Department of Medicine and Surgery

PhD program: Molecular and Translational medicine
Cycle: XXXI

Identification of molecular alterations associated with the progression of clear cell Renal Cell Carcinoma by mass spectrometric approaches

Surname: Stella **Nome:** Martina
Registration number: 81055

Tutor: Prof. Fulvio Magni **Co-tutor:** Dr. Giorgio Bovo
Coordinator: Prof. Andrea Biondi

ACADEMIC YEAR 2017-2018

“A winner is a dreamer who never give up”

Nelson Mandela

To all who support my dreams

Table of contents

CHAPTER 1

INTRODUCTION

1.1 Mass spectrometry: an introduction

Recently, translational medicine has become one of the most important areas of biomedicine aiming at reducing distance between basic research and clinical practice. To this purpose, analytical chemistry methods are often applied to pathological or clinical specimens and in this thesis, among a series of possible approaches, mass spectrometry has been used to investigate the progression of clear cell renal cell carcinoma. These two pages represent a very brief introduction to the basic of mass spectrometry, useful for a better understanding of the following chapters.

Mass spectrometry (MS) is an analytical technique born in 1886 when Eugene Goldstein discovered that positive ions are able to move in an electric field with a low pressure applied and described the possibility to separate charged species using a magnetic field [1]. Mass spectrometry underwent a consistent evolution over years but the basic mechanisms of separating molecules according to their mass-to-charge ratio (m/z) after their ionization remain the same.

The samples are at first ionized, with a process that takes places in an ion source and then separated according to their m/z ratio.

Among the large number of different ionisation techniques available, Electrospray Ionisation (ESI) and Matrix-Assisted Laser Desorption/Ionisation (MALDI) are the two mentioned in this dissertation. **ElectroSpray Ionisation** (ESI) is a soft ionisation method that takes advantage of an electric field to produce multi-charged ions [2–4]. ESI can be performed in infusion mode wherein samples are directly introduced in the mass spectrometer or in combination with liquid chromatography (LC). **Matrix-Assisted Laser Desorption/Ionisation** (MALDI) is a pulsed ionization process. During the sample preparation for a MALDI experiment, a high concentrated solution of matrix is either mixed, spotted or sprayed on the specimen. Matrices are small aromatic compounds (usually with phenolic and carboxylic groups) with a strong molar absorption at the laser wavelength. As the solvent evaporates, the matrix co-crystallises

and extract the analyte molecules. Matrix-analyte crystals are irradiated with a pulsed UV laser (e.g. Nitrogen, Nd:YAG), a process that leads to the desorption and ionization of the analytes [5].

In the analyser, magnetic or electric fields are generated under vacuum and these fields influence the ions spatial trajectories, velocity and direction allowing ions separation. The mass analysers separate the ions according to their mass-to-charge ratio and can be utilised alone or in tandem to take the advantage of the better properties of each. Detailed description of all the analyser is present in the review by Himmelsbach [6]. Finally, the detector records the number of ions at each m/z value.

The output of a mass spectrometric analysis is a mass spectrum that is used to visualise the molecular content of the sample. The X-axis represents the mass-to-charge ratio and the Y-axis the signal intensity of each single ion.

References

[1] E. Goldstein, Vorläufige Mittheilungen über elektrische Entladungen in verdünnten Gasen (Preliminary communications on electric discharges in rarefied gases), Monatsberichte Der Königlich Preuss. Akad. Der Wissenschaften Zu Berlin (Monthly Reports R. Prussian Acad. Sci. Berlin). (1876) 279–295.

[2] S.A. Carr, M.E. Hemling, M.F. Bean, G.D. Roberts, Integration of mass spectrometry in analytical biotechnology, Anal. Chem. 63 (1991) 2802–2824. doi:10.1021/ac00024a003.

[3] B.N. Pramanik, P.L. Bartner, U.A. Mirza, Y.-H. Liu, A.K. Ganguly, Electrospray ionization mass spectrometry for the study of non-covalent complexes: an emerging technology, J. Mass Spectrom. 33 (1998) 911–920. doi:10.1002/(SICI)1096-9888(1998100)33:10<911::AID-JMS737>3.0.CO;2-5.

[4] Y. Wang, M. Schubert, A. Ingendoh, J. Franzen, Analysis of non-covalent protein complexes up to 290 kDa using electrospray ionization and ion trap mass spectrometry, Rapid Commun Mass Spectrom. 14 (2000) 12–17. doi:10.1002/(SICI)1097-0231(20000115)14.

[5] K. Dreisewerd, Recent methodological advances in MALDI mass spectrometry, Anal. Bioanal. Chem. 406 (2014) 2261–2278. doi:10.1007/s00216-014-7646-6.

[6] M. Himmelsbach, 10years of MS instrumental developments – Impact on LC–MS/MS in clinical chemistry, J. Chromatogr. B. 883–884 (2012) 3–17.

1.2 The proteomic landscape of renal tumors

Clizia Chinello[a], Vincenzo L'imperio[a], Martina Stella[a], Andrew James Smith[a], Giorgio Bovo[b], Angelica Grasso[c], Marco Grasso[d], Francesca Raimondo[a], Marina Pitto[a], Fabio Pagni[a] and Fulvio Magni[a]

[a] Department of Medicine and Surgery, University Milan Bicocca, Monza, Italy
[b] Pathology unit, San Gerardo Hospital, Monza, Italy
[c] Department of Specialistic Surgical Sciences, Urology unit, Ospedale Maggiore Policlinico Foundation, Milano, Italy
[d] Department of Urology, San Gerardo Hospital, Monza, Italy

MS participate in drafting the article and in designing figures and tables.

Published: Expert Review of Proteomics, 2016, vol. 13, no. 12, 1103–1120

ABSTRACT

Introduction: Renal cell carcinoma (RCC) is the most fatal of the common urologic cancers, with approximately 35% of patients dying within 5 years following diagnosis. Therefore, there is a need for non-invasive markers that are capable of detecting and determining the severity of small renal masses at an early stage in order to tailor treatment and follow-up. Proteomic studies have proved to be very useful in the study of tumors.

Areas covered: In this review, we will detail the current knowledge obtained by the different proteomic approaches, focusing on MS-based strategies, used to investigate RCC biology in order to identify diagnostic, prognostic and predictive biomarkers on tissue, cultured cells and biological fluids.

Expert commentary: Currently, no reliable biomarkers or targets for RCC have been translated into the clinical setting. Moreover, despite the efforts of proteomics and other -omics disciplines, only a small number of them have been observed as shared targets between the different analytical platforms and biological specimens. The difficulty to define a specific molecular pattern for RCC and its subtypes highlights a peculiar profile and a heterogeneity that must be taken into account in future studies.

KEYWORDS

Renal cell carcinoma; tissue; biological fluids; biomarkers; proteomics; mass spectrometry imaging; peptidome; proteome; kidney; cancer

1. Introduction

1.1 Clinical aspects

Renal cell carcinoma (RCC) represents 2–3% of all cancers [1] and is the most fatal of the common urologic cancers, with approximately 35% of patients dying from the disease within 5 years following diagnosis [2]. The incidence of RCC worldwide has been appreciably increasing by approximately 2% per year over the last few decades [3,4]. Most renal masses, particularly small tumors, are now discovered incidentally during imaging that was prompted by nonspecific or unrelated symptoms [5] and their detection has increased to more than 50%. Symptoms associated with RCC can be the result of local tumor growth, hemorrhage, paraneoplastic syndromes, or metastatic diseases. The choices of treatment vary depending on staging and grading of the tumor, and the general performance status of the patients. Alternatives are surgical treatment (radical or partial nephrectomies), ablative therapies (cryoablation, radiofrequency ablation), chemotherapy, immunotherapy, and targeted therapies.

1.2 RCC histopathology

RCC arises from the tubular structure of the kidney and comprises many histological variants, characterized by different genetic and morphological features, each with distinct clinical behavior and responses to therapy [1]. The clear cell variant (ccRCC) represents the most prevalent of these variants, accounting for approximately 70–80% of all kidney cancers, while the papillary (10%), chromophobe (5%), medullary, and collecting duct (<1%) subtypes are rarer [6,7]. The etiology of this cancer can be associated with environmental risk factors such as smoking, obesity, and hypertension. However, the majority of the cases are related to genetic abnormalities, the most frequent of which involves the von Hippel–Lindau (VHL) tumor suppressor pathway

in ccRCC [8]. Recently, the identification of molecular mechanisms involved in RCC tumorigenesis led to the development of new therapeutic strategies based on molecular-targeted agents, e.g. those directed against the VHL pathway, such as the vascular endothelial growth factor, the platelet-derived growth factor, and the mammalian target of rapamycin (mTOR) [4,9].

From a diagnostic point of view, the differentiation among the various histotypes of renal cancer can be challenging on a pure morphological level, requiring the employment of further ancillary tools. For this reason, many immunohistochemical markers have been employed in the routine assessment of renal masses, creating immunoprofiles able to differentiate them [10,11]. The main targets are represented by structural molecules (such as cytokeratins, vimentin, e-cadherin, and claudins 7/8), transcription factors (tumor protein 63, CD117, Mast/stem cells growth factor receptor), metabolic pathway related proteins (parvalbumin, a-methylacyl coenzyme A racemase, carbonic anhydrase 9) [12], and other molecules with disparate functions (CD82, tetraspannin 27, S110A1, protein S100-A1, thrombomodulin, transcription factor E3, Neprilysin) [13,14]. Recently, many other markers that could be employed in clinical routine have been proposed, but further studies are needed in order to assess their real role in the differential diagnosis of these lesions.

In this review, we summarize the key literature published within the last 5 years, dealing with diagnostic, prognostic, and predictive biomarkers for RCC in different specimens (tissue, cell lines, urine, blood, and extracellular vesicles [EVs]). One of the most widely used proteomic approach is based on the combination of liquid chromatography to tandem mass spectrometry (LC–MS). This approach allows both the identification and the quantification of proteins to be obtained (Figure 1). In this respect, DIGE technology can also reliably quantify protein expression in a comparative way, even if it is limited in the separation of highly hydrophobic proteins. A number of other technologies have also been used for proteomic analyses (including antibody-based

methods) of renal cancer. Due to the extensive use of MS-based proteomic approaches, we will mainly focus our attention on the outcomes deriving from these platforms.

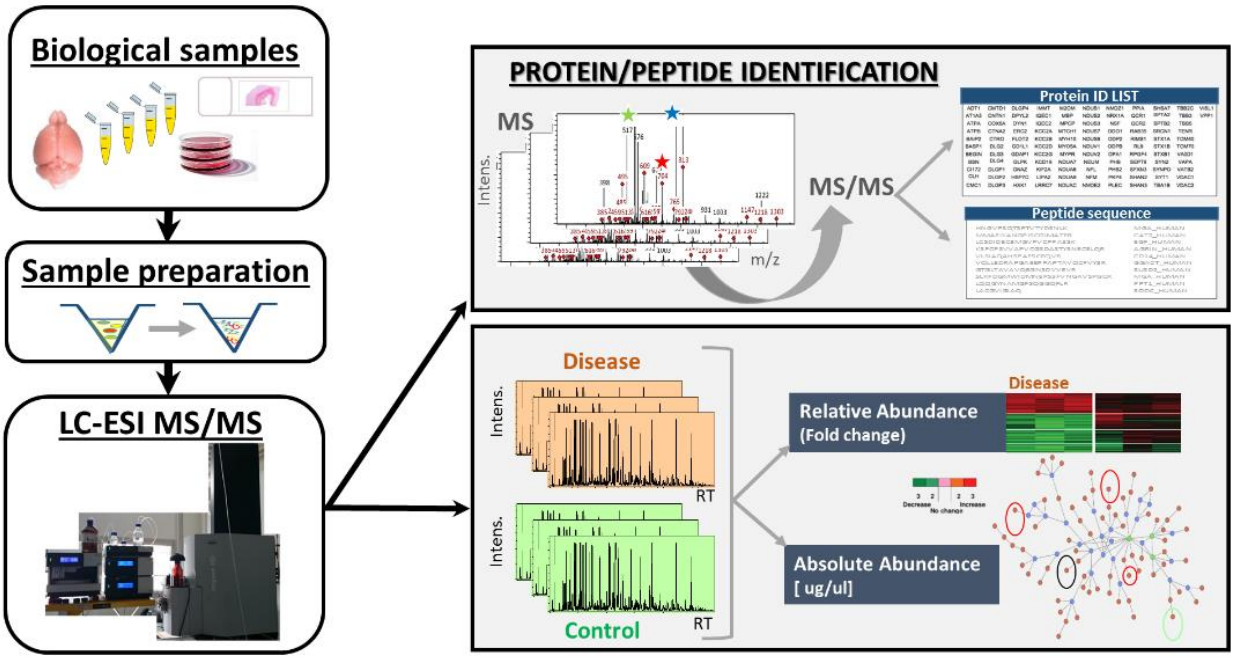


Figure 1. Typical bottom-up proteomic approach toward biomarker discovery using biological fluids. Proteins/peptides are extracted, and proteins are tryptic digested. Then, endogenous peptides or tryptic peptides are separated by LC and their identity and quantitative expression is ensured by MS/MS.

2. RCC tissue proteomics

The major benefits deriving from the use of tumor tissue in proteomics are related to the fact that the concentration of neoplasm proteins, including candidate biomarkers, is much more abundant in tumor tissue than in other specimens, and to the stronger specificity of the link between the differentially expressed proteins and the disease. On the other hand, several limitations are related to the nature of these kind of samples and its availability. Unfortunately, the studies are often retrospective, based on a small sample size with a rather short follow-up, and their applicability could be impaired by the heterogeneity of the tissue and the variability associated with the assay used for

protein detection. In particular, the reproducibility of MS-based results is a critical aspect that must be considered in this context, and that can significantly affect the clinical applicability of the biomarkers. In terms of experimental variability, protein identifications with a False Discovery Rate (FDR) below 1% have been commonly accepted, whilst for quantitative purposes, an associated variability between 10% and 30% has generally been observed for label-free approaches [15]. On the other hand, the definition of reproducibility across laboratories is very challenging due to the specificity and the heterogeneity of the biological specimens, together with the standardization of sample collection, processing, and analysis. These can also be the reasons behind the lack of concordance among different studies. However, many studies focused on tissue proteomics strongly contributed to the depiction of the molecular landscape of RCC, both in terms of biomarker discovery, biological insight, and spatial localization of molecular changes (Table 1).

2.1. Fresh-frozen tissues

The majority of investigations to discover biomarkers with a diagnostic role have been carried out on fresh-frozen (FF) tissue, comparing neoplastic tissue with the adjacent normal kidney (ANK). Various proteomic-based strategies have been employed for the proteomic profiling of RCC, mostly complementary to each other.

Gel-based technologies using conventional two-dimensional polyacrylamide gel electrophoresis (2DE) have been used to clarify the dysregulation of metabolic pathways and to highlight candidate biomarkers. One of the first attempts to obtain a comprehensive protein expression map for ccRCC was made by Seliger's group [16], performing a systematic comparison of protein expression profiles obtained by 2DE/MALDI-TOF. Using an analogous strategy, Sun et al. demonstrated that many proteins related to metabolism and cellular signaling were downregulated [17]. On the other hand, hypoxia-inducible domain family member 1A (HIF1A), together with

proteins involved in the detoxification of reactive oxygen species (ROS), was found to be upregulated. Using a gel-based approach, Giribaldi et al. identified significant differences in RCC tissue when compared with healthy tissue [19]. The overexpressed proteins included molecules related to glucose metabolism, carrier proteins, a member of the small heat shock protein family, and the calcium-binding protein reticulocalbin-1, whereas the downregulated counterparts appeared to be more related to metabolism or signal transduction, stress proteins, and ion-binding proteins. In 2012, Raimondo et al. reported that annexin-A2 (ANXA2), cyclophilin A (PPIA), fatty-acid-binding protein (FABP7), and galectin-1 (LEG1) were upregulated in comparison to ANK, through 2DDIGE/ MS. These proteins are all strongly involved in carcinogenesis and malignant behavior, as suggested by both the clinical and experimental data present in literature [25]. Very recently, Lu et al. used a comprehensive bioinformatics analysis combined with 2DE profiling in order to identify protein networks, involved in the onset or progression of ccRCC, that represent possible therapeutic targets [20]. These results show that since abnormal energy metabolism, together with the deregulation of the metabolic pathways, is a characteristic trait of RCC, metabolic control analysis could be particularly crucial in the management of kidney cancer.

Similarly, several gel-free approaches have been used with FF tissue, highlighting interesting dysregulations in the context of structural, immune system regulation, and signalling proteins. A few years ago, Siu et al. evaluated the protein expression in kidney cancer tissue versus ANK using the isobaric tags for relative and absolute quantitation (iTRAQ) labelling technique. It allows the simultaneous quantitation of up to four samples by LC–MS/MS [21].

Table 1. Proteomic/peptidomic approaches employed in RCC tissue studies.

Tissue	Target	Technique	First author (year)	Studied cohort	Analytical approach	Protein identification	Proteins of interest	Reference
FF	Diagnostic/ Speculative	2DE	Lichtenfels et al. (2009)	ccRCC ANK	2DE + MALDI MS/MS; Validation by IHC	348 DEPs spots (123 selected); 558 Ids (84 DEPs)	↑ GELS, ↓ CALB1, FABPH	[16]
			Sun et al. (2010)	ccRCC ANK	2DE + MALDI MS/MS;	16 DEPs spots (10 up and 6 down)	↑ HIF1a, GSTP1, PRDX3, PRDX6, CD2AP, ANXA5, ENOG, AL1A1, VIME, TGM2; ↓ CALB1, CK054, ADH, AMERL, ADT3, LYRM5	[17]
			Kim et al. (2010)	RCC ANK Benign masses Other tumors	2DE + MALDI MS/MS; validation by WB, IHC, ELISA	108 DEPs spots (83% down); 23 DEPs (13 up)	↑ NNMT, SEGN, PLSL, ENOG, NDKA, FRIL, PRDX4 in tissue; ↑ TBA1B, MIC60, TYPH, CAPG, RBBP7, HSPB1 in membrane fraction	[18]
			Giribaldi et al. (2013)	ccRCC ANK	2DE + MALDI MS/MS; validation by WB, IHC	18 DEPs spots; 23 DEPs (13 up)	↑ RCN1	[19]
			Lu et al. (2016)	ccRCC ANK	2DE + MS/MS; validation by WB	90 DEPs spots; 73 DEPs (47 up and 27 down)	↑ IF5A1 and KPYM	[20]
		iTRAQ + LC MS/MS	Siu et al. (2009)	ccRCC ANK CTRL (transitional cell carcinoma and end-stage glomerulonephritis)	iTRAQ labeling and LC–MS/MS; validation by WB *in silico* validation using databases	937 IDs (168 down and 156 up)	↑ SERPING1, NNMT, LDHA (in WB); ↑ MVP/LRP, ADFP, PD-ECGF, CRYAB	[21]
			Masui et al.. (2013)	RCC metastatic Primary RCC	iTRAQ labeling and LC–MS/MS; validation by WB analysis and TMA IHC	1256 IDs; 456 quantifiable IDs;29 DEPs	↑LEG1, PROF1, and 1433Z in primary ccRCC vs. kidney cortex, and in metastatic vs. primary RCC	[22]
			White et al. (2014)	ccRCC ANK	iTRAQ labeling and LC–MS/MS; validation by WB, IHC on TMA and ELISA; RT2 Glucose Metabolism Profiler Assay	1591 IDs; 55 DEPs (15 up)	↑ ENO1, LDHA, HSPE, HSPB1/Hsp27; ↑ HSPB1/Hsp27 in serum-urine and grade 3–4	[23]
		2D-LC MS/MS	Johann et al. (2010)	RCC ANK RCC plasma	2D-LC–MS; IHC validation	1275 IDs in RCC (202 only in RCC); 1281 Ids in ANK; 179 Ids in plasma; 8 DEPs	↑ CADH5, CAD11, DDX23, KPYM	[24]
		2D-DIGE	Raimondo et al. (2012)	RCC ANK	2D-DIGE + MALDI; validation by WB	41 DEPs (23 up)	↑ ANXA2, FABP7, LEG1, and PPIA	[25]
		MALDI imaging	Oppenheimer et al. (2010)	ccRCC ANK	MALDI imaging; protein identification by tissue homogenization, RP-LC fractioning + MALDI MS, in gel and in solution digestion, LC MS/MS; validation by IHC	12 DEPs	↑S10AA, S10AB	[26]
			Jones et al. (2014)	ccRCC ANK	MALDI imaging of proteins and lipids; protein identification by tissue homogenization, LC MS/MS	56 IDs	↑ FABP7 in RCC vs. normal	[27]
		Label-free LC MS/MS	Atrih et al. (2014)	ccRCC ANK	FASP + label-free LC MS/MS; validation by immunochemistry	1761IDs; 596 DEPs	↑ ADFP, COR1A	[28]
			Zhao et al. (2015)	ccRCC ANK	In solution digestion + label-free nLCMS/MS; validation by WB and IHC	1027 IDs; 213 DEPs (64 up)	↓THIL ↑ SODM	[29]
			Neely et al. (2016)	ccRCC ANK	Label-free 2D-LC–MS/MS; WB validation	Pooled samples (1551 IDs; 290 DEPs); individual samples (783 IDs; 344 DEPs)	Activation of ESRRA and ESRRG, and HIF1A, inhibition of FOXA1, MAPK1, and WISP2; ↑ COF1, PROF1, NNMT, ALDOA in late stage ccRCC	[30]
		PCT-SWATH	Guo et al. (2015)	ccRCC pRCC chRCC ANK	PCT + n2DLC -SWATH-MS	2375 IDs (317 down-296 up)	↑AMACR in pRCC vs. ccRCC; ↑GSTA1,VIME in ccRCC vs. non ccRCC	[31]
		Co-immunoprecipitationMS	Wang et al. (2015)	RCC ANK	Co-immunoprecipitation + nLC–MS/MS; validation by immunoprecipitation and WB RT-PCR, IHC	129 IDs potential YBX1 interactors	↑YBX1 ↓C1QBP	[32]
	Subcellular fractions	LC MS/MS	Raimondo et al. (2012)	ccRCC ANK	NuPAGE + nLC MS/MS; validation by WB	149 IDs (51 only in RCC, 56only in ANK)	↑ VDAC1, Thy1, BASI (EMMPRIN), Cav1 and Flot1; ↓ DPEP1 and AQP1	[33]
			Raimondo et al. (2015)	ccRCC ANK	FASP coupled to MD delipidation + nLC MS/MS; validation by immunoblot	742 IDs in ANK and 721 IDs in RCC	↑ CAH2, ANXA2; ↓ AMPN	[34]

Tissue	Target	Technique	First author (year)	Studied cohort	Analytical approach	Protein identification	Proteins of interest	Reference
FF	Other RCC variants	2DE	Valera et al. (2010)	ccRCC pRCC chRCC Oncocytomas ANK	2DE + In-gel digestion, MALDI–MS; validation by IHC	11 IDs	↑HSPB1, TPIS, PRDX2, APOA1 in ccRCC vs. ANK; ↑SODC, RD23B in chRCC; ↑ ENOA in oncocytomas	[35]
			Arai et al. (2015)	CIMP-positive ccRCC ANK	2DICAL analysis (2DE + LC MS/MS) + validation in IHC; RT-PCR; SNP microarray analysis	200 DEPs	↑ ANXA2; ↓ MSH2	[36]
		1DE	Mariño-Enríquez (2011)	Renal medullary carcinomas ANK	IHC, immunoprecipitation + WB; 1DE + LC MS/MS; RT-PCR	Targeted	Coprecipitation of strongly tyrosine phosphorylated talins with the VCL-ALK oncoprotein	[37]
		MALDI imaging	Na et al. (2015)	pRCC ANK	MALDI imaging; protein identification by tissue homogenization + RP fractioning, 1DE + LC MS/MS; validation by IHC on tissue and TMA	Targeted	↑ S10AB and FRIL in pRCC	[38]
	Prognosis/ Progression	2D-DIGE	Junker et al. (2011)	RCC pT1, pT2, pT3 ANK	2D-DIGE + MALDI; validation by WB, IHC	176 DEPs spots, 187 DEPs	Grading dependent ↑of PHB, stagerelated ↓ of PRDX3; stage- and grading ↑ of S10A9	[39]
		2DE	Ho et al. (2012)	ccRCC ANK	2DE + LC MS/MS; IHC; WB; migration analysis	Targeted	↑CRYAB in ccRCC; correlated to cell growth, migration and rapamycin-resistance	[40]
			Hosoya e al (2013)	ccRCC ANK	2DE + LC MS/MS; IHC; WB; cell invasion and proliferation analysis	73 DEPs	Nuclear NDRG1 associated to favorable prognosis and independent factor of progression-free survival; NDRG1 silencing ↑ cell proliferation and invasion	[41]
		iTRAQ + LC MS/MS	Lebdai et al. (2015)	ccRCC ANK	iTRAQ labeling and LC MALDI–MS/MS; validation WB and IHC	928 IDs; 346 quantifiable Ids	↑of TGFBI,NNMT and pentose phosphate pathway in aggressive tumors	[42]
FFPE	Prognosis/ Progression	LC MS/MS	Perroud et al. (2009)	ccRCC at different grade	LC MS/MS; validation by WB and IHC; pathway analysis	777 IDs; 105 grade-specific proteins	105 grade-specific proteins (among which AIFM1, AL1A1, ALDOA, PGK1, SERPH, NPM)	[43]
			Sprung et al. (2012)	RCC	LC MS/MS (shotgun + MRM); comparison of MRM between FF and FFPE	1982 IDs in FFPE and 2154 Ids in FF (91% shared)		[44]
			Craven et (2013)	ccRCC normal kidney tissue	Shotgun LC MS/MS	2516 IDs	No significant effect of FFPE block age on proteome	[45]
	Diagnosis		Weißer et al. (2015)	ccRCC ANK	Isotopic dimethyl labeling in LC–MS/MS analysis	1307 IDs; 189 DEPs (112 up)	↑ ANXA2, ANXA4	[46]
		MALDI imaging	Morgan et al. (2013)	TMA (ccRCC and ANK)	MALDI imaging; MS/MS sequencing; validation by IHC	Targeted	Classification cluster; ↑VIME, H2AX and ENO1	[47]
	Prognosis/ Progression + Other RCC variants		Steurer et al. (2014)	TMA (ccRCC and pRCC)	MALDI imaging on TMA	No identification	DEPs between pRCC and ccRCC; in ccRCC signals associated with tumor stage (7), Fuhrman grade (7), and presence of lymph node metastases (10), shortened patient survival (1)	[48]

Uniprot IDs of the protein of interest are specified in the table. Protein abbreviations are listed in Table 2.

↑: Increase; ↓: decrease; 2DICAL: 2-dimensional image converted analysis of Liquid chromatography mass spectrometry analysis; ANK: adjacent normal tissue; ccRCC: clear cell RCC; chRCC: chromophobe RCC; CIMP: CpG-island methylator phenotype; DEPs: differentially expressed proteins; FASP: filter-aided sample preparation; IDs: protein isoforms (Identities); IHC: immunohistochemistry; iTRAQ: isobaric tags for relative and absolute quantitation; MD: membrane microdomains; PCT: pressure cycling technology; pRCC: papillary RCC; SWATH: sequential window acquisition of all theoretical fragment-ion; TMA: tissue microarray; WB: Western blot; FF: fresh frozen; FFPE: formalin-fixed paraffin embedded.

An in-silico elaboration of the results using different tools and databases, including Serial Analysis of Gene Expression, UniGene EST ProfileViewer, Cancer Genome Anatomy Project, and Gene Ontology consortium analysis, was performed. They confirmed the role of some proteins already known to be involved in tumorigenesis, i.e. metabolism-related thymidine phosphorylase, of proteins implicated in proliferation and differentiation processes such as major vault protein, adipose differentiation-related protein (PLIN2), platelet-derived endothelial cell growth factor, and alpha-crystallin B chain (CRYAB). In 2013, Masui et al., with iTRAQ labeling and LC–MS/MS analysis, identified a dysregulated protein expression of profilin-1 (PROF1), 14-3-3 ζ/δ, and LEG1 in metastatic RCC [22].

A quantitative proteomic approach was used on FF tissue samples from primary RCC lesions and ANK using the filter aided sample preparation (FASP) method coupled with labelfree LC–MS/MS [28]. Among 596 proteins, PLIN2 and coronin 1A were further validated by immunohistochemistry (IHC) as the most interesting targets. Recently, Neely et al. performed label-free proteomics in 84 matched normal and ccRCC tissue samples to characterize the molecular phenotype of this neoplasm in terms of protein abundance and changes in biological pathways that are correlated to the presence of the tumor and its progression [30]. The proteome profiling was conducted both on pooled and individual samples in order to maximize the capacity to identify the molecular alterations without losing the information linked to the variability of ccRCC patients (stage/metastasis). In addition, they compared their outcome with a previously published transcriptomic data set from the same sample cohort. Using protein abundance data from all four stages, HIF1A and estrogen-related receptor α and γ were predicted to be activated whereas WNT1-inducible signaling pathway protein 2, hepatocyte nuclear factor 3-α, and mitogen-activated protein kinase 1 were predicted to be inhibited, confirming that many shared metabolic features among ccRCC stages were mostly related to the Warburg effect. Moreover, they also reported a relevant increase

of a small number of proteins in stage IV, such as cofilin-1, PROF1, nicotinamide N-methyltransferase (NNMT), and aldolase A (ALDOA), pinpointing late stage tumor diversity along with intratumor heterogeneity as one of the most defining traits of ccRCC. The authors also suggested that tumor aggressiveness and treatment could be likely affected by the VHL/HIF1A/ HIF2A axis, since these proteins correlated with the most relevant differences observed between tumor stages.

One key issue related to tumor biomarker research is the translation of proteomic methods and findings to applicable clinical assays, and the simultaneous analysis of tissue and peripheral fluid specimens should be beneficial when attempting to improve this aspect. Regarding RCC, Johann et al. combined the analysis of tissue and plasma samples with 2D-LC– MS, identifying four tumor-residing proteins in the blood of a patient newly diagnosed with RCC (cadherin-5, cadherin-11, DEAD-box protein-23, and pyruvate kinase) [24]. In the same year, Kim et al. starting from the hypothesis that shedding of tumor cell membrane proteins could easily release biomarkers into patient plasma, investigated RCC tissue and their membrane- enriched fractions by 2DE/MALDI [18]. They detected a potential diagnostic biomarker panel for RCC, also using samples from benign kidney masses and non-ccRCC tumors. In particular, NNMT was also highlighted by IHC and WB to be the most commonly upregulated protein for all types of RCC, even in the plasma of patients.

2.2. Mitochondrial components involvement

Particular attention needs to be paid to the role played by the dysregulation of the mitochondrial components and their tumorigenesis potential that is probably related to the increased oxidative-induced damage to cells. Notably, the Warburg effect seems to enable cancer cells to avoid excess ROS generation from mitochondrial respiration, providing a survival advantage for metastasis [49]. Zhao et al. used a nano-LC–MS/MS and label-free quantitative approach to identify two mitochondrial proteins, acetyl-CoA

acetyltransferase 1, and manganese superoxide dismutase, that were RCC related [29]. White et al. performed quantitative proteomic analysis by iTRAQ labeling and LC–MS/MS on a large cohort of approximately 200 patients, confirming the upregulation of mitochondrial metabolism proteins in RCCs, such as alpha-enolase (ENOA), L-lactate dehydrogenase A chain (LDHA), heat shock protein beta-1 (HSPB1/Hsp27), and the mitochondrial 10 kDa heat shock protein (CH10). The increase in LDHA and the activation of the pyruvate kinase pathways is likely to highlight an active anaerobic glycolysis, following the hypoxic conditions that are known to be an integral component of the functional signatures of RCC. Furthermore, Hsp27 levels were also elevated in urine and serum, with serum Hsp27 levels correlating with grade 3–4 [23].

2.3. Other technical approaches

Interestingly, Guo et al. described a new method capable of converting small amounts of tissue (similar in dimensions to a biopsy) into a digital file representing the MS-measurable proteome, combining pressure cycling technology and sequential window acquisition of all theoretical fragment ion spectra (SWATH)-MS [31]. They could detect and quantify several proteins able to differentiate distinct histomorphological kidney cancer subtypes. Another interesting approach was applied by Wang et al. to investigate the interactome of Y-box-binding protein 1 (YBOX1), a protein involved in tumorigenesis by co-immunoprecipitation and MS, to find candidate- binding partners [32]. Expression of YBOX1 was enhanced in the RCC tissues, and its nuclear expression was associated with histological T stage and metastasis. At the same time, the RCC levels of Complement component 1 Q subcomponent-binding protein, one of the proteins interacting with YBOX1, were significantly lower than in ANK tissue, and negatively correlated with the nuclear localization of YBOX1 in RCC tissue.

2.4. Subcellular fractions

Since cell surface proteins, exposed toward extracellular space, are ideal candidates to be used as biomarkers and therapeutic targets, the study of the alterations occurring in the plasma membrane is particularly intriguing [18]. Raimondo et al. investigated the proteomic profile of membrane microdomains (MD), plasma membrane supramolecular structures involved in cell signaling, transport, and neoplastic transformation, obtained by density gradient centrifugation from cancer and ANK tissue [33]. It has been shown that the global composition of MD is profoundly altered in RCC, highlighting some potential targets that can play a crucial role in ccRCC transformation. In particular, the levels of some proteins involved in signaling and adhesion, typical MD-associated processes, such as G proteins, flotillins, tubulins, annexins, and caveolins, were increased. Moreover, they optimized an enzymatic digestion protocol specific for RCC MD samples, based on FASP coupled to delipidation that enabled the identification of a larger panel of differential proteins [34].

2.5. Prognostic markers in tissues

Another interesting research field is related to the discovery of reliable prognostic markers that are able to predict the course of the disease through the analysis of one or more analytes. Junker et al. investigated the correlation of the protein expression with tumor stage (pT1, pT2, and pT3) by 2D-DIGE followed by MALDI-TOF [39]. In particular, they noticed that PCA could clearly separate not only control samples from tumor samples but also a pT1 cluster from a pT2/pT3 cluster. Ho et al. showed that CRYAB is overexpressed in ccRCC tissue, and the overall survival of patients was significantly correlated with its expression in the tumor [40]. Using a 2DE/MS approach, Hosoya et al. revealed a significant association between N-myc downstream-regulated gene-1 protein (NDRG1) and a good prognosis. They showed that NDRG1 is an independent factor of progression-free survival and its in vitro gene suppression

enhanced proliferation and invasion of RCC cells [41]. Using iTRAQ labeling and MALDI–MS, Lebdai et al. identified 928 proteins. Among them, 346 proteins had an altered expression in tumor tissue. In particular, the transforming growth factor β-induced protein (BGH3/TGFBI) showed a significant association with the stage, size, and grade, and a metabolic shift via pentose phosphate pathway in aggressive tumors has been confirmed [42].

2.6. Non-ccRCC variants

As already mentioned, ccRCC represents the most frequent histotype of kidney cancer and, for this reason, is the most commonly investigated type of kidney tumor. However, some authors also tried to identify new candidate biomarkers in less frequent RCC variants. One of the early studies was carried out by Valera et al. using a 2DE/MS approach in order to outline differences in protein levels among different subtypes of RCCs, including papillary lesions, chromophobe tumors, and renal oncocytomas [35]. They observed a positive correlation for triosephosphate isomerase and Hsp27 and the clinic-pathologic features of the patients (sex, age, histology, grade, and stage of the tumor). In an interesting study by Mariño-Enríquez et al., genetics and proteomics data were combined. They described a case of the medullary variant of RCC carrying a particular translocation, t(2;10)(p23;q22), involving the oncogene ALK tyrosine kinase receptor (ALK) [37]. The employment of MS facilitated the identification of a novel ALK oncoprotein in which the cytoskeletal protein vinculin was fused to the ALK kinase domain. In 2015, Arai et al. studied the molecular pathways involved in CpG-island methylator phenotype-positive ccRCC by a multiomic approach that also included 2-Dimensional Image Converted Analysis of Liquid Chromatography mass spectrometry analysis (2DICAL) analysis and the determination of proteins specific for the disease [36].

2.7. FFPE tissue

In addition to FF tissue, formalin-fixed paraffin-embedded (FFPE) tissue also represents a valuable sample source for investigations, even for the newest proteomic approaches [50]. The first attempt employing this form of specimen in the study of RCC was the work of Perroud et al. that correlated tumor grading to putative biomarkers [43]. All the differential proteins had in common their role in cellular metabolism and the regulation of cell proliferation: they were likely to be strongly correlated with the glycolytic and amino acid synthetic pathways, as well as with acute phase and xenobiotic metabolism signaling. Some efforts were dedicated to the comparison of the FFPE tissue specimens with non-chemically modified FF samples, even specifically for RCC, in order to evaluate the consistency of biological data and the feasibility of the approaches. Sprung et al. compared the reproducibility of quantitation for peptides deriving from FFPE and FF ccRCC tissue samples by MSbased multiple reaction monitoring (MRM) [44]. Interestingly, coefficients of variation for measurements in FFPE and FF tissue were very closed (range 18−20%). Craven et al. systematically investigated the effect that the storage time of FFPE tissue blocks in pathology archives had on the quality of data obtained by label-free MS [45]. Normal kidney and ccRCC tissue routinely collected up to 10 years prior to their analysis were profiled using LC–MS/MS. Results indicated the absence of significant effects deriving from the age of the block, thus reinforcing the use of this specimens for biomarker discovery studies. Recently, Weißer et al. verified that dimethyl labelling is applicable and sufficiently accurate for the quantitative proteomic analysis of FFPE RCC tissue specimens, without interference from the formalin fixation process [46].

2.8. Mass spectrometry imaging

Mass spectrometry imaging (MSI) is a new analytical technique used for biomarker discovery directly on tissue, due to its ability to correlate traditional morphological data

with molecular information obtained by this proteomic approach [51,52]. This method generally employs the use of a MALDITOF and FF or FFPE tissue (Figure 2) [50]. In MALDI–MSI, a mass spectrum is acquired at discrete coordinates throughout the entire tissue section. The spatial distribution of the biomolecules present in a specimen can be visualized and a molecular image of the tissue can be recomposed based on the acquired mass spectra.

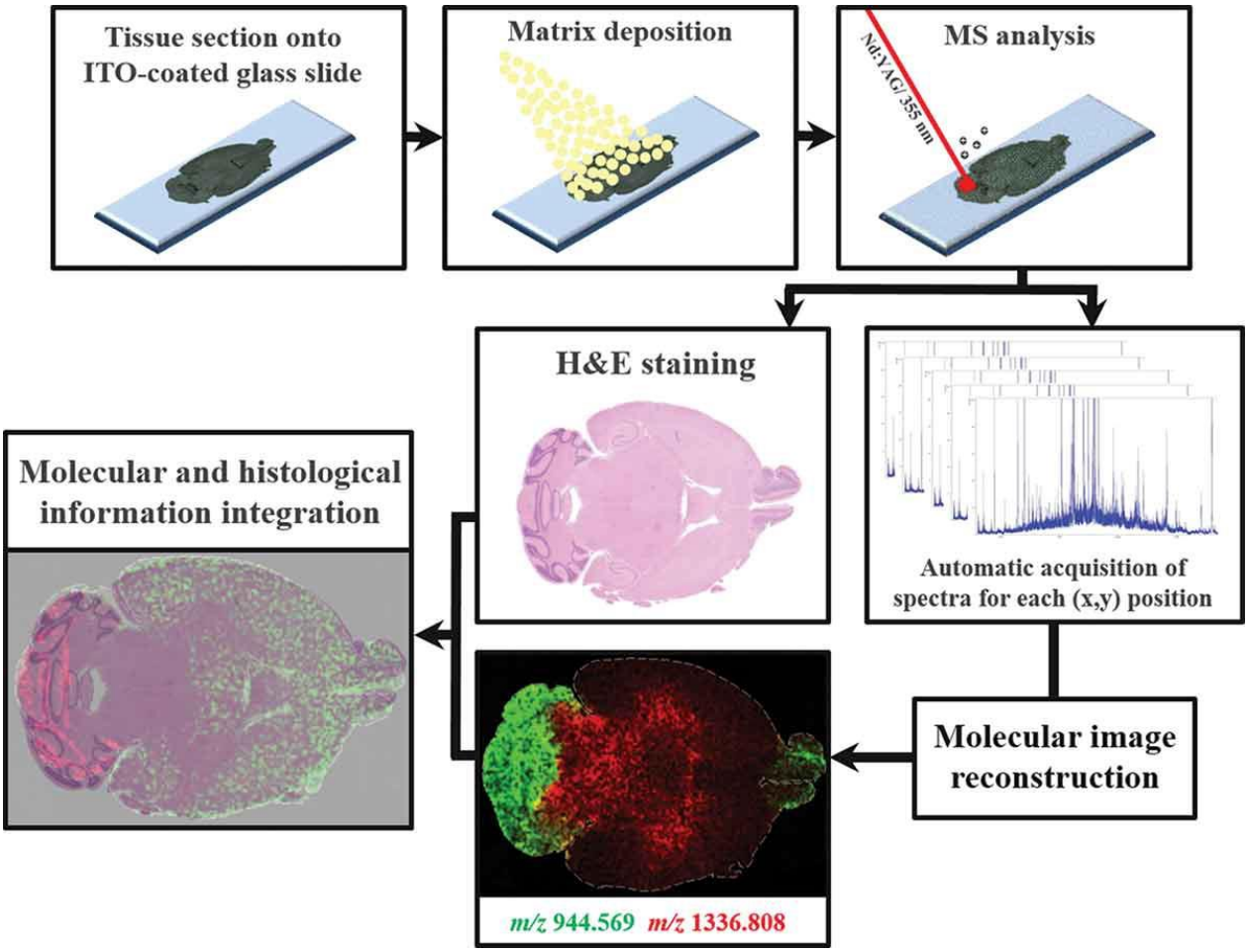


Figure 2. MALDI-TOF tissue imaging flow-chart. Thin tissue sections are cut and covered with matrix, after appropriate washing steps. Then, the entire tissue section is analysed by MALDI-TOF. Signals present in the spectra are then used to generate the molecular image of the tissue, which shows their spatial distribution. This molecular image obtained by MSI can be overlapped with the image obtained staining the tissue i.e. with Eosin and Hematoxylin (H & E).

The possibility to identify new biomarkers directly on renal tissue by MSI has also led some authors to investigate this approach in RCC. A pioneering study was performed in 2010 by Oppenheimer et al. in order to characterize the histological tumor margins of ccRCC tissue using MSI [26]. They showed that histologically, normal tissue adjacent to the tumor contains several molecular characteristics of the cancer tissue, representative of an abnormal microenvironment, which could guide the intervention on patients. This finding was further deepened by classical MS of the FF tissue specimens, showing that proteins of the mitochondrial electron transport system are an example of such altered distributions. Caprioli's group also used MSI to define the molecular signature of RCC tissue [47]. The integration of protein and lipid molecular signatures was obtained by Jones et al. [27]. Based on MSI, they detected a cluster of 108 proteins that had potential disease-specific expression patterns. Fifty-six of these proteins were identified after protein extraction by MS/MS. Lipid profiling results obtained by MALDI-FT-ICR (Fourier transform ion cyclotron resonance) were integrated with protein expression profile data. A group of 26 proteins and 39 lipid species able to distinguish either tumor from non-tumor tissue, or recurrent disease progressors from nonrecurrent disease patient tissues, was obtained. A similar attempt was later performed by comparing MSI results with clinicopathological data [48]. Significant differences were observed between papillary and ccRCC. Within the subgroup of ccRCC, 2 groups of 7 signals were noticed to have a statistically significant association with the tumor stage, Fuhrman grade, and 10 signals to the presence of lymph node metastases. Moreover, one signal was significantly linked to shortened patient survival. Finally, in the study of Na et al., surgical tissue sections of papillary RCC were analyzed using MALDI–MSI [38]. Statistical analysis revealed several distinctive cancer-specific m/z-species between normal and diseased tissues. Among these m/z-species, two particular proteins, S10AB and ferritin light chain, which are

specific for papillary RCC cancer regions, were successfully identified using LC–MS/MS following protein extraction from independent RCC samples.

3. Proteome studies in RCC in vitro cultures

Because of their easy use, cell cultures, both primary and established lines, are particularly useful for pathophysiology studies, biomarker discovery, and therapeutic investigations.

Proteomic analysis of RCC cell lines was shown to improve the understanding of the cancer biology. In particular, an important field of research is represented by the VHL-driven and undriven cancers, aimed at finding useful markers to be possibly transferred into clinical practice. Since 2006, several authors adopted different proteomic strategies in order to unravel this issue, ranging from classical 2DE/MALDI-TOF analysis of the cell proteomes [53,54] to LC–MS/MS of membrane protein-enriched fractions following stable isotopic labeling [55]. In the last 5 years, innovative and sophisticated approaches were applied and almost all were directed toward posttranslational modifications and systems biology. Combining cell surface capturing technology and SILAC-based quantification, Boysen et al. identified a pool of surface N-glycoproteins altered by VHL expression. Among these, metalloproteinase CD10 was shown to be a potential therapeutic target given that its inhibition decreased the invasiveness of ccRCC cells [56]. Moreover, Nagaprashantha et al. studied the effect of incrementing HIF2A, comparing Caki- 2 VHL-wt human and 768-O VHL-mut RCC cell lines by a labelfree LC–MS/MS approach [57]. They found differential expression of proteins involved in cancer cell energetics and oxidative stress in VHL mutant cells. Two recent papers focused on the correlation between hypoxia and VHL gene mutations. Leisz et al. compared three VHL-RCC cell lines (786-O, RCC4, and RCC10) and their VHL transfectants, under normoxia and hypoxia, in order to identify the concordant and

different genes/proteins expressed in these conditions. Using a combined approach of cDNA microarray and 2DE/MALDI-TOF analysis, the study not only showed similarities but also differences in the VHL loss- and hypoxicinduced pathways, mainly in cellular metabolism. Interestingly, an upregulation of ALDOA was observed in both the conditions, representing a potential target independent of the VHL status or oxygenation [58]. Malec et al. dealt with the same topic, comparing the signatures relative to VHL loss and/or hypoxia by a proteome and phospho-proteome analysis of isogenic 786O RCC (±VHL). They identified and quantified around 2000 proteins and a similar amount of phospho-proteome. The differential analysis confirmed the fact that some pathways are altered by both VHL and hypoxia, whilst others are not, generating a specific signature for each of these conditions. Notably, intracellular carbonic anhydrase 2 was under-expressed after both VHL loss and hypoxia; the authors correlated this finding with the NFkB signaling pathway, which is likely to respond to the pCO2 fluctuations in the microenvironment of cancer cells [59].

It has to be underlined that some differential proteins identified using cell lines were also confirmed by tissue proteomics, thus strengthening their key role in RCC. Among these, some play a crucial role in immune system regulation, such as those found by Dihazi et al. [60]. They showed that 20 different RCC cell lines had very similar protein profiles when employing a comprehensive proteomic platform (SELDI-TOF, 2DE/MALDI-TOF): LEG1 was identified as a potential RCC marker in the membrane fractions, as further confirmed by other authors [25]. Comparing a ccRCC cell line (RLC-310) and a normal one (HK-2), Yang et al. identified 31 differentially expressed proteins by 2DE/MALDI-TOF [61]. Among them, the most significant was PPIA. Its overexpression was confirmed using real-time PCR and WB analysis in ccRCC cell lines and tissue. Interestingly, IHC revealed that the overexpression of PPIA was associated with poor differentiation and decreased survival [61]. The same group updated these results using 2D-LC fractionation and LC–ESI-MS/MS, identifying 13, new,

differentially expressed proteins [62]. Among them, growth factor angiomotin (Amot) was very promising. In fact, its overexpression in ccRCC was confirmed by WB through the comparison with ANK tissue. Moreover, quantitative real-time PCR of Amot mRNA in 127 ccRCC tissue samples underlined its association with poor differentiation, venous invasion, and decreased survival, supporting that Amot may be considered as a novel prognostic factor of ccRCC. Notably, Tanaka et al., investigated the protein expression profiles of four SN12C human RCC cell clones with different metastatic potential by 2DE/LC–MS over a narrow pH range (pH 4.0–7.0) [63]. They showed a high expression of glyoxalase-1, a protein involved in detoxification, in clones with a high metastatic potential and a low expression in those with low metastatic potential.

The research of potential biomarkers was not only accomplished in cell lysates or subcellular fractions, but also by employing nonconventional strategies. In fact, Minamida et al. researched potential biomarkers in the secretome of the 769P RCC cell line in order to identify proteins that were not yet associated with RCC. The analysis was performed by 1DE and LC–MS/MS and lead to the discovery of PROF1, a structural protein involved in cytoskeletal dynamics, membrane trafficking, and nuclear transport. PROF1 overexpression was confirmed by WB, RT-PCR, and IHC of RCC tissues [64]. Another intriguing study was focused on endothelial tumor associated markers in order to characterize the angiogenesis process in RCC. Mesri et al. set up a method based on flow cytometry to isolate endothelial cells from RCC and ANK tissues of patients. By label-free LC–MS/MS, they detected enriched cell surface proteins to be investigated as potential endothelial targets for therapy [65]. In summary, all these studies were aimed at understanding the pathogenic mechanism of RCC and identifying novel prognostic and predictive markers, implicating the discovery of target cancer-specific pathways for this highly chemo-resistant tumor.

Moreover, cell cultures are the ideal model to evaluate the effect of drugs on potential targets. A series of proteomic studies were undertaken to reveal the protein expression

changes occurring in different cisplatin-resistant RCC cell lines. Vasko et al., through 2DE/MALDI-TOF, found that oxidative stress proteins were differentially regulated [66]. More recently, this result was confirmed in a paper highlighting the role of protein deglycase DJ-1, an oxidative stress scavenger, in RCC resistance to cisplatin-induced apoptosis [67]. In addition, proteomic analysis by reverse phase protein array confirmed the central role of AMP-activated protein kinase activation and mTOR pathway inhibition in RCC sensitivity/resistance to 8-chloroadenosine [68]. Another study based on proteome profiler arrays and WB analysis was performed in order to clarify the mechanism of betulin induction of caspase mediated apoptotis cell death in a multidrug-resistant human renal carcinoma cell line. Such treatment was shown to increase the expression of apoptosis-related proteins, including Poly ADP ribose polymerase (PARP) and Bcl-2 family members. Moreover, betulin might sensitize RCC cells to other anticancer drugs, since the combination of betulin with etoposide decreased the expression of multidrug-resistance protein 1, a known cause of RCC chemoresistance [69]. Fan et al. carried out an interesting investigation dealing with the erythropoietin (EPO) signaling pathway and RCC progression: they evaluated the EPO effect on two cell lines (±VHL) by LC–LTQ–Orbitrap–MS [70]. This negative EPO effect was correlated with the overexpression of the proliferating cell nuclear antigen-associated factor (KIAA0101/PAF15). On the contrary, the low expression of KIAA0101 seems to be associated with a better prognosis and a higher 5-year survival in RCC patients.

Another promising therapeutic strategy for RCC is immunotherapy. Seliger et al. used a ligandomic profiling approach to study the proteins that bind to the HLA class I antigens in MZ2733RC cell lines, in order to find targets able to elicit an immune response against RCC [71]. This approach led to the identification of approximately 50 HLA class I peptide ligands present on the cell surface of RCC cells. In particular, the authors showed that T cells directed against a specific HLA ligand, derived from sulfiredoxin-1, recognize and lyse RCC ligand-presenting target cells, thus providing

clues relevant for the development of T-cell-based immunotherapy. Moreover, Ranieri's group recently established and cloned an immunogenic cell line (Elthem) derived from the tissue of a patient affected by a nonaggressive and nonmetastatic ccRCC. This cell line was shown to be suitable for immune-stimulation. After extensive characterization by genomic and proteomic approaches, they found an enrichment of many cancer-related biological processes and pathways, such as oxidative phosphorylation and glycolysis. The authors state that novel tumor-associated antigens are currently in the process of being identified and that this will enable the evaluation of new immunotherapeutic strategies [72].

Recently, phospho-proteomic studies demonstrated their great potential to provide novel therapeutic strategies in RCC. Chen et al. showed that 17-β estradiol-induced growth inhibition in RCC ACHN cells via the upregulation of the phosphorylation level of the autophagy activation effector sequestosome 1 (SQSTM1) [73]. In a more comprehensive study, Haake et al. performed tyrosine phosphorylation profiling in 10 RCC cell lines, as well as in 15 clear cell and 15 papillary RCC human tissues specimens. Using this innovative approach, the authors identified activated signaling pathways as therapeutic targets and screened different tyrosine kinase inhibitors, highlighting the focal adhesion kinase as a potential target in RCC cells and tumors. Moreover, the comparison between clear cell and papillary RCC tissue led to a specific pathway for the papillary subtype, driven by the phosphorylation of the receptor tyrosine kinase epithelial discoidin domain-containing receptor 1 (DDR1) [74], to be discovered. Interestingly, this work pointed out that RCC cell lines do not always represent the biology of the tumor from which they were obtained. In fact, the authors showed that ~25% of all tyrosine phosphorylated proteins were exclusive to the cell lines, while several other proteins were detected only in the tumors. One of the reasons that could explain this difference is that the cell lines were grown on plastic without collagen, an essential ligand required to activate/phosphorylate DDR1, which is in fact identified

exclusively in the human tumor tissue. Therefore, studies on cell lines require the validation of their results in patient tissue, as performed in almost all the papers cited above. Regarding the therapeutic strategies, in vivo experiments are essential but are yet to be achieved until now.

4. RCC biomarker discovery in peripheral fluids

Biological fluids offer the enormous opportunity to obtain a distinctive fingerprint specific of the diseased condition through non, or very lowly, invasive methods. Unfortunately, this information is often technically difficult to be mined because the disease-related proteins are often present in very low concentrations, are frequently labile, and are hidden by high-abundance proteins such as albumin or immunoglobulin. The discovery of these molecular alterations requires powerful and very sensitive methods, which embrace the latest advance proteomic techniques, including appropriate clean-up systems as well as effective and reproducible enrichment of the detectable low-abundance proteins (e.g. hydrogel nanoparticle [75]), that can be coupled with MS. Here, we report the results of various studies, targeted to both the peptidome and the proteome, in blood and urine of RCC patients. This information is summarized in Table 2.

4.1. Serum/Plasma proteome

Blood is considered as the most valuable specimen for biomarker discovery since it is easily accessible and carries most of the tissue-derived molecules in the organism. Serum, in particular, was widely used for detecting renal cancer markers, despite the dynamic range of protein concentration being about nine orders of magnitude and could hide the alterations generated by the pathological process [94]. These molecules belong

to different classes of proteins and play different roles [95]. For example, some of them are involved in the inflammatory processes, or in apoptotic mechanisms, and/or even in RCC hypoxic response [95–97].

Very recently, several studies have been reported for the early detection of ccRCC in serum. Zhang et al. observed 16 up- and 14 downregulated serum proteins in the early stage, applying iTRAQ labeling, and LC–MS/MS analysis [96]. Among them, HSC71, a member of the heat shock protein 70 family, was selected as the most dysregulated protein and its overexpression was validated by WB and ELISA, also in comparison with other urological diseases. Using a similar approach, Zhang et al. identified 27 differentially expressed proteins in early-stage RCC, when compared with patients with benign kidney lesions, other urological tumors or noncancer patients, and with the normal controls [97]. Most of them can be connected to the dysregulated immune response of ccRCC, confirming one of the peculiar features of this neoplasm [80]. Interestingly, 11 of these proteins were cross-validated in RCC tissue against The Cancer Genome Atlas database (https://tcga-data.nci.nih.gov/tcga/), since they matched with similarly up or downregulated genes, in both the specimen types. In particular, two members of the small Ca2+-binding protein S100 family, S10A8 and S10A9, were confirmed to be potential biomarkers for the detection of RCC or even promising targets for therapeutic intervention, through in-depth bioinformatics analysis and IHC [82].

Plasma is also often used as a valuable source of candidate cancer biomarkers, instead of serum [98]. Johann et al. reported a method that combined tumor/plasma proteome analysis using 2D/LC–MS from a single patient newly diagnosed with a nonmetastatic RCC, as already discussed in the tissue proteomics section [24]. A year later, Yokomizo et al. observed an increased expression of Fibronectin-1 (FINC) in the plasma of ccRCC patients, especially in the early stage [81]. A validation in a hundred-scale cohort (77 RCC patients, 130 healthy individuals, 20 prostate cancer patients) was performed using

the Amplified Luminescent Proximity Homogeneous Assay (AlphaLISA). Results were likely to suggest a clinical relevance of monitoring the level of plasma FINC for the diagnosis of RCC patients. Another study, based on ELISA, showed the potential utility of a further three circulating plasma markers for RCC prognosis (C-reactive protein, osteopontin, and carboninc anhydrase 9), for improving the existing clinicopathological prognostic models, that are based on the stage alone [99]. Recently, Gbormittah et al. showed that the posttranslational modifications, such as glycosylation, can also be used to detect RCC in plasma samples [76]. Specifically, they searched for possible ccRCC biomarkers in the plasma of patients before and after curative nephrectomy for localized tumors, using a comparative study of the proteome, glycoproteome, and N-glycome. Through a multidimensional fractionation approach (12P-M-LAC) and LC−MS/MS, they selected a panel of 13 glycoproteins, whose glycan structure was supposed to be changed in diseased samples [77]. Later, using CID (collision-induced dissociation) tandem MS, the same group better characterized the N-linked glycan sites of immunoaffinity-purified clusterin, one of the proteins included in the previous panel. In particular, a significant increase in the levels of two specific clusterin glycoforms was observed, as further confirmed by Aleuria aurantia lectin blots [79].

4.2. Urine proteome

Although the discovery of proteins usable as renal cancer markers in urine is extremely desirable, unfortunately, for RCC, the number of possible candidates in urine seems to be less than 10 [100]. This can be noted as a general shortcoming of biomarker research in this fluid [97] and is likely to be strictly related both to different technical aspects (e.g. the lack of reproducibility, the complexity of the urine, and the difficulty to capture the changes masked from the high abundant proteins) and to biological and clinical issues (e.g. the influence of confounding factors, such as age- or lesion related kidney damage, and/or the effect of the proteolytic degradation processes).

Table 2. Proteomic/peptidomic approaches employed in RCC biofluids studies.

Biofluid	Target	Technique	First author (year)	Studied cohort	Sample collection	Analytical approach	Proteins/Peptides of interest	References
Plasma	Proteome	2D-LC MS/MS	Johann et al. (2010)	Nonmetastatic RCC Healthy controls	Presurgery	2D-LC–MS; validation by WB	↑ CADH5, CAD11, DDX23, KPYM in one nonmetastatic RCC patient, as well four other patients diagnosed with RCC	[24]
			Yokomizo et al. (2011)	ccRCC Healthy controls Prostate cancers	Presurgery	2DICAL. Validation using the Amplified Luminescent Proximity Homogeneous Assay (AlphaLISA)	↑ FINC in ccRCC	[76]
		Immunoassay	Sim et al. (2012)	ccRCC	Presurgery	ELISA	Circulating CAH9, OSTP, and CRP add value to existing clinicopathological prognostic factors/model	[77]
		Label-free LC MS/MS	Gbormittah et al (2014)	ccRCC	Presurgery Postsurgery	Multidimensional fractionation (12P-M-LAC) + label-free LC MS/MS	↑ PGBM, MUC18, ECM1, LYAM1, SYNE1, VCAM1 (low-abundant Glycosilated); glycoprotein panel (↑ CO2, APOB, CO4A, LG3BP, FINC, CLUS; ↓ KLKB1, MA2B1, C1 R, FETUA, ITIH4)	[78]
		1DE + LC MS/MS	Gbormittah et al (2015)	ccRCC	Presurgery Postsurgery	Immunoaffinity HPLC Purification + 1DE + LC MS/MS	↑ FA2G2S2 and A2G2S2 glycans in postsurgery plasma clusterin	[79]
Serum	Proteome	iTRAQ + LC MS/MS	Zhang et al. (2015)	ccRCC Healthy controls	Presurgery	Immunodepletion + iTRAQ labeling + SCX fractioning + LC MS/MS; validation by WB	↑ HSP7 C b at early stages	[80]
			Zhang et al. (2015)	RCC Transitional cell carcinoma in the kidney Benign renal masses Normal controls	Presurgery	iTRAQ labeling and LC–MS/MS + gene/microRNA expression pattern + IHC	↑ S10A8 and S10A9 at early stages	[81]
			Zhang et al. (2016)	ccRCC Other urological tumors Benign renal masses Healthy controls	Presurgery	iTRAQ for relative and absolute quantification + IHC	263 Ids; 74 DEPs (4 down and 23 up in ccRCC according to TCGA database); ↑ C1QC, C1QB, S10A8, S10A9, CERU ↓ LUM	[82]
	Peptidome	SELDI	Nuerrula et al. (2015)	ccRCC Healthy controls	Presurgery Postsurgery	SELDI-TOF–MS combined with weak cation exchange protein chip (CM10) technology	Significant difference on expression level both preoperatively and in postoperative period of five signals corresponding to Bcl-2 family apoptosis regulatory proteins, WAP four-disulfide core protein, Krueppel-like factor 8, monocyte chemotactic protein-1, serum amyloid β-protein-4	[83]
		Magnetic beads + MALDI MS	Liu et al. (2011)	RCC Benign renal space-occupying lesions	Presurgery	Magnetic beads with a WCX external surface/MALDI-TOF	A cluster of four signals differentiate malignant tumors from benign renal masses	[84]
			Gianazza et al. (2012)	ccRCC Non-ccRCC tumors Benign renal masses Healthy controls	Presurgery	Magnetic beads with a C8 hydrophobic external surface/MALDI-TOF + nLC MS/MS	A cluster of five signals differentiate malignant tumors from benign renal masses and healthy subjects. Six ions distinguish ccRCC from controls. Healthy subjects were also differentiated from non-ccRCC by three features. Seven among 12 signals considered in the discriminant patterns were identified	[85]
			Yang et al. (2014)	ccRCC Healthy controls	Presurgery Postsurgery	Magnetic beads with a WCX external surface/MALDI-TOF + LC MS/MS	Three signals of cluster were identified (RBP6, TUBB, ZFP3)	[86]
		LC MS/MS	Huang et al. (2014)	ccRCC Healthy controls	Presurgery	LC MS/MS	Cluster of peptides for RCC detection	[87]

Biofluid	Target	Technique	First author (year)	Studied cohort	Sample collection	Analytical approach	Proteins/Peptides of interest	References
Urine	Exosome	1DE + LC MS/MS	Raimondo et al. (2013)	ccRCC Healthy controls	Presurgery	Exosome isolation by ultracentrifugation; SDS-PAGE combined with LC–ESI-MS/MS; validation by WB in patient urinary exosomes	↑ CERU, PODXL, MMP9, CAH9, DKK4 ↓ NEP, BASI (EMMPRIN), DPEP1, SDCB1, AQP1	[88]
	Peptidome	SELDI	Alves et al. (2013)	ccRCC Healthy controls	Presurgery	SELDI-TOF based on CM10 chip mass SDS-PAGE combined with MALDI-TOF/TOF	↓ 22 Peaks in RCC patients: no data regarding diagnostic performance ↓ MASP2 and SCTM1	[89]
		Magnetic beads + MALDI MS	Bosso et al. (2008)	ccRCC Healthy controls	Presurgery	Magnetic beads with a C8 hydrophobic external surface/MALDI-TOF	Cluster of three ions able to discriminate patients from controls, sensitivity of 100% for patients at the primary tumor stage: potential biomarkers for early RCC diagnosis ↓ UROM fragment in RCC patient	[90]
			Chinello et al. (2014)	ccRCC Non-ccRCC tumors Benign renal masses Healthy controls	Presurgery	Magnetic beads with a C8 hydrophobic external surface/MALDI-TOF + nLC MS/MS	One cluster of 12 signals could differentiate malignant tumors from benign renal masses and controls; second cluster of 12 signals distinguished clear cell RCC from controls; 7 of them were identified	[91]
			Chinello et al. (2015)	ccRCC Healthy controls	Presurgery	Magnetic beads with a C8 hydrophobic external surface/MALDI-TOF + LC MS/MS	Peptides correlated with tumor size, pT, and grade; identification of 13 of them	[92]
		CE–MS	Frantzi et al. (2014)	ccRCC Healthy controls Other renal diseases	Presurgery	CE–MS	Cluster of 86 peptides able to discriminate patients from controls, 80% sensitivity and 87% specificity. Validation on 1153 urine samples	[93]

↑: Increase; ↓: decrease; 12P-M-LAC: 12 high-abundance proteins and multi-lectin affinity chromatography; 2DICAL: 2-dimensional image converted analysis of liquid chromatography; A2G2S2: biantennary digalactosylated disialylated glycan; ccRCC: clear cell RCC; CE: capillary electrophoresis; DEPs: differentially expressed proteins; FA2G2S2: core fucosylated biantennary digalactosylated disialylated glycan; IDs: protein isoforms (Identities); IHC: immunhystochemistry; SCX: strong cation exchange; SELDI: surface enhanced laser desorption ionization; WB: Western blot; WCX: weak cation exchange.

UNIPROT IDs abbreviations: 1433Z: 14-3-3 zeta/delta ADH: alcohol dehydrogenase; ADT3: adenosine diphosphate (ADP)/adenosine-5′-triphosphate (ATP) translocase 3; AIFM1: apoptosis-inducing factor 1, mitochondrial precursor; AL1A1: aldehyde dehydrogenase family 1 member A1 retinal dehydrogenase 1; ALDOA: fructose-bisphosphate aldolase A; ALK: ALK tyrosine kinase receptor; AMACR: alphamethylacyl-CoA racemase; AMERL: ammecr1-like protein; AMPN: aminopeptidase N; ANXA2: annexin A2; ANXA4: annexin A4; ANXA5: annexin A5; APOA1: apolipoprotein A-1 precursor; APOB: apolipoprotein B-100; AQP1: Aquaporin-1; BASI: basigin (extracellular matrix metalloproteinase inducer) (EMMPRIN); Bcl-2: B-cell CLL/lymphoma 2; BGH3: transforming growth factor, beta-induced; C1QB: complement C1q subcomponent subunit B; C1QC: complement C1q subcomponent subunit C; C1R: complement C1r subcomponent; CAD11: cadherin-11; CADH5: cadherin-5; CAH2: carbonic anhydrase 2; CAH9: carbonic anhydrase 9; CALB1: calbindin; CAPG: macrophage-capping protein; Cav1:caveolin-1; CD2AP: CD2-associated protein; CERU: ceruloplasmin; CH10: 10 kDa heat shock protein; CK054: ester hydrolase C11orf54; CLUS: clusterin; CO2: complement C2; CO4A: complement C4-A; COF1: cofilin-1; COR1A: coronin 1A; CRP: c-reactive protein; CRYAB: alpha-crystalline B chain; DDX23: DEAD-box protein-23; DKK4: Dickkopf-related protein 4; DPEP1: dipeptidase 1; ECM1: extracellular matrix protein 1; ENOA: alpha-enolase; ENOG: gamma-enolase; ESRRA: estrogen related receptor alpha; ESRRG: estrogen related receptor gamma; FABP7: fatty-acid-binding protein 7; FABPH: fatty-acid-binding protein, heart; FETUA: alpha-2-HS-glycoprotein; FINC: fibronectin; Flot1: Flotillin-1; FOXA1: forkhead box A1; FRIL: ferritin light chain; GELS: gelsolin; GSTA1: glutathione S-transferase alpha 1; GSTP1: glutathione S-transferase P; H2AX: histone H2AX; HIF1a: hypoxia inducible domain family member 1A; HSP7C: heat shock cognate 71 kDa protein; HSPB1: heat shock protein beta-1; Thy1: thy-1 membrane glycoprotein; IF5A1: eukaryotic translation initiation factor 5A-1; ITIH4: inter-alpha-trypsin inhibitor heavy chain H4; KLKB1: plasma kallikrein; KPYM: pyruvate kinase PKM; LDHA: L-lactate dehydrogenase A chain; LEG1: galectin-1; LG3BP: galectin-3-binding protein; LUM: lumican; LYAM1: L-selectin; LYRM5: leucine-tyrosine-arginine (LYR) motif-containing protein 5; MA2B1: lysosomal alpha-mannosidase; MAPK1: mitogen-activated protein kinase 1; MASP2: mannan-binding lectin serine protease 2; MIC60: MICOS complex subunit MIC60; MMP9: matrix metalloproteinase-9; MSH2: MutS homolog 2; MUC18: cell surface glycoprotein MUC18; MVP: major vault protein/lung resistance-related protein; NDKA: nucleoside diphosphate kinase A; NDRG1: N-myc downstream-regulated gene 1 protein; NEP: neprilysin; NNMT: nicotinamide *N*-methyltransferase; NPM: nucleophosmin; OSTP: osteopontin; PD-ECGF: platelet-derived endothelial cell growth factor; PGBM: basement membrane-specific heparan sulfate proteoglycan core protein; PGK1: posphoglycerate kinase 1; PHB: prohibitin; PLIN2: adipose differentiation-related protein; PLSL: plastin-2; PODXL: podocalyxin; PPIA: cyclophilin; APRDX2: peroxiredoxin 2; PRDX3: thioredoxin-dependent peroxide reductase; PRDX4: peroxiredoxin-4; PRDX6: peroxiredoxin-6; PROF1: profilin-1; RBBP7: histone-binding protein RBBP7; RBP6: RNA-binding protein 6; RCN1: reticulocalbin-1; RD23B: UV excision repair protein RAD23B; S10A8: protein S100-A8; S10A9: protein S100-A9; S10AA: protein S100-A10; S10AB: protein S100-A11; SCTM1: secreted and transmembrane protein 1; SDCB1: syntenin-1; SEGN: secretagogin; SERPH: serpinH1/HSP47; SERPING1: complement component 1 inhibitor; SODC: ZnCu-superoxide dismutase; SYNE1: nesprin-1; SODM: superoxide dismutase [Mn], mitochondrial; TBA1B: tubulin alpha-1B chain; TGM2: protein-glutamine gamma-glutamyltransferase 2; THIL: acetyl-CoA acetyltransferase, mitochondrial; TPIS: triosephosphate isomerase 1; TUBB: tubulin beta chain; TYPH: thymidine phosphorylase; UROM: uromodulin; VCAM1: vascular cell adhesion protein 1; VDAC1: voltage-dependent anion-selective channel protein 1; VIME: vimentin; VINC: vinculin; WISP2: WNT1 inducible signaling pathway protein 2; YBOX1: Y-box-binding protein 1; ZFP3: zinc finger protein 3.

The study of the RCC urinary proteome has been attempted by employing a variety of different approaches that range from the traditional proteomic analysis tools, such as 2DE, followed by MALDI–MS or LC–MS/MS, different types of immunoassays, to the most recent qualitative and quantitative LC–MS/MS strategies. These works were already comprehensively discussed in a recent review, thus they will not be further described here [101]. The emerging technologies have advanced continuously and are now trying to maximize urinary proteome coverage in ever-decreasing sample amounts and to reduce the variability. Thus, the main challenge in this field is likely to rely on the cooperation among clinicians, analytical chemists, and bioinformaticians [78]. Standardization of the analysis, sample collection, metadata elaboration, and the application of effective quality control (targeted proteomics, antibody-based validation, etc.) could improve the reproducibility and the reliability of the results, allowing studies with much wider sample sizes [79].

4.3. Serum and urine peptidome

Biomarkers are often low molecular weight proteins secreted into the bloodstream by mutated apoptotic or necrotic cells deriving from disease-affected tissues [94]. Indeed, rather than simply being protein degradation products, endogenous peptides could also reflect the state of a cell and its relationship to its surrounding environment. Generally, they are not synthesized as such but rather are the results of specific or nonspecific proteolytic cleavage by endogenous peptidases, which in turn are differentially regulated in the context of many physiological and pathologic phenomena [98]. Thus, the alteration observed at peptidome level is often not reflected at proteome level, as well as vice versa. However, the comprehensive study of peptides and small proteins (<20kDa), known as 'peptidomics,' represents an attractive alternative technological approach, allowing assumptions on stage/progression, severity, and prognosis of a wide variety of pathologies, including the one evaluated in this review. For the RCC

peptidome, MALDI combined with pre-fractionation using functionalized materials (e.g. magnetic beads [MB] [100], or surface enhanced laser desorption ionization [SELDI]) or other clean-up systems, as well as high performance separation such as LC, or CE, have all been coupled to MS for this particular line of investigation.

Only a small number of studies based on MALDI profiling have been reported thus far, three performed on serum and the others on urine (Table 2). Concerning serum evaluation, Liu et al. first proposed the use of MB combined with MALDI-TOF in order to search for a diagnostic model of renal cancer [101]. The comparison of weak cation exchange-magnetic beads (WCX-MB) protein profiles obtained from the serum of 62 RCC patients and 37 patients with benign lesions pinpointed a discriminant cluster of four signals with specificity and sensitivity above 80%. However, none of the signals was identified. One year later, Gianazza et al. applied MALDI-TOF analysis to RPC18-MB purified serum of a large cohort of ccRCC, non-ccRCC patients, and controls [102]. They were able to differentiate malignant tumors (n = 102) from benign renal masses and healthy subjects (n = 104) using a cluster of five signals with a high diagnostic performance both in the training and in the test phase. Furthermore, the power of the MALDI-profiling approach was confirmed by label-free LC–MS/MS quantification, showing no statistical difference between the peptide ratios obtained by MALDI with those obtained from ESI. Some of the signals included in the cluster were identified, among which serum deprivation response protein and ZYX (Zyxin) were found to be downregulated, while serglycin and tymosin β-4-like protein were over-concentrated in patients. More recently, using WCX-MB, Yang et al. investigated the serum peptidome of ccRCC patients (n = 58) and healthy donors (n = 64), comparing matched pre and postoperative ccRCC patients (n = 40) [84]. Three candidate peaks, whose concentration increased in preoperative patients and tended to return to the level of healthy control values after surgery, were recognized as fragments of RNA-binding protein-6, tubulin-beta-chain, and zinc finger protein-3.

The application of MALDI profiling via functionalized MB on urine samples, for the study of RCC, was first reported by Bosso et al. [85]. Their pilot study confirmed the feasibility of this approach, finding a cluster of three ions able to distinguish RCC patients (n = 29) from healthy subjects (n = 29) with high values of diagnostic efficacy, especially for patients in the first stage. One of these signals was identified as a fragment of uromodulin (UROM), which is specifically decreased in RCC urine, suggesting the implication of specific endoproteases, probably secreted by tumor cells, in determining the different RCC urinary peptidome profiles. Later, Chinello et al. utilized the same analytical workflow to study a larger patient dataset that comprised not only ccRCC patients and healthy donors but also patients affected by benign kidney lesions or nonccRCC kidney malignancies [86]. They proposed a classifier that differentiates malignant tumors (n = 137) from benign renal masses and controls (n = 153) with a sensitivity of 76% and a specificity of 87%, and a second cluster of 12 signals for the discrimination of ccRCC (n = 118) from controls (n = 137) (84% sensitivity and 91% specificity). The identity of several of these signals was assessed by LC–MS/MS. Only a few of them were assigned to highly abundant proteins such as UROM and fibrinogen α chain, while the majority were derived from proteins strongly correlated with tumor development and progression, suggesting an effective reduction of the dynamic range by the activated surface. One year later, the same group, exploring the C8RP-MB MALDI urinary profiles, showed several peptides whose urinary abundance varied according to tumor size, stage, and grade, thus serving as potential indicators of RCC progression [90].

A variation of the MALDI technology is SELDI, in which the sample is crystallized on a surface that can be selected based on its chemical structure and/or class of protein binding. Despite initial enthusiasm and several promising studies, this technique revealed evident limitations, which drastically reduced its use. However, several studies focusing on RCC, using SELDI profiling, were performed. The two most recent papers

were conducted on urine [91] and on serum [92]. In both, the WCX protein chip systems were used for identifying differentially expressed proteins/ peptides between the RCC group and controls. In the study of Alves et al., urine samples from ccRCC (n = 53) and Papillary RCC (n = 8) were compared with 29 samples of healthy control urines [91]. Twenty-two peaks were found to be statistically down-represented in the patient group, even if no data regarding the diagnostic performance were provided. Recently, Nuerrula et al. analyzed the serum of 89 patients with ccRCC (before and after surgery) and 100 controls including healthy volunteers and patients without RCC by SELDI-TOF MS [92]. They observed a significant difference in the expression levels of five protein signals related to the preoperative group and controls, and a statistically significant alteration of postoperative expression levels for each single follow-up compared to the preoperative ones. The discriminant molecules enabled a high performance in predicting ccRCC to be obtained (sensitivity of 89% and specificity of 91%).

Alterations in the peptidome associated with RCC was also studied by coupling the efficient separation of CE with the selective detection of MS. Using this technique, Frantzi et al. built a diagnostic model for RCC detection based on naturally occurring urinary peptides in very large cohorts of RCC patients and controls (healthy and not healthy) [89]. Of the 86 signals included in the classifier, 40 were assigned to specific sequences, 70% of which belonged to fragments of collagen chains. The identity of these fragments seems to poorly overlap with those found with MALDI profiling [92], but this is not completely unexpected given the fact that the peptidomic data were provided following different pre-fractionation steps and different chromatographic separation. Finally, LC–MS/MS could also be used to obtain a peptidomic profile from biological fluids [83]. In a recent study of Huang et al., this strategy was used to explore the peptidome alterations specific for RCC patient serum [93]. These authors, through a supervised holistic-variate orthogonal partial least-squares-discriminant analysis, were able to find significant differences in serum levels between the group of patients (n =

30) and healthy controls (n = 30), with 100% sensitivity and 93.3% specificity for RCC detection and excellent predictive power by the receiver-operating characteristic curve. In conclusion, as already shown [103], for both blood and urine specimens, available data are not currently translated into routine clinical application, despite the enormous advancements that have been made to the analytical methodologies, leading to improvements in specificity, dynamic range, and reproducibility. The reason for this delay is mainly the same as for other solid tumors. In fact, when screening and validation of candidate biomarkers is performed, cases and controls of similar sample size are usually compared, and the low prevalence of disease is often ignored. Moreover, it is suggested that multiple, rather than single biomarkers, should be more commonly relied upon, but the link between the markers and the underlying molecular events is not often reconstructed. In addition, the heterogeneity of disease tissue and histological differences inside the specimen are hardly taken into account, and tissue is generally homogenized for protein extraction. Finally, the issue of pre-analytical variability is also associated with the biomarker discovery pipeline and has to be addressed.

4.4. Extracellular vesicles

In the last decade, proteomic research on urinary EVs has emerged as a new frontier in the landscape of urine-biomarker discovery. The term EV comprises many different forms of secreted vesicles that are classified into two general types depending on the site of cellular production and the mode of vesicular secretion. The first includes EVs directly shed from the plasma membrane, and the second contains those originating from the inward budding of intercellular endosomes and are defined as exosomes [87]. Studies addressing the role of EVs in RCC have been mainly published in the last 3–4 years. They were generally performed by adopting RCC cell culture models, in order to analyze the effects of EV released in the culture medium on other cells, pointing out the role of some specific proteins, rather than performing protein identification and

quantification. Most of these studies revealed that the proteomic cargo of exosomes released by RCC cells is able to induce effects in three main different settings. The first is exerted on the cancer cells themselves, by cross-delivery: for example, it has been reported that RCCreleased EV determine increased levels of two important cancer-related proteins, such as chemokine receptor type 4 and matrix metalloproteinase-9 (MMP9), in other RCC cells, following EV internalization [104]. The second would act on alterations of the cancer microenvironment, both of endothelial cells, where EV seem to drive the expression of angiogenetic features [105], and of the immune system effectors, suggesting that they may be partly responsible for the well-known immune escape features of RCC *in vivo* [106,107]. A third line of research tackled the issue of the influence of human renal cancer stem cells. In fact, they were shown to secrete large amounts of EV that can transfer several molecules (proteins, lipids, and nucleic acids) and induce epigenetic changes in target cells, such as tumor-surrounding mesenchymal stromal cells [108].

Among EV, studying the proteome of the urinary exosomes (UE) seems to be of particular interest, because they are believed to provide a full molecular representation of the entire urinary system, reflecting the state of renal epithelial cells from which they are actively released. Accordingly, these nanovesicles (30–100 nm) could well represent alterations to the molecular composition of RCC cells. In fact, a comparative lipidomic analysis of UE performed by a hyphenated microLC– Q-TOF-MS platform highlighted a differential lipid composition in exosomes prepared from RCC patient urine, compared with controls, providing initial evidence of a relationship between lipid composition of UE and RCC disease [109]. Raimondo et al. performed the only proteomic profiling investigation of RCC UE to date and reported that the protein profile of RCC UE was quite different from that of controls [110]. In fact, approximately 50% of the proteins identified by 1DE and LC–MS/MS were not shared between RCC and control UE. In particular, the RCC UE differential content of some proteins, such

as MMP-9, Podocalyxin, and CAH9, was validated by immunoblotting and was suggested to represent a basis for the setup of a multimarker strategy in UE for RCC detection.

5. Expert commentary

Indications for taking a 'watch and wait' approach in small renal kidney cancers are expanding constantly and it would be strongly desirable to have a biological marker that could evaluate the aggressiveness of the tumor in order to decide whether or not to treat patients. The American Urological Association guidelines state that there is a strong need for markers which can predict aggressiveness of small renal masses in order to tailor treatment and follow-up of the affected patients [88]. The finding of proteomic markers in blood or urine might have a crucial impact, as it would mean having a non invasive test that could guide clinical decisions. Furthermore, if the proteomic markers could be related to tumor staging, they may also guide the choice between performing radical and nephron sparing surgery. It should allow the oncological risk of a conservative surgery to be evaluated prior to performing it, rather than the current approach that only relies on using the final pathologybased report postsurgery.

The research for RCC biomarkers using proteomics could greatly benefit from the recognition of common targets, revealed by experiments conducted through different platforms and starting from different kinds of samples. The recent availability of very sophisticated and integrated proteomic analytic approaches should ease this discovery, as well as the possible sharing of samples and/or databases coming from different cohorts. However, until now, this information has been rather limited. We compared the differential proteins reported by four publications to highlight the difficulties faced in identifying biomarkers of renal cancer [23,28,29,33] (Figure 3).

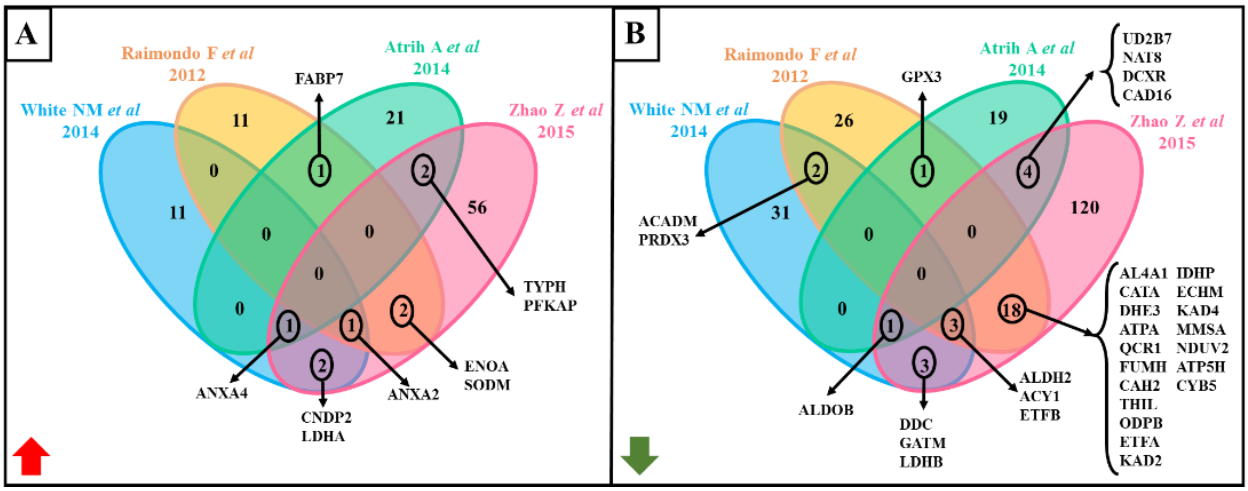


Figure 3. Venn diagram showing the number of shared protein species among four studies [23,25,28,29] reporting the differential proteomes of fresh frozen samples of RCC compared to ANK. **A)** common up-regulated proteins **B)** common down-regulated proteins. Uniprot IDs of the proteins that are present in at least two datasets are visualized.

A lack of concordance between studies was noted, with a maximum of 2 upregulated and 18 downregulated proteins in common between two studies. Obviously, this comparison is not to be considered as conclusive, but we believe that it is sufficiently representative of the actual situation. We believe that this lack of concordance between studies may derive from differences in the type of specimen (frozen, FFPE, microdissected/heterogeneous tissue), the pre-fractionation used to remove high-abundance proteins, or the enrichment used for low abundance ones.

On the other hand, even if we consider differential protein species reported in studies performed on different starting samples, only a small number of proteins is reported in more than a single paper. Taking into account that we only reviewed papers from the last 5–6 years, we found a recurrence of the family of HSP27 and CAH9, some isoforms of the glycolytic enzymes (LDHA, ENOA), ceruloplasmin, and some members of the fatty-acid-binding proteins. These possible 'shared RCC targets' were also identified in older papers and were found to be associated with RCC not only by proteomic studies but also by IHC, or as dysregulated transcript, or after gene microarray surveys. However, the relatively few biomarkers identified by proteomics are not specific for

RCC. Possible explanations for this biomarker bottleneck are (a) tissue heterogeneity and a lack of microdissection to enrich tumor cell populations prior to proteomic analysis, (b) lack of enrichment strategies to identify low abundance proteins, (c) the difficulty in measuring posttranslationally modified proteins, and (d) the possibility that different tumor types share similar high-abundant protein profiles.

It is common belief that, in general, the omics sciences failed to succeed in finding reliable markers for cancers, but in the case of RCC, a possible, more specific reason could be found in very recent observations [111]. The authors tried to integrate the TCGA transcriptomic data, available in public repositories, with the metabolomic data coming from their experiments, and demonstrated a peculiar discordance between the transcriptome and metabolome, depicting ccRCC as an outlier [112]. Strikingly, in ccRCC compared with ANK, genes and proteins driving metabolic pathways typically displayed a downregulation, while related metabolites were overexpressed. This pattern was peculiar for RCC and not evident in other tumor types, underlining a specificity for RCC, which will have to be taken into account in future studies.

6. Five-year view

In the past, investigations in RCC were mainly focused on biomarker discovery for diagnostic purposes. Nowadays, these studies are moving toward the search for indicators of disease progression that are of use to the 'watch and wait' clinical approach. However, another aspect that will be more deeply investigated in the coming years will regard markers for tailored RCC therapies. Moreover, MS data could be used to identify biomarker(s) that are predictors of response to these targeted agents. For example, analysis of RCC tumor tissue after treatment with erlotinib and bevacizumab revealed an association between poor treatment response and low AMPK expression, while it has been shown that Vorinostat enhances the activity of temsirolimus in RCC through

the suppression of survivin levels (reviewed in [113]). However, the global impact of mTOR inhibitors on RCC metabolism is currently unknown. In this respect, proteomics will play an important role due to its ability to provide an overview of the proteomic alterations that are of use in a panel of markers, taking advantage of recently optimized protocols. In fact, a recent paper showed that LC–MS analysis enabled more than 3800 proteins to be detected in FF tissue, and 3300 in FFPE, from a small amount of material obtained by human colon mucosal biopsy (sections of 1–2 mm^3) and simultaneously quantified them in a label-free manner with a very high consistency grade (about 90%) between FF and FFPE specimens [114]. Moreover, several PTMs could be also detected using ProteinPilot software that included more than 300 modifications. Furthermore, the quantification of their levels could also be done with high accuracy by MRM.

In addition, system biology is now going into the 'adult age' and the integration of the results derived from different 'omics' approaches has become more realistic, providing the opportunity to exploit the relationships between the different domains of biology. In particular, cancer proteogenomics is an emerging field of research that aims at the identification and quantification of protein sequence changes that are associated with the cancer-specific variations encountered in the genome. For example, it was recently reported that proteomic, and in particular phospho-proteomic analysis, of breast cancer samples could elucidate the functional consequences of cancer- associated somatic mutations and thus identify 'druggable' kinases as therapeutic targets beyond HER2 [115]. Unraveling the connection between cancer-related genomic variations and their effect on phenotypes and clinical behavior would be particularly useful in the case of RCC: in fact, the classification of the different RCC types, already based on strong genomic data, can further take advantage from proteomics, allowing better prediction of their natural history and progression. Moreover, these alterations could represent a target for new therapeutic interventions. These studies will benefit from the rapid development of MS instruments with new features such as ion mobility, which allows

molecules with same molecular mass but different conformational shape to be identified, or native-MS that enables protein–protein interactions to be investigated. Arguably, the possibility to perform proteomic investigations directly on tissue, thus conserving information about the spatial localization of the proteins within the specimen through MSI technology, represents an even more stimulating approach. In this respect, performing MSI analysis with a spatial resolution at the single cell level or the direct combination of the optical image of the tissue and its molecular analysis is now possible.

Key issues

- The most widely used proteomic approach is based on the coupling of liquid-chromatography to mass spectrometry. Bottom-up is now the most commonly used strategy based on enzymatic digestion that allows the identification of differentially expressed proteins to be obtained, while it is suboptimal for PTMs studies. This methodology is now integrated with the top-down approach in which intact proteins are analysed, along with their primary structures and PTMs.
- The identification of pathways involved in the genesis and progression of RCC will lead to the development of new therapeutic strategies based on molecular targets, such as those of the von Hippel-Lindau (VHL) pathway, including vascular endothelial growth factor (VEGF), platelet-derived growth factor (PDGF), and mammalian target of rapamycin (mTOR).
- The majority of investigations have been carried out in order to discover biomarkers with a diagnostic role. However, the emerging field of interesting research is related to the discovery of reliable prognostic and predictive markers that are able to predict the course of the disease and the response to specific drug treatments through the analysis of a panel of proteins.

- In addition to fresh frozen tissue, formalin fixed paraffin embedded (FFPE) specimens can also be used for proteomic investigations. Cell lines, biological fluids, and extracellular vesicles are also good specimens for diagnostic and prognostic biomarkers discovery, in particular urine and exosomes.
- Urinary exosomes, a particular type of extracellular vesicles, are widely investigated not only as possible source of markers but also for their ability to deliver crucial signals from cell to cell, such as the promotion of growth, proliferation, and inhibition of apoptosis. They can influence not only the cancer cells themselves but also the tumor microenvironment and thus contribute to renal cancer development and progression.
- Various promising marker studies targeted to both the peptidome and proteome, in blood and in urine, have been reported. For both specimens, available data is not currently translated into routine clinical application. Largescale studies are needed to validate these molecules as cancer biomarkers.
- Urine could represent a very easily-accessible specimen for finding RCC related signatures both at the peptidome and proteome level. Indeed, the protein content is quite stable and carries direct information coming from not only from the kidneys and the urogenital tract, but also from the blood stream.

Funding

The authors were supported by Ministero dell'Istruzione, dell'Università e della Ricerca (RBRN07BMCT_11), FAR 2012-2015, and in part by Fondazione Gigi & Pupa Ferrari Onlus and by FP7-PEOPLE-2013-ITN- 608332.

Declaration of interest

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. CA Cancer J Clin. 2013;63(1):11–30.

2. Itsumi M, Tatsugami K. Immunotherapy for renal cell carcinoma. Clin Dev Immunol. 2010;2010:284581.

3. Motzer RJ, Agarwal N, Beard C, et al. Kidney cancer. J Natl Compr Canc Netw. 2011;9(9):960–977.

4. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer. 2013;49(6):1374–1403.

5. Krajewski KM, Giardino AA, Zukotynski K, et al. Imaging in renal cell carcinoma. Hematol Oncol Clin North Am. 2011;25(4):687–715.

6. Moch H. An overview of renal cell cancer: pathology and genetics. Semin Cancer Biol. 2013;23(1):3–9.

7. Rini BI, Campbell SC, Escudier B. Renal cell carcinoma. Lancet (London, England). 2009;373(9669):1119–1132.

8. Young AC, Craven RA, Cohen D, et al. Analysis of VHL gene alterations and their relationship to clinical parameters in sporadic conventional renal cell carcinoma. Clin Cancer Res. 2009;15(24):7582–7592.

9. Escudier B, Eisen T, Porta C, et al. Renal cell carcinoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2012;23(SUPPL. 7):iii49–56.

10. Pan Z, Grizzle W, Hameed O, et al. Significant variation of immunohistochemical marker expression in paired primary and metastatic clear cell renal cell carcinomas. Am J Clin Pathol. 2013;140 (3):410–418.
11. Truong LD, Shen SS. Immunohistochemical diagnosis of renal neoplasms. Arch Pathol Lab Med. 2011;135(1):92–109.
12. Bing Z, Lal P, Lu S, et al. Role of carbonic anhydrase IX, α-methylacyl coenzyme a racemase, cytokeratin 7, and galectin-3 in the evaluation of renal neoplasms: a tissue microarray immunohistochemical study. Ann Diagn Pathol. 2013;17(1):58–62.
13. Klatte T, Streubel B, Wrba F, et al. Renal cell carcinoma associated with transcription factor E3 expression and Xp11. 2 translocation incidence, characteristics, and prognosis. Am J Clin Pathol. 2012;137(5):761–768.
14. Moltrasio F, Brenna A, Bovo G, et al. Pathological features of Xp11 translocation renal cell carcinoma using urine liquid-based cytology with FISH. Cytopathol Off J Br Soc Clin Cytol. 2015;26(5):325–328.
15. Nagaraj N, Mann M. Quantitative analysis of the intra- and interindividual variability of the normal urinary proteome research articles. J Proteome Res. 2011;10(2):637–645.
16. Lichtenfels R, Dressler SP, Zobawa M, et al. Systematic comparative protein expression profiling of clear cell renal cell carcinoma: a pilot study based on the separation of tissue specimens by twodimensional gel electrophoresis. Mol Cell Proteomics. 2009;8 (12):2827–2842.
17. Sun CY, Zang YC, San YX, et al. Proteomic analysis of clear cell renal cell carcinoma. Saudi Med J. 2010;86(512):525–532.
18. Kim DS, Choi YP, Kang S, et al. Panel of candidate biomarkers for renal cell carcinoma. J Proteome Res. 2010;9(7):3710–3719.
19. Giribaldi G, Barbero G, Mandili G, et al. ScienceDirect Proteomic identification of Reticulocalbin 1 as potential tumor marker in renal cell carcinoma. J Proteomics. 2013;91:385–392.

20. Lu Z, Yao Y, Song Q, et al. Metabolism-related enzyme alterations identified by proteomic analysis in human renal cell carcinoma. Onco Targets Ther. 2016;(6):1327–1337.

21. Siu KWM, Desouza LV, Scorilas A, et al. Differential protein expressions in renal cell carcinoma: new biomarker discovery by mass spectrometry research articles. J Proteome Res. 2009;8(8):3797–3807.

22. Masui O, White NMA, DeSouza LV, et al. Quantitative proteomic analysis in metastatic renal cell carcinoma reveals a unique set of proteins with potential prognostic significance. Mol Cell Proteomics MCP. 2013;12(1):132–144.

23. White NMA, Masui O, Desouza LV, et al. Quantitative proteomic analysis reveals potential diagnostic markers and pathways involved in pathogenesis of renal cell carcinoma. Oncotarget. 2014;5(2):506–518.

•• *Serum proteomic analysis by iTRAQ labeling and LC–MS/MS: a new insight on the molecular pathways that are altered and could contribute to the pathogenesis of cc*RCC.

24. Johann DJJ, Wei B, Prieto DA, et al. Combined blood/tissue analysis for cancer biomarker discovery: application to renal cell carcinoma. Anal Chem. 2010;82(5):1584–1588.

25. Raimondo F, Salemi C, Chinello C, et al. Proteomic analysis in clear cell renal cell carcinoma: identification of differentially expressed protein by 2-D DIGE. Mol Biosyst. 2012;8(4):1040–1051.

26. Oppenheimer SR, Mi D, Sanders ME, et al. Molecular analysis of tumor margins by MALDI mass spectrometry in renal carcinoma. J Proteome Res. 2010;9(5):2182–2190.

27. Jones EE, Powers TW, Neely BA, et al. MALDI imaging mass spectrometry profiling of proteins and lipids in clear cell renal cell carcinoma. Proteomics. 2014;14(7–8):924–935.

•• *MALDI-imaging approach applied to investigate proteins and lipids on the same tissue section. They found panels of proteins and lipid species able to distinguish tumor from nontumor and recurrent disease progressors from nonrecurrent disease patients.*

28. Atrih A, Mudaliar MAV, Zakikhani P, et al. Quantitative proteomics in resected renal cancer tissue for biomarker discovery and profiling. Br J Cancer. 2014;110(6):1622–1633.

29. Zhao Z, Wu F, Ding S, et al. Label-free quantitative proteomic analysis reveals potential biomarkers and pathways in renal cell carcinoma. Tumour Biol J Int Soc Oncodevelopmental Biol Med. 2015;36(324):939–951.

30. Neely BA, Wilkins CE, Marlow LA, et al. Proteotranscriptomic analysis reveals stage specific changes in the molecular landscape of clear-cell renal cell carcinoma. PLoS One. 2016;11(4):e0154074.

•• *A comprehensive proteomic analysis of ccRCC tissues grouped by histopathology stages along with proteotranscriptomic analysis. Evaluation of stage-dependent molecular changes that reflect the ccRCC progression and intratumor heterogeneity.*

31. Guo T, Kouvonen P, Koh CC, et al. Europe PMC Funders Group Rapid mass spectrometric conversion of tissue biopsy samples into permanent quantitative digital proteome maps. Nat Med. 2015;21 (4):407–413.

32. Wang Y, Yue D, Xiao M, et al. C1QBP negatively regulates the activation of oncoprotein YBX1 in the renal cell carcinoma as revealed by interactomics analysis. J Proteome Res. 2015;14 (2):804–813.

33. Raimondo F, Morosi L, Chinello C, et al. Molecular BioSystems protein profiling of microdomains purified from renal cell carcinoma and normal kidney tissue samples w z. Mol Biosyst. 2012;8 (4):1007–1016.

34. Raimondo F, Corbetta S, Savoia A, et al. Molecular BioSystems comparative membrane proteomics: a technical advancement in the search of renal cell carcinoma biomarkers. Mol Biosyst. 2015;11:1708–1716.

35. Valera VA, Li-Ning-T E, Walter BA, et al. Protein expression profiling in the spectrum of renal cell carcinomas. J Cancer. 2010;1 (1):184–196.
36. Arai E, Gotoh M, Tian Y, et al. Alterations of the spindle checkpoint pathway in clinicopathologically aggressive CpG island methylator phenotype clear cell renal cell carcinomas. Int J Cancer. 2015;137 (11):2589–2606.
37. Mariño-Enríquez A, Ou W-B, Weldon CB, et al. ALK rearrangement in sickle cell trait-associated renal medullary carcinoma. Genes Chromosomes Cancer. 2011;50(3):146–153.
38. Na CH, Hong JH, Kim WS, et al. Identification of protein markers specific for papillary renal cell carcinoma using imaging mass spectrometry. Mol Cells. 2015;38(7):624–629.
39. Junker H, Venz S, Zimmermann U, et al. Stage-related alterations in renal cell carcinoma – comprehensive quantitative analysis by 2DDIGE and protein network analysis. PLoS One. 2011;6(7):1–13.
40. Ho PY, Chueh SC, Chiou SH, et al. alphaB-Crystallin in clear cell renal cell carcinoma: tumor progression and prognostic significance. Urol Oncol. 2013;31:1367–1377.
41. Hosoya N, Sakumoto M, Nakamura Y, et al. Proteomics identified nuclear N-myc downstream-regulated gene 1 as a prognostic tissue biomarker candidate in renal cell carcinoma. Biochim Biophys Acta Proteins Proteomics. 2013;1834(12):2630–2639.
42. Lebdai S, Verhoest G, et al. Identi fi cation and validation of TGFBI as a promising prognosis marker of clear cell renal cell carcinoma. Urol Oncol Semin Orig Investig. 2015;33(2):69.e11–69.e18.
43. Perroud B, Ishimaru T, Borowsky AD, et al. Grade-dependent proteomics characterization of kidney cancer. Mol Cell Proteomics MCP. 2009;8(5):971–985.

44. Sprung RW, Martinez MA, Carpenter KL, et al. Precision of multiple reaction monitoring mass spectrometry analysis of formalin-fixed, paraffin-embedded tissue. J Proteome Res. 2012;11(6):3498–3505.
45. Craven RA, Cairns DA, Zougman A, et al. Proteomic analysis of formalin-fixed paraffin-embedded renal tissue samples by labelfree MS: assessment of overall technical variability and the impact of block age. Proteomics Clin Appl. 2013;7(3–4):273–282.
46. Weißer J, Lai ZW, Bronsert P, et al. Quantitative proteomic analysis of formalin–fixed, paraffin–embedded clear cell renal cell carcinoma tissue using stable isotopic dimethylation of primary amines. BMC Genomics. 2015;16(1):559.
47. Morgan TM, Seeley EH, Fadare O, et al. Imaging the clear cell renal cell carcinoma proteome. J Urol. 2013;189(3):1097–1103.
48. Steurer S, Seddiqi AS, Singer JM, et al. MALDI imaging on tissue microarrays identifies molecular features associated with renal cell cancer phenotype. Anticancer Res. 2014;34(5):2255–2262.
49. Lu J, Tan M, Cai Q. The Warburg effect in tumor progression: mitochondrial oxidative metabolism as an anti-metastasis mechanism. Cancer Lett. 2015;356(2):156–164.
50. De Sio G, Smith AJ, Galli M, et al. A MALDI-mass spectrometry imaging method applicable to different formalin-fixed. Mol Biosyst. 2015;11:1507–1514.
51. Lalowski M, Magni F, Mainini V, et al. Imaging mass spectrometry: A new tool for kidney disease investigations. Nephrol Dial Transpl. 2013;28(7):1648–1656.
52. Magni F, Lalowski M, Mainini V, et al. Proteomics imaging and the kidney. J Nephrol. 2013;26(3):430–436.
53. Craven RA, Hanrahan S, Totty N, et al. Proteomic identification of a role for the von Hippel Lindau tumour suppressor in changes in the expression of mitochondrial proteins and septin 2 in renal cell carcinoma. Proteomics. 2006;6(13):3880–3893.

54. Seliger B, Fedorushchenko A, Brenner W, et al. Ubiquitin COOH terminal hydrolase 1: A biomarker of renal cell carcinoma associated with enhanced tumor cell proliferation and migration. Clin Cancer Res. 2007;13(1):27–37.

55. Aggelis V, Craven RA, Peng J, et al. Proteomic identification of differentially expressed plasma membrane proteins in renal cell carcinoma by stable isotope labelling of a von Hippel-Lindau transfectant cell line model. Proteomics. 2009;9(8):2118–2130.

56. Boysen G, Bausch-Fluck D, Thoma CR, et al. Identification and functional characterization of pVHL-dependent cell surface proteins in renal cell carcinoma. Neoplasia. 2012;14(6):535–546.

57. Da Nagaprashantha L, Talamantes T, Singhal J, et al. Proteomic analysis of signaling network regulation in renal cell carcinomas with differential hypoxia-inducible factor-2α expression. PLoS One. 2013;8(8):e71654.

58. Leisz S, Schulz K, Erb S, et al. Distinct von Hippel-Lindau gene and hypoxia-regulated alterations in gene and protein expression patterns of renal cell carcinoma and their effects on metabolism. Oncotarget. 2015;6(13):11395–11406.

59. Malec V, Coulson JM, Urbé S, et al. Combined analyses of the VHL and hypoxia signaling axes in an isogenic pairing of renal clear cell carcinoma cells. J Proteome Res. 2015;14(12):5263–5272.

60. Dihazi H, Müller GA, Lindner S, et al. Characterization of diabetic nephropathy by urinary proteomic analysis: identification of a processed ubiquitin form as a differentially excreted protein in diabetic nephropathy patients. Clin Chem. 2007;53(9):1636–1645.

61. Yang J, Li A, Yang Y, et al. Identification of cyclophilin A as a potential prognostic factor for clear-cell renal cell carcinoma by comparative proteomic analysis. Cancer Biol Ther. 2011;11(5):535– 546.

62. Yang JIN, Liu P, Tian M, et al. Proteomic identification of angiomotin by ProteomeLab PF-2D and correlation with clinical outcome in human clear cell renal cell carcinoma. Int J Oncol. 2013;42:2078–2086.

63. Tanaka T, Kuramitsu Y, Wang Y, et al. Glyoxalase 1 as a candidate for indicating the metastatic potential of SN12C human renal cell carcinoma cell clones. Oncol Rep. 2013;30(5):2365–2370.
64. Minamida S, Iwamura M, Kodera Y, et al. Profilin 1 overexpression in renal cell carcinoma. Int J Urol Off J Japanese Urol Assoc. 2011;18 (1):63–71.
65. Mesri M, Birse C, Heidbrink J, et al. Identification and characterization of angiogenesis targets through proteomic profiling of endothelial cells in human cancer tissues. PLoS One. 2013;8(11): e7885.
66. Vasko R, Mueller GA, Von JA, et al. Impact of cisplatin administration on protein expression levels in renal cell carcinoma: A proteomic analysis. Eur J Pharmacol. 2011;670(1):50–57.
67. Trivedi R, Dihazi GH, Eltoweissy M, et al. The antioxidant protein PARK7 plays an important role in cell resistance to Cisplatin induced apoptosis in case of clear cell renal cell carcinoma. Eur J Pharmacol. 2016;784:99–110.
68. Kearney AY, Fan Y-H, Giri U, et al. 8-Chloroadenosine sensitivity in renal cell carcinoma is associated with AMPK activation and mTOR pathway inhibition. PLoS One. 2015;10(8):e0135962.
69. Yim NH, Jung YP, Kim A, et al. Induction of apoptotic cell death by betulin in multidrug-resistant human renal carcinoma cells. Oncol Rep. 2015;34(2):1058–1064.
70. Fan S, Li X, Tie L, et al. KIAA0101 is associated with human renal cell carcinoma proliferation and migration induced by erythropoietin. Oncotarget. 2016;7(12):13520–13537.
71. Seliger B, Dressler SP, Massa C, et al. Identification and characterization of human leukocyte antigen class I ligands in renal cell carcinoma cells. Proteomics. 2011;11(12):2528–2541.

72. Prattichizzo C, Gigante M, Pontrelli P, et al. Establishment and characterization of a highly immunogenic human renal carcinoma cell line. Int J Oncol. 2016;49(2):457–470.
73. Chen KC, Lin CM, Huang CJ et al. Dual Roles of 17-β Estradiol in Estrogen Receptor-dependent Growth Inhibition in Renal Cell Carcinoma. Cancer Genom Proteom. 2016;13(3):219–230.
74. Haake SM, Li J, Bai Y, et al. Tyrosine kinase signaling in clear cell and papillary renal cell carcinoma revealed by mass spectrometrybased phosphotyrosine proteomics. Clin Cancer Res. 2016. Doi 10.1158/1078-0432.CCR-15-1673.
75. Magni R, Espina BH, Liotta LA, et al. Hydrogel nanoparticle harvesting of plasma or urine for detecting low abundance proteins. J Vis Exp. 2014. doi 10.3791/51789.
76. Gbormittah FO, Bones J, Hincapie M, et al. Clusterin glycopeptide variant characterization reveals significant site-specific glycan changes in the plasma of clear cell renal cell carcinoma. J Proteome Res. 2015;14(6):2425–2436.
77. Gbormittah FO, Lee LY, Taylor K, et al. Comparative studies of the proteome, glycoproteome, and N-glycome of clear cell renal cell carcinoma plasma before and after curative nephrectomy. J Proteome Res. 2014;13(11):4889–4900.
78. Thomas S, Hao L, Ricke WA, et al. Biomarker discovery in mass spectrometry-based urinary proteomics. Proteomics Clin Appl. 2016;10:358–370.
79. Raimondo F, Corbetta S, Chinello C, et al. The urinary proteome and peptidome of renal cell carcinoma patients: a comparison of different techniques. Expert Rev Proteomics. 2014;11(4):503–514.
80. Geissler K, Fornara P, Lautenschläger C, et al. Immune signature of tumor infiltrating immune cells in renal cancer. Oncoimmunolog. 2015;4(1):e985082.
81. Yokomizo A, Takakura M, Kanai Y, et al. Use of quantitative shotgun proteomics to identify fi bronectin 1 as a potential plasma biomarker for clear cell carcinoma of the kidney. Cancer Biomark. 2011–2012;10(3–4):175–183.

82. Zhang L, Jiang H, Xu G, et al. Proteins S100A8 and S100A9 are potential biomarkers for renal cell carcinoma in the early stages: results from a proteomic study integrated with bioinformatics analysis. Mol Med Rep. 2015;11(6):4093–4100.

• *They profiled the differentially expressed serum proteins, by label-based quantitative LC MS–MS in order to better understand the etiology of cc*RCC *and even provide a new strategy for the early noninvasive detection of cc*RCC.

83. Mullen W, Albalat A, Gonzalez J, 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):567–574.

84. Yang J, Yang J, Gao Y, et al. Identification of potential serum proteomic biomarkers for clear cell renal cell carcinoma. PloS One. 2014;9(11):e111364.

85. Bosso N, Chinello C, Picozzi SCM, et al. Human urine biomarkers of renal cell carcinoma evaluated by ClinProt. Proteomics Clin Appl. 2008;2(7–8):1036–1046.

86. Chinello C, Cazzaniga M, De Sio G, et al. Urinary signatures of Renal Cell Carcinoma investigated by peptidomic approaches. PLoS One. 2014;9(9):e106684.

87. Witwer KW, Buzás EI, Bemis LT, et al. Standardization of sample collection, isolation and analysis methods in extracellular vesicle research. J Extracell Vesicles. 2013;2:1–25.

88. Donat SM, Diaz M, Bishoff JT, et al. Follow-up for clinically localized renal neoplasms: AUA guideline. J Urol. 2013;190(2):407–416.

89. Frantzi M, Metzger J, Banks RE, et al. Discovery and validation of urinary biomarkers for detection of renal cell carcinoma. J Proteomics. 2014;98:44–58.

•• *Building of a diagnostic model for based on naturally occurring urinary peptides in very large cohort of RCC patients and controls by CE–MS. The authors proposed a classifier with a good diagnostic capability c (80% sensitivity and 87% specificity).*

90. Chinello C, Cazzaniga M, De Sio G, et al. Tumor size, stage and grade alterations of urinary peptidome in RCC. J Transl Med. 2015;13(1):332.

• *C8 magnetic beads pre-purification and MALDI analysis was used to study the urinary peptidome on larger cohort of subjects, including ccRCC patients, healthy donors, and patients with benign kidney lesions or non-ccRCC renal malignancies. They proposed panels of peptides able to discriminate malignant tumors from benign renal masses.*

91. Alves G, Pereira DA, Sandim V, et al. Urine screening by Seldi-Tof, followed by biomarker identification, in a Brazilian cohort of patients with renal cell carcinoma (RCC). Int Braz J Urol. 2013;39 (2):228–239.

92. Nuerrula Y, Rexiati M, Liu QWYJ. Differential expression and clinical significance of serum protein among patients with clear-cell renal cell carcinoma. Cancer Biomark. 2015;15(4):485–491.

93. Huang Z, Zhang S, Hang W, et al. Liquid chromatography-mass spectrometry based serum peptidomic approach for renal clear cell carcinoma diagnosis. J Pharm Biomed Anal. 2014;100:175–183.

94. Rainer M, Sajdik C, Bonn GK. Mass spectrometric profiling of low molecular-weight proteins. Methods Mol Biol. 2013;1023:83–95.

95. Frew IJ, Moch H, Clearer A. View of the molecular complexity of clear cell renal cell carcinoma. Annu Rev Pathol. 2015;10:263–289.

96. Zhang Y, Cai Y, Yu H, et al. ITRAQ-based quantitative proteomic analysis identified HSC71 as a novel serum biomarker for renal cell carcinoma. Biomed Res Int. 2015. doi 10.1155/2015/802153.

97. Zhang L, Jiang H, Xu G, et al. iTRAQ-based quantitative proteomic analysis reveals potential early diagnostic markers of clear-cell Renal cell carcinoma. Biosci Trends. 2016;10(3):210–219.

98. Bauca JM, Martnez-Morillo E, Diamandis EP. Peptidomics of urine and other biofluids for cancer diagnostics. Clin Chem. 2014;60 (8):1052–1061.

99. Sim SH, Messenger MP, Gregory WM, et al. Prognostic utility of preoperative circulating osteopontin, carbonic anhydrase IX and CRP in renal cell carcinoma. Br J Cancer. 2012;107(7):1131–1137.
100. Magni F, Van Der Burgt YEM, Chinello C, et al. Biomarkers discovery by peptide and protein profiling in biological fluids based on functionalized magnetic beads purification and mass spectrometry. Blood Transfus. 2010;8(SUPPL. 3):92–97.
101. Liu JS, Li HZ, Zhang YS, et al. Screening for the serum differential proteins of renal cell carcinoma using magnetic beads-based matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Zhongguo Yi Xue Ke Xue Yuan Xue Bao. 2011;33 (3):287–291.
102. Gianazza E, Chinello C, Mainini V, et al. Alterations of the serum peptidome in renal cell carcinoma discriminating benign and malignant kidney tumors. J Proteomics. 2012;76:125–140.
103. Pastore AL, Palleschi G, Silvestri L, et al. Serum and urine biomarkers for human renal cell carcinoma. Dis Markers. 2015;2015. Doi 10.1155/2015/251403.
104. Wang GC, Hsieh PS, Hsu HH, et al. Expression of cortactin and survivin in renal cell carcinoma associated with tumor aggressiveness. World J Urol. 2009;27(4):557–563.
105. Zhang L, Wu X, Luo C, et al. The 786-0 renal cancer cell-derived exosomes promote angiogenesis by downregulating the expression of hepatocyte cell adhesion molecule. Mol Med Rep. 2013;8 (1):272–276.
106. Yang L, Wu X, Wang D, et al. Renal carcinoma cell-derived exosomes induce human immortalized line of Jurkat t lymphocyte apoptosis in vitro. Urol Int. 2013;91(3):363–369.
107. Gu X, Erb U, Buchler MW, et al. Improved vaccine efficacy of tumor exosome compared to tumor lysate loaded dendritic cells in mice. Int J Cancer. 2015;136(4):E74–E84.

108. Lindoso RS, Collino F, Camussi G. Extracellular vesicles derived from renal cancer stem cells induce a pro- tumorigenic phenotype in mesenchymal stromal cells. Oncotarget. 2015;6(10):7959–7969.

109. Del Boccio P, Raimondo F, Pieragostino D, et al. A hyphenated microLC-Q-TOF-MS platform for exosomal lipidomics investigations: application to RCC urinary exosomes. Electrophoresis. 2012;33(4):689–696.

110. Raimondo F, Morosi L, Corbetta S, et al. Differential protein profiling of renal cell carcinoma urinary exosomes. Mol Biosyst. 2013;9:1220–1233.

111. Hakimi AA, Reznik E, Lee C-H, et al. An integrated metabolic atlas of clear cell renal cell carcinoma. Cancer Cell. 2016;29(1):104–116.

•• *Study of system biology that compare metabolomics and transcriptomics in ccRCC: the results highlighted heterogeneity between gene expression changes and metabolomics changes in the same pathway, confirming a specificity for this neoplasm that is not evident in other tumor types.*

112. Gatto F, Nookaew I, Nielsen J. Chromosome 3p loss of heterozygosity is associated with a unique metabolic network in clear cell renal carcinoma. Proc Natl Acad Sci. 2014;111(9):E866–E875.

113. Van Der Mijn JC, Panka DJ, Geissler AK, et al. Novel drugs that target the metabolic reprogramming in renal cell cancer. Cancer Metab. 2016. doi 10.1186/s40170-016-0154-8.

114. Bennike TB, Kastaniegaard K, Padurariu S, et al. Proteome stability analysis of snap frozen, RNAlater preserved, and formalin- fixed paraffin-embedded human colon mucosal biopsies. Data Br. 2016;6:942–947.

115. Mertins P, Mani DR, Ruggles KV, et al. Proteogenomics connects somatic mutations to signalling in breast cancer. Nature. 2016;534 (7605):55–62.

1.3 Matrix Assisted Laser Desorption Ionisation - Mass Spectrometry Imaging

Extract of:

Matrix-Assisted Laser Desorption/Ionisation Mass Spectrometry Imaging in the Study of GastricCancer: A Mini Review.

Smith A*, Piga I*, Galli M*, Stella M*, Denti V, Del Puppo M, Magni F.

Department of Medicine and Surgery, University of Milano-Bicocca, Clinical Proteomics and Metabolomics Unit, 20854 Vedano al Lambro, Italy;

*These authors contributed equally to this work.

Published: Int J Mol Sci. 2017 Dec 1;18(12). pii: E2588. doi: 10.3390/ijms18122588.

1. Matrix-assisted laser desorption ionisation - mass spectrometry imaging (MALDI-MSI) in a nutshell

MALDI-MSI applied to thin mammalian tissue sections was formally introduced in 1997 and its use has increased exponentially in recent years [8]. The technique relies on the use of a MALDI matrix, which consists of small organic molecules that are designed to absorb the energy of a pulsed laser beam. These molecules commonly possess a suitable chromophore, usually in the form of an aromatic core, and it is this property of the matrix that facilitates the absorption of the UV laser energy. When this matrix is applied to the surface of a sample, it promotes the formation of a ubiquitous layer of co-crystals, which incorporates both matrix and analyte molecules in its network. When the laser beam is applied to the surface of the sample, the absorbed energy leads to rapid desorption of both the matrix and analyte crystals and subsequent "soft" ionisation [3]. Typical MALDI-MSI analysis is most commonly performed on tissue sections that have been sectioned and mounted onto electrically conductive glass slides, such as those coated with indium tin oxide (ITO) [9]. For protein, lipid, xenobiotics and metabolite imaging, the analysis is most commonly performed using fresh-frozen (FF) tissue [10,11]. Regarding the imaging of drugs and products of drug metabolism, MALDI-MSI has been readily used within the pharmaceutical community for the purpose of drug discovery and development [12]. The monitoring of the spatial distribution of drugs and their metabolites in order to evaluate a drug's absorption properties, as well as the characterisation of a drug's delivery and penetration in a target organ, represent some examples of how MALDI-MSI tools have been successfully applied in this field [13,14]. In addition to qualitative MALDI-MSI approaches, the ability to obtain absolute quantitative information by MALDI-MSI for drug analysis, by applying internal standards, has recently been further investigated [15,16]. In the case of protein imaging, formalin-fixed paraffin-embedded (FFPE) tissue is now also readily employed [17].

FFPE tissue accounts for a large percentage of the patient samples collected and stored in medical centres [18] and thus represents a potential gold mine of information for histopathological studies involving MALDI-MSI. It also facilitates multi-centric studies using tissue specimens from numerous tissue banks [19,20]. However, the sample preparation for protein imaging of FFPE tissue is more complex and requires an antigen retrieval step followed by tryptic digestion prior to MALDI-MSI analysis. Metabolite imaging has also been conducted on FFPE tissue [21]; however, it has been less extensively investigated with respect to proteins. Finally, a number of groups have focused on the analysis of N-glycans in tissue [22,23], demonstrating that it is possible to monitor the distribution of both N-glycans and proteins within the same tissue section [23]. The potential to monitor N-glycans, one of the most common post-translational modifications, may significantly advance MALDI-MSI investigations in gastric cancer given their fundamental role in many cellular processes and their establishment as clinical biomarkers [23]. A general overview of the MALDI-MSI sample preparation and analysis workflow is given in Figure 1.

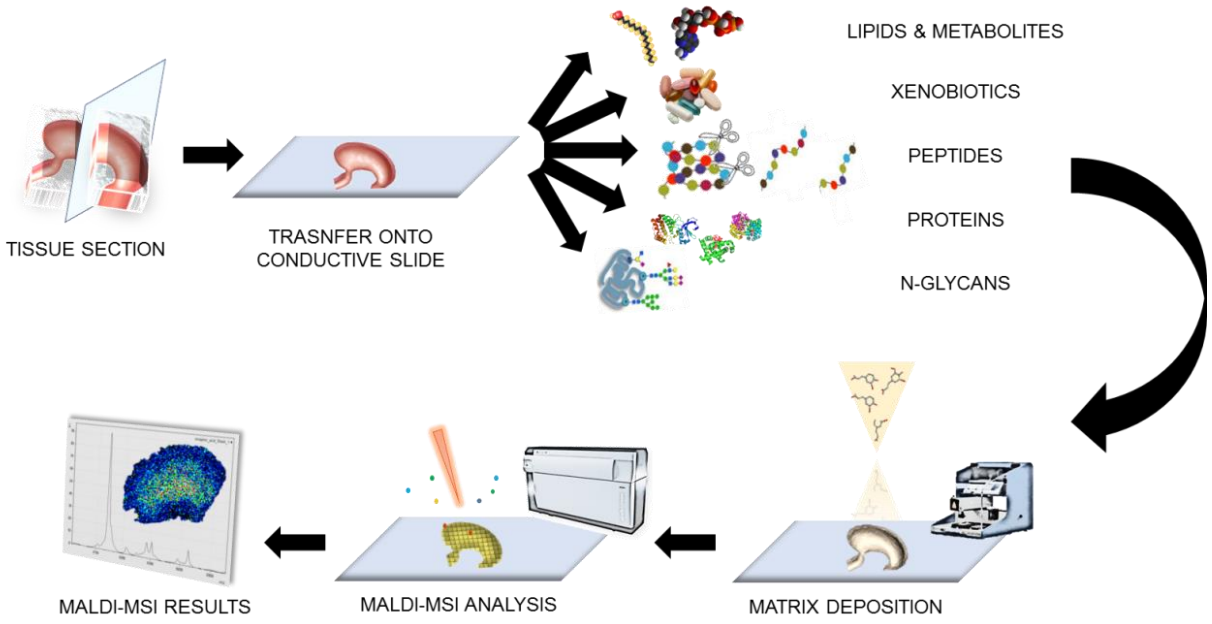


Figure 1 Illustration of the workflow for the matrix-assisted laser desorption/ionisation-mass spectrometry imaging (MALDI-MSI) analysis.

Sample Preparation

Particular attention to detail must be paid during the collection of FF tissue, as negligence during the sample collection can lead to degradation and delocalisation of the analytes of interest. The method most commonly employed during collection is snap-freezing using liquid nitrogen; however, this procedure can damage tissue morphology if it cools at different rates. This can be overcome to some degree by lightly wrapping the tissue in aluminium in order for it to cool at a more uniform rate [24]. Alternatively, Goodwin et al. recommend the use of ethanol or isopropanol solutions at temperatures of ≤-70°C [25]. Once snap-frozen, FF tissue sections can be maintained at -80 °C for up to a year without evidence of degradation [3,24,26]. Prior to matrix application, tissue washes are also performed in order to remove any molecules that may interfere with the ionisation of the target analytes, including any compounds used during the sectioning procedure. Standard protocols for protein MS-imaging recommend washing the tissue sequentially in increasing concentrations of ethanol, whilst, for example in tissue with a high lipidomic content, washing this tissue with chloroform or xylene can improve protein detection [24,25,27]. Conversely, different washing protocols should be used if the intended analytes are not proteins, e.g., ethanol (70%) with the addition of ammonium acetate (NH4Ac) is recommended for the desalting of tissue prior to lipidomic analysis [28].

Regarding FFPE tissue, metabolite MS-imaging requires tissue immersion in xylene, or a similar organic solvent, in order to remove any paraffin [21]. Protein MS-imaging, however, requires a more complex procedure [4,29,30]. Following paraffin removal, tissue rehydration is then performed prior to antigen retrieval. The antigen retrieval step is generally performed at 97°C whilst immersed in a buffer solution that most commonly contains either Tris-HCl or citric acid [31–33], and is required in order to break the methylene bridges that have formed between amino acids during the fixation process. Whilst enzymatic digestion is conventionally performed in solution for proteomic

investigations, here the spatial integrity of the proteins is required, and thus the procedure is performed in situ. However, the hydrophobic nature of certain proteins renders them proteolytically resistant to digestion and ultimately limits the peptide yield when performed in this manner. The addition of detergents to the trypsin solution can improve solubilisation by unfolding the proteins, increasing the number of possible enzymatic cleavage sites. A number of detergents have been shown to be compatible with MALDI-MSI analysis, such as N-Octanoyl-N-methylglucamin (MEGA-8) and RapiGest SF (Waters Corporation, Manchester, UK) [34,35], and significantly improved peptide yield as well as signal intensity, facilitating a greater number of peptide identifications whilst using a bottom-up approach. Alternatively, enzymatic digestion can be performed using N-glycosidase F (PNGase F) in order to visualise the distribution of N-glycans that are associated with different pathological states of tissue [23].

Matrix deposition plays a crucial role in MALDI-MSI experiments and can limit the true spatial resolution that can be achieved. The general aim of the co-crystallisation process is to maximise analyte extraction whilst at the same time limiting the degree of lateral diffusion, which is equally important to the choice of matrix [9]. Wet matrix deposition methods, involving the use of automated spotters [36] and, in particular, sprayers [37], are particularly efficient for the extraction of proteins and peptides and commonly lead to the formation of crystals of between 10 and 50 μm in diameter. On the other hand, solvent-free matrix deposition involving sublimation has surged in popularity for the analysis of lipids and metabolites due to its ability to deposit a uniform coating of fine matrix crystals that are only a few microns in diameter [38]. Therefore, sublimation represents a highly cost-effective approach to matrix deposition that is both reproducible and compatible with high spatial resolution MALDI-MSI [9]. In contrast to sublimation methods that deposit dry matrix onto the surface of the tissue section, microscope glass slides can also be pre-coated with a MALDI matrix prior to tissue

mounting [39]. This has also been shown to be a high-throughput approach that can be effective for the analysis of both proteins [39] and low molecular weight compounds, such as lipids and metabolites [40].

Depending on the target analyte of choice, a number of different matrices can be used. For example, DHB (2,5-Dihydroxybenzoic acid), sinapinic acid (SA; 3,5-dimethoxy-4-hydroxycinnamic acid) and α-CHCA (α -cyano-4-hydroxycinnamic acid) are the most common matrices of choice for the extraction of low molecular weight proteins, peptides, and lipids (1–20 kDa) [41]. However, the addition of hexafluoroisopropanol (1,1,1,3,3,3-hexaluoro-2-propanol) and 2,2,2-trifluoroethanol to the matrix solution [42], along with the use of detectors designed for the detection of higher molecular weight analytes, has been shown to enhance the potential to detect higher molecular weight proteins whilst using SA (up to 110 kDa) [43]. Alternatively, ferulic acid (3-(4-hydroxy-3-methoxy-phenyl)-prop-2-enoic acid) may also be used for the extraction of high molecular weight proteins (up to 140 kDa) [44]. Additionally, ionic matrices such as CHCA/aniline (CHCA/ANI) and CHCA/N,N-dimethylaniline (CHCA/DANI) have been employed to obtain a more ubiquitous matrix layer and enhance the detection of protein signals [45]. For metabolite imaging, 9-aminoacridine (9AA) is often employed and the mass spectrometer is set in negative-ion mode [46]. In view of the rapid evolution in mass spectrometric instrumentation, the search for novel matrices and matrix deposition protocols has also come to the fore. For example, Garate et al. demonstrated the use of MBT (2-mercaptobenzothiazole) and DAN (2,5-diaminonapthalene) as MALDI matrices that produced very small crystals and were not a limiting factor during the acquisition of MALDI-MS images with pixel sizes as low as 5 μm [47].

Instrumental Advancements

MALDI mass spectrometry instrumentation has rapidly evolved in recent years, offering ever more mass resolution and increased sensitivity. In fact, state-of-the-art MALDI-MS instrumentation enables the generation of individual spectra with intensities measured at 25,000–50,000 m/z-bins for ToF MS and even greater than 1,000,000 m/z-bins for Fourier-transform ion cyclotron resonance (FTICR) MS measurements [48]. These advancements have enabled more comprehensive analysis and the better resolution of species with similar m/z values. In fact, modern MALDI-FTICR-MS instrumentation, as well as MALDI linear ion trap (Orbitrap), can enable the unequivocal identification of certain analytes (particularly for small molecular weight compounds such as lipids, drugs and metabolites) based on their accurate mass alone [49,50].

Furthermore, the addition of a separate dimension, the drift time, to quadrupole-ToF and ion mobility instrumentation can overcome the inability of MALDI-ToF instruments to differentiate isobaric ions, enabling the detection of a higher number of peaks [51]. Notwithstanding this rapid evolution, several technical issues related to MALDI-MSI still need to be improved, such as spatial resolution and sensitivity. However, next-generation instruments are beginning to address these limiting factors [52], not only improving spatial resolution and sensitivity, but also increasing the spectral acquisition rate as well as minimising pixel-to-pixel variability, facilitating higher quality and more robust analysis. Continuing in this vein, MALDI-MSI will be able to not only analyse single cells, but also potentially delve deeper and analyse at a subcellular level, enabling the intra-cellular proteome to be investigated. Furthermore, it will also be possible to routinely generate three-dimensional MALDI-MS images in order to obtain a snapshot of the pathological state of an entire organ by combining MALDI-MS images of consecutive tissue sections and reconstructing a three-dimensional representation using the appropriate (and currently available) software [53–55].

Statistical Analysis and Data Elaboration

MALDI-MSI records the presence and relative abundance of a great variety of molecules on tissue, allowing the localisation and spatial distribution of such molecules to be visualised. For each pixel of the digitalised tissue image, a mass spectrum is acquired, generating a so-called "data cube" (Figure 2A), a tensor in which the two spatial dimensions (x and y axes) of the digitalised tissue section are combined with a third dimension consisting of the mass-to-charge ratio (m/z) of the molecules present within the tissue section. Depending on the spatial resolution and the number of data points (sampling rate) per spectrum, a MALDI-MSI dataset can be of several gigabytes, even terabytes. Therefore, efficient statistical methods for data mining must be employed in order to extract information from the spectral data [56].

Before proceeding with the statistical analysis, however, a series of pre-processing steps are required in order to remove the analytical variability connected with the chemical impurities present in the samples and the electronic nature of the mass spectrometric instrumentation [57,58]. These steps adequately prepare the MS data for statistical analysis and enhance the biological information present within the data (Figure 2B) [59]. Smoothing, performed by employing algorithms such as the Savitzky–Golay filter and the moving average window, aims at discarding the fluctuations in the spectrum mainly due to the electronic nature of the mass spectrometer: this process enhances and eases the peak detection phase, since false positive peaks corresponding to electrical noise are discarded. Baseline subtraction, performed by algorithms such as TopHat, iterative convolution and convex hull, ensures that the spectra all lie on the x-axis and all the peak intensities are estimated from the x-axis itself. Normalisation multiplies the intensity of the data points of the spectra by a scaling factor in order to bring the intensity scale (merely related with the analogue-digital conversion of the signal) within the same range and therefore make analyses more reproducible [60]: the total ion count (TIC) method divides the spectrum intensities by the sum of all the intensity values for

that spectrum; the root mean square (RMS) method divides the spectrum intensities by the square root of the sum of the intensity values for that spectrum squared; the median method divides the spectrum intensities by the median intensity of that spectrum. Finally, peak picking extracts the information regarding the peaks present within the mass spectrum, in the form of m/z and intensity pair values. After peak maxima have been aligned to each other in order to account for fluctuations in the peak values among the spectra of the dataset related with the peak picking process, the data can be submitted to statistical analysis. Mostly, machine learning algorithms are employed for statistical analysis of the data cube, and, depending on the data provided and on the aim of the data mining process, unsupervised or supervised approaches are carried out (Figure 2C,D) [61].

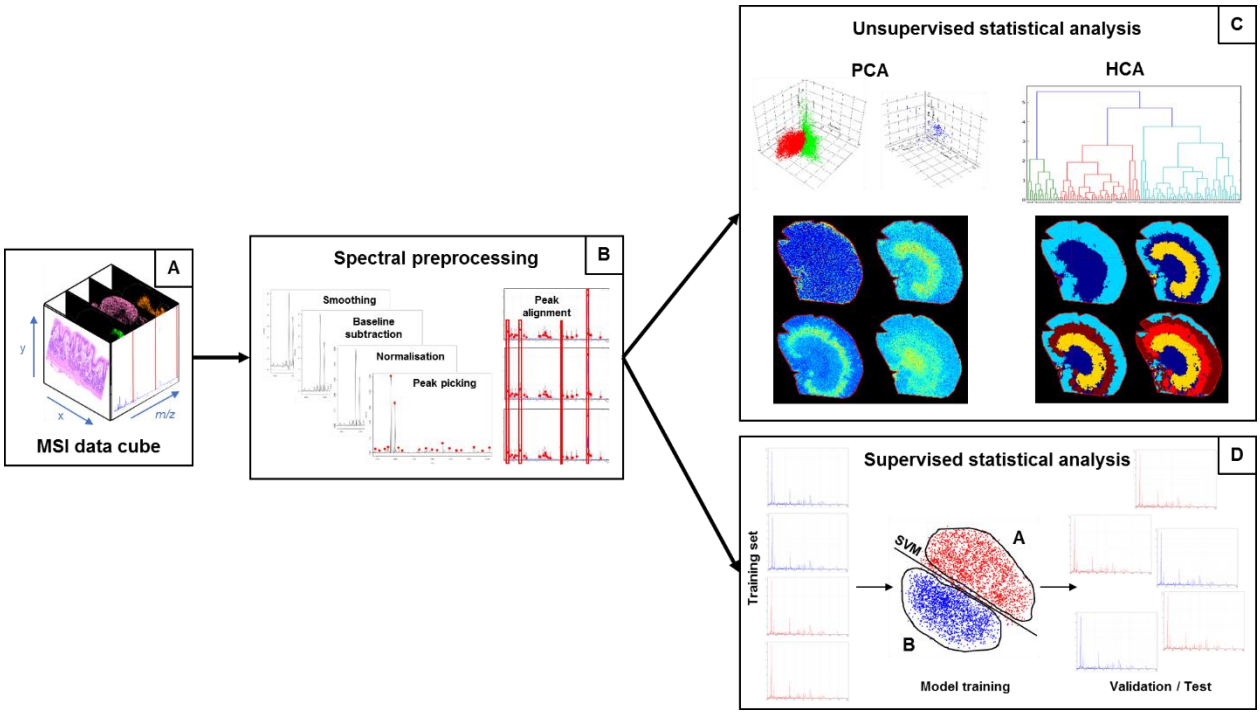


Figure 2. A schematic overview of the MSI data elaboration workflow. (A) Data cube; (B) the series of spectra pre-processing steps; (C) unsupervised and (D) supervised statistical analysis performed on a spectra dataset. MSI, mass spectrometry imaging; PCA, principal component analysis; HCA, hierarchical clustering analysis; SVM, support vector machine.

Unsupervised learning takes unlabelled data as input, i.e., data in which the outcome is not known; by the exploitation of the intrinsic information present in the data, clustering operations are performed in order to highlight hidden structures and/or patterns within the data and are achieved by estimating the similarities among data observations [62]. However, these approaches can also be used in a partially supervised manner, in such a way that the outcome of each observation is preserved during the unsupervised analysis but not taken into account by the algorithm, which performs its operations blind.

Examples of the unsupervised methods for statistical analysis that have been applied in the case of MALDI-MSI gastric cancer datasets are hierarchical clustering analysis (HCA), principal component analysis (PCA) and t-distributed stochastic neighbour embedding (t-SNE). Hierarchical clustering analysis (HCA) estimates the pairwise distance among data observations and generates a dendrogram, in which the observations are grouped under the same nodes based on their similarity to each other [62]. In mass spectrometry imaging, data observations correspond to individual spectra and pixels are associated with spectra; therefore, pixels corresponding to spectra under the same node can be coloured in the same way, generating an unsupervised segmentation tissue image, which can highlight areas of interest on a molecular basis without a priori knowledge regarding the presence of such areas in the tissue section [63]. Therefore, the MSI approach has the potential to aid the diagnostic process by bringing areas of tissue to the attention of the pathologist and highlighting the molecular alterations, even if they do not correlate with cyto-morphological features.

Principal component analysis (PCA) is a mathematical technique that aims at reducing data dimensionality whilst preserving the information present within the data [64]. PCA provides an overview of the entire spectral dataset by generating new variables (called principal components, PC) from the linear combination of the spectral features (i.e., peaks): since the PCs are generated orthogonally to one another, no redundancy among the new variables is present and PCs are sorted according to the amount of variance

that is retained from the original dataset. This is done in such a way that an overview of almost all the information present within the data can be obtained by looking at the first principal components. The output of a PCA consists of a score chart and a loadings plot: the former places data observations in a 2D or 3D graph according to the score of the PCs, allowing the degree of similarity among the spectra to be evaluated according to their distribution/clustering in the chart; the latter, by resembling the distribution of the former, allows us to evaluate which feature contributes more significantly in driving the distribution/clustering of data observations in the score chart. By combining the two plots, not only is it possible to determine whether the data is capable of discriminating among different classes, but also putative signals of interest can be highlighted for further investigation. t-SNE is a non-linear dimensionality reduction technique that aims at reducing the number of dimensions to two or three in such a way that a 2D or 3D visualisation is easily computed [65]: each n-dimensional data point is mapped to a two- or three-dimensional point in such a way that similar observations correspond to close points in the mapped space. While PCA generates new variables by computing a linear combination of features, t-SNE retains all the features as they are in order to perform the computations. In the case of spectral datasets, t-SNE can be applied by employing either all the individual peaks or only the spectral data points.

On the other hand, supervised learning aims at employing algorithms, referred to as classifiers, which learn from labelled data, i.e., data in which the outcome is known, in order to exploit known features (which correspond to peaks in the mass spectrometry imaging dataset) to make predictions about new, unknown data, resembling the classification problem [66]. The first phase, the training phase, allows classifiers to build the mathematical formula by taking labelled data as input and discriminate among the provided categories via different techniques: for example, support vector machines (SVMs) fit a hyperplane, with the additional aid of kernel functions, to maximise the distance between the closest data observations belonging to different classes [67];

random forests (RF) build a decision tree in which thresholds of feature values determine whether the observation belongs to a class or to another [68]. The following phase is the validation phase, in which the performances of the classifiers are evaluated by the predictions made in a partition of the same training set (cross validation) or in an externally labelled dataset (external validation). The discrepancy between the predicted class and the actual class yields the performance parameters of the model, such as sensitivity (TPR), specificity (TNR), positive predictive value (PPV) and negative predictive value (NPV). Finally, the classifier can be employed for making predictions regarding new data, which can also be weighed according to the performance parameters evaluated in the previous phases. In MSI, an on-tissue classification can be obtained, by generating a MS segmentation image resembling the classification by colouring pixels according to the predicted class.

Acknowledgments: This work was supported by grants from the MIUR: FIRB 2007 (RBRN07BMCT_11), FAR 2013–2016; and in part by Fondazione Gigi & Pupa Ferrari Onlus.

References

3. Chugtai K., Heeren R. Mass spectrometric imaging for biomedical tissue analysis. Chem. Rev. 2010;110:3237–3277. doi: 10.1021/cr100012c.

4. Casadonte R., Longuespée R., Kriegsmann J., Kriegsmann M. Advances in Cancer Research. Volume 134. Elsevier; Amsterdam, The Netherlands: 2017. MALDI IMS and Cancer Tissue Microarrays; pp. 173–200.

8. Caprioli R.M., Farmer T.B., Gile J. Molecular Imaging of Biological Samples: Localization of Peptides and Proteins Using MALDI-TOF MS. Anal. Chem. 1997;69:4751–4760. doi: 10.1021/ac970888i.

9. Baker T.C., Han J., Borchers C.H. Recent advancements in matrix-assisted laser desorption/ionization mass spectrometry imaging. Curr. Opin. Biotechnol. 2017;43:62–69. doi: 10.1016/j.copbio.2016.09.003.

10. Goodwin R.J.A., Nilsson A., Borg D., Langridge-Smith P.R.R., Harrison D.J., Mackay C.L., Iverson S.L., Andrén P.E. Conductive carbon tape used for support and mounting of both whole animal and fragile heat-treated tissue sections for MALDI MS imaging and quantitation. J. Proteom. 2012;75:4912–4920. doi: 10.1016/j.jprot.2012.07.006.

11. Nilsson A., Goodwin R.J.A., Shariatgorji M., Vallianatou T., Webborn P.J.H., Andrén P.E. Mass Spectrometry Imaging in Drug Development. Anal. Chem. 2015;87:1437–1455. doi: 10.1021/ac504734s.

12. Greer T., Sturm R., Li L. Mass spectrometry imaging for drugs and metabolites. J. Proteom. 2011;74:2617–2631. doi: 10.1016/j.jprot.2011.03.032. [PMC free article]

13. Giordano S., Morosi L., Veglianese P., Licandro S.A., Frapolli R., Zucchetti M., Cappelletti G., Falciola L., Pifferi V., Visentin S., et al. 3D Mass Spectrometry Imaging Reveals a very Heterogeneous Drug Distribution in Tumors. Sci. Rep. 2016;6 doi: 10.1038/srep37027.

14. Thompson C.G., Bokhart M.T., Sykes C., Adamson L., Fedoriw Y., Luciw P.A., Muddiman D.C., Kashuba A.D.M., Rosen E.P. Mass Spectrometry Imaging Reveals Heterogeneous Efavirenz Distribution within Putative HIV Reservoirs. Antimicrob. Agents Chemother. 2015;59:2944–2948. doi: 10.1128/AAC.04952-14.

15. Chumbley C.W., Reyzer M.L., Allen J.L., Marriner G.A., Via L.E., Barry C.E., Caprioli R.M. Absolute Quantitative MALDI Imaging Mass Spectrometry: A Case of Rifampicin in Liver Tissues. Anal. Chem. 2016;88:2392–2398. doi: 10.1021/acs.analchem.5b04409.

16. Groseclose M.R., Castellino S. A Mimetic Tissue Model for the Quantification of Drug Distributions by MALDI Imaging Mass Spectrometry. Anal. Chem. 2013;85:10099–10106. doi: 10.1021/ac400892z.

17. Longuespée R., Casadonte R., Kriegsmann M., Pottier C., Picard de Muller G., Delvenne P., Kriegsmann J., De Pauw E. MALDI mass spectrometry imaging: A cutting-edge tool for fundamental and clinical histopathology. Proteom. Clin. Appl. 2016;10:701–719. doi: 10.1002/prca.201500140.

18. Stauber J., MacAleese L., Franck J., Claude E., Snel M., Kaletas B.K., Wiel I.M.V.D., Wisztorski M., Fournier I., Heeren R.M.A. On-tissue protein identification and imaging by MALDI-ion mobility mass spectrometry. J. Am. Soc. Mass Spectrom. 2010;21:338–347. doi: 10.1016/j.jasms.2009.09.016.

19. De Sio G., Smith A.J., Galli M., Garancini M., Chinello C., Bono F., Pagni F., Magni F. A MALDI-Mass Spectrometry Imaging method applicable to different formalin-fixed paraffin-embedded human tissues. Mol. Biosyst. 2015;11:1507–1514. doi: 10.1039/C4MB00716F.

20. Gorzolka K., Walch A. MALDI mass spectrometry imaging of formalin-fixed paraffin-embedded tissues in clinical research. Histol. Histopathol. 2014;29:1365–1376. doi: 10.14670/HH-29.1365.

21. Ly A., Buck A., Balluff B., Sun N., Gorzolka K., Feuchtinger A., Janssen K.-P., Kuppen P.J.K., van de Velde C.J.H., Weirich G., et al. High-mass-resolution MALDI mass spectrometry imaging of metabolites from formalin-fixed paraffin-embedded tissue. Nat. Protoc. 2016;11:1428–1443. doi: 10.1038/nprot.2016.081.

22. Briggs M.T., Ho Y.Y., Kaur G., Oehler M.K., Everest-Dass A.V., Packer N.H., Hoffmann P. N-Glycan matrix-assisted laser desorption/ionization mass spectrometry imaging protocol for formalin-fixed paraffin-embedded tissues. Rapid Commun. Mass Spectrom. 2017;31:825–841. doi: 10.1002/rcm.7845.

23. Heijs B., Holst S., Briaire-de Bruijn I.H., van Pelt G.W., de Ru A.H., van Veelen P.A., Drake R.R., Mehta A.S., Mesker W.E., Tollenaar R.A., et al. Multimodal Mass Spectrometry Imaging of N-Glycans and Proteins from the Same Tissue Section. Anal. Chem. 2016;88:7745–7753. doi: 10.1021/acs.analchem.6b01739.

24. Schwartz S.A., Reyzer M.L., Caprioli R.M. Direct tissue analysis using matrix-assisted laser desorption/ionization mass spectrometry: Practical aspects of sample preparation. J. Mass Spectrom. 2003;38:699–708. doi: 10.1002/jms.505.

25. Goodwin R.J.A., Pennington S.R., Pitt A.R. Protein and peptides in pictures: Imaging with MALDI mass spectrometry. Proteomics. 2008;8:3785–3800. doi: 10.1002/pmic.200800320.

26. Patel E. Fresh Frozen Versus Formalin-Fixed Paraffin Embedded for Mass Spectrometry Imaging. Methods Mol. Biol. 2017;1618:7–14. doi: 10.1007/978-1-4939-7051-3_2.

27. Lemaire R., Wisztorski M., Desmons A., Tabet J.C., Day R., Salzet M., Fournier I. MALDI-MS direct tissue analysis of proteins: Improving signal sensitivity using organic treatments. Anal. Chem. 2006;78:7145–7153. doi: 10.1021/ac060565z.

28. Wang H.-Y.J., Liu C.B., Wu H.-W. A simple desalting method for direct MALDI mass spectrometry profiling of tissue lipids. J. Lipid Res. 2011;52:840–849. doi: 10.1194/jlr.D013060.

29. Thomas A., Chaurand P. Advances in tissue section preparation for MALDI imaging MS. Bioanalysis. 2014;6:967–982. doi: 10.4155/bio.14.63.

30. Goodwin R.J.A. Sample preparation for mass spectrometry imaging: Small mistakes can lead to big consequences. J. Proteom. 2012;75:4893–4911. doi: 10.1016/j.jprot.2012.04.012.

31. Boskamp T., Lachmund D., Oetjen J., Cordero Hernandez Y., Trede D., Maass P., Casadonte R., Kriegsmann J., Warth A., Dienemann H., et al. A new classification method for MALDI imaging mass spectrometry data acquired on formalin-fixed paraffin-embedded tissue samples. Biochim. Biophys. Acta. 2017;1865:916–926. doi: 10.1016/j.bbapap.2016.11.003.

32. Groseclose M.R., Massion P.P., Chaurand P., Caprioli R.M. High-throughput proteomic analysis of formalin-fixed paraffin-embedded tissue microarrays using MALDI imaging mass spectrometry. Proteomics. 2008;8:3715–3724. doi: 10.1002/pmic.200800495.

33. Gustafsson J.O.R., Oehler M.K., McColl S.R., Hoffmann P. Citric acid antigen retrieval (CAAR) for tryptic peptide imaging directly on archived formalin-fixed paraffin-embedded tissue. J. Proteome Res. 2010;9:4315–4328. doi: 10.1021/pr9011766.

34. Huang H.Z., Nichols A., Liu D. Direct Identification and Quantification of Aspartyl Succinimide in an IgG2 mAb by RapiGest Assisted Digestion. Anal. Chem. 2009;81:1686–1692. doi: 10.1021/ac802708s.

35. Patel E., Clench M.R., West A., Marshall P.S., Marshall N., Francese S. Alternative surfactants for improved efficiency of in situ tryptic proteolysis of fingermarks. J. Am. Soc. Mass Spectrom. 2015;26:862–872. doi: 10.1007/s13361-015-1140-z.

36. Franck J., Arafah K., Barnes A., Wisztorski M., Salzet M., Fournier I. Improving Tissue Preparation for Matrix-Assisted Laser Desorption Ionization Mass Spectrometry

Imaging. Part 1: Using Microspotting. Anal. Chem. 2009;81:8193–8202. doi: 10.1021/ac901328p.

37. Beine B., Diehl H.C., Meyer H.E., Henkel C. Tissue MALDI Mass Spectrometry Imaging (MALDI MSI) of Peptides. In: Reinders J., editor. Proteomis in Systems Biology. Volume 1394. Springer; New York, NY, USA: 2016. pp. 129–150.

38. Hankin J.A., Barkley R.M., Murphy R.C. Sublimation as a method of matrix application for mass spectrometric imaging. J. Am. Soc. Mass Spectrom. 2007;18:1646–1652. doi: 10.1016/j.jasms.2007.06.010.

39. Yang J., Caprioli R.M. Matrix pre-coated targets for high throughput MALDI imaging of proteins: Matrix pre-coated targets for MALDI imaging MS. J. Mass Spectrom. 2014;49:417–422. doi: 10.1002/jms.3354.

40. Grove K.J., Frappier S.L., Caprioli R.M. Matrix Pre-Coated MALDI MS Targets for Small Molecule Imaging in Tissues. J. Am. Soc. Mass Spectrom. 2011;22:192–195. doi: 10.1007/s13361-010-0013-8.

41. Kaletaş B.K., van der Wiel I.M., Stauber J., Dekker L.J., Güzel C., Kros J.M., Luider T.M., Heeren R.M.A. Sample preparation issues for tissue imaging by imaging MS. Proteomics. 2009;9:2622–2633. doi: 10.1002/pmic.200800364.

42. Franck J., Longuespée R., Wisztorski M., Van Remoortere A., Van Zeijl R., Deelder A., Salzet M., McDonnell L., Fournier I. MALDI mass spectrometry imaging of proteins exceeding 30,000 daltons. Med. Sci. Monit. Int. Med. J. Exp. Clin. Res. 2010;16:BR293–BR299.

43. Van Remoortere A., van Zeijl R.J.M., van den Oever N., Franck J., Longuespée R., Wisztorski M., Salzet M., Deelder A.M., Fournier I., McDonnell L.A. MALDI imaging and profiling MS of higher mass proteins from tissue. J. Am. Soc. Mass Spectrom. 2010;21:1922–1929. doi: 10.1016/j.jasms.2010.07.011.

44. Mainini V., Bovo G., Chinello C., Gianazza E., Grasso M., Cattoretti G., Magni F. Detection of high molecular weight proteins by MALDI imaging mass spectrometry. Mol. Biosyst. 2013;9:1101. doi: 10.1039/c2mb25296a.

45. Calvano C.D., Carulli S., Palmisano F. Aniline/α-cyano-4-hydroxycinnamic acid is a highly versatile ionic liquid for matrix-assisted laser desorption/ionization mass spectrometry. Rapid Commun. Mass Spectrom. 2009;23:1659–1668. doi: 10.1002/rcm.4053.

46. Fagerer S., Nielsen S., Ibáñez A., Zenobi R. Matrix-assisted laser desorption/ionization matrices for negative-mode metabolomics. Eur. J. Mass Spectrom. 2013;19:39. doi: 10.1255/ejms.1209.

47. Garate J., Fernández R., Lage S., Bestard-Escalas J., Lopez D.H., Reigada R., Khorrami S., Ginard D., Reyes J., Amengual I., et al. Imaging mass spectrometry increased resolution using 2-mercaptobenzothiazole and 2,5-diaminonaphtalene matrices: Application to lipid distribution in human colon. Anal. Bioanal. Chem. 2015;407:4697–4708. doi: 10.1007/s00216-015-8673-7.

48. Spraggins J.M., Rizzo D.G., Moore J.L., Noto M.J., Skaar E.P., Caprioli R.M. Next-generation technologies for spatial proteomics: Integrating ultra-high speed MALDI-TOF and high mass resolution MALDI FTICR imaging mass spectrometry for protein analysis. Proteomics. 2016;16:1678–1689. doi: 10.1002/pmic.201600003.

49. Cornett D.S., Frappier S.L., Caprioli R.M. MALDI-FTICR imaging mass spectrometry of drugs and metabolites in tissue. Anal. Chem. 2008;80:5648–5653. doi: 10.1021/ac800617s.

50. Römpp A., Spengler B. Mass spectrometry imaging with high resolution in mass and space. Histochem. Cell Biol. 2013;139:759–783. doi: 10.1007/s00418-013-1097-6.

51. Kettling H., Vens-Cappell S., Soltwisch J., Pirkl A., Haier J., Müthing J., Dreisewerd K. MALDI Mass Spectrometry Imaging of Bioactive Lipids in Mouse Brain with a Synapt G2-S Mass Spectrometer Operated at Elevated Pressure: Improving the

Analytical Sensitivity and the Lateral Resolution to Ten Micrometers. Anal. Chem. 2014;86:7798–7805. doi: 10.1021/ac5017248.

52. Ogrinc Potočnik N., Porta T., Becker M., Heeren R.M.A., Ellis S.R. Use of advantageous, volatile matrices enabled by next-generation high-speed matrix-assisted laser desorption/ionization time-of-flight imaging employing a scanning laser beam. Rapid Commun. Mass Spectrom. 2015;29:2195–2203. doi: 10.1002/rcm.7379.

53. Oetjen J., Veselkov K., Watrous J., McKenzie J.S., Becker M., Hauberg-Lotte L., Kobarg J.H., Strittmatter N., Mróz A.K., Hoffmann F., et al. Benchmark datasets for 3D MALDI- and DESI-imaging mass spectrometry. GigaScience. 2015;4:20. doi: 10.1186/s13742-015-0059-4.

54. Patterson N.H., Doonan R.J., Daskalopoulou S.S., Dufresne M., Lenglet S., Montecucco F., Thomas A., Chaurand P. Three-dimensional imaging MS of lipids in atherosclerotic plaques: Open-source methods for reconstruction and analysis. Proteomics. 2016;16:1642–1651. doi: 10.1002/pmic.201500490.

55. Anderson D.M.G., Van de Plas R., Rose K.L., Hill S., Schey K.L., Solga A.C., Gutmann D.H., Caprioli R.M. 3-D imaging mass spectrometry of protein distributions in mouse Neurofibromatosis 1 (NF1)-associated optic glioma. J. Proteom. 2016;149:77–84. doi: 10.1016/j.jprot.2016.02.004.

56. Trede D., Kobarg J.H., Oetjen J., Thiele H., Maass P., Alexandrov T. On the Importance of Mathematical Methods for Analysis of MALDI-Imaging Mass Spectrometry Data. J. Integr. Bioinform. 2012;9:1–11.

57. Ràfols P., Vilalta D., Brezmes J., Cañellas N., del Castillo E., Yanes O., Ramírez N., Correig X. Signal preprocessing, multivariate analysis and software tools for MA(LDI)-TOF mass spectrometry imaging for biological applications. Mass Spectrom. Rev. 2016:1–26. doi: 10.1002/mas.21527.

58. Norris J.L., Cornett D.S., Mobley J.A., Andersson M., Seeley E.H., Chaurand P., Caprioli R.M. Processing MALDI Mass Spectra to Improve Mass Spectral Direct Tissue Analysis. Int. J. Mass Spectrom. 2007;260:212–221. doi: 10.1016/j.ijms.2006.10.005.
59. Galli M., Zoppis I., Smith A., Magni F., Mauri G. Machine learning approaches in MALDI-MSI: Clinical applications. Expert Rev. Proteom. 2016;13:685–696. doi: 10.1080/14789450.2016.1200470.
60. Deininger S.-O., Cornett D.S., Paape R., Becker M., Pineau C., Rauser S., Walch A., Wolski E. Normalization in MALDI-TOF imaging datasets of proteins: Practical considerations. Anal. Bioanal. Chem. 2011;401:167–181. doi: 10.1007/s00216-011-4929-z.
61. Jones E.A., Deininger S.-O., Hogendoorn P.C.W., Deelder A.M., McDonnell L.A. Imaging mass spectrometry statistical analysis. J. Proteom. 2012;75:4962–4989. doi: 10.1016/j.jprot.2012.06.014.
62. Murtagh F., Contreras P. Algorithms for hierarchical clustering: An overview. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 2012;2:86–97. doi: 10.1002/widm.53.
63. Alexandrov T., Becker M., Deininger S.-O., Ernst G., Wehder L., Grasmair M., von Eggeling F., Thiele H., Maass P. Spatial Segmentation of Imaging Mass Spectrometry Data with Edge-Preserving Image Denoising and Clustering. J. Proteome Res. 2010;9:6535–6546. doi: 10.1021/pr100734z.
64. Jolliffe I.T., Cadima J. Principal component analysis: A review and recent developments. Philos. Trans. R. Soc. Math. Phys. Eng. Sci. 2016;374:20150202. doi: 10.1098/rsta.2015.0202.
65. Abdelmoula W.M., Balluff B., Englert S., Dijkstra J., Reinders M.J.T., Walch A., McDonnell L.A., Lelieveldt B.P.F. Data-driven identification of prognostic tumor subpopulations using spatially mapped t-SNE of mass spectrometry imaging data. Proc. Natl. Acad. Sci. USA. 2016;113:12244–12249. doi: 10.1073/pnas.1510227113.

66. Kotsiantis S.B. Supervised Machine Learning: A Review of Classification Techniques. Informatica. 2007;31:249–268.

67. Cristianini N., Shawe-Taylor J. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press; Cambridge, UK: 2000.

68. Breiman L. Bagging predictors. Mach. Learn. 1996;24:123–140. doi: 10.1007/BF00058655.

SCOPE OF THE THESIS

The body of work enclosed in this thesis present the application of complementary mass spectrometric techniques to different human specimens in order to increase our understanding of the molecular aspects of clear cell Renal Cell Carcinoma, that could be useful for its diagnosis and prognosis.

Chapter two: proteomics investigations using plasma and urine samples to evaluate which biofluid better mirrors the alterations triggered by the disease

Chapter three: analysis of formalin fixed paraffin embedded tissues in order to obtain proteomic signatures of the four ccRCC grade and information about the main biological processes involved in grade progression.

Chapter four: can proteomic signatures of different grade be also detected in urine? Study the effects of the grade and stage of the tumour on the urinary proteome.

CHAPTER 2

Proteomics of liquid biopsies: depicting RCC infiltration into the renal vein by MS analysis of urine and plasma

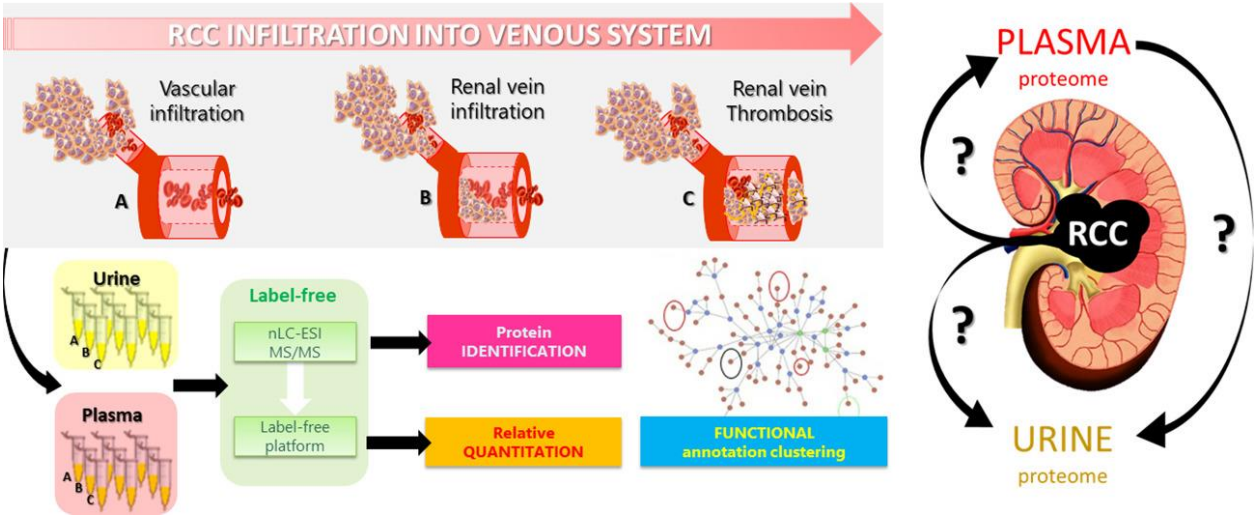
Clizia Chinello[a], Martina Stella[a], Isabella Piga[a], Andrew James Smith[a], Giorgio Bovo[b], Marta Varallo[c], Mariia Ivanova[a], Vanna Denti[a], Marco Grasso[d], Angelica Grasso[d], Marina Del Puppo[a], Apostolos Zaravinos[e], Fulvio Magni[a]

[a] Department of Medicine and Surgery, University of Milano-Bicocca, Clinical Proteomics and Metabolomics Unit, Vedano al Lambro, Italy
[b] Pathology Unit, Vimercate Hospital, Vimercate, Italy
[c] Pathology Unit, S. Gerardo Hospital, Monza, Italy
[d] Urology Unit, S. Gerardo Hospital, Monza, Italy
[e] Department of Life Sciences, School of Sciences, European University Cyprus, 1516 Nicosia, Cyprus

MS contributed to the analysis of the results and to the writing of the manuscript.

Published: Journal of Proteomics, January 2019, Volume 191, Pages 29-37

ABSTRACT Liquid biopsies, as blood and urine, could offer an invaluable, easily accessible source of biomarkers, and evidences for elucidating the pathological processes. Only few studies integrated the proteomes driven by more than one biofluid. Furthermore, it is not clear which biofluid better mirrors the alterations triggered by disease. Venous infiltrating RCC (Renal Cell Carcinoma) could represent an advantageous model for exploring this aspect. Herein, we investigate how blood and urine "proteomically" reflect the changes occurring during RCC infiltration into renal vein (RV) by label-free nLC-ESI-MS/MS. We found 574 and 58 differentially expressed proteins (DEPs) in response to vascular involvement. To the augment of vascular involvement, the abundance of only three proteins in urine (UROM,RALA,CNDP1) and two in plasma APOA1,K2C1) diminished while increased for twenty-six urinary proteins. 80 proteins were found both in urine and plasma, among which twenty-eight were DEPs. A huge overlap between the two biofluids was highlighted, as expected, being urine the filtrate of blood. However, this consistency decreases when RV-occlusion occurs suggesting alternative protein releases, and a loss of kidney architecture. Moreover, several proteomic and functional signatures were biofluid-specific. In conclusion, the complementarity between the specimens allowed to achieve a deeper level of molecular complexity of the RCC venous infiltration.



Significance: Although plasma and urine are strongly interconnected, only few proteomic studies investigated the complementarity of these fluids as bio-sources of information. Moreover, none of them was focused to their analysis and comparison in the context of vascular infiltration of renal cancer. Herein, new insights were gained regarding the impact into urinary and plasma proteome of the changes triggered by the ccRCC invasion into vascular system and renal vein. Furthermore, the integration of the information driven by the two liquid biopsies permits to unravel biological processes otherwise lost.

Keywords: Renal cell carcinoma, Proteomics, Mass spectrometry, Urine, Plasma, Renal vein invasion.

1. Introduction

Tumour-derived proteins carried out by biofluids, as blood and urine, could offer an invaluable and non-invasive source of biomarkers, as well as a font of information regarding the numerous pathological processes related to malignant lesions (primary and metastases) and their evolution.

Since blood transports most of the tissue-derived molecules in the organism, connecting all the important organs and collecting the related changes, for decades it has been gained the consensus of the researchers as an optimal biological sample for biomarker discovery. This biofluids is very rich of disease related proteins even if they are technically difficult to be mined, due to the about nine orders of magnitude dynamic range that could hide the more specific alterations generated by the pathological processes [1].

A less complex medium, such as urine, is an appreciable alternative for screening disease markers, more easy to be collected in large quantities and frequently. Urinary specimen carries a variety of set of soluble proteins and peptides that are primarily derived from kidney, bladder and prostate as well as filtrated by systemic circulation [2]. Given that urinary protein content is likely to reflect normal kidney physiology as well as systemic physiology. Therefore, alterations of the urinary proteome could be used as an indicator of disease not only for urogenital tract and kidneys but potentially also for other organs [3].

In comparison to plasma, urine can be collected over a period of time ensuring an easier monitoring of time-dependent changes of biomarker abundance, and resulted quite stable in terms of peptidome/ proteome composition since proteolytic degradation may be complete prior to collection [4]. Moreover, urine, differently from blood, is not under the strict regulation of homeostatic mechanisms [5]. In fact, blood could likely represent a worthy place to find alterations associated to disease, especially for the earliest and most sensitive biomarkers. Indeed some of these changes are contrasted and do not stay

in blood enough to be detected in time. Thus blood biomarkers are often uncompensated alterations that persist at a rather later stage of a relatively pathological stable condition (i.e. some long half-life proteins or antibody-based biomarkers) [6]. Thus urine, collecting all wastes from the body, can collect a larger number of variations, both huge and severe. Consequently, their concentration is amplified making more visible biomarkers otherwise not detectable in blood [7].

Despite the specific drawbacks, blood and urine indeed could be considered as liquid biopsies easily accessible and able to provide the proteomic landscape of the micro- and macro- changes triggered by a neoplasm. Moreover, the integration of the information driven by both biofluids can not only enrich this molecular scenario but also provide some evidences regarding the handling of tumour-derived proteins. Only a few proteomic studies, mainly investigating secreted biomarkers, have so far focused both on blood and urine [8,9]. In this context, an interesting approach was proposed by Jia L [10], who suggested an integrated strategy to explore kidney function in itemized proteomic language. In this perspective, blood, kidney and urine are investigated in the same context, as a system, instead of isolated specimens. Consequently, the related comparison of the input and output sub-proteome permits to speculate whether a particular protein is blocked, or allowed to be secreted/shed from the kidney. Thus, a similar workflow may outline a picture of the function and state of the organ in physiological conditions, and possibly, also when a modification occurs during a disease/neoplasm progression.

Beside the above mentioned studies, biomarker discovery is generally performed using serum/plasma or alternatively urine. However, which biological fluid better reflects the pathological changes caused by the disease within the cells, e.g. of the kidney, is not very clear.

One of the most distinctive features of renal cell carcinoma (RCC) is its predilection to extend into the venous system including renal vein, inferior vena cava and right atrium.

Indeed, the incidence of involvement of the renal vein (RV) and/or inferior vena cava (IVC) has been reported to range between 4 and 15% [11].

Even if the prognostic significance of venous involvement and tumour thrombus level still remain controversial, it has been observed that RCCs with venous tumour thrombus (VTT) are more aggressive and associated with poor prognosis [12,13], and the risk of cancer-specific mortality increases in VTT patients with perinephric fat invasion [14]. Moreover, VTT could represent a potential middle ground between the phenotype of primary and metastatic RCC, and it was demonstrated that it has a specific molecular trait different from the locally invading tumour and more representative of its extension [15]. For these peculiar characteristics, venous infiltrating RCC could represent a model for investigating the biological information secreted or shed by cancer cell into biofluids when tumour migrates, adapts and begins to spread into circulating system.

Therefore, we investigate by nLC-ESI MS/MS approach how blood and urine mirror the alterations of the proteome during RCC invasion into the renal vein (RV): moving from the tumour infiltration into the circulating system across the vessel wall of this vein until its complete obstruction.

2. Materials & methods

2.1. Reagents

Trifluoroacetic acid, ammonium bicarbonate, porcine trypsin, DTT (dithiothreitol), IAA (Iodoacetamide), Urea, Ammonium Bicarbonate (ABC), HPLC grade water, acetonitrile, acetone were purchased from Sigma-Aldrich (Sigma-Aldrich Chemie GmbH, Buchs, Switzerland). HPLC-grade water is used for all solutions for MS analysis. Amicon Ultra Centrifugal Filters Ultracel 4 ml 30,000 MW, and Amicon Ultra0.5 mL 30 kDa were from Millipore.

2.2. Sample collection

Urine and plasma samples were collected from patients affected by Renal Cell Carcinoma (RCC) the day before surgery at San Gerardo Hospital (Monza, Italy). All subjects had signed an informed consent prior to sample donation and analyses were carried out in agreement with the Declaration of Helsinki. Study protocols and procedures were approved by the local ethic committee (Comitato Etico Azienda Ospedaliera San Gerardo, Monza, Italy). Second morning midstream urine was collected in sterile urine tubes (Anicrin s.r.l., Italy). After centrifugation at 3000 rpm for 10 min, samples were kept at −80 °C [16]. Plasma samples were collected in Vacutainer® K3E containing EDTA (Becton Dickinson Italia S.p.A.), centrifuged at 3700 rpm for 10 min and then stocked at −80 °C.

2.3. Trypsin digestion by FASP workflow

The enzymatic digestion protocol was based on Filter Aided Sample Preparation (FASP) technique [17]. Before sample processing for LC-MS analysis, equal volumes of plasma samples were pooled according to three different levels of renal vein infiltration (A = vascular infiltration; B = RV infiltration; C = RV thrombosis). Each pool was derived from three different patients. Plasma samples were pooled using same volume before concentration and digestion. Urine samples, instead, due to the inaccuracy of the determination of protein concentration probably for the presence of interfering compounds, were pooled only after trypsinization obtaining equally represented sample in the pools. 3 ml of each plasma pool and urine sample was concentrated using 30 kDa MWCO centrifugal filter unit more than ten-fold. A buffer exchange with water was applied. Protein concentration was determined using bicinchoninic acid assay (Pierce - Thermo Fisher Scientific).

In particular, a volume corresponding to 200 µg of proteins for each sample was used both for urine and plasma specimens and mixed with an equal volume of denaturing

buffer (0.1 M DTT, 4% SDS in Tris HCl 0.1 M pH 7.6). The solutions were then incubated at 95° C for 5 min. After disulphide bond reduction, samples were transferred into the ultrafiltration units (Amicon Ultra-0.5 ml 30 kDa, Millipore), made up to 0.5 ml with 8 M urea in 100 mM Tris-HCl, pH 8.5 (UA pH 8.5 solution), and centrifuged at 14000 g for 15 min. FASP digestion was performed as already described [17,18]. Briefly, the centrifugation was repeated after adding UA pH 8.5 solution to the filter unit. For the alkylation, 200 µl of a 0.05 M iodoacetamide IAA (Sigma Aldrich) in UA pH 8.5 were added and incubated for 20 min at dark. Filter units were centrifuged at 14,000g for 10 min, and submitted to four washes, two of which adding 100 µl of UA pH 7.9 solution each, and the remaining two using 100 µl of 50 mM Ammonium Bicarbonate (ABC) for each wash (14,000g for 15 min). Protein digestion was performed overnight at 37 °C adding 2 µg of trypsin. Filtered tryptic peptides were collected in a new tube, and the filters were washed with 50 µl of 50 mM ABC and 0.5 M NaCl. Tryptic peptides were quantified by NanoDrop assay (Thermo Scientific, Sunnyvale, CA) after acidification with TFA.

2.4 nLC-ESI MS/MS label-free quantification

Digested samples were desalted and concentrated using Ziptip™ µC8 pipette tips. About 1 µg of peptide mixtures were injected into UHPLC system (Ultimate™ 3000 RSLCnano, Thermo Scientific, Sunnyvale, CA) coupled online with Impact HD™ UHR-QqToF (Bruker Daltonics, Germany). Each sample was analysed at least three times to minimize technical variability. Samples were loaded onto a pre-column (Dionex, Acclaim PepMap 100 C18, cartridge, 300 µm) followed by a 50 cm nano-column (Dionex, ID 0.075 mm, Acclaim PepMap100, C18). The separation was performed at 40 °C and at a flow rate of 300 nL/min using multistep 4 h gradients of acetonitrile as already reported [19]. The column was on-line interfaced to a nanoBoosterCaptiveSpray™ ESI source (Bruker Daltonics). Data-dependent-

acquisition mode was applied based on CID fragmentation assisted by N_2 as collision gas. Mass accuracy was improved using a specific lock mass (1221.9906 m/z) and a calibration segment (10 mM sodium formate cluster solution) before the beginning of the gradient for each single run. Acquisition parameters was set as already described [20].

Data elaboration were performed through DataAnalysis™ v.4.1 Sp4 (Bruker Daltonics, Germany) and protein identification was achieved using an in-house Mascot search engine (version: 2.4.1), through Mascot Daemon tool. Human swissprot database (accessed Feb 2017, 553,655 sequences; 198,177,566 residues) was used. Searching parameters were set as following: Trypsin as enzyme; carbamidomethyl as fixed modifications; 20 ppm as precursor mass tolerances and 0.05 Da for the product ions. Automatic decoy database search was applied for FDR calculation and a built-in Percolator algorithm for rescores peptide-spectrum matches. Only proteins with at least one unique and significant (p-value < 0.05) peptide were considered identified.

Progenesis QI for proteomics (Non-linear Dynamics, Newcastle, England) was used as label-free quantification platform as already reported [18]. Briefly, raw data were imported and the ion intensity maps of all runs (9 for each biofluid) used for the alignment process to compensate for between-run variation in the LC separation technique. For granting the maximal overlay across the data, only alignment scores above 60% were accepted. Peak peaking was performed using the default sensitivity and a peak width of 0.2 min. The survey scan data is used for the quantification of peptide ions without MS/MS data. Data is then normalized to all proteins. Protein identification was achieved using an in-house Mascot search engine as described above. Protein abundance was calculated using the sum of all unique peptide normalized ion abundances for that protein on each run. The peptide abundance was based on the sum of the intensities within the isotope boundaries. Fold changes were calculated selecting only non conflicting peptides (unique) in order to provide a more confidently

unambiguous read-out of protein abundance, preventing the overlapping of trends derived from different proteins, that shared the same peptides. Statistical tools were used to evaluate the quantitative differences between groups. To indicate the statistical significance of them in group expression data, Anova test was applied (p-value < 0.05). For the power analysis and the estimation of sample size, a threshold of 80% was chosen. Moreover, to afford the multiple testing problem, the FDR adjusted p-values, named q-value, is also provided (q-value < 0.05).

2.5. Bioinformatics analysis

The PANTHER (protein analysis through evolutionary relationship) Classification System [21] version 12.0 (released 2017-07-10) (http:// pantherdb.org) was utilized for gene ontology (GO) analysis. In particular, PANTHER Statistical overrepresentation test with GO-Slim Biological Process annotation data was applied; GO terms with $p \leq 0.05$ after Bonferroni correction were deemed significant.

Differentially expressed proteins (DEPs) were subjected to Core Expression Analysis and investigated for network interrelation by Ingenuity Pathway Analysis (IPA; Qiagen Bioinformatics). IPA scans the set of input proteins to identify networks by using Ingenuity Knowledge Base for interactions between identified "Focus Genes." The UniProt/ Swiss-Prot Accession was used as identifier in the dataset. In this study, the DEPs between Renal Vein invasion (B) and Vascular endothelium infiltration (A), as well as between Renal Vein Thrombosis (C) and Renal Vein invasion (B), along with hypothetical interacting genes stored in the knowledge base in IPA software, were used to generate a set of networks with a maximum network size of 70 genes/proteins. The ratio values in the datasets were converted to fold change values, where the negative inverse ($-1/x$) was taken for values between 0 and 1. Networks were displayed graphically as genes/gene products ("nodes") and the biological relationships between the nodes ("edges"). All edges are from canonical information stored in the Ingenuity

Pathways Knowledge Base. Networks of these genes were generated based on their connectivity and a score ranked each. This score indicates the likelihood of the focus molecules in a network from Ingenuity's knowledge base being found together due to random chance. It is based on the hypergeometric distribution, calculated with the right-tailed Fisher's Exact Test, and corresponds to the negative log of this p-value. A score of Ratio (Expression Fold Change) = 1.5 and p-value (Anova) = 0.05 were set as cutoffs for identifying networks. Furthermore, we used IPA in order to identify the top deregulated molecules and the top canonical pathways in which they participate. In addition, IPA was used to reveal the top molecular and cellular functions, as well as the top upstream regulators, top diseases and biological functions of the DEPs.

3. Results

3.1. Experimental design

A cohort of nine patients affected by clear cell RCC with vascular infiltration was studied through a quantitative proteomic approach based on nLC-ESI-MS/MS. All of the patients were subjected to surgical nephrectomy and the diagnosis was confirmed by the histological examination. Patients were classified in agreement with the 2009 TNM (tumour-node-metastasis) system classification [22] and to their clinical characteristics (Supplemental Table 1).

Assessment of vascular infiltration was achieved by CT-scan (Computed Assisted Tomography) following morphological description after surgery. Based on these examinations, the dataset was divided into 3 groups according to the level of RCC extension into renal vein: -(A) patients with the evidence of vascular invasion in renal site (not otherwise distributed); -(B) patients with the evidence of vascular invasion in

renal site and renal vein invasion; -(C) patients with evidence of renal vein thrombosis. The experimental design was illustrated in Fig. 1.

3.2. Biofluid proteome variation in response to RCC extension into renal vein

3.2.1. Urinary proteome changes

A label-free proteomic approach was applied to urine sample pools in order to identify and quantify urinary proteins whose abundance is significantly different depending on RCC infiltration level into renal vein.

From 1207 identified proteins (Supplemental Table 2), 574 proteins were observed as differentially expressed in at least one of the three conditions, using the following filters: at least 2 unique peptides; fold change ≥ 1.5; anova test p-value ≤ 0.05; power ≥ 0.8; q-values ≤ 0.05 (Supplemental Table 3). The number of proteins identified in each of the runs was reported in Supplemental Fig. 1. The proteins were then grouped according to their fold changes calculated comparing the three studied conditions. In particular among all possible combinations, four groups (i-with an ascending concentration trend; ii- with a descending concentration trend; iii-an increase of concentration from condition A to B and then a decrease from B to C; iv-a decrease of concentration from condition A to B and then a increase from B to C) were considered to better provide information of the tumour invasion (Fig. 2A).

Among these differentially expressed proteins, only three of them (uromodulin, Ras-related protein Ral-A, Beta-Ala-His dipeptidase), diminished proportionally to RCC infiltration while twenty-six proteins seems to be positively influenced by the increase of renal vein involvement (Fig. 2A). The remaining 318 proteins showed a positive (268) or negative (51) variations of their fold changes at the beginning of the invasion inside the lumen of the vein.

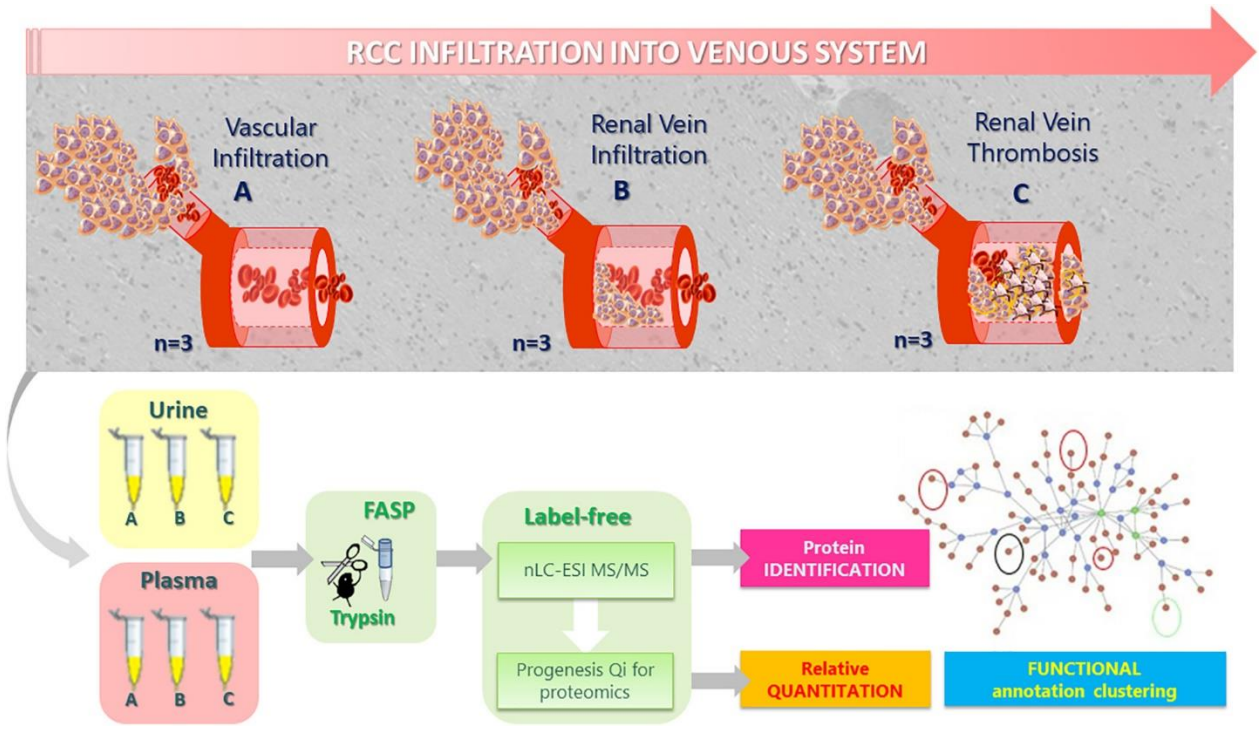


Fig. 1. Experimental design and workflow.

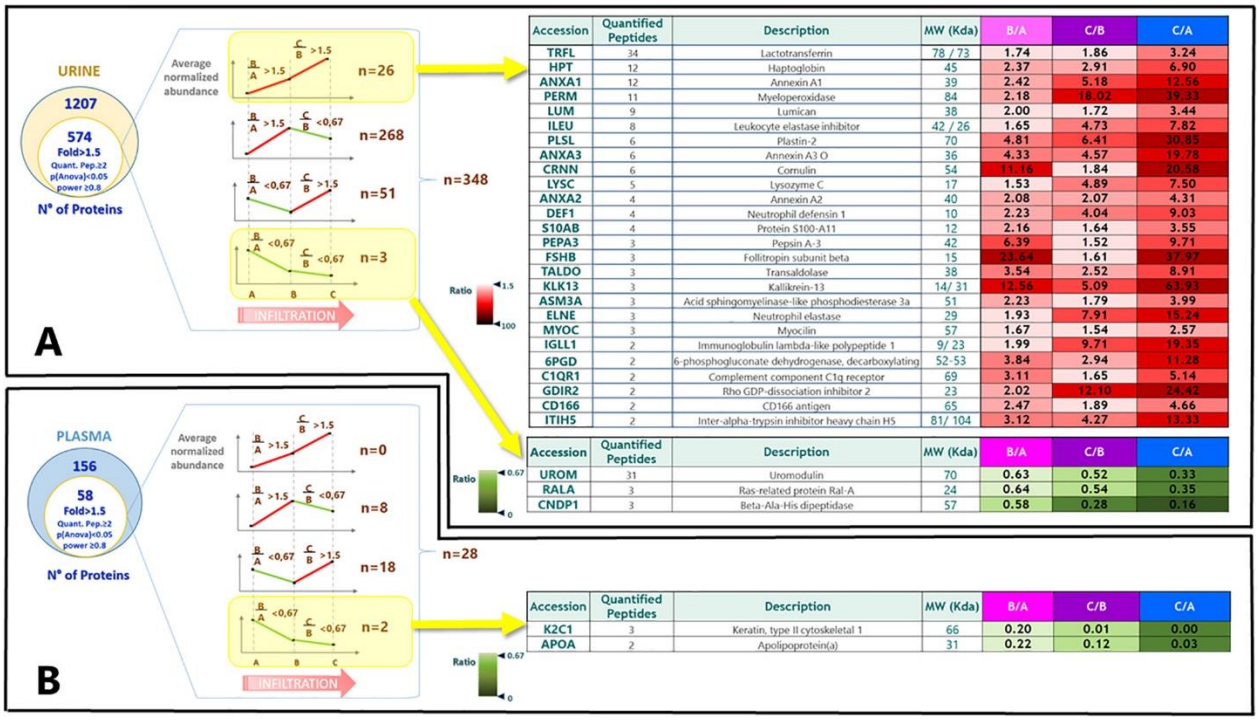
3.2.2. Plasma proteome changes

Plasma samples pools were also investigated by label-free LC-ESIMS/MS relative quantitation in order to highlight differentially expressed proteins in response to RCC vascular infiltration, similarly to previously described approach on urine.

156 different proteins were identified (Supplemental Table 4) and, among them, 58 DEPs were found filtering based on following criteria: at least 2 unique peptides; fold change ≥ 1.5; Anova test pvalue ≤ 0.05; power ≥ 0.8; q-values ≤ 0.05 (Supplemental Table 5). The number of proteins identified in each of the runs was reported in Supplemental Fig. 1.

Twenty-eight proteins were present in the fourth group according to their fold changes calculated comparing the three studied conditions as above described for urine (Fig. 2B). None of them showed an increase of their abundance consistently to the augment of RCC infiltration and only two, including Apolipoprotein A-1 and an isoform

belonging to keratin family (type II cytoskeletal 1), appeared to be inversely correlated to renal vein invasion (Fig. 2B).



Accession	Quantified Peptides	Description	MW (Kda)	B/A	C/B	C/A
TRFL	34	Lactotransferrin	78 / 73	1.74	1.86	3.24
HPT	12	Haptoglobin	45	2.37	2.91	6.90
ANXA1	12	Annexin A1	39	2.42	5.18	12.56
PERM	11	Myeloperoxidase	84	2.18	18.02	39.33
LUM	9	Lumican	38	2.00	1.72	3.44
ILEU	8	Leukocyte elastase inhibitor	42 / 26	1.65	4.73	7.82
PLSL	6	Plastin-2	70	4.81	6.41	30.85
ANXA3	6	Annexin A3 O	36	4.33	4.57	19.78
CRNN	6	Cornulin	54	11.16	1.84	20.58
LYSC	5	Lysozyme C	17	1.53	4.89	7.50
ANXA2	4	Annexin A2	40	2.08	2.07	4.31
DEF1	4	Neutrophil defensin 1	10	2.23	4.04	9.03
S10AB	4	Protein S100-A11	12	2.16	1.64	3.55
PEPA3	3	Pepsin A-3	42	6.39	1.52	9.71
FSHB	3	Follitropin subunit beta	15	23.64	1.61	37.97
TALDO	3	Transaldolase	38	3.54	2.52	8.91
KLK13	3	Kallikrein-13	14/ 31	12.56	5.09	63.93
ASM3A	3	Acid sphingomyelinase-like phosphodiesterase 3a	51	2.23	1.79	3.99
ELNE	3	Neutrophil elastase	29	1.93	7.91	15.24
MYOC	3	Myocilin	57	1.67	1.54	2.57
IGLL1	2	Immunoglobulin lambda-like polypeptide 1	9/ 23	1.99	9.71	19.35
6PGD	2	6-phosphogluconate dehydrogenase, decarboxylating	52-53	3.84	2.94	11.28
C1QR1	2	Complement component C1q receptor	69	3.11	1.65	5.14
GDIR2	2	Rho GDP-dissociation inhibitor 2	23	2.02	12.10	24.42
CD166	2	CD166 antigen	65	2.47	1.89	4.66
ITIH5	2	Inter-alpha-trypsin inhibitor heavy chain H5	81/ 104	3.12	4.27	13.33

Accession	Quantified Peptides	Description	MW (Kda)	B/A	C/B	C/A
UROM	31	Uromodulin	70	0.63	0.52	0.33
RALA	3	Ras-related protein Ral-A	24	0.64	0.54	0.35
CNDP1	3	Beta-Ala-His dipeptidase	57	0.58	0.28	0.16

Accession	Quantified Peptides	Description	MW (Kda)	B/A	C/B	C/A
K2C1	3	Keratin, type II cytoskeletal 1	66	0.20	0.01	0.00
APOA	2	Apolipoprotein(a)	31	0.22	0.12	0.03

Fig. 2. Protein expression trend correlated to RCC extension into RV (RV = Renal Vein, A = ccRCC patients that show vascular infiltration, B = ccRCC patients that show tumour infiltration into renal vein; C = ccRCC patients that show renal vein thrombosis) in urine (panel A) and in plasma (panel B).

3.2.3. Comparison between urine and plasma proteome alterations

Urinary protein content derived from patients affected by ccRCC at different vascular infiltration levels was compared with related plasma proteome belonged to the same cohort of patients.

Eighty proteins identified with at least 2 unique peptides were found to be shared between urine and plasma datasets, equivalent to about 75% of all plasmatic proteins and about 11% of urinary proteins detected in the sample pools (Fig. 3).

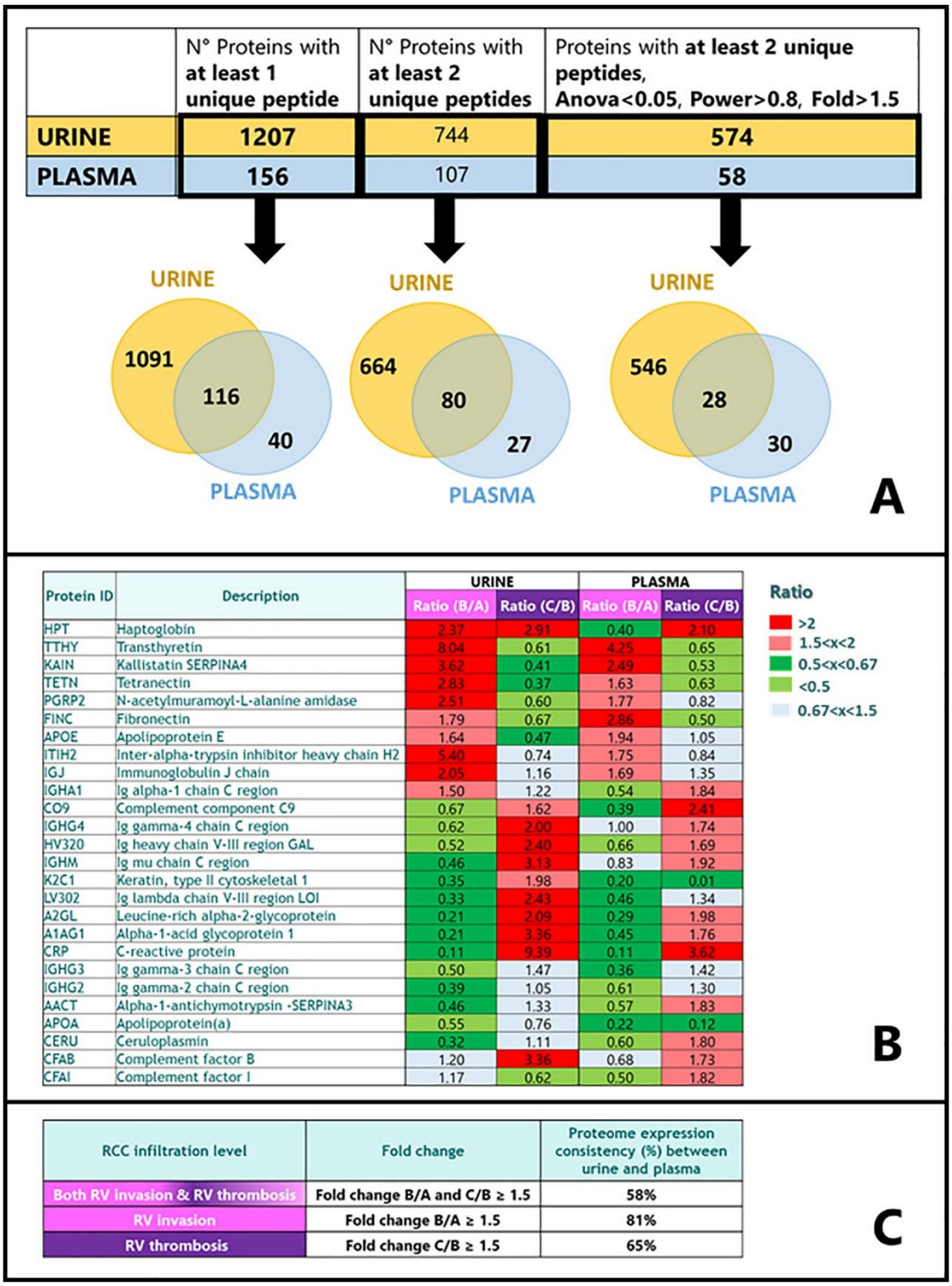


Fig. 3. A- Number of proteins isoforms identified and quantified in urine and plasma samples through nLC-ESI MS/MS. Data were elaborated through Progenesis platform and a Venn Diagram between urine and plasma outcomes is shown both for identified proteins and for differentially regulated proteins (Ratio ≥ 1.5). B- DEPs in common between urine and plasma samples. C- Proteome expression consistency between urine and plasma samples regarding shared DEPs.

Among the common protein IDs, 28 shown a significant variation of their expression (fold change ≥ 1.5), comparing the three conditions (A = vascular infiltration; B = RV infiltration; C = RV thrombosis) (Fig. 3A). In this subset, two proteins, Complement C1s subcomponent and Immunoglobulin heavy constant alpha 2, were shown to be varied in urine only in the comparison of RCC renal vein obstruction (C) respect than the initial RCC vascular invasion (A) (Ratio ≥ 1.5 or ≤ 0.67). The remaining 26 differentially modulated during RV infiltration and RV thrombosis phase were listed in Fig. 3B. A high level of concordance of ranging from 58% to 81% has been observed comparing the expression trend (up- or down-regulation) of proteins between urine and plasma (Fig. 3C). This coherence is remarkably higher (81%) considering only changes belonging to RV invasion (B/A).

From a functional point of view, proteins present in this panel are involved mainly in immune-system process and defense (Lactotransferrin, Haptoglobin, Annexin A1 Myeloperoxidase, Leukocyte elastase inhibitor, Plastin-2, Annexin A3, Lysozyme C, Annexin A2, Neutrophil defensin 1, Protein S100-A11, Neutrophil elastase, Immunoglobulin lambda-like polypeptide 1, Complement component C1q receptor, CD166 antigen). Some of them are likely to be associated also to protein binding (Annexin A1, Annexin A3, Annexin A2, Protein S100-A11), and to pentose phosphate pathway (Transaldolase and 6-phosphogluconate dehydrogenase, decarboxylating).

3.3. Functional and network analysis

3.3.1. Biological processes modulated by RCC vascular invasion in urine and plasma protein

A meta-analysis based on functional annotation tools was performed in order to highlight which biological process or pathways altered depending on the free RCC extension into renal vein (RV), are reflected by the two biofluids.

For this purpose, only proteins with significant changes in their abundance (574 IDs in urine and 58 IDs in plasma as shown in Fig. 2A) were included. Moreover, to better isolate the changes related to RV invasion, plasma and urine proteomes were grouped into four datasets for each biofluid, taking into account the three possible levels of infiltration based on the experimental design (Fig. 1). In particular, the lists of DEPs were divided considering only those proteins that resulted consistently down- or up-regulated in RV invasion (condition B respect than condition A), and in RV thrombosis (condition C respect than condition B), as displayed in Supplemental Fig. 2.

Initially, these eight lists were separately submitted to a statistical overrepresentation test on PANTHER gene analysis tool for pinpointing the most significant biological processes enriched during RV infiltration and thrombosis and for evaluating the degree of coherence between urine and plasma proteome from a functional point of view (Fig. 4). The biological processes in particular were grouped into macro-categories, according to GO-term classification, with the aim of better detecting typical functional traits characterizing tumour vascular invasion steps, and if these traits were represented similarly in urine and plasma.

As displayed in Fig. 4, most of the bioprocesses varied in urine overlapped with those found altered in plasma. Among these shared categories, no inconsistent pattern was shown.

3.3.2. Ingenuity pathway analysis of the liquid biopsies proteomes

IPA software was used to deeply explore functions and pathways that resulted differently modulated in the biofluids in response to RCC infiltration into the renal vein (Supplemental Fig. 3-6). Similarly to the previous analysis by Panther search, a comparison between urine and plasma DEPs was carried on considering both the changes occurred in patients with RV invasion and in those whose RV was obstructed. As shown in Supplemental Fig. 3 and Supplemental Fig. 4, several functional features

were shared between patients showing evidence of RCC infiltration in RV, independently from the presence of RV thrombus. However, it has to be noticed that some GO-terms, including networks and molecular functions, appear to be more specific of the level of vein invasion. On the other hand, there was a remarkable overlap in the comparison of the information gathered in urine versus the one received from plasma (Supplemental Fig. 5 and Supplemental Fig. 6). This level of overlapping is very high in the case of the physiological system development and functions section, while it tends to disappear for top networks and top upstream regulators. Moreover, in terms of pathways consideration, and disease and biofunctions, the concordance between the two liquid biopsies was slightly higher for RV invasion patients in respect to RV thrombosis subjects. On the other hand, the concordance increased in RV thrombosis samples in the case of molecular and cellular functions.

4. Discussion

The kidney, urine and plasma proteomes are not isolated compartments, rather, they are closely related and could be considered an interconnected system: kidney filtered plasma proteins and waste products into urine via excretory system, and furthermore renal cells may secrete proteins directly into blood or release them into the urine. Many large scale shotgun analyses have investigated the proteome of these specimens, but only few of them have compared the information gathered from these sources, especially for the biofluids [23]. The human proteome atlas for kidney, urine and plasma described by Farrah et al. has been built assembling proteins identified using different sources (e.g. glomerulus, urinary exosomes, urine from healthy subjects, etc.) and different analytical approach. These databases result certainly useful because they provide a reliable storage of proteins of different origin.

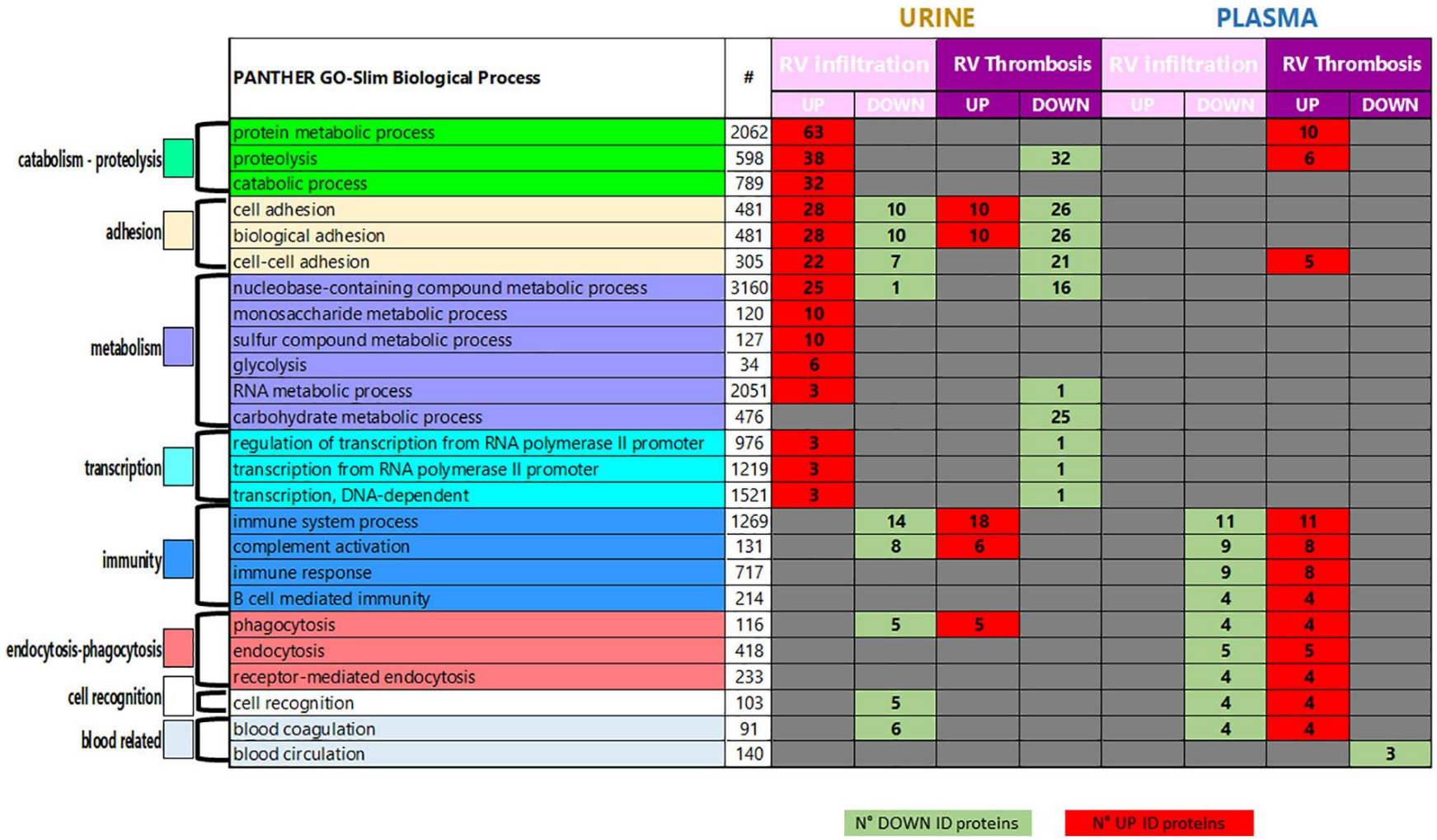


Fig. 4. Overview of the biological processes enriched in urine and plasma in RCC patients with RV infiltration or RV thrombosis, using overrepresentation test on PANTHER gene analysis tool ($p < 0.05$). The up- or down-regulation referred to condition B respect to condition A for RV infiltration label (pink), and condition C respect to condition B for RV Thrombosis label (purple). # = number of genes belonging to related PANTHER GO-Slim Biological process.

However, they do not compare the proteome from different specimens belonging to the same subject, do not provide information about the relative changes of these proteins in different conditions and do not clarify which pathways or network are more represented comparing blood and urine. Moreover, the proteome data used for comparing the specimens were often obtained by different databases and by diverse analytical methods.

One of the first work concerning the integration of more than one specimen was provided by RF Andersen and co-workers [24]. The authors through a nano-LC-MS/MS quantitative approach based on iTRAQ labelling identified DEPs in urine and plasma during childhood idiopathic nephrotic syndrome (NS) compared with remission. About 149 proteins were found to be present in both the biofluids, although none of these shared proteins were observed as significantly altered following NS remission. Li et al. investigated urinary and plasmatic proteome by LC-MS/MS to determine the best source for a more sensitive detection of protein markers characterizing the effects of two anticoagulants (heparin or argatroban) in six SD rats before and after treatments [7]. Recently, Welton et al. applying a semi-quantitative aptamer-based protein array, identified about 1000 proteins, of which almost 400 were present at comparable quantities in plasma in respect to urinary vesicles [8]. Concerning the study of kidney, data integration between different biofluids was concerned more with the study of physiology of this organ [25] than being finalized to enrich the molecular scenario of a specific disease, as renal cancer.

Herein, for the first time we applied a shotgun label-free LC MS/MS approach to compare the proteomes between urine and plasma that derives from the same ccRCC patient cohort with different levels of tumour infiltration into the renal vein, from the vascular invasion without the involvement of the RV to the complete occlusion of this vessel. Plasma and urine were collected selecting three subjects for each the three conditions in a wide cohort of RCC patients using a stringent criterion of inclusion. The

appropriateness and the risk of the biological averaging assumption in sample pooling must be seriously take into account during the choice of the proteomic workflow, especially for investigations involving class discovery and class prediction in the context of diagnostics and prognostics analysis [26]. Due to the low number of subjects showing the desired defined tumour characteristics, and to the nature of the study far from being diagnostic- or prognosticoriented, samples were pooled based on the condition and the specimen of origin. For each sample pool at least three technical replicates were analysed and statistical thresholds were considered as described in methods section. Moreover, plasma was preferred to serum and collected avoiding unnecessary manipulation (e.g. depletion), in order to make the biofluid comparison more reliable and reduce exogenous modifications.

Overall, 1207 and 156 proteins were identified in urine and in plasma, respectively; while 574 urinary and 58 plasma protein IDs were observed as differentially expressed in at least one of the three conditions. The discrepancy of identification power ~10-fold higher in urine could be ascribe to the higher protein dynamic range in blood compared to urine. A lot of approaches have been applied to overcome this issue, including depletion of high abundant proteins (such as albumin or IgG), often coupled with different strategies of off-line peptide fractionation [27,28]. However, to limit the variability of the results that could be introduced by removing the most abundant proteins, and also to keep the quantitation more reliable, we decided to analyse the two biofluids using the same protocol. The message that urine reflect more information remains valid despite the low identification power for plasma.

A panel of 26 urinary proteins were found to be directly correlated with the extension of RCC into RV, showing an increase in their abundance levels, parallel with those related to the infiltration level. This trend was not observed in plasma, probably due to the limited protein number. The panel of positive markers in urine is largely composed by proteins implicated in biological process that seem strongly related to the tumour

invasion, inflammatory process, and energetic metabolism, as described in results paragraph 3.2.3. Only three proteins (uromodulin, Ras-related protein Ral-A, Beta-Ala-His dipeptidase) in urine and two in plasma (Apolipoprotein A-1and a keratin type II cytoskeletal 1) are negatively influenced by the increase of infiltrative process. Interestingly, Ral-A GTPase was reported to be associated with advanced kidney cancer, being involved into malignancy invasion processes, through a signal pathway induced by proinflammatory cytokine prostaglandin E2 (PGE2) [29]. Similarly, Apolipoprotein A-1 (APOA1) is shown to be correlated with RCC prognosis in agreement with the findings of a recent investigation that demonstrated in a retrospectively study of 786 patients with RCC that a low APOA1 serum level has been associated to a worse overall survival and to shortened disease-free survival [30]. Moreover, activity of Beta-Ala-His dipeptidase encoded by CNDP1 gene has been observed to be correlated with a potential long term protection of complications linked to reactive metabolites accumulating, e.g. in diabetes and chronic renal failure [31].

The comparison of protein content identified in the two biofluids highlighted a huge overlap between plasma and urine, being about three-quarters of all plasmatic proteins included in urine dataset and about half of DEPs found in plasma (Fig. 3A). This could be expected since urine is mainly the result of blood filtration encompassing the most abundant and less represented proteins. However, this overlap allowed us to gain a new insight from a different perspective about the pathological processes connected to the RCC vascular invasion.

The list of biofluid-shared DEPs, included a high percentage of immunoglobulin (about 30%), components and factors of complement cascade, modulators of acute-phase response and defense involved in complement activation, innate immune-system, platelet degranulation and scavenging of heme from plasma (Supplemental Table 6).

Overall, they are consistently regulated if we compared plasma and urine, and some of them appear to be significantly related to the infiltration grade of RCC (Fig. 3B).

However, if we perform this comparison considering the different steps of the extension of RCC, the percentage of variability appear different. In fact, more advanced is the stage of vascular invasion, more discrepancies between blood and urine are present (Fig. 3C). A possible explanation for this behaviour could be found considering the 3D-development of the tumour mass. In fact, urinary proteome alterations could be influenced by the extent of RCC not only into vascular system but also into the organ itself, which can lose its architecture. If kidney structure and function is damaged, it can be supposed that non-regulated protein deliveries to renal basin can be present. If we consider the urinary albumin level, generally associated with functional status of the glomerular filtration barrier, no statistically significant proteomic variation is detectable comparing the three groups. However, the histological exams displayed a very advanced tumour progression in the three patients showing RV thrombosis. Therefore, an impairment of kidney functionality cannot be excluded in the surrounding tumour area. Alterations associated with disease generally require an entire set of effectors to be completed. By now, modern proteomic approaches despite conspicuous advancements can provide only a partial list of them. A comprehensive study of regulatory networks and pathways could compensate these lacks and effectively increase the understanding of the intricate system of functions that are turned on or off during disease process. This contribute is more evident if we apply an integrated strategy. Therefore, DEPs datasets found in urine and plasma were subjected to a functional classification and outcomes were visualized filtering both RV infiltration steps and biofluid type. Firstly, it can be noticed that the biological processes shared between urine and plasma showed a high grade of concordance in terms of positive or negative regulation (Supplemental Fig. 7). No inconsistency was revealed between urine and plasma supporting the hypothesis that urine is a good mirror of what is happening in blood. Moreover, it is likely that both urine and blood carry a specific 'biofluid functional signature'. Several processes were enriched mainly (e.g. glucose and RNA metabolism, catabolic process,

adhesion), or exclusively in urine (such as transcription related processes). Others are likely more represented in plasma (e.g. B-cell mediated immunity or blood coagulation). Furthermore, if we integrate the information carried by the two biofluids, also a functional signatures associated to biological processes could be mined. To better visualize them, the enriched biological processes were categorized into macro-groups using PANTHER gene-ontology (Fig. 4). Results suggest specific trends characterised by processes that are on or off depending on the infiltration phase. Proteins related to catabolic processes, proteolysis and cell to cell adhesion were up-represented both in RV infiltration and RV thrombosis, while energetic metabolic processes including glycolysis and regulation of transcription appeared on during RV invasion and down-represented when RV is occluded. On the other hand, in patients with the evidence of RCC thrombosis, immunity system related proteins, including those involved in the complement activation, and defense mechanisms, endocytosis and cell recognition were found significantly increased and proteins implicated in blood circulation decreased.

These data were also confirmed by the functional annotation classification provided by Ingenuity Pathway analysis (Supplemental Fig. 3- 6). Despite a certain overlap between urine and plasma, only the combination of the two datasets permits to highlight specific traits of renal cancer vascular invasion. On one side, the analysis underlined as already suggested that ccRCC is basically a metabolic disorder, since malignant cells handle a number of biological pathways to achieve their aggressive phenotype and spread into circulating system [32]. On the other side, particularly for these infiltrative forms, ccRCC is likely to behave as an immunological disease, involving immune cell trafficking, humoral immune response and positive and negative acute response.

In conclusion, the comparison between the functional classification of urine and plasma proteome confirms the complementary of the information delivered by these biofluids and shed light to those processes and pathways that are likely to be switched on or off during malignancy spreading into renal vein. On the other hand, it suggests that the loss

of kidney architecture during advanced stages of tumour growth could have a detectable repercussion to biofluids proteome. Therefore the integration of information between urine and plasma changes at a proteomic level may provide a more complete landscape of such a dynamic system as growing cancer cells are, also from a functional point of view. Supplementary data to this article can be found online at https://doi.org/10.1016/j.jprot.2018.04.029.

Acknowledgments This work was supported by grants from the MIUR: FIRB 2007 (RBRN07BMCT_11), FAR 2013–2016; and in part by Fondazione Gigi & Pupa Ferrari Onlus.

References

[1] M. Rainer, C. Sajdik, G.K. Bonn, Mass spectrometric profiling of low-molecular weight proteins, Methods Mol. Biol. 1023 (2013) 83–95, http://dx.doi.org/10. 1007/978-1-4614-7209-4-5.

[2] C. Chinello, V. L'Imperio, M. Stella, A.J. Smith, G. Bovo, A. Grasso, M. Grasso, F. Raimondo, M. Pitto, F. Pagni, F. Magni, The proteomic landscape of renal tumors, Expert Rev. Proteomics. 13 (2016), http://dx.doi.org/10.1080/14789450.2016. 1248415.

[3] A. Di Meo, I. Batruch, A.G. Yousef, M.D. Pasic, E.P. Diamandis, G.M. Yousef, An integrated proteomic and peptidomic assessment of the normal human urinome, Clin. Chem. Lab. Med. 55 (2017) 237–247, http://dx.doi.org/10.1515/cclm-2016- 0390.

[4] H. Husi, R.J.E. Skipworth, A. Cronshaw, K.C.H. Fearon, J.A. Ross, Proteomic identification of potential cancer markers in human urine using subtractive analysis, Int. J. Oncol. 48 (2016), http://dx.doi.org/10.3892/ijo.2016.3424.

[5] M. Li, Urine reflection of changes in blood, Adv. Exp. Med. Biol. 845 (2015) 13–19, http://dx.doi.org/10.1007/978-94-017-9523-4_2.

[6] Y. Gao, Urine is a better biomarker source than blood especially for kidney diseases, Adv. Exp. Med. Biol. 845 (2015) 3–12, http://dx.doi.org/10.1007/978-94-017- 9523-4_1.

[7] M.L. Li, M.D. Zhao, Y.H. Gao, Changes of proteins induced by anticoagulants can be more sensitively detected in urine than in plasma, Sci. China Life Sci. 57 (2014) 649–656, http://dx.doi.org/10.1007/s11427-014-4661-y.

[8] J.L. Welton, P. Brennan, M. Gurney, J.P. Webber, L.K. Spary, D.G. Carton, J.M. Falcón-Pérez, S.P. Walton, M.D. Mason, Z. Tabi, A. Clayton, Proteomics analysis of vesicles isolated from plasma and urine of prostate cancer patients using a multiplex, aptamer-based protein array, J. Extracell. Vesicles. 5 (2016), http://dx. doi.org/10.3402/jev.v5.31209.

[9] A. Katafigioti, I. Katafigiotis, S. Sfoungaristos, C. Alamanis, K. Stravodimos, I. Anastasiou, E. Roumelioti, M. Duvdevani, C. Constantinides, In the search of novel urine biomarkers for the early diagnosis of prostate cancer. Intracellular or secreted proteins as the target group? Where and how to search for possible biomarkers useful in the everyday clinical practice, Arch. Ital. Di Urol. E Androl. 88 (2016) 195–200, http://dx.doi.org/10.4081/aiua.2016.3.195.

[10] L. Jia, Comparing plasma and urinary proteomes to understand kidney function, Adv. Exp. Med. Biol. 845 (2015) 187–193, http://dx.doi.org/10.1007/978-94-017-9523-4_18.

[11] H. Miyake, T. Terakawa, J. Furukawa, M. Muramaki, M. Fujisawa, Prognostic significance of tumor extension into venous system in patients undergoing surgical treatment for renal cell carcinoma with venous tumor thrombus, Eur. J. Surg. Oncol. 38 (2012) 630–636, http://dx.doi.org/10.1016/j.ejso.2012.03.006.

[12] Q. Tang, Y. Song, X. Li, M. Meng, Q. Zhang, J. Wang, Z. He, L. Zhou, Prognostic outcomes and risk factors for patients with renal cell carcinoma and venous tumor thrombus after radical nephrectomy and thrombectomy: the prognostic significance of venous tumor thrombus level, Biomed. Res. Int. 2015 (2015), http://dx.doi.org/10.1155/2015/163423.

[13] M.W. Ball, M.A. Gorin, K.T. Harris, K.M. Curtiss, G.J. Netto, C.P. Pavlovich, P.M. Pierorazio, M.E. Allaf, Extent of renal vein invasion influences prognosis in patients with renal cell carcinoma, BJU Int. 118 (2016) 112–117, http://dx.doi.org/10.1111/bju.13349.

[14] V.L. Weiss, M. Braun, S. Perner, A. Harz, R. Vorreuther, G. Kristiansen, S.C. Müller, J. Ellinger, Prognostic significance of venous tumour thrombus consistency in patients with renal cell carcinoma (RCC), BJU Int. 113 (2014) 209–217, http://dx.doi.org/10.1111/bju.12322.

[15] A. Laird, F.C. O'Mahony, J. Nanda, A.C.P. Riddick, M. O'Donnell, D.J. Harrison, G.D. Stewart, Differential expression of prognostic proteomic markers in primary tumour, venous tumour Thrombus and metastatic renal cell Cancer tissue and correlation with patient outcome, PLoS One 8 (2013), http://dx.doi.org/10.1371/journal.pone.0060483.

[16] N. Bosso, C. Chinello, S.C.M. Picozzi, E. Gianazza, V. Mainini, C. Galbusera, F. Raimondo, R. Perego, S. Casellato, F. Rocco, S. Ferrero, S. Bosari, P. Mocarelli, M.G. Kienle, F. Magni, Human urine biomarkers of renal cell carcinoma evaluated by ClinProt, Proteomics: Clin. Appl. 2 (2008) 1036–1046, http://dx.doi.org/10.1002/prca.200780139.

[17] J.R. Wisniewski, A. Zougman, N. Nagaraj, M. Mann, Universal sample preparation method for proteome analysis, Nat. Methods 6 (2009) 359–362, http://dx.doi.org/10.1038/nmeth.1322\rnmeth.1322 (pii).

[18] F. Raimondo, S. Corbetta, A. Savoia, C. Chinello, M. Cazzaniga, F. Rocco, S. Bosari, M. Grasso, G. Bovo, F. Magni, M. Pitto, Comparative membrane proteomics: a technical advancement in the search of renal cell carcinoma biomarkers, Mol. BioSyst. 11 (2015), http://dx.doi.org/10.1039/c5mb00020c.

[19] C. Chinello, M. Cazzaniga, G. Sio, A.J. Smith, A. Grasso, B. Rocco, S. Signorini, M. Grasso, S. Bosari, I. Zoppis, G. Mauri, F. Magni, Tumor size, stage and grade alterations of urinary peptidome in RCC, J. Transl. Med. 13 (2015), http://dx.doi.org/10.1186/s12967-015-0693-8.

[20] X. Liu, C. Chinello, L. Musante, M. Cazzaniga, D. Tataruch, G. Calzaferri, A.J. Smith, G. De Sio, F. Magni, H. Zou, H. Holthofer, Intraluminal proteome and peptidome of human urinary extracellular vesicles, Proteomics: Clin. Appl. 9 (2015) 568–573, http://dx.doi.org/10.1002/prca.201400085.

[21] H. Mi, A. Muruganujan, P.D. Thomas, Panther in 2013: Modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees, Nucleic Acids Res. 41 (2013), http://dx.doi.org/10.1093/nar/gks1118.

[22] B. Ljungberg, K. Bensalah, S. Canfield, S. Dabestani, F. Hofmann, M. Hora, M.A. Kuczyk, T. Lam, L. Marconi, A.S. Merseburger, P. Mulders, T. Powles, M. Staehler, A. Volpe, A. Bex, EAU guidelines on renal cell carcinoma: 2014 update, Eur. Urol. 67 (2015) 913–924, http://dx.doi.org/10.1016/j.eururo.2015.01.005.

[23] T. Farrah, E.W. Deutsch, G.S. Omenn, Z. Sun, J.D. Watts, T. Yamamoto, D. Shteynberg, M.M. Harris, R.L. Moritz, State of the human proteome in 2013 as viewed through peptide atlas: comparing the kidney, urine, and plasma proteomes for the biology- and disease-driven human proteome project, J. Proteome Res. 13 (2014) 60–75, http://dx.doi.org/10.1021/pr4010037.

[24] R.F. Andersen, J. Palmfeldt, B. Jespersen, N. Gregersen, S. Rittig, Plasma and urine proteomic profiles in childhood idiopathic nephrotic syndrome, Proteomics: Clin. Appl. 6 (2012) 382–393, http://dx.doi.org/10.1002/prca.201100081.

[25] L. Jia, L. Zhang, C. Shao, E. Song, W. Sun, M. Li, Y. Gao, An attempt to understand kidney's protein handling function by comparing plasma and urine proteomes, PLoS One 4 (2009), http://dx.doi.org/10.1371/journal.pone.0005146.

[26] A.L. Oberg, O. Vitek, Statistical design of quantitative mass spectrometry-based proteomic experiments, J. Proteome Res. (2009), http://dx.doi.org/10.1021/pr8010099.

[27] H. Keshishian, M.W. Burgess, H. Specht, L. Wallace, K.R. Clauser, M.A. Gillette, S.A. Carr, Quantitative, multiplexed workflow for deep analysis of human blood plasma and biomarker discovery by mass spectrometry, Nat. Protoc. (2017), http://dx.doi.org/10.1038/nprot.2017.054.

[28] S. Chutipongtanate, S. Chatchen, J. Svasti, Plasma prefractionation methods for proteomic analysis and perspectives in clinical applications, Proteomics: Clin. Appl. (2017), http://dx.doi.org/10.1002/prca.201600135.

[29] Z. Li, Y. Zhang, W.J. Kim, Y. Daaka, PGE2 promotes renal carcinoma cell invasion through activated RalA, Oncogene 32 (2013) 1408–1415, http://dx.doi.org/10. 1038/onc.2012.161.

[30] S. Guo, X. He, Q. Chen, G. Yang, K. Yao, P. Dong, Y. Ye, D. Chen, Z. Zhang, Z. Qin, Z. Liu, Z. Li, Y. Xue, M. Zhang, R. Liu, F. Zhou, H. Han, The effect of preoperative apolipoprotein A-I on the prognosis of surgical renal cell carcinoma: a retrospective large sample study, Medicine (Baltimore) 95 (2016) e3147, , http://dx.doi.org/10. 1097/MD.0000000000003147.

[31] V. Peters, J. Zschocke, C.P. Schmitt, Carnosinase, diabetes mellitus and the potential relevance of carnosinase deficiency, J. Inherit. Metab. Dis. (2017), http://dx. doi.org/10.1007/s10545-017-0099-2.

[32] A. Zaravinos, M. Pieri, N. Mourmouras, N. Anastasiadou, I. Zouvani, D. Delakas, C. Deltas, Altered metabolic pathways in clear cell renal cell carcinoma: a metaanalysis and validation study focused on the deregulated genes and their associated networks, Oncoscience. 1 (2014) 117, http://dx.doi.org/10.18632/oncoscience.13.

CHAPTER 3

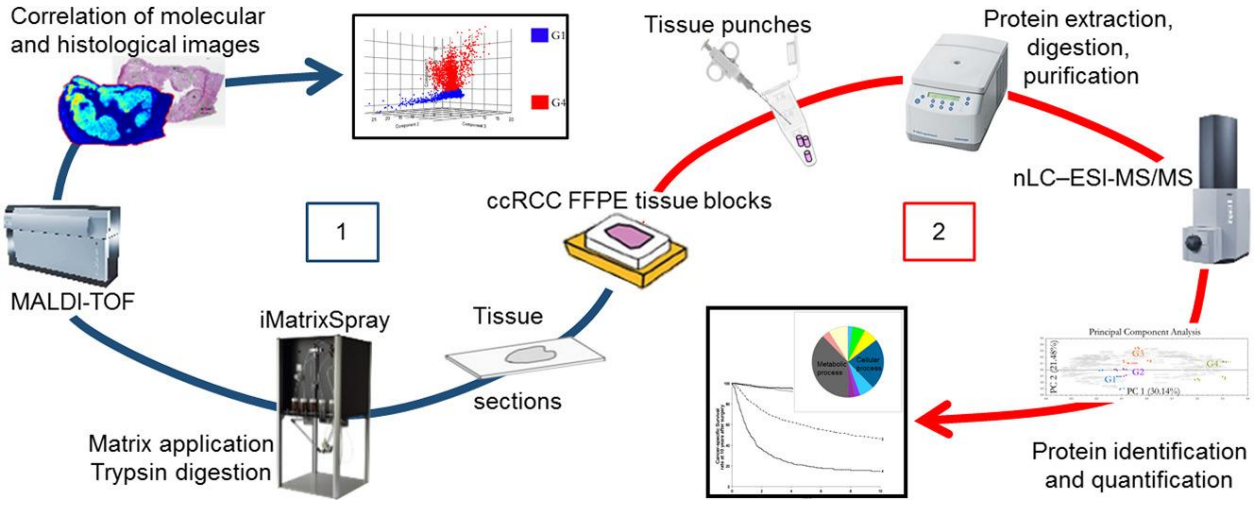
Histology-guided proteomic analysis to investigate the molecular profiles of clear cell Renal Cell Carcinoma grades

Martina Stella[a1]; CliziaChinello[a1]; Anna Cazzaniga[a]; Andrew Smith[a]; Manuel Galli[a]; Isabella Piga[a]; Angelica Grasso[b]; Marco Grasso[b]; MarinaDel Puppo[a]; Marta Varallo[c]; Giorgio Bovo[d]; Fulvio Magni[a]

[a] Department of Medicine and Surgery, University of Milano-Bicocca, Clinical Proteomics and Metabolomics Unit, Vedano al Lambro, Italy
[b] Urology Unit, S.Gerardo Hospital, Monza, Italy
[c] Pathology Unit, S.Gerardo Hospital, Monza, Italy
[d] Pathology Unit, Vimercate Hospital, Vimercate, Italy
[1]Equally contributing authors

Published: Journal of Proteomics, January 2019, Volume 191, Pages 38-47

ABSTRACT Renal Cell Carcinoma (RCC) is the most frequent form of kidney cancer and approximately 80% of cases are defined as clear cell RCC (ccRCC). Among the histopathological factors, tumour grade represents one of the most important parameters to evaluate ccRCC progression. Nonetheless, the molecular processes associated with the grading classification haven't been deeply investigated thus far. Therefore, the aim of this study was to uncover protein alterations associated with different ccRCC grade lesions. Formalin-fixed paraffin-embedded samples from ccRCC patients were analysed by histology-guided MALDI-MSI and shotgun proteomics in order to study the biological processes implicated in ccRCC. MALDI-MSI data highlighted signals able to discriminate among different grades (AUC > 0.8). The ion at m/z 1428.92 was identified as Vimentin and was overexpressed in grade 4 lesions, whereas ions at m/z 944.71, m/z 1032.78 and m/z 1325,99 were identified as histones H2A, H3, and H4, respectively. nLC-ESI-MS/MS analysis provided a further list of proteins and their abundances, showing a difference in protein content among the four grades. Moreover, the obtained molecular profiles showed a correspondence with the different Cancer-Specific Survival rate at 10 years post-surgery, as reported in literature.



Significance: Despite the generally accepted role of tumour grade in ccRCC diagnosis, the proteomic processes associated with the different tumour grades has not been extensively studied and doing so may provide insights into the development of the disease. In the current study, data obtained using MALDI-MSI was integrated with that obtained using nLC-ESI-MS/MS to highlight the proteomic alterations underlying the different ccRCC grades. The combined approach identified vimentin and three histones (H2A, H3 and H4) that were able to discriminate among the four grades whilst the nLC-ESI-MS/MS analysis alone provided a further list of proteins with an altered abundance. Furthermore, there was a good correlation between the molecular profiles generated for each grade and the different Cancer-Specific Survival rate at 10 years post-surgery. Such findings could be a valuable starting point for further studies aimed at clarifying the molecular events that occur during the development of ccRCC.

Keywords: Clear cell Renal Cell Carcinoma; Tumour grade; Mass spectrometry; Proteomics; ISUP grading system; MALDI-Imaging.

1. Introduction

Renal Cell Carcinoma (RCC) is the most frequent form of kidney cancer [1,2] and is comprised of three main subtypes: clear cell, papillary and chromophobe [3]. Approximately 80% of cases are defined as clear cell RCC (ccRCC) lesions and are further stratified by employing a combination of anatomical and histological parameters, which are also used as prognostic factors [4]. The tumour-node-metastasis (TNM) system is used to evaluate the stage of the disease by considering the dimension and position of the tumour (T), involvement of lymph nodes (N), and presence of metastasis (M) [5,6]. In addition to the anatomical features gathered in the TNM classification system, histological factors (tumour grade (G), sarcomatoid/rhabdoid differentiation, microvascular invasion,) influence the prognosis of clear cell Renal Cell Carcinoma. Among the histological parameters, tumour grade is one of the most important prognostic feature. For clear cell and papillary RCC, grade is now assessed according to the size and shape of the nucleoli: nucleolar prominence defines grades from 1 to 3, while extreme nuclear pleomorphism or sarcomatoid and/or rhabdoid differentiation determine a tumour as grade 4 [6–8]. Despite the generally accepted role of tumour grade in ccRCC prognosis, the molecular processes and protein alterations associated with each grade have not been extensively studied thus far. Most of the studies currently present in literature used the diagnostic classification that is based on the highest grade value observed in the tissue sections. However, considering the high heterogeneity of the tumour mass already demonstrated [9], their results could be affected by the most abundant grade, that could be different from that diagnosed.

Recent advancements in high-throughput technologies makes proteomics analyses applicable to a variety of samples (cell cultures, biofluids, tissues etc.) and most of them can be potentially used to profile renal tissue. Single cell populations or purified glomeruli are the samples of choice for investigating kidney filtration [10]. However,

those approaches may have limited application when looking at tumour lesions: cells in culture represent a single grade of cancer and thus they do not well represent the characteristic complexity of a tissue. Moreover, when proteins are extracted directly from a tissue, the information about altered abundance depending on the grade [11,12] could be lost due to tissue homogenisation. For these reasons, MALDI-MSI is the gold standard technique for the investigation of the inherent heterogeneity of tumour lesions, by preserving tissue integrity and protein alterations in-situ.

The aim of this work was to investigate the molecular traits of different ccRCC grades and the features that could drive grading progression, working on regions as homogeneous as possible. For this purpose, Formalin-Fixed Paraffin-Embedded (FFPE) tissue samples were analysed by complementary mass spectrometric techniques: Matrix-Assisted Laser Desorption/Ionisation Mass Spectrometry Imaging (MALDI-MSI) and nano Liquid Chromatography coupled with ElectroSpray Ionisation tandem Mass Spectrometry (nLC-ESI-MS/MS). MALDI-MSI allows the spatial distribution of proteins to be visualised within the tissue, therefore enabling the investigation of areas of tissue that are homogeneous for the grade and preventing artefacts that could occur due to the extraction of proteins from the entire tissue section. Conversely, nLC-ESI-MS/MS is employed for its greater capability to identify and quantify proteins and, thus, the combination of the two techniques enables information related with protein identity and quantity to be correlated with spatial localisation.

2. Material and methods

2.1. Materials

RapiGest SF surfactant was purchased from Waters Corporation (Waters, Milford, Massachusetts, USA). ZipTips were purchased from EMD Millipore (Billerica,

Massachusetts, USA), Indium Tin Oxide (ITO)- glass slides from Bruker Daltonics, Germany, Trypsin from porcine pancreas (Proteomics Grade, BioReagent, Dimethylated) and all the other reagents and solvents from Sigma-Aldrich (Chemie Gmbh, SBuchs, Switzerland).

2.2. Samples and sectioning

Samples from patients with a proven diagnosis of clear cell Renal Cell Carcinoma, collected between January 2011 and June 2016, were provided by San Gerardo Hospital (Monza, Italy) and included in the study. Diagnosis of the selected patients was performed after careful histopathological evaluation including grading (2009 ISUP classification system), sarcomatoid features, vascular invasion, tumour necrosis, and invasion of the collecting system and perirenal fat. The routine histological staining procedures were performed, including immunofluorescence and electron microscopy, with all these techniques required for the diagnosis. All the specimens included in the study were approved by the local ethic committee (Comitato Etico Azienda Ospedaliera San Gerardo, Monza, Italy), and analyses were carried out in agreement with the Declaration of Helsinki. After nephrectomy, the tissue was fixed according to standard routine methods, with a timing of 24 to 48 h for surgical specimens. Human ccRCC tissue samples were collected after nephrectomy from thirteen patients (Table 1) and FFPE tissue blocks were prepared using an automatic Tissue Processing Centre (TPC 15 Duo/Trio, Medite, MeBurgdorf, German). From each FFPE block, two types of samples were prepared: tissue sections for MALDI-MSI experiments and FFPE punches for nLC-ESI-MS/MS analysis. Tissue sections, with a thickness of 5 µm, were cut and mounted onto conductive ITO glass slides for MALDI-MSI analysis. Moreover, three biological replicates for each grade where used for nLC-ESI-MS/MS analysis. Each sample was made by three punches with a size of 1×1.5mm (~300,000 cells/sample), taken from the same tissue block.(G1=patients M,N,P; G2=patients

G,I,O; G3=patients E,F,I; G4=patients C,E,H; Table 1). FFPE tissue sections and punches were stored at room temperature until the day of the analysis.

Table 1. Clinical-pathological characteristic of the patients included in this study.

Sample	Patient	Age (at diagnosis) [year]	Gender	Greatest tumour dimension [cm]	pT	Stage	Grade
A	1	57	F	2.5	pT1a	1	G2
B	2	59	M	10	pT3a	4	G4
C,D	3	81	F	16.5	pT3a	3	G4
E	4	64	F	8	pT3a	3	G3
F	5	71	M	3.2	pT1a	1	G3
G	6	78	M	3.5	pT1a	1	G3
H	7	48	M	5,0	pT3a	4	G4
I	8	45	M	5.4	pT3a	4	G3
L	9	64	M	5	pT3a	3	G2
M	10	78	F	3	pT1a	1	G1
N	11	74	M	2.5	pT1a	1	G2
O	12	55	F	3.2	pT1a	1	G2
P	13	53	F	2.2	pT1a	1	G1

2.3. MALDI-MSI sample preparation and analysis

Paraffin removal, tissue rehydration and antigen retrieval were performed as previously described [13]. Briefly, tissue sections were washed with xylene, ethanol and water prior to antigen retrieval incubation. For protein digestion, uniform layers of trypsin were deposited using an iMatrix Spray automated spraying system (Tardo Gmbh, Subingen, Switzerland) with an optimised method, and tissues were then incubated overnight at 45 °C in a humid environment. After enzymatic digestion, the α-cyano-4-hydroxycinnamic acid matrix solution (10 mg/mL in 50,50 acetonitrile:water w/0.4% trifluoroacetic acid) was sprayed on the tissues using the iMatrixSpray. MALDI-MSI

analysis was performed using an UltrafleXtreme MALDI-TOF/TOF (Bruker Daltonics, Germany) equipped with a Smartbeam II laser operating at 2 kHz frequency. FlexControl software v3.4 was used to setup instrument parameters. External calibration of the instrument was performed with a mixture of standard peptides (PepMix I, Bruker Daltonics, Germany). The analyses were acquired in reflectron positive mode in the m/z range of 700 to 3500, with a laser diameter of 50 μm and a raster setting of 150 μm. Subsequently, MALDI-MS/MS analysis was performed only for signals of interest. In order to obtain the optimal mass value to be selected for dissociation during on-tissue MALDI-MS/MS, representative MALDI mass spectra were acquired in reflectron positive mode in the mass range of m/z 750 to 3500 using the random walk parameter of the sample carrier. A single precursor ion was selected by using the smallest precursor ion selector (PCIS) window possible and dissociated using laser-induced dissociation (LID) and LIFT™ technology, with the laser energy being set within a range of 40–70%. This process was performed until an MS/MS spectrum was obtained from the accumulation of ~700,000 laser shots. The resultant MALDI-MS/MS spectrum was submitted to Mascot 2.4.1 search engine (MatrixScience, London, UK) for protein identification. Then, the MALDI matrix was completely removed by washing the ITO glass with increasing concentrations of ethanol (90%, 95% and 100%) and the slide was dried before Hematoxylin and Eosin (H&E) staining. Tissue sections were then scanned at high resolution and converted to digital format using ScanScope CS Digital Scanner (Aperio, Park Center Dr., Vista, CA, USA) on which the pathologist annotated specific Regions Of Interest (ROIs) on the histological image. ROIs were selected by pathologist from regions that principally contained cells of the assigned grade (Fig. 1). However, the presence of small clusters of cell attributable to a different grade could not be completely excluded. A direct overlap of molecular and stained image was performed to integrate proteomic and morphologic data. ROIs were marked and grouped as follows: G1=8, G2=10, G3=16,

G4=10, normal cortex=8 and normal medulla=4 and were used to obtain the molecular profiles specific for each tumour grade. All the regions presenting a mixture of cells at different grades were excluded from the following analysis.

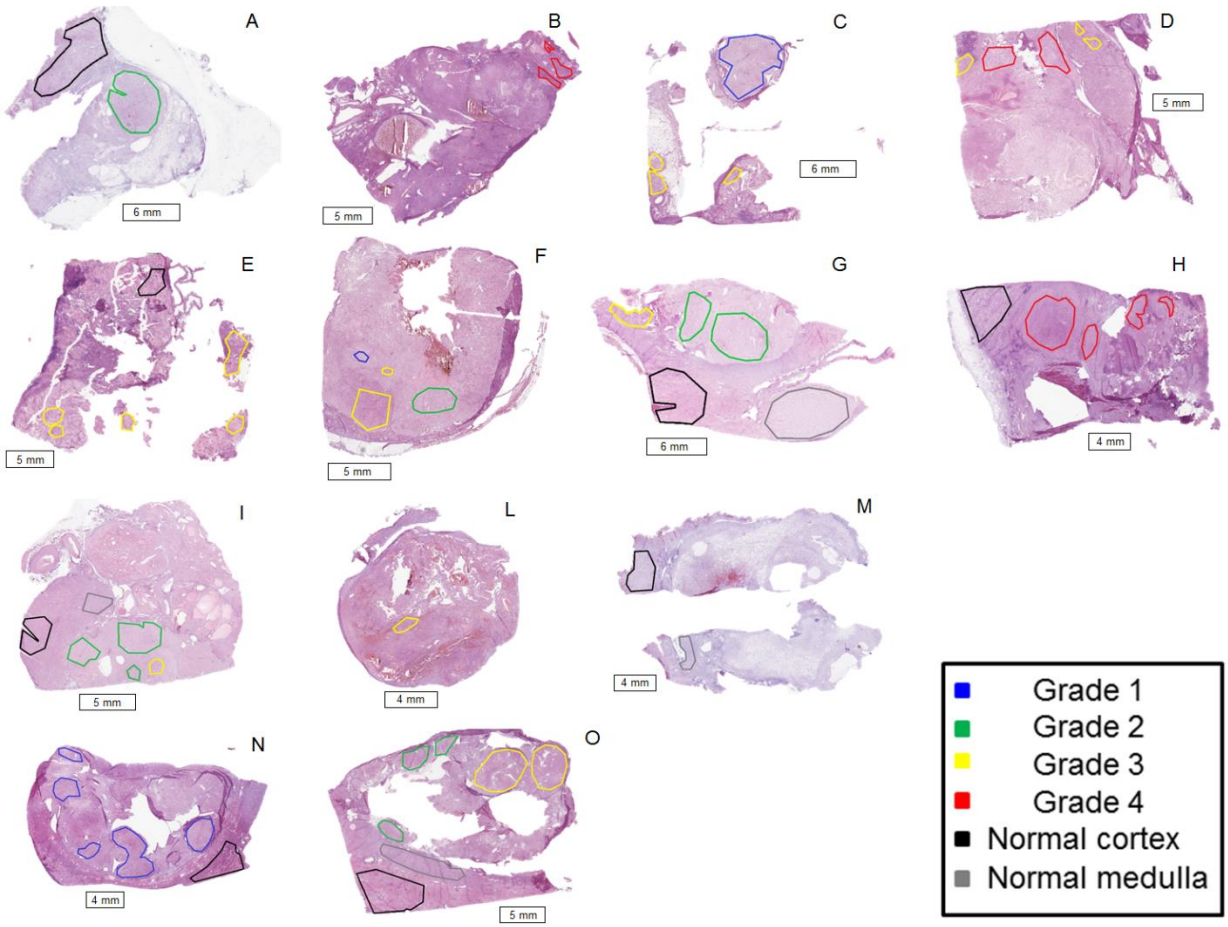


Fig. 1. Tissue samples from ccRCC patients. Hematoxylin and Eosin (H&E) stained image and annotations (grade 1: blue; grade 2: green; grade 3: yellow; grade 4: red; normal medulla: grey; normal cortex: black) performed by the pathologist on the tissue sections previously analysed by MALDI-mass spectrometry imaging.

2.3.1. MALDI-MSI data analysis

MALDI-MSI data were processed as previously described [14]. Briefly, ROIs from multiple FlexImaging 3.0 (Bruker Daltonics, Germany) analyses were imported into a single SCiLS lab 2014b file (http://scils.de/; Bremen, Germany), allowing the data from multiple acquisitions to be visualised simultaneously. Before statistical analyses, preprocessing steps were performed: baseline removal (iterative convolution) and normalization (Total ion current, TIC), peak picking (orthogonal matching pursuit algorithm), peak alignment (to align the detected ions with peak maxima), and spatial denoising. Principal Component Analysis (PCA) was also performed to reduce the high complexity of the data and visualize their variability. Finally, Receiver Operating Characteristic (ROC) analysis was performed, with an Area under the curve (AUC)>0.8 being required, as an additional criterion to the $p < 0.05$ calculated with Wilcoxon rank-sum test, for a peak to be considered as statistically significant. For MALDI-MS/MS spectra, baseline subtraction and smoothing were performed using FlexAnalysis 3.4 (Bruker Daltonics, Germany). MS/MS spectra were searched against the Swiss-Prot database (accessed July 2017; 555,100 sequences; 198,754,198 residues) with the Mascot 2.4.1 search engine. Mass tolerances were set at 100 ppm for MS and 1 Da for MS/MS. No enzymes or fixed post-translational modifications were set in the search parameters.

2.3.2. nLC-ESI-MS/MS sample preparation and analysis

Proteins were extracted from tissue punches for their identification and quantitation by nLC-ESI-MS/MS. Paraffin removal was performed by three consecutive washes in xylene (3 min each) followed by centrifugation (14,000 rpm, 5 min). Subsequently, consecutive washes in 100%, 90% and 70% of ethanol, for 4 min each, followed by centrifugation (14,000 rpm, 5 min) were performed. Then, antigen retrieval was carried out in a 10mM citric acid solution (pH=6) at 97 °C for 45 min. In order to enhance protein solubilisation and digestion, 100 µL of 0.1% RapiGest SF surfactant were added

to each sample. Reduction with 10mM DL-DiThioThreitol and alkylation with 15mM IodoAcetAmide were performed prior to enzymatic digestion done by adding 4 μg of trypsin to each specimen and by incubating samples over-night at 37 °C. The enzymatic reaction was blocked by acidification. Finally, desalting and concentration of the samples were performed using μ-C18 ZiptipTM pipette tips, following the standard protocol provided by Millipore. An aliquot of 7 μL containing 1 μg of proteins was then injected into a Dionex UltiMate 3000 rapid separation (RS) LC nano system (Thermo Scientific, Germany), coupled online to an Impact HD™ Ultra High Resolution-QqTOF (Bruker Daltonics, Germany). HPLC analyses were performed as follows: pre-column for calibrants (Dionex, Acclaim PepMap 100 C18, nanoviper, 75 μm i.d. × 2 cm, 3 μm), desalting pre-column (Dionex, Acclaim PepMap 100 C18, cartridge, 300 μm i.d. × 5 mm, 5 μm), analytical column: 50 cm nanocolumn (Dionex, ID 75 μm, Acclaim PepMap100, C18, 2 μm), mobile phase for loading pump: 98:2 H_2O:ACN w/ 0,1% TFA, loading pump isocratic flow rate: 10 μL/min. The separation was performed at 40 °C and at a flow rate of 300 nL/min, using multistep 4 h gradients from 4 to 66% B in 204 min (mobile phase A for nano pump being H_2O w/ 0,1% Formic Acid and mobile phase B 80:20 ACN:H_2O w/ 0.08% FA). The runtime was set to 240 min. The analytical column was on-line interfaced to a nanoBoosterCaptiveSpray™ ESI source (Bruker Daltonics, Germany), working in enhanced modality. Collision-Induced Dissociation, assisted by N_2, was used to generate MS/MS spectra. Mass accuracy was improved using a specific lock mass (1221.9906 m/z) and a 15 min calibration segment (10mM sodium formate cluster solution). The instrument worked in Data-Dependent-Acquisition mode, automatically alternating a complete MS scan to the acquisition of MS/MS spectra. Other parameters were set in order to optimise the analysis: capillary voltage and current (2000–2100 V and 2050–2200 nA, respectively), peak to peak voltages for Funnel 1 (RF=400 Vpp), Funnel 2 (RF=600 Vpp), Hexapole (RF=400 Vpp) and Quadrupole ion energy (5 eV).

2.3.3. LC-MS/MS data analysis

Compass DataAnalysis v4.1 software was used to calibrate, deconvolute, and convert the acquired raw data prior to protein identification and quantification. Label-free protein quantitation was achieved by Progenesis QI for proteomics (Nonlinear Dynamics, Newcastle, UK). Automatic and manual run alignment were performed to reach a score≥50%; peak picking was achieved with a default sensitivity, a minimum peak width of 0.2 min and maximum charge of 8, then the first 15 min of analysis (elution of the calibrant) were discarded for the following steps. Normalization to all proteins and protein quantification using only non-conflicting peptides were also set. In addition, anova p.value≤0.05, q value≤0.05 power analysis≥0.8 and fold change≥2 were also used as statistical filters. Protein identification was obtained with Mascot 2.4.1 search engine with the following search parameters: Swissprot database (accessed July 2017; 555,100 sequences; 198,754,198 residues); fixed modification: cysteine carbamidomethylation; no variable modification; maximum number of missed cleavages: 1; MS mass tolerance: 20 ppm; MS/MS tolerance: 0.05 Da; peptide filter: FDR≤1 using percolator algorithm. Proteins of interest were analysed for molecular functions and biological processes with Panther tools [15].

3. Results

Samples deriving from thirteen patients with a proven diagnosis of clear cell Renal Cell Carcinoma were analysed using histology-guided MALDI-MSI and nLC-ESI-MS/MS in order to investigate differences in terms of protein localisation and abundance among the different tumour grades.

3.1.1. MALDI-MSI analysis of FFPE tissue sections

Initially, the capability of MALDI-MSI to enlighten different proteomic profiles arising from pathological and control specimens was investigated by performing unsupervised PCA on the entire dataset (Fig. 2). Spectra deriving from the areas of healthy tissues are clustered together and are well separated from those obtained from neoplastic areas, as expected (Fig. 2A). By dividing the healthy tissue into normal cortex and normal medulla, it was also possible to highlight that the region of overlap between spectra of tumour and normal tissues belonged mainly to those spectra derived from ROIs of normal cortex (Fig. 2B).

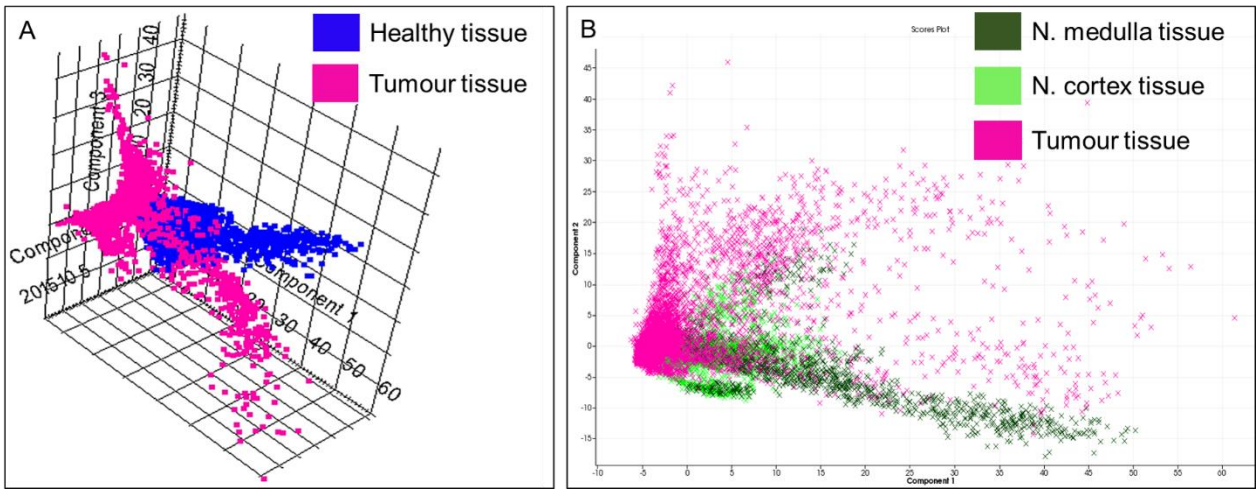


Fig. 2. Principal component analysis (PCA) score chart displaying the distribution of spectra form healthy and cancer tissue. A) Three-dimensional PCA score chart of spectra deriving from healthy tissue (blue) and cancer tissue (pink). B) Two-dimensional PCA of the spectra deriving from normal medulla (dark green), normal cortex (light green), and tumour regions (pink).

To detect the possible molecular profiles characteristic for each tumour grade, we focused our attention only on the spectra from ROIs G1, G2, G3 and G4. PCA did not show a consistent separation among spectra of different grades (Fig. 3A and B). Indeed, only G1 and G4 spectra were separated from each other (Fig. 3A), whereas those from cortex (light green), and tumour regions (pink). G2 and G3 showed a more scattered and overlapped distribution (Fig. 3B). Furthermore, the average spectra of the four

grades show a different profile (Fig. 4). In particular, there was little spectrum-to-spectrum variability observed among those obtained from the ROIs corresponding with G1 and G4 whilst a considerable amount of variability was observed in the mass spectra obtained from G2 and G3 (Figs. 3A and 4, single ROI spectrum). Therefore, more in-depth statistical analysis focused on the differently expressed signals between the two most homogeneous classes, G1 and G4, was performed. A pairwise ROC approach was employed in order to detect those peaks that could differentiate these two groups. Six signals showed a discriminatory capability, with an AUC≥0.8 and a p-value≤0.05. Three of these signals (m/z 944.71; m/z 1032.78; m/z 1325.99) were of higher intensity in G1 compared to G4, whereas the other three (m/z 872.63; m/ z 914.64 and m/z 1428.92) showed an opposite trend. A multi-ROC approach was used in order to explore the possibility of combining multiple signals to increase the discriminatory power between G1 and G4. The combination of signals with the same intensity trend within the two classes showed an AUC value of 0.93 and 0.92, respectively.

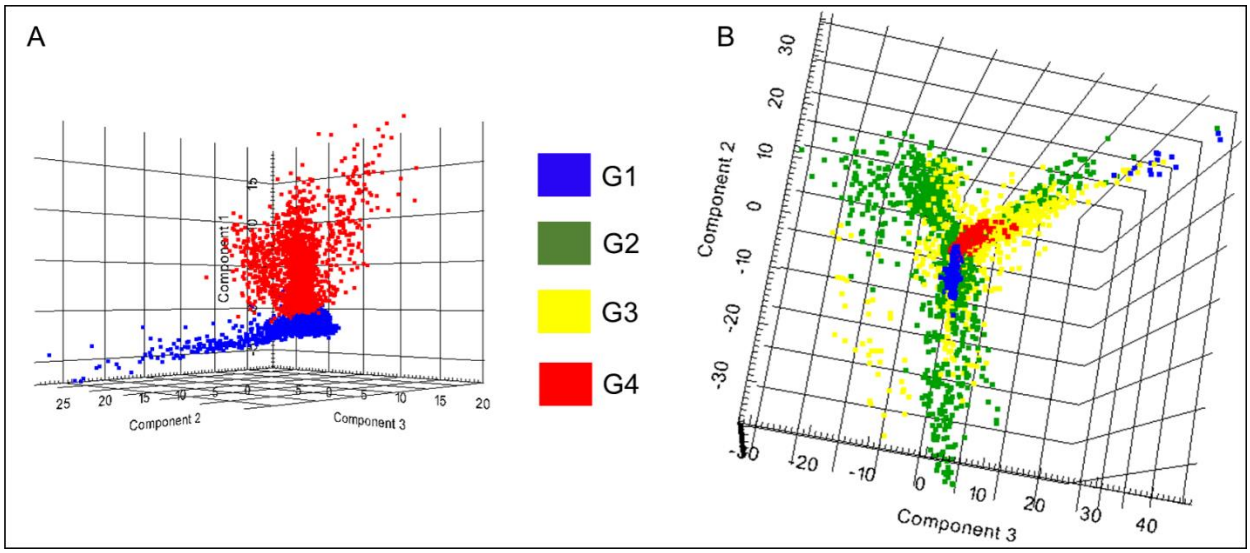


Fig. 3. Principal component analysis (PCA) score chart displaying the distribution of spectra from cancer tissue. Three-dimensional PCA score chart showing the distribution of spectra from ROIs at different grade: grade 1 (blue), grade 2 (green) (A), grade 3 (yellow), grade 4 (red) (B).

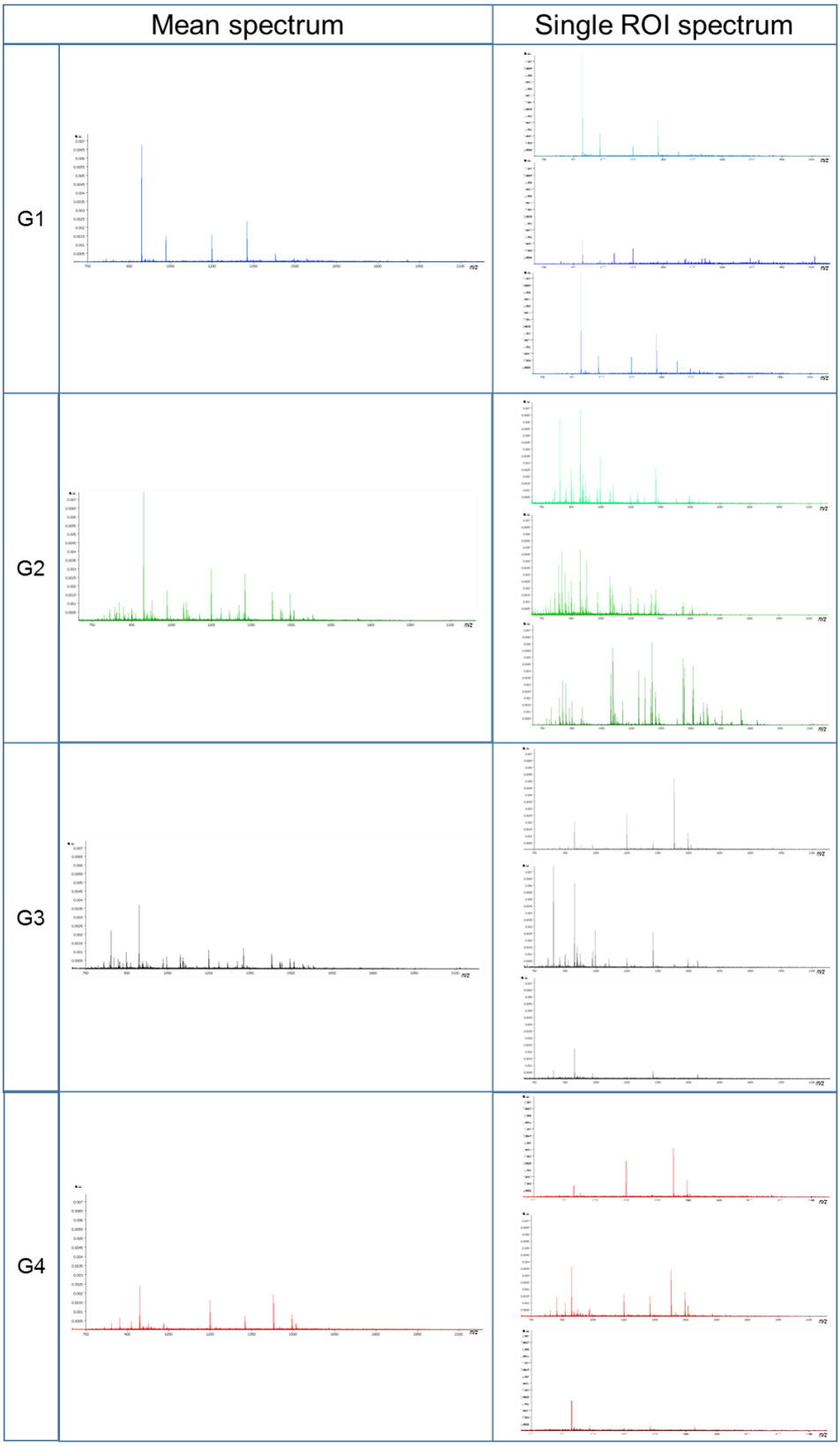


Fig. 4. Average profiles in the 700–2200 m/z mass range. Mean spectrum of the four grade (left) and average profiles of three representative single ROIs for each grade (right).

Considering the strong discriminatory power of all these m/z values, we also evaluated their expression in G2 and G3. Four of these six signals showed a good correlation, calculated using the Spearman's rho, within the grades (from G1 to G4). Three of these signals (m/z 944.71; m/z 1032.78; m/z 1325.99) had a correlation value of −0.6, −0.5 and−0.5, respectively, and a p-value lower than 0.001 whilst m/z 872.63 had a lower correlation value and higher p-value (correlation value=+0.4; p-value=0.02). On the contrary, m/z 914.64 and m/z 1428.92 did not show any correlation among the four grades. Some of the aforementioned signals were identified by on tissue MALDI-MS/MS. The ion at m/z 1428.92 was identified as a unique fragment of Vimentin (peptide sequence: SLYASSPGGVYATR) with a Mascot score of 50, whilst the ion at m/z 944.71 was identified as a fragment of Histone H2A (peptide AGLQFPVGR), with a Mascot score of 49. Finally, the signal at m/z 1325.99 was identified as a fragment of Histone H4 (unique peptide DNIQGITKPAIR) with a Mascot score of 76.

3.1.2. nLC-ESI-MS/MS analysis of FFPE tissue punches

In order to perform more extensive investigation regarding protein expression, tissue punches from three different patients for each grade were analysed in triplicate by nLC-ESI-MS/MS. Approximately 1000 proteins per patient were identified (ranging from 600 to 1300), with a total number of proteins identified per each grade of approximately 1700 (Table 2). In particular, 954 of the identified proteins were common to the four classes, whilst>200 proteins were specifically expressed by the different grades (data not shown). Protein abundance in the four grades was calculated by a label-free approach. More than 6000 peptides, with a score higher than 13, were used to evaluate the protein yield. Relative quantitation, using no conflicting peptides, was used to compare each single protein across runs and 1483 proteins were quantified.

Table 2. Number of proteins identified per injection, per sample and total number of proteins identified for each grade group (FDR ≤ 1 using percolator algorithm).

Grade	Sample	Total number of proteins identified		
		per injection	per sample	per grade
G1	M	inj1 = 650	991	1657
		inj 2 = 661		
		inj 3 = 748		
	N	inj 1 = 549	713	
		inj 2 = 432		
		inj 3 = 313		
	P	inj 1 = 988	1310	
		inj 2 = 881		
		inj 3 = 911		
G2	G	inj 1 = 729	1039	1779
		inj 2 = 697		
		inj 3 = 735		
	I	inj 1 = 920	1300	
		inj 2 = 985		
		inj 3 = 847		
	O	inj 1 = 793	1131	
		inj 2 = 623		
		inj 3 = 734		
G3	E	inj 1 = 541	1049	1774
		inj 2 = 729		
		inj 3 = 751		
	F	inj 1 = 964	1205	
		inj 2 = 790		
		inj 3 = 667		
	I	inj 1 = 810	1211	
		inj 2 = 723		
		inj 3 = 993		
G4	C	inj 1 = 582	819	1654
		inj 2 = 516		
		inj 3 = 473		
	E	inj 1 = 867	1233	
		inj 2 = 869		
		inj 3 = 824		
	H	inj 1 = 630	1009	
		inj 2 = 705		
		inj 3 = 607		

The expression of the tryptic peptide signals at m/z 1428.92 and m/z 1325.99, identified by MALDI-MS/MS as tryptic fragments of Vimentin and H4, respectively, was also confirmed by the nLC-ESI-MS/MS analysis. However, the signal at m/z 944.71, identified by MALDI-MS/MS as Histone H2A, could not be unambiguously assigned through the nLC-ESI-MS/MS approach. In fact, this signal could be attributed to three different histone species (H2A1B, H2A1D, H2AV). The signal at m/z 914.64, that was not identified by on-tissue MALDI-MS/MS, could be assigned to three different proteins containing a tryptic fragment with a mass of approximately 913.5 Da: Vimentin (VIME), Prohibitin (PHB) and Elongation factor 1-alpha 1 (EF1A1). Similarly, the signal at 1032.78 m/z was identified by nLC-ESI-MS/MS as a tryptic fragment of Histone H31T, corresponding to the peptide YRPGTVALR (m/z 344.87, triply-charged). This signal had an intensity that was similar to the one observed in the MALDI-MSI data. Finally, the signal at m/z 872.63 was identified as tryptic peptide (TAQAAAALR) uniquely assigned to Galactokinase protein (GALK1), which was over-expressed in G4 and under-expressed in G1 samples.

Afterwards, the 1483 quantified proteins were filtered (p-value≤ 0.05, q.value≤0.05, max fold change≥2 and power≥0.8) in order to select only those that had a statistically significant alteration among the four grades. A group of 393 species was obtained and only the 226 proteins quantified with at least two unique peptides were considered for the subsequent evaluation (Supplementary Tables 1 and 2). As the tumour stage of the patient cohort was not homogeneous, we decided to exclude any possible confounding factors deriving from tumour position or dimension. For this purpose, protein expression was estimated among the different stages (stage 1 vs stages 3–4: none of the studied cases presented with a stage 2 tumour). Fifty-four proteins (Supplementary Tables 3 and 4) were observed to be significantly varied between stage 1 and stages 3–4. In particular, 43 of these proteins were also present in the list of 226 proteins with a statistically significant alteration among grades and were thus excluded, given that their

variation was not associable solely to the grade (Supplementary Fig. 1). Finally, the 183 proteins that varied only in the grade comparison were retained for the PCA analysis (Fig. 5). The biological replicates for the four grades were clustered together and separated on the basis of the grade. In particular, G4 samples were well separated (Principal Component 1) from the others. Given the strongly evident difference between G4 and the lower grades, and to better highlight possible proteomic differences among the lowest grades, the data was re-elaborated excluding G4 specimens from the comparison. For this purpose, the 183 proteins, that were shown to be altered in the comparison between G1 to G4, were again filtered in order to highlight only those whose expression altered among G1, G2 and G3 (Supplementary Table 5). The PCA score chart (Supplementary Fig. 2) highlighted the presence of well-separated clusters corresponding to the three different grade classes, with higher distances between G1–2 and G3.

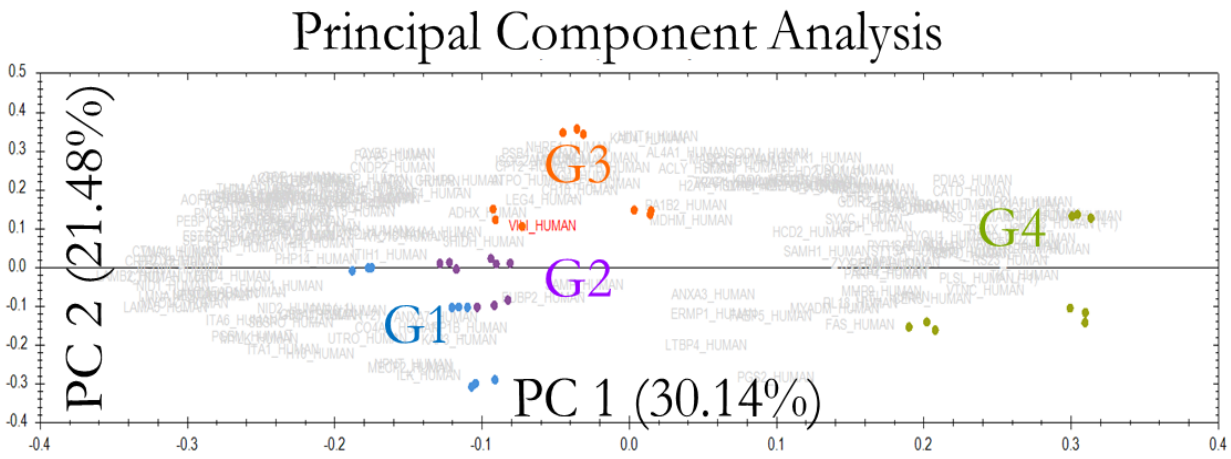


Fig. 5. Principal component analysis (PCA) score chart displaying the distribution of the twelve samples. Each coloured point represents a sample (three analytical replicates for each sample, and three biological replicates -same colour- for each grade) allocated based on the score of the Principal Components. The overlaid loadings chart shows the proteins that influence the distribution.

The biological and molecular functions of the remaining 70 proteins, with an expression varied among G1, G2 and G3, were investigated using PANTHER tools. Metabolic processes (37.8%, GO:0008152) and cellular processes (23.7%, GO:0009987) were identified as the two most represented biological processes (Fig. 6). When delving deeper into these two groups of processes, cell communication (GO:0007154) and primary metabolic (GO:0044238) processes were shown to be those most represented. Moreover, the molecular functions of these proteins (Panther- Molecular function) were investigated and a considerable part of them was represented by proteins involved in catalytic activity (59.6%, GO:0003824) and binding (28.8%, GO:0005488).

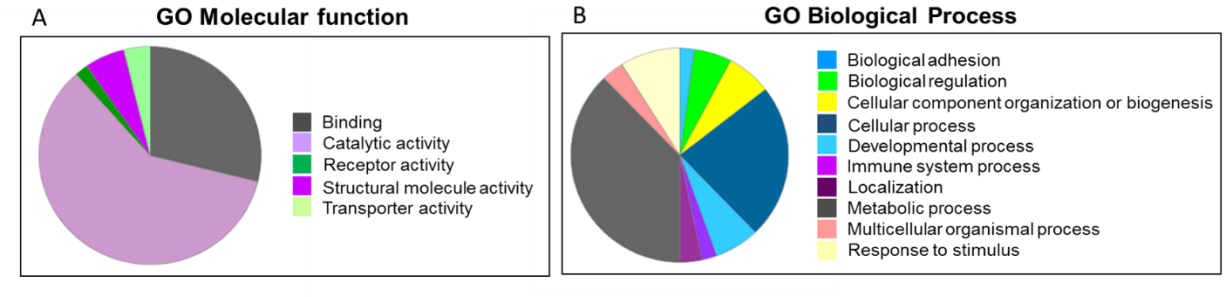


Fig. 6. Results of PANTHER database **analysis**. Pie charts of the 70 proteins statistically altered among G1, G2 and G3 samples categorized by GO classifications based on their (A) molecular function, and (B) biological process.

Those processes were also further investigated by performing the overrepresentation test, comparing our dataset with a reference dataset (Homo sapiens). In particular, fatty acid and lipid metabolic process, along with protein folding, tricarboxylic acid cycle and oxidoreductase activity resulted in being overrepresented in our dataset. (Supplementary Fig. 3). Furthermore, four proteins (Nephronectin, Isochorismatase domain-containing protein 2, Acyl-coenzyme A synthetase ACSM2A mitochondrial, Sideroflexin-1), showed a consistent differential expression among all the three comparisons (G1 vs G2, G1 vs G3 and G2 vs G3 - Supplementary Table 5).

4. Discussion

Despite the importance of tumour grade in ccRCC prognosis, no studies that investigate the molecular processes or protein alterations associated with each grade, taking into consideration the possible presence of cells ascribed to different grades within the same tissue section, have been reported so far [12,16]. The role of tumour grade as an important prognostic factor has also been demonstrated for other types of diseases, such as breast cancer [17] and prostatic adenocarcinoma [18]. On this basis, we investigated whether morphological differences, as used by the pathologist to define the tumour grade, reflect possible molecular alterations. We focused on the study of protein alterations detectable directly on tissue at different grades, by combining data obtained from two complementary mass spectrometric techniques. Initially, MALDI-MSI was used to overcome a limitation deriving from the diagnostic classification of the grade: the highest grade is assessed to the lesion even if it is present only in a minor area of the tumour. Failing that, the most abundant grade could potentially influence the results more than the diagnosed (highest) one. MALDI-MSI enabled average spectra of tumour areas, normal cortex, and normal medulla to be obtained. The PCA clustered the spectra belonging to these tissue areas in distinct groups. The small region of overlap of spectra observed in the PCA score chart was assigned partially to tumour areas and normal cortex, and this may be due to the fact that ccRCC tumours originate from the cortex and therefore may inherit part of its molecular traits. Alternatively, it may also be a result of heterogeneous tumour areas in which some non-tumoral cells are still present (Fig. 2). With the aim to investigate the molecular traits of the different ccRCC grades, a more in-depth study was performed by looking specifically at the differences among areas of a homogeneous tumour grade (ROIs) selected by the pathologist. Unsupervised PCA allows an entire imaging dataset to be represented and the output highlighted that G1 and G4 spectra were very well separated (Fig. 3A–B). This was expected based upon

their morphological and clinical features: ROIs of G1 are definitively composed only by cells attributable to G1, since only few cells of higher grade would be sufficient for a different classification to be assigned. In a similar vein, a ccRCC lesion is classified as G4 when cells already display characteristics that are evidently different from the initial stages of tumour development. Conversely, spectra derived from G2 and G3 regions displayed a more disperse distribution. The neat clustering of the MSI spectra obtained from G1 and G4 ROIs, along with the more scattered distribution of those from ROIs G2 and G3, may suggest that G1 and G4 represent the two extreme and well defined conditions, whereas G2 and 3 lesions can be better described as transitional steps in tumour progression. As also highlighted by Fig. 3A, G2 and G3 lesions seem to be characterised by proteomic profiles that are very heterogeneous but also intermediate between G1 and G4, with the proteomic profiles reflecting an ongoing process in cell dedifferentiation and tumour progression. Therefore, further investigations were at first performed comparing only the two extreme grades (G1 and G4).

The identity of the six signals that statistically varied between G1 and G4 was evaluated in order to correlate them with the pathology. Signals at m/z 944.71 and m/z 1325.99, of a higher intensity in G1, were identified as tryptic fragments of Histone H2A and Histone H4, respectively. As already known, Histones are implicated in cancer progression because of their role in the remodelling of chromatin structure and in gene accessibility [19], as well as in multiple other biological process and in countless metabolic pathways [20,21]. In addition, other research teams have found proteolytic fragments of this group of proteins overexpressed in adenocarcinoma [22] and pancreatic tumour [23]. Furthermore, the signal at m/z 1428.92 was identified as a tryptic fragment of Vimentin. This protein is usually expressed in clear cell and papillary RCC (87–100%) and approximately 50% of carcinomatous cells are Vimentin-positive in high-grade ccRCC [24]. Additionally, Vimentin is also included in the panel of antibodies used in RCC diagnosis [25]. In our study, Vimentin was also detected in

tumour areas, with a higher expression in G4 compared to G1 lesions. Although Vimentin has been detected in the later stages of many cancers, and it presence often correlates with malignancy, the role of this protein remains unclear [26,27]. Allegedly, Vimentin is involved in cell adhesion and migration as well as in the epithelial-mesenchymal transition of cancer cells [28–30]. Moreover, differences in the levels of Vimentin detected among different conditions could also be due to the presence of a tumour-specific exoprotease that cause an altered proteolytic process [31] and specific protocols to study the degradome are now under development [32]. Even though six signals were detected with a significantly different intensity between G1 and G4, it was not possible to find a panel of signals specific for each grade, most probably because of analytical limitations and biological heterogeneity. However, four of these six ions show also a good correlation among the four grades, allowing a clear separation between the extreme conditions and suggesting a possible process of grading progression among all the four grade classes.

To increase the number of proteins identified, we then proceeded to investigate the molecular profiles and factors that guide the grading progression using nLC-ESI-MS/MS as a complementary technique. The number of proteins identified per run, per patient and per grade were comparable (Table 2). Initially, we verified whether the six signals highlighted by the statistical analysis performed on MALDI-MSI data are consistent with those obtained with nLC-ESI-MS/MS. Despite the different type of ionisation process involved with these two approaches, the intensity of the ions identified by on-tissue MALDI-MS/MS showed a coherent trend with those observed by nLC-ESI-MS/MS. For the three signals that were not previously identified by MALDI-MS/MS, possible identifications were hypothesised based on results obtained by nLC-ESIMS/ MS and trends were verified. The signal at m/z 1032.78, of higher intensity in G1, was assigned to a tryptic peptide of Histone H31T and showed a similar trend. Likewise, the signal at m/z 872.63 was found to be a unique tryptic peptide of

Galactokinase (GALK1). GALK1 has recently been demonstrated to be a possible novel target for treating hepatocellular carcinoma, since its silencing inhibits the growth of HepG2 cells in culture [33].

We then proceeded to highlight the possible molecular differences among the four grade classes. First, differentially expressed proteins between stage 1 and stages 3–4 were excluded to avoid the inclusion of possible confounding factors related with the localisation and the dimension of the tumour, as well as with the presence of metastasis. Thus, proteins that only varied due to their association with different grades were used for performing the PCA analysis (Fig. 5). The three analytical replicates were extremely close to each other as well as the three biological replicates, indicating high similarity within samples of the same grade (analytical and biological CV was 16.3% (range 12.6–21.3) and 9.8% (range 8.2–11.2), respectively). Different grades generated independent clusters, in particular G4 specimens were considerably separated from lower grade samples. This distance well reflects the morphological classification, where G4 lesions are classified based on peculiar characteristics such as cell pleomorphism and presence of sarcomatoid or rhabdoid differentiation. Therefore, we investigated the possible molecular features that drive grading progression from G1 to G3.

Results highlighted that there are limited molecular differences between G1 and G2, while more pronounced differences characterised G3 specimens (Supplementary Fig. 1). These findings highlighted that the molecular profiles of different grades reflect the different survival rate of the four grades reported in literature [7]. In fact, G1 and G2 lesions, that have similar proteomic profiles, display an almost equal Cancer-Specific Survival (CSS) rate at 10 years post-surgery; 89% and 84%, respectively. G3 lesions, with different molecular profile from G1 and G2, have an intermediate CSS value, 46%, while a drastically reduced CSS rate is present for G4 patients (15%) which also show a very different protein profile from the others three grades. However, only four proteins showed a consistently different expression in the comparisons between G1 and G2, G2

and G3, and G1 and G3: Nephronectin (NPNT_HUMAN), Isochorismatase domain-containing protein 2 (ISOC2_ HUMAN), Acyl-coenzyme A synthetase ACSM2A mitochondrial (ACS2A_HUMAN), and Sideroflexin-1 (SFXN1_HUMAN) (Supplementary Table 5). NPNT was originally identified as a factor involved in tissue morphogenesis, playing a crucial role in the early stages of kidney development [34]. Our results show that NPNT has a decreasing concentration following grading progression and is in agreement with those reported by Kuphal et al. [35]. In fact, these authors, by using functional assays with stable NPNT transfection in cell lines, revealed that its expression increased cell adhesion and decreased cell migration and invasion. Moreover, they demonstrated that a loss of Nephronectin promotes neoplastic progression in malignant melanoma. On the contrary, we noticed that Isochorismatase domain-containing protein 2 shows an increasing concentration that follows the grade trend. ISOC2 is a functional, ubiquitously expressed, protein whose catalytic function is still to be clarified. ISOC2 seems to bind and co-localise with p16INK4a, a tumour suppressor that plays an important role in proliferation and tumorigenesis. ISOC2, if over-expressed, is able to inhibit p16INK4a in a dose-dependent manner, suggesting that ISOC2 may play an important role during tumour development [36]. ISOC2 was also reported to interact with SAHA, a histone deacetylase inhibitor [37]. We also noticed that Acyl-coenzyme A synthetase mitochondrial protein shows a concentration that increases with grade progression. ACSM2A is a mitochondrial medium-chain acyl-CoA synthetase [38] and expression levels of this gene in the kidney may be correlated with renal function, but the biological roles of ACSM2A have not yet been characterised in any species [39]. To the best of our knowledge, there are no studies describing the involvement of ACSM2A in cancer progression or development. Sideroflexin-1 has an atypical trend from G1 to G4. Our results showed a low expression in G2 and an over-expression in G3 samples. SFXN1 might be involved in the transport of a component required for iron utilisation, but no correlation between this protein and the possible

development of a neoplasm is currently known. However, another protein of the same family, sideroflexin-3, has been identified as a possible autoantigen in oral cancer that can be exploited for the early detection and follow-up of such patients [40]. Very recently, specific markers of immune cells that could be used to immune characterise the tumour were reported [41]. However, none of the reported markers for immune cells has been detected in our samples.

5. Conclusions

In conclusion, this study represents the first attempt to describe the proteomic profile of different ccRCC tumour grades using well defined, histology-guided, regions of tissue and analysis with two complementary MS-based proteomic approaches. MALDI-MS imaging permits a histology-guided proteomic study through the direct correlation between histological and molecular features and this approach highlighted several signals with an intensity trend associated with the different tumour grades, laying the groundwork for a more in-depth investigation of grading progression. Employing nLC-ESI-MS/MS, we were able to obtain a list of proteins whose expression altered among grades and this could represent a valuable starting point for more indepth pathway analysis, aimed at clarifying the molecular events occurring during the transition from G1 to G4, to be performed. Notwithstanding this promise, these findings need to be validated in a larger cohort of patients in order to strengthen the power of the statistical analysis.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jprot.2018.04.028.

Acknowledgments This work was supported by grants from MIUR: FIRB 2007 (RBRN07BMCT_11), the University of Milano-Bicocca (FAR 2013-2016) and from private Foundation (Fondazione Gigi & Pupa Ferrai Onlus).

References

[1] European Network of Cancer Registries, Kidney cancer (KC) Factsheet, http://www.encr.eu/images/docs/factsheets/ENCR_Factsheet_Kidney_2017.pdf, (2017).

[2] B. Ljungberg, K. Bensalah, S. Canfield, S. Dabestani, F. Hofmann, M. Hora, M.A. Kuczyk, T. Lam, L. Marconi, A.S. Merseburger, P. Mulders, T. Powles, M. Staehler, A. Volpe, A. Bex, EAU guidelines on renal cell carcinoma: 2014 update, Eur. Urol. 67 (2015) 913–924, http://dx.doi.org/10.1016/j.eururo.2015.01.005.

[3] J.R. Srigley, B. Delahunt, J.N. Eble, L. Egevad, J.I. Epstein, D. Grignon, O. Hes, H. Moch, R. Montironi, S.K. Tickoo, M. Zhou, P. Argani, The International Society of Urological Pathology (ISUP) vancouver classification of renal neoplasia, Am. J. Surg. Pathol. 37 (2013) 1469–1489, http://dx.doi.org/10.1097/PAS. 0b013e318299f2d1.

[4] C.T. Lee, J. Katz, P.A. Fearn, P. Russo, Mode of presentation of renal cell carcinoma provides prognostic information, Urol. Oncol. 7 (2002) 135–140, http://dx.doi.org/ 10.1016/S1078-1439(01)00185-5.

[5] L.H. Sobin, M. Gospodarowicz, C. Wittekind, TNM Classification of Malignant Tumours, 7th ed., Wiley, 2009, pp. 243–246.

[6] H. Moch, A.L. Cubilla, P.A. Humphrey, V.E. Reuter, T.M. Ulbright, The 2016 WHO classification of tumours of the urinary system and male genital organs—part a: renal, penile, and testicular tumours, Eur. Urol. 70 (2016) 93–105, http://dx.doi. org/10.1016/j.eururo.2016.02.029.

[7] B. Delahunt, J.K. McKenney, C.M. Lohse, B.C. Leibovich, R.H. Thompson, S.A. Boorjian, J.C. Cheville, A novel grading system for clear cell renal cell carcinoma incorporating tumor necrosis, Am. J. Surg. Pathol. 37 (2013) 311–322, http://dx.doi.org/10.1097/PAS.0b013e318270f71c.

[8] J. Dagher, B. Delahunt, N. Rioux-Leclercq, L. Egevad, J.R. Srigley, G. Coughlin, N. Dunglinson, T. Gianduzzo, B. Kua, G. Malone, B. Martin, J. Preston, M. Pokorny, S. Wood, J. Yaxley, H. Samaratunga, Clear cell renal cell carcinoma: validation of World Health Organization/International Society of Urological Pathology grading, Histopathology 71 (2017) 918–925, http://dx.doi.org/10.1111/his.13311.

[9] B. Delahunt, Advances and controversies in grading and staging of renal cell carcinoma, Mod. Pathol. 22 (2009) S24–S36, http://dx.doi.org/10.1038/modpathol.2008.183.

[10] M.M. Rinschen, T. Benzing, K. Limbutara, T. Pisitkun, Proteomic analysis of the kidney filtration barrier-problems and perspectives, Proteomics Clin. Appl. 9 (2015) 1053–1068, http://dx.doi.org/10.1002/prca.201400201.

[11] B. Perroud, T. Ishimaru, A.D. Borowsky, R.H. Weiss, Grade-dependent proteomics characterization of kidney cancer, Mol. Cell. Proteomics 8 (2009) 971–985, http://dx.doi.org/10.1074/mcp.M800252-MCP200.

[12] H.I. Wettersten, A.A. Hakimi, D. Morin, C. Bianchi, M.E. Johnstone, D.R. Donohoe, J.F. Trott, O. Abu Aboud, S. Stirdivant, B. Neri, R. Wolfert, B. Stewart, R. Perego, J.J. Hsieh, R.H. Weiss, Grade-dependent metabolic reprogramming in kidney cancer revealed by combined proteomics and metabolomics analysis, Cancer Res. 75 (2015) 2541–2552, http://dx.doi.org/10.1158/0008-5472.CAN-14-1703.

[13] G. De Sio, A.J. Smith, M. Galli, M. Garancini, C. Chinello, F. Bono, F. Pagni, F. Magni, A MALDI-mass spectrometry imaging method applicable to different formalin-fixed paraffin-embedded human tissues, Mol. BioSyst. (2015), http://dx.doi.org/10.1039/c4mb00716f.

[14] M. Galli, F. Pagni, G. De Sio, A. Smith, C. Chinello, M. Stella, V. L'Imperio, M. Manzoni, M. Garancini, D. Massimini, N. Mosele, G. Mauri, I. Zoppis, F. Magni, Proteomic profiles of thyroid tumors by mass spectrometry-imaging on tissue

microarrays, Biochim. Biophys. Acta Proteins Proteomics 1865 (2017) 817–827, http://dx.doi.org/10.1016/j.bbapap.2016.11.020.
[15] H. Mi, B. Lazareva-Ulitsky, R. Loo, A. Kejariwal, J. Vandergriff, S. Rabkin, N. Guo, A. Muruganujan, O. Doremieux, M.J. Campbell, H. Kitano, P.D. Thomas, The PANTHER database of protein families, subfamilies, functions and pathways, Nucleic Acids Res. 33 (2005), http://dx.doi.org/10.1093/nar/gki078.
[16] B. Delahunt, J.C. Cheville, G. Martignoni, P.A. Humphrey, C. Magi-Galluzzi, J. McKenney, L. Egevad, F. Algaba, H. Moch, D.J. Grignon, R. Montironi, J.R. Srigley, The International Society of Urological Pathology (ISUP) grading system for renal cell carcinoma and other prognostic parameters, Am. J. Surg. Pathol. 37 (2013) 1490–1504, http://dx.doi.org/10.1097/PAS.0b013e318299f0fb.
[17] E.A. Rakha, J.S. Reis-Filho, F. Baehner, D.J. Dabbs, T. Decker, V. Eusebi, S.B. Fox, S. Ichihara, J. Jacquemier, S.R. Lakhani, J. Palacios, A.L. Richardson, S.J. Schnitt, F.C. Schmitt, P.-H. Tan, G.M. Tse, S. Badve, I.O. Ellis, Breast cancer prognostic classification in the molecular era: the role of histological grade, Breast Cancer Res. 12 (2010) 207, http://dx.doi.org/10.1186/bcr2607.
[18] J. Gordetsky, J. Epstein, Grading of prostatic adenocarcinoma: current state and prognostic implications, Diagn. Pathol. 11 (2016) 25 , http://dx.doi.org/10.1186/s13000-016-0478-2.
[19] B. Heijs, R.J. Carreira, E.A. Tolner, A.H. De Ru, A.M.J.M. Van Den Maagdenberg, P.A. Van Veelen, L.A. McDonnell, Comprehensive analysis of the mouse brain proteome sampled in mass spectrometry imaging, Anal. Chem. 87 (2015) 1867–1875, http://dx.doi.org/10.1021/ac503952q.
[20] T.Y. Wang, Y.L. Jia, X. Zhang, Q.L. Sun, Y.C. Li, J.H. Zhang, C.P. Zhao, X.Y. Wang, L. Wang, Treating colon cancer cells with FK228 reveals a link between histone lysine acetylation and extensive changes in the cellular proteome, Sci. Rep. 5 (2015), http://dx.doi.org/10.1038/srep18443.

[21] R. Singh, S.W. Harshman, A.S. Ruppert, A. Mortazavi, D.M. Lucas, J.M. Thomas-Ahner, S.K. Clinton, J.C. Byrd, M.A. Freitas, M.R. Parthun, Proteomic profiling identifies specific histone species associated with leukemic and cancer cells, Clin. Proteomics 12 (2015), http://dx.doi.org/10.1186/s12014-015-9095-4.

[22] M.C. Djidja, S. Francese, P.M. Loadman, C.W. Sutton, P. Scriven, E. Claude, M.F. Snel, J. Franck, M. Salzet, M.R. Clench, Detergent addition to tryptic digests and ion mobility separation prior to MS/MS improves peptide yield and protein identification for in situ proteomic investigation of frozen and formalin-fixed paraffin- embedded adenocarcinoma tissue sections, Proteomics 9 (2009) 2750–2763, http://dx.doi.org/10.1002/pmic.200800624.

[23] M.C. Djidja, E. Claude, M.F. Snel, P. Scriven, S. Francese, V. Carolan, M.R. Clench, MALDI-ion mobility separation-mass spectrometry imaging of glucose-regulated protein 78 kDa (Grp78) in human formalin-fixed, paraffin-embedded pancreatic adenocarcinoma tissue sections, J. Proteome Res. 8 (2009) 4876–4884, http://dx. doi.org/10.1021/pr900522m.

[24] L.D. Truong, S.S. Shen, Immunohistochemical diagnosis of renal neoplasms, Arch. Pathol. Lab. Med. 135 (2011) 92–109, http://dx.doi.org/10.1043/2010-0478- RAR.1.

[25] P.H. Tan, L. Cheng, N. Rioux-Leclercq, M.J. Merino, G. Netto, V.E. Reuter, S.S. Shen, D.J. Grignon, R. Montironi, L. Egevad, J.R. Srigley, B. Delahunt, H. Moch, Renal tumors: diagnostic and prognostic biomarkers, Am. J. Surg. Pathol. 37 (2013) 1518–1531, http://dx.doi.org/10.1097/PAS.0b013e318299f12e.

[26] A. Satelli, S. Li, Vimentin in cancer and its potential as a molecular target for cancer therapy, Cell. Mol. Life Sci. 68 (2011) 3033–3046, http://dx.doi.org/10.1007/ s00018-011-0735-1.

[27] M.E. Kidd, D.K. Shumaker, K.M. Ridge, The role of vimentin intermediate filaments in the progression of lung cancer, Am. J. Respir. Cell Mol. Biol. 50 (2014) 1–6, http://dx.doi.org/10.1165/rcmb.2013-0314TR.

[28] J. Ivaska, H.M. Pallari, J. Nevo, J.E. Eriksson, Novel functions of vimentin in cell 9 adhesion, migration, and signaling, Exp. Cell Res. 313 (2007) 2050–2062, http://dx.doi.org/10.1016/j.yexcr.2007.03.040.

[29] M.G. Mendez, S.-I. Kojima, R.D. Goldman, Vimentin induces changes in cell shape, motility, and adhesion during the epithelial to mesenchymal transition, FASEB J. 24 (2010) 1838–1851, http://dx.doi.org/10.1096/fj.09-151639.

[30] C.-Y. Liu, H.-H. Lin, M.-J. Tang, Y.-K. Wang, Vimentin contributes to epithelialmesenchymal transition cancer cell mechanics by mediating cytoskeletal organization and focal adhesion maturation, Oncotarget 6 (2015) 15966–15983, http://dx.doi.org/10.18632/oncotarget.3862.

[31] J. Villanueva, D.R. Shaffer, J. Philip, C.A. Chaparro, H. Erdjument-Bromage, A.B. Olshen, M. Fleisher, H. Lilja, E. Brogi, J. Boyd, M. Sanchez-Carbayo, E.C. Holland, C. Cordon-Cardo, H.I. Scher, P. Tempst, Differential exoprotease activities confer tumor-specific serum peptidome patterns, J. Clin. Invest. 116 (2006) 271–284, http://dx.doi.org/10.1172/JCI26022.

[32] M.M. Rinschen, A.-K. Hoppe, F. Grahammer, M. Kann, L.A. Völker, E.-M. Schurek, J. Binz, M. Höhne, F. Demir, M. Malisic, T.B. Huber, C. Kurschat, J.N. Kizhakkedathu, B. Schermer, P.F. Huesgen, T. Benzing, N-degradomic analysis reveals a proteolytic network processing the podocyte cytoskeleton, J. Am. Soc. Nephrol. 28 (2017) 2867–2878, http://dx.doi.org/10.1681/ASN.2016101119.

[33] M. Tang, E. Etokidem, K. Lai, The Leloir pathway of galactose metabolism – a novel therapeutic target for hepatocellular carcinoma, Anticancer Res. 36 (2016) 6265–6272, http://dx.doi.org/10.21873/anticanres.11221.

[34] U. Müller, D. Wang, S. Denda, J.J. Meneses, R.A. Pedersen, L.F. Reichardt, Integrin α8β1 is critically important for epithelial-mesenchymal interactions during kidney morphogenesis, Cell 88 (1997) 603–613, http://dx.doi.org/10.1016/S0092-8674(00)81903-0.

[35] S. Kuphal, S. Wallner, A.K. Bosserhoff, Loss of nephronectin promotes tumor progression in malignant melanoma, Cancer Sci. 99 (2008) 229–233, http://dx.doi.org/10.1111/j.1349-7006.2007.00678.x.

[36] X. Huang, Z. Shi, W. Wang, J. Bai, Z. Chen, J. Xu, D. Zhang, S. Fu, Identification and characterization of a novel protein ISOC2 that interacts with p16INK4a, Biochem. Biophys. Res. Commun. 361 (2007) 287–293, http://dx.doi.org/10.1016/j.bbrc.2007.06.181.

[37] J.J. Fischer, S. Michaelis, A.K. Schrey, A. Diehl, O.Y. Graebner, J. Ungewiss, S. Horzowski, M. Glinski, F. Kroll, M. Dreger, H. Koester, SAHA capture compound - a novel tool for the profiling of histone deacetylases and the identification of additional vorinostat binders, Proteomics 11 (2011) 4096–4104, http://dx.doi.org/10.1002/pmic.201000717.

[38] G. Kochan, E.S. Pilka, F. von Delft, U. Oppermann, W.W. Yue, Structural snapshots for the conformation-dependent catalysis by human medium-chain acyl-coenzyme A synthetase ACSM2A, J. Mol. Biol. 388 (2009) 997–1008, http://dx.doi.org/10. 1016/j.jmb.2009.03.064.

[39] P.A. Watkins, D. Maiguel, Z. Jia, J. Pevsner, Evidence for 26 distinct acyl-coenzyme A synthetase genes in the human genome, J. Lipid Res. 48 (2007) 2736–2750, http://dx.doi.org/10.1194/jlr.M700378-JLR200.

[40] R. Murase, Y. Abe, T. Takeuchi, M. Nabeta, Y. Imai, Y. Kamei, L. Kagawa-Miki, N. Ueda, T. Sumida, H. Hamakawa, K. Kito, Serum autoantibody to sideroflexin 3 as a novel tumor marker for oral squamous cell carcinoma, Proteomics Clin. Appl. 2 (2008) 517–527, http://dx.doi.org/10.1002/prca.200780123.

[41] S. Chevrier, J.H. Levine, V.R.T. Zanotelli, K. Silina, D. Schulz, M. Bacac, C.H. Ries, L. Ailles, M.A.S. Jewett, H. Moch, M. van den Broek, C. Beisel, M.B. Stadler, C. Gedye, B. Reis, D. Pe'er, B. Bodenmiller, An immune atlas of clear cell renal cell carcinoma, Cell 169 (2017) 736–749.e18, http://dx.doi.org/10.1016/j.cell.2017.04.016.

CHAPTER 4

Effects of clear cell Renal Cell Carcinoma stage and grade on urinary proteomic profiles

Martina Stella[a], Marco Grasso[b], Giorgio Bovo[c], Fulvio Magni [a]

[a] Department of Medicine and Surgery, University of Milano - Bicocca, Clinical Proteomics and Metabolomics Unit, Vedano al Lambro, Italy
[b] Urology Unit, S.Gerardo Hospital, Monza, Italy
[c] Pathology Unit, Vimercate Hospital, Vimercate, Italy

Manuscript in preparation (PROTEOMICS-Clinical Applications)

ABSTRACT In this study, we investigated the urinary proteome by a label-free proteomics approach in order to obtain a deeper insight into the molecular alterations associated with the presence of clear cell Renal Cell Carcinoma (ccRCC) lesions at different stages. Nowadays, the possibility to implement a liquid biopsy in the investigation of tumour progression is of particular interest. For this reason, we evaluated the human urinary proteome of ccRCC patients at different grades and stages in order to verify if those alterations, previously detected on tissue, could also be detected in urine. In terms of protein identity, the information obtained from the urine overlaps greatly with those provided by the tissue. However, from a quantitative point of view, alterations in the urinary proteome better reflect the dimension and position of the lesion than its morphological characteristics.

KEYWORDS Clear cell renal cell carcinoma, nLC-ESI-MS/MS, tumour grade, proteomics, urine.

1. Introduction

Renal cell carcinoma (RCC) comprise a heterogeneous group of tumours with the most frequent (70-80%) and aggressive morpho-type being clear cell RCC (ccRCC) [1]. ccRCC is a malignancy of particular interest for clinical proteomic approaches because of its associated inter- and intra- tumour heterogeneity [2], high recurrence and progression rates, and lack of effective non-invasive diagnostic and prognostic tools [3]. The lesions are classified according to the recommendation of the WHO classification [4]: the stage is assigned following the TNM classification system that combines anatomical factors, such as (T) the tumour size and its location (venous invasion, renal capsular invasion or adrenal involvement), (N) lymph node and (M) distant metastasis involvement. The grade is assessed with the nuclear grading system, proposed at the International Society of Urological Pathology (ISUP) conference of 2013 and accepted by the World Health Organisation (WHO) in 2016, that replaced the previously used Fuhrman grading system [5].

Thus far, the molecular mechanism underpinning tumour development and progression remains unclear and molecular markers that can provide a more in-depth understanding of the biological process involved in this progression will be beneficial for disease management. In this context, proteomics has been extensively used [6,7] and biofluids have been samples of choice due to them being easily, and noninvasively, obtainable. Despite limitations related to a high degree of daily variation and sensitivity to physiological changes or external factors, the new concept of a liquid biopsy has been extensively investigated as a valid alternative to the classical solid biopsies in biomarker discovery [8,9]. In the last decades, different proteomics approaches based on mass spectrometry have been used to investigate the urinary peptidome/proteome for the purpose of biomarker discovery and, in particular, studying tumours that involve urinary system organs [10–12]. However, excluding some exceptions [13–15], most of the

biomarker studies in bodyfluids focused on the differences between patients with ccRCC and healthy controls. Therefore, in this work we investigated the proteome of ccRCC patients with different lesion severity in order to obtain a deeper insight into the molecular alterations associated not only with the presence but also with the progression of ccRCC that correspond with the grade classification, as previously investigated on tissue [16] as well as tumour size and stage. The results obtained by a label free-mass-spectrometry based method allowed the detection of a set of secreted proteins associated with different ccRCC lesions and potentially with the aggressiveness and the progression of the disease.

2. Experimental section

Patients selection

44 patients from San Gerardo hospital (Monza, Italy) were recruited between 2011 and 2016 and all subject had signed an informed consent. Local ethical committee (Comitato Etico Azienda Ospedaliera San Gerardo, Monza, Italy) approved the protocols and procedures. The diagnosis was performed after careful histopathological evaluation including tumour size and position (pT), grading (ISUP classification system), sarcomatoid features, vascular invasion, tumour necrosis, presence of distant metastasis and involvement of lymph nodes (stage).

Sample collection and preparation

The second urine of the morning, before total or partial nephrectomy, were collected into sterile urine tube. The samples were then centrifuged at 2400 rpm, 10 min at 4°C and stored at -80° C until the day of the analysis. 6 mL of the stored supernatant were concentrated by centrifugation on 3 kDa cut-off filter (Amicon Ultra-4mL 3kDa,

Millipore) for 10 minutes at room temperature (RT). Protein concentration was determined with the BCA (Microplate BCA ™ protein Assay Kit, Thermo Scientific) and 200 µg of urine protein samples were treated following the FASP protocol, as already described [17]. Briefly, the proteins were first reduced by incubation with 50mM DL-Dithiothreitol (Sigma Aldrich, Switzerland) and then alkylated for 30 minutes with Iodoacetamide 100mM (Sigma Aldrich, Switzerland). The digestion was performed on 30kDa filters overnight adding trypsin from porcine pancreas (Proteomics Grade, BioReagent, Dimethylated) in a ratio 1:100 to the initial protein concentration. After repeated washing of the filter, the eluted peptides were collected and acidified. Peptide solution was desalted and concentrated using µ-C18 ZiptipTM pipette tips (Millipore Corp, Bedford, MA) following the standard protocol provided by Millipore. Purified samples were resuspended in 98 H_2O: 2 ACN: 0.1% TFA and analysed by nUHPLC-MS/MS.

nUHPLC-MS/MS analysis

LC-ESI-MSMS analysis was performed using Dionex UltiMate3000 rapid separation (RS) LC nano system (Thermo Scientific, Sunnyvale, CA) coupled with an UHR-nESI-qTOF (Impact HDTM, Bruker Daltonics Germany equipped with a Captivespray nanoBooster). 1.4 µg were loaded onto a µ-precolumn (Dionex, Acclaim PepMac100 C18, cartridge, 300µm i.d. x 5 mm, 5 µm) and then separated by an analytical C18 column (Dionex, 75µm ID, Acclaim PepMac100, C18, 2 µm). A multistep gradient with phase A (0.1 FA) and phase B (0.1 FA/80%ACN) from 4% to 66% was run for 240 minutes. Eluted peptides were analysed in data-dependent acquisition mode and the MS/MS data were acquired by targeting precursors (300-2000m/z range) with a charge state between +2 and +5 and with at least 1575 counts (fixed cycle time of 5 seconds) for fragmentation, obtained by collision induced dissociation. MS scan were recorded in centroid as well as MS/MS data. Raw MS and MS/MS data were corrected using both

an internal calibration (lock mass 1221.9906 m/z) and a calibration segment (15 min. before each run), converted and deconvoluted to XML file using DataAnalysis software (Bruker Daltonics, Germany).

Data processing

Protein identification was obtained with PEAKS Studio 8.5 (Bioinformatics Solutions Inc., Waterloo, Ontario, Canada [18]). Trypsin was set as the enzyme used for the digestion, with one as a maximum acceptable missed cleavage, carbamidomethylation as fixed modification, 20 ppm and 0.05 Da as mass tolerance for MS and MS/MS tolerance, respectively. Swiss-prot (accessed July 2017; 555.100 sequences; 198.754.198 residues) was used as database. A maximum false discovery rate (FDR) for peptide spectral match was set to 1%, retention time window of 2 minutes was established for features matching between runs, and a minimum of one sequence-unique peptides was required for identification. Proteins abundance was calculated using the three most abundant unique peptides, normalised against the total ion current and only the proteins detected with at least two unique peptides and in at least 50% of the samples were used for protein quantification. Files containing the information regarding protein abundance in the form of statistical matrices (samples as rows and proteins as columns) were exported from PEAKS Studio 8.5 and imported into an in-house developed software. For each protein, the statistical evaluation was performed as follows, in order to retrieve a list of differentially expressed proteins. First, the statistical observations (patients) were grouped according to the levels of a chosen outcome attribute (e.g. the grade) than proteins were considered of different abundance when the p-value yielded by the Mann-Whitney test was less than 0.05.

Functional analysis was carried out for protein of interest: KEGG pathway term clustering using ClueGO v2.5, Clupedia v1.5 and the Cytoscape network analysis

framework was performed. The default parameters were used with an overall statistical significance value set to p-value < 0.05.

3. Results

Conventional histopathological evaluation of tumours, such as stage and grade of the lesions, has prognostic significance. However, the molecular mechanisms involved in tumour development and progression are not well understood. Therefore, the possibility to obtain information about molecular changes associated with tumour lesions has been investigated using easily-accessible urine samples.

Clinical data and study design/ identification and quantification proteins in ccRCC

Urine samples were collected from 44 patients (26 males, 18 females; median age at diagnosis 66 and 69 respectively, and mean tumour mass of 5.54 ± 3.33 cm) with a proven diagnosis of ccRCC (Table 1). Patients were classified accordingly with 2009 TNM classification [19]. In particular, a specific focus was given both to grade values (low grade (LG) = G1 and G2; high Grade (HG) = G3 and G4), pT (pT 1 and pT3) and stage (stage1 and stage3-4) in order to evaluate proteomic signatures of different lesion, as already demonstrated on tissue, using liquid biopsies.

Table 1 Clinicopathological characteristics of patients included in the analysis

Group	Number of patient	Gender [male-female]	Age (at diagnosis) [mean]	Greatest tumour dimension [cm]
pT1, Stage 1	19	11 - 7	68	3.5
pT3, Stage 3-4	25	12 - 13	67	7.09
Low grade (LG)	24	12- 12	66	4.34
High grade (HG)	20	14 - 6	69	6.99
Stage 1 - low grade (A)	15	7 - 8	65	3.44
Stage 3-4 - low grade (B)	9	5 - 4	67	5.8
Stage 1 - high grade (C)	4	4 - 0	71	3.75
Stage 3-4 - high grade (D)	16	10 - 6	67	7.8

Low grade (LG) = G1 and G2

High grade (HG) = G3 and G4

Qualitative evaluation and network analysis: relevant pathways of ccRCC

Following the FASP protocol, samples were analysed by nUHPLC-MS/MS and a total of 21640 peptides sequences, corresponding to 1609 proteins were identified using PEAKS STUDIO 8.5 with FDR Peptide-Spectrum Matches of 1% and at least one unique peptide (Supplementary Table 1). Those proteins identified in urine were compared with those identified on tissue (previously published data [16]) and more than 500 proteins were observed to be present in both of the specimens. Moreover, pathways represented by the common proteins were investigated using the KEGG database. The pathways that were associated with the presence of ccRCC lesions are shown and listed in Figure 1(A-B), and most of them were observed to be related to cancer and/or kidney damage [20–22].

A GO-Term	% Associated Genes	Nr. Genes	Associated Genes Found
Cholesterol metabolism	20.00	10.00	[APOA1, APOA4, APOB, APOC3, APOE, APOH, LRP1, LRP2, VDAC1, VDAC2]
Citrate cycle (TCA cycle)	23.33	7.00	[ACO1, ACO2, IDH1, IDH2, MDH1, MDH2, PCK1]
Complement and coagulation cascades	43.04	34.00	[A2M, C1QA, C1QB, C1QC, C2, C3, C4B, C4BPA, C5, C6, C8A, C8B, C9, CD59, CFB, CFH, CLU, F2, F9, FGA, FGB, FGG, ITGAM, ITGB2, KLKB1, KNG1, PLG, SERPINA1, SERPINC1, SERPIND1, SERPINF2, SERPING1, VTN, VWF]
Cysteine and methionine metabolism	24.44	11.00	[AHCY, AHCYL1, BHMT, GOT1, GOT2, GSS, LDHA, LDHB, MAT2A, MDH1, MDH2]
ECM-receptor interaction	30.49	25.00	[AGRN, CD44, CD47, COL1A2, COL4A1, COL4A2, COL4A3, COL4A6, COL6A1, COL6A3, DAG1, FN1, HSPG2, ITGB1, LAMA1, LAMA4, LAMA5, LAMB1, LAMB2, LAMC1, THBS1, TNC, TNXB, VTN, VWF]
Focal adhesion	14.57	29.00	[ACTN1, ACTN4, CDC42, COL1A2, COL4A1, COL4A2, COL4A3, COL4A6, COL6A1, COL6A3, FLNA, FLNC, FN1, ITGB1, LAMA1, LAMA4, LAMA5, LAMB1, LAMB2, LAMC1, MYL12A, RHOA, THBS1, TLN1, TNC, TNXB, VCL, VTN, VWF]
Fructose and mannose metabolism	27.27	9.00	[AKR1B1, ALDOA, ALDOB, ALDOC, FBP1, HK3, KHK, TKFC, TPI1]
Glutathione metabolism	21.43	12.00	[ANPEP, G6PD, GGT1, GPX3, GSR, GSS, GSTA2, GSTM3, GSTP1, IDH1, IDH2, LAP3]
Glycolysis / Gluconeogenesis	33.82	23.00	[ADH5, AKR1A1, ALDH7A1, ALDH9A1, ALDOA, ALDOB, ALDOC, ENO1, ENO2, FBP1, GALM, GAPDH, GPI, HK3, LDHA, LDHB, PCK1, PGAM1, PGK1, PGM1, PKLR, PKM, TPI1]
Glyoxylate and dicarboxylate metabolism	30.00	9.00	[ACAT1, ACO1, ACO2, CAT, GRHPR, MDH1, MDH2, SHMT1, SHMT2]
Leukocyte transendothelial migration	15.18	17.00	[ACTN1, ACTN4, CDC42, EZR, GNAI2, GNAI3, ICAM1, ITGAM, ITGB1, ITGB2, MMP9, MSN, MYL12A, PECAM1, RHOA, VCAM1, VCL]
Pentose phosphate pathway	33.33	10.00	[ALDOA, ALDOB, ALDOC, FBP1, G6PD, GPI, H6PD, PGLS, PGM1, TKT]
Phagosome	12.50	19.00	[ATP6V1A, ATP6V1B2, ATP6V1E1, C3, CALR, CANX, CORO1A, CTSL, DYNC1H1, ITGAM, ITGB1, ITGB2, LAMP1, MPO, RAB5C, RAB7A, THBS1, TUBB, TUBB4B]
PI3K-Akt signaling pathway	9.32	33.00	[COL1A2, COL4A1, COL4A2, COL4A3, COL4A6, COL6A1, COL6A3, FN1, GNB1, GNB2, GNG12, HSP90AA1, HSP90AB1, HSP90B1, ITGB1, LAMA1, LAMA4, LAMA5, LAMB1, LAMB2, LAMC1, PCK1, THBS1, TNC, TNXB, VTN, VWF, YWHAB, YWHAE, YWHAG, YWHAH, YWHAQ, YWHAZ]
Pyruvate metabolism	30.77	12.00	[ACAT1, ALDH7A1, ALDH9A1, GRHPR, HAGH, LDHA, LDHB, MDH1, MDH2, PCK1, PKLR, PKM]
Regulation of actin cytoskeleton	11.74	25.00	[ACTN1, ACTN4, ARPC2, ARPC3, ARPC4, CDC42, CFL1, EZR, F2, FN1, GNG12, GSN, IQGAP1, ITGAM, ITGB1, ITGB2, MSN, MYH9, MYL12A, PFN1, PFN2, RDX, RHOA, SCIN, VCL]

B

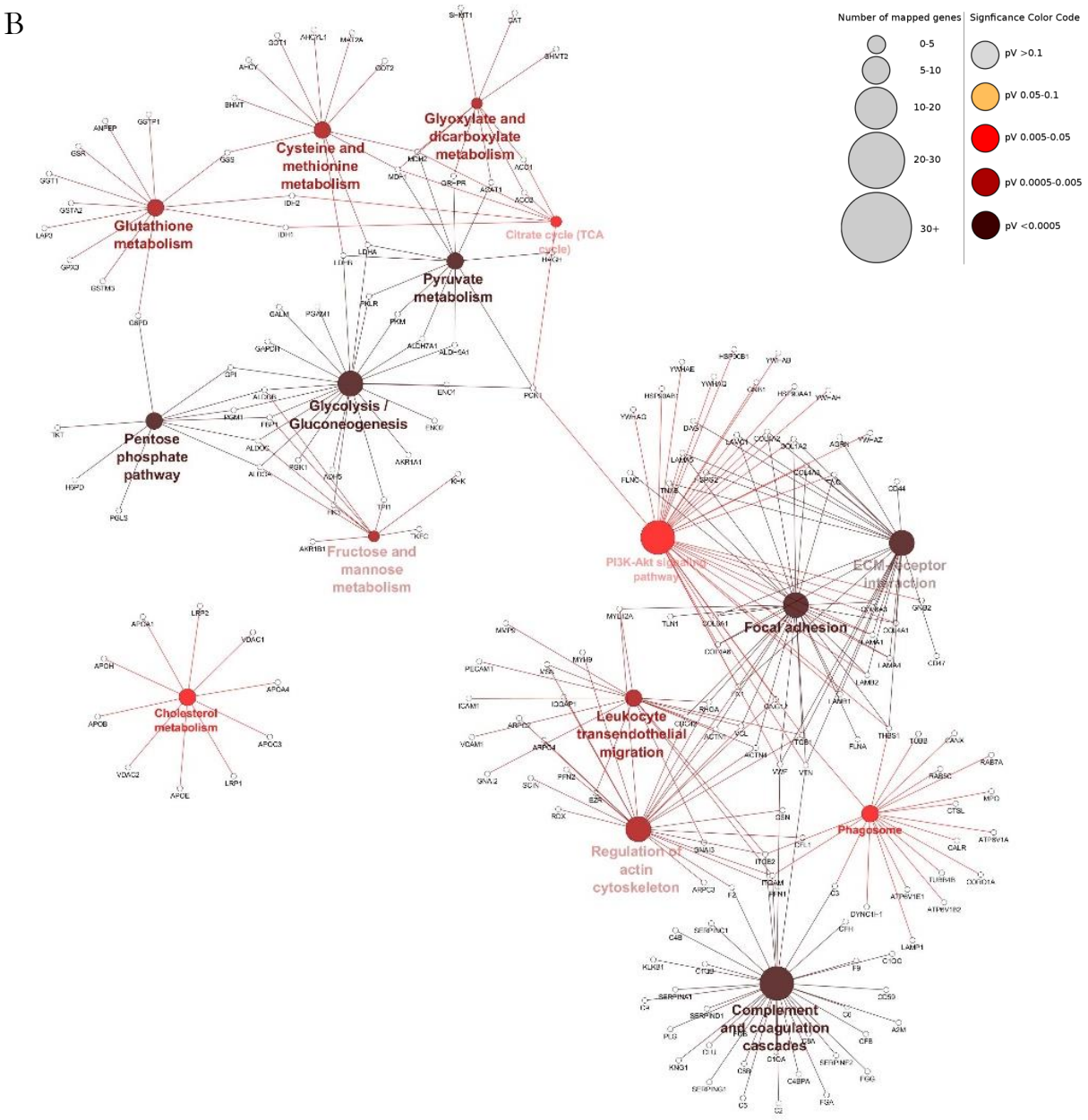


Fig. 1 Top Pathways involved in ccRCC. **A:** annotations of proteins and pathways shared by urine and tissue of ccRCC patients. **B:** Networking showing the main pathways and their significance (light to dark red, p value < 0.05).

Urinary secreted proteins varied according to different grades and stages

We further investigated the correlation between urinary protein excretion and tumour classification (Stage and Grade), with a specific attention on the possibility to also detect morphological changes in the lesion, already investigated on tissue samples, also in urine. To enhance confidence in the protein quantification, the abundances of proteins detected in at least half of the 44 urine samples and quantified with at least two unique peptides were exported from PEAKS Studio 8.5. Statistical analysis was performed with a Mann-Whitney test (p-value ≤ 0.05, max fold change ≥ 1.5). Due to the high biological variability, no further statistical correction has been performed. Initially, the influence of the grade on the urinary proteome was evaluated. Patients with different tumour grades were split into two groups, HG (G3-G4) and LG (G1-G2) tumours, and 39 proteins were identified as differentially expressed. 10 were down and 29 upregulated in HG compared to the LG group (Supplementary Table 2). Moreover, in order to highlight alterations of the urinary proteome related to different stages, we also evaluated the differences between stage 1 and stage 3-4 classes. We were able to detect 79 proteins with a statistically significant variation in their abundances: 16 proteins were downregulated, while 63 were upregulated in stage 3-4 versus stage 1 patients (Supplementary Table 3). Urinary proteins altered according to the stage (stage 1 versus stage3-4) and grade (LG versus HG) were compared and (Figure 2A) the 15 proteins commonly altered also showed the same pattern: 4 were overexpressed both in LG and in low stage samples, conversely 11 were overexpressed in HG and stage 3-4 lesions (Figure 2B). Therefore, we hypothesise that the co-presence of the two features (Stage and Grade) could have determined the intersection highlighted in Figure 2.

A

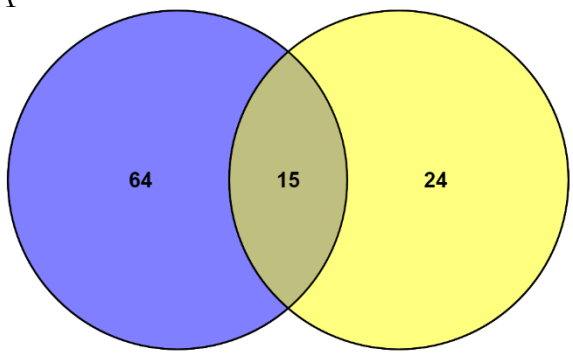


Fig.2. A: Venn diagram of the urinary proteins altered according to the stage (Stage1 vs Stage3-4, blue) and grade (G1-2 versus G3-4, yellow). **B:** List of 15 proteins commonly altered in grade and stage comparison and their expression.

B	Stage and Grade	
	Up-regulated in:	
PGD	HG	Stage 3
LAMC1	HG	Stage 3
LPA	HG	Stage 3
GDF15	HG	Stage 3
IGKV3-20	HG	Stage 3
GGCT	HG	Stage 3
MMRN2	HG	Stage 3
ACVR1B	HG	Stage 3
IGLV1-40	HG	Stage 3
AMY1A-B-C	LG	Stage 1-2
GCHFR	HG	Stage 3
THBS4	LG	Stage 1-2
IGF2R	HG	Stage 3
S100P	LG	Stage 1-2
UMOD	LG	Stage 1-2

In order to better understand which of these clinical parameters is responsible for the urinary protein alterations, we proceeded by sub-classifying our cohort of patients considering each feature a time. At first, we compared patients with low grade and stage 1 lesions with those with low grade and stage 3-4 lesions (group A and B respectively) and, then, patients with stage 1 and stage 3-4 lesions at high grade (group C and D). We identified 6 proteins upregulated in stage 3 compared to stage 1, when excluding the grade as a confounding factor (Table 2).

Subsequently, we investigated how the expression of urinary proteins reflected tumour grade. Firstly, we compared patients with lesions at different Grade (low grade -group A- and high grade - group C) and stage 1 tumours and, consecutively, group B and D of patients with stage 3-4 lesions and low and high grade, respectively (Table 3). This comparison showed no proteins correlating specifically with the grade, independently from the stage.

Table 2. Differently expressed proteins (Gene-name) in stage comparisons

Stage 1 vs Stage 3-4 Low Grade (A vs B)		**Stage 1 vs Stage 3-4 Grade independent**	**Stage 1 vs Stage 3 High Grade (C vs D)**
ACE2	IGFBP6	APOA1	ACTN4
ACVR1B	IGHG3	GDF15	AMY1A-B-C
ACY1	IGLV1-51	IGLV1-40	B2M
ADGRF5	L1CAM	PTPRG	C11orf54
APOA4	LAMC1	RBP4	CFHR1
ARSA	LGMN	STOM	CHIT1
ATP1A1	LYZ		DAG1
CAPZB	MADCAM1		ENO1
CD7	MRC2		F2
CDH11	PATJ		GDI2
CFH	PGD		GOLM1
CLEC7A	PLAUR		GPC3
CLN5	PLXDC2		GSS
COL3A1	POTEF		HRG
COLEC12	PPIA		HSPG2
CTSL	PRG2		LAIR1
ENG	RETN		LGALS3
FLG	ROR1		NEGR1
GC	S100P		PCDHGC3
GGCT	SERPINA1		PROM2
GOT1	SFN		PRTN3
GPC1	SORL1		SECTM1
GPC4	TF		SELL
GSTO1			SOD1

Table 3. Differently expressed proteins (Gene-name) in grade comparisons

Low Grade vs High Grade Stage 1 (A vs C)		Low Grade vs High Grade Stage independent	Low Grade vs High Grade Stage 3-4 (B vs D)
ACVR1B	GSS		ADGRF5
ADH5	IGLV3-19		GOLM1
ALB	LGALS3		GPC1
B4GALT1	LYPD3		HSPA5
CD58	MACF1		LAIR1
CFHR1	MUC5AC		MASP2
CLEC7A	PCDHGC3		MMRN2
COL5A3	PROM2		PATJ
COTL1	SERPINA1		PRG2
CP	STOM		SFN
CTSH	TF		
EPHA1	THBS4		
EPHB3	TYRO3		
FCER2	XPNPEP2		
FTL	ZNF469		

4. Preliminary conclusions

Data obtained in this study are still under evaluation for their biological and clinical implication in ccRCC. However, on the basis of the preliminary findings, I can summarise the conclusions made thus far:

- Those proteins detected in both ccRCC urine and tissue samples are mainly involved in metabolic pathways and in the activation of immune response. Information provided by urine and tissue have a considerable overlap.

- Six proteins whose relative abundance varied in patients with tumours of different size/stage, independently from the grade, were detected. However, no protein alterations could be associated solely to the grade of the lesion.
- These results lead us to believe that the urinary proteome of ccRCC patients is influenced primarily by the dimension and position and only partially by the grade of the tumour. To obtain information about the cell morphology of the lesion, tissues still remain the sample of choice.
- To further validate these results a larger samples cohort of patients will be enrolled. The numerosity of the samples will be estimated based on the variability of the signals and the needed statistical power.

References

[1] C. Chinello, V. L'Imperio, M. Stella, A.J. Smith, G. Bovo, A. Grasso, M. Grasso, F. Raimondo, M. Pitto, F. Pagni, F. Magni, The proteomic landscape of renal tumors, Expert Rev. Proteomics. (2016). doi:10.1080/14789450.2016.1248415.

[2] A.Q. Haddad, V. Margulis, Tumour and patient factors in renal cell carcinoma - Towards personalized therapy, Nat. Rev. Urol. (2015). doi:10.1038/nrurol.2015.71.

[3] J.J. Hsieh, M.P. Purdue, S. Signoretti, C. Swanton, L. Albiges, M. Schmidinger, D.Y. Heng, J. Larkin, V. Ficarra, Renal cell carcinoma., Nat. Rev. Dis. Prim. (2017). doi:10.1038/nrdp.2017.9.

[4] H. Moch, A.L. Cubilla, P.A. Humphrey, V.E. Reuter, T.M. Ulbright, The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs—Part A: Renal, Penile, and Testicular Tumours, Eur. Urol. 70 (2016) 93–105. doi:10.1016/j.eururo.2016.02.029.

[5] J. Dagher, B. Delahunt, N. Rioux-Leclercq, L. Egevad, J.R. Srigley, G. Coughlin, N. Dunglinson, T. Gianduzzo, B. Kua, G. Malone, B. Martin, J. Preston, M. Pokorny, S. Wood, J. Yaxley, H. Samaratunga, Clear cell renal cell carcinoma: validation of World Health Organization/International Society of Urological Pathology grading, Histopathology. (2017). doi:10.1111/his.13311.

[6] T. Geiger, S.F. Madden, W.M. Gallagher, J. Cox, M. Mann, Proteomic portrait of human breast cancer progression identifies novel prognostic markers, Cancer Res. (2012). doi:10.1158/0008-5472.CAN-11-3711.

[7] V. Sandim, D.A. Pereira, A.A. Ornellas, G. Alves, Renal cell carcinoma and proteomics, Urol. Int. (2010). doi:10.1159/000296283.

[8] D. Quandt, H. Dieter Zucht, A. Amann, A. Wulf-Goldenberg, C. Borrebaeck, M. Cannarile, D. Lambrechts, H. Oberacher, J. Garrett, T. Nayak, M. Kazinski, C. Massie, H. Schwarzenbach, M. Maio, R. Prins, B. Wendik, R. Hockett, D. Enderle,M.

Noerholm, H. Hendriks, H. Zwierzina, B. Seliger, Implementing liquid biopsies into clinical decision making for cancer immunotherapy, Oncotarget. (2015). doi:10.18632/oncotarget.17397.

[9] A. Di Meo, J. Bartlett, Y. Cheng, M.D. Pasic, G.M. Yousef, Liquid biopsy: A step forward towards precision medicine in urologic malignancies, Mol. Cancer. (2017). doi:10.1186/s12943-017-0644-5.

[10] F. Magni, Y.E.M. Van Der Burgt, C. Chinello, V. Mainini, E. Gianazza, V. Squeo, A.M. Deelder, M.G. Kienle, Biomarkers discovery by peptide and protein profiling in biological fluids based on functionalized magnetic beads purification and mass spectrometry, Blood Transfus. 8 (2010) 92–97. doi:10.2450/2010.015S.

[11] F. Raimondo, L. Morosi, C. Chinello, F. Magni, M. Pitto, Advances in membranous vesicle and exosome proteomics improving biological understanding and biomarker discovery, Proteomics. (2011). doi:10.1002/pmic.201000422.

[12] J. Klein, J.L. Bascands, H. Mischak, J.P. Schanstra, The role of urinary peptidomics in kidney disease research, Kidney Int. (2016). doi:10.1016/j.kint.2015.10.010.

[13] J.J. Morrissey, J. Mobley, J. Song, J. Vetter, J. Luo, S. Bhayani, R.S. Figenshau, E.D. Kharasch, Urinary concentrations of aquaporin-1 and perilipin-2 in patients with renal cell carcinoma correlate with tumor size and stage but not grade, Urology. 83 (2014) 256.e9-256.e14. doi:10.1016/j.urology.2013.09.026.

[14] M. Mijuskovic, I. Stanojevic, N. Milovic, S. Cerovic, D. Petrovic, D. Maksic, B. Kovacevic, T. Andjelic, P. Aleksic, B. Terzic, M. Djukic, D. Vojvodic, Tissue and urinary KIM-1 relate to tumor characteristics in patients with clear renal cell carcinoma, Int. Urol. Nephrol. (2018). doi:10.1007/s11255-017-1724-6.

[15] V. Sandim, P. Da, K. De, O.C. Al, O. Aa, G. Alves, Proteomic analysis reveals differentially secreted proteins in the urine from patients with clear cell renal cell

carcinoma. PubMed Commons, Urol. Oncol. Semin. Orig. Investig. 34 (2015) 1–2. doi:10.1016/j.urolonc.2015.07.016.

[16] M. Stella, C. Chinello, A. Cazzaniga, A. Smith, M. Galli, I. Piga, A. Grasso, M. Grasso, M. Del Puppo, M. Varallo, G. Bovo, F. Magni, Histology-guided proteomic analysis to investigate the molecular profiles of clear cell Renal Cell Carcinoma grades, J. Proteomics. (2018). doi:10.1016/j.jprot.2018.04.028.

[17] C. Chinello, M. Stella, I. Piga, A.J. Smith, G. Bovo, M. Varallo, M. Ivanova, V. Denti, M. Grasso, A. Grasso, M. Del Puppo, A. Zaravinos, F. Magni, Proteomics of liquid biopsies: Depicting RCC infiltration into the renal vein by MS analysis of urine and plasma, J. Proteomics. (2018). doi:10.1016/j.jprot.2018.04.029.

[18] B. Ma, K. Zhang, C. Hendrie, C. Liang, M. Li, A. Doherty-Kirby, G. Lajoie, PEAKS: Powerful software for peptide de novo sequencing by tandem mass spectrometry, Rapid Commun. Mass Spectrom. (2003). doi:10.1002/rcm.1196.

[19] J. Dagher, B. Delahunt, N. Rioux-Leclercq, L. Egevad, J.R. Srigley, G. Coughlin, N. Dunglinson, T. Gianduzzo, B. Kua, G. Malone, B. Martin, J. Preston, M. Pokorny, S. Wood, J. Yaxley, H. Samaratunga, Clear cell renal cell carcinoma: validation of World Health Organization/International Society of Urological Pathology grading, Histopathology. 71 (2017) 918–925. doi:10.1111/his.13311.

[20] Y. Song, L. Zhong, J. Zhou, M. Lu, T. Xing, L. Ma, J. Shen, Data-Independent Acquisition-Based Quantitative Proteomic Analysis Reveals Potential Biomarkers of Kidney Cancer, Proteomics - Clin. Appl. (2017). doi:10.1002/prca.201700066.

[21] A. Atrih, M.A. V Mudaliar, P. Zakikhani, D.J. Lamont, J.T. Huang, S.E. Bray, G. Barton, S. Fleming, G. Nabi, Quantitative proteomics in resected renal cancer tissue for biomarker discovery and profiling., Br. J. Cancer. 110 (2014) 1622–33. doi:10.1038/bjc.2014.24.

[22] A. Zaravinos, M. Pieri, N. Mourmouras, N. Anastasiadou, I. Zouvani, D. Delakas, C. Deltas, Altered metabolic pathways in clear cell renal cell carcinoma: A meta

-analysis and validation study focused on the deregulated genes and their associated networks., Oncoscience. (2014). doi:10.18632/oncoscience.13.

CHAPTER 5

Summary, Conclusion and Future Prospective

1. Summary

This final section will briefly discuss the data obtained, with a particular focus on the future prospective of this line of work. Clear cell renal cell carcinoma represents the most frequent form of kidney cancer. Given that total, or partial, nephrectomy remains the gold standard for the routine treatment, there is a strong need for the detection of prognostic biomarkers that can be translated into less invasive tools or treatments. In this context, various proteomic techniques have been used to determine the molecular changes related to disease progression and early pathological modifications [1,2].

Bottom-up proteomics, a mass spectrometric (MS)-based approach for the characterisation of peptides obtained from an in-solution protein digestion, has been used to investigate protein identity and abundance in both biofluids and tissue samples. Shotgun proteomics provides indirect information of proteins through the analysis of proteolytic digested (tryptic) peptides. The peptide mixture is then fractionated and subjected to nLC-MS/MS analysis and protein identity is achieved by comparing the obtained MS/MS spectra with those generated theoretically by in silico digestion of a protein database [3].

Matrix-Assisted Laser Desorption/Ionisation-Mass Spectrometry Imaging is a modern proteomic technology that is capable of detecting several different classes of analytes directly in situ. In this study, the spatial distribution of tryptic peptides has been used to build a molecular image of the tissue that has then been correlated with pathological alterations that are evident via routine histology (tumour lesion at different grades). Therefore, MALDI-MSI was employed due to this capability to detect new, specific, signatures of ccRCC directly on tissue.

The work presented here represents an extensive study focused on ccRCC. We initially started with a more technical approach by investigating the different information that could be obtained from two different biofluids (urine and plasma). We then investigated

the proteomic profile directly on tissue in order to better understand the molecular changes associated with grade progression. Following in this vein, we also proceeded with a more in-depth study of the urinary proteome, extending the sample cohort and focusing our attention on the correlation between protein expression and clinical features. The positive results obtained with these proteomic approaches facilitated the detection of a number of protein signals that could differentiate among various ccRCC stages and grades in both urine and tissue, and such signals could potentially represent future targets to be investigated as diagnostic or prognostic markers of the most common morphotypes of renal cancer.

2. Conclusions

2.1 Proteomics of liquid biopsies: depicting RCC infiltration into the renal vein by MS analysis of urine and plasma

The results presented in this work highlight the importance of integrating information from different samples. Not only this, the data highlighted that, even though the information delivered by the two bodily fluids are partially complementary, the urine seems to provide more information regarding how the tumour affects the kidney architecture.

This study laid the groundwork for a more in-depth study of the urinary proteome, through the analysis of a larger patient cohort, in order to better investigate the pathological processes and the proteomic changes associated with the progression of ccRCC.

2.2 Histology-guided proteomic analysis to investigate the molecular profiles of clear cell Renal Cell Carcinoma grades

This study shows a promising application of MALDI-MSI coupled with nLC-ESI-MS/MS in the discovery of the proteins or pathway alterations for the assessment of ccRCC progression. Furthermore, it highlights the potential to correlate in situ tissue findings with those obtained from tissue homogenates using complimentary mass spectrometric techniques.

MALDI-MSI was able to generate molecular signatures of different grades, with six particular signals (m/z 944.71; m/z 1032.78; m/z 1325.99; m/z 872.63; m/ z 914.64 and m/z 1428.92), three of them identified as Vimentin, Histone H2A,Histone H4, being differently expressed in areas of G1 lesions with respect to G4. The data obtained by MALDI-MSI was integrated with that obtained from quali-qantitavive analysis using nLC-ESI-MS/MS. Differences in protein content among the four grades appeared to correlate with Cancer Specific Survival rate at 10 years post-surgery, with grade 4 presenting peculiar molecular features, that were highly distinct with respect to the lower grades, thus reflecting the morphological classification. Moreover, four proteins presented a different expression in the comparisons among the three lower grades. These signals, along with the metabolic pathways that were shown to be altered, could potentially represent a starting point in a study aimed at clarifying the molecular events associated with the development of ccRCC.

2.3 Effects of clear cell Renal Cell Carcinoma stage and grade on urinary proteomic profiles

Notwithstanding limitations related to the direct translation of our findings into a tool useful for the clinics, the potential to understand how kidney alterations are reflected by urinary protein expression opens up numerous doors for a large number of clinically-relevant studies. Given that bodyfluids can be obtained with non-invasive techniques, the possibility to implement a "liquid biopsy" in clinical routine is a particularly hot topic and the potential to analyse these samples means that much larger sample cohorts can

be obtained. We demonstrate the possibility to detect in urine alterations determined by the stage of the lesion, however, to study morphological alterations within the tumour cells, tissue samples still remain essential.

3. Future perspectives

The potential of MALDI-MSI to detect disease biomarkers directly in situ is undoubted [4] and the study of renal cancer, using this modern proteomic technology, may eventually herald an era where these findings are pioneering for a clinical relevant discovery. This would be a result not only of the ability to correlate proteomic findings with morphological alterations, but also as a result of the capability to detect other molecular alterations that occur before any visible morphological alterations can be noted by the pathologist [5].

Further work related to the study of ccRCC tissue is planned, with the final aim being to better understand the biological alterations associated with ccRCC grade progression. In doing so, it is hoped that potential therapeutic targets, to be used as a valid alternative to nephrectomy, can be identified. Aiming at better clarifying proteomic signatures specific for each grade, along with the possibility to clarify the biological processes involved, the foundations for future studies has been laid.

The list of proteins found to be altered in the four grades has been further evaluated in order to explore the molecular pathways involved in cancer progression, from grade 1 to grade 4. Gene ontology (GO) and pathway analysis term clustering were performed using ClueGO v2.5 and CluePedia version 1.5 on Cytoscape v3.6. Visualisation and pathway statistical interpretation was carried out in PathVisio v3.2. The results obtained are now under evaluation, but preliminary results suggest that many metabolic pathways are involved in clear cell renal cell carcinoma progression, such as glycolysis/gluconeogenesis, Kreb's cycle, glycogen metabolism, etc. Most of those

pathways resulted to be upregulated in the first three grade, whereas, when the disease progresses to grade 4, become downregulated. A preliminary hypothesis is that in the progression to the highest grade, all cellular processes are decreasing, and cells became unable to produce energy and remove the excess of ketone bodies as normal cells do.
Moreover, considering that glycosylation is one of the most common protein modifications and, based on their recognition, contribute to protein folding, cell growth, interaction, differentiation and cancer, thus the study of their expression in different grades might be of particular interest. Since glycan changes have been shown to be a phenotypic reflection of the disease state, they can also offer a broad range of potential markers or treatment targets [6–10].
Therefore, during a six-month internship at LUMC, starting from the study published by Holst S. et al [11], a protocol for the direct analysis of N-glycans on tissue has been optimised in order to increase the detection of N-glycans and to apply this protocol using both a MALDI-TOF and a MALDI-FTICR instrument. This protocol will then be used to analyse N-glycan expression on a new, larger, cohort of 25 FFPE tissue sections from patients with ccRCC and their relative abundance will be evaluated among the different grades. However, it should also be stressed that the technical approach optimised and learnt here is relevant for application in many other studies and thus offers an additional molecular dimension in our search to better understand the patho-physiology of different diseases.
Taking the advantage of the possibility to use a high mass resolution instrument (9.4 T MALDI-FTICR), tryptic peptide analysis of a consecutive tissue section has also been performed. The data analysis has not been performed yet, but we believe that the increased chemical sensitivity offered by the MALDI-FTICR may also bring an additional level of information at the protein level to better characterise ccRCC grades. By combining the additional work proposed here, it is hoped that we can further enforce the potential role of MALDI-MSI in clinical research, from both a technical and

application standpoint. Furthermore, by taking advantage of these new possibilities related to higher spatial resolution and high mass accuracy, we hope that this can lead to the direct translation of these findings into diagnostic, or prognostic, tools that can improve the clinical management of ccRCC patients.

References

[1] J.Y. Chan, Y. Choudhury, M.H. Tan, Predictive molecular biomarkers to guide clinical decision making in kidney cancer: Current progress and future challenges, Expert Rev. Mol. Diagn. (2015). doi:10.1586/14737159.2015.1032261.

[2] Q.K. Li, C.P. Pavlovich, H. Zhang, C.R. Kinsinger, D.W. Chan, Challenges and opportunities in the proteomic characterization of clear cell renal cell carcinoma (ccRCC): A critical step towards the personalized care of renal cancers, Semin. Cancer Biol. (2018). doi:10.1016/j.semcancer.2018.06.004.

[3] Y. Zhang, B.R. Fonslow, B. Shan, M.C. Baek, J.R. Yates, Protein analysis by shotgun/bottom-up proteomics, Chem. Rev. (2013). doi:10.1021/cr3003533.

[4] J. Quanico, J. Franck, M. Wisztorski, M. Salzet, I. Fournier, Progress and potential of imaging mass spectrometry applied to biomarker discovery, in: Methods Mol. Biol., 2017. doi:10.1007/978-1-4939-6952-4_2.

[5] M. Dilillo, B. Heijs, M. Liam A., Mass spectrometry imaging: How will it affect clinical research in the future?, Expert Rev Proteomics. 15 (2018) 709–716. doi:10.1080/14789450.2018.1521278.

[6] R.L. Schnaar, Glycobiology simplified: diverse roles of glycan recognition in inflammation, J. Leukoc. Biol. (2016). doi:10.1189/jlb.3RI0116-021R.

[7] S. V. Glavey, D. Huynh, M.R. Reagan, S. Manier, M. Moschetta, Y. Kawano, A.M. Roccaro, I.M. Ghobrial, L. Joshi, M.E. O'Dwyer, The cancer glycome: Carbohydrates as mediators of metastasis, Blood Rev. (2015). doi:10.1016/j.blre.2015.01.003.

[8] J.M. Tarbell, L.M. Cancel, The glycocalyx and its significance in human medicine, J. Intern. Med. (2016). doi:10.1111/joim.12465.

[9] S.S. Pinho, C.A. Reis, Glycosylation in cancer: Mechanisms and clinical implications, Nat. Rev. Cancer. (2015). doi:10.1038/nrc3982.

[10] P. Zhang, S. Woen, T. Wang, B. Liau, S. Zhao, C. Chen, Y. Yang, Z. Song, M.R. Wormald, C. Yu, P.M. Rudd, Challenges of glycosylation analysis and control: An integrated approach to producing optimal and consistent therapeutic drugs, Drug Discov. Today. (2016). doi:10.1016/j.drudis.2016.01.006.

[11] S. Holst, B. Heijs, N. De Haan, R.J.M. Van Zeijl, I.H. Briaire-De Bruijn, G.W. Van Pelt, A.S. Mehta, P.M. Angel, W.E. Mesker, R.A. Tollenaar, R.R. Drake, J.V.M.G. Bovée, L.A. McDonnell, M. Wuhrer, Linkage-Specific in Situ Sialic Acid Derivatization for N-Glycan Mass Spectrometry Imaging of Formalin-Fixed Paraffin-Embedded Tissues, Anal. Chem. (2016). doi:10.1021/acs.analchem.6b00819.

LIST OF PUBLICATIONS

1. Smith A, L'Imperio V, Denti V, Mazza M, Ivanova M, **Stella M**, Piga I1, Chinello C, Ajello E, Pieruzzi F, Pagni F, Magni F. *High Spatial Resolution MALDI-MS Imaging in the Study of Membranous Nephropathy,* In press Proteomics Clin Appl. 2018 Dec 7:e1800016. doi: 10.1002/prca.201800016
2. Piga I*, Capitoli G*, Tettamanti S, Denti V, Smith A, Chinello C, **Stella M**, Leni D, Garancini M, Galimberti S, Magni F and Pagni F. *Feasibility study for the MALDI-MSI analysis of thyroid fine needle aspiration biopsies: evaluating the morphological and proteomic stability over time,* In press Proteomics Clin Appl. 2018 Nov 9:e1700170. doi: 10.1002/prca.201700170. *These authors contributed equally to this work.
3. L'Imperio V*, Smith A*, Ajello E, Piga I, **Stella M**, Denti V, Tettamanti S, Sinico RA, Pieruzzi F, Garozzo M, Vischini G, Nebuloni M, Pagni F and Magni F. *MALDI-MSI pilot study highlights glomerular deposits of macrophage migration inhibitory factor (MIF) as a possible indicator of response to therapy in Membranous Nephropathy,* In press Proteomics Clin Appl. 2018 Oct 25:e1800019. doi: 10.1002/prca.201800019. *These authors contributed equally to this work.
4. Raimondo F, Chinello C, **Stella M**, Santorelli L, Magni F, Pitto M. *Effects of hematuria on the proteomic profile of urinary extracellular vesicles: technical challenges,* J Proteome Res. 2018 Aug 3;17(8):2572-2580. doi: 10.1021/acs.jproteome.7b00763
5. Severi, L, Losi L, Fonda S., Taddia L, Gozzi, G, Marverti, G, Magni F, Chinello C, **Stella M**, Sheouli J, Braicu EI, Genovese F, Lauriola A, Marraccini C, Gualandi A, D'Arca D, Ferrari S, Costi MP. *Proteomic and bioinformatic studies for the characterization of response to pemetrexed in platinum drug resistant ovarian cancer*, Front Pharmacol. 2018 May 8;9:454. doi: 10.3389/fphar.2018.00454.
6. **Stella M***, Chinello C*, Cazzaniga A, Smith A, Galli M, Piga I, Grasso A, Grasso M, Del Puppo M, Varallo M, Bovo G, Magni F. *Histology-guided proteomic analysis to*

investigate the molecular profiles of clear cell Renal Cell Carcinoma grades, Journal of proteomics, J Proteomics. 2019 Jan 16;191:38-47. doi: 10.1016/j.jprot.2018.04.028 *These authors contributed equally to this work.

7. Chinello C, **Stella M**, Piga I, Smith AJ, Bovo G, Varallo M, Ivanova M, Denti V, Grasso M, Grasso A, Del Puppo M, Zaravinos A, Magni F. *Proteomics of liquid biopsies: Depicting RCC infiltration into the renal vein by MS analysis of urine and plasma,* Journal of proteomics, 2018 Apr 22. pii: S1874-3919(18)30179-9. doi: 10.1016/j.jprot.2018.04.029
8. Smith A, Galli M, Piga I, Denti V, **Stella M**, Chinello C, Fusco N, Leni D, Manzoni M, Roversi G, Garancini M, Pincelli AI, Cimino V, Capitoli G, Magni F, Pagni F. *Molecular signatures of medullary thyroid carcinoma by matrix-assisted laser desorption/ionisation mass spectrometry imaging,* J Proteomics. 2019 Jan 16;191:114-123. doi: 10.1016/j.jprot.2018.03.021
9. Smith A*, Piga I*, Galli M*, **Stella M***, Denti V, Del Puppo M, Magni F. *"Matrix-Assisted Laser Desorption/Ionisation Mass Spectrometry Imaging in the Study of Gastric Cancer: A Mini Review"* nt J Mol Sci. 2017 Dec 1;18(12). pii: E2588. doi: 10.3390/ijms18122588. * These authors contributed equally to this work.
10. L'Imperio V, Bruno I, Rabach I, Smith A, Chinello C, **Stella M**, Magni F, Pagni F. *"Histoproteomic Characterization of Localized Cutaneous Amyloidosis in X-Linked Reticulate Pigmentary Disorder."* Skin Pharmacol Physiol. 2017; 30(2):90-93. doi: 10.1159/000464336.
11. Galli M, Pagni F, De Sio G, Smith A, Chinello C, **Stella M**, L'Imperio V, Manzoni M, Garancini M, Massimini D, Mosele N, Mauri G, Zoppis I, and Magni F. "*Proteomic profiles of thyroid tumors by Mass Spectrometry-Imaging on Tissue Microarrays"*. Biochim Biophys Acta Proteins Proteom. 2017 Jul;1865(7):817-827. doi: 10.1016/j.bbapap.2016.11.020

12. Chinello C, L'Imperio V, **Stella M**, Smith A, Bovo G, Grasso A, Raimondo F, Pitto M, Pagni F, and Magni F. *"The proteomic landscape of renal tumors"*. Expert Review of Proteomics. 2016. Dec;13(12):1103-1120.
13. Govaert E, Van Steendam K, Scheerlinck E, Vossaert L, Meert P, **Stella M**, Willems S, De Clerck L, Dhaenens M, Deforce D. *"Extracting histones for the specific purpose of label-free MS"*. Proteomics. 2016 Dec; 16(23):2937-2944
14. Smith A, L'Imperio V, Ajello E, Ferrario F, Mosele N, **Stella M**, Galli M, Chinello C, Pieruzzi F, Spasovski G, Pagni F, and Magni F. *"The putative role of MALDI-MS Imaging in the study of Membranous Glomerulonephritis"*. Biochim Biophys Acta Proteins Proteom. 2017 Jul;1865(7):865-874. doi: 10.1016/j.bbapap.2016.11.013

ACKNOWLEDGEMENTS

Alla fine di questa esperienza, pronta per voltare pagina, non mi resta che ringraziare le persone con cui ho condiviso questa parte del viaggio. *At the end of this experience, ready to face the next step, it is time to acknowledge those who I have shared this chapter of my life with.*

Mamma, papà, a voi che mai avete smesso di supportarmi e a Ga, che mai ha smesso di stressarmi, il mio primo grazie e la mia immensa gratitudine.

A te che ci sei sempre stato, che non hai mai smesso di credere in me e sostenere le mie scelte, anche quando già sapevi che la delusione sarebbe potuta essere dietro l'angolo. Alla persona più paziente, generosa e affascinante che conosca, al mio compagno di avventure, a te va il mio più grande e sincero grazie.

Prof. Magni, per quanto ho imparato a livello professionale, ma anche personale non posso che esserle infinitamente grata. Fare parte del suo gruppo è stato per me da sprono, ma anche motivo di orgoglio. La possibilità di discutere intellettualmente alla pari con lei è stato uno dei gesti che ho più apprezzato: ha stimolato il mio entusiasmo e la mia innata curiosità, insegnandomi però anche ad essere (auto)critica e propositiva.

Dr. Bovo, grazie per la pazienza e il tempo che mi ha dedicato, per tutto ciò che mi ha insegnato, per non aver mai smesso di rispondere alle mie mille curiosità e domande. Anche se il tempo non è mai abbastanza, il contributo di voi medici alla ricerca è indispensabile. Le auguro di poter continuare a coltivare l'entusiasmo e la curiosità scientifica che la contraddistinguono e che mi hanno immediatamente contagiato.

Michele, con te ho condiviso (a distanza) ogni momento di questo dottorato. Grazie per avermi supportato e sopportato dal primo colloquio fino ad oggi. Ammiro la tua intelligenza e determinazione, sono certa che questo dottorato sarà per te solo la prima di tantissime soddisfazioni.

Andy e Niccolò, l'anno passato tutti e tre insieme rimane forse uno dei periodi più belli di questo dottorato. Andy, guadagnarmi la tua fiducia e rispetto prima, e la tua amicizia poi, non è stata cosa facile. Ora posso dire che averti avuto come collega (e amico) has

been one of the most enriching part of this experience. Niccolò, lavorare con te è da subito stato semplice, come lavorare con un'altra me, ma due caratteri troppo simili ad un certo punto non possono che scontrarsi. Il rispetto e l'amicizia che ho nei tuoi confronti andranno sempre oltre qualsiasi discussione (e coppino!).
Le risate fatte in lab e i momenti condivisi in U8 resteranno per sempre splendidi ricordi di questi tre anni insieme, ai colleghi e amici che ho conosciuto in questi tre anni, grazie!
Grazie Anna, ho imparato da te molto di più di quanto io ti abbia insegnato.
Grazie ai pazienti, ai medici, agli infermieri e ai tecnici del S. Gerardo. Senza di voi il mio lavoro sarebbe impossibile.
Moving on to my time in Leiden. This was a challenge both professionally and personally. I was fortunate enough to meet warm-hearted individuals, both at work and at home. Cats, having you as a housemate was more than I could have wished for. Lena, without sounding too soppy, you effectively made me feel at home and all the time spent together made my six months in Leiden full of unforgettable memories. Bram, Stephy, Hans, Renè, and Fanny, thank you for welcoming me in your group. Working on N-glycans made me struggle a lot, but at the end of the day, it is the end result that matters. The possibility to collaborate on other side projects (Gonzalo) made this experience more challenging and interesting and I really appreciate your confidence in me. Martha, you have been both colleague, as well as friend, thank you and good luck with your future plans!

From friends I have known for a lifetime, to those I have known only briefly, to my laboratory colleagyes, to "my" student, also to those I have had the pleasure of meeting in Leiden and Gent, to all of you, I thank you. You have all been part of this path, you have all given me and taught me something and, more importantly, you have made me into the person I have become today!
Agli amici da una vita, agli amici da poco, ai compagni di lab, ai "miei" studenti, agli amici di Leiden e quelli di Gent, a voi va il mio grazie. Siete tutti stati parte di questo percorso, mi avete tutti dato e insegnato qualcosa, mi avete resa quella che sono oggi!

Be yourself; everyone else is already taken *(unattributed)*