Department of Medicine and Surgery

PhD program: Molecular and Translational Medicine

Cycle: XXXII

Proteomic Analysis of Urine-based Liquid Biopsy to provide new insights into Renal Diseases

Surname: **SANTORELLI** Nome: **LUCIA**
Registration number: **823654**

Tutor: **Prof. M. Pitto** Co-tutor: **Dr. F. Raimondo**
Coordinator: **Prof. A. Biondi**

ACADEMIC YEAR 2018/2019

Table of Contents

CHAPTER 1

INTRODUCTION

1. Kidney Related Diseases

Kidney is a remarkable organ. Blood filter and endocrine organ, it removes waste, regulates electrolytes and acid-base homoeostasis, controls fluid balance and blood pressure. In addition, it regulates bone metabolism and red blood cell production and it is strictly connected with the functioning of other organs, such as the heart and liver. However, this so formidable functional unit, in many cases, can be impaired in its activity.

Kidney disease is a general term for many heterogeneous disorders affecting kidney structure and function, such as: diabetes, autoimmune disorders, congenital anomalies, many infections and diverse cancer types. The very nature of kidney means that assessments of renal function and disease are made on the basis of non-specific and proxy markers, and these estimates are often far from perfect. In fact, the area of renal diseases is a wide and complex field, and many conditions leading to them are not fully understood or curable [1]. Along with a better understanding of these pathologies, there is a need for improved risk detection, for determination of prognosis and for an improved and personalized treatment [2].

Much research focusing on biomarkers and targets for treatment of diverse renal disorders is, by necessity, preliminary and preclinical. Research in kidney disease has come a very long way, and has a very long way to go.

This work was mainly focused on two renal pathologies: clear cell Renal Cell Carcinoma (ccRCC), the most frequent and aggressive type of renal carcinoma, and Idiopathic Nephrotic Syndrome (INS), the major childhood glomerular disease.

1.1 clear cell – Renal Cell Carcinoma - subtype of RCC

Renal Cell Carcinoma.

Renal Cell Carcinoma (RCC) is one of the most frequent malignancies and it represents 2-3% of all cancers, with the highest incidence rates in developed countries. Approximately 90% of all kidney cancers are RCC; this group comprises

different types with specific histopathological and genetic characteristics [3]. There is a 1.5:1 RCC predominance in men over women, and a correlation with age; RCC represents 2% of paediatric cancer, it is rare before 20 years old and the peak of incidence is between 40 and 60 years old.

The anatomical position of the kidney causes symptomatology's delay. Adipose tissue within Gerota's fascia and retroperitoneum gives protection to the viscera and to somatic structures around the kidney, making their involvement tardive. Therefore, the appearance of persistent pain or palpable tumefaction are delayed, too. In addition, RCC tends to metastasise before giving any local sign or symptom; so, in more than one third of patients, the tumour is already metastatic at diagnosis. RCC metastasis could involve any organs, but the most commons are: lungs, liver, bones, subcutaneous tissues and central nervous system [4].

The prognosis of RCC patients is clearly correlated with anatomical (tumour size and localisation, lymph nodes or adrenal involvement, distant metastasis, etc.) and histological parameters (e.g. tumour morphotype, nuclear grading, sarcomatoid features, etc.). Nowadays, radical nephrectomy is the only effective therapy for the treatment of the major localised RCC histotypes. It consists in the removal of the kidney, the perirenal tissue contained in the Gerota's fascia, the suprarenal gland, and the loco-regional lymph nodes. Recently, active surveillance for small size neoplasms ($\leq$ 4 cm in diameter) that do not involve the excretory pathway represents a valid alternative to surgery [5].

clear cell Renal Cell Carcinoma Among the histological subtypes of RCC, clear cell Renal Cell Carcinoma (ccRCC) is the most common, representing alone the 80-90% of cases of kidney cancers [6]. ccRCC has a worse prognosis compared with the other two principal subtypes of RCC (papillary and chromophobe RCC) [7], [8]. It develops from the proximal tubule of the nephron, where the cells present their surface rich in microvilli, having the function of diverse ions and water absorption. Morphologically, the tumour appears as a yellowish mass, due to the presence of cholesterol, lipids and phospholipids sometimes with hemorrhagic, necrotic or

partially cystic areas. Differently to the non-affected cells, the neoplastic ones lost the characteristic shape, compromising the tubule functionality. The name *'clear cell'* refers to the appearance of the cancer cells when viewed with a microscope. Clear cellularity, indeed, is an artefact caused by histological procedure: after the hematoxylin and eosin staining of the renal tissue, in fact, the lipid and glycogen content of the cancerous cells is removed, therefore cells appear like empty at optical microscope. Sometimes, a limited number of ccRCC contains uniquely clear cells: usually a quota of granular cells is also present. These cells have an eosinophil cytoplasm, abundant mitochondria and appear dark at microscope; this could complicate the diagnosis [9].

Most of ccRCCs generate a weak inflammatory response, however lymphocytic or neutrophilic infiltrates could be present [10], and there is a link between strong lymphocytic infiltration and unfavourable outcome [11]. Another parameter that influences tumour aggressiveness is the presence of area of necrosis [11]. Finally, ccRCC metastasises through vein cava up to the lungs and retrograde metastasis develops through paravertebral veins, testicular or ovarian veins, intra-renal veins, or along the urethra.

In general, ccRCC is a sporadic disease, but it could arise also in familial forms. The most frequent genetic mutation is chromosome 3 deletion (LOH 3p), mapped in three possible positions: 3p14 (FHIT gene), 3p21.3 e 3p25 (VHL gene) [12].

In addition, 41% of patients with ccRCC present a mutation on PBRM1 (Protein polybromo-1), a gene located on 3p chromosome [13]. Indeed, some evidences show that an epigenetic dysregulation, mediated by methylase/demethylases action on the Histone H3, may play an important role in ccRCC biology.

So far, chemotherapy and radiotherapy give poor results, as this cancer is both radio- and chemo-resistant [14]. Effective biomarkers, able to give early diagnosis of ccRCC or to allow continuous monitoring of tumour evolution and follow-up, are not yet known.

1.2 Idiopathic Nephrotic Syndrome

Idiopathic nephrotic syndrome (INS) is the most common pediatric primary glomerular disease, affecting 16–17 per 100’000 children between the ages of 2 and 8 years, with a peak of incidence between 3 and 5 years [15].

The constellation of features that characterise INS develops from primary alterations in permeability and selectivity of the capillary walls of glomeruli, leading to proteinuria. Nephrotic-range proteinuria has been variously defined, including the increasingly popular use of spot urinary protein-to-creatinine ratio higher than 0.25 g protein per mmol creatinine (or > 2.0 mg protein per mg creatinine) [16]. Additionally, The INS is also characterised by hypoalbuminemia, generalised edema, and hyperlipidemia that taken together constitute the so-called “tetrad” [17].

Glucocorticoids (GCs) are the first-choice drugs and induce remission in approximately 85–90% of paediatric INS patients, named as Steroid Sensitive (SS); however, 10–15% of patients are steroid resistant (SR) [18].

Additionally, despite initial complete remission, almost 50% of patients show recurrence of the proteinuria and are classified as Steroid Dependent (SD): these patients, after a prolonged steroid therapy, with the possibility of severe adverse effects, often need to switch to other immunomodulating or immunosuppressive drugs. Steroid responsiveness is the key determinant for the patient prognosis and those who do not respond to therapy are subjected to aggressive treatments, often developing several complications and side effects.

Although this classification of INS based on response to GC therapy shed little light on the comprehension of the disease [19], the mechanisms involved in GC dependence and resistance are poorly understood, and studies considering GC pharmacodynamics and pharmacogenetics have been performed without conclusive results [20].

2. Proteomics

Next-generation sequencing allows the analysis of genomes, including those representing disease states. After the 2001, when the Human Genome Project delivered the first draft of the human genome sequence [21], [22], the enormous information provided by DNA sequencing technology revolutionised the biology research world-wide, with particular positive return in the field of human pathologies study. However, the causes of most disorders are multifactorial, and the only sequencing of the human genome results partial and insufficient for explaining the correlations between the genotype and the disease phenotype. Other systems-level approaches, including the analysis of proteome, are required for a more comprehensive understanding.

The proteomics world is a dynamic and unforgiving space. It involves the applications of technologies for the identification and quantification of overall proteins content present in a cell, a tissue, a fluid or an organism. It supplements the other 'omics' technologies, such as Genomics and Transcriptomics, to expound the identity of proteins of an organism and to recognise the structure and functions of a particular protein. Therefore, it is fundamental for the generation of a map of the complex, interconnected pathways, networks and molecular systems (the human proteome) that, taken together, control the function of all cells, tissues, organs and organisms [23].

To date, Mass Spectrometry (MS) has arguably become the core technology in Proteomics[24], [25]. With rapid advances in its instruments and experimental methodologies, the MS-based approach has become a reliable and essential tool for elucidating biological processes at the protein level, due to its ability to provide high-throughput information, decreasing sample analysis time and increasing the depth of proteome coverage, at the same time [26].

2.1 Clinical Proteomics: a tool for markers discovery

One key to successful treatment of most pathologies, including kidney diseases, is the early detection. Biomarkers play an important role in this area as they can provide information not only about the onset of the disease thus helping in early diagnosis, but equally in prognosis, surveillance and treatment. The term marker was introduced in 1989 and it indicates a measurable biological parameter useful to detect variations from the physiological state, or the risk rather than the presence of a pathology. In 2001, the National Institute of Health (NIH) standardised the definition of biomarker as a measurable feature capable of defining a process as physiologic or morbid, and/or capable of establishing a biological response to a pharmacological treatment [27]. An ideal marker has defined specificity and sensitivity requirements, has superior clinical utility than existing markers, and it is present in a biological sample obtainable through non-invasive procedures [28]. From a biochemical point of view, biomarkers are often proteins, measurable with antibodies, whose presence or altered amount is an index of pathology [29]. However, many studies have demonstrated that a single protein is often not sufficient to uniquely identify a disease and/or to evaluate its progression, since most diseases are derived from polygenic alterations. This suggests that the use of a combination of biomarkers, with good sensitivity and specificity, is necessary [30].

Proteomics for the study of global protein expression and analysis of post-translational modifications (PTMs) has been used for the discovery and validation of many novel potential biomarkers [31], [32], and can assist in clinical diagnosis by developing appropriate assays for assessing both normal biological processes and pathologic conditions.

Recently, the interest in applying Proteomics to clinical diagnosis and preventive medicine has considerably increased. More in details, ‘Clinical Proteomics should be defined as the application of proteomic analysis with the aim of solving a specific clinical problem within the context of a clinical study’ as reported on Proteomics –

Clinical Application in 2007 [33]. Therefore, clinical Proteomics is a proteomic application aimed at the identification of biological markers capable of characterizing a disease and predicting its onset, explaining also the mechanisms of its progression. Integrated with informatics tools for clinical diagnosis, it represents a promising area for personalised medicine with the latest high-resolution mass spectrometers capable of faster and more sensitive analysis in a high-throughput fashion. This –*omic* based approach is a powerful strategy for elucidation of disease mechanisms, detection, treatment and management of patients, and for the discovery and development of novel therapeutic modalities, concerning the kidney related disease. Indeed, the use of proteome and modified proteome analysis for the detection of clinically relevant proteins as earlier and prognostic markers may actually foster the translation of basic discoveries into clinical applications for the benefit of the patient.

2.2 Label-free Quantification

One of the challenges of the new proteomic discoveries consists in the determination of quantitative changes of proteomes between different patients, samples, conditions, etc., in order to highlight functional alterations in biological processes and molecular pathways [34]. MS-based quantitative proteomic strategies can be divided in label-based and label-free, according to the use or not of stable isotopes or fluorescent dyes to mark peptides [35]. While labeling methods give higher accuracy in the identification of proteins and in the assessment of quantification changes, they are expensive and require operative efforts. On the other hand, in the last years, label-free quantification (LFQ) methods are taking over because they are less expensive and need fewer efforts. In more detail, LFQ takes advantage from two different approaches to retrieve quantitative values for protein. The first one, known as spectral counting, correlates the abundance of a protein with the number of MS/MS spectra of all the peptides associated with that protein [36]. Then, the so-called extract ion chromatogram (XIC) is an alternative approach that measures the precursor ion current intensity at the MS1 level [37].

Since the complexity of the MS data, and regardless of the chosen approach, most of the LFQ experiments are carried out using bioinformatics tools and special software for data processing and elaboration [38], [39], like MaxQuant or Progenesis. By adopting any of the above-cited protocols in the experimental setting of "comparative proteomics", researchers have successively identified many differential proteins between healthy and diseased cells [40], tissues [41] and/or biological fluids [42], to be exploited as potential biomarkers.

2.3 Glyco-Proteomics

Once the proteins are identified and quantified, the next step to deeply investigate the proteome is the characterisation of their PTMs. PTM is an important feature of proteome, which frequently determine the specific biologic role or activity of a determinate protein. Among more than 100 different type of PTMs, the most frequent one is the glycosylation. Approximately, half of all proteins typically expressed in a cell undergo this modification. However, the role of PTMs, which extend and diversify protein function beyond gene transcription, still needs to be fully elucidated.

The protein glycosylation consists in the attachment of sugar moieties to proteins, providing greater proteomic diversity than other PTMs. It involves mainly surface proteins and it is critical for a wide range of biological processes, including cell attachment to extracellular matrix, protein-ligand interactions, monitoring activity of protein folding, molecular trafficking and clearance [43]. Various glycosidic linkages characterise this PTM, including N-, O- and C-linked glycosylation, glypiation (GPI anchor attachment) and phosphoglycosylation. Glycoproteins can be detected, purified and analysed by different strategies: glycan staining [44], purification by antibodies [45], lectins or magnetic particles [46], MS proteomic analysis [47], respectively. Therefore, Glycoproteomics is an important subdiscipline of Proteomics. Actually MS-based Glycoproteomics and Glycomics have become a vital part of biomedical research for discovery and characterisation

of disease biomarkers, due to rapid technological advances in affinity enrichment, chromatographic separation, quantitation, MS methodology and instrumentation, and bioinformatics. Quantitative comparative analyses of low-abundant, disease-related glycoforms in complex biological samples, are providing deeper understanding into the mechanism of onset and progression of disease. Moreover, these progresses are likely to provide improvements in diagnostic methods and therapeutic treatments.

3. Liquid Biopsy: the *new ambrosia* of clinical research

Increasing scientific advances in understanding circulating molecular and cellular species in diverse human bodily fluids have laid a solid foundation for the development in the research field of the routine analysis of 'liquid biopsy'.

The purpose of the biopsy, be it 'liquid' or 'solid', is to guide clinical intervention. Liquid biopsy is a non-invasive modality, alternative to tissue surgical biopsy, the solid one. Indeed, it results as a surrogate representative of tissue from which it originates, reflecting its correspondent physiological and pathophysiological status. A possible definition of liquid biopsy is every type of specimen able to provide this easily accessible window of the parental cell or tissue, such as: circulating tumour cells [48], cell-free DNA [49] and RNA [50], microparticles, extra-vesicles of diverse dimensions [51], proteins and metabolites, all isolated from many body fluids (e.g. blood, urine, saliva, faeces, ascites, pleural effusion, cerebral spinal fluid). Moreover, they can be readily obtained at different disease stages, allowing longitudinal studies for the active surveillance and/or for the monitoring of the response to drug treatment.

The high concordance between altered molecular profiles derived from liquid biopsy and their corresponding tissue biopsy is an established fact. Nowadays, the protein content of the liquid biopsy in the field of kidney related diseases has been targeted in an effort to define their possible diagnostic and prognostic course.

The renal liquid biopsy, in fact, offers the chance to investigate tumour growth,

proliferation, invasion and metastasis (for the ccRCC) and to clarify the disease etiopathogenesis and/or the pharmacoresistance (in case of INS), through minimally-invasive and simple tests. It is hoped that liquid biopsy will provide a novel inventory of disease biomarkers, acting as a potential complementary method with the tissue analysis guiding diagnosis, prognosis, treatment response and resistance in the era of personalised medicine. Biomarkers assayed in an easily available and low-cost sample, such as urine or its component (i.e. exosomes), would be highly valuable in the clinical practice.

3.1 Urine as liquid biopsy

Considering the above mentioned definition, the urine can be properly considerate a liquid biopsy. Indeed, it can be collected in large quantities and in non-invasive way, and as result of kidney filtration and activity, it represents a continuative monitoring snapshot of the physiology of the organism in general, and of the kidney in particular. Urine contains cellular elements, cell-free RNA species [52], proteins, metabolites [53] and urinary exosomes (see paragraph 3.2), derived from glomerular filtration of plasma, renal tubule excretion, and urogenital tract secretions that reflect, at a given time point, an individual's metabolic and pathophysiologic state.

Each one of the biochemical elements present in this bio-fluid is valuable as indicators of pathologies, rendering the urine a precious source of biomarkers for diseases related to kidney, primarily preferred respect to other bodily fluids, due to its lower complexity. It is noteworthy, for example, that Hippocrates linked health problems to diagnostic changes in the urine two millennia ago.

Nowadays, the systematic investigation of protein content of urine has become one of the most attractive topics in disease biomarker discovery. The urinary proteins were extensively characterised using different technological approaches; the most common and diffused is the MS-based analysis. This approach allows today the assessment of thousands of peptides and proteins in a urine sample in few hours

[54]. Moreover, several studies enlighten the ability of urinary proteomic MS-based approach to clarify and propose novel mechanistic insights for diverse kidney disorders [55]. Additionally, this strategy is able to provide an enormous wealth of information that can be exploited for clinical purpose, revealing fundamental for the discovery process of novel diagnostic and prognostic biomarkers in the area of the kidney related diseases.

3.2 Urinary Extracellular Vesicles as liquid biopsy

Unfortunately, in diverse cases, the information provided by urine can be 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 urinary proteins such as albumin or uromodulin.

In this context, the proteomic study of the urinary extracellular vesicles (UEv) represents a valid alternative to reveal and discover these hidden molecular landscape [56]. UE are nanometer-sized vesicles (> 150 nm), that can originate from endothelial cells, podocytes or tubular epithelial cells. Their molecular composition depends upon the type, and even status, of the producer cell. As such, they provide an easily accessible window to monitor the condition of the respective parental cell. Therefore, the use of UE, that contain only 3% of total urine proteins (> 3000 species), allows to reduce the complexity of the urine proteome, depleting it from the most abundant urinary proteins, such as albumin [57]. For all these reasons, they can be considerate, as the urine, a specific renal liquid biopsy, able to provide diagnostic and prognostic information about kidney impairment and disease [58].

References

[1] F. Persson and P. Rossing, “Urinary Proteomics and Precision Medicine for Chronic Kidney Disease: Current Status and Future Perspectives.,” *Proteomics. Clin. Appl.*, vol. 13, no. 2, p. e1800176, Mar. 2019.

[2] T. Lancet, “Research in kidney disease: An acute and chronic history,” *Lancet*, vol. 385, no. 9981, p. 1918, 2015.

[3] K. Cance, “The silent disease,” *Nature*, 2012.

[4] Different collaborators, “Comprehensive Molecular Characterization of,” vol. 499, no. 7456, pp. 43–49, 2014.

[5] B. I. Rini *et al.*, “Active surveillance in metastatic renal-cell carcinoma: a prospective, phase 2 trial.,” *Lancet. Oncol.*, vol. 17, no. 9, pp. 1317–1324, Sep. 2016.

[6] B. Ljungberg *et al.*, “European Association of Urology Guidelines on Renal Cell Carcinoma: The 2019 Update,” *Eur. Urol.*, vol. 75, no. 5, pp. 799–810, 2019.

[7] U. Capitanio *et al.*, “A critical assessment of the prognostic value of clear cell, papillary and chromophobe histological subtypes in renal cell carcinoma: A population-based study,” *BJU Int.*, vol. 103, no. 11, pp. 1496–1500, 2009.

[8] K. A. Keegan, C. W. Schupp, K. Chamie, N. J. Hellenthal, C. P. Evans, and T. M. Koppie, “Histopathology of surgically treated renal cell carcinoma: survival differences by subtype and stage.,” *J. Urol.*, vol. 188, no. 2, pp. 391–397, Aug. 2012.

[9] B. Kappel and S. Olsen, “Cortical interstitial tissue and sclerosed glomeruli in the normal human kidney, related to age and sex - A quantitative study,” *Virchows Arch. A Pathol. Anat. Histol.*, vol. 387, no. 3, pp. 271–277, 1980.

[10] G. Sconocchia *et al.*, "Defective infiltration of natural killer cells in MICA/B-positive renal cell carcinoma involves β2-integrin-mediated interaction," *Neoplasia*, vol. 11, no. 7, pp. 662–671, 2009.

[11] L. Morra *et al.*, "Relevance of periostin splice variants in renal cell carcinoma," *Am. J. Pathol.*, vol. 179, no. 3, pp. 1513–1521, 2011.

[12] A. H. D. Van Hout *et al.*, "Loss of heterozygosity at the short arm of chromosome 3 in renal- cell cancer correlates with the cytological tumour type," *Int. J. Cancer*, vol. 53, no. 3, pp. 353–357, 1993.

[13] M. Thompson, "Polybromo-1: the chromatin targeting subunit of the PBAF complex.," *Biochimie*, vol. 91, no. 3, pp. 309–319, Mar. 2009.

[14] S. Mikami, R. Mizuno, T. Kosaka, H. Saya, M. Oya, and Y. Okada, "Expression of TNF-alpha and CD44 is implicated in poor prognosis, cancer cell invasion, metastasis and resistance to the sunitinib treatment in clear cell renal cell carcinomas.," *Int. J. cancer*, vol. 136, no. 7, pp. 1504–1514, Apr. 2015.

[15] E. Cuzzoni *et al.*, "MIF plasma level as a possible tool to predict steroid responsiveness in children with idiopathic nephrotic syndrome," *Eur. J. Clin. Pharmacol.*, 2019.

[16] A. A. Eddy and J. M. Symons, "Nephrotic syndrome in childhood.," *Lancet (London, England)*, vol. 362, no. 9384, pp. 629–639, Aug. 2003.

[17] T. H. J. Goodship *et al.*, "Atypical hemolytic uremic syndrome and C3 glomerulopathy: conclusions from a 'Kidney Disease: Improving Global Outcomes' (KDIGO) Controversies Conference," *Kidney Int.*, vol. 91, no. 3, pp. 539–551, 2017.

[18] P. J. Barnes, "Mechanisms and resistance in glucocorticoid control of inflammation," *J. Steroid Biochem. Mol. Biol.*, vol. 120, no. 2–3, pp. 76–85, 2010.

[19] M. L. Downie, C. Gallibois, R. S. Parekh, and D. G. Noone, “Nephrotic syndrome in infants and children: Pathophysiology and management,” *Paediatr. Int. Child Health*, vol. 37, no. 4, pp. 248–258, 2017.

[20] E. Cuzzoni *et al.*, “Glucocorticoid pharmacogenetics in pediatric idiopathic nephrotic syndrome,” *Pharmacogenomics*, vol. 16, no. 14, pp. 1631–1648, 2015.

[21] E. S. Lander *et al.*, “Initial sequencing and analysis of the human genome.,” *Nature*, vol. 409, no. 6822, pp. 860–921, Feb. 2001.

[22] J. C. Venter *et al.*, “The sequence of the human genome.,” *Science*, vol. 291, no. 5507, pp. 1304–1351, Feb. 2001.

[23] K. Wang, C. Huang, and E. Nice, “Recent advances in proteomics : towards the human proteome,” pp. 848–857, 2014.

[24] K. Márton, D. K. Borbála, D. László, V. Károly, A. István, and L. Krisztina, “Application of mass spectrometry in proteomics,” *Acta Pharm. Hung.*, vol. 86, no. 2, pp. 61–67, 2016.

[25] R. Matthiesen, *Mass Spectrometry Data Analysis in Proteomics. Methods in Molecular Biology.*, vol. 1007. 2013.

[26] Z. Zhang, S. Wu, D. L. Stenoien, and L. Pasa-Tolic, “High-throughput proteomics.,” *Annu. Rev. Anal. Chem. (Palo Alto. Calif).*, vol. 7, pp. 427–454, 2014.

[27] A. J. Atkinson *et al.*, “Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework,” *Clin. Pharmacol. Ther.*, vol. 69, no. 3, pp. 89–95, 2001.

[28] D. A. Megger *et al.*, “Proteomic differences between hepatocellular carcinoma and nontumorous liver tissue investigated by a combined gel-based and label-free quantitative proteomics study,” *Mol. Cell. Proteomics*,

vol. 12, no. 7, pp. 2006–2020, 2013.

[29] E. Boja *et al.*, “Evolution of clinical proteomics and its role in medicine,” *J. Proteome Res.*, vol. 10, no. 1, pp. 66–84, 2011.

[30] K. A. Landers *et al.*, “Use of multiple biomarkers for a molecular diagnosis of prostate cancer,” *Int. J. Cancer*, vol. 114, no. 6, pp. 950–956, 2005.

[31] C. E. Parker and C. H. Borchers, “Mass spectrometry based biomarker discovery, verification, and validation - Quality assurance and control of protein biomarker assays,” *Mol. Oncol.*, vol. 8, no. 4, pp. 840–858, 2014.

[32] R. Chaerkady and A. Pandey, “Applications of proteomics to lab diagnosis.,” *Annu. Rev. Pathol.*, vol. 3, pp. 485–498, 2008.

[33] H. Mischak *et al.*, “Clinical proteomics: A need to define the field and to begin to set adequate standards,” *Proteomics - Clin. Appl.*, vol. 1, no. 2, pp. 148–156, 2007.

[34] H. Li, J. Han, J. Pan, T. Liu, C. E. Parker, and C. H. Borchers, “Current trends in quantitative proteomics – an update,” *J. Mass Spectrom.*, vol. 52, no. 5, pp. 319–341, 2017.

[35] K. Patel, M. Singh, and H. Gowda, “Chapter 12 Bioinformatics Methods to Deduce Biological,” vol. 1549, no. October 2017, pp. 1–3, 2016.

[36] J. A. Bubis, L. I. Levitsky, M. V. Ivanov, I. A. Tarasova, and M. V. Gorshkov, “Comparative evaluation of label-free quantification methods for shotgun proteomics,” *Rapid Commun. Mass Spectrom.*, vol. 31, no. 7, pp. 606–612, 2017.

[37] X. Wang, S. Shen, S. S. Rasam, and J. Qu, “MS1 ion current-based quantitative proteomics: A promising solution for reliable analysis of large biological cohorts,” *Mass Spectrom. Rev.*, pp. 1–22, 2019.

[38] L. A. Beer, P. Liu, B. Ky, K. T. Barnhart, and D. W. Speicher, “Standard

Operating Procedures for Plasma Collection in Clinical Research," pp. 339–352, 2017.

[39] D. Chokchaichamnankit *et al.*, "Urinary biomarkers for the diagnosis of cervical cancer by quantitative label‐free mass spectrometry analysis," *Oncol. Lett.*, vol. 17, no. 6, pp. 5453–5468, 2019.

[40] W. Peng, Y. Zhang, R. Zhu, and Y. Mechref, "Comparative membrane proteomics analyses of breast cancer cell lines to understand the molecular mechanism of breast cancer brain metastasis.," *Electrophoresis*, vol. 38, no. 17, pp. 2124–2134, Sep. 2017.

[41] O. Azimzadeh, M. J. Atkinson, and S. Tapio, "Qualitative and quantitative proteomic analysis of formalin-fixed paraffin-embedded (FFPE) tissue.," *Methods Mol. Biol.*, vol. 1295, pp. 109–115, 2015.

[42] Y. Sun *et al.*, "Comparative Proteomic Analysis of Exosomes and Microvesicles in Human Saliva for Lung Cancer.," *J. Proteome Res.*, vol. 17, no. 3, pp. 1101–1107, Mar. 2018.

[43] K. Ohtsubo and J. D. Marth, "Glycosylation in Cellular Mechanisms of Health and Disease," *Cell*, vol. 126, no. 5, pp. 855–867, 2006.

[44] T. H. Steinberg, *Chapter 31 Protein Gel Staining Methods. An Introduction and Overview*, 1st ed., vol. 463, no. C. Elsevier Inc., 2009.

[45] S. M. Aavula, G. Abhinay, S. V. Nimmagadda, and K. Maithal, "A novel in vitro ELISA for estimation of glycoprotein content in human rabies vaccines," *J. Immunoass. Immunochem.*, vol. 38, no. 4, pp. 400–410, 2017.

[46] W. Zhao *et al.*, "The enrichment and characterization of ginger-derived glycoprotein using magnetic particles," *Food Chem.*, vol. 244, no. September 2017, pp. 164–168, 2018.

[47] J. E. Schiel, "Glycoprotein analysis using mass spectrometry: Unraveling the

layers of complexity," *Anal. Bioanal. Chem.*, vol. 404, no. 4, pp. 1141–1149, 2012.

[48] H. I. Scher *et al.*, "Circulating tumor cell biomarker panel as an individual-level surrogate for survival in metastatic castration-resistant prostate cancer," *J. Clin. Oncol.*, vol. 33, no. 12, pp. 1348–1355, 2015.

[49] R. Cortese *et al.*, "Epigenetic markers of prostate cancer in plasma circulating DNA," *Hum. Mol. Genet.*, vol. 21, no. 16, pp. 3619–3631, 2012.

[50] A. Di Meo, J. Bartlett, Y. Cheng, M. D. Pasic, and G. M. Yousef, "Liquid biopsy: A step forward towards precision medicine in urologic malignancies," *Molecular Cancer*, vol. 16, no. 1. 2017.

[51] V. R. Minciacchi, A. Zijlstra, M. A. Rubin, and D. Di Vizio, "Extracellular vesicles for liquid biopsy in prostate cancer: where are we and where are we headed?," *Prostate Cancer Prostatic Dis.*, vol. 20, no. 3, pp. 251–258, Sep. 2017.

[52] M. Martínez-Fernández, J. M. Paramio, and M. Dueñas, "RNA Detection in Urine: From RNA Extraction to Good Normalizer Molecules," *J. Mol. Diagnostics*, vol. 18, no. 1, pp. 15–22, 2016.

[53] M. M. Khamis, D. J. Adamko, and A. El-Aneed, "Mass spectrometric based approaches in urine metabolomics and biomarker discovery.," *Mass Spectrom. Rev.*, vol. 36, no. 2, pp. 115–134, Mar. 2017.

[54] D. Fliser *et al.*, "Advances in urinary proteome analysis and biomarker discovery," *J. Am. Soc. Nephrol.*, vol. 18, no. 4, pp. 1057–1071, 2007.

[55] F. Raimondo, D. Cerra, F. Magni, and M. Pitto, "Urinary proteomics for the study of genetic kidney diseases," *Expert Rev. Proteomics*, vol. 13, no. 3, pp. 309–324, 2016.

[56] P.-G. Moon, S. You, J.-E. Lee, D. Hwang, and M.-C. Baek, "Urinary

exosomes and proteomics.," *Mass Spectrom. Rev.*, vol. 30, no. 6, pp. 1185–1202, 2011.

[57] P.-P. Chen, Y. Qin, and X.-M. Li, "Advances in the Urinary Exosomes in Renal Diseases.," *Zhongguo Yi Xue Ke Xue Yuan Xue Bao.*, vol. 38, no. 4, pp. 464–469, Aug. 2016.

[58] M. Krause, A. Samoylenko, and S. J. Vainio, "Exosomes as renal inductive signals in health and disease, and their application as diagnostic markers and therapeutic agents.," *Front. cell Dev. Biol.*, vol. 3, p. 65, 2015.

SCOPE OF THE THESIS

The body of work enclosed in this thesis proposes the in depth proteomic analysis of urine and urinary extracellular vesicles as liquid biopsy in order to increase our understanding of the molecular aspects of two renal pathologies, the clear cell Renal Cell Carcinoma and Idiopathic Nephrotic Syndrome, that could be useful for their diagnosis and prognosis.

Chapter two: The investigation of the ccRCC urine samples in order to verify the hypothesis that urine could *proteomically* reflect the two morphological tissue-related features stage and grade, usually used by the pathologist to define the tumour status.

Chapter three: Glycoproteomics analysis of urine samples of patients at different ccRCC stage in order to pinpoint a characteristic specific disease glycoprotein signature of tumour progression.

Chapter four: The investigation of the impact of different levels of haematuria on the proteome of UEv, improving the UEv preparation protocol in order to reduce the impact of the blood related interferers and to obtain properly sample for proteomic analysis.

Chapter five: The analysis of UEv isolated from the urine of children affected by INS in order to verify the feasibility to use these nanovesicles as a tool able to clarify the molecular mechanism underling the different response to drug treatment and the onset of pharmacoresistance to corticosteroids.

CHAPTER 2

Does the urinary proteome reflect ccRCC stage and grade progression?

Lucia Santorelli[1]#, Martina Stella[2]#, Clizia Chinello[3], Isabella Piga[3], Giulia Capitoli[4], Andrew Smith[3], Vanna Denti[3], Francesca Clerici[3], Angelica Grasso[5], Marco Grasso[5], Giorgio Bovo[6], Fulvio Magni[3]

[1] School of Medicine and Surgery, University of Milano - Bicocca, Clinical Proteomics and Metabolomics Unit, Vedano al Lambro, Italy; l.santorelli@campus.unimib.it

[2] School of Medicine and Surgery, University of Milano - Bicocca, Clinical Proteomics and Metabolomics Unit, Vedano al Lambro, Italy; Martina.Stella@bruker.com

#equally contributing authors

[3] School of Medicine and Surgery, University of Milano - Bicocca, Clinical Proteomics and Metabolomics Unit, Vedano al Lambro, Italy.

[4] School of Medicine and Surgery, University of Milano – Bicocca, Centre of Biostatistics for Clinical Epidemiology, Vedano al Lambro, Italy.

[5] Urology Unit, S.Gerardo Hospital, Monza, Italy.

[6] Pathology Unit, Vimercate Hospital, Vimercate, Italy.

Manuscript Submitted

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 progression of clear cell Renal Cell Carcinoma (ccRCC) lesions. 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 could be detected in urine. Several proteins presented an altered urinary expression according to stage and grade.
More in detail, when the two features were considered separately, we enlighted 6 proteins specifically related to the stage; moreover, when they were considered simultaneously, we highlighted a set of proteins able to reflect the diverse combinations of early and advanced stages and grades. These data showed that dinamic information deriving from the urinary proteome could allow overcoming the static stage and grade classification.

KEYWORDS Clear cell renal cell carcinoma; nLC-ESI-MS/MS; tumour grade and stage; proteomics; urine.

1. INTRODUCTION

Renal cell carcinoma (RCC) comprises an heterogeneous group of tumours, in which the most frequent (70-80%) and aggressive morpho-type is clear cell RCC (ccRCC) [1]. ccRCC is a malignancy of particular interest for clinical proteomic investigations because of its associated inter- and intra- tumour heterogeneity [2], high recurrence and progression rates, and lack of effective non-invasive diagnostic and prognostic indicators [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].

Molecular mechanisms underpinning tumour development and progression are still unclear and studies 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 being easily, and noninvasively obtainable. 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 kidney diseases with a special focus on tumours [10–12].

However, excluding some exceptions [13–15], most of the biomarker studies in bodily fluids focused on the differences between patients with ccRCC and healthy controls [16]. Therefore, in this work we investigated the proteome of ccRCC patients with different lesion severities, 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 [17], 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. 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.

2.1. Clinical data and study design

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).

Table 1. Clinicoplathological characteristics of patients enrolled in the study, grouped according to grade and stage.

Group	Number of patients	Gender [male - female]	Age at diagnosis [mean]	Greatest tumour dimension [cm]
Stage 1, pT1	19	11 - 7	68	3.5
Stage 3-4, pT3	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

[1] Low grade (LG) = G1 and G2; [2] High grade (HG) = G3 and G4

Patients were classified accordingly with 2009 TNM classification [5]. In particular, a specific focus was given both to grade values (low grade (LG) = G1 and G2; high Grade (HG) = G3 and G4) and stage (stage 1 and stage 3-4), in order to evaluate proteomic signatures of different lesions, as already demonstrated on tissue, using liquid biopsies.

2.2 Qualitative evaluation and network analysis: relevant pathways of ccRCC

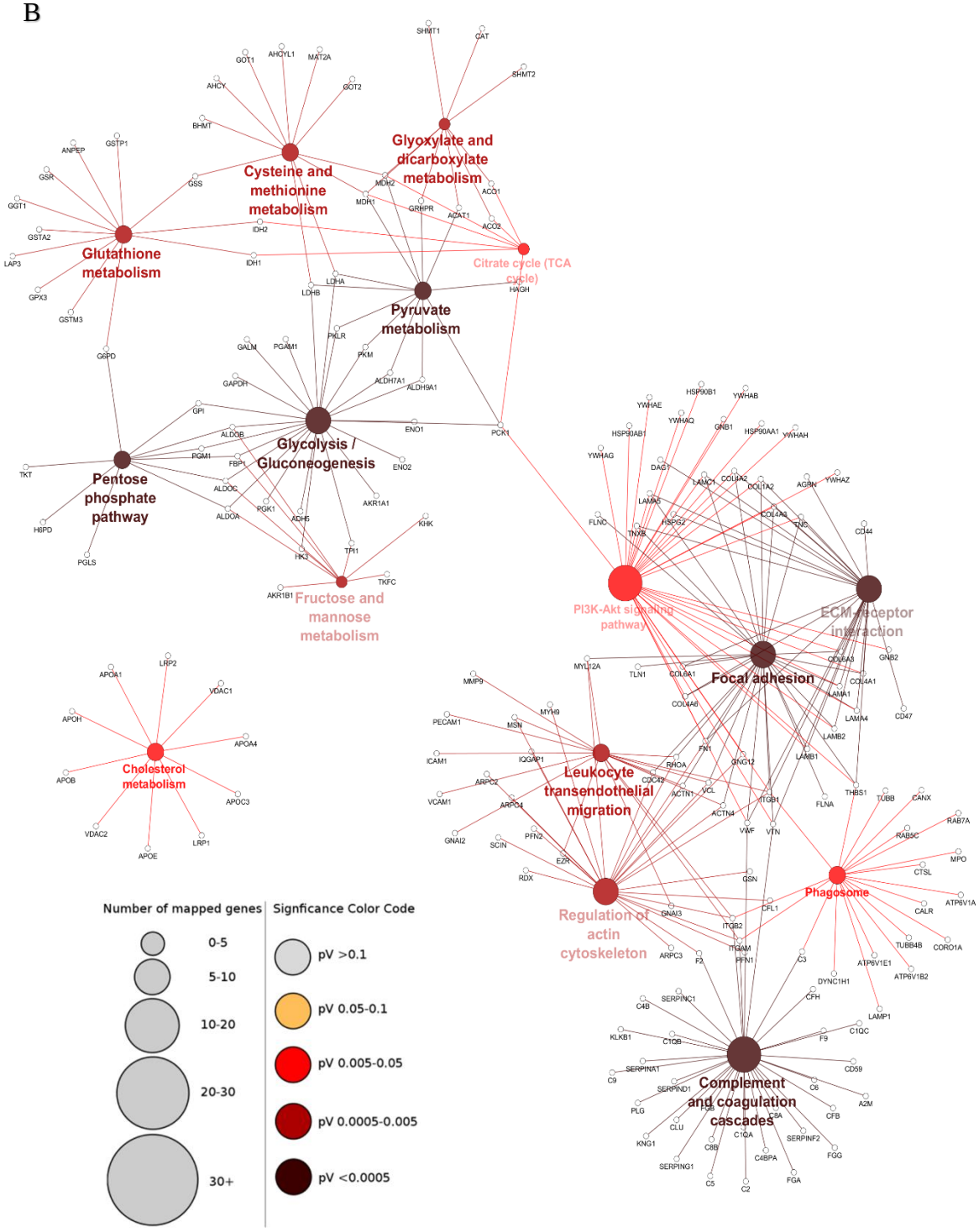
Following the FASP protocol, samples were analysed by nUHPLC-MS/MS and a total of 21640 peptides sequences, corresponding to 1609 proteins were identified with FDR Peptide-Spectrum Matches of 1% and at least one unique peptide (Supplementary Table 1). Those proteins identified in urine were compared with those identified on tissue (previously published data [17]) 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 [18–20].

Figure 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).

A

GO-Term	% Associated Genes	N° Genes	Identified Associated Genes
Complement & coagulation cascades	43.04	34	A2M, C1QA, C1QB, C1QC, C2, C3, C4B, C4BPA, C5, C6, C8A, C8B, C9, CD59, CFB, CFH, CLU, F2, FGA, FGB, FGG, ITGAM, ITGB2, KLKB1, KNG1, PLG, SERPINA1, SERPIN1, SERPINF2, SERPING1, VTN, VWFJ
Glycolysis & Gluconeogenesis	33.82	23	ADH5, AKR1A1, ALDH7A1, ALDH9A1, ALDOA, ALDOB, ALDOC, ENO1, ENO2, FBP1, GALM, GAPDH, GP1, HK3, LDHA, LDHB, PCK1, PGAM1, PGK1, PGM1, PLKLR, PKM, TPI1
Penthose phosphate pathway	33.33	10	ALDOA, ALDOB, ALDOC, FBP1, G6PD, GOI, H6PD, PLGLS, PGM1, TKTJ
Pyruvate metabolism	30.77	12	ACTN1, ACTN4, ARPC2, ARPC3, ARPC4, CDC42, CFL1, EZR, F2, FN1, GNG12, GSN, IQGAP1, ITGAM, ITGB1, ITGB2, MSN, MYH9, MYL12A, PFN1, PFN2, RDX, RHOA, SCIN, VCLJ
ECM-receptor interaction	30.49	25	AGRN, CD44, CD47, COL4A1, COL4A2, COL4A3, COL4A6, COL6A1, COL6A3, DAG1, FNI, HSPG2, ITGB1, LAMA1, LAMA4, LAMA5, LAMB1, LAMB2, LAMC1, THBS1, TNC, TNXB, VCL, VTN,VWFJ
Glyoxylate and dicarboxylate metabolism	30	9	ACAT1, ACO1, ACO2, CAT, GRHPR, MDH1, MDH2, SHMT1, SHMT2
Fructose and mannose metabolism	27.27	9	AKR1B1, ALDOA, ALDOB, ALDOC, FBP1, HK3, KHK, TKFC, TPI1
Cysteine and methionine matabolism	24.44	11	AHCY, AHCYL1, BHMT, GOT1, GOT2, GSS, LDHA, LDHB, MAT2A, MDH1, MDH2
TCA Cycle	23.33	7	ACO1, AC02, IDH1, IDH2,MDH1, MDH2, PCK1
Glutathione metabolism	21.43	12	ANPEP, G6PD, GGT1, GPX3, GSR, GSS, GSTA2, GSTM3, GSTP1, IDH1, IDH2, LAP3
Cholesterol metabolism	20	10	APOA1, APOA4, APOB, APOC3, APOE, APOH, LRP1, LPRP2, VDAC1, VDAC2
Leukocyte transendothelial migration	15.18	17	ACTN1, ACTN4, CDC42, EZR, GNAI3, ICAM1, ITGAM, ITGB1, ITGB2, MMP2, MSN MYL12A, PECAM1, RHOA, VCAM1, VCL
Focal adhesion	14.57	29	ACTIN1, 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, VLC, VTN, VWF
Phagosome	12.5	19	ATP6V1A, ATP6V1B2, ATP6VIE1, C3, CALR, CANX, CORO1A, CTSL, DYNC1H1, ITGAM, ITGB1, ITGB2, LAMP1, MPO, RAB5C, RAB7A, THBS1, TUBB, TUBB4B
P13K-Akt signaling pathway	9.32	33	COL1A2, COL4A1, COL4A2, COL4A3, COL4A6, COL6A1, COL6A3, FN1, GNB1, GNB2, GNG12, HSP90AA1, HSP90AB1, HSP90B1, ITGB1, LAMA5, LAMB2, LAMC1, PCK1, THBS1, TNC, TNXB, VTN, VWF, YWHAG, YWHAH, YWHAQ, YWHAZ

B


Glutathione metabolism
Cysteine and methionine metabolism
Glyoxylate and dicarboxylate metabolism
Citrate cycle (TCA cycle)
Pyruvate metabolism
Glycolysis / Gluconeogenesis
Pentose phosphate pathway
Fructose and mannose metabolism
Cholesterol metabolism
PI3K-Akt signaling pathway
ECM-receptor interaction
Focal adhesion
Leukocyte transendothelial migration
Regulation of actin cytoskeleton
Phagosome
Complement and coagulation cascades
Number of mapped genes
0-5
5-10
10-20
20-30
30+
Signficance Color Code
pV >0.1
pV 0.05-0.1
pV 0.005-0.05
pV 0.0005-0.005
pV <0.0005

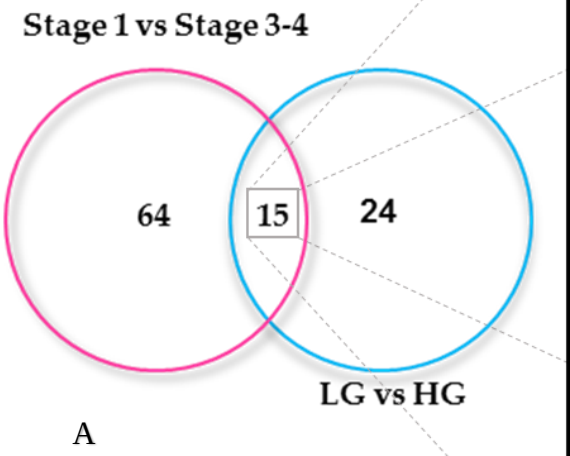
2.3 Urinary secreted proteins varied according to different grades and stages

We further investigated the correlation between urinary protein levels and tumour features (Stage and Grade). In particular, we focused the investigation on the urinary proteome alterations, possible reflecting morphological changes of the lesions, as already observed in tissue [17]. To enhance confidence in the protein quantification, the abundances of proteins detected in at least half of the 44 samples and quantified with at least two unique peptides were exported from PEAKS Studio. Statistical analysis was performed with a Mann-Whitney test (p-value ≤ 0.05 and 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 splitted into two groups, HG (G3-G4) and LG (G1-G2) tumours and 39 proteins were identified as differentially expressed. Among them, 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. Additionally, 79 proteins with a statistically significant variation in their abundances were detected: 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 also compared (Figure 2A) and the 15 proteins commonly altered 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 hypothesised that the co-presence of the two features (Stage and Grade) could have determined the intersection highlighted in Figure 2.



B

Stage and Grade		
Proteins	Up-regulated in:	
PGD	HG	Stage 3-4
LAMC1	HG	Stage 3-4
LPA	HG	Stage 3-4
GDF15	HG	Stage 3-4
IGKV3-20	HG	Stage 3-4
GGCT	HG	Stage 3-4
MMRN2	HG	Stage 3-4
ACVR1B	HG	Stage 3-4
IGLV1-40	HG	Stage 3-4
AMY1A-B-C	LG	Stage 1
GCHFR	HG	Stage 3-4
THBS4	LG	Stage 1
IGF2R	HG	Stage 3-4
S100P	LG	Stage 1
UMOD	LG	Stage 1

Figure 2. A: Venn diagram of the urinary proteins altered according to the stage (Stage1 vs Stage3-4, pink) and grade (G1-2 versus G3-4, blue).
B: List of 15 proteins commonly altered in grade and stage comparisons and their urinary abundances.

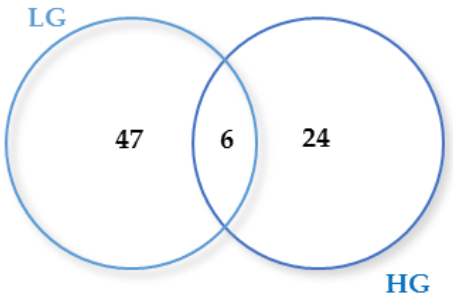
In order to better understand which of these clinical features is responsible for the urinary protein alterations, we proceeded by sub-classifying our cohort of patients considering each feature a time. At first, we compared patients with low grade and stage 1 lesions with those with low grade and stage 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 six proteins upregulated in stage 3 compared to stage 1, when excluding the grade as a confounding factor (Figure 3 A-B).

LG sub-cohort (B *vs* A)		LG & HG	HG sub-cohort (D *vs* C)
Proteins: 47		Common Proteins: 6	Proteins: 24
Stage 3-4 vs Stage 1		*Stage 3-4 vs Stage 1*	*Stage 3-4 vs Stage 1*
● 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

A

● Up-regulation

● Down-regulation



B

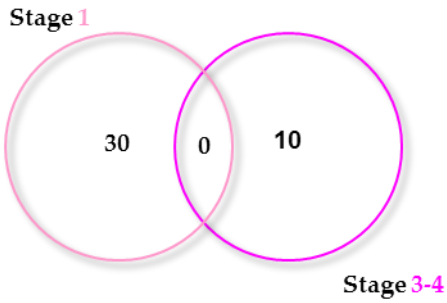
Figure 3. List (A) and Venn diagram (B) of urinary proteins altered according to HG and LG, considering only the stage as the variable feature.

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) versus high grade (group C) at stage 1 and low grade (group B) versus high grade (group D) at stage 3-4 lesions (Figure 4 A-B). This comparison showed no proteins correlating specifically with the grade, independently from the stage.

A

stage 1 sub-cohort (C *vs* A)		stage 1 & stage 3-4	stage 3-4 sub-cohort (D *vs* B)
Proteins: 30		Common Proteins: 0	Proteins: 10
HG vs LG		*HG vs LG*	*HG vs LG*
● 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		

● Up-regulation
● Down-regulation



B

Figure 4. List (A) and Venn diagram (B) of urinary proteins altered according to stage 1 and stage 3-4, considering only the grade as the variable feature.

3. DISCUSSION

Proteomics studies of ccRCC using different mass spectrometry technologies have been applied to a variety sample types, including tumour cell lines [21] and different types of human samples, such as serum [16], plasma, tissue and urine [22]. Using a combined mass spectrometric approach (MALDI MSI and nLC-ESI-MS/MS), we have previously described the proteomic profile of different ccRCC tumour grades using well defined, histology-guided, regions of tissue. Through this strategy, we highlighted several m/z signals with an intensity trend associated to the different tumour grades and we identified a list of proteins, whose expression resulted altered among the diverse tumour grade groups [17]. Starting from these evidences, we hypothesised the possibility that the urine, a readily accessible waste bodily fluid, could proteomically reflect the morphological features (stage and grade) used by the pathologist to define the tumour, in order to confirm and predict the kidney impairment and/or the cancer progression.

For this purpose, we analysed the urine of ccRCC patients, focusing the sample stratification on both values, grade (low and high) and stage (stage 1 and stage 3-4), information provided by the pathologist after the partial or total surgical removal of the tumour (Table 1). Firstly, we identified more than 1600 urinary proteins, 500 of which emerged to be shared with those identified previously on ccRCC tissues [17], confirming the hypothesis that urine could assess, also in this manner, the kidney function.

Thus, we performed a pathway analysis of the proteins, which identification is shared in both the specimens (ccRCC urine and tissue), in order to verify the possibility to intercept the molecular processes involved in the tumour etiopathogenesis and progression, also in the urine. Network connections allowed the recognition of various pathways, related to the cancer features.

Primarily, our GO term enrichment analysis intercepted as the mainly represented processes like glycolysis/gluconeogenesis, pentose phosphate pathway and pyruvate metabolism, with 52 identified proteins found associated with them (figure 1 A-B). It is well known that the glucose oxidation via glycolytic and pentose phosphate pathways is the favourite mechanism to degrade glucose, even in the presence of adequate oxygen availability for mitochondrial oxidation. This mechanism of energy production by the cell is note as Warburg's effect and it is considered the principal hallmark of RCC carcinogenesis [23].

We also detected urinary proteins coded by hub genes involved in the cellular reactions of: carbohydrates (fructose and mannose metabolism, 9 proteins), lipids (cholesterol metabolism, 10 proteins), and amino acids (cysteine and methionine metabolism, 11 proteins), further suggesting a possible dysregulation of metabolic pathways involved in molecular processes fundamental for the generation of energy (figure 1A). Taken together, our findings seem to support the widely accepted consideration of ccRCC as a metabolic disease, in which the genes typically mutated have dysregulation effects on metabolic pathways involved in oxygen or nutrient sensing [24].

The results of functional annotation analysis also indicated the presence in the urine of proteins involved in the immune response and coagulation cascade (figure 1A-B), consistent with the latest researches in the field [25, 26]. Information provided by urine and tissue had a considerable overlap; in fact, proteins detected in both ccRCC urine and tissue samples are mainly involved in metabolic pathways and in the activation of immune response. For example, in their systematic bioinformatics study, Tian et al. analysed a microarray dataset containing ccRCC and corresponding normal samples [27]. They identified especially enriched key genes,

coding for components of complement, such as C1QA, C1QB and C1QC, the same found in our GO analysis conducted on the RCC urine samples. According to their conclusions, we could consider immune response and coagulation cascade classes as indispensable in ccRCC progression.

To further investigate the correlation between the urine proteome and the tumour classification, we performed a quantitative label-free analysis. Firstly, the urine samples were classified according to grade, detecting 39 proteins (Supplementary Table 2), with different statistically significant abundance levels between the HG group and the LG one. Secondly, the quantitative analysis was conducted on the same samples, reclassifying them according to stage; this analysis permitted to pinpoint 79 dysregulated proteins between stage 1 and stage 3-4 (Supplementary Table 3). The comparison of the lists of these altered proteins coming from the two different histology-guided groups highlighted the presence of 15 proteins commonly varied in stage and grade (figure 2 A-B). The occurrence of these proteins as dysregulated in both lists suggested the possibility that their expression could be equally influenced by both histopathological features and consequently stage and grade could act as reciprocally confounding factors.

In order to in depth examine this aspect, we decided to reclassify the whole cohort of patients, considering a single histopathological feature at a time. In this way, we performed other 4 statistical analyses: patients with the same grade were compared to each other by varying only the stage feature, and patients with the same stage were compared to each other considering as variable only the grade feature (figure 3-4, Supplementary Table 4). It is noteworthy that 9 of the 15 proteins, initially found commonly altered, reappeared in those further comparisons. In particular, the abundance levels of GDF15 and IGLV1-40 resulted significantly varied in 4 of the total performed comparisons (Figure 2-3, Supplementary Table 4), while AMY A-B-C, GGCT, LAMC1, MMRN2, S100P and THBS4 resulted significantly varied in 3 of total comparisons (Figure 2, Supplementary Table 4).

Therefore, this panel of urinary proteins could be a good indicator of the ccRCC progression, enable to reflect simultaneously, but in different way, both the morphological features. In particular, GDF15 and IGLV1-40 resulted up-regulated in all statistical analysis where the features stage (early and advanced) were compared, even when adjusted for different grade. This finding suggests the hypothesis that their expression primarily reflected the stage, but the feature grade still remains relevant. In fact, these proteins are statistically significant even in the general comparison HG vs LG (Supplementary Table 4).
The same consideration regards the proteins AMY A-B-C, GGCT, LAMC1, MMRN2 and S100P. They resulted significant in 3 comparisons: 1 that involved the grade and 2 that involved the stage. AMY A-B-C and S100P are consistently down-regulated, while LAMC1, GGCT and MMRN2 are consistently up-regulated. Thus, also the abundance levels of this group of proteins seemed to reflect primarily the stage, as well as the previous ones.

To the best of our knowledge, there are no study investigating proteins changes in urine associated with ccRCC lesion at different stage and grade simultaneously. However, all the proteins that we detected as varied in both the general comparisons (stage 3-4 vs stage 1 and HG vs LG), resulted to be altered in different types of cancers when compared to control (Table 2).
Indeed, a remarkable interest is represented by laminin γ1 chain (LAMC1), a core structural protein present in the basement membrane of several organs, including the kidney. Immunohistochemical analysis of kidney specimens from patients with Chronic Kidney Disease (CKD) underlined an increased LAMC1 presence in the glomerular basement membrane, that could reflect an acceleration of its remodelling. In addition, a fragment of LAMC1, the neo-epitope peptide LG1M, has been detected in the in serum and urine from patients with CKD at high risk of progression, assessing its association with adverse outcome, including end stage renal disease and early mortality [28]. Furthermore, it is known that the microRNA-429 is down-regulated in ccRCC tissues. In a miRNAseq study is reported that the

LAMC1 gene is one of the identified target genes of miRNA-429; in particular, they observed a connection between the down-regulation of miRNA-429 and the up-regulation of the LAMC1 gene in ccRCC tissues compared to matched normal adjacent tissues [29].

All these observations imply that LACM1 is important for tumor cell function and cancer progression, consistent with our finding. In fact, in urine of patients affected by ccRCC we detected LAMC1 up-regulated in stage 3-4 (fold change=2.14, p-value=0.003) and HG group (fold change=1.62, p-value=0.017). Moreover, when the patients cohort was reclassified, considering only one histopathological feature, LAMC1 was mainly up-regulated in the urine of patients with stage 3-4 lesions, in the LG group (fold change=2.77, p-value=0.03). Therefore, this urinary protein could be a potential biomarker for clinical diagnosis and prognosis of ccRCC.

Another important player in cancer cell proliferation has been demonstrated to be the enzyme GGCT, Glutamylcyclotransferase, involved in glutathione metabolism [30] and found to be over-expressed in a variety of cancers. Concerning this protein, Gromov et al. primarily showed its high expression in different cancer types: uterine cervix (58%), lung (38%), and colon (72%) carcinoma [31]. In addition, they also detected GGCT in the extracellular fluid of mammary glands, proposing it as a marker for breast cancer. Moreover, our findings are in good agreement with that described by Li et al. that showed the overexpression of GGCT in high-grade serous ovarian cancer [32]. In fact, we found different abundance levels of GGCT in urine of ccRCC patients stratified as stage 3-4 (fold change=1.90, p-value=0.005) and as HG (fold change=1.54, p-value=0.03). Moreover, it resulted to be more up-regulated in the stage 3-4, when we considered only the LG group of patients (fold-change=1.88, p-value=0.04).

Nowadays, many researches are exploiting the possibility to use GGCT as a promising target of cancer treatments, focusing their efforts towards the design and synthesis of different kind of inhibitors that have already shown an antitumoural effect [33]. In this contest, the GGCT urinary protein is expected to become both

an easily accessible marker of tumour progression and a target for cancer treatment, including ccRCC.

Table 2. Selection of dysregulated urinary proteins already reported as altered in other diseases.

Detected Proteins	Reflected Feature	Other types of pathologies	Specimen
GDF15	*Primarily stage*	Diabetes mellitus type 1[38]	Blood, tissue
		Idiopathic nephropathy[41]	
		Membranous IgA nephropathy[39]	
		Urinary tract urothelial carcinoma[42]	
		Chronic kidney disease[40]	
		RCC[42]	
GGCT	*Primarily stage*	Uterine cervix carcinoma[31]	Tissue, extracellular fluids c mammary glands
		Lung carcinoma[31]	
		Colon carcinoma[31]	
		Breast carcinoma[31]	
		Ovarian carcinoma[32]	
LAMC1	*Primarily stage*	Chronic kidney disease[28]	Tissue, serum, urine
		ccRCC[29]	
PATJ	*Early grade & advanced stage*	ccRCC[44]	Tissue
		PRCC[44]	
		ChRCC[44]	
SNF	*Early grade & advanced stage*	Lung carcinoma[47, 48]	Tissue
GOLM1	*Advanced grade & advanced stage*	Lung squamous cell carcinoma[50]	Tissue
		Prostate cancer[51]	
		Breast cancer[52]	
LAIR1	*Advanced grade & advanced stage*	Hepatocellular carcinoma[54]	Tissue
		Oral squamous cell carcinoma[55]	
		RCC[56]	
THBS4	*Interdependence between grade & stage*	Gastric cancer[58]	Tissue
		Hepatocellular carcinoma[59,60]	
		Breast cancer[60]	
CLEC7A	*Interdependence between grade & stage*	ccRCC[63]	Tissue

The same consideration is even more relevant for the GDF15, Growth Differentiation Factor 15, a stress-inducible cytokine and member of the transforming growth factor β cytokine superfamily [34]. It is a pleiotropic protein that plays key roles in prenatal development, inflammation, regulation of cellular responses to stress signals, and tissue repair after acute injuries in adult life [35, 36]. GDF-15 has been reported as a biomarker in different kinds of pathologies, including kidney disease [37]. In fact, high circulating GDF-15 levels showed a significant correlation with a faster decline of renal function in patients with diabetes mellitus type 1 [38], IgA nephropathy [39], and CKD stages 1–4 [40]. Ham et al. displayed that circulating GDF-15 levels at the time of renal biopsy is negatively correlated with initial renal function in idiopathic membranous nephropathy [41].

In this context, this protein could be a possible indicator of development and progression also for the ccRCC. Traeger et al. already investigated this possibility, demonstrating that high GDF-15 levels in the plasma of patients suffering from urinary tract urothelial carcinomas and RCC were associated with metastasis, relapse and poor survival [42]. These evidences are in accordance with our observations, since high GDF-15 levels in the urine of ccRCC patients were associated with advanced cancerous conditions. In fact, GDF-15 was up-regulated in patients stratified as stage 3-4 (fold change=3.84, p-value=0.0002) and as HG (fold change=2.64, p-value=0.035). In addition, this protein resulted once more up-regulated in the stage 3-4, even when the cohort of patients was divided in the HG (fold change=9.22, p-value=0.03) and LG groups (fold change=1.87, p-value=0.03). Therefore, our data suggest that urinary GDF-15 could be used to monitor ccRCC development and aggressiveness and it should be evaluated as a promising prognostic marker.

Once we verified the existence of a panel of urinary proteins whose variation could be linked to the diverse histopathological features of ccRCC, each clinical parameter was considered alone in order to distinguish specific proteins for each

single component. We observed that many proteins (some already mentioned) were present only in the general comparisons (Figure 2, Supplementary Tables 2-3). However, the additional statistical analysis focused on the comparisons (Figures 2-3 and Supplementary Table 4):

- stage 3-4 vs stage 1 in LG;
- stage 3-4 vs stage 1 in HG;
- HG vs LG in stage 1;
- HG vs LG in stage 3-4;

allowed highlighting a new set of dysregulated proteins specifically related to grade and stage.

Intriguingly, it is immediately evident that existed 6 proteins whose relative abundance varied in patients with tumours at different stage (Figure 3, Supplementary Table 4) were observed. We could consider these urinary proteins able to reflect primarily the stage, given that they were present as varied in the comparison stage 3-4 vs stage 1 in both HG and LG sub-groups. Conversely, the comparison HG vs LG in the two different sub-populations of patients classified according to stage 1 and stage 3-4, did not show any protein alterations that could be associated primarily to the grade lesion (Figure 4, Supplementary Table 4).

However, the molecular scenario arisen from these comparisons appeared more intricate. We noted that a subset of proteins, such as ADGRF5, PATJ, PRG2, SFN, showed a complementary expression trend. Specifically, they were up-regulated in LG of the stage 3-4 subpopulation and up-regulated in the stage 3-4 of the LG subpopulation (Supplementary Table 4). Based on these findings, we may speculate that this panel of urinary proteins could specifically reflect the tumour conditions of early grade and advanced stage, at the same time.

Our results are consistent with literature data. In fact, it is reported that PATJ, Pals1-Associated Tight Junction protein, is a highly expressed protein in kidney epithelia that play a central role in the maintenance of the cell polarity in this kind of tissue

[43]. Through a whole exome sequencing study, Wang et al. detected in renal tissue specimens of RCC Chinese patients with primary RCC, stratified as ccRCC, PRCC and ChRCC, numerous newly somatic mutations, regarding several genes, including PATJ. They also revealed the strongly correlation between the mutated form of PATJ gene and the positive expression of PD-L1, Programmed Death Ligand 1, already known as a biomarker of response to the immunotherapy in many cancers, including RCC [44].

Among the other proteins, arouses interest the SNF, also known as stratifin. SNF is a highly conserved, ubiquitously expressed protein, associated with many different cellular processes and found most directly linked to cancer. The corresponding gene was originally identified as a p53-inducible gene, responsive to DNA-damaging agents [45, 46], defining for it a role of a negative regulator of cell cycle progression. However, Shiba-Ishii et al. in their two studies highlighted the overexpression of SNF in the lung adenocarcinoma and demonstrated that its increase stimulates tumour growth and accelerates its progression [47, 48]. Since in our analysis both PATJ and SNF resulted upregulated in urine of ccRCC patients stratified according to advanced stage and early grade (Supplementary Table 4), it is clear that these proteins could have a role in the process of tumorigenesis and spreading.

The presence of a panel of proteins in urine able to reflect the compresence of early and advanced histopathological classifications could also explain the characteristic abundance levels trend of the proteins GOLM1 and LAIR1. Both were up-regulated in stage 3-4 when considering only the HG sub-cohort and up-regulated in HG, when considering only the stage 3-4 sub-cohort (Supplementary Table 4). Together, they could reflect specifically and simultaneously the advanced stage and grade stratifications, in the urine.

GOLM1, Golgi Membrane protein 1, is a transmembrane glycoprotein of Golgi cisternae, commonly expressed in epithelial cells of normal tissues [49]. Several studies pinpointed its implication in carcinogenesis of multiple types of tumours.

Indeed, it is known that GOLM1 was significantly upregulated in lung squamous cell carcinoma tissues compared to the normal controls and its expression was associated with unfavourable overall survival [50]. In prostate and breast cancers, GOLM1 works as an oncogene by inducing cancer cell growth, migration and invasion, inhibiting cell apoptosis [51-52].

LAIR1, Leukocyte Associated Immunoglobulin like Receptor 1, is a transmembrane glycoprotein expressed in almost all cells of the immune system. It is reported that high levels of LAIR1 support leukemic cell growth [53] and promote cancer differentiation in the hepatocellular carcinoma (HCC) [54]. Additionally, the overexpression of LAIR-1 was associated with advanced pathological grade and correlated with immune suppressive features in oral squamous cell carcinoma [55]. More interestingly, LAIR1 was found significantly upregulated in RCC tissues compared to the normal ones, probably promoting cell proliferation by upregulation of the Akt phosphorylation in RCC [56]. These evidences suggest that the overexpression of LAIR1 contributes to RCC progression. In fact, we found also the urinary levels of LAIR1 up regulated in both the most advanced cancer histopathological subpopulations (Supplementary Table 4).

Moreover, the other two proteins likely to be interesting: THBS4 and XPNPEP2. They were both found down-regulated in HG in the general comparison HG vs LG, but up-regulated in the HG group, when the stage 1 feature was fixed (Supplementary Table 4). In this case, the interdependence of the two histopathological features is evident. Depending on the group of patients considered, in fact, the two tumour features differently influence the expression trend of the proteins. Nevertheless, we could still assume that the presence of a complementary expression trend might reflect the advanced grade, especially in combination with stage 1 lesion.

With this regard, it is interesting to report the case of the THBS4 protein. Thrombospondin-4 is an adhesive glycoprotein, part of the extracellular matrix. It

is involved in physiological cellular processes, such as proliferation, attachment, adhesion, and migration [57] and in pathological process of different types of malignancies. For instance, Chen et al. demonstrated that THBS4 upregulation was positively linked with an increased malignant potential and a poor clinical outcome in gastric cancer [58]. Wu et al., using immunohistochemistry and tissue microarrays, assessed THBS4 expression in HCC and concluded that its high abundance levels ware associated with shorter overall and disease-free survival for HCC patients [59]. Overexpression of THBS4 was also highly correlated with vascular invasion of advanced HCC stage [60]. Furthermore, THBS4 dysregulated expression contributes to the activated stromal response exhibited during tumour progression, facilitating the metastatic invasion in the breast cancer [61].

We found THBS4 down-regulated in the urine of RCC patients with LG lesions in the comparison regarding only the grade (fold change=-1.61, p-value=0.02), but up-regulated in HG lesions group (fold change=2.43, p-value=0.03), when the sub-cohort of stage 1 patients was evaluated. In accordance with the findings observed in the other type of cancers, we may suppose a possible role of THBS4 in advanced cc-RCC grade, only evident when the HG is co-present with the early stage lesion. This hypothesis could not be evident if our analyses were limited to the only evaluation of HG vs LG.

Finally, a similar consideration can be made for the proteins: CLEC7A, CP and CTSH. They emerged up-regulated in HG cohort of patients classified considering only the grade lesion; however, when the statistical analysis was limited to the stage 1 patients group, these proteins resulted discordantly up-regulated in LG condition (Supplementary Table 4).

Among these, the protein CLEC7A deserves a special mention. Also known as dectin-1, CLEC7A was originally identified as a pattern-recognition receptor expressed on dendritic cells [62]. Actually, it was demonstrated that its expression on dendritic cells and macrophages enhances the recognition of the N-glycans on cancerous cells, helping tumour killing by natural killer cells; this suggests its

potential role in cancer immunomodulation [65]. In this regard, Xia et al., through their immunohistochemical tissue microarrays study, showed that CLEC7A expression was strongly correlated with higher stage of ccRCC. They also highlighted that high tumoural CLEC7A expression was an independent predictor for both shorter patient recurrence free and overall survivals [63].

Our analysis pinpointed that CLEC7A levels were statistically significant in the urine of ccRCC patients. Specifically, it resulted consistently upregulated in LG when the stage 1 lesion is fixed (fold change=3.38, p-value=0.016, Supplementary Table, column 3) and in stage 3-4, when the LG lesion is fixed (fold change=2.15, p-value=0.019, Supplementary Table, column 5). According to the in tissue investigation [64], urinary CLEC7A levels appeared correlated to the higher ccRCC stage. Additionally, the presence of this protein in the ccRCC patients urine also appeared correlated to lower grade. Moreover, these evidences emerged only when the specific sub-cohort (LG and stage 1) of patients was considered.

4. MATERIALS AND METHODS

4.1. Patients selection

44 patients from San Gerardo hospital (Monza, Italy) were recruited between 2011 and 2016 and all subjects had signed an informed consent. The local ethical committee (Comitato Etico Azienda Ospedaliera San Gerardo, Monza, Italy) approved protocols and procedures. The diagnosis was performed after careful a 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).

4.2 Sample collection and preparation

The second urine of the morning, before total or partial nephrectomy, were collected into sterile urine tubes. The samples were then centrifuged at 1000 xg, 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 30 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 [65]. Briefly, 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 overnight digestion was performed on 30kDa filters (Amicon Ultra-500 µg 30 kDa, Millipore) 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 H2O: 2 ACN: 0.1% TFA and analysed by nUHPLC-MS/MS.

4.3 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). About 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, based on 10 mM sodium formiate cluster solution (15 min before each run), converted and deconvoluted to XML file using DataAnalysis software (Bruker Daltonics, Germany).

4.4 Data processing

Protein identification was obtained with PEAKS Studio 8.5 (Bioinformatics Solutions Inc., Waterloo, Ontario, Canada [66]). Trypsin was set as the enzyme used for the digestion, with 1 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.

The 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, patients were grouped according to the levels of a chosen outcome attribute (e.g. the grade), then proteins were considered of different abundance (ratio >1.5 and <0.66) 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

5. CONCLUSIONS

In summary, we enlighted that urinary proteome analysis could represent a precious source of information for the understanding of ccRCC progression. Collectively, we revealed a set of proteins with an altered urinary expression according to tumour spread. Several of them were already known to be involved in carcinogenesis, tumour progression and aggressiveness in RCC and in other cancer types. These results lead us to show that the urinary proteome of ccRCC patients is influenced primarily by the dimension and position (stage) and partially by the grade of the tumour. They also support the hypothesis that a panel of urinary proteins could be able to reflect the clinicopathological characteristics of ccRCC and the corresponding possible combinations of these two features, overcoming and complementing the static stage and grade classifications.

Author Contributions: Investigation, formal analysis and Writing— Original Draft, L. Santorelli and M. Stella (# equally contributing authors); data curation, C. Chinello, I. Piga and V. Denti; Writing—review and editing, A. Smith and F. Clerici; software, G. Capitoli; resources (patients enrollment and clinical diagnosis), A. Grasso, M. Grasso and G. Bovo; Supervision, F. Magni.

Funding: research was funded by MIUR: FI RB 2007 (RBRN07BMCT_11), Fondi Ateneo Ricerca: 2016–2018 and by Fondazione Gigi & Pupa Ferrari Onlus.

References

[1] Chinello, C.; L'imperio, V.; Stella, M.; Smith, A. J.; Bovo, G.; Grasso, A.; Grasso, M.; Raimondo, F.; Pitto, M.; Pagni, F. and Magni F. The proteomic landscape of renal tumors. *Expert Rev. Proteomics*, **2016**, vol. 13, no. 12, pp. 1103–1120. doi: 10.1080/14789450.2016.1248415.

[2] Haddad, A.; Margulis, V. Tumour and patient factors in renal cell carcinoma - Towards personalized therapy. *Nature Reviews Urology*, **2015**, pp. 253-262. DOI: 10.1038/nrurol.2015.71.

[3] van den Berg, E. Renal Cell Carcinoma. *Brenner's Encyclopedia of Genetics: Second Edition*, 2**013**, pp. 130–132. doi: 10.1016/B978-0-12-374984-0.01293-6.

[4] Moch, H.; Cubilla, A. L.; Humphrey P. A.; Reuter V. E. and Ulbright, T. M. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs—Part A: Renal, Penile, and Testicular Tumours. *Eur. Urol.*, **2016**, vol. 70, no. 1, pp. 93–105. doi: 10.1016/j.eururo.2016.02.029.

[5] Dagher, J.; Delahunt, B.; Rioux-Leclercq, N.; Egevad, L.; Srigley, J.; Coughlin, G.; Dunglinson N.; Gianduzzo, T.; Kua, B.; Malone, G.; Martin, B.; Preston, J.; Pokorny, M.; Wood, S.; Yaxley, J.; Samaratunga H. Clear cell renal cell carcinoma: validation of World Health Organization/International Society of Urological Pathology grading. *Histopathology*, vol. 71, **2017**, no. 6, pp. 918–925. doi: 10.1111/his.13311.

[6] Geiger, T.; Madden, S. F.; Gallagher, W. M.; Cox, J. and Mann, M. Proteomic portrait of human breast cancer progression identifies novel prognostic markers. *Cancer Res.*, **2012**, vol. 72, no. 9, pp. 2428–2439. doi: 10.1158/0008-5472.CAN-11-3711.

[7] Sandim, V.; Pereira, D. A.; Ornellas, A. A. and Alves, G. Renal cell carcinoma and proteomics. *Urol.* **2010**, Int., vol. 84, no. 4, pp. 373–377. doi: 10.1159/000296283.

[8] Quandt, D.; Zucht, H.; Amann, A.; Wulf-Goldenberg, A.; Borrebaeck, C.; Cannarile, M.; Lambrechts, D.; Oberacher, H.; Garrett, J.; Nayak, T.; Kazinski, M.; Massie, C.; Schwarzenbach H.; Maio, M.; Prins, R.; Wendik, B.; Hockett, R.; Enderle, D.; Noerholm, M.; Hendriks, H.; Zwierzina, H.; Seliger, B. Implementing liquid biopsies into clinical decision making for cancer immunotherapy. *Oncotarget*, **2017**, vol. 8, no. 29, pp. 48507–4852. doi 10.18632/oncotarget.17397.

[9] Di Meo, A.; Bartlett, J.; Cheng, Y.; Pasic, M. D. and Yousef G. M. Liquid biopsy: A step forward towards precision medicine in urologic malignancies. *Mol. Cancer*, **2017**, vol. 16, no. 1, pp. 1–14. doi: 10.1186/s12943-017-0644-5.

[10] Magni, F.; Van Der Burgt, Y.; Chinello, C.; Mainini, V.; Gianazza, E.; Squeo, V.; Deelder, A.; Kienle M. Biomarkers discovery by peptide and protein profiling in biological fluids based on functionalized magnetic beads purification and mass spectrometry. *Blood Transfus.*, **2010**, Suppl 3, s92-7. doi: 10.2450/2010.015S.

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

[12] Klein, J.; Bascands, J. L.; Mischak, H. and Schanstra J. P. The role of urinary peptidomics in kidney disease research. *Kidney International*, **2016**, pp. 539-545. doi: 10.1016/j.kint.2015.10.010.

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

[14] Mijuskovic, M.; Stanojevic, I.; Milovic, N.; Cerovic, S.; Petrovic, D.; Maksic, D.; Kovacevic B.; Andjelic, T.; Aleksic, P.; Terzic, B.; Djukic, M.; Vojvodic, D. Urinary concentrations of aquaporin-1 and perilipin-2 in patients with renal cell carcinoma correlate with tumor size and stage but not grade. *Urology*, **2014**, vol. 50, pp. 63-70. doi: 10.1007/s11255-017-1724-6.

[15] Sandim, V.; Pereira, D.; Kalume, D.; Oliveira-Carvalho, A.; Ornellas, A.; Soares, M.; Alves G.; Zingali, R. Proteomic analysis reveals differentially secreted proteins in the urine from patients with clear cell renal cell carcinoma.

Urol. Oncol. Semin. Orig. Investig., **2016**, vol. 34, pp. 5.e11-5.e25. doi: 10.1016/j.urolonc.2015.07.016.

[16] Gianazza, E.; Chinello, C.; Mainini, V.; Cazzaniga, M.; Squeo, V.; Albo, G.; Signorini, S.; Di Pierro, S.; Ferrero, S.; Nicolardi, S.; van der Burgt, Y.; Deelder, A.; Magni F. Alterations of the serum peptidome in renal cell carcinoma discriminating benign and malignant kidney tumors. *J. Proteomics*, **2012**, vol. 76, pp. 125–140. doi: 10.1016/j.jprot.2012.07.032.

[17] 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. *J. Proteomics*, **2019**, vol. 191, pp. 38–47. doi: 10.1016/j.jprot.2018.04.028.

[18] Song, Y.; Zhong, L; Zhou, J.; Lu, M.; Xing, T.; Ma, L.; Shen, J. Data-Independent Acquisition-Based Quantitative Proteomic Analysis Reveals Potential Biomarkers of Kidney Cancer. *Proteomics - Clin. Appl.*, **2017**, vol. 11, pp. 11-12. doi: 10.1002/prca.201700066.

[19] Atrih, A.; Mudaliar, M.; Zakikhani, P.; Lamont, D.; Huang, J.; Bray, S.; Barton, G.; Fleming, S.; Nabi, G. Quantitative proteomics in resected renal cancer tissue for biomarker discovery and profiling. *Br. J. Cancer*, **2014**, vol. 110, pp. 1622-1633. doi: 10.1038/bjc.2014.24.

[20] Zaravinos, A.; Pieri, M.; Mourmouras, N.; Anastasiadou, N.; Zouvani, I.; Delakas, D.; Deltas, C. 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**, vol. 1, no. 2, pp. 117–131. doi: 10.18632/oncoscience.13.

[21] Chan, J. Y.; Choudhury, Y. and Tan, M. H. Predictive molecular biomarkers to guide clinical decision making in kidney cancer: Current progress and future challenges, *Expert Review of Molecular Diagnostics*, **2015**, pp. 631-646. doi: 10.1586/14737159.2015.1032261.

[22] Chinello, C.; Cazzaniga, M.; Sio, G.; Smith, A.; Grasso, A.; Rocco, B.; Signorini, S.; Grasso, M.; Bosari, S.; Zoppis, I.; Mauri, G.; Magni, F. Tumor size, stage and grade alterations of urinary peptidome in RCC. *J. Transl. Med.*, **2015**, vol. 13, no. 1, pp. 1–11. doi: 10.1186/s12967-015-0693-8.

[23] Massari, F.; Ciccarese, C.; Santoni, M.; Brunelli, M.; Piva, F.; Modena, A.; Bimbatti, D.; Fantinel, E.; Santini, D.; Cheng, L.; Cascinu, S.; Montironi, R.; Tortora, G. Metabolic alterations in renal cell carcinoma. *Cancer Treat. Rev.*, **2015**, vol. 41, no. 9, pp. 767–776. doi: 10.1016/j.ctrv.2015.07.002.

[24] Linehan, W. M. Genetic *Basis and Kidney Cancer*, **2011**, vol. 7, no. 5, pp. 277–285. doi: 10.1038/nrurol.2010.47.

[25] Bockorny B. and Dasanu C. A. Intrinsic immune alterations in renal cell carcinoma and emerging immunotherapeutic approaches, *Expert Opin. Biol. Ther.*, **2013**, vol. 13, no. 6, pp. 911–925. doi: 10.1517/14712598.2013.778970.

[26] Wen, L.; Guo, L.; Zhang, W.; Li, Y.; Jiang, W.; Di, X.; Ma, J.; Feng, L.; Zhang, K.; Shou, J.; Cooperation Between the Inflammation and Coagulation Systems Promotes the Survival of Circulating Tumor Cells in Renal Cell Carcinoma Patients. *Front. Oncol.*, **2019**, vol. 9, pp. 1–13. doi: 10.3389/fonc.2019.00504.

[27] Tian, Z.-H.; Yuan, C.; Yang, K. and Gao, X.-L., Systematic identification of key genes and pathways in clear cell renal cell carcinoma on bioinformatics analysis. *Ann. Transl. Med.*, **2019**, vol. 7, no. 5, pp. 89–89. doi: 10.21037/atm.2019.01.18.

[28] Nielsen, S.; Rasmussen, D.; Brix, S.; Fenton, A.; Jesky, M.; Ferro, C.; Karsdal, M.; Genovese, F.; Cockwell, P. A novel biomarker of laminin turnover is associated with disease progression and mortality in chronic kidney disease. *PLoS One*, **2018**, vol. 13, no. 10, pp. 1–13. doi: 10.1371/journal.pone.0204239.

[29] Machackova, T.; Mlcochova, H.; Stanik, M.; Dolezel, J., Fedorko, M.; Pacik, D.; Poprach, A.; Svoboda, M.; Slaby, O. MiR-429 is linked to metastasis and poor prognosis in renal cell carcinoma by affecting epithelial-mesenchymal transition. *Tumor Biol.*, **2016**, vol. 37, no. 11, pp. 14653–14658. doi: 10.1007/s13277-016-5310-9.

[30] Liu, Y.; Hyde, A. S.; Simpson, M. A. and Barycki, J. J. Emerging regulatory paradigms in glutathione metabolism. *Elsevier Inc.*, **2014**, 1st ed., vol. 122. doi: 10.1016/B978-0-12-420117-0.00002-5.

[31] Gromov, P.; Gromova, I.; Friis, E.; Timmermans-Wielenga, V. Rank, F.; Simon, R.; Sauter, G. and Moreira J. M. A. Proteomic profiling of mammary carcinomas identifies c7orf24, a gamma-glutamyl cyclotransferase, as a potential cancer biomarker, *J. Proteome Res*., **2010**, vol. 9, no. 8, pp. 3941–3953.

[32] Li, Y.; Wu, T.; Wang, Y.; Yang, L.; Hu, C.; Chen, L.; Wu, S. γ-Glutamyl cyclotransferase contributes to tumor progression in high grade serous ovarian cancer by regulating epithelial-mesenchymal transition via activating PI3K/AKT/mTOR pathway. *Gynecol. Oncol.*, **2018**, vol. 149, no. 1, pp. 163–172. doi: 10.1016/j.ygyno.2018.01.023.

[33] Kageyama, S.; Ii, H.; Taniguchi, K.; Kubota, S.; Yoshida, T.; Isono, T.; Chano, T.; Yoshiya, T.; Ito, K.; Yoshiki, T.; Kawauchi, A.; Nakata, S. Mechanisms of tumor growth inhibition by depletion of γ-glutamylcyclotransferase (GGCT): A novel molecular target for anticancer therapy. *Int. J. Mol. Sci.*, **2018**, vol. 19, no. 7. doi: 10.3390/ijms19072054.

[34] Berezin, A. E. Cardiac biomarkers in diabetes mellitus: New dawn for risk stratification? *Diabetes Metab. Syndr. Clin. Res. Rev*., **2017**, vol. 11, pp. S201–S208. doi: 10.1016/j.dsx.2016.12.032.

[35] Preusch, M.; Baeuerle, M.; Albrecht, C.; Blessing, E.; Bischof, M.; Katus, H.; Bea, F. GDF-15 protects from macrophage accumulation in a mousemodel of advanced atherosclerosis. *Eur. J. Med. Res.*, **2013**, vol. 18, no. 1, pp. 1–6. doi: 10.1186/2047-783X-18-19.

[36] Corre, J.; Hébraud, B.; Bourin, P. Concise Review: Growth Differentiation Factor 15 in Pathology: A Clinical Role? *Stem cells transl. med*., **2013**, pp. 946–952.

[37] Desmedt, S.; Desmedt, V.; De Vos, L.; Delanghe, J. R.; Speeckaert, R. and Speeckaert, M. M. Growth differentiation factor 15: A novel biomarker with

high clinical potential. *Crit. Rev. Clin. Lab. Sci.*, **2019**, vol. 0, no. 0, pp. 1–18. doi: 10.1080/10408363.2019.1615034.

[38] Hamon, S. M.; Griffin, T. P.; Islam, M. N.; Wall, D.; Griffin, M. D. and O'Shea, P. M. Defining reference intervals for a serum growth differentiation factor-15 (GDF-15) assay in a Caucasian population and its potential utility in diabetic kidney disease (DKD) *Clin. Chem. Lab. Med.*, **2019**, vol. 57, no. 4, pp. 510–520. doi: 10.1515/cclm-2018-0534.

[39] Na, K.; Kim, Y.; Chung, H.; Yeo, M.; Ham, Y.; Jeong, J.; Kim, K.; Lee, K.; Choi, D. Growth differentiation factor 15 as a predictor of adverse renal outcomes in patients with immunoglobulin A nephropathy. *Intern. Med. J.*, **2017**, vol. 47, no. 12, pp. 1393–1399. doi: 10.1111/imj.13614.

[40] Nair, V.; Robinson-Cohen, C.; Smith, M.; Bellovich, K.; Bhat, Z.; Bobadilla, M.; Brosius, F.; de Boer, I.; Essioux, L.; Formentini, I.; Gadegbeku, C.; Gipson, D.; Hawkins, J.; Himmelfarb, J.; Kestenbaum, B.; Kretzler, M.; Magnone, M.; Perumal, K.; Steigerwalt, S.; Ju, W.; Bansal, N. Growth Differentiation Factor–15 and Risk of CKD Progression. *J. Am. Soc. Nephrol.*, **2017**, vol. 28, no. 7, pp. 2233–2240. doi: 10.1681/asn.2016080919.

[41] Ham, Y.; Jeong, J.; Kim, H.; Bae, H.; Choi, D.; Na, K.; Lee, K. Sp191Growth Differentiation Factor-15 As a Predictor of Idiopathic Membranous Nephropathy Progression. *Nephrol. Dial. Transplant.*, **2018**, vol. 33, no. suppl_1, pp. i407–i408. doi: 10.1093/ndt/gfy104.sp191.

[42] Traeger, L.; Ellermann, I., Wiethoff, H.; Ihbe, J.; Gallitz, I.; Eveslage, M.; Moritz, R.; Herrmann, E.; Schrader, A.; Steinbicker, A. Serum Hepcidin and GDF-15 levels as prognostic markers in urothelial carcinoma of the upper urinary tract and renal cell carcinoma. *BMC Cancer*, **2019**, vol. 19, no. 1, pp. 1–10. doi: 10.1186/s12885-019-5278-0.

[43] Pieczynski, J. and Margolis, B. Protein complexes that control renal epithelial polarity. *Am. J. Physiol. Physiol.*, **2011**, vol. 300, no. 3, pp. F589–F601. doi: 10.1152/ajprenal.00615.2010.

[44] Wang, J.; Xi, Z.; Xi, J.; Zhang, H.; Li, J., Xia, Y., Yi, Y. Somatic mutations in renal cell carcinomas from Chinese patients revealed by whole exome sequencing. *Cancer Cell Int.*, **2018**, vol. 18, no. 1, pp. 1–12. doi: 10.1186/s12935-018-0661-5.

[45] Hermeking, H.; Lengauer, C.; Polyak, K.; He, T. C.; Zhang, L.; Thiagalingam, S.; Kinzler, K. W. and Vogelstein, B. 14-3-3 sigma is a p53-regulated inhibitor of G2-M progression. *Mol. Cell*, **1997**, vol. 1, pp. 3–11.

[46] Vogelstein, B.; Chan, T. A.; Hermeking, H.; Lengauer, C. and Kinzler K. W. 14-3-3|[sigma]| is required to prevent mitotic catastrophe after DNA damage. *Nature*, **1999**, vol. 401, no. 6753, pp. 616–620. doi: 10.1038/44188.

[47] Shiba-Ishii, A.; Kano, J.; Morishita, Y.; Sato, Y.; Minami, Y. andNoguchi M. High expression of stratifin is a universal abnormality during the course of malignant progression of early-stage lung adenocarcinoma. *Int. J. Cancer*, 2011, vol. 129, no. 10, pp. 2445–2453. doi: 10.1002/ijc.25907.

[48] Shiba-Ishii, A.; Kim, Y.; Shiozawa, T., Iyama, S.; Satomi, K.; Kano, J.; Sakashita, S.; Morishita Y.; Noguchi, M. Stratifin accelerates progression of lung adenocarcinoma at an early stage. *Mol. Cancer*, **2015**, vol. 14, no. 1, pp. 1–6. doi: 10.1186/s12943-015-0414-1.

[49] Zhou, Y.; Li, L.; Hu, L. and Peng, T. Golgi phosphoprotein 2 (GOLPH2/GP73/GOLM1) interacts with secretory clusterin. *Mol. Biol. Rep.*, **2011**, vol. 38, no. 3, pp. 1457–1462. doi: 10.1007/s11033-010-0251-7.

[50] Liu, X.; Chen, L. and Zhang, T. Increased GOLM1 Expression Independently Predicts Unfavorable Overall Survival and Recurrence-Free Survival in Lung Adenocarcinoma. *Cancer Control*, **2018**, vol. 25, no. 1, pp. 1–10. doi: 10.1177/1073274818778001.

[51] Varambally, S.; Laxman, B.; Mehra, R.; Cao, Q.; Dhanasekaran, S.; Tomlins, S.; Granger, J.; Vellaichamy, A.; Sreekumar, A.; Yu, J.; Gu, W.; Shen, R.; Ghosh, D.; Wright, L.; Kladney, R.; Kuefer, R.; Rubin, M.; Fimmel, C., Chinnaiyan, A. Golgi Protein GOLM1 Is a Tissue and Urine Biomarker of

Prostate Cancer. *Neoplasia*, **2008,** vol. 10, no. 11, pp. 1285-IN35. doi: 10.1593/neo.08922.

[52] Yan, G.; Ru, Y.; Wu, K.; Yan, F.; Wang, Q.; Wang, J.; Pan, T.; Zhang, M.; Han, H.; Li, X.; Zou, L. GOLM1 promotes prostate cancer progression through activating PI3K-AKT-mTOR signaling. *Prostate*, **2018**, vol. 78, no. 3, pp. 166–177. doi: 10.1002/pros.23461.

[53] Zhang, R.; Zhu, Z.; Shen, W.; Li, X.; Dhoomun, D. K. and Tian Y. Golgi Membrane Protein 1 (GOLM1) Promotes Growth and Metastasis of Breast Cancer Cells via Regulating Matrix Metalloproteinase-13 (MMP13). *Med. Sci. Monit.*, **2019**, vol. 25, pp. 847–855. doi: 10.12659/msm.911667.

[54] Yang, G.; Sau, C.; Lai, W.; Cichon, J. and Li, W. The ITIM-containing receptor LAIR1 is essential for acute myeloid leukemia development. *Nat. Cell Bio.l*, **2015**, vol. 344, no. 6188, pp. 1173–1178. doi: 10.1126/science.1249098.

[55] Wu, X.; Zhang, L.; Zhou, J.; Liu, L.; Fu, Q.; Fu, A.; Feng, X.; Xin, R.; Liu, H.; Gao, Y.; Xue, J. Clinicopathologic significance of LAIR-1 expression in hepatocellular carcinoma. *Curr. Probl. Cancer*, **2019**, vol. 43, no. 1, pp. 18–26. doi: 10.1016/j.currproblcancer.2018.04.005.

[56] Yang, L.; Zhang, M.; Wu, L.; Mao, L.; Chen, L.; Yu, G.; Deng, W.; Zhang, W.; Liu, B.; Sun, W.; Sun, Z. LAIR-1 overexpression and correlation with advanced pathological grade and immune suppressive status in oral squamous cell carcinoma. *Head Neck*, **2019**, vol. 41, no. 4, pp. 1080–1086 doi: 10.1002/hed.25539.

[57] Jingushi, K.; Uemura, M.; Nakano, K.; Hayashi, Y.; Wang, C.; Ishizuya, Y.; Yamamoto, Y., Hayashi, T.; Kinouchi, T.; Matsuzaki, K.; Kato, T.; Kawashima, A.; Ujike, T.; Nagahara, A.; Fujita K.; Ueda, K.; Tsujikawa, K.; Nonomura, N. Leukocyte-associated immunoglobulin-like receptor 1 promotes tumorigenesis in RCC. *Oncol. Rep.*, **2019**, vol. 41, no. 2, pp. 1293–1303. doi: 10.3892/or.2018.6875.

[58] Liu, J.; Cheng, G.; Yang, H.; Deng, X.; Qin, C.; Hua, L.; Yin, C. Reciprocal regulation of long noncoding RNAs THBS4-003 and THBS4 control migration

and invasion in prostate cancer cell lines. *Mol. Med. Rep*., **2016**, vol. 14, no. 2, pp. 1451–1458. doi: 10.3892/mmr.2016.5443.

[59] Chen, X.; Huang, Y.; Wang, Y.; Wu, Q.; Hong, S. and Huang Z. THBS4 predicts poor outcomes and promotes proliferation and metastasis in gastric cancer. *J. Physiol. Biochem.*, **2019**, vol. 75, no. 1, pp. 117–123. Doi: 10.1007/s13105-019-00665-9.

[60] Su, F.; Zhao, J.; Qin, S.; Wang, R.; Li, Y.; Wang, Q.; Tan, Y.; Jin, H.; Zhu, F.; Ou, Y.; Cheng, Z.; Su, W.; Zhao, F.; Yang, Y.; Zhou, Z.; Zheng, J.; Li, Y., Li, Z., Wu, Q. Over-expression of Thrombospondin 4 correlates with loss of miR-142 and contributes to migration and vascular invasion of advanced hepatocellular carcinoma. *Oncotarget*, **2017**, vol. 8, no. 14, pp. 23277–23288. doi: 10.18632/oncotarget.15054.

[61] McCart Reed, A.; Song, S.; Kutasovic, J.; Reid, L.; Valle, J.; Vargas, A.; Smart, C.; Simpson, P. Thrombospondin-4 expression is activated during the stromal response to invasive breast cancer. *Virchows Arch.*, **2013,** vol. 463, no. 4, pp. 535–545. doi: 10.1007/s00428-013-1468-3.

[62] Ariizumi, K.; Shen, G.; Shikano, S.; Xu, S.; Ritter, R.; Kumamoto, T.; Edelbaum, D.; Morita, A.; Bergstresser, P.; Takashima, A. Identification of a novel, dendritic cell-associated molecule, dectin-1, by subtractive cDNA cloning. *J. Biol. Chem*., **2000**, vol. 275, no. 26, pp. 20157–20167. doi: 10.1074/jbc.M909512199.

[63] Chiba, S.; Ikushima, H.; Ueki, H.; Yanai, H.; Kimura, Y.; Hangai, S.; Nishio, J.; Negishi, H.; Tamura, T.; Saijo, S.; Iwakura, Y.; Taniguchi, T. Recognition of tumor cells by Dectin-1 orchestrates innate immune cells for anti-tumor responses. *Elife*, **2014**, vol. 3, pp. 1–20, 2014. doi: 10.7554/eLife.04177.

[64] Xia, Y.; Liu, L.; Bai, Q.; Wang, J.; Xi, W.; Qu, Y.; Xiong, Y.; Long, Q.; Xu, J.; Guo, J. Dectin-1 predicts adverse postoperative prognosis of patients with clear cell renal cell carcinoma. *Sci. Rep*., **2016**, vol. 6, pp. 2–10. Doi: 10.1038/srep32657.

[65] Chinello, C.; Stella, M.; Piga, I.; Smith, A.; 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. *J. Proteomics*, **2019**, vol. 191, pp. 29-37. doi: 10.1016/j.jprot.2018.04.029.

[66] Ma, B.; Zhang, K.; Hendrie, C.; Liang, C.; Li, M.; Doherty-Kirby, A.; Lajoie, G. PEAKS: Powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom*., **2003**, vol.17, pp. 2337-2342. doi: 10.1002/rcm.1196.

CHAPTER 3

In-depth mapping of the urinary N-Glycoproteome: distinct signatures of ccRCC related progression

Santorelli Lucia[a]#, Giulia Capitoli[c]#, Marco Grasso[b], Fulvio Magni[a]

#equally contributing authors

a) School 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) School of Medicine and Surgery, University of Milano – Bicocca, Centre of Biostatistics for Clinical Epidemiology, Vedano al Lambro, Italy.

Manuscript in preparation

ABSTRACT Protein N-glycosylation is one of the most complex and frequently occurring post-translational modifications involved in many biological processes, such as cell adhesion, molecular trafficking and clearance. Aberrant changes in N-glycosylation protein pattern are closely associated with several diseases, including cancer; indeed, it is known that glycosylation profiles are modulated consequently during tumour progression and spreading.

Thus, efficient identifications of these patterns and reliable differentiation between levels of aberrant protein glycoforms in healthy and tumour conditions would be useful for understanding the molecular mechanism of this multifactorial disease, developing specific biomarkers and finding novel therapeutic targets. In particular, this approach is likely to be promising for the most common and aggressive neoplasm in kidney, the clear cell Renal Cell Carcinoma (ccRCC).

The aim of this study was to identify potential glycomarkers in urine belonging to patients affected by ccRCC at different stages. In particular, we mapped the N-glycosylation sites of urinary proteins obtained from patients with ccRCC classified as pT1a (n=15) and pT3a (n=15), and from non-ccRCC subjects (n=15), by an N-glyco-FASP-based method. The proteome enriched in glycoproteins was identified and evaluated via label-free nLC-ESI MS/MS.

This approach permits to examine the number and the distribution of the modification sites and simultaneously to quantify stage-related changes of N-glycosylated amino acids sequences. These results expand our knowledge of this important class of post-translational modifications also to urine. The translation of this information into pre-clinical studies could gain a remarkable impact for biomarkers and therapeutic targets discovery in kidney cancer.

KEYWORDS Clear cell renal cell carcinoma, urine, Glycoproteomics, glyco-mapping, glycomarkers.

1. INTRODUCTION

Malignant transformation is a complex of heterogeneous cellular events, depending on many different factors, that regulates the growth and the survival cycle of affected cells. Among cancer-associated alterations, changes in protein N-glycosylation have recently received attention as one of the key events, able to influence, up to lead to neoplasia onset and its consequent spreading [1].

N-Glycosylation represents the most prominent post-translational modification (PTM) and its addition to the rising proteins, in the endoplasmic reticulum, plays diverse roles, including the maintenance of correct tertiary structure, the conferment of a specific function and the establishment of a well-defined cellular localisation [2]. Under physiological conditions, N-glycosylated proteins rule in the cell diverse biological fundamental processes, including: folding and quality control, cell adhesion and motility, molecular trafficking, cell signalling, immune recognition and clearance [3].

Moreover, the fundamental role in oncogenesis of this PTM was widely documented over seven decades go [4], [5]. High levels of altered N-glycosylated proteins, in fact, are demonstrated to promote tumour invasiveness and to correlate with high frequency of tumour recurrence and metastasis. Furthermore, it is well established that aberrant changes of the N-glycosylated protein patterns significantly accompany the normal tissue in its transformation towards a neoplastic sense [6], [7]. These abnormal changes essentially affect the occupancy of the glycosylation sites (glycan macro-heterogeneity) and/or the attached N-glycan structures (glycan micro-heterogeneity) [8], [9].

In recent years, a consistent group of glycoproteins have passed the discovery and validation phases and are regularly used in clinical practise as cancer biomarkers, successfully completing the research bench to the patient bedside program. Most of the approved cancer glycoprotein-based biomarkers (glycomarkers) are proteins derived from diverse bodily fluids. Prostate-specific antigen (PSA), for example, which is largely present in seminal fluids and plasma, is used to screen and monitor

prostate cancer patients since 20 years [10], [11]. Cancer Antigen 15-3 and Cancer Antigen 125, both detected in serum [12], are useful biomarkers for breast [12] and ovarian [13], [14] tumours respectively and are especially used for monitoring affected patients, evaluating the possible disease recurrence. N-glycosylated protein markers include also the Carcinoembryonic Antigen (CEA), detectable principally in blood [15], [16], which has been correlated with colorectal, bladder, breast, pancreatic and lung cancers [17], [18].

In this context, an easily accessible biological sample, such as urine, is a valuable glycomarkers source for kidney cancer related diseases. It can be collected in large quantities, non-invasive way and its molecular composition is less complex than other bodily fluids.

In the last two decades, the urine proteomics has exponentially flourished and expanded [19], [20]. Thanks to the rapid development of MS-based technologies applied to clinical research field, the urine content was in depth explored and a number of urinary glycoproteins of controls subjects was enlightened and extensively characterised [21]. The typical proteome of healthy human urine [22] counts almost 2500 proteins and about 300 of these are reported to be N-glycosylated [23]. While the urine proteome was largely investigated, limited information are viable describing the urinary N-glycoproteome of patients affected by clear cell Renal Cell Carcinoma (ccRCC), the most aggressive RCC morphotype [24].

Despite ccRCC is usually diagnosed at early stages, its aggressiveness and clinical outcomes still remains heterogeneous within each staging group, making the research of novel diagnostic and prognostic predictors an urgent priority.

Since the cancer transformation causes alterations in the synthesis and expression of specific glycosylated proteins, the evaluation of the glycoproteins content of the urine of ccRCC affected patients arises as a valid strategy to expand our knowledge about the role of this modification in the cancer onset. Moreover, we attempted to use a comprehensive Glycoproteomics approach to study ccRCC urine at different stages, early (pT1), advanced (pT3) compared to urine of non-affected ccRCC, as

controls subjects (CTRL): among these groups, we evaluated the alterations of the proteins and N-glycoproteins associated with the histological status. Our goal was to identify potential candidate indicators of biological interest in ccRCC development and progression, in order to better aid in the stage-related classification and in directing the cancer treatment.

2. RESULTS and DISCUSSION

2.1 Clinical data and study design

To identify potential protein glycoforms of interest, peptides enriched in the urine of RCC patients, divided by pT, and control subjects (patients with renal calculus) were analysed by nUHPLC-MS/MS. Three protein pools were prepared, one representative of the CTRL and other two of ccRCC, one of pT1 and one of pT3 (15 individuals each), age and sex matched (Table 1).

Table 1. Clinicopathological data regarding the patients enrolled in the study.

Group	# of patients	Gender (male - female)	Age in years at diagnosis, median (range)	Greatest tumour dimension (cm)
CTRL	15	10 ─ 5	57.9 (39 ─ 77)	/
pT1	15	8 ─ 7	67.8 (42 ─ 82)	4
pT3	15	12 ─ 3	68.9 (45 ─ 81)	12

2.3 Mapping of the N-glycosylation sites of the urinary glycoproteins

In order to enrich N-glycoproteins from urine, the Glyco-FASP method was applied [25]. An aliquot of tryptic peptides (~100 µg) was incubated in presence of a triad of lectins, proteins able to bind different types of carbohydrates, and the mixture thus obtained was loaded on the top of the filter unit. After serial centrifugation steps, the unbound peptides were eluted, while the bound ones (the N-glycopeptides) underwent to enzymatic hydrolysis with a PNGase F. This reaction

causes the disruption of the interaction lectins-peptides; the latter, released their carbohydrate component, were recovered and resulted 'marked'. In fact, the PNGase F reaction leads to deamidation of the asparagine residue to aspartic acid and to a mass increase of 0.9848 Da of the modified sites, which can be detected as a mass shift of the precursor peptide and its fragments.

Firstly, we investigated the whole N-glycoproteome of each clinical group, without considering their corresponding protein identification, in order to primarily discovery possible differences among the three sub-cohorts in terms of glycan macro-heterogeneity. Thus, we submitted the list of identified peptides in the deglycosylated fractions into an in house-made software, in order to clean up the row data and proceed to an in depth mapping of the N-glycosylated sites.

A total number of 592 glycopeptides were comprehensively identified in the three patients cohorts. Among these, the 41% resulted shared into the all groups, the 23% was exclusively present in two of the three groups, while the 14%, 13% and 9% were specific related to CTRL, pT1 and pT3, respectively.

Then, we evaluated the sites distribution of the N-glycosylation in the urinary proteome, determining the recurrence frequency of the consensus and non-consensus regions in the identified peptides. Reminding that the canonical N-glycosylation motif of proteins is N-*x*-S/T (where *x* are all the amino acids, except the proline), our analysis allowed to highlight a variation of the glycosylated sites between the CTRL and ccRCC patients.

Primarily, we compared the CTRL and the ccRCC affected subjects, without considering the pT feature as variable. We observed that the ratio between the number of asparagine modified (N) present in the glycopeptides enriched fraction and the total number of asparagine (N) present in the corresponding proteome was constant in the two groups. For both samples, threonine (T) and serine (S) were significantly overrepresented, meaning that most of the identified urinary N-glycosites were typical sites. Moreover, N-glycosites that matched with N-*x*-T (88% CTRL, 82% ccRCC) occurred more frequently than those that matched with N-*x*-S (71% CTRL, 69% ccRCC) (Figure 1a). The differences between controls and

affected subjects emerged by considering the sequence motif around the N-glycosites. For the ccRCC samples, we observed a general decrease of the frequency of the N within the consensus regions (-4%) and a corresponding increase of the frequency within the non-consensus regions (+1.2%) (Figure 1a-d).

Subsequently, we investigated the ccRCC samples cohort according to the pT feature, considering separately the two sub-groups pT1 and pT3. The N-glycosites mapping analysis showed a global decrease of the N-glycosylation levels for the pT1 compared to CTRL group, which became particularly significant in case of the consensus sequence N-x-T. In fact, we observed a reduction of the 7% of the N-modification occupancy site compared to the controls (Figure 1b-d). Considering the pT3 glycoproteome, we detected, also in this case, a decrease of glycosylation frequency for the motif N-x-T (-6%), but conversely from the pT1, the raising of the N-glycosylation levels, regarding the non-consensus sequence (1.6%) and the consensus sequence N-x-S (2%) (Figure 1c-d).

a)

	N-glycosylation frequence,%		
	CTRL	ccRCC	Δ
N / N tot	21	21	/
N in *ncSeq* / *ncSeq* tot	5,6	6,8	1,2 ↑
N in *cSeq* / *cSeq* tot	81	77	-4 ↓
N-*x*- T / N-*x*- T tot	88	82	-6 ↓
N-*x*- S / N-*x*- S tot	71	69	-2 ↓

b)

N-glycosylation frequence, %			
CTRL	pT1	Δ	
21	19	-2 ↓	N / N tot
5,6	6,4	-0,8 ↓	N in *ncSeq* / *ncSeq* tot
81	75	-6 ↓	N in *cSeq* / *cSeq* tot
88	81	-7 ↓	N-*x*- T / N-*x*- T tot
71	66	-5 ↓	N-*x*- S / N-*x*- S tot

c)

N-glycosylation frequence, %		
CTRL	pT3	Δ
21	23	2 ↑
5,6	7,2	1,6 ↑
81	79	-2 ↓
88	82	-6 ↓
71	73	2 ↑

d)

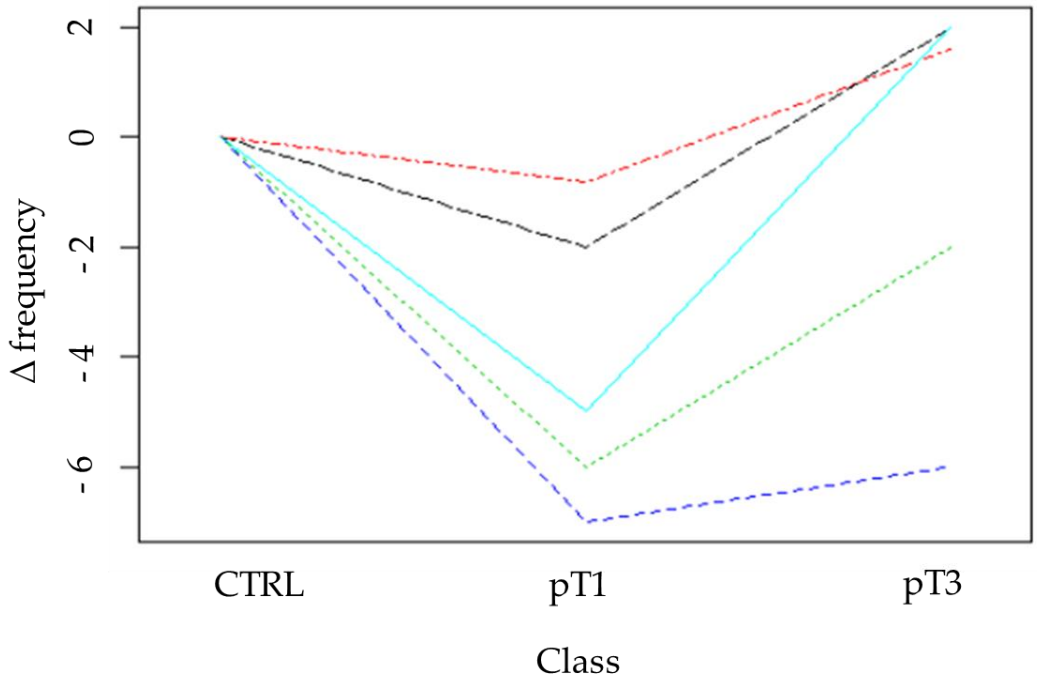


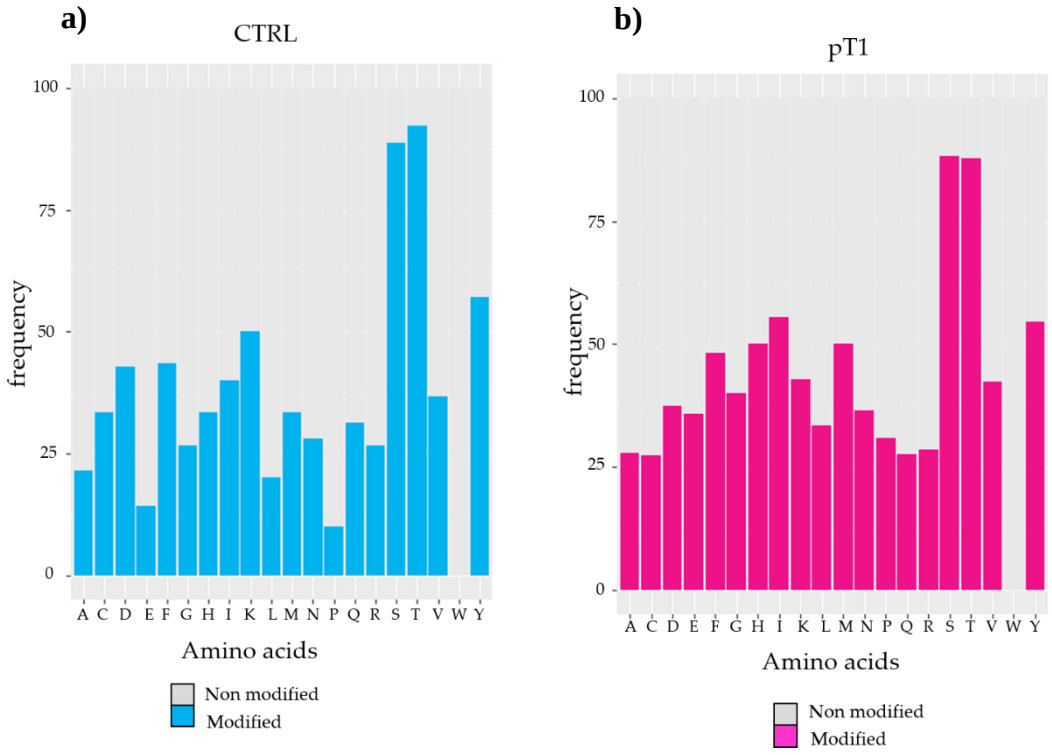
Figure 1. Comparison of the N-glycosylation frequency between CTRL and ccRCC (a), CTRL and pT1 (b) and CTRL and pT3 (c) cohorts in the urinary N-glycoproteome.
d) The trend distribution of the glycomodified asparagines within the whole glycoproteome (red line), the non-consensus regions (black line), the consensus regions (light-blue line), the N-*x*-T regions (green line), the N-*x*-S regions (blue line) estimated in CTRL, pT1 and pT3 groups.
LEGEND: **ncSeq**: non-consensus sequence; **cSeq**: consensus sequence; **N**: asparagine glycomodified; **N**: asparagine non glycomodified.

These results were also confirmed when the peptides commonly identified in all the three samples cohorts were evaluated. To further characterise the captured glycosylated peptides, we determined in details the amino acid composition of the glycosylation motif. We calculated the frequency of single amino acidic residues present in the modified triplet N-*x*-*x* compared to its corresponding non-modified triplet N-*x*-*x*.

It was immediately clear that the consensus sequences were most frequently modified in both early and advanced cancer conditions, as expected. Moreover, more than 97% of the N-glycans are attached to the consensus motifs in mammalian glycoproteins [25] and considering the specific role and/or subcellular localisation

that this modification confers to the proteins, it proper that these evidence remain constant, even in presence of tumour status.

However, some remarks can be also formulated regarding the frequency of non-canonical motif. In the pT1 group, for example, it was detectable a preference of the other amino acids, diverse from the canonical T/S, at the +3 position of the consensus region. Specifically, we recorded a considerable increase of Glutamic Acid (E), Glycine (G), Histidine (H), Leucine (L), Isoleucine (I) and Methionine (M) as third amino acid in the triplet contained the N-modified site, compared to the control group. At the same time, we also highlighted a mild inflection as glycosylated site for the non-consensus sequence containing: Cysteine (C), Aspartic Acid (D), and Lysine (K) (Figure 2a-b).



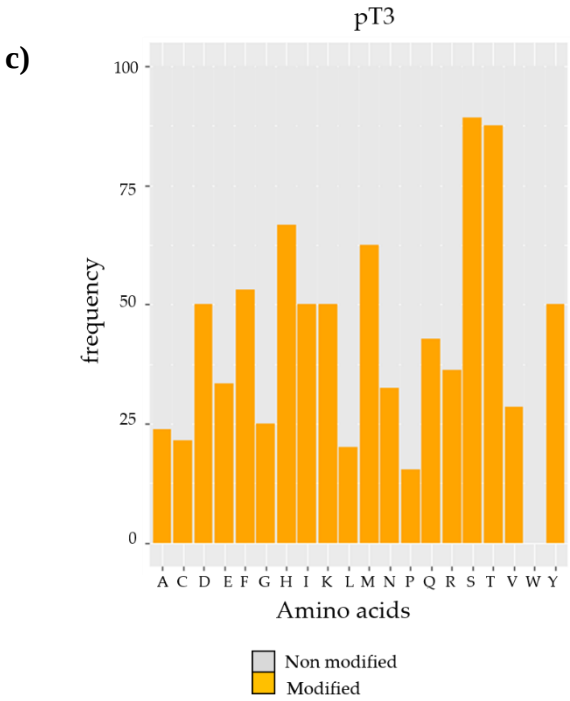


Figure 2. The frequencies of each amino acid in the 3+ position in the triplet N-*x*-*x* within the glycoproteome versus the frequency of itself within the whole proteome calculated for the CTRL (a), pT1 (b) and pT3 (c) groups. These frequencies were measured taking into account only the identify peptides present in all three subcohorts.

Also in the pT3 group, a general expansion of the N-glycosylation level regarding non-consensus sequences was observed. In this case, the non-canonical triplets included the presence at the +3 position of the amino acids: Aspartic Acid (D), Glutamine (Q) and Tyrosine (Y), in addition to those present in the pT1 group (Figure 2 b-c). However, for some sites involved in non-consensus regions, it was possible to detect a reversion to the glycosylation levels typical of control subjects (i.e. Alanine (A), G, L) (Figure 2 a-c).

The study of the occurrence of the N-glycosylation within non-consensus sequences is very fascinating. Indeed, the N-glycan attachment to the unconventional sites, as well as to the conventional ones, play an important role in the attribution to the proteins of their biological functions, crucial for the disease conditions onset and host defense mechanisms activation. Correlating the occurrence of the unconventional glycosylation with ccRCC early and advanced conditions, our

findings open questions about the mechanism that triggers such unusual modification.

2.4 Characterisation of urinary N-glycoproteins: Identification

Once characterised the number and the distribution of the N-glycomodified sites, the three groups of tryptic peptides analysed by nUHPLC-MS/MS were submitted to further investigations.

A total of 659 different N-glycoprotein species were identified with FDR Peptide-Spectrum Matches of 1% and at least one unique peptide (Supplementary Table 1). Comparing the list of the identified proteins, it was arisen that exactly half of them was shared among the three groups (50.4%), indicating that exists a glycoproteic core of urinary proteome that does not change, even in the presence of the tumour. When we considered the identified glycoproteins individually, we recorded: 11% related to pT1, 8% to pT3 and 10.7% to controls (Figure 3).

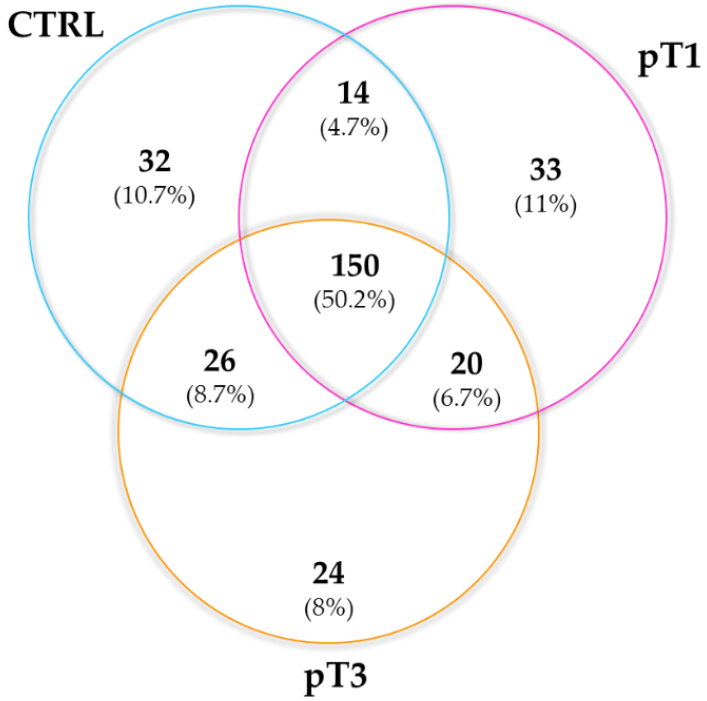


Figure 3. Distribution of the urine sample N-glycoproteins identified by shotgun LC-MS illustrated by a Venn diagram.

2.5 Functional Analysis of N-glycoproteins identified

Furthermore, glycoproteins identified were subjected to bioinformatic analysis. To obtain an overview of the primary subcellular compartments, molecular functions and biological process of N-linked glycoproteins, we used Cytoscape tool (https://cytoscape.org/ last access March 2019) to perform Gene Ontology (GO) function enrichment analysis. For all the three classes, the top enriched clusters among the subcellular compartments were extracellular region, lysosome, membrane, endocytic and extracellular vesicle, which are generally involved in the processes of N-glycoprotein synthesis and transport.

In addition, blood components and plasma lipoprotein particles were significantly enriched since the proteins related to these classes are usually the most abundant isolated from the urine [26], [27], especially in case of kidney impairment and pathological conditions (Figure 4a).

Furthermore, many molecular functions that are known to be performed by N-glycoproteins were enriched. As shown in Figure 4, the top clusters affected the binding activity (carbohydrate, chaperone, immune component), the extracellular matrix remodeling and the receptor activity (Figure 4b).

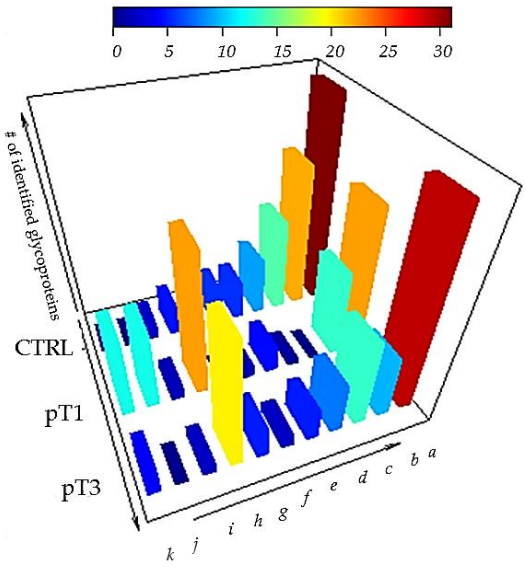
Finally, the major overrepresented biological processes included regulation of vesicle transport, tissue homeostasis, extracellular structure organization and remodeling, aminoglycan and oligosaccharide metabolisms, endocytosis and blood vessel morphogenesis (Figure 4c).

Obviously, the principal emerged clusters for each GO class differed in terms of number of enriched proteins in the controls compared to ccRCC groups. It was possible to observe the substantial increment of the N-glycoproteins as component of immunoglobulin complex for both early and advanced stages and the overrepresentation of N-glycoproteins related to cellular membrane and immune secretory vesicles in case of the pT1 status. For the process of vesiculation, in particular, the pT1 stage seemed to lose in terms of the endocytic activity and gain in the corresponding secretory activity. This evidence was confirmed either by GO terms, cellular component and biological process (Figure 4a-c). As expected, the

proteins involved in the immunological pattern were largely enriched, supporting the well-known role of immune response in the ccRCC onset and progression [28]. Moreover, it was well known that immune response plays a fundamental role in the ccRCC onset.

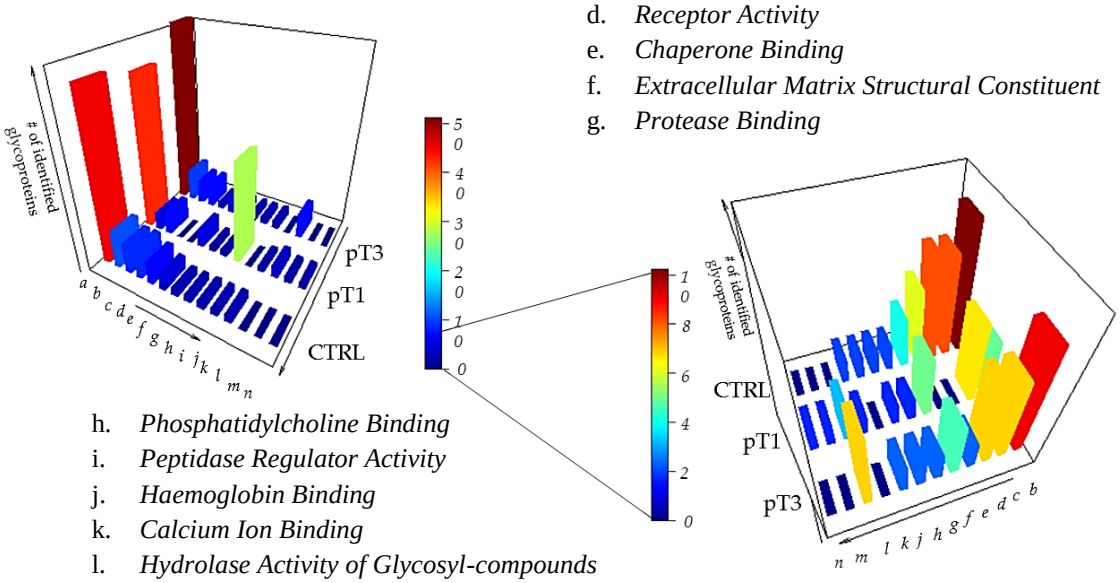
a) Cellular Component

a. *Extracellular Matrix Component*
b. *Lysosome*
c. *Blood Microparticle*
d. *Endocytic Vesicle*
e. *Plasma Lipoprotein Particle*
f. *Golgi and ER lumen*
g. *Extracellular Vesicle*
h. *Immunoglobulin Complex Circulating*
i. *Vacuolar part*
j. *Membrane Component*
k. *Immune Secretory Vesicle*



b) Molecular Functions

a. *Peptidase Activity*
b. *Glycosaminoglycan and Proteoglycan Binding*
c. *Immune Component Binding*
d. *Receptor Activity*
e. *Chaperone Binding*
f. *Extracellular Matrix Structural Constituent*
g. *Protease Binding*



h. *Phosphatidylcholine Binding*
i. *Peptidase Regulator Activity*
j. *Haemoglobin Binding*
k. *Calcium Ion Binding*
l. *Hydrolase Activity of Glycosyl-compounds*
m. *Endoribonuclease Activity*
n. *Glycolipid Binding*

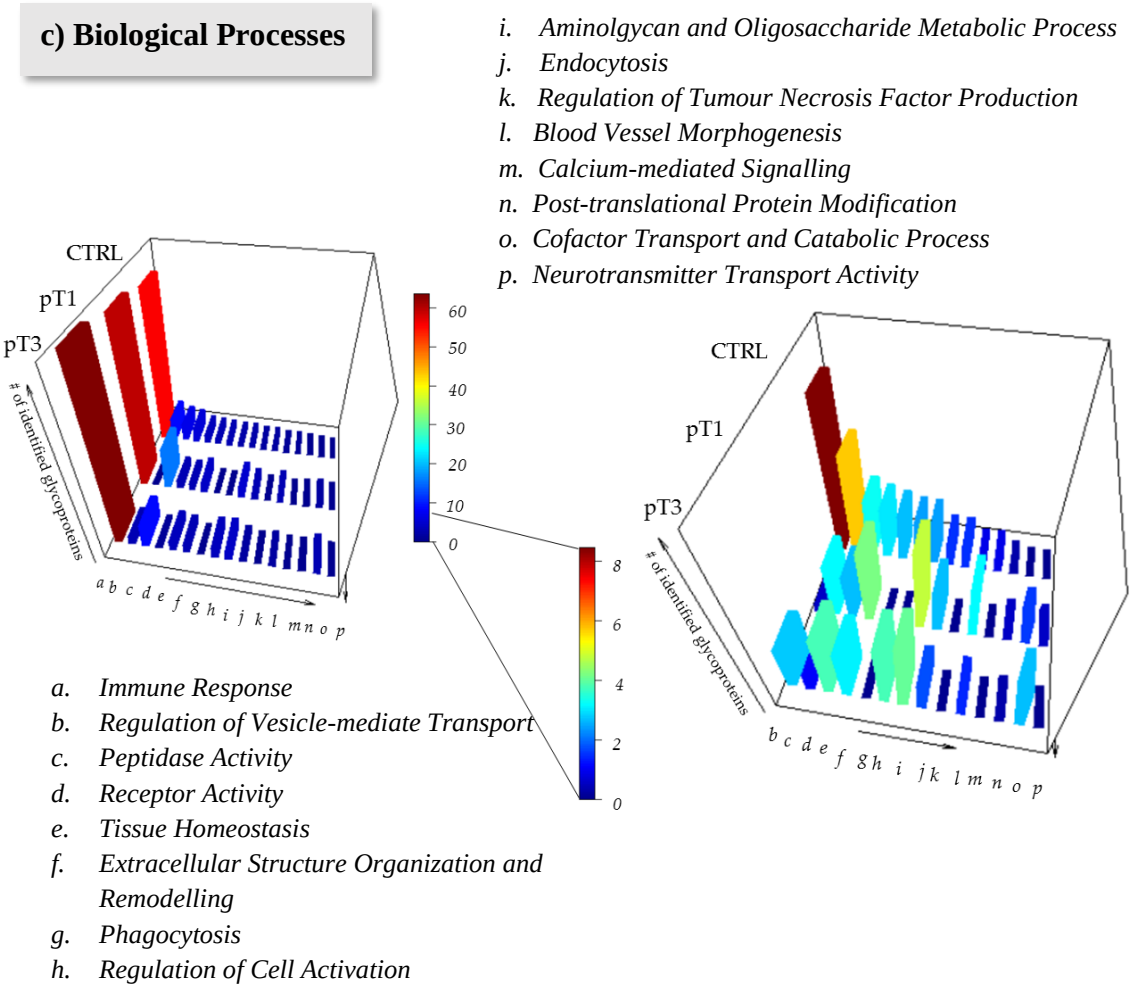


Figure 4. Gene ontology analysis of the proteins identified from all sample groups by LC-MS. Classification according to **(a)** cellular component, **(b)** molecular function, and **(c)** biological process.

2.6 Quantitative Analysis of the differentially expressed N-glycoproteins

A label-free semi-quantitation method was used to quantify glycoproteins present in different amounts in ccRCC and controls urine samples. Differential expressed proteins were selected on the basis of the fold change observed among the CTRL, pT1 and pT3 groups, after normalisation. Proteins that were detected in two analytical replicates and exhibited ≥ 2 or ≤ 2 fold changes, $p \leq 0.05$, were selected as potential candidate markers. As shown in Figure 5, the majority (approximately 70%) of potential candidate markers were upregulated in the disease proteome. These included lipid transport and metabolic process proteins (e.g., FOLR1, LRP2,

PLTP, CATD), immune system process proteins (e.g., CD97, HPT, A1AT, CD63, PTGDS, CD276), control and maintenance of the cellular shape proteins (e.g., COCH, FINC).

a)

pT1 *vs* CTRL		
Total varied proteins: **17**		
Protein	*p-value*	*Fold change*
UP: 5		
● COCH	0.001	4.78
● CD97	0.005	3.03
● FOLR1	0.004	2.29
● LRP2	0.001	2.22
● P3IP1	0.01	2.18
UP: 12		
● APOB	0.004	-34.3
● FINC	0.001	-24.5
● HPT	0.000	-14.70
● CFAH	0.001	-11.61
● PLTP	0.001	-9.76
● A1AT	0.001	-6.13
● TRFE	0.000	-3.59
● CD63	0.004	-3.47
● CERU	0.001	-3.12
● IGHA2	0.000	-2.60
● HEMO	0.001	-2.55
● APOH	0.005	-2.52

pT3 *vs* CTRL		
Total varied proteins: **17**		
Protein	*p-value*	*Fold change*
UP: 10		
● COCH	0.013	4.77
● CD97	0.000	3.82
● LAMP2	0.000	3.13
● AMBP	0.002	3.09
● P3IP1	0.001	2.60
● PTGDS	0.000	2.47
● WFDC2	0.000	2.44
● ZA2G	0.000	2.41
● CD276	0.000	2.3
● CATD	0.000	2.12
DOWN: 7		
● CERU	0.000	-2.06
● HPT	0.000	-2.41
● PERM	0.000	-2.52
● PLTP	0.000	-3.30
● FINC	0.000	-6.04
● CFAH	0.001	-6.16
● APOB	0.001	-9.59

a)

pT3 *vs* pT1		
Total varied proteins: **14**		
Protein	*p-value*	*Fold change*
UP: 13		
● KLK1	0.000	6.87
● HPT	0.003	6.10
● A1AT	0.002	4.18
● FINC	0.020	4.05
● TRFE	0.002	3.32
● PLTP	0.034	2.96
● AMBP	0.001	2.68
● ZA2G	0.000	2.54
● APOH	0.002	2.44
● PPAP	0.004	2.22
● CD63	0.002	2.11
● A1AG2	0.000	2.07
● CD276	0.002	2.06
DOWN: 1		
● PERM	0.006	-3.32

b)

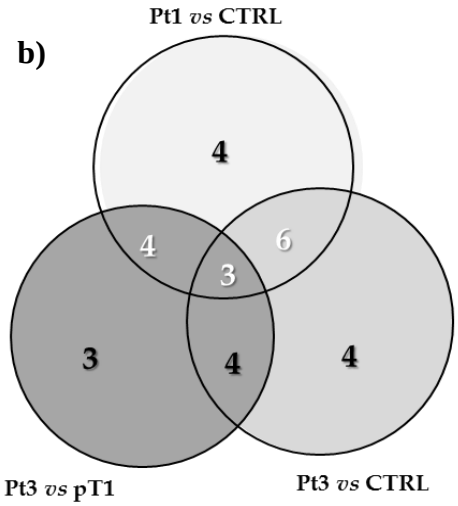


Figure 5. Lists of N-glycosylated proteins varied among the three clinical groups (a). Number of glycosylated proteins significantly varied in each group represented by Venn diagram (b).

Upon further exploitation, we established disease associations and glycosylation status. Considering the abundance levels of the proteins resulted significantly dysregulated, we noted the existence of a panel of glycosylated proteins with a specific abundance trend, directly related to the stage progression. Moreover, we detected 9 proteins of particular interest. These resulted differently expressed in the comparison pT1 *vs* CTRL and pT3 *vs* CTRL; 3 proteins shown an increasing pattern, while 6 proteins presented a decreasing trend. It is noteworthy that in the down regulated group, all the 6 glycoproteins showed a considerable decrease in the pT1, followed by a reversion to the levels of healthy subjects in the pT3 stage (Figure 6).

These N-glycoproteins could be characteristic of the early stage of tumourigenesis, constituting in that way a *glycosignature* of tumour condition. Since ccRCC is a heterogeneous disease, it seems unlikely that only one biomarker will be uniformly elevated. Therefore, a panel of biomarkers may provide more accurate diagnosis than any given single marker.

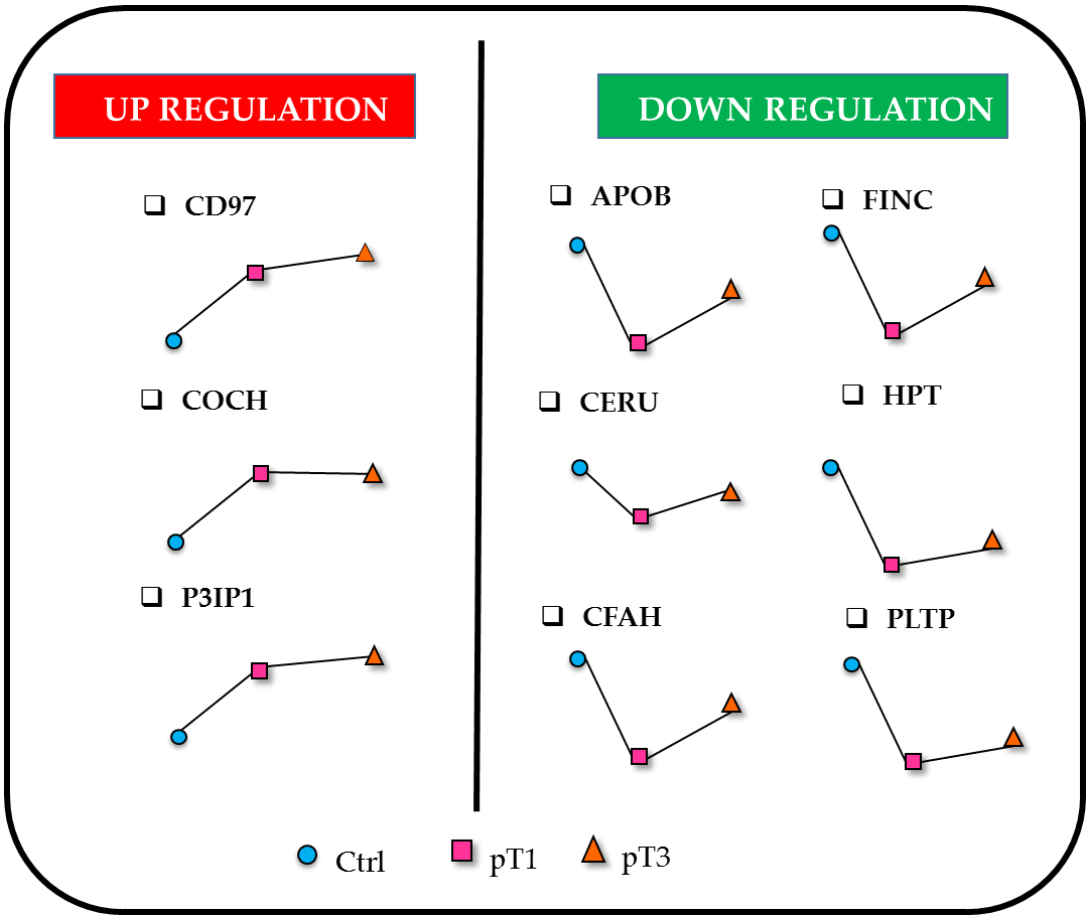


Figure 6. Panel of urinary N-glycosylated proteins with the relative abundance trend of each one in the three different sub-cohorts.

This panel of putative urinary glycomarkers certainly included Haptoglobin (HPT), a glycosylated protein, described as involved in various aspects of the acute phase response, receptor mediated endocytosis, homeostasis, and in the positive

regulation of cell apoptosis [29]. It is also one of the most commonly reported proteins affected by oligosaccharide modifications in human malignancies, including ovarian [30], liver [31], colon [32], and pancreatic cancers [33]. However, its glycosylation status is different from one type of cancer to another. Sandim et al in their proteomic study highlighted the over-secretion of HPT in the urine of ccRCC patients compared to controls [34]. However, our results were not consistent with this evidence; in fact, we observed a 14.7-fold abundance decrease of HPT in pT1 urine compared to CTRL urine (Figure 5a). Moreover, its level risen again in the pT3 urine (fold change= -2.4, p-value=0.000), but still remained under the physiological level recorded in the CTRL subjects (Figure 5a). We could explain this divergence hypothesising that the HPT detected in our analysis represented the N-glycosylated isoform, that may play a different role in the process of onset and spreading of tumour compared to the non-glycosylated one. As already described for other cancer types, also for the ccRCC the HPT could bear different kind of N-glycans and our study could indirectly give an idea of the biological complexity that accompanies this glycosylation presence. Future structural studies regarding the in depth characterisation of the glycomodification associated to this protein may provide more information in our understanding of the low amounts presence of HPT N-glycosylated isoform in the urine of affected subjects and its role in ccRCC spreading.

A similar consideration can be done for the protein Fibronectin (FINC), an extracellular matrix glycoprotein, that plays important roles in the various different stages of wound healing, with its main function being cellular adhesion and the mediation of cell migration [35]. It is largely reported that high FINC content in the extracellular space facilitates tumour cell metastasis, by promoting morphological changes and improving the motility and migratory ability of cancer cells [36]. A number of studies proposed the FINC overexpression as unfavorable prognostic indicator for diverse cancer types, such as breast [37] and pancreatic cancer [38], nasopharyngeal [39] and neck squamous cell carcinomas [40]. Therefore, this

protein seems to promote cell growth and migration also in case of RCC. Moreover, the interesting study by Kondisette et al., investigating the relationship of FINC with the clinical stage of tumour, pinpointed the increase of FINC expression levels in the RCC tissue compared to controls, that was particularly significant for the early stage condition [41]. However, in our study we did not find the same correlation; on the contrary, we observed, in a mirror fashion, the tremendous decrease of the FINC glycosylated isoform levels in the pT1 group (fold change= -24.5, p value=0.001), followed by an increase in the pT3 group (fold change= -6.04, p value=0.000), still remaining below the physiological value observed in the controls subjects. Also in this case, the hypothesis that could explain our results is that exists a sort of balance between the FINC N-glycosylated form and the non-glycosylated one: when one increases its presence, the other decreases it. According to this, we could speculate that the FINC protein detected in our analysis is an N-glycosylated isoform, whose expression could not be necessary in the preliminary phase of cancer event (pT1 status), differently from the corresponding non-glycosylated form, that is reported as a central player in the early process of ccRCC tumourigenesis [41].

3. MATERIALS AND METHODS

3.1 Urine collection and processing

45 patients, divided in ccRCC affected and non-affected, were enrolled between 2011 and 2016 at San Gerardo hospital (Monza, Italy). All participants gave their information consent before sample collection. Study protocols and procedures were approved by the local ethic committee (Comitato Etico Azienda Ospedaliera San Gerardo, Monza, Italy, BPCR 24-02-2011) and analysis were carried out in agreement with the Declaration of Helsinki.

Second morning midstream urine before total or partial nephrectomy (in case of ccRCC patients) was collected in sterile urine tubes. The collected samples were cleared by centrifuging at 4°C at 1000 *xg* for 10 minutes to remove cells and debris and stored at -80 °C until the day of the analysis.

3.2 Sample Population

Fifteen urine samples of ccRCC patients at pT1 stage, fifteen urine samples of ccRCC patients at pT3 stage and fifteen urine samples of patients affected by renal calculus as controls were included in the study.

Three protein pools were prepared, one representative of the CTRLs and the other of the RCC populations, divided by stage pT1 and pT3. Pooling was necessary for this discovery study due to the limited amount of samples and also to reduce variability among patients.

3.3 Urinary proteins digestion

2 mL of the stored supernatant were defrosted through agitation by vortexing and the protein content was concentrated by EtOH precipitation. Urinary proteins were precipitated by nine volumes of cold 90% ethanol and pelleted at 3'500 *xg* for 30 minutes [42]. After drying, proteins were dissolved in bidistilled water, and protein concentration was assessed by BCA assay (Microplate BCA ™ protein Assay Kit, Thermo Scientific), using BSA as standard.

About 400 µg of urine protein samples were treated following the FASP protocol, as already described [43]. Briefly, proteins were first reduced by incubation with 50mM DL-Dithiothreitol (Sigma Aldrich, Switzerland) and alkylated for 30 minutes with Iodoacetamide 100mM (Sigma Aldrich, Switzerland). Then, they were digested overnight on 30kDa filters (Amicon Ultra-500 µg 30 kDa, Millipore) 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 lyophilized. The resulting peptides were resuspended in steril-filtered water (Sigma Aldrich, Switzerland) and their concentration was determined by Nanodrop spectrophotometer (Thermo Scientific™).

3.4 Urinary N-glycopeptides enrichment and deglycosylation

After the whole urinary proteome digestion, the N-glycopeptides were enriched by binding to lectins on the top of a filter, used a FASP-based method (so-called N-

Glyco-FASP) described for the first time by Zielinska et al [25]. Briefly, a solution containing a pool of three different types of lectins, concavalin A (ConcA), wheat germ agglutinin (WGA) and agglutinin RCA_{120} (Sigma Aldrich, Switzerland) was mixed with approximately 100 µg of digested peptides, with a mass proportion of 1:3. The mixture was loaded on new 30kDa filter unit (Amicon Ultra-500 µg 30 kDa, Millipore) and the incubation lectins-glycopeptides was performed at 4°C for 2 hours. Then, the unbound peptides were eluted by 4 steps of centrifugations at 14'000 *xg*, each one of 10 min. The captured peptides underwent to overnight deglycosylation with PNGase F (Roche). The deglycosylated peptides were eluted with another serial of centrifugation steps (14'000 *xg*, 10 min), lyophilized and stored at -20 °C, until the following MS analysis.

3.5 Mass Spectrometric Analysis

An aliquot of 10 µL containing 2 µg of peptides was injected into a Dionex UltiMate 3000 rapid separation (RS) LC nano-system (Thermo Scientific, Germany).

HPLC analysis was 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 hours 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 minutes. The HPLC instrumentation was on-line interfaced with an Impact HD™ Ultra High Resolution-QqTOF (Bruker Daltonics, Germany), equipped with a nanoBoosterCaptiveSpray™ ESI source (Bruker Daltonics, Germany), in order to enhance the MS analysis.

The instrument worked in Data-Dependent-Acquisition mode, automatically alternating a complete MS scan to the acquisition of MS/MS spectra. The MS/MS

data were acquired by targeting precursors (300-2000 m/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 scans 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, based on 10 mM sodium formiate cluster solution (15 minutes before each run). Compass DataAnalysis v4.1 software (Bruker Daltonics, Germany) was used to calibrate, deconvolute and convert the acquired raw data prior to protein identification and quantification.

3.6 Data Processing

3.4.1 Protein identification

Mascot (v 2.4.1, Matrix Science, UK) was used for protein identification. Trypsin was chosen as the enzyme and the number of missed cleavages was set to 1. The peptide charge was set to 2+ and 3+, and the peptide tolerance and MS/MS tolerance were 20 ppm and 0.05 Da, respectively. Cysteine carbamidomethylation was set as fixed modification, while methionine oxidation and asparagine deamidation were used as variable modifications. Swiss-prot (accessed May 2017; 555.100 sequences; 198.754.198 residues) was used as database. The maximum false discovery rate (FDR) for peptide spectral match was set to 1%, using percolator algorithm and a minimum of one sequence-unique peptides was required for identification. Proteins of interest were analysed for cellular component, molecular functions and biological processes with ClueGO v2.5, Clupedia v1.5 and the Cytoscape tools [44].

3.4.2 Bioinformatics and Statistical analysis

Progenesis QI for proteomics (Nonlinear Dynamics, Newcastle, UK) was used for the label-free protein quantitation. 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 minutes 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. The Anova test was used to compare the different sample cohorts in terms of N-glycoproteins abundance and a post-hoc Tukey test, with Benjamini & Hochberg adjustment, was applied for pairwise comparisons. For the N-mapping analysis, all of the potential N-glycosylation sites were predicted applying an in house made software, able to recognise and count all the N deamidated present in sequences obtained from Uniprot database.

4. CONCLUSION

Despite recent developments in understanding glycosylation processes, there is still a gap in comprehend the full aspects of the glycoproteome in ccRCC cancer. Glycobiology enabled discoveries in many biochemical and medical studies and has contributed to the deciphering of various human diseases. Glycans are involved in numerous fundamental processes occurring during cancer, such as tumour cell dissociation and invasion, cell-matrix interactions, tumour angiogenesis, immunomodulation and metastasis formation. The roles of glycans in cancer have been emphasised as alterations in glycosylation which regulate the development and progression of cancer.

To the best of our knowledge, our study would be the first that investigates in depth the urinary N-glycoproteome in patients with early and advanced ccRCC conditions compared to non-affected subjects. We performed the evaluation of the urine glycoproteins content both qualitatively (N-glycosites distribution) and quantitatively (N-glycoprotein significant varied). Through a non-invasive urine-based liquid biopsy approach, we firstly distinguished a specific glycosites distribution related to each clinical group, pointing out a major presence of unconventional glycosylated motifs and the corresponding inflection in term of conventional site occupancy (N-*x*-T primarily) in the urinary glycoproteins content of the ccRCC patients court. Such diversification in the N-glycosylation sequons suggests that the universally acceptable N-glycosilation motif N-*x*-T/S is not necessary a general rule for cancer status, as we demonstrated in our analysis

conducted on the urine glycoprotein content of ccRCC patients. It is well known that some of the unconventional glycosylations are carried out by unique classes of glycosyltransferase enzymes. Could be very interesting in a future study verify the existence of these *special enzymes* and explore the mechanisms that they adopt to recognise both consensus e non-consensus motifs, performing a targeted study in ccRCC cells lines or tissue biopsies.

Therefore, we identified a group of urinary N-glycoproteins with a specific abundance trend, that could constitute a specific disease-related glycosignature, able to followed and underlined not only the transition from controls to ccRCC cancer status, but also to differentiate between early and advanced stage. These dysregulated proteins may have important implications in the pathogenesis of this renal malignancy, probably different from the corresponding unmodified isoform; this information can be useful for future diagnostic applications. Therefore, our data encourage further investigations in this direction; it would be interesting, for example, the in depth characterisation of the N-glycans attached to the deregulated proteins identified, in order to determine the precise role of this modification in ccRCC.

References

[1] K. Brown Chandler, C. E. Costello, and N. Rahimi, "Glycosylation in the Tumor Microenvironment: Implications for Tumor Angiogenesis and Metastasis," *Cells*, vol. 8, no. 6, p. 544, 2019.

[2] M. Aebi, "N-linked protein glycosylation in the ER," *Biochim. Biophys. Acta - Mol. Cell Res.*, vol. 1833, no. 11, pp. 2430–2437, 2013.

[3] K. Ohtsubo and J. D. Marth, "Glycosylation in Cellular Mechanisms of Health and Disease," *Cell*, vol. 126, no. 5, pp. 855–867, 2006.

[4] E. S. Lander *et al.*, "Initial sequencing and analysis of the human genome.," *Nature*, vol. 409, no. 6822, pp. 860–921, Feb. 2001.

[5] N. Taniguchi and Y. Kizuka, "Glycans and cancer: Role of N-Glycans in cancer biomarker, progression and metastasis, and therapeutics," in

Advances in Cancer Research, 2015.

[6] K. S. Lau and J. W. Dennis, "N-Glycans in cancer progression," *Glycobiology*, vol. 18, no. 10, pp. 750–760, 2008.

[7] I. Häuselmann and L. Borsig, "Altered tumor-cell glycosylation promotes metastasis," *Front. Oncol.*, vol. 4 MAR, no. February, pp. 1–15, 2014.

[8] L. Oliveira-Ferrer, K. Legler, and K. Milde-Langosch, "Role of protein glycosylation in cancer metastasis," *Semin. Cancer Biol.*, vol. 44, pp. 141–152, 2017.

[9] S. Gilgunn, P. J. Conroy, R. Saldova, P. M. Rudd, and R. J. O'Kennedy, "Aberrant PSA glycosylation - A sweet predictor of prostate cancer," *Nat. Rev. Urol.*, vol. 10, no. 2, pp. 99–107, 2013.

[10] D. Ilic, M. M. Neuberger, M. Djulbegovic, and P. Dahm, "Screening for prostate cancer.," *Cochrane database Syst. Rev.*, no. 1, p. CD004720, Jan. 2013.

[11] R. R. Drake, E. E. Jones, T. W. Powers, and J. O. Nyalwidhe, *Altered glycosylation in prostate cancer*, 1st ed., vol. 126. Elsevier Inc., 2015.

[12] M. J. Duffy, S. Shering, F. Sherry, E. McDermott, and N. O'Higgins, "CA 15-3: A prognostic marker in breast cancer," *Int. J. Biol. Markers*, vol. 15, no. 4, pp. 330–333, 2000.

[13] J. W. Choi, B. I. Moon, J. W. Lee, H. J. Kim, Y. Jin, and H. J. Kim, "Use of CA15-3 for screening breast cancer: An antibody-lectin sandwich assay for detecting glycosylation of CA15-3 in sera," *Oncol. Rep.*, vol. 40, no. 1, pp. 145–154, 2018.

[14] B. W. T. Yin and K. O. Lloyd, "Molecular cloning of the CA125 ovarian cancer antigen: Identification as a new mucin, MUC16," *J. Biol. Chem.*, vol. 276, no. 29, pp. 27371–27375, 2001.

[15] M. J. Duffy, "Carcinoembryonic antigen as a marker for colorectal cancer: Is it clinically useful?," *Clin. Chem.*, vol. 47, no. 4, pp. 624–630, 2001.

[16] G. Saito, S. Sadahiro, K. Okada, A. Tanaka, T. Suzuki, and A. Kamijo, "Relation between Carcinoembryonic Antigen Levels in Colon Cancer

Tissue and Serum Carcinoembryonic Antigen Levels at Initial Surgery and Recurrence," *Oncol.*, vol. 91, no. 2, pp. 85–89, 2016.

[17] C. G. Moertel, T. R. Fleming, J. S. Macdonald, D. G. Haller, J. A. Laurie, and C. Tangen, "An Evaluation of the Carcinoembryonic Antigen (CEA) Test for Monitoring Patients With Resected Colon Cancer," *JAMA J. Am. Med. Assoc.*, vol. 270, no. 8, pp. 943–947, 1993.

[18] S. Hammarstrom, "The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues.," *Semin. Cancer Biol.*, vol. 9, no. 2, pp. 67–81, Apr. 1999.

[19] M. T. Davis *et al.*, "Towards defining the urinary proteome using liquid chromatography-tandem mass spectrometry. II. Limitations of complex mixture analyses.," *Proteomics*, vol. 1, no. 1, pp. 108–117, Jan. 2001.

[20] W. Sun *et al.*, "Human urine proteome analysis by three separation approaches.," *Proteomics*, vol. 5, no. 18, pp. 4994–5001, Dec. 2005.

[21] Arivusudar Marimuthu, "A Comprehensive Map of the Human Urinary Proteome," *J. Proteome Res.*, vol. 23, no. 1, pp. 1–7, 2011.

[22] F. Desiere *et al.*, "The PeptideAtlas project.," *Nucleic Acids Res.*, vol. 34, no. Database issue, pp. D655-8, Jan. 2006.

[23] R. Kawahara, J. Saad, C. B. Angeli, and G. Palmisano, "Site-specific characterization of N-linked glycosylation in human urinary glycoproteins and endogenous glycopeptides," *Glycoconj. J.*, vol. 33, no. 6, pp. 937–951, 2016.

[24] Different collaborators, "Comprehensive Molecular Characterization of," vol. 499, no. 7456, pp. 43–49, 2014.

[25] D. F. Zielinska, F. Gnad, J. R. Wiśniewski, and M. Mann, "Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints," *Cell*, vol. 141, no. 5, pp. 897–907, 2010.

[26] V. Mooser, M. C. Seabra, M. Abedin, K. T. Landschulz, S. Marcovina, and H. H. Hobbs, "Apolipoprotein(a) kringle 4-containing fragments in human urine. Relationship to plasma levels of lipoprotein(a).," *J. Clin. Invest.*, vol.

97, no. 3, pp. 858–864, Feb. 1996.

[27] C. Bolenz, B. Schröppel, A. Eisenhardt, B. J. Schmitz-Dräger, and M.-O. Grimm, “Abklärung der Hämaturie,” *Dtsch. Aerzteblatt Online*, vol. 25, no. 3, pp. 127–135, 2018.

[28] Z.-H. Tian, C. Yuan, K. Yang, and X.-L. Gao, “Systematic identification of key genes and pathways in clear cell renal cell carcinoma on bioinformatics analysis,” *Ann. Transl. Med.*, vol. 7, no. 5, pp. 89–89, 2019.

[29] W. Dobryszycka, “Biological functions of haptoglobin--new pieces to an old puzzle.,” *Eur. J. Clin. Chem. Clin. Biochem.*, vol. 35, no. 9, pp. 647–654, Sep. 1997.

[30] T. Liu *et al.*, “Human plasma N-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry.,” *J. Proteome Res.*, vol. 4, no. 6, pp. 2070–2080, 2005.

[31] J. Zhu *et al.*, “Analysis of serum haptoglobin fucosylation in hepatocellular carcinoma and liver cirrhosis of different etiologies,” *J. Proteome Res.*, vol. 13, no. 6, pp. 2986–2997, 2014.

[32] S. Y. Park *et al.*, “N-glycosylation status of β-haptoglobin in sera of patients with colon cancer, chronic inflammatory diseases and normal subjects,” *Int. J. Cancer*, vol. 126, no. 1, pp. 142–155, 2010.

[33] Z. Lin *et al.*, “Mass spectrometric assay for analysis of haptoglobin fucosylation in pancreatic cancer,” *J. Proteome Res.*, vol. 10, no. 5, pp. 2602–2611, 2011.

[34] V. Sandim *et al.*, “Proteomic analysis reveals differentially secreted proteins in the urine from patients with clear cell renal cell carcinoma,” *Urol. Oncol. Semin. Orig. Investig.*, vol. 34, no. 1, pp. 5.e11-5.e25, 2016.

[35] W. S. To and K. S. Midwood, “Plasma and cellular fibronectin: Distinct and independent functions during tissue repair,” *Fibrogenes. Tissue Repair*, vol. 4, no. 1, p. 21, 2011.

[36] H. Kumra and D. P. Reinhardt, “Fibronectin-targeted drug delivery in cancer,” *Adv. Drug Deliv. Rev.*, vol. 97, pp. 101–110, 2016.

[37] Z. Han, H. Cheng, J. G. Parvani, Z. Zhou, and Z. R. Lu, “Magnetic resonance molecular imaging of metastatic breast cancer by targeting extradomain-B fibronectin in the tumor microenvironment,” *Magn. Reson. Med.*, vol. 79, no. 6, pp. 3135–3143, 2018.

[38] K. F. Ludwig *et al.*, “Small-Molecule Inhibition of Axl Targets Tumor Immune Suppression and Enhances Chemotherapy in Pancreatic Cancer.,” *Cancer Res.*, vol. 78, no. 1, pp. 246–255, Jan. 2018.

[39] L. J. Ma *et al.*, “Fibronectin overexpression is associated with latentmembrane protein 1 expression and has independent prognostic value for nasopharyngeal carcinoma,” *Tumor Biol.*, vol. 35, no. 2, pp. 1703–1712, 2014.

[40] S. Gopal *et al.*, “Fibronectin-guided migration of carcinoma collectives,” *Nat. Commun.*, vol. 8, 2017.

[41] S. Kondisetty, K. N. Menon, and G. K. Pooleri, “Fibronectin protein expression in renal cell carcinoma in correlation with clinical stage of tumour,” *Biomark. Res.*, vol. 6, no. 1, pp. 1–6, 2018.

[42] R. S. Lee, F. Monigatti, A. C. Briscoe, Z. Waldon, M. R. Freeman, and H. Steen, “Optimizing sample handling for urinary proteomics,” *J. Proteome Res.*, vol. 7, no. 9, pp. 4022–4030, 2008.

[43] C. Chinello *et al.*, “Proteomics of liquid biopsies: Depicting RCC infiltration into the renal vein by MS analysis of urine and plasma,” *J. Proteomics*, vol. 191, pp. 29–37, 2019.

[44] N. T. Doncheva, J. H. Morris, J. Gorodkin, and L. J. Jensen, “Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data,” *J. Proteome Res.*, vol. 18, no. 2, pp. 623–632, 2019.

CHAPTER 4

Effects of Hematuria on the Proteomic Profile of Urinary Exosomes: Technical Challenges

#Francesca Raimondo[a], #Clizia Chinello[a], Martina Stella[a], **Lucia Santorelli[a]**, Fulvio Magni[a] and Marina Pitto[a]

#equally contributing author

[1] School of Medicine and Surgery, University of Milano – Bicocca

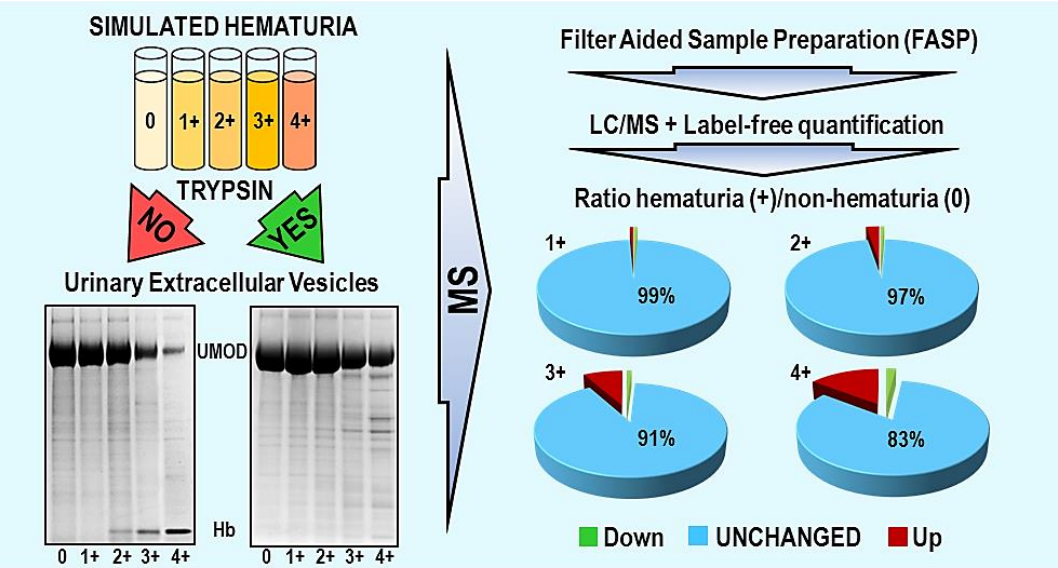
Published: Journal of Proteome Research

2018, 17(8):2572-2580

doi: 10.1021/acs.jproteome.7b00763.

ABSTRACT Hematuria is a common sign of many renal and urologic pathologic conditions and it may affect the proteomic analysis of urinary extracellular vesicles (UEv), nanovesicles released from all cells in contact with the urinary space. This condition hinders UEv based proteomic studies aiming to discover biomarkers. Therefore, we studied the effects of hematuria on the proteome of UEv and introduced a possible solution to reduce its misleading impact. We mimicked hematuria by adding increasing amount of blood to non-affected urine and investigated its effects on UEv isolation, purity, and proteomic composition. We proposed a trypsin treatment able to reduce the impact of hematuria on UEv. The effects of the treatment were investigated by evaluating the UEv proteomic profile, the enrichment of known UEv markers, and by assessing differential protein content by MS-based label-free quantification. Results showed that as the blood contamination increased, it affected both the proteome profile and the yield of UEv isolated from urine. Our treatment with trypsin was able to counteract completely these effects for low/medium levels of hematuria, which are most commonly encountered. This promising finding could lead to the reliable use of hematuria samples for UEv proteomic investigation.

KEYWORDS extracellular vesicles, hematuria, urine, proteomics


SIMULATED HEMATURIA
0
1+
2+
3+
4+
TRYPSIN
NO
YES
Urinary Extracellular Vesicles
UMOD
Hb
0 1+ 2+ 3+ 4+
0 1+ 2+ 3+ 4+
MS
Filter Aided Sample Preparation (FASP)
LC/MS + Label-free quantification
Ratio hematuria (+)/non-hematuria (0)
1+
99%
2+
97%
3+
91%
4+
83%
Down
UNCHANGED
Up

1. INTRODUCTION

Hematuria is a complex condition that can derive from a variety of etiologies, found along the entire genitourinary tract. In fact, it is a common sign of several pathologic conditions such as urolithiasis, urinary tract infection, bladder and kidney malignancies, but also glomerulonephritis.[1,2]

Recently, a promising source of biomarkers for these diseases has been identified in urinary nanovesicles/exosomes (30−150 nm of diameter), released from all cells facing the urinary tract. For simplicity, we will refer to them as urinary extracellular vesicles (UEv) because we are aware of the debate about the complexity of the extracellular vesicle samples and their nomenclature.[3−5] Proteomic analysis of UEv can be a powerful tool for biomarker discovery since their proteome is depleted of the most abundant proteins, such as albumin, and is enriched in renal proteins.[6] Moreover, since the molecular content of extracellular vesicles mirrors their cells of origin composition, UEv are more likely to reflect the cellular processes associated with the pathogenesis of urologic and renal diseases compared, to urine.[7] For this reason, several proteomic studies on UEv have been performed to identify biomarkers for urinary tract diseases, both in experimental and clinical settings. For an exhaustive review, see Moon.[8]

However, the issue of hematuria influence on proteomic analysis of UEv has been only partially addressed. In particular, a paper dealing with UEv use for the search of bladder cancer biomarkers clearly showed that microparticle protein profiles isolated from hematuric urine were altered.[9] Authors suggested that such samples should not be used at all, to avoid the discovery of biomarkers due to blood contamination of urine.

In 2012, a second work on bladder cancer arrived to similar conclusions.[10] However, they speculated that the increasing levels of some proteins in urinary microparticles of cancer patients were caused by blood-derived vesicles. Moreover, authors proposed a possible role for these proteins as noninvasive markers for the surveillance of tumor progression or recurrence, although not useful for the

diagnosis of a specific diseases. Finally, a third paper reported to have found biomarkers able to differentiate IgA nephropathy from thin basement membrane nephropathy[11] in UEv both in presence and in absence of hematuria.

These conflicting results and their different interpretations leave unresolved the issue of possible use of hematuric samples for biomarker discovery in UEv. In this study, we investigated the impact of different levels of hematuria on the proteome of UEv. Moreover, we looked for a possible improvement of the UEv preparation protocol and analysis aimed at removing, or at least reducing, hematuria effects.

2. EXPERIMENTAL SECTION

Milli-Q water was used for all solutions. BCA protein assay, methanol, CAPS, and trypsin from bovine pancreas were from SIGMA Chemical Co. (St. Louis, MO, USA); glycerol was from Merck (Darmstadt, Germany). Hybond-ECL nitrocellulose membrane was from GE (Little Chalfont, Buckinghamshire, UK). NuPAGE SDS-PAGE Gel Electrophoresis System components (mini gels, running and loading buffers, molecular weight markers, and coomassie blue staining) were supplied by Life Technologies (Paisley, Renfrewshire, UK). Antiprotease inhibitor cocktail (Complete) was from Roche (Monza, Italy). Anti-Aquaporin-2 (AQP2) mAb from Cell Signaling Technology (Beverly, MA, USA), anti-Tumor Susceptibility Gene 101 (TSG101) polyclonal antibody (pAb) from Abcam (Cambridge, UK), and anti- Na−K−Cl cotransporter (NKCC2) pAb from SIGMA Chemical Co. (St. Louis, MO, USA). Species-specific secondary peroxidase conjugated antibodies and ECL reagents were from Pierce (Rockford, IL, USA).

2.3 Urine Collection, Hematuria Mimicking, and Trypsin Treatment

Second morning urine samples were collected from healthy volunteers, according to the guidelines provided by the Human Kidney and Urine Proteome Project (HKUPP, http: www. hkupp.org).

Human blood from a healthy control was collected by venipuncture in a whole

blood Vacutainer tube (0.129 M, sodium-citrate). Pooled urines were mixed and divided into equal aliquots. Increasing amounts of exogenous blood (0, 1.4, 6.7, 67, and 133 μL/100 mL of urine) were added to urine to mimic different levels of hematuria (from 0 to 4+ according with the Combur-Test strips, Roche). After 4 h at 37°C,9 urine samples were centrifuged for sediment (P1) removal (10 min at 1000g, 4°C). The supernatant was supplemented with protease inhibitors (Complete, Roche) and stored at −80°C until UEv isolation. From an aliquot of each sample (500 μL out of about 20 mL), urinary proteins were precipitated by nine volumes of cold 90% ethanol and pelleted at 3500g for 30 min.[12] After drying, proteins were dissolved in bidistilled water, and protein concentration was assessed by BCA assay (SIGMA Chemical Co), using BSA as standard.

Before UEv isolation, the remaining stored urine samples were thawed and adjusted to pH 7.4, if needed. Each sample was divided into two 50 mL aliquots: one of them was treated with bovine trypsin dissolved in PBS (1 mg of trypsin: 50 mg of urinary proteins), and the other with PBS alone, for 1 h at 37°C. Figure 1 shows a representative scheme of the protocol.

2.2 UEv Isolation

UEv were prepared by ultracentrifugation[6] according to HKUPP (http:www.hkupp.org). All steps were performed at 20°C. Briefly, treated and untreated urine were centrifuged for 15 min at 17 000g to eliminate large membrane fragments and debris (pellet P). Supernatants were subjected to ultracentrifugation for 70 min at 200 000g: crude UEv pellets were washed in PBS and then suspended in bidistilled water, in

the presence of protease inhibitors. The UEv samples were stored at −80°C until use. Moreover, proteins from urine aliquots after sediment removal (U) and supernatants after 200 000g ultracentrifugation (Sn) were precipitated by 90% ethanol, as described above. Protein concentration was assessed by BCA assay (SIGMA Chemical Co), using BSA as standard, for all the urinary fractions obtained.

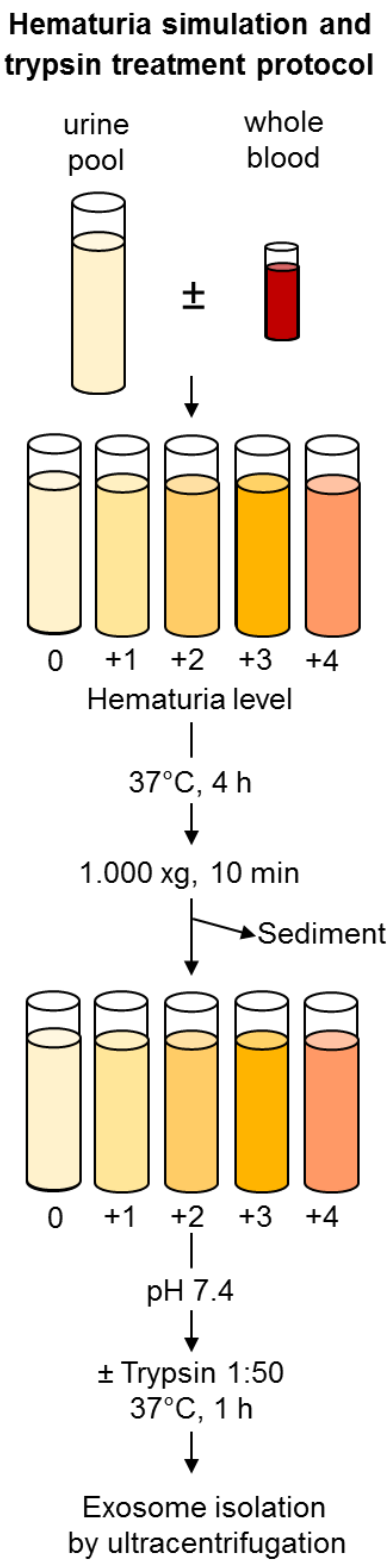


Figure 1. Schematic representation of hematuria simulation and trypsin treatment protocol.

2.3 Electrophoresis and Western Blotting

Protein separation was performed with the NuPAGE electrophoresis system (Life Technologies), using 4−12% NuPAGE and MOPS SDS buffer, as described.[6,13] Proteins were stained by Coomassie Blue to evaluate the proteome profiles of each fraction, or were transferred to nitrocellulose membranes using a "tank" electrophoretic transfer apparatus (Hoefer) to detect positive exosome markers. The blots were developed as described.[13] Densitometric analysis was performed by Image- Quant TL software (GE Healthcare).

2.4 UEv Proteome Analysis by nLC-ESI-MS/MS

For liquid chromatography-mass spectrometry (LC−MS) analysis, UEv were lysed and the protein extracts digested with trypsin with an adapted protocol of Filter Aided Sample Preparation (FASP) technique.[14,15] Briefly, UEv lysis was performed through a 30 min incubation in RIPA lysis buffer (50 mM Tris−HCl pH7.4, 1% NP40, 0.25% DOC, 150 mM NaCl, 1 mM EDTA). About 100 μg of proteins for each sample was mixed with 400 mM DTT/50 mM NH4HCO3 and heated at 95°C for 5 min for disulfide bonds reduction.

Proteins were then transferred to ultrafiltration spin units of nominal molecular weight (MW) cutoff of 30 kDa (Amicon Ultra 0.5 mL 30 K, Millipore); DTT containing lysis buffer was discharged by centrifugation, and the filters were washed with 50 mM NH4HCO3(14 100g for 15 min). Alkylation was performed by 100 mM IAA/50 mM NH_4HCO_3, incubating samples in darkness for 30 min. Then filters were centrifuged at 14 100g for 5 min and washed five times with 50 mM NH_4HCO_3. Protein digestion was conducted overnight at 37°C adding trypsin for a final trypsin/protein ratio of 1:100 (w/w).[15,16] Tryptic peptides were collected in water by centrifugation and quantified by NanoDrop assay. Peptide mixtures, after desalting by Ziptip μ-C18 pipet tips, were injected into UHPLC system (UltimateTM 3000 RSLCnano, ThermoScientific, Sunnyvale, CA) coupled online with a nESI Ion Trap AmaZon ETD (BrukerDaltonicsGmbH, Germania).

Each sample was analyzed at least three times and to minimize the variability due to sample concentration, about 1 μg of proteins was used for each run of all samples.

Peptides were loaded onto a Pepmap μ-trapping column (2 cm × 100 μm, 5 μm, Thermo Scientific) and then separated onto a 50 cm Nano column (0.075 mm, 2 μm, Acclaim PepMap100, C18, Thermo Scientific) at a flow rate of 300 nL/min. A multistep gradient ranging from 4 to 35% in 167 min and from 35% to 66% B in 32 min, with 9 min of column washing at 98% B, was used (solvent A being 0.1% formic acid in water and solvent B 0.1% formic acid in 80% acetonitrile). Runtime was set to 240 min. The column was directly connected to an online nanospray source. The iontrap mass spectrometer operated in datadependent-acquisition mode using CID MS/MS fragmentation with Helium as collision gas.

Instrumental parameters for MS and MS/MS scan were set as follows: enhanced resolution for MS and Ultrascan for MS/MS, m/z range of 400−1400 for MS, trap ICC value of 400 000, ICC target to 1 000 000 for MS/MS acquisition. CID MS/MS fragmentation was set to fragment the ten most abundant MS peaks with a intensity above 25 000 in a range of 100−2300 m/z with strict active exclusion after one spectra and released after 9 s. Isolation width of MS/MS was 2.5 m/z for charge states ≥2. A narrow range from 50 to 150% for Smart Fragmentation with MS/MS Fragmentation Amplification equal to 0.8 were used.[16] Raw MS/MS data obtained were elaborated and converted to peak lists via Compass DataAnalysis v.4.0 Sp4 (Bruker Daltonics, Germany). No averaging of MS/MS spectra and deconvolution of MS spectra were set. Data resulting from the DataAnalysis elaboration of each analysis were interrogated using in house Mascot search engine (version: 2.4.1), through Mascot Daemon. Database searching was restricted to human Swissprot (accessed Nov 2016; 553 231 sequences; 197 953 409 residues) database. Trypsin as enzyme and carbamidomethyl as fixed modifications were set in search parameters limiting to one the number of missed cleavage. Mass tolerances for all identifications were set at 2 Da for the precursor ions and 0.8 Da for the product ions. Data were filtered using a global false discovery rate of 1% for peptides using automatic decoy database search and a built-in Percolator algorithm were applied to rescore search results with a unique significance threshold. Only proteins with at

least one significant peptide (p-value < 0.05; peptide score ≥ 13) were considered identified.

2.5 Label-Free Relative Quantification

Progenesis QI (Nonlinear Dynamics, Newcastle, England) for proteomics platform was used for label-free MS data quantification, as already described.[16,17] Briefly, LC−MS/MS chromatograms were calibrated using DataAnalysis and then imported into Progenesis QI software. The ion intensity maps of all runs (3 runs for each class) were used for the automatic alignment process. All samples within the experiment were automatically aligned using a reference run to create a maximal overlay across the data. The alignment was considered accepted only if the related score was above 60%. The feature detection algorithm was set using 0.2 min as minimal peak width, and a default sensitivity. Single charged features or with more than nine charges were removed. The normalization process performed using Progenesis QI software minimized experimental variation affecting the protein expression.

Data were normalized to all proteins using the algorithm included in software. Peptides were identified using an in-house Mascot search engine as described in the previous paragraph. Protein entry was identified considering exclusively unique peptides (unique peptide ≥1).[18] Proteins quantification was performed using only non-conflicting (unique) peptides (≥2), avoiding the overlapping of trends derived from different proteins, sharing the same peptides.

2.6 Bioinformatic Analysis

Gene ontology (GO) analysis was performed using LocDB (www.rostlab.org/services/locDB/) and UniprotKB (http:// www.uniprot.org/uniprot/) databases for investigating the subcellular localization of the identified proteins. Tissue specificity was evaluated comparing the lists of the identified proteins with the Human Protein Atlas portal (www. proteinatlas.org),

in particular with the human tissue specific proteome, based on transcriptomics analysis across all major organs and tissue types in the human body.[16,19]

2.7 Statistical Analysis

UEv isolation was performed from five different pools of urine, collected and processed in different days. For each experiment, blood was freshly collected and added to the pool, as explained. We applied a multivariate analysis using one-way ANOVA, followed by Bonferroni post-test, to compare simultaneously the five groups (0−4+), and the unpaired t-test for two by two comparison. All statistical analyses were performed using GraphPad Prism 5.0 software.

3. RESULTS AND DISCUSSION

Hematuria is a very common condition displayed during the course of several nephrologic and urologic diseases. It can clearly affect urine composition, but could also impair the proteomic analysis of urinary microparticles or exosomes.[9−11]

To verify the impact of this impairment, we simulated a condition of hematuria by adding an increasing amount of blood to a unique pool of urine from healthy subjects.

3.1 UEv Protein Profile and Marker Enrichment in the Presence of Hematuria

After mimicking increasing hematuria, similar to the 1+, 2+, 3+, and 4+ clinical levels, we prepared UEv following the usual protocol and compared the protein patterns by SDS-PAGE after CBB staining. Results showed that blood contaminated samples showed a significantly lower UEv yield, as demonstrated by

the faint pattern of bands on gels (Figures 2a and S1) and by protein assay (Table 1).

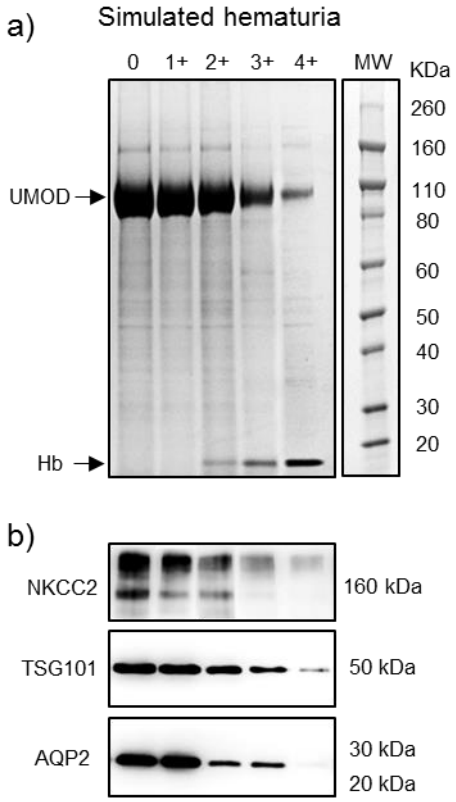


Figure 2. Comparison of protein profiles of urinary extracellular vesicles (UEv) isolated from urine with increasing levels of blood (1+, 2+, 3+, 4+). (a) UEv protein profiles obtained by control urine pool with/without blood; the lanes are loaded with UEv proteins coming from the same amount of starting urine (3 mL), to check the different UEv recovery. (b) Enrichment of UEv markers. In this case, the lanes are loaded with the same amount of UEv proteins (5 µg) to check the different enrichment. UMOD, uromodulin; Hb, hemoglobin; MW, molecular weight marker.

Table 1. Protein Content with and without Blood Contamination[a]

hematuria level (Combur Test)		0	1+	2+	3+	4+
urinary Ev proteins (μg)[b]/100 mL urine	NT	120 ± 15.1	123 ± 7.9	109.4 ± 31.3	27.5 ± 20.5	14.1 ± 7.9
	T	126 ± 15	121 ± 22	131 ± 13	127 ± 18	117 ± 23
total urinary proteins[c] (mg)/100 mL urine		10.5 ± 3.83	12.4 ± 2.81	13.7 ± 3.21	16.8 ± 2.72	23.4 ± 4.52
plasma proteins[d] (mg)/100 mL urine		0	0.098	0.47	4.7	9.3
hemoglobin from hemolysis[e] (mg)/100 mL urine		0	0.018	0.087	0.87	1.73

[a]0, without blood contamination; 1+, 2+, 3+, 4+ describe the hematuria level of the urine assessed by Combur test (Roche). [b]Protein assay performed on UEv isolated from urine treated (T) and not treated (NT) with trypsin (media ± SD of 5 different UEv preparations). [c]Protein assay performed after sediment removal, before the addition of trypsin (media ± SD of 5 different urine contamination experiments). [d]Theoretical content considering an average concentration of 7 g/dL. [e]Hemoglobin content supposing 10% hemolysis during the incubation of blood with urine, considering an average concentration of 13 g/dL.

This was especially marked in 3+ and 4+ samples (Figures 2a and S1). Furthermore, as hematuria increased, the protein profile shifted from the typical UEv of healthy subjects: the content of uromodulin, the most abundant band at 80−100 kDa,[6] decreased, while the intensity of the hemoglobin band increased (Figures 2a and S1). These effects were confirmed in different experiments, although with some variability (Figure S1).

To evaluate the purity of the UEv fractions isolated in the presence of blood, we checked the enrichment of known UEv markers, such as TSG101 and AQP2,[20] and NKCC2, a typical renal protein present in UEv,[13,21] by WB. Results showed that the intensity of the signals decreased as blood contamination increased (Figure 2b), thus suggesting that the higher was the hematuria levels, the lower the UEv purity. In other words, the high speed pellet from hematuria urine is not only reduced as quantity, but it also contains more contaminants than the native sample (0). Therefore, results strongly suggest that the UEv samples obtained from urine contaminated with high level of blood are not suitable for proteomics analysis.

3.2 Effect of Trypsin Treatment on UEv Protein Profile and Marker Enrichment in the Presence of Hematuria

To investigate the cause of such a low UEv yield in urine with higher blood levels, we examined the pellet obtained after the 17 000g centrifugation (P2) and the supernatant after the 200 000g ultracentrifugation (Sn), which are usually discarded. Results showed that a part of UEv were lost because they were not correctly pelleted at high speed. In fact, we could find them in the P2 fraction, where usually large particles and membrane fragments are pelleted, as also confirmed by markers distribution (Figures 2b, 3a, and S2A). This suggested that a process of vesicle aggregation occurred in the presence of blood, leading to the formation of heavier particles or aggregates. With the aim of preventing the reduction of UEv yield, we tried to avoid any aggregation using a trypsin treatment before UEv isolation.

We had already investigated the possible positive effect of a mild trypsin treatment of the UEv. Results had shown that not only UEv are resistant to trypsin digestion, but also that interfering proteins could be removed.[22] We therefore applied this approach to hematuric specimens, adapting the protocol to our purpose. Preliminary experiments performed with different trypsin concentrations showed that the best ratio was 1/50 (trypsin mg/proteins mg) (data not shown). In this optimized conditions, the amount of UEv recovered even in the presence of 3+ and 4+ hematuria levels increased markedly (Figure 3d and Table 1), and markers distribution was nearly restored (Figure 3b, c,e). Accordingly, marker signals were not detectable in the P2 fraction anymore (Figure 3b). In fact, after trypsin treatment, UEv markers shifted from P2 to the high speed UEv pellet in the highly contaminated samples (3+ and 4+) (Figure 3b).

We estimated the recovery of UEv after trypsin treatment, based on the enrichment of markers in UEv (Figure 3c). It was calculated that about 60−80% of the UEv are lost when blood contamination is 3/4+, but most of them could be recovered using trypsin. However, the recovery was not always full, suggesting that unknown mechanisms, other than the entrapment in the P2, influenced the loss of UEv. Furthermore, the protein profile of UEv obtained from hematuria urine treated with trypsin resembled the one without blood, except for some specific bands attributable to serum albumin (66 kDa), immunoglobulin heavy (50−55 kDa) and light (25 kDa) chains, and hemoglobin chains (15 kDa) in 3+ and 4+ UEv (Figure 3d). Their presence became visible thanks to the increased recovery of UEv in heavy contaminated samples after trypsin treatment.

To better understand these results, we took into account the theoretical amount of additional plasmatic proteins of the blood added to each sample, looking at the total protein recovery in the different conditions (Table 1). Moreover, we considered the theoretical amount of additional hemoglobin, hypothesizing a level of hemolysis of 10% during the permanence of urine in the bladder (simulated by 4 incubations at 37°C). Results showed that the measured urinary proteins agreed with theoretical values, showing increased urinary protein content in higher blood contamination

conditions (Table 1). Despite this, UEv protein content was not statistically different among the hematuria samples, when they are treated with trypsin (t test, one-tailed, p value > 0.05).

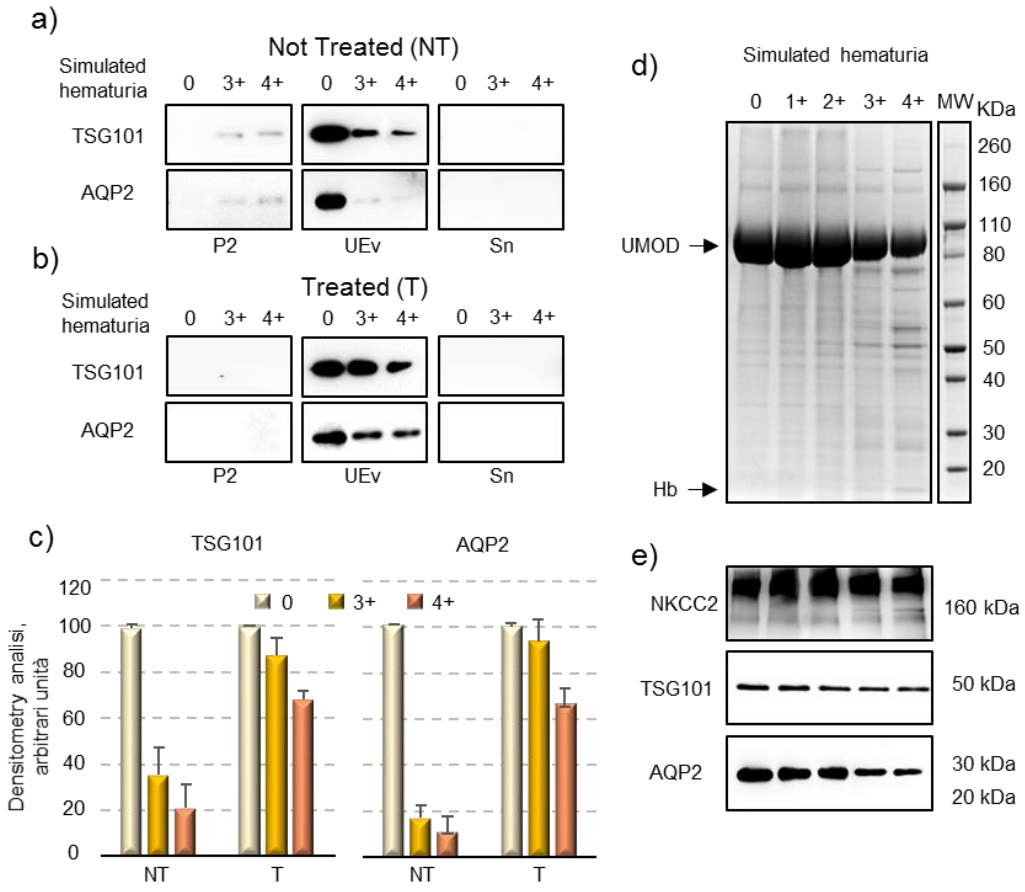


Figure 3. Effect of trypsin treatment on urinary extracellular vesicle (UEv) isolation. (a) Marker enrichment in subfractions obtained from urine without blood contamination (0) and with high level of hematuria (3+ and 4+) not treated with trypsin (NT). (b) Marker enrichment in subfractions obtained from 0 and 3+ and 4+ urine samples treated with trypsin (T). Proteins coming from the same amount of starting urine for each urine subfraction were loaded, to check the different protein recovery. P2, large membrane fragments and particles (protein corresponding to 1 mL of urine); UEv, exosome-enriched fraction (protein corresponding to 7 mL); Sn, supernatant after the 200 000g ultracentrifugation (soluble proteins corresponding to 0.2 mL). (c) Densitometric analysis of the TSG101 (left) and AQP2 (right) bands visible in UEv, from three independent experiments. (d) Trypsin treated UEv protein profiles obtained from control urine pools with (1+, 2+, 3+, 4+) or without blood (0): samples were loaded with UEv proteins coming from the same volume of starting urine (3 mL). (e) Enrichment of UEv markers. In this case, samples were loaded with the same amount of UE proteins (5 µg). For panels a, b, d, and e, one representative experiment is shown.

In fact, even if the charge of external proteins increased with increasing blood addition, almost doubling in 4+ condition, the UEv protein content did not. The trypsin treatment was hence able to free most of UEv, maintaining the peculiar compartmentalization of UEv cargo,[15] and in particular, to make UEv isolated from blood-contaminated samples suitable for the following proteomic analysis.

It is conceivable that other strategies are also able to eliminate the blood-related aggregation phenomenon. A possibility could be to treat the samples with EGTA to block calcium-mediated adhesion. Although the involvement of calcium in several mechanisms of cell−cell interaction is well-known,[23] no data are available about the possible role of calcium in interactions between vesicles and cells/other vesicles. Moreover, some papers reported that EV have specific membrane glycome, involved in modulating EV properties, among which their recognition and uptake by target cells.[24−26]

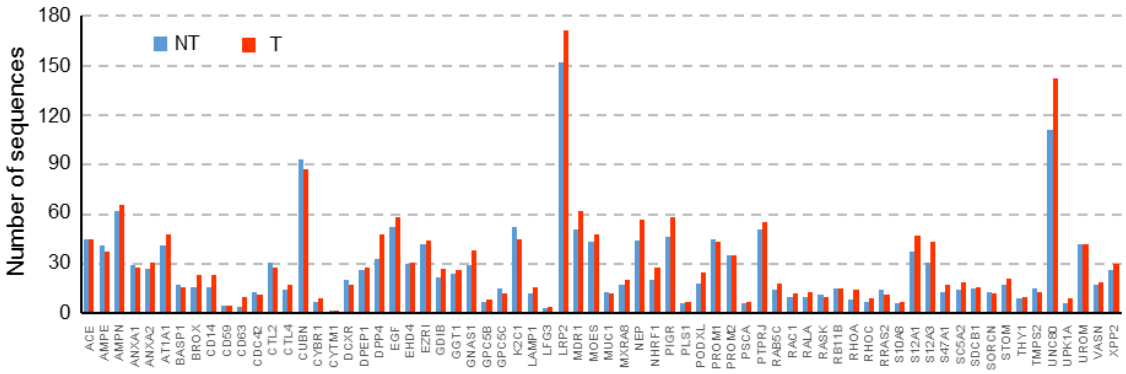
This suggests that treatment with sialidase or lectin-targeted sugars are possible approaches to prevent sialic acid interactions and aggregation at the cell/vesicle membrane level. However, assessing these alternative approaches deserves further investigation.

3.3 Proteomic Analysis of UEv from Hematuria Urine

To check the ability of the trypsin treatment to limit the negative effects of hematuria, we performed LC−MS analysis on the UEv isolated from urine samples contaminated with increasing amount of blood, after trypsin treatment. We detected about 620 proteins, by summing up all the different species identified in the different samples (Tables S1, S2, and S3). The classification of these proteins according with their subcellular localization, by UniProtKB (www.uniprot.org) and LocDB (www.rostlab.org/services/locDB), allowed us to conclude that their distribution is in good agreement with that expected for UEv, being most membrane (27%) and cytoplasmic proteins (32%) (Figure S3 and Table S1).[27]

When comparing the two samples without blood addition (0) ± trypsin treatment, we found a good correspondence between the lists of identified proteins. In fact, 332 proteins were found in both batches, 32 were present only in the not treated

(NT) sample and 80 proteins were identified only after the treatment (T) (plus about 20%), with a small increase in the detection of high hydrophobic proteins after trypsin treatment (data not shown). This was somehow a surprising result since trypsin is expected to cause a sort of membrane shaving, resulting in the loss of peptides covering protein regions outside of the vesicles. Instead, we did not find any difference in the number of peptides identified in the proteins shared between NT and T samples. On the contrary, we observed a general slight increment in the



peptide coverage of the peptides belonging to putative membrane proteins, instead of a reduction (Figure 4 and Table S1).

Figure 4. Effect of trypsin treatment on the identification of UEv surface proteins. Graphic representation of the number of the peptides found in UEv putative membrane proteins obtained from urine, not contaminated with blood, treated (T) or not treated (NT) with trypsin.

A possible explanation for this behavior is that we performed a mild treatment with trypsin diluted in a large volume of urine because when we directly submitted isolated UEv to trypsin incubation, the membrane shaving was evident (data not shown). These findings indicated that the trypsin treatment of urine did not significantly altered the whole UEv proteome; instead, it slightly improved the protein identification process.

For this reason, from here on, we will compare the lists of the identified proteins only after trypsin treatment. In each sample, we identified about 400 proteins (421 ± 45), with some variations of their number among the different conditions (Table S4). More than 330 proteins (about 80%) were present in all the samples (Figure S4).

When analyzing the tissue specificity of the identified proteins in detail (Table S4), about half of them were classified as ubiquitous. Among the remaining half, the percentage of blood-associated proteins, such as proteins deriving from bone marrow, immune system (including Ig families) and liver, were similar in all samples, regardless of blood contamination (19−21%). The same was true for proteins specific of the urinary tract (8−9%) (Table S4), indicating that the proteins identified in the blood-contaminated UEv after trypsin treatment did not differ qualitatively from not-contaminated UEv, even in the presence of higher blood levels.

Table 2. Results of Label-Free Quantification[a]

	Ratio hematuric/nonhematuric sample (0)			
	1+	2+	3+	4+
Protein name				
Exosomal Markers				
Annexin A11[b]	1.0 ± 0.3	1.1 ± 0.4	1.0 ± 0.3	1.0 ± 0.3
CD9[b]	1.3 ± 0.1	1.0 ± 0.1	1.4 ± 0.1	1.3 ± 0.1
Flotillin-1[b]	0.9 ± 0.2	1.2 ± 0.1	1.1 ± 0.2	1.4 ± 0.2
Syntenin-1[b]	0.8 ± 0.4	1.2 ± 0.4	0.8 ± 0.3	0.7 ± 0.2
Tumor susceptibility gene 101 protein[b]	1.0 ± 0.1	0.9 ± 0.1	1.0 ± 0.1	0.9 ± 0.1
Urinary Exosomal Proteins (Wang et al.)[19]				
Carbonic anhydrase 4[b]	0.8 ± 0.2	0.9 ± 0.2	0.9 ± 0.2	1.0 ± 0.2
CD59 glycoprotein[b]	0.8 ± 0.03	0.8 ± 0.03	0.8 ± 0.04	0.9 ± 0.02
Dipeptidyl peptidase 4[b]	1.0 ± 0.1	1.0 ± 0.1	1.0 ± 0.1	1.0 ± 0.1
Mucin-1[b]	1.6 ± 0.1	1.4 ± 0.2	1.7 ± 0.3	1.3 ± 0.2
Voltage-dependent anion-selective channel protein 1[b]	1.1 ± 0.3	1.2 ± 0.4	1.2 ± 0.3	1.3 ± 0.4
Urinary Tract Specific Proteins (Protein Atlas)[c]				
Dipeptidase 1[b]	1.0 ± 0.2	1.0 ± 0.1	1.1 ± 0.2	1.1 ± 0.2
Neprilysin[b]	0.9 ± 0.1	1.1 ± 0.1	0.9 ± 0.1	0.9 ± 0.2
Solute carrier family 12 member 1[b]	1.2 ± 0.3	1.0 ± 0.2	0.9 ± 0.3	1.2 ± 0.1
Uromodulin[b]	1.1 ± 0.1	1.0 ± 0.1	0.6 ± 0.04	**0.4 ± 0.02**
Xaa-Pro aminopeptidase 2[b]	1.2 ± 0.3	1.2 ± 0.3	1.1 ± 0.2	1.1 ± 0.2
Blood-Associated Proteins				
Apolipoprotein D[b]	0.9 ± 0.2	0.9 ± 0.2	0.8 ± 0.2	0.8 ± 0.2
Complement C4−B	1.0 ± 0.1	1.1 ± 0.1	1.3 ± 0.2	1.6 ± 0.2
Haptoglobin[b]	1.1 ± 0.2	1.5 ± 0.2	**4.2 ± 0.5**	**12.4 ± 1.2**
Hemoglobin subunit alpha[b]	**5.2 ± 0.6**	**19.8 ± 2.1**	**43.7 ± 7.1**	**50.1 ± 4.1**
Serum albumin[b]	1.1 ± 0.1	1.4 ± 0.3	1.9 ± 0.2	**3.7 ± 0.3**

[a]For each protein, the ratio (± absolute error) between the condition with (+) and without blood (0) is shown. The list groups proteins representative of exosome markers, urinary exosome proteins, urinary tract tissue specific proteins, and blood-associated proteins. Note: 0, without blood contamination; 1+, 2+, 3+, and 4+ describe the hematuria level of the urine assessed by Combur test (Roche). [b]Proteins identified also by Wang et al.[28] [c]We considered as urinary tract specific proteins those classified as enriched by ProteinAtlas (including the tissue-enriched, the group enriched and the tissue-enhanced ones).

We also submitted MS results to a label-free analysis, to get relative quantification (143 proteins) (Table S5). Since UEv were isolated from a unique pool of control urine, no difference in their amount should be expected. Overall, the proteins with

no alteration in their levels in the UEv (i.e., with a ratio between 0.5 and 2) reached 99% for 1+, 97% for 2+, 92% for 3+, and 84% for 4+. The number of proteins with a ratio >2 (up-proteins) increased with the level of blood contamination (Table S6), while those of down-proteins (with ratio <0.5) were irrelevant. In particular, levels of typical exosomal markers, such as Annexins (A1 and A11), Flotillins 1, Syntenin-1, TSG101, and tetraspannins (CD9),3 were not generally affected (Table 2).

Similarly, most of proteins shared with the fullest set of healthy subject UE, published by Wang et al.[28] showed unchanged ratios. Interestingly, all the specific proteins of the urinary tract, such as Neprilysin,[29] Dipeptidase 1,[30] and the proton-coupled amino acid transporter 2,[31] resulted unchanged. As expected, some blood-associated proteins, such as globin subunits, increased, but even in this group, most proteins displayed unchanged ratios: 93% in 1+, 87 in 2+, 60% in 3+. For example, apolipoprotein D seemed to be unaffected also by higher blood contamination (3+ and 4+).

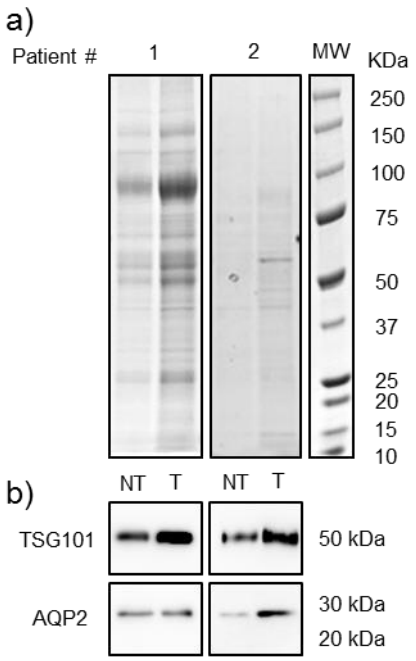


Figure 5. Effect of trypsin treatment on protein profiles of UEv of patients with hematuria. Patient #1: renal cell carcinoma, hematuria level 1+. Patient #2: renal lithiasis, hematuria level 4+.

(a) UEv protein profiles obtained from patient urine, trypsin treated (T), or not (NT): samples were loaded with UEv proteins coming from the same volume of starting urine (5 mL). (b) Enrichment of UEv markers. Samples were loaded with the same amount of UEv proteins (5 µg).

Finally, to validate the procedure further, we preliminarily tested the proposed protocol on hematuria urinary samples from two patients. Results showed that the trypsin is able to enhance the UEv recovery, as visible from the comparison between protein profiles and confirmed by maker enrichment (Figure 5).

4. CONCLUSIONS

We demonstrated that the presence of blood affects mainly the vesicle recovery in the high-speed pellet, and second its protein pattern, giving rise to very different samples both quantitatively and qualitatively. We showed that an optimized trypsin treatment performed before UEv isolation could lead to results similar to those obtained in the absence of blood. In fact, the vesicles prepared from blood contaminated urine after such treatment resembled the non-contaminated ones in terms of general protein pattern and quantity. The protein species that resulted severely increased mainly belonged to the globin family or to blood-associated species and can be easily discarded from the analysis.

Our treatment allows recovering several samples that would be otherwise discarded, such as those coming from patients affected by diseases of the urinary tract with hematuria, especially when investigating the differential abundance of proteins specific of UEv or of the urinary tract. Another possible application of this method that removes negative effects of blood contamination on UEv protein content is the case of patients with a disease affecting blood-associated protein present in UEv. In this situation, the protein abundance of the patient's UEv should be compared with those of healthy subject with a similar simulated blood contamination. The use of pathological UEv samples, even if blood-contaminated, compared to healthy ones, may therefore disclose real differential proteins specific for the investigated diseased condition.

In conclusion, our modified protocol to UEv preparation allows proteomic analysis

of vesicles also in hematuria conditions, often encountered in kidney and urological diseases, as shown by preliminary results on UEv isolated from patients presenting hematuria. By adopting this protocol, the urinary extracellular vesicles isolated from hematuria urine may reveal a differential protein cargo, a source of promising biomarkers for the diagnosis and monitoring of renal diseases.

ASSOCIATED CONTENT

The Supporting Information is available free of charge on the ACS Publications website at **DOI: 10.1021/acs.jproteome. 7b00763**.

Effect of blood contamination on UEv protein content; effect of trypsin treatment on enrichment of UEv markers in urine subfractions; subcellular localization of proteins identified in UEv; Venn diagram describing overlapping and unique proteins in Uev samples after trypsin treatment with and without blood; number and tissue specificity of identified proteins in each sample, contaminated or not with blood after trypsin treatment; distribution of up, down, and unchanged proteins in relation to blood contamination levels (PDF).

References

(1) Sharp, V. J.; Barnes, K. T.; Erickson, B. A. Assessment of asymptomatic microscopic hematuria in adults. Am. Fam. Physician 2013, 88, 747−754.

(2) Davis, R.; Jones, J. S.; Barocas, D. A.; Castle, E. P.; Lang, E. K.; Leveillee, R. J.; Messing, E. M.; Miller, S. D.; Peterson, A. C.; Turk, T. M. T.; et al. American Urological Association (AUA) Guideline ADULTS: AUA GUIDELINE American Urological Association Asymptomatic Microhematuria. J. Urol. 2012, 188 (6 Suppl), 1−30.

(3) Kowal, J.; Arras, G.; Colombo, M.; Jouve, M.; Morath, J. P.; Primdal-Bengtson, B.; Dingli, F.; Loew, D.; Tkach, M.; Théry, C. Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes. Proc. Natl. Acad. Sci. U. S. A. 2016, 113, E968−977.

(4) Webber, J.; Clayton, A. How pure are your vesicles? J. Extracell. Vesicles 2013,

2, 19861.
(5) Raimondo, F.; Morosi, L.; Corbetta, S.; Chinello, C.; Brambilla, P.; Della Mina, P.; Villa, A.; Albo, G.; Battaglia, C.; Bosari, S.; et al. Differential protein profiling of renal cell carcinoma urinary exosomes. Mol. BioSyst. 2013, 9, 1220−1233.
(6) Raimondo, F.; Cerra, D.; Magni, F.; Pitto, M. Urinary proteomics for the study of genetic kidney diseases. Expert Rev. Proteomics 2016, 13, 1−16.
(7) Pomatto, M. A. C.; Gai, C.; Bussolati, B.; Camussi, G. Extracellular Vesicles in Renal Pathophysiology. Front. Mol. Biosci. 2017, 4, 37.
(8) Moon, P. G.; You, S.; Lee, J. E.; Hwang, D.; Baek, M. C. Urinary exosomes and proteomics. Mass Spectrom. Rev. 2011, 30, 1185−1202.
(9) Smalley, D. M.; Sheman, N. E.; Nelson, K.; Theodorescu, D.Isolation and identification of potential urinary Microparticle biomarkers of bladder cancer. J. Proteome Res. 2008, 7, 2088−2096.
(10) Chen, C. L.; Lai, Y. F.; Tang, P.; Chien, K. Y.; Yu, J. S.; Tsai, C.H.; Chen, H. W.; Wu, C. C.; Chung, T.; Hsu, C. W.; et al. Comparative and targeted proteomic analyses of urinary microparticles from bladder cancer and hernia patients. J. Proteome Res. 2012, 11, 5611−5629.
(11) Moon, P. G.; Lee, J. E.; You, S.; Kim, T. K.; Cho, J. H.; Kim, I. S.; Kwon, T. H.; Kim, C. D.; Park, S. H.; Hwang, D.; et al. Proteomic analysis of urinary exosomes from patients of early IgA nephropathy and thin basement membrane nephropathy. Proteomics 2011, 11, 2459−2475.
(12) Thongboonkerd, V.; Chutipongtanate, S.; Kanlaya, R. Systematic evaluation of sample preparation methods for gel-based human urinary proteomics: Quantity, quality, and variability. J. Proteome Res. 2006, 5, 183−191.
(13) Corbetta, S.; Raimondo, F.; Tedeschi, S.; Syrèn, M. L.; Rebora, P.; Savoia, A.; Baldi, L.; Bettinelli, A.; Pitto, M. Urinary exosomes in the diagnosis of Gitelman and Bartter syndromes. Nephrol., Dial., Transplant. 2015, 30, 621−630.
(14) Wisniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M. Universal sample preparation method for proteome analysis. Nat. Methods 2009, 6, 359−362.
(15) Raimondo, F.; Corbetta, S.; Morosi, L.; Chinello, C.; Gianazza, E.; Castoldi,

G.; Di Gioia, C.; Bombardi, C.; Stella, A.; Battaglia, C.;
et al. Urinary exosomes and diabetic nephropathy: a proteomic approach. Mol. BioSyst. 2013, 9, 1139−1146.

(16) Raimondo, F.; Corbetta, S.; Savoia, A.; Chinello, C.; Cazzaniga, M.; Rocco, F.; Bosari, S.; Grasso, M.; Bovo, G.; Magni, F.; Pitto, M. Comparative membrane proteomics: a technical advancement in the search of renal cell carcinoma biomarkers. Mol. BioSyst. 2015, 11, 1708−1716.
(17) Chinello, C.; Stella, M.; Piga, I.; Smith, A. J.; 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. J. Proteomics 2018, DOI: 10.1016/j.jprot.2018.04.029.
(18) Zhang, Y.; Xu, T.; Shan, B.; Hart, J.; Aslanian, A.; Han, X.; Zong, N.; Li, H.; Choi, H.; Wang, D.; Acharya, L.; Du, L.; Vogt, P. K.; Ping, P.; Yates, J. R., 3rd ProteinInferencer: Confident protein identification and multiple experiment comparison for large scale proteomics projects. J. Proteomics 2015, 129, 25−32.
(19) Uhlén, M.; Fagerberg, L.; Hallström, B. M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, Å.; Kampf, C.; Sjöstedt, E.; Asplund, A.; et al. Proteomics. Tissue-based map of the human proteome. Science 2015, 347, 1260419.
(20) Dear, J. W.; Street, J. M.; Bailey, M. A. Urinary exosomes: a reservoir for biomarker discovery and potential mediators of intrarenal signalling. Proteomics 2013, 13, 1572−1580.
(21) Pisitkun, T.; Shen, R. F.; Knepper, M. A. Identification and proteomic profiling of exosomes in human urine. Proc. Natl. Acad. Sci. U. S. A. 2004, 101, 13368−13373.
(22) Liu, X.; Chinello, C.; Musante, L.; Cazzaniga, M.; Tataruch, D.; Calzaferri, G.; Smith, A. J.; De Sio, G.; Magni, F.; Zou, H.; et al. Intraluminal proteome and peptidome of human urinary extracellular vesicles. Proteomics: Clin. Appl. 2015,

9, 568−573.
(23) Steffen, P.; Jung, A.; Nguyen, D. B.; Müller, T.; Bernhardt, I.; Kaestner, L.; Wagner, C. Stimulation of human red blood cells leads to Ca2+-mediated intercellular adhesion. Cell Calcium 2011, 50, 54−61.
(24) Shimoda, A.; Tahara, Y.; Sawada, S. I.; Sasaki, Y.; Akiyoshi, K. Glycan profiling analysis using evanescent-field fluorescence-assisted lectin array: Importance of sugar recognition for cellular uptake of exosomes from mesenchymal stem cells. Biochem. Biophys. Res. Commun. 2017, 491, 701−707.
(25) Costa, J. Glycoconjugates from extracellular vesicles: structures, functions and emerging potential as cancer biomarkers. Biochim. Biophys. Acta, Rev. Cancer 2017, 1868, 157−166.
(26) Gerlach, J. Q.; Griffin, M. D. Getting to know the extracellular vesicle glycome. Mol. BioSyst. 2016, 12, 1071−1081.
(27) Raimondo, F.; Morosi, L.; Chinello, C.; Magni, F.; Pitto, M. Advances in membranous vesicle and exosome proteomics improving biological understanding and biomarker discovery. Proteomics 2011, 11, 709−720.
(28) Wang, Z.; Hill, S.; Luther, J. M.; Hachey, D. L.; Schey, K. L. Proteomic analysis of urine exosomes by multidimensional protein identification technology (MudPIT). Proteomics 2012, 12, 329−338.
(29) Ishida, M.; Ogawa, M.; Kosaki, G.; Mega, T.; Ikenaka, T. Purification and characterization of the neutral endopeptidase from human kidney. J. Biochem. 1983, 94, 17−24.
(30) Döring, F.; Martini, C.; Walter, J.; Daniel, H. Importance of a small N-terminal region in mammalian peptide transporters for substrate affinity and function. J. Membr. Biol. 2002, 186, 55−62.
(31) Fernández-Llama, P.; Khositseth, S.; Gonzales, P. A.; Star, R. A.; Pisitkun, T.; Knepper, M. A. Tamm-Horsfall protein and urinary exosome isolation. Kidney Int. 2010, 77, 736−742.

CHAPTER 5

Urinary Extracellular vesicles profile discriminates different clinical subgroups of children with Idiopathic Nephrotic Syndrome

Lucia Santorelli[a]#, William Morello[a,b]#**,** Elisa Barigazzi[a], Chiara Tamburello[b] , Luciana Ghio[b], Barbara Crapella[b], Giulia Capitoli[c], Stefania Galimberti[c], Francesca Raimondo[a], Giovanni Montini[b] and Marina Pitto[a]

#equally contributing authors

a) School of Medicine and Surgery, University of Milano - Bicocca, Milano, Italy
b) Paediatric Nephrology, Dialysis and Transplant Unit, Fondazione IRCCS Ca' Granda-Ospedale Maggiore Policlinico, Milan, Italy
c) School of Medicine and Surgery, University of Milano – Bicocca, Centre of Biostatistics for Clinical Epidemiology, Vedano al Lambro, Italy.

Manuscript in preparation

ABSTRACT There is growing interest regarding the use of minimally invasive "liquid biopsies" to identify new biomarkers. Urinary extracellular vesicles (UEv) are nanovesicles released into urine by cells facing the urinary space. Their molecular composition depends upon the type of the cell of origin, providing a fingerprint, capable to monitor its status. Their presence in urine makes them readily accessible, giving the possibility to investigate pathological conditions especially related to kidney.

While exosome research has flourished, few studies have specifically targeted the role of UEv in Idiopathic Nephrotic Syndrome (INS) in children the most frequent childhood glomerular disease. The pathogenesis of INS is still unknown and response to initial treatment with corticosteroids is the major indicator of long-term prognosis, as resistant patients often progressive to end-stage renal disease.

The aim of the study is to verify the feasibility to use UEv of INS patients as a source of predictive markers of response to corticosteroid treatment, and/or for clarifying the disease etiopathogenesis and/or the pharmacoresistance. Thus, we investigated the UEv protein content of a paediatric cohort of 34 patients, classified in three clinical classes, according to the corticosteroids response: steroid-Dependent, steroid-Sensitive and steroid-Resistant.

Firstly, a characteristic SDS-PAGE protein profile emerged to be associated with each class, which preserves its peculiarity also when compared to UEv protein content of non-INS patients (orthostatic proteinuria, hereditary tubulopathies) and healthy age-matched controls. Secondly, we also pinpointed different levels of specific glomerular membrane proteins (Podocin, P-glycoprotein, Neprylisin, Nephrin and Transient receptor potential cation channel subfamily C member 6), described as involved in the INS development.

These evidences confirmed the feasibility to use a UEv approach to intercept the pathophysiological differences underlying response to therapy.

KEYWORDS Idiopathic Nephrotic Syndrome, liquid byopsy, urinay extracellular vesicles, protein profile, children

1. INTRODUCTION

Idiopathic Nephrotic Syndrome (INS) is the most common glomerular disease in children and it is characterized by proteinuria, hypoalbuminaemia, and oedema. The pathogenesis is poorly understood [1]. Evidence suggests an involvement of immunological mechanisms, with B- and T-cell dysfunction [2] and a possible role of Epstein-Barr Virus (EBV) infection [3]. A yet unidentified permeability factor is believed to be involved at least in patients with INS relapsing after renal transplantation [4]. Underlying mutations in genes encoding podocyte associated proteins, resulting in structural or functional disruption of the glomerular filtration barrier, can be found in 30% of cases refractory to treatment [5].

The mainstay of therapy are steroids, and INS is usually classified according to the response to treatment, as steroid-sensitive, SSNS (S) or steroid-resistant, SRNS (R). At least 50% of SSNS will require long-term steroid or immunosuppressive treatment to maintain remission and will be further classified as steroid-dependent, SDNS (D). The prognosis is extremely different according to the subgroup, with SRNS resistant to second-line treatments progressing into end stage renal disease in virtually 100% of cases [1]. Unfortunately, established and reliable biomarkers of response are lacking and children may be exposed to unnecessary and toxic immunosuppressive therapies.

In this contest, the proteomic study of the urinary extracellular vesicles (UEv) represents a valid approach. UEv are nanometer-sized vesicles (50-200 nm), that can originate from endothelia and glomerular cells, podocytes or tubular epithelial cells. UEv act as a vehicle by which cells communicate, delivering their functional content, with biological, physiological and pathological functions [6] . Enriched in renal proteins, UEv proteome contains less than 3% of total urine proteins (>3000 species), depleting it from the most abundant urinary proteins, reducing in that way the complexity of the urine proteome. Their molecular composition depends upon the type, and even status, of the origin cell. As such, they provide an easily accessible window to monitor the status of renal tissue. For all these reasons, they can be considerate a sort of liquid biopsy, able to provide potential

pathophysiological biomarkers and possible protagonists of the disease's pathogenesis.

Few studies have specifically targeted the role of UEv in INS, maybe due to the high proteinuria typical of these patients. In fact, large amount of proteins, in particular the highly abundant ones, influences negatively UEv isolation and the following proteomic analysis

Aware of the technical challenges related to the study of UEv in case of proteinuria, we approached this issue from another point of view, focusing on UEv isolated from the urine samples of patients under treatment in remission. Here we aimed at investigating the role of UEv in differentiating patients according to the response to treatment, comparing S, R and D patients, in order to find possible implications in the pathogenesis of the disease, other than predicting the evolution of therapy response.

2. MATERIALS AND METHODS

2.1 Clinical data and study design

We performed a pilot study of UEv in pediatric patients with INS. Urine samples were prospectively collected from all INS children, attending the Pediatric Nephrology Dialysis and Transplant unit of Milan from July 2018 and July 2019 and stored at −80°C until the analysis. Patients were classified according to the response to the initial steroid therapy and the need for further immunosuppressive treatment into SSNS, SDNS, SRNS, according to international consensus. In order to prevent interferences related to the presence of serum protein in urine, patients with significant proteinuria were excluded (uPr/uCr > 1 mg/mg). A total 74 samples were collected, of which 41 were later excluded for the presence of significant proteinuria. Therefore, 33 samples were analyzed (9 S, 17 D, 6 R). Table 1 summarize the main clinical characteristics of patients included.

Table 1. Clinical characteristics of enrolled patients

Patient	Group	Sex	Age at collection	uPr/uCr	Ongoing Therapy			
					Pred	Cycl	MMF	Tacr
5	D	M	5	0.24	X		X	X
6	D	M	10	0.14			X	
7	D	M	12	0.60	X		X	
8	D	M	5	0.23			X	X
10	D	M	18	0.08			X	
13	D	F	4	0.19			X	
14	D	F	8	0.18			X	
18	D	M	8	0.15			X	
30	D	F	6	0.16			X	
32	D	F	8	0.15	X		X	
38	D	M	15	0.27	X		X	
53	D	M	10	0.84			X	
55	D	M	5	0.74		X	X	
58	D	M	12	0.15		X		
61	D	M	7	0.25			X	
66	D	F	6	0.14			X	
69	D	F	15	0.15			X	
2	S	M	7	0.14	X			
15	S	F	2	1.03	X			
16	S	M	5	0.13	X			
24	S	M	4	0.27	X			
43	S	M	17	0.08				
44	S	F	16	0.13				
67	S	F	11	0.16				
73	S	M	10	0.12	X			
74	S	M	11	0.12	X			
1	R	M	10	0.35	X	X	X	
3	R	M	4	0.18	X		X	X
11	R	F	14	0.10		X	X	
12	R	F	8	0.11		X		
39	R	F	12	0.14		X		
70	R	F	11	0.13				
4	OP	M	13	1.08				

Legend: **PRED**. Prednisone; **CYCL**. Cyclosporin; **MMF**. Mycophenolate mofetil; **TACR** Tacrolimus; OP. Orthostatic Proteinuria; uPr. urinary proteins; uCr. urinary creatinine

2.2 Urine collection

Urine samples, (mean volume = 20 ml) were centrifuged for sediment removal (10 min at 1'000 xg, 4°C) within 4 hours from the collection. The supernatant was supplemented with protease inhibitors (Complete, Roche) and stored at −80°C until exosome isolation. From an aliquot of each sample (500 µl out of about 20 ml), urinary proteins were precipitated by nine volumes of cold 90% ethanol and pelleted at 3'500 xg for 30 minutes [8]. After drying, proteins were dissolved in bidistilled water, and protein concentration was assessed by BCA assay (SIGMA Chemical Co), using BSA as standard.

Before exosomes isolation, the remaining stored urine samples were thawed, thoroughly vortexed while thawing and adjusted to pH 7.4, if needed.

2.3 UEv isolation

UEv were prepared by ultracentrifugation [9] according to HKUPP (http: www.hkupp.org), with minor modifications. All steps were performed at 20°C. Briefly, urine were added with $ZnSO_4$ 10 mM, incubated at RT for 1 hour and then centrifuged for 30 min at 3'000 *xg* at 20°C, to eliminate THP [10] (data not shown). Supernatants were then subjected to further centrifugation for 15 min at 17'000 *xg*. Supernatants were finally submitted to ultracentrifugation for 70 minutes at 200'000 *xg*: crude UEv pellets were washed and suspended in PBS, in the presence of protease inhibitors. The UEv samples were stored at -80°C until use.

2.4 UEv characterization by Nanoparticle tracking analysis

UEv size and concentration were measured by Nanoparticle Tracking Analysis (NTA) using a NanoSight NS300 (Malvern Instrument Inc., Malvern, UK) equipped with a 488 nm laser and a syringe pump system. Before injection, UEv were diluted in sterile PBS. The camera operated at 30 frames per second (fps), the threshold for the tracking system was set at 5 and the sample analysis was performed around 1 minute for 3 times. After three technical replicates, the

resulting tracking graphs were analysed by NTA 3.2 software (dev build 3.2.16).

2.5 Electrophoresis and Western Blotting

Protein separation was performed with the NuPAGE® electrophoresis system, using 4-12% NuPAGE and MOPS SDS buffer, as described [9]. Proteins were stained by SYPRO™ Ruby Protein Gel Stain to evaluate and compare the protein profiles, or were transferred to nitrocellulose membranes using a "tank" electrophoretic transfer apparatus (Hoefer), to detect typical exosome markers (TSG-101, flotillin 1) and glomerular proteins (Nephrin, podocin, TRPC6. P-glycoprotein, CD10). The blots were developed as described [9]. Densitometric analysis was performed by ImageQuant TL software (GE Healthcare).

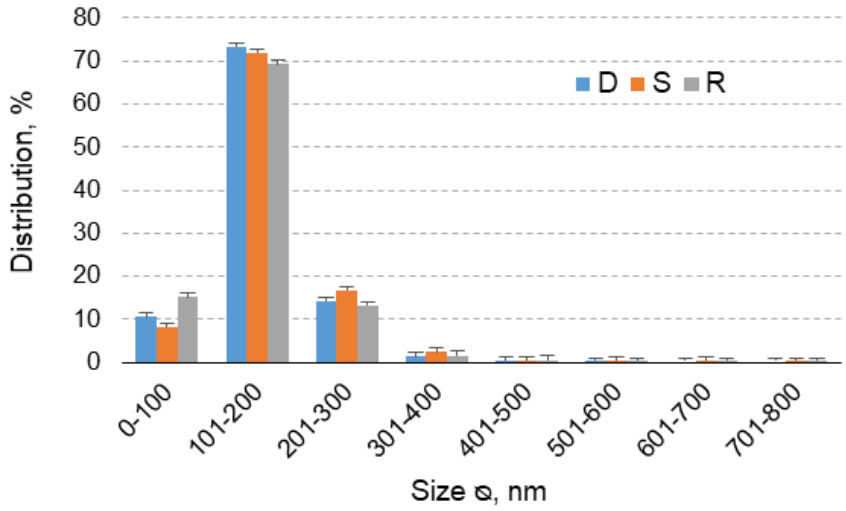
2.6 Experimental Section

Milli-Q water was used for all solutions. PBS (EuroClone), Bovine serum Albumin (BSA), BCA protein assay, methanol, CAPS / SYPRO Orange Protein Gel Stain and ZnSO4 were from SIGMA Chemical Co. (St. Louis, MO, USA); Hybond-ECL nitrocellulose membrane was from GE (Little Chalfont, Buckinghamshire, UK). NuPAGE® SDS-PAGE Gel Electrophoresis System components (mini gels, running and loading buffers, molecular weight markers and coomassie blue staining) were supplied by Life Technologies (Paisley, Renfrewshire, UK). Anti-protease inhibitor cocktail (Complete) was from Roche (Monza, Italy).
anti-Tumor Susceptibility Gene 101 (TSG101) polyclonal antibody (pAb) from Abcam (Cambridge, UK) CD9 come markers e CD10 come differenziale. Species-specific secondary peroxidase conjugated antibodies and ECL reagents were from Thermo Scientific (USA).

3. RESULTS

3.1 UEv characterization: NTA and marker enrichment

We isolated UEv from all samples by an optimized protocol for small volume of urine. NTA shows that the size distribution is typical of UEv [11] and not significantly different among the three groups (Figure 1), while UEv concentration shows different values according to the subgroup.



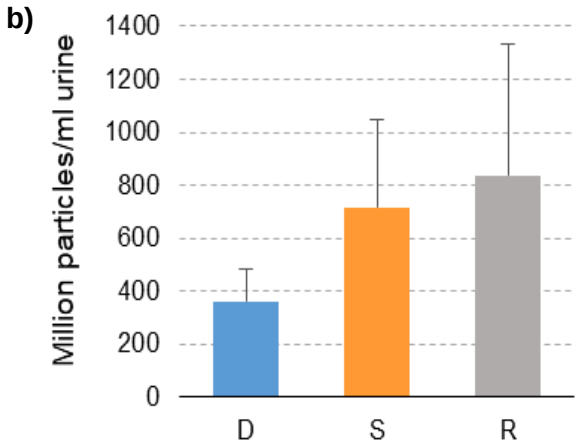


Figure 1. NTA: a) Percentage distribution of UEv isolated from the urine of ISN patients. Mean ± SE; a) UEv concentration. Mean n of 3 samples for each group.

UEv purity was checked by evaluating two commonly used UEv markers, TSG101 and CD9 [12]. Results show that markers are reproducibly enriched in the vesicle fraction, in all the three patient groups (Figure 2), although with some inter-individual variability. The assessment of UE protein markers was extended to UEv isolated from all cases and demonstrated that their purity was comparable in all the preparations.

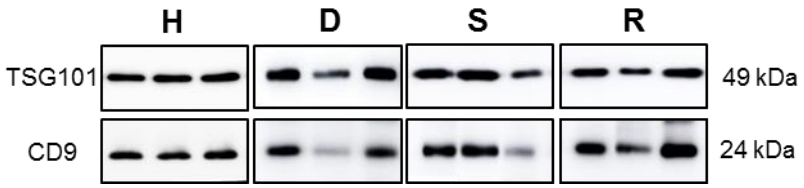


Figure 2: TSG101 e CD9 marker enrichment. Western blotting with anti-TSG101 and anti-CD9. Three representative cases for each group are shown. Lanes were loaded with equal protein amounts.

3.2 UEv protein profiling

The protein pattern of the isolated vesicles was analysed by NuPAGE, followed by SYPRO™ Ruby Protein Gel Stain (Figure 3). UEv isolated from the patients belonging to the three groups showed a peculiar protein pattern: apart from THP band (visible at 80-100 kDa) which has a highly variable level of expression in each patient, regardless of whether they belong to any of the groups, some other bands are rather specifically associated to the profile of each group (Figure 3a).

In order to highlight the specificity of the INS SDS-PAGE pattern, a panel of UEv protein profiles obtained from healthy children and paedriatic patients affected by hereditary tubulopathies (Figure 3c) are shown for comparison. Results show that each INS gel profile preserves its peculiarity and substantially differs to UEv protein content of not-INS patients and healthy subjects.

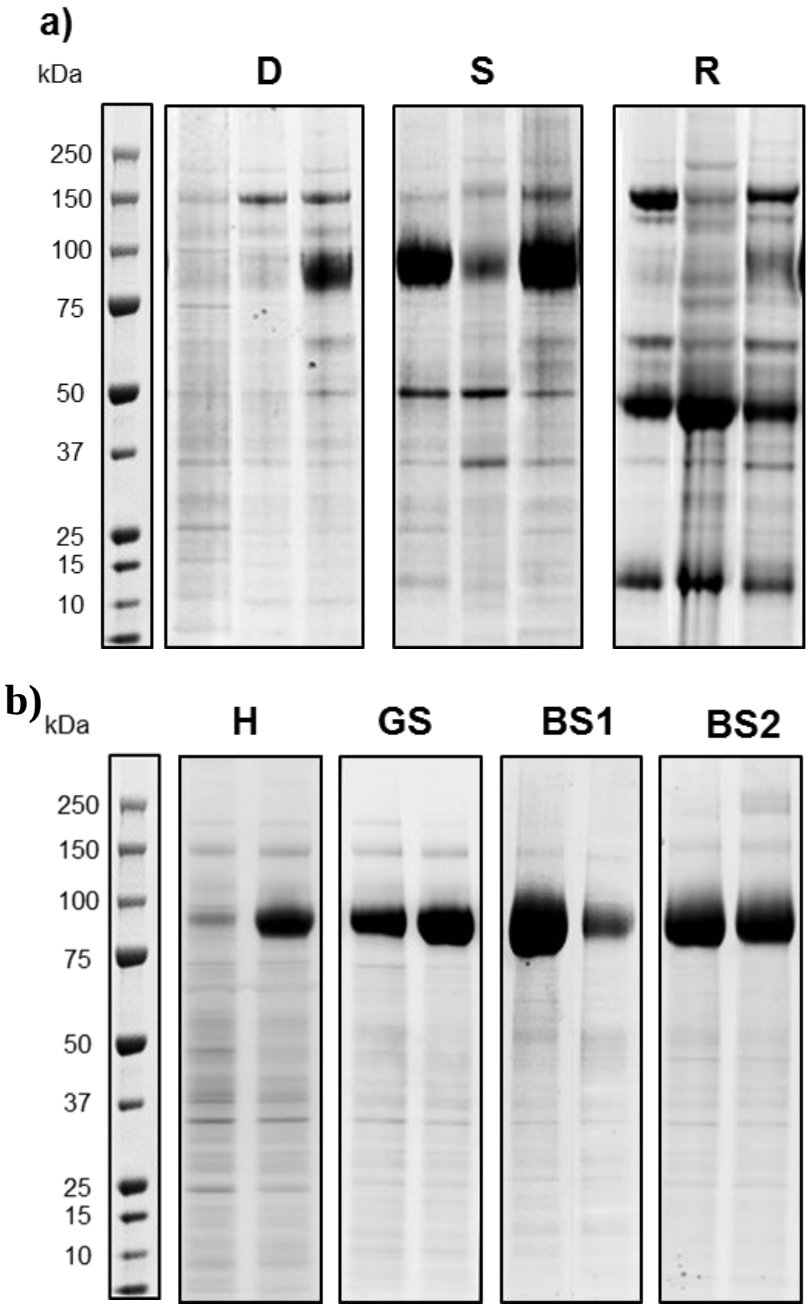


Figure 3: UEv protein profile. NuPAGE 4-12% electrophoresis and Sypro Ruby protein gel staining. a) Patients affected by INS: corticosteroid-dependent (D), corticosteroid-sensitive (S), corticosteroid-resistant (R); three representative cases for each patient group are shown. b) Healthy subjects (H), patients affected by hereditary tubulopathies: Gitelman syndrome (GS), Bartter syndrome type 1 (BS1) and type 2 (BS2); two representative cases for each patient group are shown [13]. UEv protein profiles correspond to 3 ml of starting urine.

Moreover, we were able to collect a follow-up sample for UEv isolation and protein profile analysis, in a subgroup of six patients (Table 2). At the 2nd access, the patients were mainly in the same condition as before regarding pharmacological response. The results of the profile comparison show that the 2nd profile is very similar to the first, considering some variability effects, confirming the reproducibility of UEv protein profile (Figure 4a). In two cases (41-45; 33-71), the initial response to drug changed, and the patients needed to be re-classified. Interestingly, the first profiles of these cases (41 and 33) were somehow different from that typical of their group (Figure 4b). This finding would further confirm the specificity of UEv protein signature related to patient state, as if there were UEv profile characteristics, indicating the future modification of clinical course (S to D or R to D). In fact, among the R group, some patients presented UEv 1D profile similar each other (Figure 4c- R' subgroup), but different from the typical pattern shown in figure 3a. Although these are preliminary data, since the patients 3 and 41 changed their response to therapy, becoming D, it is possible that this type of UEv profile could be predictable of this behaviour.

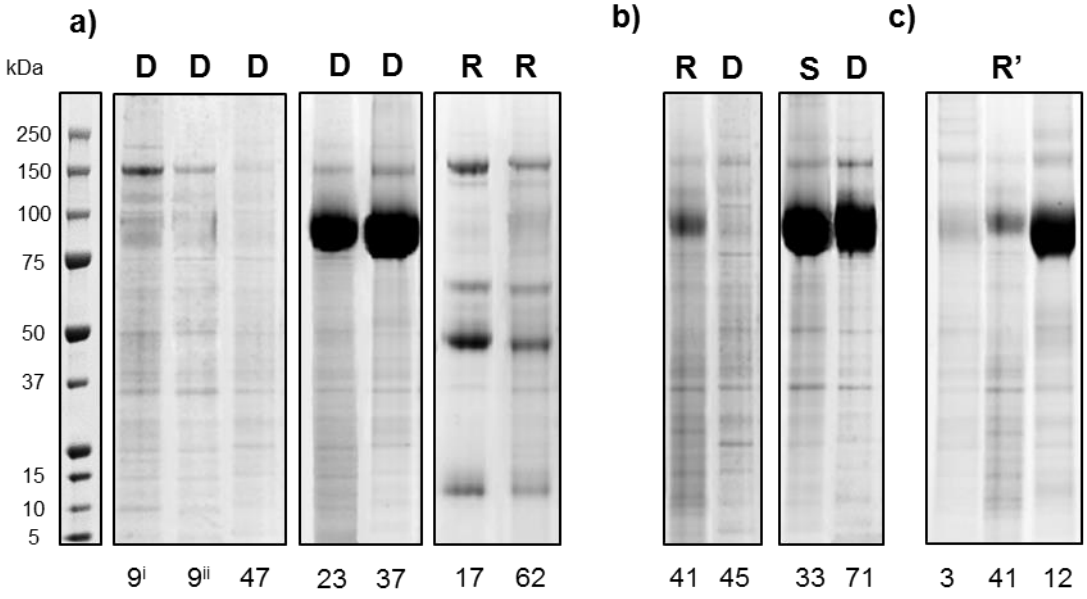


Figure 4: Reproducibility and atypical UEv protein profile. NuPAGE 4-12% electrophoresis and Sypro Ruby protein gel staining. a) UEv protein profile reproducibility: UEv isolated from patients at different time of collection; b) UEv protein profile of patients that changed their response to therapy; c) atypical UEv protein profile of R patients (R'). D, corticosteroid-dependent; S, corticosteroid-sensitive; R, corticosteroid-resistant; UEv protein profiles correspond to 3 ml of

starting urine (see Table 2).

Table 2. List of patients with a follow-up samples

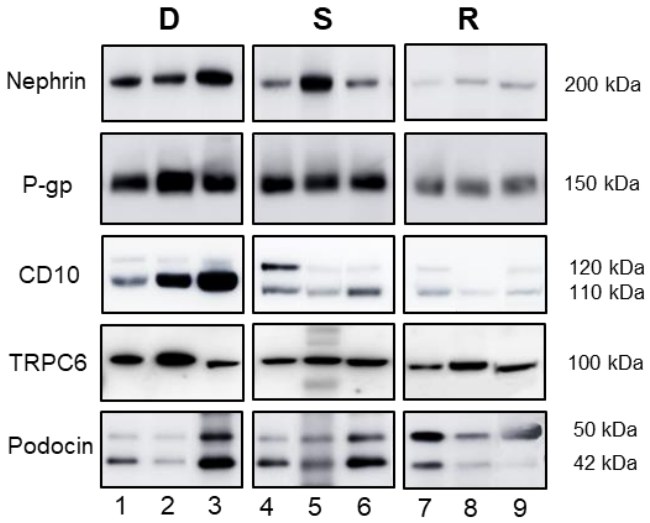
Patient	Group	Sex	Interval (months)	uPr/uCr	Ongoing Therapy			
					Pred	Cycl	MMF	Tacr
9i	D	M		0.18			X	
9ii	D		4	0.18			X	
47	D		10	0.18			X	
17	R	M	8	0.44	X			X
62	R			0.32				X
20	D	M	10	0.17			X	
72	D			0.13			X	
41	R	F	2	0.20				X
45	D			0,14				X
23	D	F	5	0.34			X	
37	D			0.30			X	
33	S	M	4	0.24	X			
71	D			0.20			X	

Legend: D, corticosteroid-dependent; S, corticosteroid-sensitive; R, corticosteroid-resistant; PRED, Prednisone; CYCL, Cyclosporin; MMF, Mycophenolate mofetil; TACR, Tacrolimus; uPr, urinary proteins; uCr, urinary creatinine

3.3 Evaluation of potential INS markers by Western blotting

To validate the differential proteomic profiles of UEv from INS patient's different group, according to their pharmacological response, we examined some specific protein levels in UEv. Protein selection was mainly based on their potential roles to INS pathogenesis, according to literature [14]–[17] and on their tissue expression, typically glomerular. Figure 5a shows the preliminary results for 5 glomerular proteins, nephrin, P-glycoprotein (P-gp), neprylisin (CD10), transient receptor potential cation channel subfamily C member 6 (TRPC6) and podocin, whose UEv differential content was investigates by Western blot analysis. After densitometric quantification of band intensity, results were expressed as OD/mg of UEv proteins in figure 5 b.

a)



b)

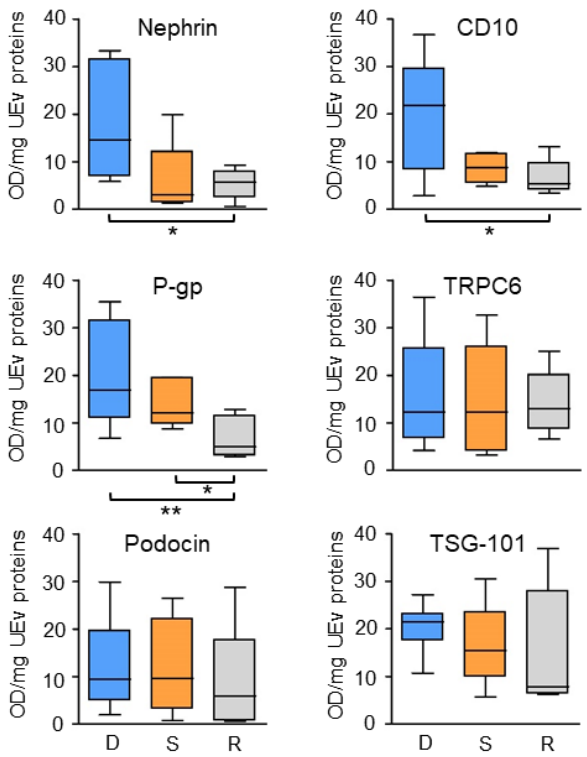


Figure 5. Glomerular protein level in INS UEv. a) Western blotting analysis; 3 representative cases are shown; b) densitometric analysis of bands (Optic Density, OD) normalized by UEv protein content (mg). Mann Whitney *t*-test was performed: * = $p < 0.05$, ** = $p < 0.005$.

Statistical analysis by Mann Whitney test shows that CD10, nephrin and P-gp UEv levels are significantly different among D and R patients, with a higher content in UEv of D patients. Moreover, P-gp seems to differentiate also S and R patients. In general, it seems to be a trend of decrease of these protein levels from D to S and to R group, suggesting the possibility of a bias in the analysis. However, TRPC6, podocin and the exosomal marker TSG-101 do not show the same comportment: these proteins result unchanged, supporting the hypothesis that the different levels of CD10, nephrin and P-gp is a specific signature of the different therapy response of INS patients.

4. DISCUSSION AND CONCLUSION

Given the potential offered by UEv, we explored their role as biomarkers source in children with INS. Nowadays, the nucleic acids present in the UEv of children affected by INS were mainly investigated. Chen et al., for example, focused on exosomal microRNA (miR), showing the alteration of specific miR (miR-194-5p and miR-23b-3p) in response to treatment in children INS affected. This finding revealed that miR could be promising biomarkers for predicting and monitoring patients undergoing severe complications [18].

While genomic analysis has attracted most attention and received major efforts, few proteomics studies have specifically targeted the role of UEv in the INS. After a period in which the EUv *proteomic potential* for the study of INS was considered but not investigated, recently two works proposed Wilms tumor 1 (WT1) transcriptor factor, a well-known marker for differentiated podocytes, as a non-invasive biomarker for the detection of podocyte injury, predicting either therapy responsiveness or monitoring progression in patients with NS (FSGS and SSNS) [19].

In addition, Rood et al. proposed a method based on ultracentrifugation and size exchange chromatography to overcome the problem, allowing detecting of lower abundant UEv proteins [7]. However, the protocol was laborious, and did not have

a follow-on, at least with regard to INS.

Since proteinuria is a negative interferer for UEv purification, we decided to investigate only patients with uPr/uCr ratio < 1mg/mg. In particular, we focused the attention to the response to initial treatment with corticosteroids, the main indicator of long-term prognosis, as steroid-resistant patients often progress to end-stage renal disease [20]. Patients were classified according to the response to the initial steroid therapy and the need for further immunosuppressive treatment into SSNS, SDNS, SRNS. This allowed us to obtain good quality UEv preparation. In fact, the isolated UEv presented the typical size distribution of urinary exosomes and the enrichment of the exosomal markers (TSG101 and CD9) [12]. The SDS-PAGE protein profile of UEv is usually characteristic and different from the total urine one: it was confirmed also for our samples. Moreover, it was clear that each patient group UEv had a peculiar protein band pattern, considering the biological variability. Furthermore, these protein profiles resulted specific for INS diseases, since they was different from the UE protein content of non-INS patients (hereditary tubulopathies) and healthy age-matched controls [13]. In addition, we confirmed that the protein profile remained constant over time, indicating a good reproducibility, as shown for the UEv isolated from urine collected from the same patients several months later. From a careful observation of the bands, the group R presented some abundant bands that correspond to the molecular weights typical of the heavy and light chains of the immunoglobulins. This could be expected since is a peculiarity of this type of patients due to the disease. However, not all the R patients showed the enrichment of the same signals: we noticed that among the 7 patients enrolled as clinically R at the time of urine collection, 3 had a different protein pattern. Interestingly, two of these patients changed their response to the therapy, becoming drug-dependent (D). Although the low number of cases, it was an intriguing finding, that gave us confidence in the usefulness of UEv in the investigation of INS.

The analysis of specific proteins also enforced this results. We checked the UEv content of some glomerular protein, known as involved in INS and present in

extracellular vesicles [21]: P-glycoprotein (P-gp), nephrin, neprylisin (CD10), transient receptor potential cation channel subfamily C member 6 (TRPC6) and podocin. These proteins were analysed with the intent to prove the specificity of INS UEv protein content, without any aim of functional or molecular explanation, that will be explored in future. However, P-gp, CD10 and nephrin seem to differentiate D from R patients, while TRPC6 and podocin show unchanged levels, further indicating a specific UEv signature of the different therapy response of INS patients.

Although preliminary, these evidences confirmed that the proteomic UEv approach is promising in the study of INS to identify physiopathological differences underlying the response to therapy. To reach this goal, we need digging deeper the proteome of the INS UEv, performing MS analysis to uncover the emerged differences in the UEv proteome of the pilot cohort. Simultaneously, the sample size will have to be enlarged in order to develop a robust classification model (in collaboration with the statistics unit), able to use UEv as predictive parameter of long term prognosis to treatment response.

References

[1] D. G. Noone, K. Iijima, and R. Parekh, "Idiopathic nephrotic syndrome in children.," *Lancet (London, England)*, vol. 392, no. 10141, pp. 61–74, Jul. 2018.

[2] M. Colucci, G. Corpetti, F. Emma, and M. Vivarelli, "Immunology of idiopathic nephrotic syndrome.," *Pediatr. Nephrol.*, vol. 33, no. 4, pp. 573–584, Apr. 2018.

[3] C. Dossier, A. Jamin, and G. Deschenes, "Idiopathic nephrotic syndrome: the EBV hypothesis.," *Pediatr. Res.*, vol. 81, no. 1–2, pp. 233–239, Jan. 2017.

[4] E. T. McCarthy, M. Sharma, and V. J. Savin, "Circulating permeability factors in idiopathic nephrotic syndrome and focal segmental glomerulosclerosis.," *Clin. J. Am. Soc. Nephrol.*, vol. 5, no. 11, pp. 2115–

2121, Nov. 2010.

[5] C. E. Sadowski *et al.*, “A single-gene cause in 29.5% of cases of steroid-resistant nephrotic syndrome.,” *J. Am. Soc. Nephrol.*, vol. 26, no. 6, pp. 1279–1289, Jun. 2015.

[6] D. Karpman, A.-L. Stahl, and I. Arvidsson, “Extracellular vesicles in renal disease.,” *Nat. Rev. Nephrol.*, vol. 13, no. 9, pp. 545–562, Sep. 2017.

[7] I. M. Rood *et al.*, “Comparison of three methods for isolation of urinary microvesicles to identify biomarkers of nephrotic syndrome.,” *Kidney Int.*, vol. 78, no. 8, pp. 810–816, Oct. 2010.

[8] R. S. Lee, F. Monigatti, A. C. Briscoe, Z. Waldon, M. R. Freeman, and H. Steen, “Optimizing sample handling for urinary proteomics,” *J. Proteome Res.*, vol. 7, no. 9, pp. 4022–4030, 2008.

[9] M. Pitto, S. Corbetta, and F. Raimondo, “Preparation of urinary exosomes: methodological issues for clinical proteomics.,” *Methods Mol. Biol.*, vol. 1243, pp. 43–53, 2015.

[10] Z. Liu *et al.*, “Isolation and characterization of human urine extracellular vesicles,” *Cell Stress Chaperones*, vol. 23, no. 5, pp. 943–953, 2018.

[11] J. Webber and A. Clayton, “How pure are your vesicles?,” *J. Extracell. Vesicles*, vol. 2, no. 1, pp. 1–6, 2013.

[12] J. Kowal *et al.*, “Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 113, no. 8, pp. E968–E977, 2016.

[13] S. Corbetta *et al.*, “Urinary exosomes in the diagnosis of Gitelman and Bartter syndromes.,” *Nephrol. Dial. Transplant*, vol. 30, no. 4, pp. 621–630, Apr. 2015.

[14] S. R. Hingorani, L. S. Finn, J. Kowalewska, R. A. McDonald, and A. A. Eddy, “Expression of nephrin in acquired forms of nephrotic syndrome in childhood.,” *Pediatr. Nephrol.*, vol. 19, no. 3, pp. 300–305, Mar. 2004.

[15] H. S. Badr, M. A. El-Hawy, and M. A. Helwa, “P-Glycoprotein Activity in Steroid-Responsive vs. Steroid-Resistant Nephrotic Syndrome.,” *Indian J.*

Pediatr., vol. 83, no. 11, pp. 1222–1226, Nov. 2016.

[16] B. B. Joshi, P. G. Koringa, K. N. Mistry, A. K. Patel, S. Gang, and C. G. Joshi, "In silico analysis of functional nsSNPs in human TRPC6 gene associated with steroid resistant nephrotic syndrome.," *Gene*, vol. 572, no. 1, pp. 8–16, Nov. 2015.

[17] S. K. N. Mulukala, R. Nishad, L. P. Kolligundla, M. A. Saleem, N. P. Prabhu, and A. K. Pasupulati, "In silico Structural characterization of podocin and assessment of nephrotic syndrome-associated podocin mutants.," *IUBMB Life*, vol. 68, no. 7, pp. 578–588, Jul. 2016.

[18] T. Chen *et al.*, "Increased urinary exosomal microRNAs in children with idiopathic nephrotic syndrome.," *EBioMedicine*, vol. 39, pp. 552–561, Jan. 2019.

[19] H. Zhou *et al.*, "Urinary exosomal Wilms' tumor-1 as a potential biomarker for podocyte injury.," *Am. J. Physiol. Renal Physiol.*, vol. 305, no. 4, pp. F553-9, Aug. 2013.

[20] A. Pasini *et al.*, "The Italian Society for Pediatric Nephrology (SINePe) consensus document on the management of nephrotic syndrome in children: Part I - Diagnosis and treatment of the first episode and the first relapse.," *Ital. J. Pediatr.*, vol. 43, no. 1, p. 41, Apr. 2017.

[21] M. C. Hogan *et al.*, "Subfractionation, characterization, and in-depth proteomic analysis of glomerular membrane vesicles in human urine.," *Kidney Int.*, vol. 85, no. 5, pp. 1225–1237, May 2014.

CHAPTER 6

SUMMARY, CONCLUSION AND FUTURE PROSPECTIVE

1. SUMMARY AND CONCLUSIONS

1.1 Distinct urinary proteomic and glycoproteomic signatures in the ccRCC progression

Approximately 90% of all kidney cancers are RCC; this group comprises different types with specific histopathological and genetic characteristics. Clear cell Renal Cell Carcinoma (ccRCC) is recognised as the most aggressive morphotype. Moreover, it tends to metastasise before giving any local sign or symptom; in more than one third of patients, the tumor is already metastatic at diagnosis [1], [2].

Furthermore, in contrast to other urological tumours, ccRCC is both chemo- and radio-resistant. Therefore, surgical resection seems to be the only effective therapy for this disease to date [3]. The internationally accepted classification WHO proposes the attribution of the stage based on the TNM system (made on tissue after partial or total nephrectomy) as the most relevant prognostic factor. However, many works, integrating together the diverse clinical, histological and molecular aspects, shown that tumours with similar characteristics could present very dissimilar evolutions and not always, the aggressiveness correlates with the stage classification [4].

For all these reasons, in several laboratories around the world and in ours, many studies were conducted in order to discovery specific "cancer features" (biomarkers), able to predict the tumour evolution after the surgery and to suggest the most appropriate clinical behaviour and pharmacological treatment to adopt. Currently, no adequate treatment or biological fluid screening tests exist for these patients. Thus, there is a clear need for the development of novel and convenient diagnostic and prognostic tests as well as for biomarkers readily accessible from biological fluids to aid in monitoring.

For our study, the model of ccRCC investigation is represented by the urine, a readily accessible waste bodily fluid, used as liquid tumour biopsy. Liquid biopsy consents the analysis of tumour material from minimally invasive or non-invasive methods utilizing biological fluids as samples, like the urine. It can include the

measurement of soluble factors, such as proteins, providing in addition a more feasible serial analysis of the cancer, including any evolution of protein landscape and development of treatment resistance over time [5], [6]. Therefore, molecular information derived from liquid biopsy is a promising area of research that has the potential to improve not only cancer detection but also tailored treatment. In this context, proteomics has emerged as a promising tool, and a growing number of groups have been using this approach in ccRCC biomarker discovery-driven research.

The work presented here is set in this context. We initially focused our investigation on the urinary proteome of ccRCC patients in order to verify the hypothesis that the urine could *proteomically* reflect the morphological features (stage and grade) used by the pathologist to define the tumour, and in the same time, in order to confirm and predict the kidney impairment and/or the cancer progression. Collectively, we revealed a set of proteins with an altered urinary expression according to tumour spread. Several of them were already known to be involved in carcinogenesis, tumour progression and aggressiveness in RCC and in other cancer types. These results lead us to show that the urinary proteome of ccRCC patients is influenced primarily by the dimension and position (stage), and partially by the grade of the tumour. They also support the hypothesis that a panel of urinary proteins could be able to reflect the clinicopathological characteristics of ccRCC and the corresponding possible combinations of these two histopathological features, overcoming and complementing the static stage and grade classifications.

Following in this vein, we in depth analyse the urine N-glycoproteome of ccRCC patients, leading the investigation level one-step forward. In this way in fact, we integrated the proteomic information with the glycoproteomic one, in order to provide a more comprehensive snapshot of the cancer landscape, directly related to the molecular content of the urine. The glycoproteomic approach, in particular, allowed evaluating the alterations of the N-glycoproteins associated with the histological cancer status (the stage feature). We essentially highlighted abnormal changes affecting the occupancy of the glycosylation sites, the so-called glycan

micro-hetereogenity, of the proteins present in the urine of ccRCC patients compared to the non-affected ones. The N-glycosylation levels in the urine seems to following the cancer progression, reflecting in a specific way the pT1, pT3 ccRCC stages. In addition, we identified and quantified a panel of N-glycoproteins, with a specific abundance trend, directly related to the stage progression as potential glycomarkers. It resulted intriguing that some of these proteins had an abundance pattern that differs in a mirror fashion from the abundance pattern of its corresponding non-glycosylated isoform (as described in literature), suggesting for these N-glycoforms a different role in the process of tumourigenesis. It is well-known that unusual glycoforms causes severe defects in protein localisation, trafficking, cellular adhesion and trasdutional pathways in disease conditions, which may lead to valuable diagnostic markers [7]–[9].

Our results represent only the tip of the iceberg. Along with the molecular pathways that we evidenced altered according to the presence of the N-glycomodification, this study can represent a starting point for future investigations aimed at clarifying mechanism triggers such unconventional glycosylation in the ccRCC development.

Recently, the UEv emerged as an appealing source for biomarker discovery. Secreted by every epithelial cell type lining the urinary tract system in human, they contain molecular constituents of their cell of origin, including proteins and genetic materials, and can be isolated in a non-invasive manner [10]. Since their capability to mirror the molecular content of the cell origin, UEv are more likely to reflect the cellular processes associated with the pathogenesis of urologic and renal diseases compared, to urine. Following the discovery of urinary exosomes in 2004, many studies have been performed using urinary exosomes as a starting material to identify biomarkers in various renal, urogenital, and systemic diseases [11]. For this reason, several proteomic studies on UEv have been performed to identify biomarkers for urinary tract diseases, both in experimental and clinical settings.

1.2 Technical challenges in the UEv proteomic study

The use of UEv in biomedical research is hampered by the lack of efficient methods for vesicle isolation and purification, representing the major critical points in the UEv proteomics applied to clinics. Moreover, in some conditions, urine samples present elements that can interfere with UEv isolation and their proteomic analysis. This can occur both in physiological conditions, given the abundant presence of Uromodulin (UMOD), the most abundant urinary glycoprotein, and in pathological conditions, in the case of patients with renal diseases, frequently presenting high levels of proteinuria and haematuria. For this reason, development and optimization of protocols for UEv isolation are fundamental of our research in order to reduce the impact of the interferers and to obtain properly sample for proteomic analysis.

We firstly focused the attention on the resolution of the haematuria negative effects on the UEv proteomic profile, proposing a modified protocol based on trypsin treatment able to reduce the impact of blood contamination on UEv isolation and purity. After the enzymatic treatment, in fact, we observed that the UEv protein profile of contaminated samples shifted to the typical healthy subjects UEv profile. In addition, the enrichment of the typical UEv markers was confirmed and the MS-based label-free analysis of the UEv protein content demonstrated that it substantially did not change, proving the feasibility of the approach. This treatment allows recovering several samples that would be otherwise discarded, such as those coming from patients affected by diseases of the urinary tract with haematuria, especially when investigating the differential abundance of proteins specific of UEv or of the urinary tract.

In conclusion, our modified protocol to UEv preparation allows proteomic analysis of vesicles also in haematuria conditions, often encountered in kidney and urological diseases, as shown by preliminary results on UEv isolated from patients presenting haematuria. By adopting this protocol, the urinary extracellular vesicles isolated from haematuria urine may reveal a differential protein cargo, a source of promising biomarkers for the diagnosis and monitoring of renal diseases.

1.3 UEv in Idiopathic Nephrotic Syndrome

Given the potential offered by UEv, we also explored the possibility to use UEv as a source of clinical information in the INS, the major childhood glomerular disease. Response to initial treatment with corticosteroids is an indicator of long-term prognosis, as resistant patients often present progressive disease [12]. While exosome research has flourished, few studies have specifically targeted the role of UEv, mainly focusing on miRNA content.

The aim of the study was to verify the feasibility to use UE of INS patients as a source of predictive markers of response to corticosteroid treatment, and/or for clarifying the disease etiopathogenesis and/or the pharmacoresistance. Thus, we investigated the UE protein content of a paediatric cohort of 34 patients, classified in three clinical classes, according to the corticosteroids response. Firstly, a characteristic SDS-PAGE protein profile emerged to be associated with each class, which preserves its peculiarity also when compared to UE protein content of non-INS patients (orthostatic proteinuria, hereditary tubulopathies) and healthy age-matched controls. This result is enforced by the reproducibility of UEv profiles demonstrated in patients who have undergone two or three repeated urine tests after a few months. Moreover, within the same class of patients, the resistant ones, it was possible to recognise two subgroups of protein profile, one defined as typical and the other one as atypical. Interestingly this type of clusterization seems to correspond to specific behaviour: in fact, two of the three patients showing the atypical UEv profile changed their response to the therapy, becoming drug-dependents. Secondly, we also pinpointed different levels of specific glomerular membrane proteins described as involved in the INS development, indicating a specific UEv signature of the different therapy response of INS patients.

These evidences confirmed that the proteomic UEv approach is promising in the study of INS to identify physiopathological differences underlying the response to therapy.

2. FUTURE PERSPECTIVE

Nowadays, we are proceeding in the study of the UEv protein content of INS, performing MS analysis to more in depth comprehend the emerged differences in the UE proteome of the pilot cohort. Simultaneously, we are enlarging the sample size in order to develop a robust classification model, able to use UEv as predictive parameter of long term prognosis to treatment response.

Contextually, we are studying the effects of proteinuria on the UEv proteome, in order to find a practical solution to reduce its misleading impact. These highly abundant proteins interfere with both the UEv isolation and the microvesicular protein identification by proteomic techniques, complicating the search for prognostic biomarkers [13]. This measure is required, given that the massive loss of the proteins in the urine is a condition closely associated with several renal pathologies, including INS at the clinical onset.

At same time, we continuing the study of ccRCC, focusing the investigation from the urine to UEv and their glycoprotein cargo as a source of prognostic biomarkers too. It is reported that UEv present a glycan signature, used as specific signalling feature. For example, a disease-specific glycosylation profile of UEv was detected in autosomal dominant polycystic kidney disease [14]. These findings suggest that better understanding the glycomics and glycoproteomics not only will provide information about the functional state of constituent proteins, but it will also highlight the comparison among proteins that are specifically targeted to UEv, especially in pathological conditions.

Nevertheless, current MS approaches for the UEv glycoproteome investigation are limited by the low amount of UEv proteins; furthermore, the presence of UMOD, the most abundant glycoprotein in urine, hampers the UEv recovery and the study of the other glycoproteins. These evidences arise a double technical challenge: primarily, the reduction of UMOD contamination, and secondly the search of a suitable strategy to investigate the UE glycoproteome. Thus, we tested different experimental procedures, starting from normal urine specimens, in order to improve

the sample preparation and achieve a better MS analysis.

3. CONCLUSION

Altogether, the presented works highlight the role of urinary components in the study of different renal diseases, from the diagnostic, prognostic and predictive point of view to the etiopathogenesis one. These results are further evidence that the urine ranges from being merely body waste to liquid biopsy, a sample readily accessible by non-invasive methods. In particular, the UEv represent a further level of investigation that best represents the cellular component of renal tissue. Enriched in renal proteins, they are functional snapshots of the kidneys and their state. There is a great deal of evidence, including our own data, to show that the UEv proteome can reflect the specific phenotype of patients, and its study could also be a promising tool for personalised medicine.

[1] C. J. Creighton *et al.*, "Comprehensivemolecular characterization of clear cell renal cell carcinoma," *Nature*, vol. 499, no. 7456, pp. 43–49, 2013.

[2] G. Shaw, "The silent disease.," *Nature*, vol. 537, no. 7620, pp. S98-9, Sep. 2016.

[3] G. Lewis and A. P. Maxwell, "Early diagnosis improves survival in kidney cancer.," *Practitioner*, vol. 256, no. 1748, pp. 2,13-16, Feb. 2012.

[4] C. Chinello *et al.*, "The proteomic landscape of renal tumors," *Expert Review of Proteomics*. 2016.

[5] D. Quandt *et al.*, "Implementing liquid biopsies into clinical decision making for cancer immunotherapy," *Oncotarget*, vol. 8, no. 29, pp. 48507–48520, 2017.

[6] A. Di Meo, J. Bartlett, Y. Cheng, M. D. Pasic, and G. M. Yousef, "Liquid biopsy: A step forward towards precision medicine in urologic malignancies," *Molecular Cancer*, vol. 16, no. 1. 2017.

[7] S. R. Stowell, T. Ju, and R. D. Cummings, "Protein glycosylation in cancer.,"

Annu. Rev. Pathol., vol. 10, pp. 473–510, 2015.

[8] S. S. Pinho and C. A. Reis, "Glycosylation in cancer: Mechanisms and clinical implications," *Nat. Rev. Cancer*, vol. 15, no. 9, pp. 540–555, 2015.

[9] N. Taniguchi and Y. Kizuka, "Glycans and cancer: Role of N-Glycans in cancer biomarker, progression and metastasis, and therapeutics," in *Advances in Cancer Research*, 2015.

[10] J. M. Street, E. H. Koritzinsky, D. M. Glispie, R. A. Star, and P. S. T. Yuen, "Urine Exosomes: An Emerging Trove of Biomarkers.," *Adv. Clin. Chem.*, vol. 78, pp. 103–122, 2017.

[11] D. Karpman, A.-L. Stahl, and I. Arvidsson, "Extracellular vesicles in renal disease.," *Nat. Rev. Nephrol.*, vol. 13, no. 9, pp. 545–562, Sep. 2017.

[12] A. Pasini *et al.*, "The Italian Society for Pediatric Nephrology (SINePe) consensus document on the management of nephrotic syndrome in children: Part I - Diagnosis and treatment of the first episode and the first relapse.," *Ital. J. Pediatr.*, vol. 43, no. 1, p. 41, Apr. 2017.

[13] P. Gonzales, T. Pisitkun, and M. A. Knepper, "Urinary exosomes: is there a future?," *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association*, vol. 23, no. 6. England, pp. 1799–1801, Jun-2008.

[14] J. Q. Gerlach *et al.*, "Surface glycosylation profiles of urine extracellular vesicles.," *PLoS One*, vol. 8, no. 9, p. e74801, 2013.