Dissertation

submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Proteomics-based discovery of novel vascular accessible markers in kidney cancer

presented by

Amelie Saskia Benk

Dissertation

submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

presented by
Diplom-Biologin (t.o.) Amelie Saskia Benk
born in Sindelfingen, Germany
Date of Oral-examination: 18 September 2014

Proteomics-based discovery of novel vascular accessible markers in kidney cancer

Referees: Prof. Dr. Andreas Trumpp

Prof. Dr. Jonathan Sleeman

The work presented in this thesis was started in October 2010 and completed June 2014 under the supervision of Dr. Christoph Rösli in the research group “Biomarker discovery”, Department of “Stem cells and cancer” at the German Cancer Research Center (DKFZ), Heidelberg, as well as at the Heidelberg Institute for Stem Cells and Experimental Medicine (HI-STEM) gGmbH, Heidelberg.

To my grandmother Ömle

ACKNOWLEDGMENTS

At this point I would like to thank all the people who contributed to the project and supported me during my doctoral studies.

First I would like to thank my PhD supervisor Dr. Christoph Rösli for his guidance and the opportunity to work in his research group on such an interesting project. In his lab I learned a lot of new techniques and got to know and appreciated the work with a mass spectrometer. In addition I would like to thank him for having always an open door for discussing further experiments and the possibility to jointly write and publish a review. Furthermore I want to recognise his active membership in my Thesis Advisory Committee (TAC) of the DKFZ Graduate School and his participation as co-examiner and referee in my doctoral studies.

In addition, I want to thank Prof. Andreas Trumpp for the possibility to perform my PhD in his laboratory as well as for his suggestions in scientific discussions in lab meetings and as part of my TAC-committee. Furthermore I am grateful for his willingness in being the first examiner and referee in my PhD defence.

My thank goes as well to Prof. Jonathan Sleeman for his kind acceptance being second PhD examiner and his helpful participation in my TAC-committee. His suggested directions were very helpful and lead to fruitful discussions from an outside view.

I would furthermore like to thank Dr. Martina Schnölzer for giving me always very helpful support and suggestions while being part of my TAC-committee.

Additionally I would like to thank Prof. Matthias Mayer for kindly accepting the chairmanship in my PhD defence.

Many thanks go my collaborators and collegues Dr. Martin Sprick and Dr. Teresa Rigo-Watermeier for sharing the orthotopic kidney cancer mouse models and answering my questions related with that. I am very grateful to Corinna Klein for her patience and lessons in all mouse related issues.

Furthemore I want to thank Alexander Kerner for his contribution in the quantification analysis of my proteomic data sets and his support in all questions related to computers.

Especially I am very thankful to Dr. Hind Medyouf for her great commitment and encouragement in the LK/LSK project as well as her willingness in sharing her knowledge in all bone marrow related issues. Although the project finally didn't work out, I learned a lot from her scientific suggestions and experimental experience.

Additionally I would like to thank the Small Animal Imaging Center from the DKFZ with Dr. Manfred Jugold as head and Sabrina Emmerich and Inna Babushkina as operators of the MRT scans for their, often spontaneous, assistance.

A thank you goes also to the animal care takers of T137c in the DKFZ animal house for their great support in nourishing my mice.

In addition, I am very grateful to my uncle Andreas Krazeisen for spending a lot of time and effort in the development of the perfusion needles without these in vivo biotinylation of mice would have been much more difficult.

Last but not least, a big thank you goes to the other colleagues in the Rösli / Trumpp lab and the neighbouring labs for inspiring discussions and their readiness to help! Especially I want mention Lisa von Paleske and Dr. Katharina Billian-Frey for their scientific suggestions as well as support and encouragement throughout my studies.

My thank also goes to the Helmholtz International Graduate School of Cancer Research and its coordinators for a great platform with interesting seminars and scientific courses as well as the opportunity to connect with other PhD students from different research fields.

Finally I would like to thank the Foundation of German Business for a three years PhD stipend, as well as the BioRN Spitzencluster “Cell-based and Molecular Medicine” which is funded by the German Bundesministerium für Bildung und Forschung (BMBF), the EU-FP7 Program “EuroSyStem”, the SFB 873 supported by the Deutsche Forschungsgemeinschaft (DFG), and the Dietmar Hopp Foundation for their financial support of the project.

TABLE OF CONTENTS

ABBREVIATIONS

ACN	acetonitrile
ACTH	adrenocorticotropic hormone
AJCC	American Joint Committee on Cancer
AKT	protein kinase B
ALDH1A1	aldehyde dehydrogenase 1 family, member A1
ALP	antibody assisted lectin profiling
ALT	alanine aminotransferase
AST	aspartate aminotransferase
AUC	area under the curve
BCA	bicinchoninic acid
BHD	Birt-Hogg-Dubé
BSA	bovine serum albumin
CA9	carbonic anhydrase 9
CAGE	cap analysis of gene expression
ccRCC	clear cell renal cell carcinoma
CD	cluster of differentiation
cDNA	complementary DNA
CE	capillary electrophoresis
CHCA	alpha-cyano-4-hydroxycinnamic acid
ChIP	chromatin immunoprecipitation
CID	collision-induced dissociation
CRF	corticotropin-releasing factor
CSC	cancer stem cell
C_T	cycle threshold
Cy	cyanine dye
Da	dalton
DAPI	4',6-diamidino-2-phenylindole
DARPins	designed ankyrin repeat proteins
DIGE	difference gel electrophoresis
DMEM/F12	dulbecco's modified eagle medium/nutrient mixture F-12
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbent assay
emPAI	exponentially modified protein abundance index
EPO	erythropoietin
ESI	electrospray ionization
EST	expressed sequence tags
Fab	fragment antigen-binding
FAK	focal adhesion kinase
Fc	fragment crystallizable
FCS	fetal calf serum
FDA	Food and Drug Administration
FDR	false discovery rate

FFPE	formalin fixed, paraffin embedded
FGF	fibroblast growth factor
FGFR	fibroblast growth factor receptor
FLCN	folliculin
FTICR	Fourier transform ion cyclotron resonance
GANAB	neutral alpha-glucosidase AB
GC	gas chromatography
GEMM	genetically engineered mouse models
GFR	glomerular filtration rate
GLUD1	glutamate dehydrogenase 1
GLUT1	glucose transporter type 1
GV-SOLAS	Gesellschaft für Versuchstierkunde
H&E	hematoxylin & eosin
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HER2	human epidermal growth factor receptor 2
HIF	hypoxia-induced factors
HPA	Human Protein Atlas
HPLC	high performance liquid chromatography
HRP	horseradish peroxidase
ICAT	isotope-coded affinity tagging
ICPL	isotope-coded protein labelling
IEF	isoelectric focusing
IFN-α	interferon-alpha
Ig	immunoglobulin
IGFBP3	insulin-like growth factor-binding protein 3
IGOT	isotope-coded tagging
IHC	immunohistochemistry
IL-2	interleukin 2
IMDM	iscove's modified dulbecco's medium
iTRAQ	isobaric tagging for relative and absolute quantitation
JNK	c-Jun N-terminal kinase
LAMA1	laminin, alpha 1
LC	liquid chromatography
LCN2	Lipocalin-2
LTBP2	latent transforming growth factor beta-binding protein 2
m/z	mass-to-charge ratio
MALDI	matrix-assisted laser desorption ionization
M-CSF-1R	macrophage colony stimulating factor 1 receptor
miRNA	micro-ribonucleic acid
MiTF	microphthalmia-associated transcription factor
MPSS	massively parallel signature sequencing
MRM	multiple reaction monitoring
mRNA	messenger ribonucleic acid
MRT	magnetic resonance tomography
MS	mass spectrometry
MS/MS	tandem mass spectrometry
mTOR	mechanistic target of rapamycin
NGS	next generation sequencing

NHS	N-hydroxysuccinimide esters
NID1	nidogen 1
NK	natural killer cell
NMR	nuclear magnetic resonance
NOD	non-obese diabetic
OD	optical density
O-GlnNAc	β-linked N-acetylglucosamine
ORR	objective response rate
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDGFR	platelet-derived growth factor receptor
PEA	proximity extension assay
PEG	polyethylene glycol
PEGF	platelet-derived growth factor
PFS	progression-free survival
PI3K	phosphatidylinositol-4,5-bisphosphate 3-kinase
PKC	protein kinase C
PNGase	peptide -N-glycosidase
PVDF	polyvinylidene fluoride
Raf	rapidly accelerated fibrosarcoma
Ras	rat sarcoma
RCC	renal cell carcinoma
RP-HPLC	reversed-phase high-pressure liquid chromatography
rpm	revolutions per minute
RPN1	dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1
RPPA	reverse phase protein array
RT-qPCR	real time quantitative polymerase chain reaction
SAGE	serial analysis of gene expression
SCID	severe combined immunodeficiency
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SILAC	stable isotope labelling with amino acids in cell culture
SLC16A1	monocarboxylate transporter 1
SNP	single-nucleotide polymorphism
SPF	specific pathogen-free
SRM	selected reaction monitoring
TBS	tris-buffered saline
TCEP	tris(2-carboxyethyl)phosphine
TCGA	The Cancer Genome Atlas
TDB	trypsin digestion buffer
TFA	trifluoroacetic acid
TFE	transcription factor E
TFRC	transferrin receptor
TGFBI	transforming growth factor, beta-induced
TMAs	tissue micro arrays
TMB	3,3'-5,5'-Tetramethylbenzidine
TMT	tandem mass tags
TNXB	tenascin-X
TOF	time-of-flight

UPLC	ultra performance liquid chromatography
VEGFR	vascular endothelial growth factor receptor
VHL	von Hippel Lindau
WTSS	whole transcriptome shotgun sequencing
XIAP	X-linked inhibitor of apoptosis protein

1 ABSTRACTS

1.1 SUMMARY

With worldwide 209 000 new cases and 102 000 deaths per year renal cell carcinoma accounts for 2-3% of all cancer incidences and shows the highest mortality rate among all genitourinary malignancies [1, 2]. Currently, surgical resection of the tumour bulk is the solely treatment strategy offering a possible long-term cure. Alternative treatments for metastasising disease targeting the mTOR and the VEGF pathways only result in tumour mass reduction while inducing severe side effects and resistances [3]. In order to improve patient survival novel, effective and well-tolerated therapeutics particularly for advanced stages are urgently needed. Thereby the antibody-based delivery of bioactive molecules directly into the tumour has proven to be a promising approach [4]. Importantly, the availability of vascular accessible tumour-associated biomarkers as target molecules is a prerequisite for the design of this innovative drug type [5].

In this thesis a proteomics based biomarker discovery approach for the identification of novel drug targets accessible from the vasculature in clear cell renal cell carcinoma is presented. Thereby, vascular accessible proteins with an increased expression in the primary kidney tumour as well as proteins specifically regulated in corresponding lung metastases were investigated. For the enrichment of vascular accessible proteins the *in vivo* biotinylation strategy was applied [6, 7]. This method relies on the terminal perfusion of tumour-bearing mice with a reactive ester derivate of biotin resulting in the covalent biotin labelling of vascular accessible proteins. However, the conventional *in vivo* perfusion technique only allows for the biotinylation of proteins accessible from the systemic circulation. In order to investigate the vascular accessible proteome of lung metastases a novel technique for the perfusion of the pulmonary circulation was developed and successfully applied. Subsequent to the *in vivo* perfusion, biotinylated proteins were captured on streptavidin sepharose followed by on resin alkylation, delipidation and tryptic digestion. Resulting modified peptides were separated by nano-UPLC and analysed by MALDI-MS and MALDI-MS/MS. Protein identification was based on the identification of the corresponding proteotypic peptides. Relative quantification of protein expression levels was carried out on representative subsets by comparative analysis of summed intensities of peptides contributing to the same protein.

In the final dataset, mass spectrometric data deriving from seven patient-matched xenograft kidney tumour models, four spontaneous patient-matched xenograft lung metastasis models, healthy murine kidney and lung as well as three human renal primary cell lines were combined. In total, more than 3000 proteins were identified, whereof more than 80 % could

be quantified. Comparative analyses of different technical replicates and various biological samples demonstrated the high reproducibility of the techniques applied in the presented study. As expected, variances between the different patient-matched xenograft models were significantly bigger than between biological and technical replicates within one model.

The six biomarker candidates IGFBP3, Lcn2, Ltbp2, SLC16A1, TGFBI and TNXB were selected for further validation studies. Candidate proteins were validated based on both mRNA and protein levels using RT-qPCR and immunofluorescence, respectively. The results obtained in the initial proteomics discovery experiment were confirmed in the course of the validation study. The upregulation of several marker candidates in tumour tissue compared to healthy kidney was confirmed when analysing a small set of patient samples. Finally, a proof-of-principle biodistribution study investigating the vascular accessibility of TNXB for a monoclonal antibody indicated TNXB to be a potential drug target. However, TNXB accessibility for intravenously injected antibodies requires optimization.

In summary, this large scale proteomics study on vascular accessible proteins in clear cell renal cell carcinoma resulted in the identification of TNXB as a possible target protein for novel antibody-based kidney cancer therapies.

1.2 ZUSAMMENFASSUNG

Das Nierenzellkarzinom ist mit jährlich 209 000 Diagnosen und 102 000 Todesfällen für 2-3% aller Krebsneuerkrankungen weltweit verantwortlich und besitzt die höchste Todesrate aller Urogenitalerkrankungen [1, 2]. Die operative Entfernung der Tumormasse ist momentan die einzige Behandlungsmethode mit kurrativem Potential. Alternative Behandlungen, die beispielsweise bei metastasierendem Nierenzellkarzinom eingesetzt werden und die die Signalwege mTOR und VEGF zielgerichtet inhibieren, erzielen lediglich eine Tumorverkleinerung, führen jedoch simultan zu schweren Nebenwirkungen und Resistenzen [3]. Um die Überlebensrate von Nierenzellkarzinom-Patienten zu erhöhen ist die Entwicklung neuer, effektiverer und verträglicherer Therapeutika – insbesondere für fortgeschrittene Krebsstadien - von entscheidender Bedeutung. Der auf Antikörpern basierende direkte Transport von bioaktiven Molekülen in den Tumor ist hierbei ein vielversprechender Ansatz [4]. Entscheidend für die Entwicklung solch eines innovativen Medikaments ist die Verfügbarkeit eines Tumor assoziierten und über das Blutsystem erreichbaren Biomarkers als Zielmolekül [5].

In dieser Doktorarbeit wird ein auf Proteomics-Methoden basierender Biomarkeridentifikationsansatz für die Ermittlung neuer vaskulär zugänglicher Zielproteine für das klarzellige Nierenkrebskarzinom vorgestellt. Dabei werden vom Blutsystem zugängliche Proteine mit erhöhten Expressionsraten in Nierenprimärtumoren sowie in daraus hervorgehenden Lungenmetastasen untersucht. Für die Anreicherung vaskulär erreichbarer Proteine wurde die Strategie der *in vivo* Biotinylierung angewendet [6, 7]. Diese Methode basiert auf der terminalen Perfusion Tumor tragender Mäusen mit einem reaktiven Biotin-Esterreagenz, wodurch eine kovalente Biotinmarkierung vaskulär erreichbarer Proteine erzielt wird. Dabei ermöglicht die publizierte *in vivo* Perfundierungssmethode lediglich die Biotinylierung von über den Körperkreislauf erreichbaren Proteinen. Um das vaskulär zugängliche Protoeme von Lungenmetastasen untersuchen zu können, wurde eine neue Technik für die Perfusion des Lungenkreislaufs entwickelt und erfolgreich angewandt. Im Anschluss an die *in vivo* Perfusion wurden die biotinylierten Proteine an Streptavidin-Sepharose gebunden, auf der Säule alkyliert, delipidiert und enzymatisch mit Trypsin verdaut. Die resultierenden modifizierten Peptide wurden auf einer nano-UPLC aufgetrennt und auf dem Massenspektrometer mit MALDI-MS und -MS/MS analysiert. Die Proteinidentifikation wurde basierenden auf der Detektion proteotypischer Peptide durchgeführt. Eine relative Quantifizierung des Proteinexpressionniveaus wurde durch vergleichende Analysen der aufsummierten Peptideintenitäten eines Proteins in repräsentativen Teildatensätzen erzielt.

Im finalen Datensatz wurden die massenspektrometrische Daten von sieben Patienten-entstammenden Nierenkrebs-Xenograftmodellen, vier spontan entstehenden

Lungenmetastasen-Xenograftmodellen, drei humanen Nierenzelllinien sowie gesunden Nieren und Lungen von Mäusen zusammengeführt. Insgesamt konnten dabei mehr als 3000 Proteine identifiziert werden. Eine Proteinquantifizierung konnte für mehr als 80% dieser Proteine erzielt werden. Bei vergleichenden Analysen von unterschiedlichen technischen und biologischen Replikaten konnte eine hohe Reproduzierbarkeit der in dieser Untersuchung angewendeten Techniken dargestellt werden. Wie erwartet, wurden unter den Patienten-entstammenden Nierenkrebs-Xenograftmodellen signifikant höhere Streuungen beobachtet als in technischen und biologischen Replikaten eines einzelnen Models.

Für weiterführende Validierungsuntersuchungen wurden die sechs Biomarkerkandidaten IGFBP3, Lcn2, Ltbp2, MOT1, TGFBI und TNXB ausgewählt. Die Proteinkandidaten wurden sowohl auf mRNA-Ebene durch RT-qPCR als auch auf Proteinebene durch Immunofluoresenz verifiziert. Die Ergebnisse des vorangegangenen Proteomics-Identifizierungsexperiments konnten durch die Validierungsstudie bestätigt werden. Die Auswertung eines kleinen Satzes Patientenproben konnte die erhöhte Regulation verschiedener Markerkandidaten im Tumorgewebe im Vergleich zum gesundem Nierengewebe bestätigen. Abschließend wurde eine Test-Biodistributionsstudie zur Überprüfung der vaskulären Erreichbarkeit von TNXB durch einen monoklonalen Antikörper durchgeführt. Dabei konnte TNXB als Zielstruktur für potentielle Medikamente bestätigt werden. Allerdings besteht die Notwendigkeit, die Moleküle (z.B. Antikörper), welche nach intravenöser Verabreichung TNXB erreichen sollen, zu optimieren.

Zusammenfassend hat dieser Protoemics-Großversuch über vaskulär erreichbare Proteine in klarzelligem Nierenzellkarzinom TNXB als mögliches Zielprotein neuer auf Antikörpern basierender Nierenkrebstherapien hervorgebracht.

2 INTRODUCTION

2.1 RENAL CELL CARCINOMA

2.1.1 Incidence

Renal cell carcinoma (RCC), also known as Grawitz tumour or hypernephroma, is the most common type of kidney cancer with an occurrence of 80 % and shows the highest mortality rate among all genitourinary malignancies [1]. With 209 000 new cases and 102 000 deaths per year worldwide RCCs account for 2-3 % of all cancer incidences and this cancer entity is the seventh most frequent cancer in men and the tenth in women (Fig. 1) [2]. In the last decades, the incidence for RCC especially in Western countries is constantly increasing by about 3 % per year [8-13].

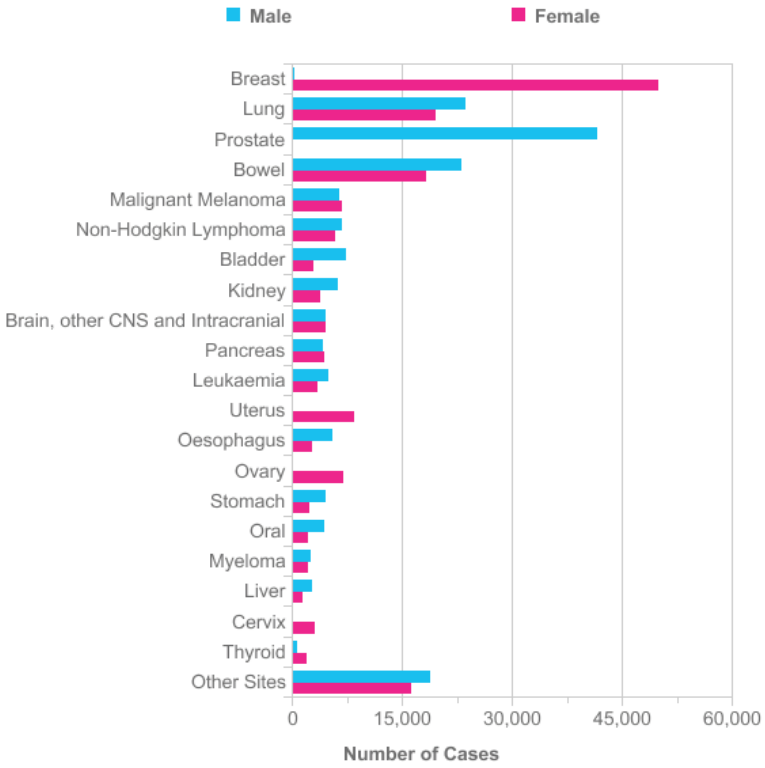


Fig. 1: Cancer incidences in the UK Kidney cancer is the 7th common cancer in men and 10th in women in the UK. Reprinted from the web page of Cancer Research UK [14]

RCC primarily affects patients in the sixth or seventh decade of life. Main potential risk factors are obesity and diabetes with 30 % in all cases [15, 16] tobacco exposure [17, 18], hypertension [19] as well as acquired cystic kidney disease, end stage renal failure and tuberous sclerosis [20, 21]. Still under discussion are other environmental causes such as the exposure to asbestos and nephrotoxic metals (As, Cd, Pb), analgesic drug use, acetaminophen, trichloroethylene, acryl amide as well as hormonal reasons (mainly in women) [9, 22-24].

Beside these risk factors physical activity and the consumption of fruits and vegetables can reduce the RCC incidence [25-27].

Renal malignancies often remain asymptomatic and are therefore usually detected coincidentally in a routine imaging examination applying magnetic resonance, computer tomography or ultrasound technologies. Only advanced stages show clinical symptoms such as haematuria, palpable abdominal masses and flank pain [28, 29]. Therefore at the stage of diagnosis, approximately 30 % of all patients already present a metastasising disease. This is quite dramatic. The five year survival rate of patients with local RCC is 91.8% while it drops to 12.1% for advanced stages [30]. The median survival of patients suffering from metastasing RCC is only 10-12 months [31, 32] (Fig. 2). In progressing disease about 25 % of all patients diagnosed with local RCC will develop metastases [33].

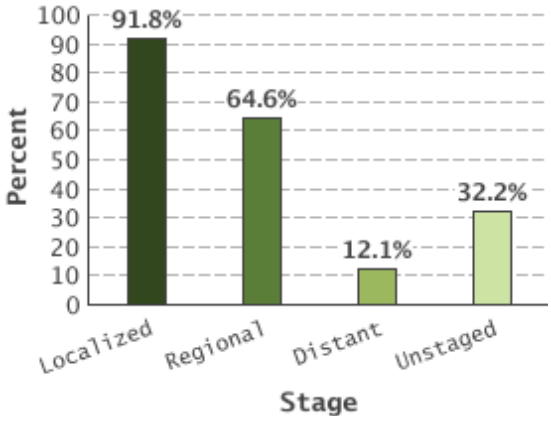


Fig. 2: Survival rates for RCC 5-year survival rates for patients with renal cell carcinoma (RCC) drop from 91.8 % for local disease to 12.1 % with metastasising RCC. Reprinted from the National Cancer Institute [30].

2.1.2 Clincial Staging and RCC subtypes

In the clinic RCC is categorized accordingly to the TNM (Tumour-Nodes-Metastasis) system of the American Joint Committee on Cancer (AJCC) [34]. Therefore different parameters such as tumour size and cancer cell infiltration of the surrounding tissue (T) as well as the affection of lymph nodes (N) and the metastatic status of other organs (M) are considered [35]. Additionally, the Fuhrmann grading system which is based on morphological characteristics of the cell nuclei is applied [36]. Depending on the percentage of healthy appearing nuclei a grade between 1 and 4, from high levels to low levels of healthy nuclei, is given. Both grading systems offer a certain prognostic value [36-38]. Low stages are associated with less aggressive tumours and longer survival rates.

Regarding immunohistological, genetic and molecular characteristics as well as clinical outcomes, RCC can be separated mainly in three different subtypes. About 3-5 % of all RCCs belong to the chromophobe, 10-15 % to the papilllary and 70-80% to the clear cell subtype (Fig. 3) [39-42]. Furthermore, a minor fraction of RCCs is grouped in the collecting duct, the

mucinous tubular, the spindle cell, the renal medullary and the MiTF-TFE translocation subtype [43].

The chromophobe RCC, which is arising from the cortical collecting duct, is characterised by large pale and eosinophilic cells with a perinuclear halo and manifests the best prognosis [44]. While most cases of the chromophobe subtype occur sporadically, it is also associated with the inherited Birt-Hogg-Dubé (BHD) syndrome connected with a mutated *FLCN* (folliculin) [45] and *BHD1* gene [44].

The papillary form can be further divided in type 1 and the more aggressive type 2 variant. Both forms derive from the proximal convoluted tubules and consist of small cuboid cells (type 1) or eosinophilic cells (type 2) covering the papillae. Papillary RCCs have been linked to hereditary mutations in the *C-Met* (type 1) [44, 46] and the fumarate hydratase gene (type 2) [44, 47] but papillary RCC arises mainly spontaneously. The most common and severe RCC subtype, the clear cell RCC, is described in the following chapter.

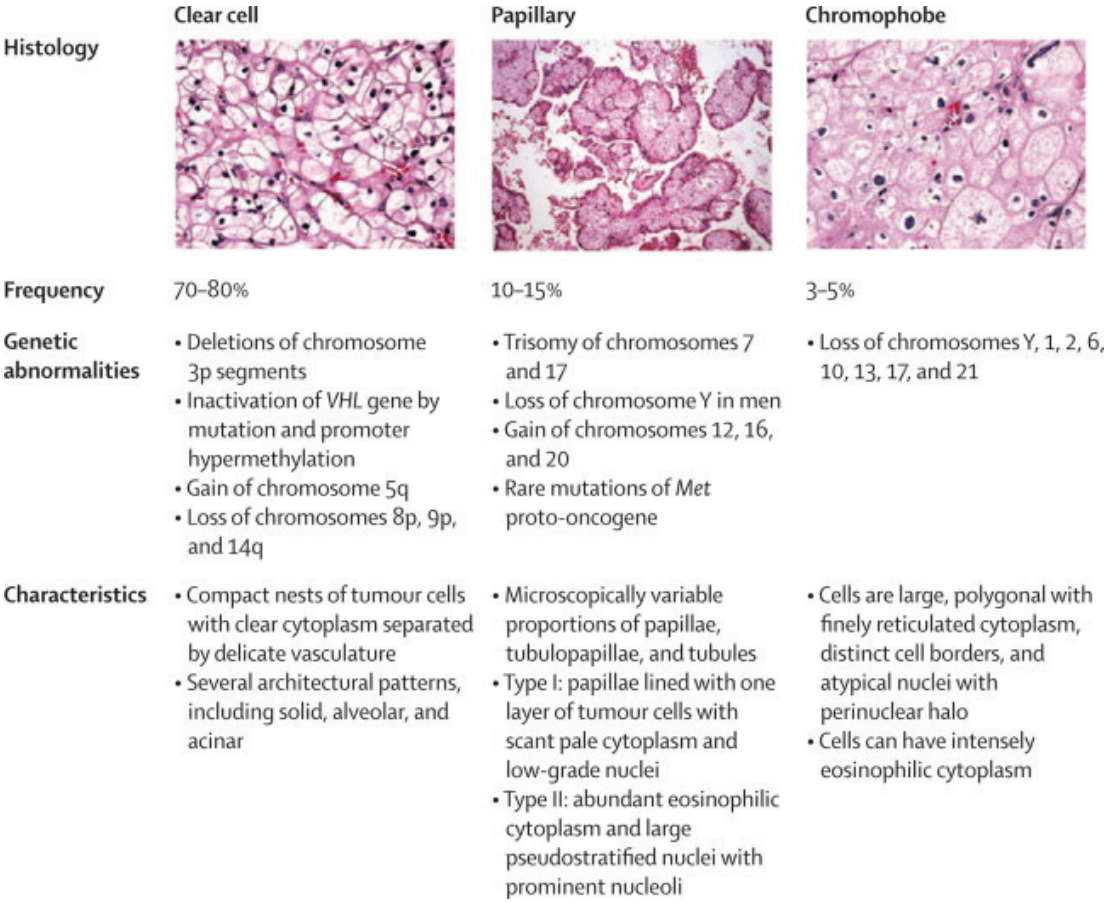


Fig. 3: Histological subtypes of RCC RCC can be classified in three main subtypes differing in histological and genetic properties and clinical prognosis. (Reprinted from Rini et al. 2009 [44])

2.1.3 Clear cell renal cell carcinoma

Clear cell RCC is the most common subtype of RCC with an occurrence of 80 % [44]. It shows a highly vascularised morphology with cells presenting a white cytoplasm due to the accumulation of glycogen and lipids. The cell of origin is still under discussion. While initially cells of the proximal tubule were considered as tumour origin [48] the distal nephron is currently discussed as the initial point of tumourigenesis [49]. Clear cell RCC has the worst prognosis of all RCCs [50, 51]. Almost half of all patients are diagnosed with metastases or develop these during therapy [31, 33]. Metastases occur primarily in the lungs (45%), the bones (30%), lymph nodes (22%) and in the liver (20%) [52]. The spreading of the tumour occurs mainly hematogenously [53] and is linked to the size of the primary tumour [54].

The vast majority of all clear cell RCC incidences (98%) occur sporadic and only very few cases are linked to hereditary mutations [55-57]. Most tumours show (80 %) a mutation in the tumour suppressor von Hippel Lindau (*VHL*) gene locus [58-60] in chromosome 3 (3p25-26) [60], which is leading to the Hippel Lindau syndrome. Patient suffering from the Hippel Lindau syndrome are predisposed to benign and malignant tumours. Transmitted in an autosomal-dominant manner [61], this mutation shows a frequency of in 1 of 36 000 [62] and is also linked to several other tumours such as haemangioblastoma in the central nervous system and pheochromocytoma [63]. Furthermore, germline mutations in *TSC1/2* genes predispose to tuberous sclerosis, which can lead to clear cell RCC [64].

In sporadic clear cell RCC incidences *VHL* alterations, either via point mutation or gene silencing (e.g. hypermethylation), are often observed [59, 65]. While inherited *VHL* defects are commonly carried on one allele only, it occasionally occurs in sporadic cases that both *VHL* alleles are altered [60]. The genetic changes in the VHL locus influence the folding properties and result in a non-functional protein [66].

2.1.4 Role of VHL

Functional VHL protein encoded by the tumour suppressor gene *VHL* forms in conjunction with with elongin–B, elongin-C and cullin-2 the E3 ubiquitin-ligase complex [67-69]. Under normoxic conditions this complex ubiquitinates several proteins including the hypoxia-induced factors (HIF) for subsequent degradation by the proteosome [70, 71]. Mutations in the *VHL* gene locus impair proper folding and assembly of the complex and influence thereby the regulation of HIFs. As a consequence, HIFs are overexpressed and bind to DNA motifs called hypoxia response elements involved in angiogenesis, cell proliferation and metabolism (e.g. glucose transporter type 1 (GLUT1), erythropoietin (EPO), vascular endothelial growth factor (VEGF), carbonic anhydrase 9 (CA9) [49, 72] (Fig. 4). Furthermore, mechanistic target of

rapamycin (mTOR) levels increase via the phosphoinositide 3-kinase (PI3K)/Akt pathway in order to compensate loss of functional VHL. mTOR can propagate cell proliferation via cyclin D1 and c-Myc acitivation as well as lead to enhanced HIF activity and therefore angiogenesis [73].

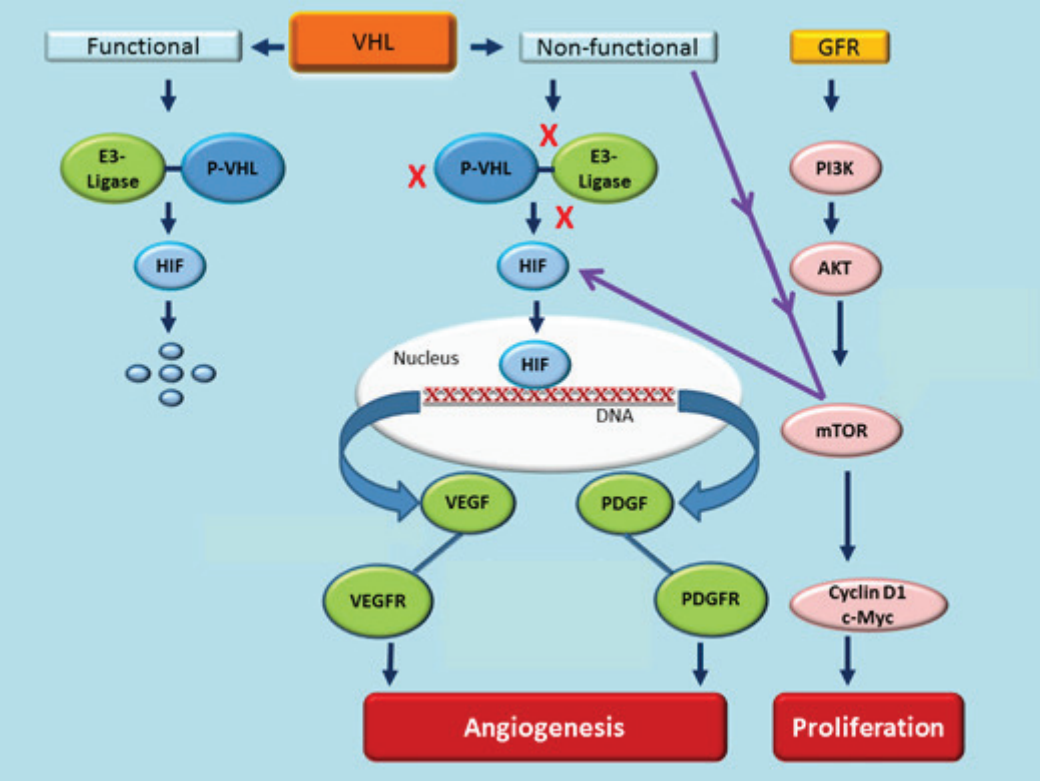


Fig. 4: VHL and mTOR signaling in clear cell RCC In normoxic conditions functional VHL forms a complex with E3-ligase and targets hypoxia-inducible factor (HIF) for degradation in the proteosome. If VHL is not functional, HIF can accumulated and initiate the transcription of hypoxia-induced genes such as vascular endothelial growth factor (VEGF) and platelet-derived growth factor (PEGF) leading to enhanced angiogenesis. Additionally, non-functional VHL increases levels of mechanistic target of rapamycin (mTOR) via the activation of the phosphoinositide 3-kinase (PI3K)/Akt pathway which propagates proliferation and angiogenesis via cyclin D1 and c-Myc as well as HIF activation. (Reprinted from Morais, C., Journal of Kidney Cancer and VHL (2014) [74] under a Creative Commons Attribution License (CC-BY 3.0): creativecommons.org/licenses/by/3.0/)

2.1.5 Treatment

2.1.5.1 Surgery

The primary treatment of clear cell RCC consists in the resection of the tumour. Thereby the tumour-bearing kidney can be removed completely in a radical nephrectomy or partially in a partial nephrectomy. The extent of the surgical intervention is based on tumour size, localisation, invasiveness and the condition of the patient.

The nephron-sparing partial nephrectomy is mainly chosen for tumours with a diameter less than 4 cm [75, 76] as well as for imperative indications such as in patients with chronic kidney

disease, bilateral tumours and other conditions affecting renal function [77, 78]. Partial nephrectomy removes only the affected tissue while preserving renal function and has therefore a positive long term effect on the patient's health [79]. That is why partial nephrectomy is still established as the gold standard for small renal masses [44].

For locally advanced tumours larger than 4 cm radical nephrectomy is the method of choice. In this surgical approach the affected kidney including the Gerota's fascia, the adrenal gland and the regional lymph nodes is completely removed [44, 78]. For non-metastatic RCC about 40-60 % of the patients can be cured with this technique [80, 81]. Patients with metastases can benefit from the surgical resection of both the primary tumour [82, 83] and individual metastases [84].

Additionally, a surgical intervention termed debulking nephrectomy can be applied for patients suffering from metastatic RCC. Following surgical nephrectomy, patients are additionally treated with cytokines such as IFN-α and IL-2 (chapter 2.1.5.2.2). The median survival is increased from 7.8 (only IFN-α treatment) to 13.6 months when IFN-α treatment is combinded with surgical nephroctomy [83, 85, 86] although there is no difference in the response rate to later treatments [83]. The survival benefit may occur due to the reduction of tumour burden, but also an immunostimulatory effect [87-89], mild renal failure and chronic low-grade metabolic acidosis are observed after nephrectomy. These effects may all contribute to reduced tumour invasiveness [90]. The removal of endogenous proangiogenic proteins (e.g. VEGF) in the course of the nephrectomy is most probably responsible for the positive effect on tumour growth and metastases and may also have an impact on the response of subsequent anti-VEGF therapy [91].

2.1.5.2 Systemic therapy

2.1.5.2.1 Chemotherapy and radiation therapy

As a standard therapy, cancer patients are usually treated with chemotherapeutics followed by radiation therapy. RCC, especially clear cell RCC, is typically resistant to this type of therapeutic approach due to the high expression of drug efflux pumps in the renal epithelium. While only 4-5 % of all patients respond to a combined chemo/radiotherapy, the response to the treatment is often short term [78, 92-95]. Therefore, these treatments have limited to no role in RCC management.

2.1.5.2.2 Cytokine therapy

The idea of immunotherapy as an effective treatment in RCC arose from the ability of some renal tumours to evoke immune responses leading to spontaneous regression [96-98]. Various

strategies to induce anti-tumour immune responses have been evaluated [99, 100], but only high dose interleukin-2 (IL-2) and interferon alpha (IFN-α) treatment resulted in improvements in overall survival. Both treatments were used as standard therapies for several decades.

IL-2 was approved in 1992 by the FDA [101]. In principle, IL-2 can be applied in three different ways a) high dose intravenously, b) continuous intravenously and c) low dose subcutaneously. The response rates ranged from 21 % over 11 % to 10 % [102], respectively, with high dose application being the most successful. Only a subset of patients (5-7 % and 8%) showed complete and partial remission [101, 103, 104], respectively. A typical patient benefiting from IL-2 treatment is young, untreated and has a disease with metastases limited to the lung [101]. But also tumours expressing carbonic anhydrase IX (G250) show a high response rate to the therapy [105, 106]. Due to the poor response rates and the high toxicities accompanying the therapy [101] , high dose IL-2 treatment is only rarely applied.

In contrast to the IL-2 monotherapy, IFN-α was never approved by the FDA as monotherapy in RCC. Clinical trials showed overall response rates of 12-15 % and complete remissions in 2-5 % of treated patients [107, 108] as well as increased survival rates [109-111]. IL-2/IFN-α combination treatment did not lead to an improved result [112, 113] whereas the combination of INF-α with medroxyprogesterone [114] or vinblastine [110] increased median survival rates. The best results were achieved in combination therapies with bevacizumab [115-118]. Currently, cytokine therapy is rarely applied in RCC due to low response rates and high toxicity. Nonetheless, treatments with cytokines have been beneficial for a selected patient cohort and resulted in durable complete remissions. The availability of predictive markers for the stratification of RCC patients according to their response to cytokine therapy would be favourable [119].

2.1.5.3 Targeted therapy

In recent years, the increased understanding of the molecular mechanisms involved in RCC resulted in the development of several therapeutic agents targeting relevant elements in the mTOR and the VEGF pathways [120]. Namely, the receptor tyrosine kinase inhibitors sorafenib (Nexavar, Bayer Healthcare), sunitinib (Sutent, Pfizer) and pazopanib (Votrient, GlaxoSmithKline) as well as the mTOR inhibitors temsirolimus (Torisel, Wyeth Pharmaceuticals) and everolimus (Afinitor, Novartis) and the anti-VEGF antibody bevacizumab (Avastin, Genentech Inc.) are applied on a regular base in clinical routine [3, 121-125]. A visual overview of the chemical structures and active sites of these reagents is given in Fig. 5. Further details are listed in Tab. 1.

Although developed with the aim of reducing the tumour size, none of these drugs achieved a durable cure. The Response Evaluation Criteria in Solid Tumours (RECIST)-defined therapy response, specifying response with 30 % as minimum for tumour size reduction, is ranging for VEGF inhibitors from about 10 to 50 % and the mTOR inhibitors from 2 to 10 % [123, 126]. Beside the very strict RECIST criteria, about 75 % of all patients profit from VEGF inhibitor drugs and 50 to 60 % from mTOR inhibitors [43]. But not only tumour size reduction is pivotal in cancer therapy. Tumour cell necrosis must not always be accompanied by size reduction and the progression-free survival, which has almost doubled compared to placebo or cytokine treatment and which is ranging from 4 to 11 months [127], is an important characteristic of successful treatments. However, therapeutic regiments which include VEGF and mTOR inhibitors exhibit severe toxicity reactions [128-130].

Up to date, no optimal treatment regimen has been identified. Treatment is therefore based on the patient's physical status and tolerability of side effects. Currently, treatment regimens are reviewed in several clinical trials. However, but due to very diverse patient cohorts and a vast variety of reagents and possible combination, it is unlikely that the RCC treatment strategy will be identified. Compared to high-dose cytokine therapy, all targeted reagents are better tolerated and seem to be less toxic. Besides being orally availability, these drugs offer patients a better quality of life and an improved clinical benefit, although initiating side effects while not achieving a durable cure. Using a combination of targeted therapy and cytokines could be an option as well. In preclinical trials antibody-cytokine conjugates were used for the treatment of advanced RCC. Thereby stable disease and no safety concerns were reported [131].

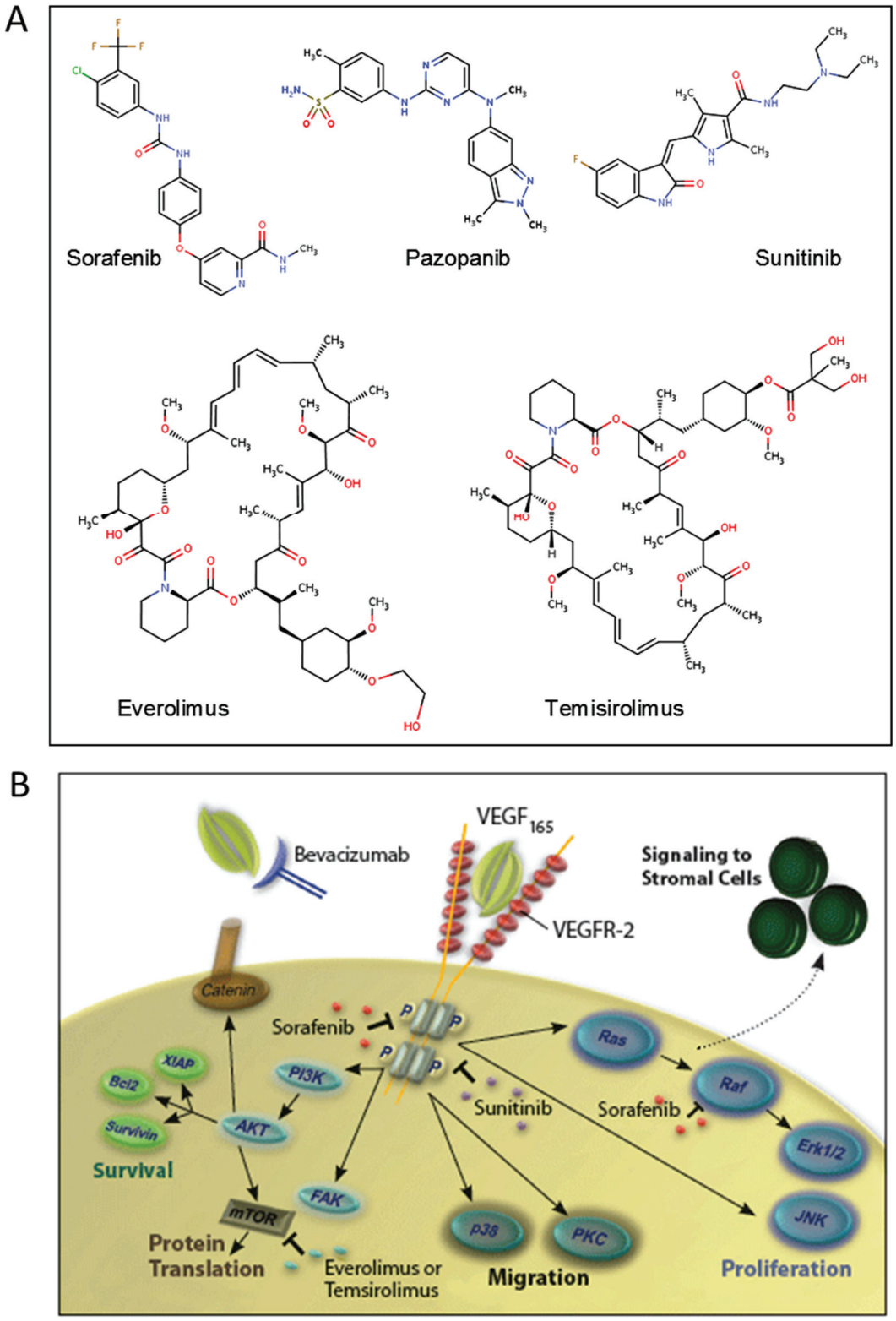


Fig. 5: Structures and molecular targets of approved drugs in RCC A: Chemical structures of VEGF and mTOR pathway inhibitors (adapted from http://www.drugbank.ca [132]). B: Point of action of approved drugs in RCC. Sorafenib, sunitinib and Pazopanib inhibit the intracellular VEGF- and PDGF-signalling while bevacizumab neutralizes soluble VEGF. The inhibitors temsirolimus and everolimus bind to mTOR influencing the PI3-kinase pathway. (Reprinted from Wu *et al., 2012* [133])

Tab. 1: Overview of approved targeted therapies in treatment of RCC. ALT: alanine aminotransferase; AST: aspartate aminotransferase; NR: not reported; ORR: objective response rate; PFS: progression-free survival. Table adapted from Yana G. Najjar & Brian I. Rini, 2012 [134]

Agent	Trade name, vendor	FDA approval	Molecular targets	ORR	PFS (months)	Complete remission	Overall survival	Major side effects	References
Sorafenib	Nexavar®, Bayer Healthcare	Dec 2005	B-Raf, c-Raf, VEGFR-1&2&3, FLT-3, PDGF-β, FGFR-1, c-Ret, c-Kit	10%	5.2	< 1%	17.8 months	Diarrhea (48 %) Rash/Desquamation (41 %) Hand-foot syndrome (33 %)	[125, 135-141]
Sunitinib	Sutent®, Pfizer	Jan 2006	PDGFR-α&β, VEGFR-1&2&3, c-Kit, FLT-3, M-CSF-1R	47%	11.0	2.9 %	26.4 months	Anemia (79 %) Lymphopenia (78 %) Neutropenia (77 %) Diarhhea (61 %) Fatigue (54 %)	[142-148]
Pazopanib	Votrient®, GlaxoSmithKline	Oct 2009	VEGFR-1&2&3, PDGFR-α&β, c-Kit, FGFR-3, Kinase EMT, FGF-1, SH2B adapter protein 3	30%	9.2	< 1 %	NR	Diarrhea (52 %) High ALT/AST levels (53 %) Hypertension (40 %)	[3, 149, 150]
Bevacimab (+IFN-α)	Avastin®, Roche Pharmaceuticals	Aug 2009	VEGF-A	31%	10.2	< 1 %	23.3 months	Pyrexia (45 %) Anorexia (36 %) Bleeding (33 %) Fatigue (33 %)	[125, 141, 151]
Temsirolimus	Torisel®, Wyeth Pharmaceuticals	May 2007	mTOR in mTORC1&2	8.6 %	3.8	NR	10.9 months	Asthenia (51 %) Rash (47 %) Anemia (45 %)	[124, 152, 153]
Everolimus	Afinitor®, Novartis	Mar 2009	mTOR in mTORC1	1.8 %	4.9	NR	14.8 months	Anemia (91 %) Hypercholesterolemia (76 %) Hypertriglyceridemia (71 %) Stomatitis (40 %)	[123, 154-156]

2.1.6 Challenges

Monitoring patients during treatment and disease progression is still a very complex task. Predictive biomarker for prognosis, therapy selection and response as well as disease progression are urgently needed for clinical practice. Both academic and pharmaceutical research is currently focusing on the development of such biomarkers.

An additional challenge in the treatment of advanced RCC are resistance mechanisms towards common chemo- and radiotherapy approaches. Also resistances against targeted therapeutics can be acquired. Usually 6 to 15 months after treatment start resistance is observed.

In many malignant diseases treatment resistances occur through mutations in a key target gene of the targeted therapeutic [157-160]. However, for VEGF inhibitors, which are applied in RCC patients, such mutations are most likely not the source of resistance since the drug targets are expressed on endothelial cells and the respective genes would have to mutate in the epithelium of each individual metastasis. Physiological changes in the tumour microenvironment allowing the reestablishment of angiogenesis during VEGFR blockade by alternative proangiogenic pathways are more likely causative [161, 162].

Increased angiogenesis can also result from hypoxia-driven upregulation of HIF-1α. Theoretically, circulating VEGF and PDGF could be increased via such a mechanism (see also Fig. 4) [161]. Additionally alternative proteins or pathways, e.g. the Tie2/Ang2 system, could be up regulated during resistance and promoting angiogenesis independently of VEGF. Resistnaces to mTOR inhibitors are likely to occur in a similar way.

Therefore, the identification of new therapeutic targets enabling a comprehensive and effective treatment of the cancer is desperately needed. Beside functional contact points also targets for simply localising potent cell toxins should be considered as potential resistance mechanism may be delayed. Thereby the application of antibody-drug conjugates is an interesting approach. By the targeting of proteins, which are upregulated in the tumour environment by monoclonal antibodies, a potent toxin can be delivered directly into the tumour. Thereby possible side effects of systemic treatments can be reduced [163].

2.1.7 Modelling disease

The availability of model systems recapitulating the human disease are indispensable for a better understanding of the underlying biological alterations and are essential for the development of novel therapeutic strategies. In the last decades a vast number of various models mimicking RCC have been established.

2.1.7.1 FCS- based cell culture

Most commonly, conventional cancer cell lines, i.e. immortalized cell lines grown in cell culture media supplied with fetal calf serum (FCS), are utilized for functional studies and preclinical screening. Due to their fast proliferation and simple manipulation, these cell lines represent a convenient tool for basic biological research. However, their use for clinical research is intensively discussed. On one hand, these cell lines are usually cultured in FCS-supplemented medium whose composition is both not clearly defined and varies from batch to batch [164]. One the other hand, such immortalized cell lines consist of a highly clonal, homogenous cell population grown for years or decades in a 2D system and are therefore unable to correctly model the heterogeneous nature and architecture of a tumour tissue [165]. Due to the long-term *in vitro* passaging of the cell lines, these cells acquired many genetic alterations changing the molecular and morphological characteristics of the parental tumour and giving the opportunity for clonal selection [166, 167]. For renal cell carcinoma several commercially available cell lines exist, e.g. UMRC2 & 3, Caki-2.Bbut only few are able to recapitulate a tumour, which is histologically similar to the patient tumour, upon transplantation in mice.

2.1.7.2 Primary cell culture

Primary cell cultures, which are generated either from enzymatically, chemically or mechanically dissociated resected patient tumour pieces or outgrown, migrating cells deriving from these pieces, are often better recapitulating the human disease [168]. Due to their origin, the cultured cells are considered to resemble the same karyotype. Another advantage is the possibility to grow the cells both in 2D and in 3D cultures. Furthermore, such cell lines can be co-cultured with other cell such as for example stromal cells. However, the culturing of primary cells requires the establishment of specific cell culture media and is therefore more laborious than FCS-based systems. The medium supplements to sustain cell viability and phenotype must be individually established [169] and regardless of the effort such systems are still only partially modelling the human disease.

2.1.7.3 Transgenic mouse models (GEMMs)

Other systems for tumour modelling are genetically engineered mouse models (GEMM). They are generated either by overexpression of an oncogene or by knockout of a tumour-suppressor gene [170] using either pronuclear injection or lentiviral transduction of embryonic stem cells.

In the course of the pronuclear injection, the additional DNA information is injected into a single cell of a developing embryo which will randomly integrate into the whole genome [171, 172]. Oliver Smithies and Mario Capecchi established the second technique where a DNA construct homologous to the gene of interest is recombined in the genomic DNA of embryonic stem cells. Successfully manipulated cells are later injected into blastocysts of the developing mice [173].

Beside the requirement of elaborate production technologies GEMMs are not able to fully represent the genetic complexity of tumour cells. Additionally, the longer telomere lengths in mice compared to humans have to be considered as the genetic instability of short human telomeres can't be examined in this setting [174]. Relying on the insertion of foreign DNA into the model organism and using a non-human model system are additional disadvantages. However, GEMMs present a good model for the examination of the disease as they allow the generation of *de novo* tumours. Also the fact that tumour cells are surrounded by stromal cells offers a lot of research opportunities including the research on different development stages of cancer. So far, several GEMMs have been established for a variety of different tumour entities. However, GEMMs for RCC [175] are not very common as the genetic inactivation of the VHL locus only results in the development of renal cysts, which do not progress to full blown tumours [176, 177]. Nevertheless, one clear cell RCC GEMM, which expresses a constitutively active mutant of HIF [178], is developing both renal cysts and renal tumours, but the model fails to recapitulate the aggressive stages in renal tumour development.

2.1.7.4 Xenograft mouse models

Currently, xenograft models from primary tumour material are considered as the most veritable tumour models. These models are generated by the transplantation of human primary tumour cells or tissue into immunocompromised mice. In former days, mainly non-obese diabetic mice with severe combined immunodeficiency (NOD/SCID mice) lacking functional T- and B-cells were used [179]. Today, mostly NOD/SCID/Il2rg$^{-/-}$ mice, which additional lack functional NK cells due to a mutated interleukin-2 γ receptor chain, are employed for transplantation as they enable a better engraftment of the human tumour cells [180]. The transplantation can be performed either orthotopic into the organ in which the tumour originally developed or ectopic (mainly subcutaneously).

The ability to grow human tissue including stromal compartments *in vivo* mimics the tumour microenvironment as close as possible. In parallel, both genetic and epigenetic modifications of the patient's tumour are preserved. Therefore, these models provide a convenient tool which can also be applied for individualised therapy approaches. Orthotopic xenograft mouse

models are useful system for drug response testing [181, 182] as the environment of the tumour cells with all hormones, growth factors and nutrients remains similar to its origin. Additionally, xenograft models often reconstitute the native metastatic behaviour observed in RCC patients. Up to date most RCC xenograft models are cell line derived [183] and the scientific community considers only few models as useful [184, 185].

2.2 Biomarker Discovery

Despite the significant advances realized during the last decade both in the molecular characterization of tumours and the evaluation of treatment options, cancer often remains an incurable disease. Therefore, the development of new therapeutic targets enabling a comprehensive and effective treatment while keeping a selection of therapy options in hand, is desperately needed. Beside functional drug targets also non-functional targets overexpressed or exclusively expressed at the site of disease available for the targeted delivery of cell toxins should be considered for therapy. For disease management not only effective drugs are needed, but also so called biomarkers for early diagnostics, for the monitoring of disease progression and to predict and monitor the therapeutic response.

According to the National Cancer Institute [186, 187] a biomarker is “a molecule detected in body fluids or tissues that is associated with special process (normal or abnormal), a condition or disease”. Therefore, biomarkers provide a great opportunity in personalized medicine. The first biomarker for cancer reported in 1848 was the light chain of immunoglobulin in the urine of 75 % of all myeloma patients [188]. In cancer research biomarkers can be proteins, metabolites, but also DNA and RNA as well as processes such as angiogenesis, apoptosis or proliferation [189]. Biomarkers can be divided in three main categories based on their application. Diagnostic biomarkers can be used for early detection of the disease. Many markers in body fluids such as blood and urine are belonging to this category. Prognostic markers form the second branch of biomarkers. These biomarkers can be used to forecast disease progression, patient survival and the risk of recurrence of the disease. This field currently comprises the majority of the clinically relevant markers. The third section of biomarkers has also a predictive nature and allows the forecasting of therapy response [190-193]. Although no fourth category officially exists, biomarkers can also serve a fourth purpose: All types of markers are associated with an abnormal status of certain molecules and if their levels are increased these biomarkers could serve as molecular target in targeted therapy. Several antibody-based therapeutics are based on this principle such as the anti-HER2 antibody trastuzumab (Herceptin, Roche) for breast cancer treatment.

2.2.1 Biomarker requirements

An ideal biomarker features several characteristics. Primarily, it needs to meet certain sensitivity and specificity requirements. Sensitivity is defined as the percentage of true positive results and specificity as the percentage of true negative results. For clinical use both parameters should be ≥0.9 [194]. Additionally, the measurement of the biomarker should be simple and cheap to perform, ideally using sample material easily obtainable in a non-invasive manner. The result of this analysis should be available quickly. In the case of a targeted therapy, it is crucial that the biomarker is a molecule highly up-regulated in the abnormal tissue and, ideally, not expressed in normal organs. For its use as drug target the biomarker has to be accessible for the drug *in vivo*, i.e. it has to be in close proximity to the vasculature in case the drug is an antibody. A biomarker, which is used as drug target, can either be utilized as an anchor point for the delivery of a payload to the site of disease or as a functional contact point, for example by blocking a certain pathway [4, 5].

2.2.2 Biomarker discovery workflow

The discovery of biomarkers is a multi-step process [195-197]. The common development pipeline for clinical markers starts with the selection of the material to analyse or the disease-representing model (e.g. mouse, cell line), on which biomarker discovery techniques can be applied [198]. Following an initial preclinical discovery phase, biomarker candidates are selected and verified using orthogonal methods such as e.g. immunohistochemistry, real time quantitative PCR (RT-qPCR) and selected reaction monitoring. Before entering the clinic or further drug development pipelines, a preclinical validation phase on retrospectively selected samples from hundreds of patients with known disease outcome is performed. Finally, the pre-validated marker candidate can enter clinical evaluation. Therefore several guidelines including quality management and reporting such as REMARK [199], STROBE-ME [200] and PROBE [201] rules are available [202].

2.2.3 Strategies in biomarker discovery

Biomarker discovery can be performed on all molecular biological levels:

- the genome using genomics and epigenomics approaches
- the transcribed RNAs in transcriptomics
- the translated proteins in proteomics
- posttranslational modified proteins in phosphoproteomics and glycomics.

Also the metabolite examination in metabonomics and lipidomics can lead to useful clinical biomarkers (chapter 2.2.3.4.2and 2.2.3.4.3). In all these fields, distinct approaches for the discovery of novel biomarkers are available and applied both in academic and pharmaceutical research labs. Novel techniques for the identification of biomarkers are regularly developed both for transcriptomics and proteomics approaches. The selection of a certain experimental technique for biomarker discovery has to be investigated beforehand with regard to advantages and disadvantages. Crucial are specificity, reliability, sensitivity and technical feasibility of the approach. Furthermore, the approach´s capability to be transferred to a high throughput screening and the cost / output ratio have to be evaluated. Finally, the sample source has to be included into any selection decision. With regard to proteomics-based approaches, the major advantage compared to transcriptomics consists in the analysis of the actual functional molecules (the protein) in the cell present at a certain condition, whereas transcriptomics relies on the precursor molecule, i.e. the mRNA. Therefore, proteomics-based approaches can also identify modified functional molecules such as alternative splicing variants and posttranslational modified proteins. Additionally, the correlation between the abundance of mRNA and the corresponding protein, a central pillar for transcriptomics analyses, has lately been described to be only minor [203-205]. In order to take advantage of the complementary nature of the above described approaches, the merging of data sets obtained with several experimental methods gets more and more common [206]. In the following chapters several approach for biomarker discovery are described.

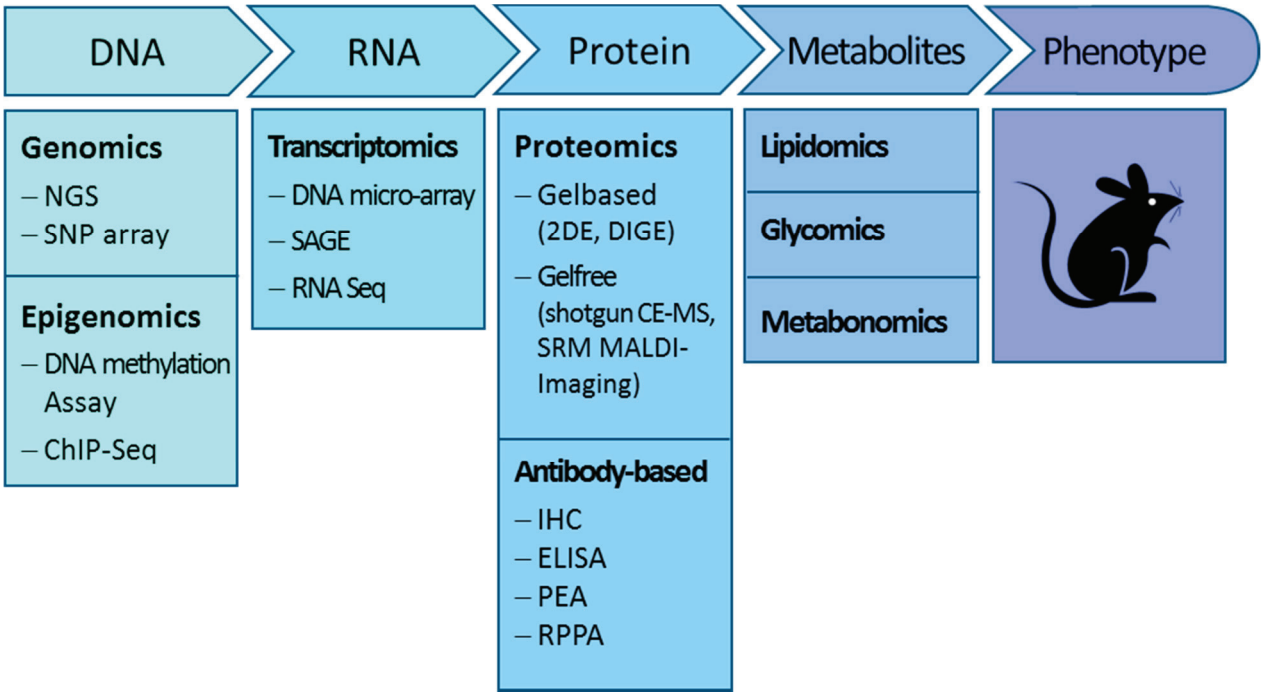


Fig. 6: Strategies in biomarker discovery All molecular biological levels can be screened for biomarkers in cancer research. Therefore, researchers can choose from a steadily growing platform of various methodologies.

2.2.3.1 Antibody-based approaches

The biomarker discovery approaches with the longest tradition are based on the screening of large antibody libraries whose targets are mainly selected by educated guessing. To date still several antibody-based approaches are applied. But despite several new multiplexing techniques like planar assays [207] and micro-bead assays [208, 209] these approaches can only be as good as their antibody applied. Both the limited specificity and the significant cross-reactivity [210] of antibodies make these techniques unreliable and difficult to reproduce [211]. Although various antibody-based techniques (e.g. ELISA) are highly sensitive, detection limits of other approaches cannot be achieved. Compared to proteomic approaches, a major drawback of antibody-based techniques is the inability to obtain detailed information about posttranslational modifications. Due to the experimental setup, only few antigens can be screened in parallel resulting in a low coverage of the proteome in one experiment. Additionally, antibodies are very expensive and the cost / output ratio is comparably low. To overcome some antibody-related issues alternative binding molecules e.g. designed ankyrin repeat proteins (DARPins) [212, 213], small organic molecules, short peptides [214] and aptamers [215] have been applied. Nevertheless, antibody-based approaches for clinical biomarker discovery are often used as their execution is relatively simple and no particular expensive equipment is needed.

2.2.3.1.1 Immunohistochemistry

A classical method to discovered tissue biomarker is immunohistochemistry (IHC). In this approach formalin fixed, paraffin embedded tissue slices are incubated with antibodies specifically binding to a target proteins such as an intracellular or membrane protein or a protein from the extracellular matrix. Enzyme linked chromogenic substrates are most commonly used for the detection of antibody binding and the respective signals are evaluated during microscopic examination by a trained person. The screening and validation process is very subjective, time consuming and can only be performed in a semi-quantitative manner. The Ventana Medical Systems for the evaluation of HER2-positive breast cancer [216], which has been approved by the FDA in 2011, has improved the automated analysis of histological pictures but did not solve it. Additionally, IHC screenings feature a bad comparability, a low sensitivity and display variable false positive and false negatives rates [217]. To control for these parameters, positive and negative control tissues should always be included in the analysis as shown in the oncology field [218, 219]. Additionally, tissue micro arrays (TMAs) consisting of several tissues on one slide are available for higher sample throughput and better comparability. In the Human Protein Atlas (HPA) a vast amount of healthy and diseased tissues

are stained with various antibodies trying to map and characterize a representative protein for each protein-coding human gene [220, 221]. IHC is an old-fashioned method widely accepted and still extensively used in the clinics. However, for biomarker discovery, it is replaced by more sensitive and quantitative methods with a higher sample processing throughput.

2.2.3.1.2 Enzyme-linked immunosorbent assay (ELISA)

Enzyme-linked immunosorbent assay (ELISA) represents the gold standard in a clinical setting for protein measurements. Several variations of the core technique are employed in the field [222]. However, the sandwich ELISA technique is mostly applied for biomarker discovery Fig. 7. Lysates derived from tissue of interest are incubated in multi-well plates pre-coated with antibodies for specific antigens. In a second step, the bound antigens are detected by a second antigen-specific antibody recognizing a different epitope. The second antibody can either be directly coupled to an enzyme or an additional anti-second antibody enzyme-linked antibody is used. Finally, a substrate is added and converted by the enzyme coupled to the detection antibody resulting usually in a detectable colour change. Typically, a sensitivity down to 10 fM can be achieved [223] making this technique attractive despite the inherent limitations of antibodies described above. Novel immunoassay technologies try to strengthen the advantages and decrease the limitations. The Erenna immunoassay system [224] uses single photon fluorescence detection and paramagnetic microparticles to increase sensitivity and precision up to three orders of magnitude [223, 225]. The digital ELISA system from Quanterix [226-229], which is based on the detection of single enzyme-linked immuno-complexes on beads, is 1000-fold more sensitive than a conventional ELISA [229, 230].

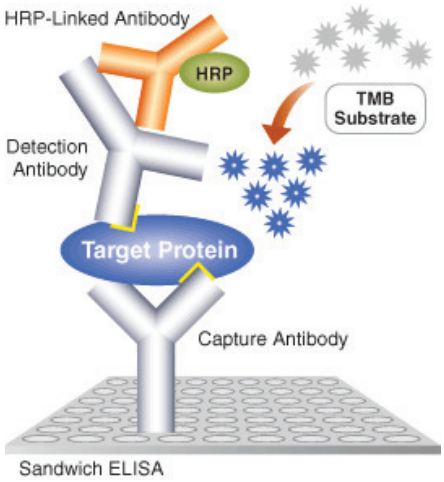


Fig. 7: Principle of a sandwich ELISA Antigen specific antibodies are immobilized on a surface and incubated with tissue lysates. Bound protein of interest is afterwards detect with an additional antibody which itself is detected by an enzyme-linked antibody. The visualisation of the protein interaction is usually achieved by a

colour change of the subsequently added substrate. Reprinted with permission from Cell Signalling Technology [231].

2.2.3.1.3 Proximity extension assay (PEA)

The antibody-based proximity extension assay (PEA, Olink Bio science) [232] is a very specific assay covering up to five orders of magnitude dynamic range and offers both high sensitivity and exact quantification [223]. Target specific antibody pairs linked to DNA strands bind to the protein of interest and a reverse transcription PCR amplicon is generated by DNA polymerase upon binding. Quantification is performed by microfluidic quantitative PCR (qPCR). Up to date a panel of 92 proteins is commercially available from Olink Biosciences.

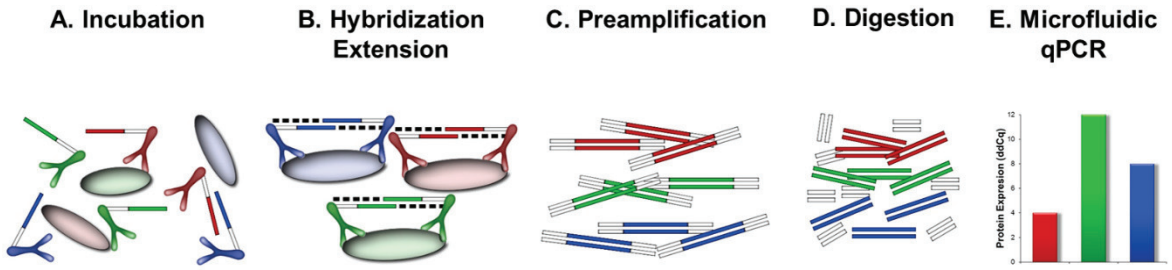


Fig. 8: Proximity extension assay (PEA) workflow Target-specific antibody pairs bind to the antigen of interest. Antibody linked to DNA strands hybridize, extend and are subsequently pre-amplified and digested. Quantification is performed by microfluidic quantitative PCR (qPCR). Reprinted from Assarsson & Lundberg et al. 2014 [233]

2.2.3.1.4 Reverse-phase protein array (RPPA)

A useful and efficient tool for tissue protein quantification and characterisation is the reverse-phase protein array (RPPA) [234-236]. Slides are prepared in dot blot manner with single spots of tissue lysates and the protein of interest is quantified by specific antibodies. Quantification is performed via immunochemical detection with enzyme linked antibodies comparable to the ELISA principle. As RPPA is a high throughput technique allowing protein quantification with a low sample input it is advantageous compared to traditional methods such as immunohistochemistry (IHC) [237, 238]. The technique can be applied in combination with both formalin fixed, paraffin embedded (FFPE) tissues and freshly frozen tissues [239].

2.2.3.2 Transcriptomics approaches

Transcriptomics is a discipline in genetics examining all kinds of RNAs including non-coding RNAs present in cell. While most cells share a similar exome, significant variations in the transcribed version of the genome can be observed. Differences in transcription are leading to different phenotypes. Therefore, the transcriptome of a cell population is a possible source

for the discovery of novel biomarkers in cancer research. However, it has to be considered that a cell not only regulates its protein expression based on mRNA expression and therefore mRNA levels must not obligatory correlate with protein expression [203]. In transcriptomics two general approaches can be performed. Hybridization-based approaches, mainly microarrays, involve the binding of fluorescently labelled cDNA to a predefined set of genes and sequence-based approaches directly determine the cDNA sequence. Thereby the traditional Sanger sequencing [240, 241] is more and more replaced by next generation sequencing (NGS) technologies. However, Sanger-based approaches are still utilized, for example in serial analysis of gene expression (SAGE) [242], cap analysis of gene expression (CAGE) [243-245] and massively parallel signature sequencing (MPSS) [246, 247]. Nevertheless, the recently developed RNA-seq [248] technology, which is based on NGS, is a very successful competitor to the old-fashioned Sanger-based sequencing approach as RNA-seq performs the alignment process in a highly quantitative manner. With respect to their basic goal hybridization- and sequencing-based transcriptomics approaches are very similar. While in sequencing-based methods RNA units are sequenced, transcribed genes are identified by hybridization of mRNA to predefined probes in hybridization-based microarrays. As a consequence no novel genes or variants can be discovered using microarrays but with sequencing-based technologies. Another advantage of the sequencing approaches is their improved quantification capability due to the direct counting of transcript numbers compared to spot intensity based measurements. Nevertheless, microarrays are the method of choice in biomarker discovery project as, especially for large-scale analyses, the high costs associated with sequencing-based methods are still decisive.

2.2.3.2.1 DNA microarray

The most common genomic profiling approach for biomarker discovery is certainly DNA microarray technology. The method originated from Southern blot analysis [249] and was first applied by Augenlicht *et al.* in 1982 screening *E. coli* for 378 sequences [250]. The miniaturized version was then introduced in 1995 [251] and only two year later the first complete eukaryotic genome (*S. cerevisiae*) on a microarray was published [252]. the hybridization between gene-specific DNA probes fixed on a solid support and cDNA sequences transcribed from tissue-extracted RNA represents the core principle of DNA microarrays [253]. The array chip displays on its surface 10 000s of spots containing DNA probes in the picomolar range coated via surface engineering techniques with covalent bonds to the chemical matrix (e.g. epoxy-silane, amino-silane, lysine, polyacrylamide). The chip material is usually glass or silica like for the Affymetrix chip [254]. The Illumina system in contrast is utilizing coated beads as solid surface [255]. The workflow applied for a DNA microarray is illustrated in Fig. 9. Briefly,

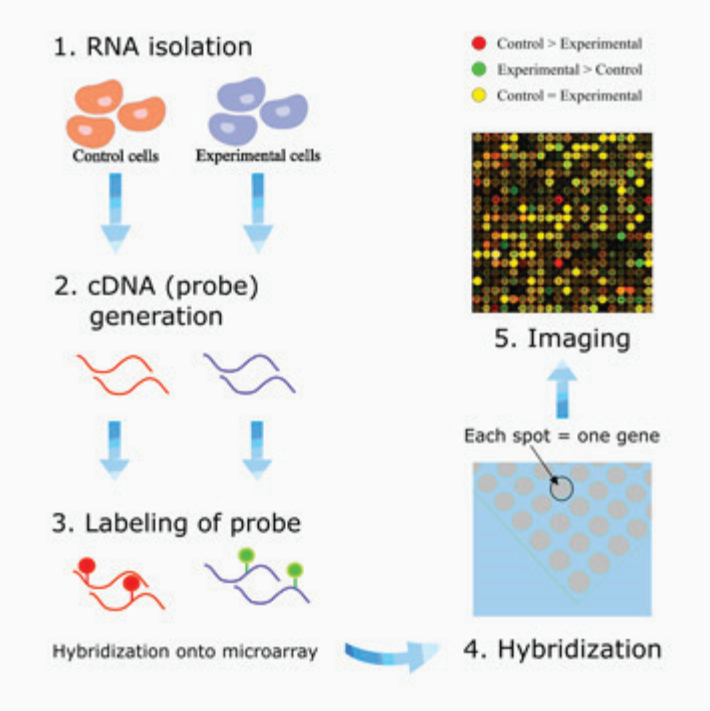


Fig. 9: DNA microarray workflow RNA is extracted from tissues of interest and reverse transcribed into cDNA. cDNA is labelled with Cy-dye (e.g. Cy3, Cy5) and hybridized on an array chip coated with DNA oligonucleotides representing unique sequences for various genes. Levels of hybridization are measured based on fluorescence signal intensities on the array chip. Subsequent bioinformatics and statistical analysis of the measured signals corresponding to RNA expression in the tissue of interest is performed. Reprinted from Pavelic et al. 2005 [264].

RNA is extracted from tissues of interest and converted to cDNA by reverse transcriptase. cDNA sequences are labelled for later detection. To this, mostly fluorophoric Cy-dyes (e.g. Cy3, Cy5) but also silver- and chemoluminsent labels are used. Subsequently to the labelling cDNA units are hybridized on the pre-coated array chip with gene-specific probes of 1000s of individual genes. Resulting signals are measured and analysed [253]. More hybridized clones yield in a higher intensity signal thereby pointing to a higher expression of the corresponding gene in the analysed tissue. Gene expression patterns obtained from DNA microarrays are highly accepted in the scientific community as they showed to be helpful in the clinical prediction of tumour behaviour [256], guided the tumour sub-classification hypothesis for example in leukaemia and breast cancer [257-259] and lead to the discovery of new diagnostic markers [260]. Besides being a fast technique with a clinically relevant output, DNA microarray technology has certain limitations. The amount of data resulting from chips with several 1000s of gene probes is tremendous and needs proper computational and statistical analysis. Furthermore, the reproducibility of the array is limited. In a meta-analysis [261] of seven microarray experiments five could not be reproduced [262]. Additionally, chips containing large numbers of probes are very expensive. That is why, often only a limited number of probes for the same gene are analysed resulting in a reduced statistical power of the experiment. As indicated above, microarrays can only analyse the known, common alleles. Usually between 500 000 and 2 million SNPs are displayed on a chip compared to 10 million

SNPs in the human genome [263]. Despite all these limitations, microarrays are useful and successful tools in biomarker discovery.

2.2.3.2.2 Serial analysis of gene expression (SAGE)

The serial analysis of gene expression (SAGE) technology was developed by Victor Velculescu in 1995 [265]. This method is based on the sequencing of mRNA transcript fragments isolated from tissues of interest. Three basic principles underlie this mRNA-based method. First, only a short sequence tag of about 10-14 base pairs is needed to identify a particular transcript, provided that the obtained tag is from a unique position of the transcript. Second, transcript fragments can be stringed together forming long serial molecules which can be cloned and sequenced. Third, the quantification of the mRNA fragments corresponding to gene expression levels is performed by counting how often a particular tag is observed. To achieve improved assignment of tags to their corresponding mRNA LongSAGE [266] followed by SuperSAGE [267] have been developed. These techniques allow analysing tag lengths of up to 26 base pairs.

2.2.3.2.3 RNA-Seq

RNA-Seq, also known as Whole Transcriptome Shotgun Sequencing (WTSS) [248], is a novel sequencing method of the entire transcribed RNA pool based on the developments in next generation sequencing technologies. It provides a deep and quantitative snapshot of all RNAs present in a certain tissue at a specific time point [268]. Compared to SAGE and other sequencing based methods it is achieving increased DNA sequence coverage and offers the opportunity to detect low frequent RNAs, alternative splice variants, posttranscriptional changes, gene fusions, mutations, single nucleotide polymorphisms (SNPs) and gene expression [269]. The latter characteristic is often addressed in cancer research [270]. Additionally, the analysis of various RNA types such as miRNA and tRNA as well as ribosomal profiling is feasible [271]. In principle every sequencing technology of the fast growing next generation sequencing (NGS) field can be used. So far, mainly the Illumina IG [272-274], Applied Biosystems SOLiD [275] and Roche 454 Life Science [203, 276] systems were applied. The general workflow consists of RNA extraction from the tissue of interest followed by reverse transcription to generate a cDNA library containing DNA fragments with end terminal adaptors (Fig. 10). Finally, this library is sequenced by the NGS method of choice [268]. Sequencing can be performed either single ended (from only one side) or pair ended (from both cDNA fragment sides). The sequenced fragments with a length between 30 and 400 base pairs are than either aligned to a reference genome using specific software (e.g. Bowtie [277],

TopHat [278] or FANse [279]) or are *de novo* translated into gene sequences. . Analysis requires significant computational power applying and dedicated software, for example Oases [280] or Trinity [281]. So far, RNA-seq has been used in several studies including mouse, human and cancer tissues [276, 282-286]. RNA-seq data are collected on two big platforms. Experimental data derived from dozens of cell lines [287] and thousands of primary tumor samples [288] are publicly available at the Encyclopedia of the regulatory elements (ENCODE) and The Cancer Genome Atlas (TCGA). In summary, RNA-seq is a quantitative high throughput technology for the characterization of the transcriptome of a tissue. It features many advantages compared to the conventional methodologies discussed above. Furthermore, its high dynamic range (theoretically no limit of quantification) needs be mentioned. In a recent study 9000-fold regulation of RNA fragments could be monitored [276] compared to a microarray with maximal 100-fold changes. However, the significant computational / bioinformatics effort as well as a reasonable cost to sequence coverage ratio needs to be considered.

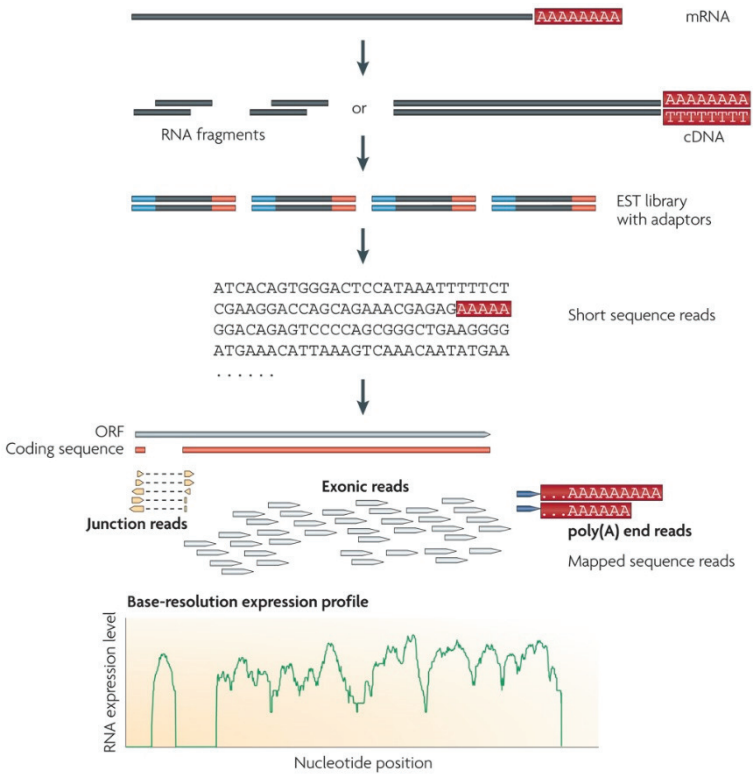


Fig. 10: General Workflow in RNA-seq RNA is extracted from tissue of interest and reverse transcribed in a cDNA fragments library and sequencing adaptors are added to each fragment. A short sequence is analysed by next generation sequencing technologies and resulting sequences are aligned to a reference genome or de novo translated and gene expression is calculated. Reprinted from Wang et al. 2009 [289].

2.2.3.3 Proteomics approaches

In biomarker discovery the application of proteomics approaches is constantly growing. Specifically, methodologies applying mass spectrometry provide a high valuable tool [187]. Various techniques have been described so fare. Due to this reason and as the focus of this thesis is based on proteomics a distinct section (chapter 2.2.3.3) following this chapter has been dedicated to this issue.

2.2.3.4 Other -omics approaches

Although transcriptomics and proteomics approaches are the most commonly applied methodologies in biomarker discovery several other, often proteomics related -omics approaches are published. Amongst these techniques, genomic [290], glycomic [291], metabonomic [292] and lipidomic [293] approaches are the most promising ones. But also peptidomics, the qualitative and quantitative multiplex analysis of endogenous peptides in a biological sample [294], immunomics, the analysis of changes in immunological responses in an organism [295], and secretomics, the analysis of the secretome [296], are applied. Additionally, various techniques for the examination of the phosphorylation status of a protein (phosphoproteomics) are available [297, 298] and used in biomarker discovery [299, 300]. Combinations of different -omics strategies are used to obtain a complete picture of a biological system and to identify the best point of attack.

2.2.3.4.1 Genomics and epigenomics

Cancer is primarily caused by genome alterations including DNA sequence changes, gene copy number aberrations, chromosomal rearrangements as well as epigenetic modifications [301]. To address this issue Sanger sequencing of the genome was applied [302] for several years as the method of choice. Due to tremendous time and financial efforts, Sanger sequencing has been largely replaced by next generation sequencing (NGS) techniques such as the Roche454 system, the ABI SOLiD system, the Illumina Solexa system and the Helicos system [303]. These methods sequence the genome 200 times faster than the conventional Sanger sequencing for a significantly reduced price. NGS techniques are based on the pyrosequencing of up to thousands of single stranded DNAs which are immobilized on a solid surface such as a glass slide or beads [304, 305]. Genomics also includes the analysis of single nucleotide polymorphisms (SNPs), which is feasable by NGS techniques [306, 307]. SNPs are the most common form of variations in the genome [308]. These variations can be addressed in a SNP array, which genotypes multiple SNPs for the presence of allelic imbalances, loss of heterozygosity and DNA copy number alterations [309, 310] in cancer [311].

A fast growing side field of genomics is the investigation of epigenetic modifications like DNA methylation, histone modification and chromatin structures [312, 313] involved in gene regulation. It involves techniques like histone modification assays, chromatin immunoprecipitation sequencing (ChIP-Chip and ChIP-Seq) as well as several DNA methylation assays [314]. In combination with genomic data valuable hints for novel biomarkers in cancer can be obtained.

2.2.3.4.2 Metabonomics

Metabonomics, not to be mistaken with metabolomics, investigates the fingerprints of biochemical perturbations caused by diseases, drugs and toxins [315]. Metabolomics in contrast focuses on the comprehensive analysis of all metabolites in a given biological system by measuring concentrations of variable metabolites. For biomarker discovery metabonomics is of greater importance. Several methods, which have to be balanced according to advantages and disadvantages, are applied for the analysis of metabolic signatures depending on the issue of interest [316, 317]. Nuclear magnetic resonance (NMR) is a non-destructive method applicable to intact cells and tissues. It can elucidate a molecular structure with low analytical variation [318]. The drawbacks of NMR are low resolution and sensitivity as well as difficulties in the metabolite quantification. In contrast, mass spectrometry (MS) based technologies are more sensitive but require the disintegration of the sample. Depending on the nature of the compound class of interest different separation techniques like gas chromatography-MS (GC-MS), liquid chromatography-MS (LC-MS), capillary electrophoresis-MS (CE-MS) and Fourier transform ion cyclotron resonance MS (FTICR-MS) have to be applied.

2.2.3.4.3 Lipidomics

Lipidomics is the study about the lipids and their in interactions in a biological system and was first described in 2003 [293]. This relatively new, fast growing field [319] gained influence since lipids not only play a central role as structural components of a cell but also have a relevant physiological function in the control of cell signalling during cell growth, differentiation and cell death [320]. Abnormalities in lipid composition can lead to several diseases [321] including cancer [322]. Lipids involved in cancer maintenance are fatty acids, triglycerides, cholesterol, phosphoglycerides, sphingomyelin, glycosphingolipid, ceramide and sphingosine phosphate [323]. To study these lipids chromatographic, biochemical, nuclear magnetic resonance (NMR) and mass spectrometry based techniques are applied. As NMR provides a low sensitivity enabling only the analysis of very high abundant lipids and biochemical approaches are experimentally very challenging as they require lipid-specific antibodies [293], mass

spectrometric approaches are commonly the methods of choice. MS-based lipidomics involves lipid extraction by organic solvents (chloroform, methanol and water [324, 325]) followed by a separation technique (thin-layer chromatography, solid-phase extraction chromatography [326, 327] and in recent days also high performance liquid chromatography (HPLC) [328, 329]). The subsequent analysis has classically been performed by gas chromatography mass spectrometry (GC-MS) [330-332], but faster and more accurate mass spectrometric analyses applying the soft ionization methods matrix-assisted laser desorption ionization (MALDI) [333, 334] and electrospray ionization (ESI) [335] are increasingly used.

2.2.3.4.4 Glycomics

Glycomics is one of the most promising approaches for biomarkers discovery. It is a branch of proteomics dealing with the changes of glycosylation patterns on proteins. About half of all mammalian proteins are supposed to be posttranslational glycosylated [336]. Changes in glycosylation patterns have fundamental consequences [337-339] as glycosylation plays an important role in development, cell division, inflammation and tumour immunology [340-345]. For the analysis of variable glycosylation patterns standard proteomic methodologies can be applied [346-348], although glycomics has to deal with additional challenges such as a high dynamic range of glycosylation modifications, instability of glycans and the isolation of modified peptides. Therefore several techniques have been developed in the last years [349-351]. Purification of glycosylated proteins can occur with lectin affinity chromatography [352] or antibody assisted lectin profiling (ALP) [353]. Furthermore, chemical approaches [354] like immobilized hydrazide [355], boronic diester [356] and huisgen cycloaddition [357] can be applied.

2.3 PROTEOMIC APPROACHES FOR BIOMARKER DISCOVERY

Parts of this chapter marked with “ ” have been excerpted from the following review [358],

Benk AS, Roesli C
Label-free quantification using MALDI mass spectrometry: considerations and perspectives
Anal Bioanal Chem. 2012 Sep; 404(4):1039-56,

by permission of Springer and Analytical and Bioanalytical Chemistry, provided by the Copyright Clearance Center.

“In recent years, profound knowledge of the cell’s total protein composition has become increasingly important not only in basic research but also for the development of diagnostics and therapeutics [359]. In particular, interest in comparing healthy versus diseased cells/tissues or treated versus untreated samples has driven the development of novel technologies [360-362]. […] In recent years, mass spectrometry based techniques often became the methods of choice to cope with these requests [363-366].” In a typical bottom-up proteomic workflow the proteins of interest are digested with a protease (e.g. trypsin) into peptides which are subsequently analysed on a mass spectrometer and the readouts are processed by bioinformatics procedures.

In the mass spectrometer the molecules of interest are analysed as gas-phase ions. A mass spectrometer is therefore composed of a) an ion source, b) a mass analyser and c) an ion detector. First the analyte molecules are ionized and transferred to the gas phase in the ion source. In the following mass analyser the ions are separated according to their mass-to-charge (m/z) ratio. Finally, the ionized molecules are detected on the ion detector and their signal is amplified and recorded. Different ionization methods and mass analysers are introduced in the following chapters (2.3.1and 2.3.2). For peptide sequence identification the peptides need to be fragmented and resulting fragments analysed by tandem mass spectrometry, which is outlined in chapter 2.3.3. Afterwards (chapter 2.3.4and 2.3.5), the two fundamentally different workflow gel-based and gel-free proteomics as well as quantification approaches will be elucidated.

2.3.1 Ionization methods

“Prerequisites for the analysis of any molecule by a mass spectrometer are its transfer to the gas phase and its ionization. Although several ionization techniques have been developed during the last century [367], only two are applicable to the ionization of biological macromolecules such as peptides and proteins: electrospray ionization (ESI) [368] and matrix-assisted laser desorption/ ionization (MALDI) [369, 370]. Both techniques were developed in the late 1980s and their inventors received the Nobel Prize for chemistry in 2002. MALDI and ESI differ substantially in their basic concept (Tab. 2) and it is, therefore, not surprising that the resulting ions are to some extent complementary [371, 372].”

In MALDI the sample is mixed with matrix (e.g. α-Cyano-4-hydroxycinnamic acid (CHCA)) and crystallized on a solid support, some of the key features of this technique (Fig. 11). By applying a laser pulse to the sample-matrix mixture, the matrix absorbs energy from the laser and converts it into excitation energy. Thereby, a small volume of sample is desorbed together with matrix leading to matrix neutrals (M), matrix ions $(MH)^+$ / $(M-H)^-$ and sample naturals (X).

The exact mode of ionization of biomolecules is still under debate and several models have been proposed [373]. Afterwards the activated matrix ionizes the sample molecules by photoexcitation followed by proton transfer ($MH^+ + X \rightarrow M + XH^+$). Upon a short pulse of high voltage applied to the sample plate and the source grids, the ionized sample is inserted into the mass spectrometer. [374].

In ESI a liquid containing the analytes and a volatile organic compound is dispersed into a fine aerosol sprayed directly into the mass spectrometer. The droplets formed thereby decrease in size until they reach the Rayleigh limit and a process termed Coulomb fission takes place. Thereby the original droplet explode resulting in many smaller, more stable droplets. This cycle continues until all the solvent has evaporated and naked charged analytes are left over [375]. The ion evaporation model [376] and the charge residue model [377] are two hypotheses describing this processes.

“Beside conceptual differences in analyte ionization the two techniques differ in their ability to be integrated into proteomic workflows. Although mass spectrometers with an ESI source can be coupled directly (i.e. online) to liquid chromatography, samples must be mixed with matrix and spotted by a robot on a target plate before their analysis on an MS with a MALDI ion source (offline coupling). For this reason, MALDI enables the uncoupling of MS^1 and MS^2 measurements and the re-analysis of samples. Furthermore, MALDI offers: 1. a wider dynamic range; 2. greater sensitivity; and 3. more stable ionization [378]. Nowadays, most mass spectrometric measurements for proteomic applications are performed on instruments with an electrospray ionization source. However, MALDI–MS also has other features, so far rarely used, which, in principle, enable increased depth of analysis.”

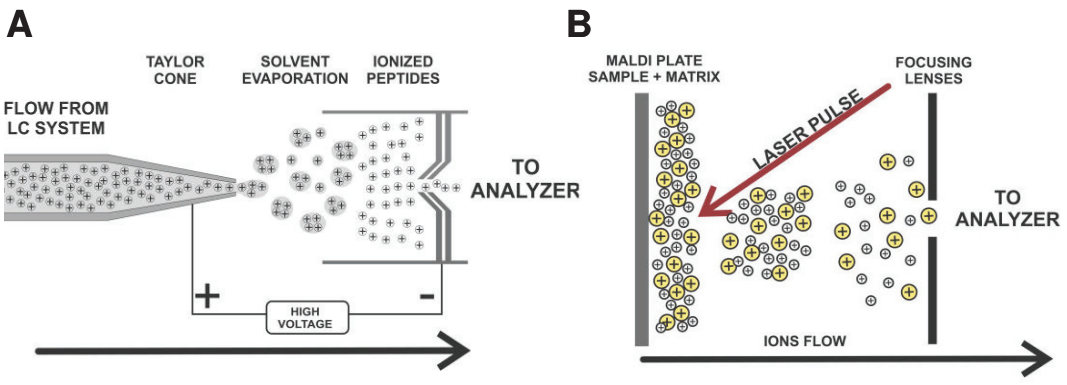


Fig. 11: Soft ionization methods A Electrospray ionization (ESI). A liquid containing the analytes is directly sprayed into the mass analyser. By evaporation the droplets decrease in size and ionize in several cycles mediated by the Coulomb fission process. B Matrix-assisted laser desorption/ionization (MALDI). The sample (yellow) co-crystallised with matrix (grey) is activated by a laser pulse and positively charged ions are accelerated by an electric field and thereby transferred into the mass analyser. Adapted from Demartini, 2013 [379, available from www.intechopen.com]

Tab. 2: Comparison of MALDI and ESI as sources for the ionization of peptides

	MALDI	ESI
Liquid chromatography	offline system	online system
Method for the generation of ions	laser shot	electro spray
Sample source consistence	solid	liquid
Peptide charge (700 to 4000 Da)	+1	+1 to +4
Separation	discontinuous (single fractions spotted by a robot)	continuous
Upper mass limit	≥2.000.000 Da	≥70.000 Da
Re-analysis of the same sample	possible	impossible
Sensitivity	high amol to low pmol	fmol to pmol
Salts	tolerance in mmolar concentrations	reduces sensitivity
Considerations	matrix background for molecules below 700 Da possibility of photo degradation possibility of sample degradation by acidic matrix	sample to sample carry-over poor analysis of complex mixtures

2.3.2 Mass analyser

In a mass analyser, the molecules of interest are separated according to their mass-to-charge (m/z) ratio [380]. In proteomics research four different basic types, which differ in size, price, resolution, mass range, and the ability to perform tandem mass spectrometry experiments [381, 382], are usually utilized: time-of-flight (TOF), quadrupole, ion trap and orbitrap analysers. Furthermore, combinations of these analysers such as quadrupole-TOF have been

developed. To avoid collisions with neutral gases the analysis is performed under high vacuum [383].

2.3.2.1 Time-of-flight (TOF)

In a time-of-flight (TOF) analyser the ions are accelerated in an electrostatic field and separated according to their m/z ratios in a field-free flight tube. As theoretically all ions leave the accelerator with the same kinetic energy their time of flight through the field-free flight tube only depends on their m/z ratio. Ions with a smaller m/z ratio will fly faster than heavier ones. TOF analysers exist in linear and reflector form. In linear mode, the ions fly through the tube directly to the detector, whereas in reflector mode the ions fly first towards an ion mirror (the so call reflector), where they get reflected. Beside prolonging the flight path, the reflector focuses ions with the same m/z values resulting in higher resolution [384]. Theoretically, the mass range of TOF analysers is unlimited [385]. However, due to limited control of the kinetic energy spread and the increased spatial distribution for heavy masses in the acceleration region, the mass range is limited [386].

2.3.2.2 Quadrupole

Quadrupole analysers are the most common analysers applied in mass spectrometry. A quadrupole can either be used as mass filter or for ion fragmentation (e.g. in selected reaction monitoring (SRM)). It consists of four roods, in which each opposing rod pair is electrically connected opposing. Upon switching of the voltages with a certain frequency, a two dimensional quadrupole field in the x-y-plane is established. Thereby, ions are pulsed in sinuous lines through the quadrupole due to the alternating attraction to the roods. Depending on the radio frequency and the voltages applied on the two rood pairs [387] only ions whose resonance fit a given ratio of voltages get a stable trajectory through the analyser and reach the detector. In order to record an MS spectrum, a range of voltage/frequency pairs is screened.

2.3.2.3 Ion trap

The basic principle of the ion trap is very similar to the one of the quadrupole described above. It is confined radially by a two dimensional radiofrequency field established between three hyperbolic electrodes and axially by so called stopping potentials. Thereby, ions can be trapped and stored in the electric field [380]. Depending on their m/z ratio ions can be ejected to the detector by changing the applied voltages, which results in a destabilization of a certain

m/z fraction in the ion trap [388]. By gradually changing the voltages of the electrodes a given m/z range can be scanned. The linear ion trap is a common variant of the conventional ion trap, which consists of four roods and which offers an improved trapping function, sensitivity and dynamic range [389]. The advantages of ion traps are their ability to tolerate higher pressures and their requirement for lower voltages. Due to the possibility of ion trapping, their sensitivity is increased [381]. Drawbacks are the required high precision in manufacturing and the relatively low resolution [390].

2.3.2.4 Orbitrap

The orbitrap is a mass analyser which is based on an orbital oscillating field [391]. Thereby, ions are trapped in an electrostatic field and circulate radially around a central spindle electrode comparable to satellite. Additionally to their rotary movements, trapped ions move along the spindle-shaped electrode forward and backward. Both oscillation and circulation frequency are dependent on the m/z ratio of the trapped ions. Using the Fourier transformation, mass-to-charge ratios for ions trapped in the orbitrap can be calculated from oscillating frequencies [392]. Therefore, the orbitrap is a very accurate, robust and high resolution mass analyser [393].

2.3.3 Fragmentation

To identify the sequence of a peptide of interest tandem mass spectrometry is applied. Thereby, a precursor peptide is selectively isolated and the peptide´s backbone is fragmented by an inert gas into smaller fragments. Subsequently, the m/z ratios of these fragments are analysed in the instrument. Various search engines such as MASCOT [394], X!Tandem [395], SEQUEST [396] or ProteinPilot (AB Sciex; incorporated Paragon algorithm [396]) read-out the fragment spectra and map these to *in silico*-generated peptide fragmentation spectra. Depending on the fragmentation method applied different peptide backbone fragments can be generated (Fig. 12) resulting in different ion types preferentially generated. The most commonly applied collision-induced fragmentation (CID) [396] produces mainly *y*-ions and *b*-ions [397] whereas the preferential occurrence of *c*-ions and *z*-ions is observed for electron capture dissociation [398] and electron transfer dissociation [399].

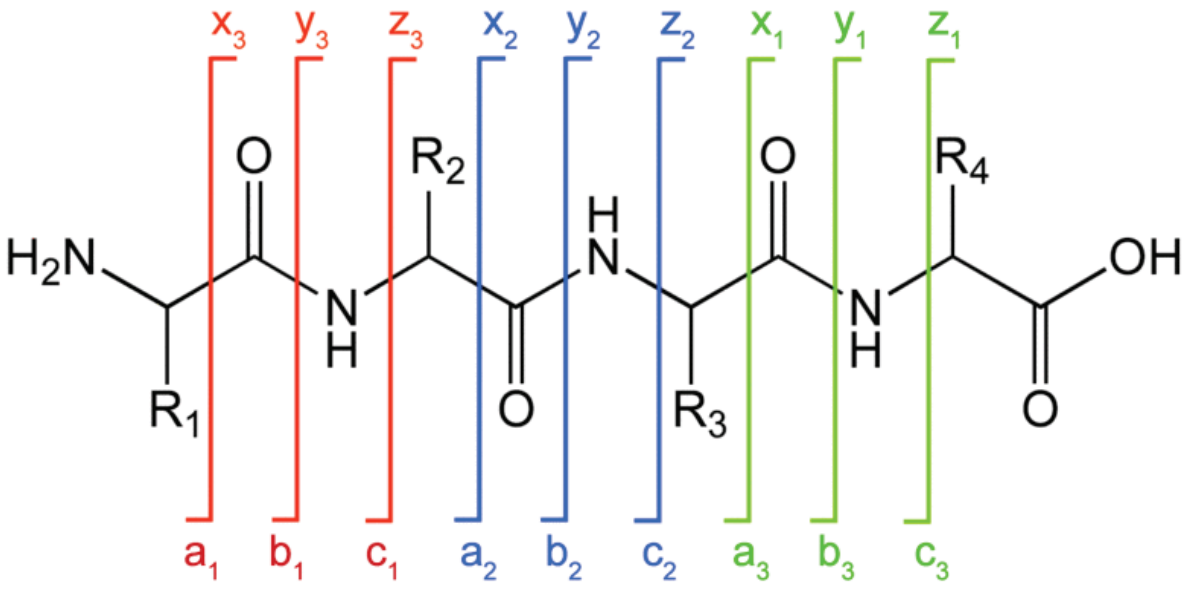


Fig. 12: Peptide fragmentation Nomenclature accordingly to Roepstorff & Fohlman [400]. Reprinted from [401] under a Creative Commons Attribution License (CC-BY 3.0): creativecommons.org/licenses/by/3.0/).

2.3.4 Gel-based proteomics

The first proteomics experiments could be conducted in 1967 following the development of the sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) technique by Shapiro and colleagues [402]. This technique set the corner stone for the area of gel-based proteomics approaches. The common workflow in gel-based proteomics consists in the separation of a protein mixture of interest in an electric field. Protein quantification is carried out optically following the staining of the gel with an appropriate reagent and is based on the spot intensities and the assumption that every spot represents one single protein. Subsequently, differentially regulated proteins are excised from the gel, digested with trypsin and MS-based protein identification is performed. In 1975, O'Farrell introduced the two-dimensional polyacrylamide gel electrophoresis approach (2D-PAGE), which replaced the one dimensional SDS-PAGE approach [403]. 2D-PAGE combines the separation of proteins by SDS-PAGE with a separation according to their charge in an immobilized pH gradient by isoelectric focusing (IEF). Unlu and colleagues developed in 1997 the two dimensional differential gel electrophoresis (2D-DIGE) as an improved method in gel-based proteomics [404]. In 2D-DIGE proteins are labelled prior to 2D-PAGE separation with different fluorescent cyanine dyes (e.g. Cy3, Cy5). Dye-labelled proteins feature the same charge and molecular mass, but differ in their fluorescent properties. Therefore, differently labelled protein mixtures can be separated on the same gel, reducing both the number of gels required for an experiment and the gel-to-gel variability [405]. Although gel-based techniques have been the gold standard in quantitative proteomics for many years, the "one spot = one protein" assumption did not always prove true leading to considerable errors in protein quantification. Further restrictions of this technology are the limited reproducibility between gels, the difficulties in handling and

detection of proteins with extreme pI values and/or molecular weights as well as the costs in labour and time [406]. Depending on the characteristics of the staining dye, the dynamic range for protein quantification using gel based methods can be varied. Using the classical silver staining a dynamic range of two orders of magnitude can be achieved [406]. With newly developed stainings such as the family of fluorescent SYPRO stains the dynamic range can be extended to about three orders of magnitude [407]. Nevertheless, after being the workhorse of proteomics for almost a quarter of century [408], gel-based proteomics can no longer keep up with gel-free proteomic approaches, which are discussed in the next chapter.

2.3.5 Gel-free proteomics

Due to various drawbacks in gel-based proteomics described in the previous chapter, efforts have been made for the development of gel-free techniques. Today, gel-free proteomic approaches have largely replaced the 2D-PAGE and 2D-DIGE methodologies. Especially the online coupling of reversed-phase liquid chromatography with mass spectrometry has led to the development of several novel gel-free proteomic methodologies. “These methods no longer rely on protein staining dyes, but on MS data. Currently, protein quantification by mass spectrometry is based either on measurement of signal intensities for samples labelled with different stable isotopes or use of label-free methods [409]” as elucidated in chapter 2.3.6.

2.3.5.1 Shotgun proteomics

Shotgun or bottom-up proteomics is the most prominent proteomic approach. Based on the analysis of protease-derived peptides with tandem mass spectrometry, it allows a global identification and systematic profiling of the dynamic proteome [410]. The general workflow consists of enzymatic protein digestion (commonly using trypsin), chromatography-based complexity reduction, peptide ionization, separation of the precursor peptide ions based on their m/z ratio, fragmentation of selected precursor ions, analysis of the product ions and finally data analysis by bioinformatics tools [411]. Protein quantification can be performed either label-based or label-free. Both approaches are discussed in chapter 2.3.6.

The aim of shotgun proteomics is the identification of all proteins in a given mixture (e.g. tissue lysates). “Peptide mixtures derived from enzymatic digestion of proteomic samples are usually much too complex to be directly analysed by mass spectrometry. For this reason, most standard proteomic procedures include a liquid chromatography-based peptide-separation step before MS analysis. Both online and the offline coupling of reversed-phase high-pressure liquid chromatography (RP-HPLC) to MS can easily be achieved because of the compatibility of the solvents deployed with MS [412, 413]. Furthermore, the offline coupling of two-

dimensional (strong cation-exchange and reversed-phase) liquid chromatography [414], capillary zone electrophoresis [415] and isoelectric focusing/reversed-phase liquid chromatography [416] to mass spectrometers […] has been demonstrated. In 1997, MacNair and colleagues presented their work on the use of extremely high pressures in liquid chromatography, termed ultrahigh-pressure liquid chromatography (UHPLC) [417]. A large increase in the application of UHPLC for proteomic applications has been observed in recent years [418-422], showing itself not only in the number of publications but also in the number of commercially available instruments. UHPLC not only provides enhanced peak resolution, and thereby an improved sensitivity [423], but at the same time makes great demands on the analytical power of the downstream mass spectrometer. Short peptide elution times require fast duty cycles during MS_1 and MS_2 analysis when using electrospray ionization, a prerequisite met only by the latest generation of mass spectrometers. By contrast, the offline peptide separation used with MALDI-based mass spectrometers can more easily cope with this demand."

Although state-of-the-art mass spectrometers allow for the identification of 1000s of proteins in a single shotgun experiment, false positive and false negative identification rates are a problem. False identifications can occur either from incorrect assignment of fragment spectra to peptide sequences or from wrong protein assignment in the peptide mapping process [187]. Therefore, several techniques such as searching against a decoy database [424] or the utilization of a scoring system [425] to evaluate the false discovery rate (FDR) have been developed. Often the FDR is estimated too optimistic and the true FDR can be up to 10-fold higher [426].

2.3.5.2 Capillary electrophoresis mass spectrometry (CE-MS)

Capillary electrophoresis mass spectrometry (CE-MS) was first described by Richard Smith in 1987 [427]. This method is usually applied for the analysis of the low molecular weight proteome (peptidome) [428-430]. In CE peptides are separated based on their molecular mass and their composition of basic, neutral or polar amino acids [431]. For CE-MS uncoated bare fused silica capillary columns at low pH showed the best results [432]. Following their separation, peptides are ionized and analysed in the mass spectrometer. Different instrument setups such as CE-MALDI-TOF/TOF, CE-ESI-QTOF and CE-ESI-orbitrap [431, 433] can be utilized. Although CE-MS is an interesting approach offering an orthogonal separation technique, it still has to cope with limited reproducibility/stability, interference problems, restricted quantification capabilities and inter-laboratory variability [434].

2.3.5.3 MALDI-Imaging

Using MALDI-imaging, the spatial distribution of analytes on thin slices of tissue can be observed [435]. In contrast to immunohistochemical approaches, MALDI-imaging does not require an antibody for detection. The method is rather based on proteomic methods using the soft ionization method MALDI and usually a time-of-flight (TOF) mass spectrometer. Therefore a thin tissue section immobilized on an oxide-coated glass slide and covered with matrix is analysed in an array-type manner by MALDI-TOF MS [436] (Fig. 13). Resulting spectra are analysed and visualized by suitable imaging software (e.g. MSiReader [437], Biomap (Novartis)). Subsequently, the images can be combined with other imaging modalities like fluorescence, immunohistochemistry and magnetic resonance imaging [438]. Single-cell imaging, a further development of this method, allows for the visualization of individual cells [439]. The spatial resolution of images generated by MALDI-imaging is limited by the focal diameter of the laser and the displacement of the analytes during sample preparation. Currently, images with a resolution of 10 µm can be acquired [439]. One of the advantages of MALDI-imaging is its ability to visualize analytes which are only expressed in a restricted area of the tissue and which would probably not be detect in whole tissue lysates due to the limited dynamic range of the MS. MALDI-imaging has been applied to several tumour entities including breast [440], lung [441] and renal cancer [442].

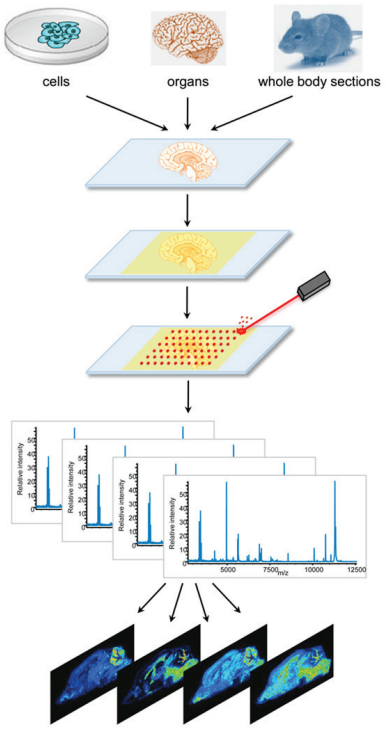


Fig. 13: MALDI-imaging principle Thin tissue sections are covered with matrix and are analysed in an array-based manner using MALDI-TOF MS. The spatial m/z data can be represented in a 2D intensity map. Reprinted from Schwamborn & Caprioli 2010 [443].

2.3.5.4 Selected Reaction Monitoring (SRM)

Selected reaction monitoring (SRM) was elected method of the year 2012 by the journal "Nature methods" [444]. It is a targeted quantitative proteomic approach preferably applied in biomarker validation studies but rather seldom in biomarker discovery. The basic principle behind SRM has been utilized in small molecule research for several decades [445]. SRM is usually performed on a triple quadrupole mass spectrometer. The first and the last quadrupole serve as mass filters, while the second quadrupole acts as fragmentation unit (Fig. 14) [380, 446]. In the first quadrupole a peptide precursor is selected based on a specific m/z value. Subsequently, these ions are fragmented in the second quadrupole. Finally, the third quadrupole filters predefined fragments of a certain m/z value, which are then detected. A precursor – product ion pair is called a transition. In general, several transitions are observed for the appropriate analysis of a protein. The dual selection process in SRM has several advantages such as high selectivity, low background signals and extended dynamic range (4-5 orders of magnitude [447, 448]. Using SRM, a reliable protein quantification can be obtained [449]. The usage of heavy labelled peptides (e.g. incorporated ^{13}C, ^{15}N in C-terminal lysines and arginines) supports quantification and enhances the reliability of the approach. Compared to antibody-based approaches, SRM method development is both faster and less expensive as it does not require the generation of antibodies.

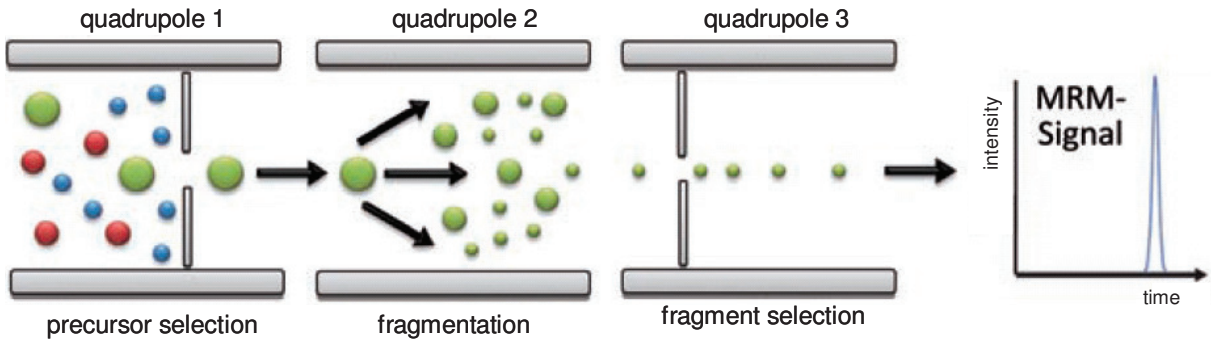


Fig. 14: SRM mode in triple quadrupole instruments In the first quadrupole the precursor peptide of interest is selected and fragmented into product ions in the second quadrupole. The third quadrupole performs the selection of a specific product ion, whose intensity is finally recorded by the detector. A certain precursor – product ion pair is called a transition. To perform reliable protein quantification, several transitions per protein are recorded. Adapted from Schmidt *et al.* 2008 [450].

2.3.6 Quantification in proteomics

"It is noteworthy that mass spectrometry on its own is, for several reasons, not quantitative [451, 452]. To analyse a molecule by mass spectrometry, it must be ionized, introduced into the analyser, separated and detected. The efficiency of this process is highly dependent on the amino acid composition of each peptide [453, 454]. Hence, different enzymatic peptides

from the same protein, although being theoretically present with the same number of molecules, have different ionization efficiencies and thus different signal intensities [455]. Furthermore, ion suppression affects precision, accuracy, and detection capability in peptide analysis. Variable background levels and further inter-LC–MS^1–MS^2 differences lead to quantification problems. Finally, the linearity of four to six log shifts achievable with state-of-the-art mass spectrometers is still way below the dynamic range of a complex proteomic sample [456, 457]. Nevertheless, mass spectrometry is often used both for protein identification and quantification, and a multitude of quantification methods have been developed in recent years."

2.3.6.1 Label-based quantification

The labelling in label-based quantification approaches can be performed metabolically, chemically or enzymatically. For every labelling strategy several techniques have been proposed. "The first gel-free technique for protein quantification, termed isotope-coded affinity tagging (ICAT), introduced in 1999 by Gygi and colleagues [409], made use of reagents with different masses due to incorporation of stable isotopes. While one sample is labelled with the heavy reagent, the other is labelled with its light version […] [458]. After combination of the two samples and proteomic analysis, protein quantification is based on MS^1 signal intensity ratios between peptide peaks with a distinct mass difference. Because both samples are analysed in conjunction their signal intensities measured during MS^1 analysis are a direct measure of the relative amounts of protein. During subsequent years, a number of isotope-based protein quantification strategies were developed […] [362]. One of the most widely used methods for introduction of stable isotopes is metabolic incorporation using the SILAC (stable isotope labelling with amino acids in cell culture) technique [459]. Although the first applications of this technology where restricted to cells cultured in vitro [459], this approach was further extended to the labelling of bacteria [460], nematodes [461], insects [461], and rodents [462]. Alternatively, isotopic labelling can be achieved by chemical or enzymatic methods. Techniques using chemical labelling include, for example, isobaric tagging for relative and absolute quantitation (iTRAQ) [463], isotope-coded protein labelling (ICPL) [464], and tandem mass tags (TMT) [465], to name only a few. Finally, absolute quantitative information can be achieved, for instance, by applying AQUA peptides [466]. Advantages of all these label-based protein quantification methods are: 1. increased reproducibility because of multiplexing (i.e. the parallel processing of all quantified samples in one run); and 2. the possibility of directly comparing measured intensities of all labelled peptides. But these methods have several disadvantages, for example: 1. high reagent costs; 2. incomplete labelling; 3. increased sample complexity and reduced dynamic range because of sample

multiplexing; and 4. restricted number of possible comparisons [452]. Today, a maximum of twelve samples can be processed in parallel by label-based protein quantification methods [467, 468]. Future developments could increase this number [469] but sample multiplexing in label-based quantification approaches will always be a limiting factor for large-scale comparative studies."

2.3.6.2 Label-free quantification

"Shortly after the development of the label-based approaches described above, label-free methods for protein quantification by mass spectrometry were introduced. […] Protein quantification is based solely on data obtained directly from the mass spectrometer. Two distinct approaches to the label-free quantification of proteins by mass spectrometry emerged: 1. techniques based on the assumption that protein abundances directly correlate with the area under the curve (AUC) of precursor ion spectra [470]; and 2. methods counting MS^2 spectra belonging to peptides derived from the same protein, thereby relying on the assumption that the number of MS^2 spectra correlates directly with protein abundance [471]." "When MALDI is used as ionization method, label-free quantification based on ion intensities is applied almost exclusively." "Samples used for label-free protein quantification are processed in parallel, thus increasing susceptibility to technical and/or handling errors. In return, no additional labelling steps are required, making these techniques both cost-effective and simple. Importantly, samples are not diluted by merging, enabling, therefore, coverage of a higher dynamic range. Last, but not least, label-free quantification theoretically enables analysis of an unlimited number of samples at the same time, the only limit being the computational power required for data analysis [472]."

2.3.6.2.1 Ion intensity measurements (AUC)

"The output of every mass spectrometer consists of two quantities: the mass-to-charge ratio and the corresponding signal (or ion) intensity. The correlation between the measured ion intensity and the concentration of the analysed analyte using ESI was observed for the first time by Voyksner and Lee in 1999 [473]." Therefore, protein quantification can be based on peak areas (AUC) of peptides of interest. "Before the label-free quantification, mass spectra have to be preprocessed to enable extraction of quantitative information [474]. Preprocessing consists of several individual operations which can be executed in different orders. So far, no optimum operating sequence has been determined. Thus, a wide range of workflows for bioinformatics analysis of data derived from label-free quantification experiments has been published [475-478]. […] In general, the following steps must be performed: 1. de-noising and

smoothening of raw mass spectra; 2. baseline subtraction; 3. peak detection; 4. retention time shift alignment; 5. ion intensity normalization; 6. feature extraction, and 7. quantification. […]

Because mass spectra are often corrupted by noise, peak detection is usually performed with smoothing and baseline correction, i.e. removal of inherent high-frequency and low-frequency noise, respectively. […] Numerous algorithms have been proposed during recent years to enable extraction of informative peaks corresponding to true peptide signals, but only lately has the performance of these algorithms in MALDI–MS data analysis been compared [479, 480]. Bauer and colleagues investigated the performance of three commonly used approaches for peak detection: signal-to-noise ratio, continuous wavelet transform, and a correlation-based approach using a Gaussian template [480]. […] Owing to the high resolution of modern mass spectrometers, isotopic pattern of peptides are resolved by default. To achieve accurate quantification these different isotope peaks must be merged, an operation called isotope deconvolution [481]. Both peak detection and isotope deconvolution lead to a dramatic decrease in data points and, therefore, not only simplify label-free quantification but also increase processing speed and reduce the computational power required. […]

Even when sample preparation and analysis are conducted as carefully as possible, minor run-to-run differences in peak intensities cannot be avoided; they are, therefore, often observed. These differences can be ascribed to technical variations, for example slight variability of chromatographic gradients or increasing accumulation of contamination in the mass spectrometer, a phenomenon particularly observed in MALDI-based mass spectrometry. Furthermore, order-dependent bias can often be observed when analysing large proteomic datasets acquired consecutively [482]. Several data-normalization algorithms" such as local regression normalisation [483] and linear regression normalisation [482] "have been used to compensate for these run-specific intensity differences. […]

After peak detection, peptides eluting in several fractions must be merged into features. Merging consecutively eluting peaks derived from the same peptide into features not only reduces sample complexity but enables subsequent sample alignment. […] Identification of peak signals belonging to the same peptide which elutes in consecutive fractions is usually achieved by analysing its profile on the RT×m/z surface. Peak signals derived from the same peptide should be found in consecutive fractions and have highly similar mass-to-charge ratios. […] Almost every software suite" (e.g. DeepQuanTR [484], MarkerView (AbSciex), Expressionist (Genedata)) "developed for label-free quantification uses a dedicated algorithm for grouping of mass spectrometric signals derived from the same peptide. However, all algorithms take both the mass-to-charge ratio and the retention time window into account for peak grouping. […]

To accurately perform comparative quantification, peak and/or feature intensities derived from identical peptides in different samples must be matched. This matching must be performed for both mass-to-charge ratio and retention time. […] Such retention-time shifts could arise from the column (column packing, column age, and contamination) and/or the system (temperature and gradient instability) [485]. Measured mass-to-charge ratios also deviate slightly between runs. […] Although these shifts may seem very small, sample alignment for large data sets is a crucial and laborious process for which different alignment algorithms have been developed. These algorithms" [486-490] "differ in: 1. the scoring of the alignments; 2. the input data; and 3. the choice of the reference sample. […]

After all these pre-processing steps, the relative quantification of peptides (and proteins) can be performed. Protein quantification based on AUC values is commonly achieved by use of two different methods [491]: 1. integration of the AUC post-peak deconvolution [484, 492]; and 2. integration of the AUC of the monoisotopic peak [493]. The AUC is calculated by simple summation of the intensity values of all the peaks grouped into a feature. Choice of one of these methods for integration of the AUC depends on the sample pre-processing, i.e. on whether or not deconvolution was performed. […] Carrillo and colleagues have shown that the second method based on the sum of intensities results in better protein quantification [494]. It not only requires less computational power but incorporates: 1. the signal intensity (and therefore data quality) into estimates of change; and 2. enables quantification data to be obtained from a small number of peptides."

2.3.6.2.2 Spectral Counting

"Spectral counting was introduced in 2004 by the Yates laboratory for estimating relative protein abundance in ESI-based shotgun proteomic experiments [471]. This technique uses the direct correlation between the abundance of a protein and the number of MS^2 spectra obtained from peptides belonging to this protein as the basis for quantification [471, 495]. Highly abundant proteins lead, therefore, to a large number of proteolytic peptides, which thus leads to increased sequence coverage [496]. Hence, it is fundamental that a large number of MS^2 spectra for each identified peptide is recorded." In principle, there are two processing methods for spectral counting in use. The first method initiates the analysis process with the interpretation of MS^2 spectra by database search. Subsequently, all spectra of one protein and the corresponding peptide ions are mapped together and the spectra are counted for quantification. The second method first clusters the gained MS^2 spectra based on their similarity using spectra comparison algorithms and finally performs the peptide/protein identification. Whereas the first processing way is quite simple and doesn't require special tools, useful software tools are crucial for spectral counting working with the second way.

2.3.6.2.3 Exponentially modified protein abundance index (emPAI)

"The exponentially modified protein abundance index (emPAI) [497-499], which is based on the number of peptides observed for a distinct protein compared with the number of peptides possibly observable for this protein, is probably the most simple label-free quantification technique so far available. It relies solely on protein identification information and can be applied to MS data obtained from any instrument equipped with both MALDI and ESI ion sources. Several protein-identification algorithms (for example Mascot [497]) have integrated this method of label-free quantification into their standard workflow. emPAI can, therefore, be flexibly combined with any other quantification technique."

2.3.6.2.4 APEX approach

"The APEX approach uses a combination of machine learning classification, for estimation of peptide detection probabilities, in combination with features of spectral counting [500, 501]. The protein's observed MS^2 total spectral count is compared with the predicted spectral count. A correction factor, termed the Oi value, is calculated for each protein to take into account variable peptide detection by MS techniques. APEX thus combines the expectation of observing a specific peptide, the total sampling depth, and confidence in protein identification in the correction factor Oi [452, 500, 501]."

2.3.7 Discovery of vascular accessible targets by proteomics

Conventional treatment options in cancer therapy such as chemotherapy are often not exclusively active at the tumour site. Reactive drugs are rather distributed in all organs causing several side effects [502]. Of all things, the drug accumulation in the tumour is the lowest and reaches only 5-10 % of the dose compared to the rest of the body [503]. The low reagent concentrations in the tumour can be explained by the high interstitial pressure in the tumour environment [504] as well as an increased active efflux of reagents through multidrug resistance transporters [505]. To overcome these issues related with conventional therapies, selective and better tolerated drugs are needed. A promising approach for the development of novel therapeutic strategies is the guided transport of bioactive molecules directly into the tumour [4]. Localising the reactive reagent in the tumour will not only increase intra-tumoural drug concentrations, but also increase the therapeutic index through a higher efficacy while reducing side effects. The design of such innovative drugs presupposes the availability of tumour-associated markers as target molecules as well as high affinity delivery molecules [5]. Currently, several categories of delivery molecules are explored: targeting peptides [506, 507], aptamers [508], small organic molecules [509], and antibodies/antibody fragments, the most

prominent delivery molecules [510, 511]. Several targeting antibody-based therapeutics are already approved by the FDA, such as Trastuzumab emtansine [512] and rituximab (Rituxan, Roche).

However, beside highly selective delivery molecules, tumour specific targets are at least as important for the success of this therapeutic approach. Ideally, a valuable target is highly abundant in the diseased area but not expressed in healthy organs. Additionally, the vascular accessibility needs to be ensured. Therefore, the target molecule should be expressed on the cell surface or in the extracellular matrix of the tumour in close proximity to the vasculature. Identifying a valuable target is a challenging task as vascular accessibility of proteins within the tumour is reduced due to physical and kinetic barriers such as interstitial hypertension, long diffusion distances in the intersitium and a significant heterogeneity of the target protein expression in different tumour regions [513]. Additionally, tumour cells are known to acquire drug resistance mechanisms, i.e. the down regulation of the target molecule. Therefore, proteins on the tumour vasculature are supposed to be useful markers in targeted therapy as the vascular endothelial cells are less heterogeneous and the development of resistance mechanisms is decelerated [514, 515].

2.3.7.1 Enrichment of vascular accessible proteins

As described above, an ideal drug target is accessible from the vasculature and therefore expressed on the cell surface of endothelial cells, cells in close proximity to endothelial cells or in the perivascular extracellular matrix. Due to the wide dynamic range of protein concentrations and the high sample complexity in tissue lysates, the identification of such proteins is a challenging task. Membrane proteins are difficult to detect due to their low abundance and the presence of extended hydrophobic stretches. In shotgun proteomics approaches membrane proteins are often underrepresented [516]. Therefore, the enrichment of membrane proteins is indispensable [517]. The classical approach for the enrichment of membrane proteins is the differential centrifugation in a density gradient separating the different cell compartments based on their particle size and shape [518]. Applied gradients can consist of sucrose [519] or a mixture of dextran and PEG [520]. Additionally, most protocols contain a carbonate wash with $NaCO_3$ to remove weakly associated proteins [521-523]. Although it was shown that these enrichment strategies can reduce the fraction of intracellular proteins [524], these techniques nevertheless suffer from significant contaminations with intracellular proteins and proteins from cell organelles. Another approach for the analysis of cell surface-exposed epitopes is the short duration proteolysis [525]. Cell surface proteins are shaved off by enzymatic treatment with trypsin, chemotrypin, pronase or proteinase κ [526] and can be analysed by proteomics techniques. While no

solubility issues have to be considered within this approach the cell membrane has to be intact during the proteolytic digestion [527].

Although the enrichment and identification of membrane proteins with the previous mentioned strategies is feasible, these techniques do not consider the additional target requirement of vascular accessibility. To address this characteristic, the *in vivo* perfusion of a living animals or the *ex vivo* perfusion of surgically removed organs is required. Several perfusion-based proteomic strategies have been developed and are described in detail in the following chapters. An additional advantage of *in vivo/ex vivo* perfusion-based techniques is the fact that only proteins expressed under native conditions are detected. Beside cell surface proteins of tumour cells, which are investigated exclusively in the classical membrane enrichment strategies, perfusion-based methods own the possibility of additionally discover extracellular matrix proteins and target molecules on vascular endothelial cells.

2.3.7.1.1 Carbohydrate-based cell surface capturing

Cell surface glycoproteins are known to change their glycosylation pattern during oncogenesis as they play an important role in cell adhesion [528]. Therefore, several enrichment methods for this protein class have been developed in the past years [529]. Lectin affinity chromatography [530] and isotope-coded glycosylation specific tagging (IGOT) [531] are two prominent examples for the simple and cost-effective enrichment of carbohydrate containing proteins. Furthermore, technologies utilising the chemical reactivity of glycoproteins have been developed. The BEMAD technique labels O-GlnNAc containing proteins via β-elimination/Michael addition reaction [532] while another technique enriches for glycoproteins with a mannose-boronic acid approach [533]. An additional possibility is the chemo-enzymatic tagging of GlcNAc sugars with a ketone-containing galactose analogue [534]. The carbohydrate-based cell surface capturing focuses on proteins which contain N-linked oligosaccharides (Fig. 15). In this approach, oxidised carbohydrate containing proteins are covalently labelled with biocytin hydrazide, which acts as a bifunctional linker between the glycoprotein and a solid support for subsequent affinity enrichment. Following cell lysis, extracted proteins are digested and tagged peptides are subsequently affinity purified via their biocytin hydrazide tag [535]. Purified peptides can be released from the affinity resin by enzymatic PNGase treatment and resulting peptides are submitted for proteomic analysis. Up to date carbohydrate-based cell surface capturing was only applied on plasma samples and *in vitro* cultured cells [535, 536]. However, this approach in principle could be used in a perfusion-based experiment for the enrichment of vascular accessible glycoproteins in living organisms.

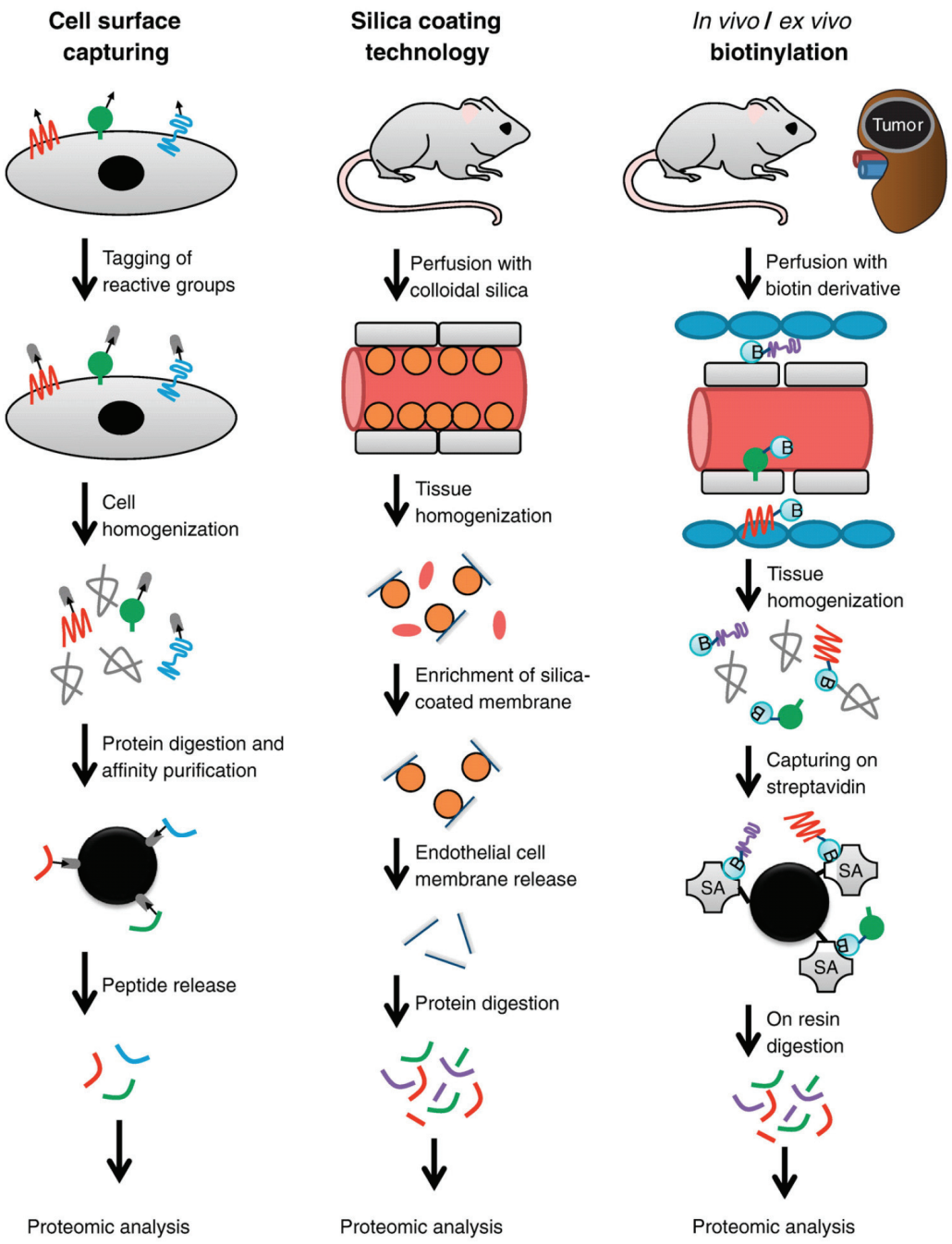


Fig. 15: In vivo applicable cell surface enrichment strategies In carbohydrate-based cell surface capturing oxidised glycoproteins are covalently labelled with biocytin hydrazide. After cell lysis proteins are digested and tagged peptides are affinity enriched. Following enzymatic treatment purified peptides are submitted for proteomic analysis. Using the silica coating technology, colloidal silica beads are attached to membrane proteins by electrostatic interactions. Subsequently, the beads are covalently cross-linked to the plasma membrane by polyacrylic acid. Pellicles formed during tissue homogenisation are enriched by centrifugation. Finally, peptides resulting from released, digested membrane proteins are analysed by proteomic techniques. The *in vivo*/*ex vivo* biotinylation strategy is based on the vascular perfusion of animals with a biotinylation reagent which covalently labels primary amines of accessible proteins. Following labelling and tissue homogenisation, biotinylated proteins are enriched on streptavidin/avidin beads, digested on resin into peptides, which are finally analysed by mass spectrometry. Reprinted from Strassberger *et al.* 2010 [4].

2.3.7.1.2 Silica coating technology

Electrostatic interactions between cationic colloidal silica beads and the anionic membrane proteins are exploited in this method [537] (Fig. 15). Subsequent to the attachment of the silica beads an anionic crosslinker (e.g polyacrylic acid) is applied in order to covalently connect the silica beads to the membrane. Following cell lysis the bead-membrane complexes form so called pellicle, which can be easily collected by centrifugation due to their increased partitioning density. Beside an *in vitro* application for the enrichment of cell surface proteins the silica coating technology has been applied *in vivo* enabling the enrichment of cell surface proteins originating from vascular endothelial cells [538, 539]. However, this method is mainly suitable for soft tissues as the shearing forces applied during harsh homogenization methods required for stronger tissues (e.g. tumours) can destroy the pellicles [540].

2.3.7.1.3 Biotinylation strategy

Another strategy for the enrichment of vascular accessible proteins is the *in vivo* biotinylation. This method relies on the covalent attachment of a biotinylation reagent to reactive residues of accessible proteins (Fig. 15) [6, 7]. A typical reagent consists of a biotin group, a sufficiently sized spacer to avoid steric hindrance and a reactive group. Reactive groups can be N-hydroxysuccinimide esters (NHS), N-iodoacetyl or maleimide groups or hydrazide groups [517]. While hydrazides react with aldehyds, maleimides and iodoactyls target sulfhydryls. NHS esters, which are mainly applied in this approach, target ε-amino groups in side chains of lysines and unsubstituted N-terminal α-amino groups of proteins. Several biotinylation reagents are commercially available such as sulfo-NHS-LC-biotin and sulfo-NHS-SS-biotin [541]. Following labelling and tissue homogenisation the biotinylated proteins are purified via the high affinity interaction ($K_d = 10^{-15}$ M [542]) with streptavidin or avidin and analysed by proteomic technologies. The method can be applied *in vitro* for the enrichment of the cell surface proteome [543, 544], but also for the enrichment of the vascular accessible proteins by *in vivo* perfusion [6, 545, 546]. Even the *ex vivo* perfusion of a surgically removed specimens (e.g. human tumour-bearing kidney) has been performed [547]. Compared to the previous described silica coating technology the *in vivo* biotinylation approach not only targets cell surface proteins of vascular endothelial cells but also labels proteins of the extracellular matrix and membrane proteins of cells in close proximity to blood vessels. The increased permeability of the tumour vasculature facilitates the extravasation of the reagent [548]. Nevertheless, enrichment of intracellular proteins cannot be completely excluded since the reagents are not totally membrane-impermeable. Also the presence of necrotic cells shedding their cytoplasm into the microenvironment can result in a larger fraction of intracellular proteins [549].

However, the *in vivo* biotinylation technique is a very promising and widely applied approach, which in theory allows the covalent biotin labelling of every vascular accessible protein.

2.3.7.1.4 *In vivo* phage display

The aim of *in vivo* phage display is the identification of peptide sequences specifically binding to tissue-specific endothelial cell markers. The method is an advancement of the conventional phage display technique described 1985 by George P. Smith [550] and patented by George Pieczenik [551]. The central point of the phage display technology, mainly used for the selection of human, monoclonal antibodies, is a library containing billions of bacteriophages. These phages display a randomized polypeptide on their surface through the coupling of the polypeptide DNA to a phage coat gene. The polypeptide can fulfil a certain function, typically the specific binding to an antigen. Thereby, the phage links the genotype (DNA sequence) with the phenotype (binding specificity). In the *in vivo* approach developed by Ruoslahti, Pasqualini and co-workers [552, 553], peptide phage libraries are injected intravenously into mice and organs or tumours are dissected in order to extract the phages bound to potential vasculature targets (Fig. 16). Subsequent sequencing, immunohistochemistry and Western blot analyses are require identifying and validating the discovered structures. Due to the intransparent patent situation and the significant costs for the generation of the phage liberary this method is barely used. Nevertheless, *in vivo* phage display is a powerful tool for the identification of specific binders to the tumour vasculature. Applying this technology, specific peptides homing to aminopeptidase N, αV integrins, NG2 and MMP-2/MMP-9 as well as tumour lymphatics have been identified [554].

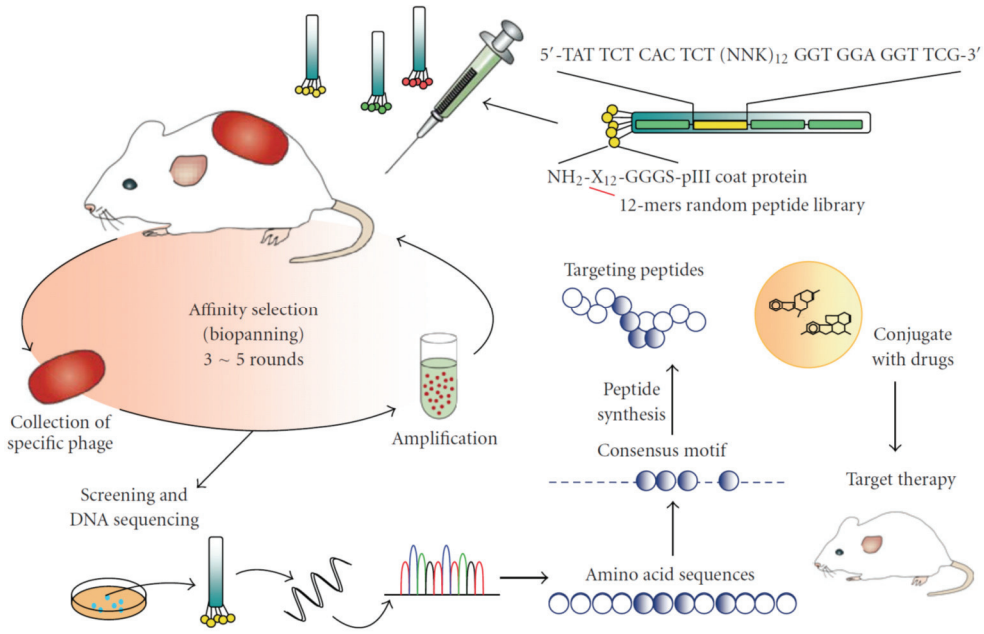


Fig. 16: Workflow of in vivo phage display A library containing phages expressing randomized peptides on their surface is injected into tumour bearing mice. The bound phages in the tissue of interest are harvest, amplified and sequenced. The obtained consensus sequence is used for the synthesis of peptide based drugs for targeted therapy. Reprinted from Wu & Chang, 2010 [555].

2.4 Abstract on selected biomarker candidates

In the following chapter a short summary on the characteristics of proteins, which are proposed in this thesis as biomarker candidates, is given.

2.4.1 Insulin-like growth factor-binding protein 3 (IGFBP3)

Insulin-like growth factor-binding protein 3 (IGFBP3) is the most abundant member of six identified LTBPs. It consists of an IGFBP domain and a thyroglobulin type-I domain. So far, IFFBP3 has mainly been observed as secreted protein [556]. However, intracellular localisations have also been reported [557, 558]. IGFBP3 is an important modulator of IGF activity [559]. A stable complex is formed upon binding of IGFBP3 to IGFI and IGFBII preventing IGFs from activating their IGFB receptors. Thereby, both the cell proliferation and the survival stimulation capabilities of IGF signalling are disrupted, leading to growth inhibition and apoptosis [560]. On the other side, also growth stimulatory effects of IGFBP3 have been reported [561]. Increased IGFBP3 expression is found in many malignancies [562, 563] including cancer and a resulting inhibition of growth and metastasis potential is reported in all cases [556, 564, 565]. Recently, the upregulation of IGFBP3 in clear cell renal cell carcinoma

has been reported and cell proliferation suppression effects for IGFBP3 have been demonstrated [566].

2.4.2 Latent-transforming growth factor beta-binding protein 2 (LTBP2)

Latent-transforming growth factor beta-binding protein 2 (LTBP2) is a large extracellular matrix protein with a multi-domain structure related to fibrillins. It consists of two types of repeating cysteine-rich domains, epidermal growth factor repeats and so called eight-cysteine domains [567]. Up to know, the function of LTBP2 has not been extensively studied. Other members of this family were reported to be extracellular matrix components and form large latent complexes with TGF-β [568]. However, LTBP2 is not part of these complexes. The protein is furthermore associated with a crucial function during embryonic development as $LTBP2^{-/-}$ mouse embryos die very early [569]. Furthermore, a strong association with elastic fibres has been identified suggesting a structural role of LTBP2 in the extracellular matrix [570]. LTBP2 is found to be expressed in the lung, heart, placenta, liver and skeletal muscle [571]. In tumours it showed a tumour suppressive function due to reduced cell migration and invasiveness [572]. LTBP2 is patented both as a biomarker for renal dysfunction (wo/2011/117371) and as a therapeutic target (EP 2102357 B1).

2.4.3 Lipocalin-2 (LCN2)

Lipocalin-2 (LCN2), also known as neutrophil gelatinase-associated lipocalin (NGAL), belongs to the calycin superfamily, whose members are associated with the transport and binding of small hydrophobic molecules [573]. LCN2 has been observed as monomer, as disulfide linked homodimer and as heterodimer with metalloproteinase 9 [574]. This 22.6 kDa secreted protein is typically expressed in the bone marrow and in tissues prone to the exposure to microorganisms such as the respiratory tract [575]. Furthermore, LCN2 expression has been observed in the uterus, prostate, salivary gland, stomach, appendix and colon [575, 576]. LCN2 expression is activated by the oncoprotein BCR-ABL through the JAK/STAT pathway, which culminates in binding of Stat5 to the 24p3 promoter [577]. LCN2 is part of the iron delivery pathway involved in apoptosis through blocking of the pro-apoptotic protein Bim, in the innate immune system with an antibacterial iron depletion strategy and in renal development [578]. As LCN2 is a protective protein in ischemic injury, its overexpression can be observed in several nephrologic disorders [578] and serves therefore as marker for kidney injury [579]. The overexpression of LCN2 is in particular linked with tubulointerstitial damage in chronic kidney disease [580]. High expression levels of LCN2 have been demonstrated in many different cancers [578] including clear cell renal cell carcinoma and corresponding metastases

[581]. However, no conclusive results of the usefulness of LCN2 as a diagnostic marker have been published so far [582].

2.4.4 Monocarboxylate transporter 1 (SLC16A1)

The 53.9 kDa protein monocarboxlate transporter 1 (SLC16A1) belongs to the family of proton coupled monocarboxylte transporters consisting of totally 14 members. All family members feature twelve transmembrane domains, from which domain 6 and 7 are connected via a large cytosolic loop [583]. SLC16A1 assures the rapid transport of many monocarboxylates such as lactate, pyruvate, branched-chain oxo acids derived from leucine, valine and isoleucine, and the ketone bodies acetoacetate, beta-hydroxybutyrate and acetate [584]. Additionally. SLC16A1 is involved in the cellular response to a high-fat diet by modulating the cellular levels of lactate and pyruvate. Typically, SLC16A is expressed in the heart and on blood lymphocytes and monocytes [585]. For its subcellular localisation in the plasma membrane [586] SLC16A1 requires ancillary proteins such as basigin and embigin [587] because the protein Otherwise accumulates in the Golgi apparatus [588]. Defects in SLC16A1 are associated with a symptomatic deficiency in lactate transport [589] and familial hyperinsulinemic hypoglycemia 7 [584]. Owing to the fact that glycolytic processes are often upregulated in cancer the inhibition of SLC16A1 was shown to reduce cell proliferation in tumour cell and to induce cell death [590].

2.4.5 Tenascin-X (TNXB)

The large glycoprotein tenascin-X (TNXB) belongs to the family of tenascins consisting of a total of four members. This 464 kDa extracellular matrix protein can be found as homotrimer and consists of 19 EGF-like domains, 32 fibronectin type-III domains and one fibrinogen C-terminal domain [591]. The expression of TNXB is largely restricted to foetal tissues. Thereby TNXB expression can be found in the adrenal, the testis and smooth, striated and cardiac muscles. The function of TNXB has so far not been extensively studied. However, the related protein tenascin C is very well characteristed in various embryonic processes and pathological conditions caused by infections, inflammation or during tumorigenesis, where it acts in cell adhesion, migration and growth [592]. Large isoforms of tenascin C have also been associated with tissue remodelling and invasiveness of carcinomas [593]. Similar functions are proposed for TNXB. It is reported to mediate interactions between cells and the extracellular matrix and accelerates the collagen fibril formation [594]. Defects in TNXB expression can lead to the Ehlers-Danlos syndrome, which is an inherited connective tissue disorder caused by defects in the structure or processing of collagens [595]. The mechanisms how TNXB affects collagen

deposition are unknown. Recently, TNXB was reported as a novel diagnostic marker for malignant mesothelioma [596]. Interestingly, tumour invasion and metastasis were demonstrated to be promoted in TNXB deficient mice [597], while also a supportive role in tumour growth of epithelial tumours has been associated with TNXB.

2.4.6 Transforming growth factor-beta-induced protein ig-h3 (TGFBI)

The protein Transforming growth factor-beta-induced protein ig-h3 (TGFBI) consists of four fascilin-1 domains and one carboxy-terminal Arg-Gly-Asp (RGD) sequence [598]. In healthy organism this 68 kDa secreted extracellular space protein is expressed in the corneal epithelium [599] where it fulfils a so far unknown function. Upon induction by TGF-β [600], TGFBI can bind to fibronectin, collagens [601], and integrins [602]. Stimulation of adhesion, spreading, migration and proliferation was observed in renal proximal tubular epithelial cells [603]. By the release of its RGD sequence TGFBI can induce apoptosis [604]. Mutations in TGFBI are associated with various different corneal dystrophies [605] and TGFBI overexpression can be observed in many different cancers [606-608]. In clear cell renal cell carcinoma TGFBI was found to be upregulated as a target of VHL [609]. In melanoma and colon cancer patient´s TGFBI overexpression is associated with high metastatic potential [610, 611]. A repression of metastases in prostate cancer could be demonstrated upon targeting TGFBI [612].

3 AIM OF THE THESIS

Clear cell renal cell carcinoma represents the most common subtype of kidney cancer [44]. The current treatment strategy comprises partial or radical nephrectomy followed by the administration of small molecule inhibitors of the mTOR and the VEGF pathways [120]. Although the prognosis for patients with local disease is quite promising, the survival rate of patients suffering from advanced clear cell renal cell carcinoma with non-dissectible tumours and metastasising disease is dramatically low. Therefore, the identification of new therapeutic strategies enabling a comprehensive and effective treatment of this malignancy is essential. A promising approach for the development of new therapeutic strategies is the guided transport of bioactive molecules directly into the tumour [4]. The design of such innovative drugs presupposes the availability of vascular accessible tumor-associated biomarkers (as target molecules) as well as high affinity targeting molecules such as monoclonal antibodies (as delivery molecules) [5].

The aim of this thesis was the identification of vascular accessible markers for clear cell renal cell carcinoma and its associated lung metastases using the *in vivo* biotinylation methodology [6, 7]. A schematic outline of this thesis represented by means of the development process of a targeted therapeutic is depicted in Fig. 17. The biomarker discovery workflow utilized recently established xenograft mouse models (Dr. Teresa Rigo-Watermeier (HI-STEM gGmbH), not published) representing patient derived clear cell renal carcinoma kidney tumours and spontaneous lung metastases. As healthy controls, both normal murine organs (kidney and lung) as well as human renal cells have been used.

The *in vivo* biotinylation approach is based on the terminal perfusion of mice with a membrane impermeable reactive ester derivative of biotin. Following tissue lysis the biotin-tagged proteins are analysed by proteomics technologies (LC-MS/MS) and quantified label-free by in house developed software tools (Alexander Kerner (HI-STEM gGmbH), not published). One aim of this thesis was thereby the expansion of the published systemic circulation perfusion method with a technique for the perfusion of the pulmonary circulation to enable the biotinylation of lung metastases.

In order to identify vascular accessible biomarker candidates exclusively expressed or extensively up-regulated at the site of disease, the data sets from xenograft tumours were compared to data obtained from healthy mouse kidney and human renal primary cells. Lung metastases data sets were levelled to protein lists obtained from xenograft tumours and healthy mouse lungs in order to identify metastases specific markers. A set of biomarker candidates was finally submitted to a first validation process using orthogonal methods such as immunofluorescence and RT-qPCR.

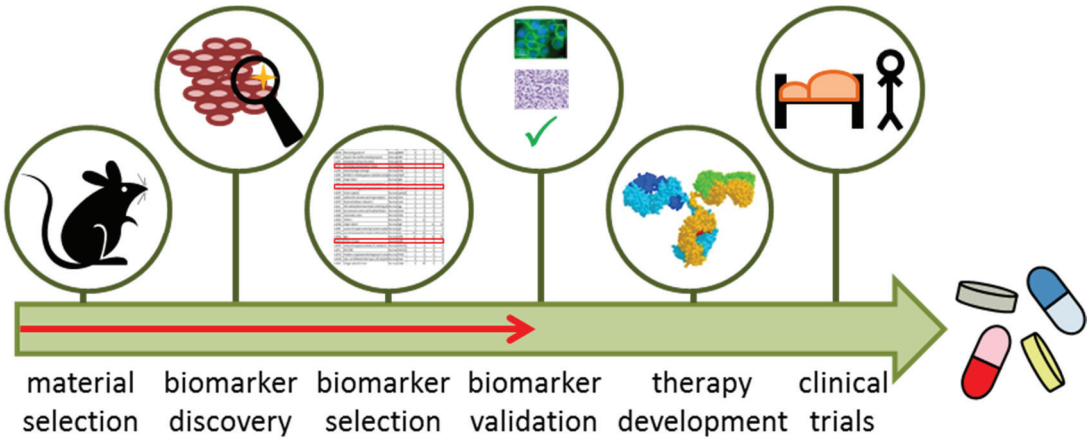


Fig. 17: Thesis outline illustrated using the development process of a targeted therapeutic The common development pipeline for a targeted therapeutic is initiated with the selection of a disease-representing (mouse) model, on which biomarker discovery techniques can be applied. Following the identification phase, biomarker candidates are selected and validated using orthogonal methods such as e.g. immunohistochemistry. Subsequently to the selection of biomarker-targeting human monoclonal antibodies, suitable bioactive compounds are coupled to the antibody and tested in a range of different disease models. Finally, the prospective therapeutic enters clinical trials. The steps executed in the course of this thesis (from material selection to potential biomarker validation) are marked in red.

4 MATERIALS AND METHODS

4.1 *IN VIVO* METHODS

4.1.1 Animals

For all animal experiments 4 to 12 weeks old female NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wj} (NSG) mice bred in the DKFZ animal facility were deployed. The mice were housed under specific pathogen-free (SPF) conditions in individually ventilated cages (IVCs) in the animal facility of the DKFZ. The experiments were conducted accordingly to the GV-SOLAS regulations and approved by the national authorities (Regierungspräsidium Karlsruhe; authorization number G134/12).

4.1.2 ccRCC xenograft generation

4 to 12 weeks old female NSG mice were anesthetised by an intraperitoneal injection of ketamin (100 µg/g) (Ketavet, Parke-Davis) and xylazin (16 µg/g) (Rompun, Bayer HealthCare). After shaving the fur a small incision was placed at the flank and the right kidney was partially exposed and via a 30 G syringe (BD, 04144150) 30 µl of the previous prepared cell suspension (chapter 4.2.3)corresponding to one million cells were injected under the kidney capsule. After injection the kidney was restored in the abdominal cavity and the gash was sutured by absorbable thread (Monosyn, Braun, 0022007) and wound clamps (Aesculup, Braun, BN507R). Kidney tumour and lung metastasis growth was monitored by palpating and magnetic resonance tomography (MRT; chapter 4.1.3). Depending on the injected tumour material, health status, tumour and lung metastasis size the mice were submitted to the final processing 3 to 40 weeks post injection.

4.1.3 Magnetic resonance tomography (MRT)

For magnetic resonance tomography the mice were anesthetized by a 1.5% Isofluran gas narcosis (Baxter, HDG9623). All measurements were performed on a 1 Tesla small animal tomograph (ICON, Bruker) controlled by the software Paravision 5.Beta-1.111 (Bruker). Two scan types, T1- and T2-RARE, depending on the visibility of the kidney tumour and the lung metastasis were operated. 10-15 min prior to T1 measurements 80 µl of a contrast reagent (0.5 mmol/ml, Magnevist, Bayer) were applied intraperitoneal. Per measurement 15 slices with a thickness of 1.5 mm and an interslice distance of 2.5 mm with a resolution of 0.0137 cm/pixel for T1 scans and 0.0182 cm/pixel for T2 scans were gathered. Afterwards the

tumour/metastasis size and volume was determined using the Paravision software. The mouse experiment was stop when reaching a tumour size of 1.4 cm or a metastasis size of 0.4 mm, in accordance with the termination criteria specified in the animal research license G134/12.

4.1.4 *In vivo* biotinylation of vascular accessible proteins - systemic circulation

The *in vivo* biotinylation by the perfusion of the body circulation was performed as described previously. [6, 7]. Following the anaesthesia of the mice using 200 µg ketamin, 20 µg xylaxin and 3 µg acepromazine (Vetranquil, Ceva) per g of body weight, the abdomen was opened and a median sternotomy was performed. For the perfusion a butterfly needle with a small hook was entered into the left ventricle of the heart and a small cut in the heart atrium was placed in order to allow an effluence of the perfusion liquids. The body circulation of the mice were now perfused with biotinylation solution (1 mg/ml sulfo-NHS-LC-biotin (Proteochem, b2103) in 10% (w/v) Dextran-40 (US biological, D6030) in PBS (Invitrogen, 18912014) and quenching solution (50 mM Tris-HCl (Sigma, T1503) in 10% (w/v) Dextran-40 in PBS, pH 7.4). Each solution was prewarmed to 42°C and applied for 11 min by a flow rate of 1 ml/min using a syringe pump system (Landgraf Laborsysteme, LA30). During the first 15 min of perfusion the thorax region of the mouse was wetted with sprinkling solution (50 mM Tris-HCl in PBS, pH 7.4) in order to directly quench the released excessive biotinylation reagent. After perfusion the kidney tumour and other organs were excised and one part was directly snap-frozen in liquid nitrogen and stored at -80°C for the preparation of tissue lysates used in the proteomic analysis and RT-qPCR. The other half was embedded in OCT-medium (TissueTek, Sakura, 4583) for later histochemical analysis. The blocks were snap-frozen in liquid nitrogen and stored at -80°C. In total 98 mice were perfused successfully and 10 RCC12, 9 RCC15, 3 RCC18, 9 RCC24, 9 RCC27, 9 RCC38, 8 RCC50 and 10 healthy female Nod-scid common gamma mice were used for the subsequent proteomic analysis.

4.1.5 *In vivo* biotinylation of vascular accessible proteins - pulmonary circulation

The *in vivo* biotinylation of the lung metastases or the lungs was performed by perfusion of the lung circulation of the mice. Therefore, mice were anesthetised as described in chapter 4.1.4and the abdomen was opened. A butterfly needle with a small hook was inserted in the inferior *vena cava* and fixed by a ligature using a surgery thread. Depending on the size and the growth direction of the kidney tumour the position of the needle was either cranial or proximal of the arborisation to the kidneys. To release the perfusion liquids a cut in the abdominal aorta proximal to the perfusion start point in the *vena cava* was introduced. The

perfusion and the processing of the lung metastases and organs were performed accordingly to chapter 4.1.4. In total, 75 mice were perfused and 2 RCC12-, 3 RCC15, 9 RCC27, 5 RCC50 and 5 healthy female NOD scid gamma mice were utilised for the subsequent proteomic analysis.

4.2 *IN VITRO* METHODS

4.2.1 Cells and Cell Culture

The three primary cell populations isolated either from the medullary, the cortical or the proximal tubules epithelium of human kidney tissue were bought from Lifeline Cell Technology (FC-0018; FC-0012; FC-0013). The cells were cultured adherently in T75 cell culture flasks (TPP, 90076) at 37°C and 5% CO_2. For amplification the cells were grown as recommended by the manufacturer in RenaLife basal medium (Lifeline Cell Technology; LM-0010) supplemented with the RenaLife Factors (Lifeline Cell Technolgy; LS-1048). When reaching a confluence of about 80% the cells were detached by incubating with accutase (StemPro Accutase, Invitrogen, A11105-01) for 5-10 min at 37°C. After transferring the detached cells to a fresh tube and diluting the accutase with IMDM (Invitrogen, 21980-032) medium, the cells were pelleted at 450 x*g* for 5 min and seeded in fresh cell culture flasks. 24 h prior to biotinylation and cell lysis for the proteomic analysis, the adherent growing cells were washed with PBS (Sigma, D8537) and cultured either in fresh RenaLife medium or in serum-free renal CSC medium (basal medium: Advanced DMEM/F12 (Invitrogen; 12634-010); supplements see Tab. 3) established by Dr. Teresa Rigo-Watermeier (HI-STEM gGmbH).

Tab. 3: Supplements for renal CSC medium added to Advanced DMEM/F12 medium

Supplement	final concentration / amount	manufacturer, product number
Bottenstein's N-2 formulation	1x	Invitrogen, 17502048
L-Glutamine	2 mM	Invitrogen, 25030-081
L-Glutathione reduced	1 µg/ml	Sigma-Aldrich, G6013
H_2O	25 ml	Invitrogen, 15230-089
D-(+)-Glucose	0.6% (w/v)	Sigma-Aldrich, G8769
Trace elements A	1x	Cellgro, 25-021-CI
Trace elements B	1x	Cellgro, 25-022-CI
Trace elements C	1x	Cellgro, 25-023-CI
HEPES for cell culture	5 mM	Invitrogen, 15630-106
Heparine	2 µg/ml	Sigma-Aldrich, H3149
Lipid Mixture-1	1x	Sigma-Aldrich, L0288
Bovine serum albumin	0.06%	Millipore, 81-068-3
Recombinant Human FGF	50 ng/ml	Peprotech, 100-18B
Recombinant Human EGF	20 ng/ml	Peprotech, AF-100-15
LONG® R3 IGF-I human (=IGFR3)	10 ng/ml	Sigma-Aldrich, I-1271
β-Mercaptoethanol	100 µM	Invitrogen, 21985-023
Hydrocortisone 21-hemisuccinate sodium	0.1 µg/ml	Sigma-Aldrich, H2270
(−)-Epinephrine (+)-bitartrate	1 µM	Sigma-Aldrich, E4375
3,3',5-Triiodo-L-thyronine sodium	10 nM	Sigma-Aldrich, T6397

4.2.2 *In vitro* biotinylation and cell lysis

For *in vitro* biotinylation the adherently growing human renal cells were washed with 10ml PBS (Invitrogen, 18912014) and incubated for 10 min at room temperature on an orbital shaker with 5 ml of biotinylation solution (0.5 mg/ml Sulfo-NHS-SS-Biotin (Proteochem, b2104) in PBS). The remaining biotinylation reagent was quenched by adding 100 µl of 1 M Tris-HCl pH 7.0. After a washing step with 10 ml PBS and 2 mg/ml L-Glutathione (ox) (GSSG)

(Serva, 23130.03) the cells were detached for about 30 min at 37°C using 10 mM EDTA (Sigma, 03609), 2 mg/ml GSSG in PBS and transferred into a 50 ml falcon. Then the cell culture flasks were flushed twice with 5 ml PBS, 2 mg/ml GSSG, the washes were combined in the same falcon tube and an aliquot was taken for cell counting (ViCell XR 2.03, Beckman Coulter). In average about 0.5 to 1.5*10^6 cells per T75 flask were obtained. After centrifugation at 450 x*g* for 5 min at 4°C the proteins were extracted by vortexing the cell pellet in 2 ml of lysis buffer (2% (w/v) Nonidet P40 (Sigma, 74385), 0.2% (w/v) SDS (Sigma, L6026), 10 mM EDTA in PBS pH 6.8, 1 tablet/50 ml cOmplete, EDTA-free Protease Inhibitor Cocktail Tablets (Roche, 04693132001), 5 mg/ml GSSG). Following an incubation of 30 min at 4°C the lysates were centrifuged in 2 ml eppendorf tubes at 21,000 x*g* for 20 min in order to remove the remaining cell debris. Afterwards the cleared supernatant was transferred in fresh eppendorf tubes and stored at -20°C. The non-biotinylated controls were processed in parallel without biotinylation reagent and supplemented GSSG. In total each of the three cell types were biotinylated at two different passages (p 5 and a mixture of p 4+5+6 (medullary epithelial cells) p 5 and p 4+6+7 (cortical epithelial cells); p 6 and p 5+7+8 (proximal tubule cells)) each in RenaLife and renal CSC medium in six to ten replicates each. In parallel two non-biotinylated replicates per sample were processed for protein concentration determination.

4.2.3 Tumour dissociation for xenograft generation

Tumours derived from xenografts were minced with a scalpel in small pieces and enzymatically digested with 1 µg/ml collagenase IV (Sigma, C5138) in CO_2-independent medium (Invitrogen, 18045-054) for 1 h on a MACSMix tube rotator (Miltenyi Biotec) at 37°C. Before and after the incubation the tissue pieces were disintegrated on a gentleMACS dissociator (Miltenyi Biotec) using gentleMACS C tubes (Miltenyi Biotec, 120-005-331) by applying the first 30 sec of the m_impTumour_01.01 program. After filtering the cell suspension via a 100 µm cell strainer (BD Biosciences, 431752) and quenching the enzyme by addition of 10 ml 10% (v/v) fetal calf serum (Invitrogen, 10270) in CO_2-independent medium, the cells were pelleted by centrifugation at 450 x*g* for 5 min at 4°C. Following a washing step with IMDM the erythrocytes were removed by a 2 min incubation at room temperature with 2 ml ACK Lysing buffer (Lonza, 10-548E). For injection into mice the cells were counted using a Neubauer chamber and the pellet was resuspended in an appropriate amount (30 µl per million cells) of 30% (v/v) matrigel (BD Biosciences, 356234) and 70% PBS (Sigma, D8537).

4.3 PROTEOMIC SAMPLE PREPARATION AND ANALYSIS METHODS

4.3.1 Protein Extraction from mouse tissues

For protein extraction snap-frozen tissues were thawed and 250 to 500 mg pieces were homogenized at level 6 (T10 standard ULTRA-TURRAX with S10N-5G, IKA), in 8 ml of lysis buffer (2% (w/v) SDS, 50 mM Tris-HCl pH 7.4, 10 mM EDTA, 1 tablet/50 ml cOmplete, EDTA-free Protease Inhibitor Cocktail Tablets (Roche) in PBS (Invitrogen, 18912014) pH 6.8) for 6 x 1 min. For improved protein extraction the samples were sonicated (Heinemann) with ninety 1 sec pulses at 35% intensity followed by a 20 min incubation at 95°C. To remove remaining cell debris the lysates were centrifuged at 21,000 x*g* for 20 min. The cleared supernatant was transferred to fresh tubes and stored at -20°C for the subsequent determination of the protein concentration as well as the concentration of the biotinylated proteins and later sample preparation for proteomic analysis.

4.3.2 Protein concentration determination

The protein concentration of the cell and tissue lysates was determined using the BCA Protein Assay Kit (Thermo Scientific Pierce; 23225) accordingly to the manufacturer's instructions for the microplate workflow. Therefore the lysates were diluted 1:5 and 1:10 in MilliQ-water and 25 µl pre-diluted lysate as well as an eight-step BSA standard (Thermo Scientific Pierce; 23209) ranging for 2 to 0.025 mg/ml was mixed in triplicates with 200 µl BCA reagent (50 parts solution A with 1 part solution B) in a 96-well plate (TPP, 92096). After an incubation time of 30 min at 37°C on an orbital shaker at 40 rpm the OD at 562 nm was measured by an ELISA plate reader (SpectraMaxM5 with SoftMaxPro V5.4 software, MolecularDevices) and the protein concentrations were determined using Excel (Microsoft).

4.3.3 Determination of biotinylated protein concentration

The amount of biotinylated protein in lysates of perfused tissues was determined by ELISA. Therefore 100µl of 53 µg /ml (corresponding with 10^{-6} M based on an average molecular weight of a protein of 53,000 g/mol) total tissue lysates diluted in PBS (Invitrogen, 18912014) was coated over night at 4°C on a Maxisorb ELISA plate (Nunc, 442404). After three washing steps with PBS the remaining binding sites were blocked with PBS-B (3% (w/v) BSA (Sigma, A2153) in PBS) for 1 h at room temperature. Following another three PBS washing steps and a 1h incubation at room temperature with streptavidin-HRP (GE-Healthcare, RPN1051) 1:1000 in PBS-B and additional washing with three times PBS-Tween (0.1% (v/v) Tween-20 (Sigma, P1379) in PBS) and three times PBS the biotinylated proteins were detected by adding 100 µl of soluble 3,3'-5,5'-Tetramethylbenzidine (TMB) (BM Blue Pod, Roche, 11484281001)

detection solution. The reaction was stopped after 10 to 15 min by adding 70 µl of 1 M H_2SO_4 (Sigma, 258105). The ODs at 450 nm and 690 nm were measured using an ELISA plate reader and the OD_{690} for the plastic of the plate was subtracted from the OD_{450} for further analysis. The amount of biotinylated protein was determined based on an equally processed BSA standard containing 66.5 µg/ml BSA with biotinylated BSA contents (biotinylated with 10x molar excess of Sulfo-NHS-LC-Biotin and washed by 1:27,000 dilution with MilliQ water using a VivaSpin column 15R (Sartorius AG, VS15RH12)) from 2.5 to 0 µg/ml in eight steps each in triplicates using Excel (Microsoft).

The mouse kidney lysates contained between 41 to 60 µg biotinylated proteins per mg of total protein lysates whereas the xenograft kidney tumours ranged from 2 to 77 µg biotinylated protein per mg total protein lysates. For the mouse lung lysates and xenograft lung metastasis samples values between 8 to 54 µg and 2 to 64 µg biotinylated protein per mg total protein lysates were measured.

4.3.4 Processing of biotinylated tissue lysates for proteomic analysis

The enrichment of biotinylated proteins for proteomic analysis was performed accordingly to [7]. Total protein lysates from perfused tissues (chapter 4.3.1) with the equivalent of 43.3 µg biotinylated proteins (determined by ELISA (chapter 4.3.3)) were mixed with 100 µl streptavidin-sepharose slurry (GE Healthcare, 17-5113-01), which was prior to that washed twice with 500 µl of buffer A (1% (v/v) Nonidet P 40 Substitute, 0.1% (w/v) SDS in PBS (Invitrogen, 18912014)). Each washing step consisted of a resuspension and a centrifugation step (30 sec, 2000 x*g*, room temperature). To capture and reduce the biotinylated proteins on the sepharose resin the samples were incubated on a revolving mixer with 2% (w/v) SDS and 2 mM TCEP (Biovision, 1202) in buffer A for 2 h at room temperature. Subsequently, the supernatant was removed and the resin was transferred to a PVDF centrifugal filter unit (Millipore, UFC30SV00) where it was washed twice with 500 µl buffer A. For the following alkylation step the resin was washed twice with 150 µl 50 mM iodoacetamide (Sigma, I1149) in PBS, transferred with 300 µl iodoacetamide solution to a fresh tube and incubated in the dark for 30 min at room temperature. After stopping the alkylation process by incubating the sample with 300 µl of 100 mM L-cysteine (Sigma, 30089) for 15 min, the sepharose resin was mixed with 400 µl of buffer A and 800 µl of delipidation solution (60% (v/v) diisopropylethylamine (Sigma, 496219), 40% (v/v) butanol (Th.Geyer, 2513.1000)) and incubated for 30 min at room temperature on a revolving mixer. Following three centrifugation steps at room temperature for 2 min at 2000 x*g* the organic top layer was removed and the sepharose pellet was transferred to a centrifugal filter unit and washed twice with 500 µl buffer A. After 10 washes with 500 µl trypsin digestion buffer (TDB) (50 mM Tris-

HCl, pH 8.0, 1 mM $CaCl_2$ (Sigma, C4901) in MilliQ water) the resin was transferred to a fresh tube and finally resuspended in 200 µl TDB with 10 µl of sequencing grade modified trypsin (stock solution of 80 ng/µl in TDB; Promega, V5111). Protease digestion was carried out for 10 h at 37°C and 1000 rpm shaking. Afterwards the peptides were separated from the sepharose resin by a centrifugal filter unit and desalted using C18 OMIX tips (Agilent, A57003100K) accordingly to the manufacturer's guidelines. Therefore a C18 resin was prepared with each two 100 µl washes of wetting (50% (v/v) acetonitrile (Biosolve, 12041) in H_2O (ULPC-grade, Biosolve, 232141) and washing solution (0.1% (v/v) trifluoroacetic acid (TFA) (Proteochem, LC6203) in H_2O (ULPC-grade)) and the peptide mixture, acidified with 0.1% (v/v) TFA, was bound by a 20 times pipetting step on the resin. After three washes with each 100 µl of washing solution, the peptides were eluted in elution solution (75% (v/v) acetonitrile, 0.1% (v/v) TFA in water (UPLC grade)) by pipetting 10 times up and down, lyophilized and stored at –20°C. In general, every tissue lysates was prepared in four individual technical replicates.

4.3.5 Processing of biotinylated cell lysates for proteomic analysis

The preparation for proteomic analysis of cell lysates containing proteins biotinylated with Sulfo-NHS-SS-biotin (chapter 4.2.2) was performed analogue to the preparation of lysates labelled with Sulfo-NHS-LC-biotin as described in 4.3.4. To avoid the reduction of the disulfide linker of the biotinylation reagent 3 mg/ml L-Glutathione (ox) (GSSG) was added in every step and the reduction and subsequent alkylation step was postponed after the tryptic digest. Briefly, 100 µl streptavidin-sepharose was washed thrice with 500 µl buffer A containing 3 mg /ml GSSG) and combined with the protein lysates (0.4 mg total protein; determined by BCA-assay from non-biotinylated reference sample), 2% (w/v) SDS and 3 mg/ml GSSG. After 2 h incubation on a revolving mixer at room temperature, the supernatant was removed and the sepharose was incubated for 30min on a revolving mixer in 400 µl buffer A, 800 µl of delipidation solution (60% (v/v) diisopropylethylamine, 40% (v/v) butanol) and 3 mg/ml GSSG at room temperature. Following three centrifugation steps at room temperature for 2 min at 2000 x*g* the organic top layer was removed and the sepharose pellet was transferred to a centrifugal filter unit and washed twice with 500 µl buffer A containing 3 mg/ml GSSG. After two washes with 500 µl buffer B (0.1% (w/v) SDS, 2 M NaCl (Sigma, 31434) in PBS, 40°C) containing 3mg /ml GSSG and ten washes with 500 µl trypsin digestion buffer (TDB) containing 3 mg/ml GSSG the resin was transferred to a fresh tube and finally resuspended in 200 µl TDB (no GSSG added) with 10 µl of sequencing grade modified trypsin (stock solution of 80 ng/µl in TDB). Protease digestion was carried out for 10 h at 37°C and 1000 rpm shaking. Subsequently, the peptides were reduced and alkylated by incubating the slurry with 2 mM TCEP at 37°C followed by a 20 mM iodoacetamide incubation at room temperature, each for

30 min at 1200 rpm shaking. Afterwards the peptides were separated from the sepharose resin by a centrifugal filter unit and desalted using C18 OMIX tips accordingly to the manufacturer's guidelines and described in chapter 4.3.4.

4.3.6 Nanocapillary reverse-phase UPLC and fractioning on MALDI target plates

Lyophilized peptides derived from the digestion of affinity purified biotinylated cell / tissue lysates were dissolved in 11 µl of UPLC dissolving solution (5% (v/v) acetonitrile (ACN), 0.1% (v/v) TFA and water (UPLC grade) by sonication for 5 min and transferred into a total recovery UPLC vial (Waters, 186000385C). The peptides were separated on an ACQUITY UPLC BEH C18 column (inner diameter: 75 µm; length: 25 cm; filled with: 1.7 µm silica particles; pore size: 130 Å) (Waters, 186003545) which was attached to a nanoACQUITY UPLC System (Waters) controlled by the software programs MassLynx V4.1 (Waters). The mobile phase A consisted of 0.1% (v/v) TFA in UPLC water whereas the mobile phase B contained 0.1% (v/v) TFA in ACN. After equilibrating the column with 95% A and 5% B and a peptide trapping time of 40 min the peptides were eluted by an applied flow rate of 350 nl/min, using a 110 min gradient ranging from 5 to 85% phase B. A diagram and a table of the elution gradient are displayed in Fig. 18. The eluting peptides were directly mixed with 3 mg/ml a-cyano-4-hydroxycinnamic acid (Proteochem, p9100) in 80% ACN, 0.1% TFA in UPLC water spiked with four internal standard peptides (sequences and concentrations in Tab. 4; PSL Peptide) and spotted 10 min post the start of the elution gradient in 1200 fractions on a blank Opti-TOF LC/MALDI target plate (AB Sciex, 1018497) using the online coupled Micro Fraction Collector MALDI Spotter (Sunchrom) controlled by the software Suncollect (Sunchrom). The flow rate of the matrix was set to 2 µl/min and a fraction was spotted every 4 sec. The final concentration of each internal standard peptide was therefore between 3.6 and 26.5 fmol per spot (refer to Tab. 4).

Tab. 4: Sequences, molecular weights and concentrations of internal standard peptides

Peptide sequence	Molecular weights (calculated (MH+)) [g/mol]	concentration in CHCA matrix [fmol/µl]	amount per spot [fmol]
CLEHMYHDLGLVRDF	1846.8732	107.2	14.3
EEQPSTPAPKVEQQEEILC	2155.0231	214.3	28.5
TGVFDEAIRTVGF	1411.7221	53.6	7.2
TVFDEAIR	951.0688	26.8	3.6

Time [min]	% Buffer A	% Buffer B	Buffer B change
0	95	5	prompt
0.33 - 1	89	11	linear
1 - 3.66	89	14	linear
3.66 - 67.66	70	30	linear
67.66 - 81	60	40	linear
81 - 90	15	85	linear
90 - 95	15	85	no change
95 - 97	95	5	linear
97 - 110	95	5	no change

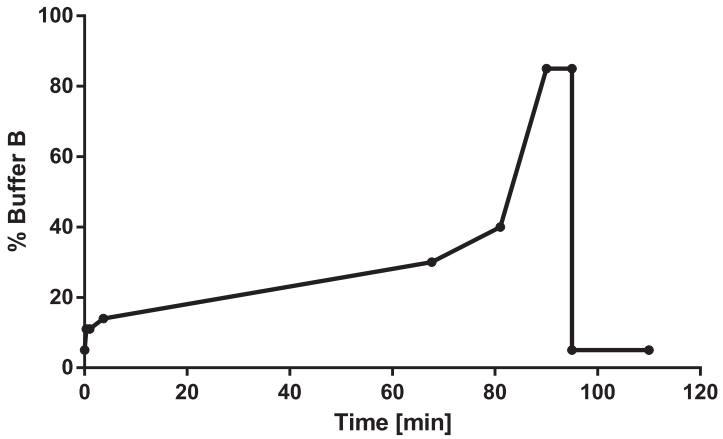


Fig. 18: Gradient applied on nanocapillary reverse-phase UPLC Buffer A: 0.1% (v/v) TFA in water (UPLC grade); Buffer B: 0.1% (v/v) TFA in acetonitrile

4.3.7 MALDI-TOF/TOF mass spectrometry

All measurements were performed on a MALDI-TOF/TOF 5800 instrument (AB Sciex) under the control of the TOF/TOF Series explorer software (AB Sciex). In the primarily measured MS reflector mode eight subspectra with each 250 laser shots were acquired per spotted fraction and averaged for the subsequent calculation of MS/MS precursor selection. Up to 35 precursors per spot with a minimum signal to noise ratio of 50 and a mass range from 750 to 4000 Da were selected and fragmented for MS/MS analysis by CID in the reverse order from weak to high signal to noise ratios. The MS/MS measurements were performed in the dynamic exit mode where the acquisition of every spectrum was stopped when the spectra quality criteria reached “high” (minimum of 250 acquired subspectra) or a total number of 3000 subspectra were measured. For ionization in all measurements a solid-state laser (OptiBeam on-Axis, AB Sciex) at 345 nm with a repetition rate of 0.4 kHz in MS and 1 kHz in MS/MS was used. Calibration prior to every measurement was performed using a calibration standard kit (ProteoChem, s6104) with an additional peptide from ACTH (PSL Peptide). The peptide

sequences with their amount per spot are listed in Tab. 5. All samples were measured in duplicates. For mouse kidney and RCC24 kidney tumour samples the measurement was carried out in quadruplicate.

Tab. 5: Sequences, molecular weights and concentrations of peptides from calibration standard

Peptide sequence	Molecular weights [g/mol]	amount per spot [pmol]
YGGFL	556.277	3.2
pEHWSYGLRPG	1182.581	1.5
DRVYIHPFHL	1296.685	1.3
pELYENKPRRPYIL	1690.928	1.1
RPVKVYPNGAEDESAEAFPLEF	2465.199	0.7

4.3.8 MALDI-TOF/TOF data analysis

From the resulting data measured in the MS reflector mode peak and spectra information was extracted using the Data Explorer software V4.10.0.124 (Applied Biosystems) and a 2D-peptide-map visualising all fractions with their corresponding m/z spectra intensities in one image was generated using 2D PepMap software V1.6 (Dr. Tim Fugmann, ETH Zuerich (unpublished)). All 2D-peptides-maps shown in this thesis are normalized to a standard peptide peak at 1411 m/z and a standard peak height of 1000. Peak heights were capped at an intensity of 500.

The spectral data obtained by MS/MS acquisition and attendant MS data was processed by the ProteinPilot Software V4.5 (AB Sciex) which matches the acquired spectra to databases containing *in silco* digested proteins using the Paragon algorithm [613]. A confidence interval of greater 95% was allowed. All modifications listed in the modus “biological modifications” from ProteinPilot together with the in the sample preparation workflow introduced alkylation of cysteines with iodoacetamide were considered. The MS/MS was searched against the reviewed swiss prot and unreviewed trembl databases from *homo sapiens* and *mus musculus* downloaded from the Uniprot-webpage (http:\\www.uniprot.org). For the data from human renal cells a human database downloaded on the 30/7/2012 consisting of 70 101 entries, for the data from mouse tissue the mouse database with 51 514 entries downloaded on 8/8/2013 and for data from xenografts a combined database from human and mouse proteins both downloaded on 10/9/2012 containing 70 101 respectively 51 514 entries was used.

In a next step the peptide summaries obtained from ProteinPilot-analysis were filtered for proteotypic peptides and their corresponding proteins using the in house developed PepSir

software (version 1.5; 28/5/2014; Alexander Kerner (unpublished)). Only peptides with a sequence confidence > 95% were considered. The mixed human and mouse database containing reviewed swissprot and unreviewed tremble entries downloaded from the UniProt-webpage (http:\\www.uniprot.org) on the 12/5/2014 and containing 88 721 human and 51 389 mouse entries was underlying the filter process. To select the potential biomarker candidates the resulting protein lists of human renal cells, mouse kidneys, mouse lungs, xenograft kidney tumours and xenograft lung metastasis were mapped against each other using Excel (Microsoft) and several proteins for which peptides were observed exclusively in the diseased tissues were selected for further validation. Clustering analysis was performed on all proteins identified with at least one proteotypic peptide by the software GeneCluster3.0 (Standford University, USA) using the Kendall's τ algorithm. Tree visualisation was done by the Java treeview software (Standford University, USA). For gene ontology analysis upregulated proteins (determined by MS_QBAT analysis, see chapter 4.3.9) in the respective dataset were submitted to the WEB-based GEne SeT AnaLysis Toolkit (Webgestalt) [614]. Analysis was based on entries of the uniprot database.

4.3.9 Relative quantification by MS_QBAT software

For relative quantification of protein abundances a sample subset containing each four healthy mouse kidney and lung samples, four of each patient derived xenograft kidney tumour sample (RCC12, RCC15, RCC18, RCC24, RCC27, RCC38, RCC50) and xenograft lung metastases sample (RCC12, RCC15, RCC27, RCC50) as well as four technical replicates of each human renal cell line was selected. The relative quantification of the protein abundances was performed using the in house developed MS_QBAT software version 17.3 (Alexander Kerner (unpublished)). Therefore, relevant peak-files containing fraction, m/z ratio and intensity information obtained from Data Explorer analysis (chapter 4.3.8) were loaded together with the corresponding proteotypic peptide summaries obtained from PepSir analysis (chapter 4.3.8) into MS_QBAT using a signal to noise ratio of > 25 and a minimal fraction length of 2. The individual peptide signals were normalized to internal standard peptide intensities contained in each sample and each sample group was normalized to a maximum signal intensity of 300 mio. Afterwards the four replicates of a sample group were aligned to a super sample using the alignment parameters determined as optimal by MS_QBAT. Afterwards samples were quantified by comparison of different super samples sets. Features with fold change above two were displayed in a volcano plot as regulated and a p-value < 0.05 was considered as significant.

4.4 BIOMARKER VALIDATION METHODS

4.4.1 RNA isolation

For RNA extraction the RNeasy Mini-Kit (Qiagen, 74104) was applied accordingly to the manufacturer's guidelines. Therefore, 10-30 mg tissue or accutase-detached *in vitro* cultured cells (see chapter 4.2.1) of one T75 flask were lysed in 600 µl of guanidine-isothiocyanate RLT lysis buffer (Qiagen, 74104) with 10 µl β-Mercaptoethanol (Sigma, 63689) by mechanical disruption using a syringe with 21 G needle. After a centrifugation of 3 min at 20,000 x*g* at room temperature, 600 µl of 70% ethanol (Sigma, 32205) were added to the cleared cell/tissue lysate and immediately mixed. For RNA purification the mixture was placed on a spin column for the RNeasy Kit and centrifuged for 15 sec at 8000 x*g*. Subsequently, a washing step with 350 µl RW1 buffer (Qiagen, 74104) and a DNase treatment with 10 µl DNase stock and 70 µl RDD buffer (both Qiagen, 79254) for 20 min at room temperature was performed. Following washing steps with 700 µl RW1 buffer (Qiagen, 74104) and twice 500 µl RPE buffer (Qiagen, 74104) the column was centrifuged for 2 min at 8000 x*g* and the RNA was eluted with 50 µl water (Qiagen, 74104) after spinning with 8000 x*g* for 1 min. The RNA quantity was determined finally using the NanoDrop 1000 spectrometer and ND-100 V3.8.1 software (both Peqlab).

4.4.2 cDNA synthesis

cDNA synthesis from RNA (chapter 4.4.1)was performed with the SuperScript VIL cDNA Synthesis Kit (Invitrogen, 11754050) according to the pipetting scheme displayed in Tab. 6. The PCR was executed during the following program (T3000 Thermocycler, biometra): After 10 min at 25°C the transcription and amplification was carried out for 2 h at 42°C. Then the reaction was terminated with 85°C for 5 min and the program finished with setting permanently 4°C. After the PCR the cDNA was stored at -20°C until use.

Tab. 6: Pipetting scheme for reverse transcription of mRNA to cDNA

Reagent	Quantity
extracted RNA	2 µg
5x VILO Reaction Mix	8 µl
10x SuperScript III Enzyme Blend	4 µl
add H_2O	to 40 µl

4.4.3 Real-time quantitative PCR (RT-qPCR)

For determining mRNA expression levels by RT-qPCR 5 ng of cDNA (chapter 4.4.2) were mixed with 5 µl SBYR green mix (Applied Biosystems, 4309155) and each 0.5 µl of 10 µM forward and reverse primer (Sigma, desalted) in a total volume of 10 µl in a 384-well plate (Thermo Scientific, AB-1384), sealed with foil (MicroAmp, Applied Biosystems, 4311971) and centrifuged at 450 g for 5 min. The RT-qPCR was performed on a ViiA7 Real-Time PCR system (Applied biosystems) controlled by the software ViiA7 V1.1 (Applied Biosystems) executing the program depicted in Tab. 8. The primer pairs used are designed to be species-specific and if possible intron-spanning. The sequences are shown in Tab. 7. For analysis obtained C_T-values (< 33 (mouse), < 31 (human)) were normalized based on expression levels of *TFRC* (human) and *Oaz1* (mouse) and relative fold changes were calculated based on the average expression in five healthy mouse kidney samples for mouse genes and three human renal cell lines for human genes using the DataAssist V3.01 software (Applied Biosystems). Visualization was performed with Prism 6 (Graphpad).

Tab. 7: Primer pairs for RT-qPCR

Gene	Species	Sequence (forward/reverse) 5' → 3'
IGFBP3	human	AACGCTAGTGCCGTCAGC CGGTCTTCCTCCGACTCAC
LCN2	human	CAGCAGAACTTCCAGGACAAC TTGCGGGTCTTTGTCTTCTC
Lcn2	mouse	CCCCATCTCTGCTCACTGTC CATCGTAAAGCTGCCTTCTG
LTBP2	human	ACCTTGGCCAGAGCACAG TCGGGTGAAGTCGGACAG
Ltbp2	mouse	ACCATCACCACCTCCATCTC TCGAGTCCTGAAGGCAAAG
Oaz1	mouse	TTTCAGCTAGCATCCTGTACTCC GACCCTGGTCTTGTCGTTAGA
TFRC	human	GTTTCTGCCAGCCCACTG CCAGTAACCGGATGCTTCAC
TGFBI	human	GTGGCTTGGAGGCTTTTATG CCATGGCTCTGTCACAATAGG
TNXB	human	ACAGTCGGTTGTGGCTCAG GCCTCTGCACCTGCTCTGTC

Tab. 8 Program utilised for RT-qPCR

Temperature [°C]	time [min:sec]	cycles	temperature adaptation speed [°C/sec]
50	2:00	1	-
95	10:00	1	1.6
95	0:15	40	1.6
60	1:00	40	1.6
95	0:15	1	1.6
60	1:00	1	1.6
95	0:15	1	0.05

4.4.4 Immunofluorescence analysis

8 µm sections from freshly frozen tissues embedded in OCT-medium were fixed for 8 min in ice-cold (-20°C) acetone (Carl-Roth, 9372.5), surrounded by a silcon ring (DAKO, S200230) and non-specific binding residues were blocked for 1 h at room temperature with 20% (v/v) donkey serum (Th.Geyer, BW/S2170/000100) in PBS (Invitrogen, 18912014) followed by 1 h Fc-blocking with a goat α-mouse IgG (H+L) Fab fragment (Jackson, 115-007-003) (0.13 mg/ml in 12% (w/v) BSA in PBS). The Fc-block was only performed for sections stained later with a primary antibody (here α-human TNXB (Novus Biologicals, NB500-513)) originating from mouse. Following 5 min of washing in PBS the sections were incubated with primary antibody solution (primary antibodies see Tab. 9, in 12% BSA in PBS) for 1 h at room temperature. For negative controls the primary antibody was omitted. Subsequently the sections were washed thrice for 5 min with PBS and secondary antibody solution (secondary antibodies see Tab. 10, DAPI (Sigma, D9542) 25 µg/ml in 12% (w/v) BSA in PBS) was applied for 1 h at room temperature for detection. Afterwards the sections were again triply washed for 5 min with PBS, dried and finally mounted with fluorescence mounting medium (DAKO, S302380). The sections were analysed with a confocal laser scanning microscope (LSM 700, Zeiss).

Tab. 9: Primary antibodies used for immunofluorescence analysis

Antigen	Species specificity	Host, Isotype	Manufacturer Product number	applied final concentration [µg/ml]
Tenascin X	human	mouse, IgG1	Novus Biologicals NB500-513	5
LTBP2	human/mouse/rat	rabbit, IgG	Proteintech 17708-1-AP	3
CD31	mouse	rat, IgG_{2a}, κ	Becton Dickinson 550274	0.078

Tab. 10: Fluorescent antibody conjugates used for immunofluorescence analysis

Antigen	Fluorophore	Host, Isotype	manufacturer product number	applied final concentration [µg/ml]
Mouse IgG (H+L)	Alexa Fluor 488	donkey, $F(ab')_2$	Abcam ab150101	8
Rabbit IgG (H+L)	Alexa Fluor 488	donkey, $F(ab')_2$	Jackson ImmunoResearch 711-546-152	2.4
Fluorescein	DyLight 488	goat, IgG	Novus Biologicals NBP1-97375	6.7
Rat IgG (H+L)	Alexa Fluor 594	goat, IgG	Invitrogen A-11007	8
Biotin	Alexa Fluor 488	-, - (streptavidin)	Invitrogen S-32354	8

4.4.5 *In vivo* biodistribution and antibody penetration

To analyse the *in vivo* targeting and biodistribution of the commercially available α-Tenascin-X antibody, the antibody was labelled with Fluorescein and injected into tumour-bearing mice. Briefly, prior to labelling the Tris component of the α-TNXB-antibody storage buffer (TBS including 50 mM Tris) was diluted 1:4500 with PBS (Invitrogen, 18912014) and the antibody solution was concentrated to a final concentration of 1 mg/ml α-human Tenascin X antibody using an Amicon Ultra 30K centrifugation tube (Millipore, UFC503096) according to the manufacturer's instructions. For labelling the antibody solution was incubated with 20-fold

molar excess of NHS-Fluorescein (Thermo Scientific, 46409) dissolved in dry DMSO (Sigma, 41648) for 1 h at room temperature in the dark. The reaction was quenched by a 2000-fold molar excess of Tris-HCl to the labelling reagent. 50 µg of the labelled α-TNXB-antibody in 100 µl PBS were injected into the tail vein of kidney tumour- and lung metastasis-bearing ccRCC xenograft mice described in chapter 4.1.2. As background control one healthy NSG mouse was injected with 100 µl labelling solution containing no antibodies. 48 h after injection the mice were sacrificed and kidney tumour, lung metastasis, kidney, liver, spleen, intestine and brain were excised and embedded in OCT-medium. Subsequently, 8 µm sections of the embedded tissues were prepared for immunofluorescence analysis as described in chapter 4.4.4. Staining was performed with an α-Fluorescein-DyLight488 antibody (Novus Biologicals, NBP1-97375). For visualizing the endothelial cells a consecutive section per tissue was stained with α-mouse CD31 antibody (Becton Dickinson, 550274) and detected with an α-rat IgG-Alexa Fluor 594 antibody (Invitrogen, A-11007) (antibodies details see Tab. 9 and Tab. 10).

5 RESULTS

Effective and well-tolerated therapies for the treatment of advanced clear cell renal carcinoma are urgently needed in order to improve patient survival. The antibody-based delivery of bioactive molecules directly into the tumour has proven to be a promising approach for the development of such a therapeutic [4]. Importantly, the availability of vascular accessible tumour-associated biomarkers as target molecules is a prerequisite for the design of this innovative drug type [5]. Several methods (elucidated in the chapter “Intro”) for the identification of such biomarkers are available. In the following sections, we present the results of a biomarker discovery study for clear cell renal cell carcinoma primary tumours and their lung metastases using the in vivo biotinylation approach. Finally, initial validation experiments for six potential drug target candidates are shown.

5.1 MATERIAL SELECTION

In order to recapitulate the characteristics of human clear cell renal cell carcinoma in mice, seven different patient derived orthotopic xenograft mouse models, which were established by Dr. Teresa Rigo-Watermeier (HI-STEM gGmbH, data not published), were used in this study. Patient derived tumour material were dissociated into single cells by enzymatic digestion and injected together with matrigel into the kidney capsule of female nod scid common gamma mice to generate the model (Fig. 19A). Each model, which represented the tumour of a distinct patient, could be propagated in mice for several passages as described for model establishment with the initial patient samples. Depending on the patient derived material the xenograft tumours were grown for three to 40 weeks in order to reach the size required for further processing (max. 1.4 cm). Tumour growth was monitored by magnetic resonance tomography (Fig. 19B). In this study seven different patient derived mouse models were analysed. RCC12, RCC18, RCC27 and RCC50 originated from primary tumours staged with T3 while RCC38 derived from a tumour with stage T4. RCC15 and RCC24 originated from lymph node metastases of clear cell renal cell carcinoma patients.

Unpublished studies conducted by Dr. Teresa Rigo-Watermeier showed that the xenograft tumors obtained were histopathologically identical to the primary material obtained from the patient (Fig. 19C). Importantly, the tumours were highly vascularised and four out of the seven mouse models (RCC12, RCC15, RCC27 and RCC50) gave rise to spontaneous lung metastases.

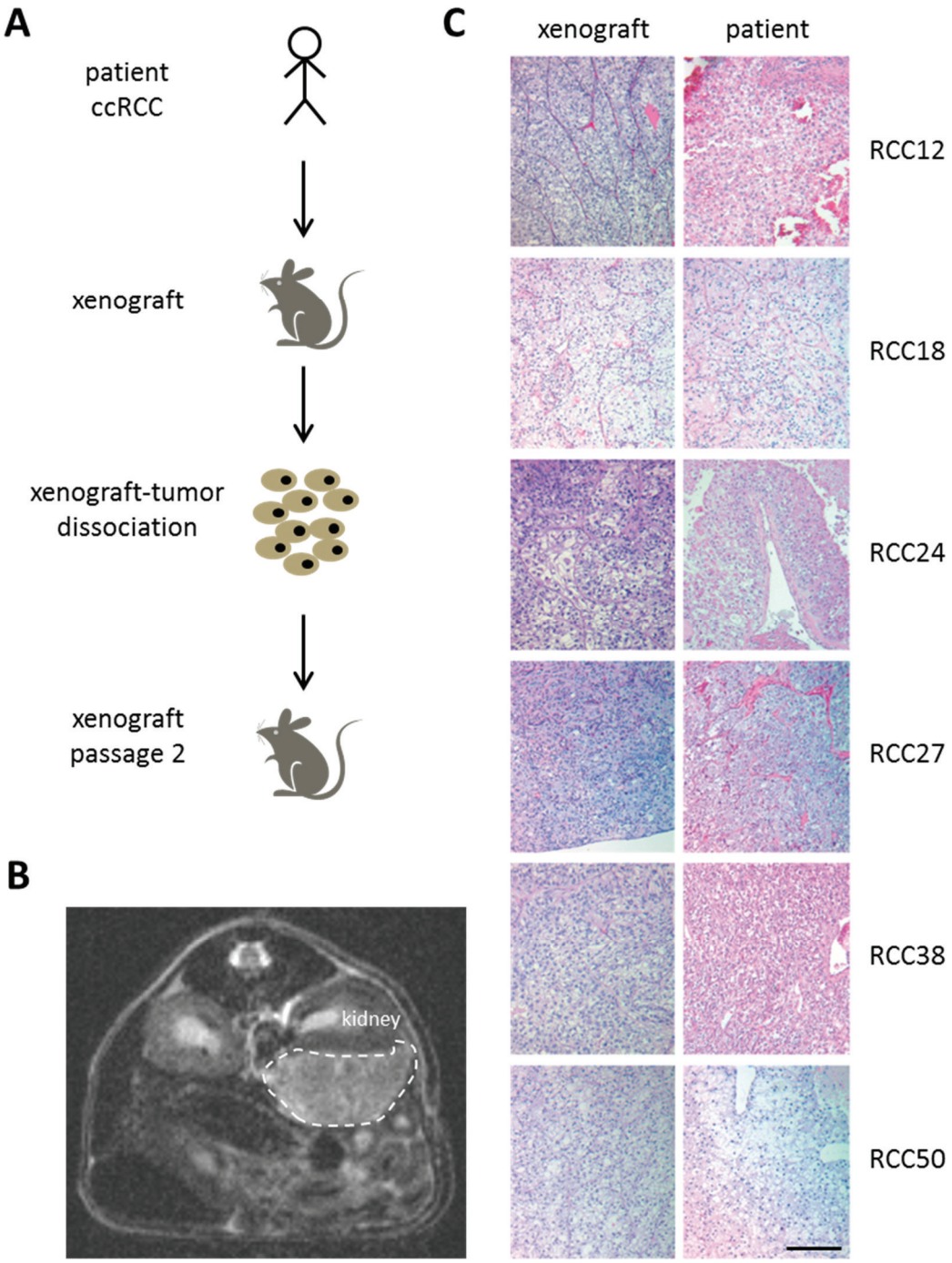


Fig. 19: Patient derived xenograft mouse models for clear cell renal cell carcinoma A: Experimental workflows for the establishment of patient-derived kidney tumor xenografts and for the subsequent passaging. B: Representative magnetic resonance tomography image of a tumour-bearing mouse (T1 scan). The kidney tumour is surrounded by a dashed line. Imaging was performed in collaboration with the "Small Animal Imaging Facility" (German Cancer Research Center, Heidelberg). C: Representative immunohistochemical hematoxylin & eosin stainings (H&E) of patient tumours and matched patient-derived xenografts (Scale bar: 100 µm). Stainings were performed by the Department of Pathology (University Clinic Heidelberg) and imaged by Dr. Teresa Rigo-Watermeier (HI-STEM gGmbH).

The three commercially available primary renal cell lines (medullary epithelial cells, cortical epithelial cells and proximal tubule cells were used as healthy control representing the human

kidney. The adherently growing primary cells were cultured in commercially available RenaLife medium. One half of the cultured cells was incubated with serum-free renal CSC medium (Dr. Teresa Rigo-Watermeier, HI-STEM gGbmH, unpublished data) for 24 h in order to remove media effects in the subsequent analysis.

Normal murine kidneys and lungs were analysed in parallel for the identification of protein expression changes in the diseased organs.

5.2 PROTEOMICS WORKFLOW

5.2.1 *In vivo* biotinylation of xenograft tumours

Molecular targets for antibody-based therapeutic approaches for cancer treatment require both a cancer specific expression profile of the target markers and its accessibility from the vasculature by the targeting antibody. Therefore, potential target proteins should either be expressed on the surface of endothelial cells or in close proximity to the vasculature. As outlined in the introduction, the *in vivo* biotinylation methodology, which was applied in the discovery approach conducted in this thesis, is one of the experimental approaches for the enrichment of these vascular accessible proteins [6, 7]. Thereby the thorax of terminal anesthetised mice was opened and the systemic circulation was perfused for 10 min by intra-cardiac administration of an aqueous solution of sulfo-NHS-LC-biotin (1 mg/ml), followed by the perfusion using a quenching solution in order to block unreacted biotinylation reagent (Fig. 20A, left panel). During this procedure the biotinylation reagent, which was composed of a reactive ester group, an appropriate linker and a biotin group, covalently labels all vascular accessible proteins. Following a successful perfusion experiment, biotin-tagged proteins are found on the cell surface of both endothelial cells and cells in close proximity to vascular structures as well as in the perivascular extracellular matrix [615, 616]. Fig. 20C (left panel) shows an immunofluorescence staining of a blood vessel (α-CD31) in a murine kidney perfused with sulfo-NHS-LC-biotin confirming accessibility of the reagent outlined above.

5.2.1.1 Establishment of a pulmonary circulation perfusion technique

During the body perfusion described above, the biotin-tagging of proteins is restricted to the systemic circulation. In order to label proteins in lung metastases, a novel technique for the perfusion of the pulmonary circulation was developed (Fig. 20A, right panel). For this, the abdomen of an anesthetised mouse was opened, a needle was inserted into the uncovered *vena cava* cranial or proximal of the arborisation to the kidneys and fixed with a ligature (Fig. 20B). To release the perfusion liquids from the circulation, the abdominal aorta was opened

proximal to the perfusion start point. The perfusion process was performed analogue to the body perfusion. Fig. 20C (right panel) presents a representative streptavidin-based immunofluorescence staining of a cryosection from a perfused murine lung demonstrating efficient biotinylation of proteins in lung tissue.

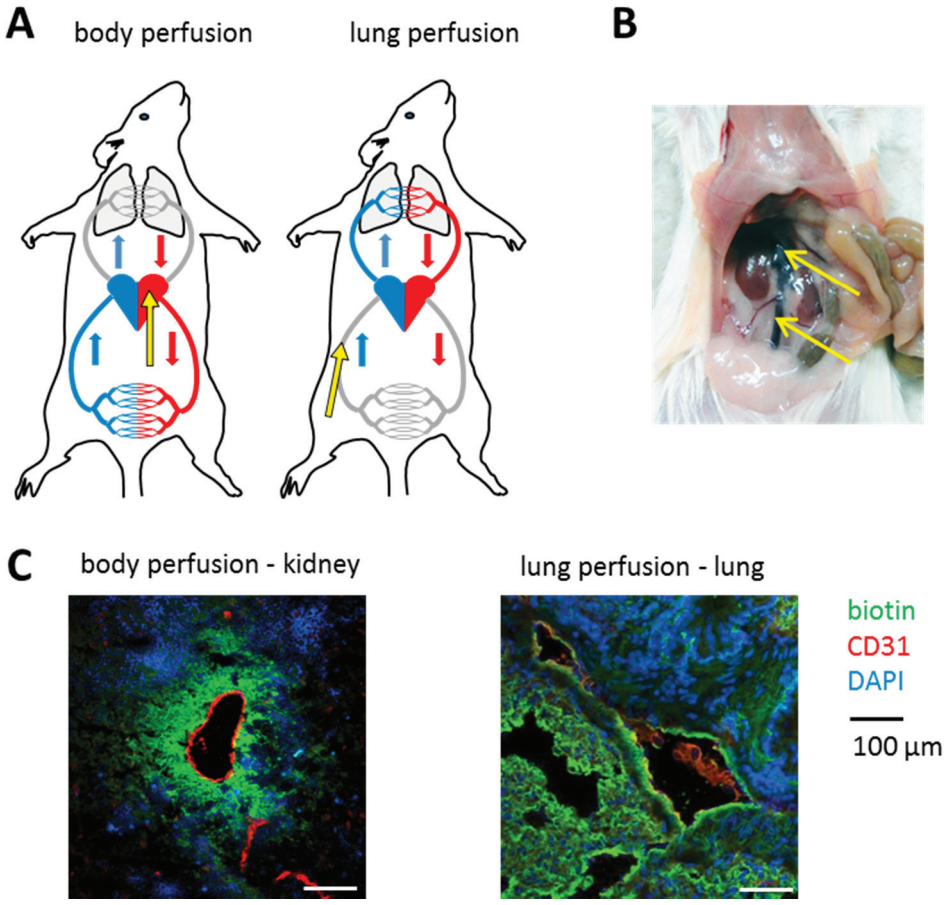


Fig. 20: In vivo biotinylation of mice A: Schematic representation of the perfused circulation and location of insertion of the perfusion needle (yellow arrow). Upon insertion of the needle into the left ventricle, the systemic circulation was perfused (body perfusion). Following the insertion of the perfusion needle into the vena caca, the pulmonary circulation was perfused (lung perfusion). B: Needle insertion position for lung perfusions. Depending on the size and the growth direction of the kidney tumours the position of the needle was either placed cranial or proximal of the arborisation to the kidneys in the vena cava (yellow arrows) C: Representative immunofluorescence images of in vivo biotinylated murine kidney (body perfusion) and lung (lung perfusion) stained with streptavidin (green), α-mouse CD31 (red) and DAPI (blue) demonstrating efficient biotinylation of vascular and perivascular structures. Scale bars: 100µm.

5.2.2 *In vitro* biotinylation of human renal cells

Cell surface proteins of *in vitro* cultured human renal primary cells were enriched using the *in vitro* biotinylation approach. This method, which is related to the *in vivo* biotinylation, is based on the incubation of *in vitro* cultured cells with a reactive ester derivate of biotin. In this study, cell surface proteins were labelled with sulfo-NHS-SS-biotin. In contrast to the reagent applied

in the *in vivo* biotinylation workflow sulfo-NHS-SS-biotin offers a lower degree of membrane permeability resulting in an increased enrichment of cell surface proteins (data not shown). Additionally, the disulfide bond introduced into the linker offers the possibility to selectively cleave the reagent.

5.2.3 Proteomics sample processing

Following the biotinylation step, proteins were extracted by tissue homogenisation or cell lysis and the protein concentration of the protein extracts was determined by BCA assay (Fig. 21). The content of biotin-tagged proteins in the total lysates was determined by ELISA using streptavidin-HRP for the detection of labelled proteins. In lysates derived from *in vivo* perfused murine kidneys, 4.1 to 6% of the total protein content was biotin tagged. In contrast, the yield of biotin-tagged proteins in *in vivo* perfused xenograft kidney tumour lysates was ranging from 0.2 to 7.7% and showed therefore a wider spreading. This finding was independent of the perfusion efficacy, which was evaluated by eye based on the colour change of the contralateral kidney. For mouse lung lysates and xenograft lung metastasis samples values between 0.8 - 5.4% and 0.2 – 6.4% were determined, respectively.

In order to standardize the protein amount for the subsequent discovery experiment, each analysis was performed with 43.3 µg biotin-tagged protein. Following the selective enrichment of biotin-tagged proteins on a streptavidin sepharose resin, disulfide bonds were reduced and alkylated with TCEP and iodoacetamide, respectively. Cysteine modification is performed in order to enhance protein unfolding and thereby to improve tryptic digestion resulting in higher sequence coverages. Following a delipidation step and extensive washings, captured proteins were digested on resin with trypsin. Resulting peptides were separated by a nano-capillary reverse-phase UPLC, mixed with matrix containing four standard peptides for later quantification and spotted in 1200 fractions on MALDI target plates. Subsequently, MALDI-MS and MALDI-MS/MS mass spectrometric analysis was performed. Resulting spectra were searched for protein identification against a chimeric database of *mus musculus* and *homo sapiens* derived from UniProt using the Paragon algorithm in Protein Pilot. Relative protein quantification was performed on a representative sample subset using the in house developed software MS_QBAT (Alexander Kerner, HI-STEM gGmbH, unpublished). MS_QBAT performs a pair-wise comparison of the averaged signal intensities of multiple tryptic peptides corresponding to the same protein in two different samples. Prior to quantification, signal intensities are normalised in each sample to internal standards, which were spiked into the sample after UPLC separation.

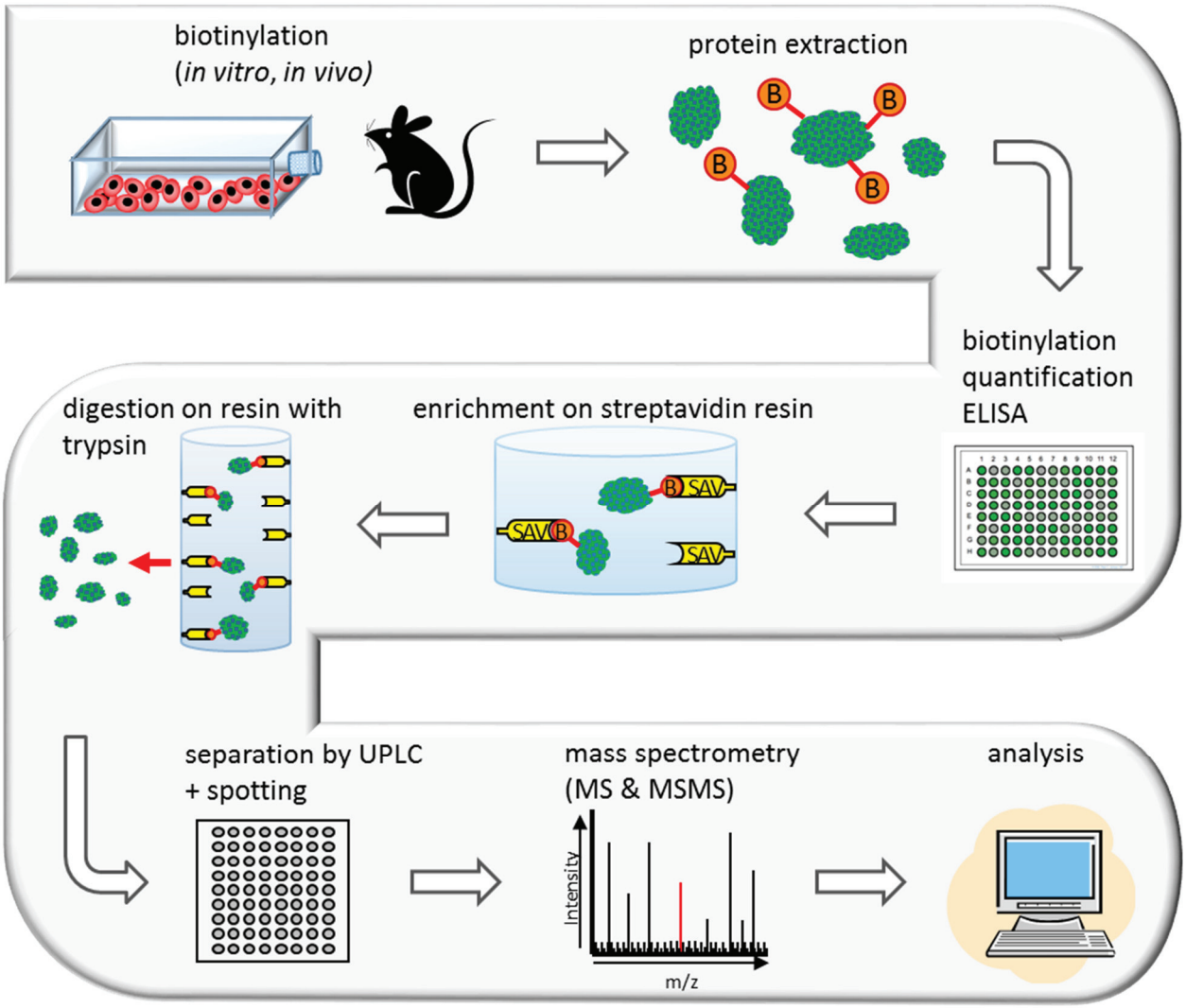


Fig. 21: Proteomics workflow applied for biomarker discovery Cell surface proteins of *in vitro* cultured cells or vascular accessible proteins of tumour-bearing mice were labelled with a reactive ester derivate of biotin. Processed cells/tissues were homogenised and proteins were extracted. The content of biotinylated proteins in the total protein extracts was determined by ELISA. Subsequently, biotin-tagged proteins were enriched on streptavidin sepharose and disulfide bonds were reduced and alkylated using TCEP and iodoacetamide, respectively. Following a delipidation step, captured proteins were digested on resin with trypsin. Resulting peptides were separated by nano-capillary reverse-phase UPLC, mixed with matrix containing a set of internal standard peptides and and spotted in 1200 fraction on a MALDI target plate. Following mass spectrometric analysis the relative abundances of identified proteins were quantified by the in house developed software MSQBAT (Alexander Kerner, HI-STEM gGmbH, unpublished).

5.3 Proteomics Analysis

Five different datasets were generated in the course of the biomarker discovery workflow described in the previous chapter. In order to allow for the comparison of human proteins expressed in the xenograft tumours to a healthy state, a dataset comprising three human renal cell lines was generated. Each of the three cell lines was analysed at two different passages and with two different culture conditions, in each case in two technical replicates (3 cell lines x 2 passages x 2 culture conditions x 2 replicates = 24 samples). Both healthy *in vivo*

biotinylated murine kidney and lungs have been analysed to obtain a dataset comprising vascular accessible proteins in healthy tissue (10 individual kidneys x 4 replicates = 39 samples; 5 individual lungs x 2 replicates = 10 samples). In order to cope with heterogenity due to patient-specific protein expression, a total of seven different, patient-specific xenograft models was analysed. Depending on the model, three to ten biological replicates (i.e. individual mice) were analysed in two to four technical replicates, resulting in a total of 135 analysed samples. Only four out of the seven xenograft models generated spontaneous lung metastases. From these four xenograft models, a total of 243 samples was analysed (two to nine biological replicates and two technical replicates).

For the subsequent data analysis, we considered only proteins identified with at least one proteotypic peptide. The restriction to proteotypic peptides was due to the inherent challenge of datasets derived from xenograft models which contain proteins of both human and murine origin. The distinction between proteins derived from the human tumour cells or the murine tumour stroma is a prerequisite for the identification of marker proteins. The tab.11 depicts an overview of the protein numbers identified in the different datasets. According to guidelines defined by the proteomic community, proteins have to be identified with at least two peptides. To cope with these guidelines, the table also contains the number of identified proteins without so called “one-hit-wonders” (i.e. proteins identified with only one proteotypic peptide). Our discovery approach resulted in the identification of a total of 3236 and 2117 proteins when counting all proteins and the proteins with at least two proteotypic peptides, respectively. In the xenograft tumour datasets 30 to 50 % of all proteins are of human origin while only 15% of all proteins from xenograft lung metastases proteins derived from human cells.

As quality control measure, MS data from every sample was interrogated concerning the performance of the streptavidin-based affinity purification, the on resin tryptic digestion, peptide desalting and UPLC separation. To this, two-dimensional peptide maps were generated displaying UPLC fractions on the *y*-axis and m/z ratios on the *x*-axis and normalised peptide intensities as a grey scale. A representative map is depicted in Fig. 22 confirming high quality performance of the above mentioned workflow steps.

To assess the variable protein expression in the different data sets not only based on the number of identified peptides but based on a label-free protein quantification technique, we performed the relative quantification of protein expression using the in house developed MS_QBAT software (Alexander Kerner, HI-STEM gGmbH, unpublished). To this, four representative samples of each subset were selected for label-free protein quantification. The selection was based on the quality of the mass spectrometric data visualized using two-dimensional peptide maps. Numbers of proteins quantified by MS_QBAT are depicted in Tab.

11. In total 75% and 69% of all proteins identified with at least one and at least two proteotypic peptides, respectively, were quantified.

Tab. 11: Overview of proteomics results Number of identified (left value) and relatively quantified (right value) proteins detected with a minimum of one and two proteotypic peptides, respectively.

sample	biological replicates analysed	total samples analysed	proteotypic proteins mouse		proteotypic proteins human		proteotypic proteins total	
			min. 1 peptide	min. 2 peptides	min. 1 peptide	min. 2 peptides	min. 1 peptide	min. 2 peptides
renal cortical ep. cells	4	8	-	-	705/591	590/347	705/591	590/347
renal medullary ep. cells	4	8	-	-	605/584	536/339	605/584	536/339
renal proximal tub. cells	4	8	-	-	629/600	536/346	629/600	536/346
kidney mouse	10	39	1193/1091	810/668	-	-	1193/1091	810/668
lung mouse	5	10	610/814	427/494	-	-	610/814	427/494
RCC12 tumour	10	20	536/556	348/289	316/336	177/144	852/892	525/433
RCC15 tumour	9	18	576/584	381/291	397/243	199/89	973/827	580/380
RCC18 tumour	3	6	343/435	227/194	557/545	335/277	900/980	562/471
RCC24 tumour	9	36	323/390	230/193	186/236	103/90	509/626	333/283
RCC27 tumour	10	20	560/657	381/312	233/335	124/129	793/992	505/441
RCC38 tumour	9	18	536/609	367/315	551/446	319/184	1087/1055	686/499
RCC50 tumour	8	16	752674	516/381	342/257	199/106	1094/931	715/487
RCC12 metastases	2	3	668/639	452/409	104/143	64/63	772/782	516/472
RCC15 metastases	3	5	701/640	459/416	49/85	25/30	750/725	484/446
RCC27 metastases	9	17	1024/754	677/466	226/132	117/44	1250/886	794/510
RCC50 metastases	6	11	677/613	450/394	183/194	107/89	860/807	557/483
total	105	243	1820/1441	1250/900	1416/1005	867/557	3236/2446	2117/1457

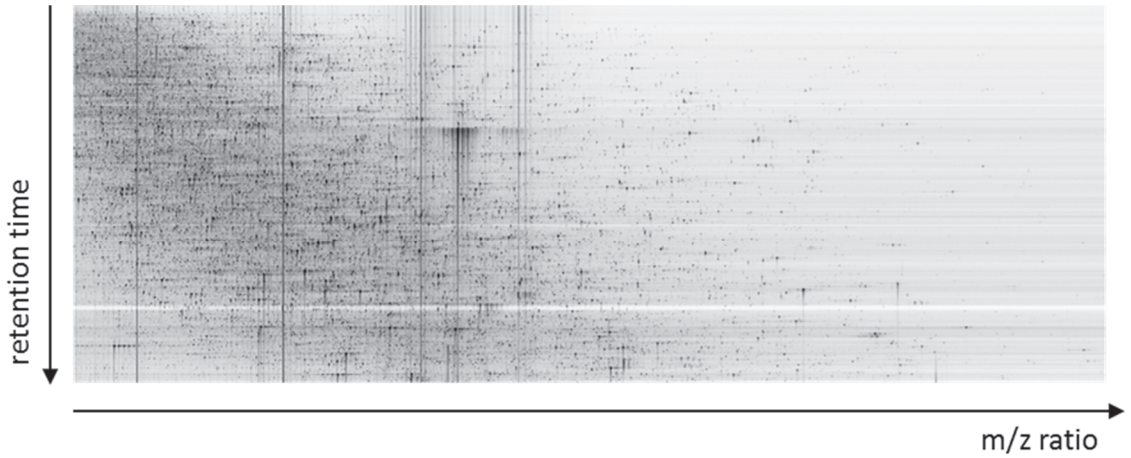


Fig. 22: Representative two dimensional peptide map The 2D peptide map displays UPLC fractions on the y-axis and m/z ratios on the x-axis. All intensities were normalized to a standard peptide peak at 1411 m/z and a standard peak height of 1000 units. For visualization purpose, the normalized intensity was capped at a value of 500 units. 2D peptide maps allow for quality evaluation of streptavidin-based affinity purification, on resin digestion, peptide desalting and UPLC separation.

5.3.1 Reproducibility among technical replicates

To assess the technical reproducibility of the proteomics workflow, protein expression ratios of two technical replicates from different subsets were compared (Fig. 23). Relative protein quantification was performed based on the added signal intensities of at least two proteotypic peptides per proteins. The volcano plot presented in Fig. 23 displays the fold ratio between the two replicates on the *x*-axis and the negative log value of the p-value on the *y*-axis. Significant regulation were defined as ratios bigger than 2 or smaller than 0.5 and with a -log(p-value) bigger than 1.3. Technical replicate datasets deriving from the *in vivo* biotinylation of healthy murine kidneys and from the *in vivo* biotinylation of xenograft tumours (RCC12) showed a striking reproducibility and only two and six significantly regulated proteins could be identified, respectively. These minuscule changes in the inter-replicate protein regulation demonstrate the technical reproducibility of the parallel sample processing. Biological replicates, i.e. replicated not originating from the same cell lysate but from cell lysates derived from two independent animals or two cell culture flasks, manifest as expected a slightly increased number of significantly regulated proteins (Fig. 23C).

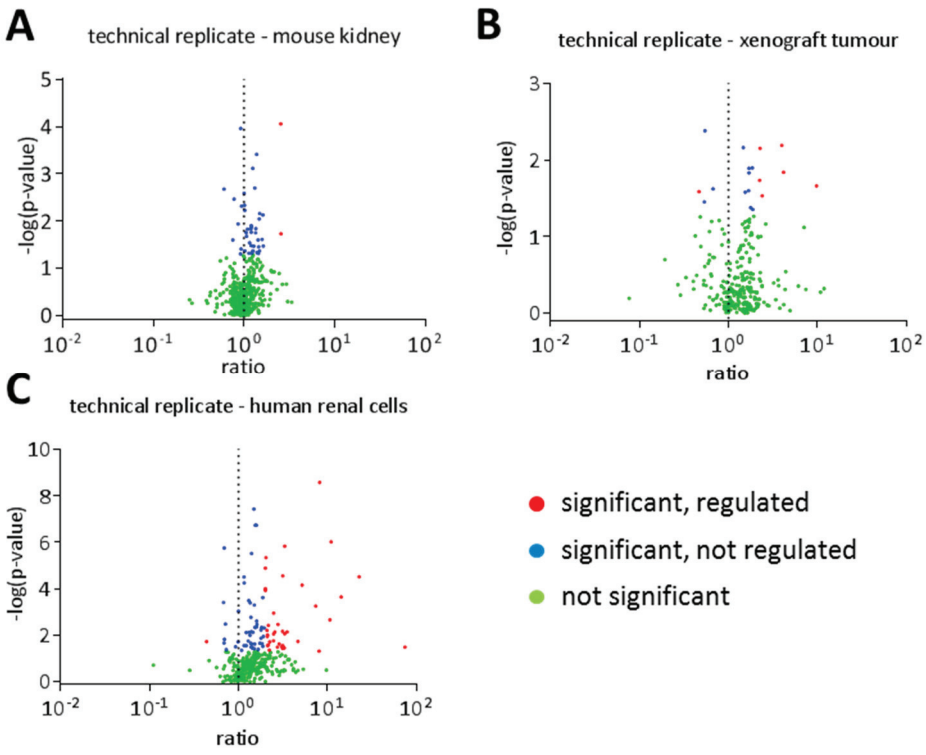


Fig. 23: Volcano-plots illustrating relative quantifications of different technical Two technical replicates from each sample subset (A: murine kidneys; B: xenograft kidney tumours; C: human renal cells) were compared. Protein regulations with a ratio bigger than two and with a p-value smaller than 0.05 (-log(p-value) > 1.3) were considered as significant. Panels A and B demonstrate the technical reproducibility of the sample preparation workflow (samples derived from identical tissue lysates). Panel C indicates in addition the biological variability introduced when using cell lysates derived from two different cell culture flasks.

5.3.2 Protein expression variances among biological replicates

Following the assessment of the technical reproducibility of the proteomic workflow, the variability introduced by biological replicates was assessed in the different datasets. For this, protein expression levels of discrete biological replicates were compated to corresponding merged dataset. Fig. 24 visualized the assessment of biological reproducibility based on counting proteins regulated with a certain relative expression ratio for six different sets of data. Biological variances among four replicates of *in vivo* biotinylated murine kidney samples are very small demonstrating a high technical and biological reproducibility. To a similar degree, the same findings were true for biological replicates of kidney tumours and lung metastases of the same xenograft model. The comparison of the relative protein expression ratios of the datasets derived from the three human renal cell lines indicated an increased variability, especially in terms of identified protein numbers. Significant changes in relative protein expression ratios were observed for the comparison of the datasets of individual xenograft tumour sets with respect to a merged dataset of all xenograft tumours. Thereby,

RCC18 showed the highest variability. Interestingly, only minor variances could be observed among individual xenograft lung metastases datasets.

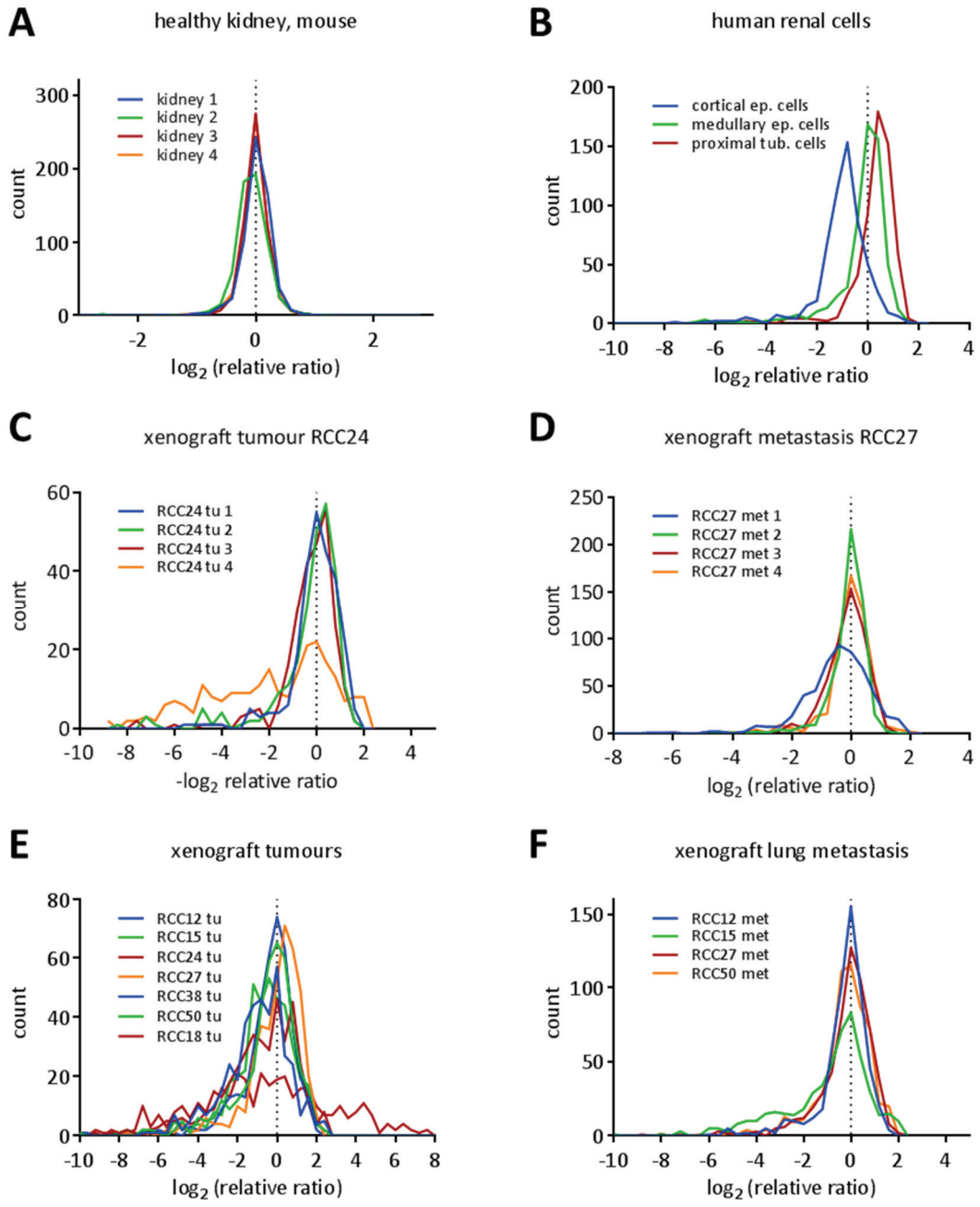


Fig. 24: Protein expression variances among different biological replicates In order to allow for a visual inspection of variances in protein expression levels between different biological replicates within a subset, data derived from discrete samples were compared to the corresponding merged dataset (termed super-set). For visualization purposes, relative protein regulation ratios were combined in bins (bin size: 0.4, central bin ranging from -0.2 to 0.2). Only proteins identified with at least two proteotypic peptides were considered for analysis. A: The comparison of a subset of four in vivo biotinylated murine kidneys deriving from different mice demonstrates a high biological and technical reproducibility. B: Within the three human renal cell lines used as healthy control, the cortical epithelial cells show a slightly divergent protein expression profile compared to the

other two cell lines. C: The comparison of a subsets of four xenograft kidney tumours (RCC24 derived) grown in individual mice indicate a high biological and technical reproducibility for three out of four samples. D: Similarly to the primary tumors, the xenograft lung metastases (RCC27 derived) grown in different mice demonstrate a high biological and technical reproducibility for three out of four samples. E: A significant biological variation between individual tumours can be observed upon comparison of all subsets derived from seven different, patient-matched xenograft kidney tumours. F: The biological variation between the four different, patient-derived spontaneous xenograft lung metastases was considerably smaller than the respective variation between the primary tumours shown in panel E.

In summary, both the reproducibility of the sample preparation / analysis workflow as well as the reproducibility of biological replicates could be considered as sufficient for the subsequent data analysis. As expected, pronounced variances among individual patient tumours (i.e. the different xenograft models) could be confirmed.

5.3.2.1 Protein expression variances among subsets of human renal cell lines

Owing to the unavailability of healthy human kidneys for *in vivo* biotinylation, three human renal primary cell lines originating from the cortical epithelium, the medullary epithelium and the proximal tubules were combined into on dataset applied as the healthy human control. Fig. 25shows the variances among the identified proteins and their expression levels between the three different cell types. Furthermore, we examined possible difference between two cultivation media (with serum and serum-free) and two different passages. Differences in protein expression for two samples with differing in cultivation passages or media are presented in the form of volcano plots (displaying relative protein expression ratios and p-values). Neither passage differences nor alterations in the medium composition resulted in major changes in protein expression. Compared to technical replicates (i.e. two samples deriving from the same passage cultured in the same medium), only minor increases in protein expressions could be observed (Fig. 23).

A Venn diagram displaying the overlapping protein identifications for the three renal cell lines is shown in Fig. 25C. While 57% of all identified proteins were present in all three datasets, about 10% were unique for each cell line.

Differences in protein expression ratios for the individual cell lines compared to the whole dataset were exemplified in the previous chapter. No major changes between the cell lines were observed.

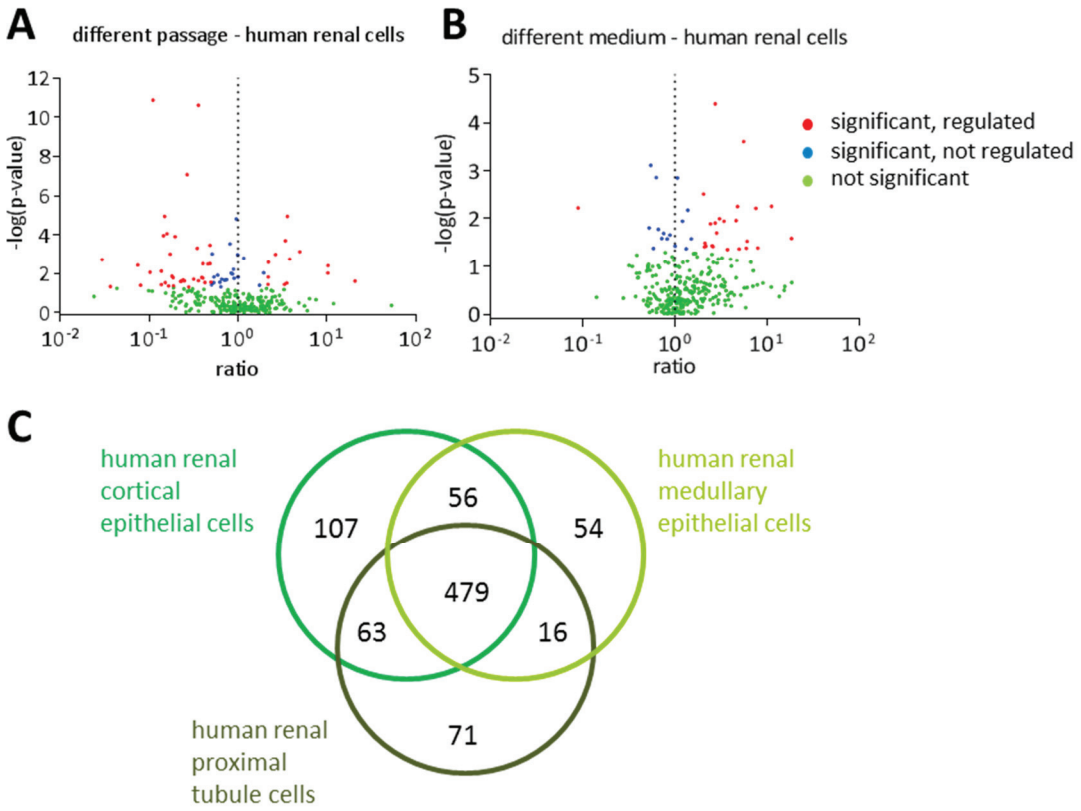


Fig. 25: Protein expression variances in the human renal cell line datasets A&B Volcano-plots depicting relative protein changes of the same cell line within different passages (A) and when cultivated in two different cell culture media (B). Quantification ratios larger than two and with a p-value smaller than 0.05 were considered significant. Neither passage differences nor alterations in the medium composition resulted in major changes in protein expression. C: Venn diagram depicting the overlap between protein identifications (at least one proteotypic peptide) in datasets from the three renal cell lines medullary epithelial cells, cortical epithelial cells and proximal tubule. While 57% of all proteins where expressed commonly in all three cell lines, 16% and 27% were expressed in two out of three and in only one cell line, respectively.

5.3.3 Subcellular localisation of identified proteins

We aimed with the *in vivo* biotinylation approach at the enrichment and identification of vascular accessible proteins located on the cell surface or in the extracellular matrix in order to discover novel antibody-targetable markers. To evaluate the degree of enrichment of desired protein classes subcellular localisation profiles of each dataset were generated based on the information available in the UniProt database (Fig. 26). In the murine kidney and xenograft kidney tumour datasets (*in vivo* biotinylation of the systemic circulation) 49 and 53%, respectively, of all proteins were annotated to cell membranes and/or the extracellular matrix. Upon *in vivo* perfusion of the pulmonary circulation, 77% and 62% of all proteins were of non-intracellular origin in murine lungs and in spontaneous lung metastases, respectively. The dataset obtained by *in vitro* biotinylation of cultured cells resulted in 55%of all proteins being associated with the cell membrane or the extracellular space. In conclusion, all datasets demonstrated a significant enrichment of cell membrane proteins and proteins of the

extracellular space. Importantly, the newly established lung perfusion technology resulted in a high content of non-intracellular proteins in the respective datasets (up to 77% of proteins with cell membrane or extracellular space annotation).

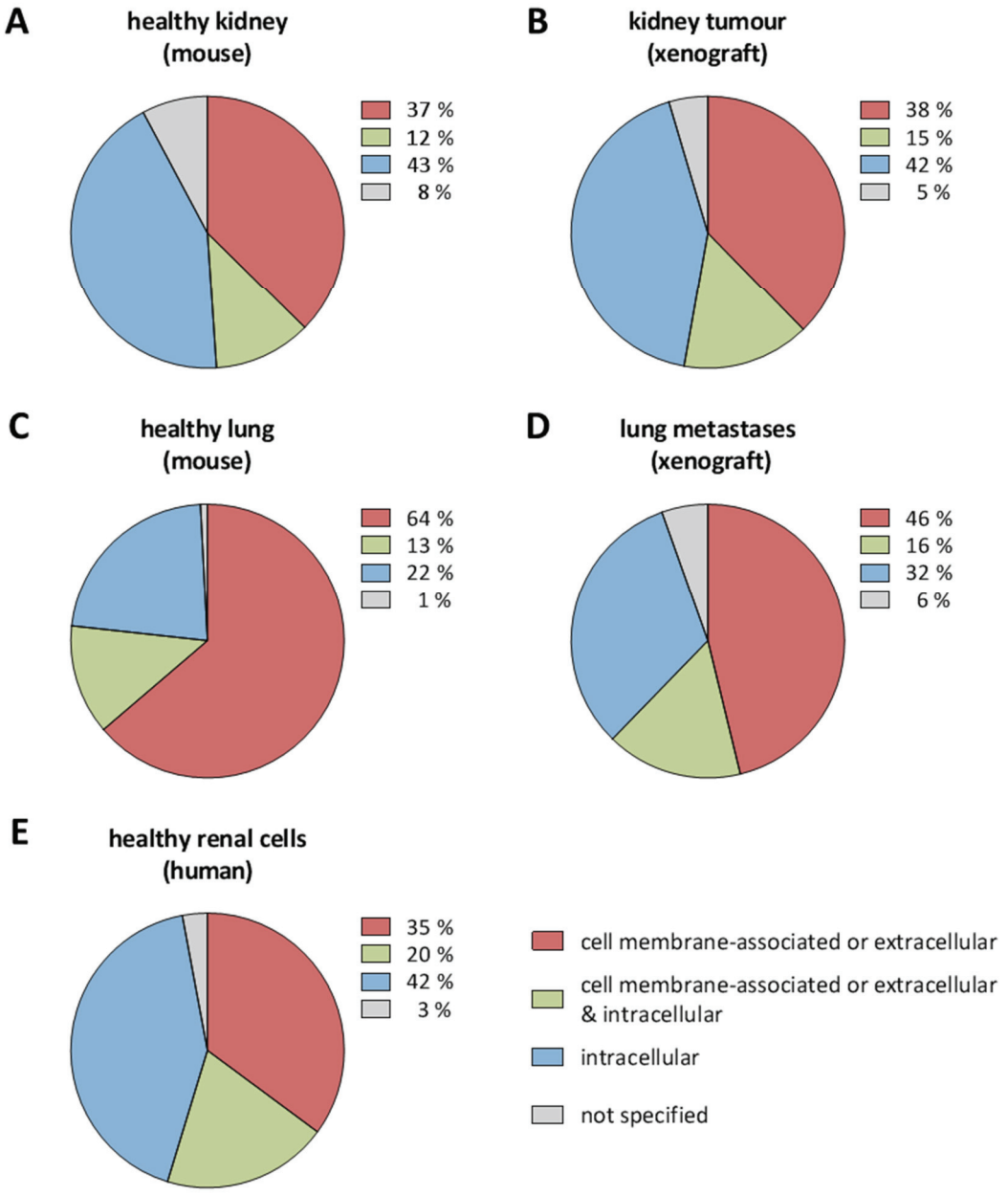


Fig. 26: Subcellular localisation of identified proteins in the different datasets Information about the subcellular localisation of identified proteins was obtained from the UniProt database. All proteins identified with at least one proteotypic peptide were considered for analysis. About 50% of all proteins identified in the datasets from murine kidneys (A), xenograft kidney tumours (B) and human renal cells (E) were annotated with locations on the cell surface or in the extracellular matrix. The datasets from murine lungs (C) and xenograft lung metastases (D), generated using the lung perfusion technology developed during this thesis, showed a higher degree of enrichment of non-intracellular proteins (62-77%).

5.3.4 Clustering analysis of different proteomics datasets

Hierarchical clustering was applied to visualize variances between protein expression levels of the different sub-dataset of xenograft kidney tumours, lung metastases and healthy materials (Fig. 27). The clustering analysis, which applied the Kendall's τ clustering algorithm, was performed based on datasets containing all proteins identified with at least one proteotypic peptide. All seven xenograft kidney tumours form a separate subcluster as expected. The same finding was true for the xenograft lung metastases subsets while the healthy material subsets (human renal cells, murine kidney and lung) clustered individual. Importantly, the murine lung set showed the closest relation to its corresponding diseased tissues. Xenograft tumour subsets clustered based on their growth rates and phenotypic appearance, while experimental parameters (i.e. the number of analysed biological replicates) influenced the clustering of the lung metastases, Pairwise clustering was observed for fast growing tumour (RCC24 and RCC38), highly metastasising tumours (RCC15 and RCC27) and tumours with a medium metastasis potential (RCC12 and RCC50).

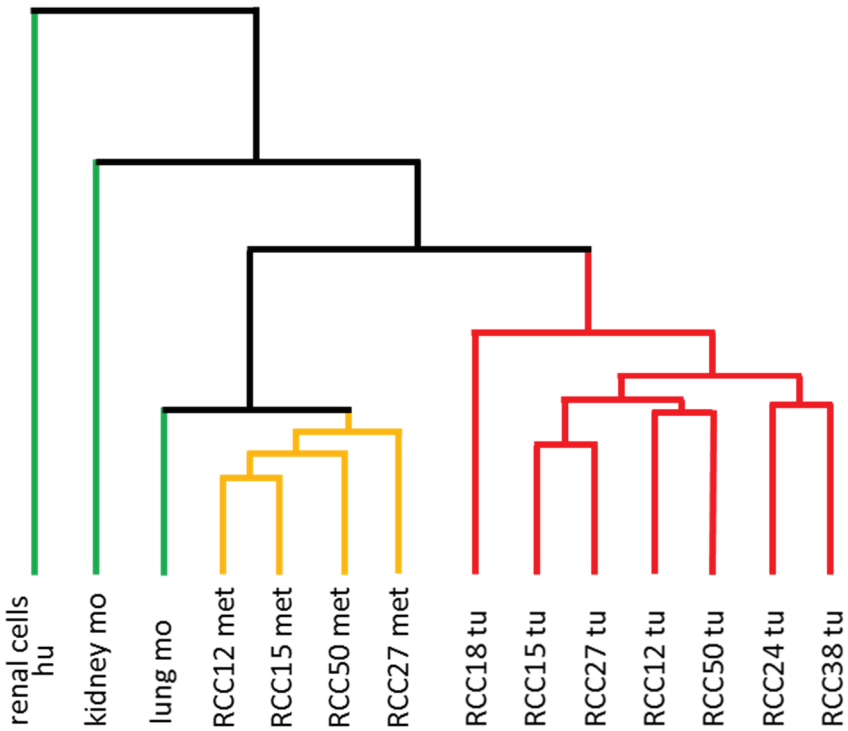


Fig. 27: Hierarchical clustering of all datasets A cluster analysis of all datasets was performed with all proteins identified with at least one proteotypic peptide using the Kendall's τ clustering algorithm. Both the xenograft tumour datasets as well as datasets from xenograft lung metastases and mouse lung form a separate cluster. The two healthy datasets "renal cells" and "kidney" are displayed individual.

5.4 COMPARATIVE ANALYSIS OF PROTEOMICS DATASETS

As outlined above, data analysis of proteomics data deriving from xenograft models is challenging. In order to identify differentially regulated proteins and to select marker candidates for further validation, all proteomics data were combined into distinct tables. The dataset deriving from human renal cells (healthy human kidney) was compared to human proteins expressed by the xenograft kidney tumours for the identification of tumour-expressed marker candidates. The dataset derived from normal murine kidney was compared to the proteins with murine origin in the xenograft kidney tumour datasets in order to identify stromal-derived marker candidates. Comparing the human proteins expressed in xenograft kidney tumours with the ones deriving from spontaneous lung metastases allowed for the identification of metastasis-specific marker candidates. Finally, murine proteins expressed in spontaneous lung metastases were compared to healthy murine lungs for the identification of stromal-derived metastasis marker candidates.

An overview of protein expression changes in the different datasets is outlined in the Venn diagram in Fig. 28. The three main datasets “healthy” (murine kidneys and lungs, human renal cells), “xenograft kidney tumours” and “xenograft lung metastases” are mapped against each other. Approximately one third of all proteins were shared between all datasets. However, unique proteins could be identfied for every dataset. Human proteins expressed in diseased tissues were mainly shared between kidney tumours and lung metastases and only few proteins where exclusively expressed in the two datasets. For biomarker selection we considered proteins which are exclusively expressed in the xenograft datasets while not being identified in the healthy subset.

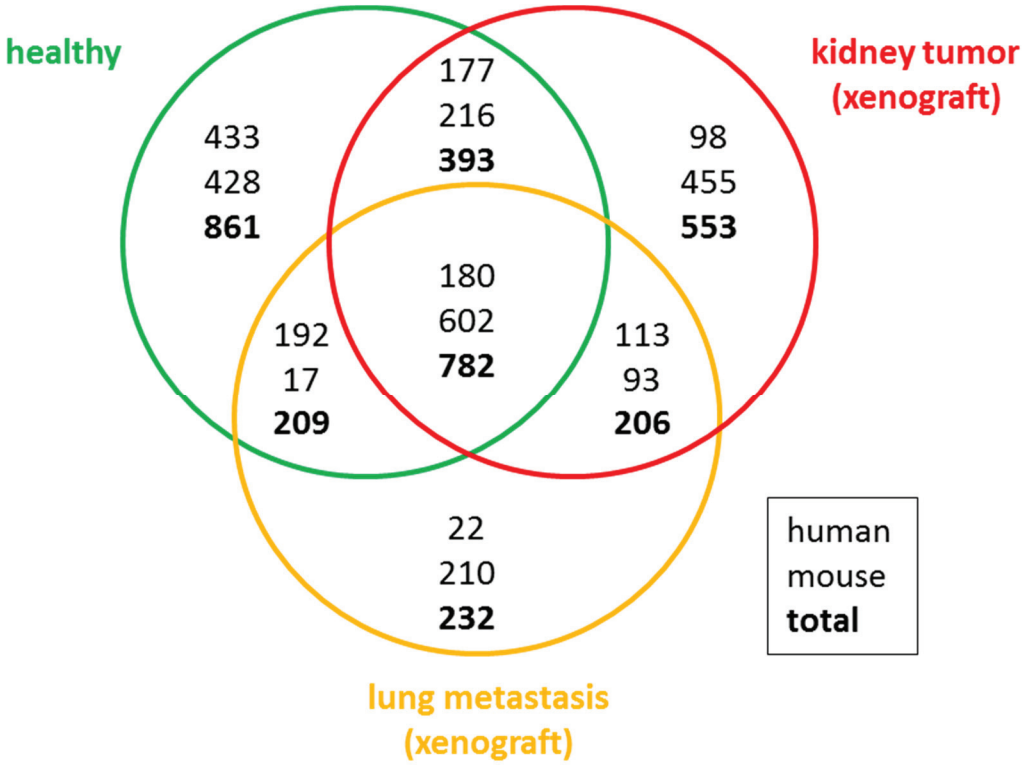


Fig. 28: Venn diagram The overlap between protein identifications (at least one proteotypic peptide) in datasets from healthy tissues, xenograft kidney tumours and xenograft lung metastases is presented in a Venn diagram. Numbers for human, murine and total protein are indicated (from top to bottom).

5.4.1 Gene ontology analysis of proteomics datasets

In order to obtain a global picture of protein regulation and the respective pathways, a gene ontology analysis was performed. Therefore, human and mouse proteins upregulated in xenograft kidney tumours (compared to proteins identified in human renal cell datasets and murine kidney datasets, repectively) were selected and analyzed using the online tool Webgestalt (Fig. 29). Human tumour-upregulated proteins were annotated to biological processes such as the carboxylic acid metabolism and organic acid metabolism. In contrast, murine tumour-upregulated proteins were annotated to biological processes involved in response to stimuli (e.g. stress and wounding) and blood coagulation.

Additionally, a gene ontology analysis using human proteins upregulated in spontaneous lung metastases compared to xenograft kidney tumours was performed (Fig. 30). Here, metastasis-upregulated human proteins were involved in cell adhesion processes, cell migration, system development and response to stress and wounding. The gene ontology analysis, which was based on murine proteins overexpressed in xenograft lung metastases compared to healthy murine lungs, resulted in the identification of biological process in cell adhesion, cell migration

and response to stress. Additionally, proteins involved in metabolic processes and blood coagulation were upregulated.

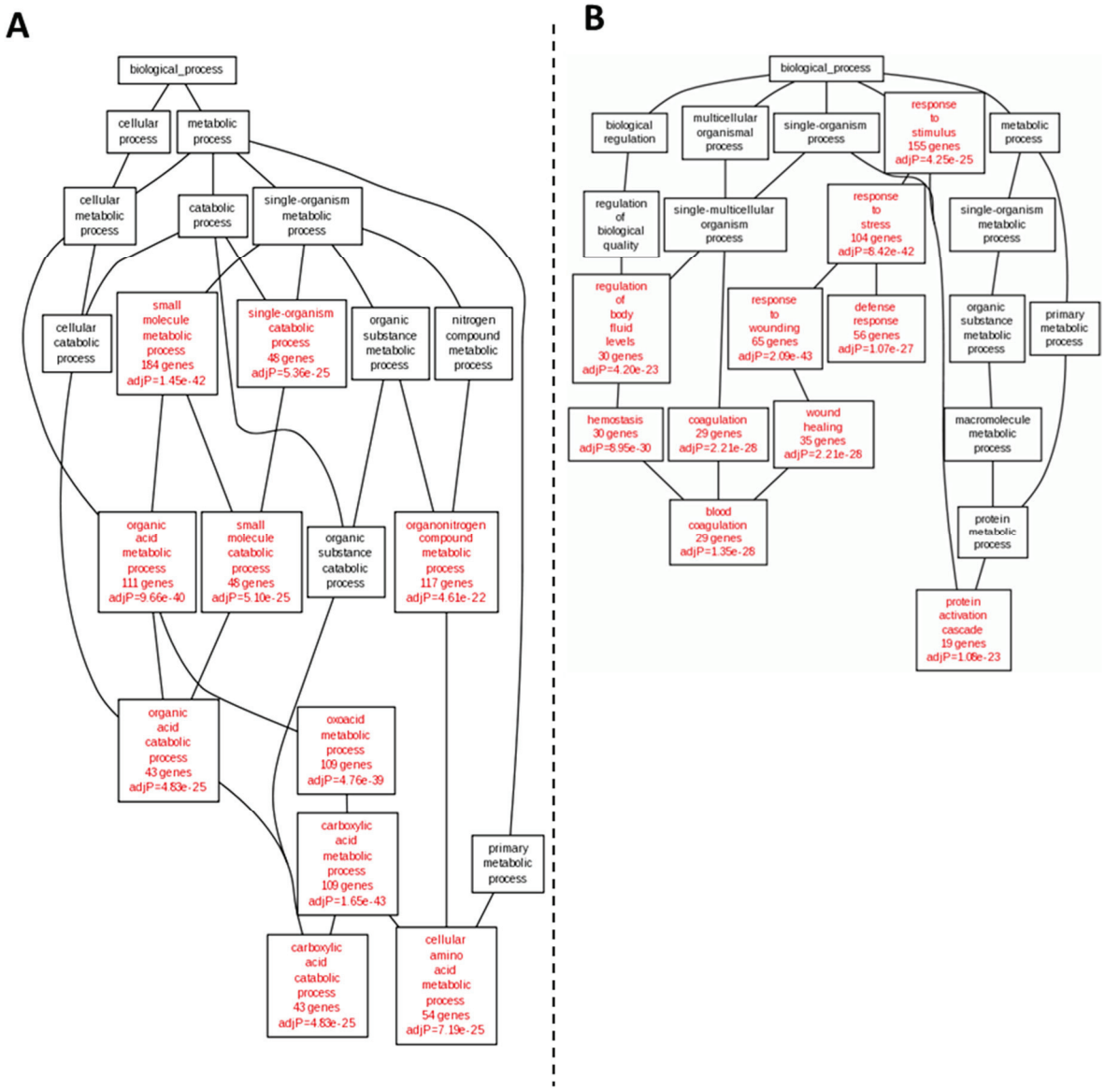
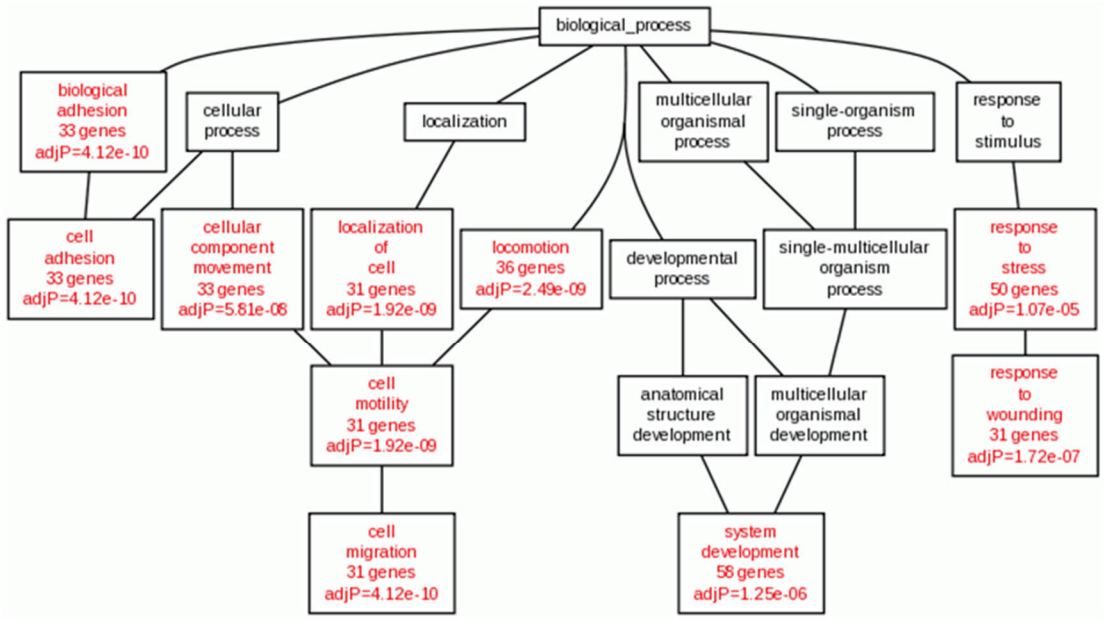


Fig. 29: Gene ontology analysis of proteins upregulated in xenograft kidney tumours Proteins overexpressed in xenograft kidney tumours in comparison to human renal cells (A) and murine kidneys (B) were submitted to gene ontology analysis using Webgestalt. While upregulated human proteins (A) were mainly annotated to metabolic processes, most murine proteins (B) were involved in blood coagulation and responses to stimuli.

A



B

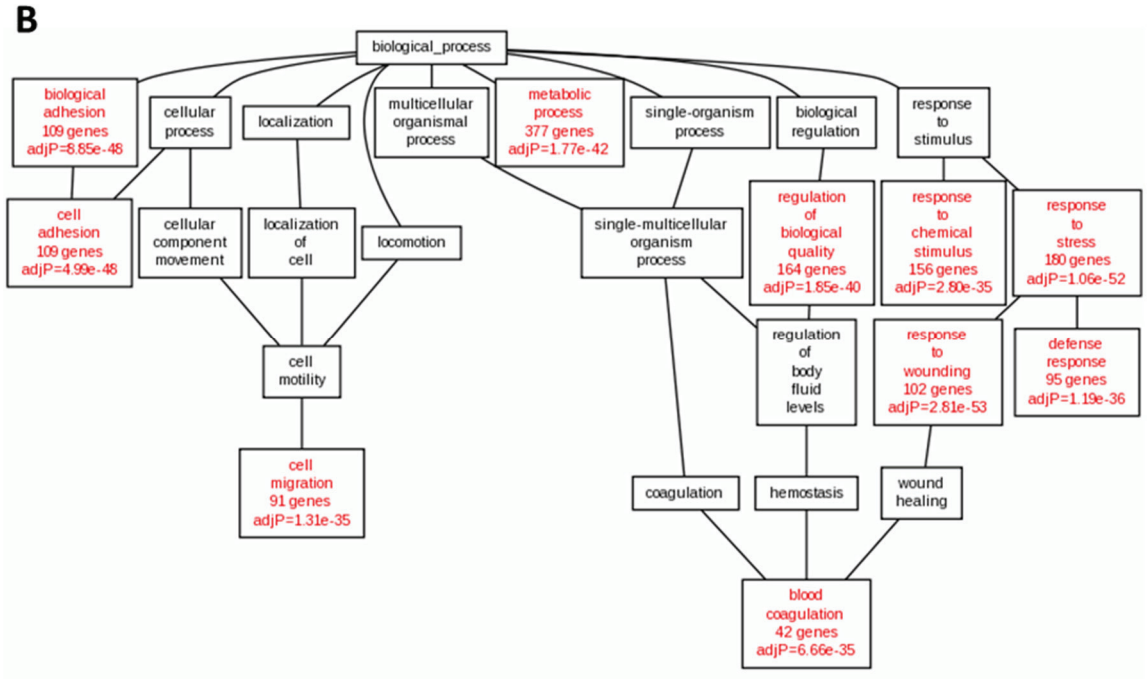


Fig. 30: Gene ontology analysis of proteins upregulated in xenograft lung metastases Proteins overexpressed in xenograft lung metastases in comparison to their corresponding xenograft tumours (A) and murine lungs (B) were submitted to gene ontology analysis using Webgestalt. While upregulated human proteins (A) were mainly annotated to system development and response to stress, most murine proteins (B) were involved in cell adhesion, cell migration and response to wounding.

5.4.2 Relative quantification of the protoemics datasets

Mass spectrometric data obtained during the initial proteomics discovery study were used for label-free relative protein quantification. The quantification strategy applied in the software

package MS_QBAT (Alexander Kerner, HI-STEM gGmbH, unpublished) is based on the AUC values of peptide intensities normalised to standard peptides. MS_QBAT allows the pairwise comparison of summed peptide intensities corresponding to individual proteins. The significance of protein regulations was calculated by an unpaired t-test. Relative protein quantification was performed on four representative samples of each sub-dataset. As outlined previously, a number of different comparisons were performed (Tab. 12 to

Tab. 15). Candidate markers for primary kidney tumours were identified by comparing human and murine proteins identified in the respective xenograft models with the datasets deriving from human renal cells and healthy murine kidneys, respectively. Candidate markers for lung metastases were identified by comparing human and murine proteins identified in spontaneous lung metastases with the datasets deriving from xenograft kidney tumours and healthy murine lung, respectively. As outlined in the gene ontology analyses (chapter 5.4.1) murine proteins upregulated in xenograft kidney tumours were mainly involved in inflammation and blood coagulation processes such as Kng1, Kng2, Apoh and Orm2. Furthermore, proteins involved in cell differentiation (e.g. Ahsg, Lama1) and proteins in connective tissues (Fbln2 and different collagens) were top ranked. Many human proteins upregulated in xenograft kidney tumours are associated with metabolic processes and hold a subcellular localization in the cytoplasm (e.g. GANAB, GLUD1 and RPN1). Additionally, proteins involved in cell adhesion and invasion such as LAMB1, LAMA5, COL1A1, COL12A1 and FSCN1were top ranked. Interestingly, the intracellularly located cancer stem cells marker ALDH1A1 was found to be significantly upregulated in all xenograft tumours. Murine proteins highly expressed in lung metastases samples were, similar to murine proteins upregulated in kidney tumours, associated with blood coagulation and immune system response (e.g. C9, F13a1 and Orm1). The integrin Itgam, the collagen Col15a1 as well as the integrin interactor protein tenascin Tnc, all invovled in cell adhesion and migration processes, showed increased expression levels in lung metastases compared to primary tumours. Only few human proteins were significantly upregulated in lung metastases compared to the corresponding primary tumours. Among them, cell adhesion proteins such as LAMA5, ITGA3 and BCAM but also immune system associated proteins (CFH, C1R and CD109) were listed as significantly upregulated.

Tab. 12: Top upregulated human proteins in xenograft kidney tumours compared to the human renal cells dataset Only proteins identified with more than one proteotypic peptide in at least five xenograft models with an average p-value of less than 0.05 were considered. Sorting was performed based on the average ratio. (p-values: < 0.001: ***; < 0.01: **; < 0.05: *; ≥ 0.05: not significant (n.s.); n.d., not determined).

swiss prot acc no	protein name	gene name	RCC12 tu			RCC15 tu			RCC18 tu			RCC24 tu			RCC27 tu			RCC38 tu			RCC50 tu		
			log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides
P00352	Retinal dehydro-genase 1	ALDH1A1	54	**	4	249	*	3	101	***	12	195	***	8	105	***	6	157	***	10	89	***	11
Q99715	Collagen alpha-1(XII) chain	COL12A1				238	*	3	18	n.s	2				87	***	3	224	***	3	85	***	11
P15121	Aldose reductase	AKR1B1	9	n.s.	3	140	*	4	28	n.s.	3	219	***	5	175	***	6	113	***	5	79	**	6
P51659	Peroxisomal multifunctional enzyme type 2	HSD17B4				15	n.d.	1							12	n.d.	1	27	*	2	340	*	4
Q16658	Fascin	FSCN1	2	n.d.	1	241	*	3							14	n.s.	3	3	n.d.	1	172	*	3
Q13011	Delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial	ECH1				41	n.d.	1	96	***	6	74	n.d.	1	78	n.s.	2	193	n.d.	1	28	n.d.	1
P42765	3-ketoacyl-CoA thiolase, mitochondrial	ACAA2				10	n.d.	1	256	***	8							20	n.d.	1	50	n.d.	1
P00367	Glutamate dehydrogenase 1, mitochondrial	GLUD1	28	n.d.	1	21	n.d.	1	259	*	2							45	n.d.	1	32	n.d.	1
Q14697	Neutral alpha-glucosidase AB	GANAB				101	n.s.	4	36	*	3				94	***	4	98	*	2	33	***	6
P07942	Laminin subunit beta-1	LAMB1	15	n.s.	3	9	n.d.	1	265	***	26	18	n.d.	1				52	**	6			
P11413	Glucose-6-phosphate 1-dehydrogenase	G6PD	54	***	4	64	**	3	56	*	3				124	**	4	45	**	4	53	n.s.	2
Q02218	2-oxoglutarate dehydrogenase, mitochondrial	OGDH				12	n.d.	1	163	**	5							3	n.d.	1	78	n.d.	1
P02452	Collagen alpha-1(I) chain	COL1A1	52	**	3	48	***	5							27	***	8	72	n.d.	1	67	**	4
Q01813	6-phosphofructokinase type C	PFKP	43	**	5	56	n.d.	1	47	***	7	83	n.d.	1	54	n.s.	2	30	***	5	26	n.s.	3
Q16401	26S proteasome non-ATPase regulatory subunit 5	PSMD5				32	*	2	11	n.s.	3				38	*	3	96	n.d.	1			
P04843	Dolichyl-diphospho-oligosaccharide--protein glycosyl-transferase subunit 1	RPN1	14	n.d.	1	41	n.d.	1	31	***	5				56	**	3	66	**	3	21	n.s.	2
Q99714	3-hydroxyacyl-CoA dehydrogenase type-2	HSD17B10							52	n.s.	2				23	*	2	44	*	3	14	n.d.	1
O43488	Aflatoxin B1 aldehyde reductase member 2	AKR7A2	2	n.d.	1	46	n.d.	1	85	***	3				4	n.d.	1				28	n.d.	1
Q7KZF4	Staphylococcal nuclease domain-containing protein 1	SND1				23	n.d.	1	61	*	3				29	n.d.	1	15	n.d.	1	25	***	3
P42224	Signal transducer and activator of transcription 1-alpha/beta	STAT1	21	**	5	43	**	2	9	n.d.	1				39	*	3				23	*	4

Tab. 13: Top upregulated murine proteins in xenograft kidney tumours compared to the murine kidney dataset
Only proteins identified with more than one proteotypic peptide in at least five xenograft models with an average p-value of less than 0.05 were considered. Sorting was performed based on the average ratio. (p-values: < 0.001: ***; < 0.01: **; < 0.05: *; ≥ 0.05: not significant (n.s.); n.d., not determined).

swiss prot acc no	protein name	gene name	RCC12 tu			RCC15 tu			RCC18 tu			RCC24 tu			RCC27 tu			RCC38 tu			RCC50 tu		
			log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides
P11370	Retrovirus-related Env polyprotein from Fv-4 locus	Fv4	11	*	4	34	*	2	10	n.d.	1	21	*	5	120	**	4	151	n.s.	2	142	*	2
Q61646	Haptoglobin	Hp	6	***	7	80	***	7		n.d.		67	***	9	11	**	8	24	***	10	107	***	8
Q61704	Inter-alpha-trypsin inhibitor heavy chain H3	Itih3	31	**	11	30	*	9	25	***	17	71	***	11	35	***	10	43	***	11	30	**	7
Q91X72	Hemopexin	Hpx	20	***	30	36	***	28	3	n.s.	19	36	***	54	26	***	28	47	***	36	25	***	31
O08677	Kininogen-1	Kng1	11	***	11	13	***	9	5	*	8	54	*	12	12	***	12	7	***	16	12	**	10
P37889	Fibulin-2	Fbln2	2	n.s.	6	2	n.s.	7	3	n.s.	4	79	***	6	5	***	14	3	***	14	4	***	11
P07724	Serum albumin	Alb	9	***	87	13	***	92	11	***	142	10	***	98	6	***	98	12	***	76	14	***	119
P29699	Alpha-2-HS-glycoprotein	Ahsg	6	***	11	6	**	8	8	**	7	12	***	10	10	***	9	7	**	8	2	n.s.	9
P11087	Collagen alpha-1(I) chain	Col1a1	5	***	36	15	***	46	2	n.s.	29	6	**	33	3	***	42	3	***	32	11	***	37
Q60847	Collagen alpha-1(XII) chain	Col12a1	4	***	24	4	**	20	7	***	34	3	*	20	5	***	24	6	***	26	10	***	30
Q61838	Alpha-2-macroglobulin	A2m	5	***	38	5	**	34	1	n.s.	41	9	***	48	3	***	43	10	***	39	3	**	35
Q01149	Collagen alpha-2(I) chain	Col1a2	3	**	53	11	***	57	2	n.s.	46	5	**	50	3	***	57	3	***	47	5	***	47
P19137	Laminin subunit alpha-1	Lama1	0	***	2	0	***	5	0	***	2	0	***	2	4	***	5	6	**	6	0	***	6

Tab. 14: Top upregulated human proteins in xenograft lung metastases compared to the corresponding xenograft kidney tumour dataset Only proteins identified with more than one proteotypic peptide in at least two xenograft models with an average p-value of less than 0.05 were considered. Sorting was performed based on the average ratio. (p-values: < 0.001: ***; < 0.01: **; < 0.05: *; ≥ 0.05: not significant (n.s.); n.d., not determined).

swiss prot acc no	protein name	gene name	RCC12 met			RCC15 met			RCC27 met			RCC50 met		
			log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides
P26006	Integrin alpha-3	ITGA3	0	*	5		n.d.		1	n.d.	1	36	*	3
Q9GZM7	Tubulointerstitial nephritis antigen-like	TINAGL1	0	n.s.	1		n.d.			n.d.		11	***	6
P08195	4F2 cell-surface antigen heavy chain	SLC3A2		n.d.		0	n.d.	1		n.d.		7	*	3
P98160	Basement membrane-specific heparan sulfate proteoglycan core protein	HSPG2	0	***	5	0	**	1	0	***	11	11	***	38
O15230	Laminin subunit alpha-5	LAMA5	0	**	6		n.d.			n.d.		3	*	15
P10909	Clusterin	CLU	1	*	3		n.d.		0	n.s.	1	3	*	3

Tab. 15: Top upregulated murine proteins in xenograft lung metastases compared to the murine lung dataset
Only proteins identified with more than one proteotypic peptide in at least three xenograft models with an average p-value of less than 0.05 were considered. Sorting was performed based on the average ratio. (p-values: < 0.001: ***; < 0.01: **; < 0.05: *; ≥ 0.05: not significant (n.s.); n.d., not determined).

swiss prot acc no	protein name	gene name	RCC12 met			RCC15 met			RCC27 met			RCC50 met		
			log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides	log_2(ratio)	p-value	peptides
Q80YX1	Tenascin	Tnc	4	n.d.	1				5	n.s.	3	3	*	3
Q60590	Alpha-1-acid glycoprotein 1	Orm1	3	*	3	5	***	5	4	*	3	4	*	4
P06802	Ectonucleotide pyrophosphatase/ phosphodiesterase family member 1	Enpp1	6	n.d.	1	2	n.d.	1	3	n.s.	2	2	**	4
Q61704	Inter-alpha-trypsin inhibitor heavy chain H3	Itih3	4	***	9	3	**	9	3	***	10	4	***	11
P06683	Complement component C9	C9	3	**	6	4	***	7	4	***	7	3	***	7
P17182	Alpha-enolase	Eno1	3	n.d.	1	2	n.s.	2	5	*	3	2	n.d.	1
O88844	Isocitrate dehydrogenase [NADP] cytoplasmic	Idh1	3	n.s.	2	3	*	2	4	n.s.	2	2	*	2
O08529	Calpain-2 catalytic subunit	Capn2	3	n.d.	1	2	n.d.	1	1	n.s.	3			
Q61702	Inter-alpha-trypsin inhibitor heavy chain H1	Itih1	4	***	14	1	**	7	1	n.s.	7	2	***	10
Q8BH61	Coagulation factor XIII A chain	F13a1	2	n.d.	1				3	n.s.	2	-1	n.d.	1
P05555	Integrin alpha-M	Itgam	1	n.s.	9	2	**	12	2	*	13	3	***	11
P97449	Aminopeptidase N	Anpep	2	n.d.	1				2	n.s.	2	3	**	4
P11370	Retrovirus-related Env polyprotein from Fv-4 locus	Fv4	2	*	4	3	*	3	0	n.d.	1	2	*	3
A6X935	Inter alpha-trypsin inhibitor, heavy chain 4	Itih4	1	*	23	2	***	29	2	*	19	2	***	27
Q91X72	Hemopexin	Hpx	2	***	44	2	n.d.	70	2	***	39	2	***	45
Q8BND5	Sulfhydryl oxidase 1	Qsox1	2	*	10	2	***	11	2	***	9	2	***	13
O35206	Collagen alpha-1(XV) chain	Col15a1	2	*	4	1	*	4	1	*	6	2	***	4
P01029	Complement C4-B	C4b	2	***	34	1	**	32	2	***	32	2	***	31
Q08761	Vitamin K-dependent protein S	Pros1	2	n.s.	2	2	n.d.	1	1	n.d.	1	2	n.s.	4
Q9QUM0	Integrin alpha-IIb	Itga2b	2	***	13	1	**	13	2	***	12	1	n.s.	11

5.4.3 Relative quantification of selected proteins

In the following chapter, an in-depth analysis of relative protein expression levels for selected proteins is performed. Importantly, one part of the quantified proteins has been selected early in the project using an alternative protein quantification approach, i.e. the counting of the number of peptides identified for a certain protein. Out of six markers selected using the alternative protein quantification approach, TGFBI, Ltbp2 and Lcn2 could be quantified in the present dataset (Fig. 31). TGFBI was the only protein showing significant upregulation in most xenograft kidney tumours. Interestingly, TGFBI expression in xenograft lung metastases samples was unchanged or down regulated when compared to the corresponding primary tumour. The extracellular protein Lcn2 was found to be highly upregulated in all xenograft kidney tumours except RCC15. Lcn2 quantification was based on a single proteotypic peptide, which is why none of the regulations obtained a p-value below 0.05. Ltbp2 was upregulated in xenograft lung metastases when compared to the corresponding xenograft tumours. Owing to the identification of two to ten proteotypic peptides, the relative quantification of Ltbp2 was more robust. Lcn2 was found to be significantly regulated in the fast growing xenograft tumour RCC24 and slightly regulated in other kidney tumours except for RCC18. Importantly, Lcn2 expression levels were additionally upregulated in lung metastases compared to the ones in the corresponding xenograft kidney tumours (quantification based on three proteotypic peptides). IGFBP3, a protein selected in the course of the inital data analyis, could be quantified in the two xenograft tumours RCC12 and RCC24 with one proteotypic peptide (relative protein regulation (log2): 5.3 and 11, respectively).

The final MS_QBAT-based data analysis resulted in the identification of additional differentially regulated proteins. Selected human proteins upregulated in xenograft kidney tumours in respect the human renal cell lines are presented in Fig. 32. Increased expression levels of EGFR could be monitored for six out of seven xenograft tumour models. However, no additional change in EGFR expression could be observed in three out of four xenograft lung metastases. LAMA1 was found to be significantly higher expressed in fast growing xenograft models (RCC24 and RCC38) and in tumours deriving from RCC18. Interestingly, CAPN1 was upregulated solely in metastasising xenograft kidney tumours (RCC12, RCC15, RCC27 and RCC50). In RCC18, CAPN1 values were significantly increased. Additionally, CAPN2, a second family member, was found to be upregulated in RCC28 xenograft kidney tumours. Both CAPN1 and CAPN2 were downregulated in their corresponding lung metastases (data not shown).

Fig. 33 presents the relative quantification of two murine proteins upregulated in xenograft kidney tumour. Ahsg expression levels were found to be significantly increased in six out of seven xenograft models (compared to levels in healthy murine kidney samples). Fbln2 showed

a minor upregulation with a very high significance in the xenograft models RCC24, RCC27, RCC38 and RCC50.

The murine proteins Igfbp3 and Plxn2 are approximately two fold upregulated in xenograft lung metastases compared to the corresponding kidney tumours (Fig. 34). However, protein levels of Igfbp3 and Plxn2 were not increased in the original xenograft tumours.

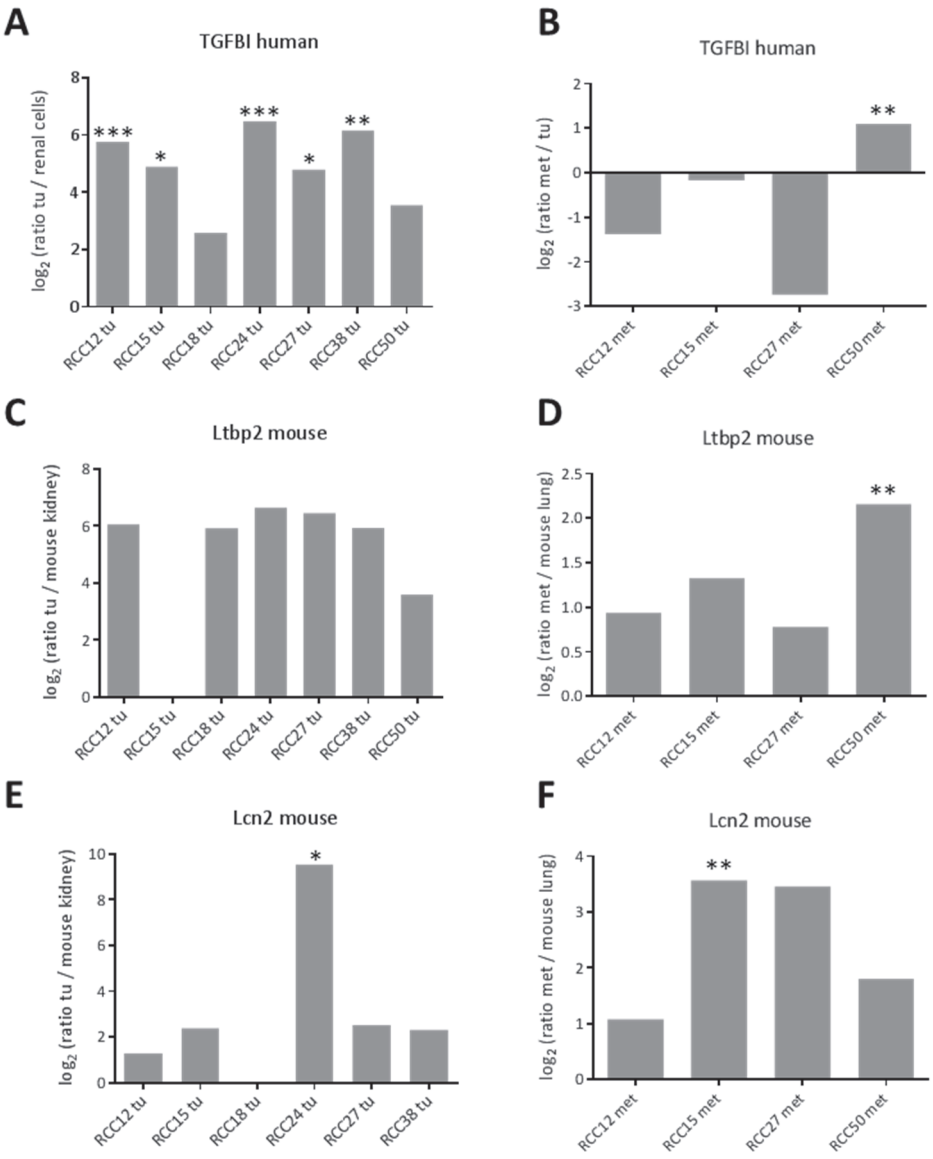


Fig. 31: Relative quantification of putative markers The relative quantification of the three marker candidates TGFBI (A&B), Ltbp2 (C&D) and Lcn2 (E&F), which were selected for subsequent validation, was obtained from the MSQBAT analysis. Summed intensities of peptides originating from TGFBI were compared between the xenograft tumour and human renal cell datasets (A). Panel B displays intensities measured in the lung metastases samples compared to the corresponding xenograft tumours. For C and E the comparison was performed between the datasets deriving from the xenograft tumours and the murine kidneys. Data from lung metastases (D&F) were compared to data obtained from murine lungs. (p-values: < 0.001: ***; < 0.01: **; < 0.05: *; ≥ 0.05: not significant). TGFBI was upregulated in all xenograft tumours, while its expression levels were not further increased in the corresponding lung metastases. The murine Ltbp2 showed an increased expression in xenograft tumours as well as one lung metastases sample. The murine Lcn2 was significantly upregulated in RCC24. Except

for metastases occurring in the RCC12 model Lcn2 expression was higher in lung metastases compared to their corresponding kidney tumours, although not significantly.

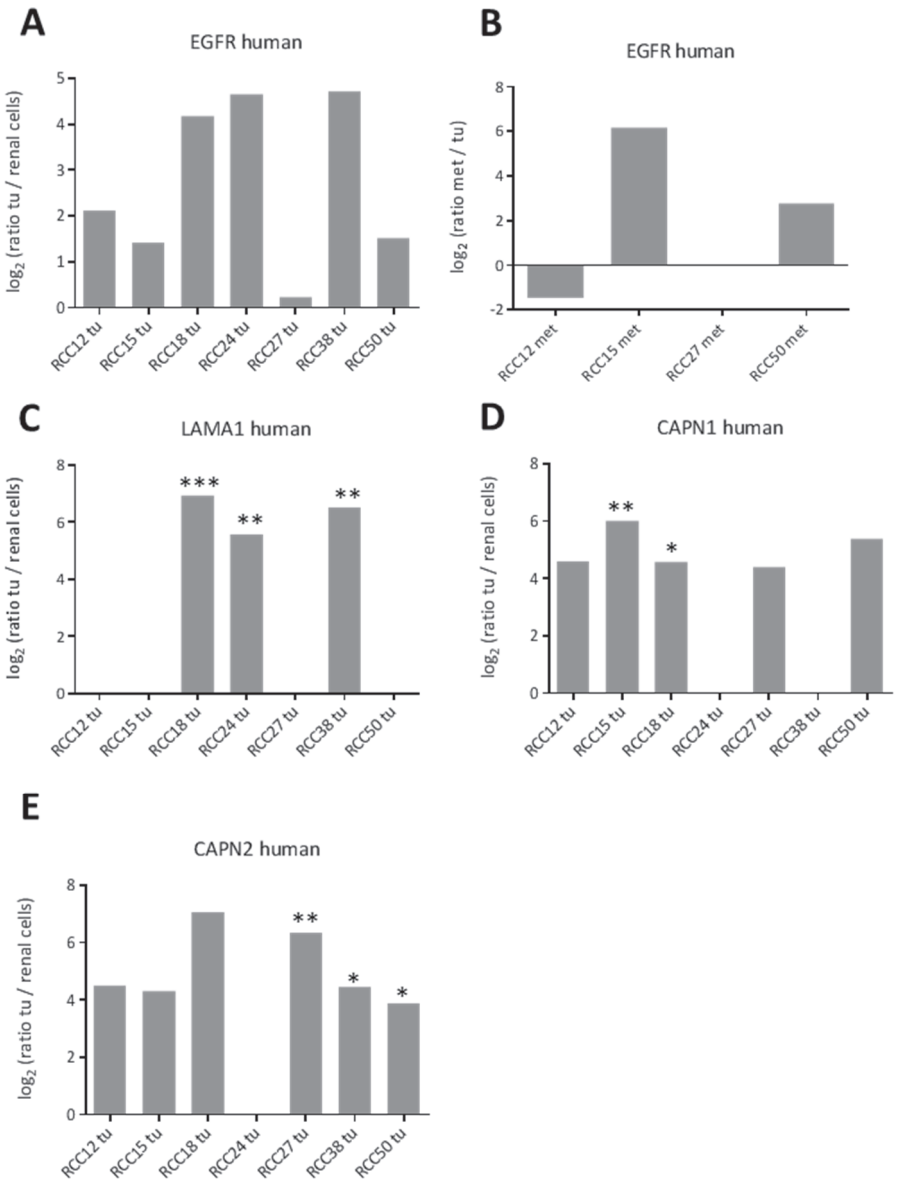


Fig. 32: Relative quantification of human proteins in xenograft kidney tumours performed by MS_QBAT analysis
Summed intensities of peptides originating from the protein of interest were compared between the xenograft tumour datasets and the human renal cell datasets. Data from lung metastases were compared to data obtained from the corresponding xenograft tumour. (p-values: < 0.001: ***; < 0.01: **; < 0.05: *; ≥ 0.05: not significant). A&B: EGFR expression was increased in xenograft tumours (except for RCC27), but expression levels were not further increased in the corresponding xenograft lung metastases. C: LAMA1 was significantly upregulated in all fast growing tumours and tumours from the RCC18 model. D: Increased levels of CAPN1 were detected in all metastasing tumours and tumours from the RCC18 model while CAPN2 levels (E) were increased in all tumours except the ones from the RCC24 model.

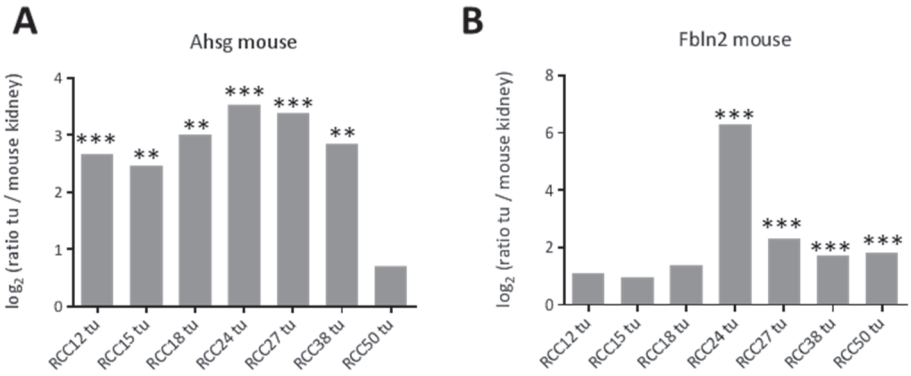


Fig. 33: Relative quantification of murine proteins in xenograft kidney tumours performed by MS$_Q$BAT analysis Summed intensities of peptides originating from the protein of interest were compared between the xenograft tumour and the murine kidney datasets. (p-values: < 0.001: ***; < 0.01: **; < 0.05: *; ≥ 0.05: not significant). A: Ahsg expression was significantly increased in all xenograft tumours except for RCC50. B: Fbln2 was upregulated in tumours of the RCC24, RCC27, RCC38 and RCC50 xenograft models.

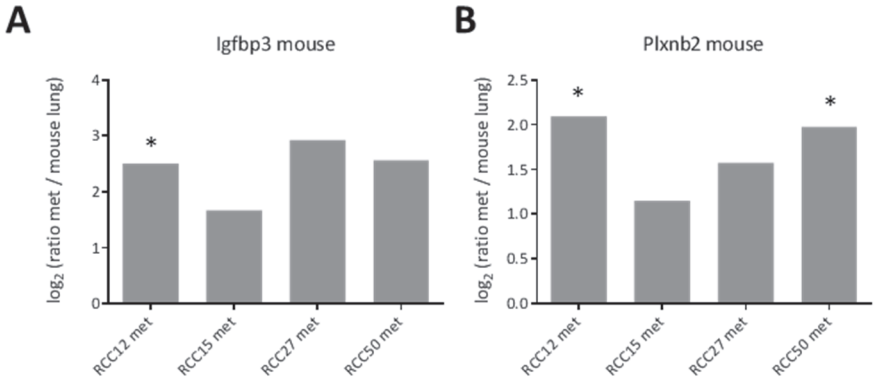


Fig. 34: Relative quantification of murine proteins in xenograft lung metastases performed by MSQBAT analysis Summed intensities of peptides originating from the protein of interest were compared between the xenograft lung metastases and murine lung datasets. (p-values: < 0.001: ***; < 0.01: **; < 0.05: *; ≥ 0.05: not significant). A&B: Expression of Igfbp3 and Plxnb2 was increased compared to their expression in healthy murine lungs.

5.5 VALIDATION OF PUTATIVE MARKERS

Following the biomarker discovery phase, selected marker candidates were validated by different orthogonal methods. To confirm both the expression levels as well as the subcellular localisations immunofluorescence analyses on cryosections were performed for the proteins TNXB and LTBP2. For other putative markers no antibody-based validation (e.g. immunofluorescence or Western blotting) could be established due to the lack of respective antibodies. Due to the properties of the deployed antibodies, immunofluorescence analyses could not be expanded to paraffin embedded samples. For that reason, selected proteins were validated based on their relative mRNA expression in RT-qPCR analyses. Tenascin-X, the most

promising marker candidate, was also validated by a biodistribution study to assess the antibody´s vascular targeting capability.

5.5.1 Validation of tenascin-X expression

The initial data analysis, which was based on counting the number of identified peptides per protein, resulted in the identification of human tenascin-X (TNXB) as protein highly upregulated in xenograft kidney tumours,. Although the final data analysis using MSQBAT did not confirm the initial findings, TNXB expression in xenograft tumours and metastases was successfully validated both on protein and on mRNA level. An initial proof-of-principle biodistribution study in a tumour bearing mouse verified TNXB as a possible marker; however the characteristics of the targeting antibody need to be improved.

5.5.1.1 Immunofluoresecene analysis of TNXB expression

The expression of the extracellular matrix protein tenascin-X in diseased tissues could be confirmed by immunofluorescence analyses. Fig. 35 presents a set of images of immunofluorescence stainings on xenograft tumours, spontaneous lung metastases in xenograft tumour-bearing mice and human patient material. Staining of tenascin-X was observed in a variable extend in all kidney tumours of the seven xenograft models and in all spontaneous lung metastases from the four metastasising models. The TNXB staining was not homogeneously distributed in the diseased tissues but rather limited to distinct sites. . No staining could be observed in any healthy control tissue. Upon omission of the primary antibody (control staining, Fig. 42), no background staining could be observed except for RCC50. Importantly, the results obtained from xenograft tissues could be transferred to primary patient material confirming TNXB as a promising marker candidate.

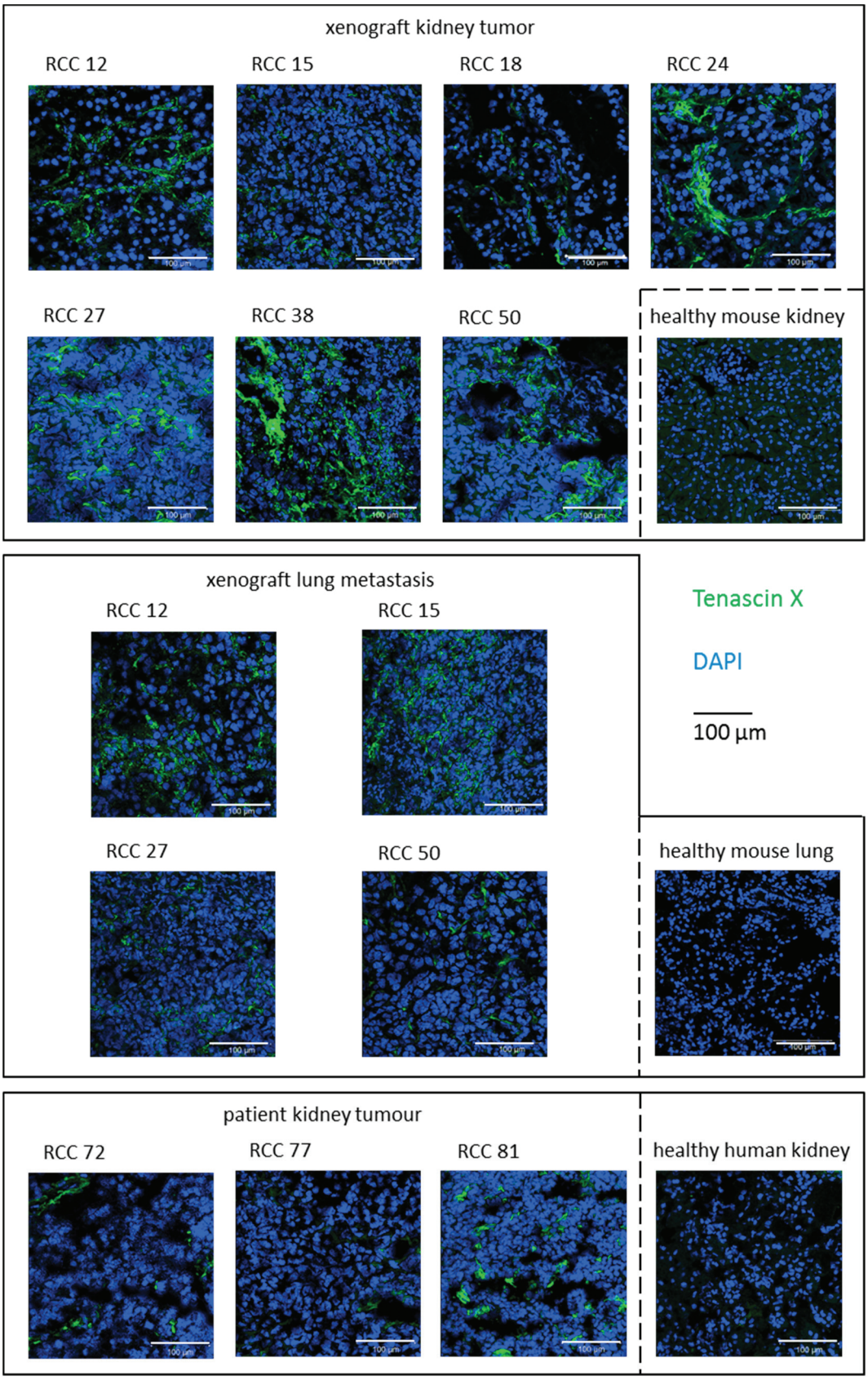


Fig. 35: Validation of TNXB expression by immunofluorescence Representative immunofluorescence images of cryosections from xenograft kidney tumours (top), xenograft lung metastases (middle) and patient derived kidney tumours (bottom) as well as their corresponding healthy controls stained with α-TNXB (green) and DAPI (blue) demonstrating the expression of TNXB in all diseased tissues. Negative controls omitting the primary antibody are shown in Fig. 42. Scale bars: 100µm.

5.5.1.2 RT-qPCR analysis of *TNXB* expression

In addition to the validation of tenascin-X protein expression by immunofluorescence analysis, the respective gene expression was analysed by RT-qPCR. The validation analysis was based on five biological replicates for each xenograft and murine tissue (primary tumour, lung metastases, kidney and lung) and three samplesfrom healthy and diseased patient kidney. All measurements were performed in triplicate (Fig. 36). In comparison to average expression levels in human renal cells (cortical epithelium, medullary epithelium and proximal tubules), *TNXB* expression was two to eight fold upregulated in xenograft tumours and the lung metastases samples RCC15, RCC27 and RCC50. However, when comparing the mRNA levels of patient tumour samples to macroscopically "healthy" kidney samples derived from ccRCC patients, no gene expression differences could be observed. These findings are contradictory to the protein expression level changes observed in the immunofluorescence analyses.

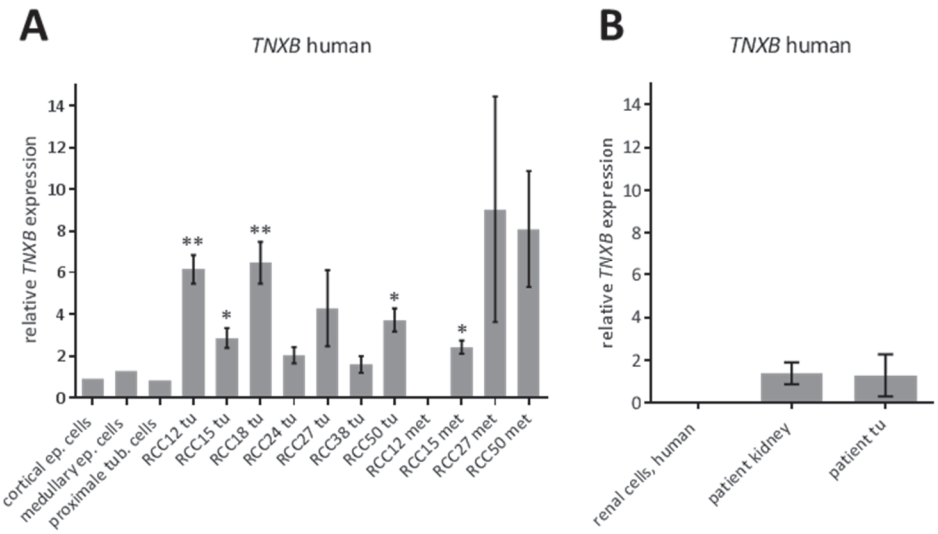


Fig. 36: Validation of TNXB expression by RT-qPCR Expression of *TNXB* in xenograft kidney tumours and lung metastases (A) as well as patient kidney tumours (B) was analysed by RT-qPCR. C_T-values were normalised to corresponding *TFRC* levels. Relative expression values were obtained by normalisation to samples deriving from the three human renal cell lines (for xenograft samples) or from macroscopically "healthy" kidney from clear cell renal cell carcinoma patients (for patient samples). Significances were calculated accordingly using an unpaired t-test (p-values: < 0.001: ***; < 0.01: **; < 0.05: *; ≥ 0.05: not significant). For xenograft samples five biological replicates were analysed in three technical replicates. For *in vitro* samples three different cell lines were analysed in three technical replicates (SEM indicated by error bars). For patient samples three individual samples from both patient kidneys and patient kidney tumours (one with metastasising status) were analysed in three technical replicates. Relative expression of *TNXB* is higher in all metastasising xenograft samples compared to healthy tissues. However, the RT-qPCR results obtained from patient kidney tumours and adjacent normal kidney tissue did not reflect the differences in staining observed in Fig. 35.

5.5.1.3 *In vivo* tumour targeting of TNXB

The vascular accessibility of an antigen utilized for antibody-based drug deliver is crucial for therapy success. Based on both the immunofluorescence and the RT-qPCR-based validation

tenascin-x was confirmed to be a potential future target in kidney tumours as well as lung metastases. To additionally validate the vascular accessibility of this protein a proof-of-principle biodistribution study was performed. Thereby fluorescein-labelled monoclonal α-TNXB antibody was injected i.v. into a xenograft kidney tumour and lung metastases bearing mouse (RCC27). 24 hours later, organs were harvested and an immunofluorescence analysis with an α-fluorescein antibody was performed (Fig. 37A). As control experiment (Fig. 37B), a healthy mouse was injected with the quenched labelling reagent NHS-fluorescein. Only very weak signals of the injected α-TNXB antibody could be detected in the xenograft tumour and no specific signal in the corresponding lung metastases. Obtained signals were so weak, that was impossible and unusually high laser intensities for the microscopic analysis were required. For this reason, signal intensities far above normal levels were found both in samples in negative control tissues. The biodistribution study confirmed the basically vascular accessibility of tenascin-X. However, the applied antibody was not able to penetrate into the tumour tissue. Due to very high background intensities no definite conclusions on the antibody distribution in normal organs could be drawn.

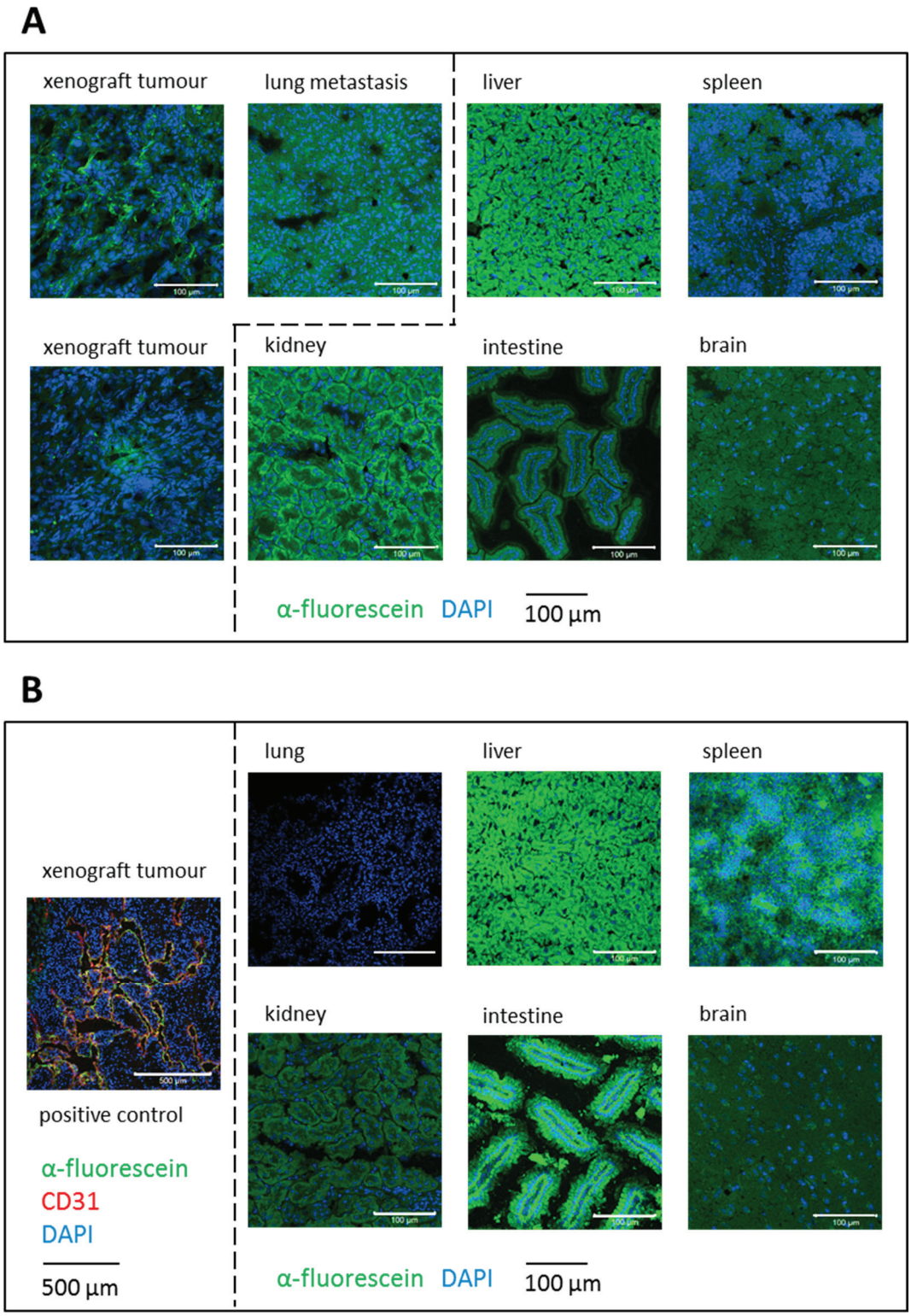


Fig. 37: *In vivo* tumour targeting analysis for an α-TNXB antibody A: Biodistribution results of a monoclonal α-TNXB antibody (clone T2H5) labelled with fluorescein. 50 µg of labelled antibody were injected in a tumour and metastases bearing xenograft mouse (RCC27). The antibody distribution in several diseased and healthy organs was analysed 24 h after injection by immunofluorescence of tissue cryosections with an α-fluorescein antibody. Exceptionally high laser intensities were required during microscopic analysis due to very low fluorescence signals in the kidney tumour resulting in high background levels in healthy organs. B: In the right panel the background analysis of the biodistribution experiment (A) is presented. Immunofluorescence staining of different murine organs after the intravenous injection of quenched antibody-labelling reagent NHS-fluoresecein. Microscope laser intensities are similar to A. In the left panel a successful antibody-based in vivo tumour

targeting (α-NID1) is presented. The injected antibody accumulated in close proximity to the vasculature (shown by α-CD31 staining). Scale bars: 100µm and 500 µm (B; left panel).

5.5.2 Validation of latent-transforming growth factor beta-binding protein 2 expression

The initial proteomics discovery study identified the murine form of latent-transforming growth factor beta-binding protein 2 (LTBP2) to be six-fold upregulated in xenograft kidney tumours compared to healthy kidney. In the following chapters initial validations of LTBP2 expression both on protein level (by immunofluorescence analysis) and on mRNA level (by RT-qPCR analysis) are presented. Additionally, *LTBP2* mRNA expression levels were determined in three patient kidney tumours.

5.5.2.1 Immunofluorescence analysis of LTBP2 expression

Immunofluorescence analysis of the extracellular matrix protein LTBP2 confirmed the expression of the murine form of the protein in diseased xenograft tissues. Fig. 38 shows representative immunofluorescence images of stainings performed for each xenograft model. According to the stainings, LTBP2 was highly expressed in RCC15, RCC18, RCC24 and RCC38 xenograft tumours and in all lung metastases from the four metastasising models. For xenograft tumours from the other three models, only weak stainings could be observed. In samples with high LTBP2 expression the staining pattern observed was forming very intense clusters. No staining could be identified in healthy control tissues. Upon omission of the primary antibody, no background staining due to the secondary α-rabbit IgG antibody could be observed (Fig. 43). Stainings of human primary patient materials confirmed the over expression of LTPB2 in kidney tumors validating the protein as a potential drug target.

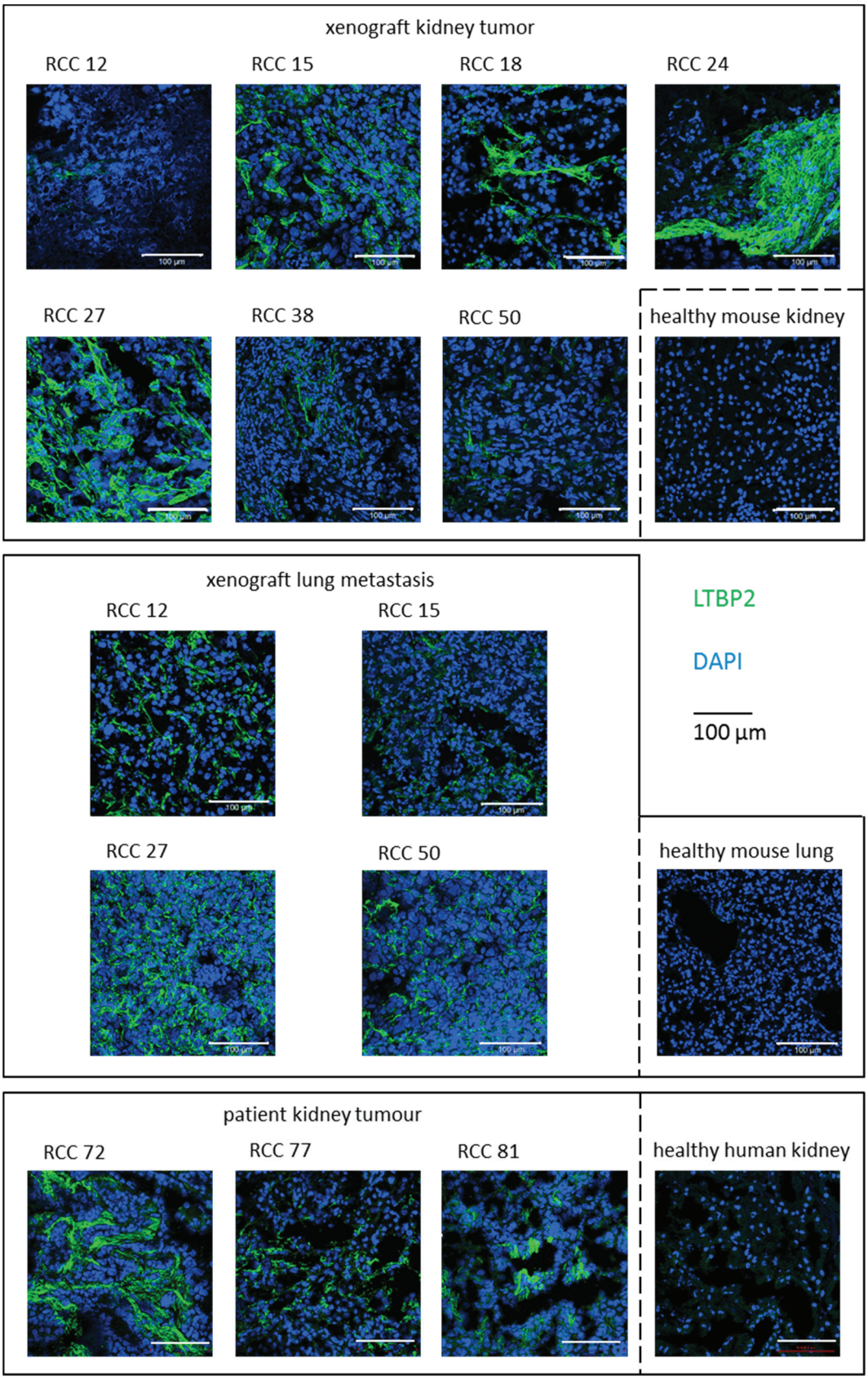


Fig. 38: Validation of Ltbp2 expression by immunofluorescence Representative immunofluorescence images of cryosections from xenograft kidney tumours (top), xenograft lung metastases (middle) and patient derived kidney tumours (bottom) as well as their corresponding healthy controls stained with α-Ltbp2 (green) and DAPI (blue) demonstrating the expression of LTBP2 in all diseased tissues. Negative controls omitting the primary antibody are shown in Fig. 43. Scale bars: 100µm.

5.5.2.2 RT-qPCR analysis of *Ltbp2/LTBP2* expression

Expression of *LTBP2* was elucidated on the mRNA level by RT-qPCR. The analysis was based on five biological replicates in case of the xenograft tissues and three samples from healthy and diseased patient kidneys. All measurements were performed in triplicates (Fig. 39). Relative expression levels were calculated in comparison to murine kidney samples (xenograft tumours), murine lungs (lung metastases) and healthy patient kidney samples (patient tumour material). mRNA expression levels of *Ltbp2* were highly upregulated in all xenograft tumours and the analysed spontaneous lung metastases. However, in contrast to the immunofluorescence validation, no increased relative *LTBP2* expression could be detected in the three examined human patient tumours.

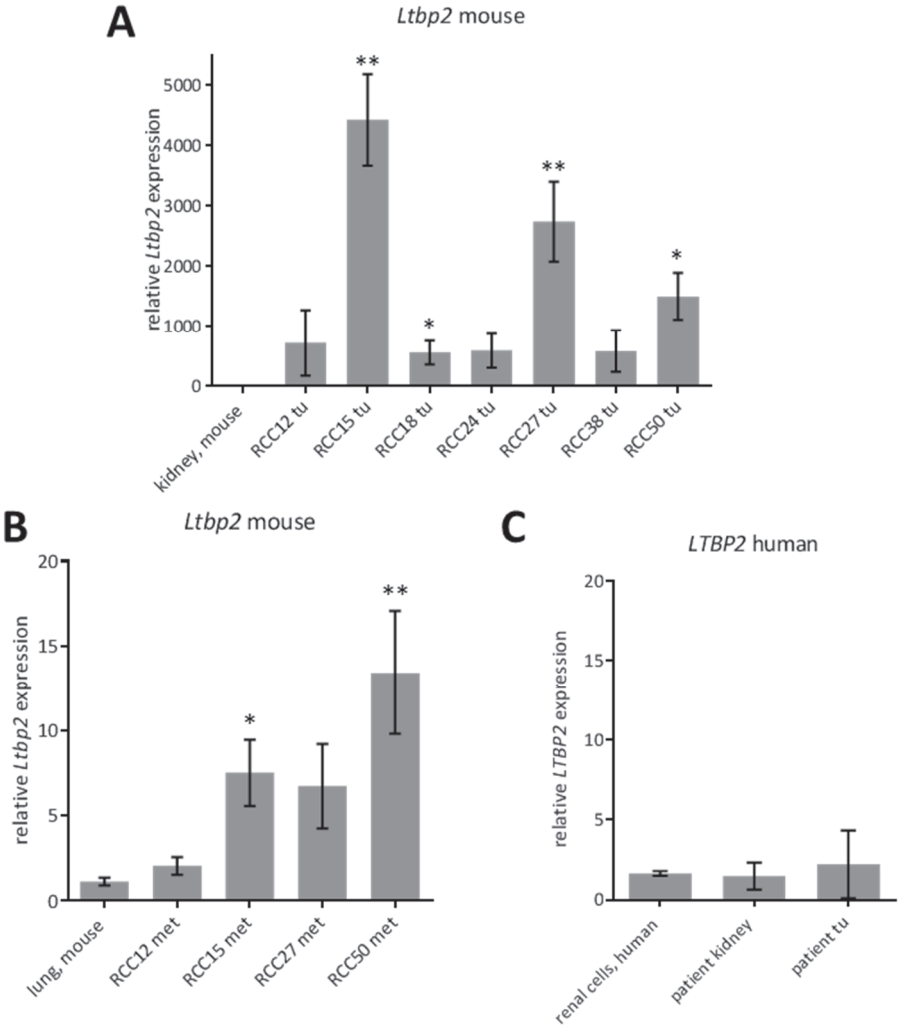


Fig. 39: Validation of Ltbp2/LTBP2 expression by RT-qPCR Expression of *Ltbp2* respectively *LTBP2* in xenograft kidney tumours (A), xenograft lung metastases (B) and patient kidney tumours (C) analysed by RT-qPCR. C_T-values were normalised to corresponding *Oaz1* (in case of *Ltbp2*) and *TFRC* (in case of *LTBP2*) levels. Relative expression values were obtained by normalisation to samples from healthy murine kidney (A), healthy murine lung (B) or macroscopically "healthy" kidney from clear cell renal cell carcinoma patients (C). Significances were calculated accordingly using an unpaired t-test (p-values: < 0.001: ***; < 0.01: **; < 0.05: *; ≥ 0.05: not significant). Five biological replicates were analysed in three technical replicates for the graphs presented in A and B. For panel C

three different human renal cell lines, three patient kidneys and three patient kidney tumours were analysed in three technical replicates (SEM indicated by error bars). Relative expression of *Ltbp2/LTBP2* is significantly higher in most metastasising xenograft kidney tumours and lung metastases compared to healthy tissues confirming the results from the initial proteomics discovery study. However, the RT-qPCR results obtained from patient kidney tumours and adjacent normal kidney tissue / renal cell lines did not reflect the differences in staining observed in Fig. 38.

5.5.3 Validation of *IGFBP3, TGFBI, SLC16A1* and *LCN2/Lcn2* expression by RT-qPCR

In this chapter, the mRNA levels of the proteins IGFBP3, TGFBI, SLC16A1 and LCN2 were analysed by RT-qPCR. The analysis was based on five biological replicates in case of the xenograft tissues and three samples from healthy and diseased patient kidneys. All measurements were performed in triplicates. The expression of the TGFβ-signalling associated *IGFBP3* and *TGFBI* genes was significantly upregulated both in the fast growing xenograft tumours RCC24 and RCC38 as well as in the low metastasising RCC12 xenograft kidney tumours (Fig. 40A-D). No increased mRNA levels were observed in xenograft lung metastases. These findings could be verified in patient samples. *IGFBP3* and *TGFBI* levels were significantly upregulated in non-metastasising kidney tumours compared to macroscopically healthy patient kidney material. Similar findings could be observed for *SLC16A1* expression (Fig. 40E&F), although no mRNA expression differences between metastasising and local kidney tumours could be determined in patient samples.

Increased mRNA levels for murine *Lcn2* were found in all seven xenograft kidney tumours and three out of four xenograft lung metastases models. The analysis of *LCN2* levels in patient tumours resulted in the identification of a 25-fold upregulation of *LCN2* mRNA in metastasising primary patient tumours compared to macroscopically “healthy” patient kidney, while no increased expression of *LCN2* could be observed in non-metastasising patient tumours.

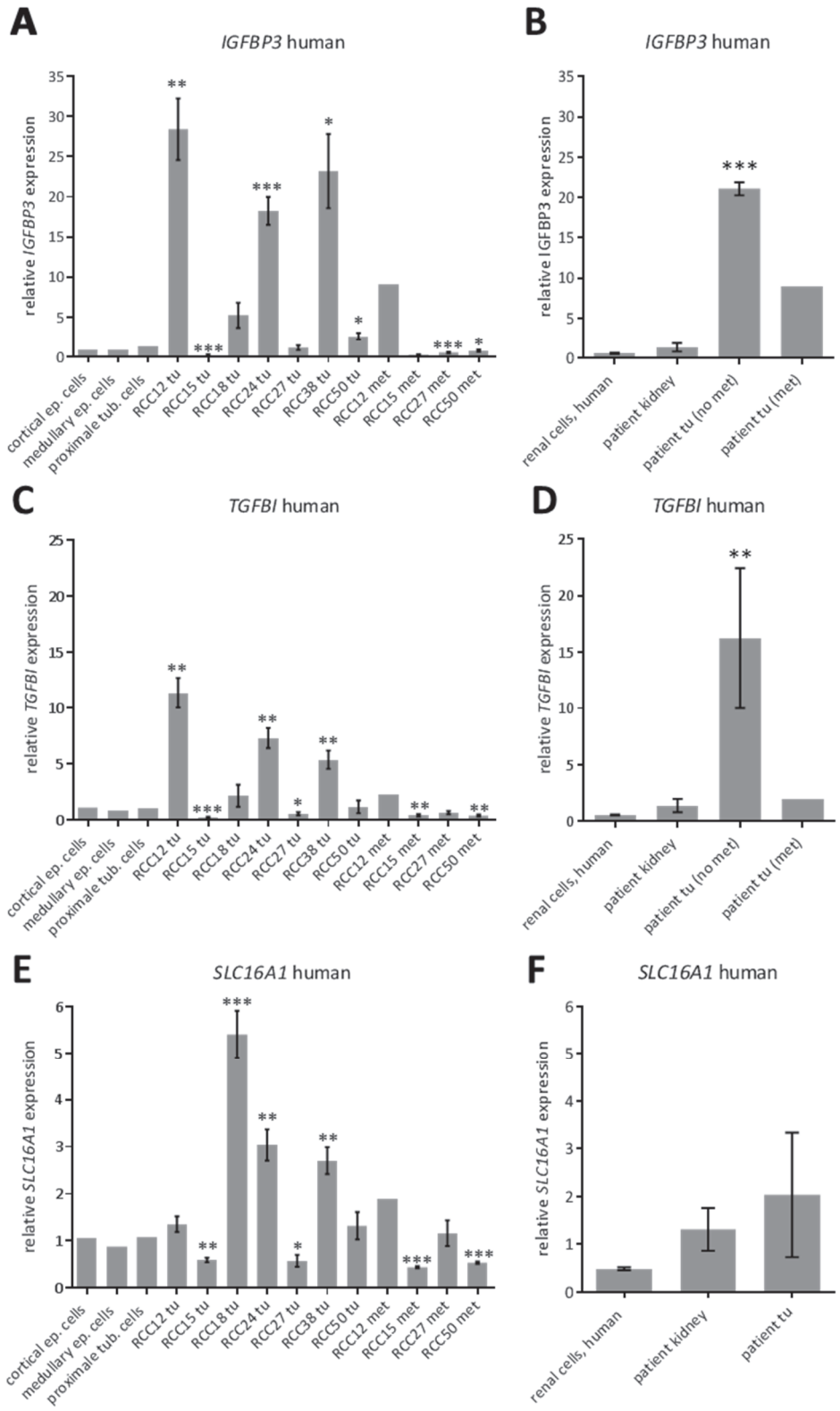


Fig. 40: Validation of *IGFBP3, TGFBI* and *SLC16A1* expression by RT-qPCR A&C&E: Expression of *IGFBP3*, *TGFBI* and *SLC16A1* in xenograft kidney tumours and xenograft lung metastases analysed by RT-qPCR. B&D&F: Expression of *IGFBP3*, *TGFBI* and *SLC16A1* in kidney tumours from patients with either local or metastasising disease status analysed by RT-qPCR. All C_T-values were normalised to the corresponding *TFRC* levels. Relative expression values were obtained by normalisation to samples of three human renal cell lines (for xenograft samples) or macroscopically "healthy" kidney from clear cell renal cell carcinoma patients (for patient tumours). Significances were calculated accordingly using an unpaired t-test (p-values: < 0.001: ***; < 0.01: **; < 0.05: *; ≥ 0.05: not significant). For xenograft samples five biological replicates were analysed in three technical replicates.

For *in vitro* samples three different cell lines were analysed in three technical replicates and for patient samples three kidneys and kidney tumours (one with metastasising status) were analysed in three technical replicates (SEM indicated by error bars). Relative expression of *IGFBP3*, *TFGBI* and *SLC16A1* is significantly higher in all non-metastasising primary kidney tumours and lung metastases compared to healthy tissues while no significant increase of expression in lung metastases compared to primary tumours could be observed. The results of the RT-qPCR analysis confirm the results obtained in the frame of the initial proteomics discovery study.

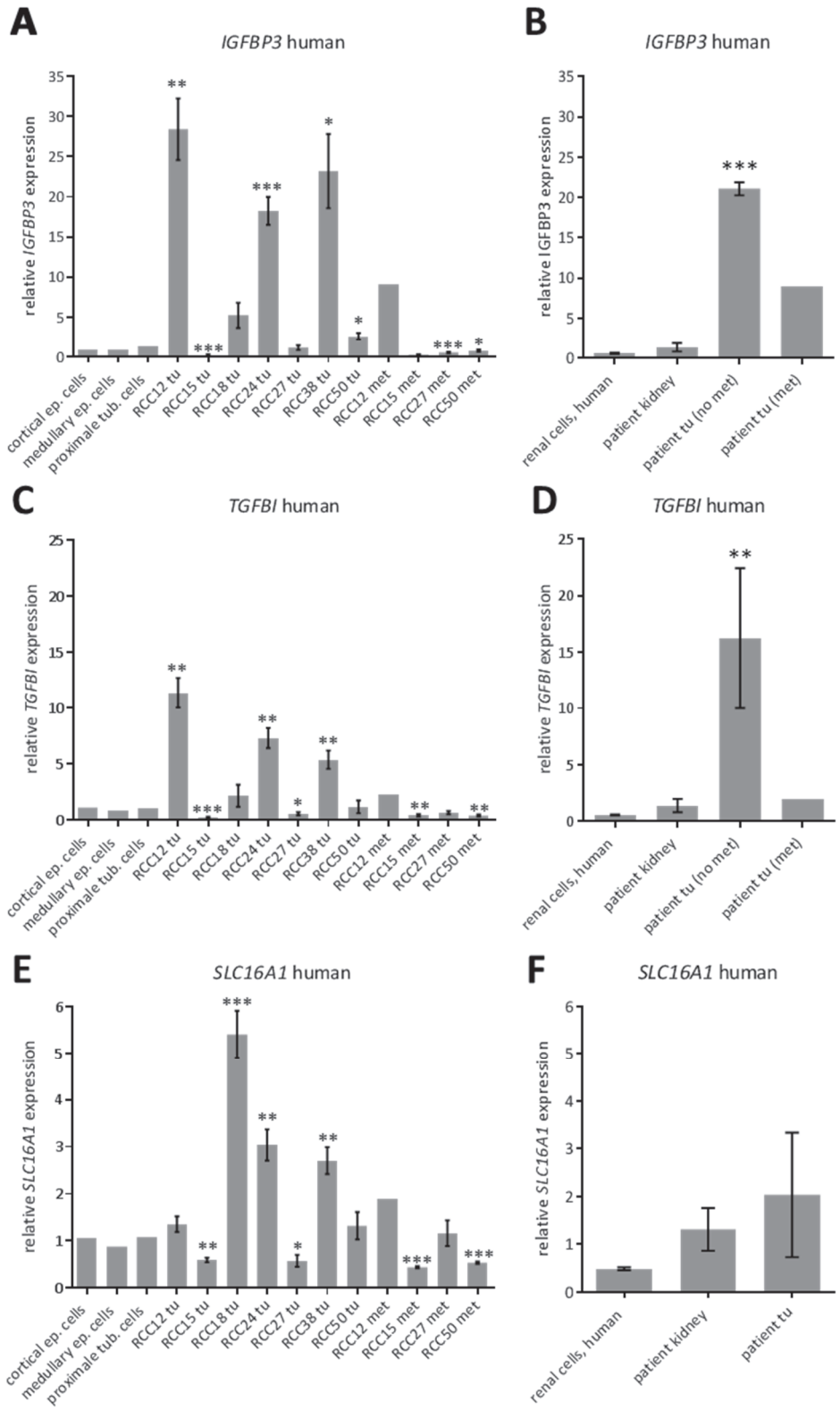


Fig. 41: Validation of Lcn2/LCN2 expression by RT-qPCR Expression of *Lcn2* respectively *LCN2* in xenograft kidney tumours (A), xenograft lung metastases (B) and kidney tumours from patients with either local or metastasising disease status (C) analysed by RT-qPCR. C_T-values were normalised to corresponding *Oaz1* (in case of *Lcn2*) and *TFRC* (in case of *LCN2*) levels. Relative expression values were obtained by normalisation to samples from healthy murine kidney (A), healthy murine lung (B) or macroscopically "healthy" kidney from clear cell renal cell carcinoma patients (C). Significances were calculated accordingly using an unpaired t-test (p-values: < 0.001: ***; < 0.01: **; < 0.05: *; ≥ 0.05: not significant). Five biological replicates were analysed in three technical replicates for the generation of the graphs displayed in panels A and B. For C three different human renal cell

lines, three patient kidneys and three patient kidney tumours (one with metastasising status) were analysed in three technical replicates (SEM indicated by error bars). Relative expression of *Lcn2/LCN2* is significantly higher in almost metastasising primary kidney tumours and lung metastases compared to healthy tissues confirming the results from the initial proteomics discovery study.

6 DISCUSSION

The work presented in this thesis aimed at the identification of novel biomarkers in clear cell renal cell carcinoma, which can be used as targets for the targeted delivery of toxic payloads directly into the tumour (e.g. by monoclonal antibodies). Both the availability of molecules exclusively expressed or highly upregulated in the target tissue and the vascular accessibility of these molecules are prerequisites for drug delivery approaches [5]. Therefore, proteins located on the cell surface of both endothelial cells and of cell in close proximity to blood vessels as well as perivascular extracellular matrix proteins are eligible candidates. In order to focus this biomarker discovery study on proteins offering these characteristics, the proteomics approach *in vivo* biotinylation was chosen. The *in vivo* biotinylation workflow is based on the vascular perfusion of tumour bearing mice with a reactive ester derivate of biotin which allows the covalent biotin-tagging of vascular accessible proteins in different organs. Following tissue homogenisation and cell lysis biotinylated proteins are enriched on streptavidin sepharose. Streptavidin-bound proteins are alkylated and delipidated on resin. Subsequently, an on resin tryptic digestion is performed and the resulting complex peptide mixture is fractionated by reverse-phase nanoUPLC. Robotically on MALDI-target plates deposited fractions are mass spectrometric analysed by MALDI-MS and MALDI-MS/MS and resulting spectra are annotated to peptide sequences which are finally mapped to their originating proteins. Performing the in vivo biotinylation approach on xenograft mouse models resulted in the identification of both human and murine proteins. In order to allow for a clear discrimination between proteins with human and murine origin, protein identifications have been only based on proteotypic peptides (i.e. peptides unique for a given protein within a given organism).

Beside the discovery of biomarkers in the primary tumour of seven patient-matched clear cell renal cell carcinoma xenograft models, the study included the analysis of spontaneous lung metastases originating in these models. Knowledge about the differential protein expression in metastases compared to primary tumours is absolutely necessary for the design of effective treatment strategies for so far untreatable advanced clear cell renal cell carcinoma. As the conventional *in vivo* biotinylation approach developed by Rybak and colleagues only allows the perfusion of the systemic circulation, the investigation of vascular accessible lung proteins was so far impossible. In this thesis, a novel *in vivo* biotinylation perfusion approach for the identification of proteins accessible from the pulmonary circulation was developed and successfully applied to four spontaneously metastasising xenograft models of clear cell renal cell carcinoma. Importantly, the newly developed lung perfusion methodology allowed for a significant enrichment of the target protein classes cell surface proteins and extracellular

matrix proteins (77% of all identified proteins). The degree of enrichment of these target protein classes was more pronounced in the lung perfusion methodology than in the original body perfusion technique, where 43% of all identified proteins were annotated to the respective protein classes.

The biomarker discovery study presented in this thesis was based on seven and four different patient derived xenograft models for the representation of primary kidney tumours and lung metastases of clear cell renal cell carcinoma, respectively. Furthermore, healthy murine kidney and lung samples were obtained from age and sex-matched control animals. In order to generate a dataset containing healthy human kidney proteins, a comparable *in vitro* biotinylation approach for the enrichment of cell surface proteins was applied to three healthy primary renal cell lines. All tissue types were analysed in several biological replicates and in at least two technical replicates. The application of the described proteomics *in vivo* biotinylation workflow resulted in the identification of 3236 proteins with at least one proteotypic peptide. Pairwise comparisons of the datasets of xenograft kidney tumours and healthy murine kidney or human renal cell samples allowed the relative quantification of identified murine or human proteins, respectively. Expression changes in lung metastases were investigated by the comparison of lung metastases datasets with either healthy murine lungs for murine proteins or with their corresponding primary xenograft kidney tumours for human proteins. In total 2446 out of the 3236 identified proteins could be quantified using the in house developed MS_QBAT software package (Alexander Kerner, HI-STEM gGmbH, unpublished).

The reliability of both the technical and the biological reproducibility of the applied proteomics workflows and the xenograft models, respectively, was thoroughly investigated. Comparative analyses of technical and biological replicates within each sub-dataset (e.g. healthy murine kidney or xenograft lung metastases) demonstrated a high reproducibility for both the workflows and the applied models. Biological variances in protein expression among the three human renal cell lines and the different culture conditions performed were shown to be negligible. As expected, the heterogeneity of the different patient-matched xenograft models was reflected in the proteomics data. Thereby, primary kidney tumour xenograft samples showed a higher variance compared to the samples derived from the spontaneous lung metastases. A hierarchical cluster analysis of the different data-subsets confirmed the relationship between kidney tumour and lung metastases samples while healthy controls subsets clustered apart.

Gene ontology analyses of murine proteins upregulated in kidney tumour and lung metastases datasets identified biological processes in response to stress and wounding as well as blood coagulation to be significantly regulated. Human proteins upregulated in xenograft tumours were mainly associated with metabolic processes. In lung metastases, cell adhesion and

migration are important upregulated biological processes both for human and murine datasets, a finding which is not surprising.

In this thesis, six biomarker candidates were selected and validated on mRNA level by RT-qPCR and two additionally on protein level by immunofluorescence analysis of tissue cryosections. The TGF-β pathway related proteins IGFBP3 and TGFBI showed an almost exclusive mRNA expression in non-metastasising kidney tumours in xenograft tumours and patient samples. Similarly, MOT1 mRNA expression was significantly increased in non-metastasising tumours. In contrast, mRNA levels of the murine protein Lcn2 were found to be increased in metastasising kidney tumours in xenograft models and patient samples. Importantly, a further upregulation of the Lcn2 expression was found in spontaneous xenograft lung metastases compared to their primary tumours. The expression of Ltbp2, a murine protein found to be upregulated in the analysed kidney tumour datasets, could be confirmed on protein level by immunofluorescence analyses as well as on mRNA level by RT-qPCR. Ltbp2 protein expression could also be observed in xenograft lung metastases and patient kidney tumours. However, mRNA levels in patient tumours were not elevated compared to the healthy controls (macroscopically unaltered tissue obtained from ccRCC patients upon radical nephrectomy). Similar findings were made for the human protein TNXB. In order to investigate the *in vivo* vascular targeting properties of the applied commercial anti-TNXB antibody, an initial biodistribution study was performed. This proof-of-principle experiment demonstrated TNXB as a potential biomarker for antibody-based therapeutics. However, major improvements on the antibody side with respect to antigen recognition in the target tissue need to be performed.

6.1 Considerations with regard to material selection

The aim of the presented biomarker discovery study was the investigation of expression changes in the vascular accessible proteome of clear cell renal carcinoma primary kidney tumours compared to the protein expression levels in healthy kidneys. Additionally, a comparison of protein expression levels between primary kidney tumours and spontaneously occurring lung metastases was performed. Ideally, such proteomic analyses are based on material recapitulating the characteristics of the human disease as close as possible while allowing the application of the *in vivo* perfusion for the biotin-labelling of vascular accessible proteins. The material, which fulfils these criteria in a best possible way while being both ethically acceptable and analysable, is obtained from the *ex vivo* perfusion of surgical resected patient kidney tumours during radical nephrectomy [484, 547]. However, work and costs of this already performed approach do not correlate with the obtained scientific benefit. The

usage of pieces originating from dissected patient tumours represents a more directly accessible alternative. Nevertheless, the quantity of material obtained from biopsies is often not sufficient to perform several technical replicates required in proteomics analyses. Additional, the vascular accessibility cannot be assessed in the analysis of tumour pieces. Artificial systems modelling distinct tumour characteristics are commonly used as a replacement of human material. An overview of such models is presented in the introduction of this thesis (chapter 2.1.7). All models are either processed *in vitro* or *in vivo* in animals. For *in vitro* cultured cells, especially when cultured conventionally in serum-based medium, genetic alterations and molecular changes, which increase with numbers of passages, are reported [617]. Additionally, an evaluation of the vascular accessibility of a given protein is impossible when using *in vitro* models. For this reason, the present study is largely based on material from *in vivo* models. The generation of the healthy human kidney protein dataset was based on *in vitro* cultured primary human renal cells due to the impossibility of acquiring healthy human kidney material. As there is no genetically engineered mouse strained available which recapitulates advanced clear cell renal cell carcinoma, orthotopic xenograft mouse models (establishes by Dr. Teresa Rigo-Watermeier, HI-STEM gGmbH, unpublished) were utilized. It has been shown that the orthotopic injection of patient derived tumour cells into animals results in the generation of xenograft tumours recapitulating the main characteristics of the disease [618]. The injection of human tumour cells into the murine kidney capsule, which offers sufficient nutrient supply, results in a highly vascularized tumour mass exposed to the same growth factors and cytokines. However, it has to be considered that murine growth factors and cytokines are not always cross-reactive to their corresponding human receptors [619]. Additionally, subsequent passaging of xenograft tumours leads to a selection process of certain features in the tumour cells which induces a selective bias. Furthermore, tumour cell engraftment is known to be correlated to the size of the primary tumour [620]. For these reasons, it might be assumed that not all different tumour characteristics of individual patients are represented in the resulting xenograft tumour. The combination of proteins expressed by human tumour cells and the murine stroma represents a major issue in sample processing as discussed in chapter 6.3. However, it allows the separate evaluation of changes in protein expression levels in tumour cells and the surrounding stroma. The usage of immunodeficient mice (NSG, no B-, T- and NK-cells [180]) facilitates on one hand tumour engraftment, but hinders on the other hand the investigation of the role of the immune system in tumour progression and metastasis [621], which should be considered for example for the development of novel immunotherapies. One of the advantages of the xenograft mouse models applied in the presented study is the spontaneous development of lung metastases. Usually, lung metastases are artificially generated by tail vein injection of tumour cells, resulting in the accumulation of these cells in the narrow vasculature of the lungs. The

analysis of spontaneous metastases allows not only the examination of cells owning the ability to settle in a new tissue (as for conventional lung metastases models) but also of cells which were able to exit the primary tumour mass, survive in the blood stream, migrate to the lungs, settle down and outgrow. In summary, xenograft models are a good alternative to primary patient materials, but will never recapitulate all characteristics of patient tumours. Importantly, the xenograft models utilized in this study were shown to recapitulate morphological, molecular and genetic properties of primary patient tumours even throughout *in vivo* passaging (data obtained from Dr. Teresa Rigo-Watermeier, HI-STEM gGmbH).

6.2 ASPECTS OF THE PROTEOMICS DISCOVERY WORKFLOW

The direct mass spectrometric analysis of peptide mixtures obtained from the enzymatic digestion of total tissue lysates is unfavourable due to the high sample complexity. Therefore, most proteomics procedures include a liquid chromatography based separation prior to mass spectrometric analysis [422]. However, due to the dynamic range in protein expression of up to seven orders of magnitude [622] the identification of low abundant proteins in total protein lysates remains a difficult task. Particularly, the identification of membrane protein is often suppressed by high abundant intracellular proteins, such as the products of many housekeeping genes. About 27% of all human proteins such as transporters and signalling receptors are predicted to be located in the plasma membrane [623]. By contrast, only 10% of the identified proteins in proteomics datasets resulting for the analyses of total protein lysates could be mapped as membrane proteins. Additionally, membrane proteins are difficult to analyse due to the presence of extended hydrophobic stretches in the amino acid sequences. To circumvent these issues additional enrichment steps for membrane proteins have to be performed [624]. Several methods such as differential centrifugation in a density gradient [518], the introduction of carbonate washings [521], and label-based enrichment strategies have been developed. In this thesis a proteomics approach based on the covalent labelling of proteins with biotin for protein enrichment was performed. As the conducted study aimed at the discovery of potential drug targets for antibody-based targeted therapy approaches in clear cell renal cell carcinoma, the *in vivo* biotinylation methodology was applied. The *in vivo* biotinylation allows not simply the enrichment of membrane proteins but enriches proteins which are accessible from the vasculature. This additional feature is crucial for the latter development of a targeted drug. The obtained datasets therefore contained mainly proteins expressed on the cell surface of murine endothelial cells or on human tumour cells in close proximity to blood vessels as well as many extracellular cellular matrix proteins. Nevertheless, also a considerable number of intracellular proteins (22-43%) involved in different metabolic processes and members of the cytoskeleton were identified. Additionally, many carboxylases, which contain biotin as a cofactor [625], were listed. A negative control

experiment performed with a complex mixture of non-biotinylated proteins resulted in the identification of several carboxylases and intracellular proteins unspecifically binding to the streptavidin sepharose resin (data not shown). Beside non-specific binding of intracellular proteins to the resin, proteins annotated to compartments without vascular accessibility could derive from necrotic and apoptotic cells. Additionally, the relocation of intracellular proteins to the cell surface and unconventional protein secretion have to be considered [626]. Finally, often subcellular protein localizations are based on predictions and not on experimental data.

The identification of intracellular proteins could also result from internalization of the charged reagent, either via biotin transporters or via passive diffusion [627]. To reduce the internalization of biotinylation reagents, chemicals differing in either their charge state and/or the linker composition have been developed. For the *in vivo* biotinylation approach performed in this work, the commercially available sulfo-NHS-LC-biotin turned out to be very useful due to its charged group and the stable and small linker. For the *in vitro* biotinylation of the three human renal cell lines three different commercially available biotin reagents with different linker characteristics (LC, SS (including a disulfide bond) and PEG_{12}) have been tested. Sulfo-NHS-SS-biotin demonstrated the highest amount of membrane proteins in the dataset (44%, 500 identified proteins, data not shown).

Following biotinylation and enrichment of vascular accessible proteins from tissues of interest, reduction/alkylation of cysteine residues with iodoacetamide and delipidation, two protein modifications often performed in proteomics experiments, were performed. Thereby disulfide bonds in proteins are reduced to allow an improved protein unfolding. Reforming of disulfide bonds is prevented by an alkylation step. Modified proteins were more efficiently digested by trypsin resulting in improved sequence coverages and higher identification rates in the subsequent mass spectrometric analysis.

Following protein modification, the captured proteins were enzymatically digested with trypsin into peptides. This protease is the most commonly used enzyme in proteomics approaches. Trypsin hydrolyses sequence specifically peptide bonds c-terminal to arginine or lysine resulting in peptides with a basic residue at the carboxyl terminus. The enzyme´s high sequence specificity is of great advantage as resulting mass spectrometric data are much easier to interpret when stringent settings can be applied for peptide generation [628]. Other sequence specific proteases, such as LysC, Asp-N and Glu-C, result in the generation of longer peptides, which are more difficult to be analysed by mass spectrometry. Due to autolysis, Ca^{2+} needs to be introduced in the digestion buffer [629]. Therefore, specific desalting steps prior to reverse phase nanoUPLC separation are required due to the fact that the presence of salts has been shown to inhibit the ionization process and suppress ion intensities in obtained mass spectra [630].

The ionization of peptides for mass spectrometric analysis can be achieved by electrospray ionization [368] or matrix assisted laser desorption/ionization [369]. In the presented proteomics discovery study MALDI ionization was applied as it allows the spiking of internal standard peptides for relative protein quantification while generating mainly singly charged ions. The possibility of uncoupling MS analysis from MS/MS acquisition allows an early evaluation of sample quality. Spectra obtained by MS acquisition were visualised in a two dimensional peptide map and samples not reaching a certain quality level were not submitted to time intensive MS/MS acquisition thereby economizing sparse mass spectrometer time.

The final dataset was based on the analysis of kidney tumours from seven different patient-matched xenograft models. Except for RCC18, eight to ten biological replicates of each model were analysed in four (RCC24) to two (others) technical replicates each. The healthy datasets were combining the analysis of ten murine kidneys in four technical replicates and five murine lungs in two technical replicates. The three human renal cell lines were investigated in two passages and culture conditions each. For the lung metastases analyses, two to nine biological replicates were analysed in two technical replicates for each of the four xenograft models yielding spontaneous lung metastases. The analysis of additional biological replicates of lung metastases samples would have strengthened the respective dataset but the perfusion of these mice has been very challenging. Reasons therefore were i) the technically highly demanding lung perfusing with a drop out rate of 30% of all animals, ii) the impossibility of performing the lung perfusion in animals with tumours growing close to the vena cava (the entrance point of the perfusion needle), and iii) the drop out of mice with numerous lung metastases in the course of the anesthesia. Post-mortem perfusion is not advised due to the immediate clogging of should small blood vessels. Last but not least, the availability of lung metastases could not always be assessed prior to the perfusion. Therefore, animals with low numbers of lung metastases not allowing for proteomics analysis had to be sorted out.

Despite the technical limitations, a total of 246 samples were analysed on the mass spectrometer requiring more than 3000 hours of instrument time. The extended analysis time resulted in an increased variability of the peptide elution profiles, identified protein numbers and sequence coverages among samples measured at different time points. Nonetheless, the comparison of technical replicates and different biological controls demonstrated a very high reproducibility of the entire workflow. For this reason, the number of technical replicates analyzed per biological sample could be reduced from four to two, significantly shortening total measurement time. The analysis of the reproducibility for *in vivo* biotinylated murine kidney samples showed that 45% of all proteins could be identified in all 39 samples. 60% of all proteins identified in one biological replicate were found in all four technical replicates and 90% of all proteins were identified in two out of the four technical replicates. These results

are excellent when compared to literature values. For a typical proteomic experiment reproducibility rates of 35-60% are reported [631]. Among the three different renal cell lines utilized as healthy kidney control, no significant biological differences on the protein level could be observed. Neither different passages nor changes in culture conditions resulted in a change in protein expression levels worth mentioning. Biological replicates derived from indiviual patient-matched xenograft models showed highly similar protein expression levels, although tumour sizes and xenograft passages varied between individual mice. Only few protein expression changes could be monitored among the seven xenograft kidney tumour models which were established from different patient tumours. Interestingly, the variances between different xenograft kidney tumour models were bigger than between different spontaneous lung metastases xenograft samples. This observation could be explained by the fact that all cells in established lung metastases feature similar characteristics which enabled them to escape from the primary tumour, to survive in the blood stream and finally to settle and outgrow in the lungs. A comparative analysis of protein expression levels in lung metastases and other metastases obtained in these xenograft tumour models should be performed in the future.

6.3 COMMENTS ON PROTEIN IDENTIFICATION AND QUANTIFICATION IN XENOGRAFT DATASETS

Xenograft mouse models are currently the best suited starting material for the discovery of vascular accessible biomarkers. However, xenograft models inherently bring along a disadvantage: xenograft tumours consist of human tumour cells surrounded by murine stroma. For this reason, both human and murine proteins with highly similar sequences will be identified in the same proteomics analysis, leading to the so called protein interference problem [632]. The protein interference problem results from the desintegration of proteins into peptides in the course of enzymatic digestions. While so called proteotypic peptides can be clearly mapped to their originating protein, the assignment of non-proteotypic peptides is impractical [632]. The handling of protein identified with a mixture of proteotypic and non-proteotypic peptides, mainly with respect to protein quantification, has to be performed with caution. It is for example not acceptable to annotate signal intensities of shared peptides with their full amount to all proteins sharing the respective peptides [633]. Today, no common rule how to deal with these non-proteotypic peptides in multi-species datasets has been established and several distinct solutions have been discussed. The first approach, which is also applied in this thesis, excludes all non-proteotypic peptides for protein identification and quantification [634]. While this approach allows to map all identified peptides to their originating protein and species, valuable data is lost. That is why the grouping of all non-proteotypic peptides corresponding to the same protein group has been proposed [635]. This

approach retains all data while every value is only counted once. Nevertheless, expression differences between proteins identified with shared peptides are difficult to detect. A third approach is based on the distribution of intensity values of shared peptides between the respective proteins according to predefined criteria such as the abundance of additional identified proteotypic peptides [636, 637]. A comparative study identified the latter approach to be the most suitable way of dealing with multi-species protein identification and quantification [633].

As indicated above, the unique usage of proteotypic peptides for protein identification and quantification was chosen within this thesis. Reasons therefore were the availability of MS_QBAT as unique quantification software for LC-MALDI-MS data which featured said approach and the need for comparative analyses of datasets deriving from a single species (healthy datasets) and two species (xenograft datasets).

6.4 Remarks on presented datasets

This thesis has been based on five datasets containing vascular accessible proteins of xenograft kidney tumours, xenograft lung metasases, murine kidney and murine lungs as well as cell surface proteins in human renal cell lines. By comparing xenograft kidney tumour datasets to their corresponding healthy control datasets (human renal cell lines for human proteins and murine kidneys for murine proteins) putative biomarkers for clear cell renal cell carcinoma could be identified. Additionally, the comparison of xenograft lung metastases to the respective kidney tumours (human proteins only) and healthy murine lungs (murine proteins only) was performed. This analysis strategy was selected in order to identify novel drug target candidates which could be used for the development of treatment strategies for advanced clear cell renal cell carcinoma. Additionally, differences in protein expression levels between metastases and primary tumours as well as proteins exclusively expressed in either of the two tissues are of interest in both clinical science and basic research.

The evaluation of protein expression level differences in the distinct datasets can be performed by different approaches. The simplest one is based on a presence-absence analysis of the identified proteins. This method may provide useful biomarker candidates, however candidates significantly regulated but present in both datasets are neglected. Another approach is counting on the evaluation of protein expression based on the number of identified peptides. Protein abundances have been correlated with peptide identifications (iBAQ values), i.e. a protein identified in one sample with less peptides than in another sample is likely regulated [204]. This method offers a reliable quantification approach for proteins producing many different mass spectrometric measurable peptides. For smaller proteins,

which produce only few tryptic peptides, this approach is less useful. Therefore, most quantification approaches are based on peptide intensities measured in mass spectrometric analyses. Different label-based and label-free methods were elucidated in the introduction of this thesis. Within the presented work, relative protein quantification was performed label-free due to the numerous sample to be processed. Importantly, in contrast to label-based methods, label-free protein quantification comes at no additional costs [358]. But foremost, today only a maximum of twelve different samples can be compared simultaneously by label-based methods [468]. Owing to sample multiplexing in label-based approaches the dynamic range of the mass spectrometer is reduced [358]. Admittedly, also label-free quantification is no easy task. The alignment of different samples is a challenging but inevitable procedure due to differences in UPLC elution profiles. Within the presented datasets, the extended period of time required for data acquisition did not facilitate the alignment process. As the software package for label-free quantification of MALDI-derived proteomics data was still under development and the presented datasets are of very complex nature only four replicates per subset were submitted to MS_QBAT-based protein quantification. Future analyses of the presented datasets should include as many biological and technical replicates as possible in order to boost the significance of the analysis. Thereby, the number of marker candidates might be enlarged.

Not surprisingly, murine proteins differentially regulated in the tumour and metastases microenvironment have been annotated to biological processes such as immune response and blood coagulation. These findings are in line with previously published *in vivo* biotinylation studys, which identified several serum components (e.g. coagulation factors and complement factors) amongst the most differentially expressed proteins [6]. It can only be speculated if these proteins originate from provisional tumour stroma which is known to contain products of blood coagulation [638] or if thrombotic events in the tumour vasculature have been biotin-tagged and enriched. However, due to the high abundance of these serum components, these proteins were not considered for pharmacological applications.

Human proteins upregulated in xenograft kidney tumours compared to human renal cell lines were mainly associated with metabolic processes and therefore with proteins mostly associated with an intracellular localisation. It might be speculated that the proteins arose due to the differences in sample preparation methods (*in vivo* biotinylation of xenograft tumors compared to *in vitro* cell surface biotinylation). *In vitro* data contained lower levels of carboxylases which are generally enriched in this approach due to their biotin cofactor. Additionally, the *in vitro* dataset does theoretically not contain any components of the extracellular matrix, whose identification is restricted to the *in vivo* biotinylation approach.

Both human and murine protein upregulated in spontaneous lung metastases have been mapped to biological processes such as cell migration, cell adhesion and system development. Cell migration as well as cell adhesion are known to be key biological processes for the development of metastases [639] and their overexpression has previously been linked to cancer metastases [640]. However, the presented proteomics data analysis did not allow to identify any human protein significantly upregulated in all lung metastases compared to the originating kidney tumour. It can be speculated if this finding is due to basic biological properties of cells in lung metastases (sharing a significant part of their characteristics and therefore also their protein expression levels with cells in the primary tumour) or if the reduced number of analysed samples did not allow to identify significantly regulated proteins. Importantly, two out of four spontaneous lung metastases xenograft models (RCC12 and RCC15) could only be analysed in biological duplicates, which was both due to time restrictions and difficulties during the perfusion experiments. Importantly, the extent between lung metastasis tissue and healthy tissue was highly variable in the analysed mice resulting in changing numbers of identified human proteins ranging from 6% (RCC15) to 21% (RCC50). Thereby, the reduced number of identified human proteins might be explained. In line with these findings were the observation of protein downregulations in RCC15 and RCC27, the two xenograft models with the lowest content of metastatic tissue. Despite all these limitations, several proteins upregulated in lung metastases compared to primary kidney tumours could be identified. EFEMP1 and EFEMP2 were found to be higher expressed in lung metastases of xenograft models deriving from RCC27 and RCC50, which are both metastasing tumours. Both proteins are binding partners of the EGF receptor inducing EGFR autophosphorylation and activation of downstream signalling pathways such as MAPK and Act [641]. Both proteins have been described to play a role in cell migration and adhesion and EFEMP1 is known to be upregulated in gliomas [642].

6.5 Discussion on biomarker validations

Several biomarker candidates for clear cell renal cell carcinoma kidney tumours were selected for detailed validation. As outlined above, candidate selection was based on an initial data analysis including the peptide counts of proteotypic and non-proteotypic peptides. The four human proteins TNXB, TGFBI, IGFBP3 and SLC16A1 and the two murine proteins Lcn2 and Ltbp2 were identified as promising candidates.

The initial validation of the selected proteins was performed by immunofluoresecence and RT-qPCR. Antibody-based immunofluorescence techniques were favoured due to the complexity of information obtained (such as expression level, subcellular localization, tissue

distribution). However, intrinsic limitations of antibody-based techniques such as antibody specificity, species cross reactivity, availability of monoclonal antibodies and high background levels have to be considered [643]. Due to these limitations, immunofluorescence validations could not be performed for SLC16A1, TGFBI and IGFBP3. Interestingly, no staining could be observed for Lcn2, although an optimized staining protocol has been developed in the lab (Laura Kuhlmann, personal communication).

The analysis of mRNA expression levels is a large-scale antibody-free validation method often used to assess protein expression levels. It has been shown lately that the correlation between mRNA expression levels and the expression levels of the respective proteins has to be carefully investigated as the mRNA translation efficacy is influenced by multiple parameters. Furthermore, mRNA turnover rates are faster than the respective rates for proteins [644].

Clear cell renal carcinoma is known for its high intratumoural heterogeneity [645]. Experimental data suggest that two third of all somatic mutations are found in restricted areas of the tumour. However, both validation approaches utilized are based on the analysis of a minor part of the diseased tissue and can therefore not extrapolated to the characteristics of the whole tumour. As for every validation study, number of samples and extent of the validation process have to be adapted accordingly. In order to circumvent locoregional expression differences, analyses based on total cell lysates of larger tumour pieces could be advantageous. Despite extensive efforts taken during this work, neither western blotting nor ELISA-based validations could be established for the selected marker candidates. Therefore, the application of an antibody-free target proteomics approach is currently considered. Single reaction monitoring would not only allow performing reliable protein quantification but also extend the sample source to serum and other body fluids.

Human tenascin X (TNXB) has been identified as the most promising protein exclusively expressed in xenograft kidney tumours and lung metastases. Immunofluorescence analyses confirmed the exclusive expression of TNXB in each xenograft model as well as in three patient kidney tumours. The respective staining was found to form distinct cluster, which could be explained by the intratumoural heterogeneity [645] of clear cell renal cell carcinoma. TNXB expression could also be confirmed on mRNA level. Interestingly, patient samples did not show an increased TNXB mRNA expression although significant protein expression level differences were confirmed by immunofluorescence analyses. Differences in mRNA and proteins expression levels have been investigated previously. Owing to the requirement of site-specifically expressed proteins for antibody-based therapies and the restricted expression of TNXB during foetal development [591], this candidate represents the most interesting biomarker for drug targeting within the dataset. Importantly, only limited findings with regard to TNXB have so far been published in comparison to the closely related protein tenascin C

(TNC). This large extracellular matrix protein is associated with adhesion modulation, motility, proliferation and differentiation [646]. Furthermore, TNC has been investigated as vascular drug target for cancer therapy [647]. It can be speculated that TNXB might also serve as a vascular target. A potential role of this protein in the growth support of epithelial tumours has already been reported [648]. Furthermore, an initial biodistribution experiment evaluating vascular targeting properties of TNXB confirmed its accessibility from the blood stream and therefore it applicability as drug target. Owing to low amounts of antibody identified in kidney tumours and the required laser intensities for imaging, which resulted in significant background signals both in the tumour and in healthy organs, exclusive targeting of the kidney tumour by the utilized α-TNXB antibody could not be stated. Antibody reformatting or the generation of a new antibody by phage-display should be considered to increase tumour infiltration rates.

Lipocalin2 (Lcn2), whereof the murine form was found, has been upregulated in xenograft tumours. Its increased expression in lung metastases as well as metastasizing patient samples could be confirmed by RT-qPCR. Antibody-based validation was not possible due to the lack of an antibody specific for the murine form and tested antibodies against human LCN2 did not show any crossreactivity with the murine form. This protein represents an interesting biomarker candidate upregulated in several cancer entities [578] in response to injury, which should be further investigated. Additionally, differential LCN2 expression has been observed in many kidney diseases for which reason this protein has been suggested as a promising drug target for several nephritic diseases [649].

Interestingly, three out of six biomarker candidates (namely Ltbp2, IGFBP3 and TGFBI) selected based on the initial proteomics discovery experiment were found to be connected to TGF-β [650]. TGF-β family members are known regulators of cell growth and extracellular matrix formation [651]. Ltbp2, whose expression was confirmed both by immunofluorescence and RT-qPCR, is a well-known biomarker for several cancer entities and other diseases including renal dysfunction [652] and has therefore been extensively patented e.g. as therapeutic target (pat. no. EP 2102357 B1).

TGFBI was the only protein in the selected marker set identified as significantly regulated in the MS_QBAT based data analysis. Subsequent validations by RT-qPCR confirmed the proteomics results. TGFBI has been previously identified as protein upregulated in clear cell renal carcinoma [609]. The two proteins TGFBI and TGF-β are furthermore targeted by dysregulated VHL n clear cell renal carcinoma. While its function in normal organs has not been fully understood, interactions with integrins [602], fibronectin and collagens [601] suggest a potential role in cell adhesion and extracellular matrix arrangements. In contrast to

the restriction of TGFBI expression to non-metastasising tumours in the presented work, a role in the promotion of metastases in colon cancer was reported [611].

IGFBP3 was found to be highly upregulated in all non-metastasing xenograft tumours and patient tumours on mRNA level, while no expression was observed in the metastasing tumours and their corresponding lung metastases. Recently, a selective inhibitory effect of IGFBP3 on extracellular matrix mediated cell adhesion but not on cell-cell interactions was observed in clear cell renal cell carcinoma [653]. Downregulation of IFGBP3 resulted in tumour cell invasion and metastases formation, confirming the findings presented in this thesis. Additionally, IGFBP3 protein levels were shown to decrease in the course of prostate cancer progression [654]. As IGFBP3 expression is upregulated in non-metastasising kidney tumours, the protein could be considered as a potential drug target. However, due to the inhibitory potential of IGFBP3 on metastases, a therapy approach which incorporates the administration of IGFBP3 could be a more successful strategy for advanced renal cell carcinoma. The targeted delivery of IGFBP3 to sites of advanced kidney cancer could result in improved patient survival. TNXB, the most promising biomarker candidate identified in this thesis, could serve as a potential target structure for such a procedure.

Several marker candidates demonstrated differential expression patterns among the different patient-matched xenograft models. LAMA1, IGFBP3, TGFBI and MOT1 were found to be highly expressed in non-metastasizing tumours and no changes in protein expression in lung metastases were observed. CAPN1, LTBP2 and Lcn2 were highly upregulated both in metastasizing kidney tumours and in their lung metastases. The LCN2 findings could be confirmed in patient samples from a metastasising kidney tumour.

Treatment response rates and survival rates of patient suffering from the same cancer entity can vary dramatically. However, tumours grouped by certain molecular markers show quite similar behaviors. Therefore the stratification of one cancer type in several subtypes has been proposed for many cancers (e.g. pancreatic cancer) [655]. Previously, also the stratification of clear cell renal cell carcinoma in the two subtypes ccA and ccB, featuring significant survival differences, was proposed [656]. Data obtained on selected marker within this thesis confirm the occurrence of two different subtypes in clear cell renal cell carcinoma. In the development process of new potential drugs this subtype stratification should be considered in order to design therapeutic strategies which can be successful in both RCC subtypes.

Above all, TNXB, which has been identified to be expressed in all analyzed xenograft tumour models and in patient tumours, might serve as a universal drug target for clear cell renal cell carcinoma.

7 CONCLUSION AND FUTURE PERSPECTIVES

In this thesis the vascular accessible proteome of primary kidney tumours and lung metastases deriving from seven and four different orthotopic xenograft mouse models, respectively, was investigated with the aim of identifying novel marker candidates applicable for antibody-based tumour targeting. In order to achieve these aims, an expansion of the recently published *in vivo* biotinylation method [6, 7], which incorporates the perfusion of the systemic circulation of tumour bearing mice with a reactive ester derivate of biotin, was developed. Based on the perfusion of the pulmonary circulation, the novel method allows for the assessment of the vascular accessible proteome in up to now unperfusable lung metastases.

The presented datasets, which have been extensively validated with respect to both biological and technical reproducibility, derive from seven different xenograft kidney tumour models, four spontaneous xenograft lung metastasis models, healthy murine kidney and lung as well as three human renal cell lines. In total, more than 3000 protein have been identified, whereof almost 2500 proteins were relatively quantified.

The initial validation of six drug target candidates was in line with data obtained in the frame of the proteomics discovery study. Human TNXB and murine Ltbp2 and Lcn2 were verified to be upregulated in xenograft kidney tumours and lung metastases. Additionally, the human protein forms LTBP2 and LCN2 could be validated on the patient tumour samples. Interestingly, LCN2 was only expressed in the metastasising patient tumour.

A proof-of-principle *in vivo* targeting experiment applying an anti-TNXB antibody lead to the conclusion that TNXB is in principle accessible for intra venously administered antibodies, confirming thereby the potential of TNXB as drug target in primary tumours of renal cell carcinoma. However, the tissue penetration properties of a clinically applicable antibody need to be improved, e.g. by developing of different antibody formats. A biodistribution study applying radionucleotide-coupled antibody and PET/CT or SPECT/CT could offer an improved and more sensitive evaluation of the antibody accumulation in different tissues.

IGFBP3 and TGFBI, two proteins connected to the TGF-β pathway, showed an increased expression level both in non-metastasising xenograft tumours and in patient material. The sixth validated protein, MOT1, was upregulated in fast growing xenograft kidney tumours and in xenograft tumours deriving from RCC18. These results could be confirmed partially in patient material.

In conclusion, the initial validation experiments demonstrated the scientific value of the obtained datasets for the selection of vascular accessible biomarker candidates for the development of antibody-based therapeutics for renal cancer.

As a result of both the ongoing development process of the quantification software MS_QBAT (Alexander Kerner, HI-STEM gGmbH, unpublished) and the high complexity of the proteomics data derived from xenograft material, the relative quantification of protein expression levels was so far only performed on a subset of the available samples. It can be speculated that upon increasing the number of samples applied for protein quantification additional marker candidates would arise.

Despite tremendous efforts taken both by academic research consortia and the industry [220, 221], high affinity monoclonal antibodies are available only for a fraction of all human proteins. In this thesis, four out of six selected candidate proteins could not be validated by antibody-based methods such as Western Blotting or immunofluorescence due to the lack of specific antibodies. Fortunately, antibody independent validation techniques have lately gained momentum and the targeted proteomics methodology “selected reaction monitoring” (SRM) [657] might allow to overcome limitations in antibody availability.

Although immunofluoresecence analyses on cryosection could be established for the two marker candidates TNXB and Ltbp2, the transfer of the stainings to formalin fixed paraffin embedded tissues was not possible. For this reason, further validations of protein expression levels in large patient cohorts were impossible since tissue micro arrays for renal cell carcinoma were only available with formalin fixed paraffin embedded tissues. In order to overcome this obstacle, the targeted proteomics approach SRM describe above could be applied. Importantly, established SRM assays could not only be used for the analysis of patient tissues but also for the screening of marker expression in serum samples.

This thesis was focusing on the identification of novel vascular accessible drug target candidates by a biomarker discovery approach. However, the obtained datasets could also be screened for the identification of predictive markers. TGFBI and IGFBP3 were found to be only expressed in non-metastasing primary kidney tumours and could therefore be validated as markers for the occurrence of metastases. MOT1, being higher expressed in fast growing, non metastasing tumours, might be a marker for this type of tumours. Future investigations should evaluate the use of the existing dataset accordingly. Importantlty, additional analyses with respect to the expression of these proteins in different renal pathologies such as nephritis and renal failure have to be conducted in order to allow for the evaluation of a potential diagnostic use of these proteins.

In the course of this biomarker discovery study, many secreted proteins were identified. One could extrapolate that proteins secreted by the tumours/metastases are also found in different body fluids and might therefore serve as predictive biomarker. Targeted proteomics approaches investigating the expression levels of such proteins in body fluids should therefore be performed. In summary, the presented proteomics datasets of vascular accessible proteins upregulated in xenograft kidney tumours and their lung metastases combined with initial validation experiments of six marker candidates confirmed their practicability for drug target discovery. TNXB was identified as the most promising future therapeutic target.

8 APPENDIX

8.1 IMMUNOFLUORESCENCE BACKGROUND STAINING OF SECONDARY ANTIBODIES

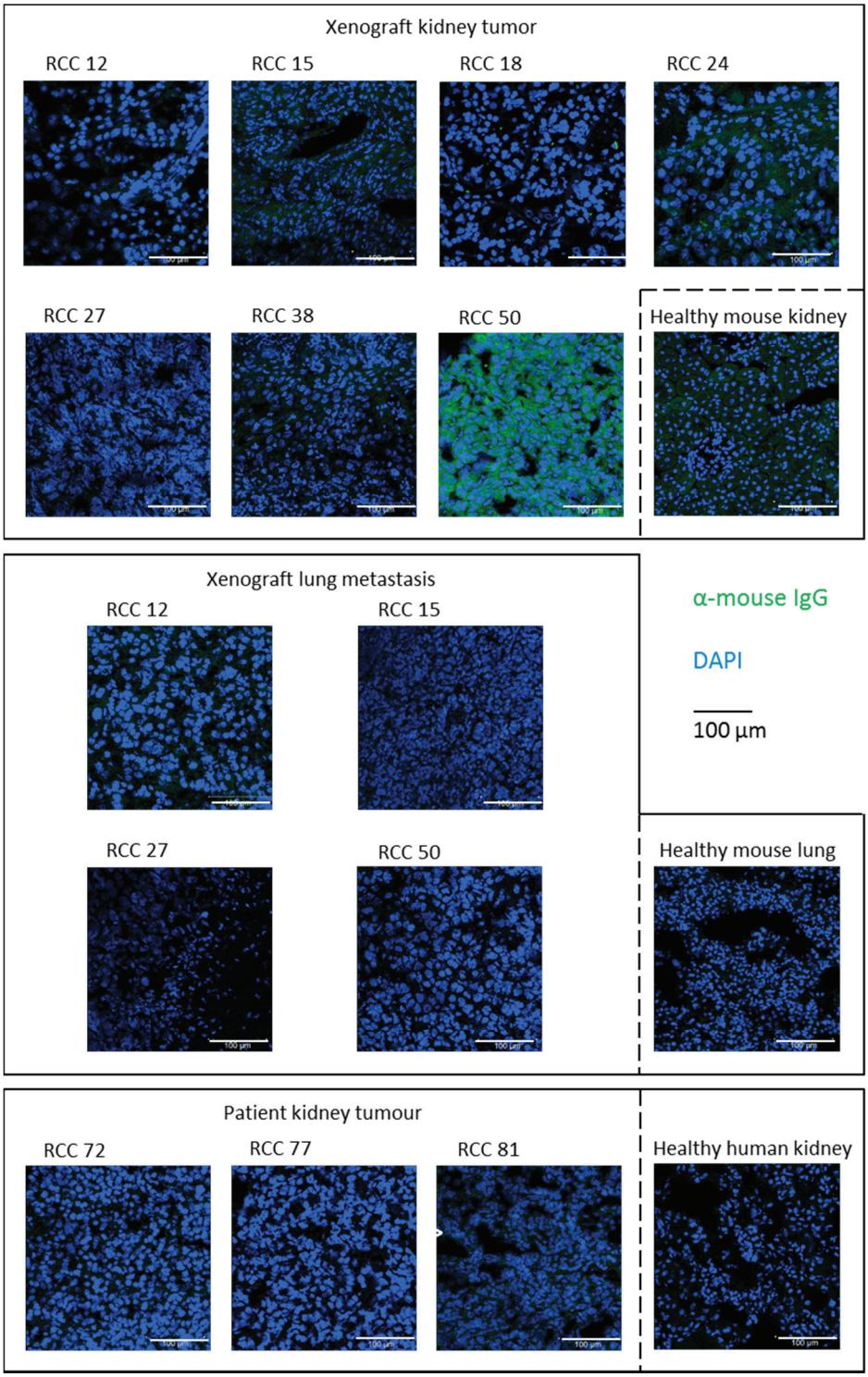


Fig. 42: Immunofluoresence staining control for α-TNXB validation Controls for α-TNXB immunofluorescence stainings shown in figure XX were obtained by omitting the primary antibody (only the secondary α-mouse IgG-Alexa488 antibody was applied). Pictures were recorded with identical microscope settings. Except for the xenograft tumour RCC50 no background staining by the secondary antibody could be observed. Scale bars: 100µm.

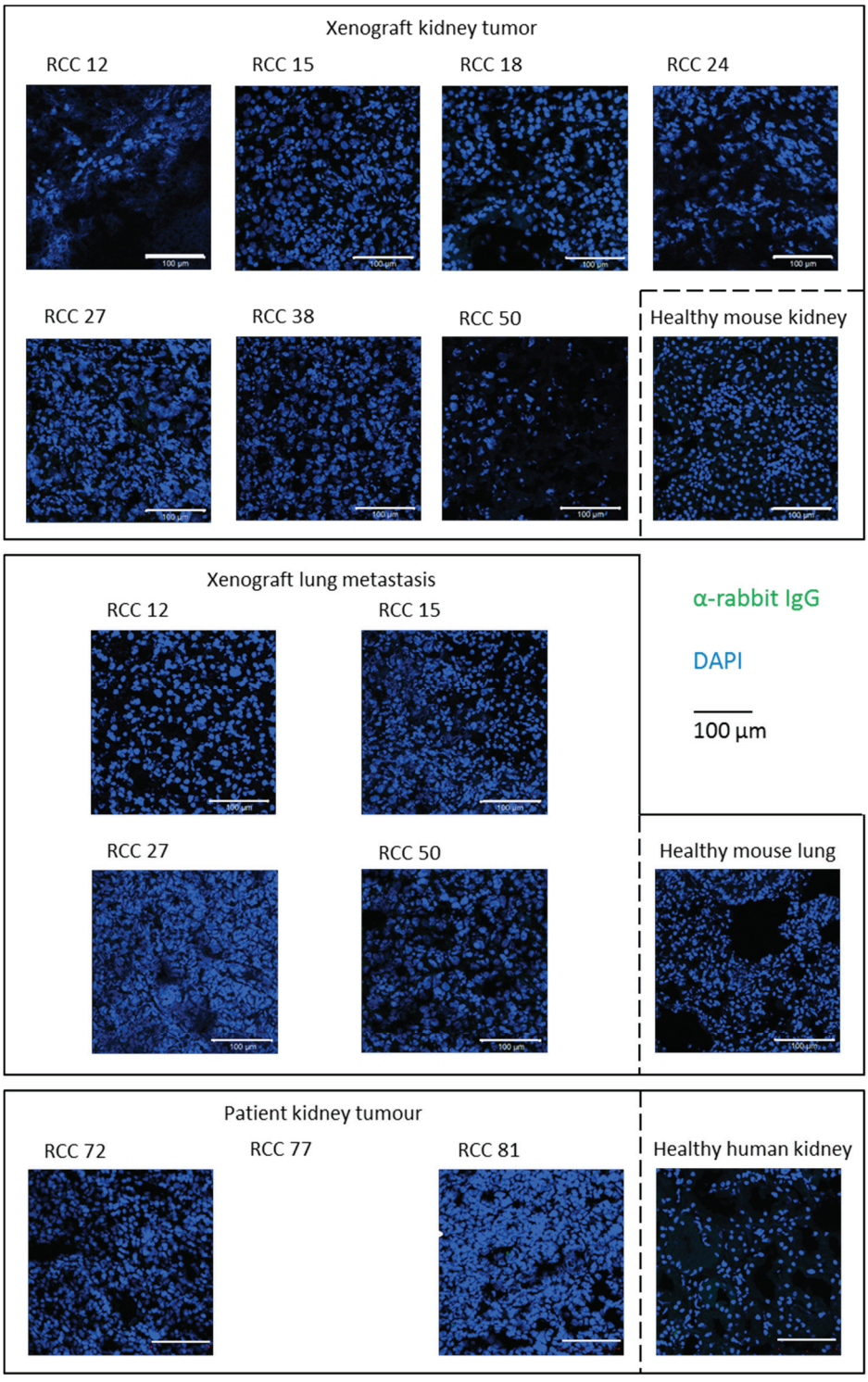


Fig. 43: Immunofluoresence staining control for α-LTBP2/Ltpb2 validation Controls for α-LTBP2/Ltbp2 immunofluorescence stainings shown in figure XX were obtained by omitting the primary antibody (only the secondary α-rabbit IgG-Alexa488 antibody was applied). Pictures were recorded with identical microscope settings. No background staining by the secondary antibody could be observed. Scale bars: 100µm.

8.2 LIST OF PROTEINS: XENOGRAFT KIDNEY TUMOURS VERSUS HEALTHY KIDNEY, MOUSE

Tab. 16: Xenograft tumours versus healthy kidney, mouse. Protein list containing all proteins quantified with at least two proteotypic peptides in at least one tumour group. Four samples from each of the seven patient groups and four samples from healthy mouse kidney are taken for quantification. Positive values indicate an upregulation in tumour tissue compared to healthy kidney tissue. Average p-value: < 0.001: ***; < 0.01: **; < 0.05: *; ≥ 0.05: not significant (n.s.).

swiss prot accession no	protein Name	gene name	$\log_2$ (average ratio)	average p-value	# tumours quantified	max. peptides tumour	max. peptides mouse kidney
Q91XL1	Leucine-rich HEV glycoprotein	Lrg1	8,5150	n.s.	7	2	0
P05366	Serum amyloid A-1 protein	Saa1	7,8500	n.s.	4	3	1
P15379	CD44 antigen	Cd44	7,4135	n.s.	2	2	0
P27659	60S ribosomal protein L3	Rpl3	7,3140	n.s.	4	2	2
P11672	Neutrophil gelatinase-associated lipocalin	Lcn2	7,2404	n.s.	5	2	2
Q60994	Adiponectin	Adipoq	6,8496	n.s.	5	2	1
P20239	Zona pellucida sperm-binding protein 2	Zp2	6,7648	n.s.	6	7	5
Q9JHH6	Carboxypeptidase B2	Cpb2	6,5570	n.s.	7	2	2
P97321	Seprase	Fap	6,5059	n.s.	3	4	0
Q60590	Alpha-1-acid glycoprotein 1	Orm1	6,2544	n.s.	7	6	3
P11370	Retrovirus-related Env polyprotein from Fv-4 locus	Fv4	6,1239	*	7	5	3
O08692	Myeloid bactenecin (F1)	Ngp	6,0827	n.s.	4	2	1
O89020	Afamin	Afm	5,6659	n.s.	7	6	3
Q61646	Haptoglobin	Hp	5,6149	***	6	10	5
P31725	Protein S100-A9	S100a9	5,5705	n.s.	7	3	2
P05367	Serum amyloid A-2 protein	Saa2	5,5626	n.s.	4	4	2
P31001	Desmin	Des	5,4510	*	3	2	1
Q62005	Zona pellucida sperm-binding protein 1	Zp1	5,3225	n.s.	2	3	2
Q61704	Inter-alpha-trypsin inhibitor heavy chain H3	Itih3	5,2428	**	7	17	6
P10761	Zona pellucida sperm-binding protein 3	Zp3	5,2379	n.s.	3	3	2
P70663	SPARC-like protein 1	Sparcl1	5,2322	*	3	2	0
P51655	Glypican-4	Gpc4	5,1424	n.s.	7	5	3
Q64726	Zinc-alpha-2-glycoprotein	Azgp1	5,0822	n.s.	4	2	1
P11859	Angiotensinogen	Agt	4,9350	n.s.	5	3	1
P03953	Complement factor D	Cfd	4,9016	n.s.	7	3	3
P31532	Serum amyloid A-4 protein	Saa4	4,8526	n.s.	3	5	2
Q91X72	Hemopexin	Hpx	4,7898	***	7	54	28
Q9JLI2	Collagen type V alpha 3 chain	Col5a3	4,7224	n.s.	3	4	2
P07361	Alpha-1-acid glycoprotein 2	Orm2	4,6195	n.s.	6	3	2
P70274	Selenoprotein P	Sepp1	4,6159	n.s.	6	6	4
P52430	Serum paraoxonase/arylesterase 1	Pon1	4,5522	n.s.	5	2	1
Q9QZJ6	Microfibrillar-associated protein 5	Mfap5	4,4875	n.s.	3	2	1
Q6S9I0	Kng2 protein	Kng2	4,4370	n.s.	7	4	2
Q91WP6	Serine protease inhibitor A3N	Serpina3n	4,4293	n.s.	7	12	5
E9PV24	Protein Fga	Fga	4,2834	n.s.	7	3	2
A2AQ53	Fibrillin-1	Fbn1	4,2658	n.s.	7	5	2
O88947	Coagulation factor X	F10	4,2642	n.s.	5	4	1
Q00897	Alpha-1-antitrypsin 1-4	Serpina1d	4,2385	n.s.	5	2	2
Q8BH61	Coagulation factor XIII A chain	F13a1	4,2290	n.s.	4	3	2
P21180	Complement C2	C2	4,1482	n.s.	7	2	2
O08677	Kininogen-1	Kng1	4,0273	**	7	16	9
P20152	Vimentin	Vim	3,9804	n.s.	6	3	3

P10107	Annexin A1	Anxa1	3,8065	n.s.	7	3	2
P37889	Fibulin-2	Fbln2	3,7992	*	7	14	10
A2AX52	Collagen alpha-4(VI) chain	Col6a4	3,7471	n.s.	5	14	6
P05977	Myosin light chain 1/3, skeletal muscle isoform	Myl1	3,7039	n.s.	1	2	0
P01831	Thy-1 membrane glycoprotein	Thy1	3,6905	n.s.	6	4	2
P33434	72 kDa type IV collagenase	Mmp2	3,6644	n.s.	6	2	2
Q07456	Protein AMBP	Ambp	3,6001	n.s.	7	11	7
P10404	MLV-related proviral Env polyprotein	null	3,5380	n.s.	5	6	1
Q8C6K9	Collagen alpha-6(VI) chain	Col6a6	3,5196	n.s.	7	7	5
Q792Z1	MCG140784	Try10	3,5030	n.s.	5	3	2
Q62000	Mimecan	Ogn	3,4876	n.s.	7	4	3
P07724	Serum albumin	Alb	3,4388	***	7	142	78
P14211	Calreticulin	Calr	3,4168	*	3	2	1
Q9QXC1	Fetuin-B	Fetub	3,4073	n.s.	6	4	2
Q9WVB4	Slit homolog 3 protein	Slit3	3,3631	n.s.	3	2	1
Q61176	Arginase-1	Arg1	3,3300	n.s.	4	2	0
P33622	Apolipoprotein C-III	Apoc3	3,3166	n.s.	5	2	2
Q6GQT1	Alpha-2-macroglobulin-P	A2mp	3,3120	n.s.	3	3	2
P12246	Serum amyloid P-component	Apcs	3,3098	n.s.	7	4	4
Q7TQ62	Podocan	Podn	3,2925	n.s.	4	2	2
Q01339	Beta-2-glycoprotein 1	Apoh	3,2544	**	7	19	13
Q02105	Complement C1q subcomponent subunit C	C1qc	3,2420	n.s.	5	2	3
P55065	Phospholipid transfer protein	Pltp	3,2377	n.s.	5	2	2
P07309	Transthyretin	Ttr	3,2075	n.s.	7	5	4
Q07968	Coagulation factor XIII B chain	F13b	3,1263	n.s.	4	2	2
O70362	Phosphatidylinositol-glycan-specific phospholipase D	Gpld1	3,1222	n.s.	7	6	6
Q99MQ4	Asporin	Aspn	3,0737	n.s.	7	4	6
P16015	Carbonic anhydrase 3	Ca3	2,9733	n.s.	4	2	1
Q61554	Fibrillin-1	Fbn1	2,9062	n.s.	7	7	3
Q80YC5	Coagulation factor XII	F12	2,9047	n.s.	7	8	7
Q921I1	Serotransferrin	Tf	2,9016	n.s.	7	3	2
Q61247	Alpha-2-antiplasmin	Serpinf2	2,8684	n.s.	7	8	7
P29699	Alpha-2-HS-glycoprotein	Ahsg	2,8590	*	7	11	6
Q62009	Periostin	Postn	2,8383	n.s.	7	9	6
P26262	Plasma kallikrein	Klkb1	2,7793	n.s.	5	2	2
Q61129	Complement factor I	Cfi	2,7699	n.s.	7	7	5
P11087	Collagen alpha-1(I) chain	Col1a1	2,7314	***	7	46	41
Q9R098	Hepatocyte growth factor activator	Hgfac	2,7060	n.s.	4	2	1
P06728	Apolipoprotein A-IV	Apoa4	2,6739	n.s.	7	15	10
P21614	Vitamin D-binding protein	Gc	2,6320	n.s.	7	11	7
Q9R0P5	Destrin	Dstn	2,5704	n.s.	7	2	2
Q8R121	Protein Z-dependent protease inhibitor	Serpina10	2,5486	n.s.	6	11	4
Q8K0E8	Fibrinogen beta chain	Fgb	2,5078	n.s.	7	27	19
Q8BND5	Sulfhydryl oxidase 1	Qsox1	2,4564	*	7	10	6
Q60847	Collagen alpha-1(XII) chain	Col12a1	2,4555	**	7	34	20
P06683	Complement component C9	C9	2,4120	n.s.	7	11	7
Q61838	Alpha-2-macroglobulin	A2m	2,3975	**	7	48	34
Q8CG14	Complement C1s-A subcomponent	C1sa	2,3817	n.s.	7	2	2
O88207	Collagen alpha-1(V) chain	Col5a1	2,3697	n.s.	7	6	5
A6X935	Inter alpha-trypsin inhibitor, heavy chain 4	Itih4	2,3571	*	7	35	22
P08121	Collagen alpha-1(III) chain	Col3a1	2,3355	**	7	42	38
Q62230	Sialoadhesin	Siglec1	2,3254	n.s.	7	13	6
P05555	Integrin alpha-M	Itgam	2,3219	n.s.	7	11	6
P35330	Intercellular adhesion molecule 2	Icam2	2,3151	n.s.	7	3	4
P19324	Serpin H1	Serpinh1	2,2241	n.s.	5	2	1
P11835	Integrin beta-2	Itgb2	2,2218	n.s.	4	2	5
Q01149	Collagen alpha-2(I) chain	Col1a2	2,2086	**	7	57	57

Q9ESB3	Histidine-rich glycoprotein	Hrg	2,1570	n.s.	7	25	20
Q08761	Vitamin K-dependent protein S	Pros1	2,1485	n.s.	5	2	2
Q9QUM0	Integrin alpha-IIb	Itga2b	2,1193	n.s.	7	5	3
Q9DBD0	Inhibitor of carbonic anhydrase	Ica	2,1138	n.s.	6	2	2
Q99K47	Fibrinogen, alpha polypeptide	Fga	2,1048	n.s.	7	39	26
Q61233	Plastin-2	Lcp1	2,1005	n.s.	3	2	3
P97807	Fumarate hydratase, mitochondrial	Fh	2,0999	n.s.	7	2	3
Q61805	Lipopolysaccharide-binding protein	Lbp	2,0805	n.s.	6	5	4
P22599	Alpha-1-antitrypsin 1-2	Serpina1b	2,0565	n.s.	7	7	5
Q9DBB9	Carboxypeptidase N subunit 2	Cpn2	2,0391	n.s.	7	10	8
P01942	Hemoglobin subunit alpha	Hba	1,9893	n.s.	6	2	2
P97298	Pigment epithelium-derived factor	Serpinf1	1,9481	n.s.	5	4	2
Q8VCM7	Fibrinogen gamma chain	Fgg	1,9398	n.s.	7	31	20
P19221	Prothrombin	F2	1,9363	n.s.	7	6	6
P07759	Serine protease inhibitor A3K	Serpina3k	1,8746	n.s.	7	12	11
Q61735	Leukocyte surface antigen CD47	Cd47	1,8718	n.s.	6	2	2
P08071	Lactotransferrin	Ltf	1,8386	n.s.	4	8	5
P29788	Vitronectin	Vtn	1,8364	n.s.	7	7	7
Q8BH35	Complement component C8 beta chain	C8b	1,8273	n.s.	6	5	4
Q3U962	Collagen alpha-2(V) chain	Col5a2	1,7799	n.s.	7	9	6
P10605	Cathepsin B	Ctsb	1,7741	n.s.	5	2	2
P49182	Heparin cofactor 2	Serpind1	1,7470	n.s.	7	5	4
A6H584	Collagen alpha-5(VI) chain	Col6a5	1,7446	n.s.	5	15	11
P20918	Plasminogen	Plg	1,7239	n.s.	7	22	19
P28654	Decorin	Dcn	1,6958	n.s.	7	10	8
Q5SWU9	Acetyl-CoA carboxylase 1	Acaca	1,6800	n.s.	7	7	5
P14106	Complement C1q subcomponent subunit B	C1qb	1,6624	n.s.	7	2	2
Q61830	Macrophage mannose receptor 1	Mrc1	1,6548	n.s.	7	6	4
Q91X70	Complement component 6	C6	1,6530	n.s.	2	2	1
P17182	Alpha-enolase	Eno1	1,6403	n.s.	7	3	4
Q00623	Apolipoprotein A-I	Apoa1	1,6377	n.s.	7	18	17
P30412	Peptidyl-prolyl cis-trans isomerase C	Ppic	1,5781	n.s.	7	2	2
P36552	Oxygen-dependent coproporphyrinogen-III oxidase, mitochondrial	Cpox	1,5777	n.s.	4	2	2
P28665	Murinoglobulin-1	Mug1	1,5332	n.s.	7	21	18
Q61702	Inter-alpha-trypsin inhibitor heavy chain H1	Itih1	1,4306	n.s.	7	26	22
Q06890	Clusterin	Clu	1,3959	n.s.	7	12	13
O70423	Membrane primary amine oxidase	Aoc3	1,3853	n.s.	7	10	10
Q80YX1	Tenascin	Tnc	1,3014	n.s.	7	11	8
P35441	Thrombospondin-1	Thbs1	1,2762	n.s.	7	5	4
P11679	Keratin, type II cytoskeletal 8	Krt8	1,2513	n.s.	7	2	2
P01029	Complement C4-B	C4b	1,2244	n.s.	7	47	39
P01898	H-2 class I histocompatibility antigen, Q10 alpha chain	H2-Q10	1,1901	n.s.	7	6	6
P82198	Transforming growth factor-beta-induced protein ig-h3	Tgfbi	1,1330	n.s.	7	7	6
P01027	Complement C3	C3	1,0894	n.s.	7	74	70
P55772	Ectonucleoside triphosphate diphosphohydrolase 1	Entpd1	1,0753	n.s.	7	7	8
P23953	Carboxylesterase 1C	Ces1c	0,9994	n.s.	7	4	5
P13020	Gelsolin	Gsn	0,9945	n.s.	7	9	8
P97290	Plasma protease C1 inhibitor	Serping1	0,9883	n.s.	7	7	5
E9Q414	Apolipoprotein B-100	Apob	0,9847	n.s.	7	13	9
P06909	Complement factor H	Cfh	0,9805	n.s.	7	26	25
Q00612	Glucose-6-phosphate 1-dehydrogenase X	G6pdx	0,9774	n.s.	2	2	1
Q6PB66	Leucine-rich PPR motif-containing protein, mitochondrial	Lrpprc	0,9683	n.s.	4	2	2
P08226	Apolipoprotein E	Apoe	0,9196	n.s.	7	12	10
Q8R2Y2	Cell surface glycoprotein MUC18	Mcam	0,8972	n.s.	7	9	9
P11276	Fibronectin	Fn1	0,8919	n.s.	7	59	46
Q61147	Ceruloplasmin	Cp	0,8918	n.s.	7	32	30

Q640N1	Adipocyte enhancer-binding protein 1	Aebp1	0,8529	n.s.	6	3	3
P13595	Neural cell adhesion molecule 1	Ncam1	0,7975	n.s.	7	3	3
P51885	Lumican	Lum	0,7588	n.s.	7	7	8
Q61406	Complement factor H-related 1	Cfhr1	0,7237	n.s.	4	2	2
P70232	Neural cell adhesion molecule L1-like protein	Chl1	0,6359	n.s.	1	3	2
P19137	Laminin subunit alpha-1	Lama1	0,6320	***	7	6	17
Q61703	Inter-alpha-trypsin inhibitor heavy chain H2	Itih2	0,5320	n.s.	7	8	9
Q91Y63	Solute carrier family 13 member 3	Slc13a3	0,4628	n.s.	7	3	4
Q08879	Fibulin-1	Fbln1	0,4465	n.s.	7	6	4
Q9JK53	Prolargin	Prelp	0,4381	n.s.	7	7	6
P04104	Keratin, type II cytoskeletal 1	Krt1	0,4198	n.s.	7	2	3
O55026	Ectonucleoside triphosphate diphosphohydrolase 2	Entpd2	0,3708	n.s.	5	2	3
O35082	Klotho	Kl	0,3609	n.s.	7	2	3
P11688	Integrin alpha-5	Itga5	0,2731	n.s.	7	4	3
Q8BPB5	EGF-containing fibulin-like extracellular matrix protein 1	Efemp1	0,2185	n.s.	7	5	3
P62962	Profilin-1	Pfn1	0,2107	n.s.	7	3	3
Q8K182	Complement component C8 alpha chain	C8a	0,1967	n.s.	7	5	4
Q3UEG6	Alanine--glyoxylate aminotransferase 2, mitochondrial	Agxt2	0,1691	n.s.	7	2	3
P32261	Antithrombin-III	Serpinc1	0,1591	n.s.	7	11	11
Q9WVT6	Carbonic anhydrase 14	Ca14	0,1374	n.s.	7	4	7
P16294	Coagulation factor IX	F9	0,1014	n.s.	6	4	2
Q99K41	EMILIN-1	Emilin1	0,0968	n.s.	7	16	16
Q8VCG4	Complement component C8 gamma chain	C8g	0,0775	n.s.	7	3	3
P56528	ADP-ribosyl cyclase 1	Cd38	0,0110	n.s.	7	3	5
Q8VCC9	Spondin-1	Spon1	-0,0236	n.s.	7	2	3
P11627	Neural cell adhesion molecule L1	L1cam	-0,0268	n.s.	7	4	5
Q61581	Insulin-like growth factor-binding protein 7	Igfbp7	-0,0401	n.s.	7	4	3
P02463	Collagen alpha-1(IV) chain	Col4a1	-0,0613	n.s.	7	3	3
P48036	Annexin A5	Anxa5	-0,0620	n.s.	6	3	2
O08573	Galectin-9	Lgals9	-0,1205	n.s.	6	2	2
P40124	Adenylyl cyclase-associated protein 1	Cap1	-0,1549	n.s.	7	2	2
O88428	Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthase 2	Papss2	-0,2035	n.s.	7	2	4
P52480	Pyruvate kinase PKM	Pkm	-0,2488	n.s.	7	3	3
Q8BTM8	Filamin-A	Flna	-0,2574	n.s.	7	2	2
P17751	Triosephosphate isomerase	Tpi1	-0,2652	n.s.	7	3	2
P21981	Protein-glutamine gamma-glutamyltransferase 2	Tgm2	-0,3595	n.s.	7	7	8
Q01279	Epidermal growth factor receptor	Egfr	-0,3941	n.s.	7	7	7
Q08481	Platelet endothelial cell adhesion molecule	Pecam1	-0,4190	n.s.	7	8	9
P28653	Biglycan	Bgn	-0,4399	n.s.	7	2	2
Q8CI59	Metalloreductase STEAP3	Steap3	-0,4570	n.s.	7	5	5
O09164	Extracellular superoxide dismutase [Cu-Zn]	Sod3	-0,4654	n.s.	7	4	7
Q8BKG3	Inactive tyrosine-protein kinase 7	Ptk7	-0,4877	n.s.	7	5	6
Q4LDF6	Protein Cfhr2	Cfhr2	-0,4981	n.s.	7	3	3
P97742	Carnitine O-palmitoyltransferase 1, liver isoform	Cpt1a	-0,5038	*	6	3	8
Q04857	Collagen alpha-1(VI) chain	Col6a1	-0,5464	n.s.	7	13	14
J3QQ16	Protein Col6a3	Col6a3	-0,5794	n.s.	7	2	3
Q9CQX2	Cytochrome b5 type B	Cyb5b	-0,5861	n.s.	7	2	2
Q8K183	Pyridoxal kinase	Pdxk	-0,5938	n.s.	7	3	7
O88322	Nidogen-2	Nid2	-0,5997	n.s.	7	42	29
Q08857	Platelet glycoprotein 4	Cd36	-0,6035	n.s.	7	4	5
Q80X19	Collagen alpha-1(XIV) chain	Col14a1	-0,6511	n.s.	7	29	31
Q61789	Laminin subunit alpha-3	Lama3	-0,6587	n.s.	7	3	6
P08122	Collagen alpha-2(IV) chain	Col4a2	-0,6591	*	7	14	17
Q9CZR2	N-acetylated-alpha-linked acidic dipeptidase 2	Naalad2	-0,6739	n.s.	7	9	10
Q8R2Z5	von Willebrand factor A domain-containing protein 1	Vwa1	-0,6940	n.s.	7	6	4
P29391	Ferritin light chain 1	Ftl1	-0,7195	n.s.	6	2	2

Q99MN9	Propionyl-CoA carboxylase beta chain, mitochondrial	Pccb	-0,7355	n.s.	5	2	6
Q02788	Collagen alpha-2(VI) chain	Col6a2	-0,7650	n.s.	7	19	19
P97927	Laminin subunit alpha-4	Lama4	-0,7658	n.s.	7	16	13
P97429	Annexin A4	Anxa4	-0,7696	n.s.	7	2	2
Q62469	Integrin alpha-2	Itga2	-0,7943	n.s.	7	2	2
Q9QZS0	Collagen alpha-3(IV) chain	Col4a3	-0,8434	n.s.	7	3	5
Q9CQW3	Vitamin K-dependent protein Z	Proz	-0,8571	n.s.	1	3	2
P51660	Peroxisomal multifunctional enzyme type 2	Hsd17b4	-0,8690	n.s.	7	4	3
Q19LI2	Alpha-1B-glycoprotein	A1bg	-0,8726	n.s.	7	10	11
O89103	Complement component C1q receptor	Cd93	-0,8790	n.s.	7	5	4
E9PWQ3	Protein Col6a3	Col6a3	-0,9144	n.s.	7	17	21
P99029	Peroxiredoxin-5, mitochondrial	Prdx5	-0,9292	n.s.	7	3	4
P10493	Nidogen-1	Nid1	-1,0347	*	7	48	42
P19157	Glutathione S-transferase P 1	Gstp1	-1,0892	n.s.	7	2	2
P70389	Insulin-like growth factor-binding protein complex acid labile subunit	Igfals	-1,1108	n.s.	7	7	10
P14824	Annexin A6	Anxa6	-1,1157	n.s.	7	2	3
Q9WVH9	Fibulin-5	Fbln5	-1,1271	n.s.	7	3	3
Q91VI7	Ribonuclease inhibitor	Rnh1	-1,1303	n.s.	7	2	2
Q9Z0K8	Pantetheinase	Vnn1	-1,1503	n.s.	7	2	3
Q8K4G1	Latent-transforming growth factor beta-binding protein 4	Ltbp4	-1,2084	n.s.	7	5	6
P09528	Ferritin heavy chain	Fth1	-1,2466	n.s.	7	3	3
P10810	Monocyte differentiation antigen CD14	Cd14	-1,2487	n.s.	7	3	4
P21956	Lactadherin	Mfge8	-1,2643	n.s.	7	9	9
Q60930	Voltage-dependent anion-selective channel protein 2	Vdac2	-1,2866	n.s.	7	3	3
Q64449	C-type mannose receptor 2	Mrc2	-1,3117	n.s.	7	5	5
P01902	H-2 class I histocompatibility antigen, K-D alpha chain	H2-K1	-1,3212	n.s.	7	9	11
Q8VCA8	Secernin-2	Scrn2	-1,3215	n.s.	7	3	4
P43406	Integrin alpha-V	Itgav	-1,3285	n.s.	7	4	6
P09055	Integrin beta-1	Itgb1	-1,3604	n.s.	7	12	13
O35206	Collagen alpha-1(XV) chain	Col15a1	-1,4043	n.s.	7	8	9
P06151	L-lactate dehydrogenase A chain	Ldha	-1,4380	n.s.	5	2	3
Q64735	Complement component receptor 1-like protein	Cr1l	-1,4475	n.s.	7	3	5
Q61490	CD166 antigen	Alcam	-1,4476	n.s.	7	2	3
Q3TNA1	Xylulose kinase	Xylb	-1,4807	n.s.	7	6	7
Q99MR8	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial	Mccc1	-1,5625	***	7	31	41
P10649	Glutathione S-transferase Mu 1	Gstm1	-1,5820	n.s.	6	2	4
Q3UQ28	Peroxidasin homolog	Pxdn	-1,6006	n.s.	7	6	6
Q9WV91	Prostaglandin F2 receptor negative regulator	Ptgfrn	-1,6268	n.s.	7	3	3
Q61508	Extracellular matrix protein 1	Ecm1	-1,6283	n.s.	7	2	3
Q91VC4	Plasmalemma vesicle-associated protein	Plvap	-1,6343	n.s.	7	8	11
Q9JII6	Alcohol dehydrogenase [NADP(+)]	Akr1a1	-1,6390	n.s.	7	6	5
P39061	Collagen alpha-1(XVIII) chain	Col18a1	-1,6519	*	7	17	19
P46935	E3 ubiquitin-protein ligase NEDD4	Nedd4	-1,6632	n.s.	2	2	1
B1B0C7	Basement membrane-specific heparan sulfate proteoglycan core protein	Hspg2	-1,6695	*	7	26	24
P56379	6.8 kDa mitochondrial proteolipid	Mp68	-1,6821	n.s.	6	2	3
Q8BY89	Choline transporter-like protein 2	Slc44a2	-1,6907	*	7	6	9
G3X922	MCG115602	Dnajc13	-1,6940	*	7	2	4
Q61001	Laminin subunit alpha-5	Lama5	-1,7217	***	7	53	67
Q8C4U3	Secreted frizzled-related protein 1	Sfrp1	-1,7519	n.s.	7	2	3
P02469	Laminin subunit beta-1	Lamb1	-1,7600	**	7	15	19
O88792	Junctional adhesion molecule A	F11r	-1,7650	n.s.	7	3	4
Q3V3R4	Integrin alpha-1	Itga1	-1,7788	n.s.	7	8	11
Q64471	Glutathione S-transferase theta-1	Gstt1	-1,8766	n.s.	7	2	4
Q61292	Laminin subunit beta-2	Lamb2	-1,8888	*	7	36	37

O35452	Protein Tnxb	Tnxb	-1,8960	*	7	17	20
Q8VDD5	Myosin-9	Myh9	-1,9024	n.s.	7	5	9
O88968	Transcobalamin-2	Tcn2	-1,9140	n.s.	7	2	2
D3YWD1	Protein Col6a3 (Fragment)	Col6a3	-1,9237	n.s.	7	22	22
Q62165	Dystroglycan	Dag1	-1,9411	n.s.	7	3	4
Q91ZA3	Propionyl-CoA carboxylase alpha chain, mitochondrial	Pcca	-1,9412	***	7	24	33
Q9CQI6	Coactosin-like protein	Cotl1	-1,9646	n.s.	7	3	3
Q6PGJ3	L1 cell adhesion molecule	L1cam	-2,0000	n.s.	6	2	3
P04919	Band 3 anion transport protein	Slc4a1	-2,0265	n.s.	7	4	4
Q6DYE8	Ectonucleotide pyrophosphatase/ phosphodiesterase family member 3	Enpp3	-2,0344	n.s.	7	2	4
Q925F2	Endothelial cell-selective adhesion molecule	Esam	-2,0847	n.s.	7	6	6
P58735	Sulfate anion transporter 1	Slc26a1	-2,0903	*	7	3	8
Q60675	Laminin subunit alpha-2	Lama2	-2,0923	**	7	12	15
P26039	Talin-1	Tln1	-2,0959	n.s.	7	2	3
Q9QZF2	Glypican-1	Gpc1	-2,1174	n.s.	7	3	3
Q9R069	Basal cell adhesion molecule	Bcam	-2,1306	n.s.	7	14	17
P09470	Angiotensin-converting enzyme	Ace	-2,1359	n.s.	7	3	3
Q05793	Basement membrane-specific heparan sulfate proteoglycan core protein	Hspg2	-2,1725	***	7	100	99
Q8VHY0	Chondroitin sulfate proteoglycan 4	Cspg4	-2,1777	***	7	15	25
P97333	Neuropilin-1	Nrp1	-2,2002	n.s.	7	4	6
P55284	Cadherin-5	Cdh5	-2,2325	n.s.	7	6	8
Q62470	Integrin alpha-3	Itga3	-2,2510	n.s.	7	9	14
Q61738	Integrin alpha-7	Itga7	-2,3446	n.s.	7	3	3
Q9DCY0	Glycine N-acyltransferase-like protein Keg1	Keg1	-2,3448	n.s.	7	4	4
O35728	Cytochrome P450 4A14	Cyp4a14	-2,3544	n.s.	7	2	5
Q91XE4	N-acyl-aromatic-L-amino acid amidohydrolase (carboxylate-forming)	Acy3	-2,3868	*	7	2	5
G5E8Q8	MCG115189	Gpr116	-2,4109	n.s.	6	2	2
O54901	OX-2 membrane glycoprotein	Cd200	-2,4197	n.s.	7	3	3
Q9QXE0	2-hydroxyacyl-CoA lyase 1	Hacl1	-2,4277	n.s.	7	2	4
Q9D1A2	Cytosolic non-specific dipeptidase	Cndp2	-2,4582	n.s.	7	2	3
Q99J99	3-mercaptopyruvate sulfurtransferase	Mpst	-2,4624	n.s.	5	2	3
G3UWD9	Protein Slc17a3	Slc17a3	-2,4624	n.s.	5	2	3
P23492	Purine nucleoside phosphorylase	Pnp	-2,4773	n.s.	5	2	2
Q9EQ20	Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial	Aldh6a1	-2,5157	n.s.	7	4	7
Q99N23	Carbonic anhydrase 15	Ca15	-2,6043	n.s.	7	3	4
O70309	Integrin beta-5	Itgb5	-2,6077	n.s.	7	3	4
Q99JR5	Tubulointerstitial nephritis antigen-like	Tinagl1	-2,6107	n.s.	7	7	7
Q78PY7	Staphylococcal nuclease domain-containing protein 1	Snd1	-2,6438	n.s.	6	2	2
P02468	Laminin subunit gamma-1	Lamc1	-2,6699	***	7	21	31
P28843	Dipeptidyl peptidase 4	Dpp4	-2,7014	*	7	5	10
Q8BGN3	Ectonucleotide pyrophosphatase/ phosphodiesterase family member 6	Enpp6	-2,7077	n.s.	7	2	5
A6H6E2	Multimerin-2	Mmrn2	-2,7149	n.s.	7	8	10
P45952	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial	Acadm	-2,7156	*	7	2	6
O35632	Hyaluronidase-2	Hyal2	-2,7538	n.s.	7	5	4
Q99JW5	Epithelial cell adhesion molecule	Epcam	-2,7793	n.s.	7	3	5
P35700	Peroxiredoxin-1	Prdx1	-2,8152	n.s.	7	2	3
Q3UUY6	Prominin-2	Prom2	-2,8235	*	7	3	6
P47911	60S ribosomal protein L6	Rpl6	-2,8530	n.s.	7	3	3
Q9R0H0	Peroxisomal acyl-coenzyme A oxidase 1	Acox1	-2,8605	n.s.	7	2	3
Q9JI33	Netrin-4	Ntn4	-2,8626	*	7	3	6
Q60928	Gamma-glutamyltranspeptidase 1	Ggt1	-2,8701	n.s.	7	10	10

P13707	Glycerol-3-phosphate dehydrogenase [NAD(+)], cytoplasmic	Gpd1	-2,8818	n.s.	7	2	3
P06802	Ectonucleotide pyrophosphatase/ phosphodiesterase family member 1	Enpp1	-2,8941	**	7	10	15
Q5FWI3	Transmembrane protein 2	Tmem2	-2,9132	n.s.	7	3	5
Q7TNG8	Probable D-lactate dehydrogenase, mitochondrial	Ldhd	-2,9557	n.s.	6	2	2
Q63961	Endoglin	Eng	-2,9869	n.s.	7	4	7
P57716	Nicastrin	Ncstn	-3,0434	n.s.	7	2	2
Q8BWN8	Acyl-coenzyme A thioesterase 4	Acot4	-3,0551	n.s.	7	2	3
P13597	Intercellular adhesion molecule 1	Icam1	-3,0594	n.s.	7	5	5
P19096	Fatty acid synthase	Fasn	-3,1000	*	7	2	3
P00920	Carbonic anhydrase 2	Ca2	-3,1209	n.s.	7	5	7
Q8CIB5	Fermitin family homolog 2	Fermt2	-3,1221	n.s.	7	2	3
B2RXS4	Plexin-B2	Plxnb2	-3,1402	**	7	8	13
P27773	Protein disulfide-isomerase A3	Pdia3	-3,1415	n.s.	7	6	7
Q9DBX3	Sushi domain-containing protein 2	Susd2	-3,1736	n.s.	7	10	12
Q8BWT1	3-ketoacyl-CoA thiolase, mitochondrial	Acaa2	-3,1742	n.s.	7	3	3
P46412	Glutathione peroxidase 3	Gpx3	-3,1820	*	7	5	6
Q8BJD1	Inter-alpha-trypsin inhibitor heavy chain H5	Itih5	-3,1924	*	7	11	14
Q91YQ5	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	Rpn1	-3,2090	n.s.	7	2	4
P10852	4F2 cell-surface antigen heavy chain	Slc3a2	-3,2173	**	7	8	13
Q62179	Semaphorin-4B	Sema4b	-3,2353	n.s.	7	2	3
P16546	Spectrin alpha chain, non-erythrocytic 1	Sptan1	-3,2553	*	7	2	7
O09172	Glutamate--cysteine ligase regulatory subunit	Gclm	-3,2643	n.s.	7	2	2
Q61598	Rab GDP dissociation inhibitor beta	Gdi2	-3,2661	n.s.	7	3	5
O08749	Dihydrolipoyl dehydrogenase, mitochondrial	Dld	-3,2878	n.s.	7	3	4
P18572	Basigin	Bsg	-3,2932	n.s.	7	4	7
O54749	Cytochrome P450 2J5	Cyp2j5	-3,3237	*	7	3	6
Q3UZZ6	Sulfotransferase 1 family member D1	Sult1d1	-3,3410	*	7	2	7
Q9Z2A9	Gamma-glutamyltransferase 5	Ggt5	-3,3503	n.s.	7	2	2
P21995	Embigin	Emb	-3,3509	n.s.	7	3	6
Q8BH95	Enoyl-CoA hydratase, mitochondrial	Echs1	-3,3675	*	7	2	2
Q9QZR9	Collagen alpha-4(IV) chain	Col4a4	-3,4273	n.s.	5	2	2
Q4VBE4	Pikachurin	Egflam	-3,4536	n.s.	7	2	3
Q02858	Angiopoietin-1 receptor	Tek	-3,4664	n.s.	4	2	2
Q8BH00	Aldehyde dehydrogenase family 8 member A1	Aldh8a1	-3,4747	*	7	3	4
A2ASQ1	Agrin	Agrn	-3,4813	***	7	14	22
Q8BFW7	Lipoma-preferred partner homolog	Lpp	-3,4866	n.s.	7	2	2
Q8BJ64	Choline dehydrogenase, mitochondrial	Chdh	-3,4910	n.s.	6	2	3
A2ARA8	Integrin alpha-8	Itga8	-3,4979	n.s.	7	3	5
Q70UZ7	von Willebrand factor A domain-containing protein 2	Vwa2	-3,5018	n.s.	5	2	5
Q9QZS7	Nephrin	Nphs1	-3,5607	n.s.	7	2	5
Q9DC69	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 9, mitochondrial	Ndufa9	-3,5763	n.s.	5	2	2
Q9WUR0	Protein Tinag	Tinag	-3,5830	**	7	16	17
P40142	Transketolase	Tkt	-3,6040	*	7	3	6
O09126	Semaphorin-4D	Sema4d	-3,6081	n.s.	7	3	7
Q9CQ65	S-methyl-5'-thioadenosine phosphorylase	Mtap	-3,6126	n.s.	5	2	2
Q91WR5	Aldo-keto reductase family 1 member C21	Akr1c21	-3,6251	*	7	6	10
Q9QXW9	Large neutral amino acids transporter small subunit 2	Slc7a8	-3,6382	n.s.	6	2	2
Q8BGV3	Tumor-associated calcium signal transducer 2	Tacstd2	-3,6688	n.s.	7	2	4
Q91WN4	Kynurenine 3-monooxygenase	Kmo	-3,6802	n.s.	7	3	6
Q9R0B6	Laminin subunit gamma-3	Lamc3	-3,6850	**	7	2	7
Q99LC3	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 10, mitochondrial	Ndufa10	-3,7127	*	7	3	5
Q9D964	Glycine amidinotransferase, mitochondrial	Gatm	-3,7164	n.s.	7	2	3

Q9CR68	Cytochrome b-c1 complex subunit Rieske, mitochondrial	Uqcrfs1	-3,8060	n.s.	7	2	2
Q91X17	Uromodulin	Umod	-3,8115	n.s.	7	5	11
Q9CQA3	Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial	Sdhb	-3,8195	n.s.	7	2	2
Q64487	Receptor-type tyrosine-protein phosphatase delta	Ptprd	-3,8654	n.s.	7	2	4
Q8K2B3	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial	Sdha	-3,8669	n.s.	7	3	6
Q62261	Spectrin beta chain, non-erythrocytic 1	Sptbn1	-3,8689	*	7	3	5
P05202	Aspartate aminotransferase, mitochondrial	Got2	-3,8938	*	7	3	5
Q99LB7	Sarcosine dehydrogenase, mitochondrial	Sardh	-3,8965	n.s.	7	2	5
P41216	Long-chain-fatty-acid--CoA ligase 1	Acsl1	-3,9271	*	7	2	8
Q9Z2V4	Phosphoenolpyruvate carboxykinase, cytosolic [GTP]	Pck1	-3,9330	**	7	2	5
P97449	Aminopeptidase N	Anpep	-3,9475	***	7	11	19
P70296	Phosphatidylethanolamine-binding protein 1	Pebp1	-3,9716	n.s.	7	2	2
Q9D8B4	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 11	Ndufa11	-3,9908	n.s.	7	2	2
Q61503	5'-nucleotidase	Nt5e	-4,0154	n.s.	7	5	7
Q99KI0	Aconitate hydratase, mitochondrial	Aco2	-4,0436	n.s.	7	3	4
P31428	Dipeptidase 1	Dpep1	-4,0459	**	7	9	13
P11352	Glutathione peroxidase 1	Gpx1	-4,0503	n.s.	7	2	3
Q99MZ7	Peroxisomal trans-2-enoyl-CoA reductase	Pecr	-4,0640	*	6	2	4
Q91WU5	Arsenite methyltransferase	As3mt	-4,0689	n.s.	5	2	2
O55111	Desmoglein-2	Dsg2	-4,0936	**	7	4	7
Q61739	Integrin alpha-6	Itga6	-4,0948	**	7	10	20
Q5U462	CUB domain-containing protein 1	Cdcp1	-4,1117	n.s.	7	3	4
Q9Z0T9	Integrin beta-6	Itgb6	-4,1276	n.s.	7	2	4
Q9CY27	Very-long-chain enoyl-CoA reductase	Tecr	-4,1490	n.s.	6	2	2
P14094	Sodium/potassium-transporting ATPase subunit beta-1	Atp1b1	-4,1969	*	7	7	12
Q8BH59	Calcium-binding mitochondrial carrier protein Aralar1	Slc25a12	-4,1978	n.s.	7	2	2
Q921G7	Electron transfer flavoprotein-ubiquinone oxidoreductase, mitochondrial	Etfdh	-4,2031	n.s.	5	2	2
Q61133	Glutathione S-transferase theta-2	Gstt2	-4,2060	**	7	4	9
Q91Y97	Fructose-bisphosphate aldolase B	Aldob	-4,2108	n.s.	7	4	5
P52760	Ribonuclease UK114	Hrsp12	-4,2134	n.s.	7	3	4
Q60805	Tyrosine-protein kinase Mer	Mertk	-4,2552	*	7	2	5
Q8K0H1	Multidrug and toxin extrusion protein 1	Slc47a1	-4,2618	n.s.	7	3	3
P09242	Alkaline phosphatase, tissue-nonspecific isozyme	Alpl	-4,2625	**	7	3	7
Q8QZT1	Acetyl-CoA acetyltransferase, mitochondrial	Acat1	-4,2707	*	7	2	8
P14152	Malate dehydrogenase, cytoplasmic	Mdh1	-4,2912	n.s.	6	3	3
O55060	Thiopurine S-methyltransferase	Tpmt	-4,3005	n.s.	6	2	2
Q9QXD1	Peroxisomal acyl-coenzyme A oxidase 2	Acox2	-4,3186	n.s.	7	4	4
O35409	Glutamate carboxypeptidase 2	Folh1	-4,3231	***	7	9	16
Q8K0L3	Acyl-coenzyme A synthetase ACSM2, mitochondrial	Acsm2	-4,3364	**	7	7	10
P58252	Elongation factor 2	Eef2	-4,3592	n.s.	7	2	3
Q3ULD5	Methylcrotonoyl-CoA carboxylase beta chain, mitochondrial	Mccc2	-4,4013	n.s.	4	2	3
P55096	ATP-binding cassette sub-family D member 3	Abcd3	-4,4081	n.s.	7	3	4
P15208	Insulin receptor	Insr	-4,4112	n.s.	7	2	2
P03911	NADH-ubiquinone oxidoreductase chain 4	Mtnd4	-4,4209	n.s.	7	2	4
P24270	Catalase	Cat	-4,4582	**	7	3	8
P41593	Parathyroid hormone/parathyroid hormone-related peptide receptor	Pth1r	-4,4723	*	7	2	2
Q60597	2-oxoglutarate dehydrogenase, mitochondrial	Ogdh	-4,5013	n.s.	7	2	4
Q80V42	Carboxypeptidase M	Cpm	-4,5051	*	7	6	14
E9Q467	Protein Abcc4	Abcc4	-4,5171	*	7	2	5
Q8VCT4	Carboxylesterase 1D	Ces1d	-4,5181	*	7	2	5
P35505	Fumarylacetoacetase	Fah	-4,5250	*	7	2	6

Q80X90	Filamin-B	Flnb	-4,5794	n.s.	7	2	4
P35822	Receptor-type tyrosine-protein phosphatase kappa	Ptprk	-4,5810	n.s.	7	3	3
Q923I7	Sodium/glucose cotransporter 2	Slc5a2	-4,5831	*	7	2	5
P56677	Suppressor of tumorigenicity 14 protein homolog	St14	-4,6062	*	7	5	8
Q9JLB4	Cubilin	Cubn	-4,6254	***	7	9	29
Q80XN0	D-beta-hydroxybutyrate dehydrogenase, mitochondrial	Bdh1	-4,6595	n.s.	7	2	2
P34914	Bifunctional epoxide hydrolase 2	Ephx2	-4,6631	n.s.	7	2	5
Q91VA0	Acyl-coenzyme A synthetase ACSM1, mitochondrial	Acsm1	-4,6780	n.s.	7	2	4
P00405	Cytochrome c oxidase subunit 2	Mtco2	-4,7021	n.s.	7	2	2
P00329	Alcohol dehydrogenase 1	Adh1	-4,7145	n.s.	7	2	2
Q9DBG6	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 2	Rpn2	-4,7255	*	7	2	4
P16125	L-lactate dehydrogenase B chain	Ldhb	-4,7262	n.s.	6	2	2
O88844	Isocitrate dehydrogenase [NADP] cytoplasmic	Idh1	-4,7369	n.s.	7	3	6
Q6NVD0	FRAS1-related extracellular matrix protein 2	Frem2	-4,7422	*	7	2	7
Q9EPX2	Papilin	Papln	-4,7519	***	7	2	7
Q9CZ13	Cytochrome b-c1 complex subunit 1, mitochondrial	Uqcrc1	-4,7722	n.s.	7	3	3
Q9CPY7	Cytosol aminopeptidase	Lap3	-4,7857	n.s.	7	2	4
Q60866	Phosphotriesterase-related protein	Pter	-4,7942	*	7	4	7
Q8VHG0	Dimethylaniline monooxygenase [N-oxide-forming] 4	Fmo4	-4,8141	*	7	3	8
P50516	V-type proton ATPase catalytic subunit A	Atp6v1a	-4,8544	n.s.	7	2	4
Q99NB1	Acetyl-coenzyme A synthetase 2-like, mitochondrial	Acss1	-4,8837	n.s.	7	3	6
Q8K209	G-protein coupled receptor 56	Gpr56	-4,8884	n.s.	7	2	4
E9PXN7	UDP-glucuronosyltransferase 1-7C	Ugt1a10	-4,8932	n.s.	7	4	4
Q9JKF6	Poliovirus receptor-related protein 1	Pvrl1	-4,8933	n.s.	7	3	5
Q62468	Villin-1	Vil1	-4,9038	***	7	5	13
Q8VEM8	Phosphate carrier protein, mitochondrial	Slc25a3	-4,9045	n.s.	7	4	5
P16406	Glutamyl aminopeptidase	Enpep	-4,9316	***	7	8	18
O88338	Cadherin-16	Cdh16	-5,0076	***	7	17	31
P55264	Adenosine kinase	Adk	-5,0094	n.s.	7	2	3
Q7TMS5	ATP-binding cassette sub-family G member 2	Abcg2	-5,0171	*	7	2	3
O35488	Very long-chain acyl-CoA synthetase	Slc27a2	-5,0241	**	7	9	17
Q8CHT0	Delta-1-pyrroline-5-carboxylate dehydrogenase, mitochondrial	Aldh4a1	-5,0729	*	5	2	5
Q9WUB6	Chloride channel protein ClC-Kb	Clcnkb	-5,0956	*	7	2	2
Q8BMS1	Trifunctional enzyme subunit alpha, mitochondrial	Hadha	-5,1033	*	7	4	5
Q9D0K2	Succinyl-CoA:3-ketoacid coenzyme A transferase 1, mitochondrial	Oxct1	-5,1243	n.s.	6	3	5
Q9NYQ2	Hydroxyacid oxidase 2	Hao2	-5,1439	*	7	3	7
Q99L13	3-hydroxyisobutyrate dehydrogenase, mitochondrial	Hibadh	-5,1822	n.s.	7	3	3
Q99MZ6	Unconventional myosin-VIIb	Myo7b	-5,1998	**	7	2	6
P42925	Peroxisomal membrane protein 2	Pxmp2	-5,3158	*	7	2	3
Q9CR62	Mitochondrial 2-oxoglutarate/malate carrier protein	Slc25a11	-5,3418	*	7	2	3
Q99KP3	Lambda-crystallin homolog	Cryl1	-5,3625	**	7	3	8
P14246	Solute carrier family 2, facilitated glucose transporter member 2	Slc2a2	-5,4091	n.s.	7	3	4
Q64462	Cytochrome P450 4B1	Cyp4b1	-5,4231	**	7	4	8
Q9CZB0	Succinate dehydrogenase cytochrome b560 subunit, mitochondrial	Sdhc	-5,4744	*	7	2	4
O88343	Electrogenic sodium bicarbonate cotransporter 1	Slc4a4	-5,4794	**	7	3	6
Q8CGC7	Bifunctional glutamate/proline--tRNA ligase	Eprs	-5,4949	*	7	2	3
P52825	Carnitine O-palmitoyltransferase 2, mitochondrial	Cpt2	-5,5243	*	6	3	4
Q8C3X8	Lipase maturation factor 2	Lmf2	-5,5401	*	7	2	3
Q8C0Z1	Protein ITFG3	Itfg3	-5,5414	*	7	3	5
Q8R0S9	BC026439 protein	Slc22a22	-5,5571	n.s.	6	2	4
Q9DBM2	Peroxisomal bifunctional enzyme	Ehhadh	-5,5695	***	7	10	28
O88909	Solute carrier family 22 member 8	Slc22a8	-5,6125	*	7	2	8

P55302	Alpha-2-macroglobulin receptor-associated protein	Lrpap1	-5,6407	*	7	2	6
P48962	ADP/ATP translocase 1	Slc25a4	-5,6563	**	7	4	6
P16331	Phenylalanine-4-hydroxylase	Pah	-5,6641	*	7	2	5
A2ARV4	Low-density lipoprotein receptor-related protein 2	Lrp2	-5,6645	***	7	22	100
Q9QZD8	Mitochondrial dicarboxylate carrier	Slc25a10	-5,6791	**	7	3	4
Q9JIL4	Na(+)/H(+) exchange regulatory cofactor NHE-RF3	Pdzk1	-5,6843	**	7	5	13
Q8VDN2	Sodium/potassium-transporting ATPase subunit alpha-1	Atp1a1	-5,6938	***	7	3	11
P28271	Cytoplasmic aconitate hydratase	Aco1	-5,7037	***	7	3	8
Q99JW2	Aminoacylase-1	Acy1	-5,7167	***	7	4	9
Q9CZ69	CKLF-like MARVEL transmembrane domain-containing protein 6	Cmtm6	-5,7271	n.s.	7	2	2
P09803	Cadherin-1	Cdh1	-5,7877	***	7	5	8
Q9QXX4	Calcium-binding mitochondrial carrier protein Aralar2	Slc25a13	-5,8012	***	7	2	11
Q9QXD6	Fructose-1,6-bisphosphatase 1	Fbp1	-5,8026	**	7	5	12
Q8R0Y6	Cytosolic 10-formyltetrahydrofolate dehydrogenase	Aldh1l1	-5,8264	***	7	7	14
P18894	D-amino-acid oxidase	Dao	-5,8410	*	7	2	5
Q9R0P3	S-formylglutathione hydrolase	Esd	-5,8473	*	7	4	5
Q9D826	Peroxisomal sarcosine oxidase	Pipox	-5,8508	*	7	2	5
Q61425	Hydroxyacyl-coenzyme A dehydrogenase, mitochondrial	Hadh	-5,9328	n.s.	7	2	2
Q8VC30	Bifunctional ATP-dependent dihydroxyacetone kinase/FAD-AMP lyase (cyclizing)	Dak	-5,9365	n.s.	6	2	2
Q64442	Sorbitol dehydrogenase	Sord	-6,0003	*	7	4	6
Q3UNZ8	Quinone oxidoreductase-like protein 2	null	-6,0069	*	7	2	3
Q9CPQ1	Cytochrome c oxidase subunit 6C	Cox6c	-6,0500	n.s.	7	2	3
A2A8L5	Receptor-type tyrosine-protein phosphatase F	Ptprf	-6,1368	*	6	2	6
P10518	Delta-aminolevulinic acid dehydratase	Alad	-6,1831	*	7	3	7
Q8VC69	Solute carrier family 22 member 6	Slc22a6	-6,1910	**	7	3	6
Q920R6	V-type proton ATPase 116 kDa subunit a isoform 4	Atp6v0a4	-6,2336	*	6	2	4
Q60825	Sodium-dependent phosphate transport protein 2A	Slc34a1	-6,2538	n.s.	7	2	5
Q9R097	Kunitz-type protease inhibitor 1	Spint1	-6,3586	*	7	2	4
Q8CFZ5	Solute carrier family 22 member 12	Slc22a12	-6,4420	**	7	2	8
Q99J94	Solute carrier organic anion transporter family member 1A6	Slco1a6	-6,4894	**	7	4	6
P17563	Selenium-binding protein 1	Selenbp1	-6,5533	n.s.	6	2	2
O54990	Prominin-1	Prom1	-6,5641	n.s.	7	2	4
P56135	ATP synthase subunit f, mitochondrial	Atp5j2	-6,7684	n.s.	7	2	2
Q9CPU4	Microsomal glutathione S-transferase 3	Mgst3	-6,9600	**	7	2	3
O88343-2	Isoform 2 of Electrogenic sodium bicarbonate cotransporter 1	Slc4a4	-7,0223	*	7	2	3
P97798	Neogenin	Neo1	-7,0936	n.s.	3	2	1

8.3 LIST OF PROTEINS: XENOGRAFT KIDNEY TUMOURS VERSUS HEALTHY RENAL CELLS, HUMAN

Tab. 17: Xenograft tumours versus healthy renal cells, human. Protein list containing all proteins quantified with at least two proteotypic peptides in at least one tumour group. Four samples from each the seven patient groups and four samples from each of the three human renal cell lines are taken for quantification. Positive values indicate an upregulation in tumour tissue compared to healthy renal cells. average p-value: < 0.001: ***; < 0.01: **; < 0.05: *; ≥ 0.05: not significant (n.s.).

swiss prot accession no	protein name	gene name	$\log_2$ (average ratio)	average p-value	# tumours quantified	max. peptides tumour	max. peptides human renal cells
Q96RQ3	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial	MCCC1	8,7272	***	7	32	3
P18428	Lipopolysaccharide-binding protein	LBP	8,4083	n.s.	2	2	0
P11498	Pyruvate carboxylase, mitochondrial	PC	8,0535	*	7	24	3
P05165	Propionyl-CoA carboxylase alpha chain, mitochondrial	PCCA	7,8245	***	7	23	4
Q13085	Acetyl-CoA carboxylase 1	ACACA	7,6503	*	7	16	2
Q8NE62	Choline dehydrogenase, mitochondrial	CHDH	7,4391	*	1	4	0
P00352	Retinal dehydrogenase 1	ALDH1A1	7,0839	**	7	12	1
Q99715	Collagen alpha-1(XII) chain	COL12A1	7,0280	*	5	11	0
P19224	UDP-glucuronosyltransferase 1-6	UGT1A6	6,9451	n.s.	4	3	0
P15121	Aldose reductase	AKR1B1	6,7669	*	7	6	1
P00918	Carbonic anhydrase 2	CA2	6,6626	n.s.	3	2	0
P08107	Heat shock 70 kDa protein 1A/1B	HSPA1A	6,6378	n.s.	3	2	1
P51659	Peroxisomal multifunctional enzyme type 2	HSD17B4	6,6242	*	4	4	1
Q9Y2S2	Lambda-crystallin homolog	CRYL1	6,6055	**	1	5	1
Q9BYF1	Angiotensin-converting enzyme 2	ACE2	6,5675	***	1	13	0
P25391	Laminin subunit alpha-1	LAMA1	6,4424	**	3	59	0
Q16658	Fascin	FSCN1	6,4347	*	5	3	1
P00387	NADH-cytochrome b5 reductase 3	CYB5R3	6,4257	n.s.	2	2	0
Q99798	Aconitate hydratase, mitochondrial	ACO2	6,4113	n.s.	2	2	1
Q13011	Delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial	ECH1	6,4103	*	6	6	0
P14868	Aspartate--tRNA ligase, cytoplasmic	DARS	6,3970	n.s.	6	2	0
P42765	3-ketoacyl-CoA thiolase, mitochondrial	ACAA2	6,3918	***	4	8	1
Q96DG6	Carboxymethylenebutenolidase homolog	CMBL	6,3768	n.s.	5	5	0
Q6IB77	Glycine N-acyltransferase	GLYAT	6,3398	n.s.	1	2	0
O00763	Acetyl-CoA carboxylase 2	ACACB	6,2760	*	3	9	0
P00367	Glutamate dehydrogenase 1, mitochondrial	GLUD1	6,2628	*	5	2	0
Q14697	Neutral alpha-glucosidase AB	GANAB	6,1785	*	5	6	2
P13010	X-ray repair cross-complementing protein 5	XRCC5	6,1781	*	3	2	1
P07942	Laminin subunit beta-1	LAMB1	6,1717	*	5	26	2
P07225	Vitamin K-dependent protein S	PROS1	6,1705	n.s.	2	3	1
Q15043	Zinc transporter ZIP14	SLC39A14	6,1550	***	3	4	0
Q99829	Copine-1	CPNE1	6,0841	*	3	2	0
P11047	Laminin subunit gamma-1	LAMC1	6,0699	*	3	24	1
Q93088	Betaine--homocysteine S-methyltransferase 1	BHMT	6,0646	*	1	3	0
P22314	Ubiquitin-like modifier-activating enzyme 1	UBA1	6,0608	n.s.	4	5	1
P24298	Alanine aminotransferase 1	GPT	6,0481	*	1	3	0
P11413	Glucose-6-phosphate 1-dehydrogenase	G6PD	6,0424	*	6	4	2
P01024	Complement C3	C3	6,0037	n.s.	6	14	1
Q02218	2-oxoglutarate dehydrogenase, mitochondrial	OGDH	5,9988	**	4	5	1
Q86UX2	Inter-alpha-trypsin inhibitor heavy chain H5	ITIH5	5,9893	**	1	5	0

Q03154	Aminoacylase-1	ACY1	5,9733	n.s.	1	2	0
P14210	Hepatocyte growth factor	HGF	5,9085	n.s.	1	2	0
O60701	UDP-glucose 6-dehydrogenase	UGDH	5,8920	n.s.	5	3	0
Q99497	Protein DJ-1	PARK7	5,8303	n.s.	1	2	1
Q9BSE5	Agmatinase, mitochondrial	AGMAT	5,8270	*	1	2	0
Q02878	60S ribosomal protein L6	RPL6	5,8034	n.s.	7	3	2
Q8IVN8	Somatomedin-B and thrombospondin type-1 domain-containing protein	SBSPON	5,7763	n.s.	2	2	0
P34896	Serine hydroxymethyltransferase, cytosolic	SHMT1	5,7461	n.s.	1	2	0
P02452	Collagen alpha-1(I) chain	COL1A1	5,7359	***	5	8	2
P14543	Nidogen-1	NID1	5,7251	***	3	21	2
Q13867	Bleomycin hydrolase	BLMH	5,6831	n.s.	5	3	1
Q6PIU2	Neutral cholesterol ester hydrolase 1	NCEH1	5,6174	*	3	3	0
P17655	Calpain-2 catalytic subunit	CAPN2	5,6167	n.s.	6	3	1
Q01813	6-phosphofructokinase type C	PFKP	5,6017	*	7	7	0
P11310	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial	ACADM	5,5968	*	1	2	0
Q96IU4	Alpha/beta hydrolase domain-containing protein 14B	ABHD14B	5,5886	*	1	2	1
P49748	Very long-chain specific acyl-CoA dehydrogenase, mitochondrial	ACADVL	5,5839	n.s.	1	2	0
P05091	Aldehyde dehydrogenase, mitochondrial	ALDH2	5,5684	***	3	4	0
P27338	Amine oxidase [flavin-containing] B	MAOB	5,5641	***	2	6	0
Q16363	Laminin subunit alpha-4	LAMA4	5,5299	***	1	8	0
O60494	Cubilin	CUBN	5,4841	***	1	26	1
Q08257	Quinone oxidoreductase	CRYZ	5,4806	***	4	5	1
Q16401	26S proteasome non-ATPase regulatory subunit 5	PSMD5	5,4593	*	4	3	1
P09467	Fructose-1,6-bisphosphatase 1	FBP1	5,4448	**	1	3	0
O14638	Ectonucleotide pyrophosphatase/phosphodiesterase family member 3	ENPP3	5,4437	**	3	6	0
P30039	Phenazine biosynthesis-like domain-containing protein	PBLD	5,4434	**	1	5	1
Q9BUT1	3-hydroxybutyrate dehydrogenase type 2	BDH2	5,4149	**	1	2	0
Q15582	Transforming growth factor-beta-induced protein ig-h3	TGFBI	5,3810	n.s.	7	12	4
Q9UJS0	Calcium-binding mitochondrial carrier protein Aralar2	SLC25A13	5,3667	n.s.	1	2	0
P31939	Bifunctional purine biosynthesis protein PURH	ATIC	5,3495	n.s.	4	3	1
P23526	Adenosylhomocysteinase	AHCY	5,2830	n.s.	6	2	2
P11216	Glycogen phosphorylase, brain form	PYGB	5,2715	n.s.	5	3	0
P04843	Dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit 1	RPN1	5,2455	*	6	5	1
P55268	Laminin subunit beta-2	LAMB2	5,2288	***	1	24	2
P50395	Rab GDP dissociation inhibitor beta	GDI2	5,1943	n.s.	5	2	0
Q96C23	Aldose 1-epimerase	GALM	5,1895	n.s.	1	2	0
P35573	Glycogen debranching enzyme	AGL	5,1650	**	1	4	0
Q02252	Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial	ALDH6A1	5,1599	n.s.	1	2	0
O00154	Cytosolic acyl coenzyme A thioester hydrolase	ACOT7	5,1385	n.s.	2	2	0
P07384	Calpain-1 catalytic subunit	CAPN1	5,1382	n.s.	5	2	1
P13489	Ribonuclease inhibitor	RNH1	5,1330	n.s.	6	4	3
Q8IUX7	Adipocyte enhancer-binding protein 1	AEBP1	5,1232	n.s.	3	2	0
P55084	Trifunctional enzyme subunit beta, mitochondrial	HADHB	5,0679	*	1	3	0
P05166	Propionyl-CoA carboxylase beta chain, mitochondrial	PCCB	5,0667	n.s.	1	2	0
Q99714	3-hydroxyacyl-CoA dehydrogenase type-2	HSD17B10	5,0561	*	4	3	1
O43488	Aflatoxin B1 aldehyde reductase member 2	AKR7A2	5,0470	***	5	3	2
Q9H0W9	Ester hydrolase C11orf54	C11orf54	5,0384	***	2	7	0
Q00796	Sorbitol dehydrogenase	SORD	5,0114	n.s.	2	2	0
P26641	Elongation factor 1-gamma	EEF1G	5,0056	n.s.	6	3	0
P34897	Serine hydroxymethyltransferase, mitochondrial	SHMT2	4,9654	n.s.	4	3	2
Q7KZF4	Staphylococcal nuclease domain-containing protein 1	SND1	4,9405	*	5	3	0

P17858	6-phosphofructokinase, liver type	PFKL	4,9116	**	2	4	1
Q9Y6R1	Electrogenic sodium bicarbonate cotransporter 1	SLC4A4	4,9115	n.s.	3	3	1
P00736	Complement C1r subcomponent	C1R	4,8833	n.s.	2	2	0
Q9HCC0	Methylcrotonoyl-CoA carboxylase beta chain, mitochondrial	MCCC2	4,8566	n.s.	1	2	0
P07814	Bifunctional glutamate/proline--tRNA ligase	EPRS	4,8474	n.s.	3	3	0
P23786	Carnitine O-palmitoyltransferase 2, mitochondrial	CPT2	4,8390	*	1	2	0
Q14117	Dihydropyrimidinase	DPYS	4,8010	**	1	4	0
O15230	Laminin subunit alpha-5	LAMA5	4,7862	*	4	65	4
O00469	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2	PLOD2	4,7798	n.s.	3	2	0
P49327	Fatty acid synthase	FASN	4,7772	n.s.	5	3	1
Q6RW13	Type-1 angiotensin II receptor-associated protein	AGTRAP	4,7707	n.s.	2	2	0
P42224	Signal transducer and activator of transcription 1-alpha/beta	STAT1	4,7488	*	5	5	1
Q9NUB1	Acetyl-coenzyme A synthetase 2-like, mitochondrial	ACSS1	4,7330	n.s.	1	2	0
Q8N135	Leucine-rich repeat LGI family member 4	LGI4	4,7024	**	2	3	0
P12821	Angiotensin-converting enzyme	ACE	4,6337	*	1	3	0
P31040	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial	SDHA	4,6275	n.s.	2	3	0
O43272	Proline dehydrogenase 1, mitochondrial	PRODH	4,6205	n.s.	1	2	0
Q14764	Major vault protein	MVP	4,6105	n.s.	5	3	1
P22234	Multifunctional protein ADE2	PAICS	4,5913	*	5	3	1
P08603	Complement factor H	CFH	4,5711	**	1	3	1
Q92597	Protein NDRG1	NDRG1	4,5073	*	5	4	2
Q9NZ08	Endoplasmic reticulum aminopeptidase 1	ERAP1	4,4750	n.s.	2	2	1
P40939	Trifunctional enzyme subunit alpha, mitochondrial	HADHA	4,4456	**	6	10	1
O43776	Asparagine--tRNA ligase, cytoplasmic	NARS	4,4375	*	4	3	0
Q13228	Selenium-binding protein 1	SELENBP1	4,4260	n.s.	1	2	0
P53004	Biliverdin reductase A	BLVRA	4,4255	**	3	2	0
P12956	X-ray repair cross-complementing protein 6	XRCC6	4,4165	n.s.	6	2	0
B0I1T2	Unconventional myosin-Ig	MYO1G	4,4100	*	2	3	0
Q08426	Peroxisomal bifunctional enzyme	EHHADH	4,3567	**	1	4	0
Q96KP4	Cytosolic non-specific dipeptidase	CNDP2	4,3558	**	4	8	1
P53041	Serine/threonine-protein phosphatase 5	PPP5C	4,3206	*	2	2	0
P98164	Low-density lipoprotein receptor-related protein 2	LRP2	4,2859	***	1	16	1
P26640	Valine--tRNA ligase	VARS	4,2770	n.s.	3	3	0
P06744	Glucose-6-phosphate isomerase	GPI	4,2436	n.s.	7	6	2
Q658P3	Metalloreductase STEAP3	STEAP3	4,2037	*	3	2	1
Q9HBJ8	Collectrin	TMEM27	4,1819	*	1	2	0
Q9UHE5	Probable N-acetyltransferase 8	NAT8	4,0509	n.s.	1	2	0
Q6XQN6	Nicotinate phosphoribosyltransferase	NAPRT1	4,0101	n.s.	2	3	1
P38606	V-type proton ATPase catalytic subunit A	ATP6V1A	3,9954	n.s.	2	2	1
Q9H4A4	Aminopeptidase B	RNPEP	3,9905	**	2	2	0
Q9HAW8	UDP-glucuronosyltransferase 1-10	UGT1A10	3,9699	**	2	2	0
Q96BI1	Solute carrier family 22 member 18	SLC22A18	3,9479	n.s.	2	2	0
Q9NSE4	Isoleucine--tRNA ligase, mitochondrial	IARS2	3,8900	n.s.	1	2	1
Q86XX4-2	Isoform 2 of Extracellular matrix protein FRAS1	FRAS1	3,8426	n.s.	1	2	0
P08294	Extracellular superoxide dismutase [Cu-Zn]	SOD3	3,8411	*	1	3	0
Q86XX4	Extracellular matrix protein FRAS1	FRAS1	3,8322	***	1	24	1
Q12931	Heat shock protein 75 kDa, mitochondrial	TRAP1	3,8169	**	1	4	0
O75309	Cadherin-16	CDH16	3,8032	n.s.	1	3	0
P08133	Annexin A6	ANXA6	3,7692	n.s.	3	3	0
P30038	Delta-1-pyrroline-5-carboxylate dehydrogenase, mitochondrial	ALDH4A1	3,7516	n.s.	1	2	2
Q93099	Homogentisate 1,2-dioxygenase	HGD	3,7320	*	2	3	0
P42704	Leucine-rich PPR motif-containing protein, mitochondrial	LRPPRC	3,7047	***	2	7	0
P30043	Flavin reductase (NADPH)	BLVRB	3,6907	n.s.	2	2	0

P48643	T-complex protein 1 subunit epsilon	CCT5	3,6396	n.s.	3	2	0
P47897	Glutamine--tRNA ligase	QARS	3,5828	n.s.	1	2	0
P00533	Epidermal growth factor receptor	EGFR	3,5254	n.s.	7	5	2
P08237	6-phosphofructokinase, muscle type	PFKM	3,5085	n.s.	2	2	0
Q14112	Nidogen-2	NID2	3,4980	***	2	10	1
P14618	Pyruvate kinase PKM	PKM	3,4862	n.s.	7	3	2
P22695	Cytochrome b-c1 complex subunit 2, mitochondrial	UQCRC2	3,4150	n.s.	5	3	0
Q16790	Carbonic anhydrase 9	CA9	3,4065	**	1	2	0
Q08AH3	Acyl-coenzyme A synthetase ACSM2A, mitochondrial	ACSM2A	3,3686	n.s.	1	2	1
P24752	Acetyl-CoA acetyltransferase, mitochondrial	ACAT1	3,3266	n.s.	4	11	2
P06737	Glycogen phosphorylase, liver form	PYGL	3,3257	**	4	2	0
Q6PI48	Aspartate--tRNA ligase, mitochondrial	DARS2	3,2402	n.s.	1	3	0
O15229	Kynurenine 3-monooxygenase	KMO	3,2322	n.s.	3	2	0
Q02790	Peptidyl-prolyl cis-trans isomerase FKBP4	FKBP4	3,1344	n.s.	2	2	0
P36871	Phosphoglucomutase-1	PGM1	3,0611	n.s.	5	3	1
O00748	Cocaine esterase	CES2	2,9998	n.s.	1	2	0
P06733	Alpha-enolase	ENO1	2,9815	n.s.	7	11	6
P19971	Thymidine phosphorylase	TYMP	2,7071	n.s.	4	2	0
P30613	Pyruvate kinase PKLR	PKLR	2,6967	**	1	2	0
P12268	Inosine-5'-monophosphate dehydrogenase 2	IMPDH2	2,6696	n.s.	1	2	0
P13861	cAMP-dependent protein kinase type II-alpha regulatory subunit	PRKAR2A	2,5738	n.s.	1	2	0
Q15274	Nicotinate-nucleotide pyrophosphorylase [carboxylating]	QPRT	2,4868	n.s.	2	3	1
P00450	Ceruloplasmin	CP	2,4577	n.s.	6	11	5
P08758	Annexin A5	ANXA5	2,3652	n.s.	7	3	2
P00338	L-lactate dehydrogenase A chain	LDHA	2,1724	n.s.	7	13	5
P31937	3-hydroxyisobutyrate dehydrogenase, mitochondrial	HIBADH	2,0239	n.s.	5	3	1
P07741	Adenine phosphoribosyltransferase	APRT	1,9828	n.s.	2	2	1
Q4KMQ2	Anoctamin-6	ANO6	1,9619	n.s.	1	2	1
P09525	Annexin A4	ANXA4	1,8579	n.s.	7	8	4
O43175	D-3-phosphoglycerate dehydrogenase	PHGDH	1,8017	n.s.	3	2	1
Q16787	Laminin subunit alpha-3	LAMA3	1,7069	n.s.	4	2	4
P14550	Alcohol dehydrogenase [NADP(+)]	AKR1A1	1,6332	n.s.	2	2	2
P24821	Tenascin	TNC	1,4829	n.s.	6	9	9
P53396	ATP-citrate synthase	ACLY	1,3818	n.s.	2	3	1
P78527	DNA-dependent protein kinase catalytic subunit	PRKDC	1,3396	n.s.	6	10	2
P30041	Peroxiredoxin-6	PRDX6	1,0790	n.s.	5	3	2
Q9UHN6	Transmembrane protein 2	TMEM2	0,9230	*	6	14	12
P09211	Glutathione S-transferase P	GSTP1	0,8854	n.s.	7	5	4
P02786	Transferrin receptor protein 1	TFRC	0,4953	n.s.	7	5	5
P29401	Transketolase	TKT	0,4783	n.s.	5	5	3
P00491	Purine nucleoside phosphorylase	PNP	0,4521	n.s.	5	2	2
P32119	Peroxiredoxin-2	PRDX2	0,4501	n.s.	5	2	3
P45880	Voltage-dependent anion-selective channel protein 2	VDAC2	0,4216	n.s.	7	3	1
P04406	Glyceraldehyde-3-phosphate dehydrogenase	GAPDH	0,4171	n.s.	7	10	8
O14773	Tripeptidyl-peptidase 1	TPP1	0,3210	n.s.	7	4	3
P21980	Protein-glutamine gamma-glutamyltransferase 2	TGM2	0,2934	n.s.	7	10	7
Q6UWP8	Suprabasin	SBSN	0,2610	n.s.	1	2	1
Q08380	Galectin-3-binding protein	LGALS3BP	0,0922	n.s.	6	4	4
Q6YHK3	CD109 antigen	CD109	0,0823	n.s.	6	13	6
O75054	Immunoglobulin superfamily member 3	IGSF3	-0,0909	n.s.	2	2	2
Q00325	Phosphate carrier protein, mitochondrial	SLC25A3	-0,1970	n.s.	6	3	3
P07737	Profilin-1	PFN1	-0,2358	n.s.	6	5	5
Q8WTV0	Scavenger receptor class B member 1	SCARB1	-0,2417	n.s.	3	2	1
P08779	Keratin, type I cytoskeletal 16	KRT16	-0,2554	n.s.	6	5	7
O43707	Alpha-actinin-4	ACTN4	-0,2718	n.s.	3	2	2
P02794	Ferritin heavy chain	FTH1	-0,3179	n.s.	6	2	1

P08581	Hepatocyte growth factor receptor	MET	-0,3321	n.s.	4	5	5
Q9BRA2	Thioredoxin domain-containing protein 17	TXNDC17	-0,3717	n.s.	5	3	1
P09622	Dihydrolipoyl dehydrogenase, mitochondrial	DLD	-0,4358	n.s.	4	2	3
O00468	Agrin	AGRN	-0,4669	n.s.	6	34	17
Q8WUM4	Programmed cell death 6-interacting protein	PDCD6IP	-0,5122	n.s.	3	2	1
P62937	Peptidyl-prolyl cis-trans isomerase A	PPIA	-0,5271	n.s.	7	3	4
Q06830	Peroxiredoxin-1	PRDX1	-0,5501	n.s.	6	2	1
P10809	60 kDa heat shock protein, mitochondrial	HSPD1	-0,6031	n.s.	7	3	3
P30086	Phosphatidylethanolamine-binding protein 1	PEBP1	-0,6124	n.s.	5	2	1
P00558	Phosphoglycerate kinase 1	PGK1	-0,6786	n.s.	6	3	3
Q9NPR2	Semaphorin-4B	SEMA4B	-0,6796	n.s.	4	2	1
P14625	Endoplasmin	HSP90B1	-0,6823	n.s.	5	2	2
P08473	Neprilysin	MME	-0,7727	n.s.	1	2	1
P30084	Enoyl-CoA hydratase, mitochondrial	ECHS1	-0,8010	n.s.	6	3	2
P35579	Myosin-9	MYH9	-0,8505	n.s.	7	4	1
O75874	Isocitrate dehydrogenase [NADP] cytoplasmic	IDH1	-0,8678	n.s.	6	2	2
Q9H3N1	Thioredoxin-related transmembrane protein 1	TMX1	-0,9423	n.s.	3	2	2
P02462	Collagen alpha-1(IV) chain	COL4A1	-0,9467	*	3	2	3
P04792	Heat shock protein beta-1	HSPB1	-1,0316	n.s.	7	4	5
Q53H82	Beta-lactamase-like protein 2	LACTB2	-1,0460	n.s.	3	2	1
P06703	Protein S100-A6	S100A6	-1,0542	n.s.	6	2	2
P30044	Peroxiredoxin-5, mitochondrial	PRDX5	-1,1027	n.s.	5	3	3
P04075	Fructose-bisphosphate aldolase A	ALDOA	-1,1050	n.s.	7	7	5
Q99685	Monoglyceride lipase	MGLL	-1,1245	n.s.	2	2	2
Q86YZ3	Hornerin	HRNR	-1,1855	n.s.	7	10	10
O43570	Carbonic anhydrase 12	CA12	-1,1869	n.s.	6	6	3
P35527	Keratin, type I cytoskeletal 9	KRT9	-1,2528	**	7	25	31
Q16698	2,4-dienoyl-CoA reductase, mitochondrial	DECR1	-1,3103	n.s.	4	5	2
P07996	Thrombospondin-1	THBS1	-1,3511	n.s.	7	14	13
P08670	Vimentin	VIM	-1,3603	n.s.	7	9	8
P30048	Thioredoxin-dependent peroxide reductase, mitochondrial	PRDX3	-1,3876	n.s.	3	2	2
O14880	Microsomal glutathione S-transferase 3	MGST3	-1,3968	n.s.	7	2	2
Q15181	Inorganic pyrophosphatase	PPA1	-1,4705	n.s.	6	2	2
O75369	Filamin-B	FLNB	-1,4998	n.s.	3	4	5
P04083	Annexin A1	ANXA1	-1,5192	n.s.	7	4	5
Q92945	Far upstream element-binding protein 2	KHSRP	-1,6046	n.s.	4	2	3
Q07075	Glutamyl aminopeptidase	ENPEP	-1,6417	n.s.	6	8	5
P22413	Ectonucleotide pyrophosphatase/phosphodiesterase family member 1	ENPP1	-1,6520	n.s.	6	4	4
Q9GZM7	Tubulointerstitial nephritis antigen-like	TINAGL1	-1,6625	n.s.	4	8	8
Q99536	Synaptic vesicle membrane protein VAT-1 homolog	VAT1	-1,6794	n.s.	1	2	1
Q14126	Desmoglein-2	DSG2	-1,7927	n.s.	2	3	3
Q92485	Acid sphingomyelinase-like phosphodiesterase 3b	SMPDL3B	-1,7965	*	5	2	6
P05783	Keratin, type I cytoskeletal 18	KRT18	-1,9442	n.s.	6	2	6
Q13308	Inactive tyrosine-protein kinase 7	PTK7	-1,9796	n.s.	6	2	4
P54886	Delta-1-pyrroline-5-carboxylate synthase	ALDH18A1	-2,0672	n.s.	3	2	1
P04264	Keratin, type II cytoskeletal 1	KRT1	-2,0803	**	7	24	26
O75915	PRA1 family protein 3	ARL6IP5	-2,0998	n.s.	6	2	2
P02533	Keratin, type I cytoskeletal 14	KRT14	-2,1075	n.s.	7	4	4
Q86UP2	Kinectin	KTN1	-2,1517	**	6	2	7
P13611	Versican core protein	VCAN	-2,1884	n.s.	5	6	7
P35908	Keratin, type II cytoskeletal 2 epidermal	KRT2	-2,2773	*	7	15	22
Q92820	Gamma-glutamyl hydrolase	GGH	-2,2831	n.s.	3	2	2
P60174	Triosephosphate isomerase	TPI1	-2,3253	n.s.	6	4	4
P13645	Keratin, type I cytoskeletal 10	KRT10	-2,3336	**	7	17	22
O14786	Neuropilin-1	NRP1	-2,3708	n.s.	3	5	3
P00505	Aspartate aminotransferase, mitochondrial	GOT2	-2,3916	n.s.	5	3	3

P21333	Filamin-A	FLNA	-2,4275	*	7	12	13
P04179	Superoxide dismutase [Mn], mitochondrial	SOD2	-2,4362	n.s.	6	4	4
Q8N1N4	Keratin, type II cytoskeletal 78	KRT78	-2,4534	n.s.	7	4	6
P02768	Serum albumin	ALB	-2,4574	***	7	6	26
P05023	Sodium/potassium-transporting ATPase subunit alpha-1	ATP1A1	-2,4903	n.s.	3	2	1
P25445	Tumor necrosis factor receptor superfamily member 6	FAS	-2,5555	***	6	5	13
P15924	Desmoplakin	DSP	-2,5885	n.s.	4	2	4
P27824	Calnexin	CANX	-2,6299	n.s.	7	4	5
P23470	Receptor-type tyrosine-protein phosphatase gamma	PTPRG	-2,6400	n.s.	1	2	2
P13647	Keratin, type II cytoskeletal 5	KRT5	-2,7230	n.s.	7	5	6
P08195	4F2 cell-surface antigen heavy chain	SLC3A2	-2,7338	*	7	10	13
P17301	Integrin alpha-2	ITGA2	-2,7385	*	6	6	12
P54289	Voltage-dependent calcium channel subunit alpha-2/delta-1	CACNA2D1	-2,9022	n.s.	3	2	3
P78417	Glutathione S-transferase omega-1	GSTO1	-2,9204	n.s.	3	2	3
P02787	Serotransferrin	TF	-2,9207	***	7	5	18
P27487	Dipeptidyl peptidase 4	DPP4	-2,9253	*	7	22	22
P54819	Adenylate kinase 2, mitochondrial	AK2	-2,9450	n.s.	5	2	4
Q9Y4K0	Lysyl oxidase homolog 2	LOXL2	-2,9483	n.s.	6	3	4
P16278	Beta-galactosidase	GLB1	-3,0240	n.s.	4	2	4
P27797	Calreticulin	CALR	-3,0281	n.s.	7	4	5
P10909	Clusterin	CLU	-3,1525	*	6	4	7
Q9Y653	G-protein coupled receptor 56	GPR56	-3,2293	*	5	2	4
P14314	Glucosidase 2 subunit beta	PRKCSH	-3,3625	**	6	2	7
Q92542	Nicastrin	NCSTN	-3,4213	*	5	2	4
Q15149	Plectin	PLEC	-3,4242	n.s.	3	3	5
Q96PD2	Discoidin, CUB and LCCL domain-containing protein 2	DCBLD2	-3,4754	*	4	6	11
P15328	Folate receptor alpha	FOLR1	-3,4922	*	3	2	6
P23229	Integrin alpha-6	ITGA6	-3,5227	n.s.	4	5	5
P98160	Basement membrane-specific heparan sulfate proteoglycan core protein	HSPG2	-3,6430	***	7	27	43
P32970	CD70 antigen	CD70	-3,6753	n.s.	4	4	5
P02751	Fibronectin	FN1	-3,7027	***	7	52	68
P37802	Transgelin-2	TAGLN2	-3,7627	n.s.	7	3	3
Q9P2B2	Prostaglandin F2 receptor negative regulator	PTGFRN	-3,7724	***	6	6	16
Q01082	Spectrin beta chain, non-erythrocytic 1	SPTBN1	-3,8857	**	2	3	5
P07237	Protein disulfide-isomerase	P4HB	-3,9035	n.s.	7	4	6
P52943	Cysteine-rich protein 2	CRIP2	-3,9393	*	2	2	2
Q09666	Neuroblast differentiation-associated protein AHNAK	AHNAK	-3,9473	***	7	11	62
Q9UIQ6	Leucyl-cystinyl aminopeptidase	LNPEP	-3,9611	n.s.	3	2	3
Q9UJW2	Tubulointerstitial nephritis antigen	TINAG	-3,9651	n.s.	3	3	4
P05787	Keratin, type II cytoskeletal 8	KRT8	-4,0489	n.s.	6	2	4
P08572	Collagen alpha-2(IV) chain	COL4A2	-4,0683	**	7	9	16
P29590	Protein PML	PML	-4,0739	***	6	2	10
P12830	Cadherin-1	CDH1	-4,1162	n.s.	2	2	2
P54709	Sodium/potassium-transporting ATPase subunit beta-3	ATP1B3	-4,1536	***	2	2	8
P30461	HLA class I histocompatibility antigen, B-13 alpha chain	HLA-B	-4,1602	n.s.	3	2	2
Q15084	Protein disulfide-isomerase A6	PDIA6	-4,1616	n.s.	5	2	3
Q10588	ADP-ribosyl cyclase 2	BST1	-4,1885	**	3	3	6
P40261	Nicotinamide N-methyltransferase	NNMT	-4,2443	n.s.	4	2	2
P53634	Dipeptidyl peptidase 1	CTSC	-4,2771	**	5	4	8
P40121	Macrophage-capping protein	CAPG	-4,3238	n.s.	2	2	2
P19256	Lymphocyte function-associated antigen 3	CD58	-4,3294	*	4	2	4
P48307	Tissue factor pathway inhibitor 2	TFPI2	-4,3491	n.s.	1	2	3
P07203	Glutathione peroxidase 1	GPX1	-4,4085	n.s.	1	2	2
P48960	CD97 antigen	CD97	-4,5512	n.s.	3	2	4
O43490	Prominin-1	PROM1	-4,6390	**	3	3	7

Q9Y624	Junctional adhesion molecule A	F11R	-4,6582	***	6	4	13
P08648	Integrin alpha-5	ITGA5	-4,6685	***	6	5	9
P39060	Collagen alpha-1(XVIII) chain	COL18A1	-4,6867	***	7	14	23
P48060	Glioma pathogenesis-related protein 1	GLIPR1	-4,7845	*	2	2	4
P19320	Vascular cell adhesion protein 1	VCAM1	-4,8110	***	7	9	25
P09382	Galectin-1	LGALS1	-4,8557	*	7	6	6
P30101	Protein disulfide-isomerase A3	PDIA3	-4,9073	**	7	6	12
Q92823	Neuronal cell adhesion molecule	NRCAM	-4,9480	**	5	2	10
O43157	Plexin-B1	PLXNB1	-4,9611	***	3	2	7
Q6EMK4	Vasorin	VASN	-4,9844	***	6	2	7
P21926	CD9 antigen	CD9	-5,0527	**	2	2	6
Q07065	Cytoskeleton-associated protein 4	CKAP4	-5,0640	**	4	4	9
Q16881	Thioredoxin reductase 1, cytoplasmic	TXNRD1	-5,1452	n.s.	1	2	1
P01892	HLA class I histocompatibility antigen, A-2 alpha chain	HLA-A	-5,1914	n.s.	3	2	4
Q9UNN8	Endothelial protein C receptor	PROCR	-5,2122	**	6	4	5
P35613	Basigin	BSG	-5,2781	***	7	7	13
P50895	Basal cell adhesion molecule	BCAM	-5,3761	***	6	9	22
P78310	Coxsackievirus and adenovirus receptor	CXADR	-5,3954	**	3	2	7
P08727	Keratin, type I cytoskeletal 19	KRT19	-5,4724	***	7	5	13
O15031	Plexin-B2	PLXNB2	-5,5154	***	6	11	24
P26006	Integrin alpha-3	ITGA3	-5,5161	***	7	17	30
Q02952	A-kinase anchor protein 12	AKAP12	-5,5604	***	6	2	17
P35052	Glypican-1	GPC1	-5,5663	***	4	3	11
Q8WUT4	Leucine-rich repeat neuronal protein 4	LRRN4	-5,5730	**	4	3	6
P04216	Thy-1 membrane glycoprotein	THY1	-5,5752	n.s.	2	3	4
Q16270	Insulin-like growth factor-binding protein 7	IGFBP7	-5,5835	***	4	4	9
P11717	Cation-independent mannose-6-phosphate receptor	IGF2R	-5,6572	***	5	4	29
P05556	Integrin beta-1	ITGB1	-5,6791	***	7	6	20
P05026	Sodium/potassium-transporting ATPase subunit beta-1	ATP1B1	-5,6862	***	5	3	9
P61769	Beta-2-microglobulin	B2M	-5,6919	**	6	3	6
P07858	Cathepsin B	CTSB	-5,7307	**	6	4	8
P05362	Intercellular adhesion molecule 1	ICAM1	-5,7408	***	7	19	28
Q04721	Neurogenic locus notch homolog protein 2	NOTCH2	-5,7444	***	5	2	9
P32004	Neural cell adhesion molecule L1	L1CAM	-5,7934	***	7	12	25
Q8IWA5	Choline transporter-like protein 2	SLC44A2	-5,8380	***	6	3	16
O43278	Kunitz-type protease inhibitor 1	SPINT1	-5,9146	***	4	3	17
Q9P2E9	Ribosome-binding protein 1	RRBP1	-6,0018	***	6	3	14
Q13813	Spectrin alpha chain, non-erythrocytic 1	SPTAN1	-6,1498	**	3	4	13
Q9H5V8	CUB domain-containing protein 1	CDCP1	-6,2208	***	6	9	23
P19022	Cadherin-2	CDH2	-6,2212	*	2	3	4
P43121	Cell surface glycoprotein MUC18	MCAM	-6,2389	***	7	5	20
Q96D42	Hepatitis A virus cellular receptor 1	HAVCR1	-6,2580	*	1	2	4
Q14118	Dystroglycan	DAG1	-6,3846	*	3	4	6
P68871	Hemoglobin subunit beta	HBB	-6,7431	n.s.	2	2	3
P16422	Epithelial cell adhesion molecule	EPCAM	-6,7545	*	3	2	6
P30447	HLA class I histocompatibility antigen, A-23 alpha chain	HLA-A	-6,9356	*	3	2	4
P04439	HLA class I histocompatibility antigen, A-3 alpha chain	HLA-A	-6,9671	**	6	6	7
P06756	Integrin alpha-V	ITGAV	-7,1281	***	7	9	27
P13746	HLA class I histocompatibility antigen, A-11 alpha chain	HLA-A	-7,4615	***	4	2	3
Q13641	Trophoblast glycoprotein	TPBG	-7,5047	*	3	2	6
P21589	5'-nucleotidase	NT5E	-7,5608	***	7	9	19
Q13740	CD166 antigen	ALCAM	-7,5798	***	6	5	15
P13987	CD59 glycoprotein	CD59	-7,6764	***	6	5	15
P14384	Carboxypeptidase M	CPM	-7,6993	***	6	5	12
Q56VL3	OCIA domain-containing protein 2	OCIAD2	-7,7226	*	2	2	4
P15144	Aminopeptidase N	ANPEP	-7,9386	***	7	17	44
P19440	Gamma-glutamyltranspeptidase 1	GGT1	-8,0234	n.s.	4	2	5

P16070	CD44 antigen	CD44	-8,2999	***	7	8	18
O75487	Glypican-4	GPC4	-8,5670	***	2	2	7

8.4 LIST OF PROTEINS: XENOGRAFT LUNG METASTASES VERSUS HEALTHY LUNG, MOUSE

Tab. 18: Xenograft metastases versus healthy lung, mouse. Xenograft metastases versus healthy lung, mouse. Protein list containing all proteins quantified with at least two proteotypic peptides in at least one metastasis group. Four samples from each of the four patient groups and four samples from healthy mouse lung are taken for quantification. Positive values indicate an upregulation in lung metastases compared to healthy lung tissue. average p-value: < 0.001: ***; < 0.01: **; < 0.05: *; ≥ 0.05: not significant (n.s.).

swiss prot accession no	protein name	gene name	$\log_2$ (average ratio)	average p-value	# metastases quantified	max. peptides metastases, mouse	max. peptides mouse lung
P05367	Serum amyloid A-2 protein	Saa2	8,2505	n.s.	2	2	0
Q8K182	Complement component C8 alpha chain	C8a	6,4417	n.s.	3	3	1
P07361	Alpha-1-acid glycoprotein 2	Orm2	6,3182	n.s.	3	2	1
P05366	Serum amyloid A-1 protein	Saa1	6,1345	n.s.	1	2	1
P97371	Proteasome activator complex subunit 1	Psme1	6,1315	n.s.	1	2	0
P58252	Elongation factor 2	Eef2	6,1219	n.s.	2	2	1
Q61646	Haptoglobin	Hp	5,7973	n.s.	4	6	5
O35744	Chitinase-like protein 3	Chil3	5,2480	n.s.	3	2	2
P27005	Protein S100-A8	S100a8	5,1217	n.s.	4	3	3
Q9R098	Hepatocyte growth factor activator	Hgfac	5,0313	n.s.	2	2	1
Q8R4V1	NFAT activation molecule 1	Nfam1	4,6654	n.s.	3	2	2
Q9D7J6	Deoxyribonuclease-1-like 1	Dnase1l1	4,6346	n.s.	2	2	0
P19324	Serpin H1	Serpinh1	4,5832	n.s.	2	2	1
P40124	Adenylyl cyclase-associated protein 1	Cap1	4,5357	n.s.	3	2	0
Q923B6	Metalloreductase STEAP4	Steap4	4,5233	n.s.	3	4	3
Q64727	Vinculin	Vcl	4,4859	n.s.	2	2	0
Q61093	Cytochrome b-245 heavy chain	Cybb	4,4184	n.s.	2	2	0
P23492	Purine nucleoside phosphorylase	Pnp	4,3290	n.s.	2	3	2
Q8R366	Immunoglobulin superfamily member 8	Igsf8	4,1532	n.s.	4	2	2
Q80YX1	Tenascin	Tnc	4,1455	n.s.	3	3	1
P09581	Macrophage colony-stimulating factor 1 receptor	Csf1r	4,1425	n.s.	3	3	2
Q60590	Alpha-1-acid glycoprotein 1	Orm1	4,1377	*	4	5	5
P31001	Desmin	Des	4,0932	n.s.	3	3	2
Q9Z100	Probable carboxypeptidase X1	Cpxm1	4,0787	*	2	3	2
Q00612	Glucose-6-phosphate 1-dehydrogenase X	G6pdx	4,0483	n.s.	4	3	2
P06802	Ectonucleotide pyrophosphatase/ phosphodiesterase family member 1	Enpp1	4,0238	n.s.	4	4	4
Q07235	Glia-derived nexin	Serpine2	3,9833	n.s.	4	3	2
Q91XL1	Leucine-rich HEV glycoprotein	Lrg1	3,9683	n.s.	4	4	3
Q63918	Serum deprivation-response protein	Sdpr	3,9442	n.s.	2	2	2
Q5U462	CUB domain-containing protein 1	Cdcp1	3,8184	n.s.	2	2	1
Q80T21	ADAMTS-like protein 4	Adamtsl4	3,7644	n.s.	1	3	0
P70232	Neural cell adhesion molecule L1-like protein	Chl1	3,7034	*	2	5	4
Q8R1F1	Niban-like protein 1	Fam129b	3,7007	n.s.	1	2	1
Q61704	Inter-alpha-trypsin inhibitor heavy chain H3	Itih3	3,6867	***	4	11	7
O35608	Angiopoietin-2	Angpt2	3,6301	n.s.	4	2	2
P31532	Serum amyloid A-4 protein	Saa4	3,5878	n.s.	1	2	1
Q00724	Retinol-binding protein 4	Rbp4	3,5735	n.s.	3	2	1
P06683	Complement component C9	C9	3,5683	***	4	7	7
P07901	Heat shock protein HSP 90-alpha	Hsp90aa1	3,5576	n.s.	2	2	2
Q00519	Xanthine dehydrogenase/oxidase	Xdh	3,5542	n.s.	2	3	1
P26039	Talin-1	Tln1	3,5361	n.s.	4	3	3

Q03734	Serine protease inhibitor A3M	Serpina3m	3,5121	n.s.	4	4	3
Q9QUR8	Semaphorin-7A	Sema7a	3,4431	n.s.	4	8	5
Q8BK62	Olfactomedin-like protein 3	Olfml3	3,4275	n.s.	4	3	3
O08692	Myeloid bactenecin (F1)	Ngp	3,4210	n.s.	4	2	2
P17182	Alpha-enolase	Eno1	3,3861	n.s.	4	3	2
Q91VB8	Alpha globin 1	Hba-a2	3,3452	n.s.	4	2	2
P99029	Peroxiredoxin-5, mitochondrial	Prdx5	3,2483	n.s.	3	3	2
Q8VCS0	N-acetylmuramoyl-L-alanine amidase	Pglyrp2	3,2102	*	1	2	2
P62962	Profilin-1	Pfn1	3,1623	n.s.	4	4	4
P27773	Protein disulfide-isomerase A3	Pdia3	3,0781	n.s.	4	4	3
P01942	Hemoglobin subunit alpha	Hba	3,0311	n.s.	3	3	4
Q3UZZ4	Olfactomedin-4	Olfm4	3,0161	n.s.	1	2	1
P97792	Coxsackievirus and adenovirus receptor homolog	Cxadr	2,9618	n.s.	2	2	1
P52480	Pyruvate kinase PKM	Pkm	2,8940	n.s.	4	2	2
Q61233	Plastin-2	Lcp1	2,8844	n.s.	4	4	4
Q8R4Y4	Stabilin-1	Stab1	2,8481	n.s.	4	6	4
O08742	Platelet glycoprotein V	Gp5	2,8360	n.s.	4	3	2
P11672	Neutrophil gelatinase-associated lipocalin	Lcn2	2,8290	n.s.	4	3	3
Q61425	Hydroxyacyl-coenzyme A dehydrogenase, mitochondrial	Hadh	2,8068	n.s.	4	2	2
Q8BWY2	C-type lectin domain family 1 member A	Clec1a	2,8021	n.s.	2	2	1
O88844	Isocitrate dehydrogenase [NADP] cytoplasmic	Idh1	2,7941	*	4	2	2
P11859	Angiotensinogen	Agt	2,7741	n.s.	4	4	4
Q61738	Integrin alpha-7	Itga7	2,7726	n.s.	2	2	1
Q91WP6	Serine protease inhibitor A3N	Serpina3n	2,7131	n.s.	4	5	4
Q61805	Lipopolysaccharide-binding protein	Lbp	2,6994	n.s.	4	4	4
Q6QLQ4	C-type lectin domain family 7 member A	Clec7a	2,6310	n.s.	2	2	2
P06800	Receptor-type tyrosine-protein phosphatase C	Ptprc	2,6197	n.s.	4	4	3
Q8K094	Poliovirus receptor	Pvr	2,6190	n.s.	1	2	1
P10810	Monocyte differentiation antigen CD14	Cd14	2,5804	n.s.	4	5	5
P29758	Ornithine aminotransferase, mitochondrial	Oat	2,5724	*	2	4	3
P09813	Apolipoprotein A-II	Apoa2	2,5624	n.s.	2	3	2
Q8R121	Protein Z-dependent protease inhibitor	Serpina10	2,5621	n.s.	4	5	4
Q61581	Insulin-like growth factor-binding protein 7	Igfbp7	2,4830	n.s.	4	5	2
P97798	Neogenin	Neo1	2,4563	n.s.	4	2	1
Q99KK7	Dipeptidyl peptidase 3	Dpp3	2,4524	n.s.	1	2	1
O08529	Calpain-2 catalytic subunit	Capn2	2,4491	n.s.	3	3	3
Q61702	Inter-alpha-trypsin inhibitor heavy chain H1	Itih1	2,3845	*	4	14	11
Q8BH61	Coagulation factor XIII A chain	F13a1	2,3431	n.s.	3	2	1
Q8VDD5	Myosin-9	Myh9	2,3316	n.s.	4	4	4
Q9CQI6	Coactosin-like protein	Cotl1	2,3093	n.s.	4	4	4
Q91VI7	Ribonuclease inhibitor	Rnh1	2,3077	n.s.	3	4	3
Q07456	Protein AMBP	Ambp	2,2911	n.s.	4	10	10
Q64277	ADP-ribosyl cyclase 2	Bst1	2,2779	n.s.	1	2	2
Q9DCN2	NADH-cytochrome b5 reductase 3	Cyb5r3	2,2733	n.s.	2	4	4
Q9CZD3	Glycine--tRNA ligase	Gars	2,2224	n.s.	1	2	1
Q60963	Platelet-activating factor acetylhydrolase	Pla2g7	2,1988	n.s.	2	2	2
P28293	Cathepsin G	Ctsg	2,1920	n.s.	4	2	2
P16301	Phosphatidylcholine-sterol acyltransferase	Lcat	2,1919	n.s.	3	2	1
P05555	Integrin alpha-M	Itgam	2,1793	*	4	13	11
P97449	Aminopeptidase N	Anpep	2,1553	n.s.	3	4	2
P11370	Retrovirus-related Env polyprotein from Fv-4 locus	Fv4	2,1434	*	4	4	4
A6X935	Inter alpha-trypsin inhibitor, heavy chain 4	Itih4	2,0928	*	4	29	25
P57780	Alpha-actinin-4	Actn4	2,0645	n.s.	2	2	2
P05622	Platelet-derived growth factor receptor beta	Pdgfrb	2,0561	n.s.	2	2	1
P11688	Integrin alpha-5	Itga5	2,0412	n.s.	4	3	3
P21614	Vitamin D-binding protein	Gc	2,0212	n.s.	4	16	16
P31725	Protein S100-A9	S100a9	2,0110	n.s.	4	7	6

Q91X72	Hemopexin	Hpx	1,9804	***	4	70	54
P51660	Peroxisomal multifunctional enzyme type 2	Hsd17b4	1,9688	n.s.	1	2	1
Q8BND5	Sulfhydryl oxidase 1	Qsox1	1,9333	**	4	13	11
Q9DBB9	Carboxypeptidase N subunit 2	Cpn2	1,9285	n.s.	4	7	6
Q9Z1Q5	Chloride intracellular channel protein 1	Clic1	1,8948	n.s.	3	2	1
P54320	Elastin	Eln	1,8872	n.s.	4	4	4
Q6S9I0	Kng2 protein	Kng2	1,8536	n.s.	4	6	5
P18181	CD48 antigen	Cd48	1,8525	n.s.	2	2	2
Q8BHN3	Neutral alpha-glucosidase AB	Ganab	1,8478	n.s.	1	3	2
Q00897	Alpha-1-antitrypsin 1-4	Serpina1d	1,8359	n.s.	2	3	3
Q9D1H9	Microfibril-associated glycoprotein 4	Mfap4	1,8302	n.s.	4	7	6
Q64471	Glutathione S-transferase theta-1	Gstt1	1,8281	n.s.	2	2	2
E9PV24	Protein Fga	Fga	1,8243	n.s.	4	2	1
O70370	Cathepsin S	Ctss	1,8227	n.s.	3	2	2
A2AVA0	Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1	Svep1	1,8038	n.s.	4	2	1
Q9R0M4	Podocalyxin	Podxl	1,7890	n.s.	4	4	4
Q80V42	Carboxypeptidase M	Cpm	1,7875	n.s.	4	4	4
P70669	Metalloendopeptidase homolog PEX	Phex	1,7807	n.s.	4	2	2
Q00651	Integrin alpha-4	Itga4	1,7698	n.s.	3	2	2
O35206	Collagen alpha-1(XV) chain	Col15a1	1,7569	*	4	6	4
O09131	Glutathione S-transferase omega-1	Gsto1	1,7546	n.s.	4	2	1
P11152	Lipoprotein lipase	Lpl	1,7466	n.s.	4	3	3
B2RXS4	Plexin-B2	Plxnb2	1,7430	n.s.	4	11	11
P01029	Complement C4-B	C4b	1,7295	***	4	34	32
Q8CJ69	BMP-binding endothelial regulator protein	Bmper	1,7223	n.s.	4	3	2
Q08761	Vitamin K-dependent protein S	Pros1	1,7138	n.s.	4	4	3
Q7TPR4	Alpha-actinin-1	Actn1	1,7050	n.s.	3	3	2
D3YXF5	Oxidation resistance protein 1	C7	1,6944	n.s.	4	4	3
Q9JHH6	Carboxypeptidase B2	Cpb2	1,6937	n.s.	4	3	2
O88968	Transcobalamin-2	Tcn2	1,6796	n.s.	4	5	5
Q9QUM0	Integrin alpha-IIb	Itga2b	1,6789	*	4	13	13
Q9CRA0	Ecto-ADP-ribosyltransferase 4	Art4	1,6789	n.s.	4	5	3
Q5SWU9	Acetyl-CoA carboxylase 1	Acaca	1,6530	n.s.	4	9	6
P97290	Plasma protease C1 inhibitor	Serping1	1,6299	n.s.	4	8	6
Q7TPD3	Roundabout homolog 2	Robo2	1,6086	*	3	2	2
P14211	Calreticulin	Calr	1,6001	n.s.	2	2	2
Q9WTI7	Unconventional myosin-Ic	Myo1c	1,5968	n.s.	4	3	3
P70663	SPARC-like protein 1	Sparcl1	1,5859	n.s.	4	10	8
P28828	Receptor-type tyrosine-protein phosphatase mu	Ptprm	1,5856	n.s.	4	3	3
Q62028	Secretory phospholipase A2 receptor	Pla2r1	1,5850	n.s.	2	2	1
P05202	Aspartate aminotransferase, mitochondrial	Got2	1,5749	n.s.	1	2	2
P43025	Tetranectin	Clec3b	1,5659	n.s.	2	2	2
P56135	ATP synthase subunit f, mitochondrial	Atp5j2	1,5602	n.s.	1	2	2
Q8BHC0	Lymphatic vessel endothelial hyaluronic acid receptor 1	Lyve1	1,5569	n.s.	4	9	10
P10404	MLV-related proviral Env polyprotein	null	1,5504	n.s.	4	3	3
P33267	Cytochrome P450 2F2	Cyp2f2	1,5464	n.s.	4	5	4
Q3TW96	UDP-N-acetylhexosamine pyrophosphorylase-like protein 1	Uap1l1	1,5451	n.s.	1	2	1
Q60710	Deoxynucleoside triphosphate triphosphohydrolase SAMHD1	Samhd1	1,5411	*	1	2	1
Q9JHI0	Matrix metalloproteinase-19	Mmp19	1,5401	n.s.	4	2	2
P14602	Heat shock protein beta-1	Hspb1	1,5359	n.s.	2	2	2
O08709	Peroxiredoxin-6	Prdx6	1,5259	n.s.	4	6	4
P21995	Embigin	Emb	1,5244	n.s.	4	4	4
Q8VCM7	Fibrinogen gamma chain	Fgg	1,5203	*	4	18	19
B2RU80	Receptor-type tyrosine-protein phosphatase beta	Ptprb	1,5169	n.s.	2	3	2
P17751	Triosephosphate isomerase	Tpi1	1,5026	n.s.	3	2	2

P19221	Prothrombin	F2	1,5016	n.s.	4	12	12
P54761	Ephrin type-B receptor 4	Ephb4	1,4990	n.s.	4	4	3
O54990	Prominin-1	Prom1	1,4970	n.s.	4	3	3
P19096	Fatty acid synthase	Fasn	1,4947	n.s.	4	6	4
O08677	Kininogen-1	Kng1	1,4939	n.s.	4	21	20
Q8VCG4	Complement component C8 gamma chain	C8g	1,4905	n.s.	4	3	3
Q01279	Epidermal growth factor receptor	Egfr	1,4674	n.s.	4	5	4
Q8CG14	Complement C1s-A subcomponent	C1sa	1,4566	n.s.	3	3	3
Q8BMS1	Trifunctional enzyme subunit alpha, mitochondrial	Hadha	1,4407	n.s.	4	4	4
O88307	Sortilin-related receptor	Sorl1	1,4345	n.s.	3	2	2
Q9JKF1	Ras GTPase-activating-like protein IQGAP1	Iqgap1	1,4334	n.s.	4	2	1
P20152	Vimentin	Vim	1,4246	n.s.	4	4	4
O08999	Latent-transforming growth factor beta-binding protein 2	Ltbp2	1,4053	n.s.	4	10	6
P35505	Fumarylacetoacetase	Fah	1,4047	n.s.	2	3	3
Q7TQ62	Podocan	Podn	1,3937	n.s.	4	6	5
P06909	Complement factor H	Cfh	1,3914	n.s.	4	18	18
P97298	Pigment epithelium-derived factor	Serpinf1	1,3889	n.s.	4	5	4
P97429	Annexin A4	Anxa4	1,3690	n.s.	4	3	3
Q9D154	Leukocyte elastase inhibitor A	Serpinb1a	1,3682	n.s.	3	3	3
P58242	Acid sphingomyelinase-like phosphodiesterase 3b	Smpdl3b	1,3546	n.s.	2	3	3
Q9ESB3	Histidine-rich glycoprotein	Hrg	1,3545	n.s.	4	20	19
Q07797	Galectin-3-binding protein	Lgals3bp	1,3404	n.s.	4	4	3
P52793	Ephrin-A1	Efna1	1,3346	n.s.	1	2	1
D3YWD1	Protein Col6a3 (Fragment)	Col6a3	1,3299	n.s.	4	7	8
Q91ZE5	EGF-like module-containing mucin-like hormone receptor-like 4	Emr4	1,3223	n.s.	1	2	1
P14106	Complement C1q subcomponent subunit B	C1qb	1,3147	n.s.	4	2	2
P70180	Atrial natriuretic peptide receptor 3	Npr3	1,3140	n.s.	4	2	2
O70362	Phosphatidylinositol-glycan-specific phospholipase D	Gpld1	1,3079	n.s.	4	3	3
P15379	CD44 antigen	Cd44	1,3005	n.s.	4	4	4
Q61838	Alpha-2-macroglobulin	A2m	1,2752	***	4	50	42
P11276	Fibronectin	Fn1	1,2656	*	4	59	59
Q03311	Cholinesterase	Bche	1,2573	n.s.	4	2	2
P12246	Serum amyloid P-component	Apcs	1,2316	n.s.	4	3	3
P97468	Chemokine-like receptor 1	Cmklr1	1,2120	n.s.	4	2	2
P47911	60S ribosomal protein L6	Rpl6	1,2068	n.s.	4	2	2
P10605	Cathepsin B	Ctsb	1,1936	n.s.	4	3	3
Q8VEM8	Phosphate carrier protein, mitochondrial	Slc25a3	1,1934	n.s.	4	3	3
P08249	Malate dehydrogenase, mitochondrial	Mdh2	1,1694	n.s.	4	2	2
Q9CQ60	6-phosphogluconolactonase	Pgls	1,1685	n.s.	2	3	1
P28665	Murinoglobulin-1	Mug1	1,1598	n.s.	4	16	16
Q640N1	Adipocyte enhancer-binding protein 1	Aebp1	1,1445	n.s.	4	2	2
P37889	Fibulin-2	Fbln2	1,1388	n.s.	4	9	8
Q61830	Macrophage mannose receptor 1	Mrc1	1,1300	n.s.	4	10	8
P24063	Integrin alpha-L	Itgal	1,1263	*	4	8	6
P26041	Moesin	Msn	1,1162	n.s.	4	4	3
Q9QXC1	Fetuin-B	Fetub	1,1061	n.s.	4	8	8
Q61129	Complement factor I	Cfi	1,1008	n.s.	4	9	6
Q8K0E8	Fibrinogen beta chain	Fgb	1,0922	n.s.	4	20	21
P08074	Carbonyl reductase [NADPH] 2	Cbr2	1,0755	n.s.	4	7	6
O70458	Oncostatin-M-specific receptor subunit beta	Osmr	1,0692	n.s.	3	2	2
P30412	Peptidyl-prolyl cis-trans isomerase C	Ppic	1,0686	n.s.	4	4	4
P16045	Galectin-1	Lgals1	1,0551	n.s.	4	2	1
P11835	Integrin beta-2	Itgb2	1,0303	n.s.	4	10	9
P10107	Annexin A1	Anxa1	1,0257	n.s.	4	8	8
Q9R0P3	S-formylglutathione hydrolase	Esd	1,0186	n.s.	4	2	1
P01027	Complement C3	C3	1,0134	***	4	69	65

O09164	Extracellular superoxide dismutase [Cu-Zn]	Sod3	0,9970	n.s.	4	14	14
Q62470	Integrin alpha-3	Itga3	0,9822	n.s.	4	26	29
Q64726	Zinc-alpha-2-glycoprotein	Azgp1	0,9781	n.s.	4	4	4
P21180	Complement C2	C2	0,9667	n.s.	4	2	2
Q8R3G9	Tetraspanin-8	Tspan8	0,9579	n.s.	4	3	3
Q8R2G4	Ecto-ADP-ribosyltransferase 3	Art3	0,9502	n.s.	4	6	6
Q61598	Rab GDP dissociation inhibitor beta	Gdi2	0,9459	n.s.	4	7	6
P53986	Monocarboxylate transporter 1	Slc16a1	0,9116	*	1	2	2
P17563	Selenium-binding protein 1	Selenbp1	0,9009	n.s.	4	2	3
Q02053	Ubiquitin-like modifier-activating enzyme 1	Uba1	0,9008	n.s.	3	2	1
Q6URW6	Myosin-14	Myh14	0,8921	n.s.	3	5	2
P14733	Lamin-B1	Lmnb1	0,8803	n.s.	3	3	3
Q8R2Q8	Bone marrow stromal antigen 2	Bst2	0,8528	n.s.	4	3	2
Q8QZT1	Acetyl-CoA acetyltransferase, mitochondrial	Acat1	0,8459	n.s.	2	7	5
P31428	Dipeptidase 1	Dpep1	0,8437	n.s.	4	12	13
Q9Z0Z4	Hephaestin	Heph	0,8412	n.s.	2	2	2
P08071	Lactotransferrin	Ltf	0,8394	n.s.	4	7	6
P51655	Glypican-4	Gpc4	0,8158	n.s.	4	3	2
P14824	Annexin A6	Anxa6	0,8100	n.s.	4	6	4
Q9QXS1	Plectin	Plec	0,8080	n.s.	2	2	2
Q99KC8	von Willebrand factor A domain-containing protein 5A	Vwa5a	0,8042	n.s.	4	6	4
P48036	Annexin A5	Anxa5	0,7953	n.s.	4	3	3
P15261	Interferon gamma receptor 1	Ifngr1	0,7908	n.s.	4	2	1
O88947	Coagulation factor X	F10	0,7557	n.s.	4	2	2
P05532	Mast/stem cell growth factor receptor Kit	Kit	0,7489	n.s.	4	10	9
Q9EQ20	Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial	Aldh6a1	0,7396	n.s.	4	3	3
Q4VBE4	Pikachurin	Egflam	0,7388	n.s.	4	11	11
Q01339	Beta-2-glycoprotein 1	Apoh	0,7360	n.s.	4	22	22
Q80YC5	Coagulation factor XII	F12	0,7359	n.s.	4	10	11
E9Q414	Apolipoprotein B-100	Apob	0,7220	n.s.	3	4	2
P16294	Coagulation factor IX	F9	0,7176	n.s.	4	2	2
Q61147	Ceruloplasmin	Cp	0,7130	n.s.	4	41	41
P35456	Urokinase plasminogen activator surface receptor	Plaur	0,7096	n.s.	4	3	2
P70387	Hereditary hemochromatosis protein homolog	Hfe	0,7073	n.s.	4	2	2
Q91XX1	Protein Pcdhgb6	Pcdhgb6	0,7023	n.s.	4	4	2
Q91ZV7	Plexin domain-containing protein 1	Plxdc1	0,6943	n.s.	4	2	2
O08553	Dihydropyrimidinase-related protein 2	Dpysl2	0,6886	n.s.	4	5	5
Q99JY3	GTPase IMAP family member 4	Gimap4	0,6878	n.s.	4	4	4
Q62230	Sialoadhesin	Siglec1	0,6809	n.s.	4	9	8
Q80X19	Collagen alpha-1(XIV) chain	Col14a1	0,6689	n.s.	3	4	3
Q9QZZ6	Dermatopontin	Dpt	0,6617	n.s.	4	2	2
Q99K41	EMILIN-1	Emilin1	0,6606	n.s.	4	19	18
P97370	Sodium/potassium-transporting ATPase subunit beta-3	Atp1b3	0,6482	n.s.	4	3	4
Q810U4	Neuronal cell adhesion molecule	Nrcam	0,6416	n.s.	4	6	5
P07759	Serine protease inhibitor A3K	Serpina3k	0,6350	n.s.	4	15	13
P47738	Aldehyde dehydrogenase, mitochondrial	Aldh2	0,6247	n.s.	4	5	5
P24549	Retinal dehydrogenase 1	Aldh1a1	0,6238	n.s.	4	9	9
P29788	Vitronectin	Vtn	0,6232	n.s.	4	10	11
P10852	4F2 cell-surface antigen heavy chain	Slc3a2	0,6214	n.s.	4	6	8
P04925	Major prion protein	Prnp	0,6178	n.s.	4	3	3
P26262	Plasma kallikrein	Klkb1	0,6137	n.s.	4	3	4
Q8VHY0	Chondroitin sulfate proteoglycan 4	Cspg4	0,6081	n.s.	4	27	20
Q63961	Endoglin	Eng	0,6032	n.s.	4	14	13
P29533	Vascular cell adhesion protein 1	Vcam1	0,5912	n.s.	4	7	6
Q61739	Integrin alpha-6	Itga6	0,5901	n.s.	4	12	11

P12367	cAMP-dependent protein kinase type II-alpha regulatory subunit	Prkar2a	0,5848	n.s.	4	2	2
Q99K47	Fibrinogen, alpha polypeptide	Fga	0,5829	*	4	27	23
Q8C310	Roundabout homolog 4	Robo4	0,5765	n.s.	4	4	3
P51885	Lumican	Lum	0,5755	n.s.	4	12	12
Q8VCT4	Carboxylesterase 1D	Ces1d	0,5754	n.s.	2	6	5
Q61703	Inter-alpha-trypsin inhibitor heavy chain H2	Itih2	0,5749	n.s.	4	11	9
Q08879	Fibulin-1	Fbln1	0,5748	n.s.	4	8	9
Q9Z0J1	Reversion-inducing cysteine-rich protein with Kazal motifs	Reck	0,5698	n.s.	4	5	5
P03953	Complement factor D	Cfd	0,5659	n.s.	4	4	3
P18572	Basigin	Bsg	0,5617	n.s.	4	3	2
Q8BKG3	Inactive tyrosine-protein kinase 7	Ptk7	0,5593	n.s.	4	11	12
P04919	Band 3 anion transport protein	Slc4a1	0,5501	n.s.	4	13	14
Q8CI59	Metalloreductase STEAP3	Steap3	0,5493	n.s.	4	7	8
Q9R097	Kunitz-type protease inhibitor 1	Spint1	0,5295	n.s.	2	3	3
B1AUH1	Receptor-type tyrosine-protein phosphatase U	Ptpru	0,5239	n.s.	4	2	1
P07356	Annexin A2	Anxa2	0,5170	n.s.	4	2	3
Q64735	Complement component receptor 1-like protein	Cr1l	0,5170	n.s.	4	9	11
P56677	Suppressor of tumorigenicity 14 protein homolog	St14	0,5137	n.s.	4	2	2
P97927	Laminin subunit alpha-4	Lama4	0,5092	n.s.	4	15	14
P22599	Alpha-1-antitrypsin 1-2	Serpina1b	0,5081	n.s.	4	8	6
Q8K183	Pyridoxal kinase	Pdxk	0,5059	n.s.	4	3	3
Q8CIZ8	von Willebrand factor	Vwf	0,5014	n.s.	4	7	6
Q9WUP0	Receptor activity-modifying protein 2	Ramp2	0,4863	n.s.	1	3	2
P23953	Carboxylesterase 1C	Ces1c	0,4823	n.s.	4	11	11
Q9QXH4	Integrin alpha-X	Itgax	0,4743	n.s.	4	4	4
Q9WVA4	Transgelin-2	Tagln2	0,4723	n.s.	4	3	2
Q9D1A2	Cytosolic non-specific dipeptidase	Cndp2	0,4690	n.s.	3	2	2
P97873	Lysyl oxidase homolog 1	Loxl1	0,4646	n.s.	4	4	3
Q9EPT5	Solute carrier organic anion transporter family member 2A1	Slco2a1	0,4636	n.s.	4	10	10
P35441	Thrombospondin-1	Thbs1	0,4632	n.s.	4	5	5
Q9QZF2	Glypican-1	Gpc1	0,4560	n.s.	4	2	2
Q8BTM8	Filamin-A	Flna	0,4448	n.s.	4	7	4
P11679	Keratin, type II cytoskeletal 8	Krt8	0,4377	n.s.	4	2	2
Q60597	2-oxoglutarate dehydrogenase, mitochondrial	Ogdh	0,4346	n.s.	1	2	2
P97333	Neuropilin-1	Nrp1	0,4325	n.s.	4	20	23
P26618	Platelet-derived growth factor receptor alpha	Pdgfra	0,4304	n.s.	4	4	4
Q02105	Complement C1q subcomponent subunit C	C1qc	0,4225	n.s.	4	3	3
Q62465	Synaptic vesicle membrane protein VAT-1 homolog	Vat1	0,4187	n.s.	4	3	3
O70318	Band 4.1-like protein 2	Epb41l2	0,4123	n.s.	1	3	3
P39061	Collagen alpha-1(XVIII) chain	Col18a1	0,4060	n.s.	4	16	16
Q61247	Alpha-2-antiplasmin	Serpinf2	0,4029	n.s.	4	12	11
P56528	ADP-ribosyl cyclase 1	Cd38	0,4015	n.s.	4	10	9
Q60930	Voltage-dependent anion-selective channel protein 2	Vdac2	0,3991	n.s.	4	2	2
Q91ZA3	Propionyl-CoA carboxylase alpha chain, mitochondrial	Pcca	0,3804	n.s.	4	30	30
P70389	Insulin-like growth factor-binding protein complex acid labile subunit	Igfals	0,3788	n.s.	4	6	5
Q62165	Dystroglycan	Dag1	0,3755	n.s.	4	6	6
P82198	Transforming growth factor-beta-induced protein ig-h3	Tgfbi	0,3745	n.s.	4	5	6
O35639	Annexin A3	Anxa3	0,3721	n.s.	4	6	6
P16406	Glutamyl aminopeptidase	Enpep	0,3692	n.s.	4	15	17
Q62181	Semaphorin-3C	Sema3c	0,3592	n.s.	4	4	4
P43406	Integrin alpha-V	Itgav	0,3572	n.s.	4	9	8
P40936	Indolethylamine N-methyltransferase	Inmt	0,3564	n.s.	4	18	18
Q8BH64	EH domain-containing protein 2	Ehd2	0,3491	n.s.	4	2	2
O89103	Complement component C1q receptor	Cd93	0,3471	n.s.	4	22	25
O70309	Integrin beta-5	Itgb5	0,3394	n.s.	4	6	6

Q4PZA2	Endothelin-converting enzyme 1	Ece1	0,3369	n.s.	4	3	2
Q923X1	EGF, latrophilin seven transmembrane domain-containing protein 1	Eltd1	0,3271	n.s.	3	2	2
P08226	Apolipoprotein E	Apoe	0,3248	n.s.	4	13	11
Q99KI0	Aconitate hydratase, mitochondrial	Aco2	0,3193	n.s.	4	3	3
P08228	Superoxide dismutase [Cu-Zn]	Sod1	0,3146	n.s.	3	2	2
Q64462	Cytochrome P450 4B1	Cyp4b1	0,3008	n.s.	4	4	4
P28654	Decorin	Dcn	0,2757	n.s.	4	11	11
P01898	H-2 class I histocompatibility antigen, Q10 alpha chain	H2-Q10	0,2753	n.s.	4	6	6
Q00623	Apolipoprotein A-I	Apoa1	0,2735	n.s.	4	26	28
Q9CZR2	N-acetylated-alpha-linked acidic dipeptidase 2	Naalad2	0,2712	n.s.	4	11	11
Q61165	Sodium/hydrogen exchanger 1	Slc9a1	0,2694	n.s.	4	2	2
Q61398	Procollagen C-endopeptidase enhancer 1	Pcolce	0,2659	n.s.	4	5	4
Q06890	Clusterin	Clu	0,2610	n.s.	4	31	30
Q99NB1	Acetyl-coenzyme A synthetase 2-like, mitochondrial	Acss1	0,2433	**	4	3	3
O54890	Integrin beta-3	Itgb3	0,2429	n.s.	4	11	8
O35632	Hyaluronidase-2	Hyal2	0,2353	n.s.	4	18	17
P55772	Ectonucleoside triphosphate diphosphohydrolase 1	Entpd1	0,2310	n.s.	4	15	18
Q60675	Laminin subunit alpha-2	Lama2	0,2283	n.s.	4	18	14
O88792	Junctional adhesion molecule A	F11r	0,1945	n.s.	4	6	5
P02469	Laminin subunit beta-1	Lamb1	0,1908	n.s.	4	11	12
P32507	Poliovirus receptor-related protein 2	Pvrl2	0,1860	n.s.	3	2	2
Q00560	Interleukin-6 receptor subunit beta	Il6st	0,1841	n.s.	4	10	10
P0CW02	Lymphocyte antigen 6C1	Ly6c1	0,1838	n.s.	4	3	3
P09470	Angiotensin-converting enzyme	Ace	0,1803	n.s.	4	63	68
Q8R2Y2	Cell surface glycoprotein MUC18	Mcam	0,1769	n.s.	4	30	29
Q08857	Platelet glycoprotein 4	Cd36	0,1750	n.s.	4	65	69
Q9Z0M6	CD97 antigen	Cd97	0,1740	n.s.	4	17	18
Q8K4G1	Latent-transforming growth factor beta-binding protein 4	Ltbp4	0,1688	n.s.	4	8	5
P40142	Transketolase	Tkt	0,1425	n.s.	4	2	3
Q61503	5'-nucleotidase	Nt5e	0,1377	n.s.	4	7	7
A6H584	Collagen alpha-5(VI) chain	Col6a5	0,1365	n.s.	1	2	1
Q61072	Disintegrin and metalloproteinase domain-containing protein 9	Adam9	0,1321	n.s.	4	3	3
Q8K2C9	Very-long-chain (3R)-3-hydroxyacyl-[acyl-carrier protein] dehydratase 3	ptplad1	0,1319	n.s.	4	3	3
Q07968	Coagulation factor XIII B chain	F13b	0,1304	n.s.	4	3	3
P28063	Proteasome subunit beta type-8	Psmb8	0,1258	n.s.	2	2	2
P33434	72 kDa type IV collagenase	Mmp2	0,1222	n.s.	4	3	3
Q99MR8	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial	Mccc1	0,1210	n.s.	4	30	30
Q91YQ5	Dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit 1	Rpn1	0,1194	n.s.	4	2	2
Q8CGC7	Bifunctional glutamate/proline--tRNA ligase	Eprs	0,1081	n.s.	4	3	2
P06728	Apolipoprotein A-IV	Apoa4	0,1064	n.s.	4	16	17
Q5FWI3	Transmembrane protein 2	Tmem2	0,1025	n.s.	4	35	37
Q8BH35	Complement component C8 beta chain	C8b	0,0938	n.s.	4	6	6
Q99MQ4	Asporin	Aspn	0,0899	n.s.	4	10	9
P35330	Intercellular adhesion molecule 2	Icam2	0,0806	n.s.	4	8	8
Q61508	Extracellular matrix protein 1	Ecm1	0,0783	n.s.	4	5	5
Q64449	C-type mannose receptor 2	Mrc2	0,0720	n.s.	4	9	9
Q921I1	Serotransferrin	Tf	0,0656	n.s.	4	3	3
Q9WVH9	Fibulin-5	Fbln5	0,0647	n.s.	4	7	6
Q8VDN2	Sodium/potassium-transporting ATPase subunit alpha-1	Atp1a1	0,0491	n.s.	4	2	2
Q3UQ28	Peroxidasin homolog	Pxdn	0,0416	n.s.	4	3	2
P35700	Peroxiredoxin-1	Prdx1	0,0392	n.s.	4	5	4
P10649	Glutathione S-transferase Mu 1	Gstm1	0,0332	n.s.	4	5	5

Q02788	Collagen alpha-2(VI) chain	Col6a2	0,0270	n.s.	4	18	17
D3Z2Q7	Protein Hmcn1	Hmcn1	0,0260	n.s.	4	4	3
P27512	Tumor necrosis factor receptor superfamily member 5	Cd40	0,0194	n.s.	4	3	3
Q61391	Neprilysin	Mme	0,0155	n.s.	4	12	14
P48962	ADP/ATP translocase 1	Slc25a4	0,0154	n.s.	4	3	3
P20918	Plasminogen	Plg	0,0083	n.s.	4	28	28
P21981	Protein-glutamine gamma-glutamyltransferase 2	Tgm2	-0,0152	n.s.	4	19	17
Q9DD06	Retinoic acid receptor responder protein 2	Rarres2	-0,0176	n.s.	3	3	2
O09118	Netrin-1	Ntn1	-0,0206	n.s.	4	2	2
Q9WVH6	Angiopoietin-4	Angpt4	-0,0268	n.s.	4	3	3
P04104	Keratin, type II cytoskeletal 1	Krt1	-0,0298	n.s.	4	2	2
A2A8L5	Receptor-type tyrosine-protein phosphatase F	Ptprf	-0,0591	n.s.	4	3	3
Q8C6B0	MCG20149, isoform CRA_a	Mettl7a1	-0,0601	n.s.	4	7	8
P32261	Antithrombin-III	Serpinc1	-0,0617	n.s.	4	19	19
Q62000	Mimecan	Ogn	-0,0785	n.s.	4	10	13
O08538	Angiopoietin-1	Angpt1	-0,0834	n.s.	4	2	1
Q8K4Q8	Collectin-12	Colec12	-0,0893	n.s.	4	5	4
Q91VC4	Plasmalemma vesicle-associated protein	Plvap	-0,0893	n.s.	4	28	29
P21956	Lactadherin	Mfge8	-0,0913	n.s.	4	14	15
Q9JK53	Prolargin	Prelp	-0,0962	n.s.	4	15	13
Q8BY89	Choline transporter-like protein 2	Slc44a2	-0,0964	n.s.	4	16	17
P13597	Intercellular adhesion molecule 1	Icam1	-0,1026	n.s.	4	16	20
B8JK39	Protein Itga9	Itga9	-0,1078	n.s.	4	2	2
P01887	Beta-2-microglobulin	B2m	-0,1156	n.s.	4	9	9
P25446	Tumor necrosis factor receptor superfamily member 6	Fas	-0,1192	n.s.	4	3	2
P29699	Alpha-2-HS-glycoprotein	Ahsg	-0,1201	n.s.	4	17	16
Q8CGA0	Protein phosphatase 1F	Ppm1f	-0,1240	n.s.	4	2	2
P09055	Integrin beta-1	Itgb1	-0,1244	n.s.	4	26	28
Q91ZX7	Prolow-density lipoprotein receptor-related protein 1	Lrp1	-0,1266	n.s.	4	8	9
Q9WV91	Prostaglandin F2 receptor negative regulator	Ptgfrn	-0,1294	n.s.	4	22	23
E9PWQ3	Protein Col6a3	Col6a3	-0,1299	n.s.	4	13	13
P07724	Serum albumin	Alb	-0,1483	n.s.	4	234	240
P00920	Carbonic anhydrase 2	Ca2	-0,1643	n.s.	4	4	4
O88322	Nidogen-2	Nid2	-0,1719	n.s.	4	31	34
G5E8Q8	MCG115189	Gpr116	-0,1926	n.s.	4	8	8
Q9DBV4	Matrix-remodeling-associated protein 8	Mxra8	-0,1932	n.s.	4	4	4
P55288	Cadherin-11	Cdh11	-0,1948	n.s.	4	3	3
P35242	Pulmonary surfactant-associated protein A	Sftpa1	-0,2079	n.s.	4	3	3
P13020	Gelsolin	Gsn	-0,2111	n.s.	4	18	21
P11680	Properdin	Cfp	-0,2367	n.s.	1	3	2
Q925F2	Endothelial cell-selective adhesion molecule	Esam	-0,2498	n.s.	4	17	18
A2AQ53	Fibrillin-1	Fbn1	-0,2517	n.s.	4	4	4
Q61490	CD166 antigen	Alcam	-0,2578	n.s.	4	12	12
P28653	Biglycan	Bgn	-0,2594	n.s.	4	5	5
P55065	Phospholipid transfer protein	Pltp	-0,2602	n.s.	4	8	8
Q3U7R1	Extended synaptotagmin-1	Esyt1	-0,2626	n.s.	1	3	2
Q61789	Laminin subunit alpha-3	Lama3	-0,2632	n.s.	4	77	77
Q64314	Hematopoietic progenitor cell antigen CD34	Cd34	-0,2690	n.s.	4	6	5
Q02858	Angiopoietin-1 receptor	Tek	-0,2709	n.s.	4	9	9
O70423	Membrane primary amine oxidase	Aoc3	-0,2710	n.s.	4	9	8
Q3V3R4	Integrin alpha-1	Itga1	-0,2756	n.s.	4	21	21
O88342	WD repeat-containing protein 1	Wdr1	-0,2900	n.s.	2	4	3
Q9WVT6	Carbonic anhydrase 14	Ca14	-0,3115	n.s.	4	6	7
Q9Z0F8	Disintegrin and metalloproteinase domain-containing protein 17	Adam17	-0,3123	n.s.	4	2	2
P70206	Plexin-A1	Plxna1	-0,3186	n.s.	4	3	3
Q9R069	Basal cell adhesion molecule	Bcam	-0,3246	*	4	62	72

P09803	Cadherin-1	Cdh1	-0,3333	n.s.	4	8	10
Q05909	Receptor-type tyrosine-protein phosphatase gamma	Ptprg	-0,3360	n.s.	4	3	3
P24270	Catalase	Cat	-0,3579	n.s.	2	4	4
Q8BXK9	Chloride intracellular channel protein 5	Clic5	-0,3657	n.s.	4	2	2
A6H6E2	Multimerin-2	Mmrn2	-0,3686	n.s.	4	12	13
Q99JW5	Epithelial cell adhesion molecule	Epcam	-0,3759	n.s.	4	2	3
Q8CFX3	Protein Pcdh1	Pcdh1	-0,3991	n.s.	4	5	6
O08530	Sphingosine 1-phosphate receptor 1	S1pr1	-0,3992	n.s.	4	3	3
Q9JIM1	Equilibrative nucleoside transporter 1	Slc29a1	-0,4012	n.s.	4	4	4
Q9DCT8	Cysteine-rich protein 2	Crip2	-0,4091	n.s.	4	2	2
Q61001	Laminin subunit alpha-5	Lama5	-0,4158	n.s.	4	70	76
Q9DBD0	Inhibitor of carbonic anhydrase	Ica	-0,4180	n.s.	4	2	2
Q8K2I3	Dimethylaniline monooxygenase [N-oxide-forming] 2	Fmo2	-0,4196	n.s.	4	11	11
Q8BWT1	3-ketoacyl-CoA thiolase, mitochondrial	Acaa2	-0,4251	n.s.	4	3	2
Q9R1Z8	Vinexin	Sorbs3	-0,4265	n.s.	4	3	2
D3Z689	ADAMTS-like 5, isoform CRA_c	Adamtsl5	-0,4338	n.s.	4	3	2
P14152	Malate dehydrogenase, cytoplasmic	Mdh1	-0,4375	n.s.	4	2	1
E9Q616	Protein Ahnak	Ahnak	-0,4391	n.s.	1	2	1
O54901	OX-2 membrane glycoprotein	Cd200	-0,4411	n.s.	4	5	5
P01901	H-2 class I histocompatibility antigen, K-B alpha chain	H2-K1	-0,4455	n.s.	4	2	2
P28843	Dipeptidyl peptidase 4	Dpp4	-0,4473	n.s.	4	20	20
Q8VCC9	Spondin-1	Spon1	-0,4644	n.s.	4	5	4
B1AYL1	Protein Scn7a	Scn7a	-0,4699	n.s.	4	8	8
Q64444	Carbonic anhydrase 4	Ca4	-0,4802	n.s.	4	15	16
P12023	Amyloid beta A4 protein	App	-0,4804	n.s.	4	3	3
Q61735	Leukocyte surface antigen CD47	Cd47	-0,4819	n.s.	4	2	2
B1B0C7	Basement membrane-specific heparan sulfate proteoglycan core protein	Hspg2	-0,4834	n.s.	4	25	28
P10493	Nidogen-1	Nid1	-0,4848	*	4	90	103
Q05793	Basement membrane-specific heparan sulfate proteoglycan core protein	Hspg2	-0,4861	n.s.	4	98	113
D3Z5G7	Protein Ces1b	Ces1b	-0,4904	n.s.	2	2	1
Q00896	Alpha-1-antitrypsin 1-3	Serpina1c	-0,5022	n.s.	4	3	3
Q62009	Periostin	Postn	-0,5147	n.s.	4	26	27
Q5SQ27	MCG140354	Sec14l3	-0,5292	n.s.	4	7	6
Q62469	Integrin alpha-2	Itga2	-0,5389	n.s.	4	15	17
Q61288	Serine/threonine-protein kinase receptor R3	Acvrl1	-0,5397	n.s.	4	4	4
O35188	Fractalkine	Cx3cl1	-0,5550	n.s.	4	2	2
Q63ZW6	Col4a5 protein	Col4a5	-0,5576	n.s.	4	4	4
Q99JR5	Tubulointerstitial nephritis antigen-like	Tinagl1	-0,5612	n.s.	4	8	9
P35969	Vascular endothelial growth factor receptor 1	Flt1	-0,5730	n.s.	4	12	12
O35452	Protein Tnxb	Tnxb	-0,5960	n.s.	4	28	26
O88662	Epithelial membrane protein 2	Emp2	-0,6050	n.s.	4	2	2
P35762	CD81 antigen	Cd81	-0,6153	*	2	2	2
Q792Z1	MCG140784	Try10	-0,6181	n.s.	2	4	5
Q19LI2	Alpha-1B-glycoprotein	A1bg	-0,6228	n.s.	4	9	11
P01902	H-2 class I histocompatibility antigen, K-D alpha chain	H2-K1	-0,6369	n.s.	4	14	15
Q8BFW7	Lipoma-preferred partner homolog	Lpp	-0,6417	n.s.	4	2	2
P49182	Heparin cofactor 2	Serpind1	-0,6489	n.s.	4	7	6
P55284	Cadherin-5	Cdh5	-0,6492	n.s.	4	31	31
Q62151	Advanced glycosylation end product-specific receptor	Ager	-0,6523	*	4	56	62
Q9R1N9	Collagen alpha-1(XIII) chain	Col13a1	-0,6597	n.s.	4	4	4
Q08481	Platelet endothelial cell adhesion molecule	Pecam1	-0,6688	n.s.	4	31	34
Q8BU59	Protein Xpnpep2	Xpnpep2	-0,6726	n.s.	4	7	6
Q8VCP9	C-type lectin domain family 14 member A	Clec14a	-0,6759	n.s.	4	12	12
O89020	Afamin	Afm	-0,6835	n.s.	4	5	4
Q59IW9	Protein Abi3bp	Abi3bp	-0,6912	n.s.	4	4	4

P82347	Delta-sarcoglycan	Sgcd	-0,6924	n.s.	4	4	5
Q8BPB5	EGF-containing fibulin-like extracellular matrix protein 1	Efemp1	-0,7093	n.s.	4	10	11
E9Q6C7	Latrophilin-2	Lphn2	-0,7122	n.s.	4	2	2
Q62009-2	Isoform 2 of Periostin	Postn	-0,7146	n.s.	4	2	2
Q9WUR0	Protein Tinag	Tinag	-0,7369	n.s.	4	3	3
P52795	Ephrin-B1	Efnb1	-0,7424	n.s.	4	2	2
Q01149	Collagen alpha-2(I) chain	Col1a2	-0,7460	*	4	71	74
A2ARA8	Integrin alpha-8	Itga8	-0,7831	n.s.	4	44	47
O88207	Collagen alpha-1(V) chain	Col5a1	-0,7876	*	4	8	8
Q06806	Tyrosine-protein kinase receptor Tie-1	Tie1	-0,8005	n.s.	4	5	4
P56857	Claudin-18	Cldn18	-0,8171	n.s.	4	4	4
P70274	Selenoprotein P	Sepp1	-0,8325	n.s.	4	6	5
P02468	Laminin subunit gamma-1	Lamc1	-0,8326	n.s.	4	29	33
C5H7W9	Advanced glycosylation end product-specific receptor	Ager	-0,8336	n.s.	4	22	23
Q61730	Interleukin-1 receptor accessory protein	Il1rap	-0,8378	n.s.	3	2	2
O55026	Ectonucleoside triphosphate diphosphohydrolase 2	Entpd2	-0,8397	n.s.	1	2	2
Q9JI33	Netrin-4	Ntn4	-0,8564	n.s.	4	8	9
O08573	Galectin-9	Lgals9	-0,8626	n.s.	4	5	4
E9Q0S6	Protein Tns1	Tns1	-0,8655	n.s.	4	9	9
O35682	Myeloid-associated differentiation marker	Myadm	-0,8718	n.s.	4	2	2
Q9WVJ9	EGF-containing fibulin-like extracellular matrix protein 2	Efemp2	-0,8741	n.s.	4	4	3
P07309	Transthyretin	Ttr	-0,8774	n.s.	4	15	15
Q04857	Collagen alpha-1(VI) chain	Col6a1	-0,8895	n.s.	4	13	13
A2ASQ1	Agrin	Agrn	-0,9071	n.s.	4	21	24
P08121	Collagen alpha-1(III) chain	Col3a1	-0,9317	n.s.	3	36	35
Q9CQ62	2,4-dienoyl-CoA reductase, mitochondrial	Decr1	-0,9402	n.s.	4	2	2
P01899	H-2 class I histocompatibility antigen, D-B alpha chain	H2-D1	-0,9675	n.s.	4	4	5
Q60994	Adiponectin	Adipoq	-0,9711	n.s.	4	4	5
Q61554	Fibrillin-1	Fbn1	-0,9799	n.s.	4	5	3
Q5ND28	Scavenger receptor class F member 1	Scarf1	-0,9891	n.s.	4	4	5
P97447	Four and a half LIM domains protein 1	Fhl1	-0,9904	n.s.	4	7	6
P70275	Semaphorin-3E	Sema3e	-0,9992	n.s.	4	2	2
F8VQJ3	Laminin subunit gamma-1	Lamc1	-1,0010	n.s.	4	2	2
Q61092	Laminin subunit gamma-2	Lamc2	-1,0081	n.s.	4	23	25
P50285	Dimethylaniline monooxygenase [N-oxide-forming] 1	Fmo1	-1,0124	n.s.	4	6	6
Q80V70	Multiple epidermal growth factor-like domains protein 6	Megf6	-1,0245	n.s.	4	2	2
Q4LDF6	Protein Cfhr2	Cfhr2	-1,0307	n.s.	4	4	4
P59222	Scavenger receptor class F member 2	Scarf2	-1,0409	n.s.	4	4	4
Q9R1W5	Calcitonin gene-related peptide type 1 receptor	Calcrl	-1,0554	*	4	5	6
Q61292	Laminin subunit beta-2	Lamb2	-1,0845	*	4	44	46
P11087	Collagen alpha-1(I) chain	Col1a1	-1,0927	n.s.	4	45	46
Q91V88	Nephronectin	Npnt	-1,0947	n.s.	4	11	12
Q61087	Laminin subunit beta-3	Lamb3	-1,1046	*	4	26	28
Q8BGT1	Fibronectin leucine rich transmembrane protein 3	Flrt3	-1,2168	n.s.	4	4	4
Q9EPX2	Papilin	Papln	-1,2215	n.s.	4	6	6
Q9Z1M0	P2X purinoceptor 7	P2rx7	-1,2590	n.s.	4	4	4
Q3U962	Collagen alpha-2(V) chain	Col5a2	-1,3525	n.s.	4	11	10
G5E874	Laminin subunit gamma-2	Lamc2	-1,3770	n.s.	4	3	3
Q60847	Collagen alpha-1(XII) chain	Col12a1	-1,4379	n.s.	4	13	12
P06869	Urokinase-type plasminogen activator	Plau	-1,4412	n.s.	4	3	3
Q9QZS0	Collagen alpha-3(IV) chain	Col4a3	-1,4748	n.s.	4	8	9
P42703	Leukemia inhibitory factor receptor	Lifr	-1,5367	n.s.	4	4	6
Q9R118	Serine protease HTRA1	Htra1	-1,6018	n.s.	4	2	2
Q8VCW8	Acyl-CoA synthetase family member 2, mitochondrial	Acsf2	-1,6066	n.s.	1	2	2
P02463	Collagen alpha-1(IV) chain	Col4a1	-1,6384	n.s.	4	4	6

Q9D1N2	Glycosylphosphatidylinositol-anchored high density lipoprotein-binding protein 1	Gpihbp1	-1,6384	n.s.	4	6	5
P08122	Collagen alpha-2(IV) chain	Col4a2	-1,6908	**	4	15	15
P15306	Thrombomodulin	Thbd	-2,0067	**	4	32	35
O35598	Disintegrin and metalloproteinase domain-containing protein 10	Adam10	-2,1162	n.s.	4	3	4
Q9QZR9	Collagen alpha-4(IV) chain	Col4a4	-2,2709	*	4	3	3
Q8BHL4	Retinoic acid-induced protein 3	Gprc5a	-2,4191	*	4	4	5
Q8R4W6	Procollagen C-endopeptidase enhancer 2	Pcolce2	-2,5708	n.s.	4	2	3
P02535	Keratin, type I cytoskeletal 10	Krt10	-2,7720	n.s.	2	2	2

8.5 List of proteins: xenograft lung metastases versus xenograft kidney tumours, human

Tab. 19: Xenograft metastases versus xenograft tumours. Protein list containing all proteins quantified with at least two proteotypic peptides in at least one group. Four samples from each of the four patient groups are quantified versus each four samples from their corresponding primary patient group. Positive values indicate an upregulation in lung metastases compared to kidney tumor tissue. average p-value: < 0.001: ***; < 0.01: **; < 0.05: *; ≥ 0.05: not significant (n.s.).

swiss prot accession no	protein name	gene name	$\log_2$ (average ratio)	average p-value	# metastases quantified	max. peptides metastases, human	max. peptides tumour, human
Q16787	Laminin subunit alpha-3	LAMA3	6,8783	n.s.	1	5	0
P13987	CD59 glycoprotein	CD59	5,9639	n.s.	3	2	4
P16070	CD44 antigen	CD44	5,4926	n.s.	3	3	3
P00533	Epidermal growth factor receptor	EGFR	4,7184	n.s.	3	4	4
P50895	Basal cell adhesion molecule	BCAM	4,4067	n.s.	2	4	3
O14880	Microsomal glutathione S-transferase 3	MGST3	4,3536	n.s.	1	2	2
P12111	Collagen alpha-3(VI) chain	COL6A3	4,3013	**	1	2	0
P14210	Hepatocyte growth factor	HGF	4,0171	n.s.	2	5	4
Q08431	Lactadherin	MFGE8	3,9967	n.s.	1	4	1
P13611	Versican core protein	VCAN	3,7983	n.s.	3	4	5
P26006	Integrin alpha-3	ITGA3	3,6115	*	3	5	10
Q9UNA0	A disintegrin and metalloproteinase with thrombospondin motifs 5	ADAMTS5	3,1719	n.s.	1	2	0
P23229	Integrin alpha-6	ITGA6	2,8504	n.s.	1	2	1
P41250	Glycine-tRNA ligase	GARS	2,8460	n.s.	3	2	2
P05556	Integrin beta-1	ITGB1	2,6507	n.s.	3	3	2
P04216	Thy-1 membrane glycoprotein	THY1	2,6103	n.s.	2	2	2
Q9UNN8	Endothelial protein C receptor	PROCR	2,5988	n.s.	1	2	1
Q9GZM7	Tubulointerstitial nephritis antigen-like	TINAGL1	2,5224	*	2	6	3
P39060	Collagen alpha-1(XVIII) chain	COL18A1	2,5215	n.s.	3	6	4
P43121	Cell surface glycoprotein MUC18	MCAM	2,5212	n.s.	1	4	2
P08581	Hepatocyte growth factor receptor	MET	2,4682	n.s.	1	3	1
P02768	Serum albumin	ALB	2,3647	n.s.	3	8	5
O00339	Matrilin-2	MATN2	2,3108	n.s.	1	3	2
P08603	Complement factor H	CFH	2,2556	n.s.	2	5	2
P05362	Intercellular adhesion molecule 1	ICAM1	2,0823	n.s.	4	9	6
P11047	Laminin subunit gamma-1	LAMC1	2,0745	n.s.	2	4	3
P14625	Endoplasmin	HSP90B1	1,9941	n.s.	3	1	2
P08195	4F2 cell-surface antigen heavy chain	SLC3A2	1,8565	*	2	3	3
Q13308	Inactive tyrosine-protein kinase 7	PTK7	1,7048	n.s.	1	2	2
P32004	Neural cell adhesion molecule L1	L1CAM	1,6708	n.s.	3	8	7
P98160	Basement membrane-specific heparan sulfate proteoglycan core protein	HSPG2	1,5053	**	4	38	29
P07996	Thrombospondin-1	THBS1	1,4118	n.s.	4	6	5
P14618	Pyruvate kinase PKM	PKM	1,3809	n.s.	3	2	3
O75915	PRA1 family protein 3	ARL6IP5	1,2946	n.s.	3	2	1
Q6YHK3	CD109 antigen	CD109	0,9256	n.s.	3	9	16
P08133	Annexin A6	ANXA6	0,9047	n.s.	1	2	2
P19320	Vascular cell adhesion protein 1	VCAM1	0,8675	n.s.	2	3	5
P00450	Ceruloplasmin	CP	0,8607	n.s.	2	7	8

P35613	Basigin	BSG	0,8464	n.s.	4	2	3
P00736	Complement C1r subcomponent	C1R	0,7887	n.s.	3	2	1
P21589	5'-nucleotidase	NT5E	0,7681	n.s.	1	2	4
P23526	Adenosylhomocysteinase	AHCY	0,7466	n.s.	3	2	2
Q16401	26S proteasome non-ATPase regulatory subunit 5	PSMD5	0,7417	n.s.	2	2	3
O15230	Laminin subunit alpha-5	LAMA5	0,7205	*	2	15	12
P36871	Phosphoglucomutase-1	PGM1	0,6748	n.s.	2	1	2
P32970	CD70 antigen	CD70	0,6269	*	1	1	3
P10909	Clusterin	CLU	0,4666	*	3	3	3
P11216	Glycogen phosphorylase, brain form	PYGB	0,4515	n.s.	3	2	3
Q02878	60S ribosomal protein L6	RPL6	0,3983	n.s.	3	3	2
P06744	Glucose-6-phosphate isomerase	GPI	0,3922	n.s.	3	1	2
P22413	Ectonucleotide pyrophosphatase/ phosphodiesterase family member 1	ENPP1	0,3904	n.s.	2	1	2
Q99715	Collagen alpha-1(XII) chain	COL12A1	0,2758	n.s.	2	13	17
Q09666	Neuroblast differentiation-associated protein AHNAK	AHNAK	0,2594	n.s.	2	2	2
P06756	Integrin alpha-V	ITGAV	0,2239	n.s.	2	2	3
Q14764	Major vault protein	MVP	0,1712	n.s.	1	1	2
Q00325	Phosphate carrier protein, mitochondrial	SLC25A3	0,1685	n.s.	4	3	3
P07814	Bifunctional glutamate/proline--tRNA ligase	EPRS	0,1664	n.s.	1	2	2
P07942	Laminin subunit beta-1	LAMB1	0,0673	n.s.	1	1	2
P17655	Calpain-2 catalytic subunit	CAPN2	0,0636	n.s.	3	3	4
P35527	Keratin, type I cytoskeletal 9	KRT9	0,0015	n.s.	4	22	24
P12955	Xaa-Pro dipeptidase	PEPD	0,0000	n.s.	0	0	2
P50395	Rab GDP dissociation inhibitor beta	GDI2	-0,0270	n.s.	3	2	2
P40939	Trifunctional enzyme subunit alpha, mitochondrial	HADHA	-0,0823	n.s.	2	1	3
P01024	Complement C3	C3	-0,1626	n.s.	4	20	18
Q15582	Transforming growth factor-beta-induced protein ig-h3	TGFBI	-0,1662	n.s.	4	12	11
P26641	Elongation factor 1-gamma	EEF1G	-0,1797	n.s.	4	2	3
P22314	Ubiquitin-like modifier-activating enzyme 1	UBA1	-0,2137	n.s.	2	3	3
P07237	Protein disulfide-isomerase	P4HB	-0,2539	n.s.	3	2	2
P18462	HLA class I histocompatibility antigen, A-25 alpha chain	HLA-A	-0,2948	n.s.	1	1	2
P08779	Keratin, type I cytoskeletal 16	KRT16	-0,3100	n.s.	4	4	3
Q9NZ08	Endoplasmic reticulum aminopeptidase 1	ERAP1	-0,3153	n.s.	2	2	3
Q08380	Galectin-3-binding protein	LGALS3BP	-0,3365	n.s.	2	1	2
Q86YZ3	Hornerin	HRNR	-0,3697	n.s.	4	8	9
Q8IUX7	Adipocyte enhancer-binding protein 1	AEBP1	-0,3853	n.s.	2	3	4
P07737	Profilin-1	PFN1	-0,4756	n.s.	4	5	5
P08107	Heat shock 70 kDa protein 1A/1B	HSPA1A	-0,6029	n.s.	2	1	2
P21980	Protein-glutamine gamma-glutamyltransferase 2	TGM2	-0,6488	n.s.	4	9	11
Q8N1N4	Keratin, type II cytoskeletal 78	KRT78	-0,6579	n.s.	1	1	2
P27487	Dipeptidyl peptidase 4	DPP4	-0,6919	n.s.	2	8	7
Q96PD2	Discoidin, CUB and LCCL domain-containing protein 2	DCBLD2	-0,6960	n.s.	1	3	6
O60701	UDP-glucose 6-dehydrogenase	UGDH	-0,7043	n.s.	2	2	2
Q14697	Neutral alpha-glucosidase AB	GANAB	-0,7295	n.s.	3	4	5
Q06830	Peroxiredoxin-1	PRDX1	-0,8113	n.s.	2	1	3
P27824	Calnexin	CANX	-0,8296	n.s.	3	2	3
Q96KP4	Cytosolic non-specific dipeptidase	CNDP2	-0,9272	n.s.	2	1	2
P04264	Keratin, type II cytoskeletal 1	KRT1	-0,9440	n.s.	4	22	20
P78527	DNA-dependent protein kinase catalytic subunit	PRKDC	-1,0370	n.s.	1	1	2
P60174	Triosephosphate isomerase	TPI1	-1,0778	n.s.	3	2	3
P49327	Fatty acid synthase	FASN	-1,0809	n.s.	2	2	2
P37802	Transgelin-2	TAGLN2	-1,1091	n.s.	2	2	2
P04843	Dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit 1	RPN1	-1,1693	n.s.	4	3	3
P02452	Collagen alpha-1(I) chain	COL1A1	-1,1979	n.s.	4	6	7
Q15274	Nicotinate-nucleotide pyrophosphorylase [carboxylating]	QPRT	-1,2113	n.s.	1	2	2

P13647	Keratin, type II cytoskeletal 5	KRT5	-1,2579	n.s.	3	1	2
P00338	L-lactate dehydrogenase A chain	LDHA	-1,3579	n.s.	3	4	8
P35908	Keratin, type II cytoskeletal 2 epidermal	KRT2	-1,3901	n.s.	4	10	14
Q7KZF4	Staphylococcal nuclease domain-containing protein 1	SND1	-1,4656	n.s.	1	3	3
P04406	Glyceraldehyde-3-phosphate dehydrogenase	GAPDH	-1,4991	n.s.	4	6	8
P51659	Peroxisomal multifunctional enzyme type 2	HSD17B4	-1,5216	n.s.	2	3	3
P08670	Vimentin	VIM	-1,5328	n.s.	4	4	5
P13645	Keratin, type I cytoskeletal 10	KRT10	-1,5870	n.s.	4	15	14
P30041	Peroxiredoxin-6	PRDX6	-1,5921	n.s.	2	2	2
P09211	Glutathione S-transferase P	GSTP1	-1,6081	n.s.	3	5	5
O00468	Agrin	AGRN	-1,6228	n.s.	1	5	10
P04179	Superoxide dismutase [Mn], mitochondrial	SOD2	-1,6235	n.s.	1	2	1
P09382	Galectin-1	LGALS1	-1,6467	n.s.	4	3	5
P07384	Calpain-1 catalytic subunit	CAPN1	-1,7032	n.s.	3	2	2
P15121	Aldose reductase	AKR1B1	-1,7275	n.s.	4	5	7
P10768	S-formylglutathione hydrolase	ESD	-1,7358	n.s.	2	1	2
P35579	Myosin-9	MYH9	-1,7981	n.s.	4	2	3
P09525	Annexin A4	ANXA4	-1,8636	n.s.	4	7	7
P42224	Signal transducer and activator of transcription 1-alpha/beta	STAT1	-1,8727	n.s.	3	3	3
P04792	Heat shock protein beta-1	HSPB1	-1,8775	n.s.	3	3	3
P00352	Retinal dehydrogenase 1	ALDH1A1	-1,8800	n.s.	4	12	11
P12956	X-ray repair cross-complementing protein 6	XRCC6	-1,9146	n.s.	2	1	2
P48643	T-complex protein 1 subunit epsilon	CCT5	-1,9341	*	1	2	2
P38606	V-type proton ATPase catalytic subunit A	ATP6V1A	-1,9383	n.s.	1	2	2
P04083	Annexin A1	ANXA1	-1,9973	n.s.	2	2	2
P17301	Integrin alpha-2	ITGA2	-2,0102	n.s.	1	1	2
P29401	Transketolase	TKT	-2,0302	n.s.	2	3	3
P27797	Calreticulin	CALR	-2,0363	n.s.	3	2	2
P19224	UDP-glucuronosyltransferase 1-6	UGT1A6	-2,0415	n.s.	1	1	2
P21333	Filamin-A	FLNA	-2,0465	n.s.	2	3	5
P40261	Nicotinamide N-methyltransferase	NNMT	-2,1018	n.s.	1	3	2
P06733	Alpha-enolase	ENO1	-2,1292	n.s.	4	6	8
P02794	Ferritin heavy chain	FTH1	-2,1676	n.s.	1	1	2
P02751	Fibronectin	FN1	-2,1912	**	4	29	37
P10809	60 kDa heat shock protein, mitochondrial	HSPD1	-2,2007	n.s.	1	2	2
P13489	Ribonuclease inhibitor	RNH1	-2,2029	n.s.	3	1	4
O00763	Acetyl-CoA carboxylase 2	ACACB	-2,2850	n.s.	1	2	3
P31939	Bifunctional purine biosynthesis protein PURH	ATIC	-2,3051	*	2	2	3
P02533	Keratin, type I cytoskeletal 14	KRT14	-2,3288	n.s.	4	2	3
O43570	Carbonic anhydrase 12	CA12	-2,4019	n.s.	1	4	5
P11413	Glucose-6-phosphate 1-dehydrogenase	G6PD	-2,4430	n.s.	3	3	3
O14773	Tripeptidyl-peptidase 1	TPP1	-2,4963	n.s.	2	1	2
Q01813	6-phosphofructokinase type C	PFKP	-2,4996	n.s.	2	2	5
P04075	Fructose-bisphosphate aldolase A	ALDOA	-2,7470	n.s.	4	2	4
Q16658	Fascin	FSCN1	-2,8700	n.s.	3	4	4
P30084	Enoyl-CoA hydratase, mitochondrial	ECHS1	-2,9170	n.s.	2	1	2
P08758	Annexin A5	ANXA5	-3,1945	n.s.	3	2	3
P05783	Keratin, type I cytoskeletal 18	KRT18	-3,2237	n.s.	1	1	2
Q13085	Acetyl-CoA carboxylase 1	ACACA	-3,2406	n.s.	4	15	17
P24821	Tenascin	TNC	-3,2509	*	1	2	5
Q9BRA2	Thioredoxin domain-containing protein 17	TXNDC17	-3,4214	n.s.	1	1	2
Q99829	Copine-1	CPNE1	-3,5030	n.s.	2	1	2
Q96DG6	Carboxymethylenebutenolidase homolog	CMBL	-3,5316	n.s.	2	1	4
P11498	Pyruvate carboxylase, mitochondrial	PC	-3,6682	n.s.	4	15	16
Q16270	Insulin-like growth factor-binding protein 7	IGFBP7	-3,6892	n.s.	1	1	2
Q13867	Bleomycin hydrolase	BLMH	-3,6941	n.s.	1	1	3

Q96RQ3	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial	MCCC1	-4,1021	**	4	29	29
P05165	Propionyl-CoA carboxylase alpha chain, mitochondrial	PCCA	-4,3494	***	4	18	20
O43707	Alpha-actinin-4	ACTN4	-4,9308	n.s.	0	0	2
P14868	Aspartate--tRNA ligase, cytoplasmic	DARS	-5,0024	n.s.	1	1	2
Q6PIU2	Neutral cholesterol ester hydrolase 1	NCEH1	-5,2384	*	1	1	2
P15144	Aminopeptidase N	ANPEP	-5,8732	n.s.	0	0	2

9 LIST OF FIGURES

10 LIST OF TABLES

11 REFERENCES

1. Siegel, R., D. Naishadham, and A. Jemal, *Cancer statistics, 2013.* CA Cancer J Clin, 2013. **63**(1): p. 11-30.
2. Gupta, K., et al., *Epidemiologic and socioeconomic burden of metastatic renal cell carcinoma (mRCC): a literature review.* Cancer Treat Rev, 2008. **34**(3): p. 193-205.
3. Sternberg, C.N., et al., *Pazopanib in locally advanced or metastatic renal cell carcinoma: results of a randomized phase III trial.* J Clin Oncol, 2010. **28**(6): p. 1061-8.
4. Strassberger, V., et al., *Chemical proteomic and bioinformatic strategies for the identification and quantification of vascular antigens in cancer.* J Proteomics, 2010. **73**(10): p. 1954-73.
5. Roesli, C. and D. Neri, *Methods for the identification of vascular markers in health and disease: from the bench to the clinic.* J Proteomics, 2010. **73**(11): p. 2219-29.
6. Rybak, J.N., et al., *In vivo protein biotinylation for identification of organ-specific antigens accessible from the vasculature.* Nat Methods, 2005. **2**(4): p. 291-8.
7. Roesli, C., D. Neri, and J.N. Rybak, *In vivo protein biotinylation and sample preparation for the proteomic identification of organ- and disease-specific antigens accessible from the vasculature.* Nat Protoc, 2006. **1**(1): p. 192-9.
8. Chow, W.H., et al., *Rising incidence of renal cell cancer in the United States.* JAMA, 1999. **281**(17): p. 1628-31.
9. Chow, W.H., L.M. Dong, and S.S. Devesa, *Epidemiology and risk factors for kidney cancer.* Nat Rev Urol, 2010. **7**(5): p. 245-57.
10. Hollingsworth, J.M., et al., *Rising incidence of small renal masses: a need to reassess treatment effect.* J Natl Cancer Inst, 2006. **98**(18): p. 1331-4.
11. Hollingsworth, J.M., et al., *Five-year survival after surgical treatment for kidney cancer: a population-based competing risk analysis.* Cancer, 2007. **109**(9): p. 1763-8.
12. Jemal, A., et al., *Global cancer statistics.* CA Cancer J Clin, 2011. **61**(2): p. 69-90.
13. *Cancer Facts & Figures 2013*. 2013, American Cancer Society, Atlanta, Georgia: www.cancer.org.
14. *Cancer incidence for common cancers*. 2014 [cited 2014 July 1]; Available from: www.cancerresearchuk.org/.
15. Bjorge, T., S. Tretli, and A. Engeland, *Relation of height and body mass index to renal cell carcinoma in two million Norwegian men and women.* Am J Epidemiol, 2004. **160**(12): p. 1168-76.
16. van Dijk, B.A., et al., *Relation of height, body mass, energy intake, and physical activity to risk of renal cell carcinoma: results from the Netherlands Cohort Study.* Am J Epidemiol, 2004. **160**(12): p. 1159-67.
17. Hunt, J.D., et al., *Renal cell carcinoma in relation to cigarette smoking: meta-analysis of 24 studies.* Int J Cancer, 2005. **114**(1): p. 101-8.
18. Yuan, J.M., et al., *Tobacco use in relation to renal cell carcinoma.* Cancer Epidemiol Biomarkers Prev, 1998. **7**(5): p. 429-33.
19. McLaughlin, J.K., et al., *International renal-cell cancer study. VIII. Role of diuretics, other anti-hypertensive medications and hypertension.* Int J Cancer, 1995. **63**(2): p. 216-21.
20. Ishikawa, I., et al., *Twenty-year follow-up of acquired renal cystic disease.* Clin Nephrol, 2003. **59**(3): p. 153-9.
21. Rakowski, S.K., et al., *Renal manifestations of tuberous sclerosis complex: Incidence, prognosis, and predictive factors.* Kidney Int, 2006. **70**(10): p. 1777-82.
22. McCredie, M., et al., *International renal-cell cancer study. II. Analgesics.* Int J Cancer, 1995. **60**(3): p. 345-9.
23. Pascual, D. and A. Borque, *Epidemiology of kidney cancer.* Adv Urol, 2008: p. 782381.

24. Rubagotti, A.M., G.; Boccardo, F.M., *Epidemiology of Kidney Cancer.* European Urology Supplements, 2006. **2006**(5): p. 558-565.
25. Lee, J.E., et al., *Intakes of coffee, tea, milk, soda and juice and renal cell cancer in a pooled analysis of 13 prospective studies.* Int J Cancer, 2007. **121**(10): p. 2246-53.
26. Rashidkhani, B., P. Lindblad, and A. Wolk, *Fruits, vegetables and risk of renal cell carcinoma: a prospective study of Swedish women.* Int J Cancer, 2005. **113**(3): p. 451-5.
27. van Dijk, B.A., et al., *Vegetable and fruit consumption and risk of renal cell carcinoma: results from the Netherlands cohort study.* Int J Cancer, 2005. **117**(4): p. 648-54.
28. Lee, C.T., et al., *Mode of presentation of renal cell carcinoma provides prognostic information.* Urol Oncol, 2002. **7**(4): p. 135-40.
29. Patard, J.J., et al., *Correlation between symptom graduation, tumor characteristics and survival in renal cell carcinoma.* Eur Urol, 2003. **44**(2): p. 226-32.
30. ; Available from: http://www.cancer.gov/.
31. Ljungberg, B., *Prognostic markers in renal cell carcinoma.* Curr Opin Urol, 2007. **17**(5): p. 303-8.
32. Motzer, R.J., et al., *Prognostic factors for survival in previously treated patients with metastatic renal cell carcinoma.* J Clin Oncol, 2004. **22**(3): p. 454-63.
33. Lam, J.S., et al., *Surveillance following radical or partial nephrectomy for renal cell carcinoma.* Curr Urol Rep, 2005. **6**(1): p. 7-18.
34. Edge, S.B. and American Joint Committee on Cancer, *AJCC cancer staging manual*. 7. ed. 2010, New York [u.a.]: Springer. XIV,648 S.
35. Bassil, B., D.E. Dosoretz, and G.R. Prout, Jr., *Validation of the tumor, nodes and metastasis classification of renal cell carcinoma.* J Urol, 1985. **134**(3): p. 450-4.
36. Fuhrman, S.A., L.C. Lasky, and C. Limas, *Prognostic significance of morphologic parameters in renal cell carcinoma.* Am J Surg Pathol, 1982. **6**(7): p. 655-63.
37. Kim, S.P., et al., *Independent validation of the 2010 American Joint Committee on Cancer TNM classification for renal cell carcinoma: results from a large, single institution cohort.* J Urol, 2011. **185**(6): p. 2035-9.
38. Novara, G., et al., *Validation of the 2009 TNM version in a large multi-institutional cohort of patients treated for renal cell carcinoma: are further improvements needed?* Eur Urol, 2010. **58**(4): p. 588-95.
39. Linehan, W.M., M.M. Walther, and B. Zbar, *The genetic basis of cancer of the kidney.* J Urol, 2003. **170**(6 Pt 1): p. 2163-72.
40. Renshaw, A.A., *Subclassification of renal cell neoplasms: an update for the practising pathologist.* Histopathology, 2002. **41**(4): p. 283-300.
41. Storkel, S., et al., *Classification of renal cell carcinoma: Workgroup No. 1. Union Internationale Contre le Cancer (UICC) and the American Joint Committee on Cancer (AJCC).* Cancer, 1997. **80**(5): p. 987-9.
42. Zambrano, N.R., et al., *Histopathology and molecular genetics of renal tumors toward unification of a classification system.* J Urol, 1999. **162**(4): p. 1246-58.
43. Jonasch, E., et al., *State of the science: an update on renal cell carcinoma.* Mol Cancer Res, 2012. **10**(7): p. 859-80.
44. Rini, B.I., S.C. Campbell, and B. Escudier, *Renal cell carcinoma.* Lancet, 2009. **373**(9669): p. 1119-32.
45. Nickerson, M.L., et al., *Mutations in a novel gene lead to kidney tumors, lung wall defects, and benign tumors of the hair follicle in patients with the Birt-Hogg-Dube syndrome.* Cancer Cell, 2002. **2**(2): p. 157-64.
46. Schmidt, L., et al., *Germline and somatic mutations in the tyrosine kinase domain of the MET proto-oncogene in papillary renal carcinomas.* Nat Genet, 1997. **16**(1): p. 68-73.
47. Tomlinson, I.P., et al., *Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer.* Nat Genet, 2002. **30**(4): p. 406-10.

48. Thoenes, W., S. Storkel, and H.J. Rumpelt, *Histopathology and classification of renal cell tumors (adenomas, oncocytomas and carcinomas). The basic cytological and histopathological elements and their use for diagnostics.* Pathol Res Pract, 1986. **181**(2): p. 125-43.
49. Mandriota, S.J., et al., *HIF activation identifies early lesions in VHL kidneys: evidence for site-specific tumor suppressor function in the nephron.* Cancer Cell, 2002. **1**(5): p. 459-68.
50. Cheville, J.C., et al., *Comparisons of outcome and prognostic features among histologic subtypes of renal cell carcinoma.* Am J Surg Pathol, 2003. **27**(5): p. 612-24.
51. Ljungberg, B., et al., *Prognostic significance of the Heidelberg classification of renal cell carcinoma.* Eur Urol, 1999. **36**(6): p. 565-9.
52. Bianchi, M., et al., *Distribution of metastatic sites in renal cell carcinoma: a population-based analysis.* Ann Oncol, 2012. **23**(4): p. 973-80.
53. Bonsib, S.M., *Renal lymphatics, and lymphatic involvement in sinus vein invasive (pT3b) clear cell renal cell carcinoma: a study of 40 cases.* Mod Pathol, 2006. **19**(5): p. 746-53.
54. Thompson, R.H., et al., *Metastatic renal cell carcinoma risk according to tumor size.* J Urol, 2009. **182**(1): p. 41-5.
55. Kim, J.J., B.I. Rini, and D.E. Hansel, *Von Hippel Lindau syndrome.* Adv Exp Med Biol, 2010. **685**: p. 228-49.
56. Linehan, W.M., et al., *Molecular diagnosis and therapy of kidney cancer.* Annu Rev Med, 2010. **61**: p. 329-43.
57. Patel, P.H., R.S. Chaganti, and R.J. Motzer, *Targeted therapy for metastatic renal cell carcinoma.* Br J Cancer, 2006. **94**(5): p. 614-9.
58. Foster, K., et al., *Somatic mutations of the von Hippel-Lindau disease tumour suppressor gene in non-familial clear cell renal carcinoma.* Hum Mol Genet, 1994. **3**(12): p. 2169-73.
59. Kim, W.Y. and W.G. Kaelin, *Role of VHL gene mutation in human cancer.* J Clin Oncol, 2004. **22**(24): p. 4991-5004.
60. Latif, F., et al., *Identification of the von Hippel-Lindau disease tumor suppressor gene.* Science, 1993. **260**(5112): p. 1317-20.
61. Maher, E.R. and W.G. Kaelin, Jr., *von Hippel-Lindau disease.* Medicine (Baltimore), 1997. **76**(6): p. 381-91.
62. McKusick, V.A., *Mendelian Inheritance in Man and its online version, OMIM.* Am J Hum Genet, 2007. **80**(4): p. 588-604.
63. Lonser, R.R., et al., *von Hippel-Lindau disease.* Lancet, 2003. **361**(9374): p. 2059-67.
64. Bjornsson, J., et al., *Tuberous sclerosis-associated renal cell carcinoma. Clinical, pathological, and genetic features.* Am J Pathol, 1996. **149**(4): p. 1201-8.
65. Moore, L.E., et al., *Von Hippel-Lindau (VHL) inactivation in sporadic clear cell renal cancer: associations with germline VHL polymorphisms and etiologic risk factors.* PLoS Genet, 2011. **7**(10): p. e1002312.
66. Gallou, C., et al., *Genotype-phenotype correlation in von Hippel-Lindau families with renal lesions.* Hum Mutat, 2004. **24**(3): p. 215-24.
67. Kibel, A., et al., *Binding of the von Hippel-Lindau tumor suppressor protein to Elongin B and C.* Science, 1995. **269**(5229): p. 1444-6.
68. Pause, A., et al., *The von Hippel-Lindau tumor-suppressor gene product forms a stable complex with human CUL-2, a member of the Cdc53 family of proteins.* Proc Natl Acad Sci U S A, 1997. **94**(6): p. 2156-61.
69. Pfaffenroth, E.C. and W.M. Linehan, *Genetic basis for kidney cancer: opportunity for disease-specific approaches to therapy.* Expert Opin Biol Ther, 2008. **8**(6): p. 779-90.
70. Ivan, M., et al., *HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing.* Science, 2001. **292**(5516): p. 464-8.
71. Kaelin, W.G., Jr., *The von Hippel-Lindau protein, HIF hydroxylation, and oxygen sensing.* Biochem Biophys Res Commun, 2005. **338**(1): p. 627-38.
72. Semenza, G.L., *Targeting HIF-1 for cancer therapy.* Nat Rev Cancer, 2003. **3**(10): p. 721-32.

73. Huang, L.E., et al., *Regulation of hypoxia-inducible factor 1alpha is mediated by an O2-dependent degradation domain via the ubiquitin-proteasome pathway.* Proc Natl Acad Sci U S A, 1998. **95**(14): p. 7987-92.
74. Morais, C., *Sunitinib resistance in renal cell carcinoma.* Journal of Kidney Cancer and VHL 2014. **1**(1): p. 1-11.
75. Simmons, M.N., C.J. Weight, and I.S. Gill, *Laparoscopic radical versus partial nephrectomy for tumors >4 cm: intermediate-term oncologic and functional outcomes.* Urology, 2009. **73**(5): p. 1077-82.
76. Leibovich, B.C., et al., *Nephron sparing surgery for appropriately selected renal cell carcinoma between 4 and 7 cm results in outcome similar to radical nephrectomy.* J Urol, 2004. **171**(3): p. 1066-70.
77. Uzzo, R.G. and A.C. Novick, *Nephron sparing surgery for renal tumors: indications, techniques and outcomes.* J Urol, 2001. **166**(1): p. 6-18.
78. Cohen, H.T. and F.J. McGovern, *Renal-cell carcinoma.* N Engl J Med, 2005. **353**(23): p. 2477-90.
79. Weight, C.J., et al., *Nephrectomy induced chronic renal insufficiency is associated with increased risk of cardiovascular death and death from any cause in patients with localized cT1b renal masses.* J Urol, 2010. **183**(4): p. 1317-23.
80. Margulis, V., et al., *Renal cell carcinoma clinically involving adjacent organs: experience with aggressive surgical management.* Cancer, 2007. **109**(10): p. 2025-30.
81. Blute, M.L., et al., *The Mayo Clinic experience with surgical management, complications and outcome for patients with renal cell carcinoma and venous tumour thrombus.* BJU Int, 2004. **94**(1): p. 33-41.
82. Franklin, J.R., et al., *Cytoreductive surgery in the management of metastatic renal cell carcinoma: the UCLA experience.* Semin Urol Oncol, 1996. **14**(4): p. 230-6.
83. Flanigan, R.C., et al., *Cytoreductive nephrectomy in patients with metastatic renal cancer: a combined analysis.* J Urol, 2004. **171**(3): p. 1071-6.
84. Thyavihally, Y.B., et al., *Management of renal cell carcinoma with solitary metastasis.* World J Surg Oncol, 2005. **3**: p. 48.
85. Mickisch, G.H., et al., *Radical nephrectomy plus interferon-alfa-based immunotherapy compared with interferon alfa alone in metastatic renal-cell carcinoma: a randomised trial.* Lancet, 2001. **358**(9286): p. 966-70.
86. Flanigan, R.C., et al., *Nephrectomy followed by interferon alfa-2b compared with interferon alfa-2b alone for metastatic renal-cell cancer.* N Engl J Med, 2001. **345**(23): p. 1655-9.
87. Tatsumi, T., et al., *Disease-associated bias in T helper type 1 (Th1)/Th2 CD4(+) T cell responses against MAGE-6 in HLA-DRB10401(+) patients with renal cell carcinoma or melanoma.* J Exp Med, 2002. **196**(5): p. 619-28.
88. Tatsumi, T., et al., *Disease stage variation in CD4+ and CD8+ T-cell reactivity to the receptor tyrosine kinase EphA2 in patients with renal cell carcinoma.* Cancer Res, 2003. **63**(15): p. 4481-9.
89. Dadian, G., et al., *Immunological parameters in peripheral blood of patients with renal cell carcinoma before and after nephrectomy.* Br J Urol, 1994. **74**(1): p. 15-22.
90. Gatenby, R.A., et al., *The possible role of postoperative azotemia in enhanced survival of patients with metastatic renal cancer after cytoreductive nephrectomy.* Cancer Res, 2002. **62**(18): p. 5218-22.
91. Rini, B.I., *Metastatic renal cell carcinoma: many treatment options, one patient.* J Clin Oncol, 2009. **27**(19): p. 3225-34.
92. Yagoda, A., B. Abi-Rached, and D. Petrylak, *Chemotherapy for advanced renal-cell carcinoma: 1983-1993.* Semin Oncol, 1995. **22**(1): p. 42-60.
93. Pizzocaro, G., et al., *Hormone treatment and sex steroid receptors in metastatic renal cell carcinoma: report of a multicentric prospective study.* Tumori, 1983. **69**(3): p. 215-20.
94. Papac, R.J., S.A. Ross, and A. Levy, *Renal cell carcinoma: analysis of 31 cases with assessment of endocrine therapy.* Am J Med Sci, 1977. **274**(3): p. 281-90.

95. Motzer, R.J. and P. Russo, *Systemic therapy for renal cell carcinoma.* J Urol, 2000. **163**(2): p. 408-17.
96. Vogelzang, N.J., E.R. Priest, and L. Borden, *Spontaneous regression of histologically proved pulmonary metastases from renal cell carcinoma: a case with 5-year followup.* J Urol, 1992. **148**(4): p. 1247-8.
97. Oliver, R.T., A.B. Nethersell, and J.M. Bottomley, *Unexplained spontaneous regression and alpha-interferon as treatment for metastatic renal carcinoma.* Br J Urol, 1989. **63**(2): p. 128-31.
98. Gleave, M.E., et al., *Interferon gamma-1b compared with placebo in metastatic renal-cell carcinoma. Canadian Urologic Oncology Group.* N Engl J Med, 1998. **338**(18): p. 1265-71.
99. Marten, A., et al., *Therapeutic vaccination against metastatic renal cell carcinoma by autologous dendritic cells: preclinical results and outcome of a first clinical phase I/II trial.* Cancer Immunol Immunother, 2002. **51**(11-12): p. 637-44.
100. McDermott, D.F. and B.I. Rini, *Immunotherapy for metastatic renal cell carcinoma.* BJU Int, 2007. **99**(5 Pt B): p. 1282-8.
101. Fyfe, G., et al., *Results of treatment of 255 patients with metastatic renal cell carcinoma who received high-dose recombinant interleukin-2 therapy.* J Clin Oncol, 1995. **13**(3): p. 688-96.
102. Yang, J.C., et al., *Randomized study of high-dose and low-dose interleukin-2 in patients with metastatic renal cancer.* J Clin Oncol, 2003. **21**(16): p. 3127-32.
103. Klapper, J.A., et al., *High-dose interleukin-2 for the treatment of metastatic renal cell carcinoma : a retrospective analysis of response and survival in patients treated in the surgery branch at the National Cancer Institute between 1986 and 2006.* Cancer, 2008. **113**(2): p. 293-301.
104. Fisher, R.I., S.A. Rosenberg, and G. Fyfe, *Long-term survival update for high-dose recombinant interleukin-2 in patients with renal cell carcinoma.* Cancer J Sci Am, 2000. **6 Suppl 1**: p. S55-7.
105. Bui, M.H., et al., *Carbonic anhydrase IX is an independent predictor of survival in advanced renal clear cell carcinoma: implications for prognosis and therapy.* Clin Cancer Res, 2003. **9**(2): p. 802-11.
106. Atkins, M., et al., *Carbonic anhydrase IX expression predicts outcome of interleukin 2 therapy for renal cancer.* Clin Cancer Res, 2005. **11**(10): p. 3714-21.
107. Wirth, M.P., *Immunotherapy for metastatic renal cell carcinoma.* Urol Clin North Am, 1993. **20**(2): p. 283-95.
108. Bukowski, R.M. and A.C. Novick, *Clinical practice guidelines: renal cell carcinoma.* Cleve Clin J Med, 1997. **64 Suppl 1**: p. SI1-44; quiz SI45-7.
109. Steineck, G., et al., *Recombinant leukocyte interferon alpha-2a and medroxyprogesterone in advanced renal cell carcinoma. A randomized trial.* Acta Oncol, 1990. **29**(2): p. 155-62.
110. Pyrhonen, S., et al., *Prospective randomized trial of interferon alfa-2a plus vinblastine versus vinblastine alone in patients with advanced renal cell cancer.* J Clin Oncol, 1999. **17**(9): p. 2859-67.
111. *Interferon-alpha and survival in metastatic renal carcinoma: early results of a randomised controlled trial. Medical Research Council Renal Cancer Collaborators.* Lancet, 1999. **353**(9146): p. 14-7.
112. Negrier, S., et al., *Medroxyprogesterone, interferon alfa-2a, interleukin 2, or combination of both cytokines in patients with metastatic renal carcinoma of intermediate prognosis: results of a randomized controlled trial.* Cancer, 2007. **110**(11): p. 2468-77.
113. McDermott, D.F., et al., *Randomized phase III trial of high-dose interleukin-2 versus subcutaneous interleukin-2 and interferon in patients with metastatic renal cell carcinoma.* J Clin Oncol, 2005. **23**(1): p. 133-41.
114. Hancock, B., et al., *Updated results of the MRC randomised controlled trial of alpha interferon vs. MPA in patients with metastatic renal carcinoma.* Proceedings of the American Society of Clinical Oncology, 2000. **19**: p. Abstract 1336.
115. Reeves, D.J. and C.Y. Liu, *Treatment of metastatic renal cell carcinoma.* Cancer Chemother Pharmacol, 2009. **64**(1): p. 11-25.

116. Kapoor, A.K. and S.J. Hotte, *Current status of cytokine therapy in management of patients with metastatic renal cell carcinoma.* Can Urol Assoc J, 2007. **1**(2 Suppl): p. S28-33.
117. Gore, M.E. and P. De Mulder, *Establishing the role of cytokine therapy in advanced renal cell carcinoma.* BJU Int, 2008. **101**(9): p. 1063-70.
118. McDermott, D.F. and M.B. Atkins, *Application of IL-2 and other cytokines in renal cancer.* Expert Opin Biol Ther, 2004. **4**(4): p. 455-68.
119. Gore, M.E., et al., *Interferon alfa-2a versus combination therapy with interferon alfa-2a, interleukin-2, and fluorouracil in patients with untreated metastatic renal cell carcinoma (MRC RE04/EORTC GU 30012): an open-label randomised trial.* Lancet, 2010. **375**(9715): p. 641-8.
120. Rini, B.I., *New strategies in kidney cancer: therapeutic advances through understanding the molecular basis of response and resistance.* Clin Cancer Res, 2010. **16**(5): p. 1348-54.
121. Rini, B.I., et al., *Bevacizumab plus interferon alfa compared with interferon alfa monotherapy in patients with metastatic renal cell carcinoma: CALGB 90206.* J Clin Oncol, 2008. **26**(33): p. 5422-8.
122. Motzer, R.J., et al., *Sunitinib versus interferon alfa in metastatic renal-cell carcinoma.* N Engl J Med, 2007. **356**(2): p. 115-24.
123. Motzer, R.J., et al., *Efficacy of everolimus in advanced renal cell carcinoma: a double-blind, randomised, placebo-controlled phase III trial.* Lancet, 2008. **372**(9637): p. 449-56.
124. Hudes, G., et al., *Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma.* N Engl J Med, 2007. **356**(22): p. 2271-81.
125. Escudier, B., et al., *Bevacizumab plus interferon alfa-2a for treatment of metastatic renal cell carcinoma: a randomised, double-blind phase III trial.* Lancet, 2007. **370**(9605): p. 2103-11.
126. Hudes, G., et al., *A phase 3, randomized, 3-arm study of temsirolimus (TEMSR) or interferon-alpha (IFN) or the combination of TEMSR + IFN in the treatment of first-line, poor-risk patients with advanced renal cell carcinoma (adv RCC)* Journal of Clinical Oncology, 2006 ASCO Annual Meeting Proceedings, 2006. **24 (18S):LBA4**.
127. Rini, B.I., et al., *Comparative effectiveness of axitinib versus sorafenib in advanced renal cell carcinoma (AXIS): a randomised phase 3 trial.* Lancet, 2011. **378**(9807): p. 1931-9.
128. Ravaud, A., *Treatment-associated adverse event management in the advanced renal cell carcinoma patient treated with targeted therapies.* Oncologist, 2011. **16 Suppl 2**: p. 32-44.
129. Patel, P.H., et al., *Phase I study combining treatment with temsirolimus and sunitinib malate in patients with advanced renal cell carcinoma.* Clin Genitourin Cancer, 2009. **7**(1): p. 24-7.
130. Feldman, D.R., et al., *Phase I trial of bevacizumab plus escalated doses of sunitinib in patients with metastatic renal cell carcinoma.* J Clin Oncol, 2009. **27**(9): p. 1432-9.
131. Johannsen, M., et al., *The tumour-targeting human L19-IL2 immunocytokine: preclinical safety studies, phase I clinical trial in patients with solid tumours and expansion into patients with advanced renal cell carcinoma.* Eur J Cancer, 2010. **46**(16): p. 2926-35.
132. [cited 2014; Available from: http://www.drugbank.ca.
133. Wu, J., S.O. Joseph, and F.M. Muggia, *Targeted therapy: its status and promise in selected solid tumors part I: areas of major impact.* Oncology (Williston Park), 2012. **26**(10): p. 936-43.
134. Najjar, Y.G. and B.I. Rini, *Novel agents in renal carcinoma: a reality check.* Ther Adv Med Oncol, 2012. **4**(4): p. 183-94.
135. Wilhelm, S., et al., *Discovery and development of sorafenib: a multikinase inhibitor for treating cancer.* Nat Rev Drug Discov, 2006. **5**(10): p. 835-44.
136. Carlomagno, F., et al., *BAY 43-9006 inhibition of oncogenic RET mutants.* J Natl Cancer Inst, 2006. **98**(5): p. 326-34.
137. Lierman, E., et al., *The ability of sorafenib to inhibit oncogenic PDGFRbeta and FLT3 mutants and overcome resistance to other small molecule inhibitors.* Haematologica, 2007. **92**(1): p. 27-34.
138. Unnithan, J. and B.I. Rini, *The role of targeted therapy in metastatic renal cell carcinoma.* ScientificWorldJournal, 2007. **7**: p. 800-7.
139. Lu, X., et al., *Sorafenib induces growth inhibition and apoptosis of human chondrosarcoma cells by blocking the RAF/ERK/MEK pathway.* J Surg Oncol, 2010. **102**(7): p. 821-6.

140. Guida, T., et al., *Sorafenib inhibits imatinib-resistant KIT and platelet-derived growth factor receptor beta gatekeeper mutants.* Clin Cancer Res, 2007. **13**(11): p. 3363-9.
141. Chen, X., Z.L. Ji, and Y.Z. Chen, *TTD: Therapeutic Target Database.* Nucleic Acids Res, 2002. **30**(1): p. 412-5.
142. Prenen, H., et al., *Efficacy of the kinase inhibitor SU11248 against gastrointestinal stromal tumor mutants refractory to imatinib mesylate.* Clin Cancer Res, 2006. **12**(8): p. 2622-7.
143. Mendel, D.B., et al., *In vivo antitumor activity of SU11248, a novel tyrosine kinase inhibitor targeting vascular endothelial growth factor and platelet-derived growth factor receptors: determination of a pharmacokinetic/pharmacodynamic relationship.* Clin Cancer Res, 2003. **9**(1): p. 327-37.
144. Guo, J., et al., *Inhibition of phosphorylation of the colony-stimulating factor-1 receptor (c-Fms) tyrosine kinase in transfected cells by ABT-869 and other tyrosine kinase inhibitors.* Mol Cancer Ther, 2006. **5**(4): p. 1007-13.
145. Roskoski, R., Jr., *Sunitinib: a VEGF and PDGF receptor protein kinase and angiogenesis inhibitor.* Biochem Biophys Res Commun, 2007. **356**(2): p. 323-8.
146. O'Farrell, A.M., et al., *An innovative phase I clinical study demonstrates inhibition of FLT3 phosphorylation by SU11248 in acute myeloid leukemia patients.* Clin Cancer Res, 2003. **9**(15): p. 5465-76.
147. Motzer, R.J., et al., *Overall survival and updated results for sunitinib compared with interferon alfa in patients with metastatic renal cell carcinoma.* J Clin Oncol, 2009. **27**(22): p. 3584-90.
148. Abrams, T.J., et al., *SU11248 inhibits KIT and platelet-derived growth factor receptor beta in preclinical models of human small cell lung cancer.* Mol Cancer Ther, 2003. **2**(5): p. 471-8.
149. Sonpavde, G. and T.E. Hutson, *Pazopanib: a novel multitargeted tyrosine kinase inhibitor.* Curr Oncol Rep, 2007. **9**(2): p. 115-9.
150. Deeks, E.D., *Pazopanib: in advanced soft tissue sarcoma.* Drugs, 2012. **72**(16): p. 2129-40.
151. Gerber, H.P. and N. Ferrara, *Pharmacology and pharmacodynamics of bevacizumab as monotherapy or in combination with cytotoxic therapy in preclinical studies.* Cancer Res, 2005. **65**(3): p. 671-80.
152. Radulovic, S. and S.K. Bjelogrlic, *Sunitinib, sorafenib and mTOR inhibitors in renal cancer.* J BUON, 2007. **12 Suppl 1**: p. S151-62.
153. Patard, J.J., et al., *Targeted therapy in renal cell carcinoma.* World J Urol, 2008. **26**(2): p. 135-40.
154. Rostaing, L. and N. Kamar, *mTOR inhibitor/proliferation signal inhibitors: entering or leaving the field?* J Nephrol, 2010. **23**(2): p. 133-42.
155. Coppin, C., *Everolimus: the first approved product for patients with advanced renal cell cancer after sunitinib and/or sorafenib.* Biologics, 2010. **4**: p. 91-101.
156. Albert, S., et al., *New inhibitors of the mammalian target of rapamycin signaling pathway for cancer.* Expert Opin Investig Drugs, 2010. **19**(8): p. 919-30.
157. Shah, N.P., et al., *Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia.* Cancer Cell, 2002. **2**(2): p. 117-25.
158. Pao, W., et al., *Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain.* PLoS Med, 2005. **2**(3): p. e73.
159. O'Hare, T., C.A. Eide, and M.W. Deininger, *Bcr-Abl kinase domain mutations, drug resistance, and the road to a cure for chronic myeloid leukemia.* Blood, 2007. **110**(7): p. 2242-9.
160. Kobayashi, S., et al., *EGFR mutation and resistance of non-small-cell lung cancer to gefitinib.* N Engl J Med, 2005. **352**(8): p. 786-92.
161. Rini, B.I. and M.B. Atkins, *Resistance to targeted therapy in renal-cell carcinoma.* Lancet Oncol, 2009. **10**(10): p. 992-1000.
162. de Bazelaire, C., et al., *Magnetic resonance imaging-measured blood flow change after antiangiogenic therapy with PTK787/ZK 222584 correlates with clinical outcome in metastatic renal cell carcinoma.* Clin Cancer Res, 2008. **14**(17): p. 5548-54.

163. Elia, G., T. Fugmann, and D. Neri, *From target discovery to clinical trials with armed antibody products.* J Proteomics, 2014.
164. Bjare, U., *Serum-free cell culture.* Pharmacol Ther, 1992. **53**(3): p. 355-74.
165. Kleinman, H.K., D. Philp, and M.P. Hoffman, *Role of the extracellular matrix in morphogenesis.* Curr Opin Biotechnol, 2003. **14**(5): p. 526-32.
166. Ertel, A., et al., *Pathway-specific differences between tumor cell lines and normal and tumor tissue cells.* Mol Cancer, 2006. **5**(1): p. 55.
167. Domcke, S., et al., *Evaluating cell lines as tumour models by comparison of genomic profiles.* Nat Commun, 2013. **4**: p. 2126.
168. Ljung, B.M., et al., *Cell dissociation techniques in human breast cancer--variations in tumor cell viability and DNA ploidy.* Breast Cancer Res Treat, 1989. **13**(2): p. 153-9.
169. Eramo, A., et al., *Identification and expansion of the tumorigenic lung cancer stem cell population.* Cell Death Differ, 2008. **15**(3): p. 504-14.
170. Adams, J.M., et al., *The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice.* Nature, 1985. **318**(6046): p. 533-8.
171. Jaenisch, R. and B. Mintz, *Simian virus 40 DNA sequences in DNA of healthy adult mice derived from preimplantation blastocysts injected with viral DNA.* Proc Natl Acad Sci U S A, 1974. **71**(4): p. 1250-4.
172. Gordon, J.W., et al., *Genetic transformation of mouse embryos by microinjection of purified DNA.* Proc Natl Acad Sci U S A, 1980. **77**(12): p. 7380-4.
173. Thomas, K.R. and M.R. Capecchi, *Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells.* Cell, 1987. **51**(3): p. 503-12.
174. Rudolph, K.L., *Telomeres and telomerase influence the course of human diseases, aging and carcinogenesis.* Curr Mol Med, 2005. **5**(2): p. 133-4.
175. Pollard, P.J., et al., *Targeted inactivation of fh1 causes proliferative renal cyst development and activation of the hypoxia pathway.* Cancer Cell, 2007. **11**(4): p. 311-9.
176. Rankin, E.B., J.E. Tomaszewski, and V.H. Haase, *Renal cyst development in mice with conditional inactivation of the von Hippel-Lindau tumor suppressor.* Cancer Res, 2006. **66**(5): p. 2576-83.
177. Kleymenova, E., et al., *Susceptibility to vascular neoplasms but no increased susceptibility to renal carcinogenesis in Vhl knockout mice.* Carcinogenesis, 2004. **25**(3): p. 309-15.
178. Fu, L., et al., *Generation of a mouse model of Von Hippel-Lindau kidney disease leading to renal cancers by expression of a constitutively active mutant of HIF1alpha.* Cancer Res, 2011. **71**(21): p. 6848-56.
179. Bosma, G.C., R.P. Custer, and M.J. Bosma, *A severe combined immunodeficiency mutation in the mouse.* Nature, 1983. **301**(5900): p. 527-30.
180. Ito, M., et al., *NOD/SCID/gamma(c)(null) mouse: an excellent recipient mouse model for engraftment of human cells.* Blood, 2002. **100**(9): p. 3175-82.
181. Kerbel, R.S., *Human tumor xenografts as predictive preclinical models for anticancer drug activity in humans: better than commonly perceived-but they can be improved.* Cancer Biol Ther, 2003. **2**(4 Suppl 1): p. S134-9.
182. Johnson, J.I., et al., *Relationships between drug activity in NCI preclinical in vitro and in vivo models and early clinical trials.* Br J Cancer, 2001. **84**(10): p. 1424-31.
183. An, Z., et al., *Development of a high metastatic orthotopic model of human renal cell carcinoma in nude mice: benefits of fragment implantation compared to cell-suspension injection.* Clin Exp Metastasis, 1999. **17**(3): p. 265-70.
184. Sivanand, S., et al., *A validated tumorgraft model reveals activity of dovitinib against renal cell carcinoma.* Sci Transl Med, 2012. **4**(137): p. 137ra75.
185. Grisanzio, C., et al., *Orthotopic xenografts of RCC retain histological, immunophenotypic and genetic features of tumours in patients.* J Pathol, 2011. **225**(2): p. 212-21.
186. *The website of the National Cancer Institute*. Available from: www.cancer.gov.
187. Frantzi, M., A. Bhat, and A. Latosinska, *Clinical proteomic biomarkers: relevant issues on study design & technical considerations in biomarker development.* Clin Transl Med, 2014. **3**(1): p. 7.

188. Jones, H.B., *On a new substance occurring in the urine of a patient with mollities ossium.* Philosophical Transactions of the Royal Society, 1848. **138**: p. 55-62.
189. Hayes, D.F., et al., *Tumor marker utility grading system: a framework to evaluate clinical utility of tumor markers.* J Natl Cancer Inst, 1996. **88**(20): p. 1456-66.
190. Kulasingam, V. and E.P. Diamandis, *Strategies for discovering novel cancer biomarkers through utilization of emerging technologies.* Nat Clin Pract Oncol, 2008. **5**(10): p. 588-99.
191. Gupta, S., et al., *Challenges and prospects for biomarker research: a current perspective from the developing world.* Biochim Biophys Acta, 2014. **1844**(5): p. 899-908.
192. Frantzi, M., M. Makridakis, and A. Vlahou, *Biomarkers for bladder cancer aggressiveness.* Curr Opin Urol, 2012. **22**(5): p. 390-6.
193. Fertig, E.J., R. Slebos, and C.H. Chung, *Application of genomic and proteomic technologies in biomarker discovery.* Am Soc Clin Oncol Educ Book, 2012: p. 377-82.
194. Brower, V., *Biomarkers: Portents of malignancy.* Nature, 2011. **471**(7339): p. S19-21.
195. Rifai, N., M.A. Gillette, and S.A. Carr, *Protein biomarker discovery and validation: the long and uncertain path to clinical utility.* Nat Biotechnol, 2006. **24**(8): p. 971-83.
196. Pavlou, M.P., E.P. Diamandis, and I.M. Blasutig, *The long journey of cancer biomarkers from the bench to the clinic.* Clin Chem, 2013. **59**(1): p. 147-57.
197. Fuzery, A.K., et al., *Translation of proteomic biomarkers into FDA approved cancer diagnostics: issues and challenges.* Clin Proteomics, 2013. **10**(1): p. 13.
198. Surinova, S., et al., *On the development of plasma protein biomarkers.* J Proteome Res, 2011. **10**(1): p. 5-16.
199. Altman, D.G., et al., *Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration.* BMC Med, 2012. **10**: p. 51.
200. Gallo, V., et al., *STrengthening the reporting of OBservational studies in Epidemiology-Molecular Epidemiology (STROBE-ME): an extension of the STROBE statement.* Eur J Epidemiol, 2011. **26**(10): p. 797-810.
201. Pepe, M.S., et al., *Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design.* J Natl Cancer Inst, 2008. **100**(20): p. 1432-8.
202. Tabb, D.L., *Quality assessment for clinical proteomics.* Clin Biochem, 2013. **46**(6): p. 411-20.
203. Greenbaum, D., et al., *Comparing protein abundance and mRNA expression levels on a genomic scale.* Genome Biol, 2003. **4**(9): p. 117.
204. Schwanhausser, B., et al., *Global quantification of mammalian gene expression control.* Nature, 2011. **473**(7347): p. 337-42.
205. Wilhelm, M., et al., *Mass-spectrometry-based draft of the human proteome.* Nature, 2014. **509**(7502): p. 582-7.
206. Hanash, S., et al., *Integration of proteomics into systems biology of cancer.* Wiley Interdiscip Rev Syst Biol Med, 2012. **4**(4): p. 327-37.
207. Kozak, K.R., et al., *Micro-volume wall-less immunoassays using patterned planar plates.* Lab Chip, 2013. **13**(7): p. 1342-50.
208. Tekin, H.C. and M.A. Gijs, *Ultrasensitive protein detection: a case for microfluidic magnetic bead-based assays.* Lab Chip, 2013. **13**(24): p. 4711-39.
209. Ostendorff, H.P., et al., *Multiplexed VeraCode bead-based serological immunoassay for colorectal cancer.* J Immunol Methods, 2013. **400-401**: p. 58-69.
210. Wilson, R., *Sensitivity and specificity: twin goals of proteomics assays. Can they be combined?* Expert Rev Proteomics, 2013. **10**(2): p. 135-49.
211. Tighe, P., et al., *Utility, reliability and reproducibility of immunoassay multiplex kits.* Methods, 2013. **61**(1): p. 23-9.
212. Stoevesandt, O. and M.J. Taussig, *Affinity proteomics: the role of specific binding reagents in human proteome analysis.* Expert Rev Proteomics, 2012. **9**(4): p. 401-14.
213. Binz, H.K., P. Amstutz, and A. Pluckthun, *Engineering novel binding proteins from nonimmunoglobulin domains.* Nat Biotechnol, 2005. **23**(10): p. 1257-68.

214. Baltzer, L., *Crossing borders to bind proteins--a new concept in protein recognition based on the conjugation of small organic molecules or short peptides to polypeptides from a designed set.* Anal Bioanal Chem, 2011. **400**(6): p. 1653-64.
215. Banerjee, J. and M. Nilsen-Hamilton, *Aptamers: multifunctional molecules for biomedical research.* J Mol Med (Berl), 2013. **91**(12): p. 1333-42.
216. *The website of Ventana Medical Systems*. Available from: www.ventana.com/.
217. Perez, E.A., et al., *Predictability of adjuvant trastuzumab benefit in N9831 patients using the ASCO/CAP HER2-positivity criteria.* J Natl Cancer Inst, 2012. **104**(2): p. 159-62.
218. Roge, R., S. Nielsen, and M. Vyberg, *Carb-3 Is the Superior Anti-CD15 Monoclonal Antibody for Immunohistochemistry.* Appl Immunohistochem Mol Morphol, 2013.
219. Cole, K.D., H.J. He, and L. Wang, *Breast cancer biomarker measurements and standards.* Proteomics Clin Appl, 2013. **7**(1-2): p. 17-29.
220. Uhlen, M., et al., *Towards a knowledge-based Human Protein Atlas.* Nat Biotechnol, 2010. **28**(12): p. 1248-50.
221. Uhlen, M., et al., *A human protein atlas for normal and cancer tissues based on antibody proteomics.* Mol Cell Proteomics, 2005. **4**(12): p. 1920-32.
222. Diamandis, E.P., *Automation of molecular diagnostics.* Clin Chem, 1996. **42**(1): p. 7-8.
223. Solier, C. and H. Langen, *Antibody-based proteomics and biomarker research - current status and limitations.* Proteomics, 2014. **14**(6): p. 774-83.
224. *The website of Singulex*. Available from: www.singulex.com/.
225. Todd, J., et al., *Ultrasensitive flow-based immunoassays using single-molecule counting.* Clin Chem, 2007. **53**(11): p. 1990-5.
226. Zetterberg, H., et al., *Hypoxia due to cardiac arrest induces a time-dependent increase in serum amyloid beta levels in humans.* PLoS One, 2011. **6**(12): p. e28263.
227. Wilson, D.H., et al., *Fifth-generation digital immunoassay for prostate-specific antigen by single molecule array technology.* Clin Chem, 2011. **57**(12): p. 1712-21.
228. Song, L., et al., *Single molecule measurements of tumor necrosis factor alpha and interleukin-6 in the plasma of patients with Crohn's disease.* J Immunol Methods, 2011. **372**(1-2): p. 177-86.
229. Rissin, D.M., et al., *Single-molecule enzyme-linked immunosorbent assay detects serum proteins at subfemtomolar concentrations.* Nat Biotechnol, 2010. **28**(6): p. 595-9.
230. Rissin, D.M., et al., *Simultaneous detection of single molecules and singulated ensembles of molecules enables immunoassays with broad dynamic range.* Anal Chem, 2011. **83**(6): p. 2279-85.
231. *The website of Cell Signaling Technology*. 2014; Available from: www.cellsignal.com/.
232. Lundberg, M., et al., *Homogeneous antibody-based proximity extension assays provide sensitive and specific detection of low-abundant proteins in human blood.* Nucleic Acids Res, 2011. **39**(15): p. e102.
233. Assarsson, E., et al., *Homogenous 96-Plex PEA Immunoassay Exhibiting High Sensitivity, Specificity, and Excellent Scalability.* PLoS One, 2014. **9**(4): p. e95192.
234. Wilson, B., L.A. Liotta, and E. Petricoin, 3rd, *Monitoring proteins and protein networks using reverse phase protein arrays.* Dis Markers, 2010. **28**(4): p. 225-32.
235. Voshol, H., et al., *Antibody-based proteomics: analysis of signaling networks using reverse protein arrays.* FEBS J, 2009. **276**(23): p. 6871-9.
236. Malinowsky, K., et al., *Deciphering signaling pathways in clinical tissues for personalized medicine using protein microarrays.* J Cell Physiol, 2010. **225**(2): p. 364-70.
237. Spurrier, B., et al., *Protein and lysate array technologies in cancer research.* Biotechnol Adv, 2008. **26**(4): p. 361-9.
238. Speer, R., et al., *Development of reverse phase protein microarrays for clinical applications and patient-tailored therapy.* Cancer Genomics Proteomics, 2007. **4**(3): p. 157-64.
239. Guo, H., et al., *An efficient procedure for protein extraction from formalin-fixed, paraffin-embedded tissues for reverse phase protein arrays.* Proteome Sci, 2012. **10**(1): p. 56.

240. Gerhard, D.S., et al., *The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC).* Genome Res, 2004. **14**(10B): p. 2121-7.
241. Boguski, M.S., C.M. Tolstoshev, and D.E. Bassett, Jr., *Gene discovery in dbEST.* Science, 1994. **265**(5181): p. 1993-4.
242. Harbers, M. and P. Carninci, *Tag-based approaches for transcriptome research and genome annotation.* Nat Methods, 2005. **2**(7): p. 495-502.
243. Shiraki, T., et al., *Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage.* Proc Natl Acad Sci U S A, 2003. **100**(26): p. 15776-81.
244. Nakamura, M. and P. Carninci, *[Cap analysis gene expression: CAGE].* Tanpakushitsu Kakusan Koso, 2004. **49**(17 Suppl): p. 2688-93.
245. Kodzius, R., et al., *CAGE: cap analysis of gene expression.* Nat Methods, 2006. **3**(3): p. 211-22.
246. Reinartz, J., et al., *Massively parallel signature sequencing (MPSS) as a tool for in-depth quantitative gene expression profiling in all organisms.* Brief Funct Genomic Proteomic, 2002. **1**(1): p. 95-104.
247. Brenner, S., et al., *Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays.* Nat Biotechnol, 2000. **18**(6): p. 630-4.
248. Morin, R., et al., *Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing.* Biotechniques, 2008. **45**(1): p. 81-94.
249. Maskos, U. and E.M. Southern, *Oligonucleotide hybridizations on glass supports: a novel linker for oligonucleotide synthesis and hybridization properties of oligonucleotides synthesised in situ.* Nucleic Acids Res, 1992. **20**(7): p. 1679-84.
250. Augenlicht, L.H. and D. Kobrin, *Cloning and screening of sequences expressed in a mouse colon tumor.* Cancer Res, 1982. **42**(3): p. 1088-93.
251. Schena, M., et al., *Quantitative monitoring of gene expression patterns with a complementary DNA microarray.* Science, 1995. **270**(5235): p. 467-70.
252. Lashkari, D.A., et al., *Yeast microarrays for genome wide parallel genetic and gene expression analysis.* Proc Natl Acad Sci U S A, 1997. **94**(24): p. 13057-62.
253. Quackenbush, J., *Microarray analysis and tumor classification.* N Engl J Med, 2006. **354**(23): p. 2463-72.
254. Dalma-Weiszhausz, D.D., et al., *The affymetrix GeneChip platform: an overview.* Methods Enzymol, 2006. **410**: p. 3-28.
255. Steemers, F.J. and K.L. Gunderson, *Illumina, Inc.* Pharmacogenomics, 2005. **6**(7): p. 777-82.
256. Perou, C.M., et al., *Molecular portraits of human breast tumours.* Nature, 2000. **406**(6797): p. 747-52.
257. Weigelt, B., et al., *Molecular portraits and 70-gene prognosis signature are preserved throughout the metastatic process of breast cancer.* Cancer Res, 2005. **65**(20): p. 9155-8.
258. Alizadeh, A.A., et al., *Towards a novel classification of human malignancies based on gene expression patterns.* J Pathol, 2001. **195**(1): p. 41-52.
259. Alizadeh, A.A., et al., *Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling.* Nature, 2000. **403**(6769): p. 503-11.
260. Pollack, J.R., *A perspective on DNA microarrays in pathology research and practice.* Am J Pathol, 2007. **171**(2): p. 375-85.
261. Michiels, S., S. Koscielny, and C. Hill, *Prediction of cancer outcome with microarrays: a multiple random validation strategy.* Lancet, 2005. **365**(9458): p. 488-92.
262. Ioannidis, J.P., *Microarrays and molecular research: noise discovery?* Lancet, 2005. **365**(9458): p. 454-5.
263. *International HapMap Project*. Available from: hapmap.ncbi.nlm.nih.gov.
264. Pavelic, K., et al., *Global Approach to Biomedicine: Functional Genomics and Proteomics.* eJIFCC - The electronic Journal of the International Federation of Clinical Chemistry and Laboratory Medicine, 2005. **16**(2).
265. Velculescu, V.E., et al., *Serial analysis of gene expression.* Science, 1995. **270**(5235): p. 484-7.

266. Saha, S., et al., *Using the transcriptome to annotate the genome.* Nat Biotechnol, 2002. **20**(5): p. 508-12.
267. Matsumura, H., et al., *SuperSAGE.* Cell Microbiol, 2005. **7**(1): p. 11-8.
268. Chu, Y. and D.R. Corey, *RNA sequencing: platform selection, experimental design, and data interpretation.* Nucleic Acid Ther, 2012. **22**(4): p. 271-4.
269. Maher, C.A., et al., *Transcriptome sequencing to detect gene fusions in cancer.* Nature, 2009. **458**(7234): p. 97-101.
270. Beane, J., et al., *Characterizing the impact of smoking and lung cancer on the airway transcriptome using RNA-Seq.* Cancer Prev Res (Phila), 2011. **4**(6): p. 803-17.
271. Ingolia, N.T., et al., *The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments.* Nat Protoc, 2012. **7**(8): p. 1534-50.
272. Trapnell, C., et al., *Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation.* Nat Biotechnol, 2010. **28**(5): p. 511-5.
273. Lu, B., Z. Zeng, and T. Shi, *Comparative study of de novo assembly and genome-guided assembly strategies for transcriptome reconstruction based on RNA-Seq.* Sci China Life Sci, 2013. **56**(2): p. 143-55.
274. Bradnam, K.R., et al., *Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species.* Gigascience, 2013. **2**(1): p. 10.
275. Trapnell, C., et al., *Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks.* Nat Protoc, 2012. **7**(3): p. 562-78.
276. Nagalakshmi, U., et al., *The transcriptional landscape of the yeast genome defined by RNA sequencing.* Science, 2008. **320**(5881): p. 1344-9.
277. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* Genome Biol, 2009. **10**(3): p. R25.
278. Trapnell, C., L. Pachter, and S.L. Salzberg, *TopHat: discovering splice junctions with RNA-Seq.* Bioinformatics, 2009. **25**(9): p. 1105-11.
279. Zhang, G., et al., *FANSe: an accurate algorithm for quantitative mapping of large scale sequencing reads.* Nucleic Acids Res, 2012. **40**(11): p. e83.
280. Schulz, M.H., et al., *Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels.* Bioinformatics, 2012. **28**(8): p. 1086-92.
281. Grabherr, M.G., et al., *Full-length transcriptome assembly from RNA-Seq data without a reference genome.* Nat Biotechnol, 2011. **29**(7): p. 644-52.
282. Wilhelm, B.T., et al., *Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution.* Nature, 2008. **453**(7199): p. 1239-43.
283. Tuch, B.B., et al., *Tumor transcriptome sequencing reveals allelic expression imbalances associated with copy number alterations.* PLoS One, 2010. **5**(2): p. e9317.
284. Murphy, J.R., et al., *Genetic assembly and selective toxicity of diphtheria-toxin-related polypeptide hormone fusion proteins.* Biochem Soc Symp, 1987. **53**: p. 9-23.
285. Mortazavi, A., et al., *Mapping and quantifying mammalian transcriptomes by RNA-Seq.* Nat Methods, 2008. **5**(7): p. 621-8.
286. Cloonan, N., et al., *Stem cell transcriptome profiling via massive-scale mRNA sequencing.* Nat Methods, 2008. **5**(7): p. 613-9.
287. *ENCODE Experiment Matrix*. Encyclopedia of DNA Elements Available from: http://genome.ucsc.edu/ENCODE/.
288. *The Cancer Genome Atlas (TCGA)*. [cited 2014 06-30]; Available from: http://cancergenome.nih.gov/.
289. Wang, Z., M. Gerstein, and M. Snyder, *RNA-Seq: a revolutionary tool for transcriptomics.* Nat Rev Genet, 2009. **10**(1): p. 57-63.
290. *A Brief Guide to Genomic*. Available from: www.genome.gov/18016863.
291. Srivastava, S., *Move over proteomics, here comes glycomics.* J Proteome Res, 2008. **7**(5): p. 1799.

292. Beckonert, O., et al., *Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts.* Nat Protoc, 2007. **2**(11): p. 2692-703.
293. Wenk, M.R., *The emerging field of lipidomics.* Nat Rev Drug Discov, 2005. **4**(7): p. 594-610.
294. Schulz-Knappe, P., M. Schrader, and H.D. Zucht, *The peptidomics concept.* Comb Chem High Throughput Screen, 2005. **8**(8): p. 697-704.
295. Heng, T.S., M.W. Painter, and C. Immunological Genome Project, *The Immunological Genome Project: networks of gene expression in immune cells.* Nat Immunol, 2008. **9**(10): p. 1091-4.
296. Hathout, Y., *Approaches to the study of the cell secretome.* Expert Rev Proteomics, 2007. **4**(2): p. 239-48.
297. Phizicky, E., et al., *Protein analysis on a proteomic scale.* Nature, 2003. **422**(6928): p. 208-15.
298. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics.* Nature, 2003. **422**(6928): p. 198-207.
299. Omenn, G.S., *Strategies for plasma proteomic profiling of cancers.* Proteomics, 2006. **6**(20): p. 5662-73.
300. Kuramitsu, Y. and K. Nakamura, *Proteomic analysis of cancer tissues: shedding light on carcinogenesis and possible biomarkers.* Proteomics, 2006. **6**(20): p. 5650-61.
301. Cancer Genome Atlas Research, N., *Comprehensive genomic characterization defines human glioblastoma genes and core pathways.* Nature, 2008. **455**(7216): p. 1061-8.
302. Lechner, A.J., *The scaling of maximal oxygen consumption and pulmonary dimensions in small mammals.* Respir Physiol, 1978. **34**(1): p. 29-44.
303. von Bubnoff, A., *Next-generation sequencing: the race is on.* Cell, 2008. **132**(5): p. 721-3.
304. Margulies, M., et al., *Genome sequencing in microfabricated high-density picolitre reactors.* Nature, 2005. **437**(7057): p. 376-80.
305. Dupont, J.M., et al., *De novo quantitative bisulfite sequencing using the pyrosequencing technology.* Anal Biochem, 2004. **333**(1): p. 119-27.
306. Parsons, D.W., et al., *An integrated genomic analysis of human glioblastoma multiforme.* Science, 2008. **321**(5897): p. 1807-12.
307. Jones, S., et al., *Core signaling pathways in human pancreatic cancers revealed by global genomic analyses.* Science, 2008. **321**(5897): p. 1801-6.
308. Nachman, M.W., *Single nucleotide polymorphisms and recombination rate in humans.* Trends Genet, 2001. **17**(9): p. 481-5.
309. Zhou, X., et al., *Concurrent analysis of loss of heterozygosity (LOH) and copy number abnormality (CNA) for oral premalignancy progression using the Affymetrix 10K SNP mapping array.* Hum Genet, 2004. **115**(4): p. 327-30.
310. Bignell, G.R., et al., *High-resolution analysis of DNA copy number using oligonucleotide microarrays.* Genome Res, 2004. **14**(2): p. 287-95.
311. Mao, X., B.D. Young, and Y.J. Lu, *The application of single nucleotide polymorphism microarrays in cancer research.* Curr Genomics, 2007. **8**(4): p. 219-28.
312. Liu, E.T., *Functional genomics of cancer.* Curr Opin Genet Dev, 2008. **18**(3): p. 251-6.
313. Gargiulo, G. and S. Minucci, *Epigenomic profiling of cancer cells.* Int J Biochem Cell Biol, 2009. **41**(1): p. 127-35.
314. Mensaert, K., et al., *Next-generation technologies and data analytical approaches for epigenomics.* Environ Mol Mutagen, 2014. **55**(3): p. 155-70.
315. Goodacre, R., *Metabolomics of a superorganism.* J Nutr, 2007. **137**(1 Suppl): p. 259S-266S.
316. Zhang, X., et al., *Mass spectrometry-based "omics" technologies in cancer diagnostics.* Mass Spectrom Rev, 2007. **26**(3): p. 403-31.
317. Shulaev, V., *Metabolomics technology and bioinformatics.* Brief Bioinform, 2006. **7**(2): p. 128-39.
318. Lindon, J.C., E. Holmes, and J.K. Nicholson, *Metabonomics and its role in drug development and disease diagnosis.* Expert Rev Mol Diagn, 2004. **4**(2): p. 189-99.
319. Han, X., *Neurolipidomics: challenges and developments.* Front Biosci, 2007. **12**: p. 2601-15.
320. Zhang, X., *Omics technologies in cancer biomarker discovery*. Molecular biology intelligence unit. 2011, Austin, Tex.: Landes Bioscience. 140 p.

321. Chisolm, G.M. and D. Steinberg, *The oxidative modification hypothesis of atherogenesis: an overview.* Free Radic Biol Med, 2000. **28**(12): p. 1815-26.
322. Ceschi, M., et al., *Epidemiology and pathophysiology of obesity as cause of cancer.* Swiss Med Wkly, 2007. **137**(3-4): p. 50-6.
323. Fahy, E., et al., *A comprehensive classification system for lipids.* J Lipid Res, 2005. **46**(5): p. 839-61.
324. Folch, J., M. Lees, and G.H. Sloane Stanley, *A simple method for the isolation and purification of total lipides from animal tissues.* J Biol Chem, 1957. **226**(1): p. 497-509.
325. Bligh, E.G. and W.J. Dyer, *A rapid method of total lipid extraction and purification.* Can J Biochem Physiol, 1959. **37**(8): p. 911-7.
326. Powell, W.S., *Rapid extraction of arachidonic acid metabolites from biological samples using octadecylsilyl silica.* Methods Enzymol, 1982. **86**: p. 467-77.
327. Powell, W.S., *Rapid extraction of oxygenated metabolites of arachidonic acid from biological samples using octadecylsilyl silica.* Prostaglandins, 1980. **20**(5): p. 947-57.
328. Lesnefsky, E.J., et al., *Separation and quantitation of phospholipids and lysophospholipids by high-performance liquid chromatography.* Anal Biochem, 2000. **285**(2): p. 246-54.
329. Hawkins, D.J., et al., *Resolution of enantiomers of hydroxyeicosatetraenoate derivatives by chiral phase high-pressure liquid chromatography.* Anal Biochem, 1988. **173**(2): p. 456-62.
330. Weintraub, S.T., C.S. Lear, and R.N. Pinckard, *Analysis of platelet-activating factor by GC-MS after direct derivatization with pentafluorobenzoyl chloride and heptafluorobutyric anhydride.* J Lipid Res, 1990. **31**(4): p. 719-25.
331. Wang, Y., et al., *Derivatization of phospholipids.* J Chromatogr B Analyt Technol Biomed Life Sci, 2003. **793**(1): p. 3-14.
332. Balazy, M., P. Braquet, and N.G. Bazan, *Determination of platelet-activating factor and alkyl-ether phospholipids by gas chromatography-mass spectrometry via direct derivatization.* Anal Biochem, 1991. **196**(1): p. 1-10.
333. Schiller, J., et al., *Matrix-assisted laser desorption and ionization time-of-flight (MALDI-TOF) mass spectrometry in lipid and phospholipid research.* Prog Lipid Res, 2004. **43**(5): p. 449-88.
334. Fujiwaki, T., et al., *Application of delayed extraction matrix-assisted laser desorption ionization time-of-flight mass spectrometry for analysis of sphingolipids in cultured skin fibroblasts from sphingolipidosis patients.* Brain Dev, 2002. **24**(3): p. 170-3.
335. Han, X. and R.W. Gross, *Shotgun lipidomics: multidimensional MS analysis of cellular lipidomes.* Expert Rev Proteomics, 2005. **2**(2): p. 253-64.
336. Apweiler, R., H. Hermjakob, and N. Sharon, *On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database.* Biochim Biophys Acta, 1999. **1473**(1): p. 4-8.
337. Zhao, Y.Y., et al., *Functional roles of N-glycans in cell signaling and cell adhesion in cancer.* Cancer Sci, 2008. **99**(7): p. 1304-10.
338. Hakomori, S., *Glycosylation defining cancer malignancy: new wine in an old bottle.* Proc Natl Acad Sci U S A, 2002. **99**(16): p. 10231-3.
339. Dube, D.H. and C.R. Bertozzi, *Glycans in cancer and inflammation--potential for therapeutics and diagnostics.* Nat Rev Drug Discov, 2005. **4**(6): p. 477-88.
340. Rudd, P.M., et al., *Glycosylation and the immune system.* Science, 2001. **291**(5512): p. 2370-6.
341. Ohtsubo, K. and J.D. Marth, *Glycosylation in cellular mechanisms of health and disease.* Cell, 2006. **126**(5): p. 855-67.
342. Lowe, J.B., *Glycosylation, immunity, and autoimmunity.* Cell, 2001. **104**(6): p. 809-12.
343. Kornfeld, S., *Diseases of abnormal protein glycosylation: an emerging area.* J Clin Invest, 1998. **101**(7): p. 1293-5.
344. Haltiwanger, R.S. and J.B. Lowe, *Role of glycosylation in development.* Annu Rev Biochem, 2004. **73**: p. 491-537.
345. Arnold, J.N., et al., *The impact of glycosylation on the biological function and structure of human immunoglobulins.* Annu Rev Immunol, 2007. **25**: p. 21-50.

346. Seeberger, P.H. and D.B. Werz, *Synthesis and medical applications of oligosaccharides.* Nature, 2007. **446**(7139): p. 1046-51.
347. Raman, R., et al., *Glycomics: an integrated systems approach to structure-function relationships of glycans.* Nat Methods, 2005. **2**(11): p. 817-24.
348. Drickamer, K. and M.E. Taylor, *Glycan arrays for functional glycomics.* Genome Biol, 2002. **3**(12): p. REVIEWS1034.
349. Meany, D.L., et al., *Glycoproteomics for prostate cancer detection: changes in serum PSA glycosylation patterns.* J Proteome Res, 2009. **8**(2): p. 613-9.
350. Lim, K.T., et al., *Clinical application of functional glycoproteomics - dissection of glycotopes carried by soluble CD44 variants in sera of patients with cancers.* Proteomics, 2008. **8**(16): p. 3263-73.
351. Block, T.M., et al., *Use of targeted glycoproteomics to identify serum glycoproteins that correlate with liver cancer in woodchucks and humans.* Proc Natl Acad Sci U S A, 2005. **102**(3): p. 779-84.
352. Fang, X. and W.W. Zhang, *Affinity separation and enrichment methods in proteomic analysis.* J Proteomics, 2008. **71**(3): p. 284-303.
353. Kuno, A., et al., *Focused differential glycan analysis with the platform antibody-assisted lectin profiling for glycan-related biomarker verification.* Mol Cell Proteomics, 2009. **8**(1): p. 99-108.
354. Zhang, H., *Glycoproteomics using chemical immobilization.* Curr Protoc Protein Sci, 2007. **Chapter 24**: p. Unit 24 3.
355. Lewandrowski, U. and A. Sickmann, *N-glycosylation site analysis of human platelet proteins by hydrazide affinity capturing and LC-MS/MS.* Methods Mol Biol, 2009. **534**: p. 225-38.
356. Zhang, Q., et al., *Enrichment and analysis of nonenzymatically glycated peptides: boronate affinity chromatography coupled with electron-transfer dissociation mass spectrometry.* J Proteome Res, 2007. **6**(6): p. 2323-30.
357. Gurcel, C., et al., *Identification of new O-GlcNAc modified proteins using a click-chemistry-based tagging.* Anal Bioanal Chem, 2008. **390**(8): p. 2089-97.
358. Benk, A.S. and C. Roesli, *Label-free quantification using MALDI mass spectrometry: considerations and perspectives.* Anal Bioanal Chem, 2012. **404**(4): p. 1039-56.
359. Ong, S.E. and M. Mann, *Mass spectrometry-based proteomics turns quantitative.* Nat Chem Biol, 2005. **1**(5): p. 252-62.
360. Simpson, K.L., A.D. Whetton, and C. Dive, *Quantitative mass spectrometry-based techniques for clinical use: biomarker identification and quantification.* J Chromatogr B Analyt Technol Biomed Life Sci, 2009. **877**(13): p. 1240-9.
361. Gramolini, A.O., S.M. Peterman, and T. Kislinger, *Mass spectrometry-based proteomics: a useful tool for biomarker discovery?* Clin Pharmacol Ther, 2008. **83**(5): p. 758-60.
362. Bantscheff, M., et al., *Quantitative mass spectrometry in proteomics: a critical review.* Anal Bioanal Chem, 2007. **389**(4): p. 1017-31.
363. Ohtsuki, S., et al., *Quantitative targeted absolute proteomics-based ADME research as a new path to drug discovery and development: methodology, advantages, strategy, and prospects.* J Pharm Sci, 2011. **100**(9): p. 3547-59.
364. Eyrich, B., A. Sickmann, and R.P. Zahedi, *Catch me if you can: mass spectrometry-based phosphoproteomics and quantification strategies.* Proteomics, 2011. **11**(4): p. 554-70.
365. Eberl, H.C., M. Mann, and M. Vermeulen, *Quantitative proteomics for epigenetics.* Chembiochem, 2011. **12**(2): p. 224-34.
366. Brownridge, P., et al., *Global absolute quantification of a proteome: Challenges in the deployment of a QconCAT strategy.* Proteomics, 2011. **11**(15): p. 2957-70.
367. Matsuo, T. and Y. Seyama, *Introduction to modern biological mass spectrometry.* J Mass Spectrom, 2000. **35**(2): p. 114-30.
368. Fenn, J.B., et al., *Electrospray ionization for mass spectrometry of large biomolecules.* Science, 1989. **246**(4926): p. 64-71.
369. Karas, M. and F. Hillenkamp, *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons.* Anal Chem, 1988. **60**(20): p. 2299-301.

370. Tanaka, K., et al., *Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry.* Rapid Commun. Mass Spectrom., 1988. **2**(8): p. 151–153.
371. Yang, Y., et al., *A comparison of nLC-ESI-MS/MS and nLC-MALDI-MS/MS for GeLC-based protein identification and iTRAQ-based shotgun quantitative proteomics.* J Biomol Tech, 2007. **18**(4): p. 226-37.
372. Bodnar, W.M., et al., *Exploiting the complementary nature of LC/MALDI/MS/MS and LC/ESI/MS/MS for increased proteome coverage.* Journal of the American Society for Mass Spectrometry, 2003. **14**(9): p. 971–979.
373. Johnson, R.E., et al., *Methods and mechanisms for producing ions from large molecules*, in *NATO ASI series*, K.G. Standing and W. Ens, Editors. 1991, New York u.a.: Plenum Press. p. 89-99.
374. Bahr, U., M. Karas, and F. Hillenkamp, *Analysis of biopolymers by Matrix-assisted Laser Desorption/Ionization (MALDI) mass spectrometry.* Fresenius' Journal of Analytical Chemistry, 1994. **348**: p. 783-791.
375. Wilm, M., *Principles of electrospray ionization.* Mol Cell Proteomics, 2011. **10**(7): p. M111 009407.
376. Iribarne, J.V. and T.B. A., *On the evaporation of small ions from charged droplets.* The Journal of Chemical Physics, 1976. **64**(6): p. 2287-2294.
377. Wilm, M.S. and M. Mann, *Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last?* International Journal of Mass Spectrometry and Ion Processes, 1994. **136**: p. 167-180.
378. Vestal, M.L., *Modern MALDI time-of-flight mass spectrometry.* J Mass Spectrom, 2009. **44**(3): p. 303-17.
379. Demartini, D.R., *A Short Overview of the Components in Mass Spectrometry Instrumentation for Proteomics Analyses*. Tandem Mass Spectrometry - Molecular Characterization, ed. A.V. Coelho and C. de Matos Ferraz Franco. 2013.
380. March, R.E., *Quadrupole ion traps.* Mass Spectrom Rev, 2009. **28**(6): p. 961-89.
381. Domon, B. and R. Aebersold, *Mass spectrometry and protein analysis.* Science, 2006. **312**(5771): p. 212-7.
382. Yates, J.R., 3rd, *Mass spectral analysis in proteomics.* Annu Rev Biophys Biomol Struct, 2004. **33**: p. 297-316.
383. Schaeffer-Reiss, C., *A brief summary of the different types of mass spectrometers used in proteomics.* Methods Mol Biol, 2008. **484**: p. 3-16.
384. Croxatto, A., G. Prod'hom, and G. Greub, *Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology.* FEMS Microbiol Rev, 2012. **36**(2): p. 380-407.
385. Lee, J., et al., *High resolution time-of-flight mass analysis of the entire range of intact singly-charged proteins.* Anal Chem, 2011. **83**(24): p. 9406-12.
386. Lee, J. and P.T. Reilly, *Limitation of time-of-flight resolution in the ultra high mass range.* Anal Chem, 2011. **83**(15): p. 5831-3.
387. Glish, G.L. and R.W. Vachet, *The basics of mass spectrometry in the twenty-first century.* Nat Rev Drug Discov, 2003. **2**(2): p. 140-50.
388. March, R.E., *Quadrupole ion trap mass spectrometry: a view at the turn of the century.* International Journal of Mass Spectrometry, 2000. **200**(1–3): p. 285-312.
389. Paul, W., *Electromagnetic traps for charged and neutral particles.* Reviews of Modern Physics, 1990. **62**(3): p. 531-540.
390. Ouyang, Z., et al., *Quadrupole ion traps and trap arrays: geometry, material, scale, performance.* Eur J Mass Spectrom (Chichester, Eng), 2007. **13**(1): p. 13-8.
391. Hu, Q., et al., *The Orbitrap: a new mass spectrometer.* J Mass Spectrom, 2005. **40**(4): p. 430-43.
392. Makarov, A., *Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis.* Anal Chem, 2000. **72**(6): p. 1156-62.
393. Kelstrup, C.D., et al., *Optimized fast and sensitive acquisition methods for shotgun proteomics on a quadrupole orbitrap mass spectrometer.* J Proteome Res, 2012. **11**(6): p. 3487-97.

394. Havilio, M., Y. Haddad, and Z. Smilansky, *Intensity-based statistical scorer for tandem mass spectrometry.* Anal Chem, 2003. **75**(3): p. 435-44.
395. Craig, R. and R.C. Beavis, *TANDEM: matching proteins with tandem mass spectra.* Bioinformatics, 2004. **20**(9): p. 1466-7.
396. Strittmatter, E.F., N. Rodriguez, and R.D. Smith, *High mass measurement accuracy determination for proteomics using multivariate regression fitting: application to electrospray ionization time-of-flight mass spectrometry.* Anal Chem, 2003. **75**(3): p. 460-8.
397. Wysocki, V.H., et al., *Mass spectrometry of peptides and proteins.* Methods, 2005. **35**(3): p. 211-22.
398. Zubarev, R.A., *Electron-capture dissociation tandem mass spectrometry.* Curr Opin Biotechnol, 2004. **15**(1): p. 12-6.
399. Kim, M.S. and A. Pandey, *Electron transfer dissociation mass spectrometry in proteomics.* Proteomics, 2012. **12**(4-5): p. 530-42.
400. Roepstorff, P. and J. Fohlman, *Proposal for a common nomenclature for sequence ions in mass spectra of peptides.* Biomed Mass Spectrom, 1984. **11**(11): p. 601.
401. Wikimedia Commons [cited 2014 July 1]; Available from: http://commons.wikimedia.org.
402. Shapiro, A.L., E. Vinuela, and J.V. Maizel, Jr., *Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels.* Biochem Biophys Res Commun, 1967. **28**(5): p. 815-20.
403. O'Farrell, P.H., *High resolution two-dimensional electrophoresis of proteins.* J Biol Chem, 1975. **250**(10): p. 4007-21.
404. Unlu, M., M.E. Morgan, and J.S. Minden, *Difference gel electrophoresis: a single gel method for detecting changes in protein extracts.* Electrophoresis, 1997. **18**(11): p. 2071-7.
405. Minden, J., *Comparative proteomics and difference gel electrophoresis.* Biotechniques, 2007. **43**(6): p. 739, 741, 743 passim.
406. Rabilloud, T. and C. Lelong, *Two-dimensional gel electrophoresis in proteomics: a tutorial.* J Proteomics, 2011. **74**(10): p. 1829-41.
407. Miller, I., J. Crawford, and E. Gianazza, *Protein stains for proteomic applications: which, when, why?* Proteomics, 2006. **6**(20): p. 5385-408.
408. Rabilloud, T., et al., *Two-dimensional gel electrophoresis in proteomics: Past, present and future.* J Proteomics, 2010. **73**(11): p. 2064-77.
409. Gygi, S.P., et al., *Quantitative analysis of complex protein mixtures using isotope-coded affinity tags.* Nat Biotechnol, 1999. **17**(10): p. 994-9.
410. Wu, C.C. and M.J. MacCoss, *Shotgun proteomics: tools for the analysis of complex biological systems.* Curr Opin Mol Ther, 2002. **4**(3): p. 242-50.
411. Zhang, Y., et al., *Protein analysis by shotgun/bottom-up proteomics.* Chem Rev, 2013. **113**(4): p. 2343-94.
412. Mueller, D.R., et al., *LC-MALDI MS and MS/MS--an efficient tool in proteome analysis.* Subcell Biochem, 2007. **43**: p. 355-80.
413. Garcia, M.C., *The effect of the mobile phase additives on sensitivity in the analysis of peptides and proteins by high-performance liquid chromatography-electrospray mass spectrometry.* J Chromatogr B Analyt Technol Biomed Life Sci, 2005. **825**(2): p. 111-23.
414. Zhao, L., et al., *A proteogenomic analysis of Shigella flexneri using 2D LC-MALDI TOF/TOF.* BMC Genomics, 2011. **12**: p. 528.
415. Bachmann, S., et al., *Peptide mapping using capillary electrophoresis offline coupled to matrix-assisted laser desorption ionization time of flight mass spectrometry.* Electrophoresis, 2011. **32**(20): p. 2830-9.
416. Maccarrone, G., C.W. Turck, and D. Martins-de-Souza, *Shotgun mass spectrometry workflow combining IEF and LC-MALDI-TOF/TOF.* Protein J, 2010. **29**(2): p. 99-102.
417. MacNair, J.E., K.C. Lewis, and J.W. Jorgenson, *Ultrahigh-pressure reversed-phase liquid chromatography in packed capillary columns.* Anal Chem, 1997. **69**(6): p. 983-9.
418. Winter, D., R. Pipkorn, and W.D. Lehmann, *Separation of peptide isomers and conformers by ultra performance liquid chromatography.* J Sep Sci, 2009. **32**(8): p. 1111-9.

419. Nagaraj, N., et al., *System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top Orbitrap.* Mol Cell Proteomics, 2012. **11**(3): p. M111 013722.
420. Motoyama, A., et al., *Automated ultra-high-pressure multidimensional protein identification technology (UHP-MudPIT) for improved peptide identification of proteomic samples.* Anal Chem, 2006. **78**(14): p. 5109-18.
421. Contrepois, K., et al., *Ultra-high performance liquid chromatography-mass spectrometry for the fast profiling of histone post-translational modifications.* J Proteome Res, 2010. **9**(10): p. 5501-9.
422. Chen, C.J., et al., *Tunnel frit: a nonmetallic in-capillary frit for nanoflow ultra high-performance liquid chromatography-mass spectrometryapplications.* Anal Chem, 2012. **84**(1): p. 297-303.
423. Sandra, K., et al., *Highly efficient peptide separations in proteomics Part 1. Unidimensional high performance liquid chromatography.* J Chromatogr B Analyt Technol Biomed Life Sci, 2008. **866**(1-2): p. 48-63.
424. Kall, L., et al., *Assigning significance to peptides identified by tandem mass spectrometry using decoy databases.* J Proteome Res, 2008. **7**(1): p. 29-34.
425. Moore, R.E., M.K. Young, and T.D. Lee, *Qscore: an algorithm for evaluating SEQUEST database search results.* J Am Soc Mass Spectrom, 2002. **13**(4): p. 378-86.
426. Li, Y.F. and P. Radivojac, *Computational approaches to protein inference in shotgun proteomics.* BMC Bioinformatics, 2012. **13 Suppl 16**: p. S4.
427. Olivares, J.A., et al., *On-line mass spectrometric detection for capillary zone electrophoresis.* Analytical Chemistry, 1987. **59**(8): p. 1230-1232.
428. Theodorescu, D., et al., *Discovery and validation of new protein biomarkers for urothelial cancer: a prospective analysis.* Lancet Oncol, 2006. **7**(3): p. 230-40.
429. Schiffer, E., et al., *Prediction of muscle-invasive bladder cancer using urinary proteomics.* Clin Cancer Res, 2009. **15**(15): p. 4935-43.
430. Frantzi, M., et al., *Discovery and validation of urinary biomarkers for detection of renal cell carcinoma.* J Proteomics, 2014. **98**: p. 44-58.
431. Zurbig, P., et al., *Biomarker discovery by CE-MS enables sequence analysis via MS/MS with platform-independent separation.* Electrophoresis, 2006. **27**(11): p. 2111-25.
432. Latosinska, A., et al., *Clinical applications of capillary electrophoresis coupled to mass spectrometry in biomarker discovery: Focus on Bladder Cancer.* Proteomics Clin Appl, 2013.
433. Mullen, W., et al., *Performance of different separation methods interfaced in the same MS-reflection TOF detector: a comparison of performance between CE versus HPLC for biomarker analysis.* Electrophoresis, 2012. **33**(4): p. 567-74.
434. Mischak, H., A. Vlahou, and J.P. Ioannidis, *Technical aspects and inter-laboratory variability in native peptide profiling: the CE-MS experience.* Clin Biochem, 2013. **46**(6): p. 432-43.
435. McDonnell, L.A. and R.M. Heeren, *Imaging mass spectrometry.* Mass Spectrom Rev, 2007. **26**(4): p. 606-43.
436. Caldwell, R.L. and R.M. Caprioli, *Tissue profiling by mass spectrometry: a review of methodology and applications.* Mol Cell Proteomics, 2005. **4**(4): p. 394-401.
437. Robichaud, G., et al., *MSiReader: an open-source interface to view and analyze high resolving power MS imaging files on Matlab platform.* J Am Soc Mass Spectrom, 2013. **24**(5): p. 718-21.
438. Sinha, T.K., et al., *Integrating spatially resolved three-dimensional MALDI IMS with in vivo magnetic resonance imaging.* Nat Methods, 2008. **5**(1): p. 57-9.
439. Boggio, K.J., et al., *Recent advances in single-cell MALDI mass spectrometry imaging and potential clinical impact.* Expert Rev Proteomics, 2011. **8**(5): p. 591-604.
440. Rauser, S., et al., *Classification of HER2 receptor status in breast cancer tissues by MALDI imaging mass spectrometry.* J Proteome Res, 2010. **9**(4): p. 1854-63.
441. Groseclose, M.R., et al., *High-throughput proteomic analysis of formalin-fixed paraffin-embedded tissue microarrays using MALDI imaging mass spectrometry.* Proteomics, 2008. **8**(18): p. 3715-24.

442. Oppenheimer, S.R., et al., *Molecular analysis of tumor margins by MALDI mass spectrometry in renal carcinoma.* J Proteome Res, 2010. **9**(5): p. 2182-90.
443. Schwamborn, K. and R.M. Caprioli, *MALDI imaging mass spectrometry--painting molecular pictures.* Mol Oncol, 2010. **4**(6): p. 529-38.
444. *Method of the Year 2012.* Nat Meth, 2013. **10**(1): p. 1-1.
445. Kondrat, R.W., G.A. McClusky, and R.G. Cooks, *Multiple reaction monitoring in mass spectrometry/mass spectrometry for direct analysis of complex mixtures.* Analytical Chemistry, 1978. **50**(14): p. 2017-2021.
446. Yost, R.A. and C.G. Enke, *Triple quadrupole mass spectrometry for direct mixture analysis and structure elucidation.* Anal Chem, 1979. **51**(12): p. 1251-64.
447. Holman, S.W., P.F. Sims, and C.E. Eyers, *The use of selected reaction monitoring in quantitative proteomics.* Bioanalysis, 2012. **4**(14): p. 1763-86.
448. Anderson, L. and C.L. Hunter, *Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins.* Mol Cell Proteomics, 2006. **5**(4): p. 573-88.
449. Boyd, R.K., C. Basic, and R.A. Bethem, in *Trace Quantitative Analysis by Mass Spectrometry*. 2008, John Wiley & Sons, Ltd.
450. Schmidt, A., P. Picotti, and R. Aebersold, *Proteomanalyse und Systembiologie.* BIOspektrum, 2008. **14**(1): p. 44-46.
451. Tang, K., J.S. Page, and R.D. Smith, *Charge competition and the linear dynamic range of detection in electrospray ionization mass spectrometry.* J Am Soc Mass Spectrom, 2004. **15**(10): p. 1416-23.
452. Elliott, M.H., et al., *Current trends in quantitative proteomics.* J Mass Spectrom, 2009. **44**(12): p. 1637-60.
453. Krause, E., H. Wenschuh, and P.R. Jungblut, *The dominance of arginine-containing peptides in MALDI-derived tryptic mass fingerprints of proteins.* Anal Chem, 1999. **71**(19): p. 4160-5.
454. Gay, S., et al., *Peptide mass fingerprinting peak intensity prediction: extracting knowledge from spectra.* Proteomics, 2002. **2**(10): p. 1374-91.
455. Eyers, C.E., et al., *CONSeQuence: prediction of reference peptides for absolute quantitative proteomics using consensus machine learning approaches.* Mol Cell Proteomics, 2011. **10**(11): p. M110 003384.
456. Ludwig, C., et al., *Estimation of absolute protein quantities of unlabeled samples by selected reaction monitoring mass spectrometry.* Mol Cell Proteomics, 2012. **11**(3): p. M111 013987.
457. Anderson, N.L. and N.G. Anderson, *The human plasma proteome: history, character, and diagnostic prospects.* Mol Cell Proteomics, 2002. **1**(11): p. 845-67.
458. Shiio, Y. and R. Aebersold, *Quantitative proteome analysis using isotope-coded affinity tags and mass spectrometry.* Nat Protoc, 2006. **1**(1): p. 139-45.
459. Ong, S.E., et al., *Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics.* Mol Cell Proteomics, 2002. **1**(5): p. 376-86.
460. Neher, S.B., et al., *Proteomic profiling of ClpXP substrates after DNA damage reveals extensive instability within SOS regulon.* Mol Cell, 2006. **22**(2): p. 193-204.
461. Krijgsveld, J., et al., *Metabolic labeling of C. elegans and D. melanogaster for quantitative proteomics.* Nat Biotechnol, 2003. **21**(8): p. 927-31.
462. Kruger, M., et al., *SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function.* Cell, 2008. **134**(2): p. 353-64.
463. Ross, P.L., et al., *Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents.* Mol Cell Proteomics, 2004. **3**(12): p. 1154-69.
464. Schmidt, A., J. Kellermann, and F. Lottspeich, *A novel strategy for quantitative proteomics using isotope-coded protein labels.* Proteomics, 2005. **5**(1): p. 4-15.
465. Thompson, A., et al., *Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS.* Anal Chem, 2003. **75**(8): p. 1895-904.
466. Gerber, S.A., et al., *Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS.* Proc Natl Acad Sci U S A, 2003. **100**(12): p. 6940-5.

467. Zeng, D. and S. Li, *Improved CILAT reagents for quantitative proteomics.* Bioorg Med Chem Lett, 2009. **19**(7): p. 2059-61.
468. Li, S. and D. Zeng, *CILAT--a new reagent for quantitative proteomics.* Chem Commun (Camb), 2007(21): p. 2181-3.
469. Xie, F., et al., *Liquid chromatography-mass spectrometry-based quantitative proteomics.* J Biol Chem, 2011. **286**(29): p. 25443-9.
470. Bondarenko, P.V., D. Chelius, and T.A. Shaler, *Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography-tandem mass spectrometry.* Anal Chem, 2002. **74**(18): p. 4741-9.
471. Liu, H., R.G. Sadygov, and J.R. Yates, 3rd, *A model for random sampling and estimation of relative protein abundance in shotgun proteomics.* Anal Chem, 2004. **76**(14): p. 4193-201.
472. Neilson, K.A., et al., *Less label, more free: approaches in label-free quantitative mass spectrometry.* Proteomics, 2011. **11**(4): p. 535-53.
473. Voyksner, R.D. and H. Lee, *Investigating the use of an octupole ion guide for ion storage and high-pass mass filtering to improve the quantitative performance of electrospray ion trap mass spectrometry.* Rapid Commun Mass Spectrom, 1999. **13**(14): p. 1427-37.
474. Listgarten, J. and A. Emili, *Statistical and computational methods for comparative proteomic profiling using liquid chromatography-tandem mass spectrometry.* Mol Cell Proteomics, 2005. **4**(4): p. 419-34.
475. Zhu, W., J.W. Smith, and C.M. Huang, *Mass spectrometry-based label-free quantitative proteomics.* J Biomed Biotechnol, 2010. **2010**: p. 840518.
476. Wong, J.W., M.J. Sullivan, and G. Cagney, *Computational methods for the comparative quantification of proteins in label-free LCn-MS experiments.* Brief Bioinform, 2008. **9**(2): p. 156-65.
477. Panchaud, A., et al., *Experimental and computational approaches to quantitative proteomics: status quo and outlook.* J Proteomics, 2008. **71**(1): p. 19-33.
478. Mueller, L.N., et al., *An assessment of software solutions for the analysis of mass spectrometry based quantitative proteomics data.* J Proteome Res, 2008. **7**(1): p. 51-61.
479. Yang, C., Z. He, and W. Yu, *Comparison of public peak detection algorithms for MALDI mass spectrometry data analysis.* BMC Bioinformatics, 2009. **10**: p. 4.
480. Bauer, C., R. Cramer, and J. Schuchhardt, *Evaluation of peak-picking algorithms for protein mass spectrometry.* Methods Mol Biol, 2011. **696**: p. 341-52.
481. Cappadona, S., et al., *Deconvolution of overlapping isotopic clusters improves quantification of stable isotope-labeled peptides.* J Proteomics, 2011. **74**(10): p. 2204-9.
482. Kultima, K., et al., *Development and evaluation of normalization methods for label-free relative quantification of endogenous peptides.* Mol Cell Proteomics, 2009. **8**(10): p. 2285-95.
483. Bolstad, B.M., et al., *A comparison of normalization methods for high density oligonucleotide array data based on variance and bias.* Bioinformatics, 2003. **19**(2): p. 185-93.
484. Fugmann, T., D. Neri, and C. Roesli, *DeepQuanTR: MALDI-MS-based label-free quantification of proteins in complex biological samples.* Proteomics, 2010. **10**(14): p. 2631-43.
485. McMaster, M.C., *HPLC, a practical user's guide*. WILEY INTERSCIENCE. 2007: John Wiley & Sons, Inc.
486. Wang, P., et al., *A statistical method for chromatographic alignment of LC-MS data.* Biostatistics, 2007. **8**(2): p. 357-67.
487. Vahamaa, H., et al., *PolyAlign: A Versatile LC-MS Data Alignment Tool for Landmark-Selected and -Automated Use.* Int J Proteomics, 2011. **2011**: p. 450290.
488. Tang, Z., et al., *A new method for alignment of LC-MALDI-TOF data.* Proteome Sci, 2011. **9 Suppl 1**: p. S10.
489. Christin, C., et al., *Optimized time alignment algorithm for LC-MS data: correlation optimized warping using component detection algorithm-selected mass chromatograms.* Anal Chem, 2008. **80**(18): p. 7012-21.
490. Lange, E., et al., *A geometric approach for the alignment of liquid chromatography-mass spectrometry data.* Bioinformatics, 2007. **23**(13): p. i273-81.

491. Podwojski, K., et al., *Peek a peak: a glance at statistics for quantitative label-free proteomics.* Expert Rev Proteomics, 2010. **7**(2): p. 249-61.
492. Mueller, L.N., et al., *SuperHirn - a novel tool for high resolution LC-MS-based peptide/protein profiling.* Proteomics, 2007. **7**(19): p. 3470-80.
493. Kohlbacher, O., et al., *TOPP--the OpenMS proteomics pipeline.* Bioinformatics, 2007. **23**(2): p. e191-7.
494. Carrillo, B., et al., *Methods for combining peptide intensities to estimate relative protein abundance.* Bioinformatics, 2010. **26**(1): p. 98-103.
495. Zhang, B., et al., *Detecting differential and correlated protein expression in label-free shotgun proteomics.* J Proteome Res, 2006. **5**(11): p. 2909-18.
496. Washburn, M.P., D. Wolters, and J.R. Yates, 3rd, *Large-scale analysis of the yeast proteome by multidimensional protein identification technology.* Nat Biotechnol, 2001. **19**(3): p. 242-7.
497. Shinoda, K., M. Tomita, and Y. Ishihama, *emPAI Calc--for the estimation of protein abundance from large-scale identification data by liquid chromatography-tandem mass spectrometry.* Bioinformatics, 2010. **26**(4): p. 576-7.
498. Rappsilber, J., et al., *Large-scale proteomic analysis of the human spliceosome.* Genome Res, 2002. **12**(8): p. 1231-45.
499. Ishihama, Y., et al., *Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein.* Mol Cell Proteomics, 2005. **4**(9): p. 1265-72.
500. Lu, P., et al., *Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation.* Nat Biotechnol, 2007. **25**(1): p. 117-24.
501. Braisted, J.C., et al., *The APEX Quantitative Proteomics Tool: generating protein quantitation estimates from LC-MS/MS proteomics results.* BMC Bioinformatics, 2008. **9**: p. 529.
502. van der Veldt, A.A. and A.A. Lammertsma, *In vivo imaging as a pharmacodynamic marker.* Clin Cancer Res, 2014. **20**(10): p. 2569-77.
503. Bosslet, K., et al., *Elucidation of the mechanism enabling tumor selective prodrug monotherapy.* Cancer Res, 1998. **58**(6): p. 1195-201.
504. Jain, R.K., *Transport of molecules in the tumor interstitium: a review.* Cancer Res, 1987. **47**(12): p. 3039-51.
505. Szakacs, G., et al., *Targeting multidrug resistance in cancer.* Nat Rev Drug Discov, 2006. **5**(3): p. 219-34.
506. Maeshima, Y., et al., *Tumstatin, an endothelial cell-specific inhibitor of protein synthesis.* Science, 2002. **295**(5552): p. 140-3.
507. Hamano, Y., et al., *Physiological levels of tumstatin, a fragment of collagen IV alpha3 chain, are generated by MMP-9 proteolysis and suppress angiogenesis via alphaV beta3 integrin.* Cancer Cell, 2003. **3**(6): p. 589-601.
508. Brody, E.N. and L. Gold, *Aptamers as therapeutic and diagnostic agents.* J Biotechnol, 2000. **74**(1): p. 5-13.
509. Arkin, M.R. and J.A. Wells, *Small-molecule inhibitors of protein-protein interactions: progressing towards the dream.* Nat Rev Drug Discov, 2004. **3**(4): p. 301-17.
510. Winter, G., et al., *Making antibodies by phage display technology.* Annu Rev Immunol, 1994. **12**: p. 433-55.
511. Kohler, G. and C. Milstein, *Continuous cultures of fused cells secreting antibody of predefined specificity.* Nature, 1975. **256**(5517): p. 495-7.
512. Verma, S., et al., *Trastuzumab emtansine for HER2-positive advanced breast cancer.* N Engl J Med, 2012. **367**(19): p. 1783-91.
513. van Osdol, W., K. Fujimori, and J.N. Weinstein, *An analysis of monoclonal antibody distribution in microscopic tumor nodules: consequences of a "binding site barrier".* Cancer Res, 1991. **51**(18): p. 4776-84.
514. Thorpe, P.E., *Vascular targeting agents as cancer therapeutics.* Clin Cancer Res, 2004. **10**(2): p. 415-27.
515. Neri, D. and R. Bicknell, *Tumour vascular targeting.* Nat Rev Cancer, 2005. **5**(6): p. 436-46.

516. Speers, A.E. and C.C. Wu, *Proteomics of integral membrane proteins--theory and application.* Chem Rev, 2007. **107**(8): p. 3687-714.
517. Macher, B.A. and T.Y. Yen, *Proteins at membrane surfaces-a review of approaches.* Mol Biosyst, 2007. **3**(10): p. 705-13.
518. DePierre, J.W. and M.L. Karnovsky, *Plasma membranes of mammalian cells: a review of methods for their characterization and isolation.* J Cell Biol, 1973. **56**(2): p. 275-303.
519. Maddalo, G., et al., *Systematic analysis of native membrane protein complexes in Escherichia coli.* J Proteome Res, 2011. **10**(4): p. 1848-59.
520. Schindler, J. and H.G. Nothwang, *Aqueous polymer two-phase systems: effective tools for plasma membrane proteomics.* Proteomics, 2006. **6**(20): p. 5409-17.
521. Sun, L., et al., *Ionic liquid 1-butyl-3-methyl imidazolium tetrafluoroborate for shotgun membrane proteomics.* Anal Bioanal Chem, 2011. **399**(10): p. 3387-97.
522. Nagaraj, N., et al., *Detergent-based but gel-free method allows identification of several hundred membrane proteins in single LC-MS runs.* J Proteome Res, 2008. **7**(11): p. 5028-32.
523. Blonder, J., et al., *Enrichment of integral membrane proteins for proteomic analysis using liquid chromatography-tandem mass spectrometry.* J Proteome Res, 2002. **1**(4): p. 351-60.
524. Zhang, H., et al., *Differential recovery of membrane proteins after extraction by aqueous methanol and trifluoroethanol.* Proteomics, 2007. **7**(10): p. 1654-63.
525. De Palma, A., et al., *Extraction methods of red blood cell membrane proteins for Multidimensional Protein Identification Technology (MudPIT) analysis.* J Chromatogr A, 2010. **1217**(33): p. 5328-36.
526. Solis, N. and S.J. Cordwell, *Current methodologies for proteomics of bacterial surface-exposed and cell envelope proteins.* Proteomics, 2011. **11**(15): p. 3169-89.
527. Vuckovic, D., et al., *Membrane proteomics by high performance liquid chromatography-tandem mass spectrometry: Analytical approaches and challenges.* Proteomics, 2013. **13**(3-4): p. 404-23.
528. Wei, X. and L. Li, *Comparative glycoproteomics: approaches and applications.* Brief Funct Genomic Proteomic, 2009. **8**(2): p. 104-13.
529. Bond, M.R. and J.J. Kohler, *Chemical methods for glycoprotein discovery.* Curr Opin Chem Biol, 2007. **11**(1): p. 52-8.
530. Lotan, R. and G.L. Nicolson, *Purification of cell membrane glycoproteins by lectin affinity chromatography.* Biochim Biophys Acta, 1979. **559**(4): p. 329-76.
531. Kaji, H., et al., *Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins.* Nat Biotechnol, 2003. **21**(6): p. 667-72.
532. Vosseller, K., et al., *Quantitative analysis of both protein expression and serine / threonine post-translational modifications through stable isotope labeling with dithiothreitol.* Proteomics, 2005. **5**(2): p. 388-98.
533. Sparbier, K., T. Wenzel, and M. Kostrzewa, *Exploring the binding profiles of ConA, boronic acid and WGA by MALDI-TOF/TOF MS and magnetic particles.* J Chromatogr B Analyt Technol Biomed Life Sci, 2006. **840**(1): p. 29-36.
534. Khidekel, N., et al., *Exploring the O-GlcNAc proteome: direct identification of O-GlcNAc-modified proteins from the brain.* Proc Natl Acad Sci U S A, 2004. **101**(36): p. 13132-7.
535. Wollscheid, B., et al., *Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins.* Nat Biotechnol, 2009. **27**(4): p. 378-86.
536. Zhang, H., et al., *Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry.* Nat Biotechnol, 2003. **21**(6): p. 660-6.
537. Chaney, L.K. and B.S. Jacobson, *Coating cells with colloidal silica for high yield isolation of plasma membrane sheets and identification of transmembrane proteins.* J Biol Chem, 1983. **258**(16): p. 10062-72.
538. Oh, P., et al., *Subtractive proteomic mapping of the endothelial surface in lung and solid tumours for tissue-specific therapy.* Nature, 2004. **429**(6992): p. 629-35.

539. Durr, E., et al., *Direct proteomic mapping of the lung microvascular endothelial cell surface in vivo and in cell culture.* Nat Biotechnol, 2004. **22**(8): p. 985-92.
540. Arjunan, S., et al., *Limitations of the colloidal silica method in mapping the endothelial plasma membrane proteome of the mouse heart.* Cell Biochem Biophys, 2009. **53**(3): p. 135-43.
541. Elia, G., *Biotinylation reagents for the study of cell surface proteins.* Proteomics, 2008. **8**(19): p. 4012-24.
542. Green, N.M., *Avidin.* Adv Protein Chem, 1975. **29**: p. 85-133.
543. Rugg-Gunn, P.J., et al., *Cell-surface proteomics identifies lineage-specific markers of embryo-derived stem cells.* Dev Cell, 2012. **22**(4): p. 887-901.
544. Garcia, J., et al., *Comprehensive profiling of the cell surface proteome of Sy5Y neuroblastoma cells yields a subset of proteins associated with tumor differentiation.* J Proteome Res, 2009. **8**(8): p. 3791-6.
545. Schliemann, C., et al., *In vivo biotinylation of the vasculature in B-cell lymphoma identifies BST-2 as a target for antibody-based therapy.* Blood, 2010. **115**(3): p. 736-44.
546. Rybak, J.N., et al., *The extra-domain A of fibronectin is a vascular marker of solid tumors and metastases.* Cancer Res, 2007. **67**(22): p. 10948-57.
547. Castronovo, V., et al., *A chemical proteomics approach for the identification of accessible antigens expressed in human kidney cancer.* Mol Cell Proteomics, 2006. **5**(11): p. 2083-91.
548. Konerding, M.A., E. Fait, and A. Gaumann, *3D microvascular architecture of pre-cancerous lesions and invasive carcinomas of the colon.* Br J Cancer, 2001. **84**(10): p. 1354-62.
549. Weekes, M.P., et al., *Comparative analysis of techniques to purify plasma membrane proteins.* J Biomol Tech, 2010. **21**(3): p. 108-15.
550. Smith, G.P., *Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface.* Science, 1985. **228**(4705): p. 1315-7.
551. Pieczenik, G., *Method and means for sorting and identifying biological information*. 1999, US Patent 5866363.
552. Rajotte, D., et al., *Molecular heterogeneity of the vascular endothelium revealed by in vivo phage display.* J Clin Invest, 1998. **102**(2): p. 430-7.
553. Pasqualini, R. and E. Ruoslahti, *Tissue targeting with phage peptide libraries.* Mol Psychiatry, 1996. **1**(6): p. 423.
554. Laakkonen, P., et al., *A tumor-homing peptide with a targeting specificity related to lymphatic vessels.* Nat Med, 2002. **8**(7): p. 751-5.
555. Wu, H.C. and D.K. Chang, *Peptide-mediated liposomal drug delivery system targeting tumor blood vessels in anticancer therapy.* J Oncol, 2010. **2010**: p. 723798.
556. Ingermann, A.R., et al., *Identification of a novel cell death receptor mediating IGFBP-3-induced anti-tumor effects in breast and prostate cancer.* J Biol Chem, 2010. **285**(39): p. 30233-46.
557. Bhattacharyya, N., et al., *Nonsecreted insulin-like growth factor binding protein-3 (IGFBP-3) can induce apoptosis in human prostate cancer cells by IGF-independent mechanisms without being concentrated in the nucleus.* J Biol Chem, 2006. **281**(34): p. 24588-601.
558. Butt, A.J., et al., *IGF-binding protein-3-induced growth inhibition and apoptosis do not require cell surface binding and nuclear translocation in human breast cancer cells.* Endocrinology, 2002. **143**(7): p. 2693-9.
559. Jones, J.I. and D.R. Clemmons, *Insulin-like growth factors and their binding proteins: biological actions.* Endocr Rev, 1995. **16**(1): p. 3-34.
560. Firth, S.M. and R.C. Baxter, *Cellular actions of the insulin-like growth factor binding proteins.* Endocr Rev, 2002. **23**(6): p. 824-54.
561. Oh, Y., et al., *Insulin-like growth factor (IGF)-independent action of IGF-binding protein-3 in Hs578T human breast cancer cells. Cell surface binding and growth inhibition.* J Biol Chem, 1993. **268**(20): p. 14964-71.
562. Kettunen, E., et al., *Differentially expressed genes in nonsmall cell lung cancer: expression profiling of cancer-related genes in squamous cell lung cancer.* Cancer Genet Cytogenet, 2004. **149**(2): p. 98-106.

563. Chuang, S.T., et al., *Over expression of insulin-like growth factor binding protein 3 in clear cell renal cell carcinoma.* J Urol, 2008. **179**(2): p. 445-9.
564. Torng, P.L., et al., *Insulin-like growth factor binding protein-3 (IGFBP-3) acts as an invasion-metastasis suppressor in ovarian endometrioid carcinoma.* Oncogene, 2008. **27**(15): p. 2137-47.
565. Liu, B., et al., *Insulin-like growth factor-binding protein-3 inhibition of prostate cancer growth involves suppression of angiogenesis.* Oncogene, 2007. **26**(12): p. 1811-9.
566. Cheung, C.W., et al., *The roles of IGF-I and IGFBP-3 in the regulation of proximal tubule, and renal cell carcinoma cell proliferation.* Kidney Int, 2004. **65**(4): p. 1272-9.
567. Davis, C.G., *The many faces of epidermal growth factor repeats.* New Biol, 1990. **2**(5): p. 410-9.
568. Hyytiainen, M., C. Penttinen, and J. Keski-Oja, *Latent TGF-beta binding proteins: extracellular matrix association and roles in TGF-beta activation.* Crit Rev Clin Lab Sci, 2004. **41**(3): p. 233-64.
569. Shipley, J.M., et al., *Developmental expression of latent transforming growth factor beta binding protein 2 and its requirement early in mouse development.* Mol Cell Biol, 2000. **20**(13): p. 4879-87.
570. Gibson, M.A., et al., *Bovine latent transforming growth factor beta 1-binding protein 2: molecular cloning, identification of tissue isoforms, and immunolocalization to elastin-associated microfibrils.* Mol Cell Biol, 1995. **15**(12): p. 6932-42.
571. Moren, A., et al., *Identification and characterization of LTBP-2, a novel latent transforming growth factor-beta-binding protein.* J Biol Chem, 1994. **269**(51): p. 32469-78.
572. Chan, S.H., et al., *The ECM protein LTBP-2 is a suppressor of esophageal squamous cell carcinoma tumor formation but higher tumor expression associates with poor patient outcome.* Int J Cancer, 2011. **129**(3): p. 565-73.
573. Flower, D.R., *The lipocalin protein family: structure and function.* Biochem J, 1996. **318 (Pt 1)**: p. 1-14.
574. Tong, Z., et al., *Neutrophil gelatinase-associated lipocalin: a novel suppressor of invasion and angiogenesis in pancreatic cancer.* Cancer Res, 2008. **68**(15): p. 6100-8.
575. Cowland, J.B. and N. Borregaard, *Molecular characterization and pattern of tissue expression of the gene for neutrophil gelatinase-associated lipocalin from humans.* Genomics, 1997. **45**(1): p. 17-23.
576. Yang, J., et al., *An iron delivery pathway mediated by a lipocalin.* Mol Cell, 2002. **10**(5): p. 1045-56.
577. Sheng, Z., S.Z. Wang, and M.R. Green, *Transcription and signalling pathways involved in BCR-ABL-mediated misregulation of 24p3 and 24p3R.* EMBO J, 2009. **28**(7): p. 866-76.
578. Candido, S., et al., *Roles of neutrophil gelatinase-associated lipocalin (NGAL) in human cancer.* Oncotarget, 2014. **5**(6): p. 1576-94.
579. Devarajan, P., *Review: neutrophil gelatinase-associated lipocalin: a troponin-like biomarker for human acute kidney injury.* Nephrology (Carlton), 2010. **15**(4): p. 419-28.
580. Nickolas, T.L., et al., *NGAL (Lcn2) monomer is associated with tubulointerstitial damage in chronic kidney disease.* Kidney Int, 2012. **82**(6): p. 718-22.
581. Barresi, V., et al., *Neutrophil gelatinase-associated lipocalin immunoexpression in renal tumors: correlation with histotype and histological grade.* Oncol Rep, 2010. **24**(2): p. 305-10.
582. A, D.I.C., *Evaluation of neutrophil gelatinase-associated lipocalin (NGAL), matrix metalloproteinase-9 (MMP-9) and their complex MMP-9/NGAL in sera and urine of patients with kidney tumors.* Oncol Lett, 2013. **5**(5): p. 1677-1681.
583. Halestrap, A.P., *The SLC16 gene family - structure, role and regulation in health and disease.* Mol Aspects Med, 2013. **34**(2-3): p. 337-49.
584. Otonkoski, T., et al., *Physical exercise-induced hypoglycemia caused by failed silencing of monocarboxylate transporter 1 in pancreatic beta cells.* Am J Hum Genet, 2007. **81**(3): p. 467-74.

585. Merezhinskaya, N., et al., *Presence and localization of three lactic acid transporters (MCT1, -2, and -4) in separated human granulocytes, lymphocytes, and monocytes.* J Histochem Cytochem, 2004. **52**(11): p. 1483-93.
586. Manoharan, C., et al., *The role of charged residues in the transmembrane helices of monocarboxylate transporter 1 and its ancillary protein basigin in determining plasma membrane expression and catalytic activity.* Mol Membr Biol, 2006. **23**(6): p. 486-98.
587. Poole, R.C. and A.P. Halestrap, *Interaction of the erythrocyte lactate transporter (monocarboxylate transporter 1) with an integral 70-kDa membrane glycoprotein of the immunoglobulin superfamily.* J Biol Chem, 1997. **272**(23): p. 14624-8.
588. Kirk, P., et al., *CD147 is tightly associated with lactate transporters MCT1 and MCT4 and facilitates their cell surface expression.* EMBO J, 2000. **19**(15): p. 3896-904.
589. Merezhinskaya, N., et al., *Mutations in MCT1 cDNA in patients with symptomatic deficiency in lactate transport.* Muscle Nerve, 2000. **23**(1): p. 90-7.
590. Sonveaux, P., et al., *Targeting lactate-fueled respiration selectively kills hypoxic tumor cells in mice.* J Clin Invest, 2008. **118**(12): p. 3930-42.
591. Bristow, J., et al., *Tenascin-X: a novel extracellular matrix protein encoded by the human XB gene overlapping P450c21B.* J Cell Biol, 1993. **122**(1): p. 265-78.
592. Chiquet-Ehrismann, R. and M. Chiquet, *Tenascins: regulation and putative functions during pathological stress.* J Pathol, 2003. **200**(4): p. 488-99.
593. Jones, P.T., et al., *Replacing the complementarity-determining regions in a human antibody with those from a mouse.* Nature, 1986. **321**(6069): p. 522-5.
594. Egging, D., et al., *Interactions of human tenascin-X domains with dermal extracellular matrix molecules.* Arch Dermatol Res, 2007. **298**(8): p. 389-96.
595. Bristow, J., et al., *Tenascin-X, collagen, elastin, and the Ehlers-Danlos syndrome.* Am J Med Genet C Semin Med Genet, 2005. **139C**(1): p. 24-30.
596. Yuan, Y., et al., *Tenascin-X is a novel diagnostic marker of malignant mesothelioma.* Am J Surg Pathol, 2009. **33**(11): p. 1673-82.
597. Matsumoto, K., et al., *Tumour invasion and metastasis are promoted in mice deficient in tenascin-X.* Genes Cells, 2001. **6**(12): p. 1101-11.
598. Kawamoto, T., et al., *Structural and phylogenetic analyses of RGD-CAP/beta ig-h3, a fasciclin-like adhesion protein expressed in chick chondrocytes.* Biochim Biophys Acta, 1998. **1395**(3): p. 288-92.
599. Escribano, J., et al., *cDNA from human ocular ciliary epithelium homologous to beta ig-h3 is preferentially expressed as an extracellular protein in the corneal epithelium.* J Cell Physiol, 1994. **160**(3): p. 511-21.
600. Skonier, J., et al., *cDNA cloning and sequence analysis of beta ig-h3, a novel gene induced in a human adenocarcinoma cell line after treatment with transforming growth factor-beta.* DNA Cell Biol, 1992. **11**(7): p. 511-22.
601. Billings, P.C., et al., *The transforming growth factor-beta-inducible matrix protein (beta)ig-h3 interacts with fibronectin.* J Biol Chem, 2002. **277**(31): p. 28003-9.
602. Kim, J.E., et al., *Identification of motifs in the fasciclin domains of the transforming growth factor-beta-induced matrix protein betaig-h3 that interact with the alphavbeta5 integrin.* J Biol Chem, 2002. **277**(48): p. 46159-65.
603. Park, S.W., et al., *Beta ig-h3 promotes renal proximal tubular epithelial cell adhesion, migration and proliferation through the interaction with alpha3beta1 integrin.* Exp Mol Med, 2004. **36**(3): p. 211-9.
604. Kim, J.E., et al., *RGD peptides released from beta ig-h3, a TGF-beta-induced cell-adhesive molecule, mediate apoptosis.* Oncogene, 2003. **22**(13): p. 2045-53.
605. Kannabiran, C. and G.K. Klintworth, *TGFBI gene mutations in corneal dystrophies.* Hum Mutat, 2006. **27**(7): p. 615-25.
606. Sasaki, H., et al., *Beta IGH3, a TGF-beta inducible gene, is overexpressed in lung cancer.* Jpn J Clin Oncol, 2002. **32**(3): p. 85-9.

607. Schneider, D., et al., *Induction and expression of betaig-h3 in pancreatic cancer cells.* Biochim Biophys Acta, 2002. **1588**(1): p. 1-6.
608. Becker, J., et al., *Keratoepithelin suppresses the progression of experimental human neuroblastomas.* Cancer Res, 2006. **66**(10): p. 5314-21.
609. Ivanov, S.V., et al., *Two novel VHL targets, TGFBI (BIGH3) and its transactivator KLF10, are up-regulated in renal clear cell carcinoma and other tumors.* Biochem Biophys Res Commun, 2008. **370**(4): p. 536-40.
610. Lauden, L., et al., *TGF-beta-induced (TGFBI) protein in melanoma: a signature of high metastatic potential.* J Invest Dermatol, 2014. **134**(6): p. 1675-85.
611. Ma, C., et al., *Extracellular matrix protein betaig-h3/TGFBI promotes metastasis of colon cancer by enhancing cell extravasation.* Genes Dev, 2008. **22**(3): p. 308-21.
612. Zhu, M., et al., *lncRNA H19/MiR-675 axis represses prostate cancer metastasis by targeting TGFBI.* FEBS J, 2014.
613. Shilov, I.V., et al., *The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra.* Mol Cell Proteomics, 2007. **6**(9): p. 1638-55.
614. Zhang, B., S. Kirov, and J. Snoddy, *WebGestalt: an integrated system for exploring gene sets in various biological contexts.* Nucleic Acids Res, 2005. **33**(Web Server issue): p. W741-8.
615. Brack, S.S., et al., *Tumor-targeting properties of novel antibodies specific to the large isoform of tenascin-C.* Clin Cancer Res, 2006. **12**(10): p. 3200-8.
616. Marlind, J., et al., *Antibody-mediated delivery of interleukin-2 to the stroma of breast cancer strongly enhances the potency of chemotherapy.* Clin Cancer Res, 2008. **14**(20): p. 6515-24.
617. McQueen, H.A., et al., *Stability of critical genetic lesions in human colorectal carcinoma xenografts.* Br J Cancer, 1991. **63**(1): p. 94-6.
618. Fidler, I.J., S. Naito, and S. Pathak, *Orthotopic implantation is essential for the selection, growth and metastasis of human renal cell cancer in nude mice [corrected].* Cancer Metastasis Rev, 1990. **9**(2): p. 149-65.
619. Mestas, J. and C.C. Hughes, *Of mice and not men: differences between mouse and human immunology.* J Immunol, 2004. **172**(5): p. 2731-8.
620. Siolas, D. and G.J. Hannon, *Patient-derived tumor xenografts: transforming clinical samples into mouse models.* Cancer Res, 2013. **73**(17): p. 5315-9.
621. DeNardo, D.G., et al., *CD4(+) T cells regulate pulmonary metastasis of mammary carcinomas by enhancing protumor properties of macrophages.* Cancer Cell, 2009. **16**(2): p. 91-102.
622. Zubarev, R.A., *The challenge of the proteome dynamic range and its implications for in-depth proteomics.* Proteomics, 2013. **13**(5): p. 723-6.
623. Almen, M.S., et al., *Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin.* BMC Biol, 2009. **7**: p. 50.
624. Wu, C.C. and J.R. Yates, 3rd, *The application of mass spectrometry to membrane proteomics.* Nat Biotechnol, 2003. **21**(3): p. 262-7.
625. Baez-Saldana, A., et al., *Effects of biotin on pyruvate carboxylase, acetyl-CoA carboxylase, propionyl-CoA carboxylase, and markers for glucose and lipid homeostasis in type 2 diabetic patients and nondiabetic subjects.* Am J Clin Nutr, 2004. **79**(2): p. 238-43.
626. Mellman, I., G. Coukos, and G. Dranoff, *Cancer immunotherapy comes of age.* Nature, 2011. **480**(7378): p. 480-9.
627. Strassberger, V., et al., *A novel reactive ester derivative of biotin with reduced membrane permeability for in vivo biotinylation experiments.* Proteomics, 2010. **10**(19): p. 3544-8.
628. Steen, H. and M. Mann, *The ABC's (and XYZ's) of peptide sequencing.* Nat Rev Mol Cell Biol, 2004. **5**(9): p. 699-711.
629. Bier, M., F.F. Nord, and L. Terminiello, *On the mechanism of enzyme action. LXI. The self digestion of trypsin, calcium-trypsin and acetyltrypsin.* Arch Biochem Biophys, 1956. **65**(1): p. 120-31.

630. Constantopoulos, T.L., G.S. Jackson, and C.G. Enke, *Effects of salt concentration on analyte response using electrospray ionization mass spectrometry.* J Am Soc Mass Spectrom, 1999. **10**(7): p. 625-34.
631. Tabb, D.L., et al., *Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry.* J Proteome Res, 2010. **9**(2): p. 761-76.
632. Nesvizhskii, A.I. and R. Aebersold, *Interpretation of shotgun proteomic data: the protein inference problem.* Mol Cell Proteomics, 2005. **4**(10): p. 1419-40.
633. Zhang, Y., et al., *Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins.* Anal Chem, 2010. **82**(6): p. 2272-81.
634. Usaite, R., et al., *Characterization of global yeast quantitative proteome data generated from the wild-type and glucose repression saccharomyces cerevisiae strains: the comparison of two quantitative methods.* J Proteome Res, 2008. **7**(1): p. 266-75.
635. Jin, S., et al., *The effects of shared peptides on protein quantitation in label-free proteomics by LC/MS/MS.* J Proteome Res, 2008. **7**(1): p. 164-9.
636. Liu, W.L., et al., *Structural changes in TAF4b-TFIID correlate with promoter selectivity.* Mol Cell, 2008. **29**(1): p. 81-91.
637. Zybailov, B., et al., *Sorting signals, N-terminal modifications and abundance of the chloroplast proteome.* PLoS One, 2008. **3**(4): p. e1994.
638. Dvorak, H.F., *Tumors: wounds that do not heal. Similarities between tumor stroma generation and wound healing.* N Engl J Med, 1986. **315**(26): p. 1650-9.
639. Kawauchi, T., *Cell Adhesion and Its Endocytic Regulation in Cell Migration during Neural Development and Cancer Metastasis.* Int J Mol Sci, 2012. **13**(4): p. 4564-90.
640. McSherry, E.A., et al., *Breast cancer cell migration is regulated through junctional adhesion molecule-A-mediated activation of Rap1 GTPase.* Breast Cancer Res, 2011. **13**(2): p. R31.
641. Camaj, P., et al., *EFEMP1 binds the EGF receptor and activates MAPK and Akt pathways in pancreatic carcinoma cells.* Biol Chem, 2009. **390**(12): p. 1293-302.
642. Etchason, J.A., et al., *Niacin-induced hepatitis: a potential side effect with low-dose time-release niacin.* Mayo Clin Proc, 1991. **66**(1): p. 23-8.
643. Saper, C.B., *A guide to the perplexed on the specificity of antibodies.* J Histochem Cytochem, 2009. **57**(1): p. 1-5.
644. de Sousa Abreu, R., et al., *Global signatures of protein and mRNA expression levels.* Mol Biosyst, 2009. **5**(12): p. 1512-26.
645. Gerlinger, M., et al., *Intratumor heterogeneity and branched evolution revealed by multiregion sequencing.* N Engl J Med, 2012. **366**(10): p. 883-92.
646. Chiquet-Ehrismann, R. and R.P. Tucker, *Tenascins and the importance of adhesion modulation.* Cold Spring Harb Perspect Biol, 2011. **3**(5).
647. Berndt, A., et al., *A comparative analysis of oncofetal fibronectin and tenascin-C incorporation in tumour vessels using human recombinant SIP format antibodies.* Histochem Cell Biol, 2010. **133**(4): p. 467-75.
648. Chen, R., et al., *Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry.* J Proteome Res, 2009. **8**(2): p. 651-61.
649. Bolignano, D., et al., *Neutrophil gelatinase-associated lipocalin (NGAL) as a marker of kidney damage.* Am J Kidney Dis, 2008. **52**(3): p. 595-605.
650. Miyazono, K., et al., *Latent high molecular weight complex of transforming growth factor beta 1. Purification from human platelets and structural characterization.* J Biol Chem, 1988. **263**(13): p. 6407-15.
651. Massague, J., *The transforming growth factor-beta family.* Annu Rev Cell Biol, 1990. **6**: p. 597-641.
652. Vehvilainen, P., M. Hyytiainen, and J. Keski-Oja, *Latent transforming growth factor-beta-binding protein 2 is an adhesion protein for melanoma cells.* J Biol Chem, 2003. **278**(27): p. 24705-13.
653. Massoner, P., et al., *Novel mechanism of IGF-binding protein-3 action on prostate cancer cells: inhibition of proliferation, adhesion, and motility.* Endocr Relat Cancer, 2009. **16**(3): p. 795-808.

654. Miyake, H., M. Pollak, and M.E. Gleave, *Castration-induced up-regulation of insulin-like growth factor binding protein-5 potentiates insulin-like growth factor-I activity and accelerates progression to androgen independence in prostate cancer models.* Cancer Res, 2000. **60**(11): p. 3058-64.
655. Collisson, E.A., et al., *Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy.* Nat Med, 2011. **17**(4): p. 500-3.
656. Brannon, A.R., et al., *Molecular Stratification of Clear Cell Renal Cell Carcinoma by Consensus Clustering Reveals Distinct Subtypes and Survival Patterns.* Genes Cancer, 2010. **1**(2): p. 152-163.
657. *Method of the Year 2012.* Nat Methods, 2013. **10**(1): p. 1.