www.elsevier.com/locate/talanta

Glycan affinity magnetic nanoplatforms for urinary glycobiomarkers discovery in bladder cancer

Rita Azevedo, Janine Soares, Cristiana Gaiteiro, Andreia Peixoto, Luís Lima, Dylan Ferreira, Marta Relvas-Santos, Elisabete Fernandes, Ana Tavares, Sofia Cotton, Ana Luísa Daniel-da-Silva, Lúcio Lara Santos, Rui Vitorino, Francisco Amado, José Alexandre Ferreira

PII: S0039-9140(18)30271-6
DOI: https://doi.org/10.1016/j.talanta.2018.03.028
Reference: TAL18462

To appear in: *Talanta*

Received date: 27 December 2017
Revised date: 7 March 2018
Accepted date: 11 March 2018

Cite this article as: Rita Azevedo, Janine Soares, Cristiana Gaiteiro, Andreia Peixoto, Luís Lima, Dylan Ferreira, Marta Relvas-Santos, Elisabete Fernandes, Ana Tavares, Sofia Cotton, Ana Luísa Daniel-da-Silva, Lúcio Lara Santos, Rui Vitorino, Francisco Amado and José Alexandre Ferreira, Glycan affinity magnetic nanoplatforms for urinary glycobiomarkers discovery in bladder cancer, *Talanta,* https://doi.org/10.1016/j.talanta.2018.03.028

Glycan affinity magnetic nanoplatforms for urinary glycobiomarkers discovery in bladder cancer

Rita Azevedo[a,b*], Janine Soares[a*], Cristiana Gaiteiro[a,b], Andreia Peixoto[a,b,c], Luís Lima[a,c,d,e], Dylan Ferreira[a], Marta Relvas-Santos[a], Elisabete Fernandes[a,b,c], Ana Tavares[a], Sofia Cotton[a], Ana Luísa Daniel-da-Silva[f], Lúcio Lara Santos[a,b], Rui Vitorino[g,h], Francisco Amado[i] and José Alexandre Ferreira[a,b,c,d,j]

[a]Experimental Pathology and Therapeutics Group, Research Centre, Portuguese Oncology Institute of Porto (IPO-Porto), R. Dr. António Bernardino de Almeida 62, 4200-162 Porto, Portugal; [b]Institute of Biomedical Sciences Abel Salazar, University of Porto, R. Jorge de Viterbo Ferreira 228, 4050-013 Porto, Portugal; [c]Institute for Research and Innovation in Health (i3S), University of Porto, R. Alfredo Allen, 4200-135 Porto, Portugal; [d]Glycobiology in Cancer, Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP), R. Júlio Amaral de Carvalho 45, 4200-135 Porto, Portugal; [e]Porto Comprehensive Cancer Centre (P.ccc), R. Dr. António Bernardino de Almeida 62, 4200-162 Porto, Portugal; [f]CICECO - Aveiro Institute of Materials, Department of Chemistry, University of Aveiro, 3810-193 Aveiro, Portugal; [g]iBiMED - Institute of Biomedicine, Department of Medical Sciences, University of Aveiro, Campus Universitario de Santiago, Agra do Crasto – Edificio 30, 3810-193 Aveiro, Portugal; [h]Research Unit of the Physiology and Cardiothoracic Surgery Department of the Faculty of Medicine, University of Porto, Alameda Prof. Hernâni Monteiro, 4200-319 Porto, Portugal; [i]Mass Spectrometry Centre, Organic Chemistry and Natural Products Unit, Department of Chemistry, University of Aveiro, 3810-193 Aveiro, Portugal; [j]International Iberian Nanotechnology Laboratory (INL), Avda. Mestre José Veiga, 4715 Braga, Portugal.

Corresponding author:
José Alexandre Ferreira
Experimental Pathology and Therapeutics Group, Research Centre
Portuguese Oncology Institute of Porto (IPO-Porto)
R. Dr. António Bernardino de Almeida 62, 4200-162 Porto, Portugal
Office phone: +351225084000 (ext. 5111); Email: jose.a.ferreira@ipoporto.min-saude.pt

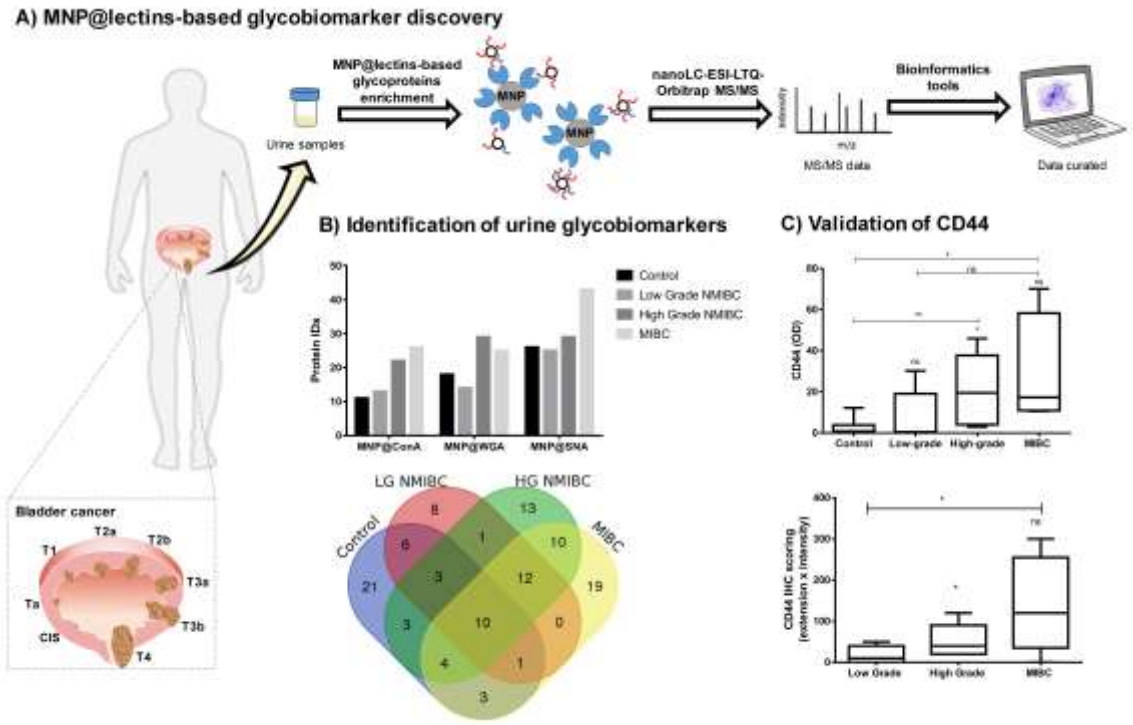
Declarations of interest: none

*Equal contribution

Abstract

Bladder Cancer (BC) presents one of the highest recurrence rates amongst solid tumours and constitutes the second deadliest disease of the genitourinary track. Non-invasive identification of patients facing disease recurrence and/or progression remains one of the most critical and challenging aspects in disease management. To contribute to this goal, we demonstrate the potential of glycan-affinity glycoproteomics nanoplatforms for urinary biomarkers discovery in bladder cancer. Briefly, magnetic nanoprobes (MNP) coated with three broad-spectrum lectins, namely *Concanavalin A* (ConA; MNP@ConA), Wheat Germ Agglutinin (WGA; MNP@WGA), and *Sambucus nigra* (SNA; MNP@SNA), were used to selectively capture glycoproteins from the urine of low-grade and high-grade non-muscle invasive as well as muscle-invasive BC patients. Proteins were identified by nano-LC MALDI-TOF/TOF and data was curated using bioinformatics tools (UniProt, NetOGlyc, NetNGlyc, ClueGO app for Cytoscape and Oncomine) to highlight clinically relevant species. Accordingly, 63 glycoproteins were exclusively identified in cancer samples compared with healthy controls matching in age and gender. Specific glycoprotein sets exclusively found in low-grade non-muscle invasive bladder tumours may aid early diagnosis, while those only found in high-grade non-invasive and muscle-invasive tumours hold potential for accessing progression. Amongst these proteins is bladder cancer stem-cell marker CD44, which has been associated with poor prognosis. Orthogonal validation studies by slot-blotting demonstrated an elevation in urine CD44 levels of high-grade patients, which became more pronounced upon muscle-invasion, in mimicry of the primary tumour. These observations demonstrate the potential of MNP@lectins for identification of clinically relevant glycoproteomics signatures in bladder cancer. Future clinical validation in a larger and well characterized patient subset is required envisaging clinical translation of the results.

Graphical abstract



Keywords:

Glycoproteomics, urine biomarkers, bladder cancer, glycosylation, nanoprobes; nanoparticles

Introduction

Bladder cancer (BC) is the second most deadly malignancy of the urinary tract [1]. Patients diagnosed with low-grade non-muscle invasive BC (NMIBC) generally face better prognosis, whereas those with high-grade lesions are frequently burdened by significant recurrences, generally accompanied by progression to muscle invasion (MIBC) and metastasis [2]. As such, these patients are subjected to several post-operative cycles of intravesical therapy proceeded by intensive and invasive follow-up interventions [2]. The development of non-invasive tools supporting the differentiation between low and high-grade lesions, as well as the early identification of patients at risk of progression, remains a critical aspect for disease management [3].

Urine is easily accessible and, in bladder cancer cases, is in direct contact with the tumour, thereby providing a key source of biomarkers for addressing the above mentioned difficulties [4]. Moreover, the human urine proteome is well characterized, supporting proteomics-based biomarker studies [5]. Interestingly, the most relevant targets arising from these studies include several glycoproteins, such as members of the apolipoprotein family, fibrinogen chains, alpha-1-antitrypsin, alpha-2-macroglobulin, and uromodulin [5, 6]. Nevertheless, these biomarkers may be reflective of haematuria and have also been associated with kidney disease (e.g. apolipoproteins, uromodulin, or alpha-1-antitrypsin) and other pathologies [6]. As such, current urine biomarkers lack the necessary sensitivity and specificity for clinical use; however, their incorporation in multi-biomarker panels may provide the necessary molecular context for more effective applications. However, the lack of efficient enrichment strategies remains one of the most critical aspects underlying the identification of cancer-specific biomarkers, generally present in minute amounts (nano-femtomolar range) in bodily fluids such as urine [7].

Targeting glycan moieties in cancer-associated glycoproteins holds tremendous potential to overcome these limitations, while providing an important source of clinically relevant biomarkers [8]. Moreover, the identification of disease-associated glycoforms may further improve the biomarker potential of already proposed urine glycobiomarkers. Accordingly, several decades of research, mostly using lectins and antibodies for tissue screening, have disclosed a plethora of alterations in membrane proteins glycosylation associated with bladder cancer [3]. Namely, we have reported that disease progression to

invasion and dissemination is accompanied by the overexpression of short-chain sialylated *O*-glycans, associated with poor prognosis [9-11]. Lectin-enrichment targeting these glycans further enabled the identification of clinically relevant glycoproteins in patient samples, including several integrins, cadherins, the stem cell marker CD44 and MUC16 [12, 13]. Taken together, our studies demonstrated the importance of pre-enrichment strategies for downstream comprehensive glycoproteomics approaches envisaging biomarker discovery [12-14]. Concomitantly to *O*-glycan alterations, advanced stage bladder cancer cells also experience *N*-glycome remodelling, including changes in mannose core, branching, elongation and degree of fucosylation and sialylation [3]. However, the nature of the glycoproteins exhibiting these alterations is still poorly understood. Nevertheless, we hypothesize that abnormally glycosylated proteins may be secreted or shed from the primary tumour into bodily fluids, constituting an important source of potentially relevant urine glycobiomarkers.

Building on these insights, our group has developed lectin functionalized magnetic nanoprobes for glycoproteomics research, which have demonstrated enhanced sensitivity in comparison to conventional enrichment strategies [15, 16]. Herein, we aim to explore the potential of these nanoplatforms to highlight the existence of glycoprotein signatures associated with bladder cancer stage and grade. We envisage that this preliminary approach may provide the necessary analytical rationale for more in depth translational studies.

Material and Methods

Patient Samples

Between 2010 and 2011, thirty-one first void urine samples without visible signs of haematuria were prospectively collected before surgery of bladder cancer male patients with mean age of 70 (range, 45–89) years, attending the Portuguese Institute of Oncology of Porto (IPO-Porto; Portugal). Only patients that had not been previously submitted to neoadjuvant therapy were included. Corresponding formalin-fixed paraffin embedded (FFPE) tumours were also obtained for this study. Based on the World Health Organization urothelial carcinoma grading and staging criteria, three different groups were considered in this study, namely low-grade (n=15) and high grade (n=9) non-muscle-invasive bladder tumours (NMIBC) and muscle-invasive (n=7) tumours (MIBC). An additional 15 urines

were obtained from healthy control male volunteers, mean age of 68 (range 41-82) years. All procedures were performed under the approval of the institution ethics committee and upon patients' informed consent. All clinicopathological information was obtained from patients' clinical records.

Isolation and Quantification of Urine Proteins

Five to forty millilitres of collected urine were centrifuged at 5000*g* for 40 min at 4 °C to remove cells and debris, desalted on Amicon Ultra 10 kDa centrifugal filters (Merck KGaA, Darmstadt, Germany) and proteins were resuspended on 50 mM Ammonium bicarbonate (pH 7.8, Sigma-Aldrich, St. Louis, MO, EUA). Protein quantification was accessed using the DC Protein assay (Bio-Rad, Hercules, CA, USA) and protein content was normalized in relation to creatinine, as previously described [17, 18]. Creatinine content was determined using a colorimetric Creatinine Assay Kit (Abcam, Cambridge, UK). Protein/Creatinine ratio was used to disclose sample proteinuria as result of disease.

Total Sialic acids

Total sialic acids (TSA) were determined by fluorimetry as previously described [15, 16]. All solutions were precooled in an ice bath. Twenty microliters of sodium periodate solution (10 mM, Sigma-Aldrich) were added to 30 μl of glycoconjugate sample (10–200 μg) placed in a 2 mL polypropylene test tube. The solution was chilled in the ice bath for 45 min. The reaction was stopped by adding 100 μl of 50 mM sodium thiosulfate (Sigma-Aldrich), 500 μl of 4.0 M ammonium acetate (pH 7.5, Sigma-Aldrich) and 400 μl of an ethanolic solution of 100 mM acetoacetanilide (Sigma-Aldrich), followed by incubation for 10 min at room temperature. The fluorescence intensities of the solution were measured at 471 nm with an excitation wavelength of 388 nm. Detection limits of the determination were obtained according to the relative fluorescence of sample solution to a sample blank prepared under identical conditions. TSA content was normalized in relation to the creatinine content in the samples.

Synthesis of lectin functionalized magnetic nanoparticles (MNP@ConA; MNP@WGA; MNP@SNA)

Lectin functionalized ferromagnetic nanoparticles were synthetized as previously described by Ferreira *et al.* (2011) [15] and Cova M *et al.* (2015) [16]. Briefly, the iron oxide magnetic core of the MNPs was synthesized by coprecipitation of $FeCl_2$ and $FeCl_3$ (Sigma-Aldrich) under alkaline conditions and coated with amorphous silica to prevent particle clustering and ensure their chemical stability. Encapsulation of the magnetic core was achieved by hydrolysis and condensation of tetraethyl orthosilicate (TEOS, Sigma-Aldrich) under alkaline conditions with trimethylamine (Sigma-Aldrich) as catalyst. Subsequently, the silica-coated nanoparticle surface was functionalized with amine groups with 3-aminopropyltrimethoxysilane (APS, Sigma-Aldrich) to reduce nonspecific interactions with the target proteins. Successful synthesis and surface modification of the MNP matrix was ensured by step-by-step monitorization with Fourier transform infrared (FT-IR) spectroscopy. The size and shape of the MNPs was determined by transmission electron microscopy. The MNPs exhibited relatively narrow size distribution and an average diameter of 14±3 nm, thus confirming their nanoscale dimensions. The MNPs were then dispersed in dimethyl sulfoxide (DMSO, Sigma-Aldrich) and their surface was activated with suberic acid bis-N-hydroxysuccinimide ester (DSS, Sigma-Aldrich), washed thoroughly with DMSO, and incubated with *Concanavalin A* (ConA; Vector Laboratories, Burlingame, CA, USA), Wheat Germ Agglutinin (WGA; Vector Laboratories), and *Sambucus nigra* (SNA; Vector Laboratories) (1 mg/mL in binding buffer; Table 1S) at 4 °C for 12 h to produce MNP@ConA, MNP@WGA, and MNP@SNA particles, respectively. Binding of the lectins at the cells surface was confirmed by FT-IR and UV spectroscopy at 280 nm. The amount of bounded lectin was indirectly estimated from the conjugation supernatant.

Urine Glycoprotein enrichment with MNP@lectins

Proteins isolated from the urine of male individuals (n=5) were polled together according to the following groups: i) Controls; ii) low-grade NMIBC; iii) high-grade NMIBC; iv) MIBC. Protein extracts (100 μg) were diluted in 500 μL of lectin binding buffers and incubated with 1 mg of MNP@lectins at 4°C for 30 min. Specifically, ConA binding buffer was 20 mM Tris, 500 mM NaCl, 5 mM $MgCl_2$, 5 mM $MnCl_2$, and 5 mM $CaCl_2$, pH 7.45. WGA binding buffer was 20 mM Tris, 150 mM NaCl, and 0.05% Tween-

20, pH 7.45. SNA binding buffer was 20 mM Tris, 150 mM NaCl, and 0.05% Tween-20, pH 7.45. The MNP were then pulled to the tube walls under the influence of a strong magnetic field, allowing an easy and complete removal of the supernatant. The unbound proteins were removed by washing three times with 200 μL of wash buffer (0.05% Tween-20 in binding buffer). The glycoproteins were recovered with specific elution buffers for each ligand as previously reported [15, 16].

Protein identification and data curation

The proteins recovered from MNP@lectins were first reduced upon incubation with 5 mM dithiothreitol (DTT, Sigma-Aldrich) at 75ºC for 1 h, then alkylated with 20 mM iodoacetamide (Sigma-Aldrich) at room temperature in the dark for 2 h and digested with trypsin (Promega, Madison, WI, USA) for 37.0°C overnight. Protein identification was performed as previously described by our group [16]. Tryptic digests were separated with a C18 Pepmap (Dionex) column on an Ultimate 3000 (Dionex/LC Packings, Sunnyvale, CA) nano-HPLC, and fractions were collected with a Probot (Dionex/LC Packings, Sunnyvale, CA) directly onto a matrix-assisted laser desorption ionization (MALDI) plate. The MALDI-TOF/TOF (time-of-flight) mass spectrometry (MS) analysis was performed on a 4800 MALDI-TOF/TOF Analyzer (Applied Biosystems, Foster City, CA). The MS and MS/MS spectra acquired were processed and analyzed by the Global Protein Server Workstation (Applied Biosystems). LC-MALDI-MS/MS runs were done in triplicates for each MNP@lectin (ConA, WGA, SNA) and experimental condition (control, low-grade NMIBC, high-grade NMIBC; MIBC). Protein identification was achieved with a search performed against the Swiss-Prot protein database (March 2009, 428 650 entries) for *Homo sapiens*. The final list includes proteins common to all three independent LC-MALDI-MS/MS runs, which were then queried using the "Retrieve ID/mapping" tool of UniProtKB [19] for membrane glycoproteins with extracellular domain or secreted glycoproteins [14]. The identified glycoproteins were then screened for putative *N*-glycosylation sites using NetNglyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc), an artificial neural network that examines the sequence context of Asn-X-Ser/Thr (where X is not Pro) sequons [20]. *O*-glycosylation was predicted using the NetOglyc 4.0 server (http:// www.cbs.dtu.dk/services/NetOGlyc) that produces neural network predictions of mucin-

type GalNAc *O*-glycosylation sites in mammalian proteins [21]. Online bioinformatics tools Oncomine [22] and the ClueGO app for Cytoscape (http://www.cytoscape.org/) [23, 24] allowed narrowing the identified species to glycobiomarkers with non-invasive clinical potential, as previously described by us [14].

Slot-Blot analysis

Urine proteins (10 μg) were slot-blotted on a nitrocellulose membrane (Whatman, Protan; pore size 0.45 μm) using the Hybri-slot apparatus (21052-014; Gibco BRL, Life Technologies, Waltham, MA, USA). Protein loads in the nitrocellulose membranes was determined by Ponceau S (Sigma-Aldrich). Nonspecific binding was blocked with Carbo-Free Blocking Solution (SP-5040, Vector Laboratories) for 30 minutes at room temperature. The samples were incubated with biotinylated ConA, WGA and SNA lectins (Vector Laboratories) for 30 minutes at room temperature. Membranes were then washed with TBS-T (TBS with 0.5 % Tween 20) and incubated with VECTASTAIN® ABC (Peroxidase, PK-6100, Vector Laboratories) for 30 minutes at room temperature. Lectin-affinity was determined using biotin/avidin interaction. Albumin from chicken egg white (Sigma-Aldrich) was used as positive control for ConA [25] and fetuin from fetal calf serum (Sigma-Aldrich) was used as positive control for WGA [26] and SNA [27]. Positive controls were used for normalization of results amongst different blots. Deglycosylated bovine serum albumin (Sigma-Aldrich) was used as negative control. Urine proteins were also screened for CD44 using a recombinant monoclonal antibody (anti-CD44, 1:150 in PBS; EPR1013Y; Abcam, Cambridge, UK) and a goat anti-rabbit IgG (H+L) secondary antibody, HRP (Invitrogen, Carlsbad, CA, USA). Reactive bands were detected by enhanced chemiluminescence ECL (Amersham Pharmacia Biotech Inc, Piscataway, NJ, USA) according to the manufacturer's instructions. Images were recorded using X-ray films (Kodak Biomax light Film, Sigma-Aldrich). The films were scanned in Molecular Imager Gel Doc XR+ System (Bio-Rad) and analyzed with QuantityOne software (v 4.6.3, Bio-Rad). The results reflected the average of at least three independent replicates.

Immunohistochemistry

Formalin-fixed, paraffin-embedded tissue sections (FFPE) were screened for CD44 expression by immunohistochemistry using the strepavidin/biotin peroxidase method. Briefly, 3 μm sections were deparaffinized with xylene, rehydrated with graded ethanol series, microwaved for 15 min in boiling citrate buffer (10mM Citric Acid, 0.05% Tween 20, pH 6.0, Sigma-Aldrich), and exposed to 3% hydrogen peroxide for 25 min. CD44 was detected using a recombinant monoclonal antibody (anti-CD44, 1:4000 in PBS; ab157107; Abcam, Cambridge, UK) after incubation overnight at 4°C. The antigens were identified with UltraVision HRP Detection System Kit (Thermo Fisher Scientific, Waltham, MO, USA) followed by incubation with 3,3-diaminobenzidine tetrahydrochloride (Impact Dab, Vector Laboratories) for chromogenic development. Finally, the slides were counterstained with Harris's haematoxylin for 1 min. Negative control sections were performed by adding BSA (5% in PBS) devoided of primary antibody. CD44 negative muscle tissue sections were also analysed in parallel as negative controls. The immunostained sections were blindly assessed using light microscopy by two independent observers and validated by an experienced pathologist. Briefly, a semi-quantitative approach was established to score immunoreactivity based on the intensity and extension of the staining. The extension of staining was rated in cutoffs of 10%, and staining intensity was rated as follows: negative-0, weak-1, moderate-2, strong-3. The tumours were then classified based on the multiplication of extension evaluation and intensity. Disaccording readings were re-analyzed using a double-headed microscope and consensus was reached.

Statistical analysis

Statistical analysis was performed using Graphpad prism7 by GraphPad Software, Inc. Differences between continuous variables among the evaluated groups were accessed by Mann-Whitney non- parametric test for independent samples. Differences were considered significant when $p<0.05$.

Results

To highlight the potential of glycan-affinity glycoproteomics nanoplatforms for biomarker discovery in BC, we have prospectively collected urine samples from low-grade and high-grade NMIBC and MIBC patients. A significant increase in protein content was

observed in these urines in comparison to healthy controls matching in age and gender (**Figure S1A**-Supporting Information). Moreover, proteinuria increased with the severity of the lesions, in accordance with previous reports [28]. Concomitantly, the urines of cancer patients presented higher sialic acids content in comparison to the controls, which also was more pronounced in high-grade NMIBC and MIBC (**Figure S1B** – Supporting Information), suggesting profound alterations in the urine glycome. Urine glycoproteins were then isolated after dialysis and pooled together in accordance with the histopathological natures of the tumours.

As a proof-of-concept, we have incubated urine proteins from each group with MNP functionalized with three broad spectra lectins: i) ConA, targeting α-linked mannose present in *N*-glycan core oligosaccharides; ii) WGA, targeting terminal GlcNAc units common to many glycans found in membrane proteins; iii) SNA, targeting preferentially *O*-6 linked and, to less extent, *O*-3 linked sialic acids. Proteins showing affinity for these lectins were recovered from the beads and identified based on a bottom-up shotgun proteomics by nanoLC-MALDI-TOF/TOF. The resulting data was sorted for membrane or secreted glycoproteins showing putative *N*- and/or *O*-glycosylation sites, based on bioinformatics predictions using UniProt, NetNGlyc and NetOGlyc, respectively.

Overall, we have identified 114 glycoproteins, 63 of which were exclusively found in cancer patients' urines and 21 solely in the controls (**Table S1 and S2**-Supporting Information). Considering the lectins individually, a higher number of proteins were identified with SNA (n=70) in comparison to ConA (n=56) and WGA (n=52), for both controls and cancer groups (**Figure S2A** – Supporting Information). It was also highlighted the existence of lectin-specific proteomic fingerprints exclusively associated to the disease state (**Figure S2B** – Supporting Information), which may result from significant alterations in the glycome and/or glycoproteins abundance. However, such differences were more pronounced for advanced disease (high grade NMIBC+MIBC) in comparison to superficial lesions and the controls. In particular, ConA and SNA retrieved a higher number of proteins for MIBC, suggesting significant alterations in *N*-glycans mannose backbone oligosaccharide and α2,6 sialylation of glycans, respectively.

We then merged this information in an attempt to provide a broader vision on the lectin-based proteomic fingerprint associated with disease. Amongst bladder cancer patient

urine samples, commonly observed proteins were apolipoprotein family members (APOA1, APOE), fibrinogen chains (FGA, FGB, FGG), alpha-1-antitrypsin (SERPINA1) and alpha-2-macroglobulin (A2M) (**Table S1**-Supporting Information). A higher number of proteins were identified in high-grade NMIBC (n=80) and MIBC (n=96) in comparison to low-grade NMIBC (n=52) and control urines (n=55), considering all three lectins and accounting the repetitive capture of some proteins by the three lectins (**Figure 1A**). Nevertheless, cancer-specific glycoproteins could be observed in all cancer groups, including low-grade tumours (**Figure 1B**). Moreover, supporting the alterations observed in the glycome, slot-blot analysis of individual samples demonstrated that all three lectins presented higher affinity for proteins isolated from high-grade lesions, particularly in patients facing muscle-invasion, when compared to the controls (**Figure S3**-Supporting Information). Taken together, these findings suggest the existence of significant alterations in the nature of urine glycome and glycoproteome, accompanying the transition from low-grade to high-grade lesions; however, these alterations become more pronounced upon muscle invasion.

Considering cancer-associated glycoproteins, 8 protein species were exclusively found in low-grade tumours, suggesting potential for early diagnosis of patients with superficial lesions. Conversely, 13 were exclusively found in high-grade NMIBC and may aid non-invasive detection of progression of low-grade lesions. Moreover, 19 glycoproteins have been exclusively found in MIBC, holding potential to determine further progression to invasion. On the other hand, 12 glycoproteins have been detected in all groups and lack potential for patient discrimination (**Figure 2 and Table S2**-Supporting Information). We have then comprehensively matched these findings against previous reports in bladder cancer tissues using Oncomine thresholds (p=0.001, 2-fold upregulation). Interestingly, 81% of the cancer-associated urine glycoproteins had been previously detected in both healthy human urothelium and bladder cancer tissues (**Figure 3 and Table S2**-Supporting Information). Nevertheless, it is possible that the net loss of integrity of the urothelium accompanying malignant transformations may ultimately dictate their release into urine in pathological situations. Moreover, many of these glycoproteins may experience glycome remodeling, favoring affinity for the lectins chosen for enrichment. Moreover, 19% of these glycoproteins have been found overexpressed in cancer tissues (IGHG3, SERPINA1,

CD44, IGHM, HSPG2, GGH, APOE, MUC16, ATP13A2, ADCK2, PRSS2, and LAMA5) (**Figure 3**). More importantly, we note that SERPINA1, IGHM and CD44, which are significantly overexpressed in MIBC (according to Oncomine), can be detected in the urine of both high-grade NMIBC as well as MIBC patients. These observations suggest that these proteins may further aid on non-invasive differentiation between low-grade and high-grade tumors and ultimately be useful in risk stratification. In accordance, literature demonstrated that SERPINA1 is frequently found in higher stage and grade patients and was associated with tumour progression and shorter disease-free survival in both urines and tumour tissues from bladder cancer patient's analysis [29, 30]. Regarding IGHM, this is the first report of this protein in bladder cancer as a non-invasive differentiator between low-grade and high-grade tumors. Conversely, CD44 has been frequently reported as a bladder cancer stem-cell biomarker expressed by more aggressive subpopulations [13, 31]. Previous studies have demonstrated the presence of CD44 in exfoliated cancer cells in urine and its potential as a urinary marker for bladder cancer detection [32, 33]. This is the first report identifying CD44 as a protein that can be released into urine and suggesting its potential as a non-invasive stratification marker between low and high-grade bladder tumours.

A comprehensive integration of cancer-associated urine glycobiomarkers using Cytoscape's ClueGo plug-in was conducted by showing only significant pathways ($p \leq 0.05$). This analysis further demonstrated an overrepresentation of MIBC-associated pathways, including regulation of immune responses (humoral and complement), platelet degranulation, which has been considered a key event for tumour growth and metastasis [34], and cholesterol transport, which may play a critical role in cancer progression [35]. However, there are also common biological features to both NMIBC and MIBC, concerning the promotion of heterotypic regulation of cell-cell adhesion, mostly mediated by CD44 (**Figure 4**). Given the key role played by CD44, its presence in urine samples and overexpression in more aggressive and invasive bladder tumours (Oncomine), further emphasis was devoted to validate these results and its potential biomarker value in the context of advanced disease.

CD44 expression in urine and bladder tumours

Orthogonal validation of CD44 expression in urine samples was performed by slot blotting. According to **Figure 5A**, CD44 urinary levels showed a trend increase with the severity of the lesions; however, this effect is only significant for high-grade tumours in comparison to controls and low-grade NMIBC.

In parallel, we have screened the corresponding bladder tissues for CD44 expression by immunohistochemistry. CD44 antigen was mainly found at the cancer cells membrane, in accordance with expected cellular location. Again, we observed a significant overexpression of CD44 with the severity of the lesions, which is more pronounced for high-grade tumours, thus, mimicking urine analysis (**Figure 5B**). In superficial tumours, CD44 was predominantly found in basal layer cells, which have been described to harbor more malignant clones. Contrastingly, MIBC cells presented an extensive and intense CD44 staining without a defined pattern throughout the tumour (**Figure 5C**).

These findings agree with MNP@ lectins-based glycoproteomics and reinforce the role of CD44 in the context of advanced disease (high-grade NMIBC and MIBC). However, even though both CD44 urinary and tissue levels increase with the severity of the lesions, no statistically significant correlation was observed between both events. Studies involving a higher number of samples are required to fully disclose this matter.

Discussion and concluding remarks

The establishment of proteomics workflows to address clinically relevant urine glycoproteins in bladder cancer remains challenging due to the lack of efficient enrichment strategies. Herein, we have explored the potential of targeting glycans using MNP@lectins and/or MNP@moAbs. Nevertheless, we have previously demonstrated that MNP@lectins significantly increase glycoprotein yields when compared to conventional immobilization strategies (sepharose/agarose) and microspheres [15]. This is most likely associated to superior dynamics provided by large surface volume ratios, while displaying low unspecific binding. Moreover, we have used this approach to address the serum, saliva, and urine glycoproteomes, which resulted in the identification of proteins spanning large dynamic ranges, including low abundant species [16]. Likewise, we have hypothesized that this approach could be ideal for addressing cancer-biomarkers, generally present in minute amounts (nano-fentomolar range) in urine samples. Accordingly, we have combined

information from MNP functionalized with three large spectra lectins (ConA - *N*-glycans structures; WGA - GlcNAc in *N*- and *O*-glycans; and SNA - sialic acids also present in *N*- and *O*-glycans) envisaging biomarker panels associated with different aspects of the disease.

Even though explorative, this study has highlighted significant alterations in the urine glycome and/or glycoproteome, warranting comprehensive and more in-depth evaluation in future studies. These changes were significantly more pronounced in MIBC in comparison to initial stages of the disease and healthy controls, supporting a disease-associated nature of these findings. In this context, we observed that ConA and, particularly, SNA-affinity led to the identification of higher number of glycoproteins associated with muscle-invasion. These findings suggest structural alterations in the mannose backbone of *N*-glycans (translated by ConA) and α2,6 oversialylation (translated by SNA). Reinforcing these observations, both structural changes have been found in advanced bladder cancer and implicated in increased invasion and metastasis potential of bladder cancer cells [3, 36, 37]. Amongst the identified proteins were known bladder cancer-associated biomarkers apolipoprotein family, fibrinogen chains, alpha-1-antitrypsin, alpha-2-macroglobulin and uromodulin, previously reported as urinary biomarkers for bladder cancer [38, 39]. However, it has been extensively discussed that, individually, these proteins lack the necessary sensitivity and specificity to face the main clinical challenges raised by bladder cancer management, i.e. early identification of patients facing disease progression to high-grade disease and invasion. Giving this challenge, our strategy has generated a broader panel of 42 glycoproteins associated to high-grade disease (NMIBC and MIBC) that may hold clinical potential, 19 of which exclusively found in MIBC. In addition, the comprehensive integration of data using protein-protein interaction networks (ClueGO plug-in for Cytoscape) demonstrated the existence of common pathways to both NMIBC and MIBC (humoral response, complement activation, cell-adhesion) even though driven by different protein players. Amongst these glycoproteins is CD44, which has been found to modulate cell-cell adhesion in both NMIBC and MIBC (oncomine) and has been previously reported to drive invasion and metastasis [40, 41]. In agreement with these observations, we found high levels of CD44 in the urine of patients diagnosed with high-grade NMIBC and MIBC, possibly reflecting the molecular nature of the tumour. The

CD44 antigen has been previously found overexpressed in more aggressive bladder tumours and bladder cancer cells exfoliated into urine, and is frequently associated to poor prognosis [32]. Nevertheless, this is the first report demonstrating the shedding of this glycoprotein into urine, supporting its potential as biomarkers of disease progression. However, CD44 may present several isoforms that may change according to the origin of the tumour as well as the severity of the lesions. In bladder cancer, the most described CD44 isoforms are CD44s, CD44v6 and CD44v9, both associated with poor prognosis [42-44]. Variations in CD44 expression might be connected to the derangement of differentiation in bladder tumour cells [45]. Moreover, these cells are usually located in the basal layer of normal urothelium, the potential location of initiation tumour bladder cells, thereby reinforcing the value of CD44 as bladder cancer stem cell marker [46]. Future studies should therefore devote to a comprehensive clarification of the nature of CD44 isoforms patterns as well as its glycoforms in bladder cancer envisaging highly specific biomarkers. The identification of clinically relevant CD44 glycoforms may further enhance the biomarker value of this glycoprotein, has previously demonstrated for other glycoproteins, namely PSA in prostate cancer [47]. Further supporting our observations, previous reports have also identified CD44 in the urine of bladder cancer patients [48, 49]. However, we must note that this is a preliminary study focusing on a small number of samples of each subgroup for discovery (n=5 pooled). Therefore, biomarker candidates require validation in additional samples and cohorts. Moreover, even though we have focused on CD44 due to previously reported clinical significant, a significant number of other glycobiomarkers have been identified that also warrant future validation. Moreover, proteomic fingerprint provided by the lectins holds potential for multiplex disease detection towards improved sensitivity and specificity. Ultimately, lectins proteomics fingerprints maybe indicators of altered glycosylation associated with malignant transformation and disease dissemination, constituting starting points for glycomics studies envisaging clinical intervention.

Accordingly, we believe that the MNP@lectins-based enrichment strategy may also constitute a key starting point for glycome characterization by narrowing down identified glycoproteins to its clinically relevant glycoforms.

In summary, we believe that the future of biomarker discovery in bladder cancer should include a comprehensive interrogation of the glycome and glycoproteome, as well as the integration of relevant glycobiomarkers into broader panomics data. The generalization of glycan-based enrichment nanoplatforms may provide the necessary means to achieve this goal.

Acknowledgments

The authors wish to acknowledge the Portuguese Foundation for Science and Technology (FCT) for the human resources grants: PhD grant SFRH/BD/105355/2014 (Rita Azevedo), SFRH/BD/111242/2015 (Andreia Peixoto), SFRH/BD/103571/2014 (Elisabete Fernandes), SFRH/BD/127327/2016 Cristiana Gaiteiro and Postdoctoral grants SFRH/BPD/101827/2014 (Luis Lima) and SFRH/BPD/111048/2015 (José Alexandre Ferreira) as well as for the contract IF/00405/2014 (Ana Daniel-da-Silva). FCT is co-financed by European Social Fund (ESF) under Human Potential Operation Programme (POPH) from National Strategic Reference Framework (NSRF). The authors also acknowledge FCT the funding for CICECO research unit (POCI-01–0145-FEDER-007679 and UID/CTM /50011/2013), the Portuguese Oncology Institute of Porto Research Centre (CI-IPOP-29-2014; CI-IPOP-58-2015) and PhD Programs in Biomedicine and Pathology and Molecular Pathology of ICBAS-University of Porto.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

[1] J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D.M. Parkin, D. Forman, F. Bray, Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012, Int J Cancer 136(5) (2015) E359-86.

[2] M. Babjuk, A. Bohle, M. Burger, O. Capoun, D. Cohen, E.M. Comperat, V. Hernandez, E. Kaasinen, J. Palou, M. Roupret, B.W. van Rhijn, S.F. Shariat, V. Soukup, R.J. Sylvester,

R. Zigeuner, EAU Guidelines on Non-Muscle-invasive Urothelial Carcinoma of the Bladder: Update 2016, Eur Urol 71(3) (2017) 447-461.

[3] R. Azevedo, A. Peixoto, C. Gaiteiro, E. Fernandes, M. Neves, L. Lima, L.L. Santos, J.A. Ferreira, Over forty years of bladder cancer glycobiology: Where do glycans stand facing precision oncology?, Oncotarget 8(53) (2017) 91734-91764.

[4] P. Bassi, V. De Marco, A. De Lisa, M. Mancini, F. Pinto, R. Bertoloni, F. Longo, Non-invasive diagnostic tests for bladder cancer: a review of the literature, Urol Int 75(3) (2005) 193-200.

[5] M. Zhao, M. Li, Y. Yang, Z. Guo, Y. Sun, C. Shao, M. Li, W. Sun, Y. Gao, A comprehensive analysis and annotation of human normal urinary proteome, Sci Rep 7(1) (2017) 3024.

[6] Y.T. Chen, H.W. Chen, D. Domanski, D.S. Smith, K.H. Liang, C.C. Wu, C.L. Chen, T. Chung, M.C. Chen, Y.S. Chang, C.E. Parker, C.H. Borchers, J.S. Yu, Multiplexed quantification of 63 proteins in human urine by multiple reaction monitoring-based mass spectrometry for discovery of potential bladder cancer biomarkers, J Proteomics 75(12) (2012) 3529-45.

[7] M. Frantzi, A. Latosinska, L. Fluhe, M.C. Hupe, E. Critselis, M.W. Kramer, A.S. Merseburger, H. Mischak, A. Vlahou, Developing proteomic biomarkers for bladder cancer: towards clinical application, Nat Rev Urol 12(6) (2015) 317-30.

[8] O.H. Hashim, J.J. Jayapalan, C.S. Lee, Lectins: an effective tool for screening of potential cancer biomarkers, PeerJ 5 (2017) e3784.

[9] L. Lima, M. Neves, M.I. Oliveira, L. Dieguez, R. Freitas, R. Azevedo, C. Gaiteiro, J. Soares, D. Ferreira, A. Peixoto, E. Fernandes, D. Montezuma, A. Tavares, R. Ribeiro, A. Castro, M. Oliveira, A. Fraga, C.A. Reis, L.L. Santos, J.A. Ferreira, Sialyl-Tn identifies muscle-invasive bladder cancer basal and luminal subtypes facing decreased survival, being expressed by circulating tumor cells and metastases, Urol Oncol 35(12) (2017) 675 e1-675 e8.

[10] L. Lima, P.F. Severino, M. Silva, A. Miranda, A. Tavares, S. Pereira, E. Fernandes, R. Cruz, T. Amaro, C.A. Reis, F. Dall'Olio, F. Amado, P.A. Videira, L. Santos, J.A. Ferreira, Response of high-risk of recurrence/progression bladder tumours expressing sialyl-Tn and sialyl-6-T to BCG immunotherapy, Br J Cancer 109(8) (2013) 2106-14.

[11] C. Costa, S. Pereira, L. Lima, A. Peixoto, E. Fernandes, D. Neves, M. Neves, C. Gaiteiro, A. Tavares, R.M. Gil da Costa, R. Cruz, T. Amaro, P.A. Oliveira, J.A. Ferreira, L.L. Santos, Abnormal Protein Glycosylation and Activated PI3K/Akt/mTOR Pathway: Role in Bladder Cancer Prognosis and Targeted Therapeutics, PLoS One 10(11) (2015) e0141253.

[12] A. Peixoto, E. Fernandes, C. Gaiteiro, L. Lima, R. Azevedo, J. Soares, S. Cotton, B. Parreira, M. Neves, T. Amaro, A. Tavares, F. Teixeira, C. Palmeira, M. Rangel, A.M. Silva, C.A. Reis, L.L. Santos, M.J. Oliveira, J.A. Ferreira, Hypoxia enhances the malignant nature of bladder cancer cells and concomitantly antagonizes protein O-glycosylation extension, Oncotarget 7(39) (2016) 63138-63157.

[13] S. Cotton, R. Azevedo, C. Gaiteiro, D. Ferreira, L. Lima, A. Peixoto, E. Fernandes, M. Neves, D. Neves, T. Amaro, R. Cruz, A. Tavares, M. Rangel, A.M.N. Silva, L.L. Santos, J.A. Ferreira, Targeted O-glycoproteomics explored increased sialylation and identified MUC16 as a poor prognosis biomarker in advanced-stage bladder tumours, Mol Oncol 11(8) (2017) 895-912.

[14] R. Azevedo, A.M.N. Silva, C.A. Reis, L.L. Santos, J.A. Ferreira, In silico approaches for unveiling novel glycobiomarkers in cancer, J Proteomics 171 (2018) 95-106.

[15] J.A. Ferreira, A.L. Daniel-da-Silva, R.M. Alves, D. Duarte, I. Vieira, L.L. Santos, R. Vitorino, F. Amado, Synthesis and optimization of lectin functionalized nanoprobes for the selective recovery of glycoproteins from human body fluids, Anal Chem 83(18) (2011) 7035-43.

[16] M. Cova, R. Oliveira-Silva, J.A. Ferreira, R. Ferreira, F. Amado, A.L. Daniel-da-Silva, R. Vitorino, Glycoprotein enrichment method using a selective magnetic nano-probe platform (MNP) functionalized with lectins, Methods Mol Biol 1243 (2015) 83-100.

[17] M. Miyake, Y. Morizawa, S. Hori, Y. Tatsumi, S. Onishi, T. Owari, K. Iida, K. Onishi, D. Gotoh, Y. Nakai, S. Anai, Y. Chihara, K. Torimoto, K. Aoki, N. Tanaka, K. Shimada, N. Konishi, K. Fujimoto, Diagnostic and prognostic role of urinary collagens in primary human bladder cancer, Cancer Sci 108(11) (2017) 2221-2228.

[18] C.J. Rosser, S. Ross, M. Chang, Y. Dai, L. Mengual, G. Zhang, J. Kim, V. Urquidi, A. Alcaraz, S. Goodison, Multiplex protein signature for the detection of bladder cancer in voided urine samples, J Urol 190(6) (2013) 2257-62.

[19] C. The UniProt, UniProt: the universal protein knowledgebase, Nucleic Acids Res 45(D1) (2017) D158-D169.

[20] R. Gupta, E. Jung, S. Brunak, Prediction of N-glycosylation sites in human proteins, (In preparation, 2004).

[21] C. Steentoft, S.Y. Vakhrushev, H.J. Joshi, Y. Kong, M.B. Vester-Christensen, K.T. Schjoldager, K. Lavrsen, S. Dabelsteen, N.B. Pedersen, L. Marcos-Silva, R. Gupta, E.P. Bennett, U. Mandel, S. Brunak, H.H. Wandall, S.B. Levery, H. Clausen, Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology, The EMBO journal 32(10) (2013) 1478-88.

[22] D.R. Rhodes, S. Kalyana-Sundaram, V. Mahavisno, R. Varambally, J. Yu, B.B. Briggs, T.R. Barrette, M.J. Anstet, C. Kincead-Beal, P. Kulkarni, S. Varambally, D. Ghosh, A.M. Chinnaiyan, Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles, Neoplasia 9(2) (2007) 166-80.

[23] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome Res 13(11) (2003) 2498-504.

[24] G. Bindea, B. Mlecnik, H. Hackl, P. Charoentong, M. Tosolini, A. Kirilovsky, W.H. Fridman, F. Pages, Z. Trajanoski, J. Galon, ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks, Bioinformatics 25(8) (2009) 1091-3.

[25] A. Lee, M. Nakano, M. Hincapie, D. Kolarich, M.S. Baker, W.S. Hancock, N.H. Packer, The lectin riddle: glycoproteins fractionated from complex mixtures have similar glycomic profiles, OMICS 14(4) (2010) 487-99.

[26] Y. Itakura, S. Nakamura-Tsuruta, J. Kominami, H. Tateno, J. Hirabayashi, Sugar-Binding Profiles of Chitin-Binding Lectins from the Hevein Family: A Comprehensive Study, Int J Mol Sci 18(6) (2017).

[27] E. Katorcha, I.V. Baskakov, Analyses of N-linked glycans of PrP(Sc) revealed predominantly 2,6-linked sialic acid residues, FEBS J 284(21) (2017) 3727-3738.

[28] D.Y. Yang, R.H. Thompson, H.B. Zaid, C.M. Lohse, A.D. Rule, S.A. Boorjian, B.C. Leibovich, J.C. Cheville, M.K. Tollefson, Severity of Preoperative Proteinuria is a Risk

Factor for Overall Mortality in Patients Undergoing Nephrectomy, J Urol 198(4) (2017) 795-802.

[29] M. Linden, U. Segersten, M. Runeson, K. Wester, C. Busch, U. Pettersson, S.B. Lind, P.U. Malmstrom, Tumour expression of bladder cancer-associated urinary proteins, BJU Int 112(3) (2013) 407-15.

[30] G. Zhang, E. Gomes-Giacoia, Y. Dai, A. Lawton, M. Miyake, H. Furuya, S. Goodison, C.J. Rosser, Validation and clinicopathologic associations of a urine-based bladder cancer biomarker signature, Diagn Pathol 9 (2014) 200.

[31] T. Hofner, S. Macher-Goeppinger, C. Klein, A. Schillert, C. Eisen, S. Wagner, T. Rigo-Watermeier, I. Baccelli, V. Vogel, A. Trumpp, M.R. Sprick, Expression and prognostic significance of cancer stem cell markers CD24 and CD44 in urothelial bladder cancer xenografts and patients undergoing radical cystectomy, Urol Oncol 32(5) (2014) 678-86.

[32] M. Sugiyama, A. Woodman, T. Sugino, S. Crowley, K. Ho, J. Smith, Y. Matsumura, D. Tarin, Non-invasive detection of bladder cancer by identification of abnormal CD44 proteins in exfoliated cancer cells in urine, Clin Mol Pathol 48(3) (1995) M142-7.

[33] A.C. Woodman, S. Goodison, M. Drake, J. Noble, D. Tarin, Noninvasive diagnosis of bladder carcinoma by enzyme-linked immunosorbent assay detection of CD44 isoforms in exfoliated urothelia, Clin Cancer Res 6(6) (2000) 2381-92.

[34] M. Yan, P. Jurasz, The role of platelets in the tumor microenvironment: From solid tumors to leukemia, Biochim Biophys Acta 1863(3) (2016) 392-400.

[35] S. Silvente-Poirot, M. Poirot, Cholesterol metabolism and cancer: the good, the bad and the ugly, Curr Opin Pharmacol 12(6) (2012) 673-6.

[36] M. Przybylo, D. Hoja-Lukowicz, A. Litynska, P. Laidler, Different glycosylation of cadherins from human bladder non-malignant and cancer cell lines, Cancer Cell Int 2 (2002) 6.

[37] J.M. Guo, X.Y. Zhang, H.L. Chen, G.M. Wang, Y.K. Zhang, Structural alterations of sugar chains in urine fibronectin from bladder cancer patients and its enzymatic mechanism, J Cancer Res Clin Oncol 127(8) (2001) 512-9.

[38] M. Frantzi, A. Vlahou, Ten Years of Proteomics in Bladder Cancer: Progress and Future Directions, Bladder Cancer 3(1) (2017) 1-18.

[39] V. Sandim, A. Pereira Dde, D.E. Kalume, A.L. Oliveira-Carvalho, A.A. Ornellas, M.R. Soares, G. Alves, R.B. Zingali, Proteomic analysis reveals differentially secreted proteins in the urine from patients with clear cell renal cell carcinoma, Urol Oncol 34(1) (2016) 5 e11-25.

[40] G. Yu, W. Yao, W. Xiao, H. Li, H. Xu, B. Lang, MicroRNA-34a functions as an anti-metastatic microRNA and suppresses angiogenesis in bladder cancer by directly targeting CD44, J Exp Clin Cancer Res 33 (2014) 779.

[41] K. Wu, Z. Ning, J. Zeng, J. Fan, J. Zhou, T. Zhang, L. Zhang, Y. Chen, Y. Gao, B. Wang, P. Guo, L. Li, X. Wang, D. He, Silibinin inhibits beta-catenin/ZEB1 signaling and suppresses bladder cancer metastasis via dual-blocking epithelial-mesenchymal transition and stemness, Cell Signal 25(12) (2013) 2625-33.

[42] O.M. Omran, H.S. Ata, CD44s and CD44v6 in diagnosis and prognosis of human bladder cancer, Ultrastruct Pathol 36(3) (2012) 145-52.

[43] T. Klatte, D.B. Seligson, J.Y. Rao, H. Yu, M. de Martino, I. Garraway, S.G. Wong, A.S. Belldegrun, A.J. Pantuck, Absent CD44v6 expression is an independent predictor of poor urothelial bladder cancer outcome, J Urol 183(6) (2010) 2403-8.

[44] K. Kobayashi, H. Matsumoto, H. Matsuyama, N. Fujii, R. Inoue, Y. Yamamoto, K. Nagao, Clinical significance of CD44 variant 9 expression as a prognostic indicator in bladder cancer, Oncol Rep 36(5) (2016) 2852-2860.

[45] J. Kuncova, Z. Kostrouch, M. Viale, R. Revoltella, V. Mandys, Expression of CD44v6 correlates with cell proliferation and cellular atypia in urothelial carcinoma cell lines 5637 and HT1197, Folia Biol (Praha) 51(1) (2005) 3-11.

[46] Y.M. Yang, J.W. Chang, Bladder cancer initiating cells (BCICs) are among EMA-CD44v6+ subset: novel methods for isolating undetermined cancer stem (initiating) cells, Cancer Invest 26(7) (2008) 725-33.

[47] E. Llop, M. Ferrer-Batalle, S. Barrabes, P.E. Guerrero, M. Ramirez, R. Saldova, P.M. Rudd, R.N. Aleixandre, J. Comet, R. de Llorens, R. Peracaula, Improvement of Prostate Cancer Diagnosis by Detecting PSA Glycosylation-Specific Changes, Theranostics 6(8) (2016) 1190-204.

[48] P. Kreunin, J. Zhao, C. Rosser, V. Urquidi, D.M. Lubman, S. Goodison, Bladder cancer associated glycoprotein signatures revealed by urinary proteomic profiling, J Proteome Res 6(7) (2007) 2631-9.

[49] S.R. Ambrose, N.S. Gordon, J.C. Goldsmith, W. Wei, M.P. Zeegers, N.D. James, M.A. Knowles, R.T. Bryan, D.G. Ward, Use of Aleuria alantia Lectin Affinity Chromatography to Enrich Candidate Biomarkers from the Urine of Patients with Bladder Cancer, Proteomes 3(3) (2015) 266-282.

FIGURE CAPTIONS

Figure 1. Urine glycoproteins identified in controls and low-grade NMIBC, high-grade NMIBC and MIBC using MNP@ConA, MNP@WGA and MNP@SNA lectins. **A.** A higher number of proteins were identified for high-grade NMIBC (n=80) and MIBC (n=96) in comparison to low-grade NMIBC (n=52) and control urines (n=55), considering all three lectins and accounting the repetitive capture of some proteins by the three lectins. **B.** Venn diagram highlighting the existence of distinct lectin-based glycoprotein signatures for each group, holding potential for early detection (low-grade tumours group); disease progression to undifferentiating (high-grade NMIBC group); progression to invasion (MIBC) and poor prognosis (high-grade NMIBC+MIBC group).

Figure 2. Graphical representation of cancer-specific urine glycobiomarkers with affinity to MNP@ConA, MNP@WGA and MNP@SNA lectin nanoprobes distributed according to their expression in bladder cancer. Accordingly, 63 glycoproteins were exclusively found in cancer samples (top right Venn diagram). The left panel represents the distribution of these glycoproteins amongst the cancer groups (low-grade and high-grade NMIBC and MIBC). The number of predicted *O*- and *N*-glycosylation sites were also represented according to results from NetOGlyc and NetNGlyc bioinformatics tools, respectively.

Figure 3. Graphical representation of cancer-specific urine glycobiomarkers distributed according to their overexpression in bladder cancer by Oncomine. Accordingly, 63 glycoproteins that were exclusively found in cancer samples were analyzed *in silico* using Oncomine and associations of these glycoproteins with overexpression with bladder cancer compared to normal tissues is represented. The number of predicted *O*- and *N*-glycosylation sites were also represented according to results from NetOGlyc and NetNGlyc bioinformatics tools, respectively.

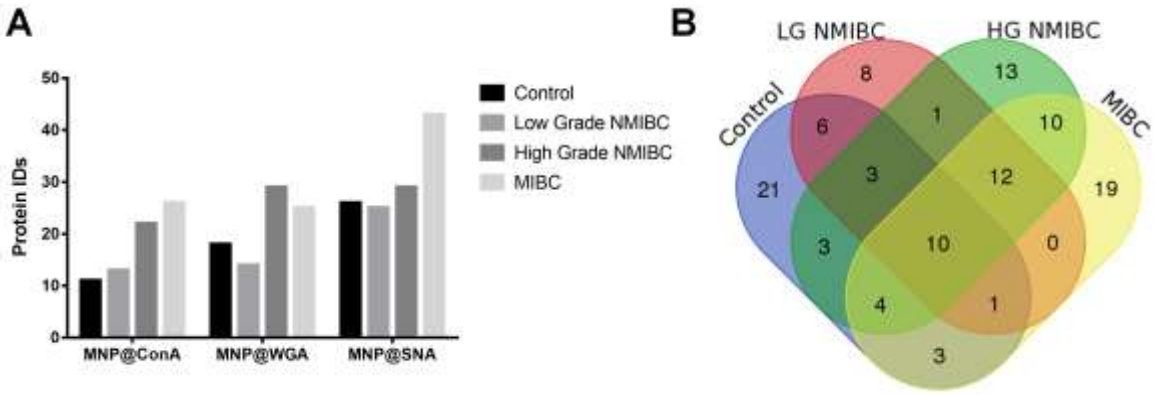
Figure 4. Network presenting the most significant biological functions of the candidate glycobiomarkers in bladder cancer. This network mainly reflects biological functions of

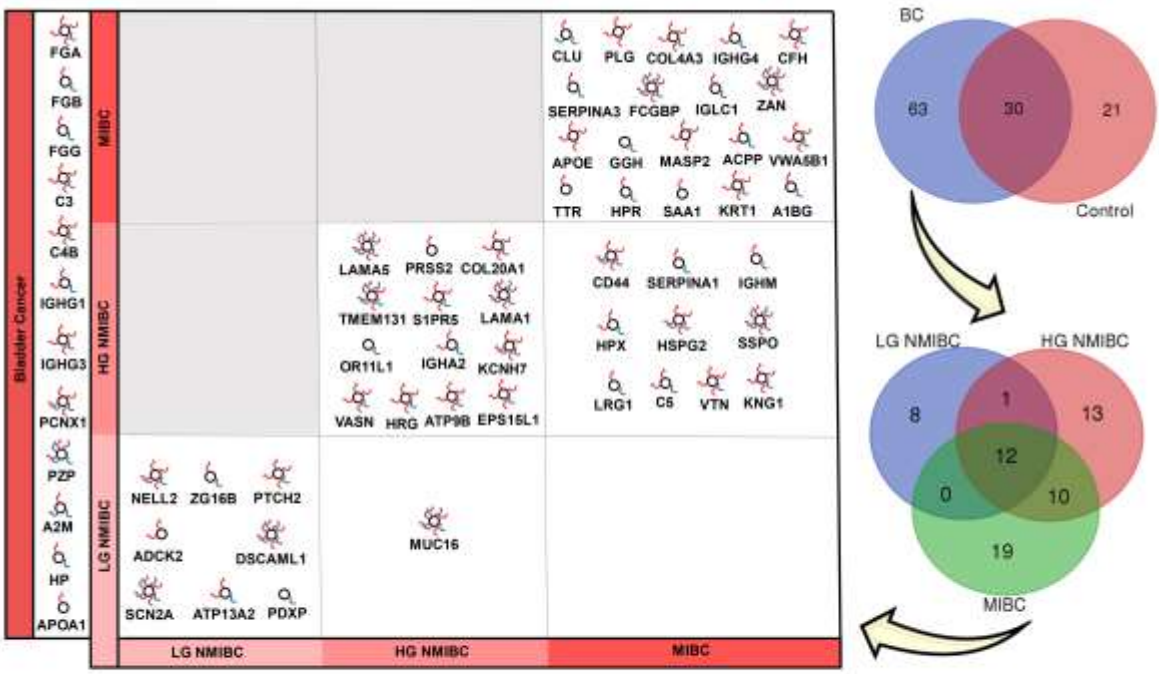
immune system regulation (humoral response and complement), platelet degranulation and cholesterol transport as also heterotypic cell-cell adhesion, which is essentially played by CD44 glycoprotein in both NMIBC and MIBC. The network reproduces significant protein-protein interactions and biological functions ($p \leq 0.05$). Blue nodes reflect events mainly associated to NMIBC, red nodes with MIBC and grey nodes with both NMIBC and MIBC. The size of the nodes reflects the statistical significance of the terms. The degree of connectivity between terms (edges) is calculated using kappa statistics.

Figure 5. Validation of CD44 expression in urine samples and immunohistochemistry of thirty-one bladder cancer patients. A. CD44 urinary levels were determined by slot-blot. CD44. CD44 showed a trend increase with the severity of the lesions; however this effect is only significant for high-grade tumours in comparison to the control and low-grade NMIBC groups. Analyses were conducted in triplicates. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. **B.** CD44 antigen was mainly found at membrane of cancer cells, in accordance with expected cellular location. Again, we observed a significant overexpression of CD44 with the severity of the lesions, which is more pronounced for high-grade tumours, thus, mimicking urine analyses. **C.** In superficial tumours, CD44 was predominantly found in cells underlining the basal layer, which have been described to harbor more malignant clones. Contrastingly, MIBC cells presented an extensive and intense CD44 expression without a defined pattern throughout the tumour.

Highlights

- MNP@Lectins hold potential for urine glycobiomarker discovery in bladder cancer.
- Urine glycome and glycoproteome changes according the nature of bladder cancer.
- More alterations in the urinary glycoproteome are observed for aggressive lesions.
- Advanced bladder cancer patients show Increased CD44 urinary and tumour levels.





Key:

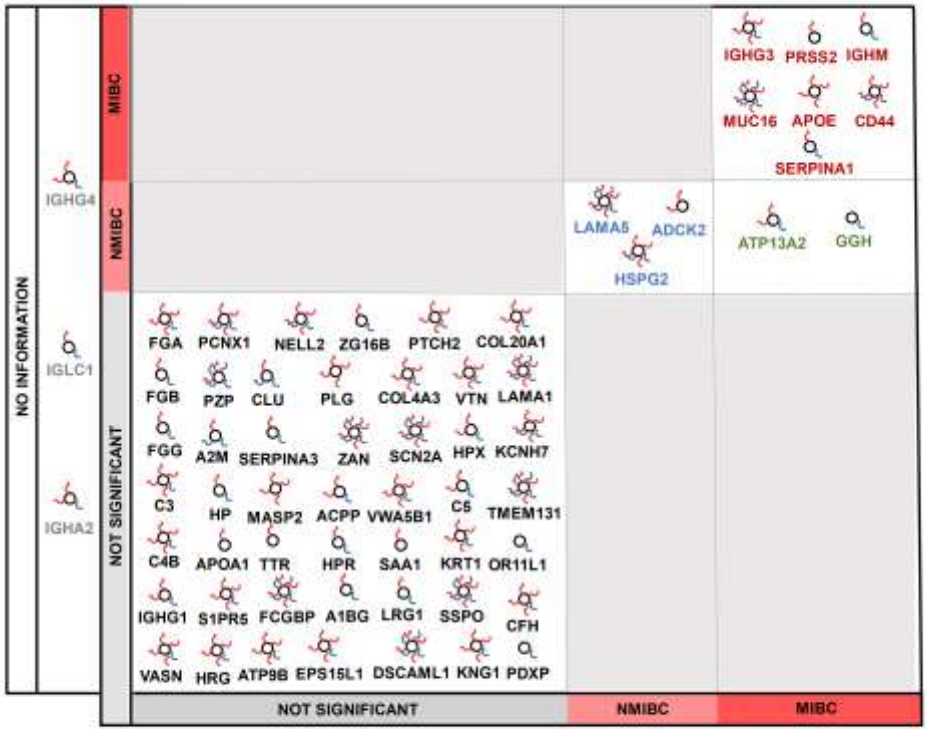
Predicted glycosylation sites

0 1-5 6-9 ≥10

Predicted sites of *O*-glycosylation (Data from *NetOGlyc*)

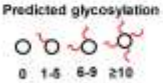
0 1-5 6-9 ≥10

Predicted sites of *N*-glycosylation (Data from *NetNGlyc*)



Key:

Predicted glycosylation sites



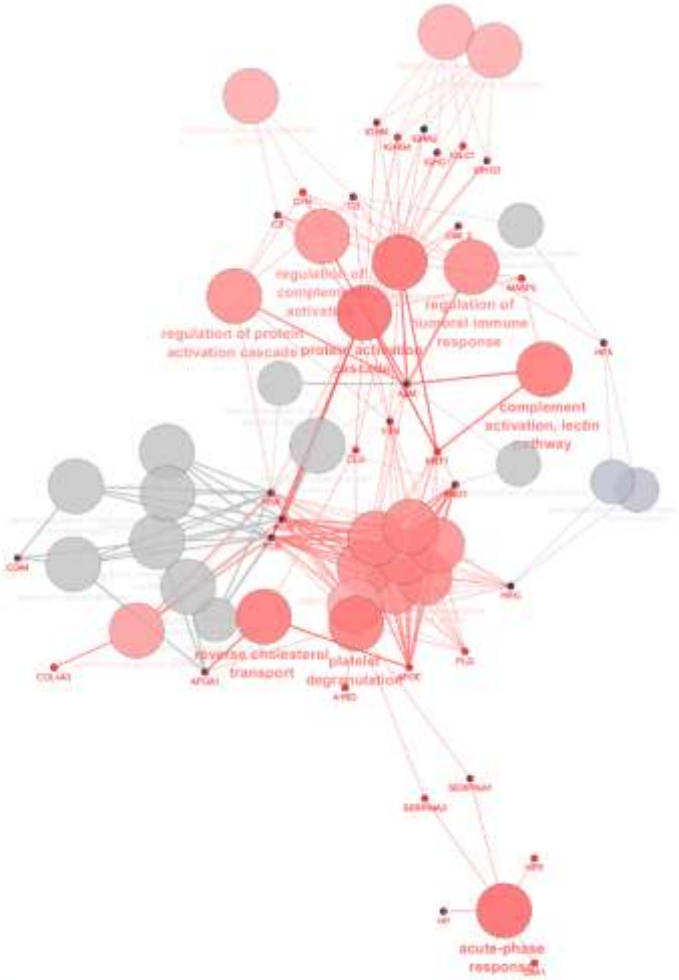
Predicted sites of O-glycosylation
(Data from NetOGlyc)

Predicted sites of *N*-glycosylation
(Data from NetNGlyc)

Association with overexpression in bladder cancer (vs normal tissues)

- **BLACK** protein name: present in same amounts in both BC and normal tissues, not significant
- **RED** protein name: overexpressed in MIBC
- **BLUE** protein name: overexpressed in NMIBC
- **GREEN** protein name: overexpressed in both NMIBC and MIBC
- GREY protein name: no information

(Data from Oncomine - threshold: p=0.001, 2-fold upregulation, All rank)



RED proteins: expressed in MIBC
BLUE proteins: expressed in NMIBC
(Data from ClueGo plug-in for Cytoscape - Only pathways with p≤0.05 are shown)