellierung der extrazellulären Matrix ⁹¹. Das proteolytische MMP-Netzwerk wird hauptsächlich über das eben beschriebene Cathepsin B reguliert, kann aber auch durch Furin aktiviert werden, einer Serinprotease, welche die in einem Drittel der MMPs vorkommende Sequenz RXKR oder RRRK spaltet ^{91,104}. MMPs können auch auto-proteolytisch oder durch unspezifische Oxidation aktiviert werden ^{105,106}. Die in der MMP-Forschung zuerst genutzte neue Technologie der Kombination aus Isotopen-Markierung endständiger Aminosäuren (TRAIL) und der iTRAQ Quantifizierung erlaubte die Identifizierung neuer Substrate für MMP2 wie Cystatin C sowie der Cathepsine B und L in Zusammenhang mit VEGF ¹⁰⁷.



TRENDS in Cell Biology

Abb. 8: Proteolytisches Interaktionsnetzwerk um Cathepsin B, MMP2 und MMP9. Details siehe Text (verändert nach ⁹¹).

Die Interaktion von Proteasen mit nicht-proteolytischen Netzwerken wird häufig über Phosphorylierung durch Kinasen reguliert ¹⁰⁸, wie bei der *A Disintegrin And Metalloproteinase* (ADAM)-Proteasenfamilie ⁹¹. Der *Platelet-Derived Growth Factor Receptor* (PDGFR) wird über ADAM10 stimuliert und aktiviert ADAM17, welches die Externalisierung des *Vascular Endothelial Growth Factor Receptor* (VEGFR) 2 fördert, die Phosphorylierung von ADAM9 durch PKC und SRC beeinflusst und die ADAM12-induzierte PI3K-Aktivierung fördert ¹⁰⁸. Insgesamt wurden mehr als 100 Proteasen

identifiziert, die durch Phosphorylierung reguliert werden, und ebenso viele Kinasen werden durch Proteasen gespalten ¹⁰⁸. Die wichtigsten Beispiele für solche bidirektionalen Interaktionen sind das Wechselspiel zwischen Kinasen und Caspasen während der Apoptose sowie die Beeinflussung der Rezeptortyrosinkinase (RTK)-Aktivität durch Deubiquitinasen im Zusammenhang mit proteasomalem Abbau ¹⁰⁸. Das Ubiquitin-Proteasom-System (UPS) ist ein permanent aktives, fein abgestimmtes Abbausystem, welches die selektive Proteindegradation ermöglicht und somit durch Kontrolle der Proteinmengen die zelluläre Aktivität reguliert ¹⁰⁹. Teil des Systems sind die Ubiquitin Ligasen E1, E2 und E3, die mittels Poly-Ubiquitinierung Proteine zum Abbau durch das Proteasom markieren, sowie eine Reihe von deubiquitinierenden Enzymen (DUBs), die im Gleichgewicht miteinander das System feinregulieren und so nicht nur fehlgefaltete Proteine entfernen, sondern insbesondere wichtige zelluläre Funktionen steuern ¹⁰⁹. Das 26S Proteasom besteht aus einer zentralen 20S- und zwei 19S-Untereinheiten, die Fass-artig angeordnet sind und im Inneren eine Art Kammer bilden, in welcher die Proteine schließlich degradiert werden ¹⁰⁹. Die Aktivität des Proteasoms selbst kann wiederum über die dynamische Phosphorylierung der Untereinheiten reguliert werden ¹¹⁰. Das UPS ermöglicht so insbesondere auch unidirektionale zelluläre Prozesse wie sie u.a. für den Zellzyklus essentiell sind. Dabei werden gezielt Proteine phosphoryliert, was die Voraussetzung zu deren Ubiquitinierung ist und den weiterführenden Abbau durch das Proteasom zur Folge hat ¹⁰⁸. Durch die Abwesenheit der abgebauten Proteine wird die Initiation der nächsten Phase erst möglich, und über ein Wechselspiel aus Transkriptions-Stimulation und Proteinabbau in Feedback-Schleifen gesteuert (im Detail in ¹¹¹ beschrieben). Fast alle Serin- und Threonin-Kinasen wie MAP-Kinasen, AKT, Protein Kinase C (PKC) und TGF β werden über das UPS abgebaut und können zusätzlich durch den Abbau oder die Hochregulation ihrer Inhibitoren reguliert werden ¹¹². Die Deubiquitinase USP8 steuert zudem die Proteinmengen der Rezeptortyrosinkinasen (RTK), wie am Beispiel des *Epidermal Growth Factor Receptor* (EGFR), MET und ERBB3 gezeigt werden konnte ¹¹³. Die überlebensfördernden Eigenschaften des EGFR kann durch Caspasen über die Abspaltung der C-terminalen Region aufgehoben werden, und die Caspase-vermittelte Spaltung von RET, MET und ERBB2 generiert sogar pro-apoptotische Fragmente ^{114,115}. Caspasen wiederum können von Kinasen phosphoryliert und so in ihrer pro-apoptotischen Wirkung inhibiert werden, wie erstmals bereits 1998 anhand von AKT-vermittelter Phosphorylierung von Caspase 9 gezeigt werden konnte ^{108,116}. Die meisten RTKs können auch durch das Abspalten essentieller extrazellulärer Rezeptorbausteine (*Ectodomain-Shedding*) und nachfolgendem Verlust der Ligandenbindungskapazität reguliert werden, z.B. durch ADAMs und MMPs ^{114,117}. ERBB4 kann in seiner Transmembrandomäne durch γ -Sekretase gespalten werden,

womit ein intrazelluläres Fragment generiert wird, das im Zellkern akkumuliert und dort als Chaperon oder Kofaktor für verschiedene Transkriptionsfaktoren fungieren kann ¹¹⁸. Verschiedene ADAMs sowie MMP2 und MMP9 sind in der Lage, die Aktivität von RTKs wie FGFR1, EPHB2, AXL, VEGFR2 und dem EGFR durch Spaltung zu modifizieren ¹⁰⁸.

2.4.2 Threonin Aspartase 1

Innerhalb der Protease-Familie nimmt die Threonin Aspartase 1 (Taspase1) eine pathophysiologische Sonderstellung ein. Sie wird auf Grund ihrer Asparaginase 2-Homologie-Domäne den Typ2-Asparaginasen zugeordnet, hat jedoch mit den üblichen Vertretern der Familie sonst nicht viel gemein. Taspase1 wurde 2003 erstmals als die MLL-spaltende Protease identifiziert und in diesem Zusammenhang bereits in Kapitel 2.2 beschrieben ¹¹⁹. Untersuchungen an *Knockout*-Mäusen zeigten bei homozygot defizienten *Taspase1*^{-/-}-Tieren eine geringere Körpermasse, erhöhte postnatale Sterblichkeit und schwerwiegende Skelett-Anomalien, wie eine unvollständige Segmentierung und die Ausbildung zusätzlicher Rippenpaare ¹²⁰. Die beschriebenen Anomalien in der Ausbildung der korrekten Segmentidentität von *Taspase1*^{-/-}-*Knockout*-Mäusen sind gravierender als die von *MLL*^{-/-}-*Knockout*-Mäusen, was den Schluss zulässt, dass Taspase1 weitere Substrate prozessiert und so weitläufiger in die Regulation des Zellzyklus involviert ist ¹²⁰.

Taspase1 wird als Pro-Enzym von 420 Aminosäuren und einem Molekulargewicht von 50 kDa synthetisiert und auto-katalytisch zwischen Aspartat 233 und Threonin 234 in eine α -Untereinheit zu 28 kDa und eine β -Untereinheit zu 22 kDa prozessiert ⁴⁵. Durch diese *cis*-Aktivität wird das aktive Zentrum am N-terminalen Threonin 234 der entstehenden β -Untereinheit konformativ freigelegt und damit die *trans*-Aktivität, also die Spaltung anderer Substrate erst ermöglicht ^{45,121}. Aus kristallographischen Untersuchungen der beiden Untereinheiten wurden die evolutionär konservierten Aminosäuren G51, Y52, K57, Y61, F98, N100, N106, S252, G254, L255 und G292 identifiziert, die an der spezifischen Substratbindung beteiligt sein könnten - mit einem Chlorid-Ion im aktiven Zentrum zwischen T234, G49 und N100 ¹²¹. Physiologische Chlorid-Konzentrationen (IC_{50} = 26 mM) wirken hierbei inhibitorisch, was eine wenigstens partielle Hemmung der Taspase1 wahrscheinlich macht ¹²¹. Wie in Kapitel 2.2 beschrieben, schneidet Taspase1 MLL an den Spaltstellen CS1 und CS2 (**Abb. 4**). *In vitro*-Untersuchungen der Bindekapazität von Taspase1 mit ihrem Substrat MLL ergaben eine klassische Michaelis-Menten Kinetik mit $K_m = 12 \mu M$ ⁴⁵.

Das Taspase1-Degradom wurde mittels der erweiterten *Cleavage Site*-Konsensussequenz (Q³[F,I,L,V]²D¹↓G^{1'}X^{2'}D^{3'}D^{4'}) *in silico* auf 27 potentielle Zielproteine mit unterschiedlicher zellulärer Lokalisation eingegrenzt ¹²². Verifiziert werden konnten bisher MLL1, MLL4, Myosin-1F (MYO1F), TFIIA, der *Nuclear RNA Export Factor* (NXF)2, NXF5 und USF2 ⁹⁴. Da sich unter den vorhergesagten Substraten auffällig viele große Proteine mit jeweils mehr als 800 Aminosäuren finden, liegt die Vermutung nahe, dass diese häufig Bausteine größerer Komplexe sind und die proteolytische Spaltung einzelner Komponenten durch Taspase1 der Feinregulation ihrer biologischen Funktion dient ¹²².

Der in Kapitel 2.1 beschriebene Transkriptionsfaktor TFIIA wird als αβ-Zymogen von Taspase1 spezifisch innerhalb der Aminosäuren 272-275 im QVDG-Motif geschnitten ^{19,123}. RNAi-*knockdown*-Experimente legen nahe, dass die Spaltung durch Taspase1 zwar nicht für die transkriptionelle Aktivität von TFIIA notwendig ist, jedoch den proteasomalen Abbau des Proteins beeinflusst ^{19,123} und zur Gewebe-spezifischen Genexpression beiträgt ^{124,125}. Letztere scheint eine besondere Rolle in der Entwicklung zu spielen, wie im Rahmen der Embryonalentwicklung in Mäusen an der Morphogenese des Schädels gezeigt werden konnte ¹²⁵.

Die 27 potentiellen Zielproteine von Taspase1 lokalisieren in unterschiedlichen zellulären Kompartimenten. Grundsätzlich stellt die Kompartimentierung von Zellen eine wesentliche evolutionäre Errungenschaft der eukaryotischen Zellen dar, da diese die unabhängige Regulation komplexer biochemische Reaktionen in verschiedenen Reaktionsumgebungen ermöglicht ¹²⁶. Insbesondere Proteasen können so die Zugangsmöglichkeiten zu ihren Substraten durch spezielle, aktive Transportmechanismen sicher regulieren, wie beispielsweise an MMP-12 gezeigt werden konnte ¹²⁷. Für den Transport in und aus dem Kern ist vor allem die nukleäre Doppellipid-Membran entscheidend. Sie wird durch nukleäre Porenkomplexe durchbrochen und ermöglicht so zum einen die Diffusion biologisch kleiner Moleküle (< ca. 20kDa), erfordert aber den aktiven Transport größerer Substrate durch die Kernporen. Dieser wird über spezifische Signale und Transportrezeptoren wie Importin-α oder Exportin1/Crm1 reguliert ¹²⁸. Zu diesen spezifischen Signalen auf Seite der zu transportierenden Proteine gehören die nukleären Lokalisationssequenzen (NLS), sowie die nukleären Exportsignale (NES), die eine Interaktion mit den Transportrezeptoren ermöglichen ¹²⁹. Taspase1 enthält ein evolutionär konserviertes NLS und wird mittels Bindung an Importin-α aktiv in den Kern transportiert (**Abb. 9**)¹³⁰.

Im Kern akkumuliert Taspase1 durch Interaktion mit Nucleophosmin 1 (NPM1) in den Nukleoli, auch Kernkörperchen genannt, besonders dicht gepackten Regionen des Kerns ¹³⁰. NPM1 ist in verschiedene zelluläre Prozesse wie die Ribosomen-Biogenese, die Duplikation der Zentrosomen und den Erhalt der genomischen Integrität involviert und kann zwischen dem Kern und dem Zytoplasma hin- und herwandern ^{131,132}. Taspase1 ist trotz fehlendem nukleären Exportsignal (NES) über die Interaktion mit NPM1/Crm1 in der Lage, auch aus dem Kern ins Zytoplasma zu gelangen und so auch dort verschiedene Zielproteine zu spalten, wie das in Kapitel 2.2 beschriebene pathobiologisch relevante AF4·MLL-Fusionsprotein ¹³⁰.

Die höchste Taspase1-Spaltaktivität wurde in Fraktionierungsexperimenten in der Fraktion detektiert, welche die leichten Membranen wie das endoplasmatische Retikulum enthält, das direkt mit der nukleären Membran verbunden ist ⁴⁵. Möglicherweise spaltet Taspase1 also ihre Substrate also auf dem Weg in den Kern, ermöglicht den Proteinen so ein Ausnutzen der starken NLS von Taspase1 und damit einen Transport in den Zellkern ⁹⁴. Aus dieser Hypothese wurde ein Modell entwickelt, das einen Feinregulations-Mechanismus der proteolytischen Aktivität über Importin- α /NPM1 ermöglicht ^{94,130}. In Krebszellen könnte diese Regulation ausgenutzt werden, um die Taspase1-Spaltaktivität in Richtung eines bestimmten Kompartiments sowie zu unterschiedlichen Substraten hin zu verschieben und auf diese Weise zur Tumorgenese beizutragen ^{94,126}.

Taspase1 wurde von Chen *et al.* 2010 als *Non-Oncogene-Addiction-Protease* beschrieben ¹³³. Das Prinzip der *Oncogene Addiction*, also der Abhängigkeit eines Tumors von einem bestimmten Onkogen, wurde von Bernard Weinstein entwickelt. In der Hypothese wird davon ausgegangen, dass der Tumor ohne ein bestimmtes Onkogen nicht weiterexistieren kann und somit den besten Ansatzpunkt für Therapien darstellt ¹³⁴. Weiterführend wird davon ausgegangen, dass Tumore nicht nur von Onkogenen, sondern auch von Nicht-Onkogenen abhängen können und somit neue Interventionsmöglichkeiten zur Therapie-Sensibilisierung bei einer malignen Erkrankung ermöglichen ¹³⁵. Dadurch, dass Taspase1 in adulten, gesunden Zellen nur sehr gering exprimiert wird, in verschiedenen malignen Erkrankungen jedoch hoch exprimiert ist, bietet sie sich als erstrebenswertes Therapieziel an ^{133,136}. Taspase1 war bei der tumorigenen Transformation von embryonalen Maus-Fibroblasten durch unterschiedliche Onkogene nicht nur für die Transformation selbst, sondern auch für den Erhalt des transformierten Tumorpotentials notwendig ^{120,133}. Taspase1-*knockdown* in Taspase1-überexprimierenden Tumorzellen führte zu einer Proliferationsinhibition, die abhängig vom Expressionsniveau variierte ¹³³.

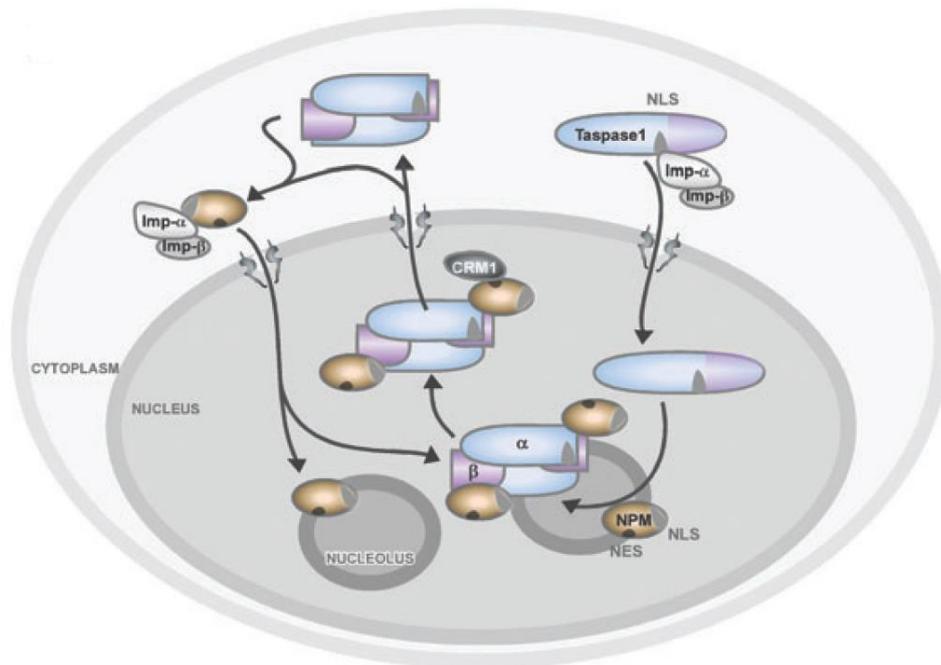


Abb. 9: Schematische Darstellung des Taspase1-Transportmechanismus über Importin- α und NPM1/Crm1.

Das Taspase1-Proenzym wird über das zweigeteilte NLS mittels der Importin- α (Imp- α)/Importin- β (Imp- β)-Achse (hell-/mittelgrau) in den Kern transportiert. Dort wird es autoproteolytisch in seine Untereinheiten α (hellblau) und β (violett) gespalten und bindet als aktives Enzym an NPM1 (braun). Basierend auf dieser Interaktion akkumuliert Taspase1 an den Nukleoli (dunkelgrau), ist aber auch in der Lage, den Crm1-vermittelten Kernexport durch NPM1 für sich zu nutzen (aus ¹³⁰).

In Maus-Fibroblasten konnte bei Taspase1-*knockout*-Experimenten gezeigt werden, dass die Expression der Cycline E, A und B sank, während der CDK Inhibitor p16^{INK4A} und die Zellzyklus-Regulatoren p21^{CIP1}, p27^{KI1} und p14^{ARF} vermehrt vorhanden waren, was auf eine Blockade des G₁/S-Checkpoints schließen ließ ¹²⁰. Die Erhöhung der Zellzyklus-Regulatoren p16^{INK4A}, p21^{CIP1} und p14^{ARF} war direkt auf die Spaltung von TFIIA durch Taspase1 und damit auf eine Spezifitätsänderung von TFIID und der Rekrutierung von Pol II zurückzuführen ¹²⁵. Ein Fehlen von Taspase1 resultierte zudem in einer erhöhten Resistenz gegenüber der malignen Transformation durch transiente Überexpression der Onkogene Ras und Myc, oder durch einen *Knockout* von p53 ^{120,133}. Zusätzlich wurden die Zellen sensitiver gegenüber apoptotischen Signale, was auf reduzierte Mengen des anti-apoptischen Proteins Mcl-1, vermittelt durch einen vermehrten USP9X-vermittelten proteasomalen Abbau zurückzuführen war ¹³³.

Die klinische Relevanz der Taspase1 wurde in Kapitel 2.2 bereits für Leukämien beschrieben; Taspase1 spielt jedoch auch in verschiedenen soliden Tumoren eine große Rolle und konnte in Kopf-Hals Tumoren¹³⁰, Glioblastom^{120,137}, Brustkrebs^{137,138} und Darmkrebs¹³⁰ als relevant nachgewiesen werden. Kopf-Hals Plattenepithelkarzinome (HNSCC) entwickeln sich dabei aus verschiedenen anatomischen Sublokalisationen, darunter Mundhöhle, Hypopharynx, Oropharynx, Larynx und Nasopharynx mit über 500.000 neuen Fällen pro Jahr weltweit mit steigender Tendenz^{2,139}. Die Nutzung von Tabak, insbesondere zusammen mit Alkoholmissbrauch, sowie eine Infektion mit den onkogenen Viren HPV-16, HPV-18, HPV-31 und HPV-33 stellen die Hauptrisikofaktoren dar¹³⁹⁻¹⁴¹. Bei früher Erkennung sind HNSCC gut mit einer operativen Entfernung und/oder Radiotherapie zu behandeln. Für fortgeschrittenere Stadien ist die Behandlung oft multimodal mit einer Operation – wenn möglich –, die für gewöhnlich mit einer darauffolgenden (adjuvanten) Chemoradiatio kombiniert wird, oder es wird direkt mit einer primären Chemoradiatio begonnen¹³⁹. Für wiederkehrende oder metastasierende Neoplasmen ist die Standardtherapie ein chemotherapeutischer Ansatz mit der Option einer alternativen Immuntherapie bei Unverträglichkeiten gegenüber regulären Chemotherapeutika mit Cetuximab, einem Antikörper zur Hemmung des EGFR¹³⁹. Jegliche Behandlungsentscheidung wird auf Basis der primären Lokalisation, der TNM-Klassifikation und histopathologischen Merkmalen unter Berücksichtigung möglicher Limitationen durch den Performance-Status des Patienten (bspw. Karnofsky-Index) in einem multidisziplinären Tumorboard getroffen^{139,142}.

2.5 Personalisierte Therapie

Generell verändern sich Krebstherapien im klinischen Alltag derzeit graduell von systemisch wirkenden zytotoxischen Medikamenten zu individuellen Ansätzen mit zielgerichteten Therapien, die auf jeden einzelnen Tumor zugeschnitten sind. Essentiell für diese personalisierten Therapien sind diagnostische Verfahren, die in der Lage sind, den klinischen Verlauf bestimmter Therapien vorherzusagen und das bestehende molekularbiologische Wissen aus der Tumorgenese zu nutzen¹⁴³. Die Anzahl von Mutationen erhöht sich in Patienten mit Lymphknotenmetastasen im Vergleich zu Patienten ohne Lymphknotenmetastasen nicht¹⁴⁴. Es sind also wahrscheinlich andere Änderungen als Mutationen für die Progression eines Primärtumors zu einem invasiven und metastatischem Verhalten hin verantwortlich¹⁴⁵. Expressionsprofile sind eine wertvolle Methode, um zwischen

Krebs-Subtypen zu unterscheiden und auf längere Sicht prognostische Vorhersagen in Kombination mit klinischen Daten tätigen zu können ¹⁴⁶⁻¹⁵⁰. Eine Studie in HNSCC erlaubte eine Kategorisierung in Subtypen ¹⁵¹, während andere Studie eine Gensignatur mit Rezidiven in HNSCC in Verbindung bringen ¹⁵² oder allgemein Gene mit diagnostischem oder prognostischem Potential identifizieren konnten ¹⁵³⁻¹⁵⁸. Microarray Analysen, in denen normales Epithelgewebe der Mundhöhle mit HNSCC verglichen wurde, konnten Hinweise auf die Mikroumgebungsveränderungen und die Immunantwort bei HNSCC eruieren ¹⁵⁹. Dennoch war für lange Zeit der einzige im klinischen Alltag verwendete prognostische Marker für HNSCC der HPV Status des Tumors, während prädiktive Biomarker, die in der Lage wären, Behandlungsentscheidungen zu beeinflussen, komplett fehlten ^{139,160}. Obwohl HPV-positive Tumore stark in ihren Expressionsprofilen von HPV-negativen abweichen, waren die Behandlungsoptionen die gleichen und hingen nicht vom HPV-Status des Tumors ab ¹³⁹. Erst seit der 2017 veröffentlichten 8. Ausgabe der TNM-Klassifikation wird in der klinischen Routine bei Oropharynxkarzinomen immunhistochemisch der p16^{INK4a}-Status ermittelt und in die prognostische Bewertung einbezogen ^{161,162}.

Mehrere neue, zielgerichtete Therapien gegen Wachstumsrezeptoren oder in den jeweils korrespondierenden Signalwegen nachgeschaltete Faktoren werden derzeit in vielversprechenden klinischen Studien getestet ¹⁶³. Mediziner werden so demnächst die effektivste Therapie mit den geringsten Nebenwirkungen und Resistenzmechanismen für den jeweiligen Patienten aus einer Vielzahl verfügbarer Optionen ohne begleitende Diagnostika wählen müssen. Totale Überlebensstatistiken (*overall survival*, OS) und qualitätsadjustierte Lebensjahre (*quality adjusted life years*, QALY) werden für die Evaluation hochspezialisierter Krebstherapien in Zukunft möglicherweise nicht mehr ausreichend sein ^{164,165}. Genomsequenzierungen konnten bereits wertvolle Hinweise auf potentielle Biomarker geben und dienen als Grundlage für weitere Untersuchungen ^{139,166,167}. Die Expressionsanalyse möglicher Resistenzmechanismen gegen z.B. Cetuximab-behandelte Tumore zeigte mehrere vielversprechende Ansätze für weitere Biomarkerstudien auf ¹⁶⁸. Die Kosten derartiger 'Omics-Technologien sind in den letzten Jahren drastisch gesunken, während die Datenqualität gleichzeitig gesteigert wurde. Dennoch sind sie immer noch zu teuer, um sie in der klinischen Routine anzuwenden. Hochdurchsatzsequenzierungstechniken werden vor allem in Kohortenstudien verwendet, was eine geringe statistische Relevanz und das Risiko falsch-positiver Resultate zur Folge hat und somit die technische Reproduzierbarkeit einschränkt ¹⁶⁹. Evidenzbasierte Medizin, die als Grundlage für die Entscheidung der Zulassungsbehörden dient, benötigt große

Mengen an verlässlichen empirischen Daten, die aus robusten, reproduzierbaren Assays in regelgerecht durchgeführten, klinischen Studien gewonnen wurden¹⁷⁰. Verwendete Methoden müssen also stark vereinfacht und standardisiert werden, um Fehlerquellen zu minimieren und signifikante Daten für eine effektive Entwicklung von Biomarkern parallel zu klinischen Studien zu generieren¹⁶⁵. Die Microarray-Technologie wird derzeit auf nanotechnologischen und mikrofluidischen Chips miniaturisiert und ermöglicht so einen automatisierten und kosteneffektiven Ansatz für die klinische Diagnostik¹⁷¹, was u.a. bereits für die Pathogenidentifikation getestet wird¹⁷².

2.6 Zielsetzung

Krebs ist die zweithäufigste Todesursache weltweit und wird vermutlich sogar Herz-Kreislauf-Erkrankungen als Haupttodesursache in näherer Zukunft überholen². Auch wenn in der Kanzerogenese vielfältige Mechanismen involviert sind³, konnten mehrere Studien der letzten Jahre zeigen, dass insbesondere Proteasen bei Tumorwachstum und -progression eine wichtige Rolle innehaben^{173,174}.

In dieser Arbeit sollen deshalb am Beispiel von Kopf-Hals Tumorpatienten mittels Expressionsanalysen die Besonderheiten und Unterschiede von Lymphknotenmetastasen zu ihren Primärtumoren, sowie der Primärtumore zu ihrem Normalgewebe herausgearbeitet werden. Dabei sollte der Fokus auf der Auswertung der proteolytischen Netzwerke und deren Auswirkungen auf maligne Mechanismen und potentielle Biomarker liegen. Die vielfältigen Interaktionen zwischen Proteasen und Kinasen könnten auch neue Hinweise für Kombinationstherapien und diagnostische Möglichkeiten liefern.

Da Taspase1 eine Protease mit onkologischer Relevanz für sowohl spezielle Formen der Leukämie sowie auch für solide Tumoren ist, sollten ihre Wirkmechanismen im evolutionären und epigenetischen Zusammenhang evaluiert werden. Um weiterführend die pathobiologische Bedeutung von Taspase1 in der Klinik untersuchen zu können, sollte zudem eine Basis geschaffen werden, um gezielt epidemiologische und histopathologische Daten zu den Laboruntersuchungen an Patientenmaterial hinzuziehen zu können. So könnten in Zukunft prognostische Aussagen anhand

EINLEITUNG

bestimmter Biomarker getroffen und so wiederum Hinweise auf potentielle neue Zielstrukturen für Medikamente und diagnostische Methoden identifiziert werden.

3 ERGEBNISSE

3.1 Expressionsprofiling zur Identifikation neuer krankheitsrelevanter Signalwege und potenzieller Biomarker

Die Haupttodesursache bei Krebs ist die Verdrängung gesunden Gewebes und damit der Verlust von Organaktivität durch Metastasen ¹⁷⁵. Die Mechanismen, die zu invasivem und disseminierendem Verhalten und damit zur Malignität des Tumors beitragen, sind also von größtem Interesse für die Krebsforschung. Die Akkumulation von Mutationen in Tumorpatienten mit Lymphknotenmetastasen ist im Vergleich zu Patienten ohne Lymphknotenbefall nicht erhöht ¹⁴⁴. Es ist also anzunehmen, dass darüber hinaus andere Zellveränderungen für die Progression von Primärtumor zu Metastasen mitverantwortlich sind ¹⁴⁵. Expressionsprofiling wurde bereits in verschiedenen Krebsarten sowohl zur Unterscheidung von Krebssubtypen als auch zur Vorhersage des Krankheitsverlaufs verwendet ¹⁴⁶⁻¹⁵⁰. Eine frühere Studie konnte Kopf-Hals Plattenepithelkarzinome (HNSCC) in verschiedene Subtypen unterteilen ¹⁵¹, während andere Studien Gen-Signaturen mit der Rückfallwahrscheinlichkeit in Verbindung bringen ¹⁵² oder Gene mit diagnostischem oder prognostischem Potenzial vorschlagen konnten ¹⁵³⁻¹⁵⁸. Frühere Microarray-Analysen von normaler Schleimhaut im Vergleich zu Mundhöhlenkarzinomen konnten den Einfluss der Mikroumgebung und Immunantwort im Zusammenhang mit dem Überleben darstellen ¹⁵⁹.

In dieser Arbeit wurde das Expressionsmuster von Gewebeproben aus HNSCC-Primärtumoren (PT) erstmals mit den korrespondierenden Lymphknotenmetastasen (M) verglichen. Eine Probe aus der Mundschleimhaut des jeweiligen Patienten wurde als Normalgewebe-Kontrolle (N) mitgeführt. Insgesamt wurden 45 Proben aus 15 Patienten in die Studie eingeschlossen, die histopathologisch als HNSCC diagnostiziert wurden, bei denen wenigstens eine Lymphknotenmetastase resiziert werden konnte, aber klinisch noch keine Fernmetastasen evident waren (**Tab. 1**). Die Gewebeproben wurden aus chirurgischen Resektaten von Kopf-Hals-Tumorpatienten entnommen, die sich in den Universitätskliniken in Frankfurt a.M. und Mainz einer Krebsbehandlung unterzogen und ihr schriftliches Einverständnis gaben. Die Studie wurde gemäß den ethischen Standards der Deklaration

ERGEBNISSE

von Helsinki durchgeführt und von der zuständigen Ethikkommission geprüft und zugelassen (#83756604).

Tab. 1: Klinische Charakteristika der HNSCC Patienten.

#	N/PT/M	Primärtumor	Alter	Geschlecht	HPV	T/A**	pT	pN	Grading	cM
1	N1/PT1/M1	Hypopharynx	47	W	-		2	2	2	0
2	N2/PT2/M2	Oropharynx	58	W	-		3	2	2	0
3	N3/PT3/M3	Hypopharynx	46	M	-		4	1	2	0
4	N4/PT4/M4	Larynx	68	M	-		4	2	2	0
5	N5/PT5/M5	Oropharynx	56	M	-		3	2	2	0
6	N6/PT6/M6	Hypopharynx	48	M	-		3	1	2	0
7	N7/PT7/M7	Larynx	56	M	-		4	3	3	0
8	N8/PT8/M8	Hypopharynx	58	M	-		3	2a	2	0
9	N9/PT9/M9	Hypopharynx	72	M	-		3	2	2	0
10	N10/PT10/M10	Oropharynx	49	M	-		3	1	2	0
11	N11/PT11/M11	Hypopharynx	57	M	-		1	2	3	0
12	N12/PT12/M12	Oropharynx	53	M	-		2	2b	2	0
13	N13/PT13/M13	Hypopharynx	57	M	-		2	3	2	0
14	N14/PT14/M14	Hypopharynx	47	M	-		2	2	2	0
15	N15/PT15/M15	Hypopharynx	56	M	-		2	3	2	0
16*	PT16	Larynx	49	M	-		3	1	2	0
17*	N17/PT17	Hypopharynx	75	W	-		1	2	1	0
18*	PT18	Oropharynx	53	M	-		2	2	2	0
19*	PT19/M19	Hypopharynx	63	M	-		2	3	2	0
20*	PT20	Oropharynx	57	M	-		1	2	2	0

*auf Grund unzureichender RNA-Qualität ausgeschlossen.

**T/A: Tabak- / Alkoholkonsum

Bei der Mehrheit der Tumore war der Primärtumor im Hypopharynx lokalisiert (60%), bei 13,3% im Larynx und bei 26,6% im Oropharynx. Das mediane Alter aller in die Studie eingeschlossenen Patienten war 56 Jahre, bei 13% Frauen und 86% Männern. Alle Patienten hatten positive Lymphknoten, wie vom Studiendesign her vorgegeben, wobei bei den meisten (80%) ein Lymphknotenstatus von N1 oder N2 und bei nur drei Patienten (20%) N3 diagnostiziert wurde. Der Differenzierungsgrad der Tumore wurde bei sechs Patienten als Grade 2 (40%), bei neun Patienten als Grade 3 (60%) eingeteilt.

Die RNA der Gewebeproben wurde mit Microarray Gene Chips (Affymetrix HG-U133A) analysiert. Um die Microarray-Genexpressionsdaten zu verifizieren, wurden einzelne repräsentative Werte stichprobenartig auf RNA-Ebene mittels quantitativer Real-Time-PCR (qPCR) und Reverser Transkriptions-PCR (RT-PCR), auf Protein-Ebene mittels Immunhistochemie (IHC) und Western Blots überprüft. Die RT-PCR wurde an den im Primärtumor gegenüber Normalgewebe (PTvs.N) differenziell exprimierten Genen CLCA4, FN1, POSTN, KRT24 und PRR4, für die Lymphknotenmetastasen gegenüber dem korrespondierenden Primärtumor (Mvs.PT) die ebenfalls differenziell exprimierten Gene ARHGAP25, FCMR, RASGRP2, LYPD3, SERPINB4 und TP73L durchgeführt. Als Kontrolle wurde jeweils GAPDH mitgeführt (**Abb. 10**).

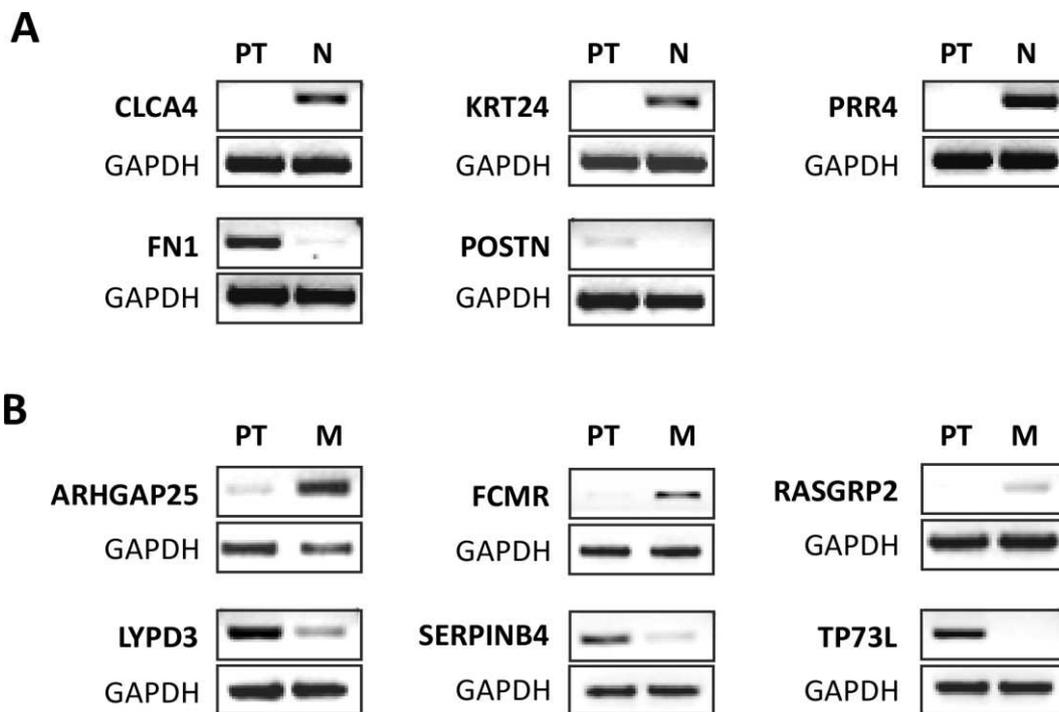


Abb. 10: Validierung der Microarray-Genexpressionsdaten mittels RT-PCR.

A. Die RT-PCR Amplifikationsprodukte der niedriger exprimierten (obere Reihe) oder höher exprimierten (untere Reihe) Gene in PTvs.N. **B.** Die RT-PCR Amplifikationsprodukte der niedriger exprimierten (obere Reihe) oder höher exprimierten (untere Reihe) Gene in Mvs.PT.

Für die Gene FN1 und POSTN konnte im Vergleich PTvs.N eine Hochregulierung, während für die Gene CLCA4, KRT24 und PRR4 eine Herabregulation bestätigt werden konnte. Im Vergleich Mvs.PT konnte auch für die Gene ARHGAP25, FCMR und RASGRP2 die Hochregulation, für LYPD3, SERPINB4 und TP73L die Herabregulierung verifiziert werden. Nachdem diese erste Validierung erfolgreich war,

wurden die Gene FN1 (**Abb. 11**), KRT24 (**Abb. 12**), OLR1 (**Abb. 13**) und PRR4 (**Abb. 14**) für PTvs.N sowie für Mvs.PT ARHGAP25 (**Abb. 15**) und SERPINB4 (**Abb. 16**) zusätzlich mittels qPCR überprüft und zeigten sich in guter Übereinstimmung mit den erhaltenen Microarray-Daten. Alle validierten Microarray-Daten sind in **Tab. 34** für PTvs.N und in **Tab. 35** für Mvs.PT aufgelistet.

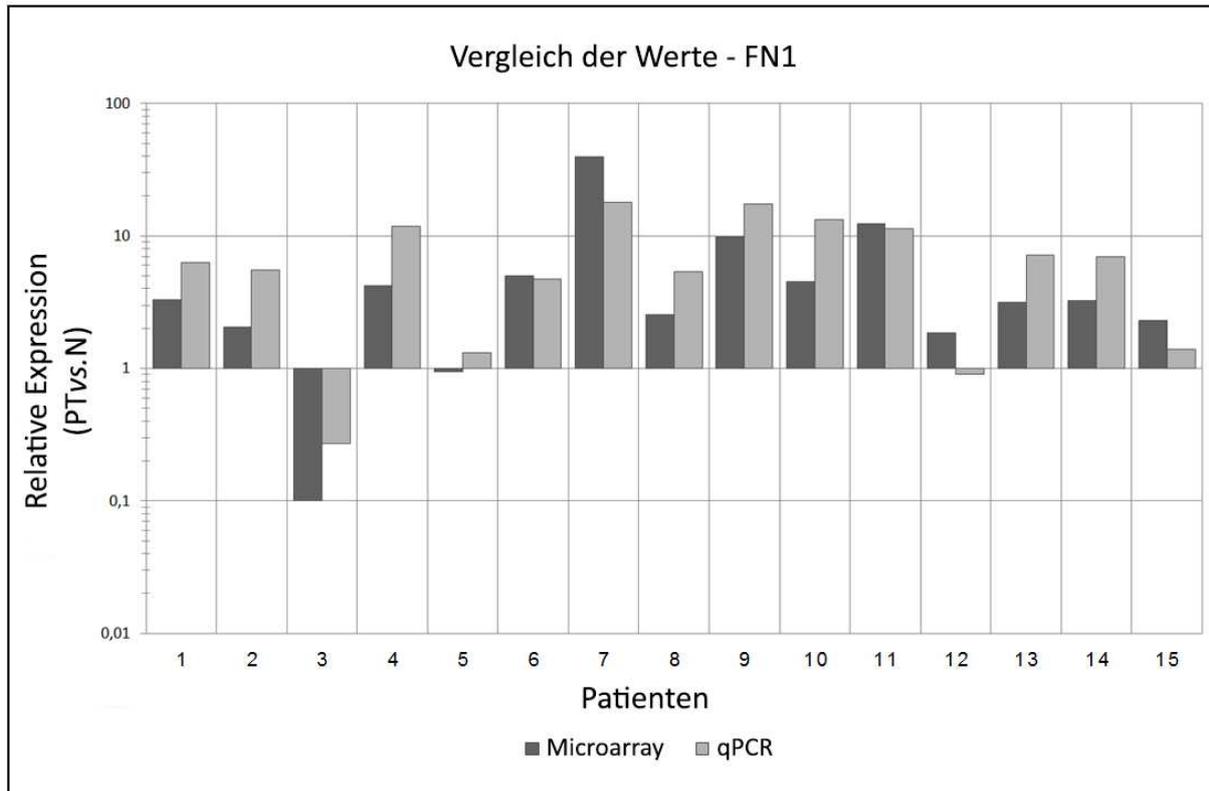


Abb. 11: Vergleich der relativen Expression des Gens FN1 in PTvs.N.

Die mittels Microarray und qPCR erhaltenen Expressionswerte wurden als Verhältnis von Primärtumor (PT) gegenüber Normalgewebe (N) für die einzelnen Patienten aufgetragen. Die Mehrzahl der Patienten zeigte eine Hochregulation von FN1 im Tumorgewebe.

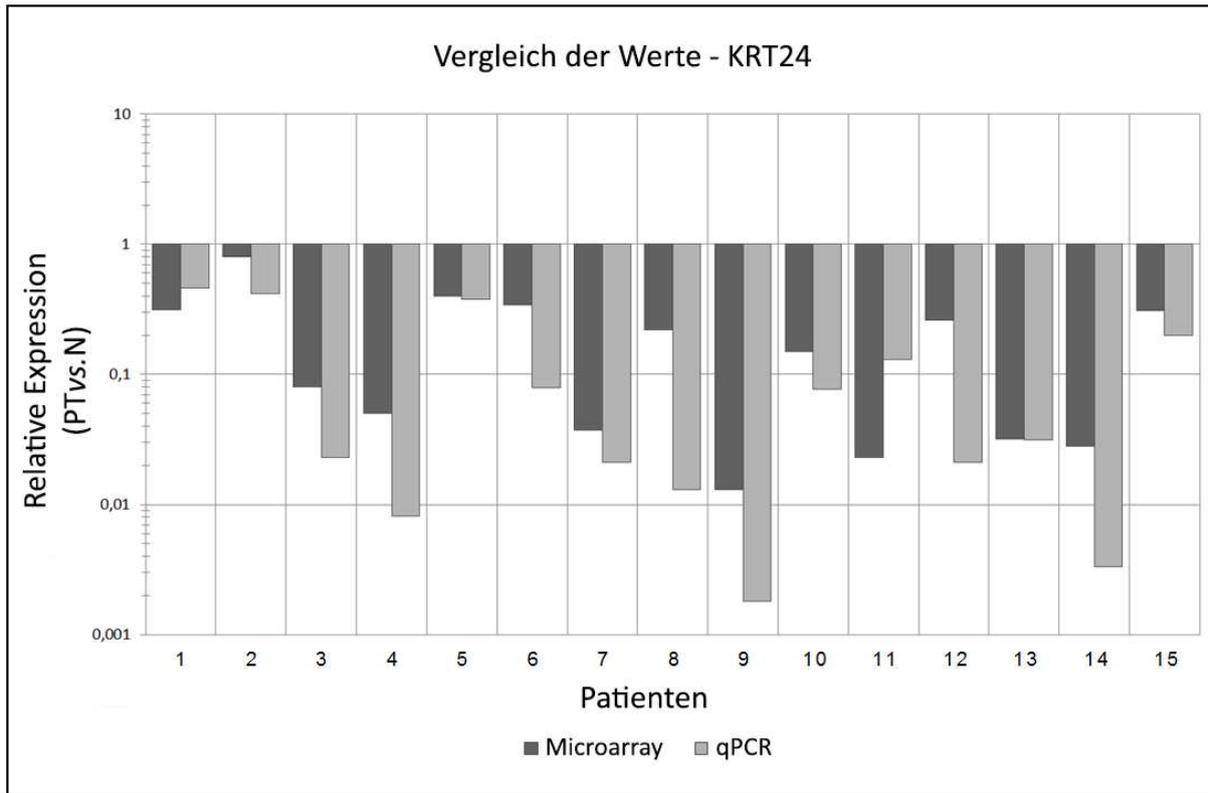


Abb. 12: Vergleich der relativen Expression für das Gen KRT24 in PTvs.N.

Die mittels Microarray und qPCR erhaltenen Expressionswerte wurden als Verhältnis von Primärtumor (PT) gegenüber Normalgewebe (N) für die einzelnen Patienten aufgetragen. Eine Herabregulation von KRT24 konnte für alle Patienten bestätigt werden.

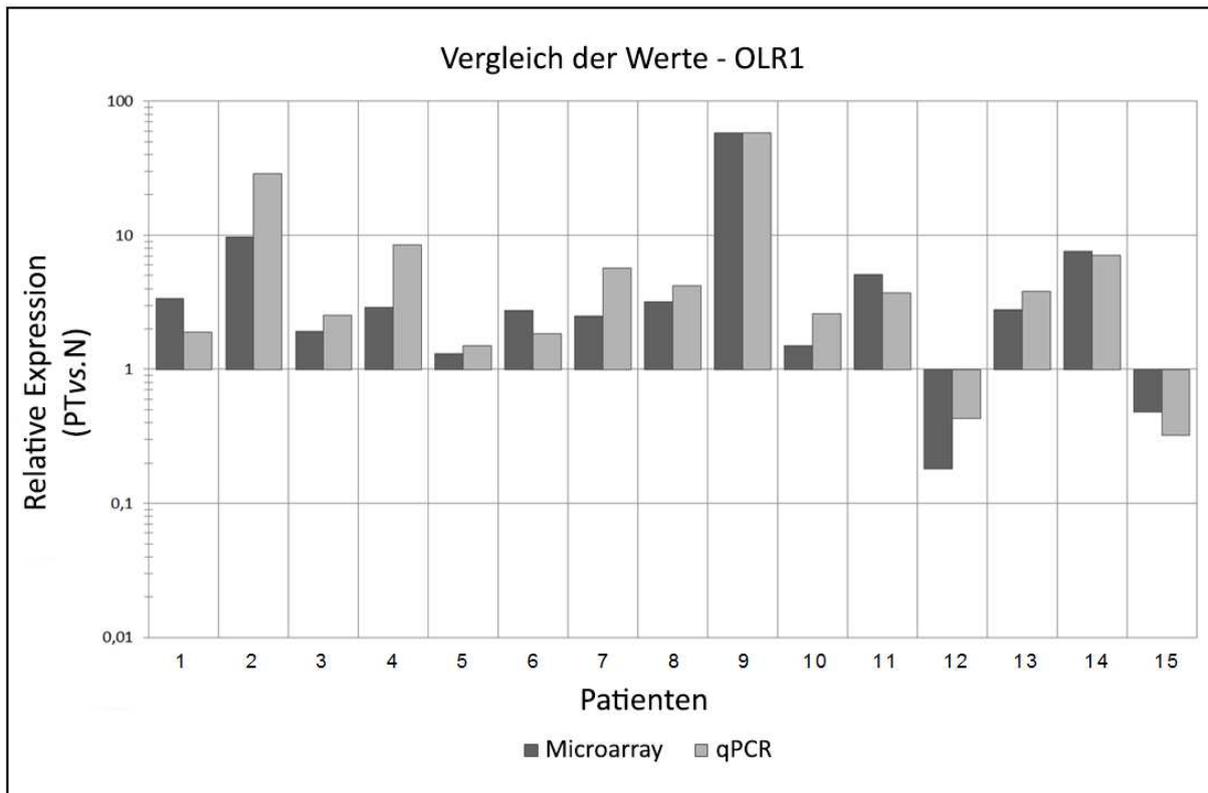


Abb. 13: Vergleich der relativen Expression für das Gen OLR1 in PTvs.N.

Die mittels Microarray und qPCR erhaltenen Expressionswerte wurden als Verhältnis von Primärtumor (PT) gegenüber Normalgewebe (N) für die einzelnen Patienten aufgetragen. In der Mehrzahl der Patienten wurde OLR1 als hochreguliert nachgewiesen.

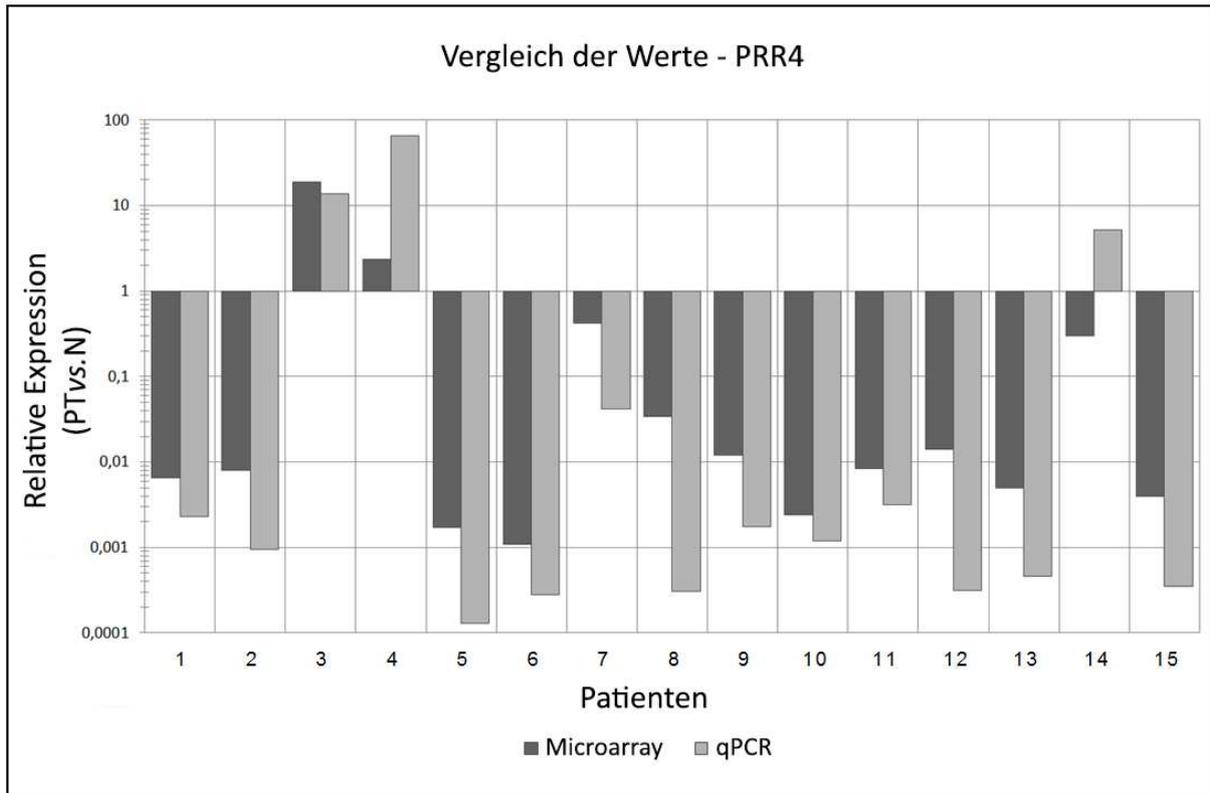


Abb. 14: Vergleich der relativen Expression für das Gen PRR4 in PTvs.N.

Die mittels Microarray und qPCR erhaltenen Expressionswerte wurden als Verhältnis von Primärtumor (PT) gegenüber Normalgewebe (N) für die einzelnen Patienten aufgetragen. Für fast alle Patienten bis auf Patient 14, bei dem eine sehr kleine Ratio und damit ein geringer Unterschied zwischen PT und N errechnet wurde, konnten die Microarray Daten durch die qPCR Ergebnisse bestätigt werden.

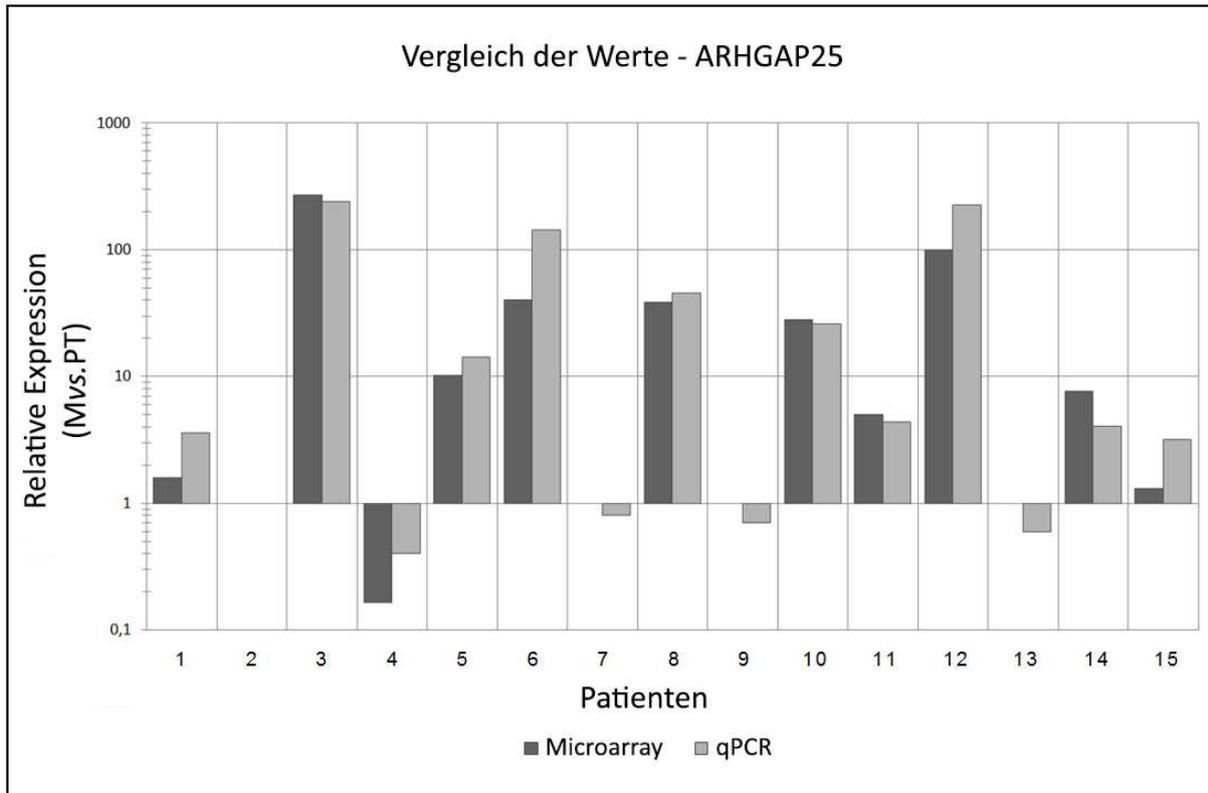


Abb. 15: Vergleich der relativen Expression für das Gen ARHGAP25 in Mvs.PT.

Die mittels Microarray und qPCR erhaltenen Expressionswerte wurden als Verhältnis von Metastase (M) gegenüber Primärtumor (PT) für die einzelnen Patienten aufgetragen. Für die meisten Patienten konnte in Metastasen eine deutlich höhere Expression von ARHGAP25 gemessen werden. Die Microarray-Daten wurden für alle Patienten durch die qPCR bestätigt.

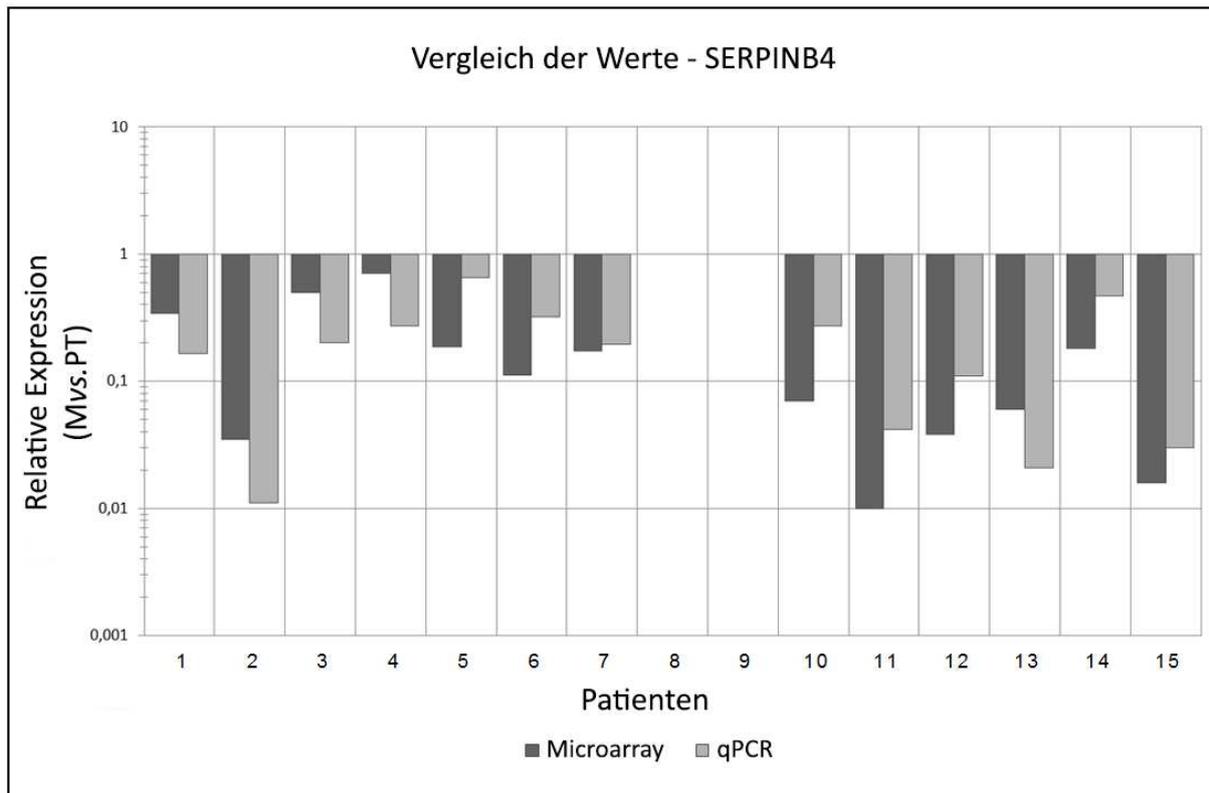


Abb. 16: Vergleich der relativen Expression für das Gen SERPINB4 in Mvs.PT.

Die mittels Microarray und qPCR erhaltenen Expressionswerte wurden als Verhältnis von Metastase (M) gegenüber Primärtumor (PT) für die einzelnen Patienten aufgetragen. Für alle Patienten, bei denen SERPINB4 detektiert werden konnte, konnte eine Herunterregulation sowohl durch Microarray als auch durch qPCR gezeigt werden.

Mit der erfolgreichen Validierung der Einzeldaten wurde die Qualität und Zuverlässigkeit der erhaltenen Chip-Array Daten unterstrichen. Damit wurde die Voraussetzung geschaffen, die aus den Microarray-Analysen erhaltenen Daten weiter bioinformatisch zu bearbeiten. Dies umfasst sowohl die detaillierte funktionelle wie klinische Auswertung sowie die drauf aufbauende Erstellung entsprechender mechanistischer Hypothesen. Die folgenden Berechnungen und Auswertungen wurden mit dem *limma*-Paket in der Statistiksoftware *R* durchgeführt. Da trotz Durchsicht und Bearbeitung der Proben durch einen versierten Pathologen und sorgfältiger Durchführung der RNA-Isolation immer leichte Unterschiede im RNA-Gehalt wie auch der Bindeaffinität der Proben auf den Arrays bestehen, müssen die erhaltenen Intensitätswerte mit Hilfe der auf den Chips enthaltenen Positiv- und Negativkontrollen bioinformatisch geglättet werden. Die Daten wurden zunächst mit der

für Affymetrix-Chips gängigen und robusten RMA-Methode als Korrektur für das Hintergrundrauschen normalisiert (**Abb. 17**).

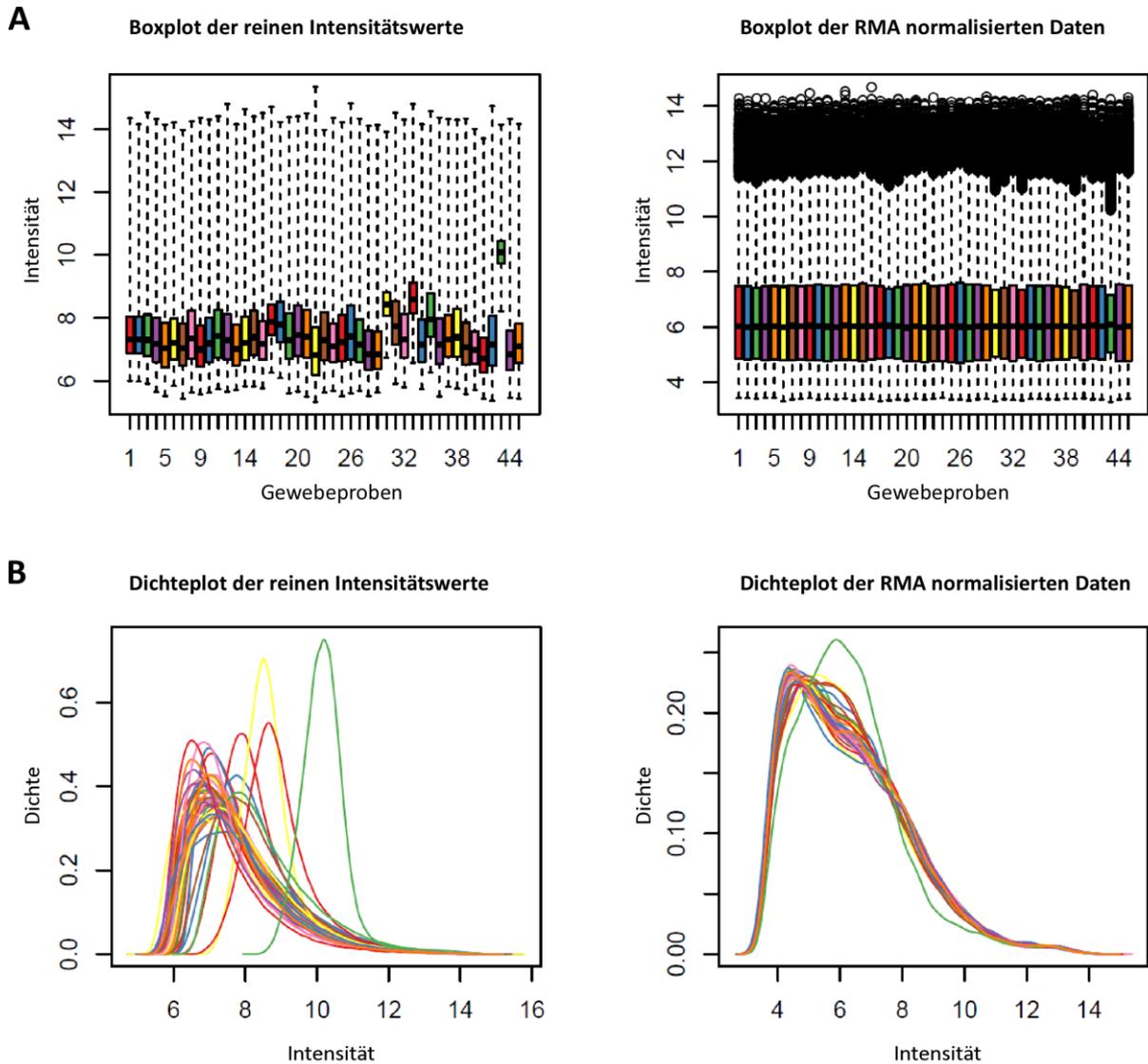


Abb. 17: Grafische Darstellung der RMA-Normalisierung der rohen Microarray-Intensitätswerte.

A. Die Verteilung der Daten vor (links) und nach (rechts) der RMA Normalisierung wurden mit Box Plots visualisiert. **B.** Hier ist die Verteilung der Daten vor (links) und nach (rechts) der RMA Normalisierung als Dichteplot dargestellt.

Um die Variabilität zwischen den Array-Daten der 45 verschiedenen Gewebeproben zu eliminieren und die Proben miteinander vergleichbar zu machen, gibt es im *limma*-Paket zwei verschiedene Quantilnormalisierungen („*aquantile*“ und „*quantile*“) innerhalb der Funktion

„normalizeBetweenArrays“. Beide Optionen der Funktion wurden getestet und zeigten keine signifikanten Unterschiede (**Abb. 18**). Die Voreinstellung „aquantil“ wurde für die weiteren Berechnungen gewählt.

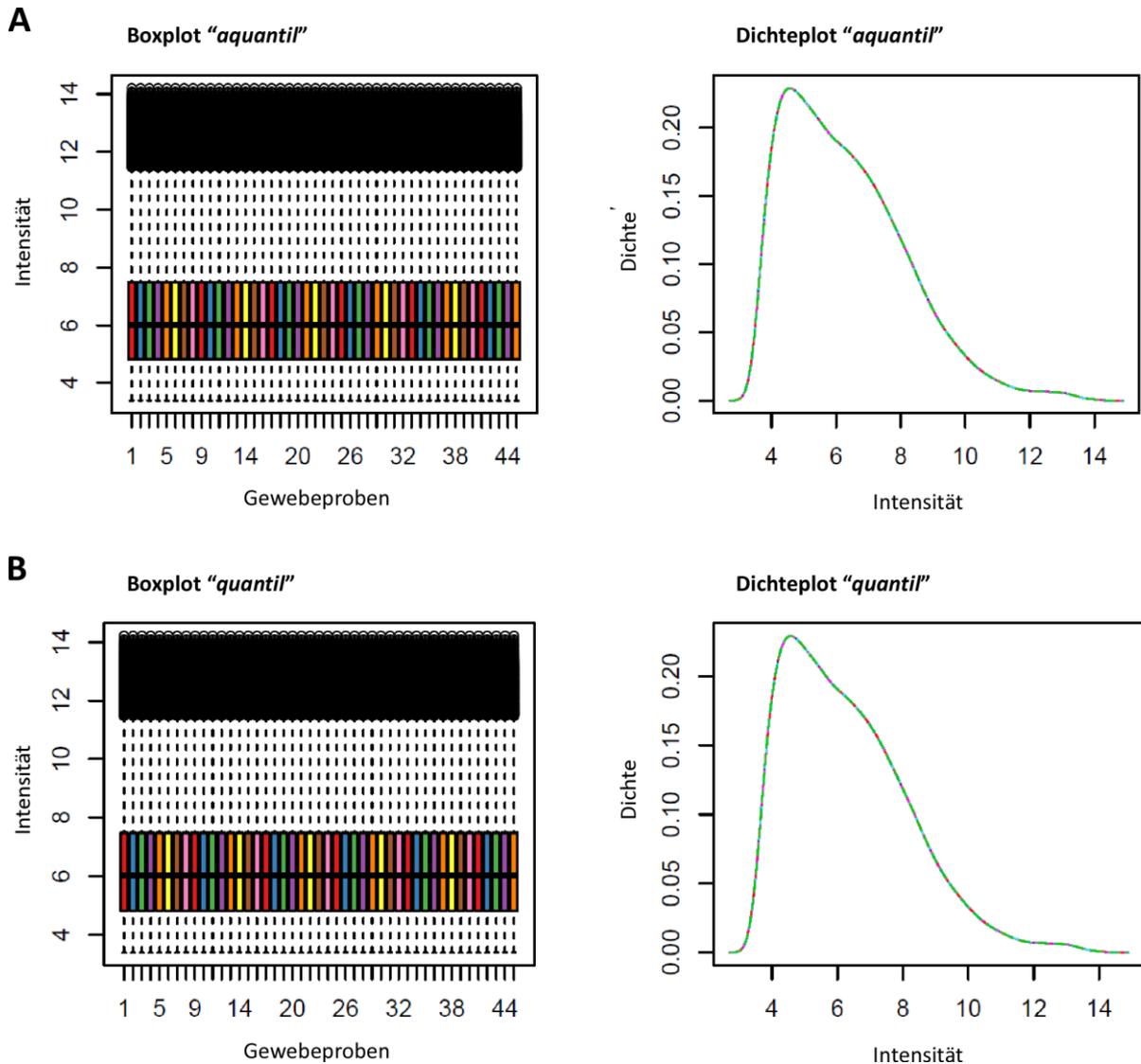


Abb. 18: Grafische Darstellung der Quantilnormalisierung der RMA-normalisierten Arraydaten.

A. Boxplot (links) und Dichteplot (rechts) der Quantil-normalisierten Daten nach Nutzung der Voreinstellung „aquantil“ der Funktion „normalizeBetweenArrays“. **B.** Boxplot (links) und Dichteplot (rechts) der Quantil-normalisierten Daten nachdem die Funktion mit der Voreinstellung „quantil“ verwendet wurde.

Im nächsten Schritt wurden mit Hilfe des *limma*-Pakets gepaarte Analysen über alle Patienten und für jede einzelne Affy-ID durchgeführt. Verglichen wurden die Gewebekombinationen Primärtumor gegenüber Normalgewebe (PTvs.N), Lymphknotenmetastase gegenüber Normalgewebe (Mvs.N) und

Lymphknotenmetastase gegenüber Primärtumor (Mvs.PT). Die Zuordnung der Gen-Namen erfolgte anhand des von der *Core Facility Bioinformatics* eigens entwickelten R-Pakets „*BiumNGSBase*“. Die Tabellen beinhalten die jeweilige ID des geprüften RNA-Abschnitts auf den Chips (Affy-ID), das dazugehörige Gen-Symbol, die durchschnittliche Veränderung (Fold Change, FC) der Expression über der Basis 2 in logarithmierter Form (\log_2FC), die mittlere Expression (AveExpr), die moderate t-Statistik (t), der p-Wert und der adjustierte p-Wert nach der Benjamini-Hochberg-Prozedur (adj. p-Wert). Eine dreifache Überexpression würde dabei einem \log_2FC von etwa 1,6 entsprechen. Die TOP 50 Listen für PTvs.N (**Tab. 36**), Mvs.PT (**Tab. 37**), und Mvs.N (**Tab. 38**) befinden sich im Anhang.

Um einen Überblick über die Verwandtschaft der Proben bei solch großen Datenmengen zu bekommen, ist es sinnvoll, die Daten in sog. Heatmaps zu visualisieren, welche mittels Cluster-Analysen erstellt werden. Die Ähnlichkeit der Proben wird dabei durch die räumliche Anordnung und durch Verbindungslinien gekennzeichnet. Eine erste Übersicht der Verteilung der 45 Proben ohne jegliche Filter, aber nach der Quantilnormalisierung ist in **Abb. 19** dargestellt. Durch den Farbschlüssel lässt sich auf den ersten Blick erkennen, dass die meisten Gene eine niedrige Intensität (grün) aufweisen und sich erst etwa im unteren Achtel Gene häufen, welche mehr als eine mittlere Intensität (schwarz) zeigen. Erst die untersten Zeilen sind rot, beschreiben also Gene, die mit einer hohen Intensität in den Gewebeproben exprimiert wurden.

Zur weiteren Veranschaulichung wurde die gleiche Methode nur auf die Gewebeproben aus Primärtumor und Normalgewebe (**Abb. 20**) sowie Lymphknotenmetastasen und Primärtumor (**Abb. 21**) angewendet. Während beim Vergleich aller Proben (**Abb. 19**) keine Gruppierung zu erkennen war, ordnen sich die Gewebeproben im PTvs.N Clustering (**Abb. 20**) bereits untereinander, wie aus der Beschriftung der x-Achse unter der Heatmap erkenntlich wird, was auf eine relativ distinkte Unterscheidung der Expression von PT zu N schließen lässt. Der Vergleich Mvs.PT (**Abb. 21**) ist dagegen eher ungeordnet, was auf eine größere Verwandtschaft der verglichenen Entitäten M und PT untereinander schließen lässt.

ERGEBNISSE

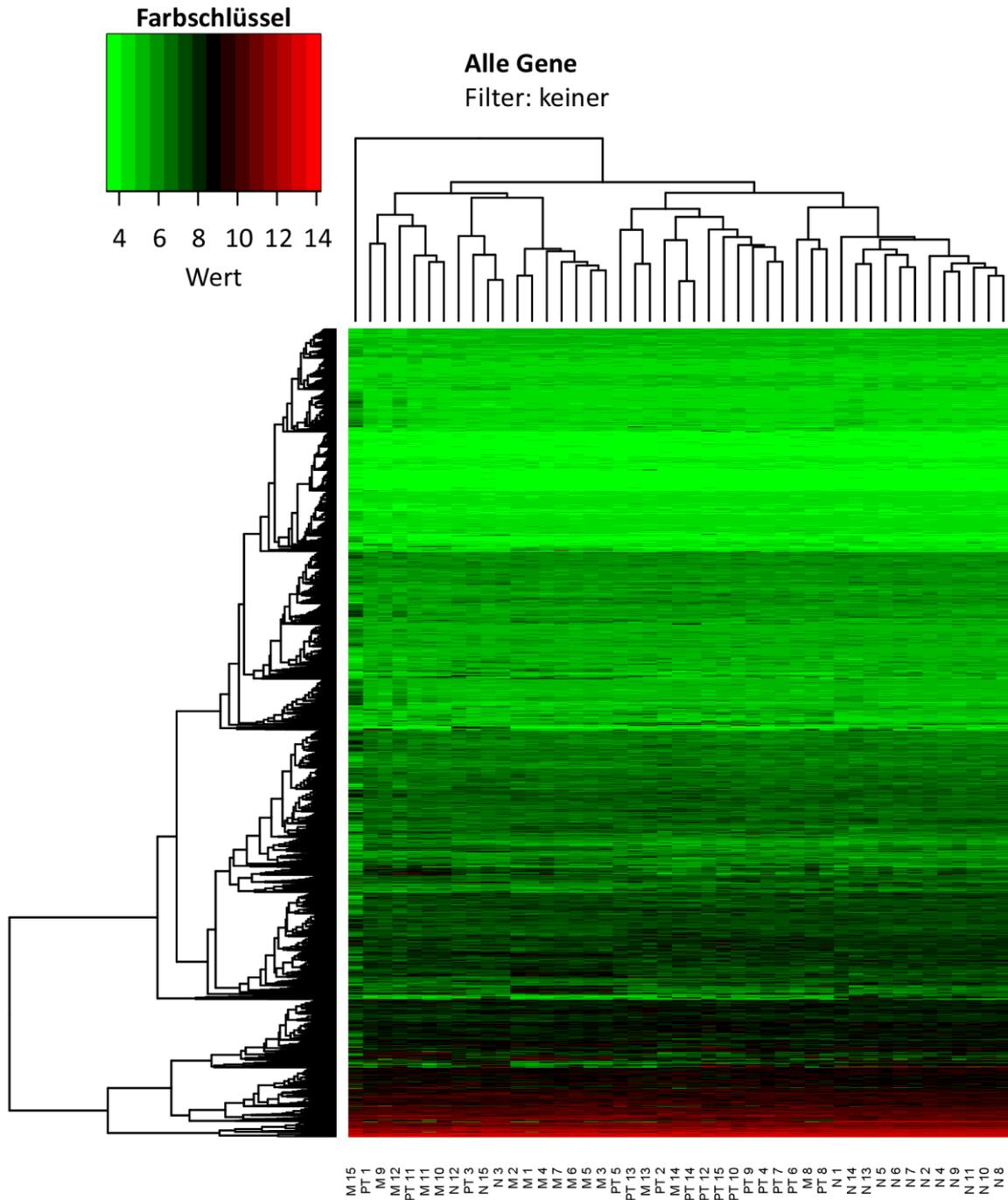


Abb. 19: Hierarchisches Clustering aller Affy-IDs (alle Gene) über alle Gewebeprobe. Visualisierung der Ähnlichkeiten des Genmusters aller Gewebeprobe (x-Achse) zu den Ähnlichkeiten aller einzelnen Affy-IDs untereinander über die Gewebeprobe (y-Achse). Farbschlüssel: Grün: Niedrige Intensität der Expression. Rot: Hohe Intensität der Expression. Schwarz: Mittlere Intensität der Expression.

ERGEBNISSE

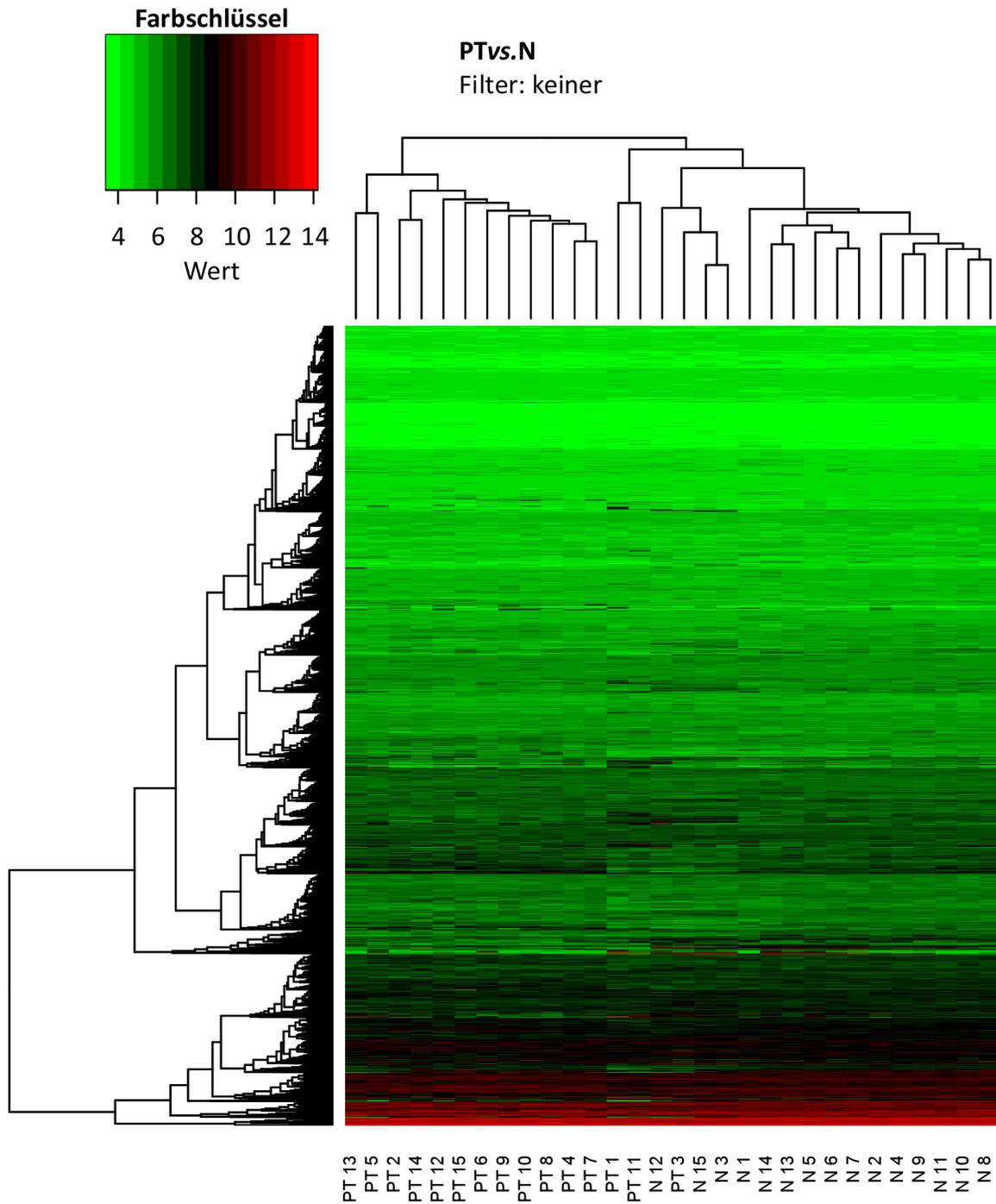


Abb. 20: Hierarchisches Clustering aller Affy-IDs (alle Gene) über die Gewebeproben aus Primärtumoren (PT) und Normalgewebe (N).

Visualisierung der Ähnlichkeiten des Genmusters der Gewebeproben aus PTvs.N (x-Achse) zu den Ähnlichkeiten aller einzelnen Affy-IDs untereinander über die Gewebeproben (y-Achse). Farbschlüssel: Grün: Niedrige Intensität der Expression. Rot: Hohe Intensität der Expression. Schwarz: Mittlere Intensität der Expression.

ERGEBNISSE

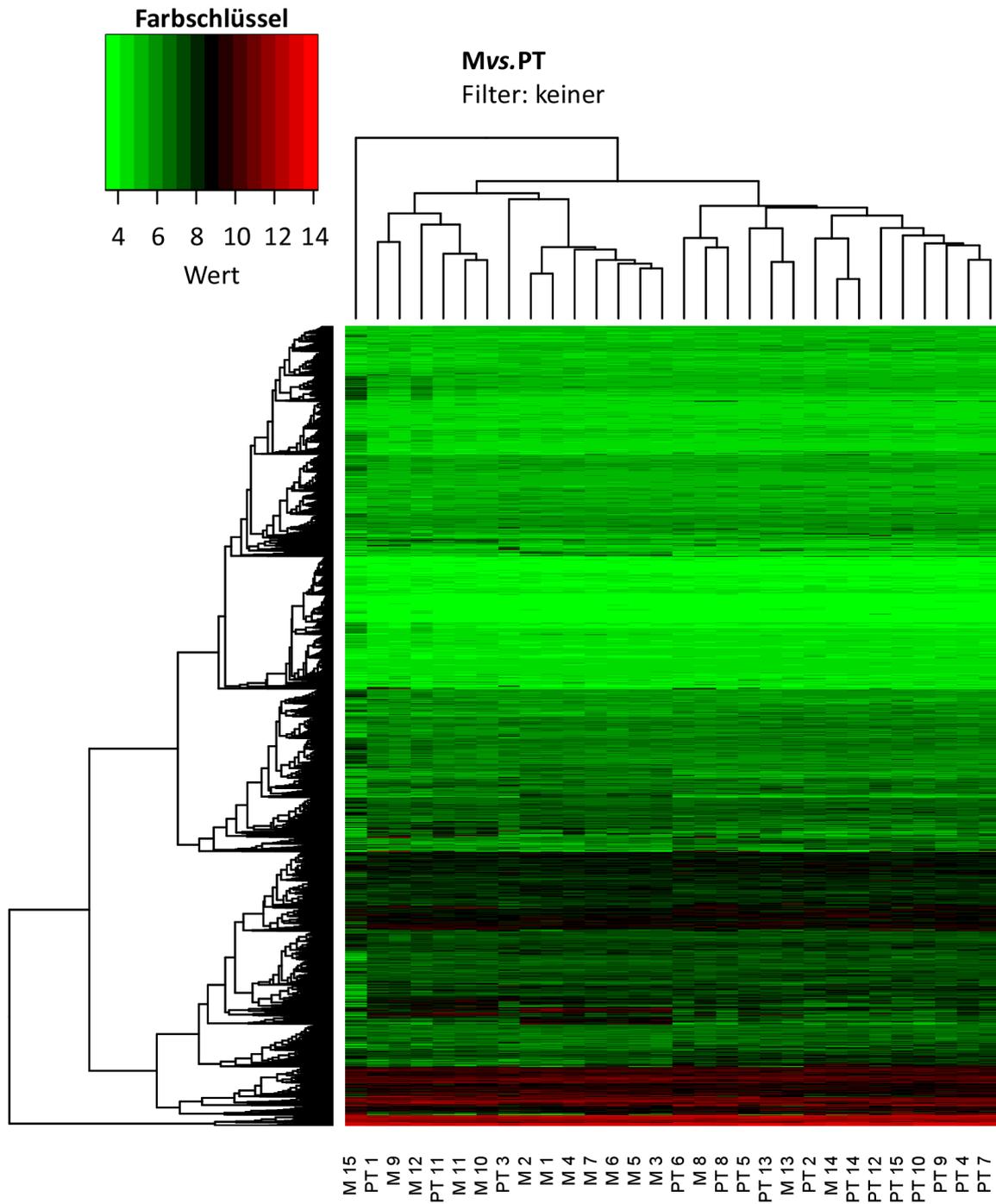


Abb. 21: Hierarchisches Clustering aller Affy-IDs (alle Gene) über die Gewebeproben aus Metastasen (M) und Primärtumoren (PT).

Visualisierung der Ähnlichkeiten des Genmusters der Gewebeproben aus Mvs.PT (x-Achse) zu den Ähnlichkeiten aller einzelnen Affy-IDs untereinander über die Gewebeproben (y-Achse). Farbschlüssel: Grün: Niedrige Intensität der Expression. Rot: Hohe Intensität der Expression. Schwarz: Mittlere Intensität der Expression.

Eine derart große Anzahl an Affy-IDs miteinander zu vergleichen gestaltet detailliertere Analysen als schwierig. Ziel der nächsten Berechnungen war also, die Anzahl der Gene auf die signifikantesten zu reduzieren und idealerweise eine Gensignatur von etwa 30 Genen (bei möglicherweise mehr Affy-IDs durch Mehrfachbesetzungen) für PTvs.N und Mvs.PT zu erstellen, um diese weiterführend auf ihr Potential als prognostische Biomarker zu untersuchen. Dafür werden Gene gesucht, die bei möglichst allen Patienten gleichmäßig hoch- oder herabreguliert wurden. Nach ersten Vorversuchen wurden die Clustering-Ergebnisse auf die verschiedenen Gewebetypen justiert (Larynx, Hypopharynx, Oropharynx), der \log_2FC auf mind. 1,5 festgelegt und als Signifikanzniveaus der adjustierte p-Wert nach der Benjamini-Hochberg-Prozedur (False Discovery Rate, FDR) als auch der p-Wert in separaten Clustering-Analysen auf ihre Nutzbarkeit für diese Fragestellung überprüft.

Bei den Filterkriterien $\log_2FC \geq 1,5$ und $FDR < 0,025$ war weder für PTvs.N (**Abb. 22**) noch für Mvs.PT (**Abb. 23**) eine distinkte Unterscheidung der Entitäten erreicht. In **Abb. 22** ordnen sich die Gensignaturen von N3, N12 und N15 bei den Primärtumoren auf der linken Seite ein, und in **Abb. 23** sind mehrere Proben Lymphknotenmetastasen unter die Primärtumore gemischt. Während bei den Gensignaturen von Lymphknotenmetastasen zu Primärtumoren ein fließender Übergang der Expressionsmuster zu erwarten ist, wie bereits in **Abb. 21** bei Inklusion aller Gene zu erkennen, ist eine distinkte Unterscheidung von Primärtumoren zu Normalgewebe wie in **Abb. 20** (mit Ausnahme von PT3) für eine weitere Analyse unerlässlich und anzustreben.

Für die Filterkriterien $\log_2FC \geq 1,5$ und p-Wert $< 0,0001$ konnte für PTvs.N eine sehr ähnliche Aufteilung erreicht werden wie in **Abb. 20**, ebenfalls mit Ausnahme von PT3, der zwischen dem Normalgewebe einsortiert ist (**Abb. 24**). Für den Vergleich Mvs.PT war der Übergang wie erwartet eher fließend von PT zu M (**Abb. 25**). Der Vollständigkeit halber wurde die Clustering-Analyse mit den gleichen Filterkriterien auch für Mvs.N durchgeführt, was ebenfalls zu einer zufriedenstellenden Separation der Entitäten führte (**Abb. 26**). Da aus diesen Filterkriterien die stringenteste Aufteilung folgte, wurden sie nach Möglichkeit für alle folgenden Auswertungen angesetzt.

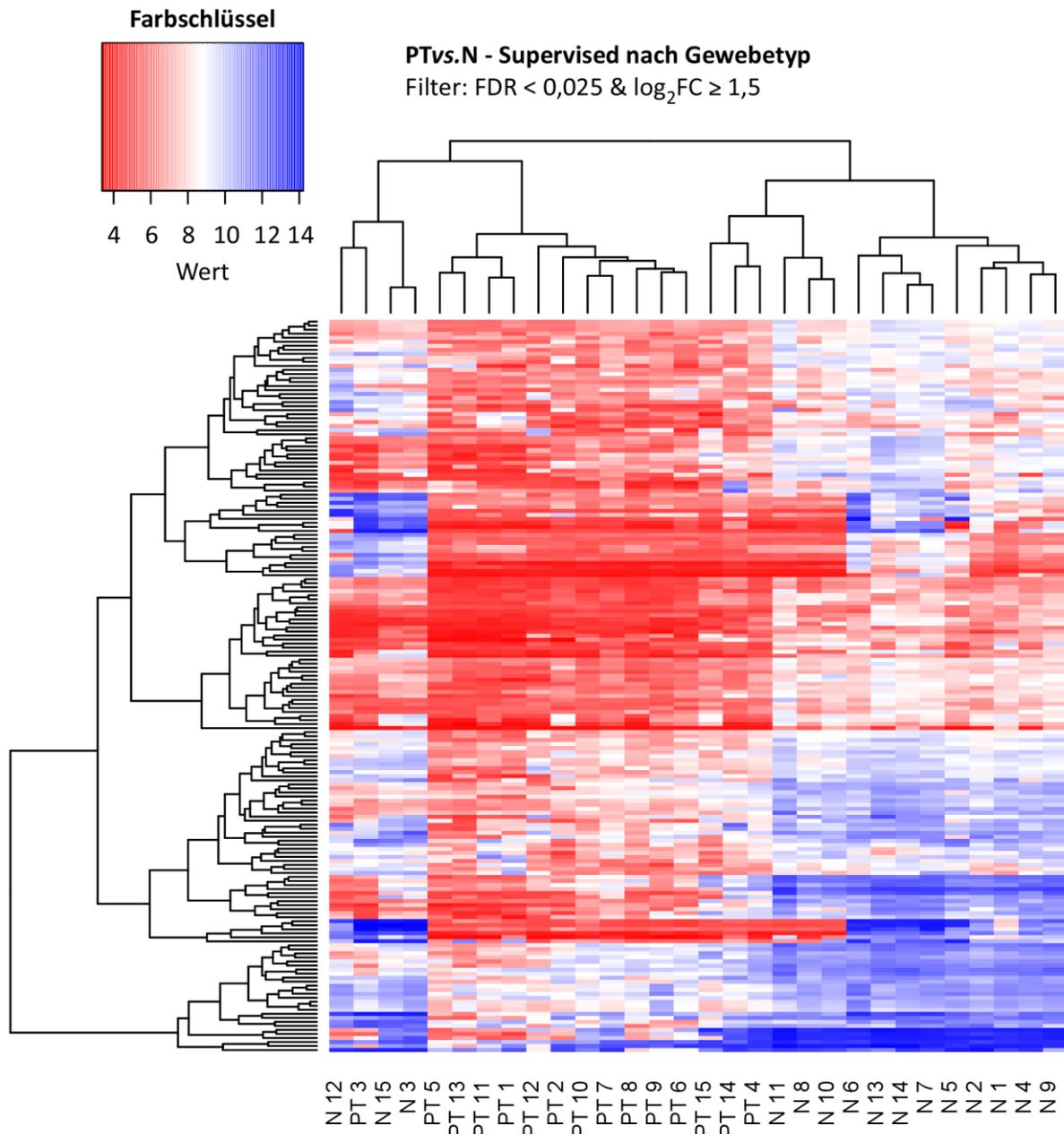


Abb. 22: Hierarchisches Clustering gefilterter Affy-IDs über die Gewebeproben aus Primärtumoren (PT) und Normalgewebe (N).

Visualisierung der Ähnlichkeiten des Genmusters der Gewebeproben aus PTvs.N (x-Achse) zu den Ähnlichkeiten der einzelnen Affy-IDs untereinander über die Gewebeproben (y-Achse). Stringente Filterkriterien ($\log_2FC \geq 1,5$ und $FDR < 0,025$) wurden zur Reduktion der Affy-IDs zu Gunsten eines höheren Signifikanzniveaus verwendet. Farbschlüssel: Rot: Niedrige Intensität der Expression. Blau: Hohe Intensität der Expression. Weiß: Mittlere Intensität der Expression.

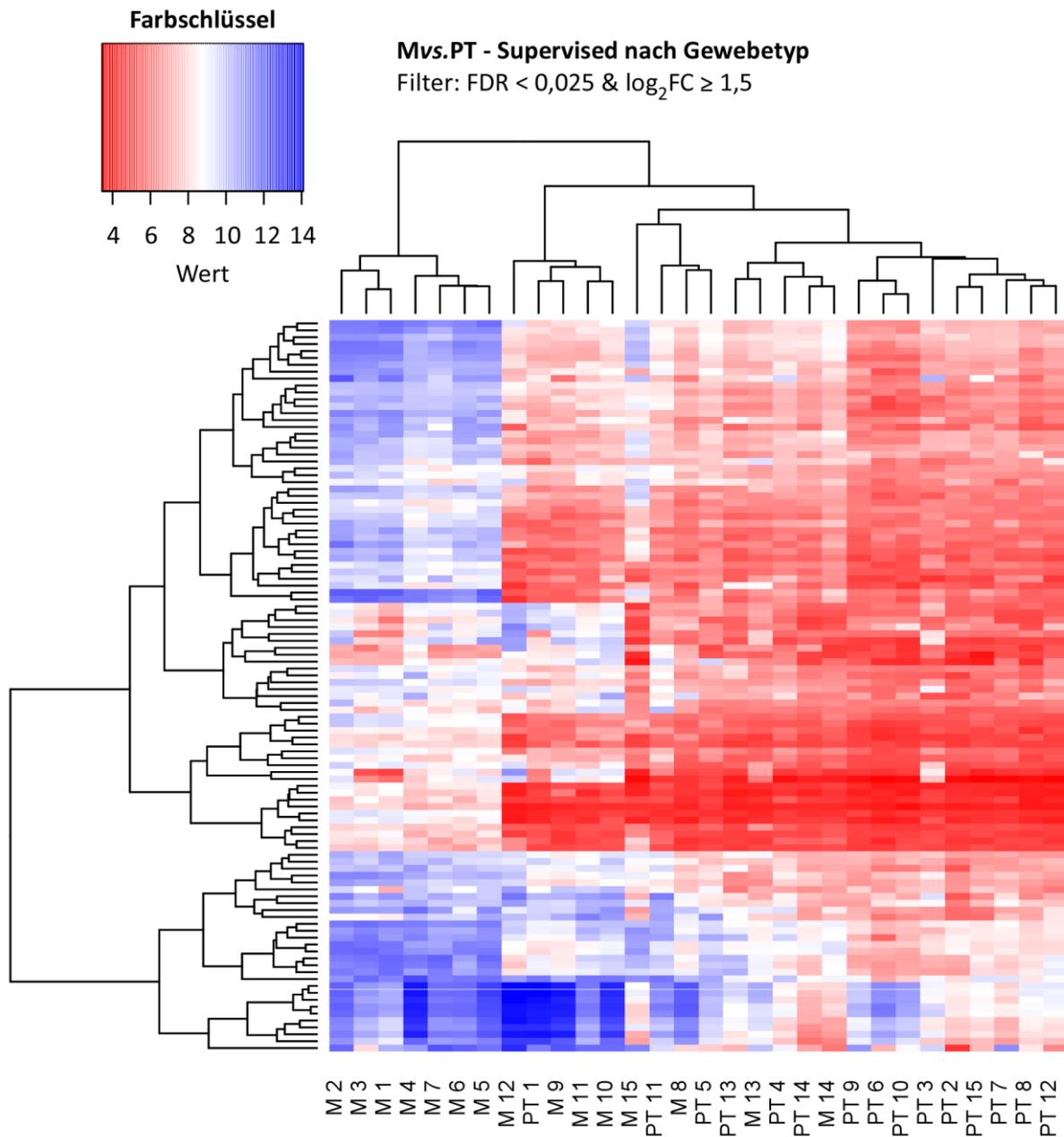


Abb. 23: Hierarchisches Clustering gefilterter Affy-IDs über die Gewebeproben aus Metastasen (M) und Primärtumoren (PT).

Visualisierung der Ähnlichkeiten des Genmusters der Gewebeproben aus Mvs.PT (x-Achse) zu den Ähnlichkeiten der einzelnen Affy-IDs untereinander über die Gewebeproben (y-Achse). Stringente Filterkriterien ($\log_2FC \geq 1,5$ und $FDR < 0,025$) wurden zur Reduktion der Affy-IDs zu Gunsten eines höheren Signifikanzniveaus verwendet. Farbschlüssel: Rot: Niedrige Intensität der Expression. Blau: Hohe Intensität der Expression. Weiß: Mittlere Intensität der Expression.

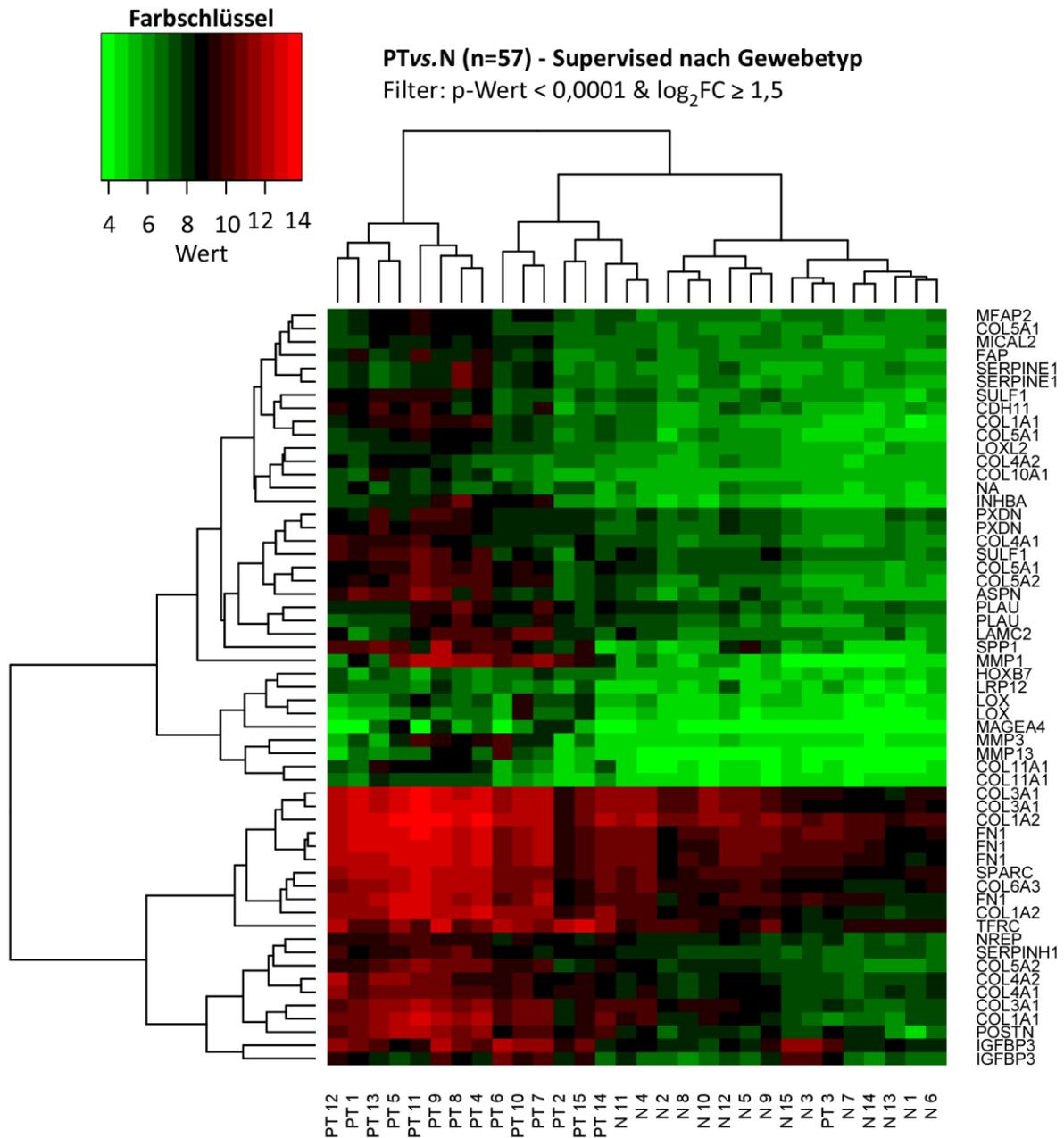


Abb. 24: Hierarchisches Clustering der Gensignatur von Primärtumoren (PT) und Normalgewebe (N).

Visualisierung der Ähnlichkeiten des Genmusters der Gewebeproben aus PTvs.N (x-Achse) zu den Ähnlichkeiten der einzelnen Affy-IDs untereinander über die Gewebeproben (y-Achse). Stringente Filterkriterien ($\log_2FC \geq 1,5$ und p-Wert < 0,0001) wurden zur Reduktion der Affy-IDs zu Gunsten eines höheren Signifikanzniveaus verwendet und ergaben eine Gensignatur von 38 differenziell exprimierten Genen bei 57 Affy-IDs. Grün: Niedrige Intensität der Expression. Farbschlüssel: Rot: Hohe Intensität der Expression. Schwarz: Mittlere Intensität der Expression.

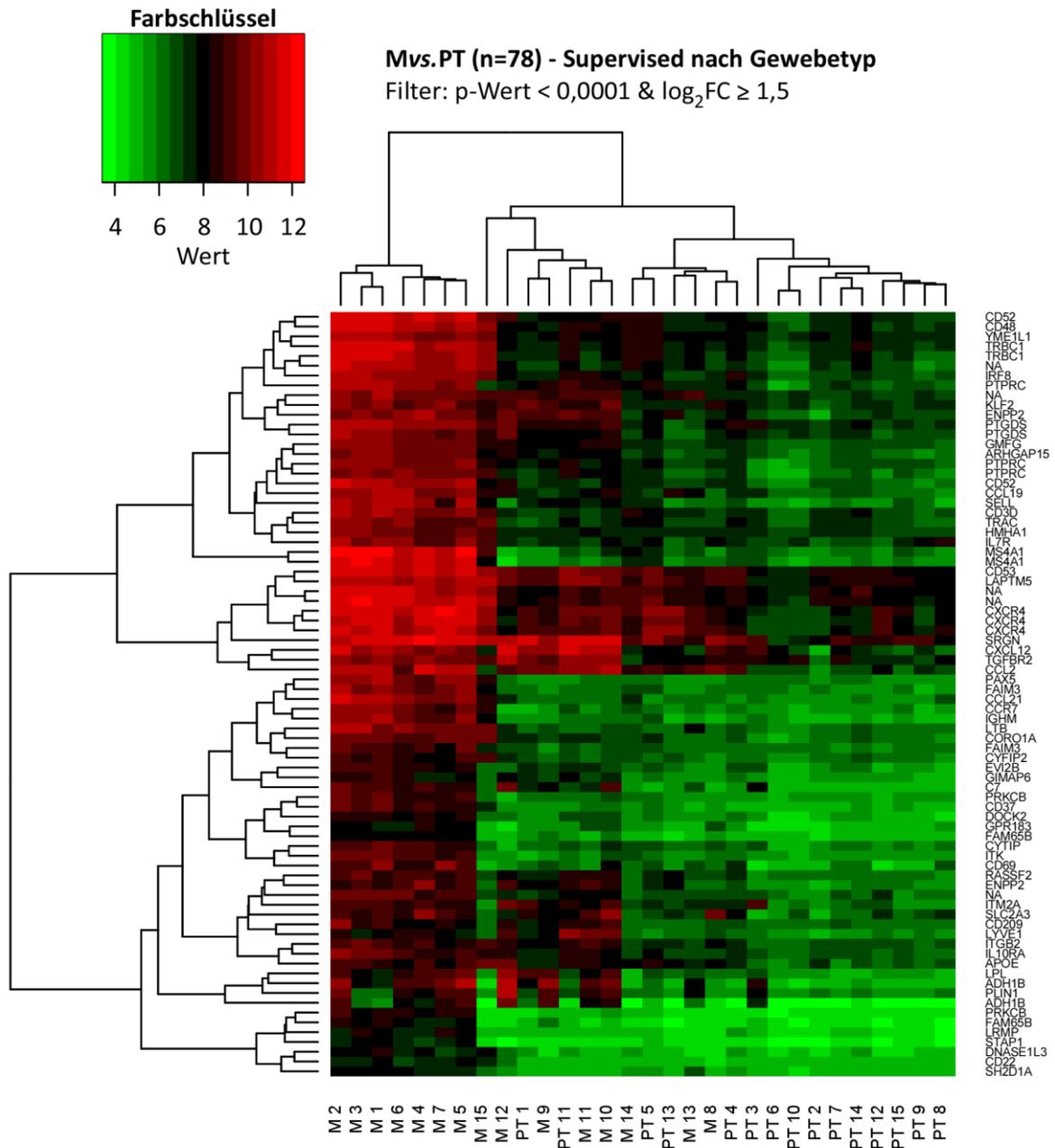


Abb. 25: Hierarchisches Clustering der Gensignatur von Metastasen (M) und Primärtumoren (PT).

Visualisierung der Ähnlichkeiten des Genmusters der Gewebeproben aus Mvs.PT (x-Achse) zu den Ähnlichkeiten der einzelnen Affy-IDs untereinander über die Gewebeproben (y-Achse). Stringente Filterkriterien ($\log_2FC \geq 1,5$ und p-Wert < 0,0001) wurden zur Reduktion der Affy-IDs zu Gunsten eines höheren Signifikanzniveaus verwendet und ergaben eine Gensignatur von 62 differenziell exprimierten Genen bei 78 Affy-IDs. Farbschlüssel: Grün: Niedrige Intensität der Expression. Rot: Hohe Intensität der Expression. Schwarz: Mittlere Intensität der Expression.

ERGEBNISSE

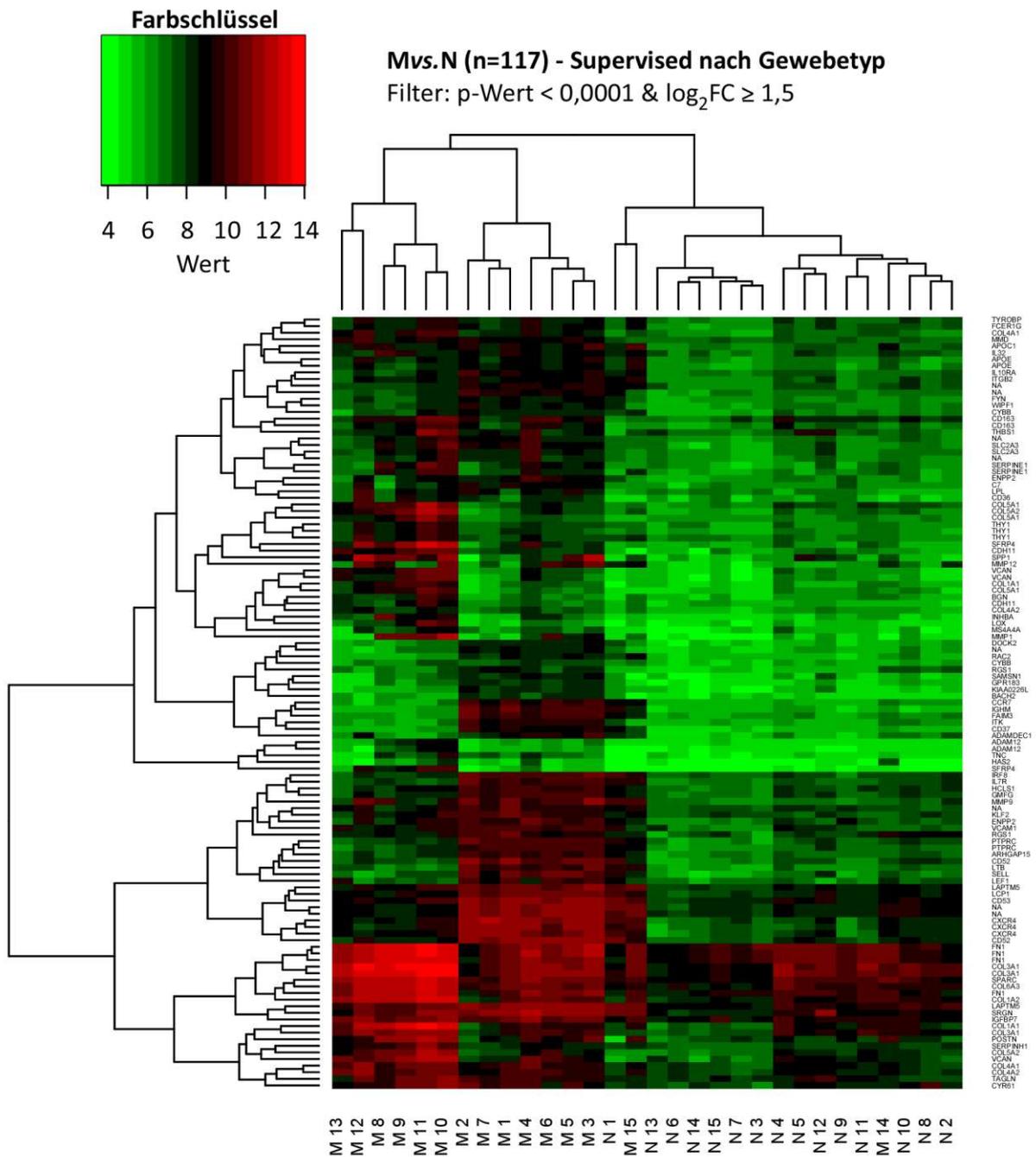


Abb. 26: Hierarchisches Clustering der Gensignatur von Metastasen (M) und Normalgewebe (N). Visualisierung der Ähnlichkeiten des Genmusters der Gewebeproben aus Mvs.N (x-Achse) zu den Ähnlichkeiten der einzelnen Affy-IDs untereinander über die Gewebeproben (y-Achse). Stringente Filterkriterien ($\log_2FC \geq 1,5$ und p-Wert < 0,0001) wurden zur Reduktion der Affy-IDs zu Gunsten eines höheren Signifikanzniveaus verwendet und ergaben eine Gensignatur von 96 differenziell exprimierten Genen bei 117 Affy-IDs. Grün: Niedrige Intensität der Expression. Farbschlüssel: Rot: Hohe Intensität der Expression. Schwarz: Mittlere Intensität der Expression.

Die erhaltenen Gensignaturen für PTvs.N (**Tab. 2**) und Mvs.PT (**Tab. 3**) wurden im Detail aufgelistet und zur Herstellung biologischer Zusammenhänge mit allgemeinen Informationen aus der Online-Datenbank Uniprot.org*¹ gefüllt.

Mehrere Proteine, die in die Organisation und den Zusammenbau der Extrazellulären Matrix (ECM) involviert sind, waren in PTvs.N (**Tab. 2**) überexprimiert: Matrix-Metallopeptidasen (MMP1, MMP3, MMP13), Kollagene (COL1A1, COL1A2, COL3A1, COL4A1, COL4A2, COL5A1, COL5A2, COL6A3, COL10A1, COL11A1), Asporin (ASPN), Fibronectin (FN1), Laminin (LAMC2), Lysyl Oxidasen (LOX, LOXL2), Microfibrillar-Associated Protein (MFAP2), Periostin (POSTN), Serpin Inhibitoren (SERPINE1, SERPINH1), Osteonectin (SPARC) und Osteopontin (SPP1).

Einige differenziell exprimierte Gene unterschiedlicher Funktion konnten auf den ersten Blick keiner speziellen Kategorie zugeordnet werden: Fibroblast Activation Protein (FAP), Cadherin 11 Type 2 (CDH11), Homeobox B7 (HOXB7), Insulin-Like Growth Factor Binding Protein 3 (IGFBP3), Inhibin (INHBA), Low Density Lipoprotein Receptor-Related Protein 12 (LRP12), Melanoma Antigen A4 (MAGEA4), a monooxygenase (MICAL2), Neuronal Regeneration Related Protein (NREP), Urokinase (PLAU), Peroxidasin Homolog (PXDN), Schwannomin (SCHIP1) und Transferrin (TFRC).

Die höhere Expression von KIF23, UBE2C, MCM7, NUP107, RSRC1, MAD2, BTAK, RFC5, ECT2 und CENPA wurde bereits in metastasierenden Tumoren oder Tumoren mit malignem Potenzial in einer früheren Studie identifiziert ¹⁵¹. Von den letzten fünf Genen wurde auch bereits in van't Veers Brustkrebs-Studie berichtet ¹⁷⁶.

*¹ <http://www.uniprot.org/> abgerufen im Juni 2015.

Tab. 2: Gensignatur des Vergleichs PTvs.N (Filter: $\log_2FC \geq 1,5$ und p-Wert $< 0,0001$)

Gen	\log_2FC	p-Wert	Protein	Lokalisation	Molekulare Funktion	Biologischer Prozess	Ligand	Reaktom
ASPN	2,7333	6,72E-06	Asporin	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix	Calcium Ion Bindung	Mineralisierung	Calcium	ECM Proteoglykane
CDH11	2,1314	7,13E-05	Cadherin 11, Type 2 (OSF-4)	Zellmembran	Calcium Ion Bindung	Zelladhäsion	Calcium Metall-Bindung	Adhärenz junctions Interaktionen
COL1A1	2,7697	7,79E-07	Collagen, Type I, Alpha 1	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix	Metall Ion Bindung	-	Calcium Metall-Bindung	ECM Organisation; Integrin und Non-Integrin Membran-ECM Interaktionen;
COL1A2	2,0748	2,74E-05	Collagen, Type I, Alpha 2	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix	Metall Ion Bindung	-	Calcium Metall-Bindung	ECM Organisation; Integrin und Non-Integrin Membran-ECM Interaktionen;
COL3A1	2,0382	4,13E-06	Collagen, Type III, Alpha 1	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix	Metall Ion Bindung	-	Calcium Metall-Bindung	ECM Organisation; Integrin und Non-Integrin Membran-ECM Interaktionen; Signaling von PDGF
COL4A1	1,9220	1,58E-08	Collagen, Type IV, Alpha 1	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix Basalmembran	Metall Ion Bindung	Angiogenese	Calcium Metall-Bindung	ECM Organisation; Integrin und Non-Integrin Membran-ECM Interaktionen; Signaling von PDGF;
COL4A2	1,7942	1,34E-07	Collagen, Type IV, Alpha 2	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix Basalmembran	Metall Ion Bindung	Angiogenese	Calcium Metall-Bindung	ECM Organisation; Integrin und Non-Integrin Membran-ECM Interaktionen; Signaling von PDGF;
COL5A1	1,9556	3,05E-05	Collagen, Type V, Alpha 1	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix	Metall Ion Bindung	-	Calcium Heparin-Bindung Metall-Bindung	ECM Organisation; Integrin und Non-Integrin Membran-ECM Interaktionen; Signaling von PDGF;
COL5A2	2,4991	8,50E-07	Collagen, Type V, Alpha 2	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix	Metall Ion Bindung	-	Calcium Metall-Bindung	ECM Organisation; Integrin und Non-Integrin Membran-ECM Interaktionen; Signaling von PDGF;
COL6A3	1,7597	1,03E-06	Collagen, Type VI, Alpha 3	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix	Serine Protease Inhibitor	-	Zelladhäsion	ECM Organisation; Integrin Zelloberfläche Interaktionen; Signaling von PDGF;
COL10A1	1,7102	3,33E-06	Collagen, Type X, Alpha 1	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix	Metall Ion Bindung	-	Calcium Metall-Bindung	Integrin und Non-Integrin Membran-ECM Interaktionen
COL11A1	2,2924	5,46E-05	Collagen, Type XI, Alpha 1	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix	Metall Ion Bindung	-	Calcium Metall-Bindung	Non-Integrin Membran-ECM Interaktionen

ERGEBNISSE

FAP	1,6717	9,01E-05	Fibroblast Activation Protein Alpha	Zellmembran	Hydrolase Serine Protease	Angiogenese Apoptose Zelladhäsion	-	
FN1	1,8254	6,56E-06	Fibronectin 1	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix	Acute Phase Angiogenese Zelladhäsion Zellform	-	Heparin-Bindung	Zersetzung der Extrazelluläre Matrix; Integrin Zell Oberfläche Interaktionen; MAPK Signaling für Integrine
HOXB7	1,5093	7,80E-09	Homeobox B7	Nucleus	Entwicklungsprotein	Transkription Transkription Regulation	DNA-Bindung	
IGFBP3	1,6046	7,30E-05	Insulin-Like Growth Factor Binding Protein 3	Sekretiert Nucleus	Apoptose	Wachstumsfaktor Bindung		Regulation von IGF Transport und Uptake von IGFBPs
INHBA	2,7222	4,22E-07	Inhibin, Beta A	Sekretiert	Growth Factor Hormone	-	-	Signaling von Activin
LAMC2	1,9688	2,06E-05	Laminin, Gamma 2	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix Basalmembran	Heparin-Bindung	Zelladhäsion	Heparin-Bindung	Zersetzung von ECM; Non-Integrin Membran-ECM Interaktionen; Anchoring Fibril Formation
LOX	1,8267	3,43E-05	Lysyl Oxydase	Sekretiert Extrazellulärer Raum	Oxidoreduktase		Metall-Bindung	Crosslinking von Collagen Fasern
LOXL2	1,7868	8,94E-08	Lysyl Oxydase-like 2	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix Basalmembran Nucleus	Chromatin Regulator Oxidoreduktase Repressor	Transkription Regulation	Metall-Bindung	Crosslinking von Collagen Fasern
LRP12	1,5531	1,18E-06	Low Density Lipoprotein Receptor-Related Protein 12	Membran	Receptor	Endozytose Host-virus Interaktion		Retinoid Metabolismus und Transport
MAGEA4	1,7378	5,33E-06	Melanoma Antigen Family A, 4	-	-	-	-	-
MFAP2	1,6469	3,23E-07	Microfibrillar-Associated Protein 2	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix	-	ECM Organisation		Elastische Faserformation
MICAL2	1,5042	9,79E-07	Microtubule Associated Monooxygenase, Calponin and LIM Domain containing2	Zytoplasma (Zytoskelett)	Monooxygenase Oxidoreduktase	Aktin-Bindung FAD Metall-Bindung NADP		Zytoskelett Organisation; Oxidation-Reduktion Prozess
MMP1	4,8582	2,38E-09	Matrix	Sekretiert	Hydrolase,	Collagen Zersetzung	Calcium	EC No. 3.4.24.7.; Zersetzung von ECM; Collagen

ERGEBNISSE

			Metallopeptidase 1	Extrazellulärer Raum Extrazelluläre Matrix	Metalloprotease	Host-virus Interaktion	Metall-Bindung Zink	Zersetzung; Regulation von IGF Transport und Uptake
MMP3	2,7054	2,00E-07	Matrix Metallopeptidase 3	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix	Hydrolase, Metalloprotease	Collagen Zersetzung	Calcium Metall-Bindung Zink	Zersetzung von ECM, Collagen Zersetzung, Assemblierung von Collagen Fasern, EGFR Transaktivierung
MMP13	2,3693	6,55E-07	Matrix Metallopeptidase 13	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix	Hydrolase, Metalloprotease	Collagen Zersetzung	Calcium Metall-Bindung Zink	EC No. 3.4.24.x.; Zersetzung von ECM; Collagen Zersetzung; Zusammensetzung von Collagen Fasern
NREP	1,5573	2,80E-07	Neuronal Regeneration Related Protein	Zytoplasma	-	-	-	Regulation von TGF β Signaling
PLAU	1,5493	9,87E-06	Plasminogen Activator, Urokinase	Sekretiert	Hydrolase Serine Protease	Blood Koagulation Fibrinolyse Hämostase Plasminogen Aktivierung		EC No. 3.4.21.73.; Auflösen eines Fibrin Klumpens
POSTN	3,0609	9,40E-09	Periostin, Osteoblast Specific Factor	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix	Zelladhäsion	Heparin-Bindung		ECM Organisation, Regulation von Notch Signaling, Zelladhäsion, Gewebeentwicklung
PXDN	1,5851	2,84E-05	Peroxidasin Homolog (Drosophila)	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix	Oxidoreduktase Peroxidase	Hydrogenperoxid	Calcium Heme Iron Metall-Bindung	EC No. 1.11.1.7
SCHIP1	1,5408	3,04E-06	Schwannomin interacting protein 1	Zytoplasma	-	-	-	-
SERPINE1	1,6819	5,97E-05	Serpin Peptidase Inhibitor	Sekretiert	Protease Inhibitor Serine Protease Inhibitor	-	-	Auflösung eines Fibrin Klumpens; ECM Proteoglykane; Regulation von Transkription
SERPINH1	1,6014	2,01E-07	Serpin Peptidase Inhibitor Clade H	Sekretiert	Chaperon	Stressantwort	-	Collagen Biosynthese und Modifikation von Enzymen
SPARC	1,6408	3,77E-07	Osteonectin	Sekretiert Extrazellulärer Raum Extrazelluläre Matrix	-	Calcium Kupfer Metall-Bindung	-	ECM Proteoglykane
SPP1	3,2720	2,01E-05	Osteopontin	Sekretiert	Zytokine	Biomineralisation Zelladhäsion	Sialische Säure	Zersetzung der Extrazelluläre Matrix; Integrin Zell-Oberflächen-Interaktionen; Signaling von PDGF
SULF1	1,6696	6,12E-05	Sulfatase1	Zelloberfläche	Hydrolase	Hydrolase	Calcium Metall-Bindung	-
TFRC	1,5982	3,76E-06	Transferrin Receptor (CD71)	Zellmembran Sekretiert	Receptor	Endozytose Host-virus Interaktion	-	Golgi Vesikel Biogenese; Transferrin Endozytose und Recycling

ERGEBNISSE

Tab. 3: Gensignatur des Vergleichs Mvs.PT (Filter: $\log_2FC \geq 1,5$ und p-Wert $< 0,0001$)

Gen	\log_2FC	p-Wert	Protein	Lokalisation	Molekulare Funktion	Biologischer Prozess	Ligand	Reaktom
ADH1B	3,02	3,04E-06	Alkohol Dehydrogenase 1B	Zytoplasma	Oxidoreduktase	Ethanol Oxidation	Metall-Bindung NAD Zink	EC No. 1.1.1.1; Xenobiotisch Metabolischer Prozess
APOE	1,62	2,96E-07	Apolipoprotein E	Sekretiert	Antioxidante Aktivität	Cholesterol-/Lipid-/Steroid-Metabolismus Lipid Transport	Heparin-Bindung	HDL-vermittelter Lipid Transport, Chylomicron-vermittelter Lipid Transport; Retinoid Metabolismus und Transport
ARHGAP15	2,12	1,40E-07	Rho GTPase-activating prot. 15	Zytoplasma Membran	GTPase Aktivierung	Regulation der Zell Form	-	Rho GTPase Zyklus
C7	2,47	2,14E-08	Complement Component C7	Sekretiert	-	Komplement Pfad Zytolyse Angeborene Immunität	-	Regulation der Komplement Kaskade; Terminaler Signalweg von Chemokinen
CCL19	2,34	4,73E-06	C-C motif Chemokine 19	Sekretiert	Zytokin	Chemotaxis Inflammatorische Antwort	-	Chemokine Rezeptoren binden Chemokine, G alpha(i) Signaling Events
CCL2	2,29	5,51E-05	C-C motif Chemokine 2	Sekretiert	Zytokin	Chemotaxis Inflammatorische Antwort	-	Chemokine Rezeptoren binden Chemokine, ATF4 aktivieren Gene
CCL21	2,70	6,60E-07	C-C motif Chemokin 21	Sekretiert	Zytokin	Chemotaxis Inflammatorische Antwort	-	Chemokine Rezeptoren binden Chemokine, G alpha(i) Signaling
CCR7	2,22	2,56E-05	C-C Chemokine Receptor Type 7	Zellmembran	G-protein gekoppelter Rezeptor Transducer	Aktivierung von GTPasen Zelluläre Antwort auf Zytokin Stimulus	-	Chemokine Rezeptoren binden Chemokine, G alpha(i) Signaling Events
CD22	1,52	3,27E-06	B-cell receptor CD22	Zellmembran	Kohlehydrat-Bindung	Zelladhäsion	Lectin	-
CD37	2,13	1,05E-05	Leukocyte Antigen CD37	Membran	-	-	-	Regulation of Zell Proliferation, Regulation der Humoralen Immun Antwort, Regulation of Immunglobulin Produktion
CD48	2,17	1,49E-06	CD48 Antigen	Zellmembran	Rezeptor	Blut Koagulation Leukozyten Migration	-	Zell Oberflächen Interaktionen an der vaskulären Wund
CD52	2,58	3,47E-07	CAMPATH-1 antigen	Zellmembran	-	-	-	Regulation der zytosolischen Calcium Ion Konzentration, Atmung
CD53	1,84	7,22E-07	Leukocyte Antigen CD53	Zell Junction Zellmembran	-	-	-	Regulation der Myoblasten Funktion, Signal Transduktion
CD69	2,10	2,17E-05	Early Activation Antigen CD69	Membran	Rezeptor	Zellular Antwort auf Substanzen	Calcium Lectin	Signal Transduktion

ERGEBNISSE

CD209	1,58	5,31E-07	CD209 antigen	Zellmembran Sekretiert	Rezeptor	Immunität Zelladhäsion Endozytose	Calcium Lectin Mannose-Bindung Metall-Bindung	CD209 (DC-SIGN) Signaling
CD3D	1,91	2,01E-05	T-cell surface glycoprotein CD3 delta chain	Membran	Rezeptor	T-Zell Regulation	-	Immunregulatorische Interaktionen zwischen a Lymphoid und a non-Lymphoid Zell, TCR Signaling, PD- 1 Signaling
CORO1A	2,09	5,86E-06	Coronin-1A	Zytoplasma Zytoskelett Zytoplasmatisch Vesikel Membran	-	Zytoskelett Organisation	Aktin-Bindung Myosin-Bindung PI3K-Bindung	Bewegung der Zelle oder subzellulären Komponenten
CXCL12	1,92	6,23E-05	Stromal Cell-derived Factor 1	Sekretiert	Zytokin Growth Factor	Chemotaxis	-	Nukleäres Signaling von ERBB4, Chemokine Bindung, G alpha (i) Signaling Events
CXCR4	1,97	6,04E-05	C-X-C Chemokine Receptor Type 4	Zell Verbindung Zellmembran	G-protein gekoppelter Rezeptor Transducer	Host-Virus Interaktion	-	Chemokine Bindung, G alpha (i) Signaling Events
CYFIP2	1,57	3,33E-06	Cytoplasmic FMR1- interacting Protein 2	Zell Verbindung Zytoplasma	-	Apoptose Zelladhäsion	-	VEGFA-VEGFR2 Pathway, Rho GTPase aktivieren WASPs und WAVES, Regulation von Aktin Dynamics in Phagozytose
CYTIP	1,65	4,14E-05	Cytohesin- interacting Protein	Zytoplasma Endosom	-	Zelladhäsion	-	Aktivierung von ARFs durch CYTH1
DNASE1L3	1,56	5,83E-05	Deoxyribonuclease Gamma	Nucleus	Endonuklease Hydrolase Nuklease	Apoptose	Calcium	Apoptotische DNA Fragmentierung, Programmierter Zelltod in Entwicklungsprozessen
DOCK2	1,81	8,53E-07	Dedicator of Cytokinesis Protein 2	Zytoplasma Zytoskelett Membran	Guanin-Nukleotid Releasing Faktor	Zytoskelett Organisation Chemotaxis Reg. of GTPase Aktivität	-	Nef und Signal Transduktion, Faktoren involviert in Megakaryozyten Entwicklung und Blutplättchen Produktion
ENPP2	1,90	1,05E-05	Ectonucleotide Pyrophosphatase Family Member 2	Sekretiert	Hydrolase	Chemotaxis Lipid Degradation	Calcium Metall-Bindung Zink	EC No. 3.1.4.39; LPA Produktion in extrazellulären Flüssigkeiten, Modulation der Zellbeweglichkeit
EVI2B	1,54	6,94E-05	Protein EVI2B	Membran	-	-	-	-
FAIM3	1,86	1,49E-05	Fas Apoptotic Inhibitory Molecule 3	Membran Sekretiert	-	Immunität	-	Schutz der Zellen FAS-, TNF alpha- und FADD- induzierter Apoptose
FAM65B	1,73	2,03E-05	Protein FAM65B	Zellmembran Zytoskelett Mitochondrien	Entwicklungsprotein	Differenzierung Hearing Myogenese	-	Zelldifferenzierung, Muskelentwicklung, Sensorische Wahrnehmen von Geräuschen

ERGEBNISSE

GIMAP6	1,58	1,30E-05	GTPase IMAP Family Member 6	Zytoplasma	-	-	GTP-Bindung Nucleotide-Bindung	-
GMFG	2,01	7,13E-08	Glia Maturation Factor Gamma	Intrazellulär	Growth Factor	-	-	Zytoskelett Reorganisation, Regulation der Protein Kinase Aktivität
GPR183	1,69	1,73E-06	G-protein Coupled Receptor 183	Zellmembran	G-protein gekoppelter Rezeptor Transducer	Immunität	-	Immunantwort, Regulation der ERK1 und ERK2 Kaskade
HMHA1	1,53	4,43E-06	Minor Histocompatibility Protein HA-1	Zytoplasma Membran	GTPase Aktivierung	-	Metall-Bindung Zink	Rho GTPase Zyklus
ICAM2	1,77	2,08E-07	Intracellular Adhesion Molecule2	Membran	-	Zelladhäsion	-	Immunoregulatorische Interaktionen zwischen einer Lymphoiden und einer Non-Lymphoiden Zelle, Integrin Interaktionen, CD209 (DC-SIGN) Signaling
IGHM	2,42	1,81E-05	Ig Mu Chain C Region	Zellmembran Sekretiert	Antigen-Bindung	Immunität	-	BCR Aktivierung von Antigenen
IL10RA	1,64	1,54E-07	Interleukin-10 Receptor Subunit Alpha	Membran	Rezeptor	-	-	Zytokin-vermittelter Signaling Pathway, Antwort auf Lipopolysaccharide
IL7R	1,63	3,86E-05	Interleukin-7 Receptor Subunit Alpha	Zellmembran Sekretiert	Rezeptor	-	-	Interleukin-7 Signaling
IRF8	2,08	1,52E-06	Interferon Regulatory Factor 8	Nucleus	-	Transkription Regulation	DNA-Bindung	Interferon Signaling
ITGB2	1,56	6,51E-07	Integrin Beta-2	Membran	Integrin, Rezeptor	Zelladhäsion	Metall-Bindung	Immunoregulatorische Interaktionen zwischen einer Lymphoiden und einer non-Lymphoiden Zell, Integrin Interaktionen, Toll Like Rezeptor 4 (TLR4) Kaskade
ITK	1,98	8,38E-06	Tyrosin-Protein Kinase ITK/TSK	Zytoplasma	Transferase, Tyrosin-Protein Kinase	Immunität	ATP-Bindung Metall-Bindung Nucleotide-Bindung Zink	EC No. 2.7.10.2; Generation eines Second Messenger Moleküls, FCERI vermittelte Ca ²⁺ Mobilisation
ITM2A	1,50	6,82E-05	Integral Membrane Protein 2A	Membran	-	-	-	-
KLF2	1,86	8,83E-08	Krueppel-Like Factor 2	Nucleus	Activator	Transkription Regulation	DNA-Bindung Metall-Bindung Zink	Aktivierung of Beta-Globin Transkription
LAPTM5	1,73	8,94E-07	Lysosomal-associated Transmembrane Protein 5	Lysosom Membran	-	Transport	-	Rolle in Embryogenese und in hämaopoetischen Zellen
LPL	2,04	1,86E-05	Lipoprotein Lipase	Zellmembran	Hydrolase	Lipid Degradation	Heparin-Bindung	EC No. 3.1.1.34; Retinoid Metabolismus und

ERGEBNISSE

				Sekretiert					
LRMP	1,62	7,55E-05	Lymphoid-restricted Membrane Protein	Zytoplasma	-	Fertilisation Immunität	-		Transport, Transkriptionelle Regulation der weißen Adipozytendifferenzierung, Chylomicron-vermittelter Lipid Transport
LTB	2,38	8,21E-07	Leukotriene B4 Receptor 2	Zellmembran	G-protein gekoppelter Rezeptor, Transducer	Chemotaxis	-		Lieferung von Peptiden zu MHC Class I Molekülen, Geschmackssignal Transduktion Via ITPR3, Pronukleus Kongression und Fusion
LYVE1	1,67	8,50E-06	Lymphatic Vessel Endothelial Hyaluronic Acid Receptor 1	Membran	Rezeptor	Transport	-		G alpha (q) Signaling Events, Leukotriene Rezeptoren
MS4A1	2,80	8,83E-05	B-Lymphocyte CD20	Zellmembran	-	B-Zell Aktivierung	-		Hyaluron Uptake und Degradation
PAX5	2,21	7,71E-05	Paired Box Protein Pax-5	Nucleus	Entwicklungsprotein	Differenzierung Neurogenesis Transkription Regulation	DNA-Bindung		Regulation of B-Zell Aktivierung und Proliferation
PLIN1	1,88	6,42E-05	Perilipin-1	Endoplasmatisches Retikulum Lipid Tropfen	-	Lipid Metabolismus	-		Immunantwort, Nervensystementwicklung, Spermatogenese, Skelettmuskel Zelldifferenzierung, Entwicklung eines Multizellulärer Organismus
PRKCB	1,97	1,19E-05	Proteinkinase C Beta Type	Zytoplasma Membran Nucleus	Chromatin Regulator Serine/Threonin-Protein Kinase Transferase	Immunität Apoptose Transkription Regulation	ATP-Bindung Calcium Metall-Bindung Nucleotide-Bindung Zink		Transkriptionelle Regulation der weißen Adipozytendifferenzierung, Hormone-sensitive Lipase (HSL)-vermittelte Traclyglycerol Hydrolyse EC No. 2.7.11.13; Aktivierung of NF-kappaB in B Zellen, G alpha (z) Signaling Events, Depolymerisation der Nukleären Lamina, VEGFR2-vermittelte Zellproliferation
PTGDS	1,89	1,15E-06	Prostaglandin-H2 D-Isomerase	Zytoplasma Sekretiert	Isomerase	Fettsäure-metabolismus Lipid Biosynthesis Prostaglandin Biosynthesis Transport	-		Synthesis von Prostaglandinen (PG) und Thromboxanen (TX)
PTPRC	2,20	8,53E-06	Receptor-type Tyrosine-protein Phosphatase C	Membran	Hydrolase Protein Phosphatase	Immunität	-		Phosphorylierung von CD3 und TCR Zeta Ketten, Semaphorin Interaktionen
RASSF2	1,61	2,47E-05	Ras-association Domain-containing Protein 2	Zytoplasma Nucleus	-	Zell Zyklus	-		Knochen Remodellierung, EGFR Signaling via NF-kappaB Kaskade, Ossifikation, Entwicklung des Skelettsystems
SELL	2,41	1,37E-05	L-Selectin	Membran	-	Zelladhäsion	Lectin		Immunregulatorische Interaktionen zwischen einer Lymphoiden und a non-Lymphoiden Zelle;

ERGEBNISSE

SEPT6	1,71	7,68E-06	Septin-6	Zytoplasma	-	Zell Zyklus, Zell Division	GTP-Bindung Nucleotide-Bindung	Zytokinese, Virale Prozesse
SH2D1A	1,55	4,48E-06	SH2 Domain-containing Prot. 1A	Zytoplasma	-	Immunität	-	Inhibitor der SLAM Selbstassoziierung, Mediation der Interaktion zwischen FYN und SLAMF1
SLC2A3	1,80	1,46E-05	Solute Carrier Family 2, Facilitated Glucose Transporter Member 3	Membran	-	Zucker Transport	-	Na+-unabhängiger Glucose Transporter
SRGN	1,81	3,05E-05	Serglycin	Sekretiert	-	Apoptose Biominalisation	-	Blutplättchen Degranulation
STAP1	1,65	8,24E-05	Signal-Transducing Adaptor Protein 1	Zytoplasma Nucleus	-	Immunität	-	BCR Aktivierung von Antigenen
TGFBR2	1,52	2,43E-05	TGF-beta Receptor Type-2	Zellmembran	Rezeptor Serine/Threonin-Protein Kinase Transferase	Apoptose Differenzierung Wachstumsregulation	ATP-Bindung Magnesium Metall-Bindung Nucleotide-Bindung	EC No. 2.7.10.2; TGF-beta Receptor Signaling in EMT, SMAD2/3 Mutanten in Krebs, TGFBR Mutanten in Krebs;
TRAC	1,78	2,00E-05	T Cell Receptor Alpha Constant	Membran	Rezeptor	Immunität	-	Immunoregulatorische Interaktionen zwischen einer Lymphoiden und einer Non-Lymphoiden Zelle, TCR Signaling, PD-1 Signaling
TRBC1	2,24	1,85E-05	T-Cell Receptor Beta-1 Chain C Region	Membran	Rezeptor	Immunität	-	Immunoregulatorische Interaktionen zwischen einer Lymphoiden und einer Non-Lymphoiden Zelle, TCR Signaling, PD-1 Signaling
YME1L1	1,65	8,37E-06	ATP-dependant Zinc Metalloprotease YME1L1	Mitochondriale Innere Membran	Hydrolase Metalloprotease	Mitochondriales Protein Metabolismus	ATP-Bindung Metall-Bindung Nucleotide-Bindung Zink	Erhalt der Cristae Morphologie und Komplex I Respirationsaktivität, Schutz vor Akkumulation vor oxidativ geschädigten Membranproteinen

Die Gensignatur für Mvs.PT (**Tab. 3**) enthält mehrere Gene, die bei Immunantwort aktiviert werden: Complement Component C7 (C7), C-C motif Chemokine (CCL19, CCL2, CCL21) und ein C-C Chemokine Rezeptor (CCR7), Interleukine (IL10RA, IL8R), Interferon Regulatory Factor 8 (IRF8), Krueppel-Like Factor 2 (KLF2), Lymphatic Vessel Endothelial Hyaluronic Acid Receptor 1 (LYVE1), Lymphoid-restricted Membrane Protein (LRMP), Paired Box Protein Pax-5 (PAX5), Proteinkinase C (PRKCB), Receptor-type Tyrosin-protein Phosphatase C (PTPRC), Septin-6 (SEPT6), SH2 Domain-containing Prot. 1A (SH2D1A), Serglycin (SRGN), Signal-Transducing Adaptor Protein 1 (STAP1), Tyrosin-Protein Kinase ITK/TSK, sowie mehrere Membranproteine (CD22, CD37, CD53, CD69, CD209, CD3D, FAIM3, GPR183, ICAM2, IGHM, ITGB2, LTB, MS4A1, TRAC, TRBC1).

GTPasen scheinen in Metastasen ebenfalls eine große Rolle zu spielen. Mehrere Gene, die mit GTPasen interagieren oder diese regulieren, sind Teil der Gensignatur für Mvs.PT: Rho GTPase-activating Protein 15 (ARHGAP15), Stromal Cell-derived Factor 1 (CXCL12), C-X-C Chemokine Receptor Type 4 (CXCR4), Cytoplasmic FMR1-interacting Protein 2 (CYFIP2), Deducator of Cytokinesis Protein 2 (DOCK2), Protein FAM65B (FAM65B), GTPase IMAP Family Member 6 (GIMAP6) und Minor Histocompatibility Protein HA-1 (HMHA1).

Einige Gene, die in den zellulären Metabolismus involviert sind, waren ebenfalls hochreguliert: der Zuckertransporter SLC2A3, die mitochondriale ATP-abhängige Zink Metalloprotease YME1L1, Apolipoprotein E (APOE), das CAMPATH-1 Antigen (CD52), Lipoprotein Lipase (LPL), Perilipin-1 (PLIN1) und Prostaglandin-H2 D-Isomerase (PTGDS), und die Oxidoreduktase Alkohol-Dehydrogenase 1B (ADH1B), die in den xenobiotischen metabolischen Prozess der Alkoholoxidation involviert ist.

Zellmobilitäts- oder Zelladhäsionsproteine waren durch Coronin-1A (CORO1A), CD48 Antigen, Cytohesin-interacting Protein (CYTIP), Ectonucleotide Pyrophosphatase Family Member 2 (ENPP2), Glia Maturation Factor Gamma (GMFG) und L-Selectin (SELL) vertreten.

Nicht weiter in eine größere Gruppe einzuordnende Proteine aus der Gensignatur von Mvs.PT waren Deoxyribonuclease Gamma (DNASE1L3), EVI2B, Integral Membrane Protein 2A (ITM2A), Lysosomal-associated Transmembrane Protein 5 (LAPTM5), Ras-association Domain-containing Protein 2 (RASSF2) und TGF-beta Rezeptor Type-2 (TGFB2).

Auch mit dem Wissen über die häufig vorkommenden Protein-Gruppen sind komplexe „Omics“-Daten nur schwer zu interpretieren und in einen größeren Zusammenhang zu setzen. Zur weiteren Auswertung der Daten werden entsprechende bioinformatische Programme benötigt, die in der Lage sind, das bereits vorhandene molekularbiologische Wissen über Signalwege, Upstream/Downstream-Faktoren und andere Zusammenhänge mit den Daten zu vernetzen und damit zur Interpretation der Daten nutzbar zu machen. Nach ersten Vorversuchen mit der Software *GeneSpring*[®] (Agilent) und *Ingenuity*[®] (Qiagen) wurde die Software *Ingenuity*[®] zur weiteren Bearbeitung der Daten gewählt.

Von vordringlichem Interesse waren die Signalwege und Netzwerke, welche durch die Signatur-Gene in PTvs.N und Mvs.PT aktiviert werden. Das Werkzeug *Ingenuity Pathway Analysis*[®] (IPA) ist ein Programm innerhalb von *Ingenuity*[®], in welchem detaillierte Informationen aus aktuellen Publikationen von Wissenschaftlern in eine Datenbank übertragen wurden. Durch unterschiedliche Visualisierungen können die zu interpretierenden Daten so in Zusammenhang mit molekularbiologischen Netzwerken gebracht werden. Die vollständigen Genlisten wurden folglich in *Ingenuity*[®] eingespeist und der gleiche Filter eingestellt, welcher auch für die Heatmaps verwendet wurde, diesmal auch unter Einbeziehung der herabregulierten Gene ($\log_2FC \geq |1,5|$ und p-Wert < 0,0001).

Tab. 4: Top Netzwerke PTvs.N (*Ingenuity*[®] Pathway Analysis)

ID	Score	Aktivierte Gene	Wahrscheinlichste Krankheiten und Funktionen
1	57	27	Zelluläre Entwicklung, Infektionskrankheit, Krebs
2	47	24	Erkrankung des Bindegewebes, Dermatologische Erkrankungen, Zelluläre Anordnung und Organisation
3	34	19	Molekularer Transport, Zelluläres Wachstum und Proliferation, Embryonale Entwicklung
4	31	17	Embryonale Entwicklung, Entwicklung des Organismus, Gewebeentwicklung
5	28	16	Entzündliche Erkrankung, Organverletzung und –anomalien, Renale und urologische Erkrankungen
6	26	15	Dermatologische Erkrankungen, Entzündliche Antwort, Gastrointestinalerkrankungen
7	19	12	Verletzungen und Anomalien des Organismus, Zelluläre Bewegung, Entwicklung und Funktion des hämatologischen Systems

ERGEBNISSE

ID	Score	Aktivierte Gene	Wahrscheinlichste Krankheiten und Funktionen
8	17	12	Zelluläre Funktion und Erhaltung, "Small Molecule" Biochemie, Aminosäure-Metabolismus
9	13	9	Zelluläre Anordnung und Organisation, Zelluläre Funktion und Erhaltung, Zelltod und -überleben
10	2	1	Krebs, Gastrointestinale Erkrankungen, Systemische hepatische Erkrankungen

Für den Datensatz PTvs.N wurden über IPA mehrere krankheits-relevante Netzwerke für Krebs und Erkrankungen des Bindegewebes als am wahrscheinlichsten (*Score* > 20) angezeigt, welche ausschließlich über die Gensignatur identifiziert werden konnten (**Tab. 4**). Der *Score* ist dabei ein Wert für die Wahrscheinlichkeit, dass diese Netzwerke auf Grund dieser Expressionsveränderung angeschaltet wurden. Für Mvs.PT wurden Netzwerke für die Immunantwort, Zellmobilität und Zell-Zell-Signalwege und –Interaktionen mit einem *Score* höher als 20 bewertet (**Tab. 5**).

Tab. 5: Top Netzwerke Mvs.PT (Ingenuity® Pathway Analysis)

ID	Score	Aktivierte Gene	Wahrscheinlichste Krankheiten und Funktionen
1	38	18	Humorale Immunantwort, Proteinbiosynthese, Entwicklung und Funktion des hämatologischen Systems
2	27	14	Zelluläre Bewegung, Entwicklung und Funktion des hämatologischen Systems, Immunzell-Transport
3	25	14	Zell-Zell-Signalwege und –Interaktionen, Entwicklung und Funktion des hämatologischen Systems, Gewebemorphologie
4	19	11	Zellzyklus, Zellmorphologie, Zelluläre Beeinträchtigung
5	12	8	Krebs, Dermatologische Erkrankungen, Gastrointestinale Erkrankungen
6	12	8	Lipidmetabolismus, "Small Molecule" Biochemie, Kardiovaskuläre Erkrankungen
7	11	7	Krebs, Erkrankungen des endokrinen Systems, Entwicklungsstörungen
8	2	1	Kohlehydratmetabolismus, Zelluläre Anordnung und Organisation, Zelluläre Bewegung
9	2	1	Zellmorphologie, Zelluläre Anordnung und Organisation, Zelluläre Beeinträchtigung

Anschließend wurden sowohl die Top-Netzwerke visualisiert als auch diejenigen, die mehr als ein aktiviertes Fokusbmolekül aufwiesen (Abb. 27 bis Abb. 42). Die Ergebnisse werden in der Diskussion in den Kontext bisheriger Publikationen gesetzt.

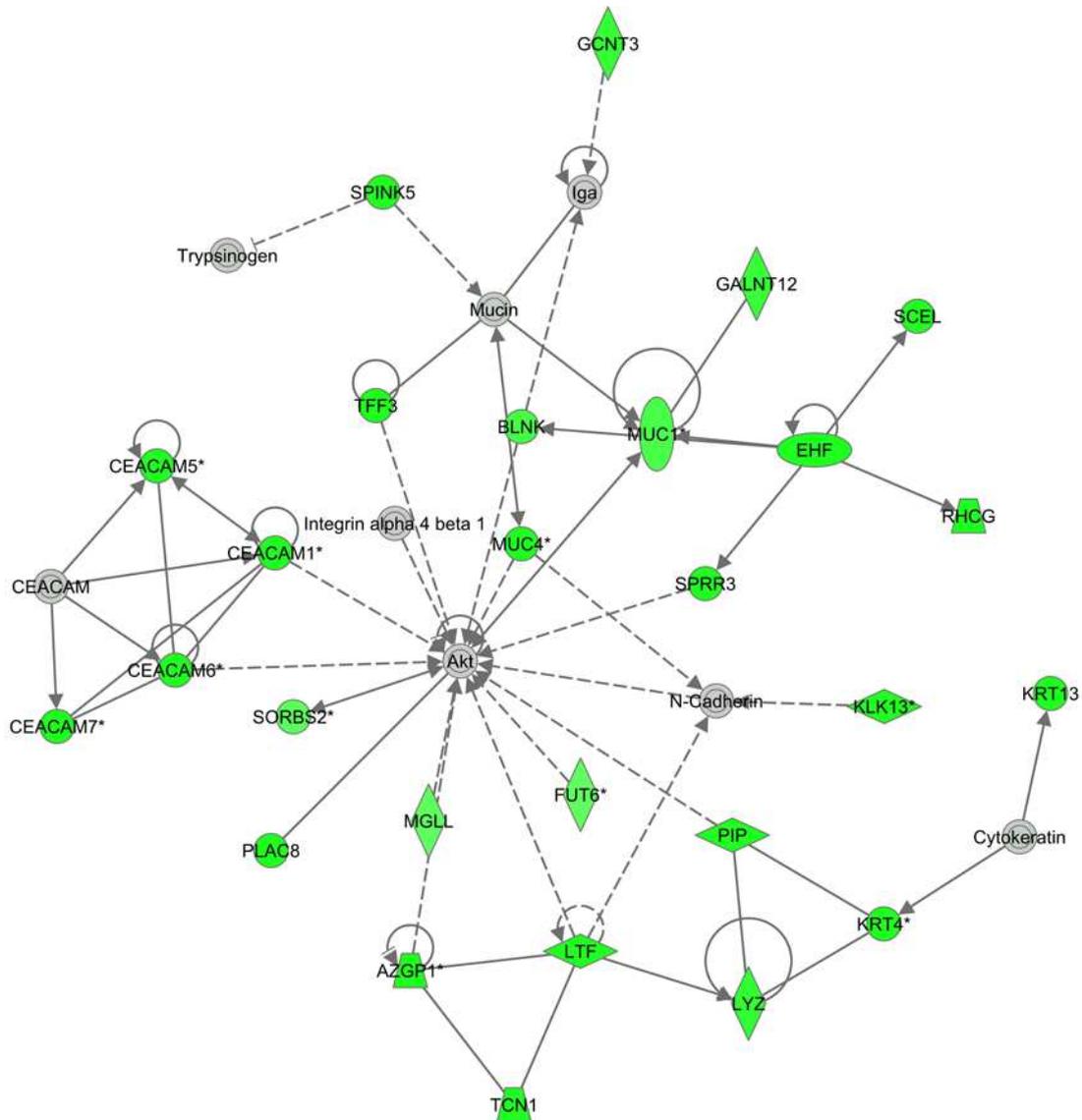


Abb. 27: Top 01-Netzwerk (*Ingenuity Pathway Analysis*[®]) differenziell exprimierter Gene in PTvs.N.

In diesem Netzwerk, das mit zellulärer Entwicklung, Infektionskrankheiten und Krebs assoziiert wird, sind 27 Gene signifikant überexprimiert (bei dem benutzten Filter: $\log_2FC \geq |1,5|$ und $p\text{-Wert} < 0,0001$).

Die Symbolfarbe visualisiert hochregulierte (rote), herabregulierte (grüne) und nicht-signifikant veränderte (graue) Gene. Linien und Pfeile zwischen den Symbolen repräsentieren direkte (durchgehende Linien) und indirekte (gestrichelte Linien) Interaktionen zwischen den Proteinen. Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

ERGEBNISSE

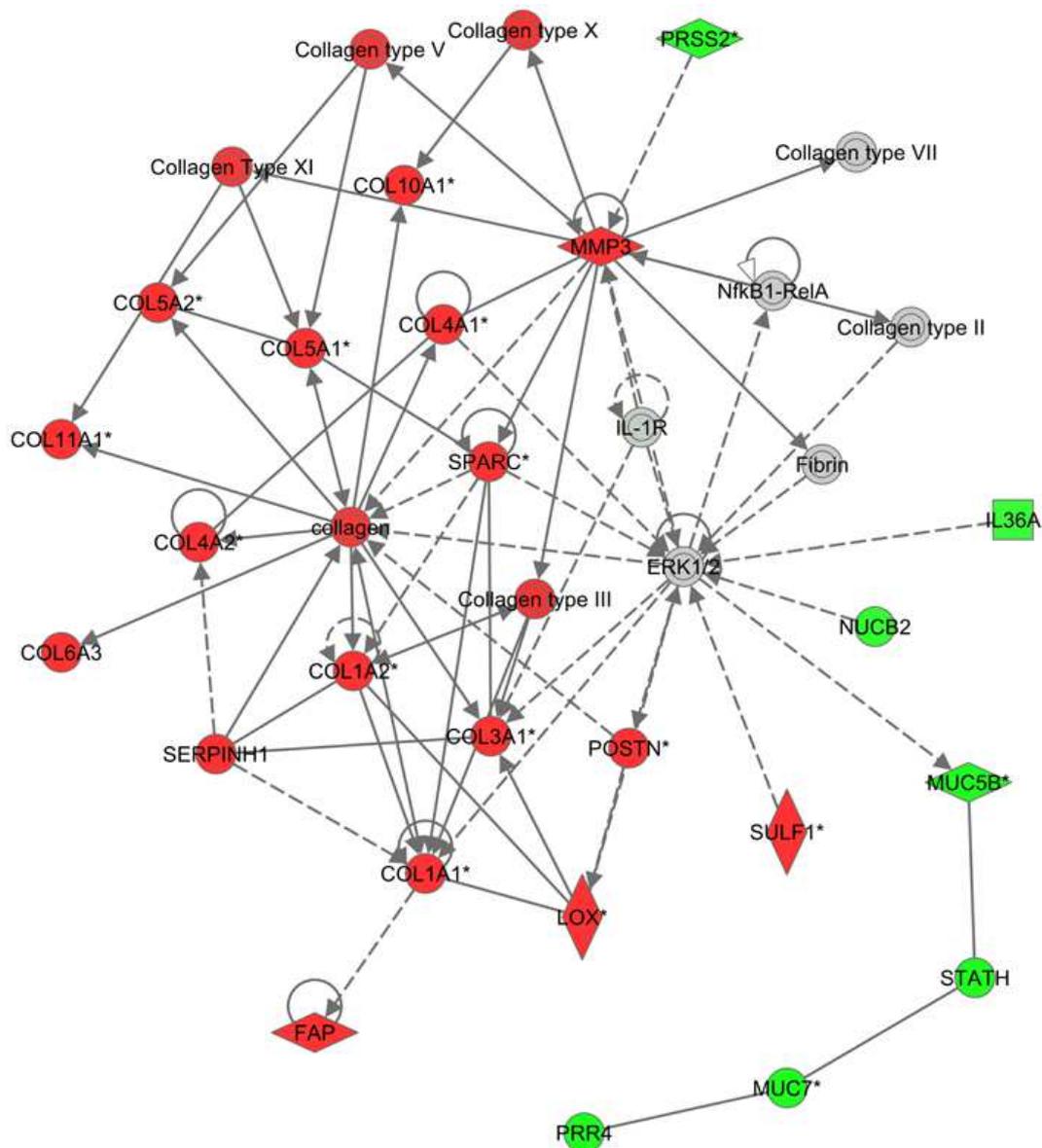


Abb. 28: Top 02-Netzwerk (*Ingenuity Pathway Analysis*[®]) differenziell exprimierter Gene in PTvs.N.

In diesem Netzwerk, das mit Krankheiten des Bindegewebes, dermatologischen Erkrankungen und zellulärer Anordnung und Organisation assoziiert wird, sind 24 Gene signifikant überexprimiert (bei dem benutzten Filter: $\log_2FC \geq |1,5|$ und $p\text{-Wert} < 0,0001$).

Die Symbolfarbe visualisiert hochregulierte (rote), herabregulierte (grüne) und nicht-signifikant veränderte (graue) Gene. Linien und Pfeile zwischen den Symbolen repräsentieren direkte (durchgehende Linien) und indirekte (gestrichelte Linien) Interaktionen zwischen den Proteinen. Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

ERGEBNISSE

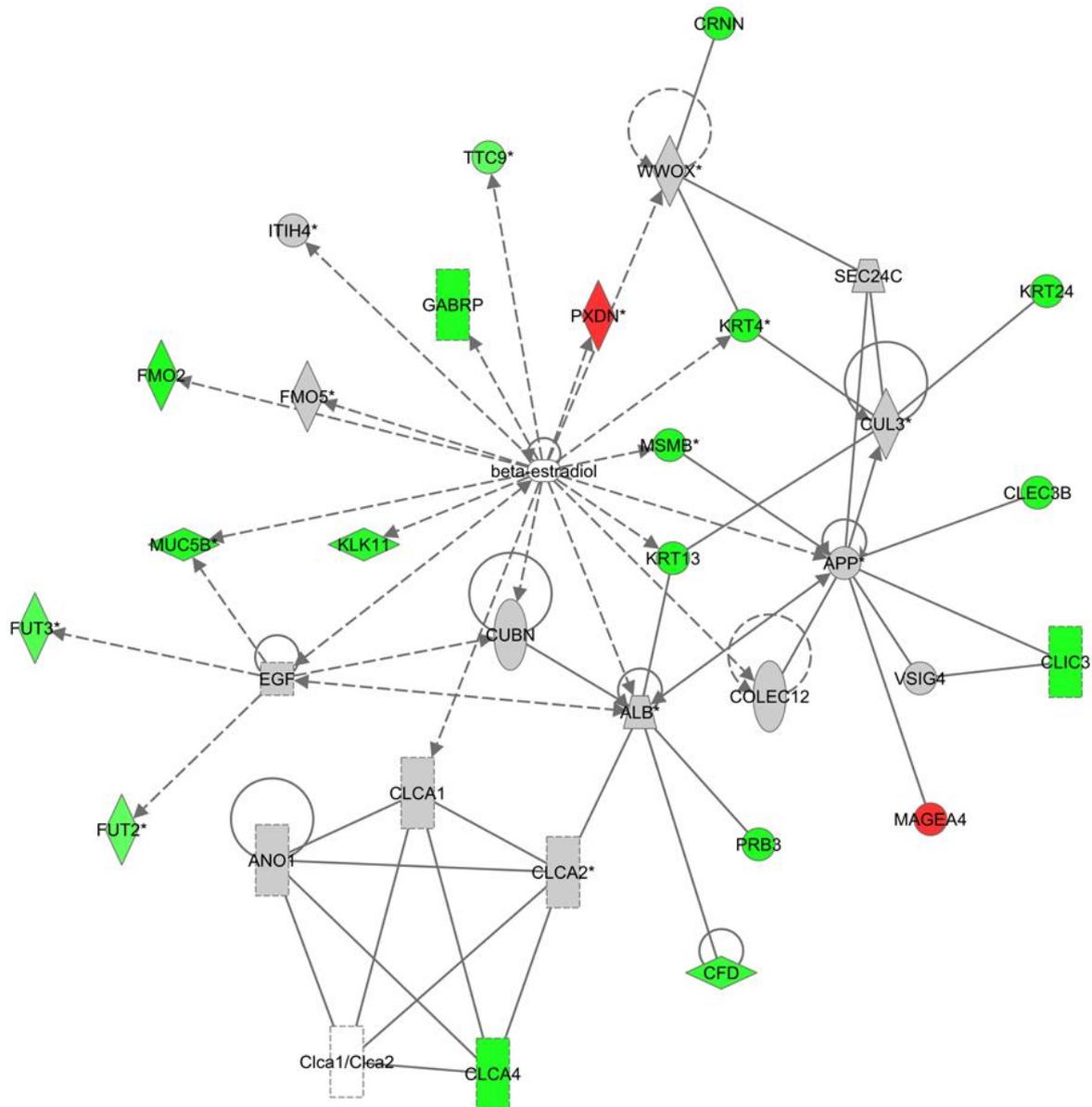


Abb. 29: Top 03-Netzwerk (*Ingenuity Pathway Analysis*[®]) differenziell exprimierter Gene in PTvs.N.

In diesem Netzwerk, das mit molekularem Transport, zellulärem Wachstum, Proliferation und embryonaler Entwicklung assoziiert wird, sind 19 Gene signifikant überexprimiert (bei dem benutzten Filter: $\log_2FC \geq |1,5|$ und $p\text{-Wert} < 0,0001$). Die Symbolfarbe visualisiert hochregulierte (rote), herabregulierte (grüne) und nicht-signifikant veränderte (graue) Gene. Linien und Pfeile zwischen den Symbolen repräsentieren direkte (durchgehende Linien) und indirekte (gestrichelte Linien) Interaktionen zwischen den Proteinen. Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

ERGEBNISSE

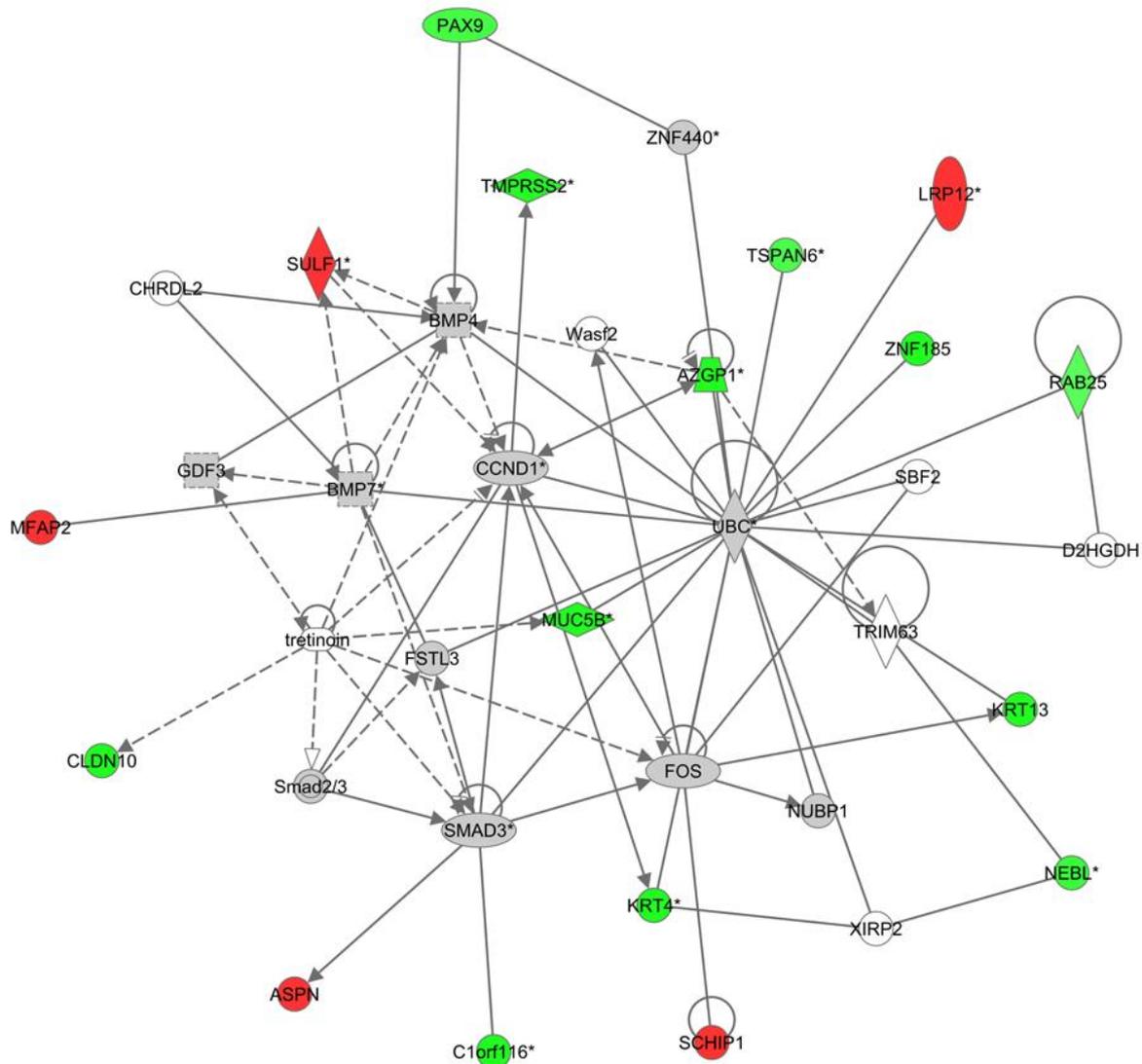


Abb. 30: Top 04-Netzwerk (*Ingenuity Pathway Analysis*[®]) differenziell exprimierter Gene in PTvs.N.

In diesem Netzwerk, das mit embryonaler Entwicklung, Entwicklung des Organismus und Gewebeerwicklung assoziiert wird, sind 17 Gene signifikant überexprimiert (bei dem benutzten Filter: $\log_2FC \geq |1,5|$ und $p\text{-Wert} < 0,0001$).

Die Symbolfarbe visualisiert hochregulierte (rote), herabregulierte (grüne) und nicht-signifikant veränderte (graue) Gene. Linien und Pfeile zwischen den Symbolen repräsentieren direkte (durchgehende Linien) und indirekte (gestrichelte Linien) Interaktionen zwischen den Proteinen. Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

ERGEBNISSE

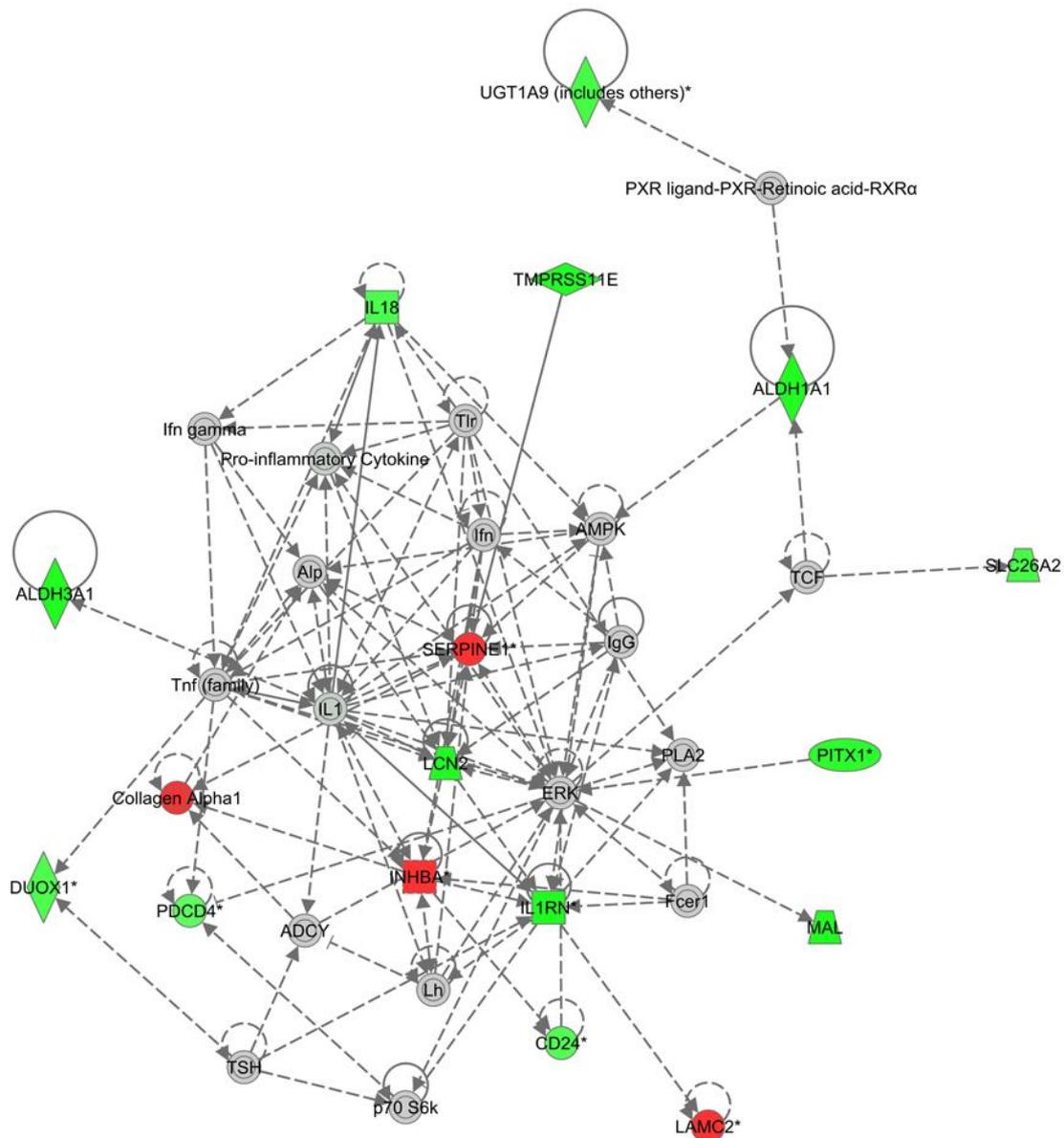


Abb. 31: Top 05-Netzwerk (*Ingenuity Pathway Analysis*[®]) differenziell exprimierter Gene in PTvs.N.

In diesem Netzwerk, das mit entzündlicher Erkrankung, Organverletzung und -anomalien, renalen und urologischen Erkrankungen assoziiert wird, sind 16 Gene signifikant überexprimiert (bei dem benutzten Filter: $\log_2FC \geq |1,5|$ und p-Wert $< 0,0001$).

Die Symbolfarbe visualisiert hochregulierte (rote), herabregulierte (grüne) und nicht-signifikant veränderte (graue) Gene. Linien und Pfeile zwischen den Symbolen repräsentieren direkte (durchgehende Linien) und indirekte (gestrichelte Linien) Interaktionen zwischen den Proteinen. Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

ERGEBNISSE

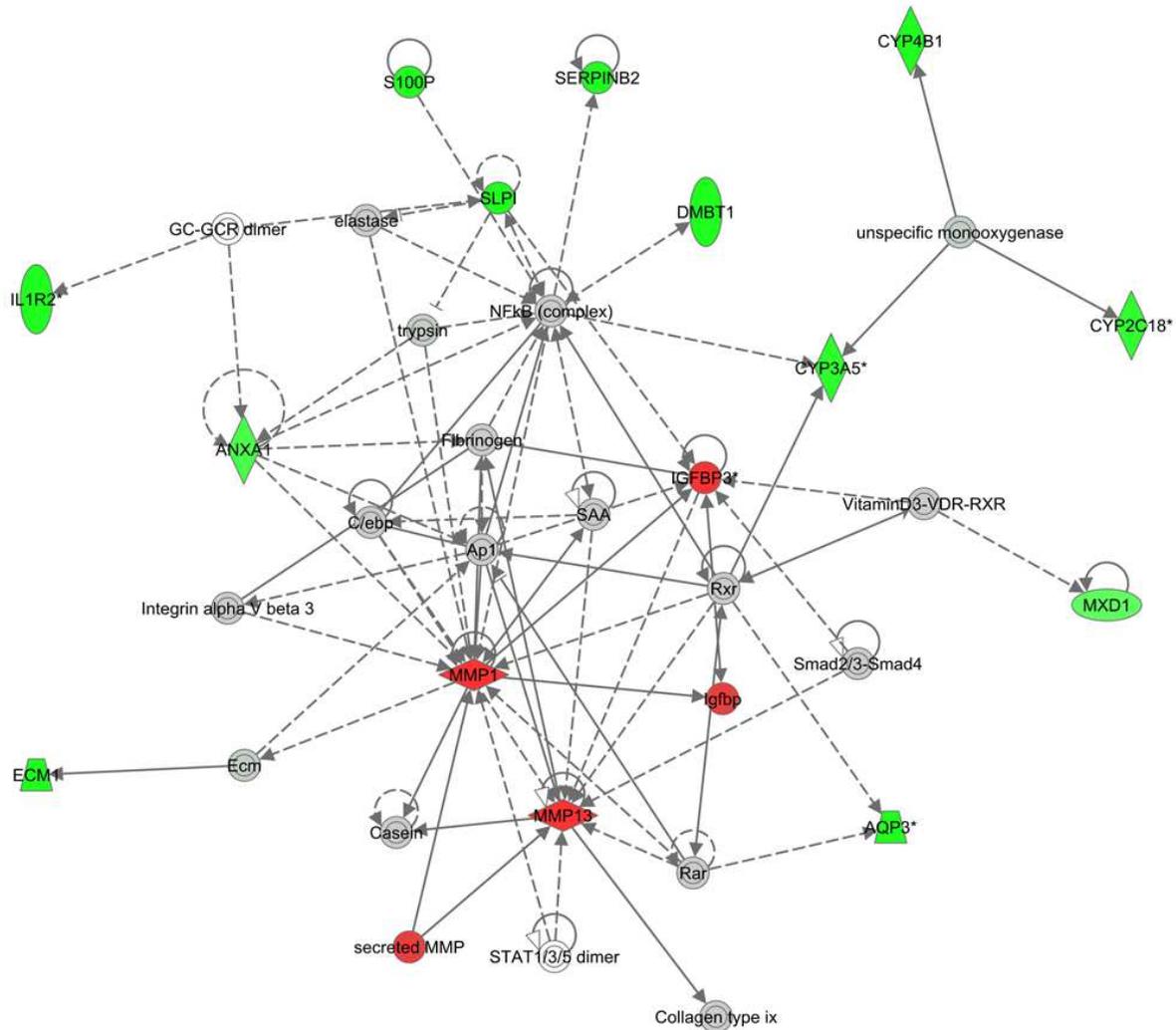


Abb. 32: Top 06-Netzwerk (*Ingenuity Pathway Analysis*[®]) differenziell exprimierter Gene in PTvs.N.

In diesem Netzwerk, das mit dermatologischen Erkrankungen, entzündlicher Antwort und gastrointestinalen Erkrankungen assoziiert wird, sind 15 Gene signifikant überexprimiert (bei dem benutzten Filter: $\log_2FC \geq |1,5|$ und $p\text{-Wert} < 0,0001$).

Die Symbolfarbe visualisiert hochregulierte (rote), herabregulierte (grüne) und nicht-signifikant veränderte (graue) Gene. Linien und Pfeile zwischen den Symbolen repräsentieren direkte (durchgehende Linien) und indirekte (gestrichelte Linien) Interaktionen zwischen den Proteinen. Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

ERGEBNISSE

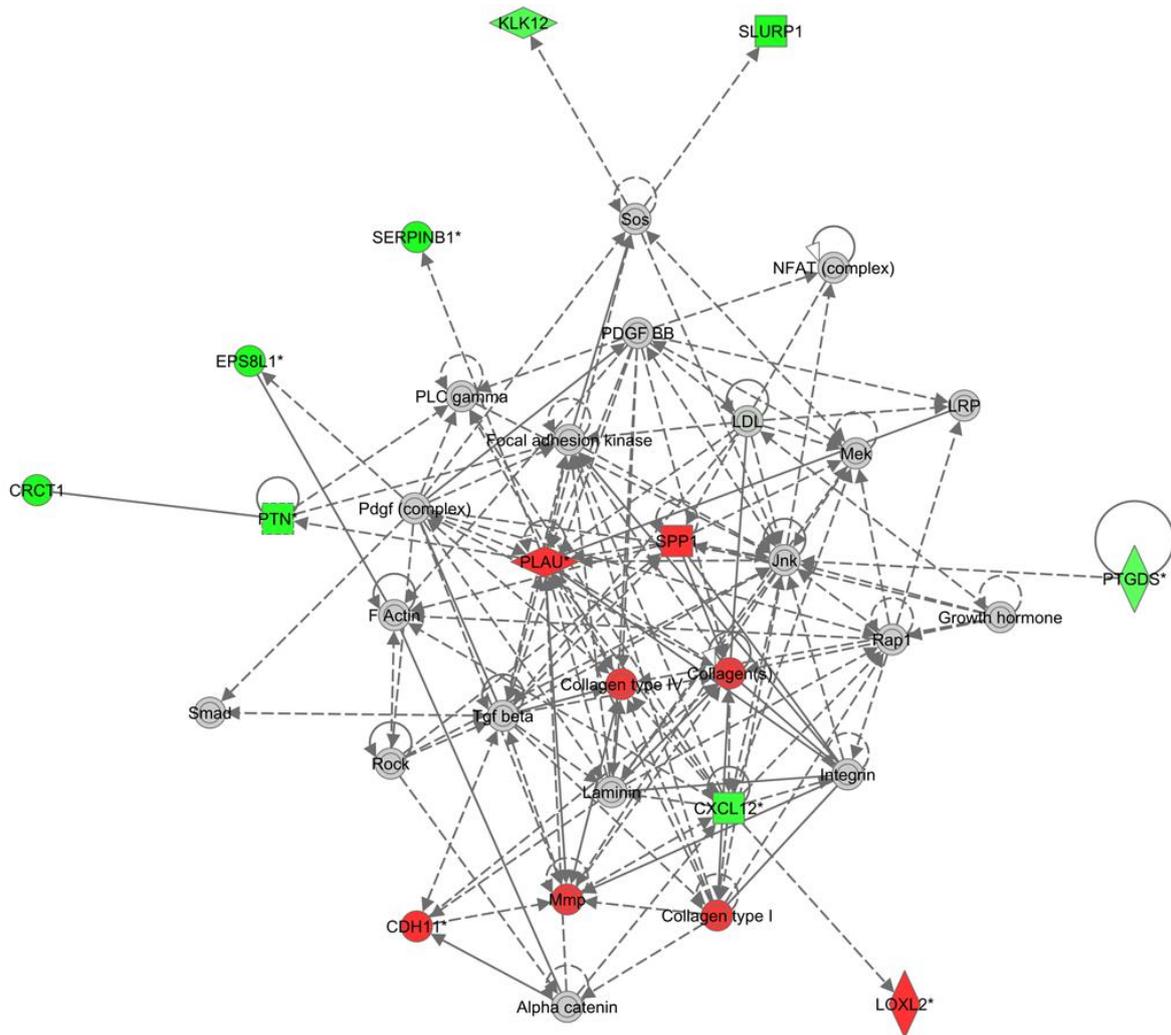


Abb. 33: Top 07-Netzwerk (*Ingenuity Pathway Analysis*[®]) differenziell exprimierter Gene in PTvs.N.

In diesem Netzwerk, das mit Verletzungen und Anomalien des Organismus, zellulärer Bewegung sowie der Entwicklung und Funktion des hämatologischen Systems assoziiert wird, sind 12 Gene signifikant überexprimiert (bei dem benutzten Filter: $\log_2FC \geq |1,5|$ und $p\text{-Wert} < 0,0001$).

Die Symbolfarbe visualisiert hochregulierte (rote), herabregulierte (grüne) und nicht-signifikant veränderte (graue) Gene. Linien und Pfeile zwischen den Symbolen repräsentieren direkte (durchgehende Linien) und indirekte (gestrichelte Linien) Interaktionen zwischen den Proteinen. Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

ERGEBNISSE

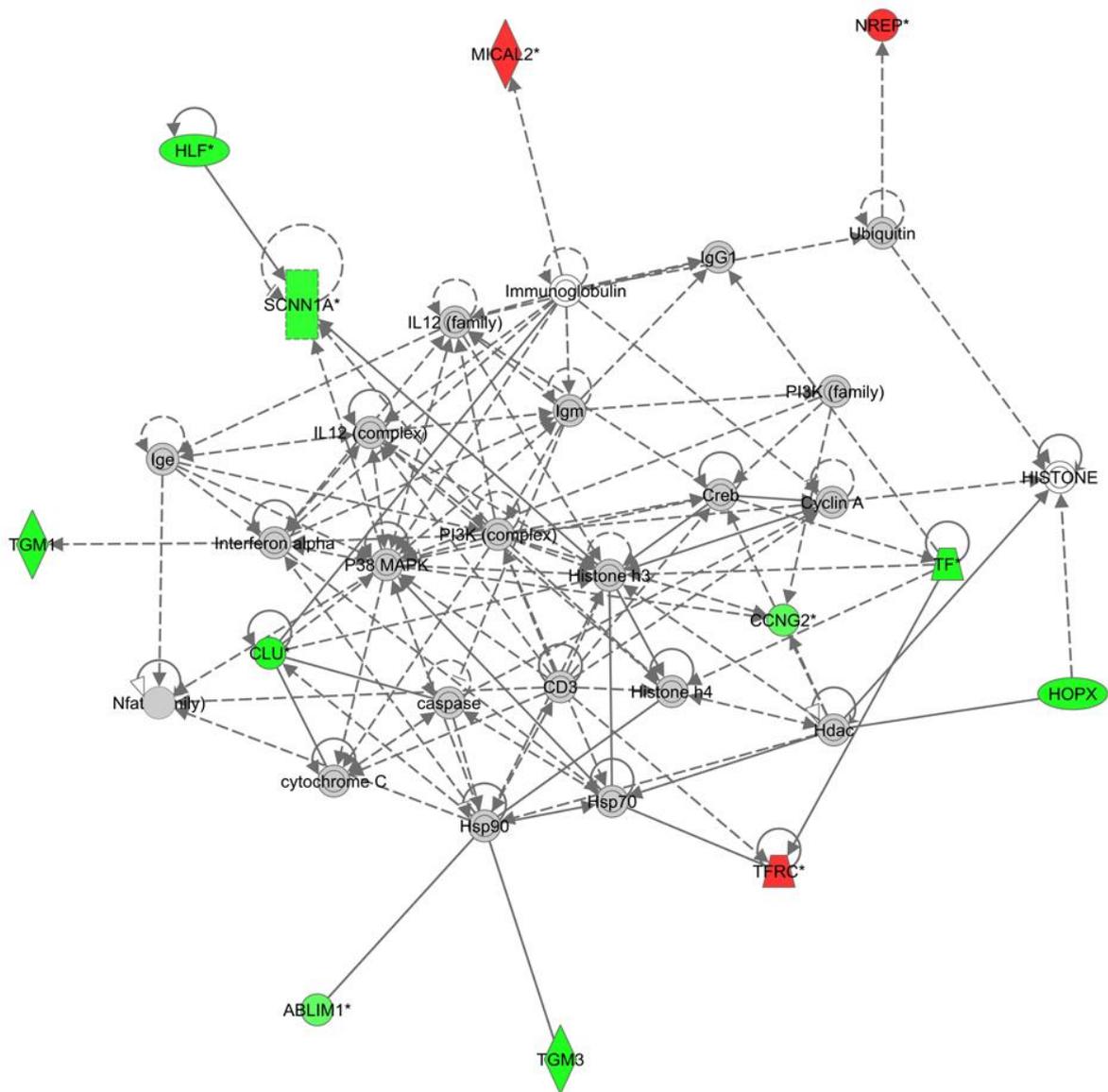


Abb. 34: Top 08-Netzwerk (*Ingenuity Pathway Analysis*[®]) differenziell exprimierter Gene in PTvs.N.

In diesem Netzwerk, das mit zellulärer Funktion und Erhaltung, „Small Molecule“ Biochemie und dem Aminosäure-Metabolismus assoziiert wird, sind 12 Gene signifikant überexprimiert (bei dem benutzten Filter: $\log_2FC \geq |1,5|$ und $p\text{-Wert} < 0,0001$).

Die Symbolfarbe visualisiert hochregulierte (rote), herabregulierte (grüne) und nicht-signifikant veränderte (graue) Gene. Linien und Pfeile zwischen den Symbolen repräsentieren direkte (durchgehende Linien) und indirekte (gestrichelte Linien) Interaktionen zwischen den Proteinen. Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

ERGEBNISSE

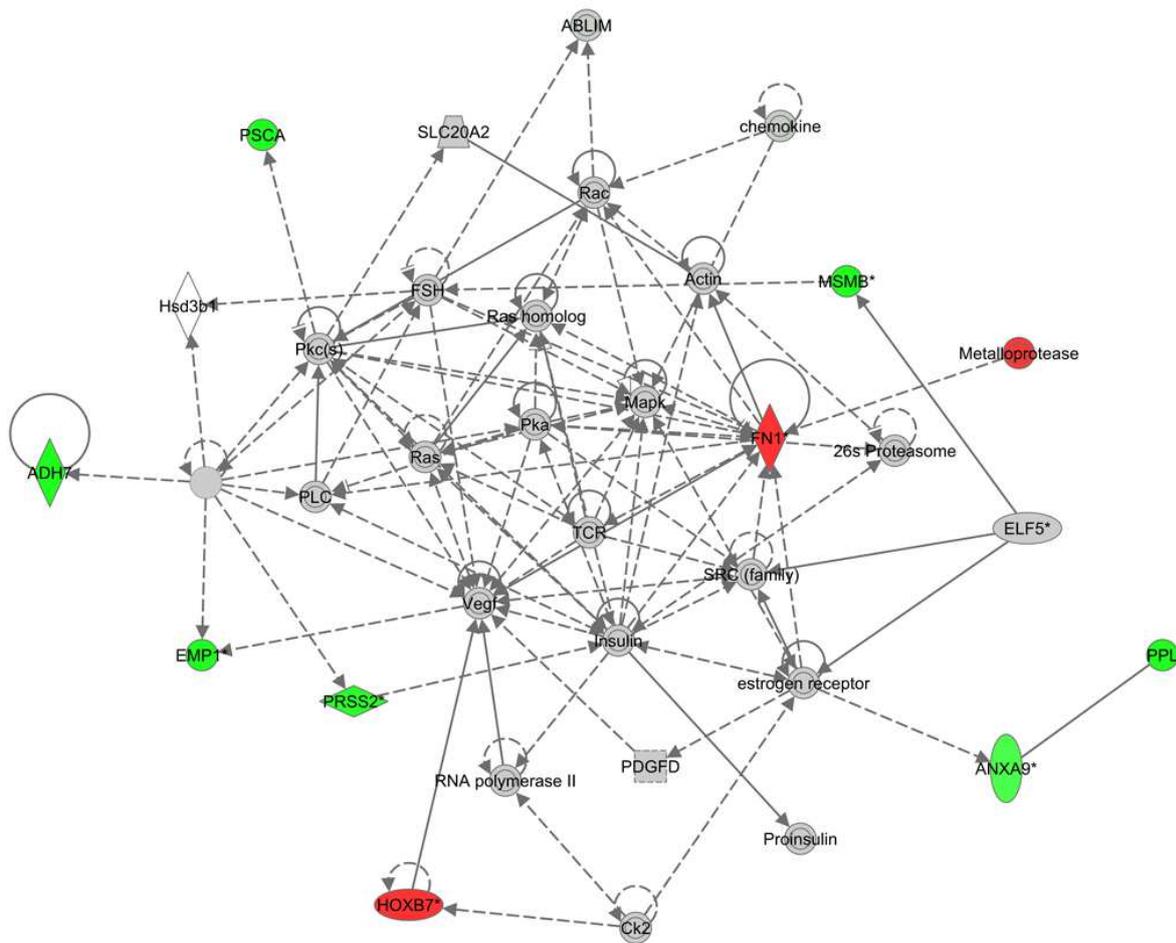


Abb. 35: Top 09-Netzwerk (*Ingenuity Pathway Analysis*[®]) differenziell exprimierter Gene in PTvs.N.

In diesem Netzwerk, das mit zellulärer Anordnung und Organisation, zellulärer Funktion und Erhaltung sowie Zelltod und -überleben assoziiert wird, sind neun Gene signifikant überexprimiert (bei dem benutzten Filter: $\log_2FC \geq |1,5|$ und p-Wert < 0,0001).

Die Symbolfarbe visualisiert hochregulierte (rote), herabregulierte (grüne) und nicht-signifikant veränderte (graue) Gene. Linien und Pfeile zwischen den Symbolen repräsentieren direkte (durchgehende Linien) und indirekte (gestrichelte Linien) Interaktionen zwischen den Proteinen. Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

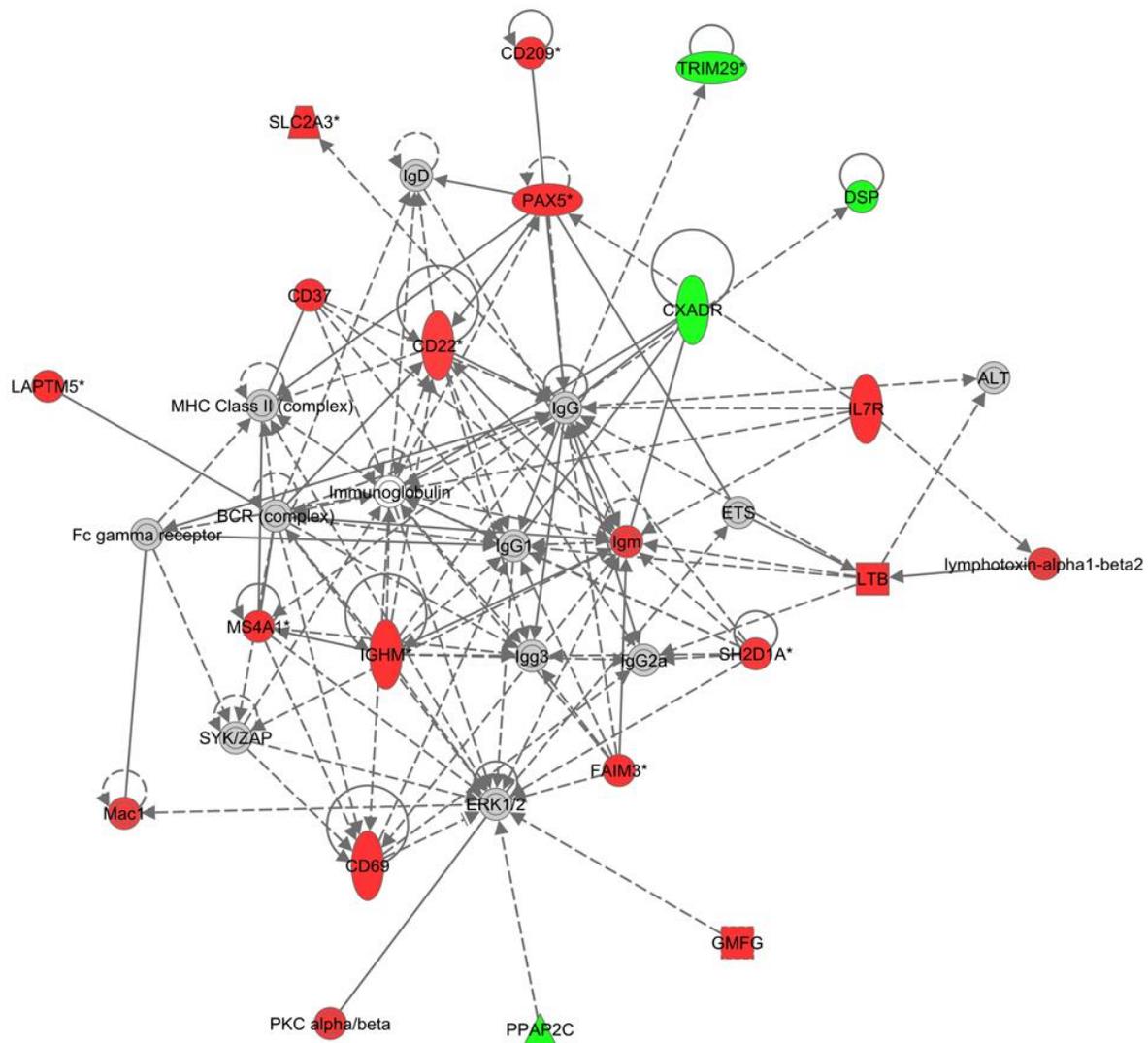


Abb. 36: Top 01-Netzwerk (*Ingenuity Pathway Analysis*[®]) differenziell exprimierter Gene in Mvs.PT.

In diesem Netzwerk, das mit humoraler Immunantwort, Proteinbiosynthese, Entwicklung und Funktion des hämatologischen Systems assoziiert wird, waren 18 Gene signifikant überexprimiert (bei dem benutzten Filter: $\log_2FC \geq |1,5|$ und $p\text{-Wert} < 0,0001$).

Die Symbolfarbe visualisiert hochregulierte (rote), herabregulierte (grüne) und nicht-signifikant veränderte (graue) Gene. Linien und Pfeile zwischen den Symbolen repräsentieren direkte (durchgehende Linien) und indirekte (gestrichelte Linien) Interaktionen zwischen den Proteinen. Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

ERGEBNISSE

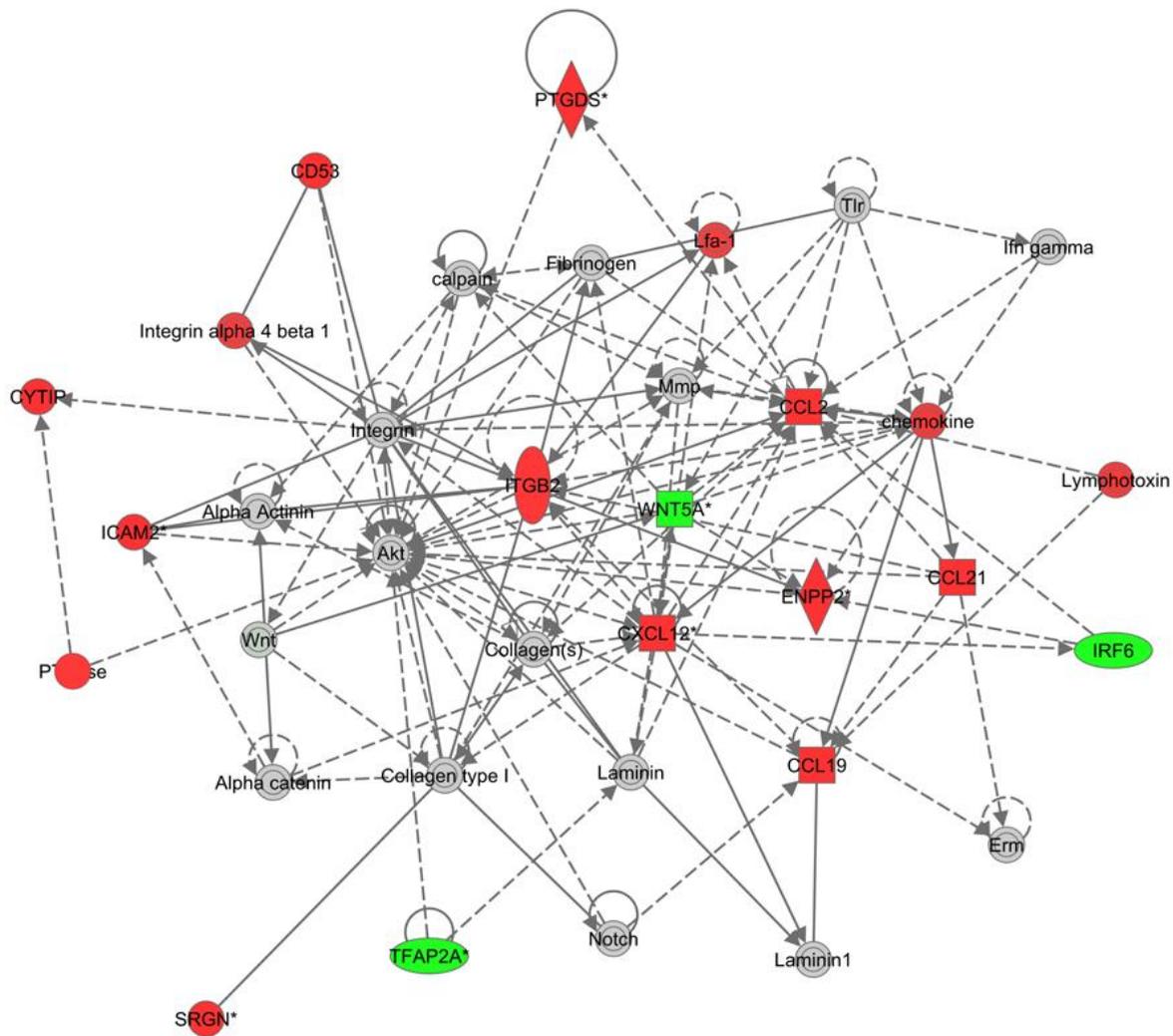


Abb. 37: Top 02-Netzwerk (*Ingenuity Pathway Analysis*[®]) differenziell exprimierter Gene in Mvs.PT.

In diesem Netzwerk, das mit zellulärer Bewegung, Entwicklung und Funktion des hämatologischen Systems und Immunzell-Transport assoziiert wird, waren 14 Gene signifikant überexprimiert (bei dem benutzten Filter: $\log_2FC \geq |1,5|$ und $p\text{-Wert} < 0,0001$).

Die Symbolfarbe visualisiert hochregulierte (rote), herabregulierte (grüne) und nicht-signifikant veränderte (graue) Gene. Linien und Pfeile zwischen den Symbolen repräsentieren direkte (durchgehende Linien) und indirekte (gestrichelte Linien) Interaktionen zwischen den Proteinen. Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

ERGEBNISSE

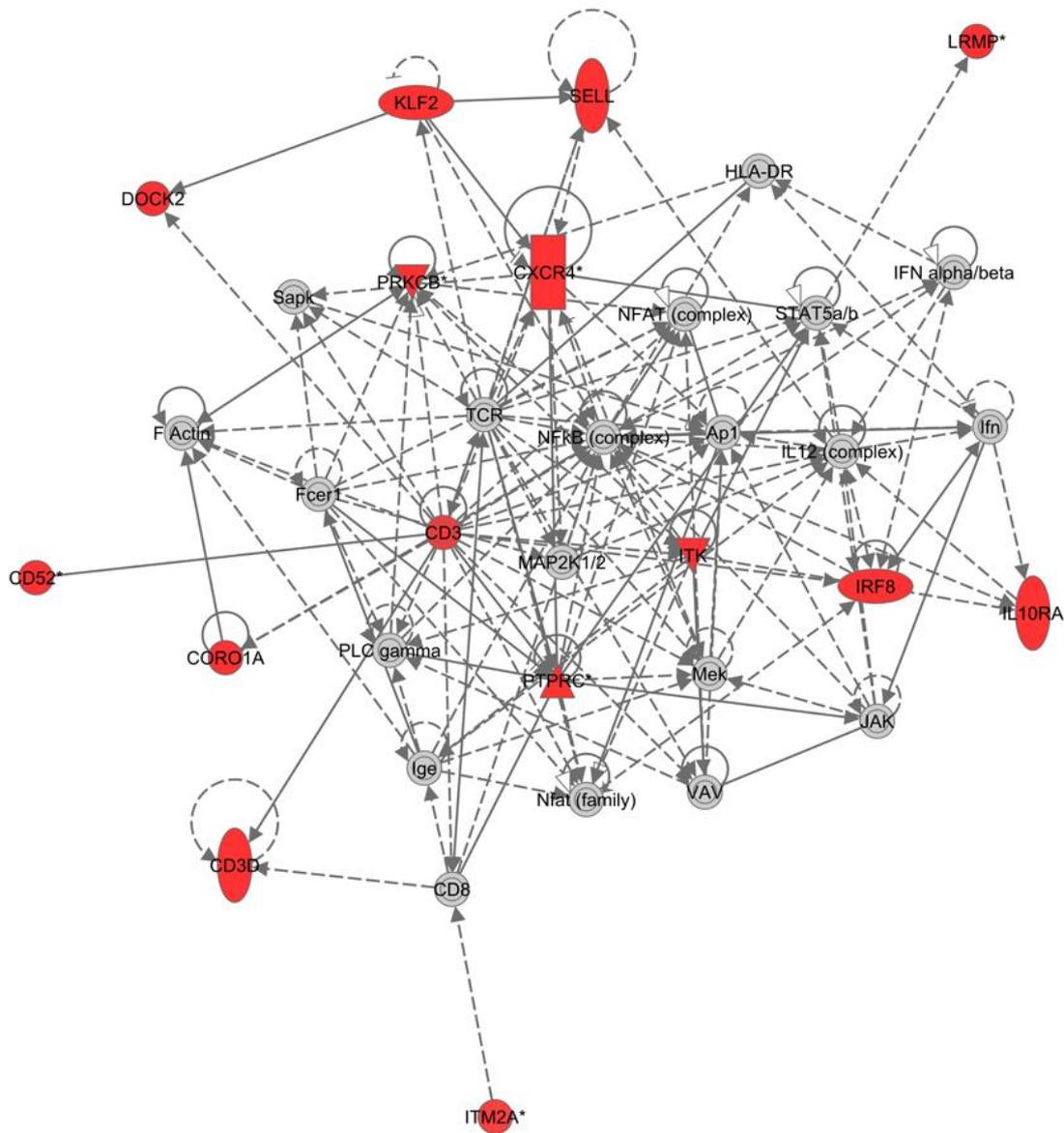


Abb. 38: Top 03-Netzwerk (*Ingenuity Pathway Analysis*[®]) differenziell exprimierter Gene in Mvs.PT.

In diesem Netzwerk, das mit Zell-Zell-Signalwegen und -Interaktionen, Entwicklung und Funktion des hämatologischen Systems und Gewebemorphologie assoziiert wird, waren 14 Gene signifikant überexprimiert (bei dem benutzten Filter: $\log_2FC \geq |1,5|$ und $p\text{-Wert} < 0,0001$).

Die Symbolfarbe visualisiert hochregulierte (rote), herabregulierte (grüne) und nicht-signifikant veränderte (graue) Gene. Linien und Pfeile zwischen den Symbolen repräsentieren direkte (durchgehende Linien) und indirekte (gestrichelte Linien) Interaktionen zwischen den Proteinen. Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

ERGEBNISSE

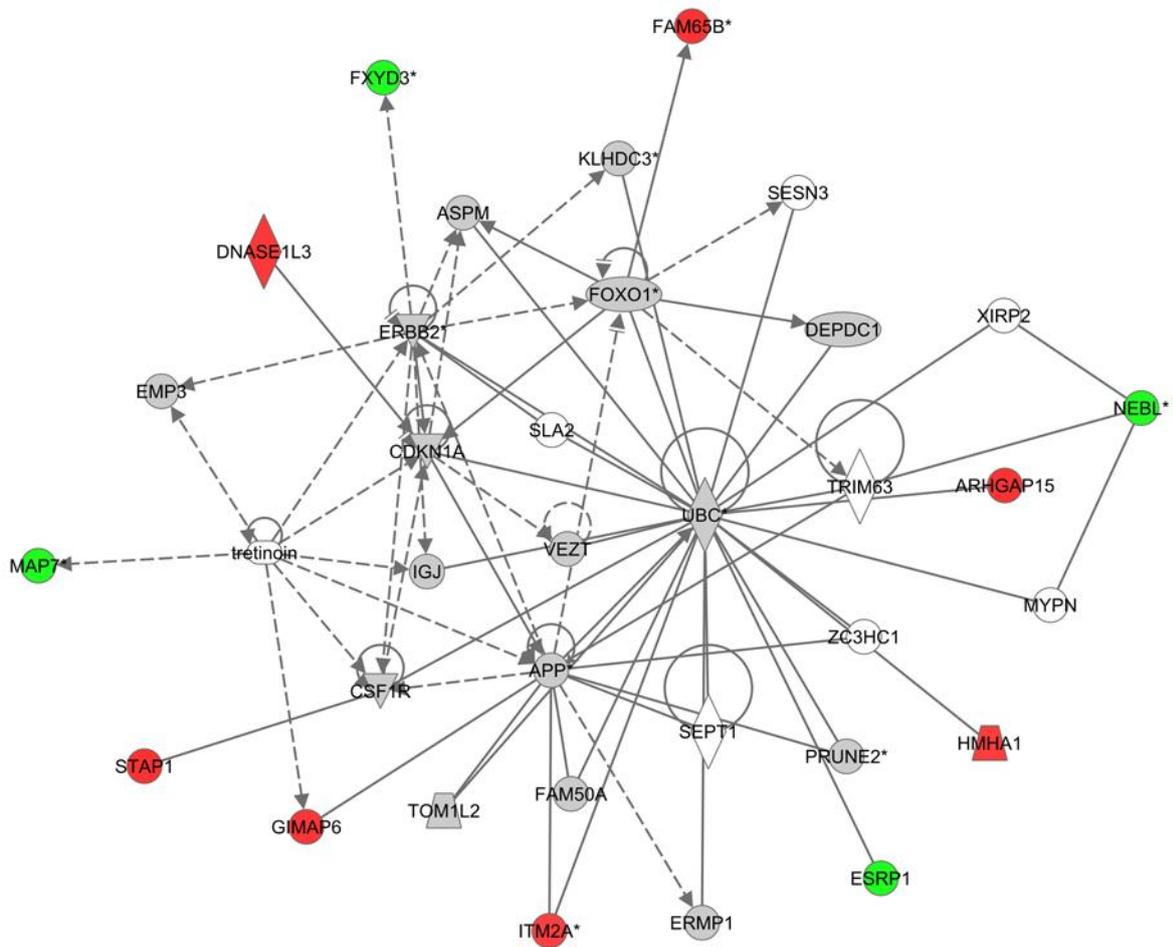


Abb. 39: Top 04-Netzwerk (*Ingenuity Pathway Analysis*[®]) differenziell exprimierter Gene in Mvs.PT.

In diesem Netzwerk, das mit Zellzyklus, Zellmorphologie und zellulärer Beeinträchtigung assoziiert wird, waren elf Gene signifikant überexprimiert (bei dem benutzten Filter: $\log_2FC \geq |1,5|$ und $p\text{-Wert} < 0,0001$).

Die Symbolfarbe visualisiert hochregulierte (rote), herabregulierte (grüne) und nicht-signifikant veränderte (graue) Gene. Linien und Pfeile zwischen den Symbolen repräsentieren direkte (durchgehende Linien) und indirekte (gestrichelte Linien) Interaktionen zwischen den Proteinen. Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

ERGEBNISSE

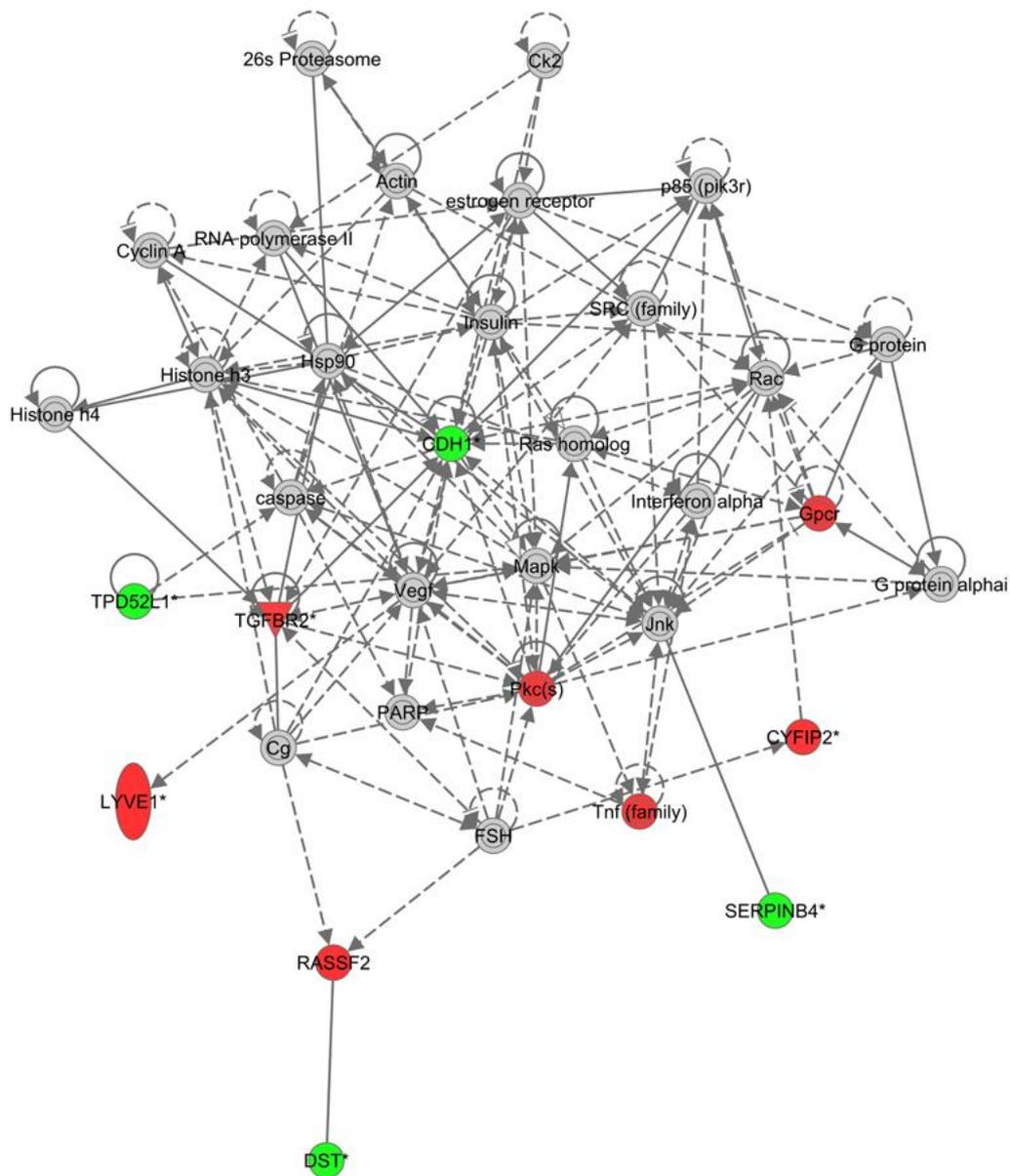


Abb. 40: Top 05-Netzwerk (*Ingenuity Pathway Analysis*[®]) differenziell exprimierter Gene in Mvs.PT.

In diesem Netzwerk, das mit Krebs, dermatologischen und gastrointestinalen Erkrankungen assoziiert wird, waren acht Gene signifikant überexprimiert (bei dem benutzten Filter: $\log_2FC \geq |1,5|$ und $p\text{-Wert} < 0,0001$).

Die Symbolfarbe visualisiert hochregulierte (rote), herabregulierte (grüne) und nicht-signifikant veränderte (graue) Gene. Linien und Pfeile zwischen den Symbolen repräsentieren direkte (durchgehende Linien) und indirekte (gestrichelte Linien) Interaktionen zwischen den Proteinen. Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

ERGEBNISSE

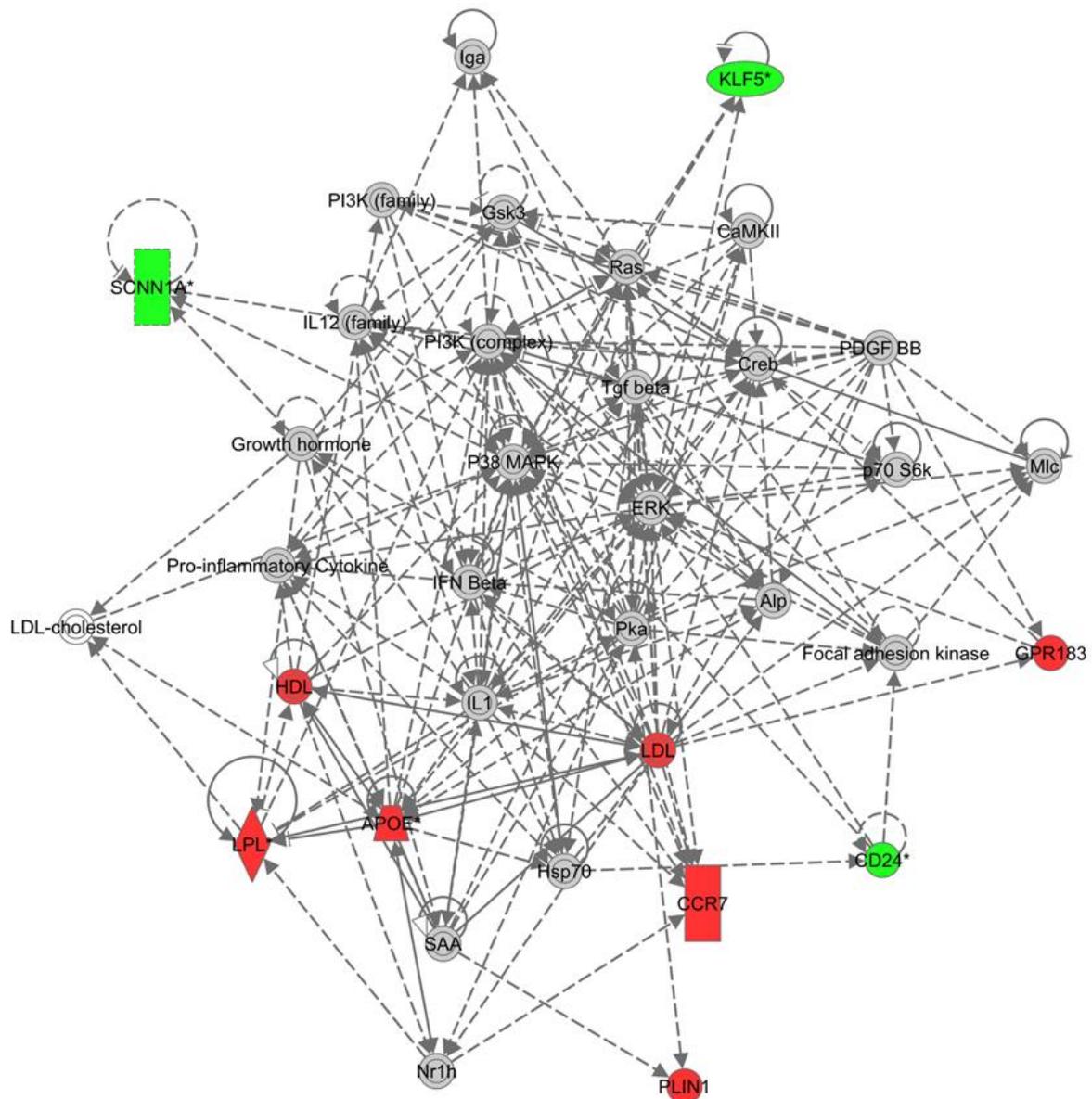


Abb. 41: Top 06-Netzwerk (*Ingenuity Pathway Analysis*[®]) differenziell exprimierter Gene in Mvs.PT.

In diesem Netzwerk, das mit dem Lipidmetabolismus, „Small Molecule“ Biochemie und kardiovaskulären Erkrankungen assoziiert wird, waren acht Gene signifikant überexprimiert (bei dem benutzten Filter: $\log_2FC \geq |1,5|$ und $p\text{-Wert} < 0,0001$). Die Symbolfarbe visualisiert hochregulierte (rote), herabregulierte (grüne) und nicht-signifikant veränderte (graue) Gene. Linien und Pfeile zwischen den Symbolen repräsentieren direkte (durchgehende Linien) und indirekte (gestrichelte Linien) Interaktionen zwischen den Proteinen. Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

ERGEBNISSE

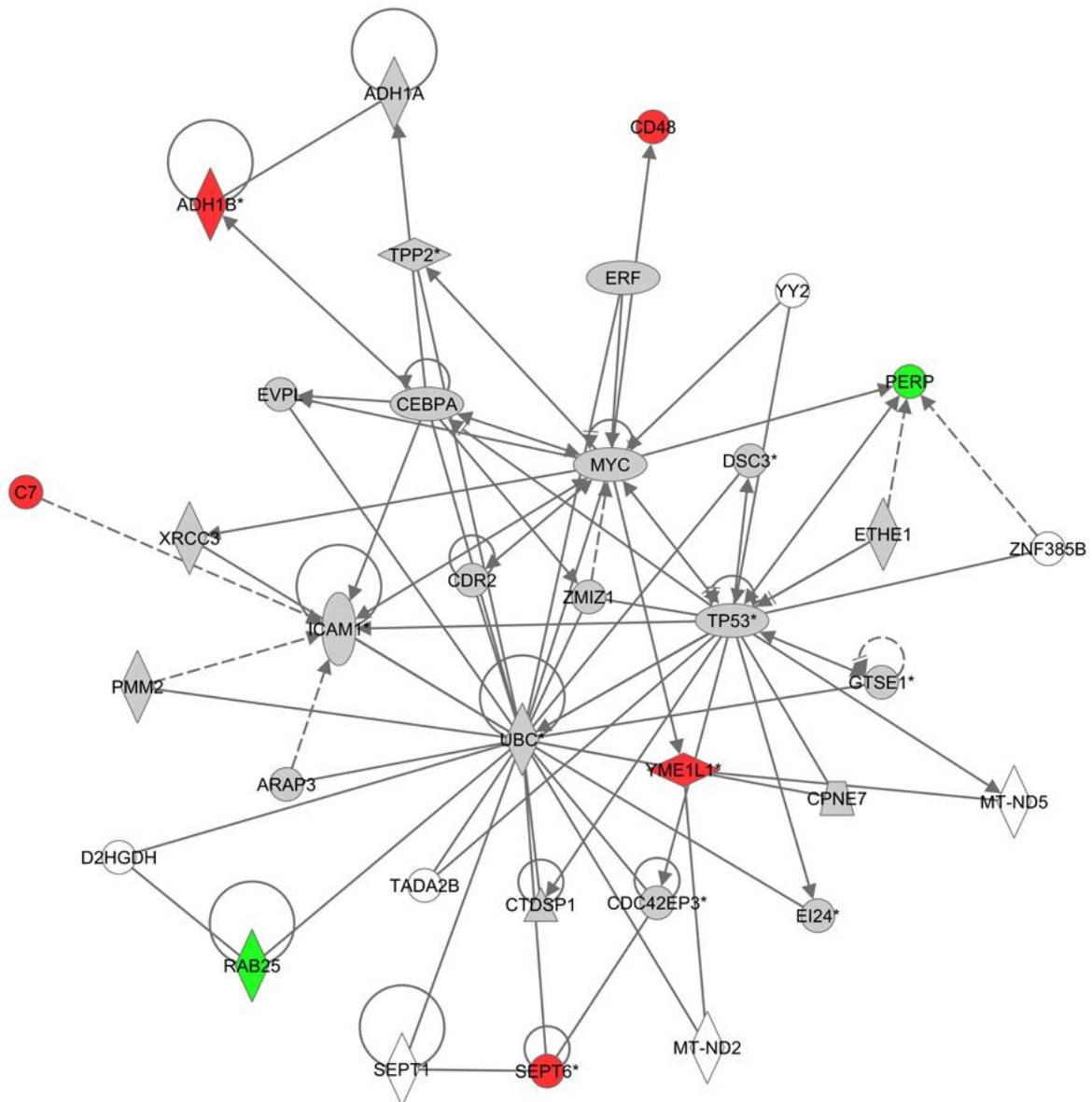


Abb. 42: Top 07-Netzwerk (*Ingenuity Pathway Analysis*[®]) differenziell exprimierter Gene in Mvs.PT.

In diesem Netzwerk, das mit Krebs, Erkrankungen des endokrinen Systems und Entwicklungsstörungen assoziiert wird, waren sieben Gene signifikant überexprimiert (bei dem benutzten Filter: $\log_2FC \geq |1,5|$ und $p\text{-Wert} < 0,0001$).

Die Symbolfarbe visualisiert hochregulierte (rote), herabregulierte (grüne) und nicht-signifikant veränderte (graue) Gene. Linien und Pfeile zwischen den Symbolen repräsentieren direkte (durchgehende Linien) und indirekte (gestrichelte Linien) Interaktionen zwischen den Proteinen. Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

Die Darstellung der Netzwerke, in denen besonders viele beteiligte Gene überexprimiert sind, gibt insbesondere die Möglichkeit, die direkte Wirkung möglicher Fehlregulationen im Tumor und den dazugehörigen Lymphknotenmetastasen sichtbar zu machen. Die Fehlregulationen selbst, die zu einer Expressionsänderung führen, werden jedoch nur durch folgende Upstream-Analysen sichtbar, in denen das bioinformatische Werkzeug der *Upstream Regulator Analysis* innerhalb der *Ingenuity Pathway Analysis*[®] die am wahrscheinlichsten beteiligten Regulatoren errechnet. Mit Hilfe dieser Regulatoren lassen sich die beteiligten Netzwerke zum besseren Verständnis in einen größeren Kontext setzen. Der „z-Score“ beschreibt hier die Wahrscheinlichkeit einer Aktivierung. Diese Aktivierung kann durch Mutationen, aber auch durch andere transkriptionelle Veränderungen oder posttranslationale Modifikationen verursacht sein. Einen Hinweis darauf, ob transkriptionelle Veränderungen verantwortlich sind, gibt die Expressions-Ratio (Log Ratio) des Regulators selbst. Für alle folgenden Upstream Analysen wurde ein z-Score mit einem Betrag von mindestens 2 für eine signifikante Aktivierung oder Inhibition vorausgesetzt, und nur diese Regulatoren in die Tabellen übernommen.

Bei den analysierten Upstream Regulatoren für PTvs.N mit einem z-Score >2 und damit aktiviert waren mehrere Wachstumsfaktoren (FGF2, TGFB3, CTGF und TGFB1), sowie ein für HNSCC bekanntes Oberflächenglykoprotein (CD44)¹⁷⁷, der Komplex Ap1 und das Zytokin IL1A, mit einem z-Score von unter -2 und damit inhibiert waren mehrere Transkriptionsregulatoren (EHF, MYC und SPDEF), aber auch ein Wachstumsfaktor (WISP2), ein Transporter (ABCB4), der Kernrezeptor AHR und Andere (CR1L, FBN1 und der Östrogenrezeptor) (**Tab. 6**).

Tab. 6: Aktivierte/Inhibierte Upstream Regulatoren PTvs.N (*Ingenuity Pathway Analysis*[®])

Upstream Regulator	Log Ratio	Molekültyp	Vorhergesagter Aktivierungszustand	z-Score	p-Wert der Überlappung
CD44	0,382	Enzym	Aktiviert	2,137	6,39E-09
FGF2	-0,133	Wachstumsfaktor	Aktiviert	2,261	1,08E-10
IL1A	0,090	Zytokin	Aktiviert	2,298	1,51E-11
TGFB3	0,471	Wachstumsfaktor	Aktiviert	2,416	3,72E-09
Ap1		Komplex	Aktiviert	2,581	9,43E-08
CTGF	0,357	Wachstumsfaktor	Aktiviert	2,930	2,80E-08
TGFB1	0,688	Wachstumsfaktor	Aktiviert	3,316	2,75E-09

ERGEBNISSE

Upstream Regulator	Log Ratio	Molekültyp	Vorhergesagter Aktivierungszustand	z-Score	p-Wert der Überlappung
EHF	-2,432	Transkriptionsregulator	Inhibiert	-2,333	6,02E-08
MYC	-0,071	Transkriptionsregulator	Inhibiert	-2,433	4,50E-10
WISP2	-0,000	Wachstumsfaktor	Inhibiert	-2,630	1,64E-07
ABCB4	-0,054	Transporter	Inhibiert	-2,631	1,16E-08
CR1L		Andere	Inhibiert	-2,804	4,24E-10
FBN1	0,528	Andere	Inhibiert	-2,804	5,11E-12
estrogen receptor		Gruppe	Inhibiert	-2,937	2,28E-10
SPDEF	-0,558	Transkriptionsregulator	Inhibiert	-3,000	3,99E-08
AHR	0,269	Ligand-abhängiger Kernrezeptor	Inhibiert	-3,063	7,40E-10

Unerwarteterweise gab es für Mvs.PT nur insgesamt zwei signifikante Einträge: eine Aktivierung der PI3K Familie und das Zytokin IFNG (**Tab. 7**). Es wurden keine signifikant inhibierten Upstream-Faktoren für Mvs.PT vorhergesagt.

Tab. 7: Aktivierte/Inhibierte Upstream Regulatoren Mvs.PT (*Ingenuity Pathway Analysis*®)

Upstream Regulator	Log Ratio	Molekültyp	Vorhergesagter Aktivierungszustand	z-Score	p-Wert der Überlappung
PI3K (family)		Gruppe	Activated	2,216	3,43E-05
IFNG	-0,038	Zytokin	Activated	2,449	1,48E-05

Weiterführend bietet IPA die Möglichkeit, die Aktivierung nachfolgender Downstream-Ziele der aktivierten Upstream-Regulatoren nachzuverfolgen und zu visualisieren, wobei sich diese Arbeit auf die auffälligsten Beispiele beschränkt.

Da CD44 für Kopf-Hals-Tumore als Stammzellmarker beschrieben wurde ^{177,178} und zudem zur Detektion zirkulierender Tumorzellen in Flüssigbiopsien als Marker getestet wird ¹⁷⁹, hat dieser Upstream-Regulator für den Datensatz PTvs.N große Bedeutung. Der Rezeptor für Hyaluronsäure bindet Osteopontin (SPP1), Kollagene und Matrix-Metalloproteinasen und verstärkt die Signalwege

von PI3K und ERK, während inkonsistente Regulationsmechanismen für PTvs.N vor allem NFκB-Signale, STAT3 und den Ap1-Komplex involvierten (**Abb. 43**).

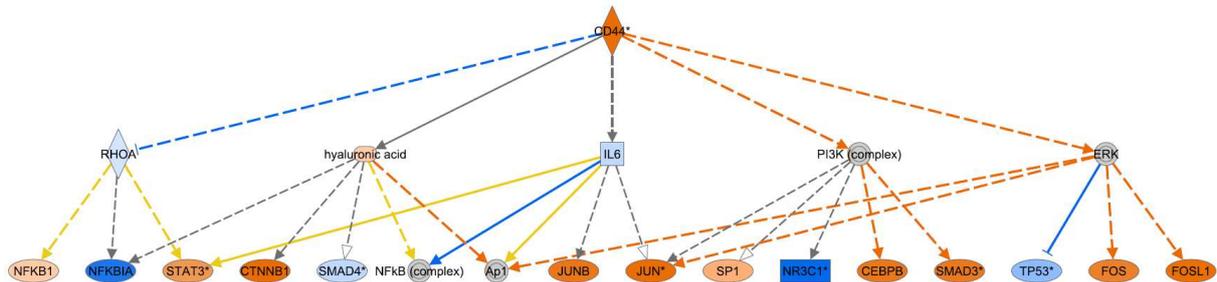


Abb. 43: Der Upstream-Regulator CD44 in PTvs.N (*Ingenuity Pathway Analysis*[®]).

Regulierte Downstream-Ziele von CD44 sind als aktiviert (orange) oder inhibiert (blau) vorhergesagt (grau: nicht-signifikant verändert). Orangene Linien und Pfeile zeigen eine vorausgesagte Aktivierung, blaue eine Inaktivierung, und gelbe symbolisieren einen Widerspruch mit dem Status des Downstream-Moleküls an (durchgehende Linien: direkte Interaktion; gestrichelte Linien: indirekte Verbindung). Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

Inkonsistenzen der vorhergesagten Regulationsmechanismen waren auch beim potentiellen Upstream-Regulator *Connective Tissue Growth Factor* (CTGF), der vor allem von vaskulären Endothelzellen sekretiert wird, in PTvs.N erkennbar, insbesondere bei p53 und – erneut – dem NFκB-Signalweg. Von CTGF ausgehend wurde eine Aktivierung von VEGFA, MAPK14 und FN1 vorhergesagt, die über diverse molekulare Faktoren wie den Ap-1-Transkriptionskomplex mit Jun/Fos (Regulation der Genaktivität einer Wachstumsfaktorantwort folgend) und β-Catenin (CTNNB1, als wichtigstem Faktor im Wnt-Signalweg) Einfluss auf diverse zelluläre Prozesse (**Abb. 44**).

ERGEBNISSE

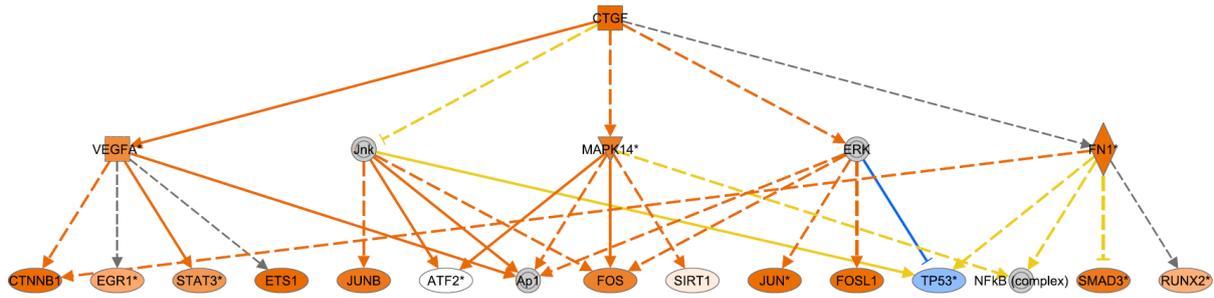


Abb. 44: Der Wachstumsfaktor CTGF in PTvs.N (*Ingenuity Pathway Analysis*[®]).

Regulierte Downstream-Ziele von CTGF sind als aktiviert (orange) oder inhibiert (blau) vorhergesagt (grau: nicht-signifikant verändert). Orangene Linien und Pfeile zeigen eine vorausgesagte Aktivierung, blaue eine Inaktivierung, und gelbe zeigen einen Widerspruch mit dem Status des Downstream-Moleküls an (durchgehende Linien: direkte Interaktion; gestrichelte Linien: indirekte Verbindung). Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

Da der CTGF-Downstreamfaktor Fibronectin (FN1) auch als eines der Gene aus der PTvs.N Gensignatur identifiziert, in der Validierung mittels RT-PCR überprüft und in der Downstream-Analyse als signifikant aktiviert vorhergesagt wurde, werden die vorhergesagten Zusammenhänge hier noch im Detail exemplarisch dargestellt. FN1 ist in viele zelluläre Prozesse involviert, trägt indirekt steuernd über VEGFA, die Jnk und Tfgβ-Familie sowie über HRAS und ERK kollektiv über den Transkriptionsfaktor ETS1 (Expression von Zytokinen und Chemokinen), β-Catenin (CTNNB1), dem Transkriptionskomplex Ap-1, dem korrespondierenden nukleären Phosphoprotein c-Fos und Ap-1, als auch dem Transkriptionsfaktor SP1 bei, die alle als aktiviert oder hochreguliert im selben Datensatz gefunden wurden. Die gleichen Widersprüche bei der Regulation von p53 wurden konsistent zur CTGF-Vorhersage auch für FN1 gefunden (**Abb. 45**).

ERGEBNISSE

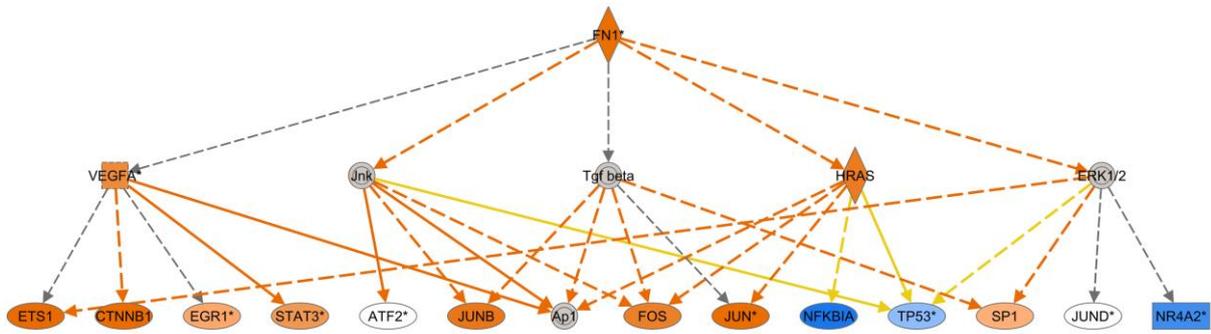


Abb. 45: Der Upstream-Regulator FN1 in PTvs.N (*Ingenuity Pathway Analysis*[®]).

Regulierte Downstream-Ziele von FN1 sind als aktiviert (orange) oder inhibiert (blau) vorhergesagt (grau: nicht-signifikant verändert). Orangene Linien und Pfeile zeigen eine vorausgesagte Aktivierung und gelbe zeigen einen Widerspruch mit dem Status des Downstream-Moleküls an (durchgehende Linien: direkte Interaktion; gestrichelte Linien: indirekte Verbindung). Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

Die vorhergesagte Inaktivierung von MYC widersprach den üblichen Erwartungen, da MYC in vielen Tumoren als Onkogen fungiert und entsprechend aktiviert und/oder überexprimiert ist^{180,181}. MYC befindet sich zudem downstream von diversen Wachstumsfaktor-aktivierten Signalwegen, stellt jedoch nur eine der Möglichkeiten der Zelle dar, die Transkription als Antwort auf Wachstumsfaktoren zu verändern. Eine Detailansicht der involvierten Mechanismen zeigte, dass die Vorhersagen viele Widersprüche erzeugten (**Abb. 46**).

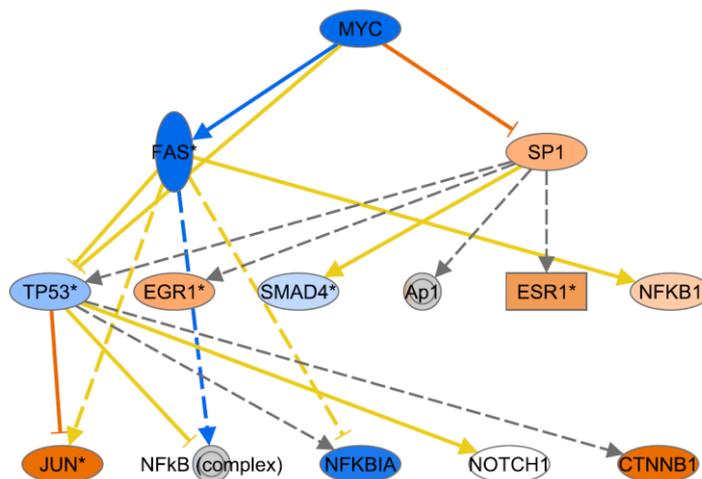


Abb. 46: Der Upstream-Regulator MYC in PTvs.N (*Ingenuity Pathway Analysis*[®]).

ERGEBNISSE

Regulierte Downstream-Ziele von MYC sind als aktiviert (orange) oder inhibiert (blau) vorhergesagt (grau: nicht-signifikant verändert). Orangene Linien und Pfeile zeigen eine vorausgesagte Aktivierung, blaue eine Inaktivierung, und gelbe zeigen einen Widerspruch mit dem Status des Downstream-Moleküls an (durchgehende Linien: direkte Interaktion; gestrichelte Linien: indirekte Verbindung). Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

Ein weiterer Fokus wurde auf Wachstumsfaktoren und involvierte Rezeptortyrosinkinasen (RTKs) gelegt, da insbesondere in Hochdurchsatzsequenzierungs-Studien des Genoms von zahlreichen Kopf-Hals-Tumoren häufig Amplifikationen im Bereich der RTKs gefunden werden konnten ¹⁶⁷. Detailansichten der vorgesagten involvierten Wachstumsfaktoren konzentrieren sich erneut auf den Ap1-Komplex mit JUN/FOS und Wnt-Signaling über CTNNB1. Widersprüchlich waren wie in vorherigen Abbildungen die Regulationen insbesondere von p53 und NFκB (**Abb. 47** und **Abb. 48**).

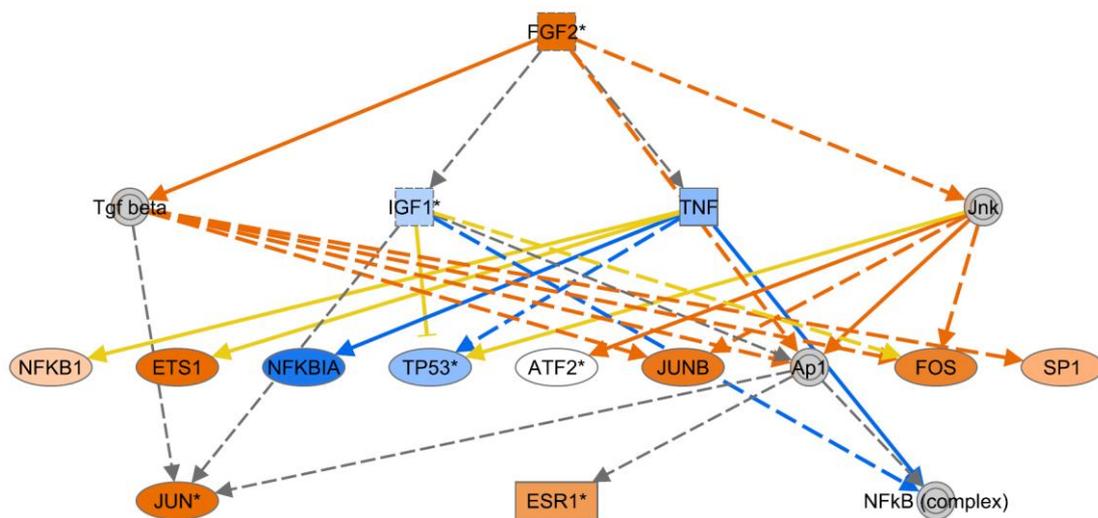


Abb. 47: Der Wachstumsfaktor FGF2 in PTvs.N (*Ingenuity Pathway Analysis*[®]).

Regulierte Downstream-Ziele von FGF2 sind als aktiviert (orange) oder inhibiert (blau) vorhergesagt (grau: nicht-signifikant verändert). Orangene Linien und Pfeile zeigen eine vorausgesagte Aktivierung, blaue eine Inaktivierung, und gelbe zeigen einen Widerspruch mit dem Status des Downstream-Moleküls an (durchgehende Linien: direkte Interaktion; gestrichelte Linien: indirekte Verbindung). Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

ERGEBNISSE

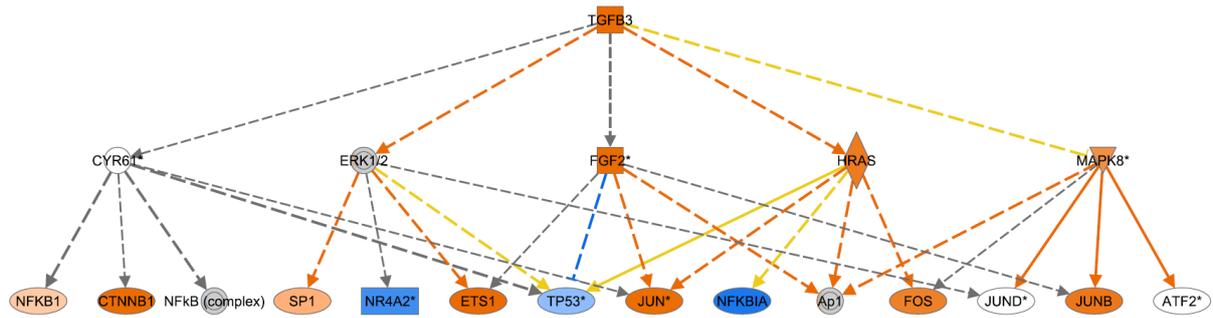


Abb. 48: Der Wachstumsfaktor TGFβ3 in PTvs.N (*Ingenuity Pathway Analysis*[®]).

Regulierte Downstream-Ziele von TGFβ3 sind als aktiviert (orange) oder inhibiert (blau) vorhergesagt (grau: nicht-signifikant verändert). Orangene Linien und Pfeile zeigen eine vorausgesagte Aktivierung, blaue eine Inaktivierung, und gelbe zeigen einen Widerspruch mit dem Status des Downstream-Moleküls an (durchgehende Linien: direkte Interaktion; gestrichelte Linien: indirekte Verbindung). Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

Der Fibroblast Growth Factor FGF2 aktiviert entsprechende Rezeptoren wie den FGFR2 und FGFR3 und kommt in Keratinozyten, aber auch verschiedenen Tumoren vor ¹⁸². Da die Rezeptoren sich jedoch in Tumoren durch verschiedene Mutationen auch selbst aktivieren können ¹⁸³, ist es erforderlich, die Rezeptoren im Detail selbst zu untersuchen. Eine Betrachtung der grafischen Darstellung von Wachstumsfaktor-Rezeptoren durch Ingenuity im *Upstream Regulator Tool* war unzureichend, da nur drei epidermale Wachstumsfaktor-Rezeptoren in die Vorhersage eingeschlossen wurden, EGFR, ERBB2 und ERBB3 (**Abb. 49**). Um weiterführend ein detaillierteres Bild der Wachstumsfaktorrezeptoren und der beeinflussenden Faktoren zu gewinnen, wurde eine eigene Anzeigemethodik entworfen, auf welche im Folgenden näher eingegangen wird (**Abb. 53**).

ERGEBNISSE

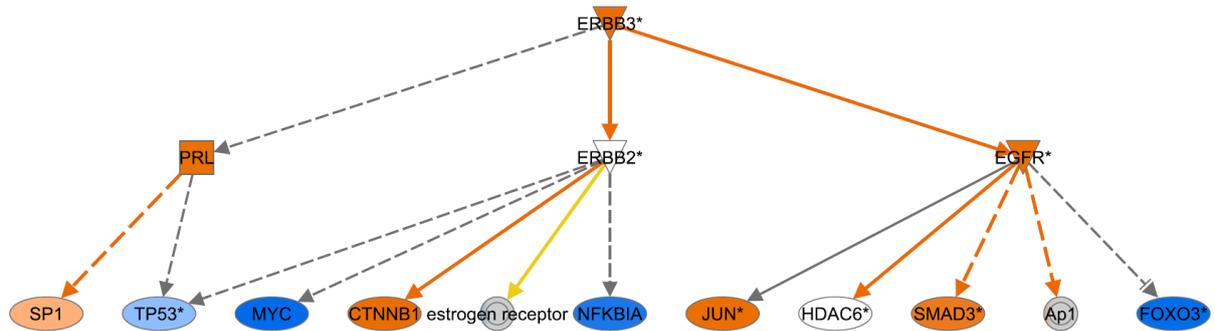


Abb. 49: Der Rezeptor ERBB3 in PTvs.N (*Ingenuity Pathway Analysis*[®]).

Regulierte Downstream-Ziele von ERBB3 sind als aktiviert (orange) oder inaktiviert (blau) vorhergesagt (grau: nicht-signifikant verändert). Orangene Linien und Pfeile zeigen eine vorausgesagte Aktivierung, blaue eine Inaktivierung, und gelbe zeigen einen Widerspruch mit dem Status des Downstream-Moleküls an (durchgehende Linien: direkte Interaktion; gestrichelte Linien: indirekte Verbindung). Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

Das für Mvs.PT mit einem z -Score >2 aktivierte Interferon Gamma (IFNG) war einer von nur zwei signifikant aktivierten Upstream-Faktoren. Interferone sind Zytokine, die über die gezielte Beeinflussung von Zellreaktionen und der Aktivierung von Immunreaktionen vor Krankheiten schützen sollen. Über Jak/STAT-Signalwege wird die Transkription innerhalb der Zelle angeregt, über die Proliferation, Apoptose, Überleben, Zellmigration und die Immunantwort reguliert werden können¹⁸⁴. In der genaueren Upstream-Analyse für Mvs.PT wurden einige Diskrepanzen in der Vorhersage für IFNG deutlich (**Abb. 50**). Insbesondere die Aktivierung von IL10 und PPARG war für diesen Signalweg unüblich.

ERGEBNISSE

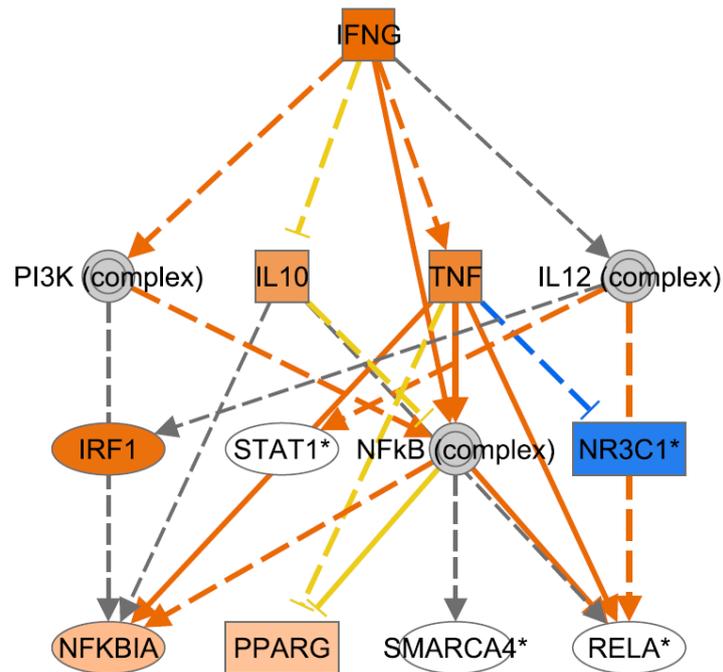


Abb. 50: Das Zytokin Interferon Gamma (IFNG) in Mvs.PT (*Ingenuity Pathway Analysis*[®]).

Regulierte Downstream-Ziele von IFNG sind als aktiviert (orange) oder inhiert (blau) vorhergesagt (grau: nicht-signifikant verändert). Orangene Linien und Pfeile zeigen eine vorausgesagte Aktivierung, blaue eine Inaktivierung, und gelbe zeigen einen Widerspruch mit dem Status des Downstream-Moleküls an (durchgehende Linien: direkte Interaktion; gestrichelte Linien: indirekte Verbindung). Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

Die PI3K-Familie stimuliert die Sekretion von MMP9, aktiviert den AKT und NFkB Signalweg und ist die in Mvs.PT die am signifikantesten hochregulierte Gruppe (**Tab. 7**). Inkonsistent war in der Analyse der Downstream-Ziele der PI3K-Familie in **Abb. 51** die Aktivierung von FOXO1 sowie PPARG. Konsistent war die Inaktivierung von NR3C1, welches in **Abb. 50** durch TNF inhiert wurde. Zur vertieften Betrachtung wurden die weiterführenden Downstream-Ziele des NFkB-Weges mit dargestellt (**Abb. 52**). Dort war vor allem die Aktivierung von PPARG unüblich, welches regulär von zwei aktivierten Zielen des NFkB Komplexes, TNF und IKBKB, inaktiviert werden müsste.

ERGEBNISSE

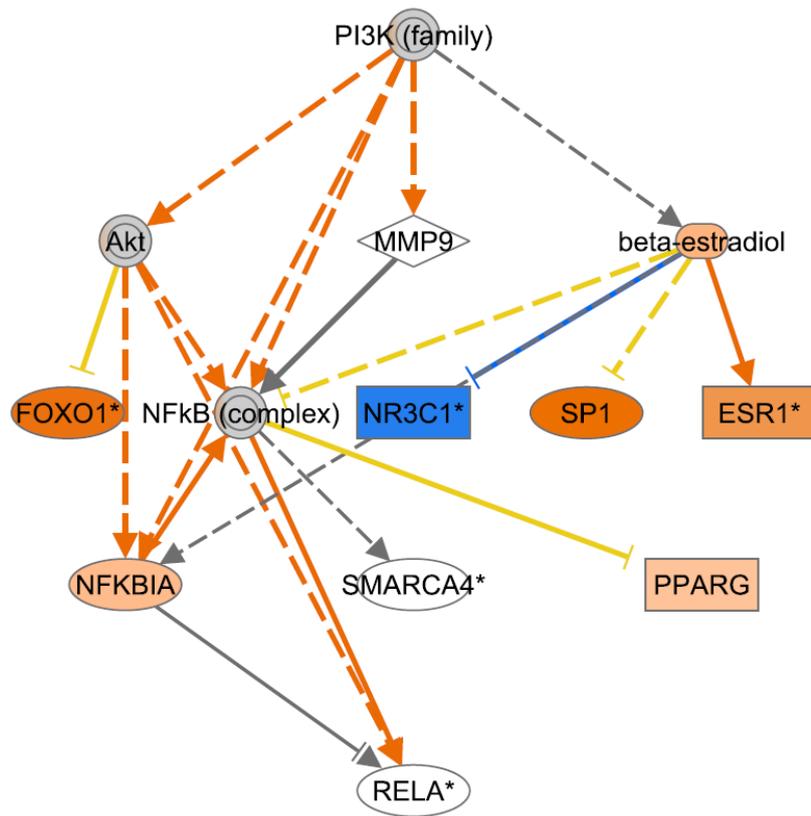


Abb. 51: Die PI3K-Familie in Mvs.PT (*Ingenuity Pathway Analysis*[®]).

Regulierte Downstream-Ziele der PI3K Familie sind als aktiviert (orange) oder inhibiert (blau) vorhergesagt (grau: nicht-signifikant verändert). Orangene Linien und Pfeile zeigen eine vorausgesagte Aktivierung, blaue eine Inaktivierung, und gelbe zeigen einen Widerspruch mit dem Status des Downstream-Moleküls an (durchgehende Linien: direkte Interaktion; gestrichelte Linien: indirekte Verbindung). Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

ERGEBNISSE

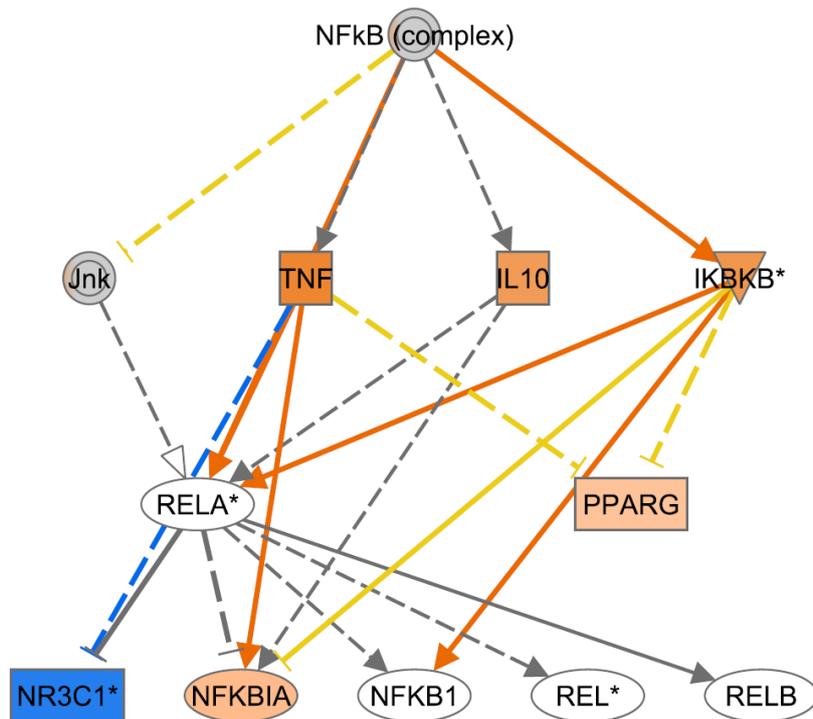


Abb. 52: Der NFκB-Komplex in Mvs.PT (*Ingenuity Pathway Analysis*[®]).

Regulierte Downstream-Ziele der PI3K-Familie sind als aktiviert (orange) oder inhibiert (blau) vorhergesagt (grau: nicht-signifikant verändert). Orangene Linien und Pfeile zeigen eine vorausgesagte Aktivierung, blaue eine Inaktivierung, und gelbe zeigen einen Widerspruch mit dem Status des Downstream-Moleküls an (durchgehende Linien: direkte Interaktion; gestrichelte Linien: indirekte Verbindung). Alle grafischen Darstellungen werden von mindestens einer Referenz aus der *Ingenuity Knowledge Base*[®] unterstützt. Symbolformen zeigen die funktionelle Klasse der Genprodukte an: Quadrat → Zytokin; Dreieck → Kinase; Rechteck → nukleärer Rezeptor; Kreis im Kreis → Gruppe oder Komplex; Vertikale Raute → Enzym; Horizontale Raute → Protease; Trapez → Transporter; Vertikale Ellipse → Transmembranrezeptor; Horizontale Ellipse → Transkriptionsregulator; Kreis → Andere.

Im Vergleich der im Rahmen dieser Arbeit gewonnenen Microarray-Expressionsdaten (MEx) mit den Amplifikationsdaten aus *Next-Generation-Sequenzierungen* (NGS) des *Cancer Genome Atlas*-Projekts ergaben sich viele Übereinstimmungen in der Tendenz bei Rezeptortyrosinkinasen wie die Hochregulation von FGFR1 (13% MEx in PTvs.N und 7% MEx in Mvs.PT zu 10% NGS), von FGFR3 (7% MEx zu 2% in NGS), bei DDR2 in der Entwicklung zu Lymphknotenmetastasen (0% MEx in PTvs.N, 13% MEx in Mvs.PT zu 3% in NGS), von MET (7% MEx zu 2% NGS), aber auch bei Überlebensfaktoren wie BIRC2 (7% MEx zu 7% NGS) und FADD (33% MEx zu 32% NGS) sowie TP63 (13% MEx zu 19% NGS) (Tab. 8). Bei anderen, erwarteten Hochregulationen durch in statistisch ermittelten Amplifikationsereignissen aus NGS zeigte sich keine Übereinstimmung in der Tendenz, wie beim EGFR (0% MEx zu 15% NGS) PIK3CA (7% MEx zu 34% NGS) und NFE2L2 (0% MEx zu 14% NGS). Eine

ERGEBNISSE

20% höhere Regulation als im korrespondierenden Normalgewebe wurde dabei als „tendenziell hochreguliert“ kategorisiert und prozentual über alle Patienten berechnet.

Tab. 8: Vergleich der Expressionsdaten mit Amplifizierungsdaten aus Next Generation Sequencing.

Gruppe	Gen	Hochregulierung von 20% oder mehr		Vorhergesagt in NGS	
		% in PTvs.N	% in Mvs.N	% in HPV(-)	% in HPV(+)
(R)TK	EGFR	0	0	15↑	6↑
	ERBB2	0	0	4↑	3↑
	FGFR1	13	7	10↑	0
	FGFR3	7	7	2↑	11↑
	IGF1R	0	0	4↑	0
	EPHA2	0	0	4m	3m
	DDR2	0	13	3↑	6↑
	MET	7	7	2↑	0
Onkogene	CCND1	7	0	31↑	3↑
	MYC	7	7	14↑	3↑
	HRAS	13	13	5m	0
PI(3)K	PIK3CA	7	0	34↑/m	56↑/m
Überleben	BIRC2	7	0	7↑	3↑
	FADD	33	13	32↑	6↑
	E2F1	0	7	2↑	19↑
Andere	TP63	13	0	19↑	28↑
	NFE2L2	0	0	14↑	0
	NF1	0	0	3↑/m	0

↑: amplifiziert, m: mutiert.

Um weiterführend ein realistischeres und direktes Bild der eigentlichen Expressionsveränderung der verschiedenen Rezeptoren zu bekommen, wurde eine eigene Anzeigemethodik entworfen. Die Expressionsintensitäten der einzelnen Gewebeproben aus Primärtumor und Metastasen wurden für jeden Patienten ins Verhältnis mit dem dazugehörigen Normalgewebe (als Ratio: PT/N und M/N) gestellt, über die Basis 2 logarithmiert und in *Microsoft Excel® (2010)* mit einem Farbfilter versehen, so dass rote Felder eine Überexpression von wenigstens 20% (entspricht $\log_2FC \geq 0,2263$), grüne

Felder eine Herabregulation von -20% ($\log_2FC \leq -0.263$) und schwarze Felder eine gleichbleibende Expression darstellen. Waren mehrere Affy-IDs für ein Gen vorhanden, wurden erst die Verhältnisse berechnet und anschließend daraus der Mittelwert in die Anzeige miteinbezogen (**Abb. 53**). Die Auswahl der betrachteten Gene richtete sich nach den publizierten *Next-Generation-Sequenzierungsdaten* des *Cancer Genome Atlas* Projektes und bereits bekannten Verwandten aus der jeweils gleichen Gruppe oder mit ähnlicher Funktion, die auf Grund bisher publizierter Studien potentiell in HNSCC auf Grund differenzieller Expression eine Rolle spielen könnten ¹⁶⁷.

Der zu 4% in NGS amplifizierte Rezeptor Her-2 (ERBB2) war in MEx fast ausschließlich eher herabreguliert. Auch der zu 15% amplifizierte EGFR in NGS war in keinem Patienten der MEx signifikant überexprimiert und zeigte nur in den Tumoren der Patienten 1, 5, 9 und 15 eine leicht höhere Expression. Die in den NGS fehlenden Rezeptoren VEGFR2, VEGFR3, PDGFRB und ABL2 waren in MEx bei einigen, aber unterschiedlichen Patienten signifikant erhöht (**Abb. 53**).

Bei Genen aus anderen Gruppen, die in NGS amplifiziert waren, gab es in MEx bei IAP-2 (BIRC2, 7% MEx in PTvs.N zu 7% in NGS), FADD (33% MEx in PTvs.N zu 32% in NGS), E2F1 (0% MEx in PTvs.N und 7% in Mvs.PT zu 2% in NGS), TP63 (13% MEx in PTvs.N zu 19% NGS). Cyclin D1 (CCND1) war in knapp einem Drittel der NGS amplifiziert, aber nur in 7% der MEx (PTvs.N) überexprimiert. Auch MYC war nur in prozentual halb so vielen Tumoren der MEx hochreguliert (7%) wie es in NGS amplifiziert war (14%). HRAS war in NGS bei 5% der Tumore als mutiert (nicht als amplifiziert) angegeben, während es in 13% der in MEx untersuchten Primärtumore signifikant hochreguliert war, bei vielen anderen allerdings eher herabreguliert. Bei PI3KCA war in vielen MEx sowohl in PTvs.N als auch in Mvs.N tendenziell eher hochreguliert, jedoch nur in einem Tumor über 20%, im Gegensatz zu den NGS, in denen eine Amplifikation oder Mutation bei 34% der Tumore vorlag. Da bei 84% der Tumore in NGS eine Mutation in TP53 vorlag, wurde auch der Expressionsstatus in MEx gemessen und zeigte sich in Patient 11 signifikant hochreguliert, in allen anderen Patienten aber eher herabreguliert (**Abb. 53**).

3.1.1 “Identification of novel disease-relevant signaling-pathways in head and neck cancer by genome-wide transcriptional profiling”

Dorothee Goesswein, Angelina Hahlbrock, Max Deichelbohrer, Cecilia Vallet, Désirée Wünsch, Dana Westmeier, Günter Schneider, Andreas Hildebrandt, Shirley K. Knauer, Negusse Habtemichael und Roland H. Stauber.

*Oncotarget (in revision).*²

Zusammenfassung

Kopf-Hals-Tumore (HNC) sind mit jährlich etwa 640.000 neuen Fällen weltweit als sechsthäufigste Krebsart seit Jahren auf dem Vormarsch ². Während zum Milleniumswechsel global etwa 300.000 Menschen an Kopf-Hals-Tumoren starben, waren es im Jahr 2015 bereits über 400.000 ¹. Üblicherweise werden HNC erst in fortgeschrittenen Stadien mit loko-regionalen und/oder Fernmetastasen diagnostiziert, was zu einer 5-Jahres-Überlebensrate von weniger als 50% führt ^{139,140}. Die vorwiegende Form von Kopf-Hals-Tumoren ist mit über 90% das Kopf-Hals Plattenepithelkarzinom (HNSCC), das zu den aggressivsten und invasivsten Krebsarten gehört ^{139,140}. Verschiedene Gewebe des Aerodigestiv-Traktes wie die Nasenhöhle, der Sinus, Lippen, Mund, Speicheldrüsen, Larynx und Pharynx können davon betroffen sein ^{140,185}. Die Hauptrisikofaktoren, die zur Entwicklung von HNSCC beitragen können, sind Tabak- und/oder Alkoholkonsum sowie Infektionen mit den Hochrisiko-Typen 16, 18, 31 und 33 des humanen Papillomavirus (HPV)^{139,186,187}. Trotz großer Wissensfortschritte ist die Behandlung von HNSCC immer noch häufig mit kosmetischen Deformationen und funktionellen Beeinträchtigungen des Aerodigestivtraktes verbunden, die das Schlucken und Atmen erschweren können, aber auch die Stimme, das Hören und Riechen negativ betreffen können ^{139,185,188,189}. Die Überlebensraten von HNSCC Patienten haben sich seit Jahren nicht wesentlich verbessert, hauptsächlich auf Grund einer späten Vorstellung und Diagnose des Patienten sowie dem Fehlen von probaten Biomarkern und dazugehörigen Therapien für individuelle Behandlungsansätze ¹⁸⁹⁻¹⁹¹.

² Eine Kopie des Manuskripts kann auf Grund der aktuellen Lizenzierung nicht abgedruckt werden.

Zur Identifikation neuer Biomarker wurden in dieser Studie die Genexpressionsprofile von primären Kopf-Hals-Plattenepithelkarzinomen mit ihren korrespondierenden Lymphknoten-Metastasen gegenüber dem jeweiligen Normalgewebe des einzelnen Patienten untersucht. Die erhaltenen Daten wurden durch unabhängige qPCR-, RT-PCR- und Western Blot-Untersuchungen validiert. Differenziell regulierte Gene wurden bioinformatisch mittels strenger Filterkriterien identifiziert, um zwischen normaler Schleimhaut, Primärtumor und Lymphknotenmetastasen unterscheiden zu können. Involvierte Signalwege waren thematisch angelehnt an die epitheliale-zu-mesenchymale Transition, metabolische Veränderungen, Immunreaktivität wie auch Protease-Netzwerke.

Das bisher nicht mit Kopf-Hals-Tumoren assoziierte Cytokeratin 24 (KRT24), das in epithelialen Zellen maßgeblich zur Zellpolarität und als Intermediärfilament zum Zytoskelett beiträgt, war in den Primärtumoren im Vergleich zum Normalgewebe stark herabreguliert. Diese verminderte Expression wurde sowohl in bisher vorhandenen, als auch in neuen Primärtumorproben auf mRNA und Proteinlevel mit Hilfe eines eigens kreierte Antikörpers bestätigt. Die KRT24 Sequenz wurde kloniert und in HNSCC Zelllinien überexprimiert, welche in der Folge ein deutlich reduziertes Wachstum zeigten. Um die Wirkung als potentiell Tumor suppressor weiterhin zu untersuchen, wurden entweder Leervektoren oder Vektoren mit KRT24 in FaDu Zellen transfiziert, auf ihre Expression mittels FACS sortiert und in die Flanken von Mäusen injiziert. In diesen Xenograft-Experimenten zeigte sich eine deutliche Reduktion des Wachstums der KRT24-positiven Tumore im Vergleich zur Kontrolle.

Die Ergebnisse legen zusammenfassend nahe, dass eine KRT24-Überexpression die Proliferation und potentiell auch das metastatische Potential in Kopf-Hals-Tumoren deutlich reduziert und KRT24 demnach als Tumorsuppressor agieren könnte.

3.1.2 “Expressional analysis of disease-relevant signalling pathways in primary tumours and metastasis of head and neck cancers”

Dorothee Goesswein, Negusse Habtemichael, Aslihan Gerhold-Ay, Johanna Mazur, Désirée Wunsch, Shirley K. Knauer, Julian Künzel, Christoph Matthias, Sebastian Strieth, Roland H. Stauber.

*Scientific Reports, 2018.*³

Zusammenfassung

Kopf-Hals Plattenepithelkarzinome (HNSCC) entstehen an unterschiedlichen anatomischen Bereichen, darunter die Mundhöhle, der Hypopharynx, Oropharynx, Larynx und Nasopharynx^{2,139}. Tabak- und Alkoholkonsum sowie eine Infektion mit den onkogenen Varianten des Humanen Papillomavirus HPV-16, HPV-18, HPV-31 und HPV-33 sind die Hauptrisikofaktoren dieser Erkrankung¹³⁹⁻¹⁴¹. Früh diagnostiziert können HNSCC oft kurativ resiziert werden, zusätzlich besteht die Option einer lokalen Radiotherapie. Für fortgeschrittene Stadien ist die Behandlung multimodal und kann Operationen, meistens kombiniert mit Chemoradiatio beinhalten, oder definitiver Chemoradiatio. Für Rezidive und metastatische Erkrankungen sind Chemo- und/oder EGFR-Immuntherapien empfohlen. Alle Behandlungen sind primär durch das Gewebe, in dem der Tumor entstanden ist, die TNM Klassifikation und andere pathologische Eigenschaften des Tumors indiziert, aber limitiert durch den allgemeinen Gesundheitszustand des Patienten^{139,142}.

Grundsätzlich verändern sich Krebstherapien derzeit mehr und mehr weg von systemisch wirkenden, zytotoxischen Agenzien hin zu individuellen Ansätzen, die spezifische Eigenschaften des jeweiligen Tumors angreifen. Für solche personalisierten Therapien sind diagnostische Tests essentiell, welche die Wirksamkeit bestimmter Therapien sowie deren Nebenwirkungen vorhersagen und dabei idealerweise den molekularbiologischen Wissensschatz über die Kanzerogenese mit einbeziehen¹⁴³. Für Kopf-Hals-Plattenepithelkarzinome gibt es derzeit noch keine Biomarker mit prognostischer oder

³ Eine Kopie des Artikels wird im Folgenden abgedruckt unter der *Creative Commons Attribution 4.0 International License*. <https://creativecommons.org/licenses/by/4.0/>

prädiktiver Signifikanz, obwohl bereits einige neue zielgerichtete Therapien in klinischen Studien der Phase II und III getestet werden, die behandelnde Ärzte früher oder später in Entscheidungsnöte bringen werden.

Der Mangel an diagnostischen Biomarkern lässt sich vor allem darauf zurückführen, dass Daten aus evidenzbasierten, klinischen Studien fehlen, welche die diagnostische Validität eines Biomarkers für eine Zulassung hinreichend untermauern können. Das statistische Gesamtüberleben und auf die Lebensqualität adjustierte Lebensjahre (QALY) sind nicht ausreichend, um hochspezifische neue Krebstherapien und ihre begleitenden Diagnostika abschließend zu bewerten ^{164,165}.

In dieser Expressions-basierten Microarray Studie wurden Primärtumore und Lymphknotenmetastasen von 15 Patienten mit korrespondierendem Normalgewebe jedes Patienten verglichen und hinsichtlich ihres diagnostischen und prognostischen Potentials bewertet. Kürzlich veröffentlichte Next-Generation-Sequenzierungsdaten wurden mit den Expressionsdaten verglichen und in einen größeren Zusammenhang gebracht. Im Speziellen wurden Kombinationen hochregulierter Rezeptortyrosinkinasen für einzelne Patienten untersucht, die verschiedene Behandlungsoptionen für jeden Patienten anregen. Für mehrere Patienten besteht auf Grund bisheriger molekularbiologischer Erkenntnisse eine hohe Wahrscheinlichkeit, dass die verschiedenen Tumore großes kompensatorisches Potential und diverse Resistenzmechanismen gegen bestimmte zielgerichtete Therapien besitzen.

Evidenzbasierte Medizin erfordert eine große Quantität valider empirischer Daten aus klinischen Studien, welche durch stringente Datensammlung und robuste, reproduzierbare Assays erzeugt werden muss ¹⁷⁰. Folgerichtig müssen Methoden stark vereinfacht und standardisiert werden, um die gleichzeitige Entwicklung von Biomarkern parallel zu klinischen Studien der dazugehörigen Krebsmedikamente zu ermöglichen ¹⁶⁵. Die Microarray-Technologie wird derzeit über nanotechnologische und mikrofluidische Ansätze für ein On-Chip-Verfahren miniaturisiert und für die Nutzung in Automaten für einen kosteneffektiven Gebrauch in der klinischen Diagnostik weiterentwickelt ¹⁷¹. Entsprechende Ansätze wurden bereits zur Pathogen-Detektion getestet ¹⁷².

In dieser Studie konnte das Potential der DNA-Hybridisierung für eine diagnostische Evaluation dargestellt werden. Die Nutzung standardisierter, zielgerichteter Microarray-Analysen zur Auswahl potentieller diagnostischer Marker könnte die parallele Entwicklung von Biomarkern neben klinischen Studien von zielgerichteten Medikamenten deutlich beschleunigen.

SCIENTIFIC REPORTS

OPEN

Expressional analysis of disease-relevant signalling-pathways in primary tumours and metastasis of head and neck cancers

Dorothee Goesswein¹, Negusse Habtemichael¹, Aslihan Gerhold-Ay², Johanna Mazur², Désirée Wünsch¹, Shirley K. Knauer³, Julian Künzel¹, Christoph Matthias¹, Sebastian Strieth¹ & Roland H. Stauber¹

Head and neck squamous cell carcinoma (HNSCC) often metastasize to lymph nodes resulting in poor prognosis for patients. Unfortunately, the underlying molecular mechanisms contributing to tumour aggressiveness, recurrences, and metastasis are still not fully understood. However, such knowledge is key to identify biomarkers and drug targets to improve prognosis and treatments. Consequently, we performed genome-wide expression profiling of 15 primary HNSCCs compared to corresponding lymph node metastases and non-malignant tissue of the same patient. Differentially expressed genes were bioinformatically exploited applying stringent filter criteria, allowing the discrimination between normal mucosa, primary tumours, and metastases. Signalling networks involved in invasion contain remodelling of the extracellular matrix, hypoxia-induced transcriptional modulation, and the recruitment of cancer associated fibroblasts, ultimately converging into a broad activation of PI3K/AKT-signalling pathway in lymph node metastasis. Notably, when we compared the diagnostic and prognostic value of sequencing data with our expression analysis significant differences were uncovered concerning the expression of the receptor tyrosine kinases *EGFR* and *ERBB2*, as well as other oncogenic regulators. Particularly, upregulated receptor tyrosine kinase combinations for individual patients varied, implying potential compensatory and resistance mechanisms against specific targeted therapies. Collectively, we here provide unique transcriptional profiles for disease predictions and comprehensively analyse involved signalling pathways in advanced HNSCC.

With over 500,000 new cases per year worldwide, head and neck squamous cell carcinoma (HNSCC) develops from multiple anatomic subsites, including oral cavity, hypopharynx, oropharynx, larynx and nasopharynx^{1,2}. Main risk factors include tobacco exposure, together with alcohol use, as well as infection with potentially oncogenic viruses, such as HPV-16, HPV-18, HPV-31, and HPV-33^{1,3,4}. Moreover, the increasing (ab)use of so called 'e-liquids', as a more 'healthy' way of smoking, might need to be considered as an additional risk factor. However, conclusive studies are missing so far. When discovered in early-stage, HNSCC are treated with surgery and/or radiotherapy; for advanced stages treatment is multimodal¹. Surgery is used when possible, usually combined with adjuvant chemoradiation, or definitive chemoradiation¹. For recurrent and metastatic diseases, chemotherapy is applied with the option of alternative immunotherapy when chemotherapy is not tolerated¹. All treatment is indicated mainly by primary site, TNM classification and other pathologic features limited by performance status^{1,5}. Cancer therapies are changing gradually, shifting from systemic cytotoxic drugs to individual approaches. Essential for these personalized therapies are diagnostic tests predicting the clinical outcome of certain therapies using the vast knowledge of cancer origins and maintenance obtained by molecular biology⁶.

¹Molecular and Cellular Oncology, ENT/University Hospital of Mainz, Mainz, 55131, Germany. ²Institute for Medical Biostatistics, Epidemiology and Informatics (IMBEI), University of Mainz Medical Center, Mainz, 55101, Germany. ³Institute for Molecular Biology, Centre for Medical Biotechnology, University of Duisburg-Essen, Essen, 45117, Germany. Correspondence and requests for materials should be addressed to D.G. (email: goesswein@uni-mainz.de) or R.H.S. (email: Roland.Stauber@unimedizin-mainz.de)

As the accumulation of mutations in tumours from patients with lymph node metastasis is not increased compared to patients without lymph node involvement⁷, other alterations have been suggested⁸. Expression profiling has been used to distinguish between cancer subtypes and predict disease outcome in different types of cancer^{9–13}. Previous studies allowed the categorization of HNSCC tumours into distinct subtypes¹⁴, discovered a gene expression signature associated with recurrent disease¹⁵, or suggested genes with diagnostic or prognostic potential^{16–23}. Microarray analysis performed on normal oral epithelium compared to HNSCC elucidated the roles of microenvironment remodelling and immune responses in survival of HNSCC²⁴. The only prognostic marker for HNSCC is the HPV status of the tumour, and predictive biomarkers to influence treatment selection are completely lacking^{1,25}. Although HPV positive tumours differ significantly in their expression profiles, treatment options do not depend on the HPV status of the tumour¹. Several new therapies targeting growth factor receptors or downstream signalling pathways are currently tested in clinical trials²⁶. Without matching companion diagnostics, physicians will soon be challenged in choosing the most suitable therapy for each patient.

In this study, we analysed differentially expressed genes of HNSCC primary tumours compared to lymph node metastases in order to explore new biomarker candidates and involved signalling pathways. In contrast to other studies concentrating only on the analysis of primary tumours, we additionally focused on identifying the changes that lead to the dissemination of the primary tumour to lymph node metastasis. Gene expression microarray technology was applied to a group of 15 HNSCC tumours, their corresponding normal tissue mucosa and lymph node metastases resected by surgery. To facilitate data interpretation, statistical gene-set-enrichment methods and Qiagen's software Ingenuity (Pathway Analysis, Comparison and Upstream Analysis, Biomarker Discovery) were used to intersect expressed genes with sets of genes associated with particular biological functions or pathways (as reviewed in²⁷). We focused on obtaining a global view of differential gene expression in HNSCC compared to normal tissue mucosa, the primary tumour and corresponding lymph node metastases. The identification of characteristic gene expression signatures may potentially be exploited for diagnostic, prognostic or therapeutic purposes and promotes the reliability of the corresponding method.

Results

Validation of the gene expression data generated by oligonucleotide chip-arrays. In order to analyse gene expression, RNA from primary tumours (PT), lymph node metastases (M), and corresponding normal tissue mucosa (N) was subjected to microarray analysis using Affymetrix HG-U133A chips. The verification of the obtained data quality by independent experimental approaches is a prerequisite for further analytical approaches in all 'omics' technologies. Accordingly, we confirmed our chip-array data by quantitative real-time PCR (qRT-PCR) and reverse transcription PCR (RT-PCR) for selected representative genes and samples. RT-PCR was performed to analyse expression patterns of randomly chosen *CLCA4*, *FNI*, *POSTN*, *KRT24* and *PRR4* in primary HNSCC vs. their corresponding normal tissue mucosa (PTvs.N, Fig. S1A) and *ARPGAP25*, *FAIM3*, *RASGRP2*, *LYPD3*, *SERPINB4* and *TP73L* in lymph node metastasis vs. their corresponding primary HNSCC (Mvs.PT, Fig. S1B). *FNI*, *POSTN* and *OLR1* genes were confirmed as up-regulated, *KRT24*, *KRT4* and *PRR4* as down-regulated in PTvs.N. *ARPGAP25*, *FAIM3* and *RASGRP2* genes were up-regulated, whereas *LYPD3*, *SERPINB4* and *TP73L* were down-regulated in Mvs.PT. Additionally, quantitative PCR analyses were performed for *ARGHAP25* (Fig. S2), *SERPINB4* (Fig. S3) and *PRR4* (Fig. S4). All results of qRT-PCR analyses were in accordance to the microarray data. In summary, the conducted controls underline the quality and reliability of our obtained GeneChip array data, fulfilling all prerequisites for further bioinformatics exploitation, model building and in depth functional and clinical analysis.

Identification of significantly upregulated genes in primary tumours and corresponding lymph node metastases. After normalization, we calculated the expression differences over all patients to compare primary tumours to normal mucosa (PTvs.N) and lymph node metastasis to primary tumours (Mvs.PT). By sorting the resulting lists using stringent filter criteria (first by p-value, second by \log_2 FC), we obtained four TOP 30 lists for highly upregulated PTvs.N (Table S4), downregulated PTvs.N (Table S5), upregulated Mvs.PT (Table S6) and downregulated Mvs.PT (Table S7).

The 45 data sets were subjected to supervised (adjusted to the origin of the primary site) average-linkage hierarchical clustering. Normal tissue mucosa and primary tumours grouped separately on a clustering dendrogram (Fig. 1A) as well as normal tissue mucosa and metastasis (data not shown). Comparison of primary tumours vs. metastasis showed only marginal differences, displaying a rather gradual transition (Fig. 2A). As expected, these results show that metastasis and primary tumour were closely related and could not be clearly distinguished from each other.

In order to pinpoint the most important upregulated genes for either PTvs.N or Mvs.PT, we restricted the dataset to a p-value of lower than 0.0001 (corresponds to a FDR of 0.003 in PTvs.N and 0.004 in Mvs.PT) and a \log_2 fold change (FC) higher than 1.5 (equates to a FC of 2.83). This data restriction resulted in a signature gene set of 38 upregulated genes for PTvs.N (Fig. 1B and Table S8), and 62 upregulated genes for Mvs.PT respectively (Fig. 2B and Table S9).

Several proteins involved in the (re-)organization and assembly of the extracellular matrix were significantly overexpressed in all primary tumours compared to normal tissue mucosa (Table S8): Matrix Metalloproteinases, Collagens, Asporin, Fibronectin, Laminin, Lysyl Oxidases, Microfibrillar-Associated Protein, Periostin, Serpin Inhibitors, Osteonectin and Osteopontin. Higher expression of genes like *KIF23*, *UBE2C*, *MCM7*, *NUP107*, *RSRC1*, *MAD2*, *BTAK*, *RFC5*, *ECT2* and *CENPA* was also found in metastatic tumours reports from other studies¹⁴. The latter 5 genes were also present in the distant recurrence predictor of van't Veer *et al.* for breast cancer²⁸. Regarding the comparison of Mvs.PT (Table S9), we identified upregulation of several genes related to the immune response, cellular metabolism, several membrane proteins, as well as genes interacting with or regulating GTPases.

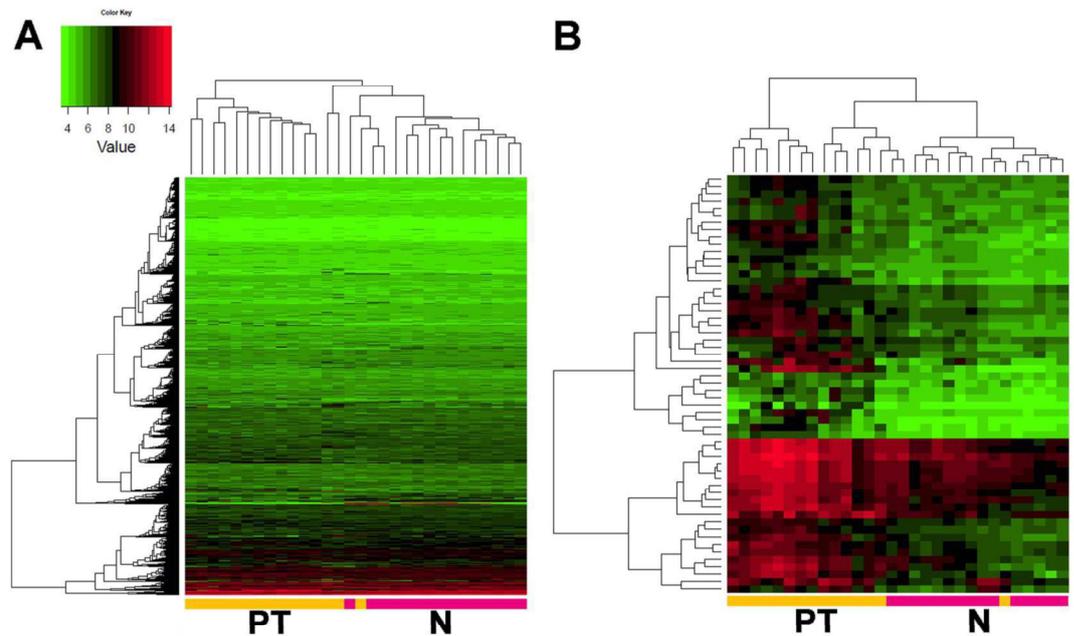


Figure 1. Gene set cluster analysis of 15 HNSCC samples and their 15 corresponding normal mucosa allows profiling of disease progression by identifying a condensed list of 38 most significant genes. Supervised two-way hierarchical clustering and gene tree representation adjusted to the primary site of all genes (A) and differentially regulated genes (B) allow separating primary tumours and normal mucosa. X-axis' represent patient samples; y-axis' represent the list of probe sets grouped by similarity using Pearson correlation. Intensities recorded for genes are colour coded: Green – low expression, Red – high expression, Black – medium expression. Orange bar: primary tumour samples; pink bar: normal mucosa samples. (A) Heatmap of all probesets shows the comparison of primary tumours and normal tissue mucosa (PT vs. N). (B) Stringent filter criteria ($\log_2FC \geq 1.5$ and $p\text{-value} < 0.0001$, corresponding to a FDR of 0.003) were used to reduce the probesets to the most significant for PT vs. N, leading to a condensed list of 38 genes able to distinguish diseased from non-malignant tissue. Signature genes are listed separately in Table S8.

Ingenuity Analysis of significantly up-/downregulated genes. Previously reported correlations were investigated thoroughly using the Ingenuity Pathway Analysis tool to reinforce the validity of our analysis. We applied the same restrictions as before, but included downregulated genes ($p\text{-value}$ of lower than 0.0001 and higher than a \log_2FC of $|1.5|$) to our Ingenuity analysis. Comparing gene expression profiles of PT vs. N, we identified several disease/function networks for cancer and connective tissue disorders with a score higher than 20 (Fig. 3 and Table S10). For the dataset M vs. PT, top networks included one network for immune response, as well as several networks for cell movement, cell-to-cell signalling and interaction with a score higher than 20 (Fig. 4 and Table S11). All top networks are illustrated in detail in Supplementary Figures S12–S27.

Ingenuity Analysis of upstream regulators. To further understand the context of activated networks, we used the upstream analysis tool of the Ingenuity software. Among the analysed activated upstream regulators for PT vs. N ($z\text{-score}$ higher than 2) we found glycoprotein CD44, a stem cell marker for HNSCC²⁹ and several growth factors like FGF2, TGFB1 and TGFB3 (Table 1A). Notably, there were only two significantly activated upstream regulators in M vs. PT, namely the PI3K family which has been implicated with a more aggressive and invasive tumour type^{8,30} and the cytokine $IFN\gamma$ which has previously been suggested as a salivary biomarker for early diagnosis of oral cancer³¹ (Table 2).

Surprisingly, inhibited upstream regulators ($z\text{-score}$ lower than -2) for PT vs. N included the oncogene *MYC* (Table 1). There were no significantly inhibited upstream regulators for M vs. PT (Table 2). The Ingenuity Comparison Analysis of selected canonical pathways showed a high relevance for cancer pathways and tissue morphology in both primary tumours as well as metastases, whereas cell movement and inflammatory response were predicted more prominent just in metastases (Figs S8 and S9).

Further downstream analysis revealed several interesting networks, one of them being activated by Fibronectin 1 (FN1) in PT vs. N (Fig. S5). Since FN1 is involved in several cellular processes, such as remodelling of the extracellular matrix, it is of general note to investigate the expression and activation status of its downstream targets. This will allow the identification of potential activation processes and prediction of corresponding cellular processes. Downstream targets of FN1 include VEGFA, the Jnk Family, the Tgf beta family, HRAS and the ERK1/2 family, which collectively activate diverse cellular functions via the transcription factor ETS1 (expression of cytokine and chemokine genes), beta-Catenin (key downstream component of Wnt signalling), the transcription factor Jun-B (regulation of gene activity following growth factor response), the corresponding nuclear phosphoprotein c-Fos and Ap-1, as well as the transcription factor SP1. All of these genes were activated and/or overexpressed in our dataset PT vs. N (Fig. S5).

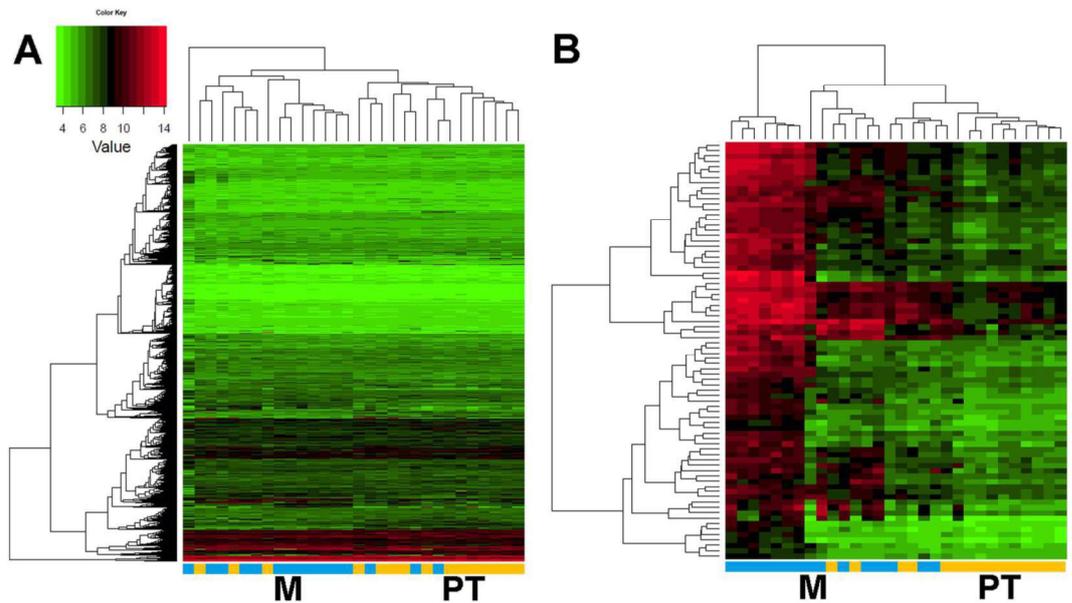


Figure 2. Gene set cluster analysis of 15 lymph node metastasis and their 15 corresponding primary tumours allows profiling of disease progression by identifying a condensed list of 62 most significant genes. Supervised two-way hierarchical clustering and gene tree representation adjusted to the primary site of all genes (A) and differentially regulated genes (B) allow distinguishing lymph node metastasis and primary tumours. X-axis' represent patient samples; y-axis' represent the list of probe sets grouped by similarity using Pearson correlation. Intensities recorded for genes are colour coded: Green – low expression, Red – high expression, Black – medium expression. Orange bar: primary tumour samples; blue bar: lymph node metastasis samples. (A) Heatmap of all probesets shows the comparison of lymph node metastasis to primary tumours (Mvs.PT). (B) Stringent filter criteria ($\log_2FC \geq 1.5$ and $p\text{-value} < 0.0001$, corresponding to a FDR of 0.004) were used to reduce the probesets for Mvs.PT, leading to a condensed list of 62 genes most significantly changed in primary tumours to their corresponding lymph node metastasis. Signature genes are listed separately in Table S9.

The PI3K family has been shown to be essentially important for metastatic transformation³⁰ and is the most significant activated upstream activator in Mvs.PT (Table 2). The PI3K group activates NFkB via Akt and promotes the release of MMP9, which is the most important matrix metalloprotease in degrading the basement membrane, which is also predicted by Ingenuity (Fig. S6). The NFkB complex usually activates DNA damage response via RELA and RELB, blocking PPARG and the Jnk family but this could not be confirmed by our Ingenuity prediction for Mvs.PT (Fig. S7). The nuclear receptor PPAR gamma is activated, which should block pro-inflammatory NFkB-mediated responses³². In contrast, tumour-promoting cytokines, such as TNF- α and IL10 are upregulated (Fig. S7) underlining the complexity and crosslinking of malignant signalling pathways as previously described³³.

Deregulated canonical pathways are essential for tumour maintenance and important factors of malignancy. For both primary tumours (Fig. S8) and metastases (Fig. S9), the pathways of Nicotine Degradation and Ethanol Degradation were activated (Figs S10 and S11). Since all patients were smokers and alcohol consumers (Table S1), those findings are in accordance to our expectations.

Comparison of amplification profiles of next generation genome sequencing data to our analysis.

Since next generation sequencing (NGS) has had a major impact on cancer research, we compared recent NGS amplification³⁴ and PARADIGM prediction data³⁴ based on RNA sequencing data from the Broad Institute TCGA Genome database³⁵ to our microarray analysis. Here, a \log_2FC higher than 0.263 was used as an overexpression tendency (corresponding to a 20% higher expression). For many genes, we found a concordant tendency of their regulation when we compared the genomic amplification and PARADIGM percentages (AMP) published by The Cancer Genome Atlas (TCGA) project³⁶ and our expressional data (Table 3). Especially *FGFR1* (13% in PTvs.N and 7% in Mvs.PT to 10% in AMP), *FGFR3* (7% in PTvs.N and Mvs.PT to 2% in AMP), *DDR2* in lymph node metastasis (13% in Mvs.PT to 3% in AMP) and *MET* (7% in PTvs.N and Mvs.PT to 2% AMP, FC 2.4; Fig. S28) seem to be similarly overexpressed. Remarkably, we did not find a significant overexpression of *EGFR*, neither in the primary tumour, nor in their corresponding lymph node metastasis. In contrast, RNA-Seq-based amplification data suggested an AMP of 15% for *EGFR*. The receptor *Her-2 (ERBB2)*, which was amplified in 4% of specimens in TCGA cohort, was mostly downregulated in our expression data as well as the RNA-Seq data (FC 0.479; Fig. S28). *MYC*, which was shown to be inactivated in our Ingenuity upstream analysis, was overexpressed in 7% of patients, compared to 14% amplification in TCGA samples. *E2F1* and *NFE2L2* overexpression values were both higher than expected by TCGA data, and even more elevated in metastasis than in primary tumours. We also found a higher expression for *EPHA2*, *DDR2* and *PIK3R1* compared to AMP published by TCGA. No significant differences between the fold changes of the PTvs.N to the Mvs.N dataset could be detected for *FGFR1*,

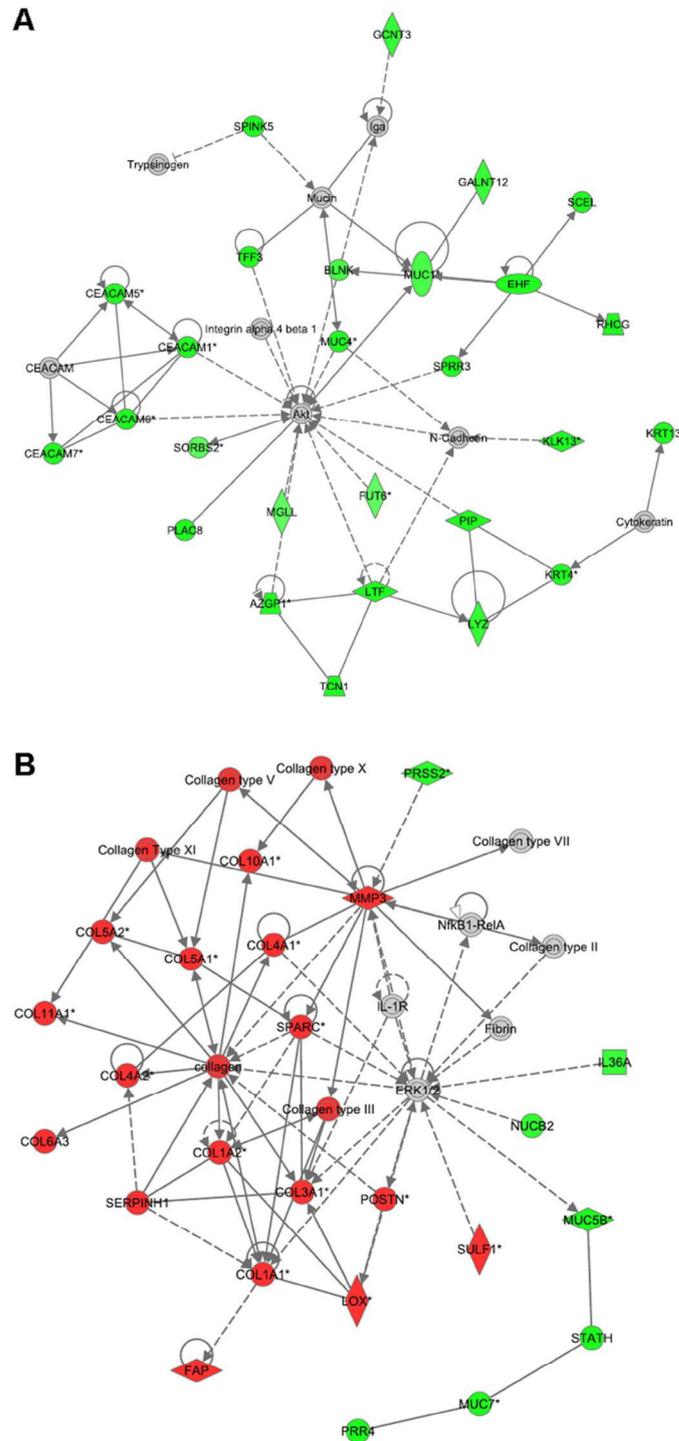


Figure 3. Top networks of differentially expressed genes in PT vs. N as identified by Ingenuity Pathway Analysis show strong association to cancer and connective tissue disorders. For network analyses a threshold of $\log_2FC > |1.5|$ with a p -value < 0.0001 (corresponding to a FDR of 0.003) was used, giving an overview of involved genes within the networks and their differential expression. Node colour indicates up-regulated (red), down-regulated (green) or not significantly changed according to threshold (grey) genes. Lines and arrows between nodes represent direct (solid lines) and indirect (dashed lines) interactions between molecules. All edges are supported by at least one reference from the literature or from canonical information stored in the Ingenuity Knowledge Base. Node shapes represent functional classes of gene products: square \rightarrow cytokine, triangle \rightarrow kinase, rectangle \rightarrow nuclear receptor, concentric circle \rightarrow group or complex, vertical diamond \rightarrow enzyme, horizontal diamond \rightarrow peptidase, trapezium \rightarrow transporter, vertical ellipse \rightarrow transmembrane receptor horizontal ellipse \rightarrow transcription regulator, circle \rightarrow other. **(A)** Top 01 significant PT vs. N, network (Cellular Development, Infectious Disease, Cancer) contains 27 focus genes. **(B)** Top 02 significant PT vs. N network (Connective Tissue Disorders, Dermatological Diseases and Conditions, Cellular Assembly and Organization) contains 24 focus genes.

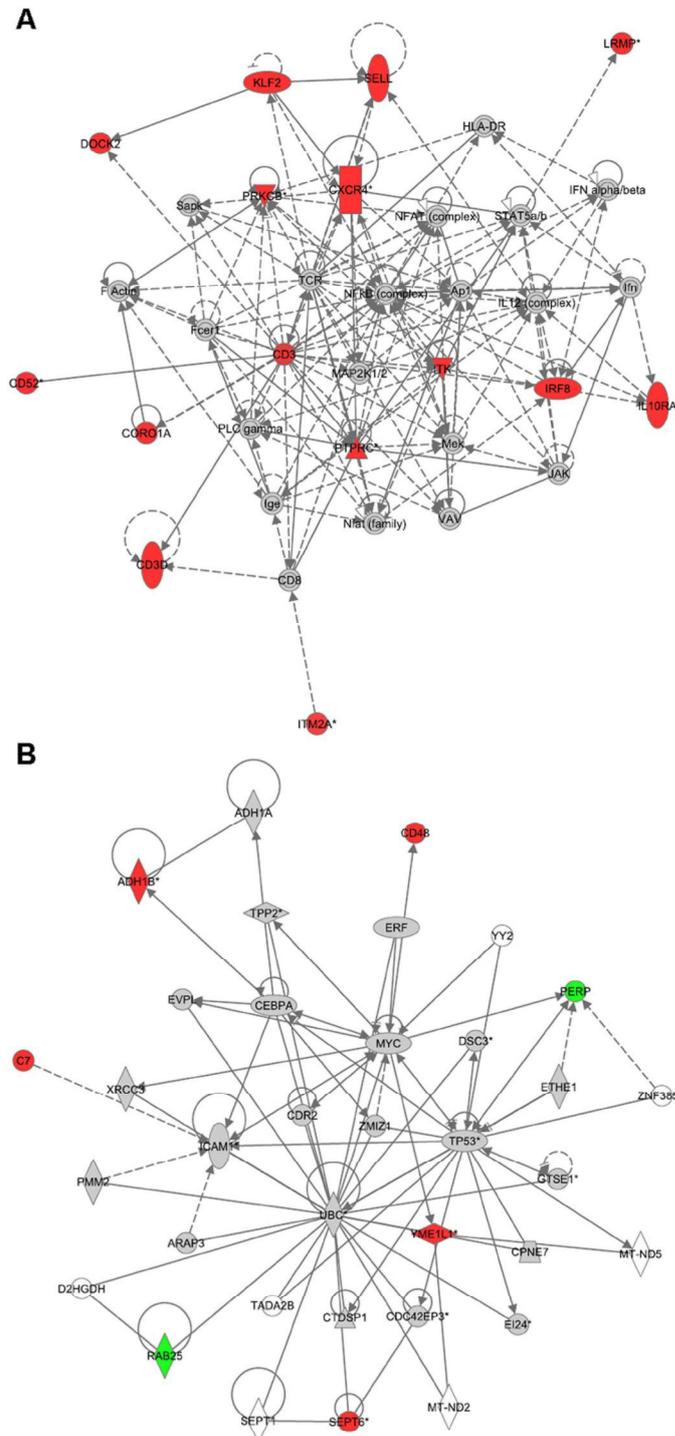


Figure 4. Top networks of differentially expressed genes in *Mvs.PT* as identified by Ingenuity Pathway Analysis show strong association to cancer and connective tissue disorders. For network analyses a threshold of $\log_2FC > |1.5|$ with a p -value < 0.0001 (corresponding to a FDR of 0.004) was used, giving an overview of involved genes within the networks and their differential expression. Node colour indicates up-regulated (red), down-regulated (green) or not significantly changed according to threshold (grey) genes. Lines and arrows between nodes represent direct (solid lines) and indirect (dashed lines) interactions between molecules. All edges are supported by at least one reference from the literature or from canonical information stored in the Ingenuity Knowledge Base. Node shapes represent functional classes of gene products: square \rightarrow cytokine, triangle \rightarrow kinase, rectangle \rightarrow nuclear receptor, concentric circle \rightarrow group or complex, vertical diamond \rightarrow enzyme, horizontal diamond \rightarrow peptidase, trapezium \rightarrow transporter, vertical ellipse \rightarrow transmembrane receptor horizontal ellipse \rightarrow transcription regulator, circle \rightarrow other. (A) Top 03 significant *Mvs.PT* network (Cell-To-Cell Signalling and Interaction, Haematological System Development and Function, Tissue Morphology) contains 14 focus genes. (B) Top 07 significant *Mvs.PT* network (Cancer, Endocrine System Disorders, Developmental Disorder) contains 7 focus genes.

Upstream Regulator	Log Ratio	Molecule Type	Predicted Activation State	Activation z-score	p-value of overlap
CD44	0.382	enzyme	Activated	2.137	6.39E-09
FGF2	-0.133	growth factor	Activated	2.261	1.08E-10
IL1A	0.090	cytokine	Activated	2.298	1.51E-11
TGFB3	0.471	growth factor	Activated	2.416	3.72E-09
Ap1		complex	Activated	2.581	9.43E-08
CTGF	0.357	growth factor	Activated	2.930	2.80E-08
TGFB1	0.688	growth factor	Activated	3.316	2.75E-09
EHF	-2.432	transcription regulator	Inhibited	-2.333	6.02E-08
MYC	-0.071	transcription regulator	Inhibited	-2.433	4.50E-10
WISP2	-0.000	growth factor	Inhibited	-2.630	1.64E-07
ABCB4	-0.054	transporter	Inhibited	-2.631	1.16E-08
CRIL		other	Inhibited	-2.804	4.24E-10
FBN1	0.528	other	Inhibited	-2.804	5.11E-12
estrogen receptor		group	Inhibited	-2.937	2.28E-10
SPDEF	-0.558	transcription regulator	Inhibited	-3.000	3.99E-08
AHR	0.269	ligand-dependent nuclear receptor	Inhibited	-3.063	7.40E-10

Table 1. Activated/Inhibited Upstream Regulators (Ingenuity) PTvs.N.

Upstream Regulator	Log Ratio	Molecule Type	Predicted Activation State	Activation z-score	p-value of overlap
PI3K (family)		group	Activated	2.216	3.43E-05
IFNG	-0.038	cytokine	Activated	2.449	1.48E-05

Table 2. Activated/Inhibited Upstream Regulators (Ingenuity) Mvs.PT.

Group	Gene	Upregulation of 20% or higher		TCGA RNA-Seq* AMP**	
		% in PTvs.N	% in Mvs.N	HPV(-)	HPV(+)
(R)TK	EGFR	0	0	15↑	6↑
	ERBB2	0	0	4↑	3↑
	FGFR1	13	7	10↑	0
	FGFR3	7	7	2↑	11↑
	IGF1R	0	0	4↑	0
	EPHA2	0	0	4m	3m
	DDR2	0	13	3↑	6↑
	MET	7	7	2↑	0
Oncogenes	CCND1	7	0	31↑	3↑
	MYC	7	7	14↑	3↑
	HRAS	13	13	5m	0
PI(3)K	PIK3CA	7	0	34↑/m	56↑/m
Survival	BIRC2	7	0	7↑	3↑
	FADD	33	13	32↑	6↑
	E2F1	0	7	2↑	19↑
Other	TP63	13	0	19↑	28↑
	NFE2L2	0	0	14↑	0
	NF1	0	0	3↑/m	0

Table 3. Comparison of Microarray Expression Data with RNA-Seq Data. *TCGA RNA-Seq: The RNA sequencing data as published by The Cancer Genome Atlas Network³⁶; the Broad Institute TCGA Genome Data Analysis Center³⁵ and the tumorportal.org⁹¹. ** AMP: genomic amplification percentage (AMP) as published by The Cancer Genome Atlas Network³⁶. ↑ predicted as amplified; ↑/m predicted as amplified or mutated.

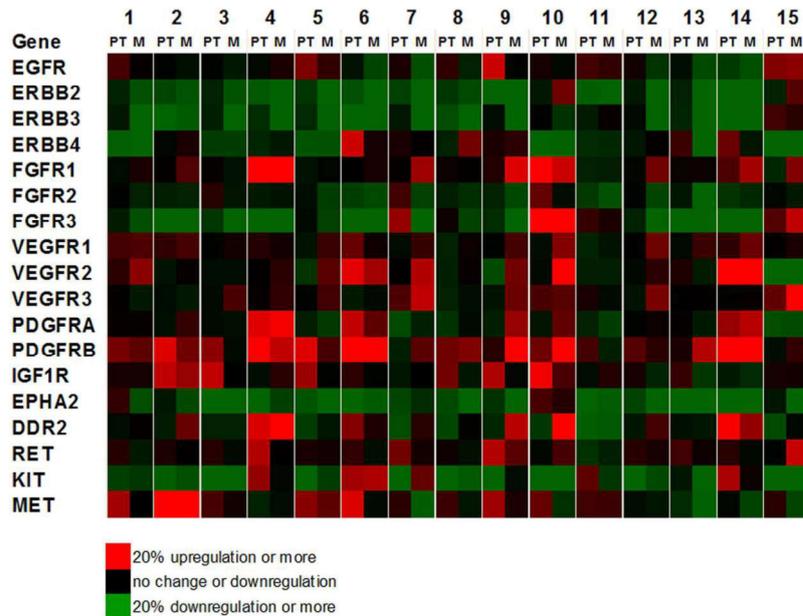


Figure 5. Comparison of Receptor Tyrosine Kinase (RTK) Expression of all patients. Differential expression values for selected genes were colour coded (bright red for an overexpression of $\geq 20\%$, bright green for a downregulation of $\geq 20\%$). In a ratio to their corresponding normal tissue mucosa, primary tumour values (PT) and metastasis values (M) are listed for each patient 1–15.

IGF1R, *FGFR2*, *MET*, *HRAS*, *PTEN* and *NF1*. *ERBB2*, *FGFR3* and *MYC* overexpression percentages were significantly higher in primary tumours than in metastases.

Comparison of individual Receptor Tyrosine Kinase (RTK) expression patterns. Due to their crucial role in HNSCC as upstream regulators of e.g. the phosphoinositol-3-kinase (PI3K) pathway mediating invasiveness and malignancy³⁰, we concentrated on further exploring the expression of growth factor receptors (receptor tyrosine kinases, RTKs). Remarkably, in our study PI3K was one of only two significantly activated upstream regulators in *Mvs.PT* (Table 2). Based on the list of amplified genes from the HNSCC TCGA project, we created an overview of relative RTK expression for *PTvs.N* and *Mvs.N* (Fig. 5). With the exception of two patients, all patients showed overexpression tendency of at least one RTK indicated by an upregulation of at least 20%. Most patients showed an upregulation of random combinations of at least two RTKs. For example, patient 6 showed an upregulation combination of *ERBB4*, *VEGFR2* and *PDGFRB*, and a slight upregulation of *MET* in *PTvs.N*. In patient 14, *VEGFR2*, *PDGFRB* and *DDR2* were higher expressed in comparison to normal mucosa. An expanded overview of RTK and oncogene expression can be found in the Supplementary Material (Fig. S29).

Discussion

Precision medicine requires individualized therapeutic approaches that fulfill the needs of each patient. Therefore, both prognostic as well as predictive markers are a crucial prerequisite. The thorough understanding of underlying mechanisms of tumorigenicity is not only the basis for the development of new drugs, but also for biomarker development.

In this study we analysed and compared the expression profiles of 15 primary tumours to their corresponding normal tissue mucosa and paired lymph node metastasis. In contrast, previous microarray-based studies of HNSCC have focused only on primary tumours versus normal mucosa or primary tumour patterns of expression^{16–21,37}. Since the number of patients is relatively low, p-values have to be interpreted with caution³⁸. We are aware of the limitations of this study and recommend replicate tests for their verification.

We have identified and verified several genes and pathways that are significantly overexpressed in HNSCC tumours with lymph node metastases. Based on this study, we here provide a model how genes involved in tumorigenesis and metastatic progression are differentially expressed in HNSCC, involving changes in the extracellular matrix, hypoxia and EMT (overview in Fig. 6). Our finding of highly upregulated matrix metalloproteases, integrins, collagens, fibronectin, laminin and other ECM proteins supports previous studies showing that degradation of the basement membrane and remodelling of the extracellular matrix are essential during the invasion process of a tumour^{39–41}.

Molecules Interacting with CasL-2 Proteins (MICALs) are monooxygenases that cause actin to depolymerize and are implicated in the regulation of cytoskeletal dynamics, intracellular trafficking, the docking and fusion of exocytotic vesicles and also have anti-apoptotic properties⁴². MICAL2, which was significantly upregulated in our *PTvs.N* dataset, has been identified as one of the most significantly upregulated genes in smoking-induced lung cancer⁴³ and has been associated with prostate cancer progression⁴⁴ as well as EMT, cancer growth and invasion⁴⁵, but has not been implicated with head and neck cancer, yet.

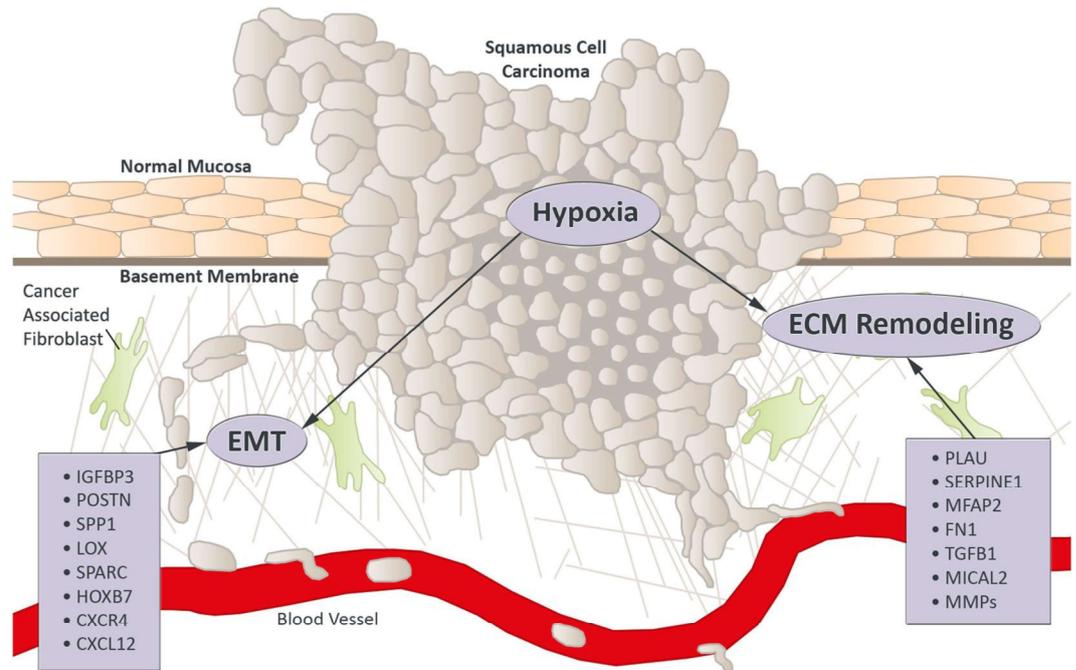


Figure 6. Potential prognostic biomarkers for HNSCC linked to hypoxia, EMT and ECM remodelling are directly connected. Hypoxic environments induce genes promoting an aggressive, invasive phenotype, mediated by epithelial to mesenchymal transition (EMT) and remodelling of the extracellular matrix (ECM). Matrix metalloproteinases (MMPs) and other proteases degrade the basement membrane and new collagen fibres are aligned by modifying enzymes to facilitate local invasion and recruitment of cancer associated fibroblasts.

The destruction and remodelling of the extracellular matrix not only drives the mechanical and proteolytic invasion of tumour cells, but also favours metastatic cell reprogramming and behaviour^{46–48}. Specialized Cancer Associated Fibroblasts (CAFs) can degrade various ECM proteins like collagens, basement membrane laminin and fibronectin⁴⁶, leading to a more aggressive behaviour of the cancer, lymph node metastasis and poor prognosis, as shown in oral cancer^{49,50}. When the Chemokine C-X-C Motif Receptor Type 4 (CXCR4), whose expression can be enhanced by hypoxia, binds Stromal Cell-derived Factor 1 (CXCL12), secreted by CAFs, release of MMP-9 to the tumour microenvironment is stimulated^{51,52}. In our data, we identified a switch from MMP1, 3 and 13 in primary tumours to MMP1, 9 and 12 in metastases, as well as a significant upregulation of CXCL12/CXCR4. Resulting MMP levels and the ratio of activated MMPs to total MMP concentration compared to adjacent normal tissue is positively correlated with lymph node involvement^{51,53}. MMP9 is the most versatile of the MMPs and can degrade various elements of the tumour microenvironment, like elastin, fibrillin, laminin, gelatin and collagens of type IV, V, XI and XVI^{52,54}.

Hypoxia is very common in solid tumours, especially in HNSCC⁵⁵, which has repeatedly been associated with poor prognosis⁵⁶. Additionally, hypoxic environments induce genes supporting stem cell maintenance, angiogenesis and epithelial to mesenchymal transition (EMT), leading to an aggressive, invasive phenotype^{57–59}. CXCR4 inhibitors have most recently been suggested as treatment for advanced and metastasizing HNSCC⁶⁰, and have previously been reported to have synergistic effects with anti-PDL1 therapies⁶¹.

Homeobox B7 (HOXB7) is a developmental gene we found highly upregulated in our P1 vs. N dataset. HOXB7 confers epithelial-mesenchymal transition in breast cancer⁶², promotes malignant progression by activation of TGF signalling⁶³ and lung metastasis⁶⁴ and renders breast cancer cells resistant to tamoxifen through activation of the EGFR pathway⁶⁵. HOXB7 overexpression is associated with poor prognosis⁶⁶ and has been suggested as a prognostic factor for oral squamous cell carcinoma⁶⁷. Migration of endothelial progenitor cells is mediated by CXCL12/CXCR4 via PI3K/Akt/eNOS signalling⁶⁸. The CXCL12/CXCR4 system also facilitates lymph node metastatic potential in oral squamous cell carcinoma by enabling EMT⁶⁹.

The phosphoinositol-3-kinase (PI3K) pathway is the most frequently altered oncogenic driver in head and neck cancer and its upstream regulators and downstream targets might be the most interesting targets for predictive biomarkers^{30,70}. While we could not show a single significantly upregulated growth factor receptor for all patients, we could confirm the global upregulation of the pathway in our Ingenuity Analysis. Interestingly, most patients showed a tendency for upregulation of at least two RTKs in varying combinations (Fig. 5). Several correlations involving sets of RTKs have previously been shown in different tumour entities. A meta-analysis for breast cancer has indicated the association of an overexpression of various RTKs with poor outcomes and suggested a clinical evaluation of combination agents against RTKs or relevant oncogenic nodes⁷¹. In breast cancer, MEK inhibition resulted in acute ERK activity loss and subsequent c-MYC degradation that induced expression and activation of PDGFRB, DDR2 and VEGFR2⁷², an RTK combination found in patients 6 (PT) and 10 (M) of our dataset (Fig. 5). It has been proposed that upregulation of FGFRs (as seen in patients 4, 9 and 10, as well

as to a minor degree in patients 7, 12, 14 and 15, Fig. 5) could be involved in the development of resistance to anti-VEGFR therapies, since both pathways are regulating tumour neoangiogenesis (as reviewed in⁷³). In breast cancer patients with *FGFR1* or *FGF3* amplification, multikinase FGFR/VEGFR inhibitors have shown promising activity⁷⁴. It has been suggested to treat RTK-driven tumours by hitting signalling nodes interconnecting core pathways, targeting c-Abl as a signalling node connecting MET and TP53^{75,76} (cp. patients 2, 5, 6 and 9, Figs 5 and S29).

In comparison with recent high throughput sequencing data, we showed a comprehensive list of upregulated genes for each patient. Regarding numerous targeted therapies for HNSCC currently in clinical studies²⁶, it will be essential to develop predictive biomarkers allowing the selection of an optimal therapy for each patient. Currently, molecular companion diagnostics are still lacking for all therapeutic approaches in HNSCC. For Cetuximab, the only targeted therapy currently in clinical use for HNSCC, the development of acneiform rash is still the best indicator for therapy response⁷⁷. *EGFR*-expression status, determined by immunohistochemistry, could not be verified as a predictive marker for *EGFR*-targeted therapy response^{78,79}, even though mechanisms of action and resistance mechanisms have been thoroughly investigated (as reviewed in^{80,81}). *EGFR* is amplified in normal mucosa of tobacco using patients, and might thereby be not the oncogenic driver, but rather a response to carcinogenic agents⁸¹. This might explain our results for the comparison of *EGFR* expression in primary tumour vs. normal mucosa (Table 3 and Fig. 5).

The combination of NGS with other approaches is most promising for the development of clinically relevant, molecular subgroups that will guide treatment (as reviewed in⁸²). Whole genome expression profiling, including mutational analysis by RNA-Seq might surpass the limitations of microarray technology, but currently is still too expensive to be included in large multi-centered clinical studies. This experimental approach might also lack reproducibility due to different platforms and handling. Additionally, the large amount of data produced by high throughput sequencing techniques has to be processed and interpreted carefully. Microarrays may no longer be the standard for basic scientific research, but are currently miniaturized and automated as a cost-effective on-chip nanotechnology for clinical diagnostic approaches^{83,84}. We suggest analysing a selection of biomarker candidates by customized automated microarrays, which should be narrowed down by previous next generation sequencing approaches. These analyses should be performed parallel to clinical trials for HNSCC and other types of cancer. Correlating a small list of chosen potential biomarkers to clinical data of outcomes and treatment efficiency might facilitate following health technology assessment for drug approval as well as biomarker development. Such approaches should also be employed for investigating the pathobiological relevance of so called 'e-liquids', currently widely considered as a more 'healthy' way of smoking by the population.

In conclusion, we here shed new light on the expression differences in lymph node metastasis compared to primary HNSCC tumours, and provide the data for further analyses.

Material and Methods

Reagents. All reagents were purchased from Sigma unless stated otherwise.

Patient characteristics and material. Tissue samples were obtained from 15 patients undergoing surgical resection at the departments of otolaryngology of the Universities of Frankfurt and Mainz. Investigation has been conducted in accordance with the ethical standards according to the Declaration of Helsinki and according to national and international guidelines. The study protocol has been approved by the ethics committees "Ethik-Kommission des Fachbereichs Medizin, Universitätsklinikum der Goethe-Universität" (#83756604) and "Ethik-Kommission der Landesärztekammer Rheinland-Pfalz" (#83748515) after obtaining the patients' informed consent to participate in the study and was processed anonymously. All cases were diagnosed histopathologically as HNSCC and staged according to the TNM classification of malignant tumours recommended by the 'Union International contre le Cancer' UICC. In this study, tumour specimens, corresponding non-malignant tissue from the oral cavity (buccal mucosa), and lymph node metastasis from each patient were analysed. The median age of patients was 56 years and 86% were males. Six of the resected tumours were classified as grade 2 (40%), nine as grade 3 (60%). Twelve patients had a lymph node status of N1/2 (80%) and three of N3 (20%). The majority of cancers were located in the hypopharynx (60%), 13.3% in the larynx and 26.6% in the oropharynx (Table S1). Specimens included oropharyngeal and laryngopharyngeal carcinoma of different tumour size (T1–T4) and lymph node status (N1–3), without distant metastasis (M0). Additional inclusion criteria were positive alcohol and tobacco use, HPV-negative status of the primary tumour, a histopathological grading of 2–3 and the patients did not have previous cancer treatment before surgery. Upon resection, samples were immediately placed on ice and snap-frozen in liquid nitrogen within 30 min. Histological analyses were performed to ensure that each specimen of primary tumours and metastasis contained >70% tumour cells and <10% necrotic debris. Normal mucosa was histopathologically checked for non-malignancy. All criteria were cross-checked and approved by the independent analyses of at least two qualified head and neck pathologists. Samples not meeting these criteria were rejected.

RNA extraction. Frozen tissue samples (30–50 mg) were collected in 1 ml Trizol (Invitrogen, Karlsruhe, Germany) and dispersed using an Ultra-Turrax T25 tissue homogenizer (IKA Werke, Staufen, Germany). Total RNA was extracted according to the recommendations given by the manufacturer's Trizol protocol and further purified on RNeasy Mini spin columns (Qiagen, Hilden, Germany). Integrity and purity of total RNA were assessed on a Bioanalyzer 2100 (Agilent Technologies, Boeblingen, Germany) using RNA 6000 Nano LabChip Kit (Agilent) according to the manufacturer's instructions as previously described⁸⁵, accepting absorbance ratios of 1.8–2.2 for A_{260}/A_{280} and a A_{260}/A_{230} ratio of lower than 1.7. Samples not meeting these criteria were rejected.

Reverse transcription (RT)-PCR. RT-PCR was performed using specific primer sets for OSF2, FN1, KRT24, CLCA4, PRR4, SCCA2b, TP73L, LYPD3, FCMR, ARHGAP25 and RASGRP2 genes (Table S9). GAPDH was used to control the amounts of cDNA generated from each sample as described^{85–87}. First-strand cDNA synthesis was carried out using a cDNA synthesis kit for RT-PCR (Superscript II, Invitrogen life technologies). 1 µg of total RNA was converted to cDNA and 1 µl of the 20 µl RT product was amplified for 30 cycles (initial denaturation at 95 °C for 3 min, 30 sec at 95 °C, 30 sec at a variable temperature for annealing and 1 min at 72 °C) followed by an extension of 5 min at 72 °C. RT-PCR amplification products were analysed on 2% agarose gel stained with ethidium bromide.

Quantitative Real-Time PCR analysis. For ARHGAP25, SERPINB4 and PRR4, changes in mRNA levels detected in microarray experiments were evaluated by quantitative real-time PCR analysis, using the iCycler (BioRad, Munich, Germany). One µg of total RNA was converted to cDNA using Superscript II (Invitrogen Corporation) and oligo(dT) primer according to manufacturer's specifications. PCR reaction mixtures consisted of 12.5 µl of 2× iQ SYBER Green Supermix (Abgene, UK), 0.5 µl of each 10 µM target primer, 1 µl (1:10) cDNA template in a final scaled down reaction volume of 25 µl. Thermal cycling conditions comprised an initial denaturation step at 95 °C for 15 min, 40 cycles at 95 °C for 30 sec and variable annealing temperatures for 30 sec depending on the respective set of target primers. Cumulative fluorescence was measured at the end of the extension phase of each cycle. Specific amplicon formation with each primer pair was confirmed by dissociation curve analysis and by visualization of a single band on a 2% agarose gel. To define the relative expression of the genes, the PCR product results from each tumour sample were compared with the results from normal tissue or lymph node metastasis from the same patient. The relative expression ratio (R) of target gene is calculated using the equation: $\text{Ratio} = (E_{\text{target}})^{\Delta C_p \text{ target}(\text{control-sample})} / (E_{\text{ref}})^{\Delta C_p \text{ ref}(\text{control-sample})}$ based on its real-time PCR efficiencies (E) and the crossing point (C_p) difference of an unknown sample versus a control and expressed relative to the non-regulated housekeeping gene glyceraldehyde-3-phosphate dehydrogenase gene (GAPDH), as described^{86,88}.

Target preparation and hybridization for Affymetrix GeneChip Arrays. Total RNA (5 µg) was used to prepare biotinylated cRNAs for hybridization, following the guidelines given in the Affymetrix GeneChip Expression Analysis Technical Manual. cRNA clean-up was performed on RNeasy Mini filters (Qiagen). In all, 10 µg of fragmented, labelled cRNA were hybridized to Affymetrix HG-U133A arrays (Affymetrix, Santa Clara, CA, USA) using standard conditions (16 h, 45 °C). Arrays were washed and stained in a Fluidics Station 400 (Affymetrix) and scanned on a Gene Array Scanner 2500 (Agilent), as recommended by Affymetrix.

Statistical Analysis. Analysis of the data was performed using R 2.15.2 with the *limma* package 3.14.4. Raw fluorescence intensities from all hybridizations were normalized, applying robust multichip average (RMA) normalization for the CEL data, followed by a quantile normalization to compare expression results across specimens^{89,90}. For the comparison of primary tumours to normal mucosa (PT vs. N), metastasis to normal mucosa (M vs. N) and metastasis to primary tumours (M vs. PT) of differentially expressed genes, we also utilized the *limma* package and adjusted our model for each tissue type (supervised clustering). For multiple testing a Bonferroni correction was performed. Due to the large number of tests applied in this study, p-values have to be interpreted with caution.

For comparative analysis of our microarray analysis to NGS amplification data³⁴ and PARADIGM prediction data (including RNA-Seq)³⁴, we used genomic amplification percentages (AMP). AMP specifies the percentage of samples for which amplification of a respective gene was published by The Cancer Genome Atlas (TCGA) project³⁶. It should be noted that in our study we just compared data of HPV-negative tumours.

Ingenuity analysis. The Ingenuity Pathway Analysis (IPA) tool (Ingenuity Systems) was used to identify pathways related to enriched genes in primary tumours (PT vs. N) or metastasis (M vs. N), as well as primary tumours compared to metastasis (M vs. PT). Data was filtered to meet the criteria $p < 0,0001$ and a $\log_2 \text{FC} > |1,5|$ threshold in our comparison analysis. For each probeset, its Affymetrix probeset ID and its p-value from the comparison analysis were loaded into the analysis tool and mapped to its corresponding gene object in the Ingenuity Pathways Knowledge Base.

Data availability statement. The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1. Marur, S. & Forastiere, A. A. Head and Neck Squamous Cell Carcinoma: Update on Epidemiology, Diagnosis, and Treatment. *Mayo Clinic proceedings* **91**, 386–396, <https://doi.org/10.1016/j.mayocp.2015.12.017> (2016).
2. Torre, L. A. *et al.* Global cancer statistics, 2012. *CA: a cancer journal for clinicians* **65**, 87–108, <https://doi.org/10.3322/caac.21262> (2015).
3. Leemans, C. R., Braakhuis, B. J. & Brakenhoff, R. H. The molecular biology of head and neck cancer. *Nature reviews. Cancer* **11**, 9–22, <https://doi.org/10.1038/nrc2982> (2011).
4. Vidal, L. & Gillison, M. L. Human papillomavirus in HNSCC: recognition of a distinct disease type. *Hematol Oncol Clin North Am* **22**, 1125–1142, vii, <https://doi.org/10.1016/j.hoc.2008.08.006> (2008).
5. Pfister, D. G. *et al.* Head and neck cancers, version 2.2013. Featured updates to the NCCN guidelines. *J Natl Compr Canc Netw* **11**, 917–923 (2013).
6. van't Veer, L. J. & Bernards, R. Enabling personalized cancer medicine through analysis of gene-expression patterns. *Nature* **452**, 564–570, <https://doi.org/10.1038/nature06915> (2008).
7. Onken, M. D. *et al.* A surprising cross-species conservation in the genomic landscape of mouse and human oral cancer identifies a transcriptional signature predicting metastatic disease. *Clinical cancer research: an official journal of the American Association for Cancer Research* **20**, 2873–2884, <https://doi.org/10.1158/1078-0432.CCR-14-0205> (2014).

8. Markwell, S. M. & Weed, S. A. Tumor and stromal-based contributions to head and neck squamous cell carcinoma invasion. *Cancers* **7**, 382–406, <https://doi.org/10.3390/cancers7010382> (2015).
9. Alizadeh, A. A. *et al.* Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403**, 503–511 (2000).
10. Golub, T. R. *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**, 531–537 (1999).
11. Bittner, M. *et al.* Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* **406**, 536–540 (2000).
12. Huang, E. *et al.* Gene expression predictors of breast cancer outcomes. *Lancet* **361**, 1590–1596 (2003).
13. MacDonald, T. J. *et al.* Expression profiling of medulloblastoma: PDGFRA and the RAS/MAPK pathway as therapeutic targets for metastatic disease. *Nature genetics* **29**, 143–152 (2001).
14. Chung, C. H. *et al.* Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. *Cancer cell* **5**, 489–500 (2004).
15. Ginos, M. A. *et al.* Identification of a gene expression signature associated with recurrent disease in squamous cell carcinoma of the head and neck. *Cancer research* **64**, 55–63 (2004).
16. Al Moustafa, A. E. *et al.* Identification of genes associated with head and neck carcinogenesis by cDNA microarray comparison between matched primary normal epithelial and squamous carcinoma cells. *Oncogene* **21**, 2634–2640 (2002).
17. Alevizos, I. *et al.* Oral cancer *in vivo* gene expression profiling assisted by laser capture microdissection and microarray analysis. *Oncogene* **20**, 6196–6204 (2001).
18. Cromer, A. *et al.* Identification of genes associated with tumorigenesis and metastatic potential of hypopharyngeal cancer by microarray analysis. *Oncogene* **23**, 2484–2498 (2004).
19. El-Naggar, A. K. *et al.* Differential expression profiling of head and neck squamous carcinoma: significance in their phenotypic and biological classification. *Oncogene* **21**, 8206–8219 (2002).
20. Leethanakul, C. *et al.* Distinct pattern of expression of differentiation and growth-related genes in squamous cell carcinomas of the head and neck revealed by the use of laser capture microdissection and cDNA arrays. *Oncogene* **19**, 3220–3224 (2000).
21. Mendez, E. *et al.* Transcriptional expression profiles of oral squamous cell carcinomas. *Cancer* **95**, 1482–1494 (2002).
22. Knauer, S. K. *et al.* An otoprotective role for the apoptosis inhibitor protein survivin. *Cell Death Dis* **1**, e51 (2010).
23. Fetzi, V. *et al.* Inducible NO synthase confers chemoresistance in head and neck cancer by modulating survivin. *International journal of cancer* **124**, 2033–2041, <https://doi.org/10.1002/ijc.24182> (2009).
24. Thurlow, J. K. *et al.* Spectral clustering of microarray data elucidates the roles of microenvironment remodeling and immune responses in survival of head and neck squamous cell carcinoma. *J Clin Oncol* **28**, 2881–2888, <https://doi.org/10.1200/JCO.2009.24.8724> (2010).
25. Fakhry, C. *et al.* Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. *Journal of the National Cancer Institute* **100**, 261–269, <https://doi.org/10.1093/jnci/djn011> (2008).
26. Suh, Y., Amelio, I., Guerrero Urbano, T. & Tavassoli, M. Clinical update on cancer: molecular oncology of head and neck cancer. *Cell death & disease* **5**, e1018, <https://doi.org/10.1038/cddis.2013.548> (2014).
27. Kramer, A., Green, J., Pollard, J. Jr & Tugendreich, S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* **30**, 523–530 (2013).
28. van't Veer, L. J. *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530–536 (2002).
29. Faber, A. *et al.* CD44 as a stem cell marker in head and neck squamous cell carcinoma. *Oncology reports* **26**, 321–326, <https://doi.org/10.3892/or.2011.1322> (2011).
30. Simpson, D. R., Mell, L. K. & Cohen, E. E. Targeting the PI3K/AKT/mTOR pathway in squamous cell carcinoma of the head and neck. *Oral oncology* **51**, 291–298, <https://doi.org/10.1016/j.oraloncology.2014.11.012> (2015).
31. Prasad, G. & McCullough, M. Chemokines and cytokines as salivary biomarkers for the early diagnosis of oral cancer. *Int J Dent* **2013**, 813756, <https://doi.org/10.1155/2013/813756> (2013).
32. Lawrence, T. The nuclear factor NF-kappaB pathway in inflammation. *Cold Spring Harb Perspect Biol* **1**, a001651, <https://doi.org/10.1101/cshperspect.a001651> (2009).
33. Karin, M. NF-kappaB as a critical link between inflammation and cancer. *Cold Spring Harb Perspect Biol* **1**, a000141, <https://doi.org/10.1101/cshperspect.a000141> (2009).
34. Cancer Genome Atlas Research, N. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519–525, <https://doi.org/10.1038/nature11404> (2012).
35. Center, B. I. T. G. D. A. In *Broad Institute of MIT and Harvard* (2016).
36. Cancer Genome Atlas, N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* **517**, 576–582, <https://doi.org/10.1038/nature14129> (2015).
37. Kuriakose, M. A. *et al.* Selection and validation of differentially expressed genes in head and neck cancer. *Cellular and molecular life sciences: CMLS* **61**, 1372–1383 (2004).
38. Schlingemann, J. *et al.* Patient-based cross-platform comparison of oligonucleotide microarray expression profiles. *Laboratory investigation; a journal of technical methods and pathology* **85**, 1024–1039 (2005).
39. Paz, H., Pathak, N. & Yang, J. Invading one step at a time: the role of invadopodia in tumor metastasis. *Oncogene* **33**, 4193–4202, <https://doi.org/10.1038/ncr.2013.393> (2014).
40. Rosenthal, E. L. & Matrisian, L. M. Matrix metalloproteases in head and neck cancer. *Head & neck* **28**, 639–648, <https://doi.org/10.1002/hed.20365> (2006).
41. Wunsch, D. *et al.* Taspase1: a 'Misunderstood' protease with translational cancer relevance. *Oncogene* **35**, 3351–3364, <https://doi.org/10.1038/ncr.2015.436> (2016).
42. Zhou, Y., Gunput, R. A., Adolfs, Y. & Pasterkamp, R. J. MICALs in control of the cytoskeleton, exocytosis, and cell death. *Cellular and molecular life sciences: CMLS* **68**, 4033–4044, <https://doi.org/10.1007/s00018-011-0787-2> (2011).
43. Fang, X., Netzer, M., Baumgartner, C., Bai, C. & Wang, X. Genetic network and gene set enrichment analysis to identify biomarkers related to cigarette smoking and lung cancer. *Cancer treatment reviews* **39**, 77–88, <https://doi.org/10.1016/j.ctrv.2012.06.001> (2013).
44. Ashida, S. *et al.* Expression of novel molecules, MICAL2-PV (MICAL2 prostate cancer variants), increases with high Gleason score and prostate cancer progression. *Clinical cancer research: an official journal of the American Association for Cancer Research* **12**, 2767–2773, <https://doi.org/10.1158/1078-0432.CCR-05-1995> (2006).
45. Mariotti, S. *et al.* MICAL2 is a novel human cancer gene controlling mesenchymal to epithelial transition involved in cancer growth and invasion. *Oncotarget* **7**, 1808–1825, <https://doi.org/10.18632/oncotarget.6577> (2016).
46. Gaggioli, C. *et al.* Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells. *Nature cell biology* **9**, 1392–1400, <https://doi.org/10.1038/ncb1658> (2007).
47. Goetz, J. G. *et al.* Biomechanical remodeling of the microenvironment by stromal caveolin-1 favors tumor invasion and metastasis. *Cell* **146**, 148–163, <https://doi.org/10.1016/j.cell.2011.05.040> (2011).
48. Uchida, D. *et al.* Acquisition of lymph node, but not distant metastatic potentials, by the overexpression of CXCR4 in human oral squamous cell carcinoma. *Laboratory investigation; a journal of technical methods and pathology* **84**, 1538–1546, <https://doi.org/10.1038/labinvest.3700190> (2004).

49. Lim, K. P. *et al.* Fibroblast gene expression profile reflects the stage of tumour progression in oral squamous cell carcinoma. *The Journal of pathology* **223**, 459–469, <https://doi.org/10.1002/path.2841> (2011).
50. Marsh, D. *et al.* Stromal features are predictive of disease mortality in oral cancer patients. *The Journal of pathology* **223**, 470–481, <https://doi.org/10.1002/path.2830> (2011).
51. Ishikawa, T. *et al.* Hypoxia enhances CXCR4 expression by activating HIF-1 in oral squamous cell carcinoma. *Oncology reports* **21**, 707–712 (2009).
52. Koontongkaew, S., Amornphimoltham, P., Monthanpisut, P., Saensuk, T. & Leelakriangsak, M. Fibroblasts and extracellular matrix differently modulate MMP activation by primary and metastatic head and neck cancer cells. *Medical oncology* **29**, 690–703, <https://doi.org/10.1007/s12032-011-9871-6> (2012).
53. Uchida, D. *et al.* Possible role of stromal-cell-derived factor-1/CXCR4 signaling on lymph node metastasis of oral squamous cell carcinoma. *Experimental cell research* **290**, 289–302 (2003).
54. Price, S. J., Greaves, D. R. & Watkins, H. Identification of novel, functional genetic variants in the human matrix metalloproteinase-2 gene: role of Sp1 in allele-specific transcriptional regulation. *The Journal of biological chemistry* **276**, 7549–7558, <https://doi.org/10.1074/jbc.M010242200> (2001).
55. Isa, A. Y., Ward, T. H., West, C. M., Slevin, N. J. & Homer, J. J. Hypoxia in head and neck cancer. *The British journal of radiology* **79**, 791–798, <https://doi.org/10.1259/bjr/17904358> (2006).
56. Janssen, H. L., Haustermans, K. M., Balm, A. J. & Begg, A. C. Hypoxia in head and neck cancer: how much, how important? *Head & neck* **27**, 622–638, <https://doi.org/10.1002/hed.20223> (2005).
57. Jiang, J., Tang, Y. L. & Liang, X. H. EMT: a new vision of hypoxia promoting cancer progression. *Cancer biology & therapy* **11**, 714–723 (2011).
58. Semenza, G. L. Oxygen sensing, hypoxia-inducible factors, and disease pathophysiology. *Annual review of pathology* **9**, 47–71, <https://doi.org/10.1146/annurev-pathol-012513-104720> (2014).
59. Sullivan, R. & Graham, C. H. Hypoxia-driven selection of the metastatic phenotype. *Cancer metastasis reviews* **26**, 319–331, <https://doi.org/10.1007/s10555-007-9062-2> (2007).
60. Knopf, A., Fritsche, M. K. & Li, Y. CXCR4: A New Player in an Old Scene? *ORL J Otorhinolaryngol Relat Spec* **79**, 34–42, <https://doi.org/10.1159/000455706> (2017).
61. Garrido-Laguna, I. & Hidalgo, M. Pancreatic cancer: from state-of-the-art treatments to promising novel therapies. *Nature reviews. Clinical oncology* **12**, 319–334, <https://doi.org/10.1038/nrclinonc.2015.53> (2015).
62. Wu, X. *et al.* HOXB7, a homeodomain protein, is overexpressed in breast cancer and confers epithelial-mesenchymal transition. *Cancer research* **66**, 9527–9534, <https://doi.org/10.1158/0008-5472.CAN-05-4470> (2006).
63. Liu, S. *et al.* HOXB7 promotes malignant progression by activating the TGFbeta signaling pathway. *Cancer research* **75**, 709–719, <https://doi.org/10.1158/0008-5472.CAN-14-3100> (2015).
64. Chen, H. *et al.* Hoxb7 inhibits transgenic HER-2/neu-induced mouse mammary tumor onset but promotes progression and lung metastasis. *Cancer research* **68**, 3637–3644, <https://doi.org/10.1158/0008-5472.CAN-07-2926> (2008).
65. Jin, K. *et al.* The HOXB7 protein renders breast cancer cells resistant to tamoxifen through activation of the EGFR pathway. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 2736–2741, <https://doi.org/10.1073/pnas.1018859108> (2012).
66. De Souza Setubal Destro, M. F. *et al.* Overexpression of HOXB7 homeobox gene in oral cancer induces cellular proliferation and is associated with poor prognosis. *International journal of oncology* **36**, 141–149 (2010).
67. Bitu, C. C. *et al.* HOXB7 expression is a prognostic factor for oral squamous cell carcinoma. *Histopathology* **60**, 662–665, <https://doi.org/10.1111/j.1365-2559.2011.04102.x> (2012).
68. Zheng, H., Fu, G., Dai, T. & Huang, H. Migration of endothelial progenitor cells mediated by stromal cell-derived factor-1alpha/CXCR4 via PI3K/Akt/eNOS signal transduction pathway. *Journal of cardiovascular pharmacology* **50**, 274–280, <https://doi.org/10.1097/FJC.0b013e318093ec8f> (2007).
69. Onoue, T. *et al.* Epithelial-mesenchymal transition induced by the stromal cell-derived factor-1/CXCR4 system in oral squamous cell carcinoma cells. *International journal of oncology* **29**, 1133–1138 (2006).
70. Lui, V. W. *et al.* Frequent mutation of the PI3K pathway in head and neck cancer defines predictive biomarkers. *Cancer discovery* **3**, 761–769, <https://doi.org/10.1158/2159-8290.CD-13-0103> (2013).
71. Templeton, A. J. *et al.* Prognostic relevance of receptor tyrosine kinase expression in breast cancer: a meta-analysis. *Cancer treatment reviews* **40**, 1048–1055 (2014).
72. Duncan, J. S. *et al.* Dynamic reprogramming of the kinome in response to targeted MEK inhibition in triple-negative breast cancer. *Cell* **149**, 307–321 (2012).
73. Dieci, M. V., Arnedos, M., Andre, F. & Soria, J. C. Fibroblast growth factor receptor inhibitors as a cancer treatment: from a biologic rationale to medical perspectives. *Cancer discovery* **3**, 264–279 (2013).
74. Dienstmann, R. *et al.* Genomic aberrations in the FGFR pathway: opportunities for targeted therapies in solid tumors. *Ann Oncol* **25**, 552–563 (2014).
75. Furlan, A. *et al.* Abl interconnects oncogenic Met and p53 core pathways in cancer cells. *Cell death and differentiation* **18**, 1608–1616 (2011).
76. Brandl, A. *et al.* Dynamically regulated sumoylation of HDAC2 controls p53 deacetylation and restricts apoptosis following genotoxic stress. *J Mol Cell Biol* (2012).
77. Trivedi, S. *et al.* Immune biomarkers of anti-EGFR monoclonal antibody therapy. *Ann Oncol* **26**, 40–47 (2015).
78. Ang, K. K. *et al.* Randomized phase III trial of concurrent accelerated radiation plus cisplatin with or without cetuximab for stage III to IV head and neck carcinoma: RTOG 0522. *J Clin Oncol* **32**, 2940–2950 (2014).
79. Harrington, K. *et al.* Postoperative Adjuvant Lapatinib and Concurrent Chemoradiotherapy Followed by Maintenance Lapatinib Monotherapy in High-Risk Patients With Resected Squamous Cell Carcinoma of the Head and Neck: A Phase III, Randomized, Double-Blind, Placebo-Controlled Study. *J Clin Oncol* **33**, 4202–4209 (2015).
80. Chong, C. R. & Janne, P. A. The quest to overcome resistance to EGFR-targeted therapies in cancer. *Nature medicine* **19**, 1389–1400, <https://doi.org/10.1038/nm.3388> (2013).
81. Chapman, C. H., Saba, N. F. & Yom, S. S. Targeting epidermal growth factor receptor for head and neck squamous cell carcinoma: still lost in translation? *Annals of translational medicine* **4**, 80 (2016).
82. Sethi, N., MacLennan, K., Wood, H. M. & Rabbitts, P. Past and future impact of next-generation sequencing in head and neck cancer. *Head & neck* **38**(Suppl 1), E2395–2402 (2015).
83. Gulley, M. L. & Morgan, D. R. Molecular oncology testing in resource-limited settings. *J Mol Diagn* **16**, 601–611, <https://doi.org/10.1016/j.jmoldx.2014.07.002> (2014).
84. Docter, D. *et al.* No king without a crown—impact of the nanomaterial-protein corona on nanobiomedicine. *Nanomedicine (Lond)* **10**, 503–519 (2015).
85. Knauer, S. K. *et al.* Functional characterization of novel mutations affecting survivin (BIRC5)-mediated therapy resistance in head and neck cancer patients. *Hum Mutat* **34**, 395–404, <https://doi.org/10.1002/humu.22249> (2013).
86. Knauer, S. K. *et al.* The survivin isoform survivin-3B is cytoprotective and can function as a chromosomal passenger complex protein. *Cell cycle Georgetown, Tex* **6**, 1502–1509 (2007).
87. Rosorius, O. *et al.* Direct observation of nucleo-cytoplasmic transport by microinjection of GFP-tagged proteins in living cells. *BioTechniques* **27**, 350–355 (1999).

88. Engels, K. *et al.* NO signaling confers cytoprotectivity through the survivin network in ovarian carcinomas. *Cancer Res* **68**, 5159–5166 (2008).
89. Bolstad, B. M., Irizarry, R. A., Astrand, M. & Speed, T. P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185–193 (2003).
90. Irizarry, R. A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249–264, <https://doi.org/10.1093/biostatistics/4.2.249> (2003).
91. Lawrence, M. S. *et al.* Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505**, 495–501, <https://doi.org/10.1038/nature12912> (2014).

Acknowledgements

We thank Sandra Schneider for excellent technical assistance. Grant support: BMBF-*NanoBEL*, *Impulsfonds RLP*, *BioMATICS*, *SFB1093*, and *NanoTransMed*, which is cofounded by the European Regional Development Fund (ERDF) in the framework of the INTERREG V Upper Rhine program, the Swiss Confederation and the Swiss cantons of Aargau, Basel-Landschaft and Basel-Stadt. This study was supported by grants from the Friedrich Naumann Foundation for Freedom, Else-Kröner Foundation, Zeiss-ChemBioMed, Stiftung Tumorforschung Kopf-Hals, Deutsche Forschungsgemeinschaft and the Support Program of the University Hospital Mainz.

Author Contributions

D.G., N.H., A.G.-A. and J.M. conceived, designed and conducted experiments and analysed the data; D.W., S.K., J.K., C.M. and S.S. co-designed the study and revised the manuscript; D.G., D.W. and R.S. conceived and designed the study and wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-25512-7>.

Competing Interests: The authors declare no competing interests.

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/>.

© The Author(s) 2018

3.2 Proteasen und Steuerungsmechanismen in Karzinogenese und Entwicklung

Proteasen sind maßgeblich für die Remodellierung der extrazellulären Matrix verantwortlich, die eine der wichtigsten Voraussetzung für das invasive Potential eines Tumors darstellt. Matrix-Metalloproteinasen waren ein großer Bestandteil der im vorangegangenen Teil der Arbeit identifizierten Gensignatur. Jedoch war auch dort zu erkennen, dass nicht nur extrazelluläre Enzyme eine große Rolle in der Tumor- und Karzinogenese spielen, sondern auch intrazelluläre Proteasen, die durch posttranslationale Modifikation bestimmte Signalwege steuern und so in vielfältige biologische Prozesse wie Genexpression und Differenzierung, aber auch Abbauprozesse, Autophagie und Zelltod eingreifen.

3.2.1 „Protein Translocation Assays to probe protease function and screen for inhibitors“

Angelina Hahlbrock, Dorothee Goesswein, Roland H. Stauber.

*Springer Protocols, 2017.*⁴

Zusammenfassung

Proteasen sind nicht nur für den unspezifischen Abbau von Proteinen verantwortlich, sondern funktionieren auch als hochspezifische molekulare Scheren, die durch proteolytische Prozessierung neue Proteinprodukte herstellen können und so irreversible posttranslationale Modifizierungen vermitteln⁸⁹. Um ihre Rolle in biologischen Prozessen zu untersuchen, bedarf es komplexer und

⁴ Eine Kopie des Artikels kann auf Grund der aktuellen Lizenzierung nicht abgedruckt werden.

raffinierter molekularer Werkzeuge, die in der Lage sind, die Aktivität einer Protease gezielt zu untersuchen. Dabei sind die wertvollsten Versuchsanordnungen solche, die die natürliche Umgebung der Protease mit einbeziehen und so einen realistischen Eindruck ihres Aktionspotentials geben. Fluoreszierende Protein-Biosensoren stellen hierbei vielseitige wie robuste Werkzeuge dar, um die Komplexität pathobiologischer Prozesse abzubilden. Zudem können sie mit automatisierten Screening-Plattformen kombiniert werden, um unterschiedlichste Substanzen nach ihrem modulatorischen Potential hinsichtlich proteolytischer Aktivität aber auch bezüglich des Interaktionsspektrums mit anderen Proteinen zu untersuchen. Das Grundgerüst des hier verwendeten Biosensors besteht dabei aus einem nukleären Lokalisationssignal (NLS) und einem dominanten nukleären Exportsignal (NES) an den jeweiligen Protein-Termini, GFP als fluoreszierendem Marker, einer Schnittstelle (PCS) oder Interaktionsdomäne (BP) je nach Untersuchung, einem Myc-Marker und Gluthathion-S-Transferase (GST), um unspezifischen Transport durch passive Diffusion zu unterbinden: NLS-GFP-GST-(PCS/BP)-Myc-NES. Da das NES stärker ist als das NLS, wird das fluoreszierende Signal vorwiegend im Zytoplasma detektiert, solange der Sensor unmodifiziert bleibt. Wird der Biosensor nun an der PCS von einer Protease geschnitten, lokalisiert der N-Terminus NLS-GFP-GST-(PCS1) durch das NLS im Kern, während der kleinere C-terminale Teil (PCS2)-Myc-NES im Zytoplasma verbleibt. Im Fall einer Protein-Protein-Interaktionsstudie wird durch Bindung an ein nukleär/nukleolär lokalisiertes Partnerproteins der gesamte Komplex im Zellkern detektiert.

3.2.2 „Overexpression of the catalytically impaired Taspase1 T234V or Taspase1 D233A variants does not have a dominant negative effect in T(4;11) leukemia cells“

Carolin Bier, Rouven Hecht, Lena Kunst, Sabine Scheiding, Désirée Wunsch, **Dorothee Goesswein**, Günter Schneider, Oliver H. Krämer, Shirley K. Knauer, Roland H. Stauber.

Plos One, 2012.⁵

Zusammenfassung

Chromosomale Translokationen, die das MLL-Gen mit anderen genetischen Partnern zu einem Onkogen verbinden, sind häufig in Akuter Myeloider Leukämie (AML) und Akuter Lymphoblastischer Leukämie (ALL) vorhanden¹⁹². Die häufigste ist dabei die t(4;11)(q21;q23)-Translokation, welche MLL mit AF4 (AFF1) verbindet und zur Expression der Fusionsproteine AF4·MLL und MLL·AF4 führt¹⁹³. Die Prognose für diese Form der ALL ist noch immer sehr schlecht, und die Krankheit gilt als nahezu unheilbar^{194,195}. Selbst mit neuesten CAR T-Zell Therapien erleiden Patienten mit MLL-vermittelten ALL häufig auf Grund von Resistenzmechanismen einen Rückfall in Form einer Akuten Myeloiden Leukämie¹⁹⁶. Die pathobiologische Relevanz des AF4·MLL Fusionsproteins wird nach wie vor kontrovers diskutiert. Es konnte jedoch gezeigt werden, dass AF4·MLL in Mäusen ALL ohne den Einfluss von MLL·AF4 auslösen kann⁵⁹. Unter normalen physiologischen Bedingungen wird MLL posttranslational von der Protease Threonin-Aspartase 1 (Taspase1) in zwei größere Fragmente MLL.N und MLL.C geschnitten, die intramolekular miteinander interagieren^{45,119,195}. Das Fusionsprotein AF4·MLL enthält die Spaltstelle für Taspase1 und wird, sofern es nicht durch Taspase1 prozessiert wird, durch das Proteasom abgebaut⁵¹. Die Prozessierung des AF4·MLL Genproduktes könnte also durch die Inhibition der Taspase1 verhindert werden und so einen wertvollen Beitrag zur Behandlung t(4;11)(q21;q23)-vermittelter Leukämien leisten.

⁵ Eine Kopie des Artikels wird im Folgenden abgedruckt unter der *Creative Commons Attribution 4.0 International License*. <https://creativecommons.org/licenses/by/4.0/>

Das Taspase1-Proenzym wird zur Aktivierung autokatalytisch in eine α - und eine β - Untereinheit prozessiert und soll, wie in früheren Publikationen postuliert wurde, zu einem aktiven $\alpha\beta\alpha$ -Heterodimer zusammengelagert werden^{45,121}. Da es keine chemischen Inhibitoren für Taspase1 gibt, wurde die Heterodimerisierung als potentiell Therapieziel zur Behandlung von AF4·MLL-vermittelten Leukämien untersucht.

Die katalytisch inaktive Taspase1^{T234V} und die proteolytisch stark eingeschränkte Taspase1^{D233A} Variante wurden in leukämischen Zellen überexprimiert und ein speziell entwickelter Translokations-Biosensor-Assay wurde zur Überprüfung der Aktivität *in vivo* angewendet. Selbst eine neunfache Überexpression der jeweiligen Mutante konnte weder die *trans*- noch die *cis*-Aktivität der Taspase1 einschränken und eine Prozessierung des AF4·MLL Onkoproteins verhindern. Auch eine Überexpression der einzelnen Taspase1-Untereinheiten vermochte die Aktivität der Protease nicht zu beeinflussen, und in weiteren Versuchen konnte in der Tat keine Zusammenlagerung zu einem aktiven Enzymkomplex nachgewiesen werden. Die These der Zusammenlagerung von je vier Untereinheiten zu einem aktiven $\alpha\beta\alpha$ -Heterodimers muss also als widerlegt gelten, da gezeigt werden konnte, dass Taspase1 ausschließlich als $\alpha\beta$ -Homodimer aktiv ist und auf die Aktivierung durch die autokatalytische Spaltung *in vivo* angewiesen ist. Für AF4·MLL-vermittelte Leukämien scheint es sinnvoller zu sein, weiterhin auf eine chemische Inhibition der Taspase1 zu setzen als eine Dimerisierung zu verhindern.

Overexpression of the Catalytically Impaired Taspase1^{T234V} or Taspase1^{D233A} Variants Does Not Have a Dominant Negative Effect in T(4;11) Leukemia Cells

Carolin Bier^{1*}, Rouven Hecht², Lena Kunst², Sabine Scheiding³, Désirée Wünsch¹, Dorothee Goesswein¹, Günter Schneider⁴, Oliver H. Krämer³, Shirley K. Knauer^{2*}, Roland H. Stauber^{1*}

1 Molecular and Cellular Oncology, Mainz Screening Center (MSC), University Hospital of Mainz, Mainz, Germany, **2** Institute for Molecular Biology, Center for Medical Biotechnology (ZMB), University Duisburg-Essen, Essen, Germany, **3** Institute for Biochemistry and Biophysics/Centre for Molecular Biomedicine (CMB), Friedrich-Schiller-University Jena, Jena, Germany, **4** Department of Internal Medicine II, Technical University of Munich, Munich, Germany

Abstract

Background: The chromosomal translocation t(4;11)(q21;q23) is associated with high-risk acute lymphoblastic leukemia of infants. The resulting AF4•MLL oncoprotein becomes activated by Taspase1 hydrolysis and is considered to promote oncogenic transcriptional activation. Hence, Taspase1's proteolytic activity is a critical step in AF4•MLL pathophysiology. The Taspase1 proenzyme is autoproteolytically processed in its subunits and is assumed to assemble into an $\alpha\beta\alpha$ -heterodimer, the active protease. Therefore, we investigated here whether overexpression of catalytically inactive Taspase1 variants are able to interfere with the proteolytic activity of the wild type enzyme in AF4•MLL model systems.

Methodology/Findings: The consequences of overexpressing the catalytically dead Taspase1 mutant, Taspase1^{T234V}, or the highly attenuated variant, Taspase1^{D233A}, on Taspase1's processing of AF4•MLL and of other Taspase1 targets was analyzed in living cancer cells employing an optimized cell-based assay. Notably, even a nine-fold overexpression of the respective Taspase1 mutants neither inhibited Taspase1's *cis*- nor *trans*-cleavage activity *in vivo*. Likewise, enforced expression of the α - or β -subunits showed no *trans*-dominant effect against the ectopically or endogenously expressed enzyme. Notably, co-expression of the individual α - and β -subunits did not result in their assembly into an enzymatically active protease complex. Probing Taspase1 multimerization in living cells by a translocation-based protein interaction assay as well as by biochemical methods indicated that the inactive Taspase1 failed to assemble into stable heterocomplexes with the wild type enzyme.

Conclusions: Collectively, our results demonstrate that inefficient heterodimerization appears to be the mechanism by which inactive Taspase1 variants fail to inhibit wild type Taspase1's activity in *trans*. Our work favours strategies targeting Taspase1's catalytic activity rather than attempts to block the formation of active Taspase1 dimers to interfere with the pathobiological function of AF4•MLL.

Citation: Bier C, Hecht R, Kunst L, Scheiding S, Wünsch D, et al. (2012) Overexpression of the Catalytically Impaired Taspase1^{T234V} or Taspase1^{D233A} Variants Does Not Have a Dominant Negative Effect in T(4;11) Leukemia Cells. PLoS ONE 7(5): e34142. doi:10.1371/journal.pone.0034142

Editor: Matthew Bogoy, Stanford University, United States of America

Received: October 14, 2011; **Accepted:** February 22, 2012; **Published:** May 3, 2012

Copyright: © 2012 Bier et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Grant support: German Cancer Aid (FKZ102362), Head and Neck Cancer Foundation (CB), donation by R. Patzke, Rheinland-Pfalz für Innovationen, Mainz Screening Center, DFG KN973/1-1 and DFG INST371/5-1FUGG, and the Inneruniversity Funds and MAIFOR Mainz Support Program (RS and SK). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

* E-mail: bier@uni-mainz.de (CB); shirley.knauer@uni-due.de (SK); rstauber@uni-mainz.de (RS)

Introduction

Chromosomal rearrangements of the mixed lineage leukemia (*MLL*) gene with numerous partner genes are frequently found in acute myeloid (AML) and acute lymphoblastic leukemia (ALL) [1,2]. Fused on chromosome4 (AF4) is the most common translocation partner in *MLL*-mediated leukemia, resulting in the expression of the AF4•*MLL* and *MLL*•AF4 fusion proteins. Although the pathomechanism of t(4;11)-mediated leukemia is still discussed controversially, expression of the AF4•*MLL* fusion enhanced the hematopoietic repopulating potential of CD34⁺ cells, and led to the development of predominantly proB ALL in a mouse model [2,3]. Similar to the prototypic *MLL* protein, the AF4•*MLL* fusion protein contains cleavage-sites for Threonine

Aspartase 1 (Taspase1, Tasp) and, is considered a bona fide substrate for this protease [4,5,6,7,8].

In leukemic cells, the unprocessed AF4•*MLL* fusion protein is a substrate for the E3-ubiquitin ligases SIAH1 and SIAH2 [9,10]. SIAH1 and SIAH2, belonging to the family of the seven in absentia homolog (SIAH), are E3 ligases involved in ubiquitination and proteasome-mediated degradation of specific proteins [9,10]. Therefore the presence of these ligases leads to a low steady-state level of the AF4•*MLL* fusion, hampering its detection and experimental analysis. However, following its proteolytic processing by Taspase1, the AF4•*MLL* cleavage products AF4•*MLL*.N and *MLL*.C heterodimerize through their FYRN and FYRC interaction domains, forming a high molecular-weight protein

complex resistant to SIAH-mediated degradation [9,10]. Thus, chemico-genetic interference with Taspase1's activity is expected to promote AF4•MLL degradation, thereby precluding the activation of oncogenic programs and disease development. Moreover, a total of at least 64 MLL chromosomal-fusion partners have been characterized so far at the molecular level, associated with disease (**Table S1**) [1,11,12]. Hence, Taspase1's proteolytic processing of MLL fusions in general may contribute to various pathologies.

The human Taspase1 gene encodes a protein of 420 amino acids (aa) and is able to cleave other substrates *in trans* by recognizing a conserved peptide motif ($Q^3[F,I,L,V]^2D^1 \downarrow G^1x^2D^3D^4$) with an aspartate at the P1 position [6,7]. The discovery of Taspase1 founded a new class of endopeptidases that utilize the N-terminal threonine of its mature β -subunit as the active site, which is generated by autoproteolysis of the proenzyme (*cis*-cleavage) [6]. Mutation of the catalytic nucleophile, Thr²³⁴, completely abolishes Taspase1's catalytic activity [6,13]. Based on data mainly derived from analyzing bacterially expressed Taspase1, it is assumed that the proenzyme assembles into an asymmetric $\alpha\beta\beta\alpha$ -heterodimer following autoproteolysis, representing the active protease [6,13].

Protein-protein interactions (PPIs) in general are key players for multiple (patho)biological cellular processes [14]. Thus, interfering with disease-relevant interactomes *via* enforced expression of dominant-negative mutants and/or small molecules has emerged as a promising, though challenging strategy for human therapeutics [15,16]. Blocking the p53-mdm2 interaction with synthetic molecules had been shown to induce p53 activation and thereof tumor cell death [17]. Likewise, the peptide-mediated disruption of the AF4-AF9 protein complex, or targeting the oligomerization domain of RUNX1/ETO interfered with the activity of the fusion proteins in leukemic cells [18,19]. Recently, we also showed that it is in principle possible to specifically inhibit and to destroy the AF4•MLL oncoprotein by genetic PPIs inhibitors [9].

Consequently, as currently no effective synthetic Taspase1 inhibitors are available, we here investigated whether selectively interfering with the formation of the proposed active Taspase1 $\alpha\beta\beta\alpha$ -heterodimer would block processing of the AF4•MLL fusion and thus, may prevent leukemogenesis [20,21]. Studies investigating the biological consequences of overexpressing inactive Taspase1 variants have not been performed before. Hence we here developed and employed novel cell-based assays allowing to functionally monitor the effects of overexpressing catalytically inactive or attenuated Taspase1 mutants on Taspase1's processing of AF4•MLL in living cancer cells in real time. To our knowledge, this is the first comprehensive study addressing Taspase1 multimerization and genetic interference thereof *in vivo*.

Materials and Methods

Antibodies (Ab), Reagents and Compounds

Ab used: anti-(α)-GapDH (Glyceraldehyde-3-phosphate dehydrogenase, sc-47724; Santa Cruz Biotechnology, Heidelberg, Germany); α -GFP (green fluorescent protein, sc-8334; Santa Cruz Biotechnology, Heidelberg, Germany); α -GST (glutathione S-transferase, sc-57753; Santa Cruz Biotechnology, Heidelberg, Germany); α -NPM1 (nucleophosmin, #3542 Cell Signaling); α -Tasp_C (directed against the C-terminus of Taspase1, AP1330b BioCat GmbH, Heidelberg, Germany); α -Tasp_N (directed against the N-terminus of Taspase1, sc-85945; Santa Cruz Biotechnology, Heidelberg, Germany). Appropriate HRP-conjugated secondary antibodies (Santa Cruz Biotechnology, Heidelberg, Germany) were used.

Plasmids

AF4•MLL, transcription factor IIA (TFIIA) and upstream stimulatory factor2 (USF2) indicator protein expression plasmids were derived from pNLS-GFP/GST-CS3-RevNES (pCasp3-Clev), encoding a fusion composed of the SV40 large T-antigen nuclear localization signal (NLS), GST, GFP, the Caspase3 cleavage site (CS3), and a strong nuclear export signal (NES) of the human immunodeficiency virus type 1 (HIV-1) Rev protein. [7,22] In p_{NLS-GFP/GST-AF4•MLL_S1/2-NES}_{Rev} (pA•M_S1/2), p_{NLS-GFP/GST-AF4•MLL_S1-NES}_{Rev} (pA•M_S1) and p_{NLS-GFP/GST-AF4•MLL_S2-NES}_{Rev} (pA•M_S2), CS3 was replaced by the Taspase1 cleavage site from AF4•MLL (S1/2 – both cleavage sites aa 1582–1710; S1 – first cleavage site: ¹⁶⁰⁰AEGQVDGADD¹⁶⁰⁹; S2 – second cleavage site: ¹⁶⁵²KISQLDGVDD¹⁶⁶¹), as well as by a GSGS linker following the cleavage site. To determine the linker sequences allowing optimal processing by Taspase1, plasmids containing the sequences ¹⁶⁵⁵QLDGVDD¹⁶⁶¹, GSGS¹⁶⁵⁵QLDGVDD¹⁶⁶¹ or ¹⁶⁵²KISQLDGVDD¹⁶⁶¹G were established likewise. pA•M_S1_{mut} and pA•M_S2_{mut} encode fusions containing mutated Taspase1 cleavage sites, in which P1 and P1' were exchanged by alanin (S1_{mut}: ¹⁶⁰⁰AEGQVAAADD¹⁶⁰⁹ and S2_{mut}: ¹⁶⁵²KISQLAAVDD¹⁶⁶¹), precluding their processing by Taspase1. pTFIIA_S, pUSF2_S, pTFIIA-GFP and pUSF2-GFP were described. [7] Besides the plasmids expressing green fluorescent protein fusions, we also constructed versions in which GFP was replaced by the red-fluorescent protein mCherry or RFP allowing performing dual- or triple-color *in vivo* assays.

Expression constructs encoding untagged Taspase1, Taspase1 fusion with autofluorescent proteins, including the red-fluorescent protein mCherry, a cytoplasmatic version of GFP-tagged Taspase1 (pTasp_{Cyt}), and NPM1 as untagged or fusions with autofluorescent proteins were described. [7,23] Expression constructs encoding Taspase1 as a fusion with the small HA-tag was generated as described. [9,24] The Taspase1 expression plasmid was used as template to amplify the Taspase1 α - and β -subunits. Cloning of the subunits into expression vectors pc3-GFP and pc3-BFP using BamHI/NheI-restriction sites, respectively, allowed the expression of Taspase1 subunits as fusions with green and blue fluorescent proteins as described. [25] A cytoplasmatic version of a GFP-tagged Taspase1 β -subunit is encoded by pTasp- β _{Cyt}, which was described. [20] Plasmids pTasp^{T234V}- and pTasp^{D233A}-GFP/-mCherry/-BFP or their untagged versions were generated by splice overlap extension PCR as reported. [7] pRevM10BL-RFP was generated by replacing BFP by RFP in RevM10BL-BFP using NheI- and EcoRI-restriction sites. pF143 encoding GFP, F145 encoding BFP and pBluescript (BSK) were described. [23,26] Bacterial expression plasmids pGEX_GST-Tasp-GFP and pGEX_GST-GFP were described. [20] A detailed overview of plasmids and oligonucleotides used for PCR amplification and cloning can be found in **Table S2** and **S3**.

Cells, Transfection and Microscopy

Leukemic and solid cancer cell lines used in the study were maintained and transfected as described [9,23]. Observation and image analysis of living or fixed cells were performed as described [7,27]. To determine the average intracellular protein localization, at least 200 fluorescent cells from three separate images were examined in three independent experiments, and representative images are shown. The number of cells exhibiting cytoplasmic (C; cytoplasmic signal >80% of the total cellular signal), cytoplasmic and nuclear (C/N), or nuclear (N; nuclear signal >80% of the total cellular signal) fluorescence was counted. As standards for this semiquantitative determination, the total cellular BFP (blue), GFP

(green) or mCherry/RFP (red) signal was quantitated by calculating the integrated pixel intensity in the imaged cell multiplied by the area of the cell in 100 fluorescent cells using a digital AxioCam CCD camera (Carl Zeiss, Jena, Germany) as described [7,27]. The nuclear signal was similarly obtained by measuring the pixel intensity in the respective nuclei. Nuclei were marked by Hoechst 33258 staining as described [20,27].

Criteria for efficient *in vivo* protein interaction using the protein interaction assay was that in >80% of 200 mCherry- and GFP-positive cells, mCherry and GFP co-localized at the nucleolus. Colocalization was further quantitated by confocal laser scanning microscopy using the “Overlap coefficient according to Manders” co-localization algorithm as described [24,27]. Colocalization coefficients represent “Overlap coefficient according to Manders” [9,24,27]. $R = \frac{\sum_i S1_i S2_i}{\sqrt{\sum_i (S1_i)^2 \sum_i (S2_i)^2}}$ R represents the calculated overlap coefficient, $S1$ and $S2$ the measured signal in the two channels, and i a specific pixel of the taken image. Therefore, the calculated R value indicates an overlap of the signals and thus represents the true degree of colocalization. Other presented coefficients resembling Pearson’s correlation coefficient R . Overlap coefficients k_1 and k_2 and colocalization coefficients m_1 and m_2 were described [27,28].

Protein Extraction and Immunoblot Analysis

Preparation of whole lysates from cells and immunoblotting were carried out as described [23]. Equal loading of lysates was controlled by reprobing blots for GapDH as described [7].

Gel Filtration Chromatography

MV4;11 cells were harvested by centrifugation at 350°g, 4°C, 5°min, washed with ice-cold phosphate-buffered saline and lysed by the addition of NETN buffer (100 mM NaCl; 10 mM Tris, pH 8; 10% Glycerol; 1 mM EDTA; 0.5% v/v NP40; 1 mM DTT; 1 mM PMSF; 1 fold Complete Protease Inhibitor – Roche, Germany) as described in [7]. After sonication lysates were centrifuged (14.000 rpm, 4°C, 30 min) and supernatant filtered (micro-centrifuge-filtered tubes 0.2 µm – Laborservice Onken GmbH) prior to gel filtration chromatography. The extracts were fractionated using size exclusion chromatography with Superose-6 10/300 GL columns (GE Healthcare FPLC system, optimal separation range from 3 MDa to 5 kDa). For system calibration purified aprotinin (6.5 kDa), ribonuclease A (13.7 kDa), carbonic anhydrase (29 kDa) and ovalbumin (44 kDa) were purchased from Sigma Aldrich (Sigma Aldrich, Munich, Germany) and used as standards. Loading and elution of the FPLC system was carried out in 50 mM Phosphate with 0.15 M NaCl pH 7. 500 µl fractions were collected and stored at –70°C until analyzed by SDS-PAGE as described [29].

Immunoprecipitation (IP)

IP of GFP-tagged proteins was performed using α-GFP magnetic beads and µ-MACS columns (MiltenyiBiotec, Bergisch Gladbach, Germany) [7,20]. Briefly, whole cell lysates were incubated with 50 µl α-GFP magnetic beads for 30 min on ice. Lysates with magnetic beads were applied onto the columns, washed, 20 µl elution buffer was applied and incubated for 5 min. To complete protein elution, 50 µl elution buffer were applied. 30 µl of the eluate, as well as 3% of input were analyzed as outlined [23].

Statistical Analysis

For experiments stating p -values, a paired Student’s t -test was performed. Unless stated otherwise, p -values represent data

obtained from three independent experiments done in triplicate. p -values <0.05 were considered significant [23].

Results

Monitoring Taspase1 Processing of AF4•MLL Substrates in Living Cells

As the AF4•MLL fusion is a substrate for SIAH1/SIAH2, its steady-state level is low, thereby hampering its detection and experimental analysis [8]. In fact, studies visualizing the intracellular localization of uncleaved or processed AF4•MLL protein are still missing [3]. Also, as biochemical data and *in vitro* interference strategies must be effective at the cellular level, they have to be verified *in vivo*. Hence, we set out to establish a suitable cell-based assay in the most relevant test tube, the living cell (**Figure 1**).

To analyze Taspase1’s processing of AF4•MLL substrates in living cells, we exploited a two component autofluorescent indicator protein system [22]. Therefore, the AF4•MLL residues surrounding the two Taspase1 cleavage sites (A•M_S1/2: aa 1582–1710 of the AF4•MLL fusion protein) were inserted into a backbone composed of GST, GFP, a N-terminal nuclear import (NLS) and a C-terminal nuclear export signal (NES) (**Figure 1a**). As second element the Taspase1 open reading frame was cloned from the Taspase1-expressing acute monocytic leukemia (AMoL) cell lines MV4;11 and THP-1, carrying a t(4;11) or t(9;11) translocation, respectively. Both sequences are identical to the one described by Hsieh et al. [6]. The rationale of this specific assay set-up was that the resulting NLS-GFP/GST-AF4•MLL_S1/2-NES fusion protein (A•M_S1/2) localizes predominantly to the cytoplasm, whereas Taspase1-mediated cleavage liberates the NES triggering nuclear accumulation (**Figure 1b**). However, due to the presence of two cleavage-sites, the A•M_S1/2 indicator protein was already (partially) cleaved by endogenous Taspase1 (**Figure 1c**), which was especially evident in cell lines with high protease expression levels (**Figure S1a**).

As this efficient processing precludes the use of the indicator protein in its current set-up, we engineered proteins harboring only individual AF4•MLL cleavage sites, A•M_S1 (¹⁶⁰⁰AEGQVD-GADD¹⁶⁰⁹) or A•M_S2 (¹⁶⁵²KISQLDGVDD¹⁶⁶¹), which are expected to be less efficiently processed. As shown in **Figure S1b–e**, we found that the addition of a linker sequence was crucial for the performance of the A•M_S1 and A•M_S2 indicator proteins. Both indicator proteins localize predominantly to the cytoplasm in cancer cells, whereas ectopic expression of biologically active Taspase1 promoted their cleavage and complete nuclear accumulation (**Figure 1d**). As a control, constructs containing non-functional Taspase1 cleavage sites (A•M_S1_{mut}, aa ¹⁶⁰⁰AEGQVAAADD¹⁶⁰⁹ or A•M_S2_{mut}, ¹⁶⁵²KISQ-LAAVDD¹⁶⁶¹) remained cytoplasmic (data not shown).

Also in leukemic cells Taspase1 localizes predominantly to cellular nucleus and both indicator proteins localize predominantly to the cytoplasm, while co-expression of either indicator protein and the active protease promotes the indicator protein’s nuclear accumulation (**Figure 2a**). Mutation of Thr²³⁴ into Val (Tasp^{T234V}) or Asp²³³ into Ala (Tasp^{D233A}) of Taspase1 affected autoprocessing as well as the protease’s *trans*-cleavage activity. Both mutants showed a nuclear but not nucleolar localization, but in contrast to the wild type protease co-expression of the catalytically inactive Tasp^{T234V}- or Tasp^{D233A}-GFP mutants did not result in effective cleavage and nuclear translocation of the indicator proteins. Similar results were obtained upon co-expression of untagged Tasp^{T234V} or Tasp^{D233A} as well as of their fusions to the red- (mCherry) or blue-fluorescent (BFP) protein, allowing their independent detection in living cells

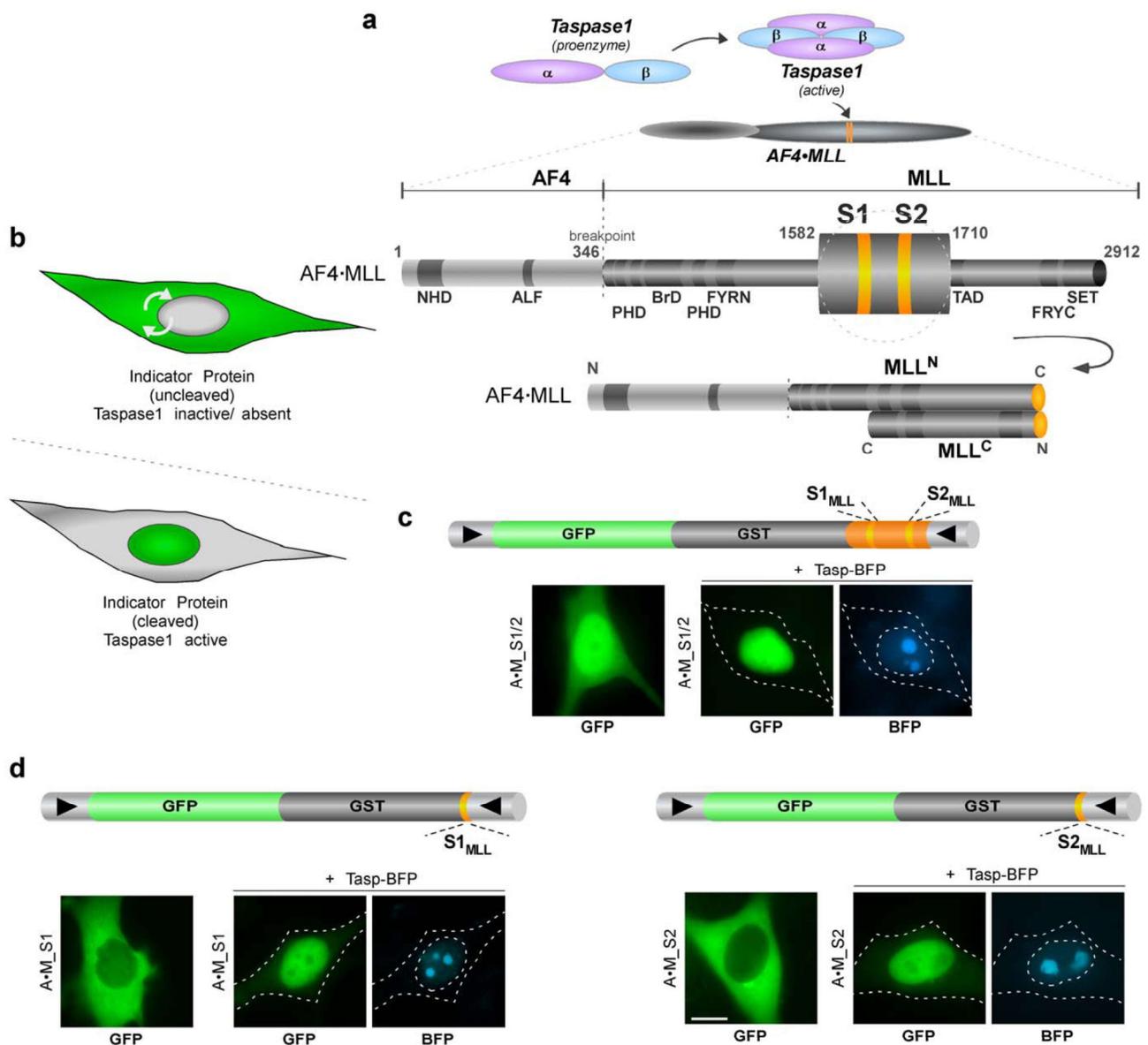


Figure 1. Analyzing Taspase1's processing of AF4-MLL substrates in living cells. A. Autoproteolysis of the Taspase1 proenzyme is assumed to trigger formation of the active $\alpha\beta\alpha$ -heterodimer, which hydrolyses the AF4-MLL fusion protein. Following processing, the cleavage products AF4-MLL^N and MLL^C heterodimerize, forming a high molecular-weight protein complex resistant to degradation. Domain organization of the AF4-MLL fusion. Taspase1 cleavage sites, S1 (QVDGADD) and S2 (QLDGVDD), are highlighted. NHD: N-terminal homology domain; ALF: AF4/LAF4/FMR2 homology domain; PHD: plant homeodomain; BrD: bromodomain; FYRN: F/Y rich domain N-terminal; TAD: transactivation domain; FRYC: F/Y rich domain C-terminal; SET: suppressor of variegation, enhancer of zeste and trithorax. Domains are not drawn to scale. **B.** Principle of the cell-based biosensor assay to analyze Taspase1-mediated AF4-MLL processing. The indicator protein localizes predominantly to the cytoplasm but is continuously shuttling between the nucleus and the cytoplasm. Co-expression of active Taspase1 results in the proteolytic removal of the NES, thereby triggering nuclear accumulation of the green fluorescent indicator. **C-D.** Domains of the indicator protein, composed of GST, GFP, combinations of a nuclear import (? : NLS) and an export (? : NES) signal, combined with the indicated cleavage sites of AF4-MLL. **c.** A-M_{S1/2} containing both cleavage sites is already partially processed by endogenous Taspase1 (left panel), but is completely nuclear upon expression of Taspase1-BFP (right panel). **D.** Indicator proteins containing only one cleavage site (A-M_{S1} or A-M_{S2}) are cytoplasmic in their uncleaved state, whereas ectopic expression of active Taspase1 triggers their cleavage and complete nuclear accumulation. GFP/BFP were visualized by fluorescence microscopy in living HeLa transfectants 24 h after transfection. Scale bars, 10 μ m. Dashed lines mark cytoplasmic/nuclear cell boundaries obtained from the corresponding phase contrast images. doi:10.1371/journal.pone.0034142.g001

(Figure 2a, Figure S2 and Table S4), which was confirmed by immunoblot analysis (Figure 2b) [6,13].

These results demonstrate the specificity of the assay system for Taspase1 and underline the advantage of using rational combi-

nations of multi-color autofluorescent proteins to study Taspase1's biological activity in living cells. Interestingly, we found that Tasp^{D233A} showed cleavage-site specificity, being able to process A-M_{S2R}, albeit with a highly attenuated activity, but not

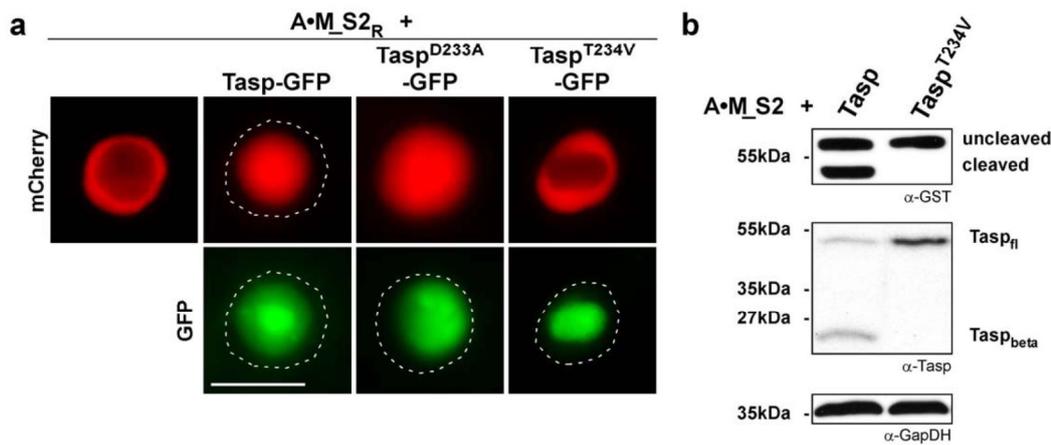


Figure 2. Activity and complex formation of Taspase1 and catalytically inactive mutants. **A.** Taspase1 processing of AF4•MLL substrates in leukemic cells. Co-transfection of Tasp-GFP resulted in proteolytic cleavage and nuclear accumulation of the red fluorescent biosensor, A•M_S2_R, in K562 cells. In contrast, co-expression of Tasp^{D233A}-GFP leads to partial processing and nuclear translocation, while Tasp^{T234V}-GFP was completely inactive. Localization was analyzed 24 h post transfection. GFP/mCherry were visualized by fluorescence microscopy. Scale bars, 10 μm. **B.** Processing of AF4•MLL substrates. Co-transfection of Tasp resulted in proteolytic cleavage of the biosensor A•M_S2_R in 293T cells as indicated by immunoblot. In contrast, Tasp^{T234V} was inactive in *cis* and *trans*. Proteins were visualized using α-GST or α-Taspase1 Abs. GapDH served as loading control. fl, unprocessed Taspase1; Tasp_β, Taspase1 β-subunit. doi:10.1371/journal.pone.0034142.g002

A•M_S1_R (Figure S2). Although fusions of Taspase1 with autofluorescent proteins have been shown to be fully functional, we confirmed these results by employing untagged or HA-tagged Taspase1 variants (data not shown) [23]. Hence, our system is also applicable to assess Taspase1 *trans*-cleavage activity on the individual AF4•MLL cleavage sites independently from each other.

Targeting Taspase1 Function in Trans by Catalytically Inactive Mutants

Subsequently, we used the established bioassay to investigate the consequences of overexpressing catalytically impaired Taspase1 mutants on the activity of the wild type (WT) enzyme *in trans*. We reasoned if inactive Taspase1 mutants are capable of forming heterodimers with WT Taspase1 (heterodimerization model), enforced overexpression of these mutants should have a dominant-negative effect. Besides the catalytically dead Tasp^{T234V}-GFP mutant, we also included Tasp^{D233A}-GFP in the analysis, as this variant exists in a biologically active though highly attenuated conformation. Notably, our assay demonstrated that even co-transfecting a nine-fold excess of the Tasp^{T234V}- or of the Tasp^{D233A}-GFP mutants over the WT Taspase1 expression plasmid did not affect Taspase1's processing of either the first or the second AF4•MLL cleavage site in solid as well as in leukemic cancer cell lines. These results could be independently verified in several solid as well as leukemic cancer cell lines (Figure 3a/b and Table 1). Immunoblot analysis confirmed that the mutants were efficiently overexpressed (Figure 3c). Similar results were obtained when using HA-tagged or untagged Taspase mutants (data not shown). To further exclude the formal possibility that our results are only valid for ectopically expressed Taspase1, we used the SaOs and SW480 cell lines expressing high levels of endogenous Taspase1 [23]. Upon expression in these cells, the A•M_S1/2 indicator protein is already fully or partially cleaved by endogenous Taspase1 resulting in its predominant nuclear localization (Figure S1a and Table 1). As expected, overexpression of the inactive Taspase1 variants did not inhibit the

endogenous enzyme and thus, did not affect cleavage of the indicator protein in *trans* (Table 1).

Next, we further analyzed whether *cis*-cleavage of WT Taspase1 could be affected *in trans*. As shown in Figure 3d, co-transfection of the WT protease with GFP-tagged or untagged mutants did not inhibit Taspase1's *cis*-cleavage activity, since the processed Taspase1 β-subunit was detectable in all plasmid combinations used. Immunoblot analysis verified that the Tasp^{T234V}- or Tasp^{D233A}-GFP proenzymes are impaired in their activation by autoproteolytic *cis*-cleavage (Figure 3d).

Also, we tested whether overexpression of the individual Taspase1 α- or β-subunit, which are clearly proteolytically inactive, affects Taspase1's *trans* cleavage. In line with the results obtained upon overexpression of full-length inactive Taspase1 variants, no inhibition of Taspase1's processing was detectable (Figure S3a/b).

To additionally exclude the possibility that the lack of a *trans*-dominant phenotype was restricted to the AF4•MLL protein, we tested the ability of the mutants to interfere with the processing of indicator proteins containing the cleavage-sites from the *bona fide* Taspase1 targets TFIIA (NLS-mCherry/GST-TFIIA_S-NES = TFIIA_S_R) and USF2 (NLS-mCherry/GST-USF2_S-NES = USF2_S_R) [7]. No inhibition of processing occurred for these substrates as well as for the full length TFIIA or USF2 proteins (Table 1).

Analysing Taspase1 Heterocomplex-formation

In general, interruption of pathological relevant protein complexes *via* enforced expression of *trans*-dominant negative mutants critically depends on efficient heterocomplex formation [9,30]. Thus, the lack of a *trans*-dominant negative effect upon overexpression of inactive Taspase1 mutants may be explained by inefficient heterocomplex formation *in vivo*. Expression of Taspase1-GFP in bacteria showed protein aggregation (Figure S3c), which had been previously reported [13]. Co-immunoprecipitation studies of overexpressed Taspase1 and GFP-fusions of the Taspase1 variants also indicated that the WT protein is in principle able to interact with biologically impaired mutants

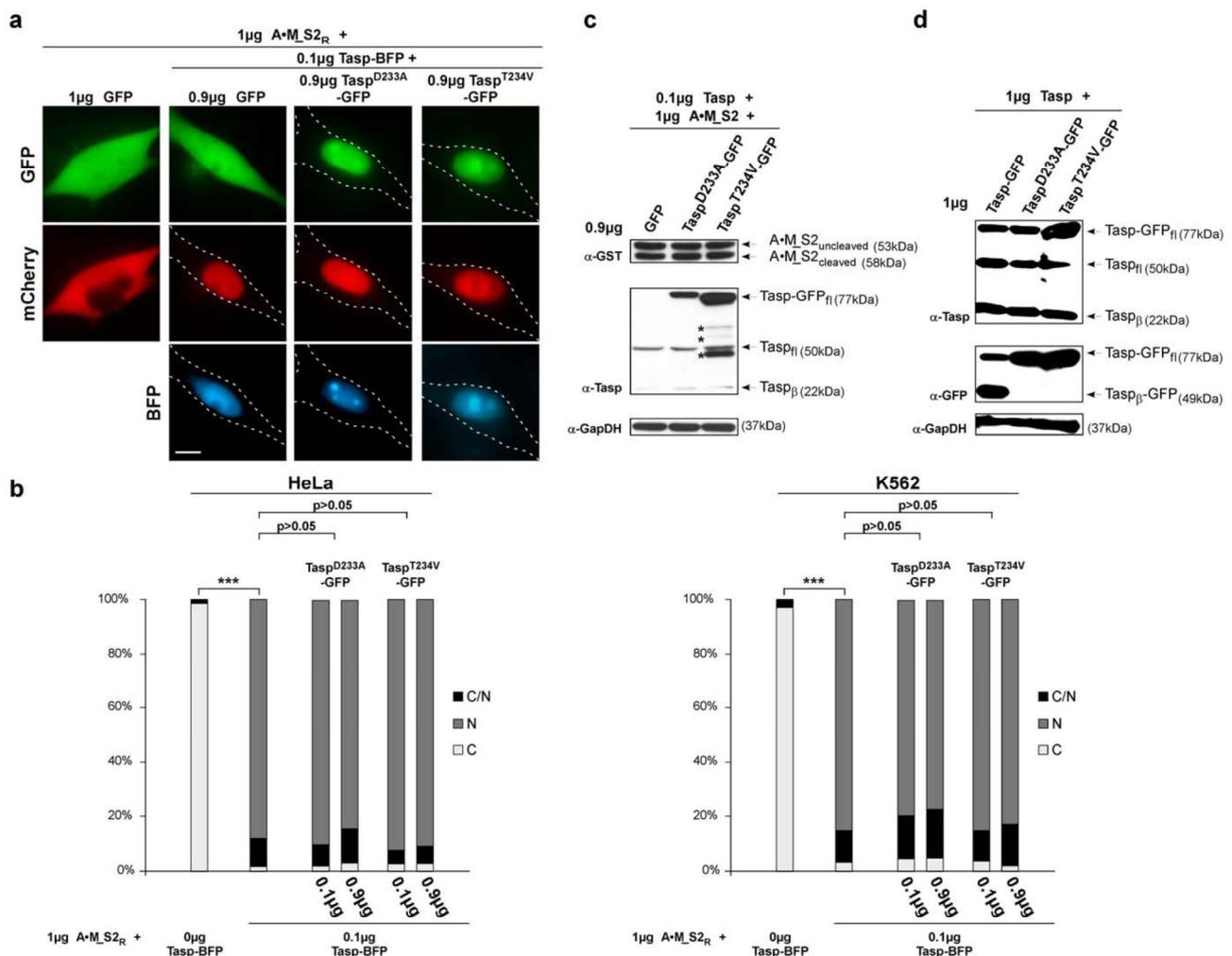


Figure 3. Overexpression of inactive Taspase1 mutants does not inhibit Taspase1's *cis*- or *trans*-cleavage activity. **A.** Cells were transfected with 1 µg of A•M_{S2}_R, 0.1 µg Tasp-BFP together with the indicated amounts of inactive Taspase1 mutants or GFP expression plasmid, and analyzed 24 h later. Even co-transfection of a nine-fold excess of plasmids encoding the inactive Taspase1 variants did not affect A•M_{S2}_R processing in living HeLa cells. **B.** The number of HeLa (left panel) or leukemic K562 cells (right panel) showing cytoplasmic (C), cytoplasmic and nuclear (N/C) or nuclear (N) fluorescence was counted in at least 200 A•M_{S2}_R-expressing cells. Results from one representative experiment of each indicated cell line are shown. Whereas the number of cell displaying cytoplasmic fluorescence significantly decreased by *trans*-cleavage upon co-transfection of 0.1 µg Tasp-BFP expression plasmid (***: $p < 0.0001$), no significant *trans*-dominant negative effect was evident for Taspase1 mutants. **C.** Taspase1 *trans*-cleavage of A•M_{S2}_R is unaffected by inactive Taspase1 mutants as shown by immunoblot analysis of 293T cells transfected with the indicated expression plasmids. Proteins and cleavage products were visualized using α-GST and α-Tasp Ab. GapDH served as loading control. **D.** *Cis*-cleavage of Taspase1 is not inhibited by inactive Taspase1 mutants as shown by immunoblot analysis of 293T cells transfected with 1 µg of the indicated expression plasmids. doi:10.1371/journal.pone.0034142.g003

(Figure 4a). However, when compared to complex formation of Taspase1 with a *bona fide* interaction partner, the nucleolar protein NPM1, the observed interaction was rather weak (Figure S3d) [23].

To further exclude that these results might be valid only for ectopically overexpressed Taspase1, we additionally examined the endogenous protein in MV4;11 human leukemia cells. These cells were isolated from a patient containing a t(4;11) translocation and thus, express the AF4•MLL fusion protein, which is processed by endogenous Taspase1. Employing gel filtration chromatography of cell lysates isolated under native conditions, we detected endogenous Taspase1 predominantly as an αβ-monomer (Figure S3e).

Probing Taspase1 Heterocomplex-formation in Living Cells by a Translocation-based Protein Interaction Assay

Subsequently, we applied a dual color translocation assay that allows visualization of protein complex formation in living cells (Figure 4b) to test our hypothesis. This principle has been successfully employed in several studies to assess protein interaction in living cells, including the t(4;11) leukemia relevant MLL-FYRN and -FYRC proteins [9,22,23,31]. Here, GFP-tagged Taspase1 was engineered to localize predominantly to the cytoplasm by C-terminal fusion of a strong nuclear export signal (NES) (Tasp_{Cyt}). Due to Taspase1's intrinsic nuclear import signal, Tasp_{Cyt} is continuously shuttling between the nucleus and the cytoplasm, and still catalytically active (Figure 4b/c) [23]. The red-fluorescent Taspase1 variants (Tasp-mCherry, prey) however

Table 1. Effects of overexpressing inactive Taspase1 mutants in trans on Taspase1's processing of various target proteins.

1 μ g indicator +	0.1 μ g BFP + 0.9 μ g GFP		0.1 μ g Tasp-BFP + 0.9 μ g GFP		0.1 μ g Tasp-BFP + 0.9 μ g Tasp ^{T234V} -GFP		0.1 μ g Tasp-BFP + 0.9 μ g Tasp ^{D233A} -GFP	
	Localization	C	N	C	N	C	N	C
A-M_S1 _R (HeLa)	>90%	<1%	<5%	>80%	<5%	>80%	<5%	>80%
A-M_S1 _R (K652)	>85%	<1%	<7%	>75%	<7%	>75%	<7%	>75%
A-M_S2 _R (HeLa)	>90%	<1%	<5%	>80%	<5%	>80%	<5%	>80%
A-M_S2 _R (K562)	>85%	<1%	<5%	>80%	<5%	>80%	<5%	>80%
TFIIA_S _R (HeLa)	>90%	<1%	<5%	>80%	<5%	>80%	<5%	>80%
USF2_S _R (HeLa)	>90%	<1%	<5%	>80%	<5%	>80%	<5%	>80%
	1 μ g BFP		1 μ g Tasp ^{T234V} -BFP		1 μ g Tasp ^{D233A} -BFP			
	C	N	C	N	C	N		
A-M_S1/2 (SaOs)	<2%	>90%	<2%	>90%	<2%	>90%		
A-M_S1/2 (SW480)	<5%	>80%	<5%	>80%	<5%	>80%		

Leukemic (K562) and solid tumor cells were transfected with the indicated amounts of the different indicator plasmids, together with respective control plasmids, or expression plasmids encoding active or inactive Taspase1 mutants, and analyzed 24 h later. The number of cells showing cytoplasmic (C) or nuclear (N) fluorescence was counted in at least 200 indicator protein-expressing cells. Results from one representative experiment are shown. Whereas the number of transfectants displaying cytoplasmic fluorescence, i.e., uncleaved indicator protein, significantly decreased upon co-transfection of 0.1 μ g Tasp-BFP expression plasmid (***) ($p < 0.0001$), no inhibition of cleavage was observed even upon co-transfection of 0.9 μ g expression plasmids encoding for the inactive Taspase1 mutants.

In transfectants with high (SaOs) or intermediate (SW480) levels of endogenous Taspase1, the A-M_S1/2 indicator protein (0.2 μ g expression plasmid) is already fully or partially cleaved in absence of ectopically expressed protease resulting in its predominant nuclear localization. A similar localization was observed upon co-expression of the inactive Taspase1 variants (1 μ g expression plasmid), indicating that the activity of endogenous Taspase1 is not inhibited in *trans*.

doi:10.1371/journal.pone.0034142.t001

accumulate in the nucleus/nucleolus (Figure 4c/d). Upon co-expression and efficient heterocomplex formation, the GFP-tagged Tasp_{Cyt} is expected to co-localize with the Tasp-mCherry prey variants in the nucleus/nucleolus. Therefore, nuclear translocation serves as a reliable indicator for efficient protein-protein interaction in living cells. This approach allows analyzing complex formation between the WT and the inactive mutant enzymes (Figure 4b). Co-expression of the positive control, NPM1-RFP, significantly triggered nuclear/nucleolar translocation of GFP-Tasp_{Cyt}, whereas co-expression of the non-interacting nucleolar RevM10BL-RFP protein (negative control) showed no effect (Figure 4d), confirming the assays specificity. As already expected from the functional data (Figure 3), co-expression of mutant Taspase1 variants did not result in strong nuclear/nucleolar translocation of Tasp_{Cyt}, indicative of only weak heterocomplex formation (Figure 4d). Similar results were obtained upon expression of untagged WT or mutant Taspase1 by immunofluorescence analysis in fixed cells (data not shown). To objectively quantitate the degree of co-localization, we employed confocal laser scanning microscopy revealing a colocalization R-value of 0.74 for NPM1-RFP, 0.19 for RevM10BL-RFP and R-values of 0.38–0.39 for WT and Taspase1 mutants, respectively (Table S4 and Figure S4). Hence, although WT or mutant Taspase1 variants are in principle able to form (hetero)complexes, these complexes appear to be rather unstable.

Subsequently, we investigated whether the individual α - or β -subunits efficiently bound to WT or mutant Taspase1 variants. We found that the nuclear Taspase1 α -BFP protein (Figure S5a, upper picture) was unable to efficiently multimerize with Tasp_{Cyt} and to recruit it to the nucleus (Figure S5b). Second, co-expression of Taspase1- or Tasp^{T234V}-mCherry did not induce nuclear/nucleolar translocation of Tasp β -GFP (Figure S5a, lower picture and S5c).

Of note, although the subunits were unable to efficiently interact with full length Taspase1, we though observed hetero-

complex formation when both subunits were co-expressed. As shown in Figure S5d, Tasp α -BFP or Tasp α -HA recruited Tasp β -GFP to the nucleus. Also, an engineered cytoplasmic Tasp β protein (Tasp β _{Cyt}), accumulated in the nucleus due to complex formation with nuclear Tasp α -BFP or Tasp α -HA (Figure S5e). These results are somehow unexpected as overexpression of the individual Taspase1 α - or β -subunits showed no *trans*-dominant negative effect (Figure S3c/d). Thus, we examined whether the complex formed upon coexpression of the individual subunits resembles biologically active Taspase1 generated by autoproteolytic cleavage of the proenzyme. Upon co-expression of Tasp α -BFP with Tasp β -GFP neither the AF4•MLL- nor the TFIIA-indicator protein was cleaved (Figure S5f and data not shown). Similar results were obtained by co-expressing Tasp α -HA with untagged Tasp β (data not shown). Hence, co-expression of the individual α - and β -subunit does not allow their assembly into an enzymatically active protease complex. Formation of the $\alpha\beta$ -monomer by *cis*-cleavage of the proenzyme seems to occur by a regulated step-wise process.

Discussion

Recent advances towards the understanding of cancer system biology inspired to consider cancer-related protein-protein interaction networks as potential therapeutic targets [15,16,17]. Recently, we used our *in vivo* protein interaction assay to also demonstrate that it is in principle possible to specifically inhibit the AF4•MLL oncoprotein by genetic PPIs inhibitors [9]. As the steady-state amount of the AF4•MLL protein is critically controlled through its processing by Taspase1, it is tempting to target the formation of the active protease by interfering with its multimerization as a novel strategy to block the pathological function of AF4•MLL. However, in order to potentially transfer such approaches into the clinics it is imperative to know whether the protein of interest indeed efficiently forms multimers *in vivo* causally required for its pathological functions.

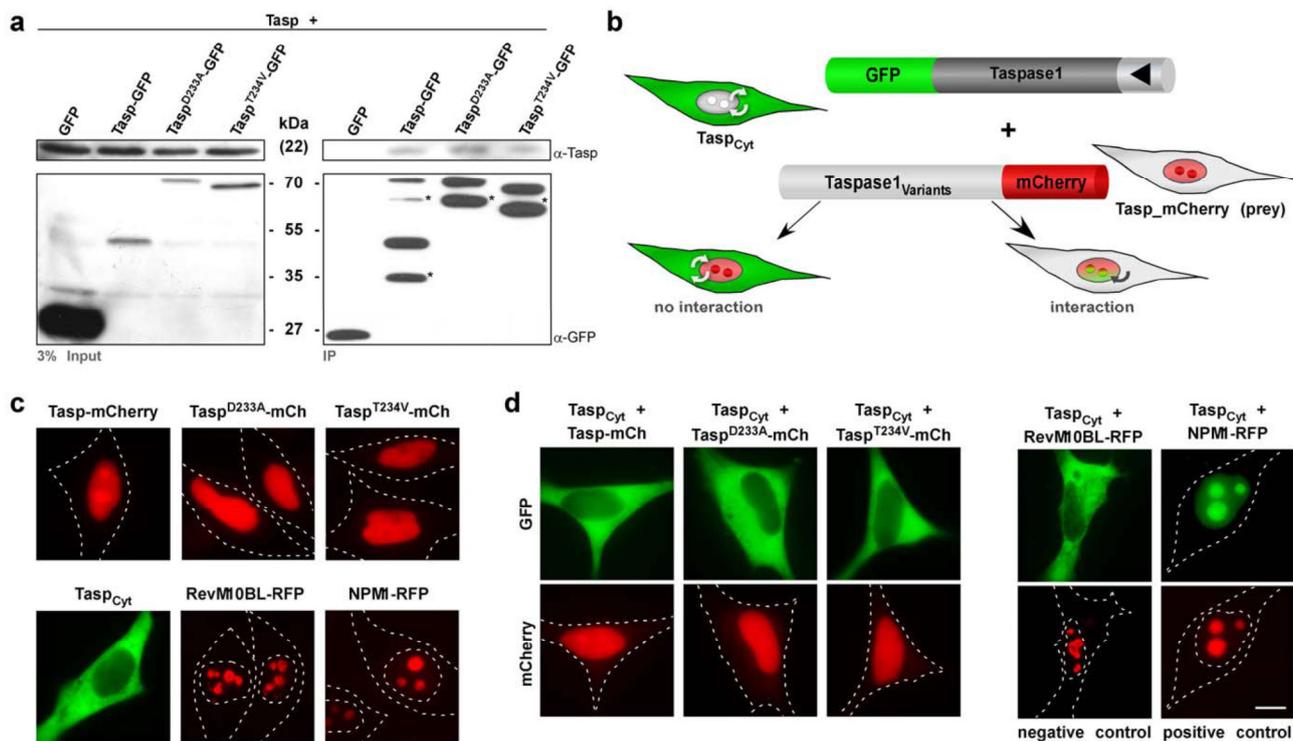


Figure 4. Probing Taspase1 multimerization in living cells. A. Heterocomplex formation of Taspase1 and Taspase1 variants shown by co-immunoprecipitation (IP). IPs of 293T cell extracts co-transfected with the indicated expression constructs were carried out using α -GFP Ab-coated magnetic beads and μ -MACS columns. Precipitated proteins were identified by immunoblot using the indicated antibodies. Input: Total amount of cell lysate. IP: immunoprecipitated proteins. *: GFP-degradation products [33]. **B.** Principle of the translocation based protein-protein interaction assay. The Tasp_{Cyt} fusion is composed of GFP, Taspase1 and a NES (?) and thus, continuously shuttling between the nucleus and the cytoplasm. The red fluorescent Taspase1 variants (Tasp mCherry prey) accumulate at the nucleus/nucleolus. Upon efficient protein-protein interaction, the GFP-tagged cytoplasmic Tasp_{Cyt} co-localizes with the Tasp-mCherry prey to the nucleus/nucleolus in living cells. **C.** Localization of indicated proteins in the absence of potential interaction partners. **D.** Neither co-expression of WT nor inactive Taspase1 variants resulted in strong nuclear/nucleolar translocation of Tasp_{Cyt}. Co-expression of NPM1-RFP, known to strongly interact with Taspase1, triggered nuclear/nucleolar translocation of Tasp_{Cyt} (positive control). In contrast, co-expression of the non-interacting nucleolar RevM10BL-RFP protein showed no effect (negative control) as visualized by fluorescence microscopy in living HeLa transfectants. Scale bars, 10 μ m. doi:10.1371/journal.pone.0034142.g004

For Taspase1 it is assumed that following autoproteolysis of the zymogen, its subunits assemble into an asymmetric $\alpha\beta\beta\alpha$ -heterodimer, representing the active protease [6,13]. This model is mainly based on the crystal structures of other type 2 asparaginases, as well as on the structure obtained from bacterially expressed Taspase1 [13]. Thus, it was concluded that the enzymes consist as a four-layered $\alpha\beta\beta\alpha$ structure, with a central, mostly anti-parallel β -sandwich that is surrounded by α -helices on both faces [6,13]. However, experimental evidence convincingly demonstrating that not only Taspase1 but also other type 2 asparaginases do exist in their natural environment as heterodimers, and that multimerization is indeed essential for their biological activities is still missing. Clearly, the structure resolved by Khan et al. provided important insights into Taspase1 function, albeit some limitations may exist [13]. For example, the position of critical functional domains, such as the bipartite NLS can't be deduced from the current computational model of Taspase1 as these residues are disordered [13,23]. Also, the structure of the $\alpha\beta\beta\alpha$ -heterodimer was obtained by co-crystallizing the individual subunits rather than the autoproteolytically processed zymogen. As shown in our study, co-expression of the individual Taspase1 subunits was unable to assemble into a functional protease *in vivo*.

Based on our data it is thus conceivable to speculate that *in vivo* a complex equilibrium between Taspase1 dimers and already active

$\alpha\beta$ -monomers might exist (Figure 5). According to the "heterodimer model", the full length Taspase1 zymogen dimerizes, and upon autoproteolysis assembles into an asymmetric Taspase1 $_{\alpha\beta\beta\alpha}$ -heterodimer, representing the active protease. Hence, Taspase1 is expected to exist in equilibrium of full length Taspase1 monomers, unprocessed Taspase1 dimers as well as active processed Taspase1 $_{\alpha\beta\beta\alpha}$ -heterodimers. The Taspase1 $_{\alpha\beta\beta\alpha}$ -heterodimers may further dissociate into free Taspase1 $_{\alpha}$ and Taspase1 $_{\beta}$ subunits. The formation of these forms is regulated by their association (k_1) and dissociation constants (k_{-1}) as well as by the kinetics of autoproteolysis, which have not been determined yet (Figure 5a-c).

Interruption of pathobiological relevant protein complexes *via* enforced expression of *trans*-dominant negative mutants has been employed in several disease models and requires efficient heterocomplex formation [15,32]. Assuming that inactive Taspase1 variants are capable of interacting efficiently with the wild type enzyme, a nine-fold overexpression of inactive Taspase1 variants would strongly shift the equilibrium towards the formation of catalytically impaired heterodimers, resulting in a significant *trans*-dominant negative phenotype *in vivo*. For the cases reported, inhibition was already evident upon equimolar co-expression of WT protein and *trans*-dominant mutants, in contrast to what we observed for Taspase1 and inactive Taspase1 variants.

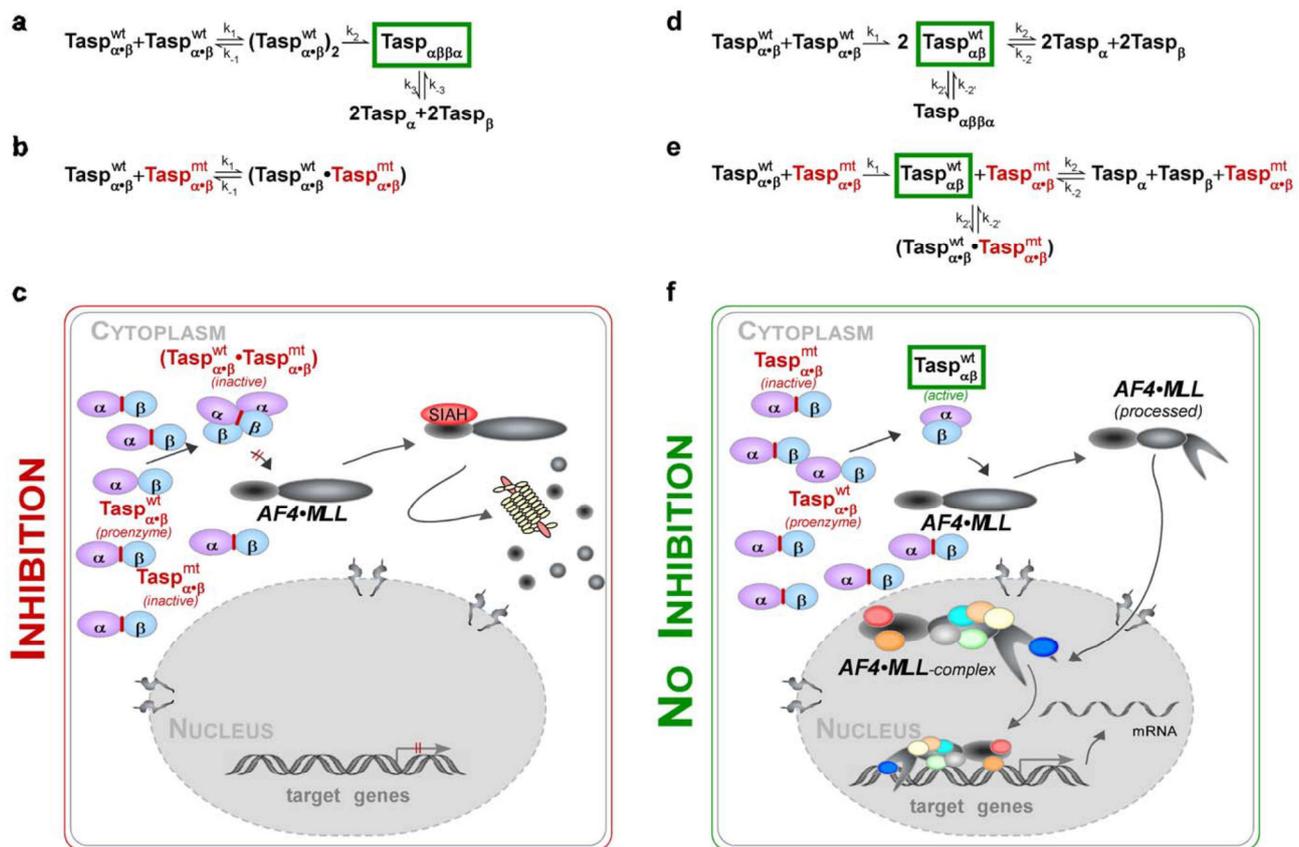


Figure 5. Models illustrating how Taspase1 heterocomplex formation determines the biological effects of overexpressing inactive Taspase1 mutants. **A–C:** Heterodimer model - allowing inhibition of Taspase1 function by *trans* dominant mutants. **A.** Upon translation, the Taspase1 zymogen dimerizes and following autoproteolysis matures into an asymmetric Taspase1_{αβ}-heterodimer, representing the active protease. Taspase1 exist in equilibrium of unprocessed Taspase1 monomers, unprocessed Taspase1 dimers, and active processed Taspase1_{αβ}-heterodimers. The Taspase1_{αβ}-heterodimers may further dissociate into free Taspase1_α and Taspase1_β subunits. **B.** Co-expression of an excess of inactive Taspase1 variants results in the formation of catalytically impaired heterodimers, reducing the concentration of active Taspase1 molecules. **C.** Consequently, AF4-MLL processing is inhibited allowing its degradation by SIAH1/2, thereby preventing the activation of cellular proliferation programs. **D–F:** Monomer model - predicting Taspase1's resistance to enforced expression of inactive mutants. **D.** The Taspase1_{αβ} proenzyme is autoproteolytically cleaved, forming an active Taspase1_{αβ} monomer. The processed Taspase1_{αβ} monomer seems to exist also as a Taspase1_{αβ}-heterodimer, and potentially in equilibrium with its subunits. **E.** Overexpression of inactive Taspase1 variants does not affect the concentration and activity of Taspase1_{αβ} monomers. **F.** Hence, Taspase1_{αβ} monomers are able to cleave the AF4-MLL fusion protein, resulting in the formation of a SIAH-resistant AF4-MLL complex allowing the activation of target genes driving oncogenesis. doi:10.1371/journal.pone.0034142.g005

Albeit the current literature does not indicate that tagged- and untagged-Taspase1 behaves differently [7,23; this report], we are aware that the results of the *in vivo* protein interaction assays might be affected by the use of Taspase1 variants fused to autofluorescent proteins.

Alternatively, our data could be interpreted by a “monomer model”, in which the Taspase1_{αβ} proenzyme is autoproteolytically cleaved forming a Taspase1_{αβ} monomer, already representing the active protease (Figure 5d–f). According to this model, the relative concentrations of these forms and thus, protease activity are regulated by the kinetics of autoproteolysis. As such, it is expected that even an excess of inactive Taspase1 variants would not affect the formation and biological activity of the Taspase1_{αβ} monomers, which is in line with our experimental evidence. Such a model though does not exclude that the cellular pool is composed of Taspase1_{αβ} monomers as well as Taspase1_{αβ}-heterodimers. Whether such multimerization has additional biological implications, such as preventing Taspase1 degradation, conferring cleavage-site selectivity or specificity remains to be

resolved. Theoretically, processed Taspase1_{αβ} may also exist in equilibrium with its subunits. However our observation that co-expression of the individual α- and β-subunit does not allow their assembly into an enzymatically active protease complex, argues against a highly dynamic exchange of the subunits. Hence, assembly of the αβ-monomer by *cis*-cleavage of the proenzyme seems to occur by an ordered, stepwise process, which might be guided by molecular chaperones, such as NPM1 [23]. In line with our findings, Khan et al. reported significant differences in the structures obtained by co-expression of the individual Taspase1 subunits *versus* the full-length enzyme [13].

As one might argue that the catalytically inactive Tasp^{T234V} variant is unable to achieve a conformation allowing efficient complex formation with active Taspase1, we included the highly attenuated Tasp^{D233A} mutant in the analysis. The molecular mechanism explaining why this mutant showed cleavage-site specificity is not yet resolved. Although this variant thus exists, at least transiently, in a biologically relevant conformation, we did not observe an inhibitory effect.

In conclusion, we provide first evidence that (i) overexpression of biologically inactive Taspase1 mutants as well as of the α - or β -subunits does not interfere with Taspase1's *cis*- and *trans*-cleavage activity; and (ii) complexes formed by the individual Taspase1 α and β subunit are inactive and do not reflect biologically active Taspase1 generated by autoproteolytic cleavage of the zymogen.

We are aware that besides the models presented here, our results may be explained by alternative models, in which the tetramer only needs one active dimer or has such a high turn over that the effects are not measurable. Clearly, comprehensive experimental work is required to discriminate between these molecular scenarios *in vivo*. However, this would imply that targeting Taspase1 heteromultimerization by genetic or chemical decoys is unlikely to interfere with its (patho)biological activity, including activation of the AF4•MLL oncoprotein (Figure 5f). As stabilization of protein complexes by chemical decoys is currently considered an alternative approach to inhibit disease-relevant pathways, it is tempting to investigate such an approach also for Taspase1. Based on our work and studies reported so far, strategies to dissect and selectively inhibit the (patho)biological activity of Taspase1 in oncogenesis are still advised to focus on the (high-throughput) identification of chemicals targeting Taspase1's catalytic activity.

Supporting Information

Figure S1 Optimization of indicator proteins to monitor AF4•MLL processing. **A.** Nuclear localization of the A•M_S1/2 indicator protein in cancer cell lines expressing high levels of endogenous Taspase1. **B–E.** Optimization of the AF4•MLL cleavage indicator proteins by addition of linker sequences shown in HeLa cells. Integration of the Taspase1 AF4•MLL recognition site alone does not allow processing and nuclear accumulation of the indicator protein by ectopically expressed Taspase1-BFP (**B**) Processing was improved by the integration of a GSGS- (**C**) or KIS-linker (**D**) N-terminal to the cleavage site. The A•M_S2 indicator protein containing the KISQLDGVDDGSGS cleavage site (spacer sequence underlined) showed optimal performance, cytoplasmic in the absence of ectopic Taspase1, whereas co-expression of Taspase1-BFP triggered proteolytic cleavage and complete nuclear translocation (**E**). BFP/GFP-fusion was visualized by fluorescence microscopy in living transfectants. Scale bars, 10 μ m. Dashed lines mark cytoplasmic/nuclear cell boundaries obtained from the corresponding phase contrast images. (PDF)

Figure S2 Taspase1 *trans* processing of AF4•MLL substrates shows cleavage site-specificity. **A.** Whereas the indicator protein A•M_S1_R, containing the first cleavage-site from AF4•MLL, was efficiently processed by Tasp-GFP, both Taspase1 mutants, Tasp^{T234V}- or Tasp^{D233A}-GFP, were inactive. **B.** In contrast, Tasp^{D233A}-GFP was able to partially process A•M_S2_R, containing the second cleavage-site from AF4•MLL, whereas Tasp^{T234V}-GFP was inactive. Proteins were visualized by fluorescence microscopy in living HeLa cell transfected with the indicated expression plasmids 24 h after transfection. Scale bar, 10 μ m. **C–D.** Cytoplasmic (C), cytoplasmic and nuclear (N/C) or nuclear (N) fluorescence was counted in at least 200 A•M_S1_R (**C**) or A•M_S2_R (**D**) -expressing HeLa cell co-transfected with the indicated expression plasmids. Results from a representative experiment are shown. The number of cells displaying cytoplasmic fluorescence significantly decreased upon cotransfection of 0.1 μ g Tasp-BFP expression plasmid (***: $p < 0.0001$). Neither Tasp^{T234V}- nor Tasp^{D233A}-GFP cleaved A•M_S1_R, but

Tasp^{D233A}-GFP was able to partially process A•M_S2_R, containing the second AF4•MLL cleavage-site. (PDF)

Figure S3 **A.** Expression of GST-Tasp1-GFP (upper panel) in BL21 bacteria shows extensive protein aggregation. In contrast, GST-GFP showed no aggregation (lower panel). Images were taken with identical CCD camera settings. Scale bar, 1 μ m. **B.** NPM1 strongly interacts with Tasp-GFP. IPs of 293T cell extracts co-transfected with the indicated expression. Precipitated proteins were identified by immunoblot using the indicated antibodies. Input: Total amount of cell lysate. IP: Immunoprecipitated proteins. #: GFP-degradation products. **C–D.** Taspase1 *trans*-cleavage is unaffected by overexpression of the Taspase1 α -subunit. HeLa cells were co-transfected with the indicated expression plasmid and analyzed 24 h later. **C.** Even co-transfection of a nine-fold excess of the nuclear Tasp α -GFP did not affect A•M_S2_R processing and its nuclear translocation. The cleaved red-fluorescent indicator protein, Tasp α -GFP, and active Tasp-BFP fusions were independently visualized by fluorescence microscopy in living cells. A representative cell is shown. Scale bar, 10 μ m. **D.** The number of cells showing cytoplasmic (C), cytoplasmic and nuclear (N/C) or nuclear (N) fluorescence was counted in at least 200 A•M_S2_R-expressing cells. Results from a representative experiment are shown. Whereas the number of cell displaying cytoplasmic fluorescence significantly decreased upon co-transfection of 0.1 μ g Tasp-BFP expression plasmid, overexpression of Tasp α -GFP or GFP alone did not inhibit the activity of Tasp-BFP in *trans*. **E.** Endogenous Taspase1 is detectable predominantly as an $\alpha\beta$ -monomer. Cell lysates isolated under native conditions from MV4;11 human leukemia cells were separated by gel filtration chromatography and resolved by 1D-SDS PAGE. Immunoblot analysis of FPLC of MV4;11 cell lysates. Endogenous Taspase1 was visualized in the fractions (49 to 94 kDa) by immunoblot using α -Tasp Ab. *: degradation products. (PDF)

Figure S4 Quantitating Taspase1 protein-interaction in living cells by confocal microscopy. HeLa cells were transfected with the indicated expression plasmids and protein localization as well as co-localization analyzed by confocal microscopy 24 h post transfection. Scale bars, 10 μ m. **A–B.** Localization of RevM10BL- (negative control), NPM1-RFP (positive control), Tasp_{Cyt}, and the red-fluorescent Taspase1 variants (Tasp_{mCherry}-prey) in the absence of potential interaction partners in living cells. **C–D.** Quantitation of protein co-localization shown as scatter gram with the gained Manders overlap coefficient indicated (R values). **C.** Co-expression of RevM10BL-RFP had no effect on Tasp_{Cyt} localization (R = 0.1928), whereas efficient nuclear/nucleolar translocation was observed upon co-expression of NPM1-RFP (R = 0.7354). **D.** In contrast, neither co-expression of WT (R = 0.3867) nor mutant Taspase1 variants (Tasp^{D233A}-mCherry, R = 0.3942; Tasp^{T234V}-mCherry, R = 0.3876) resulted in strong nuclear/nucleolar translocation of Tasp_{Cyt}, indicative of only weak heterocomplex formation in living cells. (PDF)

Figure S5 Translocation assay to analyze complex formation of Taspase1 subunits. **A–C.** The Taspase1 α - or β -subunits do not form stable heterocomplexes with WT Taspase1. **A.** Localization of Taspase1 α - or β -subunits in HeLa transfectants. Tasp α -BFP localizes to the nucleus, whereas Tasp β -GFP is nuclear and cytoplasmic. **B.** Co-expression of Tasp α -BFP did not trigger nuclear/nucleolar translocation of full length

Tasp_{Cyt}. **C.** Also, co-expression of nuclear/nucleolar Tasp-mCh did not translocate Tasp_β-GFP to the nucleolus. Autofluorescent fusion proteins were visualized in the same cells by fluorescence microscopy. **D–E.** Co-expression of the isolated Taspase1 subunits results in complex formation. **D.** Upon co-expression, nuclear Tasp_α-BFP associates with Tasp_β-GFP and recruits to the nucleus. **E.** Also, a cytoplasmatic GFP-Tasp_β protein (Tasp-β_{Cyt}), generated by fusion of a strong nuclear export signal (left panel), accumulated in the nucleus by binding to nuclear Tasp_α-BFP (right panel). **F.** Upon co-expression the isolated Taspase1 subunits do not assemble into an enzymatically active protease complex. Co-expression of Tasp_α-BFP with Tasp_β-GFP does not result in processing of the A•M_{S2R} indicator protein. The uncleaved red-fluorescent indicator protein, Tasp_α-BFP, and Tasp_β-GFP were independently visualized by fluorescence microscopy in living cells. A representative cell is shown. Scale bar, 10 μm.

(PDF)

Table S1 List of described disease-associated MLL fusions. *Abbreviations:* ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; CML, chronic myeloid leukemia; JMML, juvenile myelomonocytic leukemia; AUL/ANL, acute undifferentiated leukemia/acute nonlymphocytic leukemia; MDS, myelodysplastic syndromes; tALL/tAML/tMDS, therapy related ALL/AML/MDS; tT-ALL, therapy related T-cell ALL. X: indicates the presence of a putative Taspase1 cleavage site, based on the Taspase1 recognition sequence (Q³[F,I,L,V]²D¹ ↓ G¹x²D³D⁴) [1].

(PDF)

References

- Meyer C, Kowarz E, Hofmann J, Renneville A, Zuna J, et al. (2009) New insights to the MLL recombinome of acute leukemias. *Leukemia* 23: 1490–1499.
- Montes R, Ayllon V, Gutierrez-Aranda I, Prat I, Hernandez-Lamas MC, et al. (2011) Enforced expression of MLL-AF4 fusion in cord blood CD34+ cells enhances the hematopoietic repopulating cell function and clonogenic potential but is not sufficient to initiate leukemia. *Blood* 117: 4746–4758.
- Bursen A, Schwabe K, Ruster B, Henschler R, Ruthardt M, et al. (2010) The AF4.MLL fusion protein is capable of inducing ALL in mice without requirement of MLL.AF4. *Blood* 115: 3570–3579.
- Hsieh JJ, Ernst P, Erdjument-Bromage H, Tempst P, Korsmeyer SJ (2003) Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. *Mol Cell Biol* 23: 186–194.
- Yokoyama A, Kitabayashi I, Ayton PM, Cleary ML, Ohki M (2002) Leukemia proto-oncoprotein MLL is proteolytically processed into 2 fragments with opposite transcriptional properties. *Blood* 100: 3710–3718.
- Hsieh JJ, Cheng EH, Korsmeyer SJ (2003) Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. *Cell* 115: 293–303.
- Bier C, Knauer SK, Klaphor A, Schweitzer A, Reikik A, et al. (2011) Cell-based Analysis of Structure-Function Activity of Threonine Aspartase 1. *J Biol Chem* 286: 3007–3017.
- Bursen A, Moritz S, Gaussmann A, Moritz S, Dingermann T, et al. (2004) Interaction of AF4 wild-type and AF4.MLL fusion protein with SIAH proteins: indication for t(4;11) pathobiology? *Oncogene* 23: 6237–6249.
- Pless B, Oehm C, Knauer S, Stauber RH, Dingermann T, et al. (2011) The heterodimerization domains of MLL-FYRN and FYRC-are potential target structures in t(4;11) leukemia. *Leukemia* 25: 663–670.
- Benedikt A, Baltruschat S, Scholz B, Bursen A, Arrey TN, et al. (2011) The leukemogenic AF4-MLL fusion protein causes P-TEFb kinase activation and altered epigenetic signatures. *Leukemia* 25: 135–144.
- Liu H, Cheng EH, Hsieh JJ (2009) MLL fusions: pathways to leukemia. *Cancer Biol Ther* 8: 1204–1211.
- Marschalek R (2010) Mixed lineage leukemia: roles in human malignancies and potential therapy. *Febs J* 277: 1822–1831.
- Khan JA, Dunn BM, Tong L (2005) Crystal structure of human Taspase1, a crucial protease regulating the function of MLL. *Structure* 13: 1443–1452.
- Huang YJ, Hang D, Lu IJ, Tong L, Gerstein MB, et al. (2008) Targeting the human cancer pathway protein interaction network by structural genomics. *Mol Cell Proteomics* 7: 2048–2060.
- Arkin MR, Whitty A (2009) The road less traveled: modulating signal transduction enzymes by inhibiting their protein-protein interactions. *Curr Opin Chem Biol* 13: 284–290.
- Tanaka T, Rabbits TH (2008) Interfering with protein-protein interactions: potential for cancer therapy. *Cell Cycle* 7: 1569–1574.
- Weber L (2010) Patented inhibitors of p53-Mdm2 interaction (2006–2008). *Expert Opin Ther Pat* 20: 179–191.
- Srinivasan RS, Nesbit JB, Marrero L, Erfurth F, LaRussa VF, et al. (2004) The synthetic peptide PFWT disrupts AF4-AF9 protein complexes and induces apoptosis in t(4;11) leukemia cells. *Leukemia* 18: 1364–1372.
- Wichmann C, Chen L, Heinrich M, Baus D, Pfizner E, et al. (2007) Targeting the oligomerization domain of ETO interferes with RUNX1/ETO oncogenic activity in t(8;21)-positive leukemic cells. *Cancer Res* 67: 2280–2289.
- Knauer SK, Fetz V, Rabenstein J, Friedl S, Hofmann B, et al. (2011) Bioassays to monitor taspase1 function for the identification of pharmacogenetic inhibitors. *PLoS One* 6: e18253.
- Lee JT, Chen DY, Yang Z, Ramos AD, Hsieh JJ, et al. (2009) Design, syntheses, and evaluation of Taspase1 inhibitors. *Bioorg Med Chem Lett* 19: 5086–5090.
- Knauer SK, Moodt S, Berg T, Liebel U, Pepperkok R, et al. (2005) Translocation biosensors to study signal-specific nucleocytoplasmic transport, protease activity and protein-protein interactions. *Traffic* 6: 594–606.
- Bier C, Knauer SK, Docter D, Schneider G, Krämer OH, et al. (2011) The Importin-Alpha/Nucleophosmin Switch Controls Taspase1 Protease Function. *Traffic* 12: 703–714.
- Krämer OH, Knauer SK, Greiner G, Jandt E, Reichardt S, et al. (2009) A phosphorylation-acetylation switch regulates STAT1 signaling. *Genes Dev* 23: 223–235.
- Knauer SK, Bier C, Schlag P, Fritzmeyer J, Dietmaier W, et al. (2007) The survivin isoform survivin-3B is cytoprotective and can function as a chromosomal passenger complex protein. *Cell Cycle* 6: 1502–1509.
- Knauer SK, Carra G, Stauber RH (2005) Nuclear export is evolutionarily conserved in CVC paired-like homeobox proteins and influences protein stability, transcriptional activation, and extracellular secretion. *Mol Cell Biol* 25: 2573–2582.
- Knauer SK, Bier C, Habtemichael N, Stauber RH (2006) The Survivin-Crm1 interaction is essential for chromosomal passenger complex localization and function. *EMBO Rep* 7: 1259–1265.
- Schlingemann J, Habtemichael N, Itrich C, Toedt G, Kramer H, et al. (2005) Patient-based cross-platform comparison of oligonucleotide microarray expression profiles. *Lab Investigation* 85: 1024–1039.

Table S2 List of plasmids used in the study. Plasmid name, encoded protein, and function are indicated. *: tag used for detection.

(PDF)

Table S3 Oligonucleotides used for PCR amplification and cloning. Oligonucleotide name and nucleotide sequence are indicated.

(PDF)

Table S4 Quantitation of Taspase1 heterocomplex formation by confocal laser scanning microscopy in living cells. HeLa cells were co-transfected with 1 μg of Tasp_{Cyt} and 1 μg of the indicated mCherry-prey expression plasmids, and analyzed 24 h later. Colocalization coefficients as an indicator of complex formation were calculated using the colocalization pro software. Rr, Pearson's correlation coefficient; R, overlap coefficient according to Manders; k1/k2, overlap coefficients; m1/m2, colocalization coefficients. Results from a representative experiment are shown. A colocalization R-value of 0.74 for NPM1-RFP with Tasp_{Cyt} indicates 74% of colocalization.

(PDF)

Acknowledgments

We thank Sandra Friedl for excellent technical assistance.

Author Contributions

Conceived and designed the experiments: CB SK RS. Performed the experiments: CB RH LK DW DG SS. Analyzed the data: CB OK GS SK RS. Contributed reagents/materials/analysis tools: RS OK GS. Wrote the paper: CB SK RS.

29. Tenzer S, Docter D, Rosfa S, Wlodarski A, Kuharev J, et al. (2011) Nanoparticle size is a critical physico-chemical determinant of the human blood plasma corona - A comprehensive quantitative proteomic analysis. *ACS Nano* 5: 2011–01950e.R1.
30. Heger P, Rosorius O, Hauber J, Stauber RH (1999) Titration of cellular export factors, but not heteromultimerization, is the molecular mechanism of transdominant HTLV-1 rex mutants. *Oncogene* 18: 4080–4090.
31. Knauer SK, Stauber RH (2005) Development of an autofluorescent translocation biosensor system to investigate protein-protein interactions in living cells. *Anal Chem* 77: 4815–4820.
32. Stauber RH, Afonina E, Gulnik S, Erickson J, Pavlakis GN (1998) Analysis of intracellular trafficking and interactions of cytoplasmic HIV-1 Rev mutants in living cells. *Virology* 251: 38–48.
33. Landgraf P, Wahle P, Pape HC, Gundelfinger ED, Kreuz MR (2008) The survival-promoting peptide Y-P30 enhances binding of pleiotrophin to syndecan-2 and -3 and supports its neuritogenic activity. *J Biol Chem* 283: 25036–25045.

Die molekularen Mechanismen, die für die Aktivierung der Taspase1 durch die autoproteolytische Spaltung verantwortlich sind, sind bisher noch nicht vollständig verstanden. Es konnte jedoch gezeigt werden, dass sie evolutionär hoch konserviert sind.

3.2.3 „Fly versus man: evolutionary impairment of nucleolar targeting affects the degradome of *Drosophila*'s Taspase1“

Désirée Wunsch, Angelina Hahlbrock, Christina Heiselmayer, Sandra Bäcker, Patrick Heun, **Dorothee Goesswein**, Walter Stöcker, Tanja Schirmeister, Günter Schneider, Oliver H. Krämer, Shirley K. Knauer, Roland H. Stauber.

The FASEB Journal, 2015.⁶

Zusammenfassung

Durch Proteasen regulierte Prozesse spielen nicht nur in der Karzinogenese, sondern auch in vielen anderen essentiellen physiologischen Prozessen des Organismus eine große Rolle, u.a. bei der intra- und interzellulären Proteindegradation, im Metabolismus, bei der Geweberemodellierung, in der Apoptose, Proliferation und Migration^{95,197}. Insbesondere auf Grund der Irreversibilität der posttranslationalen Proteolyse sind Proteasen äußerst strikt reguliert. Taspase1 ist als für die Spaltung des Mixed-Lineage Leukemia (MLL)-Proteins verantwortliches Enzym essentiell für die embryonale Entwicklung humaner Zellen^{45,120}. Das Ortholog von MLL in *Drosophila melanogaster*, trithorax, wird durch die *Drosophila* Taspase1 (dTaspase1) gespalten und zeugt von einer funktionellen Ko-Evolution. Die weiterführende Analyse der dTaspase1 erlaubte den Vergleich zur humanen Taspase1. So besitzt die dTaspase1 ein essentielles, nukleophiles Threonin¹⁹⁵, welches die *cis* Spaltung in die aktive α - und β -Untereinheit katalysiert. Ein bereits bewährter zellbasierter Assay wurde so modifiziert, dass mittels substitutiver Alanin-Mutationen das Spaltungsmotiv

⁶ Eine Kopie des Artikels wird im Folgenden abgedruckt unter der *Creative Commons Attribution 4.0 International License*. <https://creativecommons.org/licenses/by/4.0/>

(Q³[F,I,L,M]²D¹↓G¹X²X³) für dTaspase1 identifiziert werden konnte. Das daraus resultierende, *in silico* vorhergesagte Degradom ist mit 70 potentiellen Substraten für die dTaspase1 deutlich größer als das der humanen Taspase1 mit 27¹²². Die Aktivität der humanen Taspase1 ist durch die Interaktion mit dem ein Kernexportsignal tragenden Protein Nucleophosmin1 und einen Importin- α -vermittelten Kernimport geprägt, welche zum einen die dominierende Lokalisation der Taspase1 im Nukleolus verursachen, aber auch die Spaltung zytoplasmatischn Substrate ermöglicht¹³⁰. dTaspase1 zeigt hingegen keine spezifische Lokalisation innerhalb der Zelle und auch keine Interaktion mit Nucleophosmin1 oder Importin- α . Dies liegt in einer fehlenden Interaktionsdomäne (NoLS) begründet, welche erst in Vertebraten konserviert ist, nicht aber in *Drosophila*. Nukleäre Substrate, wie der Upstream Stimulatory Factor 2 (USF2) konnten von der dTaspase1 nicht prozessiert werden. Eine durch genetische Manipulation eingebrachte NoLS in die dTaspase1 konnte die nukleoläre Lokalisation herstellen und ermöglichte in der Folge auch eine effektivere Spaltung nukleärer Substrate. Der wesentliche evolutionäre Unterschied zwischen Vertebraten und Invertebraten besteht für Taspase1 daher in einem Mechanismus, der über geregelten intrazellulären Transport sowie über die dynamische Regulation spezifischer Proteininteraktion in zur Spaltung unterschiedlicher zellulärer Substrate nötigen Lokalisationswechsel resultiert.

Fly versus man: evolutionary impairment of nucleolar targeting affects the degradome of *Drosophila's* Taspase1

Désirée Wunsch,^{*,1} Angelina Hahlbrock,^{*} Christina Heiselmayer,[†] Sandra Bäcker,[†] Patrick Heun,[‡] Dorothee Goesswein,^{*} Walter Stöcker,[§] Tanja Schirmeister,[¶] Günter Schneider,^{||} Oliver H. Krämer,[#] Shirley K. Knauer,[†] and Roland H. Stauber^{*,1}

^{*}Molecular and Cellular Oncology, Ear, Nose and Throat Department, University Medical Center of Mainz, Mainz, Germany; [†]Institute for Molecular Biology, Centre for Medical Biotechnology, Mainz Scientific Screening Center UG & Co. KG, University of Duisburg-Essen, Essen, Germany; [‡]Max Planck Institute of Immunobiology and Epigenetics, Freiburg, Germany; [§]Cell and Matrix Biology, Institute for Zoology, and [¶]Institute for Pharmacy and Biochemistry, University of Mainz, Mainz, Germany; ^{||}II. Medizinische Klinik, Technische Universität München, Munich, Germany; and [#]Department of Toxicology, University Medical Center, Mainz, Germany

ABSTRACT Human Taspase1 is essential for development and cancer by processing critical regulators, such as the mixed-lineage leukemia protein. Likewise, its ortholog, trithorax, is cleaved by *Drosophila* Taspase1 (dTaspase1), implementing a functional coevolution. To uncover novel mechanism regulating protease function, we performed a functional analysis of dTaspase1 and its comparison to the human ortholog. dTaspase1 contains an essential nucleophile threonine¹⁹⁵, catalyzing *cis* cleavage into its α - and β -subunits. A cell-based assay combined with alanine scanning mutagenesis demonstrated that the target cleavage motif for dTaspase1 (Q³[F/I/L/M]²D¹↓G¹X²X³) differs significantly from the human ortholog (Q³[F,I,L,V]²D¹↓G¹x²D³D⁴), predicting an enlarged degradome containing 70 substrates for *Drosophila*. In contrast to human Taspase1, dTaspase1 shows no discrete localization to the nucleus/nucleolus due to the lack of the importin- α /nucleophosmin1 interaction domain (NoLS) conserved in all vertebrates. Consequently, dTaspase1 interacts with neither the *Drosophila* nucleoplasmin-like protein nor human nucleophosmin1. The impact of localization on the protease's degradome was confirmed by demonstrating that dTaspase1 did not efficiently process nuclear substrates, such as upstream stimulatory factor 2. However, genetic introduction of the NoLS into dTaspase1 restored its nucleolar localization, nucleophosmin1 interaction, and efficient cleavage of nuclear substrates. We report that evolutionary functional divergence separating vertebrates from invertebrates can be achieved for proteases by a transport/localization-regulated mechanism.—Wunsch, D., Hahlbrock, A.,

Heiselmayer, C., Bäcker, S., Heun, P., Goesswein, D., Stöcker, W., Schirmeister, T., Schneider, G., Krämer, O. H., Knauer, S. K., Stauber, R. H. Fly versus man: evolutionary impairment of nucleolar targeting affects the degradome of *Drosophila's* Taspase1. *FASEB J.* 29, 1973–1985 (2015). www.fasebj.org

Key Words: cancer • development • leukemia • protease • threonine aspartase

PROTEOLYSIS IS A CRITICAL requirement for life. Approximately 700 proteases are present in the human genome, and proteases are key enzymes throughout evolution, bridging “simple” viruses with higher eukaryotes (1–3). Hence, proteases can be considered “molecular fossils” that serve to mark evolutionary milestones in the history of life. Based on the analysis of complete genomes, it is assumed that ~2% of proteins in all kinds of organisms are proteases. The MEROPS database classifies proteolytic enzymes based on structural similarity or sequence features supporting (functional) homology (1, 3).

In higher organisms, proteases are key regulators in homeostasis and disease, involved in numerous (patho) physiologic processes including intra- and intercellular protein degradation, metabolism, tissue remodeling, apoptosis, cell proliferation, and migration (4, 5). Thus, these enzymes are not only accepted as important signaling molecules but also as key targets for therapeutic strategies. However, to effectively and specifically modulate proteases, a profound knowledge of their molecular mechanistic function, of their degradome, *i.e.*, the sum of their substrates,

Abbreviations: Ab, antibody; BioTasp, fluorescent biosensor containing Taspase1 cleavage site; CRMI, chromosome region maintenance 1 export receptor; dNLP, *Drosophila* nucleoplasmin-like protein; GFP, green/blue/red fluorescent protein; GI, GenInfo identifier; GST, glutathione S-transferase; MLL, mixed lineage leukemia;

(continued on next page)

¹ Correspondence: Molecular and Cellular Oncology/ENT/ University Medical School Mainz, Langenbeckstr. 1, D-55101 Mainz, Germany. E-mail: wunsch@uni-mainz.de (D.W.); rstauber@uni-mainz.de (R.H.S.)

doi: 10.1096/fj.14-262451

This article includes supplemental data. Please visit <http://www.fasebj.org> to obtain this information.

as well as of the involved pathways in health and disease, are important requirements.

Whereas much is known about key proteolytic enzymes such as matrix metalloproteinases (6) and the proteasome (7), our knowledge on the mechanistic function, *in vivo* regulation, and (patho)biologic impact of threonine aspartase 1 (Taspase1) in humans and particularly in other organisms is still incomplete.

Human Taspase1, a protease composed of 420 amino acids, is responsible for the activation and stabilization of the mixed-lineage leukemia protein (MLL) and oncogenic MLL fusions at conserved ($Q^3[F,I,L,V]^2D^1\downarrow G^1x^2D^3D^4$) sites (8, 9). It belongs to a family of enzymes possessing an asparaginase-2 homology domain. In contrast to the other exclusively *cis* active type 2 asparaginases, only Taspase1 is able to cleave other substrates in *trans* (9). Thus, the discovery of Taspase1 founded a new class of endopeptidases that use the N-terminal threonine of its mature β -subunit as the active site. In humans, this N-terminal threonine (Thr²³⁴) is generated by autoproteolysis of the Taspase1 proenzyme (referred to as *cis* cleavage) into the 2 subunits α and β , which assemble into an $\alpha\beta$ -monomer representing the active protease (10, 11). Human Taspase1 is accumulating at the nucleus/nucleolus but is exploiting the dynamic importin- α (IMP- α)/nucleophosmin (NPM1) axis, allowing the processing of substrates in both the nucleus and the cytoplasm (12). As such, human Taspase1 is one of the few examples for which protease signaling is combined with active nucleo-cytoplasmic transport allowing an additional level of regulatory complexity (12).

Prediction of Taspase1's degradome (13) together with recent experimental evidence further implies that the protease plays not only a crucial role in leukemia, but is also important in solid tumors (9, 12, 14). Besides MLL fusions, other nuclear and cytoplasmic regulatory proteins, such as the precursor of the transcription factor IIA (TFIIA) or the upstream stimulating factor 2 (USF2), are *bona fide* Taspase1 targets (13, 15). Taspase1's catalytic activity is not affected by common protease inhibitors. Currently, neither effective synthetic small molecule nor natural inhibitors are available for this enzyme, counteracting not only to further dissect Taspase1's (patho)biologic functions *in vivo*, but also precluding the assessment of its clinical and therapeutic relevance (9, 13, 16–19).

Although MEROPS classification provides a context to decipher peptidase function, the wide diversity of proteolytic enzymes calls for investigation of how these proteins and their degradomes were shaped over time. It is expected that changes in the molecular structure have accompanied the evolution of proteolytic enzymes from more primitive organisms, serving relatively primitive functions, to the diverse and more complex functions that they fulfill in man. Hence, uncovering such structural

adaptions may deepen our understanding of the full repertoire of proteases' functional regulation.

The fruit fly *Drosophila melanogaster* has been extensively used as an important model organism (20). Many basic biologic functions and pathways are conserved between mammals and *Drosophila*, and nearly 75% of human disease-causing genes are believed to have a functional homolog in the fly (20), impressively underlined by the elucidation of the Ras signaling cascade (21).

After demonstrating that human Taspase1 is responsible for cleavage and stabilization of MLL proteins (9, 12), it was shown that the MLL homolog in *Drosophila*, called Trithorax (Trx), is proteolytically cleaved by the *Drosophila* Taspase1 homolog (dTaspase1) (22). Hence, a tight co-evolution of molecular mechanisms regulating Taspase1's proteolytic activity among these taxa was postulated (9, 22). However, because comprehensive experimental data supporting this hypothesis have been missing thus far, we performed the first detailed *in vivo* characterization and functional analysis of *Drosophila*'s Taspase1.

Unexpectedly, our study not only provides novel insights into the evolutionary development of Taspase1 “from fly to man,” but uncovers for the first time that evolutionary functional divergence separating vertebrates from invertebrates can be achieved for proteases by a transport/localization-regulated mechanism.

MATERIALS AND METHODS

Antibodies, reagents, and compounds

The antibodies (Abs) used were as follows: α -glyceraldehyde-3-phosphate dehydrogenase (GAPDH, sc-47724; Santa Cruz Biotechnology, Dallas, TX, USA); α -green fluorescent protein (GFP, sc-8334; Santa Cruz Biotechnology); α -glutathione S-transferase (GST, sc-57753; Santa Cruz Biotechnology); α -NPM1 (#3542; Cell Signaling Technology, Danvers, MA, USA); α -myc-tag (#2276; Cell Signaling Technology); α -Taspase1 (directed against the C terminus, AP1330b; BioCat GmbH); and α -hemagglutinin (HA) (MMS-101R; Covance). Appropriate horseradish peroxidase-, Cy3-, or FITC-conjugated secondary antibodies (Sigma-Aldrich, St. Louis, MO, USA, and Santa Cruz Biotechnology) were used. Reagents were from Sigma-Aldrich unless stated otherwise. Cells were treated with the export inhibitor Leptomycin B (LMB; 10 nM) as previously described (23).

Cell culture, microinjection, microscopy, and fluorescence imaging of cells

Mammalian and *Drosophila* Schneider 2 (S2) cell lines used in the study were maintained and transfected as previously described (13, 24, 25). Observation, quantitation, image analysis, and presentation were performed as previously described (12, 25, 26). Cell nuclei were stained with 0.5 μ g/ml Hoechst 33342 (excitation: 405 nm; emission: 430–480 nm).

Confocal laser scanning microscopy

Confocal laser scanning microscopy images were acquired using Leica LAS AF software on a Leica SP5 II system equipped with 4 lasers and a HCX-PL APO 63x/1.20 water lens (Leica, Buffalo Grove, IL, USA) as described previously in detail (27, 28). To determine the average intracellular protein localization, cells from 3 separate images were examined in 3 independent

(continued from previous page)

GAPDH, glyceraldehyde-3-phosphate dehydrogenase; Imp- α , Importin- α ; LMB, Leptomycin B; NES, nuclear export signal; NLS, nuclear import signal; NoLS, nuclear import/nucleolar localization signal; NPM1, nucleophosmin1; SEMP, small ubiquitin-like modifier (SUMO)-specific protease; TFIIA, transcription factor 2A; Trx, trithorax; USF2, upstream stimulatory factor 2; WT, wild-type

experiments, and representative images are shown. The percentage of cells exhibiting cytoplasmic and nuclear (C/N; cytoplasmic/nuclear signal > 80% of the total cellular signal) or nuclear/nucleolar (N/Ncl; nuclear/nucleolar signal > 80% of the total cellular signal) fluorescence was determined by counting ≥ 50 cells.

Plasmids

Drosophila Taspase1 (dTaspase1) coding sequence was cloned from the *Drosophila* cell line S2 as previously described (12). Cloning of the dTaspase1 coding sequence into expression vectors pc3-GFP, pc3-mCherry, pc3-HA, and the *Drosophila* expression vector pAc5.1V5/HisA (Invitrogen/Life Technologies, Carlsbad, CA, USA) using *KpnI/NheI* or *KpnI/EcoRI* restriction sites (primers 5'-TTTGGTACCATGGCAGGCTTCGTGCGAGTG-3' and 5'-TTTGCTAGCAATGCGCAGGTGGAAGTTGTGG-3'), respectively, allowed the expression of dTaspase1, alone or as a fusion with fluorescent proteins, as previously described (29, 30). Plasmid p_{dTasp}TV-GFP encoding the catalytically inactive Taspase1 mutant, dTaspT195V, was generated by splice overlap extension PCR as previously reported (31, 32). To generate the N-terminal myc-tagged dTaspase1, a construct myc-tag coding sequence flanked by *KpnI* restriction sites was produced by oligo annealing (30) and cloned into pC3-dTasp-GFP and pAc5.1dTaspase1-GFP. The human NPM1 coding sequence was amplified using pC3-NPM1-GFP and cloned into pAc5.1 vector using *KpnI/NheI* restriction sites. Expression constructs encoding Taspase1 (hTasp), hTaspIV, and NPM1 as fusions with green, blue, or red fluorescent proteins or HA-tag were described (12, 13). The USF2 coding sequence was derived by PCR from plasmids that were previously described (33, 34). Generation of plasmids p_{NLS}-GFP/GST-CS2-NES_{REV} (*BioTasp*), and biosensors containing cleavage sites of TFIIA, USF2, and other described Taspase1 substrates as described previously (12). p_{BioTasp}_R encodes a red fluorescent biosensor (NI S-mCherry/GST-CS2-NES_{REV}), in which GFP was replaced by mCherry (31). The biosensor assay was adapted for the expression in *Drosophila* cell line S2R+ by cloning NLS-GFP/GST-CS2-NES_{REV} into pAc5.1V5/HisA via blunted *KpnI/EcoRI* restriction sites. Expression plasmids for BioTasp variants, in which CS2 was mutated, were previously described (12). pF143 plasmid expressing GFP alone was used as a control (35). pIB_{dNLP}-mCherry encoding the NMP-like protein from *Drosophila* was previously described (25). For cloning of the dTaspase1-NoLS mutant, the amino acid sequence ₁₅₅NHYKSK-EESV₁₇₈ was mutated into ₁₅₅KRNKRK-KKRR₁₇₈ by splice overlap extension PCR as previously reported (31, 32). Plasmids were verified by sequence analysis as previously described (36).

Protein extraction, immunoblot analysis, and immunofluorescence

Preparation of whole cell lysates was carried out as described using a physiologic lysis buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, 1 mM DTT, 1 mM PMSF, and Complete Protease Inhibitor Cocktail EDTA-free; Roche Diagnostics, Mannheim, Germany) (37). Equal loading of lysates was controlled by reprobing blots for GAPDH as previously described (38). Immunofluorescence was performed as reported previously in detail (12, 23).

Immunoprecipitation

Immunoprecipitation of GFP-tagged proteins was performed using α -GFP magnetic beads and μ -MACS columns (Miltenyi Biotec) according to the supplier's recommendation. Briefly, whole cell lysates were incubated with 50 μ l α -GFP magnetic

beads for 30 minutes on ice. μ -MACS columns were placed into magnetic field and equilibrated with 200 μ l lysis buffer. Lysates with magnetic beads were applied onto the columns and washed 4 times with 200 μ l lysis buffer and once with 100 μ l wash buffer-2 (Miltenyi Biotec). After complete flow through, 20 μ l preheated elution buffer (Miltenyi Biotec) was applied onto each column and incubated for 5 minutes. To complete protein elution, 50 μ l preheated elution buffer was applied. Thirty microliters of the eluate and 3% of input were analyzed.

Bioinformatics and statistical analysis

dTaspase1 protein was analyzed using PredictNLS algorithm (<https://www.predictprotein.org>) and cNLS Mapper (39, 40). For protein alignment of Taspase1 protein sequences, the program GeneDoc (www.psc.edu/biomed/genedoc) was used, and protein sequences were obtained from National Center for Biotechnology Information databases [*Homo sapiens* (GenInfo identifier (GI): 19263670); *Macaca mulatta* (GI: 109092934); *Mus musculus* (GI: 123233857); *Rattus norvegicus* (GI: 112982912); *Bos taurus* (GI: 164448634); *Canis familiaris* (GI: 73991167); *Oryctolagus cuniculus* (GI: 291389018); *Sus scrofa* (GI: 350594655); *Danio rerio* (GI: 169636455); *Xenopus laevis* (GI: 148227000); *Gallus gallus* (GI: 61098282); *Ciona intestinalis* (GI: 198426192); *Drosophila melanogaster* (GI: 23093366); *Amphimedon queenslandica* (GI: 340375835)]. For the *in silico* identification of potential *Drosophila* Taspase1 targets, ScanProsite searches (<http://ca.expasy.org/tools/scanprosite>) were performed using the pattern Q-[FILM]-D-G as the query. Searches were performed in the *Drosophila* taxon of the UniProtKB/SwissProt database, and all other parameters were used at default settings. For experiments stating *P* values, a paired Student *t* test was performed. Unless stated otherwise, *P* values represent data obtained from 3 independent experiments done in triplicate. *P* < 0.05 was considered significant.

RESULTS

Evolutionary conservation: Threonine¹⁹⁵ is the essential nucleophile for the *cis* and *trans* cleavage activity of *Drosophila* Taspase1

The function of human Taspase1 critically depends on the active nucleophile threonine²³⁴, essential not only for the protease's self-activating *cis* but also *trans* cleavage activity (9, 12). Sequence alignments of different Taspase1 homologs (Fig. 1A) indicate that activation occurs most likely by autocatalytic cleavage between aspartate¹⁹⁴ and threonine¹⁹⁵ for dTaspase1 (dTasp: ¹⁹³LD↓TVGAV¹⁹⁹). In fact, whereas protein alignment of human and *Drosophila* Taspase1 revealed only 36% overall amino acid identity (Supplemental Fig. S1A), the region required for their *cis* cleavage is fully conserved. *cis* cleavage would result in an α -subunit of about 20 kDa and a β -subunit of 18 kDa, thereby generating an N-terminal threonine at the β -subunit as the key nucleophile. To verify this prediction, we cloned dTaspase1 from *Drosophila* cells, and the predicted sequence could be verified (22). On expression as a GFP-tagged or myc-dTaspase1-GFP protein in *Drosophila* or human cells, dTaspase1 was detectable by immunoblot as a protein of 44 kDa, representing the β -subunit (18 kDa) in fusion with GFP (27 kDa) and the α -subunit (20 kDa)

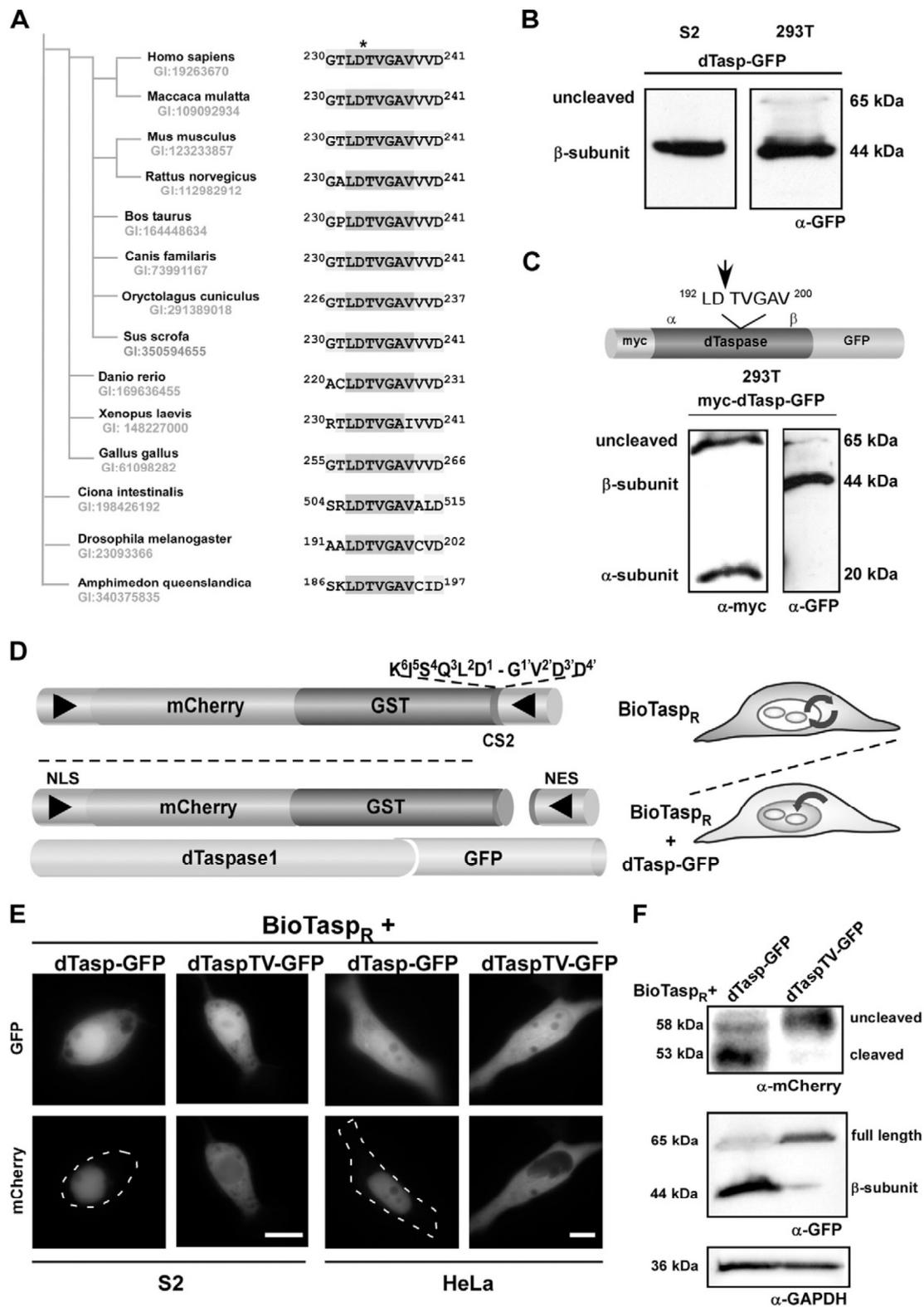


Figure 1. Threonine¹⁹⁵ is the essential nucleophile for the *cis* and *trans* cleavage activity of *Drosophila* Taspase1. **A)** Alignment of the predicted *cis* cleavage sequences in Taspase1 homologs. Phylogram was constructed on the basis of amino acid sequence similarities illustrating the evolutionary relationships among Taspase1 proteins from different species (TreeDomViewer, <http://www.bioinformatics.nl/tools/treedom/>). Sequence alignment demonstrates that the predicted active nucleophile threonine (gray) is conserved in all Taspase1 homologs. *cis* cleavage site is marked by asterisk. Organisms, GI, and amino acid positions are indicated. **B, C)** Verification of the predicted dTaspase1 *cis* cleavage site. *Drosophila* S2 or human 293T cells were transfected with GFP (**B**) or an N-terminal myc-tagged dTaspase1 (**C**) expression plasmid, and expression was analyzed in whole-cell lysates by immunoblot. (continued on next page)

(Fig. 1B, C). Hence, *cis* cleavage of the proenzyme occurs in both *Drosophila* and human cells. Of note, the unprocessed protease with an expected size of 65 kDa was hardly detectable in the more physiologic environment of *Drosophila* cells (Fig. 1B). In contrast, we detected the precursor in human 293T cells, where proteolysis does not occur to 100% (Fig. 1C). Also, detection of the full-length protein was more efficient using an α -myc Ab in contrast to using the α -GFP Ab. These variations might be because of differences in the accessibility of the antibodies to the uncleaved/cleaved protein, antibody specificity, and unequal stabilities of the different dTaspase1 cleavage products.

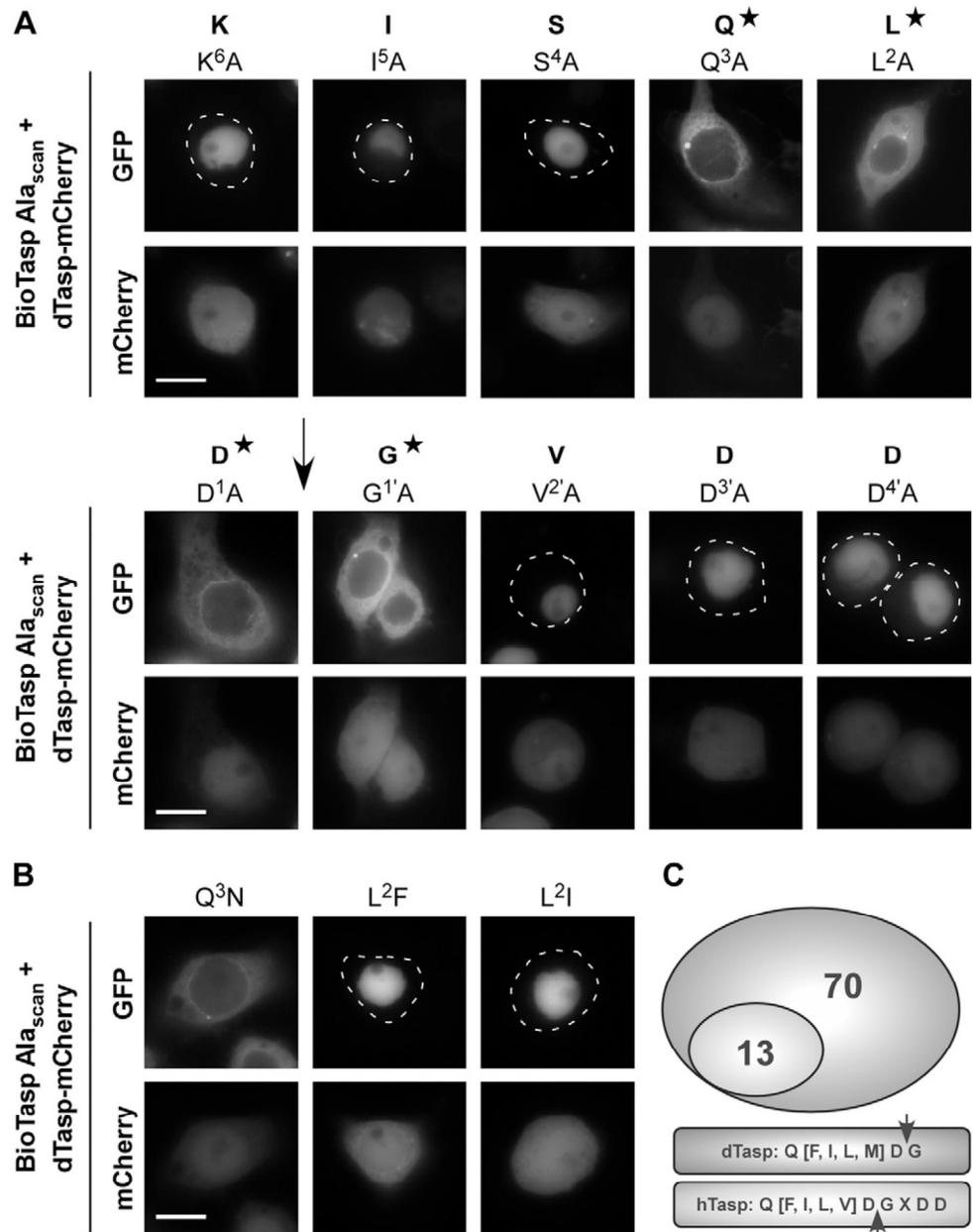
To experimentally demonstrate that the presumed nucleophile is indeed required for dTaspase1's proteolytic activity, Thr¹⁹⁵ was mutated into Val, in analogy to the catalytically inactive human Taspase1^{T134V} mutant (10, 13). To investigate protease function in living cells, we first tested whether our translocation biosensor assay could be adapted to *Drosophila* cells (8, 13) (Fig. 1D). The fluorescent biosensor BioTaspR protein is composed of GST, mCherry, combinations of a nuclear import (NLS) and an export (NES) signal, and a Taspase1 cleavage site. BioTaspR localizes predominantly to the cytoplasm, but is continuously shuttling between the nucleus and the cytoplasm. Coexpression of BioTaspR with biologically active dTaspase1-GFP results in the proteolytic cleavage of the NES, causing nuclear accumulation of the red fluorescent sensor triggered by the NLS. As shown in Fig. 1D, E, expression of the biosensor BioTaspR in the context of a *Drosophila*-specific vector was followed by its cleavage and a cytoplasmic to nuclear accumulation on coexpression of tagged- or untagged biologically active wild-type (WT) dTaspase1. In contrast, the inactive dTaspTV¹⁹⁵ mutant showed no proteolytic activity (Fig. 1E), neither in *Drosophila* nor human cells, supporting a conserved proenzyme-to-active-protease coevolution. We independently verified the cleavage activity of HA-tagged dTaspase1 variants against BioTaspR by immunoblot analysis (Fig. 1F). Of note, we observed that the expression of the dTaspTV¹⁹⁵-mutant is lower compared with the WT enzyme. Because the expression plasmids are under the control of the same regulatory elements, these differences might be caused by reduced protein stability of the uncleaved protease. However, the molecular details need to be confirmed in follow-up studies.

Evolutionary protease divergence I: Taspase1's degradome is enlarged in *Drosophila*

As the MLL ortholog in *Drosophila*, trithorax (Trx), is cleaved by dTaspase1, a functional coevolution of these type 2 asparaginases is currently expected by the field. To experimentally test this hypothesis, we first performed an *in vivo* mapping of residues critical for the *trans* cleavage activity of dTaspase1 in the appropriate biologic environment, *i.e.*, *Drosophila* cells. Using our translocation biosensor assay, which was efficiently processed in *Drosophila* S2 cells (Fig. 1E), alanine scanning mutagenesis of the cleavage site motif (amino acids KISQLD↓GVDD) was performed. As depicted in Fig. 2, coexpression of the indicated *BioTasp Ala_{scan}*-mutants with the catalytically active dTaspase1-mCherry fusion resulted in proteolytic cleavage and nuclear accumulation of only those biosensor proteins, in which we mutated nonessential residues in the cleavage site. In contrast, mutation of critical residues completely prevented cleavage and nuclear accumulation of the fluorescent sensor proteins in *Drosophila* cells (Fig. 2; data not shown). We observed cytoplasmic localization of the biosensor for Gln³, Leu², Asp¹, and Gly^{1'}, unambiguously identifying those residues as essential for proteolysis. Interestingly, replacing Leu² by the chemically similar amino acid Phe or Ile restored protease recognition and cleavage, whereas substitution of Gln at P3 with the connatural amino acid Asn did not allow cleavage (Fig. 2B). By cotransfection of the inactive dTaspTV¹⁹⁵-mCherry mutant, we verified assay specificity, which did not result in nuclear accumulation of all biosensor variants (data not shown). Also, nuclear accumulation of all variants on treatment with the export inhibitor LMB further excluded the formal possibility that mutagenesis affected the shuttling capability of the biosensors in *Drosophila* cells (data not shown). Of note, we obtained similar results in human HeLa cells, indicating that cleavage site specificity is inherently controlled by the enzyme itself and not by additional organism-specific cofactors (data not shown). Collectively, our analysis identified the following *trans* cleavage site consensus sequence for dTaspase1: Q³[F/I/L/M]²D¹↓G^{1'}x^{2'}x^{3'}. Hence, compared with the consensus sequence defined for human Taspase1 (Q³[F,I,L,V]²D¹↓G^{1'}x^{2'}x^{3'}D^{4'}/D^{3'}x^{4'}) (8, 13) recognition and productive *trans* cleavage activity of *Drosophila* Taspase1 lacks specificity for Asp in the P3' and P4' positions and thus allows more flexibility in substrate recognition (Fig. 2C).

In both cell lines, a band of 44 kDa could be detected representing the β -subunit (18 kDa) fused with GFP (27 kDa). The α -subunit produced by autocatalytic cleavage of full-length dTaspase1 is detectable at 20 kDa. Proteins were detected using α -GFP or α -myc Ab. Molecular weight of proteins is indicated. D) Principle and domain organization of the translocation biosensor assay, BioTaspR, to probe dTaspase1 activity in living cells. For further information see Results section. →, NLS; ←, NES. E) WT dTaspase1 (dTasp), but not the catalytically inactive dTasp^{T195V} mutant (dTaspTV), is able to cleave BioTaspR in cells. The red fluorescent sensor is predominantly cytoplasmic, whereas coexpression of biologically active dTasp-GFP, but not of inactive dTaspTV-GFP, results in the proteolytic cleavage of the NES and nuclear accumulation of BioTaspR in *Drosophila* S2 or HeLa cells. Scale bars, 5 μ m. Dashed lines mark cytoplasmic cell boundaries obtained from the corresponding phase contrast images. F) Proteolytic cleavage of BioTaspR shown by immunoblot analysis of whole-cell lysates. 293T cells were transfected with BioTaspR together with the indicated Taspase1 expression plasmids (HA-tag about 1 kDa). A 2-fold higher amount of dTaspTV expression plasmid was transfected to obtain equal protein expression levels. In contrast to dTaspase1, the dTaspTV mutant showed neither *cis*- nor *trans* cleavage activity. Expression of proteins and cleavage products in cell lysates was visualized using α -GST and α -HA Ab. GAPDH served as loading control.

Figure 2. Taspase1's degradome is enlarged in *Drosophila*. A, B) *In vivo* characterization of residues critical for the *trans* cleavage activity of dTaspase1. *Drosophila* S2 cells were transfected with the indicated cleavage site mutants (BioTasp Ala_{scan}) together with the active dTaspase1-mCherry expression plasmid. Localization of the BioTasp Ala_{scan} variants was analyzed 48 hours later in ≥ 200 fluorescent living cells. Representative images are shown. Scale bars, 5 μm . A) Analysis of transfectants revealed that mutagenesis of Gln, Lys, Asp, and Gly (P3-P1') inhibits cleavage of the BioTasp sensor. *Residues essential for dTaspase1 cleavage; ↓, position of dTaspase1 *trans* cleavage; B) Gln at P3 is critical, whereas exchange of Leu² with the unpolar aa Phe or Ile restored cleavage activity. C) Bioinformatical prediction of potential *Drosophila* targets containing respective cleavage sites. Although the *in silico* analysis applying the previously identified human Taspase1 consensus sequence resulted in only 13 predicted proteins (gray), the search using the identified specific cleavage site motif of dTaspase1 (red) suggested 70 substrates. Consensus sequences used for the *in silico* predictions are indicated, position of cleavage marked by an arrow.



This finding motivated us to next perform an *in silico* identification of dTaspase1's degradome in *Drosophila*. Scanning of the UniProt/Swiss-Prot database resulted in the prediction of 70 potential dTaspase1 substrates in contrast to the 27 targets predicted for human Taspase1 (8, 13) (Fig. 2C; Table 1; Supplemental Table S1). These comprise not only proteins specifically expressed in *Drosophila*, but also homologs already verified as *bona fide* targets for the human protease, such as the transcription factor TFIIA or the proteasomal regulatory subunit 4. As summarized in Table 1 and Supplemental Table S1, bioinformatic analyses of the predicted dTaspase1 substrates indicated various molecular functions, as well as diverse intracellular localization patterns of these targets. Notably, using the consensus sequence motif found for human Taspase1 (8, 13) resulted in a prediction of only 13 candidates in *Drosophila* (Fig. 2C). Thus, one would have

missed 80% of the potential *Drosophila* degradome, underlining the need to perform a species-specific mapping of proteases' degradomes in general.

However, *in silico* analyses and predictions of a protease's degradome obtained by systematic biochemical methods, such as proteomics, have to be confirmed in detail for each of the proposed targets. Although it is beyond the scope of the current study to analyze all of the predicted 70 targets of dTaspase1, we investigated whether Asp at positions 3' and 4' is indeed not essential for cleavage by dTaspase1 by testing additional targets. Based on our prediction (Supplemental Table S1), we analyzed *Drosophila* histone-binding protein Caf1 and cyclin-dependent kinase 8, both lacking an Asp at positions 3' and 4', in our biosensor system. As shown in Supplemental Fig. S2F, coexpression of dTaspase1, but not of hTaspase1-mCherry, with the biosensor constructs

TABLE 1. Characteristics of selected dTaspase1 targets predicted by ScanProsite

Gene locus	Swiss Prot	Gene name	Alias	Length (amino acids)	Sequence	GO: biologic process	
						ID	Description
TRX	P20659	Histone-lysine N-methyltransferase trithorax	KMT2A Trx-g Rg-bx	3726	2232–2283 QMDGvDD	0006350 0006355 0016571	Transcription, regulation of transcription, histone methylation
TF2AA	P52654	Transcription initiation factor IIA subunit 1	TFIIA TFIIA-L	366	257–263 QLDGaD	0006350 0006355	Transcription, regulation of transcription
PRS4	P48601	26S protease regulatory subunit 4	Pros26.4 P26s4 S4 P56	439	314–319 QLDGfD	0000022 0006508 0006511 0008283	Mitotic spindle elongation, proteolysis, ubiquitin-dependent protein catabolic process, cell proliferation
XPC	Q24595	DNA repair protein complementing XP-C cells homolog	XPCDM mus210 Xpcc	1293	886–892 QLDGgDD	0006281 0006289 0006974	DNA repair, nucleotide-excision repair, response to DNA damage stimulus
RG190	Q9VX32	Rho GTPase-activating protein 190	RhoGAPp190 D-p190 RhoGAP	1561	1157–1162 QLDGeD	0007165 0043547 0050770	Signal transduction, positive regulation of gtpase activity, regulation of axonogenesis
Hcf	Q9V4C8	Host cell factor	dHCF CG1710	1500	1011–1017 QLDGagD	0007049 0035065 0045893 0045927	Cell cycle, regulation of histone acetylation, positive regulation of transcription DNA-dependent, positive regulation of growth
JHD1	Q9VHH9	JmjC domain-containing histone demethylation protein 1	JHDM1 CG11033	1345	774–780 QLDGteD	0006350 0006355 0016568 0070544	Transcription, regulation of transcription DNA-dependent, chromatin modification, histone H3K36 demethylation
INSL3	Q9VT52	Probable insulin-like peptide 3	ILP3 CG14167	120	65–71 QIDGfeD	0008286 0005158 0060180	Insulin receptor signaling pathway Insulin receptor binding Female mating behavior
TRXR2	Q9VNT5	Thioredoxin reductase 2, mitochondrial	CG11401	516	296–301 QMDGsD	0045454 0055114	Cell redox homeostasis, oxidation-reduction process
DPN	Q26263	Protein deadpan	44C bHLHe50 CG8704	435	133–138 QMDGiD	0006351 0007399 0007530 0030154 0048813	Transcription DNA-dependent, nervous system development, sex determination, cell differentiation, dendrite morphogenesis
MICAL	Q86BA1	Protein-methionine sulfoxide oxidase Mical	CG33208	4723	1742–1748 QMDGaD	0030042 0007411 0070995 0019417	Actin filament depolymerization Axon guidance NADPH oxidation Sulfur oxidation

Potential targets were identified using the consensus sequence: Q³[F/I/L/M]²D¹↓G¹X²X³, Q³[F/I/L/M]²D¹↓G¹X²X³D⁴ and Q³[F/I/L/M]²D¹↓G¹X²D³X⁴ (for complete list, see Supplemental Table S1). Gene loci, SwissProt accession numbers, gene names, and aliases in accordance with the HUGO Gene Nomenclature Committee are given together with the amino acid sequence containing the predicted dTaspase1 cleavage site (consensus marked in bold). Gene ontology (GO) identifiers and descriptions are listed.

containing the cleavage site of histone-binding protein Caf1 (H³Y⁴) or cyclin-dependent kinase 8 (G³L⁴), resulted in their complete cleavage. As an additional example, we investigated the effects of mutating the MLL-CS2 from D³D⁴ to A³A⁴ (Supplemental Fig. S2G). As predicted from our data, we found that *Drosophila* but not human Taspase1 was still proteolytically active. Thus, these results support our finding of an extended cleavage site motif for dTaspase1 lacking amino acid specificity for positions 3' and 4'.

Evolutionary protease divergence II: Active nuclear transport and localization of dTaspase1 differ from its human relative

To investigate additional evolutionary divergence of the Taspase1 homologs, we analyzed in detail their intracellular localization and transport. Whereas protein alignment revealed high similarities in the region required for Taspase1's *cis* cleavage, we noticed that the sequence we previously identified as a functionally important

nuclear/nucleolar localization signal (NoLS) in human Taspase1 (12) is absent in *Drosophila* (Fig. 3A). Moreover, whereas the NoLS is highly conserved in all known vertebrate Taspase1 proteins, the sequence is missing in invertebrates, such as *Drosophila*. Applying 2 independent NoLS prediction tools did not identify other potential NoLS (data not shown).

The *in vivo* relevance of this finding became apparent when we studied the subcellular localization of dTaspase1 in living cells. In contrast to the nuclear/nucleolar localization of the human protease, ectopic expression of mCherry-, GFP-, or myc-tagged dTaspase1 in *Drosophila* S2 or human cell lines rather showed a predominantly cytoplasmic and only partial nuclear localization (Fig. 3B, C; Supplemental Fig. S2E; data not shown). Notably, quantitation of intracellular protein localization in ≥ 200 fluorescent cells confirmed that dTaspase1 was never detectable at the nucleoli (Fig. 3C). Treatment of cells with the export inhibitor

LMB did not induce nuclear accumulation dTaspase1 (Supplemental Fig. S2A), demonstrating that the localization of dTaspase1 is the result of passive diffusion rather than regulated by active transport signals, such as NLS or NES. Also, treatment with Actinomycin D, resulting in dissolution of the nucleolus, did not affect the intracellular distribution of dTaspase1 (data not shown).

Because localization and function of human Taspase1 is regulated by an Imp- α /NPM1 switch (12), we subsequently examined the interaction of dTaspase1 not only with human NPM1, but also with its *Drosophila* homolog dNLP (25). The dNLP protein is a member of the nucleophosmin/nucleoplamin family and shares $\sim 30\%$ sequence similarity with its human homologs (25) (Supplemental Fig. S1B). We recently showed that dNLP plays an important role in centromere clustering and nucleolar anchoring (25). On coexpression of dTaspase1 with dNLP-mCherry or with human NPM1-GFP, no detectable

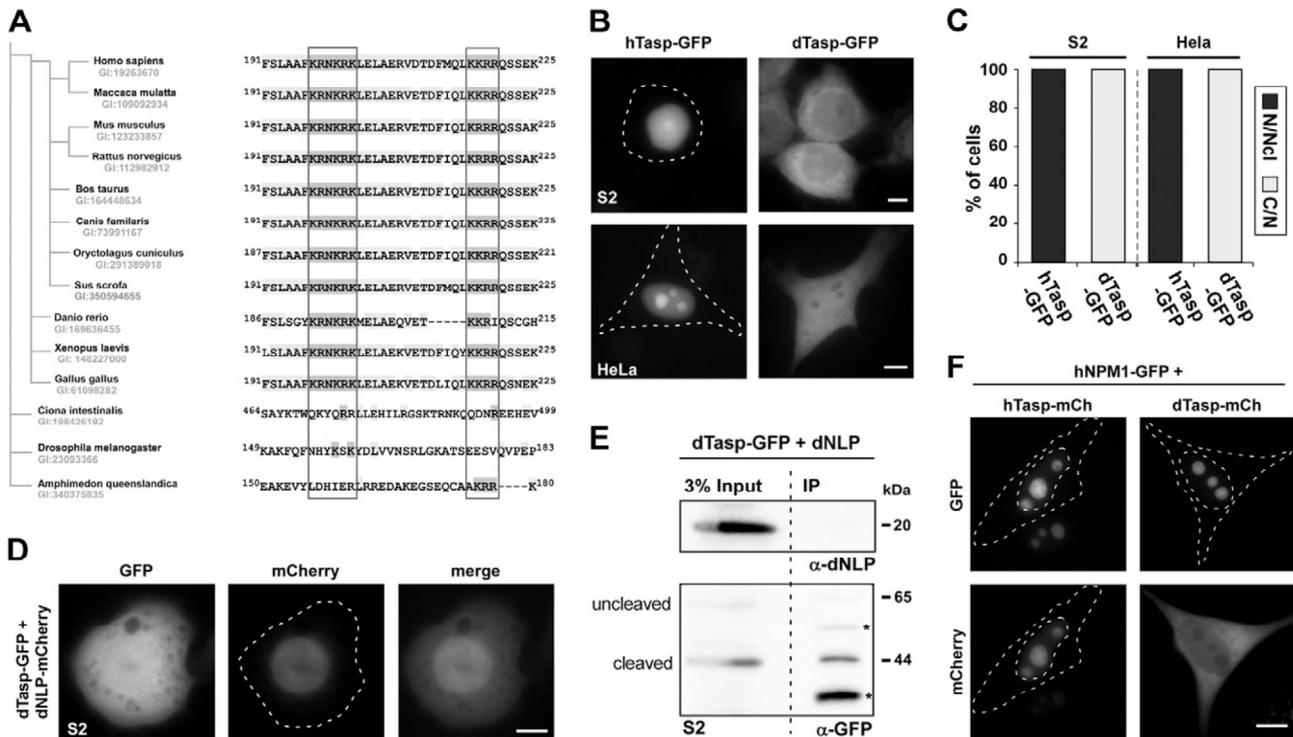


Figure 3. Active nuclear transport and localization of dTaspase1 differ from its human relative. **A**) NoLS of Taspase1 enzymes is conserved in vertebrates but lacking in invertebrates. Alignment of NoLS (boxed) in different Taspase1 homologs. Phylogram constructed on the basis of amino acid sequence homology depicting the evolutionary relationships between Taspase1 proteins of different species. Organisms, GenInfo identifier (GI), and amino acid positions are indicated. **B**) Subcellular localization of dTaspase1 differs from hTaspase1 because of the lack of NoLS. Ectopic expression of hTaspase1- and dTaspase1-GFP in the *Drosophila* S2 cell line (upper) or HeLa cells (lower). In contrast to nuclear/nucleolar localization of hTaspase1, dTaspase1 localized to the cytoplasm and nucleus. **C**) Quantification of protein localization in S2 and HeLa cells confirmed that in contrast to hTaspase1, dTaspase1 was never detectable at the nucleoli. Cells were either transfected with human or *Drosophila* Taspase1-GFP expression constructs and analyzed 24 h later. Nuclear/nucleolar (N/Ncl) or cytoplasmic/ nuclear (C/N) fluorescence was quantified in ≥ 200 transfected living cells by microscopic observation. **D–F**) dTaspase1 neither interacts with the dNLP nor with human NPM1. Scale bars: 5 μ m. Dashed lines mark nuclear/cytoplasmic cell boundaries obtained from the corresponding phase contrast images. **D**) Cotransfection of GFP-tagged dTaspase1 and dNLP-mCherry in S2 cell line revealed that both proteins did not change their subcellular localization upon coexpression. **E**) Coimmunoprecipitation assay confirms the lack of dTaspase1/dNLP protein interaction. Immunoprecipitation of S2 cell extracts cotransfected with the indicated expression constructs was carried out using α -GFP Ab-coated magnetic beads and μ -MACS columns (Materials and Methods). Proteins were identified by immunoblot using the indicated Abs. Input, total amount of cell lysate added to the immunoprecipitation reaction. *GFP-degradation products as reported (12). **F**) dTaspase1 does not colocalize at the nucleolus on overexpression of NPM1. Coexpression of human NPM1-GFP protein with human or *Drosophila* Taspase1-mCherry proteins in HeLa cells. Although hTaspase1 accumulates at the nucleoli, dTaspase1 localizes throughout the cell sparing the nucleolus.

changes in cellular distribution of dTaspase1, indicative of complex formation in living cells, could be observed (Fig. 3D, F; Supplemental Fig. S2B). We independently confirmed the lack of protein interaction biochemically by coimmunoprecipitation assays (Fig. 3E; Supplemental Fig. S2C). As a control, a strong interaction of human Taspase1 with NPM1 was verified under identical experimental conditions (Supplemental Fig. S2C).

Impairment of nuclear/nucleolar targeting affects the proteolytic activity of *Drosophila* Taspase1

These results further supported our hypothesis that *Drosophila*'s Taspase1 proteolytic function is differentially regulated and affected by transport processes. As the NoLS-mediated interaction with Imp- α and the binding to NPM1 promote and maintain rapid and efficient nuclear targeting of human Taspase1, the lack of both interactions prompted us to investigate dTaspase1's proteolytic activity. Based on our findings that the nuclear steady-state concentration of dTaspase1 is reduced compared with the human enzyme, we expected that dTaspase1 should be less active in processing nuclear substrates. As shown in Fig. 4A, we indeed found that dTaspase1 showed reduced proteolytic activity compared with the human homolog for the nuclear substrate USF2. In case our hypothesis was correct, we reasoned that restoration of nuclear import and NPM1 interaction in dTaspase1 should, at least partially, restore dTaspase1's ability to efficiently process nuclear substrates. Hence, we genetically introduced the importin/NPM1 interaction domain into dTaspase1 (Fig. 4B). Indeed, on ectopic expression in *Drosophila* and human cells, this humanized dTaspase1-NoLS protein not only localized predominantly to the nucleus but also revealed an efficient nucleolar targeting (Fig. 4B; Supplemental Fig. S2D). Moreover, we found dTaspase1-NoLS to be more active in processing the nuclear substrate USF2 compared with WT dTaspase1 lacking the efficient nuclear/nucleolar targeting sequence (Fig. 4C; USF forms are marked by asterisks). Collectively, these results strongly support our conclusion that in contrast to invertebrates, the Imp- α /NPM1 axis has evolved in vertebrates as a transport-based evolutionary mechanism to fine-tune the proteolytic cleavage activity of Taspase1 (Fig. 4D).

DISCUSSION

Besides their numerous physiologic functions, protease deregulation often contributes to disease development that could be treated by specific inhibitors (5, 41, 42). However, many protease inhibitors failed in the clinics for multiple reasons (5). One prerequisite to potentially improve clinical success is to deepen our holistic knowledge of a protease's mechanistic function to understand how its activity and degradome are involved in health and disease pathways.

Besides the structure of a protease's catalytic site and its chemical proteolytic mechanism, additional functional and structural domains contribute to the (fine) regulation of the enzyme's biologic activity and impact on cellular homeostasis and/or the survival advantage of tumors as

a micro-organ. In general, it is assumed that changes in the molecular structure have accompanied the evolution of proteolytic enzymes from more primitive organisms, serving relatively primitive functions, to the diverse and more complex functions that they fulfill in humans. Hence, uncovering such structural adaptations during evolution may deepen our understanding of the full repertoire of proteases' functional regulation allowing the development of more potent targeting strategies.

Sequences encoding Taspase1 enzymes have been conserved in vertebrates and insects but not in the worm *Caenorhabditis elegans* (9, 22). As both proteins, MLL and Trx, are developmental regulators in vertebrates and insects, these are also *bona fide* Taspase1 targets, albeit *Drosophila* Taspase1 seems to be more active on its cognate Trx substrate (22). Nevertheless these studies strongly suggested a close coevolution of these enzymes, including their degradomes, as well as their structure-activity relationships (9, 22).

Thus, it was indeed unexpected that when analyzing the function of *Drosophila* Taspase1, we uncovered significant differences in its localization and substrate recognition. For one, we not only verified that the predicted threonine¹⁹⁵ is the essential nucleophile in dTaspase1 for *cis* cleavage, but also found that the target cleavage motif for dTaspase1 ($Q^3[F/I/L/M]^2D^1\downarrow G^1x^2x^3$) is more flexible compared with the human ortholog ($Q^3[F,I,L,V]^2D^1\downarrow G^1x^2x^3D^4/D^3x^4$), implying an enlarged degradome in *Drosophila*. As dTaspase1 is an aspartase, Asp at the P1 position is mandatory for cleavage, and the unique chemical characteristics of Gly at P1 do not tolerate its replacement by Ala. Also, Gln at position P3 is required for substrate recognition, and substitution of this uncharged polar amino acid with the smaller hydrophobic residue Ala or the similar amino acid Asn completely blocks *trans* cleavage. Similar to the human enzyme, position P2 is also critical for function (10, 12). Although Met, present in the Trx cleavage site, can be substituted by hydrophobic residues of similar size (Phe, Ile, or Leu), other amino acids such as the smaller hydrophobic amino acid Ala are not tolerated. Hence, hydrophobicity in combination with size is required for productive cleavage. Whereas in both enzymes, position P2' is highly flexible, specific residues at positions P3' and P4' seem to be only required for human Taspase1 (10, 12), explaining why the human Taspase1 degradome is much more restricted. In contrast to previous predictions (9, 22), the identified cleavage motif for dTaspase1 strongly implies an enlarged degradome containing 70 potential substrates in *Drosophila* in contrast to the rather limited 27 targets predicted for humans (13). As knowledge of a protease's degradome is key to understanding its biologic relevance, our findings suggest that Taspase1 may modulate additional pathways in *Drosophila*. The obtained list of the 70 most likely Taspase1 targets in *Drosophila* not only contains known Taspase1 substrates such as Trx or the transcription factor TFIIA but also several candidates, which thus far have not been considered as potential targets for this protease (Supplemental Table S1). Moreover, our results also provide an explanation for previous findings. It was shown that the host cell factor is cleaved by *Drosophila* Taspase1 but not by the human ortholog (22). Based on the cleavage motif defined here, only the

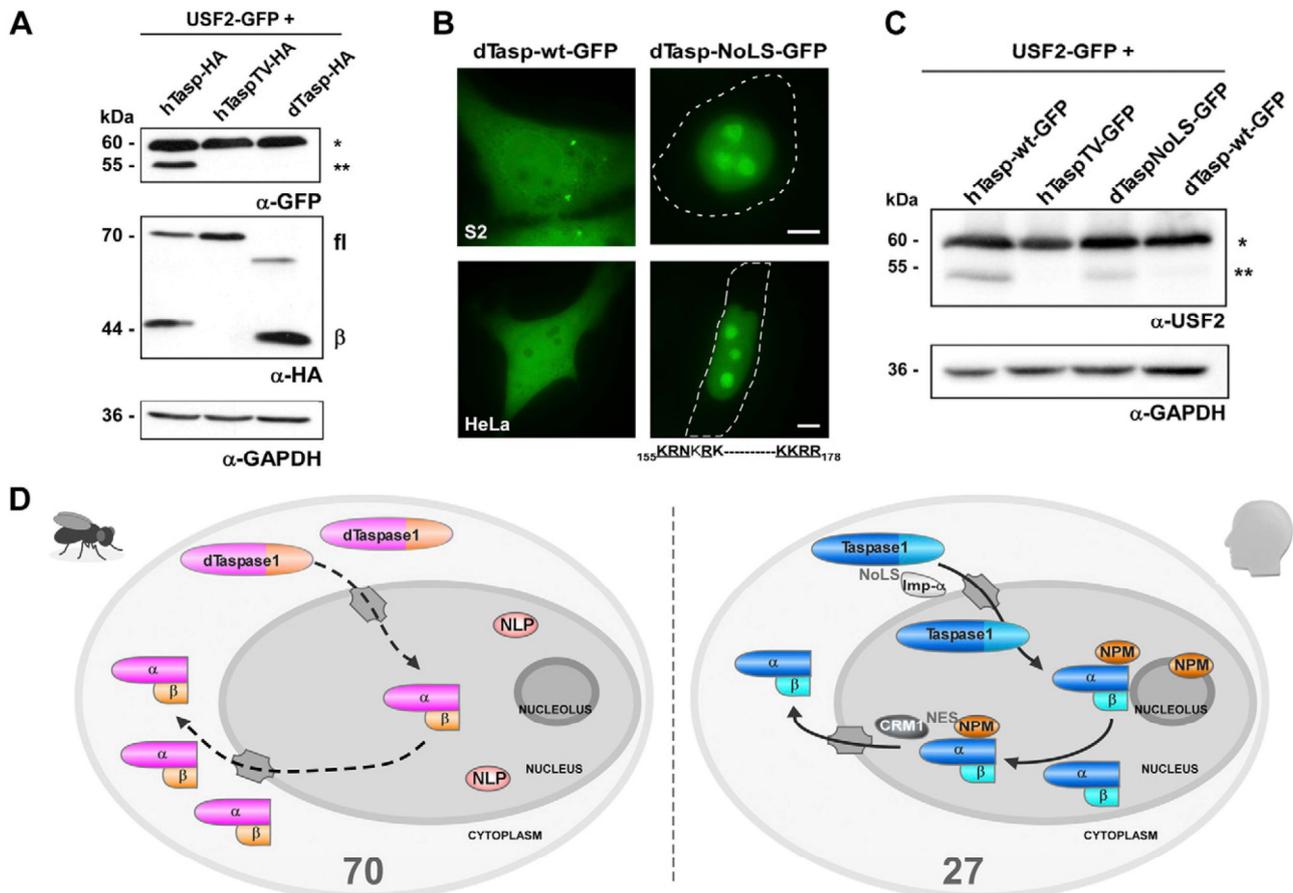


Figure 4. Impairment of nuclear/nucleolar targeting affects the proteolytic activity of *Drosophila* Taspase1. **A)** Compared with the human homolog, dTaspase1 is less active in cleaving the nuclear substrate USF2. Immunoblot demonstrates that USF2 was inefficiently processed by dTaspase1 compared with hTaspase1. Catalytically inactive hTaspase1^{T234V} mutant (hTaspTV) served as the negative control. 293T cells were transfected with USF2-GFP together with the indicated Taspase1 expression plasmids. Expression of proteins and cleavage products was visualized using α -GFP and -HA Ab. GAPDH served as loading control. *Full-length USF2-GFP protein; **USF2-GFP cleavage fragment; fl, Taspase1-GFP full-length protein; β , Taspase1 β -subunit GFP. Molecular weight of proteins is indicated. **B)** Introduction of the NoLS restores nucleolar targeting of dTaspase1. Localization in living S2 or HeLa transfectants was analyzed by fluorescence microscopy. Mutation of dTaspase1 amino acids into the NoLS is indicated; critical amino acids are underlined. Scale bars, 5 μ m. **C)** Nuclear/nucleolar targeting of dTaspase1 increases its proteolytic activity against the nuclear substrate USF2. Immunoblot analysis demonstrates that dTaspase1-NoLS was more active in processing USF2 compared with WT dTaspase1. Cells were transfected with USF2-GFP together with the indicated Taspase1 expression plasmids. Expression of proteins and cleavage products was visualized using α -GFP and -HA Ab. GAPDH served as loading control. *Uncleaved USF2-GFP protein; **USF2-GFP cleavage products; fl, Taspase1-GFP full-length protein; β , Taspase1 β -subunit GFP. **D)** Model illustrating how the importin- α /NPM1 axis was acquired during evolution from invertebrates (left) to vertebrates (right) as a transport-based mechanism to fine-tune the proteolytic activity of Taspase1. In invertebrates, Taspase1's intracellular localization is not actively regulated by transport signals. Hence, not only the proenzyme but also the active protease localizes to the cytoplasm and the nucleus, allowing the processing of an enlarged degradome (70 predicted targets). Lack of the NoLS prevents active nuclear accumulation and interaction with nucleolar proteins, resulting in reduced nuclear concentration of the enzyme and thus reduced proteolytic processing of nuclear substrates. In contrast, all vertebrate Taspase1 proteins acquired the highly conserved NoLS, mediating efficient nuclear transport *via* the importin- α / β -axis. Autoproteolytic *cis* processing of the proenzyme results in protease activation and binding to the nucleolar NPM1 protein, critical for the efficient cleavage of nuclear targets. Taspase1 exploits the CRM1-mediated nuclear export activity of NPM1 to gain transient access to the cytoplasm to process cytoplasmic substrates. Number of predicted Taspase1 substrates is indicated.

flexibility of *Drosophila* Taspase1 permits recognition and processing of the host cell factor protein, in contrast to the restricted human enzyme. However, *in silico* experiments and predictions of a protease's degradome obtained by systematic biochemical methods, such as proteomics, have to be confirmed in detail for each of the proposed targets. Clearly, future studies need to investigate whether the predicted targets are indeed cleaved by dTaspase1 and what are the biologic consequences thereof.

Overall, the structure of Taspase1 enzymes allowing substrate recognition and *trans* cleavage seem to differ among these ortholog proteases. Because to date no effective active site-directed inhibitors are available for human Taspase1, we suggest that resolving the structure of *Drosophila* Taspase1 may provide additional mechanistic insights, potentially aiding the rational design of inhibitory compounds for the human enzyme. However, our findings also advise caution. Although *Drosophila* is an important

TABLE 2. Overview of proteases in humans and *Drosophila* for which (active) nuclear transport/localization was suggested

Protease		Nuclear transport mechanism	Nuclear transport signal	Reference
Human	<i>Drosophila</i> homologs			
hTaspase1		Imp- α /NPM1 switch	NLS/NoLS	(12)
	dTaspase1	—	—	This work
MMP-12		?	—	(6)
SENP3		Signal-dependent transport	NLS/NoLS	(48)
	DmUlp1	?	?	(49)
Cath-D		Interaction with nuclear proteins?	—	(50)
	cathD (lysosomal/extracellular)	—	—	(51)
GranzymeA		Via perforin	—	(52)
Cath-L (isoform)		Indirect (loss of ER-signal)	—	(53)
	cathL/Cp1 (lysosomal/fusome)	—	—	(54)

model organism, results or targeting strategies obtained for *Drosophila* Taspase1 may not be directly translatable to the human enzyme.

During evolution, eukaryotes have acquired an additional way of controlling a protease's degradome by regulating its (dynamic) localization to specific cellular compartments. Depending on their biological function proteases are thus secreted, found on the outside of the cell membrane (43), in the mitochondria (44), the cytoplasm, in the nucleus as well as at the nucleolus (10, 12, 45–47). However, inspection of the literature reveals that there are only a few examples of proteases in humans, and even less in *Drosophila*, for which active nuclear transport and localization have been unambiguously shown. However, protease signaling may be combined with active nucleo-cytoplasmic transport, allowing an additional level of regulatory complexity (Table 2) (12). Strikingly, for the classic membrane-associated proteolytic enzymes, such as matrix metalloproteinases, it was recently shown that even these proteases can execute biologically relevant functions in the nucleus (6). However, with the exception of Taspase1 and SUMO-specific proteases (SENPs) in humans, the detailed mechanisms mediating active nuclear transport and localization are not completely resolved (Table 2). Eukaryotes do have the molecular repertoire to combine protease signaling with active nucleo-cytoplasmic transport. However, previous examples demonstrating that such mechanisms have indeed been acquired or lost during evolution were to our knowledge missing. In addition to the evolutionary protease divergence concerning the enlarged degradome of dTaspase1, we found that, in *Drosophila*, active nuclear transport and regulated localization of Taspase1 differ from its human relative. In fact, protein alignment revealed that the functionally important NoLS, which is highly conserved in all known vertebrate Taspase1 proteins, is absent in *Drosophila* and other species. Such divergence has implications on protease function. For one, dTaspase1 shows no distinct localization to the nucleus/nucleolus and fails to interact with the nucleoplasmin-like protein from *Drosophila* or with human NPM1. Clearly, in contrast to vertebrate Taspase1 enzymes, the subcellular localization of dTaspase1 is the result of passive diffusion rather than being regulated by active transport signals and protein interaction motifs. Second, as the NoLS-mediated interaction with Imp- α and the binding to NPM1 maintains high nuclear concentration of human Taspase1, the lack of both interactions explains why

dTaspase1 is less active in processing nuclear substrates. Here, we wish to emphasize that not all nucleolar localization signals, identified in a variety of regulatory proteins, such as NPM1, ARF, or p53, are capable of mediating both active nuclear import and nucleolar localization.

Based on our findings for Taspase1, we therefore suggest a model of how the Imp- α /NPM1 axis was acquired during evolution from invertebrates to vertebrates as a transport-based mechanism to fine-tune the proteolytic function of Taspase1 enzymes (Fig. 4D). In invertebrates, such as flies, Taspase1's intracellular localization is not actively regulated by transport signals. Hence, not only the proenzyme but also the active protease localizes to the cytoplasm and the nucleus, allowing the processing of an enlarged degradome, potentially required in these species. Lack of the NoLS precludes active nuclear accumulation and interaction with nucleolar proteins, resulting in a lower nuclear protease concentration and thus reduced proteolytic processing of nuclear substrates. In contrast, all vertebrate Taspase1 proteins acquired the highly conserved NoLS, mediating efficient nuclear transport *via* the Imp- α / β axis. Autoproteolytic *cis* processing of the proenzyme results in the active protease and binding to the nucleolar NPM1 protein, which is critical for the efficient cleavage of nuclear targets. Human Taspase1 further exploits the CRM1-mediated nuclear export activity of NPM1 to gain transient access to the cytoplasmic compartment to also process cytoplasmic substrates. However, besides increasing the protease's steady-state concentration in a given subcellular compartment, it is not known whether, particularly for human Taspase1, its nucleolar localization and NPM1 interaction are also fine-tuning the enzyme's proteolytic activity *per se*.

Collectively, our study describes a hitherto unknown evolutionary principle of how a protease exploits a transport-based mechanism to fine-tune its degradome and proteolytic activity in different species. However, we emphasize that both of our findings, the broader substrate specificity and difference in localization of the *Drosophila* enzyme, are interconnected and thus contribute to the protease's biologic functions. Although we are well aware that further experimental confirmation is required, it is suffice to hypothesize that the principle identified here is of general biologic relevance. We propose that divergent transport-based mechanisms have evolved for other regulatory proteins in different species, supporting the adaption of multicellular organisms to master evolutionary challenges. **FJ**

The authors thank Sandra Friedl for excellent technical assistance. This work was supported by Fritz Thyssen Stiftung, MAIFOR Mainz Support Program, Zeiss Foundation, and Interdisziplinärer Arbeitskreis (IAK) ChemBioMed. The authors declare no conflicts of interest.

REFERENCES

- Rawlings, N. D., Waller, M., Barrett, A. J., and Bateman, A. (2014) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **42**, D503–D509
- Fortelny, N., Cox, J. H., Kappelhoff, R., Starr, A. E., Lange, P. F., Pavlidis, P., and Overall, C. M. (2014) Network analyses reveal pervasive functional regulation between proteases in the human protease web. *PLoS Biol.* **12**, e1001869
- Buller, A. R., and Townsend, C. A. (2013) Intrinsic evolutionary constraints on protease structure, enzyme acylation, and the identity of the catalytic triad. *Proc. Natl. Acad. Sci. USA* **110**, E653–E661
- Overall, C. M., and Dean, R. A. (2006) Degradomics: systems biology of the protease web. Pleiotropic roles of MMPs in cancer. *Cancer Metastasis Rev.* **25**, 69–75
- Turk, B. (2006) Targeting proteases: successes, failures and future prospects. *Nat. Rev. Drug Discov.* **5**, 785–799
- Marchant, D. J., Bellac, C. L., Moraes, T. J., Wadsworth, S. J., Dufour, A., Butler, G. S., Bilawchuk, L. M., Hendry, R. G., Robertson, A. G., Cheung, C. T., Ng, J., Ang, L., Luo, Z., Heilbron, K., Norris, M. J., Duan, W., Bucyk, T., Karpov, A., Devel, L., Georgiadis, D., Hegele, R. G., Luo, H., Granville, D. J., Dive, V., McManus, B. M., and Overall, C. M. (2014) A new transcriptional role for matrix metalloproteinase-12 in antiviral immunity. *Nat. Med.* **20**, 493–502
- Beck, P., Dubiella, C., and Groll, M. (2012) Covalent and non-covalent reversible proteasome inhibition. *Biol. Chem.* **393**, 1101–1120
- Knauer, S. K., Fetz, V., Rabenstein, J., Friedl, S., Hofmann, B., Sabiani, S., Schröder, E., Kunst, L., Proschak, E., Thines, E., Kindler, T., Schneider, G., Marschalek, R., Stauber, R. H., and Bier, C. (2011) Bioassays to monitor Taspase1 function for the identification of pharmacogenetic inhibitors. *PLoS ONE* **6**, e18253
- Hsieh, J. J., Cheng, E. H., and Korsmeyer, S. J. (2003) Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. *Cell* **115**, 293–303
- Bier, C., Hecht, R., Kunst, L., Scheiding, S., Wünsch, D., Goesswein, D., Schneider, G., Krämer, O. H., Knauer, S. K., and Stauber, R. H. (2012) Overexpression of the catalytically impaired Taspase1 T234V or Taspase1 D233A variants does not have a dominant negative effect in T(4;11) leukemia cells. *PLoS ONE* **7**, e34142
- Bier, C., Knauer, S. K., Wünsch, D., Kunst, L., Scheiding, S., Kaiser, M., Ottmann, C., Krämer, O. H., and Stauber, R. H. (2012) Allosteric inhibition of Taspase1's pathobiological activity by enforced dimerization in vivo. *FASEB J.* **26**, 3421–3429
- Bier, C., Knauer, S. K., Docter, D., Schneider, G., Krämer, O. H., and Stauber, R. H. (2011) The importin- α /nucleophosmin switch controls taspase1 protease function. *Traffic* **12**, 703–714
- Bier, C., Knauer, S. K., Klapthor, A., Schweitzer, A., Reik, A., Krämer, O. H., Marschalek, R., and Stauber, R. H. (2011) Cell-based analysis of structure-function activity of threonine aspartase 1. *J. Biol. Chem.* **286**, 3007–3017
- Chen, D. Y., Liu, H., Takeda, S., Tu, H. C., Sasagawa, S., Van Tine, B. A., Lu, D., Cheng, E. H., and Hsieh, J. J. (2010) Taspase1 functions as a non-oncogene addiction protease that coordinates cancer cell proliferation and apoptosis. *Cancer Res.* **70**, 5358–5367
- Zhou, H., Spicuglia, S., Hsieh, J. J., Mitsiou, D. J., Høiby, T., Veenstra, G. J., Korsmeyer, S. J., and Stunnenberg, H. G. (2006) Uncleaved TFIIA is a substrate for taspase 1 and active in transcription. *Mol. Cell Biol.* **26**, 2728–2735
- Khan, J. A., Dunn, B. M., and Tong, L. (2005) Crystal structure of human Taspase1, a crucial protease regulating the function of MLL. *Structure* **13**, 1443–1452
- Chen, D. Y., Lee, Y., Van Tine, B. A., Searleman, A. C., Westergard, T. D., Liu, H., Tu, H. C., Takeda, S., Dong, Y., Piwnica-Worms, D. R., Oh, K. J., Korsmeyer, S. J., Hermone, A., Gussio, R., Shoemaker, R. H., Cheng, E. H., and Hsieh, J. J. (2012) A pharmacologic inhibitor of the protease Taspase1 effectively inhibits breast and brain tumor growth. *Cancer Res.* **72**, 736–746
- Wünsch, D., Fetz, V., Heider, D., Tenzer, S., Bier, C., Kunst, L., Knauer, S., and Stauber, R. (2012) Chemo-genetic strategies to inhibit the leukemic potential of threonine aspartase-1. *Blood Cancer J* **2**, e7
- Stauber, R. H., Bier, C., and Knauer, S. K. (2012) Targeting Taspase1 for cancer therapy—letter. *Cancer Res.* **72**, 2912, author reply 2913
- Pandey, U. B., and Nichols, C. D. (2011) Human disease models in *Drosophila melanogaster* and the role of the fly in therapeutic drug discovery. *Pharmacol. Rev.* **63**, 411–436
- Olivier, J. P., Raabe, T., Henkemeyer, M., Dickson, B., Mbamalu, C., Margolis, B., Schlessinger, J., Hafen, E., and Pawson, T. (1993) A *Drosophila* SH2-SH3 adaptor protein implicated in coupling the sevenless tyrosine kinase to an activator of Ras guanine nucleotide exchange, Sos. *Cell* **73**, 179–191
- Capotosti, F., Hsieh, J. J., and Herr, W. (2007) Species selectivity of mixed-lineage leukemia/trithorax and HCF proteolytic maturation pathways. *Mol. Cell Biol.* **27**, 7063–7072
- Knauer, S. K., Heinrich, U. R., Bier, C., Habtemichael, N., Docter, D., Helling, K., Mann, W. J., and Stauber, R. H. (2010) An optoprotective role for the apoptosis inhibitor protein survivin. *Cell Death Dis.* **1**, e51
- Fetz, V., Knauer, S. K., Bier, C., von Kries, J. P., and Stauber, R. H. (2009) Translocation Biosensors - Cellular System Integrators to Dissect CRM1-Dependent Nuclear Export by Chemicogenomics. *Sensors (Basel)* **9**, 5423–5445
- Padeken, J., Mendiburo, M. J., Chlamydas, S., Schwarz, H. J., Kremmer, E., and Heun, P. (2013) The nucleoplasmic homolog NLP mediates centromere clustering and anchoring to the nucleolus. *Mol. Cell* **50**, 236–249
- Habtemichael, N., Heinrich, U. R., Knauer, S. K., Schmidtman, I., Bier, C., Docter, D., Brochhausen, C., Helling, K., Brieger, J., Stauber, R. H., and Mann, W. J. (2010) Expression analysis suggests a potential cytoprotective role of Birc5 in the inner ear. *Mol. Cell Neurosci.* **45**, 297–305
- Docter, D., Distler, U., Storck, W., Kuharev, J., Wünsch, D., Hahlbrock, A., Knauer, S. K., Tenzer, S., and Stauber, R. H. (2014) Quantitative profiling of the protein coronas that form around nanoparticles. *Nat. Protoc.* **9**, 2030–2044
- Tenzer, S., Docter, D., Kuharev, J., Musyanovych, A., Fetz, V., Hecht, R., Schlenk, F., Fischer, D., Kiouptsi, K., Reinhardt, C., Landfester, K., Schild, H., Maskos, M., Knauer, S. K., and Stauber, R. H. (2013) Rapid formation of plasma protein corona critically affects nanoparticle pathophysiology. *Nat. Nanotechnol.* **8**, 772–781
- Knauer, S. K., Bier, C., Schlag, P., Fritzmann, J., Dietmaier, W., Rödel, F., Klein-Hitpass, L., Kovács, A. F., Döring, C., Hansmann, M. L., Hofmann, W. K., Kunkel, M., Brochhausen, C., Engels, K., Lippert, B. M., Mann, W., and Stauber, R. H. (2007) The survivin isoform survivin-3B is cytoprotective and can function as a chromosomal passenger complex protein. *Cell Cycle* **6**, 1502–1509
- Knauer, S. K., and Stauber, R. H. (2005) Development of an autofluorescent translocation biosensor system to investigate protein-protein interactions in living cells. *Anal. Chem.* **77**, 4815–4820
- Knauer, S. K., Moodt, S., Berg, T., Liebel, U., Pepperkok, R., and Stauber, R. H. (2005) Translocation biosensors to study signal-specific nucleo-cytoplasmic transport, protease activity and protein-protein interactions. *Traffic* **6**, 594–606
- Knauer, S. K., Krämer, O. H., Knösel, T., Engels, K., Rödel, F., Kovács, A. F., Dietmaier, W., Klein-Hitpass, L., Habtemichael, N., Schweitzer, A., Brieger, J., Rödel, C., Mann, W., Petersen, I., Heinzl, T., and Stauber, R. H. (2007) Nuclear export is essential for the tumor-promoting activity of survivin. *FASEB J.* **21**, 207–216
- Jun, L., Frints, S., Duhamel, H., Herold, A., Abad-Rodriguez, J., Doti, C., Izaurralde, E., Marynen, P., and Froyen, G. (2001) NXF5, a novel member of the nuclear RNA export factor family, is lost in a male patient with a syndromic form of mental retardation. *Curr. Biol.* **11**, 1381–1391

34. Van Deursen, D., van Leeuwen, M., Vaulont, S., Jansen, H., and Verhoeven, A. J. (2009) Upstream Stimulatory Factors 1 and 2 activate the human hepatic lipase promoter via E-box dependent and independent mechanisms. *Biochim. Biophys. Acta* **1791**, 229–237
35. Rosorius, O., Heger, P., Stelz, G., Hirschmann, N., Hauber, J., and Stauber, R. H. (1999) Direct observation of nucleocytoplasmic transport by microinjection of GFP-tagged proteins in living cells. *Biotechniques* **27**, 350–355
36. Schlingemann, J., Habtemichael, N., Ittrich, C., Toedt, G., Kramer, H., Hambek, M., Knecht, R., Lichter, P., Stauber, R., and Hahn, M. (2005) Patient-based cross-platform comparison of oligonucleotide microarray expression profiles. *Lab. Invest.* **85**, 1024–1039
37. Fetz, V., Bier, C., Habtemichael, N., Schuon, R., Schweitzer, A., Kunkel, M., Engels, K., Kovács, A. F., Schneider, S., Mann, W., Stauber, R. H., and Knauer, S. K. (2009) Inducible NO synthase confers chemoresistance in head and neck cancer by modulating survivin. *Int. J. Cancer* **124**, 2033–2041
38. Krämer, O. H., Knauer, S. K., Greiner, G., Jandt, E., Reichardt, S., Gührs, K. H., Stauber, R. H., Böhmer, F. D., and Heinzel, T. (2009) A phosphorylation-acetylation switch regulates STAT1 signaling. *Genes Dev.* **23**, 223–235
39. Cokol, M., Nair, R., and Rost, B. (2000) Finding nuclear localization signals. *EMBO Rep.* **1**, 411–415
40. Kosugi, S., Hasebc, M., Matsumura, N., Takashima, H., Miyamoto-Sato, E., Tomita, M., and Yanagawa, H. (2009) Six classes of nuclear localization signals specific to different binding grooves of importin alpha. *J. Biol. Chem.* **284**, 478–485
41. Drag, M., and Salvesen, G. S. (2010) Emerging principles in protease-based drug discovery. *Nat. Rev. Drug Discov.* **9**, 690–701
42. Clausen, T., Kaiser, M., Huber, R., and Ehrmann, M. (2011) HTRA proteases: regulated proteolysis in protein quality control. *Nat. Rev. Mol. Cell Biol.* **12**, 152–162
43. Ugalde, A. P., Ordóñez, G. R., Quirós, P. M., Puente, X. S., and López-Otín, C. (2010) Metalloproteases and the degradome. *Methods Mol. Biol.* **622**, 3–29
44. Koppen, M., Bonn, F., Ehses, S., and Langer, T. (2009) Autocatalytic processing of m-AAA protease subunits in mitochondria. *Mol. Biol. Cell* **20**, 4216–4224
45. Belanger, K. D., Walter, D., Henderson, T. A., Yelton, A. L., O'Brien, T. G., Belanger, K. G., Geier, S. J., and Fahrenkrog, B. (2009) Nuclear localisation is crucial for the proapoptotic activity of the HtrA-like serine protease Nma111p. *J. Cell Sci.* **122**, 3931–3941
46. Blink, E. J., Jiansheng, Z., Hu, W., Calanni, S. T., Trapani, J. A., Bird, P. I., and Jans, D. A. (2005) Interaction of the nuclear localizing cytolitic granule serine protease granzyme B with importin alpha or beta: modulation by the serpin inhibitor PI-9. *J. Cell. Biochem.* **95**, 598–610
47. Yun, C., Wang, Y., Mukhopadhyay, D., Backlund, P., Kolli, N., Yergey, A., Wilkinson, K. D., and Dasso, M. (2008) Nucleolar protein B23/nucleophosmin regulates the vertebrate SUMO pathway through SENP3 and SENP5 proteases. *J. Cell Biol.* **183**, 589–595
48. Yeh, E. T. (2009) SUMOylation and De-SUMOylation: wrestling with life's processes. *J. Biol. Chem.* **284**, 8223–8227
49. Smith, M., Bhaskar, V., Fernandez, J., and Courey, A. J. (2004) Drosophila Ulp1, a nuclear pore-associated SUMO protease, prevents accumulation of cytoplasmic SUMO conjugates. *J. Biol. Chem.* **279**, 43805–43814
50. Masson, O., Bach, A. S., Derocq, D., Prébois, C., Laurent-Matha, V., Pattingre, S., and Liaudet-Coopman, E. (2010) Pathophysiological functions of cathepsin D: Targeting its catalytic activity versus its protein binding activity? *Biochimie* **92**, 1635–1643
51. Eissenberg, J. C., Ilvarsonn, A. M., Sly, W. S., Waheed, A., Krzyzanek, V., Pohlmann, R., Waschkau, D., Kretzschmar, D., and Dennes, A. C. (2011) Drosophila GGA model: an ultimate gateway to GGA analysis. *Traffic* **12**, 1821–1838
52. Jans, D. A., Briggs, L. J., Jans, P., Froelich, C. J., Parasivam, G., Kumar, S., Sutton, V. R., and Trapani, J. A. (1998) Nuclear targeting of the serine protease granzyme A (fragmentin-1). *J. Cell Sci.* **111**, 2645–2654
53. Goulet, B., Baruch, A., Moon, N. S., Poirier, M., Sansregret, L. L., Erickson, A., Bogoy, M., and Nepveu, A. (2004) A cathepsin L isoform that is devoid of a signal peptide localizes to the nucleus in S phase and processes the CDP/Cux transcription factor. *Mol. Cell* **14**, 207–219
54. Lighthouse, D. V., Buszczak, M., and Spradling, A. C. (2008) New components of the Drosophila fusome suggest it plays novel roles in signaling and transport. *Dev. Biol.* **317**, 59–71

Received for publication September 9, 2014.
Accepted for publication December 29, 2014.

3.2.4 „TFIIA transcriptional activity is controlled by a ‚cleave-and-run‘ Exportin1 / Taspase1-switch“

Christian Schrenk, Verena Fetz, Cecilia Vallet, Christina Heiselmayer, Elisabeth Schröder, Astrid Hensel, Angelina Hahlbrock, Désirée Wünsch, **Dorothee Goesswein**, Carolin Bier, Negusse Habtemichael, Günther Schneider, Roland H. Stauber, Shirley K. Knauer.

*Journal of Molecular Cell Biology, 2017.*⁷

Zusammenfassung

Die Regulation der Genexpression ermöglicht Zellen, auf Umwelteinflüsse zu reagieren, unterschiedlich zu differenzieren und so spezifische Zelltypen und Gewebe auszubilden. Transkriptionsfaktoren sind die Proteinfamilie, die diese differenzielle Expression maßgeblich vermitteln und machen dabei etwa 10% der Gene des gesamten Genoms aus¹⁹⁸. Der Transkriptionsfaktor TFIIA wird zur basalen und aktiven Transkription einer Zelle rekrutiert und ist in der Lage, den TBP-DNA Komplex zu stabilisieren, indem die inhibitorische TAND2 Domäne von TAF1 innerhalb von TFIID verdrängt wird^{7,14,16-18}. Taspase1 spaltet TFIIA, fördert die Bindung zu dem Prostata-spezifischen Transkriptionsfaktor TRF2 und ermöglicht damit die Transkription entwicklungspezifischer Gene in der Spermatogenese¹²⁴. Entgegen bisheriger Annahmen verbleibt TFIIA jedoch auf Grund eines evolutionär konservierten nukleären Exportsignals (NES, aa²¹VINDVRDIFL³⁰), prädominant im Zytoplasma. Dabei interagiert TFIIA mit dem nukleären Exportrezeptor Exportin-1/Crm-1. Wird das NES genetisch inaktiviert (durch Umwandlung in²¹VINDVRDAFA³⁰) oder chemisch der Export inhibiert, lokalisiert TFIIA im Kern und verändert seine transkriptionelle Aktivität. Die nukleäre Akkumulation kann entweder durch eine Bindung mit TBP und einer damit verbundenen Maskierung des NES vermittelt werden, die durch p300 gefördert wird, oder durch die Abspaltung des NES durch die Taspase1. Dabei sind die Auswirkungen auf die transkriptionelle Aktivität unterschiedlich: Durch die Bindung mit TBP werden vor allem Gene

⁷ Eine Kopie des Artikels kann auf Grund der aktuellen Lizenzierung nicht abgedruckt werden.

ERGEBNISSE

bevorzugt transkribiert, die Zellzyklus- und Embryo-spezifisch sind, darunter der CDKN2A-Genlocus, wie an der verstärkten Transkription des p16^{INK4a}-Reporters demonstriert und bestätigt werden konnte. Die Prozessierung durch Taspase1 dagegen fördert unter anderem die Transkription von Genen, die spezifisch für die Spermiogenese sind, da die Spaltung von TFIIA $\alpha\beta/\gamma$ in TFIIA $\alpha/\beta/\gamma$ eine Bindung mit dem TBP-assoziierten Faktor TRF2 ermöglicht und die Affinität von TFIIA zu TBP selbst verringert. Die Bindung von TFIIA mit entweder TBP oder Taspase1 ist kompetitiv und fungiert als Regulationsmechanismus, der durch Kerntransport und Veränderung der Bindeaffinität eine transkriptionelle Feinregulation ermöglicht.

3.3 Systematische, wissenschaftliche und epidemiologische Erfassung von Kopf-Hals-Tumordaten

Translationale Krebsforschung ist seit über einem Jahrzehnt wesentlich für die Fortschritte in Diagnose und Behandlung verantwortlich. Genomsequenzierungsdaten verschiedener Krebsarten aus internationalen Projekten vermochten die erfolgreiche Entwicklung neuer, zielgerichteter Medikamente als auch neuer diagnostischer Möglichkeiten insbesondere im Bereich der Biomarkerforschung maßgeblich zu beeinflussen ¹⁹⁹. Die standardisierte molekulare Charakterisierung der Interaktion verschiedener Onkogene und Tumorsuppressoren in klinischer Anwendung wird durch die Heterogenität der Tumore erschwert. Lösungsansätze scheitern oft an der Komplexität der biologischen Fragestellungen. Der erste Schritt zur Erarbeitung sinnvoller Versuchsanordnungen ist das Sammeln großer Datenmengen zu spezifischen Tumorarten, die Erarbeitung daraus resultierender, statistisch signifikanter, medizinisch relevanter Korrelationen sowie die nachfolgende, gezielte, molekularbiologische Untersuchung möglicher kausaler Zusammenhänge.

3.3.1 Struktureller Aufbau der Datenbank „TUBA“

Ein solches Projekt zur wissenschaftlichen Sammlung und Auswertung vergleichbarer Daten von Kopf-Hals-Tumorpatienten ist im Rahmen dieser Arbeit in Kooperation mit aus dem Institut für Bioinformatik (Prof. Hildebrandt, Universität Mainz), dem Institut für Pathologie (Prof. Wilfried Roth, Universitätsmedizin Mainz), der Hals-, Nasen-, Ohrenklinik (Prof. Christoph Matthias, Universitätsmedizin Mainz) und der Arbeitsgruppe von Prof. Stauber (HNO Labore, Universitätsmedizin Mainz) entstanden. Die Tumordatenbank (TUBA) wurde von Maximilian Deichelbohrer im Rahmen seiner Masterarbeit ²⁰⁰ programmiert und umfasst derzeit Fall- und Tumor-bezogene klinische, histopathologische und immunhistochemische Daten von Kopf-Hals-Tumorpatienten der HNO-Klinik der Universitätsmedizin Mainz aus den Jahren 2008-2017. Die Systemarchitektur der TUBA baut auf einem *Apache Tomcat Server* mit einem *Spring Framework* als Kommunikationsschnittstelle zu den einzelnen *Controllern* mit dem *Spring Security Modul* zur

Kommunikationsüberwachung und Authentifizierung der Nutzer. Die Informationen selbst werden in einer PostgreSQL-Datenbank gespeichert. Die Kommunikation der in Java programmierten Webapplikation mit der PostgreSQL-Datenbank wird über das *Hibernate Persistence Framework* mit Hilfe von objektrationalem Mapping (ORM) zur Projizierung relationaler Datenmodelle reguliert ²⁰⁰.

Die Eingabe der Daten erfolgt über individuelle Eingabemasken zu Fall, Tumor und immunhistochemischen Schnitten in der Webapplikation. Die Entitäten werden jeweils über 1 zu n Beziehungen in der Datenbank gespeichert, was bedeutet, dass zu jedem Patienten (Fall) mehrere Tumore angelegt werden können und zu jedem Tumor mehrere Schnitte (**Abb. 54**).

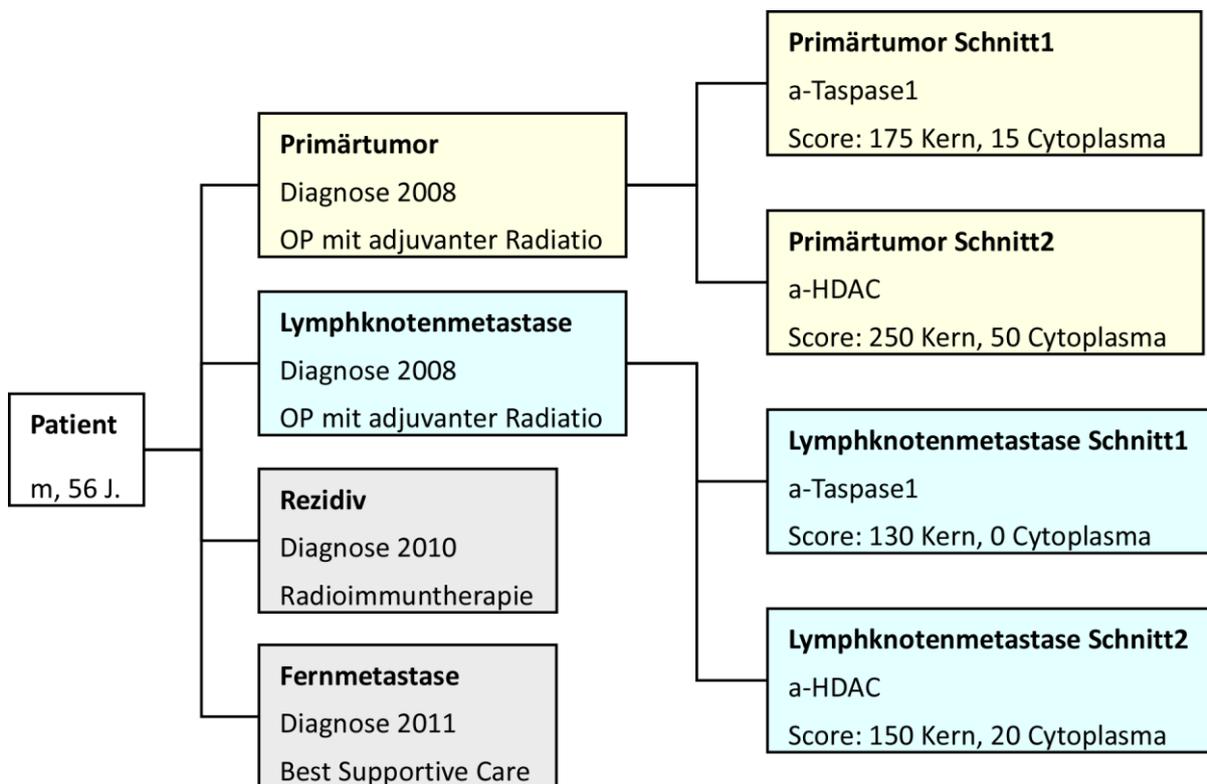


Abb. 54: Stark vereinfachte, beispielhaft dargestellte Datenbankstruktur.

In diesem fiktiven Fallbeispiel wird die Datenbankstruktur deutlich: Zu jedem Fall (Patient) können mehrere Tumore (Primärtumor, Lymphknotenmetastase, Rezidiv, Fernmetastase) angelegt werden und dazugehörige Informationen wie das Datum der Diagnose und Therapien (OP mit adjuvanter Radiatio, Radioimmuntherapie, Best Supportive Care) gespeichert werden. Tumore, die operativ entfernt werden, werden in der Biobank der Pathologie gelagert und können immunhistochemisch untersucht werden. Mit Hilfe eines aus Färbeintensität und Anteil der gefärbten Zellen ermittelten Scores kann die Intensität der jeweiligen Färbung (α -Taspase1, α -HDAC) für einzelne Lokalisationen (Kern, Zytoplasma) verdeutlicht und über mehrere Patienten und/oder zwischen Tumoren verglichen werden.

3.3.2 Epidemiologische und wissenschaftliche Daten

Die medizinische Evaluierung, Aufarbeitung und Eingabe der Daten erfolgt über derzeit vier medizinische Doktoranden der HNO und zwei medizinische Doktoranden der Pathologie unter der Supervision von Dr. med. habil. Künzel auf Seiten der HNO sowie Dr. med. Zimmer und Dr. med. Försch auf Seiten der Pathologie. Da die Dateneingabe kontinuierlich erfolgt, wurde ein momentäres Abbild der Datenbank vom 18.09.2017 für alle hier dargestellten Ergebnisse verwendet und soll lediglich als Zwischenübersicht der bereits eingegebenen Daten und ihrer hohen Qualität dienen. Der Fokus liegt derzeit auf der Eingabe aller Oropharynx-Plattenepithelkarzinome aus dem Zeitraum 2008-2012. Alle Daten sind ungefiltert und nur rudimentär von klaren Ausreißern bereinigt, die nicht in den Datenzeitraum oder zur Indikation Plattenepithelkarzinom passen. Auch andere Lokalisationen als Oropharynx werden mit eingegeben und so für zukünftige wissenschaftliche Vorhaben vorgehalten. Diese sind daher in der folgenden Darstellung mitberücksichtigt.

Für die Tumore aller Patienten wurde die 7. Auflage der TNM-Klassifikation ²⁰¹ zur histopathologischen Einstufung des Tumors in der klinischen Routine verwendet und wird derzeit mit der 8. Auflage der TNM-Klassifikation vergleichend überarbeitet ²⁰². Für die Datentabellen wurde zuerst über die Primärtumore selektiert (auch für Patienten-bezogene Werte), und im Anschluss über die Existenz des jeweiligen Wertes. Dadurch können sich kleinere Abweichungen der Gesamtzahlen ergeben, da bislang nicht für jeden Patienten und jeden Tumor alle Werte eingegeben oder herausgefiltert wurden. Insgesamt ist die Datenkohärenz jedoch durch eine äußerst sorgfältige Dateneingabe sehr gut und bedarf nur kleiner Bereinigungen für die weitere wissenschaftliche Auswertung.

Tab. 9 zeigt eine Übersicht über alle in der Datenbank aufgezeichneten Primärtumore, aufgeteilt nach Lokalisation. Die Mehrheit der Patienten war mit 72% männlich und unabhängig vom Geschlecht zu 87% über einem Alter von 50 Jahren. Die meisten Primärtumore waren im Oropharynx lokalisiert, wie aus dem bisherigen retrospektiven Studiendesign zu erwarten war. 59% der Tumore wurden einem UICC Stadium von IV zugeordnet, was dem allgemein beobachteten klinischen Alltag entspricht und zu großen Teilen der allgemein eher späten Vorstellung von Kopf-Hals-Tumorpatienten zuzuschreiben ist ¹³⁹.

ERGEBNISSE

Tab. 9: Übersicht über alle in der Datenbank aufgezeichneten Tumore, aufgeteilt nach Lokalisation.

	Alle	Oropharynx	Mundhöhle	Larynx	Hypopharynx	Nasopharynx
Geschlecht						
männlich	420	294	52	21	32	2
weiblich	160	110	24	4	8	0
Alter						
≤ 50	77	47	10	5	3	0
>50	500	357	65	19	37	2
UICC						
I	49	31	13	2	2	0
II	37	26	5	4	1	0
III	85	58	11	7	8	0
IV	245	193	23	6	18	1
T						
1/2	272	198	49	9	8	0
3/4	253	184	21	14	31	1
X/0	9	2	1	1	0	1
N						
0	151	100	26	14	6	0
1/2/3	102	67	14	4	13	1
X	12	4	5	1	0	1
M						
0	377	278	47	16	24	0
1	51	35	8	3	5	0
X	80	51	11	5	10	2
G						
1/2	336	243	50	18	19	0
3/4	170	128	18	2	16	2
X	6	3	2	0	0	0
L						
0	27	18	8	0	1	0
1	59	46	5	0	8	0

ERGEBNISSE

	Alle	Oropharynx	Mundhöhle	Larynx	Hypopharynx	Nasopharynx
V						
0	34	21	8	0	5	0
1	15	11	2	0	0	0
2	2	1	0	0	1	0
R						
0	191	138	36	4	5	0
1	67	46	13	4	2	0
2	17	12	1	0	4	0

Um die UICC-Verteilung besser im Verhältnis zum Überleben der Patienten sichtbar zu machen, wurde eine Überlebenskurve in Abhängigkeit der UICC-Einteilung generiert (**Abb. 55**).

ERGEBNISSE

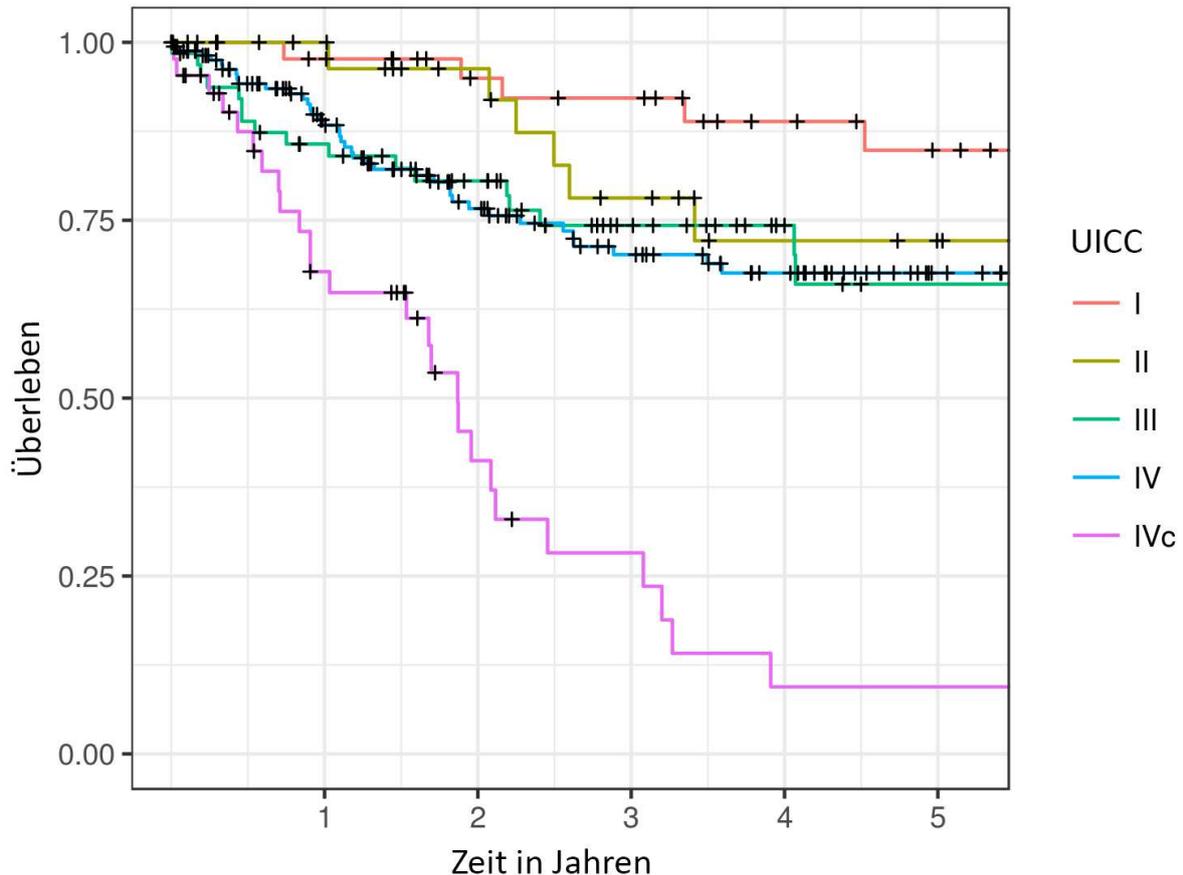


Abb. 55: Kaplan-Meier Überlebenskurven für die UICC Stadien I, II, III, IVab (IV) und IVc. Die einzelnen Stadien sind, wie in der Legende angegeben, farblich unterschieden.

Auffällig ist, dass die Überlebenskurve für das Stadium IVc, das durch Fernmetastasen (M1) gekennzeichnet ist, deutlich schlechter ist als das aller anderen Stadien, inklusive der Stadien IVa und IVb (in **Abb. 55** als „IV“ zusammengefasst). Der Behandlungserfolg der UICC-Stadien I-IVab war also an der Universitätsmedizin in Mainz während der Jahre 2008-2012 tendenziell besser als bei einem UICC-Stadium von IVc und ist nach Therapieerfolg in **Tab. 10** dargestellt. Erwartungsgemäß wurden mit einem UICC-Stadium von I mit 74% relativ die meisten Vollremissionen verzeichnet, während bei einem Stadium von IVc mit 66% der größte Anteil dennoch in einer partiellen Remission (34%) oder Progress (32%) resultierten.

Tab. 10: Therapieerfolge in Abhängigkeit der UICC-Stadien.

	alle	unbekannt	keine Änderung	Voll- remission	partielle Remission	Abbruch	Progress
I	74	6	0	55	8	2	3
II	52	6	1	28	9	6	2
III	145	12	3	73	33	20	4
IV	391	44	7	155	119	29	37
IVC	200	30	11	8	68	19	64

Mit welchen Therapien welche Erfolge verzeichnet werden konnten, ist in **Tab. 11** aufgelistet. Insgesamt zeigten sich Operationen als die vielversprechendste Option, wobei diese insbesondere in Kombination mit nachfolgender Strahlen- und gleichzeitiger Chemotherapie in den anteilig meisten Vollremissionen (68%) resultierte. Die primären Therapien waren durchschnittlich zu 44% mit einer Vollremission erfolgreich und konnten zu 19% immerhin noch eine partielle Remission erwirken. Erwartungsgemäß endeten die palliativen Radio- und Chemotherapien sowie alle Behandlungen, die unter *Best Supportive Care* (BSC) zusammengefasst werden, zu durchschnittlich 15% in keiner Änderung und zu 60% in einem Fortschreiten der Erkrankung.

Diese Zusammenhänge spiegeln sich auch in der Überlebenskurve in **Abb. 56D** wider, in der mit operativen Therapien, Induktions- und primären Therapien sowie palliativen/BSC Behandlungen drei sehr unterschiedliche Erfolgfelder zu erkennen sind.

Tab. 11: Therapieerfolge in Abhängigkeit der Behandlung nach Primärtumoren (Alle).

	alle	Voll- remission	partielle Remission	Abbruch	Progress	keine Änderung
Induktions- chemotherapie	180	11	98	33	22	7
Operationen gesamt	304	163	44	11	12	1
OP ohne adjuvante Therapie	107	47	28	0	6	1
OP mit adjuvanter Radiatio	87	40	7	6	6	0
OP mit adjuvanter	112	76	9	5	0	0

ERGEBNISSE

	alle	Voll-remission	partielle Remission	Abbruch	Progress	keine Änderung
Radiochemo						
primäre Therapien gesamt	207	91	40	27	18	1
primäre Radiatio	38	15	5	4	8	1
primäre simultane Radiochemo	98	43	23	13	3	0
primäre Radioimmuntherapie	71	33	12	10	7	0
andere	47	2	2	2	28	7
Tumorebulking	1	0	1	0	0	0
TACE	1	0	0	0	0	1
palliative Chemo-/Immuntherapie	26	1	1	2	17	4
<i>Best Supportive Care</i>	19	1	0	0	11	2

In **Abb. 56** sind zusätzlich die Überlebenskurven in Abhängigkeit von Alter (**A**), Geschlecht (**B**) und Lokalisation (**C**) dargestellt.

ERGEBNISSE

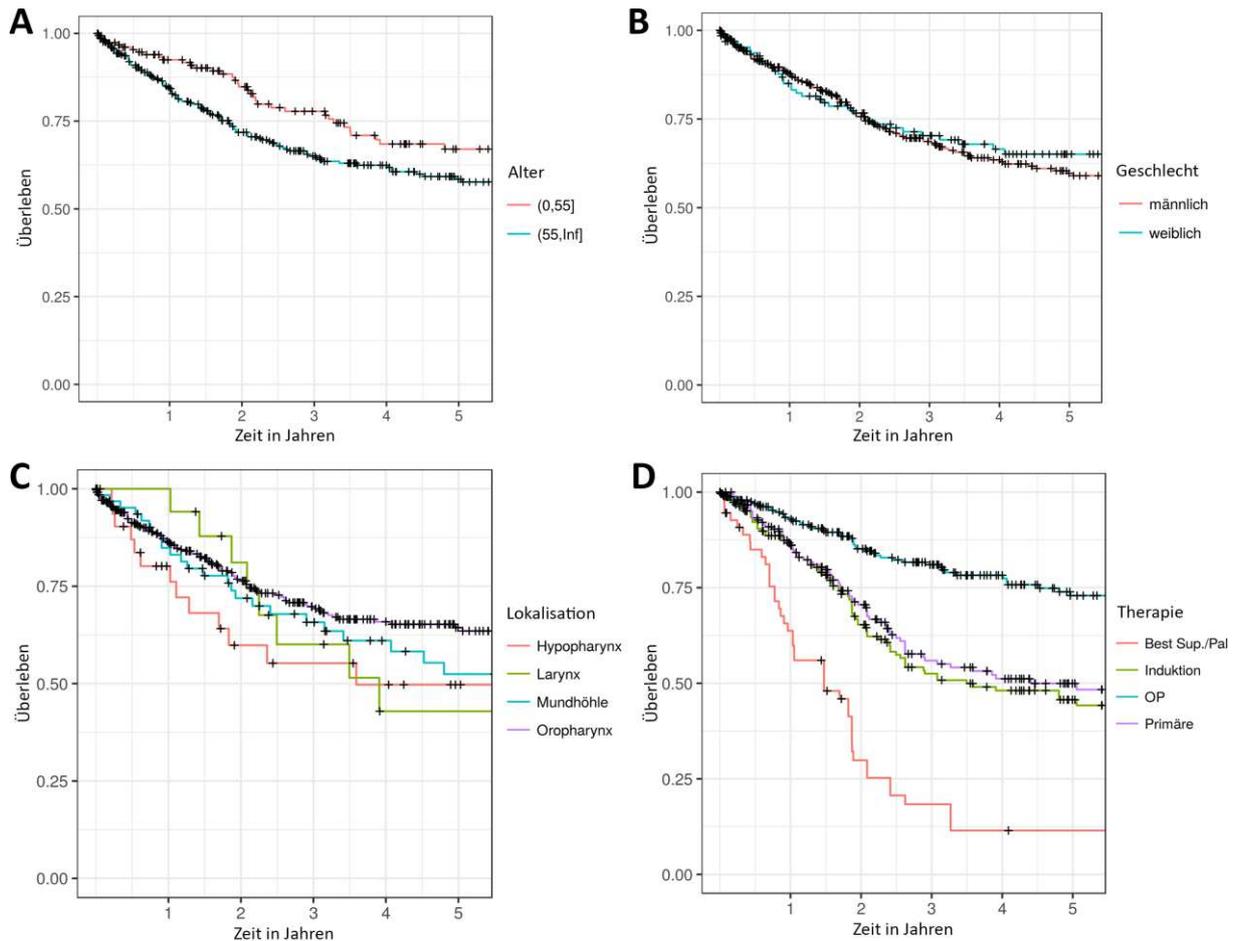


Abb. 56: Kaplan-Meier-Überlebenskurven nach Alter, Geschlecht, Lokalisation des Primärtumors und Therapie.

A. Überlebenskurven in Abhängigkeit des Alters für alle Patienten bis oder gleich 55 Jahre und älter als 55 Jahre. B. Überlebenskurven in Abhängigkeit des Geschlechtes der Patienten. C. Überlebenskurven in Abhängigkeit der Lokalisation des Tumors. D. Überlebenskurven in Abhängigkeit der Therapie. OP: alle operativen Therapien, inklusive solcher mit adjuvanter Radio/Chemo/Immuntherapie, Induktion: alle Induktionstherapien, Primäre: alle primären Radio/Chemo/Immuntherapien, Best Sup./Pal: Best Supportive Care und palliative Behandlungen. Die unterschiedlichen Kategorien sind jeweils, wie in den Legenden angegeben, farblich unterschieden.

Diese bisher fast ausschließlich klinischen und histopathologischen Daten bilden die Basis für spätere Korrelationsanalysen, um daraus die pathobiologische Wirkung der Taspase1, ihrer Interaktionspartner und damit zusammenhängende zellbiologische Signalwege mit epidemiologischen Daten verknüpfen zu können. Derzeit werden Daten und Scores aus immunhistochemischen Färbungen ermittelt und in die Datenbank eingespeist, so dass prognostische und prädiktive Korrelationsaussagen möglich werden und neue Biomarker auf der Basis solider evidenzbasierter Daten entwickelt werden können.

ERGEBNISSE

Im pathologischen Institut der Universitätsmedizin Mainz wurden zu den jeweiligen Patienten die rezidierten Tumore routinemäßig in Formalin fixiert, in Paraffin eingebettet und histopathologisch untersucht. Dieses wertvolle Biomaterial steht für das Projekt im Rahmen der Biobank-Initiative zur Verfügung und wird unter der Supervision von Dr. Zimmer im pathologischen Institut bearbeitet. Zu Primärtumoren und vorhandenen Metastasen werden Tissue Microarray (TMA)-Schnitte angefertigt und nach etablierten Standards im Rahmen verschiedener medizinischer Dissertationen immunhistochemisch gefärbt.

Zur eigentlichen Auswertung und Vergleichbarkeit im Rahmen der medizinischen Doktorarbeiten wird zusätzlich eine objektive Datenselektion erfolgen müssen, um die epidemiologische Repräsentierbarkeit zu gewährleisten, da die hier aufgezeigten Rohdaten ungefiltert nahezu alle derzeit in der Datenbank gespeicherten Datensätze enthalten.

4 DISKUSSION UND AUSBLICK

Präzisionsmedizin setzt eine Vielfalt an therapeutischen und begleitenden diagnostischen Möglichkeiten voraus, um die individuellen Bedürfnisse eines jeden Patienten zu ermitteln und eine persönlich angepasste Therapie zu ermöglichen. Prognostische Biomarker, die den Fortgang der Erkrankung vorhersagen können, und prädiktive Biomarker, die die Wirkungsweise und Nebenwirkungen spezifischer Medikamente am Patienten einschätzen lassen, sind für die Effektivität der Behandlung essentiell. Das tiefe Verständnis der zu Grunde liegenden Mechanismen der Malignität sind dabei sowohl die Basis für die Entwicklung neuester Medikamente als auch für eine sinnvolle Biomarker-Selektion.

Im ersten Teil dieser Arbeit wurden Expressionsprofile von primären Kopf-Hals-Plattenepithelkarzinomen mit ihren korrespondierenden Lymphknotenmetastasen und gesunder Mundschleimhaut untersucht und verglichen, während frühere Studien sich maßgeblich auf den Vergleich von Primärtumoren mit Normalgewebe und damit der Tumorgenese konzentrierten^{153-158,203}. Hier konnten mehrere Gene und Signalwege identifiziert und verifiziert werden, die in verschiedenen malignen Erkrankungen als überexprimiert beschrieben wurden, jedoch bislang noch nicht alle mit Kopf-Hals Tumoren in Verbindung gebracht wurden oder aber nicht mit der Entwicklung von Lymphknotenmetastasen assoziiert wurden. Da in dieser Studie nur gepaarte Proben von 15 Patienten zur Verfügung standen, ist die Signifikanz der erhaltenen Daten mit Vorsicht zu interpretieren. Positiv zu werten ist jedoch, dass die Patienten auf Grund des Studiendesigns hohe Ähnlichkeiten zueinander aufwiesen: Alle waren N-positiv bei keinen detektierbaren Fernmetastasen (cM0), alle konsumierten Tabakprodukte und Alkohol, und bei allen wurden die Primärtumore als HPV-negativ eingestuft. Auch der pathologisch ermittelte Differenzierungsgrad der Tumore war mit Grad II und III sehr ähnlich. Die größten Unterschiede betrafen die Größe der Primärtumore (T1-4) als auch die Anzahl und Größe der betroffenen Lymphknoten (N1-3).

Eine Übersicht der aus der Gensignatur ermittelten Zusammenhänge ist in Abb. 57 dargestellt.

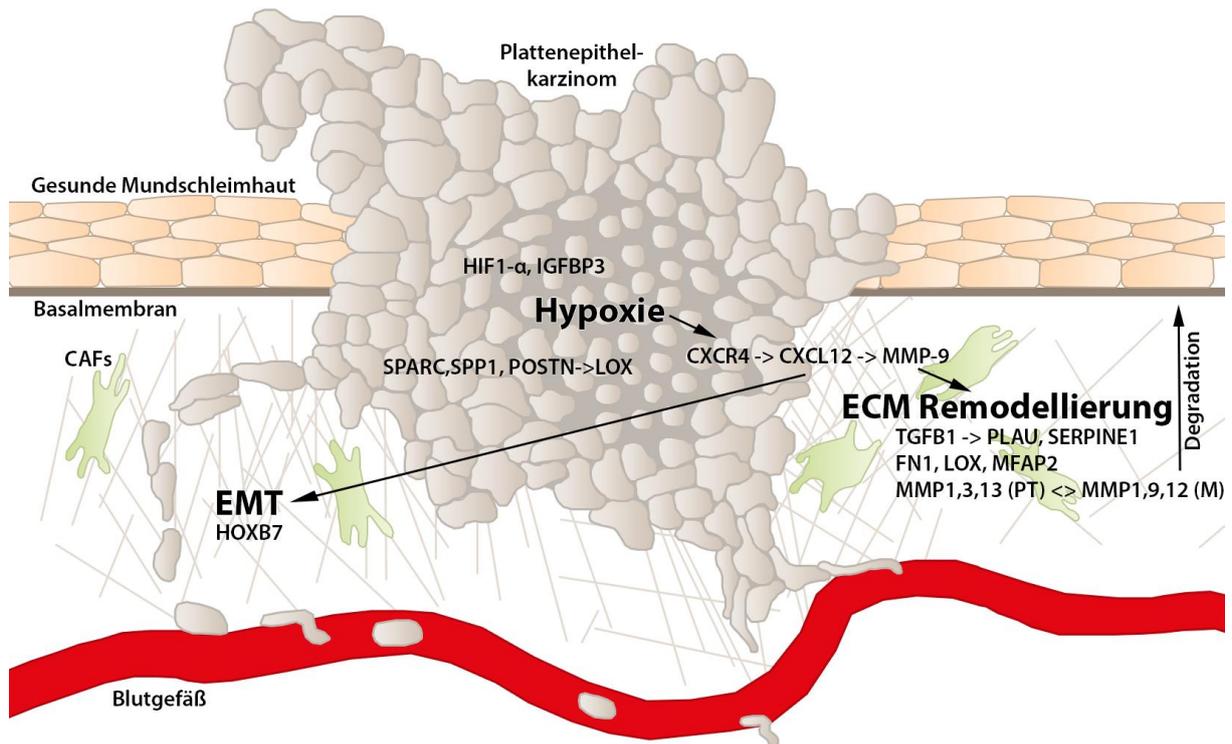


Abb. 57: Schematische Darstellung der Zusammenhänge aus Hypoxie, epithelialer-zu-mesenchymaler Transition (EMT) und Remodellierung der extrazellulären Matrix (ECM).

Essentiell für den invasiven Prozess eines Tumors und damit auch für dessen Malignität ist der Abbau der Basalmembran und ein Remodellieren der umgebenden extrazellulären Matrix (ECM), um so die Invasion der Tumorzellen einzeln oder im Verbund zu ermöglichen^{204,205}. Diese Aussage wird durch die signifikant hochregulierten Matrixmetalloproteasen, Integrine, Kollagene, Fibronectin, Laminine und andere ECM-assoziierte Proteine in Primärtumoren wie zuvor beschrieben unterstützt²⁰⁶.

Der *Urokinase-type Plasminogen Activator* (PLAU) und der *Serpin Peptidase Inhibitor* (SERPINE1) wurden zuvor als prognostische Marker für Brustkrebs und andere maligne Erkrankungen beschrieben^{207,208}. Eine erhöhte Expression jedes dieser Proteine in Primärtumoren wird mit einer schlechten Prognose assoziiert²⁰⁷. Die nicht-proteolytischen Funktionen des Proteins von SERPINE1 sind maßgeblich mit der Assoziation einer schlechten Prognose verknüpft²⁰⁹ und scheinen durch den *Transforming Growth Factor β 1* (TGF β 1) zusammen mit PLAU reguliert zu werden^{210,211}. Die Hochregulierung von TGF β 1 und dazugehörigen Rezeptoren in PTvs.N deutet zudem auf eine

Aktivierung tumorfördernder Eigenschaften von Tumor-Assoziierten-Makrophagen (TAMs) hin und damit auf ein erhöhtes invasives Verhalten des Tumors^{68,208}.

MFAP2 und FN1 sind Interaktionspartner von Kollagenen, die in zelluläre Prozesse wie der Adhäsion und der Kollagenfaserzusammensetzung involviert sind. Netzartige Fibroblasten (FCRs) sind Teil der Stromazellen von Lymphknoten und produzieren hohe Mengen von LOX und MFAP2²¹², die beide bereits im PTvs.N Datensatz signifikant überexprimiert waren. Die Expression von MFAP2 korreliert erstaunlicherweise ebenfalls mit fortgeschrittenen pathologischen Stufen von Magenkrebs²¹³. *Molecules Interacting with CasL-2 Proteins* (MICALs) sind Monooxygenasen, welche die Depolymerisierung von Aktinfilamenten katalysieren und mit der Regulation der Dynamik des Zytoskeletts, intrazellulärem Transport, dem Docken und der Fusion von exozytotischen Vesikeln verbunden sind und anti-apoptotische Eigenschaften besitzen²¹⁴. MICAL2, das in PTvs.N signifikant überexprimiert war, wurde als eines der am signifikantesten hochregulierten Proteine in Noxen-induziertem Lungenkrebs identifiziert²¹⁵. Es wurde sowohl mit der Progression von Prostatakrebs²¹⁶ als auch mit Epithelialer-zu-mesenchymaler-Transition (EMT), Krebswachstum und -invasion²¹⁷ assoziiert, aber bisher noch nicht im Speziellen mit Kopf-Hals-Tumoren in Zusammenhang gebracht. MICAL2 stellt also ein neues Ziel für weitere Biomarkerforschung dar und könnte bei weiteren mechanistischen Untersuchungen als prognostischer Biomarker für Kopf-Hals-Tumore an Bedeutung gewinnen.

Der Abbau und die folgende Remodellierung der extrazellulären Matrix treibt nicht nur die mechanische und proteolytische Invasion der Tumorzellen voran, sondern auch die metastatische Umprogrammierung von Zellen und ihrem Verhalten im Zellverband²¹⁸⁻²²⁰. Spezialisierte Krebs-assoziierte Fibroblasten (CAFs) können mehrere ECM-Proteine wie Kollagene, Laminin und Fibronectin degradieren²¹⁸. Dies führt zu einem aggressiveren Verhalten des Tumors sowie zu Lymphknotenmetastasen, was schließlich eine schlechte Prognose zur Folge hat, wie in Mundhöhlenkrebs gezeigt werden konnte^{221,222}. Die Expression des *Chemokin C-X-C Motif Receptor Typ 4* (CXCR4) wird durch Hypoxie erhöht und bindet den *Stromal-derived Factor 1* (CXCL12), der durch CAFs sekretiert wird. Dadurch wird die Sekretion von MMP9 in die Tumor-Mikroumgebung stimuliert^{73,75}. Im Vergleich des Datensatzes von PTvs.N zu Mvs.N verändert sich die Expression von MMP1, 3 und 13 in Primärtumoren zu MMP1, 9 und 12 in Metastasen. Nur in Metastasen waren CXCR4 und CXCL12 signifikant überexprimiert. Grundsätzlich scheint das Verhältnis von aktivierten MMPs zur totalen MMP-Konzentration im Vergleich zu verwandtem Normalgewebe positiv mit dem

Lymphknotenbefall zu korrelieren ^{75,223}. Dabei ist MMP9 die vielseitigste der Metalloproteasen und u.a. in der Lage, verschiedenartige Elemente der Tumor-Mikroumgebung abzubauen, wie z.B. Elastine, Fibrilline, Laminine, Gelatine und Kollagene des Typs IV, V, XI und XVI ^{73,224}. Die Migration endothelialer Progenitorzellen wird über die CXCL12/CXCR4-Achse zudem über PI3K/Akt/eNOS Signale vermittelt ²²⁵. Zudem erleichtert das CXCL12/CXCR4 System durch die Aktivierung der EMT die Ausbildung von metastatischem Potential in Mundhöhlenkrebs ²²⁶.

Hypoxie ist in soliden Tumoren, im Speziellen in Kopf-Hals Tumoren ²²⁷, sehr verbreitet und wurde wiederholt mit einer schlechten Prognose assoziiert ²²⁸. Zudem stimuliert eine sauerstoffarme Tumor-Mikroumgebung die Expression von Genen, die den Stammzellerhalt, die Angiogenese und die EMT unterstützen, was zu einem aggressiven, invasiven Phänotyp führt ²²⁹⁻²³¹.

Die Transkription des *Insulin-like Growth Factor Binding Protein 3* (IGFBP3) wird durch HIF1- α ²³² aktiviert, stimuliert das Ösophaguskrebswachstum in hypoxischen Mikroumgebungen ²¹¹ und fördert TGF- β -vermittelte EMT ²³³. *Periostin* (POSTN), das sowohl in Primärtumoren als auch in Metastasen hochsignifikant überexprimiert war, wird im Normalgewebe zur Gewebereparatur bei mechanischem Stress ausgeschüttet ²³⁴ und wird bei Hypoxie zusammen mit *Osteopontin* (SPP1) hochreguliert ²³⁵. Letzteres wurde in Kopf-Hals-Tumoren bereits mit der Förderung von invasivem Verhalten und Lymphangiogenese in Verbindung gebracht ^{236,237}. *Periostin* stimuliert zudem metastatisches Wachstum von Dickdarmkrebs und Nicht-Kleinzelligem Lungenkrebs durch die Aktivierung des Akt/PKB-Signalwegs ^{238,239}. Es stimuliert dabei die proteolytische Aktivität der *Lysyl Oxidase* (LOX) durch die Interaktion mit dem *Bone Morphogenic Protein 1* (BMP1) ²⁴⁰ und wurde mit EMT in mehreren Krebsarten assoziiert ²⁴¹⁻²⁴⁵. Folgerichtig wurde *Periostin* als Tumormarker für Invasion in Ösophagus-Krebs vorgeschlagen ²⁴⁶ und ebenso als Ziel für die Entwicklung neuer zielgerichteter Therapien ²⁴⁷⁻²⁴⁹. *Osteopontin* (SPP1) wird durch die Kinase Aurora-A stimuliert, verstärkt invasives Verhalten des Tumors und wurde als prognostischer Biomarker bereits für Brustkrebs vorgeschlagen ^{250,251}. *Osteonectin* (SPARC) ist ein sekretiertes Glykoprotein, das mit metastatischer Progression in Prostatakrebs und Melanomen assoziiert wurde ²⁵²⁻²⁵⁴. Für SPARC wurde zudem eine Funktion als Anti-Stress Faktor vermutet, der p53 durch AKT-vermittelte MDM2 Phosphorylierung inaktiviert und die EMT in Melanomen aktiviert ^{255,256}.

Homeobox B7 (HOXB7) ist ein Entwicklungsgen, das im Datensatz PTvs.N signifikant hochreguliert war. HOXB7 vermittelt EMT in Brustkrebs ²⁵⁷, verstärkt eine maligne Progression durch die

Aktivierung des TGF-Signalwegs²⁵⁸, fördert Lungenmetastasen²⁵⁹ und ermöglicht Brustkrebszellen über die Aktivierung des EGFR-Signalwegs die Entwicklung von Resistenzen gegenüber Tamoxifen²⁶⁰. Eine HOXB7-Überexpression wird darüber hinaus mit einer schlechten Prognose in Verbindung gebracht²⁶¹ und wurde als prognostischer Faktor für Plattenepithelkarzinome der Mundhöhle vorgeschlagen²⁶².

Der *Phosphoinositol-3-Kinase* (PI3K)-Signalweg ist der am häufigsten aktivierte onkogene Treiber in Kopf-Hals-Tumoren¹⁶⁷. Seine Upstream-Regulatoren und Downstream-Ziele sind mitunter die interessantesten Kandidaten für prädiktive Biomarker^{263,264}. Während kein einzelner Faktor des PI3K-Signalwegs im untersuchten Datensatz über alle Patienten gleichermaßen signifikant überexprimiert war, konnte die globale Aktivierung des Signalwegs hochsignifikant in der *Ingenuity* Analyse bestätigt werden. Die weiterführende Einzelanalyse verschiedener Rezeptortyrosinkinasen (RTK) zeigte, dass mit der Ausnahme von zwei Patienten bei allen wenigstens eine RTK signifikant hochreguliert war. Interessanterweise zeigten die meisten Patienten wenigstens eine Hochregulation von zwei RTKs in unterschiedlichen Kombinationen. Mehrere Korrelationen, die verschiedene Sets von RTKs beinhalten, konnten bisher in einigen Tumorentitäten gezeigt werden: In einer Meta-Analyse für Brustkrebs konnten ebenfalls verschiedene Kombinationen von überexprimierten RTKs mit schlechten Prognosen korreliert werden²⁶⁵. Die Autoren schlugen die klinische Evaluierung von Kombinationsagenzien gegen mehrere RTKs oder relevante onkogene Knoten vor²⁶⁵. Ebenfalls in Brustkrebs konnte gezeigt werden, dass eine MEK-Inhibition in spontanem Verlust der ERK-Aktivität resultiert und folgend c-MYC degradiert wird, was wiederum zu einer induzierten Expression von PDGFRB, DDR2 und VEGFR2 führt²⁶⁶ – einer RTK Kombination, die in den Patienten 6 (PT) und 10 (M) gefunden wurde. Die Hochregulierung von FGFRs, die in den Patienten 4, 7, 9, 10, 12, 14 und 15 festgestellt werden konnte, wurde als Resistenz-Mechanismus gegen anti-VEGFR Therapien spekuliert, da beide Signalwege in die Tumorneoangiogenese involviert sind (zusammengefasst in²⁶⁷). In Brustkrebspatienten mit FGFR1 oder FGF3 Amplifikation zeigten FGFR- und VEGFR-Inhibitoren vielversprechende Aktivität²⁶⁸. In RTK-induzierten und -abhängigen Tumoren wurde grundsätzlich vorgeschlagen, die Erkrankung durch eine Konzentration auf die Verbindungen zwischen den Signalwegen zu behandeln, z.B. durch eine Inhibition von c-Abl als Verbindung zwischen MET und p53²⁶⁹ (vgl. Patienten 2, 5, 6 und 9).

Im Vergleich mit vorhandenen genomischen Sequenzierungsdaten des *Cancer Genome Atlas*-Projektes konnten die vorliegenden Ergebnisse individuelle RKT-Expressionsdaten für jeden

Patienten zeigen. Obwohl große Gemeinsamkeiten zwischen den Patienten in der Expressionssignatur bestand, lagen die Upstream-befindlichen, möglicherweise verursachenden RTKs in sehr unterschiedlicher Kombination vor. Genau gegen diese individuell hochexprimierten RTKs werden derzeit mehrere zielgerichtete Medikamente in klinischen Studien getestet¹⁶³. Für diese wird es essentiell sein, dass prädiktive Biomarker entwickelt werden, die den behandelnden Onkologen in die Lage versetzen, die für den Patienten optimale Therapie zu wählen. Derzeit fehlen jegliche molekulare Begleitdiagnostika im Kopf-Hals-Tumorbereich, welche die Therapieentscheidung beeinflussen könnten. Für Cetuximab, die einzige zielgerichtete Therapie, die derzeit für HNSCC zugelassen ist, ist der alleinige Indikator für eine positive EGFR-gerichtete Therapieantwort ein Akne-artiger Ausschlag^{270,271}, obwohl bereits seit Jahren die molekularen Funktions- und Resistenzmechanismen intensiv beforscht werden (zusammengefasst in^{272,273}). EGFR konnte in gesunder Mundschleimhaut Tabakprodukte-nutzender Patienten als amplifiziert nachgewiesen werden und könnte daher als Antwort auf Karzinogene überexprimiert werden, fungiert jedoch wahrscheinlich nicht als onkogener Treiber²⁷³. Dieser Zusammenhang könnte unter anderem eine Erklärung für die fehlende differenzielle EGFR-Überexpression bei allen untersuchten Patienten sein, sowohl im Vergleich von Normalgewebe zu Primärtumoren, als auch zu Lymphknotenmetastasen.

Für eine erfolgreiche Behandlung müssen also alle onkogenen Treiber gleichzeitig ausgeschaltet werden, wie in präklinischen Untersuchungen bereits gezeigt werden konnte^{274,275}. Ohne eine Entwicklung geeigneter Biomarker wird eine solche Umsetzung von Präzisionsmedizin im klinischen Alltag nicht möglich sein²⁷⁶. Erst kürzlich wurden für Kopf-Hals Tumore einige vielversprechende, prognostische Expressionssignaturen vorgeschlagen²⁷⁷⁻²⁸¹; es fehlen jedoch weiterhin prädiktive Ansätze. Um eine Individualisierung von Therapien zu ermöglichen, wird das molekulare Profiling des Tumors an erster Stelle stehen, gefolgt von der Analyse prognostischer und prädiktiver molekularer Biomarker und schlussendlich darauf basierend der Auswahl der effektivsten Therapie für den Patienten. Entsprechend haben Biomarker das Potential, zum Wohl des Patienten eine Brücke zwischen Grundlagenforschung und klinischer Anwendung zu schlagen. Leider erreichen jedoch nur 3-5% der an die Entwicklung eines Medikaments gekoppelten Biomarker eine Zulassung für die klinische Anwendung²⁸²⁻²⁸⁶. Es besteht in diesem Bereich also weiterhin ein großer Bedarf an neuen Methoden und Ansätzen.

Die Entwicklung von Biomarkern für die klinische Diagnostik ist noch immer sehr kostenintensiv (geschätzte \$100Mio oder mehr), da die Effizienz der diagnostischen Methode ähnlich eines

Arzneimittels evidenzmedizinisch evaluiert werden muss²⁸². Einige Unternehmen entwickeln deshalb bereits passende Biomarker parallel zu dem jeweiligen Arzneimittelevaluierungsprozess und nutzen so die Synergieeffekte zum Zeitpunkt der Einführung des Medikaments²⁸⁷. Zusätzlich könnte das Screening auf multiple Biomarker auf einer Plattform oder einem Assay-Set die ausreichende Anzahl an Patientenproben sichern und so die Entwicklungseffizienz steigern²⁸⁸. Grundsätzlich sind die meisten Krebstherapien, die in unselektierten Patientenpopulationen evaluiert werden, weniger effektiv als solche aus Biomarker-selektierten Populationen²⁸², was den Bedarf an Subgruppendefinitionen oder Biomarker-Selektionen für neue Agenzien erneut unterstreicht.

Da die Standardisierung von Assays für die Biomarkerentwicklung essentiell ist²⁸², bieten sich automatisierte Microarrays als einfach zu reproduzierende Assays mit klinischen Nutzen an²⁷⁶. Gerade die Eingrenzung auf bestimmte, sorgfältig ausgewählte Gene könnte sich auch in der klinischen Routine etablieren. In Brustkrebs wurden Expressions-Microarrays mit reduzierten Gensets bereits für diagnostische Verfahren genutzt und werden derzeit als klinische Begleitdiagnostik etabliert²⁸⁹⁻²⁹². In Magenkrebs wurden NGS und immunhistochemische Ansätze kombiniert und als begleitendes, personalisiertes Therapiekonzept vorgeschlagen²⁹³.

Als einer der vielversprechendsten Ansätze für die Entwicklung klinisch relevanter, molekularer Subgruppen hat sich die Kombination von genomischen Next Generation Sequencing (NGS)-Methoden auf DNA-Ebene mit anderen, davon unabhängigen Methoden erwiesen (zusammengefasst in²⁹⁴). Expressionsprofiling des gesamten Genoms, das die Mutationsanalyse mittels RNA Sequenzierung (RNAseq) mit einbezieht, wird höchstwahrscheinlich die Limitierung der Microarray-Technologie überwinden und sich langfristig als erfolgreicher durchsetzen, ist jedoch derzeit noch zu teuer, um in multizentrischen Studien angewendet zu werden. Zudem fehlt die Reproduzierbarkeit auf Grund verschiedener Plattformen und den unterschiedlichen Versuchsabläufen der zuständigen Personen. Die große Menge an Daten, die bei jeglichen Hochdurchsatzsequenzierungsmethoden produziert wird, muss zudem akribisch nach den gleichen Abläufen prozessiert werden, um eine vergleichbare Interpretation zu ermöglichen. Microarrays mögen nach heutigen Maßstäben in der Grundlagenforschung nicht mehr dem Standard entsprechen, werden jedoch derzeit miniaturisiert und automatisiert als kostengünstige On-Chip Alternative für die klinische Diagnostik weiterentwickelt¹⁷¹. Dementsprechend bietet sich an, auch im onkologischen Bereich eine Liste von für die Diagnostik interessanten Genen auf ihre Expression in automatisierten – und damit standardisierten – Arrays parallel zu klinischen Studien zu untersuchen. Diese Listen könnten auf

Basis der in dieser Arbeit vorgestellten Methoden eingegrenzt werden und durch stichprobenartige Hochdurchsatzsequenzierung bestätigt werden. Durch die Limitierung auf wenige Gene wäre es möglich, klinische, evidenzmedizinisch validierte Ergebnisse mit Daten potentieller Biomarker statistisch relevant zu korrelieren und so eine gesundheitsökonomische Validierung im diagnostischen Zulassungsprozess deutlich zu vereinfachen.

Grundsätzlich ist ein tiefes molekulares Verständnis von elementarer Bedeutung für die Entwicklung neuer Therapieansätze und begleitender Diagnostika gleichermaßen. Je besser der dahinterstehende Mechanismus verstanden ist, desto stärker können die Methoden vereinfacht werden. Dabei ist ebenso grundsätzlich in Frage zu stellen, ob große empirisch-statistische Datenmengen zielführend sind oder lediglich Hinweise auf molekulare Zusammenhänge und Gemeinsamkeiten liefern können. Die großen Hochdurchsatzsequenzierungsprojekte des *Cancer Genome Atlas* und anderer Gruppen verdeutlichten eingehend die Komplexität der Krebserkrankungen schon auf genomischer (DNAseq), aber auch bereits durch RNAseq auf transkriptioneller Ebene ^{167,295-299}. Um die Komplexität der schnelllebigen Proteinebene und der Mikroumgebung von Tumoren mit einzubeziehen, wird es zukünftig notwendig sein, zusätzlich zur DNA- und RNA-Ebene auch das Proteom und das Sekretom mit in die Untersuchungen zu integrieren, was für einige Tumorarten bereits begonnen wurde ^{145,300-302}.

Gerade für die Betrachtung der Zusammenhänge innerhalb des Proteoms ist jedoch noch eine weitere Informationsebene entscheidend: die subzelluläre Lokalisation der jeweiligen Proteine, die bei einer rein quantitativen Betrachtung durch massenspektrometrische Untersuchungen verlorengeht ³⁰³. Denn damit Enzyme ihre Aktivität entfalten können, müssen sie mit ihren Substraten und Aktivatoren interagieren können - räumlich wie zeitlich aufeinander abgestimmt. Dabei kann die An- oder Abwesenheit jedes einzelnen Beteiligten der Reaktion reguliert werden und maßgebliche Vorgänge in der Zelle beeinflussen ^{129,130,304}.

Um Kontrolle über Kompartimentwechsel ausüben zu können, nutzen Zellen aktive Transportmechanismen, die Energie verbrauchen ³⁰⁴. Vor allem der nukleo-zytoplasmatische Transport ist essentiell für verschiedene Zellfunktionen und eingehend untersucht worden ^{128,129,305}. Die zugrundeliegenden Mechanismen, die über NLS- und NES-Signalsequenzen gesteuert werden, können auch dazu genutzt werden, um die Aktivität von Proteasen näher zu untersuchen. Hierfür wurde im Rahmen dieser Arbeit ein Biosensor-System für Hochdurchsatzscreenings verfeinert

(Kapitel 3.2.1), welches ursprünglich zur Evaluation von Protein-Protein-Interaktionen entworfen wurde³⁰⁶ und zur Untersuchung der proteolytischen Taspase-Aktivität weiterentwickelt wurde³⁰⁷. Da eine transdominante Hemmung der Taspase1 wie in Kapitel 3.2.2 gezeigt nicht möglich ist, ist die Suche nach chemischen Inhibitoren mit Hilfe des Translokationssensorsystems therapeutisch umso relevanter¹³⁶.

Durch die systematische Modifikation des Spaltmotivs innerhalb des Sensors konnte durch ein gezieltes Ausschlussverfahren die Konsensussequenz der *Drosophila* Taspase1 (dTaspase1) identifiziert werden (Kapitel 3.2.3). Mit der Konsensussequenz (Q³[F,I,L,M]²D¹↓G¹X²X³) konnte das Degradom der dTaspase1 *in silico* als signifikant gegenüber dem der humanen Taspase1 erweitert ermittelt werden¹²². Zusätzlich fehlt der dTaspase1 die funktionell wichtige nukleoläre Lokalisationssequenz (NoLS), die über die Interaktion mit Imp-α die nukleäre und über NPM1 die nukleoläre Lokalisation der Protease vermittelt¹³⁰. Zwar lokalisiert die dTaspase1 auch im Nukleus, akkumuliert jedoch nicht in den Nukleoli und wird wahrscheinlich durch passive Diffusion nur ineffizient in den Kern transportiert, was ihre nukleäre und insbesondere nukleoläre Aktivität einschränkt. Das NoLS ist dabei in Vertebraten hochkonserviert. Die Imp-α/NPM1-Achse, mit der die humane Taspase1 über das NoLS aktiv zu den Nukleoli transportiert wird, hat sich also vermutlich in der Evolution von Invertebraten zu Vertebraten als ein Transport-basierter Mechanismus zur Feinregulation entwickelt.

Für Proteasen ist diese Form der Regulation durch intrazelluläre Kompartimentwechsel besonders kritisch, da ihre Aktivität in der Regel irreversibel ist und weitreichende Veränderungen in der Zelle hervorrufen kann, die zu Erkrankungen wie Krebs führen können^{89,108}. Wie entscheidend ein Kompartimentwechsel für die Aktivität eines Proteins sein kann, zeigte sich auch in Kapitel 3.2.4, bei der Untersuchung des Transkriptionsfaktors TFIIA. Da die Aktivität von Transkriptionsfaktoren auf den Nukleus beschränkt ist, resultiert ein Kernexport in einer Inaktivierung, so dass hier eine Regulationsmöglichkeit über die Zeit der Anwesenheit des Transkriptionsfaktors im Nukleus besteht³⁰⁸. Wie in Kapitel 3.2.4 gezeigt werden konnte, wird TFIIA durch das N-terminale NES über das Exportin-1/Crm1-System nach der Proteinsynthese im Zytoplasma gehalten und akkumuliert erst nach der proteolytischen Prozessierung durch Taspase1 im Zellkern. TFIIA verdrängt die inhibitorische TAND2-Domäne von TAF1 und stabilisiert so den TBP-DNA Komplex^{15,18}. Allerdings wird seine Bindeaffinität zum TBP-DNA Komplex nach der Prozessierung durch Taspase1 verringert. Die Langzeit-Expression von TFIIA führte unabhängig von einer Spaltung zu einer nukleären

Akkumulation, was möglicherweise auf die Bildung der PICs und der Bindeaffinität zum TBP-DNA Komplex zurückzuführen sein könnte. Die Hauptauswirkung der Proteolyse von TFIIA macht sich letztendlich in der resultierenden Veränderung der Transkription bemerkbar und konnte an der Beeinflussung der Transkription von p16^{INK4a} veranschaulicht werden. Wird die Prozessierung verhindert, ist die Expression dieses zellzyklusinhibitorischen Tumorsuppressors höher, während sie nach der Spaltung durch Taspase1 sinkt, was die Funktion der Taspase1 als *Non-Oncogene-Addiction-Protease* unterstreicht.

Hinsichtlich der weiteren Erforschung von krebisrelevanten Prozessen und daraus resultierenden Zielen für neue medizinisch wirksame Substanzen oder Biomarker ist dabei eine evidenzbasierte Evaluation in Kombination mit klinischen und histopathologischen Daten unabdingbar. Dieser translationale biomedizinische Ansatz mit dem Ziel, Daten aus der klinischen Routine und molekularbiologischer Forschung miteinander zu vereinen, ist ein relativ neues Feld für die Bioinformatik ³⁰⁹. Viele Kliniken sind schon vor längerer Zeit dazu übergegangen, die medizinischen Daten ihrer Patienten in elektronischen Systemen zu dokumentieren ³¹⁰. Die Nutzung dieser Daten für wissenschaftliche Zwecke ist jedoch häufig durch verschiedene Umstände erschwert. Zum einen sind die medizinischen Dokumentationssysteme nicht dazu ausgelegt, nach bestimmten Indikationen zu filtern und entsprechende detaillierte Daten zur weiteren Auswertung aus dem System extrahieren zu können. Weiterhin sind zahlreiche regulatorische Voraussetzungen zu beachten, wie der Datenschutz und verschiedene Datenstandards, organisatorische Hürden und die einfache Tatsache, dass oftmals eine Vielzahl an Personen diese Daten generieren und somit eine Vergleichbarkeit nicht automatisch gewährleistet ist. Die Systemarchitektur eines wissenschaftlich zu nutzenden Systems, das klinische mit histopathologischen und forschungsbezogenen Daten vereint und auswertbar macht, muss also dazu ausgelegt sein, nach verschiedenen Kriterien filtern und sortieren zu können, verwertbare Strukturen zu erzeugen und so statistische Korrelationen möglich zu machen. Dafür müssen Daten in erster Linie gut aufbereitet werden und nach immer gleichen Standards eingegeben werden, so dass keine personenbezogene Verzerrung der Daten (*Bias*) entsteht. Die Übertragung retrospektiver Daten, die zuvor auf Papier festgehalten wurden, in elektronische Form ist nicht ohne Risiko und hat in der Vergangenheit bereits zu teuren Fehlern geführt ³¹¹. Viele dieser Herausforderungen können über geschickte Eingabemasken mit festgelegten Auswahlmöglichkeiten, einer guten Schulung und Absprache der eingebenden medizinischen

Mitarbeiter und klar festgelegten Standards abgefangen werden, so dass eine hohe Datenqualität gewährleistet werden kann.

Besonders wertvoll sind die Biobanken in pathologischen Instituten, in denen routinemäßig Tumormaterial von Patienten untersucht und anschließend zur medizinischen Dokumentation gelagert wird ^{309,312,313}. Dieses Material retrospektiv mit evidenzmedizinischer Relevanz wissenschaftlich zu untersuchen, ist Ziel des TUBA-Projektes, das in Kapitel 3.3.1 beschrieben ist. Die dargestellten Daten sind mit Vorsicht zu interpretieren, da sie in der unselektierten Form keine Aussagen zur Prävalenz oder Inzidenz zulassen. Ziel ist es im ersten Schritt, für alle Oropharynx-Plattenepithelkarzinome aus 2008 bis 2012 retrospektiv den HPV-Status und die Expression des Zellzyklusregulators p16^{INK4a} zu bestimmen und so eine epidemiologische Grundlage für weitere Auswertungen zu gewinnen. Im weiteren Verlauf werden immunhistochemische Färbungen gegen Taspase1, verschiedene RTKs (EGFR, ERBB2, ERBB3, ERBB4, MET), Mcl-1, Cyclin D, Cyclin E2, RB1 und HNF4A durchgeführt und auf prognostische und prädiktive Relevanz überprüft. Diese immunhistochemischen Färbungen werden anschließend mit der Software HALO ausgewertet, die über verschiedene Funktionen in der Lage ist, Tumor- von Stromagewebe zu unterscheiden, die Lokalisation und Färbeintensität zu erkennen und entsprechende Zahlen für die TMAs standardisiert auszugeben. Der prozentuale Anteil positiver Zellen gewichtet nach verschiedenen Färbeintensitäten geht dabei in die Berechnung des *TMA Scores S* ein, der nach der Formel für jede subzelluläre Lokalisation berechnet wird:

$$S = (1 * \%Z_n + 2 * \%Z_m + 3 * \%Z_h)$$

Mit

$\%Z_n$ = Anteil niedrig exprimierender Zellen im Tumorgewebe

$\%Z_m$ = Anteil mittel exprimierender Zellen im Tumorgewebe

$\%Z_h$ = Anteil hoch exprimierender Zellen im Tumorgewebe

Die errechneten Werte zwischen 0 und 300 sind so für jede Sublokalisierung vergleichbar und werden direkt durch die TUBA-Datenbank errechnet.

Die immunhistochemischen Auswertungen erfolgen auf Basis dieses in die Systemarchitektur implementierten *TMA Scores S*, der den prozentualen Anteil positiver Zellen gewichtet mit der

Färbeintensität für die Sublokalisationen (Kern, Zytoplasma, Membran) statistisch vergleichbar macht. Zukünftig soll es so ermöglicht werden, über Korrelationsanalysen Hinweise auf potentielle pathobiologische Zusammenhänge zu erhalten, welche Taspase1 in das Signal-Netzwerk der Tumorzellen einordnen lassen und die Grundlage für weiterführende Analysen proteolytischer Netzwerke bieten.

Zusätzlich ist TUBA eine ideale Grundlage für neue bioinformatische Ansätze, die in der Lage sind, auch große Datensätze auszuwerten. Während eine kostengünstige Möglichkeit, statistische Signifikanz zu erzeugen, wie oben beschrieben in der Reduktion der betrachteten Parameter liegt, liegt die Zukunft der Präzisionsmedizin vermutlich in der Analyse von *Big Data*. Eine der Hauptschwierigkeiten bei der Analyse großer Datenmengen ist, dass Korrelationen bei einer Masse an Daten auch immer zufällig auftreten werden, egal welcher Parameter analysiert wird³¹⁴. Diesen zufällig auftretenden Korrelationen wird statistisch mit der *False Discovery Rate* (FDR)³¹⁴ oder der gewichteten FDR, die speziell für klinische Studien entwickelt wurde, Rechnung getragen³¹⁵. Diese bezieht jedoch nicht die biologischen Abhängigkeiten und synergistischen Effekte biologischer Systeme mit ein. Letztendlich ist auch die Bestimmung des p-Werts für jeden einzelnen Marker für biologische Systeme nur mit Vorsicht zu interpretieren³¹². Sinnvoller, jedoch kostenintensiver als die Reduktion der betrachteten Werte, wäre die Nutzung selbstlernender bioinformatischer Systeme, die in der Lage sind, bereits bestehendes molekularbiologisches Wissen aus zellulären Netzwerken und Zusammenhängen wie der subzellulären Lokalisation mit einzukalkulieren und so Abhängigkeiten zu erkennen und entsprechend zu gewichten^{316,317}. Für solche Fragestellungen wurde ein rückwärtiger Eliminierungsalgorithmus vorgeschlagen und bereits für die Vorhersage von Fernmetastasen in Brustkrebs genutzt³¹⁸.

Die Wichtigkeit der Betrachtung der subzellulären Lokalisation eines Proteins, wie ebenfalls bereits oben erwähnt, wurde durch die dargestellten Mechanismen und wertvolle Beispiele für die Feinregulationen durch Taspase1 unterstrichen. Durch seit langem in der klinischen Routine verwendete immunhistochemische Färbungen ist es möglich, nicht nur die Gewebestruktur aus Tumor und Stroma mit in die Untersuchung mit einzubeziehen, sondern vor allem auch die subzelluläre Lokalisation der betrachteten Proteine, die durch Verwendung einer Software speziell für pathologische Institute standardisiert und somit reproduzierbar wird. TUBA bietet durch ihre individuelle Systemarchitektur also die ideale Grundlage für weiterführende klinische, molekularbiologische und bioinformatische Fragestellungen wie beispielsweise der Untersuchung

proteolytischer Netzwerke um Taspase1 in Kopf-Hals Tumoren, soll aber sicherlich nicht darauf beschränkt bleiben. Andere tumor-relevante Proteasen, aber auch weitere feinregulierte Transkriptionsfaktoren wären nächste, interessante Ziele zur subzellulären Beobachtung und statistischen Auswertung. Darüber hinaus eröffnet eine fortlaufende Eingabe und Dokumentation prospektiver Daten weiterführend die Möglichkeit, kontinuierlich hochaktuelle Forschungsstände zu ermitteln und zu veröffentlichen.

5 MATERIAL UND METHODEN

5.1 Material

5.1.1 Allgemeine Verbrauchsmaterialien

Alle Verbrauchsmaterialien wurden bei den Firmen Sigma Aldrich (Steinheim), Gibco/Invitrogen GmbH (Karlsruhe), Fisher-Scientific GmbH (Schwerte), Greiner Bio-One GmbH (Frickenhausen) und BD Bioscience Europe (Heidelberg) bezogen, sofern nicht anders angegeben. Zellkulturschälchen mit Glasbodeneinsatz für die Fluoreszenzmikroskopie stammten von der Firma MatTek (Ashland, MA, USA).

5.1.2 Chemikalien und Reagenzien

Tab. 12: Verwendete Chemikalien und Reagenzien

Bezeichnung	Hersteller	Ort
3,3'-Diaminobenzidin	Dako GmbH	Glostrup
Acrylamid (Rotiphorese 30)	Carl Roth GmbH & Co. KG	Karlsruhe
Agarose	Invitrogen GmbH	Karlsruhe
Ammoniumpersulfat	AppliChem GmbH	Darmstadt
Ampicilin	AppliChem GmbH	Darmstadt
Bioterm Taq-DNA-Polymerase	Genecraft GmbH	Lüdinghausen
Bovines Serumalbumin (BSA 100x)	New England Biolabs Inc.	Ipswich
Bovines Serumalbumin (BSA)	AppliChem GmbH	Darmstadt
Calf Intestinal Phosphatase (CIP)	New England Biolabs Inc.	Ipswich
CHAPS	Carl Roth GmbH & Co. KG	Karlsruhe
Coomassie-Brilliant Blau	Carl Roth GmbH & Co. KG	Karlsruhe
Dimethylsulfoxid	Sigma Aldrich Chemie GmbH	München
ECL	Amersham Bioscience	Buckinghamshire

MATERIAL UND METHODEN

Bezeichnung	Hersteller	Ort
EGTA	Carl Roth GmbH & Co. KG	Karlsruhe
Enthelan	Merck KG	Darmstadt
Entwickler-Lösung	Imagine Medical	Sierra Blanca
Essigsäure, 100%	Carl Roth GmbH & Co. KG	Karlsruhe
Ethanol 100% (absolut)	AppliChem GmbH	Darmstadt
Ethanol 99% (vergällt mit 1% MEK)	AppliChem GmbH	Darmstadt
Ethidiumbromid	Sigma Aldrich Chemie GmbH	Karlsruhe
Ethylendiamintetracetat	Sigma Aldrich Chemie GmbH	München
Fetales Kälberserum (FCS)	Gibco, Invitrogen GmbH	Karlsruhe
Fixierer	Imagine Medical	Sierra Blanca
Giemsa	Sigma Aldrich Chemie GmbH	Darmstadt
Glycin	AppliChem GmbH	Darmstadt
Größenstandard (DNA, 100 bp)	New England Biolabs Inc.	Ipswich
Größenstandard (DNA, 1 kb)	New England Biolabs Inc.	Ipswich
Größenstandard (Protein)	Peqlab Biotechnologie GmbH	Erlangen
HCl, rauchend	Merck KG	Darmstadt
Hoechst 33342	Sigma Aldrich Chemie GmbH	München
Hydroxurea	Sigma Aldrich Chemie GmbH	München
Isopropanol	Fisher Scientific GmbH	Schwerte
KCl	Merck KG	Darmstadt
KH₂PO₄	AppliChem GmbH	Darmstadt
Leptomycin B	Axxora Deutschland GmbH	Lörrach
L-Glutamin	Gibco, Invitrogen GmbH	Karlsruhe
Lipofectamine 2000	Invitrogen GmbH	Karlsruhe
Magermilchpulver	AppliChem GmbH	Darmstadt
Mercaptoethanol	Carl Roth GmbH & Co. KG	Karlsruhe
Methanol ≥99%	Carl Roth GmbH & Co. KG	Karlsruhe
Na₂HPO₄	Carl Roth GmbH & Co. KG	Karlsruhe
NaCl	Carl Roth GmbH & Co. KG	Karlsruhe
Natriumazid	Merck KG	Darmstadt
Natriumcitrat	Carl Roth GmbH & Co. KG	Karlsruhe
Natriumdeoxycholat	Sigma Aldrich Chemie GmbH	München
Nucleotide-Mix	New England Biolabs Inc.	Ipswich
Paraformaldehyd	Serva Elektrophoresis GmbH	Heidelberg

Bezeichnung	Hersteller	Ort
PCR-Puffer, 10x	Genecraft GmbH	Lüdinghausen
Penicillin/Streptomycin	Gibco, Invitrogen GmbH	Karlsruhe
Phenylmethylsulfonyl Fluorid	Carl Roth GmbH & Co. KG	Karlsruhe
Phosphate buffered saline (PBS), steril	Gibco, Invitrogen GmbH	Karlsruhe
Polyethylenimin (PEI)	Sigma Aldrich Chemie GmbH	München
Ponceau S	Sigma Aldrich Chemie GmbH	München
Protease-Inhibitor-Mix (Complete 50x)	Roche Diagnostics GmbH	Mannheim
Proteinase K	Dako	Kopenhagen
ReBlotplus strong (10x)	Chemicon International	Hofheim
Sodiumdodecylsulfat	Carl Roth GmbH & Co. KG	Karlsruhe
Takara	Cambrex Bioscience	Lonza
Tetramethylethyldiamin (TEMED)	Serva Elektrophoresis GmbH	Heidelberg
Tris Xtrapure	Neolab	Heidelberg
Triton-X-100	Sigma Aldrich Chemie GmbH	München
Trypsin/EDTA	Gibco, Invitrogen GmbH	Karlsruhe
Tween 20	AppliChem GmbH	Darmstadt
Xylol	Merck KG	Darmstadt
Zitronensäure	Carl Roth GmbH & Co. KG	Karlsruhe

5.1.3 Puffer und Lösungen

Tab. 13: Verwendete Puffer und Lösungen

Bezeichnung	Zusammensetzung
6x Lämmli-Ladepuffer	60 mM Tris 2% SDS 10% Glycerin 5% β -Mercaptoethanol 0,01% Bromphenolblau <i>ad</i> H ₂ O, pH 6,8
Blockierungspuffer	5% Magermilchpulver in PBS 0,001% NaN ₃

MATERIAL UND METHODEN

Bezeichnung	Zusammensetzung
Coomassie-Gelfärbelösung	0,08% Coomassie Brilliant Blue 1,6% Phosphorsäure 8% Ammoniumsulfat 20% Methanol <i>ad</i> H ₂ O
Coomassie-Gelentfärbelösung	7% Eisessig 20% Methanol <i>ad</i> H ₂ O
DNA-Probenpuffer 5x (Bromphenolblau)	20% (w/v) Glycerin 100 mM EDTA 0,25% (w/v) Bromphenolblau <i>ad</i> H ₂ O
DNA-Probenpuffer 5x (Xylencyanol)	20% (w/v) Glycerin 100 mM EDTA 0,25% (w/v) Xylencyanol <i>ad</i> H ₂ O
Dulbecco's Phosphate Buffered Saline (DPBS)	8,1 mM Na ₂ HPO ₄ <i>ad</i> PBS, pH 7,4
EDTA-Lösung	0,5 M EDTA <i>ad</i> H ₂ O, pH 8,0
Ethylenediaminetetraacetic Acid (EDTA)	0,5 M Ethylenediaminetetraacetic Acid <i>ad</i> H ₂ O, pH 8,0
GST-Elutionspuffer	150 mM NaCl 50 mM Tris pH 8,0 30 mM Glutathion <i>ad</i> H ₂ O
IP-Lysepuffer	25 mM Tris 150 mM NaCl 1% NP-40 1% SDS 1x Proteaseinhibitor-Mix 1% NaDeoxycholat <i>ad</i> H ₂ O, pH 7,0

MATERIAL UND METHODEN

Bezeichnung	Zusammensetzung
Lyse-Puffer (humanes Tumorgewebe)	10 mM Tris-HCl 1 mM MgCl ₂ 1 mM EGTA 5 mM Mercaptoethanol 0,1 mM Phenylmethylsulfonyl Fluorid 0,5% CHAPS 10% Glycerol (v/v) Protease-Inhibitor Complete ad H ₂ O, pH 7,5
MES-Puffer (Ni-NTA Regeneration)	20 mM 2-(N-morpholine)ethanesulfonic Acid (MES) 100 mM NaCl ad H ₂ O, pH 5,0
PBS mit Tween (PBS-T)	0,1% Tween in PBS
Phosphate buffered saline (PBS), unsteril	4,3 mM Na ₂ HPO ₄ 1,4 mM KH ₂ PO ₄ 137 mM NaCl 2,7 mM KCl ad H ₂ O, pH 7,4
RIPA-Lysepuffer	25 mM Tris 150 mM NaCl 1% NP-40 0,1% SDS 1x Proteaseinhibitor-Mix 1% NaDeoxycholat ad H ₂ O, pH 7,0
SDS-Laufpuffer	25 mM Tris 250 mM Glycin 0,1% SDS ad H ₂ O
TAE-Puffer	40 mM Tris 1 mM EDTA ad H ₂ O, pH 8,0 (mit Essigsäure)

Bezeichnung	Zusammensetzung
Transferpuffer	25 mM Tris
	192 mM Glycin
	20% Methanol

5.1.4 Geräte

Tab. 14: Verwendete Geräte

Gerät	Bezeichnung	Hersteller	Ort
Autoklav	5050 ELV	Tuttnauer	Breda
Bakterienbrutschrank	Gasboy C40	Labotec	Midrand
Bioanalyser (zur DNA/RNA Analyse)	Bioanalyser 2100	Agilent Technologies	Böblingen
Einfrierhilfe	Mr. Frosty	Nalgene Labware	Neerijse
Elektrophoresekammer (DNA)	PerfectBlue Gelsystem Mini S	Peqlab Biotechnologie GmbH	Erlangen
Elektrophoresekammer (Protein)	Elektrophoresesystem	Bio-Rad Laboratories GmbH	München
Entwicklermaschine	Curix 60	Agfa Healthcare GmbH	Köln
Entwicklermaschine	Chemidoc	Bio-Rad Laboratories GmbH	München
Falkonzentrifuge	Universal 16R	Andreas Hettich GmbH & Co. KG	Tuttlingen
Gefrierschrank (-80°)	Hera freeze	Heraeus Holding GmbH	Hanau
Geldokumentation DNA	Multigenius Bioimagine	Syngene	Cambridge
Gelgießstand	Gießschiene PerfectBlue Mini S	Peqlab Biotechnologie GmbH	Erlangen
Heizblock	Htermomoxer comfort	Eppendorff GmbH	Wesseling
Homogenisierer	Ultra-Turrax 25	IKA Labortechnik	Staufen
Kühlzentrifuge	Biofuge firesco	Heaeus Holding GmbH	Hanau
Magnetrührer	Ikamag RCT	Ika	Staufen
Mikroskop (Fluoreszenz)	Nikon Eclipse TE2000-U AxioVert 200M	Nikon GmbH, Zeiss	Düsseldorf, Jena
Mikroskop (Zellkultur)	Nikon TMS	Nikon GmbH	Düsseldorf

MATERIAL UND METHODEN

Gerät	Bezeichnung	Hersteller	Ort
Mini-Tischzentrifuge	C13000-RT	Fisherbrand	Schwerte
PCR-Thermocycler	Tpersonal	Biometra biomedizinische Analytik GmbH	Göttingen
PCR-Thermocycler	ThermalUno	VWR	Darmstadt
Rollmixer	RollerMixer SRT6	Stuart Scientific	Asbach
Schüttelinkubator	Certomat H	B.Braun Biotech GmbH	Melsungen
Schüttler	WS10	Edmund Bühler GmbH	Hechingen
Sterilbank	Hera Safe	Heraeus Holding GmbH	Hanau
Stickstoff-Tank	GT 38	Omnilab AG	Mettmenstetten
Stromgeber (Elektrophorese)	Power Pac	Bio-Rad Laboratories GmbH	München
Tank-Blotter	PerfectBlue Tank Elektro Blotter Web S	Peqlab Biotechnologie GmbH	Erlangen
Tischzentrifuge	Mikro 120	Andreas Hettich GmbH & Co. KG	Tuttlingen
Trockenschrank	Function Line	Heraeus Holding GmbH	Hanau
UV-Tisch	TFX-351	Biometra biomedizinische Analytik GmbH	Göttingen
Vortexer	VWR VV3	VWR Int. GmbH	Darmstadt
Waage	ABT 120-SDM	Kern & Sohn GmbH	Balingen
Wasserbad	Polystat CC3	P.Huber Kältemaschinenbau GmbH	Offenburg
Zellkulturbrutschrank	Hera Cell	Heraeus Holding GmbH	Hanau
Zentrifuge	Avanti JE-30-I	Beckman Coulter GmbH	Krefeld

5.1.5 Antikörper

Tab. 15: Verwendete Antikörper

Antigen	Organismus	Eigenschaft	Hersteller	Verwendung		
				Western Blot	Immunfluoreszenz	Immunhistochemie
Flag-tag	Maus	monoklonal	Sigma	1:1000 ü.N. 4°C	1:200 1 h RT	

MATERIAL UND METHODEN

Antigen	Organismus	Eigenschaft	Hersteller	Verwendung		
				Western Blot	Immun- fluoreszenz	Immun- histochemie
HA-tag	Maus	monoklonal	Covance	1:1000 ü.N. 4°C	1:300 1 h RT	
GAPDH (sc-47724)	Maus	monoklonal	Santa Cruz	1:3000 1 h RT		
GFP (FL)	Kaninchen	polyklonal	Santa Cruz	1:5000 ü.N. 4°C		
GST (3D4)	Maus	monoklonal	Santa Cruz	1:1000 ü.N. 4°C		
KRT10 (ab76318)	Kaninchen	monoklonal	Abcam	1:10.000 ü.N. 4°C	1:150 1 h RT	1:2500 2 h RT
KRT24 (HPA022978)	Kaninchen	polyklonal	Sigma	1:1000 ü.N. 4°C	1:150 1 h RT	1:250 2 h RT
mCherry	Ziege	polyklonal	Antibodies online	1:1000 ü.N. 4°C		
Taspase1 (D-20)	Kaninchen	polyklonal	Santa Cruz	1:3000 ü.N. 4°C		
Taspase1 (C-term)	Kaninchen	polyklonal	Abgent	1:1000 ü.N. 4°C		
TF2A-α (T20)	Ziege	polyklonal	Santa Cruz	1:500 ü.N. 4°C		
TF2A-β (C-15)	Ziege	polyklonal	Santa Cruz	1:500 ü.N. 4°C		
TF2A-β (A15)	Ziege	polyklonal	Santa Cruz	1:500 ü.N. 4°C		
Vimentin (MS-129-P)	Maus	monoklonal	Thermo Fisher	1:1000 ü.N. 4°C		1:250 2 h RT

Zudem wurde ein individuell hergestellter KRT24-Antikörper verwendet, der polyklonal in Kaninchen hergestellt wurde (Pineda, Berlin). Folgende KRT24-Peptide wurden zur Immunisierung der Kaninchen verwendet: Aminosäuren $^3\text{NH}_2\text{-CSSRASSSRAGGSSAR-COOH}^{19}$ und $^{486}\text{NH}_2\text{-CSGQGRDSSKTRVTK-COOH}^{500}$. Drei der 4 gelieferten Immunseri wurden positiv getestet, im

Folgenden gepoolt und anschließend im Western Blot bei 1:1500 (ü.N. 4°C), in der IHC bei 1:250 (2 h RT) verwendet.

Tab. 16: Verwendete Sekundärantikörper

Antigen	Organismus	Konjugiert mit	Hersteller	Verwendung		
				Western Blot	Immun- fluoreszenz	Immun- histochemie
α-Kaninchen	Ziege	Cy3	Dianova		1:400 1 h RT	
α-Maus	Ziege	Cy3	Dianova		1:400 1 h RT	
α-Ziege	Esel	Cy3	Mobitec		1:300 1 h RT	
α-Maus	Ziege	HRP	Santa Cruz	1:5000 1 h RT		1:500 1 h RT
α-Kaninchen	Ziege	HRP	Santa Cruz	1:5000 1 h RT		1:500 1 h RT
α-Ziege	Esel	HRP	Santa Cruz	1:10.000 1 h RT		1:500 1 h RT

Für die Immunhistochemische Färbung mit Primärantikörpern aus Maus oder Kaninchen wurde zudem das *Envision-Kit* von Dako-Cytomation (Dänemark) verwendet.

5.1.6 Oligonukleotide

Alle verwendeten Oligonukleotide wurden von den Firmen Eurofins MWG Operon (Frankfurt) und Sigma-Aldrich (Staufen) synthetisiert und bezogen.

Tab. 17: Primer zur Sequenzanalyse

Bezeichnung	Sequenz
5'-pc3_seq	5'-GGAGGTCTATATAAGCAGAGCTC-3'

Bezeichnung	Sequenz
3'-pc3_seq	5'-CAACTAGAAGGCACAGTCGAG-3'
3'-GST_seq	5'-GCCTTTGCAGGGCTGGCAAG-3'
5'-GFP_seq	5'-GTGGCATCGCCCTCGCCCTC-3'
3'-GFP_seq	5'-GTGCTGCTGCCCCGACAACCAC-3'
5'-mCherry_seq	5'-CTCCATGTGCACCTTGAAGC-3'
3'-mCherry_seq	5'-CACCATCGTGGAACAGTACG-3'
3'-SP6_seq	5'-ATTTAGGTGACACTATAG-3'
5'-T7_seq	5'-TAATACGACTCACTATAG-3'
3'-Tasp1_seq	5'-GAGCAGTTGGAGCACTGAGTG-3'
3'-TF2A_seq	5'-CAAACATGAGTGCTGCTGCT-3'

Doppelsträngige Oligonukleotide, die für eine MLL-Schnittstelle bzw. eine mutierte MLL-Schnittstelle kodieren und mittels Überhängen der Schnittstellen NotI und XhoI in den entsprechend aufgeschnittenen pcDNA 3.1-Biovektor (Invitrogen) kloniert wurden, wurden durch Oligoannealing der folgenden Sequenzen hergestellt:

Tab. 18: Primer für Oligoannealing

Bezeichnung	Sequenz
Cl2 ⁺ -NotI_sense	5'-GGCCGCAAAAATCAGCCAGCTTGATGGTGTGATGATGGCTCTGGCAGCGC-3'
Cl2 ⁺ -XhoI_antisense	5'-TCGAGCGCTGCCAGAGCCATCATCAACACCATCAAGCTGGCTGATTTTTGC-3'
Cl2 ⁺ mut-NotI_sense	5'-GGCCGCACAGCTTGCTGCAGTTGATGATGGCGC-3'
Cl2 ⁺ mut-XhoI_antisense	5'-TCGAGCGCCATCATCAACTGCAGCAAGCTGTGC-3'
CS1-NotI_sense	5'-GGCCGCAGCTGAAGGACAGGTGGATGGGGCCGATGACTTAAGCACTTCAGC-3'
CS1-XhoI_antisense	5'-TCGAGCTGAAGTGCTTAAGTCATCGGCCCATCCACCTGTCCTTCAGCTGC-3'
Cl2 ⁺ _K1A_sense	5'-GGCCGCAGCAATCAGCCAGCTTGATGGTGTGATGATGGCTCTGGCAGCGC-3'
Cl2 ⁺ _K1A_antisense	5'-TCGAGCGCTGCCAGAGCCATCATCAACACCATCAAGCTGGCTGATTGCTGC-3'
Cl2 ⁺ _I2A_sense	5'-GGCCGCAAAAGCCAGCCAGCTTGATGGTGTGATGATGGCTCTGGCAGCGC-3'
Cl2 ⁺ _I2A_antisense	5'-TCGAGCGCTGCCAGAGCCATCATCAACACCATCAAGCTGGCTGGCTTTTTGC-3'
Cl2 ⁺ _S3A_sense	5'-GGCCGCAAAAATCGCCAGCTTGATGGTGTGATGATGGCTCTGGCAGCGC-3'
Cl2 ⁺ _S3A_antisense	5'-TCGAGCGCTGCCAGAGCCATCATCAACACCATCAAGCTGGGCGATTTTTGC-3'
Cl2 ⁺ _Q4A_sense	5'-GGCCGCAAAAATCAGCGCGCTTGATGGTGTGATGATGGCTCTGGCAGCGC-3'
Cl2 ⁺ _Q4A_antisense	5'-TCGAGCGCTGCCAGAGCCATCATCAACACCATCAAGCGCGCTGATTTTTGC-3'
Cl2 ⁺ _Q4N_sense	5'-GGCCGCTAAGATCAGCAATCTTGATGGTGTGATGATGGCTCTGGCAGCGC-3'

MATERIAL UND METHODEN

Bezeichnung	Sequenz
Cl2 ⁺ _Q4N_antisense	5'-TCGAGCGCTGCCAGAGCCATCATCAACACCATCAAGATTGCTGATCTTAGC-3'
Cl2 ⁺ _L5A_sense	5'-GGCCGCAAAAATCAGCCAGGCTGATGGTGTGATGATGGCTCTGGCAGCGC-3'
Cl2 ⁺ _L5A_antisense	5'-TCGAGCGCTGCCAGAGCCATCATCAACACCATCAGCTGGCTGATTTTTGC-3'
Cl2 ⁺ _L5F_sense	5'-GGCCGCAAAAATCAGCCAGTTCGATGGTGTGATGATGGCTCTGGCAGCGC-3'
Cl2 ⁺ _L5F_antisense	5'-TCGAGCGCTGCCAGAGCCATCATCAACACCATCGAACTGGCTGATTTTTGC-3'
Cl2 ⁺ _L5I_sense	5'-GGCCGCTAAGATCAGCCAGATCGATGGTGTGATGATGGCTCTGGCAGCGC-3'
Cl2 ⁺ _L5I_antisense	5'-TCGAGCGCTGCCAGAGCCATCATCAACACCATCGATCTGGCTGATCTTAGC-3'
Cl2 ⁺ _L5Q_sense	5'-GGCCGCAAAAATCAGCCAGCAAGATGGTGTGATGATGGCTCTGGCAGCGC-3'
Cl2 ⁺ _L5Q_antisense	5'-TCGAGCGCTGCCAGAGCCATCATCAACACCATCTTGCTGGCTGATTTTTGC-3'
Cl2 ⁺ _L5A_sense	5'-GGCCGCAAAAATCAGCCAGGCTGATGGTGTGATGATGGCTCTGGCAGCGC-3'
Cl2 ⁺ _L5A_antisense	5'-TCGAGCGCTGCCAGAGCCATCATCAACACCATCAGCTGGCTGATTTTTGC-3'
Cl2 ⁺ _D6A_sense	5'-GGCCGCAAAAATCAGCCAGCTTGCTGGTGTGATGATGGCTCTGGCAGCGC-3'
Cl2 ⁺ _D6A_antisense	5'-TCGAGCGCTGCCAGAGCCATCATCAACACCAAGCTGGCTGATTTTTGC-3'
Cl2 ⁺ _G7A_sense	5'-GGCCGCAAAAATCAGCCAGCTTGATGCTGTTGATGATGGCTCTGGCAGCGC-3'
Cl2 ⁺ _G7A_antisense	5'-TCGAGCGCTGCCAGAGCCATCATCAACAGCATCAAGCTGGCTGATTTTTGC-3'
Cl2 ⁺ _V8A_sense	5'-GGCCGCAAAAATCAGCCAGCTTGATGGTGTGATGATGGCTCTGGCAGCGC-3'
Cl2 ⁺ _V8A_antisense	5'-TCGAGCGCTGCCAGAGCCATCATCAGACCATCAAGCTGGCTGATTTTTGC-3'
Cl2 ⁺ _D9A_sense	5'-GGCCGCAAAAATCAGCCAGCTTGATGGTGTGCTGATGGCTCTGGCAGCGC-3'
Cl2 ⁺ _D9A_antisense	5'-TCGAGCGCTGCCAGAGCCATCAGCAACACCATCAAGCTGGCTGATTTTTGC-3'
Cl2 ⁺ _D10A_sense	5'-GGCCGCAAAAATCAGCCAGCTTGATGGTGTGATGCTGGCTCTGGCAGCGC-3'
Cl2 ⁺ _D10A_antisense	5'-TCGAGCGCTGCCAGAGCCAGCATCAACACCATCAAGCTGGCTGATTTTTGC-3'

Tab. 19: Primer zur quantitativen Analyse via RT-PCR

Bezeichnung	Sequenz
5'-OSF-2	5'-ATTAGGCTTGGCATCTGCTC-3'
3'-OSF-2	5'-CTCGCGGAATATGTGAATCG-3'
5'-KRT24	5'-TGACACAGAAGCTGGCTACG-3'
3'-KRT24	5'-CCTTGACCAGAACAGCTTCC-3'
5'-CLCA4	5'-GGGTAAGGACCGCCTAAATC-3'
3'-CLCA4	5'-TATCCTCCCATCAGTCAGC-3'
5'-FN 1	5'-GCCATGACAATGGTGTGAAC-3'
3'-FN 1	5'-ATACTGGTTGTAGGACTGGC-3'
5'-PRR4	5'-GCTCTCAGTGGTCCTTCTGG-3'

MATERIAL UND METHODEN

Bezeichnung	Sequenz
3'-PRR4	5'-CGGGGTAGAGAGAGTTGACG-3'
5'-TOSO	5'-TGTGGTACCGTGGTATCCACC-3'
3'-TOSO	5'-GGAGTCTCAGGCATTGGCTG-3'
5'-RASGRP2	5'-GTACTGGATCTCCGCCTTCC-3'
3'-RASGRP2	5'-TAGGTGAGATGCTCCGCCAG-3'
5'-C4.4A	5'-CAATGCGCTCAGGATCGCTG-3'
3'-C4.4A	5'-GTTACTCCATCCCGAGTGCAG-3'
5'-SCCA2b	5'-GGATGTACAGGCCAAGGTCC-3'
3'-SCCA2b	5'-CATGCCTGAGAGGTCTGCATC-3'
5'-TP73L	5'-TCCTTCCAGCAGTCGAGCAC-3'
3'-TP73L	5'-CATACTGGGCATGGCTGTTCC-3'
5'-ARHGAP25	5'-AAGCTCTGTAGGCTGGGATG-3'
3'-ARHGAP25	5'-GAGGAAGGCGACCAGAATTC-3'

Tab. 20: Primer zur PCR-Synthese

Bezeichnung	Sequenz
5'-bam-KRT24	5'-TTTGGATCCACATGTCTTGCTCGTCTCGCGCC-3'
3'-nhe-KRT24	5'-TTTGCTAGCTTAACTTTCACCTCAGAAATACTGCTGAC-3'
5'-bam-KRT3	5'-TTTGGATCCACATGAGCAGACAAGCCAGCAA-3'
3'-nhe-KRT3	5'-TTTGCTAGCTCTGGAGTAGCGCTGGGA-3'
5'-bam-KRT4	5'-TTTGGATCCACATGACTTCTGTAGGTGTGTTAG-3'
3'-nhe-KRT4	5'-TTTGCTAGCTATCGTCTCTTGTTCCAGGGTG-3'
5'-bam-KRT6a	5'-TTTGGATCCACATGGCCAGCACATCCACC-3'
3'-nhe-KRT6a	5'-TTTGCTAGCGTGCTTATAGCTCTCCTGC-3'
5-bn_csmut1a	5'-GATCCACAAAATCAGCCAGCTTGCTGGTGTGATGATGGCTCTGGCAGCG-3'
3-bn_csmut1a	5'-CTAGCGCTGCCAGAGCCATCATCAACACCAGCAAGCTGGCTGATTTTGTG-3'
5-bn_csmut3a	5'-GATCCACAAAATCAGCGCGCTTGCTGCTGTTGATGATGGCTCTGGCAGCG-3'
3-bn_csmut3a	5'-CTAGCGCTGCCAGAGCCATCATCAACAGCAGCAAGCGCGCTGATTTTGTG-3'
5-bn_cs1	5'-GATCCACGCTGAAGGACAGGTGGATGGGGCCGATGACTTAAGCACTAGCG-3'
3-bn_cs1	5'-CTAGCGCTAGTGCTTAAAGTCATCGGCCCATCCACCTGTCCTCAGCGTG-3'
5-xba_tasptv	5'-AAATCTAGAATGACCATGGAGAAGGGG-3'
3-xho_tasptv	5'-TTTCTCGAGGTTCACTGGGCTCTC-3'

5.1.7 Vektoren

Um eine hohe Kopienzahl eines bestimmten Proteins in einer Wirtszelle herstellen zu können, wurden entsprechende Expressionssysteme hergestellt. Zu diesem Zweck wird ein Expressionsvektor in die Wirtszelle transformiert (Prokaryoten) oder transfiziert (Eukaryoten), der die genetische Information zur Produktion des Proteins, einen Promoter, einen Transkriptionsterminator und eine Sequenz zur Ribosomenbindung enthält. Um selektionieren zu können, welche Zellen den Vektor aufgenommen haben und exprimieren, enthalten die hier verwendeten Vektoren jeweils ein bakterielles Ampicillin-Resistenz-Gen, sowie der eukaryotische Expressionsvektor eine Neomycin-Resistenz. Das Gen, das für das gewünschte Protein kodiert, kann an der sogenannten Multiple Cloning Site (MCS) mit Hilfe von Restriktionsenzymen und nachfolgender Ligation eingebracht werden. Eine genaue Beschreibung des Ablaufs der für diese Arbeit erstellten Plasmide folgt detailliert im Abschnitt Methoden (Kapitel 5.2.1).

Nachfolgend sind die für diese Arbeit verwendeten Ausgangs-Vektoren aufgeführt.

Tab. 21: Ausgangsvektoren

Bezeichnung	Relevante Methode	Quelle
pGEM Teasy	amp ^r – Klonierungs-Vektor	Promega, Mannheim
pcDNA3.1 (+)	amp ^r , neo ^r – eukaryotischer Expressionsvektor	Invitrogen, Karlsruhe
pc3_GFP	amp ^r , neo ^r – eukaryotischer Expressionsvektor zur Klonierung c-terminaler GFP-Fusionsproteine, basierend auf pcDNA3.1(+)	Labor Prof. Stauber ³¹⁹
pc3_BFP	amp ^r , neo ^r – eukaryotischer Expressionsvektor zur Klonierung c-terminaler BFP-Fusionsproteine, basierend auf pcDNA3.1(+)	Labor Prof. Stauber ³¹⁹
pc3_mCherry	amp ^r , neo ^r – eukaryotischer Expressionsvektor zur Klonierung c-terminaler mCherry-Fusionsproteine, basierend auf pcDNA3.1(+)	Labor Prof. Stauber ³¹⁹
pc3_GST/GFP	amp ^r , neo ^r – eukaryotischer Expressionsvektor zur Klonierung n-terminaler GST- und c-terminaler GFP-Fusionsproteine, basierend auf pcDNA3.1(+)	Labor Prof. Stauber ³¹⁹
pAC 5.1/V5 His-A	amp ^r – eukaryotischer Expressionsvektor für <i>Drosophila melanogaster</i>	Invitrogen, Karlsruhe

MATERIAL UND METHODEN

Bezeichnung	Relevante Methode	Quelle
pIB/V5-His	amp ^r – eukaryotischer Expressionsvektor für <i>Drosophila melanogaster</i>	Life Technologies, Darmstadt

Tab. 22: Humane Expressionsvektoren

Bezeichnung	Kodierend für	Quelle
pc3-Tasp	Volllänge-hTaspase1	Bier <i>et al.</i> 2011
pc3-TaspTV	Volllänge hTaspase1 mit Mutation von T234 nach V	Bier <i>et al.</i> 2011
pc3-Tasp-GFP	Volllänge hTaspase1 in Fusion mit GFP	Bier <i>et al.</i> 2011
pc3-Tasp-mcherry	Volllänge hTaspase1 in Fusion mit mCherry	Bier <i>et al.</i> 2011
pc3-TaspTV-mcherry	Volllänge hTaspase1 in Fusion mit mCherry mit Mutation von T234 nach V	Bier <i>et al.</i> 2011
pc3-BioTasp	Fusionsprotein aus SV30NLS-GFP-GST-Taspase-Schnittstelle2 des MLL Proteins-myc-RevNES	Bier <i>et al.</i> 2011
pc3-BioTaspmut	BioTasp mit Mutation von D1 und G1' nach A	Bier <i>et al.</i> 2011
pc3-BioCasp3	Fusionsprotein aus SV30NLS-GFP-GST-Caspase3-Schnittstelle-myc-RevNES	Bier <i>et al.</i> 2011
pc3-KRT4-GFP	Volllänge KRT4 in Fusion mit GFP	diese Arbeit
pc3-KRT4-mcherry	Volllänge KRT4 in Fusion mit mCherry	diese Arbeit
pc3-KRT4-BFP	Volllänge KRT4 in Fusion mit BFP	diese Arbeit
pc3-KRT6a-GFP	Volllänge KRT6a in Fusion mit GFP	diese Arbeit
pc3-KRT6a-mcherry	Volllänge KRT6a in Fusion mit mCherry	diese Arbeit
pc3-KRT6a-BFP	Volllänge KRT6a in Fusion mit BFP	diese Arbeit
pc3-KRT24-GFP	Volllänge KRT24 in Fusion mit GFP	diese Arbeit
pc3-KRT24-mcherry	Volllänge KRT24 in Fusion mit mCherry	diese Arbeit
pc3-KRT24-BFP	Volllänge KRT24 in Fusion mit BFP	diese Arbeit

Tab. 23: Drosophila-Expressionsvektoren

Bezeichnung	Kodierend für	Quelle
pAC-dTasp-GFP	Volllänge-Drosophila Taspase1 in Fusion mit GFP	Wünsch <i>et al.</i> 2015
pAC-dTasp-mCherry	Volllänge Drosophila Taspase1 in Fusion mit mCherry	Wünsch <i>et al.</i> 2015
dBioTasp _{CS2}	Fusionsprotein aus SV40NLS-GFP-GST-Taspase-	Wünsch <i>et al.</i> 2015

MATERIAL UND METHODEN

Bezeichnung	Kodierend für	Quelle
	Schnittstelle2 des MLL-Proteins-myc-RevNES im pAC-Vektor	
pAC-hTasp-GFP	Volllänge humane Taspase1 in Fusion mit GFP	Wünsch et al. 2015
pAC-dTaspTV-GFP	Volllänge Drosophila Taspase1 in Fusion mit GFP mit Mutation von Thr195 nach Val	Wünsch et al. 2015
pAC-dTF2A-GFP	Volllänge Drosophila TF2A in Fusion mit GFP	Wünsch et al. 2015
pAC-dBioTaspdTf2A	Fusionsprotein aus SV40NLS-GFP-GST-dTF2A-CS-myc-RevNES	Wünsch et al. 2015
pAC-dBioTasp _{CSmut(K1A)}	Fusionsprotein aus SV40NLS-GFP-GST-Taspase-Schnittstelle2 des MLL-Proteins mit Mutation von K6 nach Ala-myc-RevNES	Wünsch et al. 2015
pAC-dBioTasp _{CSmut(I2A)}	Fusionsprotein aus SV40NLS-GFP-GST-Taspase-Schnittstelle2 des MLL-Proteins mit Mutation von I5 nach Ala-myc-RevNES	Wünsch et al. 2015
pAC-dBioTasp _{CSmut(S3A)}	Fusionsprotein aus SV40NLS-GFP-GST-Taspase-Schnittstelle2 des MLL-Proteins mit Mutation von S4 nach Ala-myc-RevNES	Wünsch et al. 2015
pAC-dBioTasp _{CSmut(Q4A)}	Fusionsprotein aus SV40NLS-GFP-GST-Taspase-Schnittstelle2 des MLL-Proteins mit Mutation von Q3 nach Ala-myc-RevNES	Wünsch et al. 2015
pAC-dBioTasp _{CSmut(Q4N)}	Fusionsprotein aus SV40NLS-GFP-GST-Taspase-Schnittstelle2 des MLL-Proteins mit Mutation von Q3 nach Asn-myc-RevNES	Wünsch et al. 2015
pAC-dBioTasp _{CSmut(L5A)}	Fusionsprotein aus SV40NLS-GFP-GST-Taspase-Schnittstelle2 des MLL-Proteins mit Mutation von L2 nach Ala-myc-RevNES	Wünsch et al. 2015
pAC-dBioTasp _{CSmut(L5I)}	Fusionsprotein aus SV40NLS-GFP-GST-Taspase-Schnittstelle2 des MLL-Proteins mit Mutation von L2 nach Iso-myc-RevNES	Wünsch et al. 2015
pAC-dBioTasp _{CSmut(L5F)}	Fusionsprotein aus SV40NLS-GFP-GST-Taspase-Schnittstelle2 des MLL-Proteins mit Mutation von L2 nach Phe-myc-RevNES	Wünsch et al. 2015
pAC-dBioTasp _{CSmut(D6A)}	Fusionsprotein aus SV40NLS-GFP-GST-Taspase-Schnittstelle2 des MLL-Proteins mit Mutation von D1 nach Ala-myc-RevNES	Wünsch et al. 2015
pAC-dBioTasp _{CSmut(G7A)}	Fusionsprotein aus SV40NLS-GFP-GST-Taspase-Schnittstelle2 des MLL-Proteins mit Mutation	Wünsch et al. 2015

Bezeichnung	Kodierend für	Quelle
	von G1' nach Ala-myc-RevNES	
pAC-dBioTasp ^{CSmut(V8A)}	Fusionsprotein aus SV40NLS-GFP-GST-Taspase-Schnittstelle ² des MLL-Proteins mit Mutation von V2' nach Ala-myc-RevNES	Wünsch et al. 2015
pAC-dBioTasp ^{CSmut(D9A)}	Fusionsprotein aus SV40NLS-GFP-GST-Taspase-Schnittstelle ² des MLL-Proteins mit Mutation von D3' nach Ala-myc-RevNES	Wünsch et al. 2015
pAC-dBioTasp ^{CSmut(D10A)}	Fusionsprotein aus SV40NLS-GFP-GST-Taspase-Schnittstelle ² des MLL-Proteins mit Mutation von D4' nach Ala-myc-RevNES	Wünsch et al. 2015

Weitere Vektoren können den jeweiligen Publikationen entnommen werden.

5.1.8 Zelllinien

Tab. 24: Zelllinien und deren Ursprung.

Zelllinie	Gewebe	Organismus	ATCC Nr./Referenz
293T / HEK	Embryonales Nierengewebe	<i>Homo sapiens</i>	CRL-11268
Fadu	Rachen, squamöses Zellkarzinom (HNSCC)	<i>Homo sapiens</i>	HTB-43
H441	Lunge, papilares Adenokarzinom (HNSCC)	<i>Homo sapiens</i>	HTB-174
HeLa	Cervix Adenokarzinom	<i>Homo sapiens</i>	HTB-38
NeHab	Rachen, squamöses Zellkarzinom (HNSCC)	<i>Homo sapiens</i>	Universitätsmedizin HNO, Mainz
SaOs	Knochen, Osteosarkom	<i>Homo sapiens</i>	HTB-85
StauMa	Stirnhöhle, squamöses Zellkarzinom (HNSCC)	<i>Homo sapiens</i>	Universitätsmedizin HNO, Mainz
SW480	Kolorektales Adenokarzinom	<i>Homo sapiens</i>	CCL-228
UMB-SCC-969	Mandel, squamöses Zellkarzinom (HNSCC)	<i>Homo sapiens</i>	Universität Marburg

Im Folgenden sind die Zusammensetzungen der verwendeten Medien zur Kultivierung der jeweiligen Zelllinie aufgelistet.

Tab. 25: Medien zur Kultivierung von Zellen.

Medium	Beschreibung	Zusatz	für Zelllinie	Hersteller
DMEM+++	Glucose 4500 mg/l	10% FCS	293T / HEK	Gibco, Invitrogen (Karlsruhe) #21969
	NaPyruvat	1% Pen/Strep	Fadu	
	ohne L-Glutamin	1% Glutamin	H441	
			HeLa	
			SW480	
			UMB-SCC-969	
RPMI20%	Glucose 2000 mg/l	20% FCS	NeHab	Gibco, Invitrogen (Karlsruhe) #21870-084
	ohne L-Glutamin	1% Pen/Strep	StauMa	
		1% Glutamin		
DMEM/F12+++	1:1 DMEM/F12	10% FCS	SaOs	Gibco, Invitrogen (Karlsruhe) #21331-020
	ohne L-Glutamin	1% Pen/Strep		
		1% Glutamin		

5.1.9 Patienten-Charakteristika und -Gewebe

Alle Gewebeproben stammten aus HNO-Patienten der Universitätskliniken Frankfurt und Mainz wurden in Übereinstimmung mit den ethischen Standards aus der Deklaration von Helsinki sowie nationalen und internationalen Richtlinien entnommen, archiviert und weiterbearbeitet. Das Studienprotokoll wurde von den Ethikkomitees „Ethik-Kommission des Fachbereichs Medizin, Universitätsklinikum der Goethe-Universität“ (#83756604) und der „Ethik-Kommission der Landesärztekammer Rheinland-Pfalz“ (#83748515) zugelassen. Alle Patienten wurden protollgemäß informiert, gaben ihr Einverständnis und wurden anonymisiert. Die Fälle wurden histopathologisch als Kopf-Hals-Plattenepithelkarzinome (HNSCC) diagnostiziert und gemäß den Richtlinien der „Union International contre le Cancer“ (UICC) und der TNM (2010) klassifiziert. Von jedem in die Studie aufgenommenen Patienten war Material aus Primärtumor, Normalgewebe aus der

Wangenschleimhaut sowie Lymphknotenmetastasen verfügbar. Unter den Primärtumorproben waren Oropharynx, Larynx und Hypopharynx Karzinome von unterschiedlicher Größe (T1-T4), der Lymphknoten-Status war gemäß Studiendesign in allen Fällen positiv (N1-3) ohne Fernmetastasen (M0). Zusätzliche Inklusionskriterien waren positive Antworten auf Alkohol- und Tabakgebrauch, HPV-negativer Status der Primärtumore, histopathologisches Grading von 2-3 und keine Krebstherapien vor der Resektion.

Nach Resektion wurden die Gewebeproben sofort auf Eis gelegt und innerhalb von 30 min in flüssigem Stickstoff eingefroren. In histologische Analysen wurde bestätigt, dass die Tumor- und Metastasenproben mindestens 70% Tumorzellen und weniger als 10% nekrotisches Gewebe enthielten. Wangenschleimhaut wurde auf die Abwesenheit von Tumorgeschehen überprüft. Alle Kriterien wurden von wenigstens zwei Pathologen überprüft und bestätigt.

Proben, die diesen Anforderungen nicht entsprachen, wurden von der Studie ausgeschlossen.

5.1.10 Software

Tab. 26: Verwendete Software

Programm/Seite	Hersteller	Referenz
BLAST	National Center for Biotechnology Information (NCBI)	http://www.ncbi.nlm.nih.gov/blast/
MS Word 2010	Microsoft	https://products.office.com/de-de/microsoft-word-2010
MS Excel 2010	Microsoft	https://products.office.com/de-de/microsoft-excel-2010
Photoshop CE / CS6	Adobe	http://www.adobe.com/de/products/photoshop.html
Canvas X16	ACD Systems	http://www.canvasgfx.com/
Endnote X7	Clarivate Analytics	http://endnote.com/product-details
GeneSpring GX	Agilent	http://www.agilent.com/en/products/software-informatics/life-sciences-informatics/genespring-gx
Ingenuity	Qiagen	https://analysis.ingenuity.com/
BioEdit	Tom Hall	http://www.mbio.ncsu.edu/BioEdit/bioedit.html
Gene Construction Kit	TextCo Bio Software	http://www.textco.com/gene-construction-kit.php
Axiovision 3.0	Zeiss	https://www.zeiss.de/mikroskopie/produkte/mikroskopssoftware/axiovision-fuer-materialien.html

MATERIAL UND METHODEN

ExpASy	SIB Swiss Institute of Bioinformatics	https://www.expasy.org/
TUBA	Max Deichelbohrer	http://tuba.zdv.uni-mainz.de:8080/tuba/
HALO	Indica Labs	http://www.indicalab.com/

5.2 Methoden

5.2.1 DNA- und RNA-Techniken

5.2.1.1 RNA Extraktion aus Gewebeproben

Um die durchschnittlich transkribierten Gene eines Gewebes zu analysieren, muss zuerst die gesamte RNA aus den Zellen isoliert werden. 30-50 mg gefrorene Gewebeproben wurden in 1 ml Trizol aufgenommen und in einem Ultra-Turrax T25-Homogenisierer aufgelöst. Die RNA wurde gemäß den Anleitungen des Herstellers für das Trizol-Protokoll extrahiert und auf RNeasy Mini-Säulchen (Qiagen) isoliert. Die Integrität und Reinheit der RNA wurde an einem Bioanalyzer 2100 mittels des RNA 6000 NanoLabChip Kit (Agilent) gemäß den Herstellerangaben überprüft. Absorptionsverhältnisse von 1,8-2,2 für A_{260}/A_{280} und kleiner als 1,7 für A_{260}/A_{230} wurden als qualitativ hinreichend akzeptiert.

Proben, die diesen Anforderungen nicht entsprachen, wurden von der Studie ausgeschlossen.

5.2.1.2 Polymerasekettenreaktion (PCR)

Bei der Polymerasekettenreaktion werden Oligonukleotide benutzt, um von einer Vorlage (Templat) DNA exponentiell zu amplifizieren. Die DNA wird hierzu durch Erhitzen auf über 95°C denaturiert und liegt so in Einzelsträngen vor. An diese können nun nach Abkühlen die spezifischen Primer binden (*annealing*). Die optimale Temperatur der *annealing*-Phase ist vom GC-Gehalt und der Länge der Primer abhängig. Anschließend werden die Stränge durch eine hitzebeständige DNA-Polymerase aufgefüllt, z.B. die Taq-DNA-Polymerase (aus dem thermophilen Bakterium *Thermus aquaticus*). Bei 72°C repliziert sie 1000 Nukleotide pro Sekunde, besitzt jedoch keine Exonukleaseaktivität, so dass sie eine hohe Fehlerquote besitzt (etwa 1 pro 9000 Nukleotide). Eine ähnlich schnelle Alternative mit sogenannter *Proofreading*-, also Korrektur-Funktion bietet die Pfu-DNA-Polymerase (aus dem thermophilen Archebakterium *Pyrococcus furiosus*).

Ein reguläres PCR-Programm besteht aus mehreren sich wiederholenden Zyklen, die aus Denaturierungs- (*denaturation*), Anlagerungs- (*annealing*) und Verlängerungsphase (*elongation*) bestehen. Es wurden folgende Programme in einem Thermocycler verwendet (**Tab. 14**).

Tab. 27: PCR Programme.

PCR Programm	Schritt	Temperatur	Zeit
Standard (bis 1 kbp)	1 (<i>denaturation</i>)	95°C	10 min
	2 (<i>denaturation</i>)	95°C	30 s
	3 (<i>annealing</i>)	54°C	60 s
	4 (<i>elongation</i>)	72°C	60 s
		30x Wiederholung der Schritte 2-4	
	5 (<i>elongation</i>)	72°C	7 min
	6 (<i>cooling</i>)	4°C	unbegrenzt
Expand (größer 1kbp)	1 (<i>denaturation</i>)	94°C	2 min
	2 (<i>denaturation</i>)	94°C	10 s
	3 (<i>annealing</i>)	56°C	30 s
	4 (<i>elongation</i>)	68°C	4 min
		9x Wiederholung der Schritte 2-4	
	5 (<i>denaturation</i>)	94°C	15 s
	6 (<i>annealing</i>)	56°C	30 s
	7 (<i>elongation</i>)	68°C	4 min
		26x Wiederholung der Schritte 5-7	
8 (<i>elongation</i>)	68°C	7 min	
9 (<i>cooling</i>)	4°C	unbegrenzt	

Für einen präparativen Ansatz von 50 µl wurden folgende Komponenten auf Eis in ein PCR-Reaktionsgefäß pipettiert:

Tab. 28: Pipettierschema für einen 50µl PCR-Ansatz.

Komponente	Volumen pro Reaktion [µl]	Endkonzentration im Ansatz
10x PCR Puffer	5	1x
10 mM dNTP-Mix	1	200 µM
10 µM forward Primer	1	0,1 µM

Komponente	Volumen pro Reaktion [μl]	Endkonzentration im Ansatz
10 μM reverse Primer	1	0,1 μM
Templat-DNA	je nach Konzentration	0,2-0,5 μg
DNA-Polymerase	1	5 U
H ₂ O bidest, steril	ad 50 μl	

5.2.1.3 Reverse Transkriptions (RT-)PCR

Je 1 μg wurde pro Ansatz verwendet, um cDNA zu synthetisieren und die Qualität der RNA zu überprüfen. Hierfür wurde das *Transcriptor First Strand cDNA Kit* (Roche, Mannheim) gemäß den Herstellerangaben verwendet.

5.2.1.4 Quantitative / Real-Time-PCR (qPCR)

Die relative Menge verschiedener Transkripte wurde mittels quantitativer PCR mit Hilfe des SYBR-Green I-Systems (Roche, Mannheim) bestimmt. Der Fluoreszenzfarbstoff SYBR Green lagert sich in die kleine Furche der doppelsträngigen DNA ein und kann im LightCycler (**Tab. 14**) detektiert werden. Der Anstieg der cDNA während der Amplifikation korreliert mit der Zunahme des Fluoreszenzfarbstoffes. In Relation zu einem Referenz-Gen (GAPDH) kann daraus die ursprünglich vorhandene Kopienzahl des jeweiligen Transkripts berechnet werden.

Der 20 μl Reaktionsansatz bestand aus 2 μl eines Mastermix (10 μl FastStartDNAMasterSYBRGreenIa zu 54 μl FastStartDNAMasterSYBRGreenIb), 1 μl 10 μM Primer, 2 μl cDNA und 15 μl H₂O bidest. Folgendes Programm wurde verwendet:

Tab. 29: qPCR Programm

Phase	Schritt	Temperatur	Zeit
Vorinkubation	1 (<i>denaturation</i>)	95° C	10 min
Quantifizierung	2 (<i>denaturation</i>)	95° C	10 s
	3 (<i>annealing</i>)	60° C	10 s
	4 (<i>elongation</i>)	72° C	10 s
	45x Wiederholung der Schritte 2-4		
Schmelzkurve	5 (<i>denaturation</i>)	95° C	10 s

Phase	Schritt	Temperatur	Zeit
	6 (<i>annealing</i>)	65° C	15 s
	7 (<i>melting</i>)	von 95° absteigend 0,1°C/s	
Abkühlen	8 (<i>cooling</i>)	40° C	30 s

In diesem System wird am Ende jedes Zyklus die kumulative Fluoreszenz gemessen (Schritt 4) und nach dem letzten Zyklus die Schmelzkurve (Schritt 7) zur Überprüfung der spezifischen Amplifikation der DNA-Fragmente, die einen idealerweise sigmoiden Verlauf haben sollte. Der Schwellenwert wird in den linearen Verlauf der Kurve gesetzt, und die Software errechnet für jede Probe die Zyklenzahl, die benötigt wurde, um den Schwellenwert zu erreichen (*cycle threshold*, CT-Wert). Falls mehrere Peaks in der Schmelzkurve vorhanden sind, wurde mehr als ein Fragment amplifiziert. Schmelztemperaturen unter 70° C kennzeichnen zu kurze Fragmente oder Primer-Dimere. In beiden Fällen ist das Ergebnis unbrauchbar und zu verwerfen.

Zur Berechnung der Relativen Expression (R) eines Zielgens wurde folgende Formel verwendet:

$$R = (E_{\text{ZIEL}})^{\Delta \text{CT ZIEL(KONTROLLE-PROBE)}} / (E_{\text{REF}})^{\Delta \text{CT REF(KONTROLLE-PROBE)}}$$

Basierend auf der Real-Time PCR-Effizienz (E) und der Schwellenwert (CT)-Differenz einer unbekannt Probe zu einer Kontrolle im Vergleich zu einem Referenzgen wird die Relative Expression errechnet.

5.2.1.5 Splice Overlap Extension(SOE)-PCR

Mittels der Methode der *Splice Overlap Extension*(SOE)-PCR³²⁰ können gezielt Punktmutationen in DNA-Fragmente eingebaut und so spezifisch einzelne Aminosäuren innerhalb eines Proteins ausgetauscht werden (**Abb. 58**). Die DNA wird dafür im ersten Schritt in zwei Segmenten (AB und CD) amplifiziert, in welchen die jeweilige Mutation bereits über die inneren Primer (b und c) eingebaut wird. Im zweiten Schritt werden die Segmente AB und CD sowie die Außen-Primer (a und d) verwendet, um schließlich das gewünschte DNA-Fragment (AD) zu amplifizieren.

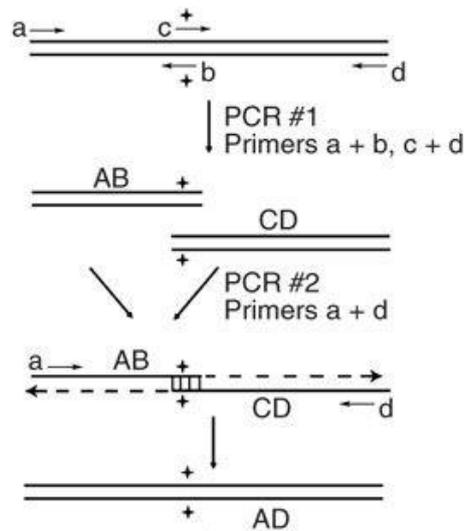


Abb. 58: Schematische Darstellung einer Splice Overlap Extension PCR.
Verändert nach ³²⁰.

5.2.1.6 Oligoannealing

Für sehr kurze Fragmente bietet sich die synthetische Herstellung in Verbindung mit anschließendem Oligoannealing an. Um ein entsprechendes Insert zu generieren, werden komplementäre Oligonukleotide mit überhängenden Enden verwendet. Zur Phosphorylierung der 5'-Enden wird folgendes Pipettierschema verwendet:

Tab. 30: Pipettierschema für einen 10µl Ansatz zur Phosphorylierung der 5'-Enden.

Komponente	Volumen pro Reaktion [µl]	Endkonzentration im Ansatz
Oligonukleotid 1	je nach Konzentration	10 µg
Oligonukleotid 2	je nach Konzentration	10 µg
PNK-Puffer (10x)	1 µl	1x
2,5 mM dNTP-Mix	1 µl	250 µM
Phosphonukleotidkinase	1 µl	5 U
H ₂ O bidest, steril	ad 10 µl	

Der Ansatz wird für 20 min bei 37°C inkubiert. Für die eigentliche Hybridisierung (*annealing*) der Oligonukleotide wird 2 µl 10xNEB-Puffer³ hinzugegeben und mit sterilem, bidest. Wasser auf 20 µl

Reaktionsvolumen aufgefüllt, in einem Wasserbad auf 100°C erhitzt und anschließend bei Raumtemperatur inkubiert, bis das Gemisch abgekühlt ist.

5.2.1.7 Restriktionsendonukleaseverdau

Der Verdau mit Restriktionsendonukleasen wurde in dieser Arbeit zum einen zur Charakterisierung von DNA-Fragmenten durch anschließende Auftrennung auf einem Agarose-Gel und zum anderen zur Präparation von Vektoren und Inserts für eine anschließende Ligation herangezogen.

Restriktionsendonukleasen, kurz Restriktionsenzyme, stammen ursprünglich aus Bakterien oder Archaeen und dienen dort dem Abbau von aufgenommener und möglicherweise gefährlicher Fremd-DNA, die durch unterschiedliche Methylierungsmuster von eigener DNA unterschieden wird. Es gibt vier unterschiedliche Typen von Restriktionsenzymen mit diversen Subtypen. Entscheidend für die vorliegende molekularbiologische Arbeit sind diejenigen vom Typ II, die kein ATP benötigen und sequenzspezifisch schneiden. Dieser Schnitt kann glatte Enden (blunt ends) oder überhängende Enden (sticky ends) erzeugen, wobei letztere das Einsetzen in einen Vektor an bestimmter Stelle erleichtern.

Für den Kontrollansatz (10 µl) und den präparativen Ansatz (50 µl) wurde folgendes Pipettierschema verwendet. Bei einem Einzelverdau mit nur einem Enzym ist die doppelte Menge einzusetzen. Der Einsatz von BSA ist nicht bei allen Restriktionsenzymen nötig - für diese sind die Werte in Klammern vorgesehen.

Tab. 31: Pipettierschema DNA-Restriktionsverdau: Kontrollansatz (10µl) und präparativer Ansatz (50µl).

Komponente	Kontrollansatz	Präparativer Ansatz
Template	5 µg	40 µg
BSA	(0,5 µl)	(0,5 µl)
10x Puffer (NEB 1-4)	1 µl	5 µl
Enzym 1	0,5 µl	2,5 µl
Enzym 2	0,5 µl	2,5 µl
H₂O bidest.	ad 10 µl	ad 10 µl

Das Gemisch wurde bei 37°C für 2 h inkubiert und anschließend auf ein Agarosegel aufgetragen.

5.2.1.8 Dephosphorylierung

Nach dem Verdau wurden die 5'-Enden der DNA mit alkalischer Phosphatase dephosphoryliert, um eine Religation des Vektors zu verhindern. Es wurden die Angaben des Herstellers NEB beachtet.

5.2.1.9 Agarose-Gelelektrophorese

Mit Hilfe der Agarose-Gelelektrophorese wurden DNA-Fragmente zur Analyse oder Präparation durch Anlegen einer Stromspannung aufgetrennt. In der Gelmatrix aus vernetztem Polysaccharid können kleine Fragmente schneller passieren als größere und aufgrund der negativen Ladung der DNA zum Pluspol wandern.

Zur Herstellung des Gels wird das Agarose-Pulver, das aus den Rotalgen *Gelidium* und *Gracillaria* gewonnen wird, in TAE-Puffer durch Aufkochen gelöst und in die gewünschte Form gegossen. Ein eingesetzter Kamm bildet die Taschen, in die später die DNA-Probe hineingegeben wird. Je höher die Konzentration an Agarose, desto kleiner sind die Poren später im erkalteten Gel, so dass auch kleinere Fragmente gut aufgetrennt werden können.

Für diese Arbeit wurden 1-2%-Gele verwendet, die für je 40 ml flüssigen Gels mit 12,5 µl Ethidiumbromid (2 mg/ml) versetzt wurden, um die DNA unter UV-Licht sichtbar zu machen. Die Beladung der Gele erfolgt nach der Versetzung der Proben mit Ladepuffer (Bromphenolblau oder Xylencyanol, je nach Größe der gesuchten Fragmente). Die Größe der Fragmente wird mit Hilfe eines DNA-Größenstandards ermittelt, der parallel zu den Proben läuft. Die Auftrennung erfolgt nach ca. 45-60 min bei einer angelegten Spannung von 75 V und wird anschließend mit einem Transilluminator unter UV-Licht dokumentiert.

5.2.1.10 Gel-Elution

Um ein bestimmtes DNA-Fragment aufzureinigen, wurde die entsprechende Bande auf dem Gel ausgeschnitten und mit Hilfe des *QIAquick Gel Extraction Kits* nach Herstellerangaben (Quiagen, Düsseldorf) isoliert.

5.2.1.11 Ligation

Die Ligation verbindet DNA-Fragmente mit passenden Enden („sticky“ oder „blunt ends“, s. Restriktionsendonukleaseverdau), so dass ein Insert in einen vorbereiteten Vektor eingepasst werden kann. Die Reaktion wird durch eine Ligase katalysiert, die die 3'-5'-Phosphodiesterbindung zwischen beiden DNA-Fragmenten herstellt. Um eine Religation des geschnittenen, leeren Vektors zu vermeiden, wird eine niedrige Reaktionstemperatur von 16°C angesetzt, sowie das Insert in 4-fachem Überschuss hinzugegeben. Außerdem kann eine Dephosphorylierung der 5'-Enden des Vektors durchgeführt werden, so dass nur das Insert die benötigte Phosphatgruppe für die Reaktion bereitstellt (s. Dephosphorylierung).

Für die Ligation werden 0,5 µl Vektor (1 µg/µl), 2 µl Insert (1 µg/µl) und 2,5 µl Takara-Lösung für zwei Stunden bei 16°C inkubiert.

5.2.1.12 Herstellung kompetenter Bakterien

Um die Zellen durch Calciumchlorid kompetent zu machen, wurde eine Kultur des gewünschten *E. coli* Stammes über Nacht bei 37°C in LB-Medium unter Schütteln inkubiert, neu überimpft und bei Raumtemperatur auf eine optische Dichte von 0,6 herangezogen. Anschließend wird die Kultur durch Zentrifugation geerntet (4°C, 10 min, 3700 g), in 80 ml Transformation Buffer Japanese (TBjap: 10 mM Pipes, 15 mM CaCl₂, 250 mM KCl, mit KOH auf pH 6,7), 2% DMSO (v/v) resuspendiert und erneut unter gleichen Bedingungen abzentrifugiert. Der Überstand wird verworfen, das Pellet in 19 ml TBjap, 2% DMSO (v/v) gelöst und 1 ml DMSO (EK 7% (v/v)) hinzugegeben. Aliquots der nun Calciumchlorid-kompetenten Zellen werden zu je 100 µl in flüssigem Stickstoff eingefroren und bei -80°C gelagert.

5.2.1.13 Transformation

Als Transformation bezeichnet man die Aufnahme freier DNA durch kompetente Bakterienzellen (s. Herstellung kompetenter Bakterien). Dieses Verfahren dient in der Molekularbiologie dem Einbringen von Vektoren z.B. zur Herstellung rekombinanter Proteine oder zur Amplifikation des Vektors in der Wirtszelle. In dieser Arbeit wurden ausschließlich Calciumchlorid-kompetente *E. coli*-Stämme verwendet und mittels Hitzeschock transformiert.

Pro Ligationsansatz werden 50 µl kompetenter Zellen (auf Eis aufgetaut, s. Ligation) hinzugegeben, für 30 min auf Eis inkubiert (Anlagerung des Plasmids) und für 45 s auf 42°C zur Aufnahme der DNA gebracht. Anschließend werden die Bakterien erneut auf Eis gekühlt, auf angewärmten LB-Agarplatten mit entsprechendem Antibiotikum ausgestrichen und über Nacht bei 37°C inkubiert.

5.2.1.14 Sequenzierung

Einzelne Kolonien der aus der Transformation inkubierten Platten werden mit einem sterilen Zahnstocher in 5 ml LB Medium überführt, über Nacht bei 37°C inkubiert und mit Hilfe des DNA Miniprep-Kits von Qiagen (Düsseldorf) nach Herstellerangaben aufgereinigt.

400-700 ng der zu sequenzierenden DNA werden mit 10 pmol eines Sequenzierprimers vermischt und auf 10 µl mit sterilem H₂O bidest aufgefüllt. Die Sequenzierung wurde von der Firma StarSeq GmbH (Mainz) durchgeführt.

5.2.1.15 Plasmidpräparation

Die mittels Sequenzierung verifizierten Klone werden in 250 ml LB-Medium mit Ampicillin über Nacht bei 37°C vermehrt und die Plasmide mit dem DNA Maxiprep-Kit von Qiagen (Düsseldorf) nach Angaben des Herstellers isoliert. Die DNA-Konzentration wurde am NanoDrop Spektrophotometer (**Tab. 14**) gemessen, mit sterilem H₂O bidest auf 1 µg/µl eingestellt und bei -20°C gelagert.

5.2.1.16 Herstellung von Bakterien-Stocks

Zur Langzeitlagerung werden die positiven Klone in Stocks konserviert. Hierzu werden 10ml einer über Nacht gewachsenen Kultur bei 4°C, 4000 g für 10 min zentrifugiert, das Pellet in 1 ml sterilem 87% Glycerin resuspendiert und in einem Kryoröhrchen bei -80°C eingefroren.

5.2.2 Zellkulturtechniken

5.2.2.1 Kultivierung und Passage

Als Kultivierung wird die Schaffung und Aufrechterhaltung von Bedingungen bezeichnet, die das Wachstum von Zellen in einer künstlichen Umgebung ermöglicht. Primäre als auch immortalisierte

Zellen können so in einem auf den Zelltyp abgestimmten Nährmedium vermehrt werden und stehen für weitere Versuche zur Verfügung. Die Kulturen werden in sterilen Flaschen angeimpft und in Brutschränken gehalten. Für die in dieser Arbeit verwendeten humanen Zelllinien waren Bedingungen von 37°C und 5% CO₂ in wasserdampfgesättigter Umgebung ideal. Um Kontaminationen mit Pilzen, Bakterien oder anderen Zelllinien zu vermeiden, wird beim Arbeiten an der Zellkultur penibel auf Sauberkeit geachtet und die Kulturflaschen nur unter der Sterilbank geöffnet und bearbeitet.

Drei Mal pro Woche werden Vitalität und Dichte der Zellen unter dem Mikroskop kontrolliert und unter der Sterilbank das alte Medium entfernt, um Stoffwechselprodukte zu entfernen. Anschließend werden die Zellen entweder dem Zelltyp entsprechend zu Teilen in eine neue Flasche überführt (passagiert) oder das Medium gewechselt, wenn eine Passage aufgrund der Dichte der Zellen noch nicht angebracht ist (vgl. **Tab. 24**).

Zur Passage der Zellen wird das Medium vorsichtig abgesaugt, die Zellen mit sterilem PBS gewaschen und je nach Größe der Kulturflasche 1-3 ml Trypsin hinzugegeben. Der Waschschrift ist notwendig, da das im Medium enthaltene FCS die proteolytische Aktivität des Trypsins inhibieren würde. Die geschlossene Flasche wird auf einem Heizblock auf 37°C inkubiert, bis sich alle Zellen abgelöst haben. Um die Ablösereaktion zu stoppen, werden 1-3 ml Medium zugegeben und die Zellen vorsichtig resuspendiert und vereinzelt. Die gewünschte Menge je Passage wird in eine neue Zellkulturflasche gegeben und mit frischem, vorgewärmtem Medium aufgefüllt.

5.2.2.2 Zellzählung

Zur Ermittlung der Zellzahl wurde das CASY 1 TT (vgl. **Tab. 14**) verwendet, das mittels elektronischer Pulsflächenanalyse die Anzahl der Zellen in einem Gemisch aus Zellsuspension und Casyton, einer isotonischen Salzlösung, bestimmt. Durch die Messung des elektrischen Widerstands der Zellen kann auch deren Durchschnittsvolumen, -durchmesser sowie deren Vitalität ermittelt werden.

5.2.2.3 Transfektion

Im Gegensatz zur Transformation bei Bakterien wird das Einbringen von DNA oder RNA in eukaryotische Zellen als Transfektion bezeichnet. Die so eingebrachte genetische Information kann

vielfältigen Zwecken dienen, z. B. um in der Wirtszelle Proteine herzustellen oder mittels RNA-Interferenz bestimmte mRNAs stillzulegen.

5.2.2.4 Transfektion mittels Polyethylenimin (PEI)

Polyethylenimin ist ein stark basisches Polymer, das in wässriger Lösung Polykationen bildet, somit DNA einschließen und als Micelle von der Zelle aufgenommen werden kann. Im Inneren der Zelle wird die DNA freigegeben. Diese Transfektionsmethode wurde in dieser Arbeit ausschließlich für 293T/HEK Zellen zur Herstellung von Lysaten oder zur schnellen fluoreszenzmikroskopischen Kontrolle von Klonierungserfolgen verwendet.

Für eine 100 mm Schale werden 12 µg DNA mit 180 µl PBS, sowie 32,4 µl PEI mit 180 µl PBS erst getrennt angesetzt, gut vermischt und anschließend zusammengegeben. Die beiden Ansätze werden miteinander durch Vortexen vermischt und 5 min bei Raumtemperatur inkubiert, bevor das Gemisch vorsichtig auf die Zellen pipettiert wird. Nach 6 Stunden wird das Medium gewechselt. Die Zellen werden nach 48 Stunden geerntet.

5.2.2.5 Transfektion mittels Lipofektion

Bei der Lipofektion wird die DNA in Liposomen eingeschlossen und über Membranfusion in die Zelle eingebracht. Als Transfektionsreagenz wurde Lipofectamine 2000 (**Tab. 12**) verwendet, das aus kationischen Lipiden besteht.

Für eine 35 mm Schale werden 2 µg DNA mit 125 µl OptiMEM, sowie 5 µl Lipofectamine mit 125 µl OptiMEM getrennt angesetzt, gemischt und für 5 min bei Raumtemperatur inkubiert, bevor die Ansätze miteinander durch Vortexen vermischt werden. Nach 20-minütiger Inkubation wird den 60-80% konfluent gewachsenen Zellen das Medium abgesaugt, 1 ml OptiMEM vorgelegt und das vorbereitete Transfektionsgemisch hinzugegeben. Nach 6 Stunden wird das Medium gewechselt.

5.2.2.6 Fluoreszenzmikroskopie

Die Eigenschaft von Fluorochromen, absorbiertes Licht wieder durch Emission eines Lichtquants anderer Wellenlänge abzugeben, kann dazu genutzt werden, Proteine in lebenden Zellen sichtbar zu machen. Fluoreszierende Proteine wie das grün-fluoreszierende Protein GFP aus *Aequorea victoria*

und seine mittlerweile diversen Varianten haben sich in Fusion mit Zielproteinen als Markierung bewährt.

Die Fluoreszenzanalysen dieser Arbeit wurden an einem inversen Fluoreszenzmikroskop (Axiovert 200M, vgl. **Tab. 14**) aufgenommen, mittels der computergestützten Bilderfassung AxioVision 4.7 (Zeiss, Jena) visualisiert und in der Bildbearbeitungssoftware Adobe Photoshop CE (Adobe, München) weiterbearbeitet (**Tab. 26**).

5.2.3 Proteintechniken

5.2.3.1 Gesamtzellysate

Um den relativen zellulären Proteingehalt zu ermitteln, müssen die Zellen zunächst aufgeschlossen werden, und die lösliche von der unlöslichen Fraktion getrennt werden. Die gesamte Bearbeitung erfolgt auf Eis oder bei 4°C. Die Zellen wurden nach Abschaben einer dicht gewachsenen 100 mm Schale in einem Falcon-Röhrchen bei 4°C, 500 g für 5 min abzentrifugiert, in PBS gewaschen und erneut bei 4°C, 800 g für 5 min abzentrifugiert. Der Überstand wird verworfen und je nach Größe des Pellets 70-400 µl RIPA-Puffer (**Tab. 13**) hinzugegeben. Der Aufschluss erfolgt über die im Puffer enthaltenen Detergenzien SDS, NP-40 und Triton-X-100. Das Gemisch wird durch anschließendes Sonifizieren (3x, 30 s je 0,5 s Pulse/Pause) weiter solubilisiert und bei 4°C, 17.700 g, 30 min zentrifugiert. Der Überstand wird in ein neues Reaktionsgefäß überführt und bei -80°C gelagert.

5.2.3.2 Aufreinigung von GST-Fusionsproteinen

Ein *E. coli*-Klon wird in 25 ml LB Medium mit entsprechendem Antibiotikum (LB_{AB}) angeimpft, für 4-6 h auf 37°C inkubiert und in 250 ml LB_{AB} überimpft. Nach Erreichen einer Optischen Dichte (OD) von 0,4-0,6 wird mit 0,1 mM Isopropyl-β-D-1-thiogalactopyranosid (IPTG) induziert und über Nacht bei Raumtemperatur unter Schütteln inkubiert. Die Bakterien werden durch Zentrifugation bei 4°C, 3700 g, 30 min geerntet und das Pellet in 10 ml PBS sauber homogenisiert. Nach Zugabe von 10 µl DNaseI (Endkonz. 100 µg/ml), 200 µl Proteaseinhibitor-Mix-Stocklösung (50x Complete in PBS), 26 µl 1 M MnCl₂, 260 µl 1 M MgCl₂, sowie einer Spatelspitze Lysozym wird 15 min auf Eis inkubiert. Anschließend wird 3x 30 s sonifiziert (0,5 s Pulse/Pause), 1 ml 5 M NaCl und 500 µl 20% Triton X-100 hinzugefügt und 10 min auf Eis inkubiert. Nach Zentrifugation bei 4°C, 17.700 g, 30 min wird der

Überstand in ein neues Falcontube überführt und für je 300 Fluoreszenz-Einheiten 300 µl Glutathion-Sepharose-Beads (Amersham Pharmacia, GE Healthcare, Freiburg) zugegeben. Zur Bindung der GST-Fusionsproteine an die Beads wird 2 h auf dem Drehrad bei 4°C inkubiert. Die Beads werden bei 4°C, 2000 g, 10 min abzentrifugiert, der Überstand abgesaugt, erst mit 50 ml 1% Triton X-100 in PBS, dann 4x mit 50 ml PBS gewaschen. Zur Herstellung der Konstrukt-Bead-Aliquots für den Fluoreszenzassay wird beim letzten Waschschrift nur auf 5 ml abgesaugt und je 150 µl in PCR-Tubes überführt, diese in flüssigem Stickstoff eingefroren und bei -80°C gelagert.

Zur Aufreinigung der GST-Fusionsproteine werden die Beads nach dem letzten Waschschrift mit 1 ml GST-Elutionspuffer (vgl. **Tab. 13**) für 15 min unter Schütteln inkubiert, und anschließend die Flüssigkeit über ein PolyPrep-Säulchen (Bio-Rad, München) von den Beads getrennt. Das Eluat wird in einen Dialyseschlauch überführt und über Nacht im Kühlraum gegen 2-5l PBS dialysiert. Das gereinigte GST-Fusionsprotein wird in 50 µl Aliquots aufgeteilt, in flüssigem Stickstoff eingefroren und bei -80°C gelagert.

5.2.3.3 Proteinquantifizierung nach Bradford

Zur Bestimmung des Proteingehalts wurde die Methode nach Bradford ³²¹ angewendet. Das Proteinfärbereagenz (vgl. **Tab. 12**) wurde 1:5 mit bidest. Wasser verdünnt. Je 1-2 µl der zu bestimmenden Proteinlösung wurden zu 1ml der Bradfordlösung pipettiert und bei 595 nm gegen den verwendeten Puffer in einem Photometer gemessen. Als Standard dient eine Eichkurve von im verwendeten Puffer gelösten BSA.

5.2.3.4 SDS-Polyacrylamidgelelektrophorese

Sind die Proteine in Lösung verfügbar, können sie mit Hilfe der Natriumdodecylsulfat-Polyacrylamid-Gelelektrophorese (SDS-PAGE) analytisch nach ihrer Größe aufgetrennt werden. Ein Polyacrylamid-Gel besteht aus vernetztem Acrylamid und Bisacrylamid und weist wesentlich kleinere Poren als das im Kapitel 5.2.1.9 beschriebene Agarose-Gel auf. Bei der hier verwendeten denaturierenden PAGE werden die Proteine mit SDS umschlossen und so auf eine gleiche Ladung überführt.

Die Acrylamidgele setzten sich wie folgt zusammen:

Tab. 32: Zusammensetzung der SDS Acrylamidgele

Substanz	Sammelgel	Trenngel
Acrylamid	4%	7,5%-15% je nach Proteingröße
0,5 M Tris/HCl pH 6,8	12,5 ml	-
1,5 M Tris/HCl pH 8,9	-	12,5 ml
SDS 10%	0,5 ml	0,5 ml
H ₂ O bidest.	ad 50 ml	ad 50 ml

Von den vorbereiteten Lösungen werden für pro Ansatz 8 ml Trenngel mit 8 µl Tetramethylethylenediamine (TEMED) und 80 µl Ammoniumpersulfat (APS, 10% (w/v) in H₂O) versetzt, durch Invertieren vermischt und in die mit Isopropanol gereinigten Gelkammern gegossen. Um eine gerade Kante zu erzeugen, wird das noch flüssige Gel mit Isopropanol überschichtet. Sobald das Trenngel ausgehärtet ist, wird das Isopropanol abgegossen. Es wird mit etwa 3 ml Sammelgel aufgefüllt, das mit 3 µl TEMED und 30 µl APS vermischt wurde, und der Kamm wird eingesetzt, um die Taschen auszusparen. Nach etwa 30 min ist das Gel so weit ausgehärtet, dass es beladen werden kann.

Vor der Beladung des Gels werden die Proben mit Lämmli-Ladepuffer versetzt, 5 min bei 95°C denaturiert und 3 min bei 17.000 g zentrifugiert. Zur Molekularmassenbestimmung der Proteine wird parallel ein Protein-Größenstandard (**Tab. 12**) aufgetragen. Während die Proben das Sammelgel passieren, wird eine Spannung von 50 V angelegt, bei Erreichen des Trenngels wird diese auf 90 V erhöht bis die Lauffront das Ende des Gels erreicht hat.

5.2.3.5 Immunoblot

Beim Blotting werden durch SDS-PAGE aufgetrennte Proteine von einem Gel auf eine Trägermembran aus Nitrocellulose, Nylon oder Polyvinylidendifluorid (PVDF) elektrophoretisch übertragen. Das Auftrennungsmuster, das durch die SDS-PAGE entstanden ist, bleibt so erhalten, aber die Proteine werden für Antikörper zugänglich, so dass sie spezifisch detektiert werden können. Für diese Arbeit erfolgte der Transfer in einem Wet-Blot-System auf PVDF-Membranen.

Nach dem Transfer wurden unspezifische Bindestellen durch Inkubation mit Blockierungspuffer (vgl. **Tab. 13**) für 1 h abgesättigt. Anschließend wurde der Primärantikörper in Blockierungspuffer auf die Membran gegeben und über Nacht bei 4°C unter Schwenken inkubiert. Nach dem Waschen (1x 5 min mit PBS-T, 2x 5 min mit PBS, vgl. **Tab. 12**) wurde der Sekundärantikörper in Blockierungspuffer hinzugegeben, für 2 h bei Raumtemperatur inkubiert und erneut gewaschen.

Um die Proteinbanden nun sichtbar zu machen sind die hier benutzten Sekundärantikörper an die Meerrettichperoxidase (*horseradish peroxidase*, HRP) gekoppelt, die die Oxidation von Luminol katalysiert. Die Detektion der HRP-Aktivität erfolgte mit ECL-, bzw. ECLplus-Reagenz gemäß den Angaben des Herstellers (Thermo Fisher, Dreieich).

Sollte eine Membran auf die Expression mehrerer unterschiedlicher Proteine hin getestet werden, wurden die zuerst gebundenen Antikörper mit der *ReBlot Plus Strong Antibody Stripping Solution* (Millipore, Schwalbach) entfernt. Hierzu wird die Membran 15 min bei Raumtemperatur in der Lösung inkubiert und anschließend 3x 5min in PBS gewaschen. Danach ist eine Inkubation mit einem weiteren Erstantikörper möglich.

5.2.4 Bioinformatik

5.2.4.1 Structured Query Language (SQL)

Die Systemarchitektur der TUBA-Datenbank und des dazugehörigen Frontends wurde in der Masterarbeit von Maximilian Deichelbohrer ausführlich beschrieben ²⁰⁰. Zur Bearbeitung und für Abfragen aus der PostgreSQL-Datenbank wurde die Datenbanksprache *Structured Query Language* (SQL) verwendet, die von fast allen gängigen Datenbanksystemen unterstützt wird. In SQL gibt es drei Kategorien von Befehlen: Befehle zur Datenmanipulation und dem lesenden Zugriff (DML), zur Definition des Datenbankschemas (DDL) und zur Rechteverwaltung und Transaktionskontrolle (DCL). Zur Datenextraktion wird vor allem der DML-Befehl SELECT verwendet, mit den Ergebnistabellen definiert und mit Informationen aus der Datenbank gefüllt werden können. Die Selektion der Datensätze basiert auf den Grundsätzen der Relationsmathematik, welche es ermöglicht, die Systemarchitektur mathematisch abzubilden und Tabellenbeziehungen auswertbar zu machen.

Zusätzlich können in den Befehl SELECT einige Hilfsfunktionen eingebunden werden, die es ermöglichen, Operationen wie zusätzliche Berechnungen und Sortierungen vor der Generierung der Ergebnistabelle durchzuführen. Durch die Verschachtelung solcher Befehle und Operationen können hochkomplexe Abfragen formuliert werden. Die Ergebnistabelle kann in verschiedenen Formaten ausgegeben werden, hier wurde das .csv-Dateiformat gewählt.

Hinter dem SELECT Befehl werden die angezeigten Spalten über zu bestimmende Parameter definiert, während der weitere Hauptbestandteil des Befehls 'FROM' die Tabelle nominiert, aus der die Daten stammen. Zusätzlich kann 'WHERE' als Filter eingesetzt werden, um Relationen zu definieren. Um mehrere Tabellen zusammenzuführen, wird 'JOIN' verwendet.

Über die Ubuntu-Konsole wurden in PostgreSQL die Befehle über das Terminal ausgeführt. Die Rootrechte auf den Server wurden über SSH gesichert und verifiziert. Um den aktuellsten Stand der Daten zur Verfügung zu haben, wurde jeweils ein SQL-Dump vom Server heruntergeladen und ein lokales Abbild der Datenbank erstellt.

5.2.4.1.1 Datenselektion über Primärtumore

Anzahl betrachteter Fälle: 492

Basierend auf dem SQL-QUERY:

```
SELECT c.*, t.* FROM cases as c  
  
JOIN tumors as t on t.c_id = c.id WHERE t.number  
= 1 and t.kind = 'Primärtumor' and (c.deceased is  
not null or c.lastcontactdate is not null) and  
(c.remark not like '%löschen%' and c.remark not  
like '%test%' or c.remark is null) and (t.remark  
not like '%löschen%' and t.remark not like  
'%test%' or t.remark is null)
```

In diesem SQL-QUERY werden alle Daten aus Fällen (c.*) und Tumoren (t.*) miteinander vereint, in denen ein Primärtumor eingegeben wurde, entweder der Überlebensstatus (lebendig oder tot) oder der letzte Kontakt eingegeben wurde, und der Fall nicht zum Löschen oder Testen markiert wurde.

5.2.4.1.2 Datenselektion über Fälle

Anzahl betrachteter Fälle: 551

Basierend auf dem SQL-QUERY:

```
SELECT c.*, t.* FROM cases as c
JOIN tumors as t on t.c_id = c.id WHERE
(c.deceased is not null or c.lastcontactdate is
not null) and t.number= 1 and (c.remark not
like '%löschen%' and c.remark not like '%test%' or
c.remark is null) and (t.remark not
like '%löschen%' and t.remark not like '%test%' or
t.remark is null)
```

In diesem SQL-QUERY werden alle Daten aus Fällen (c.*) und Tumoren (t.*) miteinander vereint, in denen entweder der Überlebensstatus (lebendig oder tot) oder der letzte Kontakt eingegeben wurde, und der Fall nicht zum Löschen oder Testen markiert wurde.

5.2.4.2 Statistische Programmiersprache R

R ist eine Statistik-orientierte Programmiersprache mit spezifischen, für die Statistik entworfenen Datenstrukturen und Funktionen, die nicht kompiliert werden muss und in Skripten ausgeführt werden kann. Sofern nicht anders angegeben, wurde für alle Auswertungen dieser Arbeit das *limma*-Paket verwendet.

6 REFERENZEN

- 1 WHO. (World Health Organisation, 2017).
- 2 Torre, L. A. *et al.* Global cancer statistics, 2012. *CA: a cancer journal for clinicians* **65**, 87-108, doi:10.3322/caac.21262 (2015).
- 3 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).
- 4 Alberts B, J. A., Lewis J, et al. *Molecular Biology of the Cell*. 4th edition edn, (Garland Science, 2002).
- 5 Knippers, R. *Molekulare Genetik*. (Georg Thieme Verlag, 2006).
- 6 Dinger, M. E., Pang, K. C., Mercer, T. R. & Mattick, J. S. Differentiating protein-coding and noncoding RNA: challenges and ambiguities. *PLoS Comput Biol* **4**, e1000176, doi:10.1371/journal.pcbi.1000176 (2008).
- 7 Sainsbury, S., Bernecky, C. & Cramer, P. Structural basis of transcription initiation by RNA polymerase II. *Nature reviews* **16**, 129-143, doi:10.1038/nrm3952 (2015).
- 8 Charles Molnar, J. G. *Concepts of Biology*. 1st Canadian Edition edn, (Rice University, 2015).
- 9 Nogales, E., Fang, J. & Louder, R. K. Structural dynamics and DNA interaction of human TFIID. *Transcription* **8**, 55-60, doi:10.1080/21541264.2016.1265701 (2017).
- 10 Basehoar, A. D., Zanton, S. J. & Pugh, B. F. Identification and distinct regulation of yeast TATA box-containing genes. *Cell* **116**, 699-709 (2004).
- 11 Rhee, H. S. & Pugh, B. F. Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature* **483**, 295-301, doi:10.1038/nature10799 (2012).
- 12 Kim, J. L., Nikolov, D. B. & Burley, S. K. Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **365**, 520-527, doi:10.1038/365520a0 (1993).
- 13 Bieniossek, C. *et al.* The architecture of human general transcription factor TFIID core complex. *Nature* **493**, 699-702, doi:10.1038/nature11791 (2013).
- 14 Kang, J. J., Auble, D. T., Ranish, J. A. & Hahn, S. Analysis of the yeast transcription factor TFIIA: distinct functional regions and a polymerase II-specific role in basal and activated transcription. *Molecular and cellular biology* **15**, 1234-1243 (1995).
- 15 Hoiby, T., Zhou, H., Mitsiou, D. J. & Stunnenberg, H. G. A facelift for the general transcription factor TFIIA. *Biochim Biophys Acta* **1769**, 429-436, doi:10.1016/j.bbaexp.2007.04.008 (2007).
- 16 Imbalzano, A. N., Zaret, K. S. & Kingston, R. E. Transcription factor (TF) IIB and TFIIA can independently increase the affinity of the TATA-binding protein for DNA. *The Journal of biological chemistry* **269**, 8280-8286 (1994).
- 17 Bleichenbacher, M., Tan, S. & Richmond, T. J. Novel interactions between the components of human and yeast TFIIA/TBP/DNA complexes. *Journal of molecular biology* **332**, 783-793 (2003).
- 18 Ozer, J., Mitsouras, K., Zerby, D., Carey, M. & Lieberman, P. M. Transcription factor IIA derepresses TATA-binding protein (TBP)-associated factor inhibition of TBP-DNA binding. *The Journal of biological chemistry* **273**, 14293-14300 (1998).
- 19 Hoiby, T. *et al.* Cleavage and proteasome-mediated degradation of the basal transcription factor TFIIA. *The EMBO journal* **23**, 3083-3091, doi:10.1038/sj.emboj.7600304 (2004).

REFERENZEN

- 20 Aperlo, C., Boulukos, K. E. & Pognonec, P. The basic region/helix-loop-helix/leucine repeat transcription factor USF interferes with Ras transformation. *Eur J Biochem* **241**, 249-253 (1996).
- 21 Deng, W., Malecova, B., Oelgeschlager, T. & Roberts, S. G. TFIIIB recognition elements control the TFIIA-NC2 axis in transcriptional regulation. *Molecular and cellular biology* **29**, 1389-1400, doi:10.1128/MCB.01346-08 (2009).
- 22 Plaschka, C. *et al.* Transcription initiation complex structures elucidate DNA opening. *Nature* **533**, 353-358, doi:10.1038/nature17990 (2016).
- 23 Reina, J. H. & Hernandez, N. On a roll for new TRF targets. *Genes & development* **21**, 2855-2860, doi:10.1101/gad.1623207 (2007).
- 24 Zehavi, Y., Kedmi, A., Ideses, D. & Juven-Gershon, T. TRF2: TRansForming the view of general transcription factors. *Transcription* **6**, 1-6, doi:10.1080/21541264.2015.1004980 (2015).
- 25 Teichmann, M. *et al.* Human TATA-binding protein-related factor-2 (hTRF2) stably associates with hTFIIA in HeLa cells. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 13720-13725 (1999).
- 26 Nakadai, T., Shimada, M., Shima, D., Handa, H. & Tamura, T. A. Specific interaction with transcription factor IIA and localization of the mammalian TATA-binding protein-like protein (TLP/TRF2/TLF). *The Journal of biological chemistry* **279**, 7447-7455, doi:10.1074/jbc.M305412200 (2004).
- 27 Isogai, M., Suzuki, H., Maeda, R. & Tamura, T. A. Ubiquitin-proteasome-dependent degradation of TBP-like protein is prevented by direct binding of TFIIA. *Genes Cells* **21**, 1223-1232, doi:10.1111/gtc.12441 (2016).
- 28 Schones, D. E. & Zhao, K. Genome-wide approaches to studying chromatin modifications. *Nat Rev Genet* **9**, 179-191, doi:10.1038/nrg2270 (2008).
- 29 Turner, B. M. Histone acetylation and an epigenetic code. *Bioessays* **22**, 836-845, doi:10.1002/1521-1878(200009)22:9<836::AID-BIES9>3.0.CO;2-X (2000).
- 30 Martens, J. H. *et al.* The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *The EMBO journal* **24**, 800-812, doi:10.1038/sj.emboj.7600545 (2005).
- 31 Gardiner-Garden, M. & Frommer, M. CpG islands in vertebrate genomes. *Journal of molecular biology* **196**, 261-282 (1987).
- 32 Ioshikhes, I. P. & Zhang, M. Q. Large-scale human promoter mapping using CpG islands. *Nature genetics* **26**, 61-63, doi:10.1038/79189 (2000).
- 33 Bird, A. DNA methylation patterns and epigenetic memory. *Genes & development* **16**, 6-21, doi:10.1101/gad.947102 (2002).
- 34 Kim, T. H. *et al.* A high-resolution map of active promoters in the human genome. *Nature* **436**, 876-880, doi:10.1038/nature03877 (2005).
- 35 Roh, T. Y., Cuddapah, S. & Zhao, K. Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping. *Genes & development* **19**, 542-552, doi:10.1101/gad.1272505 (2005).
- 36 Shah, N. & Sukumar, S. The Hox genes and their roles in oncogenesis. *Nature reviews. Cancer* **10**, 361-371, doi:10.1038/nrc2826 (2010).
- 37 Sparmann, A. & van Lohuizen, M. Polycomb silencers control cell fate, development and cancer. *Nature reviews. Cancer* **6**, 846-856, doi:10.1038/nrc1991 (2006).
- 38 Krivtsov, A. V. & Armstrong, S. A. MLL translocations, histone modifications and leukaemia stem-cell development. *Nature reviews. Cancer* **7**, 823-833, doi:10.1038/nrc2253 (2007).
- 39 Schuettengruber, B., Bourbon, H. M., Di Croce, L. & Cavalli, G. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell* **171**, 34-57, doi:10.1016/j.cell.2017.08.002 (2017).

REFERENZEN

- 40 Biondi, A. *et al.* Unique genotypic features of infant acute lymphoblastic leukaemia at presentation and at relapse. *Br J Haematol* **80**, 472-479 (1992).
- 41 Huret, J. L., Dessen, P. & Bernheim, A. An atlas of chromosomes in hematological malignancies. Example: 11q23 and MLL partners. *Leukemia* **15**, 987-989 (2001).
- 42 Butler, L. H., Slany, R., Cui, X., Cleary, M. L. & Mason, D. Y. The HRX proto-oncogene product is widely expressed in human tissues and localizes to nuclear structures. *Blood* **89**, 3361-3370 (1997).
- 43 Guenther, M. G. *et al.* Global and Hox-specific roles for the MLL1 methyltransferase. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 8603-8608, doi:10.1073/pnas.0503072102 (2005).
- 44 Milne, T. A. *et al.* MLL associates specifically with a subset of transcriptionally active target genes. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 14765-14770, doi:10.1073/pnas.0503630102 (2005).
- 45 Hsieh, J. J., Cheng, E. H. & Korsmeyer, S. J. Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. *Cell* **115**, 293-303 (2003).
- 46 Birke, M. *et al.* The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. *Nucleic acids research* **30**, 958-965 (2002).
- 47 Zeleznik-Le, N. J., Harden, A. M. & Rowley, J. D. 11q23 translocations split the "AT-hook" cruciform DNA-binding region and the transcriptional repression domain from the activation domain of the mixed-lineage leukemia (MLL) gene. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 10610-10614 (1994).
- 48 Yokoyama, A. *et al.* Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Molecular and cellular biology* **24**, 5639-5649, doi:10.1128/MCB.24.13.5639-5649.2004 (2004).
- 49 Ernst, P., Wang, J., Huang, M., Goodman, R. H. & Korsmeyer, S. J. MLL and CREB bind cooperatively to the nuclear coactivator CREB-binding protein. *Molecular and cellular biology* **21**, 2249-2258, doi:10.1128/MCB.21.7.2249-2258.2001 (2001).
- 50 Milne, T. A. *et al.* MLL targets SET domain methyltransferase activity to Hox gene promoters. *Molecular cell* **10**, 1107-1117 (2002).
- 51 Pless, B. *et al.* The heterodimerization domains of MLL-FYRN and FYRC--are potential target structures in t(4;11) leukemia. *Leukemia* **25**, 663-670, doi:10.1038/leu.2010.308 (2011).
- 52 Meyer, C. *et al.* The MLL recombinome of acute leukemias in 2017. *Leukemia*, doi:10.1038/leu.2017.213 (2017).
- 53 Thomas, M. *et al.* Targeting MLL-AF4 with short interfering RNAs inhibits clonogenicity and engraftment of t(4;11)-positive human leukemic cells. *Blood* **106**, 3559-3566, doi:10.1182/blood-2005-03-1283 (2005).
- 54 Andersson, A. K. *et al.* The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. *Nature genetics* **47**, 330-337, doi:10.1038/ng.3230 (2015).
- 55 Ballabio, E. & Milne, T. A. Epigenetic control of gene expression in leukemogenesis: Cooperation between wild type MLL and MLL fusion proteins. *Mol Cell Oncol* **1**, e955330, doi:10.1080/23723548.2014.955330 (2014).
- 56 Okuda, H. *et al.* Cooperative gene activation by AF4 and DOT1L drives MLL-rearranged leukemia. *The Journal of clinical investigation* **127**, 1918-1931, doi:10.1172/JCI91406 (2017).
- 57 Godfrey, L. *et al.* MLL-AF4 binds directly to a BCL-2 specific enhancer and modulates H3K27 acetylation. *Exp Hematol* **47**, 64-75, doi:10.1016/j.exphem.2016.11.003 (2017).
- 58 Kowarz, E. *et al.* Complex MLL rearrangements in t(4;11) leukemia patients with absent AF4.MLL fusion allele. *Leukemia* **21**, 1232-1238, doi:10.1038/sj.leu.2404686 (2007).

REFERENZEN

- 59 Bursen, A. *et al.* The AF4.MLL fusion protein is capable of inducing ALL in mice without requirement of MLL.AF4. *Blood* **115**, 3570-3579, doi:10.1182/blood-2009-06-229542 (2010).
- 60 Benedikt, A. *et al.* The leukemogenic AF4-MLL fusion protein causes P-TEFb kinase activation and altered epigenetic signatures. *Leukemia* **25**, 135-144, doi:10.1038/leu.2010.249 (2011).
- 61 Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57-70 (2000).
- 62 Horne, S. D., Pollick, S. A. & Heng, H. H. Evolutionary mechanism unifies the hallmarks of cancer. *International journal of cancer* **136**, 2012-2021, doi:10.1002/ijc.29031 (2015).
- 63 Joyce, J. A. & Pollard, J. W. Microenvironmental regulation of metastasis. *Nature reviews. Cancer* **9**, 239-252, doi:10.1038/nrc2618 (2009).
- 64 Hanahan, D. & Coussens, L. M. Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer cell* **21**, 309-322, doi:10.1016/j.ccr.2012.02.022 (2012).
- 65 Grivennikov, S. I., Greten, F. R. & Karin, M. Immunity, inflammation, and cancer. *Cell* **140**, 883-899, doi:10.1016/j.cell.2010.01.025 (2010).
- 66 Vajdic, C. M. & van Leeuwen, M. T. Cancer incidence and risk factors after solid organ transplantation. *International journal of cancer* **125**, 1747-1754, doi:10.1002/ijc.24439 (2009).
- 67 Schulz, T. F. Cancer and viral infections in immunocompromised individuals. *International journal of cancer* **125**, 1755-1763, doi:10.1002/ijc.24741 (2009).
- 68 Cook, J. & Hagemann, T. Tumour-associated macrophages and cancer. *Curr Opin Pharmacol* **13**, 595-601, doi:10.1016/j.coph.2013.05.017 (2013).
- 69 Hagemann, T. *et al.* "Re-educating" tumor-associated macrophages by targeting NF-kappaB. *J Exp Med* **205**, 1261-1268, doi:10.1084/jem.20080108 (2008).
- 70 Mosser, D. M. & Edwards, J. P. Exploring the full spectrum of macrophage activation. *Nat Rev Immunol* **8**, 958-969, doi:10.1038/nri2448 (2008).
- 71 Escribese, M. M., Casas, M. & Corbi, A. L. Influence of low oxygen tensions on macrophage polarization. *Immunobiology* **217**, 1233-1240, doi:10.1016/j.imbio.2012.07.002 (2012).
- 72 Fukumura, D. *et al.* Tumor induction of VEGF promoter activity in stromal cells. *Cell* **94**, 715-725 (1998).
- 73 Koontongkaew, S., Amornphimoltham, P., Monthanpisut, P., Saensuk, T. & Leelakriangsak, M. Fibroblasts and extracellular matrix differently modulate MMP activation by primary and metastatic head and neck cancer cells. *Medical oncology* **29**, 690-703, doi:10.1007/s12032-011-9871-6 (2012).
- 74 Zhang, X. H. *et al.* Selection of bone metastasis seeds by mesenchymal signals in the primary tumor stroma. *Cell* **154**, 1060-1073, doi:10.1016/j.cell.2013.07.036 (2013).
- 75 Ishikawa, T. *et al.* Hypoxia enhances CXCR4 expression by activating HIF-1 in oral squamous cell carcinoma. *Oncology reports* **21**, 707-712 (2009).
- 76 Mammoto, T. & Ingber, D. E. Mechanical control of tissue and organ development. *Development* **137**, 1407-1420, doi:10.1242/dev.024166 (2010).
- 77 Pickup, M. W., Mouw, J. K. & Weaver, V. M. The extracellular matrix modulates the hallmarks of cancer. *EMBO reports* **15**, 1243-1253, doi:10.15252/embr.201439246 (2014).
- 78 Hynes, R. O. The extracellular matrix: not just pretty fibrils. *Science* **326**, 1216-1219, doi:10.1126/science.1176009 (2009).
- 79 Frantz, C., Stewart, K. M. & Weaver, V. M. The extracellular matrix at a glance. *J Cell Sci* **123**, 4195-4200, doi:10.1242/jcs.023820 (2010).
- 80 Parri, M. & Chiarugi, P. Rac and Rho GTPases in cancer cell motility control. *Cell Commun Signal* **8**, 23, doi:10.1186/1478-811X-8-23 (2010).

REFERENZEN

- 81 Hamidi, A. *et al.* TGF-beta promotes PI3K-AKT signaling and prostate cancer cell migration through the TRAF6-mediated ubiquitylation of p85alpha. *Sci Signal* **10**, doi:10.1126/scisignal.aal4186 (2017).
- 82 Bakin, A. V., Tomlinson, A. K., Bhowmick, N. A., Moses, H. L. & Arteaga, C. L. Phosphatidylinositol 3-kinase function is required for transforming growth factor beta-mediated epithelial to mesenchymal transition and cell migration. *The Journal of biological chemistry* **275**, 36803-36810, doi:10.1074/jbc.M005912200 (2000).
- 83 Grassian, A. R., Coloff, J. L. & Brugge, J. S. Extracellular matrix regulation of metabolism and implications for tumorigenesis. *Cold Spring Harb Symp Quant Biol* **76**, 313-324, doi:10.1101/sqb.2011.76.010967 (2011).
- 84 Samuels, Y. & Ericson, K. Oncogenic PI3K and its role in cancer. *Curr Opin Oncol* **18**, 77-82 (2006).
- 85 Fruman, D. A. & Rommel, C. PI3K and cancer: lessons, challenges and opportunities. *Nature reviews. Drug discovery* **13**, 140-156, doi:10.1038/nrd4204 (2014).
- 86 Massague, J. & Obenauf, A. C. Metastatic colonization by circulating tumour cells. *Nature* **529**, 298-306, doi:10.1038/nature17038 (2016).
- 87 Nagrath, S. *et al.* Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature* **450**, 1235-1239, doi:10.1038/nature06385 (2007).
- 88 Braun, S. *et al.* A pooled analysis of bone marrow micrometastasis in breast cancer. *The New England journal of medicine* **353**, 793-802, doi:10.1056/NEJMoa050434 (2005).
- 89 Lopez-Otin, C. & Bond, J. S. Proteases: multifunctional enzymes in life and disease. *The Journal of biological chemistry* **283**, 30433-30437, doi:10.1074/jbc.R800035200 (2008).
- 90 Li, J. & Yuan, J. Caspases in apoptosis and beyond. *Oncogene* **27**, 6194-6206, doi:10.1038/onc.2008.297 (2008).
- 91 Mason, S. D. & Joyce, J. A. Proteolytic networks in cancer. *Trends in cell biology* **21**, 228-237, doi:10.1016/j.tcb.2010.12.002 (2011).
- 92 Mian, B. M. *et al.* Fully human anti-interleukin 8 antibody inhibits tumor growth in orthotopic bladder cancer xenografts via down-regulation of matrix metalloproteases and nuclear factor-kappaB. *Clinical cancer research : an official journal of the American Association for Cancer Research* **9**, 3167-3175 (2003).
- 93 Gocheva, V. *et al.* IL-4 induces cathepsin protease activity in tumor-associated macrophages to promote cancer growth and invasion. *Genes & development* **24**, 241-255, doi:10.1101/gad.1874010 (2010).
- 94 Wunsch, D. *et al.* Taspase1: a 'misunderstood' protease with translational cancer relevance. *Oncogene* **35**, 3351-3364, doi:10.1038/onc.2015.436 (2016).
- 95 Overall, C. M. & Dean, R. A. Degradomics: systems biology of the protease web. Pleiotropic roles of MMPs in cancer. *Cancer metastasis reviews* **25**, 69-75, doi:10.1007/s10555-006-7890-0 (2006).
- 96 Varfolomeev, E. E. *et al.* Targeted disruption of the mouse Caspase 8 gene ablates cell death induction by the TNF receptors, Fas/Apo1, and DR3 and is lethal prenatally. *Immunity* **9**, 267-276 (1998).
- 97 Droga-Mazovec, G. *et al.* Cysteine cathepsins trigger caspase-dependent cell death through cleavage of bid and antiapoptotic Bcl-2 homologues. *The Journal of biological chemistry* **283**, 19140-19150, doi:10.1074/jbc.M802513200 (2008).
- 98 Caglic, D., Pungercar, J. R., Pejler, G., Turk, V. & Turk, B. Glycosaminoglycans facilitate procathepsin B activation through disruption of propeptide-mature enzyme interactions. *The Journal of biological chemistry* **282**, 33076-33085, doi:10.1074/jbc.M705761200 (2007).

REFERENZEN

- 99 Sloane, B. F. *et al.* Cathepsin B and tumor proteolysis: contribution of the tumor microenvironment. *Semin Cancer Biol* **15**, 149-157, doi:10.1016/j.semcancer.2004.08.001 (2005).
- 100 Skrzydlewska, E., Sulkowska, M., Koda, M. & Sulkowski, S. Proteolytic-antiproteolytic balance and its regulation in carcinogenesis. *World journal of gastroenterology : WJG* **11**, 1251-1266 (2005).
- 101 Kobayashi, H. *et al.* Cathepsin B efficiently activates the soluble and the tumor cell receptor-bound form of the proenzyme urokinase-type plasminogen activator (Pro-uPA). *The Journal of biological chemistry* **266**, 5147-5152 (1991).
- 102 Sotiropoulou, G., Pampalakis, G. & Diamandis, E. P. Functional roles of human kallikrein-related peptidases. *The Journal of biological chemistry* **284**, 32989-32994, doi:10.1074/jbc.R109.027946 (2009).
- 103 He, Y. *et al.* Interaction between cancer cells and stromal fibroblasts is required for activation of the uPAR-uPA-MMP-2 cascade in pancreatic cancer metastasis. *Clinical cancer research : an official journal of the American Association for Cancer Research* **13**, 3115-3124, doi:10.1158/1078-0432.CCR-06-2088 (2007).
- 104 Ra, H. J. & Parks, W. C. Control of matrix metalloproteinase catalytic activity. *Matrix Biol* **26**, 587-596, doi:10.1016/j.matbio.2007.07.001 (2007).
- 105 Makowski, G. S. & Ramsby, M. L. Autoactivation profiles of calcium-dependent matrix metalloproteinase-2 and -9 in inflammatory synovial fluid: effect of pyrophosphate and bisphosphonates. *Clin Chim Acta* **358**, 182-191, doi:10.1016/j.cccn.2005.03.012 (2005).
- 106 Gu, Z. *et al.* S-nitrosylation of matrix metalloproteinases: signaling pathway to neuronal cell death. *Science* **297**, 1186-1190, doi:10.1126/science.1073634 (2002).
- 107 Dean, R. A. *et al.* Identification of candidate angiogenic inhibitors processed by matrix metalloproteinase 2 (MMP-2) in cell-based proteomic screens: disruption of vascular endothelial growth factor (VEGF)/heparin affinity regulatory peptide (pleiotrophin) and VEGF/Connective tissue growth factor angiogenic inhibitory complexes by MMP-2 proteolysis. *Molecular and cellular biology* **27**, 8454-8465, doi:10.1128/MCB.00821-07 (2007).
- 108 Lopez-Otin, C. & Hunter, T. The regulatory crosstalk between kinases and proteases in cancer. *Nature reviews. Cancer* **10**, 278-292, doi:10.1038/nrc2823 (2010).
- 109 Weathington, N. M. & Mallampalli, R. K. Emerging therapies targeting the ubiquitin proteasome system in cancer. *The Journal of clinical investigation* **124**, 6-12, doi:10.1172/JCI71602 (2014).
- 110 Zhang, F., Paterson, A. J., Huang, P., Wang, K. & Kudlow, J. E. Metabolic control of proteasome function. *Physiology (Bethesda)* **22**, 373-379, doi:10.1152/physiol.00026.2007 (2007).
- 111 Bertoli, C., Skotheim, J. M. & de Bruin, R. A. Control of cell cycle transcription during G1 and S phases. *Nature reviews* **14**, 518-528, doi:10.1038/nrm3629 (2013).
- 112 Lu, Z. & Hunter, T. Degradation of activated protein kinases by ubiquitination. *Annual review of biochemistry* **78**, 435-475, doi:10.1146/annurev.biochem.013008.092711 (2009).
- 113 Niendorf, S. *et al.* Essential role of ubiquitin-specific protease 8 for receptor tyrosine kinase stability and endocytic trafficking in vivo. *Molecular and cellular biology* **27**, 5029-5039, doi:10.1128/MCB.01566-06 (2007).
- 114 Ancot, F., Foveau, B., Lefebvre, J., Leroy, C. & Tulasne, D. Proteolytic cleavages give receptor tyrosine kinases the gift of ubiquity. *Oncogene* **28**, 2185-2195, doi:10.1038/onc.2009.88 (2009).

REFERENZEN

- 115 Strohecker, A. M., Yehiely, F., Chen, F. & Cryns, V. L. Caspase cleavage of HER-2 releases a Bad-like cell death effector. *The Journal of biological chemistry* **283**, 18269-18282, doi:10.1074/jbc.M802156200 (2008).
- 116 Cardone, M. H. *et al.* Regulation of cell death protease caspase-9 by phosphorylation. *Science* **282**, 1318-1321 (1998).
- 117 Murphy, G. The ADAMs: signalling scissors in the tumour microenvironment. *Nature reviews. Cancer* **8**, 929-941, doi:10.1038/nrc2459 (2008).
- 118 Ni, C. Y., Murphy, M. P., Golde, T. E. & Carpenter, G. gamma -Secretase cleavage and nuclear localization of ErbB-4 receptor tyrosine kinase. *Science* **294**, 2179-2181, doi:10.1126/science.1065412 (2001).
- 119 Hsieh, J. J., Ernst, P., Erdjument-Bromage, H., Tempst, P. & Korsmeyer, S. J. Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. *Molecular and cellular biology* **23**, 186-194 (2003).
- 120 Takeda, S. *et al.* Proteolysis of MLL family proteins is essential for taspase1-orchestrated cell cycle progression. *Genes & development* **20**, 2397-2409, doi:10.1101/gad.1449406 (2006).
- 121 Khan, J. A., Dunn, B. M. & Tong, L. Crystal structure of human Taspase1, a crucial protease regulating the function of MLL. *Structure* **13**, 1443-1452, doi:10.1016/j.str.2005.07.006 (2005).
- 122 Bier, C. *et al.* Cell-based analysis of structure-function activity of threonine aspartase 1. *The Journal of biological chemistry* **286**, 3007-3017, doi:10.1074/jbc.M110.161646 (2011).
- 123 Zhou, H. *et al.* Uncleaved TFIIA is a substrate for taspase 1 and active in transcription. *Molecular and cellular biology* **26**, 2728-2735, doi:10.1128/MCB.26.7.2728-2735.2006 (2006).
- 124 Oyama, T. *et al.* Cleavage of TFIIA by Taspase1 activates TRF2-specified mammalian male germ cell programs. *Dev Cell* **27**, 188-200, doi:10.1016/j.devcel.2013.09.025 (2013).
- 125 Takeda, S. *et al.* Taspase1-dependent TFIIA cleavage coordinates head morphogenesis by limiting Cdkn2a locus transcription. *The Journal of clinical investigation* **125**, 1203-1214, doi:10.1172/JCI77075 (2015).
- 126 Stauber, R. H., Mann, W. & Knauer, S. K. Nuclear and cytoplasmic survivin: molecular mechanism, prognostic, and therapeutic potential. *Cancer research* **67**, 5999-6002, doi:10.1158/0008-5472.CAN-07-0494 (2007).
- 127 Marchant, D. J. *et al.* A new transcriptional role for matrix metalloproteinase-12 in antiviral immunity. *Nature medicine* **20**, 493-502, doi:10.1038/nm.3508 (2014).
- 128 Gorlich, D. & Mattaj, I. W. Nucleocytoplasmic transport. *Science* **271**, 1513-1518 (1996).
- 129 Knauer, S. K., Carra, G. & Stauber, R. H. Nuclear export is evolutionarily conserved in CVC paired-like homeobox proteins and influences protein stability, transcriptional activation, and extracellular secretion. *Molecular and cellular biology* **25**, 2573-2582 (2005).
- 130 Bier, C. *et al.* The importin-alpha/nucleophosmin switch controls taspase1 protease function. *Traffic (Copenhagen, Denmark)* **12**, 703-714, doi:10.1111/j.1600-0854.2011.01191.x (2011).
- 131 Falini, B. *et al.* Altered nucleophosmin transport in acute myeloid leukaemia with mutated NPM1: molecular basis and clinical implications. *Leukemia* **23**, 1731-1743, doi:10.1038/leu.2009.124 (2009).
- 132 Chan, N. & Lim, T. M. Cytoplasmic nucleophosmin has elevated T199 phosphorylation upon which G2/M phase progression is dependent. *Sci Rep* **5**, 11777, doi:10.1038/srep11777 (2015).
- 133 Chen, D. Y. *et al.* Taspase1 functions as a non-oncogene addiction protease that coordinates cancer cell proliferation and apoptosis. *Cancer research* **70**, 5358-5367, doi:10.1158/0008-5472.CAN-10-0027 (2010).

REFERENZEN

- 134 Weinstein, I. B. & Joe, A. Oncogene addiction. *Cancer research* **68**, 3077-3080; discussion 3080, doi:10.1158/0008-5472.CAN-07-3293 (2008).
- 135 Luo, J., Solimini, N. L. & Elledge, S. J. Principles of cancer therapy: oncogene and non-oncogene addiction. *Cell* **136**, 823-837, doi:10.1016/j.cell.2009.02.024 (2009).
- 136 Wunsch, D. *et al.* Chemico-genetic strategies to inhibit the leukemic potential of threonine aspartase-1. *Blood cancer journal* **2**, e77, doi:10.1038/bcj.2012.22 (2012).
- 137 Chen, D. Y. *et al.* A pharmacologic inhibitor of the protease Taspase1 effectively inhibits breast and brain tumor growth. *Cancer research* **72**, 736-746, doi:10.1158/0008-5472.CAN-11-2584 (2012).
- 138 Dong, Y. *et al.* Taspase1 cleaves MLL1 to activate cyclin E for HER2/neu breast tumorigenesis. *Cell Res* **24**, 1354-1366, doi:10.1038/cr.2014.129 (2014).
- 139 Marur, S. & Forastiere, A. A. Head and Neck Squamous Cell Carcinoma: Update on Epidemiology, Diagnosis, and Treatment. *Mayo Clinic proceedings* **91**, 386-396, doi:10.1016/j.mayocp.2015.12.017 (2016).
- 140 Leemans, C. R., Braakhuis, B. J. & Brakenhoff, R. H. The molecular biology of head and neck cancer. *Nature reviews. Cancer* **11**, 9-22, doi:10.1038/nrc2982 (2011).
- 141 Vidal, L. & Gillison, M. L. Human papillomavirus in HNSCC: recognition of a distinct disease type. *Hematol Oncol Clin North Am* **22**, 1125-1142, vii, doi:10.1016/j.hoc.2008.08.006 (2008).
- 142 Pfister, D. G. *et al.* Head and neck cancers, version 2.2013. Featured updates to the NCCN guidelines. *J Natl Compr Canc Netw* **11**, 917-923 (2013).
- 143 van't Veer, L. J. & Bernards, R. Enabling personalized cancer medicine through analysis of gene-expression patterns. *Nature* **452**, 564-570, doi:10.1038/nature06915 (2008).
- 144 Onken, M. D. *et al.* A surprising cross-species conservation in the genomic landscape of mouse and human oral cancer identifies a transcriptional signature predicting metastatic disease. *Clinical cancer research : an official journal of the American Association for Cancer Research* **20**, 2873-2884, doi:10.1158/1078-0432.CCR-14-0205 (2014).
- 145 Markwell, S. M. & Weed, S. A. Tumor and stromal-based contributions to head and neck squamous cell carcinoma invasion. *Cancers* **7**, 382-406, doi:10.3390/cancers7010382 (2015).
- 146 Alizadeh, A. A. *et al.* Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403**, 503-511 (2000).
- 147 Golub, T. R. *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**, 531-537 (1999).
- 148 Bittner, M. *et al.* Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* **406**, 536-540 (2000).
- 149 Huang, E. *et al.* Gene expression predictors of breast cancer outcomes. *Lancet* **361**, 1590-1596 (2003).
- 150 MacDonald, T. J. *et al.* Expression profiling of medulloblastoma: PDGFRA and the RAS/MAPK pathway as therapeutic targets for metastatic disease. *Nature genetics* **29**, 143-152 (2001).
- 151 Chung, C. H. *et al.* Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. *Cancer cell* **5**, 489-500 (2004).
- 152 Ginos, M. A. *et al.* Identification of a gene expression signature associated with recurrent disease in squamous cell carcinoma of the head and neck. *Cancer research* **64**, 55-63 (2004).
- 153 Al Moustafa, A. E. *et al.* Identification of genes associated with head and neck carcinogenesis by cDNA microarray comparison between matched primary normal epithelial and squamous carcinoma cells. *Oncogene* **21**, 2634-2640 (2002).
- 154 Alevizos, I. *et al.* Oral cancer in vivo gene expression profiling assisted by laser capture microdissection and microarray analysis. *Oncogene* **20**, 6196-6204 (2001).

REFERENZEN

- 155 Cromer, A. *et al.* Identification of genes associated with tumorigenesis and metastatic potential of hypopharyngeal cancer by microarray analysis. *Oncogene* **23**, 2484-2498 (2004).
- 156 El-Naggar, A. K. *et al.* Differential expression profiling of head and neck squamous carcinoma: significance in their phenotypic and biological classification. *Oncogene* **21**, 8206-8219 (2002).
- 157 Leethanakul, C. *et al.* Distinct pattern of expression of differentiation and growth-related genes in squamous cell carcinomas of the head and neck revealed by the use of laser capture microdissection and cDNA arrays. *Oncogene* **19**, 3220-3224 (2000).
- 158 Mendez, E. *et al.* Transcriptional expression profiles of oral squamous cell carcinomas. *Cancer* **95**, 1482-1494 (2002).
- 159 Thurlow, J. K. *et al.* Spectral clustering of microarray data elucidates the roles of microenvironment remodeling and immune responses in survival of head and neck squamous cell carcinoma. *J Clin Oncol* **28**, 2881-2888, doi:10.1200/JCO.2009.24.8724 (2010).
- 160 Fakhry, C. *et al.* Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. *Journal of the National Cancer Institute* **100**, 261-269, doi:10.1093/jnci/djn011 (2008).
- 161 Beltz, A. *et al.* [Staging of oropharyngeal carcinomas : New TNM classification as a challenge for head and neck cancer centers]. *HNO* **66**, 375-382, doi:10.1007/s00106-018-0499-0 (2018).
- 162 Doescher, J., Veit, J. A. & Hoffmann, T. K. [The 8th edition of the AJCC Cancer Staging Manual : Updates in otorhinolaryngology, head and neck surgery]. *HNO* **65**, 956-961, doi:10.1007/s00106-017-0391-3 (2017).
- 163 Suh, Y., Amelio, I., Guerrero Urbano, T. & Tavassoli, M. Clinical update on cancer: molecular oncology of head and neck cancer. *Cell death & disease* **5**, e1018, doi:10.1038/cddis.2013.548 (2014).
- 164 Chau, N. G. & Hammerman, P. S. Heads Up! Predictive Gene Signatures in Head and Neck Cancer May Be Coming Soon. *Clinical cancer research : an official journal of the American Association for Cancer Research* **22**, 3710-3712, doi:10.1158/1078-0432.CCR-16-0582 (2016).
- 165 Simon, R. Designs and adaptive analysis plans for pivotal clinical trials of therapeutics and companion diagnostics. *Expert Opin Med Diagn* **2**, 721-729, doi:10.1517/17530059.2.6.721 (2008).
- 166 Chau, N. G. *et al.* Incorporation of Next-Generation Sequencing into Routine Clinical Care to Direct Treatment of Head and Neck Squamous Cell Carcinoma. *Clinical cancer research : an official journal of the American Association for Cancer Research* **22**, 2939-2949, doi:10.1158/1078-0432.CCR-15-2314 (2016).
- 167 Cancer Genome Atlas, N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* **517**, 576-582, doi:10.1038/nature14129 (2015).
- 168 Bossi, P. *et al.* Functional Genomics Uncover the Biology behind the Responsiveness of Head and Neck Squamous Cell Cancer Patients to Cetuximab. *Clinical cancer research : an official journal of the American Association for Cancer Research* **22**, 3961-3970, doi:10.1158/1078-0432.CCR-15-2547 (2016).
- 169 Vargas, A. J. & Harris, C. C. Biomarker development in the precision medicine era: lung cancer as a case study. *Nature reviews. Cancer* **16**, 525-537, doi:10.1038/nrc.2016.56 (2016).
- 170 Gonzalez de Castro, D., Clarke, P. A., Al-Lazikani, B. & Workman, P. Personalized cancer medicine: molecular diagnostics, predictive biomarkers, and drug resistance. *Clinical pharmacology and therapeutics* **93**, 252-259, doi:10.1038/clpt.2012.237 (2013).
- 171 Gulley, M. L. & Morgan, D. R. Molecular oncology testing in resource-limited settings. *J Mol Diagn* **16**, 601-611, doi:10.1016/j.jmoldx.2014.07.002 (2014).

REFERENZEN

- 172 Lam, B. *et al.* Solution-based circuits enable rapid and multiplexed pathogen detection. *Nature communications* **4**, 2001, doi:10.1038/ncomms3001 (2013).
- 173 Eatemadi, A. *et al.* Role of protease and protease inhibitors in cancer pathogenesis and treatment. *Biomed Pharmacother* **86**, 221-231, doi:10.1016/j.biopha.2016.12.021 (2017).
- 174 Perez-Silva, J. G., Espanol, Y., Velasco, G. & Quesada, V. The Degradome database: expanding roles of mammalian proteases in life and disease. *Nucleic acids research* **44**, D351-355, doi:10.1093/nar/gkv1201 (2016).
- 175 Mehlen, P. & Puisieux, A. Metastasis: a question of life or death. *Nature reviews. Cancer* **6**, 449-458, doi:10.1038/nrc1886 (2006).
- 176 van 't Veer, L. J. *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530-536 (2002).
- 177 Chen, J. *et al.* Significance of CD44 expression in head and neck cancer: a systemic review and meta-analysis. *BMC Cancer* **14**, 15, doi:10.1186/1471-2407-14-15 (2014).
- 178 Seino, S. *et al.* CD44(high) /ALDH1(high) head and neck squamous cell carcinoma cells exhibit mesenchymal characteristics and GSK3beta-dependent cancer stem cell properties. *J Oral Pathol Med* **45**, 180-188, doi:10.1111/jop.12348 (2016).
- 179 Khoo, B. L. *et al.* Clinical validation of an ultra high-throughput spiral microfluidics for the detection and enrichment of viable circulating tumor cells. *PloS one* **9**, e99409, doi:10.1371/journal.pone.0099409 (2014).
- 180 Eilers, M. & Eisenman, R. N. Myc's broad reach. *Genes & development* **22**, 2755-2766 (2008).
- 181 Meyer, N. & Penn, L. Z. Reflecting on 25 years with MYC. *Nature reviews. Cancer* **8**, 976-990, doi:10.1038/nrc2231 (2008).
- 182 Ornitz, D. M. & Itoh, N. The Fibroblast Growth Factor signaling pathway. *Wiley Interdiscip Rev Dev Biol* **4**, 215-266, doi:10.1002/wdev.176 (2015).
- 183 Gschwind, A., Fischer, O. M. & Ullrich, A. The discovery of receptor tyrosine kinases: targets for cancer therapy. *Nature reviews. Cancer* **4**, 361-370, doi:10.1038/nrc1360 (2004).
- 184 Parker, B. S., Rautela, J. & Hertzog, P. J. Antitumour actions of interferons: implications for cancer therapy. *Nature reviews. Cancer* **16**, 131-144, doi:10.1038/nrc.2016.14 (2016).
- 185 Gregoire, V., Lefebvre, J. L., Licitra, L., Felip, E. & Group, E.-E.-E. G. W. Squamous cell carcinoma of the head and neck: EHNS-ESMO-ESTRO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* **21 Suppl 5**, v184-186, doi:10.1093/annonc/mdq185 (2010).
- 186 Benson, E., Li, R., Eisele, D. & Fakhry, C. The clinical impact of HPV tumor status upon head and neck squamous cell carcinomas. *Oral oncology* **50**, 565-574, doi:10.1016/j.oraloncology.2013.09.008 (2014).
- 187 Tomasetti, C., Li, L. & Vogelstein, B. Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. *Science* **355**, 1330-1334, doi:10.1126/science.aaf9011 (2017).
- 188 Lefebvre, J. L. Current clinical outcomes demand new treatment options for SCCHN. *Ann Oncol* **16 Suppl 6**, vi7-vi12 (2005).
- 189 Denaro, N., Russi, E. G., Adamo, V. & Merlano, M. C. State-of-the-art and emerging treatment options in the management of head and neck cancer: news from 2013. *Oncology* **86**, 212-229, doi:10.1159/000357712 (2014).
- 190 Rothenberg, S. M. & Ellisen, L. W. The molecular pathogenesis of head and neck squamous cell carcinoma. *The Journal of clinical investigation* **122**, 1951-1957 (2012).
- 191 Cohen, R. B. Current challenges and clinical investigations of epidermal growth factor receptor (EGFR)- and ErbB family-targeted agents in the treatment of head and neck squamous cell carcinoma (HNSCC). *Cancer treatment reviews* **40**, 567-577, doi:10.1016/j.ctrv.2013.10.002 (2014).

REFERENZEN

- 192 Meyer, C. *et al.* The MLL recombinome of acute leukemias in 2013. *Leukemia* **27**, 2165-2176, doi:10.1038/leu.2013.135 (2013).
- 193 Zeleznik-Le, N. J. One Step Forward in the Challenging Arena of MLL-AF4 Leukemia. *Cancer cell* **30**, 657-658, doi:10.1016/j.ccell.2016.10.017 (2016).
- 194 Schultz, K. R. *et al.* Risk- and response-based classification of childhood B-precursor acute lymphoblastic leukemia: a combined analysis of prognostic markers from the Pediatric Oncology Group (POG) and Children's Cancer Group (CCG). *Blood* **109**, 926-935, doi:10.1182/blood-2006-01-024729 (2007).
- 195 Slany, R. K. The molecular mechanics of mixed lineage leukemia. *Oncogene* **35**, 5215-5223, doi:10.1038/onc.2016.30 (2016).
- 196 Gardner, R. *et al.* Acquisition of a CD19-negative myeloid phenotype allows immune escape of MLL-rearranged B-ALL from CD19 CAR-T-cell therapy. *Blood* **127**, 2406-2410, doi:10.1182/blood-2015-08-665547 (2016).
- 197 Turk, B. Targeting proteases: successes, failures and future prospects. *Nature reviews. Drug discovery* **5**, 785-799, doi:10.1038/nrd2092 (2006).
- 198 Babu, M. M., Luscombe, N. M., Aravind, L., Gerstein, M. & Teichmann, S. A. Structure and evolution of transcriptional regulatory networks. *Curr Opin Struct Biol* **14**, 283-291, doi:10.1016/j.sbi.2004.05.004 (2004).
- 199 Doroshow, J. H. & Kummar, S. Translational research in oncology--10 years of progress and future prospects. *Nature reviews. Clinical oncology* **11**, 649-662, doi:10.1038/nrclinonc.2014.158 (2014).
- 200 Deichelbohrer, M. *TuBa - A head and neck cancer database providing input assistance, visualization and R-based analysis*, Johannes Gutenberg University, (2016).
- 201 Wittekind, C. M., Hans-Joachim. *TNM: Klassifikation Maligner Tumoren, 7. Auflage.* (Wiley-VCH, 2013).
- 202 Wittekind, C. *TNM: Klassifikation Maligner Tumoren, 8. Auflage.* (Wiley-VCH, 2017).
- 203 Kuriakose, M. A. *et al.* Selection and validation of differentially expressed genes in head and neck cancer. *Cellular and molecular life sciences : CMLS* **61**, 1372-1383 (2004).
- 204 Paz, H., Pathak, N. & Yang, J. Invading one step at a time: the role of invadopodia in tumor metastasis. *Oncogene* **33**, 4193-4202, doi:10.1038/onc.2013.393 (2014).
- 205 Lohmer, L. L., Kelley, L. C., Hagedorn, E. J. & Sherwood, D. R. Invadopodia and basement membrane invasion in vivo. *Cell adhesion & migration* **8**, 246-255 (2014).
- 206 Rosenthal, E. L. & Matrisian, L. M. Matrix metalloproteases in head and neck cancer. *Head & neck* **28**, 639-648, doi:10.1002/hed.20365 (2006).
- 207 Manders, P. *et al.* Complex of urokinase-type plasminogen activator with its type 1 inhibitor predicts poor outcome in 576 patients with lymph node-negative breast carcinoma. *Cancer* **101**, 486-494 (2004).
- 208 Ohba, K. *et al.* Expression of urokinase-type plasminogen activator, urokinase-type plasminogen activator receptor and plasminogen activator inhibitors in patients with renal cell carcinoma: correlation with tumor associated macrophage and prognosis. *J Urol* **174**, 461-465, doi:10.1097/01.ju.0000165150.46006.92 (2005).
- 209 Chazaud, B. *et al.* Promigratory effect of plasminogen activator inhibitor-1 on invasive breast cancer cell populations. *Am J Pathol* **160**, 237-246 (2002).
- 210 Wakahara, K. *et al.* Transforming growth factor-beta1-dependent activation of Smad2/3 and up-regulation of PAI-1 expression is negatively regulated by Src in SKOV-3 human ovarian cancer cells. *J Cell Biochem* **93**, 437-453, doi:10.1002/jcb.20160 (2004).

REFERENZEN

- 211 Natsuizaka, M. *et al.* IGFBP3 promotes esophageal cancer growth by suppressing oxidative stress in hypoxic tumor microenvironment. *American journal of cancer research* **4**, 29-41 (2014).
- 212 Malhotra, D. *et al.* Transcriptional profiling of stroma from inflamed and resting lymph nodes defines immunological hallmarks. *Nature immunology* **13**, 499-510, doi:10.1038/ni.2262 (2012).
- 213 Cui, J. *et al.* Gene-expression signatures can distinguish gastric cancer grades and stages. *PLoS one* **6**, e17819, doi:10.1371/journal.pone.0017819 (2011).
- 214 Zhou, Y., Gunput, R. A., Adolfs, Y. & Pasterkamp, R. J. MICALs in control of the cytoskeleton, exocytosis, and cell death. *Cellular and molecular life sciences : CMLS* **68**, 4033-4044, doi:10.1007/s00018-011-0787-2 (2011).
- 215 Fang, X., Netzer, M., Baumgartner, C., Bai, C. & Wang, X. Genetic network and gene set enrichment analysis to identify biomarkers related to cigarette smoking and lung cancer. *Cancer treatment reviews* **39**, 77-88, doi:10.1016/j.ctrv.2012.06.001 (2013).
- 216 Ashida, S. *et al.* Expression of novel molecules, MICAL2-PV (MICAL2 prostate cancer variants), increases with high Gleason score and prostate cancer progression. *Clinical cancer research : an official journal of the American Association for Cancer Research* **12**, 2767-2773, doi:10.1158/1078-0432.CCR-05-1995 (2006).
- 217 Mariotti, S. *et al.* MICAL2 is a novel human cancer gene controlling mesenchymal to epithelial transition involved in cancer growth and invasion. *Oncotarget* **7**, 1808-1825, doi:10.18632/oncotarget.6577 (2016).
- 218 Gaggioli, C. *et al.* Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells. *Nature cell biology* **9**, 1392-1400, doi:10.1038/ncb1658 (2007).
- 219 Goetz, J. G. *et al.* Biomechanical remodeling of the microenvironment by stromal caveolin-1 favors tumor invasion and metastasis. *Cell* **146**, 148-163, doi:10.1016/j.cell.2011.05.040 (2011).
- 220 Uchida, D. *et al.* Acquisition of lymph node, but not distant metastatic potentials, by the overexpression of CXCR4 in human oral squamous cell carcinoma. *Laboratory investigation; a journal of technical methods and pathology* **84**, 1538-1546, doi:10.1038/labinvest.3700190 (2004).
- 221 Lim, K. P. *et al.* Fibroblast gene expression profile reflects the stage of tumour progression in oral squamous cell carcinoma. *The Journal of pathology* **223**, 459-469, doi:10.1002/path.2841 (2011).
- 222 Marsh, D. *et al.* Stromal features are predictive of disease mortality in oral cancer patients. *The Journal of pathology* **223**, 470-481, doi:10.1002/path.2830 (2011).
- 223 Uchida, D. *et al.* Possible role of stromal-cell-derived factor-1/CXCR4 signaling on lymph node metastasis of oral squamous cell carcinoma. *Experimental cell research* **290**, 289-302 (2003).
- 224 Price, S. J., Greaves, D. R. & Watkins, H. Identification of novel, functional genetic variants in the human matrix metalloproteinase-2 gene: role of Sp1 in allele-specific transcriptional regulation. *The Journal of biological chemistry* **276**, 7549-7558, doi:10.1074/jbc.M010242200 (2001).
- 225 Zheng, H., Fu, G., Dai, T. & Huang, H. Migration of endothelial progenitor cells mediated by stromal cell-derived factor-1alpha/CXCR4 via PI3K/Akt/eNOS signal transduction pathway. *Journal of cardiovascular pharmacology* **50**, 274-280, doi:10.1097/FJC.0b013e318093ec8f (2007).

REFERENZEN

- 226 Onoue, T. *et al.* Epithelial-mesenchymal transition induced by the stromal cell-derived factor-1/CXCR4 system in oral squamous cell carcinoma cells. *International journal of oncology* **29**, 1133-1138 (2006).
- 227 Isa, A. Y., Ward, T. H., West, C. M., Slevin, N. J. & Homer, J. J. Hypoxia in head and neck cancer. *The British journal of radiology* **79**, 791-798, doi:10.1259/bjr/17904358 (2006).
- 228 Janssen, H. L., Haustermans, K. M., Balm, A. J. & Begg, A. C. Hypoxia in head and neck cancer: how much, how important? *Head & neck* **27**, 622-638, doi:10.1002/hed.20223 (2005).
- 229 Jiang, J., Tang, Y. L. & Liang, X. H. EMT: a new vision of hypoxia promoting cancer progression. *Cancer biology & therapy* **11**, 714-723 (2011).
- 230 Semenza, G. L. Oxygen sensing, hypoxia-inducible factors, and disease pathophysiology. *Annual review of pathology* **9**, 47-71, doi:10.1146/annurev-pathol-012513-104720 (2014).
- 231 Sullivan, R. & Graham, C. H. Hypoxia-driven selection of the metastatic phenotype. *Cancer metastasis reviews* **26**, 319-331, doi:10.1007/s10555-007-9062-2 (2007).
- 232 Natsuzaka, M. *et al.* Hypoxia induces IGFBP3 in esophageal squamous cancer cells through HIF-1 α -mediated mRNA transcription and continuous protein synthesis. *FASEB J* **26**, 2620-2630, doi:10.1096/fj.11-198598 (2012).
- 233 Natsuzaka, M. *et al.* Insulin-like growth factor-binding protein-3 promotes transforming growth factor- β 1-mediated epithelial-to-mesenchymal transition and motility in transformed human esophageal cells. *Carcinogenesis* **31**, 1344-1353, doi:10.1093/carcin/bgq108 (2010).
- 234 Nakazawa, T. *et al.* Gene expression of periostin in the early stage of fracture healing detected by cDNA microarray analysis. *Journal of orthopaedic research : official publication of the Orthopaedic Research Society* **22**, 520-525, doi:10.1016/j.orthres.2003.10.007 (2004).
- 235 Li, P., Oparil, S., Feng, W. & Chen, Y. F. Hypoxia-responsive growth factors upregulate periostin and osteopontin expression via distinct signaling pathways in rat pulmonary arterial smooth muscle cells. *Journal of applied physiology* **97**, 1550-1558; discussion 1549, doi:10.1152/jappphysiol.01311.2003 (2004).
- 236 Kudo, Y. *et al.* Periostin directly and indirectly promotes tumor lymphangiogenesis of head and neck cancer. *PLoS one* **7**, e44488, doi:10.1371/journal.pone.0044488 (2012).
- 237 Kudo, Y. *et al.* Periostin promotes invasion and anchorage-independent growth in the metastatic process of head and neck cancer. *Cancer research* **66**, 6928-6935, doi:10.1158/0008-5472.CAN-05-4540 (2006).
- 238 Bao, S. *et al.* Periostin potently promotes metastatic growth of colon cancer by augmenting cell survival via the Akt/PKB pathway. *Cancer cell* **5**, 329-339 (2004).
- 239 Ouyang, G. *et al.* Upregulated expression of periostin by hypoxia in non-small-cell lung cancer cells promotes cell survival via the Akt/PKB pathway. *Cancer letters* **281**, 213-219, doi:10.1016/j.canlet.2009.02.030 (2009).
- 240 Maruhashi, T., Kii, I., Saito, M. & Kudo, A. Interaction between periostin and BMP-1 promotes proteolytic activation of lysyl oxidase. *The Journal of biological chemistry* **285**, 13294-13303, doi:10.1074/jbc.M109.088864 (2010).
- 241 Hu, Q. *et al.* Periostin Mediates TGF- β -Induced Epithelial Mesenchymal Transition in Prostate Cancer Cells. *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology* **36**, 799-809, doi:10.1159/000430139 (2015).
- 242 Liu, Y. & Liu, B. A. Enhanced proliferation, invasion, and epithelial-mesenchymal transition of nicotine-promoted gastric cancer by periostin. *World journal of gastroenterology : WJG* **17**, 2674-2680, doi:10.3748/wjg.v17.i21.2674 (2011).

REFERENZEN

- 243 Kanno, A. *et al.* Periostin, secreted from stromal cells, has biphasic effect on cell migration and correlates with the epithelial to mesenchymal transition of human pancreatic cancer cells. *International journal of cancer* **122**, 2707-2718, doi:10.1002/ijc.23332 (2008).
- 244 Schramm, A. *et al.* Prognostic significance of epithelial-mesenchymal transition in malignant pleural mesothelioma. *European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery* **37**, 566-572, doi:10.1016/j.ejcts.2009.08.027 (2010).
- 245 Soltermann, A. *et al.* Prognostic significance of epithelial-mesenchymal and mesenchymal-epithelial transition protein expression in non-small cell lung cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **14**, 7430-7437, doi:10.1158/1078-0432.CCR-08-0935 (2008).
- 246 Heidari, P. *et al.* Imaging of the secreted extracellular periostin, an important marker of invasion in tumor microenvironment in esophageal cancer. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, doi:10.2967/jnumed.115.156216 (2015).
- 247 Nuzzo, P. V. *et al.* Periostin: a novel prognostic and therapeutic target for genitourinary cancer? *Clinical genitourinary cancer* **12**, 301-311, doi:10.1016/j.clgc.2014.02.005 (2014).
- 248 Mikheev, A. M. *et al.* Periostin is a novel therapeutic target that predicts and regulates glioma malignancy. *Neuro-oncology*, doi:10.1093/neuonc/nou161 (2014).
- 249 Kudo, Y., Siriwardena, B. S., Hatano, H., Ogawa, I. & Takata, T. Periostin: novel diagnostic and therapeutic target for cancer. *Histology and histopathology* **22**, 1167-1174 (2007).
- 250 Anborgh, P. H. *et al.* Role of plasma osteopontin as a biomarker in locally advanced breast cancer. *American journal of translational research* **7**, 723-732 (2015).
- 251 Zduniak, K. *et al.* Nuclear osteopontin-c is a prognostic breast cancer marker. *British journal of cancer* **112**, 729-738, doi:10.1038/bjc.2014.664 (2015).
- 252 Botti, G. *et al.* SPARC/osteonectin is involved in metastatic process to the lung during melanoma progression. *Virchows Archiv : an international journal of pathology* **465**, 331-338, doi:10.1007/s00428-014-1616-4 (2014).
- 253 Derosa, C. A. *et al.* Elevated osteonectin/SPARC expression in primary prostate cancer predicts metastatic progression. *Prostate cancer and prostatic diseases* **15**, 150-156, doi:10.1038/pcan.2011.61 (2012).
- 254 Tichet, M. *et al.* Tumour-derived SPARC drives vascular permeability and extravasation through endothelial VCAM1 signalling to promote metastasis. *Nature communications* **6**, 6993, doi:10.1038/ncomms7993 (2015).
- 255 Fenouille, N. *et al.* SPARC functions as an anti-stress factor by inactivating p53 through Akt-mediated MDM2 phosphorylation to promote melanoma cell survival. *Oncogene* **30**, 4887-4900, doi:10.1038/onc.2011.198 (2011).
- 256 Fenouille, N. *et al.* The epithelial-mesenchymal transition (EMT) regulatory factor SLUG (SNAI2) is a downstream target of SPARC and AKT in promoting melanoma cell invasion. *PLoS one* **7**, e40378, doi:10.1371/journal.pone.0040378 (2012).
- 257 Wu, X. *et al.* HOXB7, a homeodomain protein, is overexpressed in breast cancer and confers epithelial-mesenchymal transition. *Cancer research* **66**, 9527-9534, doi:10.1158/0008-5472.CAN-05-4470 (2006).
- 258 Liu, S. *et al.* HOXB7 promotes malignant progression by activating the TGFbeta signaling pathway. *Cancer research* **75**, 709-719, doi:10.1158/0008-5472.CAN-14-3100 (2015).
- 259 Chen, H. *et al.* Hoxb7 inhibits transgenic HER-2/neu-induced mouse mammary tumor onset but promotes progression and lung metastasis. *Cancer research* **68**, 3637-3644, doi:10.1158/0008-5472.CAN-07-2926 (2008).

REFERENZEN

- 260 Jin, K. *et al.* The HOXB7 protein renders breast cancer cells resistant to tamoxifen through activation of the EGFR pathway. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 2736-2741, doi:10.1073/pnas.1018859108 (2012).
- 261 De Souza Setubal Destro, M. F. *et al.* Overexpression of HOXB7 homeobox gene in oral cancer induces cellular proliferation and is associated with poor prognosis. *International journal of oncology* **36**, 141-149 (2010).
- 262 Bitu, C. C. *et al.* HOXB7 expression is a prognostic factor for oral squamous cell carcinoma. *Histopathology* **60**, 662-665, doi:10.1111/j.1365-2559.2011.04102.x (2012).
- 263 Simpson, D. R., Mell, L. K. & Cohen, E. E. Targeting the PI3K/AKT/mTOR pathway in squamous cell carcinoma of the head and neck. *Oral oncology* **51**, 291-298, doi:10.1016/j.oraloncology.2014.11.012 (2015).
- 264 Lui, V. W. *et al.* Frequent mutation of the PI3K pathway in head and neck cancer defines predictive biomarkers. *Cancer discovery* **3**, 761-769, doi:10.1158/2159-8290.CD-13-0103 (2013).
- 265 Templeton, A. J. *et al.* Prognostic relevance of receptor tyrosine kinase expression in breast cancer: a meta-analysis. *Cancer treatment reviews* **40**, 1048-1055 (2014).
- 266 Duncan, J. S. *et al.* Dynamic reprogramming of the kinome in response to targeted MEK inhibition in triple-negative breast cancer. *Cell* **149**, 307-321 (2012).
- 267 Dieci, M. V., Arnedos, M., Andre, F. & Soria, J. C. Fibroblast growth factor receptor inhibitors as a cancer treatment: from a biologic rationale to medical perspectives. *Cancer discovery* **3**, 264-279 (2013).
- 268 Dienstmann, R. *et al.* Genomic aberrations in the FGFR pathway: opportunities for targeted therapies in solid tumors. *Ann Oncol* **25**, 552-563 (2014).
- 269 Furlan, A. *et al.* Abl interconnects oncogenic Met and p53 core pathways in cancer cells. *Cell death and differentiation* **18**, 1608-1616 (2011).
- 270 Ang, K. K. *et al.* Randomized phase III trial of concurrent accelerated radiation plus cisplatin with or without cetuximab for stage III to IV head and neck carcinoma: RTOG 0522. *J Clin Oncol* **32**, 2940-2950 (2014).
- 271 Harrington, K. *et al.* Postoperative Adjuvant Lapatinib and Concurrent Chemoradiotherapy Followed by Maintenance Lapatinib Monotherapy in High-Risk Patients With Resected Squamous Cell Carcinoma of the Head and Neck: A Phase III, Randomized, Double-Blind, Placebo-Controlled Study. *J Clin Oncol* **33**, 4202-4209 (2015).
- 272 Chong, C. R. & Janne, P. A. The quest to overcome resistance to EGFR-targeted therapies in cancer. *Nature medicine* **19**, 1389-1400, doi:10.1038/nm.3388 (2013).
- 273 Chapman, C. H., Saba, N. F. & Yom, S. S. Targeting epidermal growth factor receptor for head and neck squamous cell carcinoma: still lost in translation? *Annals of translational medicine* **4**, 80 (2016).
- 274 Neal, J. W. & Sledge, G. W. Decade in review-targeted therapy: successes, toxicities and challenges in solid tumours. *Nature reviews. Clinical oncology* **11**, 627-628 (2014).
- 275 Flaherty, K. T. *et al.* Combined BRAF and MEK inhibition in melanoma with BRAF V600 mutations. *The New England journal of medicine* **367**, 1694-1703 (2012).
- 276 Jane Bayani, J. M. S. B. in *Molecular Pathology* (ed Abeer Shaaban John M. S. Bartlett, Fernando Schmitt) Ch. 6, (2015).
- 277 Eschrich, S. A. *et al.* A gene expression model of intrinsic tumor radiosensitivity: prediction of response and prognosis after chemoradiation. *International journal of radiation oncology, biology, physics* **75**, 489-496, doi:10.1016/j.ijrobp.2009.06.014 (2009).

REFERENZEN

- 278 Toustrup, K. *et al.* Development of a hypoxia gene expression classifier with predictive impact for hypoxic modification of radiotherapy in head and neck cancer. *Cancer research* **71**, 5923-5931, doi:10.1158/0008-5472.CAN-11-1182 (2011).
- 279 Lohavanichbutr, P. *et al.* A 13-gene signature prognostic of HPV-negative OSCC: discovery and external validation. *Clinical cancer research : an official journal of the American Association for Cancer Research* **19**, 1197-1203, doi:10.1158/1078-0432.CCR-12-2647 (2013).
- 280 De Cecco, L., Bossi, P., Locati, L., Canevari, S. & Licitra, L. Comprehensive gene expression meta-analysis of head and neck squamous cell carcinoma microarray data defines a robust survival predictor. *Ann Oncol* **25**, 1628-1635, doi:10.1093/annonc/mdu173 (2014).
- 281 De Cecco, L. *et al.* Head and neck cancer subtypes with biological and clinical relevance: Meta-analysis of gene-expression data. *Oncotarget* **6**, 9627-9642, doi:10.18632/oncotarget.3301 (2015).
- 282 de Gramont, A. *et al.* Pragmatic issues in biomarker evaluation for targeted therapies in cancer. *Nature reviews. Clinical oncology* **12**, 197-212 (2015).
- 283 DiMasi, J. A., Feldman, L., Seckler, A. & Wilson, A. Trends in risks associated with new drug development: success rates for investigational drugs. *Clinical pharmacology and therapeutics* **87**, 272-277 (2010).
- 284 Arrowsmith, J. Trial watch: Phase II failures: 2008-2010. *Nature reviews. Drug discovery* **10**, 328-329 (2011).
- 285 Arrowsmith, J. Trial watch: phase III and submission failures: 2007-2010. *Nature reviews. Drug discovery* **10**, 87 (2011).
- 286 Huriez, A. *Personalized medicine, introduction, business impact on the healthcare sector and regulatory aspects*, <<http://www.epsa-online.org/doc/aa13/Symposium%20-%20Alain%20Huriez%20-%20Regulatory.pdf>> (2013).
- 287 Paul, S. M. *et al.* How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nature reviews. Drug discovery* **9**, 203-214 (2002).
- 288 Redman, M. W. & Allegra, C. J. The Master Protocol Concept. *Seminars in oncology* **42**, 724-730 (2015).
- 289 Chia, S. K. *et al.* A 50-gene intrinsic subtype classifier for prognosis and prediction of benefit from adjuvant tamoxifen. *Clinical cancer research : an official journal of the American Association for Cancer Research* **18**, 4465-4472 (2012).
- 290 Nielsen, T. O. *et al.* A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **16**, 5222-5232 (2010).
- 291 Paik, S. *et al.* Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol* **24**, 3726-3734 (2006).
- 292 Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747-752 (2000).
- 293 Kuboki, Y. *et al.* Comprehensive analyses using next-generation sequencing and immunohistochemistry enable precise treatment in advanced gastric cancer. *Ann Oncol* **27**, 127-133 (2016).
- 294 Sethi, N., MacLennan, K., Wood, H. M. & Rabbitts, P. Past and future impact of next-generation sequencing in head and neck cancer. *Head & neck* **38 Suppl 1**, E2395-2402 (2015).
- 295 Cancer Genome Atlas Research, N. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061-1068, doi:10.1038/nature07385 (2008).

REFERENZEN

- 296 Cancer Genome Atlas Research, N. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519-525, doi:10.1038/nature11404 (2012).
- 297 Yan, X. *et al.* Comprehensive Genomic Characterization of Long Non-coding RNAs across Human Cancers. *Cancer cell* **28**, 529-540, doi:10.1016/j.ccell.2015.09.006 (2015).
- 298 Zheng, S. *et al.* Comprehensive Pan-Genomic Characterization of Adrenocortical Carcinoma. *Cancer cell* **29**, 723-736, doi:10.1016/j.ccell.2016.04.002 (2016).
- 299 Cancer Genome Atlas Research Network. Electronic address, w. b. e. & Cancer Genome Atlas Research, N. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell* **169**, 1327-1341 e1323, doi:10.1016/j.cell.2017.05.046 (2017).
- 300 Pereira, E. R., Jones, D., Jung, K. & Padera, T. P. The lymph node microenvironment and its role in the progression of metastatic cancer. *Semin Cell Dev Biol* **38**, 98-105, doi:10.1016/j.semcdb.2015.01.008 (2015).
- 301 Gromov, P. & Gromova, I. Characterization of the Tumor Secretome from Tumor Interstitial Fluid (TIF). *Methods Mol Biol* **1459**, 231-247, doi:10.1007/978-1-4939-3804-9_16 (2016).
- 302 Marimuthu, A. *et al.* Identification of head and neck squamous cell carcinoma biomarker candidates through proteomic analysis of cancer cell secretome. *Biochim Biophys Acta* **1834**, 2308-2316, doi:10.1016/j.bbapap.2013.04.029 (2013).
- 303 Aebersold, R. & Mann, M. Mass-spectrometric exploration of proteome structure and function. *Nature* **537**, 347-355, doi:10.1038/nature19949 (2016).
- 304 Boonstra, J. & Verkleij, A. J. Regulation of enzyme activity in vivo is determined by its cellular localization. *Adv Enzyme Regul* **44**, 61-73, doi:10.1016/j.advenzreg.2003.11.008 (2004).
- 305 Stauber, R. H. *et al.* Nucleocytoplasmic shuttling and the biological activity of mouse survivin are regulated by an active nuclear export signal. *Traffic (Copenhagen, Denmark)* **7**, 1461-1472 (2006).
- 306 Knauer, S. K. & Stauber, R. H. Development of an autofluorescent translocation biosensor system to investigate protein-protein interactions in living cells. *Analytical chemistry* **77**, 4815-4820 (2005).
- 307 Knauer, S. K. *et al.* Bioassays to monitor Taspase1 function for the identification of pharmacogenetic inhibitors. *PLoS one* **6**, e18253, doi:10.1371/journal.pone.0018253 (2011).
- 308 Hogan, P. G. & Rao, A. Transcriptional regulation. Modification by nuclear export? *Nature* **398**, 200-201, doi:10.1038/18327 (1999).
- 309 Prokosch, H. U. & Ganslandt, T. Perspectives for medical informatics. Reusing the electronic medical record for clinical research. *Methods Inf Med* **48**, 38-44 (2009).
- 310 Brandt, C. A. *et al.* Informatics tools to improve clinical research study implementation. *Contemp Clin Trials* **27**, 112-122, doi:10.1016/j.cct.2005.11.013 (2006).
- 311 Schmier, J. K., Kane, D. W. & Halpern, M. T. Practical applications of usability theory to electronic data collection for clinical trials. *Contemp Clin Trials* **26**, 376-385, doi:10.1016/j.cct.2005.01.012 (2005).
- 312 Borrebaeck, C. A. Precision diagnostics: moving towards protein biomarker signatures of clinical utility in cancer. *Nature reviews. Cancer* **17**, 199-204, doi:10.1038/nrc.2016.153 (2017).
- 313 Friedman, A. A., Letai, A., Fisher, D. E. & Flaherty, K. T. Precision medicine for cancer with next-generation functional diagnostics. *Nature reviews. Cancer* **15**, 747-756, doi:10.1038/nrc4015 (2015).
- 314 Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289-300 (1995).

REFERENZEN

- 315 Benjamini, Y. & Cohen, R. Weighted false discovery rate controlling procedures for clinical trials. *Biostatistics* **18**, 91-104, doi:10.1093/biostatistics/kxw030 (2017).
- 316 Swanton, C. Intratumor heterogeneity: evolution through space and time. *Cancer research* **72**, 4875-4882, doi:10.1158/0008-5472.CAN-12-2217 (2012).
- 317 McGranahan, N. & Swanton, C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell* **168**, 613-628, doi:10.1016/j.cell.2017.01.018 (2017).
- 318 Carlsson, A. *et al.* Molecular serum portraits in patients with primary breast cancer predict the development of distant metastases. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 14252-14257, doi:10.1073/pnas.1103125108 (2011).
- 319 Rosorius, O. *et al.* Direct observation of nucleocytoplasmic transport by microinjection of GFP-tagged proteins in living cells. *BioTechniques* **27**, 350-355 (1999).
- 320 Heckman, K. L. & Pease, L. R. Gene splicing and mutagenesis by PCR-driven overlap extension. *Nature Protocols* **2**, 924, doi:10.1038/nprot.2007.132 (2007).
- 321 Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-254 (1976).

7 PUBLIKATIONEN

7.1 Artikel

D. Goesswein, A. Hahlbrock, M. Deichelbohrer, C. Vallet, D. Wünsch, D. Westmeier, G. Schneider, A. Hildebrandt, S. K. Knauer, N. Habtemichael, R. H. Stauber. "Identification of novel disease-relevant signaling-pathways in head and neck cancer by genome-wide transcriptional profiling." *Oncotarget (in revision)*.

D. Goesswein, N. Habtemichael, A. Gerhold-Ay, J. Mazur, D. Wünsch, S. Knauer, J. Künzel, C. Matthias, S. Strieth, R. H. Stauber (2018). "Expressional analysis of disease-relevant signalling pathways in primary tumours and metastasis of head and neck cancers." *Scientific Reports, Volume 8, Issue 1, Article 7326*.

A. Beltz, **D. Goesswein**, S. Zimmer, R.H. Stauber, J. Hagemann, S. Strieth, C. Matthias, J. Künzel (2018). "Staging von Oropharynxkarzinomen. Neue TNM-Klassifikation als Herausforderung für Kopf-Hals-Tumorzentren". *HNO, Volume 66, Issue 5, Pages 375-382*.

C. Schrenk, V. Fetz, C. Vallet, C. Heiselmayer, E. Schröder, A. Hensel, A. Hahlbrock, D. Wünsch, **D. Goesswein**, C. Bier, N. Habtemichael, G. Schneider, R.H. Stauber, S.K.Knauer (2017). „TFIIA transcriptional activity is controlled by a ‚cleave-and-run‘ Exportin-1/Taspase1-switch.“ *Journal of Molecular Cell Biology, Volume 10, Issue 1, Pages 33-47*.

O. Schilling (2017). „Protein Terminal Profiling, Methods and Procols.“ *Methods in Molecular Biology* 1574, Springer Protocols. Chapter 17: A. Hahlbrock, **D. Goesswein**, R. H. Stauber. „Protein Translocation Assays to Probe Protease Function and Screen for Inhibitors.“ *Springer Science+Business Media LLC*.

Hahlbrock, A., **D. Goesswein**, J. Künzel, D. Wünsch, R. H. Stauber (2016). "Threonine Aspartase1: An unexplored protease with relevance for head and neck cancer?" *Oral Oncology, Volume 54, Article e10-2*.

Wunsch, D., A. Hahlbrock, C. Heiselmayer, S. Backer, P. Heun, **D. Goesswein**, ... , R. H. Stauber (2015). "Fly versus man: Evolutionary impairment of nucleolar targeting affects the degradation of Drosophila's Taspase1." *FASEB Journal, Volume 29, Issue 5, Pages 1973-85.*

Bier, C., R. Hecht, L. Kunst, S. Scheiding, D. Wunsch, **D. Goesswein**, ... , R. H. Stauber (2012). "Overexpression of the catalytically impaired Taspase1 T234V or Taspase1 D233A variants does not have a dominant negative effect in T(4;11) leukemia cells." *PLoS One, Volume 7, Issue 5, Article e34142.*

7.2 Vorträge

D. Goesswein. "Systematic transcriptomics and proteomics analyses to identify cancer relevant protease networks". *International Young Scientists' Forum of Proteases in Disease 2016, Mainz.*

D. Goesswein. "Detecting circulating tumor cells using nanotechnology". *International Nano(bio)medicine Forum for Young Scientists 2016, Mainz.*

J. Künzel, **D. Goesswein**, A. Gerhold-Ay, R. H. Stauber. "Komparatives Expressions-Profilung von Primärtumoren und korrespondierenden Lymphknotenmetastasen". *Onkologie-Sitzung, 87. Jahresversammlung, HNO Kongress 2016, Congress Center Düsseldorf.*

7.3 Posterbeiträge

D. Goesswein, J. Künzel, A. Gerhold-Ay, ... , R. H. Stauber. "Analyse der differentiellen Expression von metastasierenden Kopf-Hals-Plattenepithelkarzinomen". *Onkologie (experimentell), 88. Jahresversammlung, HNO Kongress 2017, Erfurt.*

E. Deuss, **D. Goesswein**, S. Strieth, J. Künzel. "Analyse des Wachstumsrezeptormusters von HNSCC unter Therapie mit Cetuximab". *Onkologie (experimentell), 88. Jahresversammlung, HNO Kongress 2017, Erfurt.*

C. S. Eger, S. Strieth, J. Künzel, **D. Goesswein**. "Analyse der pathobiologischen Relevanz der Taspase1 in Kopf-Hals Tumoren". *Onkologie (experimentell), 88. Jahresversammlung, HNO Kongress 2017, Erfurt.*

D. Goesswein, D. Westmeier, J. Künzel, ... , R. H. Stauber. "Applied nanomedicine – Nanomagnetic detection and characterization of circulating tumor cells in HNSCC". *Onkologie-Sitzung, 87. Jahresversammlung, HNO Kongress 2016, Congress Center Düsseldorf.*

I. Limburg, **D. Gößwein**, J. Künzel, R. H. Stauber. "Einfluss des HPV-Status auf das Metastasierungsmuster bei Kopf-Hals Tumoren". *Onkologie-Sitzung, 87. Jahresversammlung, HNO Kongress 2016, Congress Center Düsseldorf.*

S. K. Knauer, **D. Gößwein**, S. Bäcker, ..., R. H. Stauber. "Therapy resistance in HNSCC – enhanced levels of the cytoprotective protein survivin is mediated by down-regulation of microRNA-542-3p." 1. *International Symposium on Tumor-Host-Interaction in Head and Neck Cancer, 5. Annual Meeting on Experimental and Translational Head and Neck Oncology, Essen.*

D. Gößwein, V. Fetz, R.H. Stauber. "Identification of compounds with anti-tumoral activity – exploiting natural products to target diseases." 6. *UCT Science Day 2014, Universitäres Centrum für Tumorerkrankungen, Frankfurt.*

D. Docter, V. Fetz, **D. Gößwein**, A. Hahlbrock, R. H. Stauber. "Nanoparticle-based drug delivery – Impact of the protein corona for efficacy and target-specificity." 5. *AIO Arbeitsgemeinschaft Internistische Onkologie Symposium 2013, Berlin.*

D. Docter, N. Habtemichael, S. Tenzer, V. Fetz, **D. Goesswein**, ... , R. H. Stauber. "The dynamic Nanoparticle-Protein Corona – Implications for Nano-Toxicology and Biomedicine". 3. *Clustertreffen NanoCare/NanoNature 2013, Frankfurt.*

D. Gößwein, D. Wünsch, C. Bier, R.H. Stauber. "Genetic inhibition of Taspase1 heterodimerization will not aid to target t(4;11) leukemias." *CGT Perspectives in Cell- and Gene-Based Medicines 2012, Frankfurt.*

8 APPENDIX

8.1 Abkürzungen

Tab. 33: Abkürzungsverzeichnis

Abkürzung	Beschreibung
Abb	Abbildung
ADAM	<i>A Disintegrin And Metalloprotease</i>
ADH1B	Oxidoreduktase Alkohol-Dehydrogenase 1B
ALL	Akute Lymphoblastische Leukämie
AML	Akute Myeloide Leukämie
APOE	Apolipoprotein E
ARHGAP15	<i>Rho GTPase-activating Protein 15</i>
AS	Aminosäuren
ASPN	Asporin
ATH1-3	AT-Hook Domänen
AveExpr	Mittlere Expression
bp	Basenpaare
BSC	<i>Best Supportive Care</i>
C7	<i>Complement Component C7</i>
CAFs	Krebs-Assoziierte-Fibroblasten
CBP	<i>CREB-binding Protein</i>
CCL	C-C Motiv Chemokin Ligand
CCR	C-C Motiv Chemokin Rezeptor
CD52	CAMPATH-1 Antigen
CDH11	<i>Cadherin 11 Type 2</i>
CTCs	Zirkulierende Tumorzellen (<i>Circulating Tumour Cells</i>)
CTGF	<i>Connective Tissue Growth Factor</i>
CXCL12	<i>Stromal Cell-derived Factor 1</i>
CXCR4	C-X-C Motiv Chemokin Rezeptor Typ 4
CYFIP2	<i>Cytoplasmic FMR1-interacting Protein 2</i>
CYTIP	<i>Cytohesin-interacting Protein</i>
DCL	Rechteverwaltung und Transaktionskontrolle
DDL	Definition des Datenbankschemas
DML	Befehle zur Datenmanipulation und dem lesenden Zugriff
DNASE1L3	<i>Deoxyribonuclease Gamma</i>
DOCK2	<i>Dedicator of Cytokinesis Protein 2</i>

APPENDIX

Abkürzung	Beschreibung
DPE	<i>Downstream Promotor Element</i>
dTaspase1	<i>Drosophila Taspase1</i>
DTCs	Disseminierte Tumorzellen (<i>Disseminated Tumour Cells</i>)
DUBs	Deubiquitinierende Enzyme
E	Real-Time PCR Effizienz
ECM	Extrazelluläre Matrix
EGFR	<i>Epidermal Growth Factor Receptor</i>
EMT	Epitheliale-zu-mesenchymale-Transition
ENPP2	<i>Ectonucleotide Pyrophosphatase Family Member 2</i>
ERBB2	Rezeptor Tyrosin-Protein Kinase Her2
FAP	<i>Fibroblast Activation Protein</i>
FC	<i>Fold Change</i>
FCRs	Fibroblasten
FDR	<i>False Discovery Rate</i>
FN1	Fibronectin
GIMAP6	<i>GTPase IMAP Family Member 6</i>
GMFG	<i>Glia Maturation Factor Gamma</i>
GST	<i>Gluthathion-S-Transferase</i>
HMHA1	<i>Minor Histocompatibility Protein HA-1</i>
HNC	Kopf-Hals-Tumore (<i>Head and Neck Cancer</i>)
HNSCC	Kopf-Hals Plattenepithelkarzinome (<i>Head and Neck Squamous Cell Carcinoma</i>)
HOXB7	Homeobox B7
HP1	<i>Heterochromatic Protein 1</i>
HPV	Humanes Papillomavirus
IFNG	Zytosin Interferon Gamma
IGFBP3	<i>Insulin-Like Growth Factor Binding Protein 3</i>
IHC	Immunhistochemie
IL-4	Interleukin-4
Inr	Initiator-Motiv
IPA	<i>Ingenuity Pathway Analysis</i>
IRF8	<i>Interferon Regulatory Factor 8</i>
ITM2A	<i>Integral Membrane Protein 2A</i>
KLK	Kallikreine
KRT24	Zytokeratin24 / <i>Cytokeratine24</i>
LAMC2	Laminin
LAPTM5	<i>Lysosomal-associated Transmembrane Protein 5</i>
LOX, LOXL2	Lysyl Oxidasen
LPL	Lipoprotein Lipase
LRMP	<i>Lymphoid-restricted Membrane Protein</i>
LRP12	<i>Low Density Lipoprotein Receptor-Related Protein 12</i>
M	Metastase(n)

APPENDIX

Abkürzung	Beschreibung
MAGEA4	Melanoma Antigen A4
MCS	<i>Multiple Cloning Site</i>
MEN1	Menin
MEx	Microarray Expressionsdaten
MFAP2	<i>Microfibrillar-Associated Protein</i>
MICALs	<i>Molecules Interacting with CasL-2 Proteins</i>
MLL	<i>Mixed Lineage Leukemia</i>
MMP	Matrix Metalloprotease
mRNA	<i>Messenger RNA</i>
MTE	<i>Motif Ten Element</i>
mTOR	<i>Mammalian Target of Rapamycin</i>
MYO1F	Myosin-1F
N	Normalgewebe
NES	Nukleäres Exportsignal
NGS	<i>Next-Generation-/ HochdurchsatzSequenzierung</i>
NLS	Nukleäres Lokalisationssignal
NoLS	Nukleoläres Lokalisationssignal
NPM1	Nukleophosmin1
NREP	<i>Neuronal Regeneration Related Protein</i>
NXF	<i>Nuclear RNA Export Factor</i>
ORO1A	Coronin-1A
OS	Gesamtüberleben / <i>Overall Survival</i>
PAX5	<i>Paired Box Protein Pax-5</i>
PcG	Polycomb
PCR	Polymerasekettenreaktion
PDGFR	<i>Platelet-Derived Growth Factor Receptor</i>
PHD	Pflanzenhomologiedomäne
PI3K	<i>Phosphoinositol-3-Kinase</i>
PIC	Pre-Initiationskomplex
PKC	Protein Kinase C
PLAU	Urokinase
PLIN1	Perilipin-1
Pol I	RNA Polymerase I
POSTN	Periostin
PRKCB	Proteinkinase C
PT	Primärtumore
PTEN	<i>Phosphatase And Tensin Homolog</i>
PTGDS	Prostaglandin-H2 D-Isomerase
PTPRC	<i>Receptor-type Tyrosin-protein Phosphatase C</i>
PVDF	Polyvinylidendifluorid
PXDN	Peroxidasin Homolog

APPENDIX

Abkürzung	Beschreibung
QALY	<i>Quality adjusted life years</i>
qPCR	Quantitative Real-Time PCR
RASSF2	<i>Ras-association Domain-containing Protein 2</i>
rRNA	Ribosomale RNA
RTK	Rezeptortyrosinkinase
RT-PCR	Reverse Transkriptions PCR
SCHIP1	Schwannomin
SDS-PAGE	Natriumdodecylsulfat-Polyacrylamidgelelektrophorese
SELL	L-Selectin
SEPT6	Septin-6
SERPINH1	Serpin Inhibitor
SET	(Su(var)3-9, enhancer-of-zeste, trithorax)-Methyltransferasedomäne
SH2D1A	SH2 Domain-containing Protein 1A
SIAH1	E3 ubiquitin-protein ligase SIAH1
SIAH2	E3 ubiquitin-protein ligase SIAH2
SLC2A3	<i>Solute Carrier Family 2 Member 3</i>
snRNA	<i>Small nuclear RNA</i>
SOE-PCR	<i>Splice Overlap Extension PCR</i>
SPARC	Osteonectin
SPP1	Osteopontin
SQL	<i>Structured Query Language</i>
SRGN	Serglycin
STAP1	<i>Signal-Transducing Adaptor Protein 1</i>
TA	Transkriptionelle Aktivierungsdomäne
TAFs	TBP-assozierte Faktoren
TAMs	Tumor-assozierte Makrophagen
Taspase1	Threonin-Aspartase 1
TBP	TATA Box bindendes Protein
TEMED	Tetramethylethylenediamine
TFIIA	Transkriptionsfaktor IIA
TFIID	Transkriptionsfaktor IID
TFRC	Transferrin
TGFB1	<i>Transforming Growth Factor β1</i>
TGFBR2	<i>TGF-beta Receptor Type-2</i>
TGF- β	Transforming Growth Factor- β
TMA	<i>Tissue Microarray</i>
tPA	<i>Tissue-Type Plasminogen Activator</i>
TRAIL	<i>TNF-related apoptosis-inducing ligand</i>
TRD	Transkriptionelle Repressionsdomäne
TrxG	Trithorax
UICC	<i>Union International Contre le Cancer</i>

Abkürzung	Beschreibung
uPA	<i>Urokinase-Type Plasminogen Activator</i>
UPS	Ubiquitin-Proteasom System
USF2	<i>Upstream Stimulatory Factor 2</i>
VEGF	<i>Vascular Endothelial Growth Factor</i>

8.2 Zusätzliche Tabellen und Listen

Tab. 34: Durchschnittliche Microarray-Daten validierter Gene (PTvs.N)

Affy-ID	Gen	log ₂ FC	AveExpr	t	p-Wert	FDR
220026_at	CLCA4	-4,05	7,26	-7,13	1,39E-08	4,85E-06
210495_x_at	FN1	1,78	11,20	5,25	5,65E-06	3,71E-04
211719_x_at	FN1	1,83	11,02	5,21	6,36E-06	4,01E-04
216442_x_at	FN1	1,83	11,12	5,20	6,67E-06	4,13E-04
212464_s_at	FN1	1,87	10,53	5,16	7,56E-06	4,49E-04
214701_s_at	FN1	1,00	6,33	4,04	2,41E-04	5,46E-03
214702_at	FN1	0,37	4,19	1,72	9,40E-02	3,14E-01
210809_s_at	POSTN	3,06	9,31	7,26	9,40E-09	3,43E-06
220267_at	KRT24	-2,97	6,39	-6,14	3,26E-07	4,40E-05
204919_at	PRR4	-5,24	7,55	-6,62	7,11E-08	1,54E-05

Tab. 35: Durchschnittliche Microarray Daten validierter Gene (Mvs.PT)

Affy-ID	Gen	log ₂ FC	AveExpr	t	p-Wert	FDR
211906_s_at	SERPINB4	-2,14	6,19	-4,41	7,96E-05	3,50E-03
220804_s_at	TP73	0,20	4,57	1,27	2,12E-01	4,14E-01
204952_at	LYPD3	-1,32	8,73	-3,17	2,92E-03	3,07E-02
221601_s_at	FAIM3	2,23	6,69	5,08	9,52E-06	1,18E-03
221602_s_at	FAIM3	1,50	6,75	4,84	2,03E-05	1,74E-03
38149_at	ARHGAP25	1,23	6,80	4,67	3,54E-05	2,31E-03
204882_at	ARHGAP25	1,15	6,88	3,72	6,16E-04	1,19E-02
214369_s_at	RASGRP2	1,10	5,04	5,07	9,98E-06	1,21E-03
208206_s_at	RASGRP2	1,42	6,16	4,92	1,61E-05	1,54E-03
214368_at	RASGRP2	0,43	5,61	4,41	7,89E-05	3,49E-03
214367_at	RASGRP2	0,12	5,90	1,26	2,15E-01	4,17E-01

APPENDIX

Tab. 36: TOP100 Liste differenziell exprimierter Gene in PTvs.N, sortiert nach der FDR.

Affy-ID	Gen	log ₂ FC	AveExpr	t	p-Wert	FDR	B
211726_s_at	FMO2	-2,69	7,36	-10,95	1,74E-13	2,04E-09	20,30
211883_x_at	CEACAM1	-1,82	6,76	-10,94	1,83E-13	2,04E-09	20,26
211889_x_at	CEACAM1	-1,86	6,15	-10,31	1,04E-12	7,72E-09	18,64
91826_at	EPS8L1	-2,30	6,85	-10,09	1,89E-12	1,05E-08	18,08
211372_s_at	IL1R2	-2,29	6,52	-9,83	3,98E-12	1,44E-08	17,38
221665_s_at	EPS8L1	-2,39	6,26	-9,80	4,37E-12	1,44E-08	17,29
205403_at	IL1R2	-2,62	6,29	-9,72	5,54E-12	1,44E-08	17,07
209498_at	CEACAM1	-2,32	7,31	-9,71	5,64E-12	1,44E-08	17,05
204755_x_at	HLF	-2,11	5,91	-9,70	5,83E-12	1,44E-08	17,02
221655_x_at	EPS8L1	-1,91	7,01	-9,06	3,82E-11	8,26E-08	15,25
220090_at	CRNN	-5,18	7,32	-9,03	4,09E-11	8,26E-08	15,18
AFFX-LysX-5_at	NA	-2,13	9,25	-9,00	4,45E-11	8,26E-08	15,10
211689_s_at	TMPRSS2	-2,15	5,62	-8,92	5,69E-11	9,75E-08	14,87
207802_at	CRISP3	-4,94	5,95	-8,69	1,13E-10	1,70E-07	14,22
AFFX-r2-Bs-lys-5_at	NA	-2,08	9,56	-8,68	1,16E-10	1,70E-07	14,19
218779_x_at	EPS8L1	-2,06	7,01	-8,66	1,22E-10	1,70E-07	14,15
AFFX-LysX-M_at	NA	-1,90	9,76	-8,64	1,30E-10	1,71E-07	14,08
217846_at	QARS	-0,81	9,65	-8,61	1,43E-10	1,77E-07	14,00
206576_s_at	CEACAM1	-1,95	6,57	-8,59	1,51E-10	1,77E-07	13,95
204754_at	HLF	-1,43	5,61	-8,31	3,59E-10	3,85E-07	13,12
218885_s_at	GALNT12	-1,94	6,81	-8,31	3,63E-10	3,85E-07	13,11
207059_at	PAX9	-1,83	6,46	-8,25	4,34E-10	4,24E-07	12,94
202730_s_at	NA	-1,32	7,88	-8,23	4,61E-10	4,24E-07	12,89
211885_x_at	FUT6	-1,54	5,48	-8,22	4,69E-10	4,24E-07	12,87
AFFX-r2-Bs-lys-M_at	NA	-2,09	9,10	-8,22	4,76E-10	4,24E-07	12,86
209368_at	EPHX2	-1,09	5,94	-8,19	5,08E-10	4,32E-07	12,79
210461_s_at	ABLIM1	-1,27	7,42	-8,19	5,23E-10	4,32E-07	12,77
206539_s_at	CYP4F12	-1,06	6,19	-8,16	5,68E-10	4,51E-07	12,69
205765_at	CYP3A5	-2,07	5,78	-8,13	6,17E-10	4,51E-07	12,61
210398_x_at	FUT6	-1,53	5,38	-8,13	6,25E-10	4,51E-07	12,60
202770_s_at	CCNG2	-1,37	7,94	-8,12	6,38E-10	4,51E-07	12,58
204753_s_at	HLF	-1,92	5,94	-8,12	6,47E-10	4,51E-07	12,56
219476_at	C1orf116	-2,34	7,41	-8,03	8,51E-10	5,75E-07	12,30
208126_s_at	CYP2C18	-2,01	6,11	-7,90	1,27E-09	8,12E-07	11,92
211882_x_at	FUT6	-1,45	5,81	-7,90	1,28E-09	8,12E-07	11,92
214399_s_at	KRT4	-1,68	5,05	-7,87	1,36E-09	8,44E-07	11,85
AFFX-r2-Bs-lys-3_at	NA	-1,79	8,87	-7,82	1,60E-09	9,65E-07	11,70
210399_x_at	FUT6	-1,47	7,45	-7,78	1,82E-09	1,07E-06	11,58

APPENDIX

Affy-ID	Gen	log ₂ FC	AveExpr	t	p-Wert	FDR	B
210608_s_at	FUT2	-1,55	7,12	-7,71	2,29E-09	1,31E-06	11,36
204475_at	MMP1	4,86	7,38	7,69	2,38E-09	1,33E-06	11,32
210096_at	CYP4B1	-2,44	5,21	-7,68	2,51E-09	1,36E-06	11,27
204777_s_at	MAL	-4,33	8,94	-7,65	2,77E-09	1,47E-06	11,18
206884_s_at	SCEL	-4,06	6,79	-7,63	2,88E-09	1,48E-06	11,14
200965_s_at	ABLIM1	-1,55	9,64	-7,63	2,94E-09	1,48E-06	11,12
212586_at	CAST	-1,22	9,53	-7,62	2,98E-09	1,48E-06	11,11
216210_x_at	TRIOBP	-0,64	8,26	-7,61	3,13E-09	1,52E-06	11,06
203961_at	NEBL	-1,96	7,15	-7,54	3,81E-09	1,80E-06	10,87
220119_at	EPB41L4A	-0,74	5,25	-7,45	5,05E-09	2,32E-06	10,61
AFFX-LysX-3_at	NA	-1,57	10,16	-7,45	5,10E-09	2,32E-06	10,59
211980_at	COL4A1	1,90	9,28	7,41	5,81E-09	2,56E-06	10,47
205464_at	SCNN1B	-1,28	6,00	-7,41	5,85E-09	2,56E-06	10,46
216010_x_at	FUT3	-1,68	5,57	-7,36	6,77E-09	2,90E-06	10,32
201768_s_at	CLINT1	-0,89	8,71	-7,35	7,06E-09	2,93E-06	10,29
220197_at	ATP6V0A4	-1,21	4,81	-7,34	7,11E-09	2,93E-06	10,28
204779_s_at	HOXB7	1,51	5,76	7,31	7,80E-09	3,09E-06	10,19
206835_at	STATH	-5,39	6,41	-7,31	7,85E-09	3,09E-06	10,18
215103_at	CYP2C18	-1,69	4,83	-7,31	7,90E-09	3,09E-06	10,18
39249_at	AQP3	-2,24	8,94	-7,30	8,28E-09	3,16E-06	10,13
201892_s_at	IMPDH2	-0,89	9,74	-7,29	8,42E-09	3,16E-06	10,12
214088_s_at	FUT3	-1,43	7,06	-7,29	8,51E-09	3,16E-06	10,11
210809_s_at	POSTN	3,06	9,31	7,26	9,40E-09	3,43E-06	10,01
35148_at	TJP3	-1,01	4,89	-7,21	1,07E-08	3,85E-06	9,89
213174_at	TTC9	-1,56	6,85	-7,17	1,23E-08	4,34E-06	9,76
220026_at	CLCA4	-4,05	7,26	-7,13	1,39E-08	4,85E-06	9,64
211712_s_at	ANXA9	-1,75	5,21	-7,10	1,52E-08	5,16E-06	9,55
205319_at	PSCA	-2,67	5,67	-7,10	1,53E-08	5,16E-06	9,55
202185_at	PLOD3	0,91	6,99	7,07	1,67E-08	5,56E-06	9,46
211559_s_at	CCNG2	-1,59	7,44	-7,06	1,74E-08	5,69E-06	9,42
218928_s_at	SLC37A1	-0,67	6,00	-7,00	2,12E-08	6,83E-06	9,24
203962_s_at	NEBL	-1,81	7,20	-6,99	2,19E-08	6,88E-06	9,20
218035_s_at	RBM47	-1,47	9,03	-6,99	2,19E-08	6,88E-06	9,20
218552_at	ECHDC2	-0,96	7,59	-6,94	2,56E-08	7,78E-06	9,05
203845_at	KAT2B	-1,46	6,81	-6,94	2,56E-08	7,78E-06	9,05
211981_at	COL4A1	1,95	7,90	6,94	2,58E-08	7,78E-06	9,04
211465_x_at	FUT6	-1,32	7,76	-6,92	2,67E-08	7,94E-06	9,01
202731_at	NA	-1,57	8,08	-6,90	2,94E-08	8,61E-06	8,92
206605_at	ENDOU	-2,28	5,34	-6,89	2,99E-08	8,62E-06	8,91
202027_at	TMEM184	0,73	6,77	6,88	3,05E-08	8,62E-06	8,89
	B						
201884_at	CEACAM5	-2,82	8,69	-6,88	3,06E-08	8,62E-06	8,88
219296_at	ZDHHC13	-1,27	7,96	-6,86	3,31E-08	9,21E-06	8,81

APPENDIX

Affy-ID	Gen	log ₂ FC	AveExpr	t	p-Wert	FDR	B
219100_at	OBFC1	-1,32	7,32	-6,85	3,39E-08	9,31E-06	8,79
212593_s_at	NA	-1,18	9,72	-6,83	3,61E-08	9,81E-06	8,73
204034_at	ETHE1	-1,03	8,29	-6,83	3,65E-08	9,81E-06	8,71
203675_at	NUCB2	-2,02	7,68	-6,79	4,04E-08	1,07E-05	8,62
205198_s_at	ATP7A	-0,66	5,85	-6,79	4,12E-08	1,08E-05	8,60
204867_at	GCHFR	-1,44	6,67	-6,78	4,19E-08	1,09E-05	8,58
203747_at	AQP3	-2,02	8,05	-6,77	4,34E-08	1,11E-05	8,55
222127_s_at	EXOC1	-0,73	8,76	-6,76	4,51E-08	1,14E-05	8,51
220962_s_at	PADI1	-1,13	7,48	-6,75	4,64E-08	1,16E-05	8,49
219856_at	C1orf116	-1,48	5,97	-6,74	4,80E-08	1,19E-05	8,45
203878_s_at	MMP11	1,48	6,94	6,71	5,27E-08	1,28E-05	8,36
208771_s_at	LTA4H	-0,77	10,16	-6,71	5,29E-08	1,28E-05	8,36
209466_x_at	PTN	-1,83	7,53	-6,71	5,34E-08	1,28E-05	8,35
211964_at	COL4A2	1,96	9,29	6,70	5,42E-08	1,29E-05	8,34
217110_s_at	MUC4	-2,45	5,38	-6,70	5,51E-08	1,29E-05	8,32
216244_at	IL1RN	-2,59	5,27	-6,67	5,99E-08	1,39E-05	8,24
205185_at	SPINK5	-3,70	8,22	-6,66	6,16E-08	1,42E-05	8,21
209109_s_at	TSPAN6	-1,76	8,13	-6,66	6,28E-08	1,43E-05	8,19
201272_at	AKR1B1	0,98	8,55	6,65	6,50E-08	1,46E-05	8,16
205272_s_at	NA	-5,18	7,30	-6,63	6,73E-08	1,48E-05	8,13

APPENDIX

Tab. 37: TOP100 Liste differenziell exprimierter Gene in Mvs.PT, sortiert nach der FDR.

Affy-ID	Gen	log ₂ FC	AveExpr	t	p-Wert	FDR	B
219452_at	DPEP2	1,16	5,03	7,55	3,76E-09	8,38E-05	10,73
202992_at	C7	2,47	6,62	7,00	2,14E-08	0,00024	9,13
204220_at	GMFG	2,01	7,69	6,62	7,13E-08	0,00035	8,01
219371_s_at	KLF2	1,86	8,04	6,55	8,83E-08	0,00035	7,81
212829_at	PIP4K2A	1,28	8,16	6,43	1,27E-07	0,00035	7,47
204661_at	CD52	2,53	8,37	6,42	1,34E-07	0,00035	7,42
218870_at	ARHGAP15	2,12	7,33	6,41	1,40E-07	0,00035	7,38
202003_s_at	ACAA2	1,21	7,80	6,38	1,50E-07	0,00035	7,32
204912_at	IL10RA	1,64	7,41	6,38	1,54E-07	0,00035	7,29
221989_at	NA	0,63	7,73	6,36	1,61E-07	0,00035	7,25
203962_s_at	NEBL	-1,64	7,20	-6,34	1,73E-07	0,00035	7,19
213620_s_at	NA	1,77	8,10	6,28	2,08E-07	0,00039	7,02
206377_at	FOXF2	-0,94	6,47	-6,22	2,49E-07	0,00041	6,85
206949_s_at	RUSC1	-0,68	7,29	-6,20	2,70E-07	0,00041	6,77
203381_s_at	APOE	1,62	7,26	6,17	2,96E-07	0,00041	6,68
204683_at	ICAM2	1,42	6,61	6,16	3,05E-07	0,00041	6,66
212884_x_at	APOE	0,94	7,70	6,16	3,09E-07	0,00041	6,64
215051_x_at	AIF1	1,02	8,32	6,13	3,39E-07	0,00042	6,56
204789_at	FMNL1	0,79	5,88	6,09	3,81E-07	0,00045	6,45
207238_s_at	PTPRC	2,35	7,25	6,06	4,21E-07	0,00046	6,36
218424_s_at	STEAP3	-0,99	7,87	-6,04	4,51E-07	0,00046	6,29
204416_x_at	APOC1	1,36	7,86	6,04	4,55E-07	0,00046	6,28
207277_at	CD209	1,58	7,02	5,99	5,31E-07	0,00046	6,14
203760_s_at	SLA	1,03	7,04	5,98	5,46E-07	0,00046	6,11
34210_at	CD52	2,63	7,44	5,97	5,60E-07	0,00046	6,09
207571_x_at	THEMIS2	1,07	5,76	5,97	5,65E-07	0,00046	6,08
203961_at	NEBL	-1,54	7,15	-5,95	6,09E-07	0,00046	6,01
209734_at	NCKAP1L	1,41	7,07	5,93	6,39E-07	0,00046	5,97
202803_s_at	ITGB2	1,56	7,44	5,92	6,51E-07	0,00046	5,95
204606_at	CCL21	2,70	7,14	5,92	6,60E-07	0,00046	5,94
211105_s_at	NFATC1	1,12	7,51	5,90	7,00E-07	0,00046	5,88
203416_at	CD53	1,84	9,21	5,89	7,22E-07	0,00046	5,85
213285_at	TMEM30B	-1,44	7,28	-5,88	7,55E-07	0,00046	5,81
211663_x_at	PTGDS	1,33	7,29	5,87	7,79E-07	0,00046	5,78
37117_at	NA	-1,12	5,79	-5,86	7,91E-07	0,00046	5,77
201130_s_at	CDH1	-1,56	7,03	-5,86	8,00E-07	0,00046	5,76
207339_s_at	LTB	2,38	7,09	5,85	8,21E-07	0,00046	5,73
40420_at	STK10	0,99	6,44	5,84	8,49E-07	0,00046	5,70
213160_at	DOCK2	1,81	5,94	5,84	8,53E-07	0,00046	5,70
201656_at	ITGA6	-1,01	8,90	-5,83	8,70E-07	0,00046	5,68
213309_at	PLCL2	1,13	6,54	5,83	8,83E-07	0,00046	5,67

APPENDIX

Affy-ID	Gen	log ₂ FC	AveExpr	t	p-Wert	FDR	B
201720_s_at	LAPTM5	1,73	8,98	5,83	8,94E-07	0,00046	5,66
218223_s_at	PLEKHO1	1,21	6,62	5,83	8,94E-07	0,00046	5,65
208824_x_at	CDK16	-0,60	7,21	-5,82	9,25E-07	0,00046	5,62
203523_at	LSP1	0,93	6,99	5,81	9,42E-07	0,00046	5,61
212187_x_at	PTGDS	1,84	8,27	5,81	9,45E-07	0,00046	5,60
218261_at	AP1M2	-0,88	6,71	-5,80	9,85E-07	0,00046	5,56
214084_x_at	NA	1,29	5,54	5,79	9,93E-07	0,00046	5,56
203047_at	STK10	1,05	6,50	5,79	1,01E-06	0,00046	5,55
214790_at	SENP6	0,64	4,57	5,77	1,07E-06	0,00047	5,49
203920_at	NR1H3	0,85	6,05	5,76	1,10E-06	0,00047	5,46
212587_s_at	PTPRC	2,26	7,50	5,76	1,11E-06	0,00047	5,45
205611_at	NA	0,48	6,08	5,75	1,13E-06	0,00047	5,44
212750_at	PPP1R16B	1,41	6,43	5,75	1,15E-06	0,00047	5,42
201394_s_at	RBM5	0,63	7,96	5,73	1,22E-06	0,00049	5,37
204265_s_at	GPSM3	1,00	6,67	5,71	1,29E-06	0,00051	5,32
211748_x_at	PTGDS	1,93	8,74	5,69	1,36E-06	0,00051	5,26
209539_at	ARHGEF6	1,39	7,72	5,69	1,40E-06	0,00051	5,23
209614_at	ADH1B	0,60	4,78	5,68	1,44E-06	0,00051	5,21
209901_x_at	AIF1	0,94	8,00	5,68	1,45E-06	0,00051	5,21
215177_s_at	ITGA6	-1,18	8,80	-5,67	1,45E-06	0,00051	5,20
204588_s_at	SLC7A7	0,69	7,64	5,67	1,46E-06	0,00051	5,20
204118_at	CD48	2,17	8,38	5,67	1,49E-06	0,00051	5,18
221526_x_at	PARD3	-0,96	8,56	-5,66	1,51E-06	0,00051	5,17
204057_at	IRF8	2,08	7,92	5,66	1,52E-06	0,00051	5,16
202663_at	WIPF1	1,16	4,64	5,65	1,55E-06	0,00051	5,14
216218_s_at	PLCL2	0,63	4,27	5,65	1,55E-06	0,00051	5,14
202890_at	MAP7	-1,62	6,49	-5,65	1,56E-06	0,00051	5,13
205709_s_at	CDS1	-1,02	5,63	-5,63	1,66E-06	0,00054	5,08
205419_at	GPR183	1,69	5,63	5,62	1,73E-06	0,00055	5,04
203382_s_at	APOE	1,37	7,13	5,59	1,90E-06	0,00060	4,95
201288_at	ARHGDIB	1,41	9,33	5,58	1,94E-06	0,00060	4,93
210644_s_at	LAIR1	0,71	6,71	5,58	1,97E-06	0,00060	4,92
214181_x_at	LST1	1,06	6,77	5,56	2,07E-06	0,00062	4,87
218942_at	PIP4K2C	-0,67	7,81	-5,56	2,10E-06	0,00062	4,86
202838_at	FUCA1	1,45	8,98	5,55	2,14E-06	0,00062	4,84
209337_at	PSIP1	1,13	8,41	5,54	2,20E-06	0,00062	4,81
202617_s_at	MECP2	0,46	5,79	5,54	2,23E-06	0,00062	4,80
41577_at	PPP1R16B	1,38	5,28	5,54	2,23E-06	0,00062	4,80
209442_x_at	ANK3	-1,08	7,82	-5,54	2,24E-06	0,00062	4,80
205789_at	CD1D	0,56	4,21	5,52	2,34E-06	0,00064	4,76
203508_at	TNFRSF1B	1,28	7,08	5,52	2,39E-06	0,00064	4,74
203458_at	SPR	-0,60	7,06	-5,52	2,41E-06	0,00064	4,73
202597_at	IRF6	-1,69	9,02	-5,51	2,42E-06	0,00064	4,73

APPENDIX

Affy-ID	Gen	log ₂ FC	AveExpr	t	p-Wert	FDR	B
210086_at	HR	-0,59	5,92	-5,51	2,47E-06	0,00065	4,71
209613_s_at	ADH1B	3,04	5,77	5,50	2,56E-06	0,00066	4,67
202747_s_at	ITM2A	1,42	8,03	5,50	2,57E-06	0,00066	4,67
203786_s_at	TPD52L1	-1,49	9,52	-5,49	2,59E-06	0,00066	4,66
203761_at	SLA	1,27	7,67	5,49	2,63E-06	0,00066	4,65
212706_at	NA	0,89	6,04	5,47	2,78E-06	0,00069	4,60
201721_s_at	LAPTM5	1,41	9,98	5,46	2,87E-06	0,00070	4,57
211796_s_at	NA	2,46	7,85	5,46	2,91E-06	0,00070	4,55
203923_s_at	CYBB	1,40	6,81	5,45	2,96E-06	0,00070	4,54
217838_s_at	EVL	1,20	7,13	5,44	3,04E-06	0,00070	4,51
219506_at	C1orf54	0,79	7,67	5,44	3,06E-06	0,00070	4,51
202947_s_at	GYPC	1,08	7,21	5,44	3,10E-06	0,00070	4,49
210582_s_at	LIMK2	-0,69	7,42	-5,43	3,17E-06	0,00070	4,48
209933_s_at	CD300A	0,90	7,02	5,43	3,20E-06	0,00070	4,46
220704_at	IKZF1	0,87	4,13	5,43	3,21E-06	0,00070	4,46
204581_at	CD22	1,52	5,51	5,42	3,27E-06	0,00070	4,45

APPENDIX

Tab. 38: TOP100 Liste differenziell exprimierter Gene in Mvs.N, sortiert nach der FDR.

Affy-ID	Gen	log ₂ FC	AveExpr	t	p-Wert	FDR	B
203961_at	NEBL	-3,50	7,15	-13,49	2,68E-16	2,95E-12	26,73
91826_at	EPS8L1	-3,04	6,85	-13,35	3,72E-16	2,95E-12	26,42
203962_s_at	NEBL	-3,45	7,20	-13,33	3,97E-16	2,95E-12	26,36
211883_x_at	CEACAM1	-2,19	6,76	-13,19	5,53E-16	3,08E-12	26,04
211889_x_at	CEACAM1	-2,23	6,15	-12,36	4,40E-15	1,86E-11	24,08
219476_at	C1orf116	-3,59	7,41	-12,31	5,00E-15	1,86E-11	23,96
209498_at	CEACAM1	-2,92	7,31	-12,22	6,16E-15	1,96E-11	23,76
211372_s_at	IL1R2	-2,81	6,52	-12,04	9,82E-15	2,74E-11	23,31
205403_at	IL1R2	-3,18	6,29	-11,77	2,00E-14	4,96E-11	22,64
221665_s_at	EPS8L1	-2,80	6,26	-11,51	3,94E-14	8,79E-11	21,99
210608_s_at	FUT2	-2,28	7,12	-11,35	5,95E-14	1,21E-10	21,59
205765_at	CYP3A5	-2,81	5,78	-11,05	1,35E-13	2,51E-10	20,81
208126_s_at	CYP2C18	-2,80	6,11	-10,99	1,59E-13	2,60E-10	20,65
207059_at	PAX9	-2,44	6,46	-10,98	1,63E-13	2,60E-10	20,63
204755_x_at	HLF	-2,38	5,91	-10,94	1,82E-13	2,70E-10	20,52
206539_s_at	CYP4F12	-1,41	6,19	-10,83	2,43E-13	3,32E-10	20,24
221655_x_at	EPS8L1	-2,28	7,01	-10,80	2,63E-13	3,32E-10	20,17
211689_s_at	TMPRSS2	-2,61	5,62	-10,80	2,68E-13	3,32E-10	20,15
218779_x_at	EPS8L1	-2,55	7,01	-10,71	3,35E-13	3,93E-10	19,93
205470_s_at	KLK11	-3,40	7,57	-10,68	3,68E-13	4,10E-10	19,84
218885_s_at	GALNT12	-2,48	6,81	-10,58	4,90E-13	5,16E-10	19,57
220090_at	CRNN	-6,06	7,32	-10,56	5,09E-13	5,16E-10	19,53
39249_at	AQP3	-3,21	8,94	-10,46	6,71E-13	6,50E-10	19,27
211726_s_at	FMO2	-2,53	7,36	-10,32	1,01E-12	9,39E-10	18,87
207802_at	CRISP3	-5,85	5,95	-10,29	1,08E-12	9,58E-10	18,81
205190_at	PLS1	-2,48	6,36	-10,28	1,12E-12	9,58E-10	18,78
203431_s_at	ARHGAP3 2	-2,25	7,73	-10,20	1,38E-12	1,14E-09	18,57
203747_at	AQP3	-3,04	8,05	-10,18	1,49E-12	1,16E-09	18,50
218180_s_at	EPS8L2	-2,15	7,40	-10,17	1,51E-12	1,16E-09	18,49
206576_s_at	CEACAM1	-2,30	6,57	-10,13	1,70E-12	1,24E-09	18,37
201884_at	CEACAM5	-4,15	8,69	-10,13	1,72E-12	1,24E-09	18,36
206884_s_at	SCEL	-5,37	6,79	-10,08	1,94E-12	1,35E-09	18,24
209109_s_at	TSPAN6	-2,66	8,13	-10,06	2,09E-12	1,41E-09	18,17
216836_s_at	ERBB2	-1,77	7,91	-10,04	2,22E-12	1,46E-09	18,11
218026_at	COA3	-1,11	9,73	-9,99	2,55E-12	1,63E-09	17,98
204416_x_at	APOC1	2,24	7,86	9,92	3,05E-12	1,89E-09	17,81
204034_at	ETHE1	-1,47	8,29	-9,75	5,09E-12	3,07E-09	17,31
216025_x_at	CYP2C9	-1,06	5,97	-9,71	5,57E-12	3,24E-09	17,22
203453_at	SCNN1A	-3,71	7,55	-9,71	5,67E-12	3,24E-09	17,21
210544_s_at	ALDH3A2	-1,85	8,25	-9,69	5,96E-12	3,32E-09	17,16

APPENDIX

Affy-ID	Gen	log ₂ FC	AveExpr	t	p-Wert	FDR	B
214234_s_at	CYP3A5	-1,93	5,40	-9,64	6,82E-12	3,70E-09	17,03
205709_s_at	CDS1	-1,71	5,63	-9,46	1,16E-11	6,15E-09	16,51
202454_s_at	ERBB3	-2,52	6,39	-9,44	1,22E-11	6,33E-09	16,46
205185_at	SPINK5	-5,22	8,22	-9,40	1,39E-11	7,02E-09	16,34
219296_at	ZDHC13	-1,73	7,96	-9,39	1,44E-11	7,12E-09	16,31
213929_at	EXPH5	-2,12	5,61	-9,37	1,51E-11	7,14E-09	16,26
215103_at	CYP2C18	-2,17	4,83	-9,37	1,51E-11	7,14E-09	16,26
202053_s_at	ALDH3A2	-2,13	8,84	-9,34	1,66E-11	7,71E-09	16,17
208407_s_at	NA	-1,60	8,02	-9,29	1,90E-11	8,65E-09	16,04
205783_at	KLK13	-3,99	6,35	-9,28	1,96E-11	8,75E-09	16,00
203407_at	PPL	-4,02	9,06	-9,24	2,22E-11	9,57E-09	15,88
218792_s_at	BSPRY	-1,25	5,64	-9,24	2,23E-11	9,57E-09	15,88
220026_at	CLCA4	-5,23	7,26	-9,21	2,42E-11	1,02E-08	15,80
219856_at	C1orf116	-2,02	5,97	-9,20	2,50E-11	1,03E-08	15,77
211885_x_at	FUT6	-1,72	5,48	-9,19	2,57E-11	1,04E-08	15,74
209108_at	TSPAN6	-2,65	7,98	-9,15	2,87E-11	1,14E-08	15,64
203704_s_at	RREB1	-0,87	8,53	-9,14	2,97E-11	1,16E-08	15,60
211737_x_at	PTN	-2,88	7,28	-9,11	3,20E-11	1,21E-08	15,53
214070_s_at	ATP10B	-2,05	6,56	-9,11	3,22E-11	1,21E-08	15,52
204753_s_at	HLF	-2,15	5,94	-9,11	3,26E-11	1,21E-08	15,51
217110_s_at	MUC4	-3,33	5,38	-9,08	3,54E-11	1,29E-08	15,43
202054_s_at	ALDH3A2	-1,94	8,43	-9,05	3,88E-11	1,37E-08	15,35
214088_s_at	FUT3	-1,78	7,06	-9,05	3,88E-11	1,37E-08	15,34
202096_s_at	TSPO	-1,31	9,44	-9,03	4,15E-11	1,44E-08	15,28
203021_at	SLPI	-4,22	10,01	-8,98	4,80E-11	1,65E-08	15,14
219850_s_at	EHF	-4,17	7,42	-8,94	5,44E-11	1,84E-08	15,02
209465_x_at	PTN	-1,75	5,94	-8,88	6,48E-11	2,15E-08	14,85
218186_at	RAB25	-3,28	8,15	-8,86	6,80E-11	2,20E-08	14,80
209466_x_at	PTN	-2,41	7,53	-8,86	6,86E-11	2,20E-08	14,79
214734_at	EXPH5	-2,65	7,67	-8,85	6,92E-11	2,20E-08	14,78
210398_x_at	FUT6	-1,66	5,38	-8,83	7,42E-11	2,33E-08	14,71
220030_at	STYK1	-1,39	4,63	-8,77	8,91E-11	2,75E-08	14,54
203382_s_at	APOE	2,15	7,13	8,77	9,02E-11	2,75E-08	14,52
213285_at	TMEM30B	-2,14	7,28	-8,73	1,01E-10	3,03E-08	14,42
210644_s_at	LAIR1	1,11	6,71	8,71	1,06E-10	3,15E-08	14,37
213159_at	PCNX	0,58	6,04	8,67	1,22E-10	3,57E-08	14,23
200764_s_at	CTNNA1	-1,15	8,27	-8,65	1,26E-10	3,66E-08	14,20
204754_at	HLF	-1,49	5,61	-8,65	1,28E-10	3,66E-08	14,18
211657_at	CEACAM6	-4,01	8,06	-8,64	1,31E-10	3,70E-08	14,16
204351_at	S100P	-3,80	8,47	-8,62	1,38E-10	3,84E-08	14,11
204519_s_at	PLL	-1,24	5,98	-8,61	1,44E-10	3,95E-08	14,07
217109_at	MUC4	-3,56	6,16	-8,59	1,51E-10	4,05E-08	14,03
212884_x_at	APOE	1,31	7,70	8,59	1,51E-10	4,05E-08	14,02

APPENDIX

Affy-ID	Gen	log ₂ FC	AveExpr	t	p-Wert	FDR	B
215471_s_at	MAP7	-1,18	6,51	-8,59	1,53E-10	4,05E-08	14,01
218035_s_at	RBM47	-1,80	9,03	-8,58	1,56E-10	4,09E-08	13,99
211882_x_at	FUT6	-1,57	5,81	-8,56	1,70E-10	4,40E-08	13,91
203757_s_at	CEACAM6	-4,24	8,31	-8,55	1,73E-10	4,43E-08	13,89
216243_s_at	IL1RN	-3,03	7,90	-8,54	1,77E-10	4,48E-08	13,87
219267_at	GLTP	-1,98	7,50	-8,53	1,86E-10	4,65E-08	13,82
203381_s_at	APOE	2,24	7,26	8,51	1,93E-10	4,79E-08	13,78
208336_s_at	TECR	-1,39	8,10	-8,46	2,24E-10	5,47E-08	13,64
209772_s_at	CD24	-3,11	9,12	-8,46	2,26E-10	5,47E-08	13,63
214399_s_at	KRT4	-1,81	5,05	-8,46	2,28E-10	5,47E-08	13,62
202889_x_at	MAP7	-1,26	6,13	-8,45	2,34E-10	5,55E-08	13,60
213572_s_at	SERPINB1	-2,78	8,52	-8,43	2,46E-10	5,76E-08	13,55
202525_at	PRSS8	-1,71	6,69	-8,43	2,48E-10	5,76E-08	13,54
221696_s_at	STYK1	-0,74	4,51	-8,41	2,68E-10	6,15E-08	13,47
210461_s_at	ABLIM1	-1,31	7,42	-8,40	2,71E-10	6,16E-08	13,46
212268_at	SERPINB1	-2,95	9,27	-8,39	2,83E-10	6,33E-08	13,41
210096_at	CYP4B1	-2,67	5,21	-8,39	2,84E-10	6,33E-08	13,41

8.4 Eidesstattliche Versicherung

Hiermit erkläre ich an Eides statt, dass ich meine Dissertation selbstständig und nur unter Verwendung der angegebenen Hilfsmittel angefertigt habe. Ich habe keinen anderen Promotionsversuch unternommen.

Mainz, den