scientific reports

Check for updates

OPEN

# Profiling of urinary extracellular vesicle protein signatures from patients with cribriform and intraductal prostate carcinoma in a cross-sectional study

Rui Bernardino[1,2]✉, Ana Sofia Carvalho[1]✉, Michael J. Hall[1], Liliana Alves[1], Ricardo Leão[3], Rashid Sayyid[2], Hermínia Pereira[4], Hans Christian Beck[5], Luís Campos Pinheiro[6], Rui Henrique[7,8], Neil Fleshner[2] & Rune Matthiesen[1]✉

**Prognostic tests and treatment approaches for optimized clinical care of prostatic neoplasms are an unmet need. Prostate cancer (PCa) and derived extracellular vesicles (EVs) proteome changes occur during initiation and progression of the disease. PCa tissue proteome has been previously characterized, but screening of tissue samples constitutes an invasive procedure. Consequently, we focused this study on liquid biopsies, such as urine samples. More specifically, urinary small extracellular vesicle and particles proteome profiles of 100 subjects were analyzed using liquid chromatography coupled to high-resolution mass spectrometry (LC–MS/MS). We identified 171 proteins that were differentially expressed between intraductal prostate cancer/cribriform (IDC/Crib) and non-IDC/non-Crib after correction for multiple testing. However, the strong correlation between IDC/Crib and Gleason Grade complicates the disentanglement of the underlying factors driving this association. Nevertheless, even after accounting for multiple testing and adjusting for ISUP (International Society of Urological Pathology) grading, two proteins continued to exhibit significant differential expression between IDC/Crib and non-IDC/non-Crib. Functional enrichment analysis based on cancer hallmark proteins disclosed a clear pattern of androgen response down-regulation in urinary EVs from IDC/Crib compared to non-IDC/non-Crib. Interestingly, proteome differences between IDC and cribriform were more subtle, suggesting high proteome heterogeneity. Overall, the urinary EV proteome reflected partly the prostate pathology.**

**Keywords** Liquid chromatography mass spectrometry, Cribriform, Intraductal prostate carcinoma, Proteomics, Urinary extracellular vesicles and particles

Prostate cancer (PCa) is the most frequently diagnosed cancer and the second leading cause of cancer-related death among men in developed countries[1]. Cribriform pattern (Crib) and intraductal prostate cancer (IDC) are two histological subtypes of PCa related to poor clinical outcomes and a worse prognosis, yet there are limited clinical tools for their early detection[2,3]. Studies that have investigated the histologic correlation between

[1]Computational and Experimental Biology Group, iNOVA4Health, NOVA Medical School, Faculdade de Ciências Médicas, Universidade NOVA de Lisboa, Lisboa, Portugal. [2]Division of Urology, Department of Surgical Oncology, Princess Margaret Cancer Centre, University of Toronto, Toronto, ON, Canada. [3]Cuf Hospitais, Lisbon, Portugal. [4]Department of Pathology, Centro Hospitalar E Universitário Lisboa Central, Lisbon, Portugal. [5]Centre for Clinical Proteomics, Department of Clinical Biochemistry, Odense University Hospital, 5000 Odense, Denmark. [6]Department of Urology, Centro Hospitalar e Universitário Lisboa Central, Lisbon, Portugal. [7]Department of Pathology and Cancer Biology and Epigenetics Group - Research Center of IPO Porto (CI-IPOP) / RISE@CI-IPOP (Health Research Network), Portuguese Oncology Institute of Porto (IPO Porto) / Porto Comprehensive Cancer Center Raquel Seruca (Porto.CCC), R. Dr. António Bernardino de Almeida, 4200-072 Porto, Portugal. [8]Department of Pathology and Molecular Immunology, ICBAS – School of Medicine and Biomedical Sciences, University of Porto (ICBAS-UP), Rua Jorge Viterbo Ferreira 228, 4050-513 Porto, Portugal. ✉email: ruimmbernardino@gmail.com; ana.carvalho@nms.unl.pt; rune.matthiesen@nms.unl.pt

multiparametric magnetic resonance imaging (mpMRI) observation and Crib PCa have shown inconsistent results, as those are often less visible on mpMRI than non-Crib predominant tumors[4]. Moreover, the positive predictive value of an abnormal mpMRI remains extremely low, being less than 30% in most studies, meaning that if one has an abnormal mpMRI lesion, IDC/Crib is found in only 30% of patients at radical prostatectomy (RP). Overall, the use of mpMRI and mpMRI-guided biopsies is not sufficient for accurately categorizing these abnormalities in a way that can be reliably applied in medical practice[5].

Furthermore, prostate biopsies still have a significant number of false negatives (about 50%) for the accurate diagnosis of those patterns[6]. The routine failure to accurately identify IDC/Crib in needle biopsies and mpMRI outcomes leads to the misclassification of potentially aggressive prostate tumors. This clinical scenario takes on even more significance due to the increasing trend of recommending active surveillance for patients with Gleason 7 (3 + 4) disease.

Given the above clinical challenges, it is an urgent need to develop specific biomarkers for IDC/Crib as a diagnostic tool in the clinical setting. Extracellular vesicles (EVs), particularly exosomes, which are 50–150 nm in size and enclosed by a lipid bilayer membrane, are shed by various mammalian cell types, including cancerous cells, and are formed in the endosomal network before being released by the fusion of multi-vesicular bodies with the plasma membrane[7]. These vesicles contain various biomolecules, including proteins, lipids, DNA, and RNA, that reflect the molecular composition of their tissue of origin[8]. The proteome of urinary EVs appears to constitute a promising source of biomarkers for cancers of the urinary tract and prostate[9]. In this project, we focus on liquid biopsies (urine samples) from benign prostatic disease (BPD) and PCa patients. BPD group constitute non-cancer, and benign prostatic hyperplasia. More specifically, the proteome profiles of small EV enriched fractions from 100 individuals were analyzed using liquid chromatography coupled to high-resolution mass spectrometry (LC–MS/MS). EVs from three cohorts: BPD (N = 24) patients without Crib and IDC histological prostate patterns (non-IDC/non-Crib, N = 21); and patients with Crib and/or IDC histological prostate patterns (IDC/Crib, N = 55) at biopsy were interrogated by LC–MS/MS-based proteomics.

Based on our results, we hypothesize that urinary EVs constitute a non-invasive target to unveil proteomic signatures of IDC and Crib PCa patterns, allowing the identification of a urinary biomarker protein signature for patients with Crib and IDC. Indeed, we identified 171 proteins differentially expressed between IDC/Crib and non-IDC/Crib. Furthermore, functional analysis suggests that proteins involved in androgen responses are overall downregulated in IDC/Crib compared to BPD in urinary sEVs. However, androgen response was less relevant when comparing non-IDC/non-Crib and BPD.

## Results

### Patient samples

Clinical characteristics of the whole study population with urinary EV samples available are described in Table 1. Table S1 provides clinical and demographic variables linked to MS raw data files. Figure S1 displays age and pre-PSA concentration across major clinical subgroups. The study population consisted of 100 males, with a median age of 72 years at diagnosis. Most patients (N = 55, 55%) were diagnosed with either cribriform and/or intraductal carcinoma of prostate (IDC/Crib). Non-IDC/non-Crib (N = 21, 21%) and BPD (N = 24, 24%) were collected as control groups. BPD group is constituted of patients with prostate pathologies other than cancer. Non-IDC/non-Crib consisted of patients who underwent radical prostatectomy and were subsequently verified

| Variable | N | $N_{BPD} = 24^1$ | $N_{non\text{-}IDC/non\text{-}Crib} = 21^1$ | $N_{IDC/Crib} = 55^1$ | *p*-value[2] |
|---|---|---|---|---|---|
| prePSA | 92 | 2.8 (0.7, 4.2) | 7.0 (5.0, 8.8) | 11.4 (7.7, 31.8) | < 0.001 |
| Cribriform | 100 | | | | < 0.001 |
| NA | | 24 (100%) | 0 (0%) | 0 (0%) | |
| No | | 0 (0%) | 21 (100%) | 11 (20%) | |
| Yes | | 0 (0%) | 0 (0%) | 44 (80%) | |
| IDC | 100 | | | | < 0.001 |
| NA | | 24 (100%) | 0 (0%) | 0 (0%) | |
| No | | | 21 (100%) | 36 (65%) | |
| Yes | | | 0 (0%) | 19 (35%) | |
| #positive cores | 74 | | 5 (3, 6) | 7 (4, 12) | 0.011 |
| Gleason grade | 74 | | | | < 0.001 |
| 3 + 3 | | | 6 (32%) | 0 [0/0/0] (0%) | |
| 3 + 4 | | | 10 (53%) | 14 [4/2/8] (25%) | |
| 4 + 3 | | | 3 (16%) | 14 [1/0/13] (25%) | |
| 4 + 4 | | | 0 (0%) | 7 [1/0/6] (13%) | |
| > = 9 | | | 0 (0%) | 20 [2/9/9] (36%) | |

**Table 1**. Clinicopathological features of 100 subjects. [1]Median (IQR); n (%). [2]Kruskal-Wallis rank sum test; Pearson's Chi-squared test; Fisher's exact test. Numbers in square brackets indicate the number of subjects with IDC and Crib, with IDC only, and with Crib only, respectively. *BPD* benign prostatic disease (non-cancer, benign prostatic hyperplasia).

to be non-IDC/non-Crib. Clinical parameters such as pre-biopsy, number of positive cores, Gleason grade, and maximum percentage of core involvement are all significantly and positively correlated with histological patterns with either cribriform and/or intraductal carcinoma of prostate.

## Workflow for analyzing urinary EV proteome

Figure S2 displays the workflow used for the analysis of urinary extracellular vesicles and particles (EVs), including the steps of EV isolation, trypsin digestion, and liquid chromatography mass spectrometry (LC–MS) analysis. Urine samples were collected and immediately frozen within two hours of collection, and stored at − 80 °C until further processing. EVs were isolated by differential centrifugation and small urinary EVs were further digested with trypsin and analyzed by liquid chromatography mass spectrometry (LC–MS). The acquired data were analyzed using multivariate statistics and functional enrichment analysis.

## Quality control of urinary small EV preparations from prostate *cancer* patients

Multiple subsets of urinary small EV (sEV) preparations were analyzed by western blotting for EV and non-EV markers (Fig. 1A,B). The immunoblots assays of sEV preparations isolated from all major clinical subgroups such as urine of BPD, patients without Crib and IDC histological prostate patterns (non-IDC/non-Crib) and

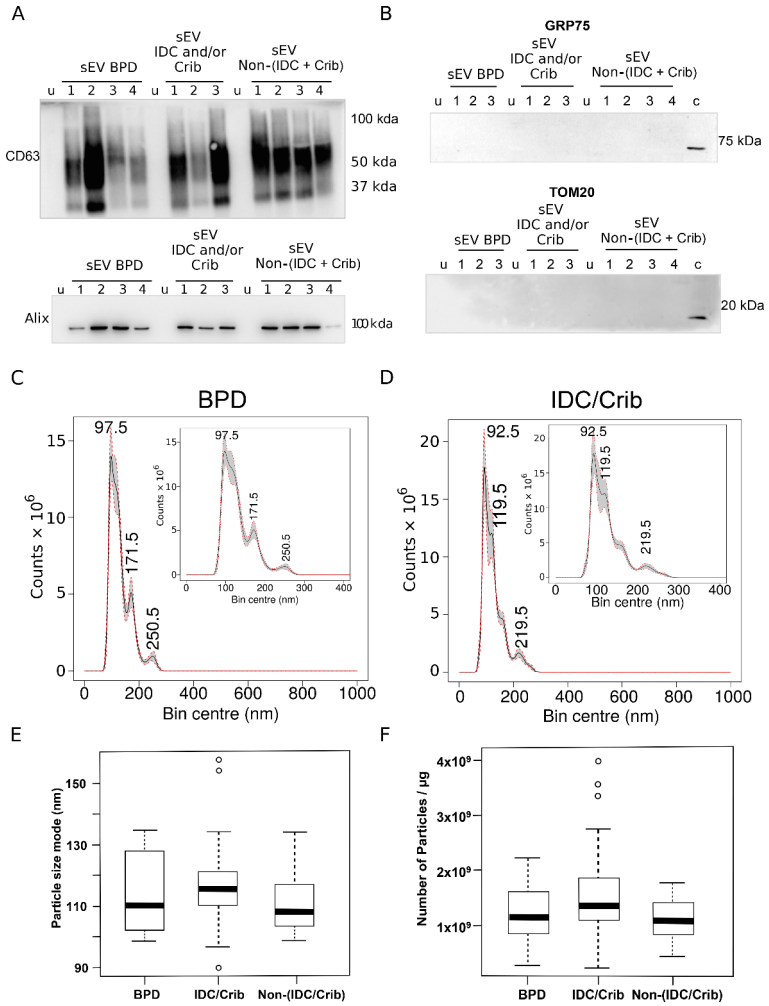


**Fig. 1**. Urinary EV characterization by western blot and NTA. (**A**) CD63 and Alix western blot of representative urinary EV samples from each major clinical groups. (**B**) Non-EV markers GRP75 (endoplasmic reticulum marker) and TOM20 (mitochondrial marker) as markers of transmembrane, lipid-bound and soluble proteins associated to other intracellular compartments than plasma membrane and endosomes. The process of transforming the original blot image into the final annotated image is outlined at the end of the supplementary documentation. NTA particle size distribution analysis of a representative urinary EV sample from (**C**) BPD and (**D**) IDC/Crib (insert represents a zoom in). Particle size mode distribution (**E**) and number of particles per protein amount (**F**) across major clinical subgroups. The central line inside each box represents the median value. Whiskers extend from the edges of the box to the minimum and maximum values within 1.5 times the interquartile range (IQR). Individual data points beyond the whiskers are considered outliers and are plotted individually. "c" indicates a mitochondrial enriched fraction from the large diffuse B cell lymphoma cell line HT.

patients with Crib and/or IDC histological prostate patterns (IDC/Crib) were performed. Consistent with the MISEV 2018 guidelines[10], which recommend characterizing EVs based on protein content, our results show that CD63 (Category 1, Fig. 1A) and ALIX (Category 2, Fig. 1A) both indicate the presence of extracellular vesicles. As for specificity of small EV subtypes (Category 4) we have used GRP75 and TOM20 as markers of transmembrane, lipid-bound and soluble proteins associated to other intracellular compartments than plasma membrane and endosomes (Fig. 1B). Detailed immunoblot methods are described in the methods section and MISEV 2018 check list is available in the supplementary data.

Nanoparticle tracking analysis (NTA) was performed on all urinary sEV samples. Figure 1C and D displays two representative particle size distributions from two distinct urinary sEV samples. The highest peak is close to 100 nm as expected for sEVs. Figure 1E displays the distribution of the most abundant particle size across all the major clinical subgroups. The size distributions by NTA for urinary sEV samples were highly reproducible across multiple measurements. Figure 1F shows the distributions of the number of particles per protein amount across all the major clinical subgroups. The obtained particles per protein amount is consistent with moderate to high quality sEVs preparations[11].

MS-based proteome data from urinary sEV preparations of 100 individuals were quality checked using principal component analysis (PCA) and linear discriminant analysis (LDA) based on all quantitative data. PCA exhibited reasonable separation with minor overlaps between groups when the first two components were plotted. Supervised LDA resulted in excellent separation except for three data points, one from each of the three subgroups analyzed (Figure S3). The scatter plot shows the projection of the data points onto the first two linear discriminant functions (LD1 and LD2), which are derived from the LDA analysis. The discriminant functions (LD1 and LD2) are plotted on the x-axis and y-axis, respectively, with the scale indicated on the corresponding axes. The scatter plot reveals the separation of the three groups in two-dimensional space based on the LDA analysis.

MS-based proteome data based on urinary sEV preparations from 100 patients were quality checked for EV markers (Fig. 2A). The urinary sEVs from the current study had the highest expression of established small EV markers compared to previous studies based on sEV samples from other sources, such as cell lines[12–14], lung fluids[15], and blood plasma[16]. Urinary sEVs were devoid of microsomes or large EVs protein markers. For markers indicating contamination from other subcellular organelles, a low level of calnexin was detected in

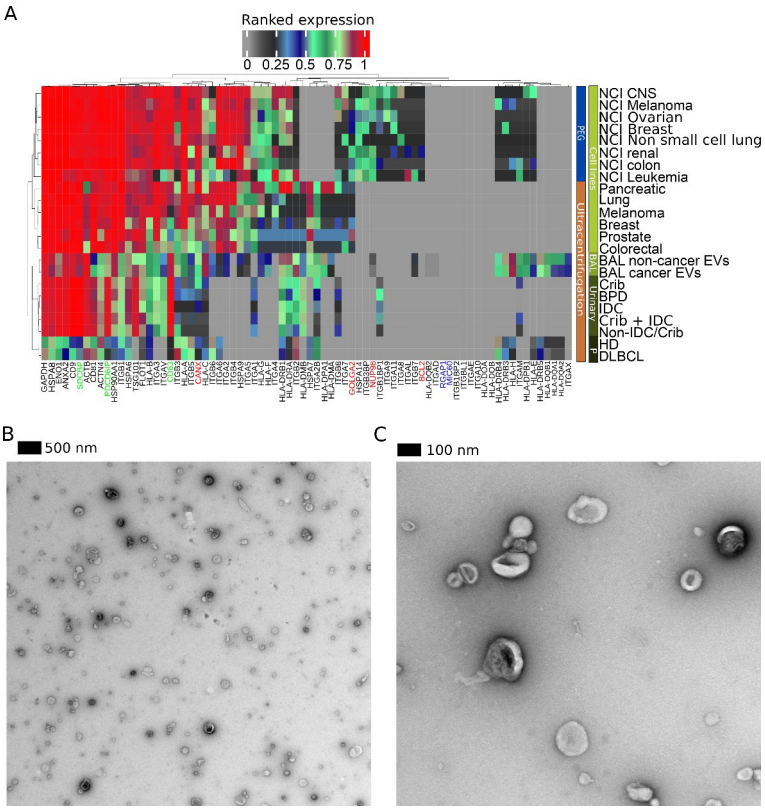


**Fig. 2**. Urinary EV characterization by MS and TEM. (**A**) Small EV protein marker expression comparison with expression observed in published studies on small EVs isolated from cell lines (based on ultracentrifugation or polyethylene glycol (PEG)), bronchoalveolar lavage (BAL), plasma (P), diffuse large B cell lymphoma (DLBCL), healthy donor in DLBCL cohort (HD), and urinary samples (current study). General exosomes or small EV markers are color-coded in black, established exosome markers are in green, potential contaminant markers from other subcellular organelles are in red and markers from large EVs or microsomes are indicated in blue. (**B**) TEM 4000 × magnification, (**C**) TEM 20,000 × magnification.

urinary sEVs. However, the expression levels were lower than what had been observed in previous studies. On the other hand, the contaminant Tamm-Horsfall Protein (THP, uromodulin) was abundant in urinary sEVs isolated in this study. Despite THP abundance, high protein coverage by mass spectrometry analysis was achieved as well as a high number of identified proteins. Moreover, there was no regulation of THP across clinical subgroups (Fig. 2A). In conclusion, sEV preparation was enriched in small EVs, according to our analysis which focused on known EV and contaminant markers.

Transmission electron microscopy (TEM) analysis at 4000× magnification was consistent with NTA analysis in that the majority of the particles were around 100 nm (Fig. 2B). Magnification at 20,000× revealed characteristic cup-shape of sEVs as a known artifact result of the staining procedure (Fig. 2C).

## Comparison of prostate tissue and urinary sEV proteome

It has been proposed that urinary EVs are mainly enriched in EV cargo of the surrounding tissues of origin, therefore we have analyzed the proteome of prostate, kidney, and bladder tissues. Urinary sEVs proteome obtained in this study was benchmarked against different human datasets reported in the literature. Based on the bottom-up proteomics proteome profile of prostate tissue samples and baseline protein expression in human tissues extracted from previous publications[17,18] in comparison with the urinary sEV proteome, we estimated that approximately 75% of the observed proteins in urinary sEVs were also detected in prostate tissue (Figure S4A). Assuming the entire human proteome as background, this overlap was estimated to be highly significant based on the hypergeometric probability function. Nevertheless, the overlap between the proteomes of urinary sEVs and kidney and bladder were similar (Figure S4B and C). Kidney tissue displayed the largest overlap with urinary sEVs, although the difference was marginal compared to urinary sEVs overlap with the other two organ tissues (Figure S4D). The high significance in overlap was also identified for all other tissues tested from Prakash et al[17]. Restricting the analysis to the 50% most abundant proteins expressed in each tissue and ranking based on the significance of the overlap the five highest ranked tissues were kidney, adipose tissue, pancreas, bladder, and prostate. The largest overlap in Figure S4D is between the three tissue proteomes. The second largest is between all four proteomes (urinary sEVs and tissue proteomes). Finally, each of the three tissue proteomes has unique overlap with urinary sEVs ranging from 11 to 43 proteins. Although most proteins in urinary sEVs are present in all three tissues, there are 367 proteins from other tissue sources. Overall, the large coverage of prostate tissue proteins in urinary sEVs suggests that urinary sEVs are a promising source of markers for prostate-related pathologies.

## Unique identified proteins across clinical subgroups

Data-dependent acquisition with two technical replicates was applied for the identification of proteins. Figure S5A summarizes the overlaps of proteins identified for each clinical subgroup. The unique proteins for each clinical subgroup were typically not consistently identified throughout the subgroup (Figure S5B-F). Nevertheless, the proteins unique to non-IDC/non-Crib and Crib displayed unique proteins that were shared between four to six patients. For example, proteins such as NPTN (Neuroplastin), KRTAP11-1 (Keratin associated protein 11-1) and CD99 (CD99 molecule) (Figure S5B,D).

## Significant differentially expressed proteins

Three pairwise comparisons were performed using the R package limma[19]. Figure 3 provides an overview table and volcano plots visualizing regulated proteins for three pairwise comparisons. The comparison between non-IDC/non-Crib and BPD resulted in the most regulated proteins (Fig. 3A,B and Table S2). For this comparison, 238 of the proteins were found down-regulated in non-IDC/non-Crib compared to BPD, with a $\log_2$ range of regulation from −6.57 to −0.25 (Fig. 3B). The range for the 118 up-regulated proteins ranged from 0.19 to 4.25. The comparison between IDC/Crib and BPD exhibited less regulated proteins (Fig. 3A,C and Table S3). The range of significantly regulated proteins was 0.58–3.89 for up-regulated proteins and −3 to −0.23 for down-regulated proteins. Comparing the merged non-IDC/non-Crib and IDC/Crib into the group cancer for comparison with BPD resulted in a similar pattern of regulation as for the IDC/Crib and BPD comparison (Fig. 3D, Table S4). Suggesting that IDC/Crib and non-IDC/non-Crib are identified as distinct entities based on urinary EV proteome. Therefore, merging IDC/Crib and non-IDC/non-Crib results in large variance for statistical comparisons. The regulated proteins parsed for functional analysis were additionally filtered to be at least twofold regulated (indicated in Fig. 3B-F sub-title). Grouping patients into significant PCa (ISUP ≥ 2) and non-significant PCa (ISUP = 1) also resulted in regulated proteins after correction of multiple testing (Fig. 3E, Table S5). Adjusting the linear models for prePSA, age and batch number resulted in only minor differences in the list of regulated proteins (Fig. 3F, Table S6 versus Fig. 3E, Table S5).

To identify significant differences between IDC/Crib versus non-IDC/non-Crib and IDC versus cribriform at EV proteome level, the IDC/Crib group characteristic of aggressive PCa was divided into more precise subgroups. Pairwise comparisons of different combinations of cribriform and IDC samples versus non-IDC/non-Crib patient samples resulted in many significantly regulated proteins after correction for multiple testing (Fig. 4) compared to the comparisons using BPD group as a reference. Comparison between cribriform and IDC patient samples also revealed significantly regulated proteins, but not after correction for multiple testing (Fig. 4F). Again model adjustments by prePSA, age and batch effect resulted in only minor differences. However, including adjustment for ISUP or Gleason grade eliminated almost all significantly regulated proteins. Nevertheless, for the comparison IDC/Crib versus non-IDC/non-Crib two proteins (S100A10 and PTGES3) showed significant dysregulation after correction for multiple testing and including adjustment for ISUP in the linear model (Table S7).

A

| Comparison | Up | Down |
|---|---|---|
| Non-IDC/non-Crib vs BPD | 18(118) | 85(238) |
| IDC/Crib vs BPD | 1(49) | 0(158) |
| Cancer vs BPD | 1(38) | 0(170) |
| ISUP ≥ 2 vs ISUB = 1 | 11(78) | 22(64) |
| ISUP ≥ 2 vs ISUB = 1 * | 10(76) | 28(70) |

*) Adjusted model

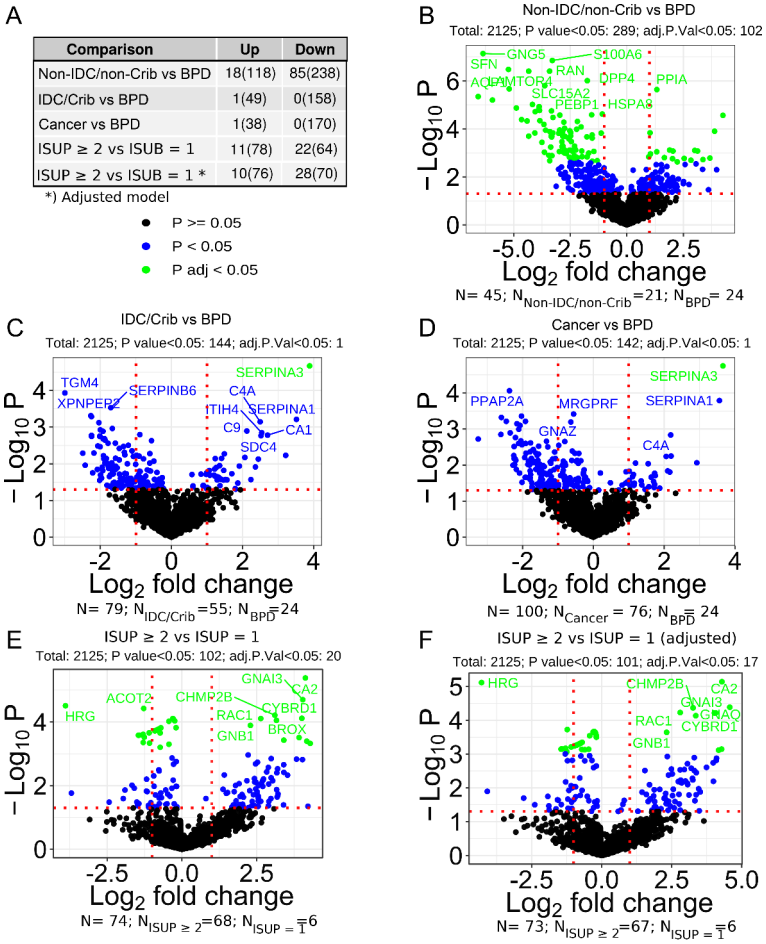


**Fig. 3**. Overview of regulated proteins for three pairwise comparisons. (**A**) Table summary of up- and down-regulated proteins applying a 0.05 adjusted P value and no threshold on effect size. The numbers in parenthesis represent regulated proteins with similar thresholds but applying regular P values. Volcano plots for the pairwise comparisons (**B**) non-IDC/non-Crib versus BPD, (**C**) IDC/Crib versus BPD, (**D**) cancer versus BPD, (**E**) ISUP ≥ 2 versus ISUP = 1 and (**C**) ISUP ≥ 2 versus ISUP = 1 considering adjustment of other clinical variables. Red horizontal line indicates 0.05 P value threshold. Red vertical lines indicate two-fold differential regulation. The number of regulated proteins indicated in the titles is based on P values and adjusted P values with an effect size greater than twofold.

## Proteins correlated with Gleason score and pre-PSA

Several protein expression patterns in urinary EVs were observed to correlate with clinical parameters such as pre-biopsy PSA, Gleason grade, and number of positive cores. Figure 5 displays box plots and *P values* calculated by Jonckheere's test for the four significant increased protein expressions as Gleason score severity increase. These proteins include histone cluster 1, H2be (HIST1H2BE), immunoglobulin J chain (JCHAIN), HPX (hemopexin) and alpha-1-antitrypsin (SERPINA1). In addition, to JCHAIN and a number of other immunoglobulin related proteins displayed significant increased trends as a severity of Gleason score increase (not shown).

prePSA measurements were also significantly correlated with increasing Gleason score (Jonckheere's test *P value* < 0.001). Therefore the proteins correlated with prePSA were similar to the ones correlating with Gleason score. The four most correlated proteins to prePSA were SERPINA1, IGLV3-21 (Immunoglobulin lambda variable 3-21), SERPINA3 and C9 (Complement component C9) were also among the most significantly correlated to Gleason score.

## Functional analysis of regulated proteins

Significantly regulated proteins with an effect size of at least two fold were subjected to functional enrichment analysis (Figs. 6). In the first analysis, all regulated proteins with at least two-fold regulation were submitted for each of the comparisons using the BPD group as reference (Figures S3A-C). Fatty acid metabolism appeared as the most relevant functional group when comparing non-IDC/non-Crib and BPD groups (Fig. 6A, highlighted in dashed box). When comparing IDC/Crib group or whole PCa cancer group versus BPD, androgen response related proteins appeared as the most relevant functional group (Fig. 6A and C, highlighted in dashed box).

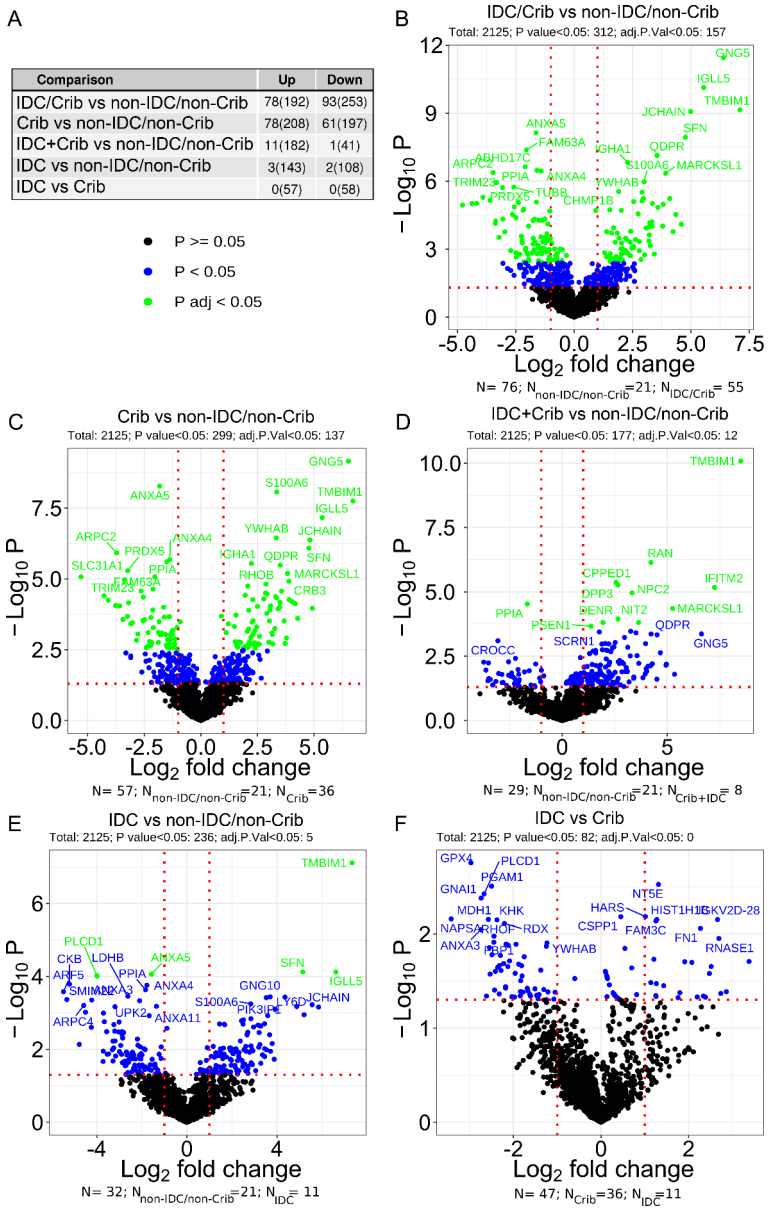


**Fig. 4**. Volcano plots for the pairwise comparisons. (**A**) Number of significantly regulated proteins for each of the pairwise comparisons given as $N_{adjust\text{-}p\text{-}value}$ ($N_{regular\text{-}p\text{-}value}$). (**B**) IDC/Crib versus non-IDC/non-Crib, (**C**) crib versus non-IDC/non-Crib, (**D**) cribriform and IDC versus non-IDC/non-Crib, and (**E**) IDC versus non-IDC/ non-Crib. (**F**) Crib versus IDC. Red horizontal line indicates 0.05 P value thresholds. Red vertical lines indicate two-fold differential regulation. The number of regulated proteins indicated in the titles is based on P values and adjusted P value with an effect size greater than twofold.

To provide additional information on the overall direction of regulation in the functional groups, heatmaps summarizing *p value* enrichment based on all regulated proteins with a twofold effect size for all comparisons with BPD as reference group (Fig. 6A), all regulated proteins with two-fold up-regulation (Fig. 6B) and all regulated proteins with two-fold down-regulation were plotted (Fig. 6C). Androgen response appears overall down-regulated for cancer group compared to BPD group. Furthermore, IDC/Crib group has the most significant down-regulation of androgen response compared to non-IDC/non-Crib group (Fig. 6C). For up-regulated proteins, fatty acid metabolism showed a correlation in non-IDC/non-Crib group whereas epithelial-mesenchymal transition is involved in the IDC/Crib group (Fig. 6B). In the detailed analysis, androgen response, epithelial-mesenchymal transition, and reactive oxygen species appeared as the main functional entities playing a role in distinguishing the subgroups (Figures S6A-C).

## Discussion

There are some previous clinical proteomics studies on urinary small EVs for the diagnosis of PCa[20–23], which were recently reviewed in Bernardino et al[9]. Nonetheless, analyzed cohorts were typically small, and the clinical

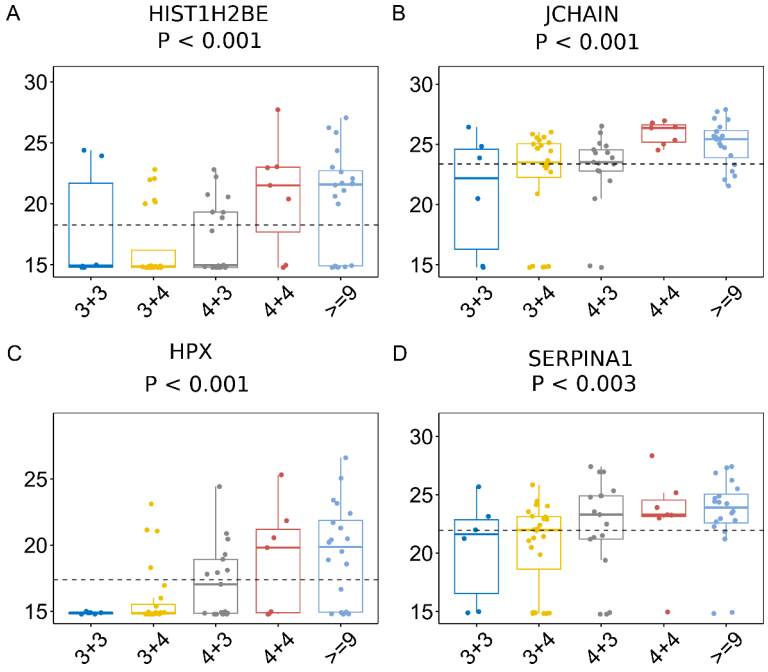


**Fig. 5**. Box plot displaying the distribution of iBAQ values across increasing levels of Gleason score. The box plot showcases the median (represented by the horizontal line inside the box), interquartile range (IQR; the box's height), and the minimum and maximum values (whiskers) within each level. Increasing trends are displayed for (**A**) HIST1H2BE, (**B**) JCHAIN, (**C**) HPX, and (**D**) SERPINA1. P value calculated by Jonckheere's test.

groups considered were either PCa versus controls or based on Gleason score classification only. Specific histopathological patterns with prognostic impact, such as IDC and Crib, were not considered in those reports.

Herein, we isolated small urinary EVs from 100 individuals (Table 1). The isolated urinary EVs displayed the highest level of established EV markers compared to previous cell line studies and studies performed on other types of biofluids (Fig. 2A). Although we identified THP as an abundant contaminant protein, THP was not identified as significantly regulated among samples. According to previous reports, 70% of all urinary proteins originate from kidney tissue[24]. According to our analysis of the proteome of small urinary EVs, the overlap of identified proteins between prostate, kidney, and bladder tissue with small urinary EVs is quite similar (Figure S4). Only 367 proteins (3%) could not be explained by the three main organs in more close contact with urine (Figure S4D). We consequently argue that small urinary EVs are promising sources of biomarkers for diseases affecting the kidney, bladder, or prostate. Our findings also parallel those of Dhondt et al[20], who compared urinary EV proteomes with prostate tissue. Concerning the number of uniquely identified proteins per clinical subgroup, IDC + Crib and Crib displayed the highest number of identified proteins. This is concordant with a previously published hypothesis that most advanced cancers display a higher richness in proteins[15]. Most of the uniquely identified proteins in the current study were identified in a few individuals, but for crib and non-IDC/non-Crib, one protein was consistently identified only within the specific clinical group (Figures S5B & D).

TMBIM1, also known as Bax inhibitor 1 (BI-1), was among those most upregulated in Crib and IDC compared to prostate cancer without IDC/Crib (Fig. 4B-E). Bax inhibitor 1 (BI-1), plays a crucial role in regulating apoptosis and calcium homeostasis by inhibiting the activity of the pro-apoptotic protein Bax and protecting cells from apoptosis induced by various stimuli. In PCa, Bax inhibitor-1 is overexpressed[25]. Bax inhibitor-1 specific down-regulation by RNA interference leads to cell death in human PCa cells[25]. TMBIM1 has also been previously identified as up-regulated in urinary EVs in patients with high-grade prostate cancer (poor prognosis)[23]. Bax inhibitor-1 has also been implicated in glioblastoma multiforme, colon cancer[26] and drug resistance in hepatocellular carcinoma[27].

Another protein factor consistently up-regulated in Crib and IDC found in our study is GNG5 (guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-5) and IGLL5 (immunoglobulin lambda-like polypeptide 5). The higher expression level of heavy immunoglobulins (Igs) in the urine has been associated with several diseases of the urinary tract. Moreover, in gammopathies such as multiple myeloma, in which plasma cells subpopulation in the bone marrow increases to 10%, from the normal 2–3%, immunoglobulins are found in the urine. In this study we have specifically identified an immunoglobulin subclass that correlated with the Gleason score. GNG5 and IGLL5 are involved in immune regulation and B-cell activation, respectively, suggesting regulated immune function. Recently, fusion transcripts of GNG5 have been identified in PCa which might explain GNG5 dysregulation protein in urinary EVs from PCa patients[28]. GNG5 has also been recently described as a novel oncogene associated with cell migration, proliferation, and poor prognosis in gliomas[29].

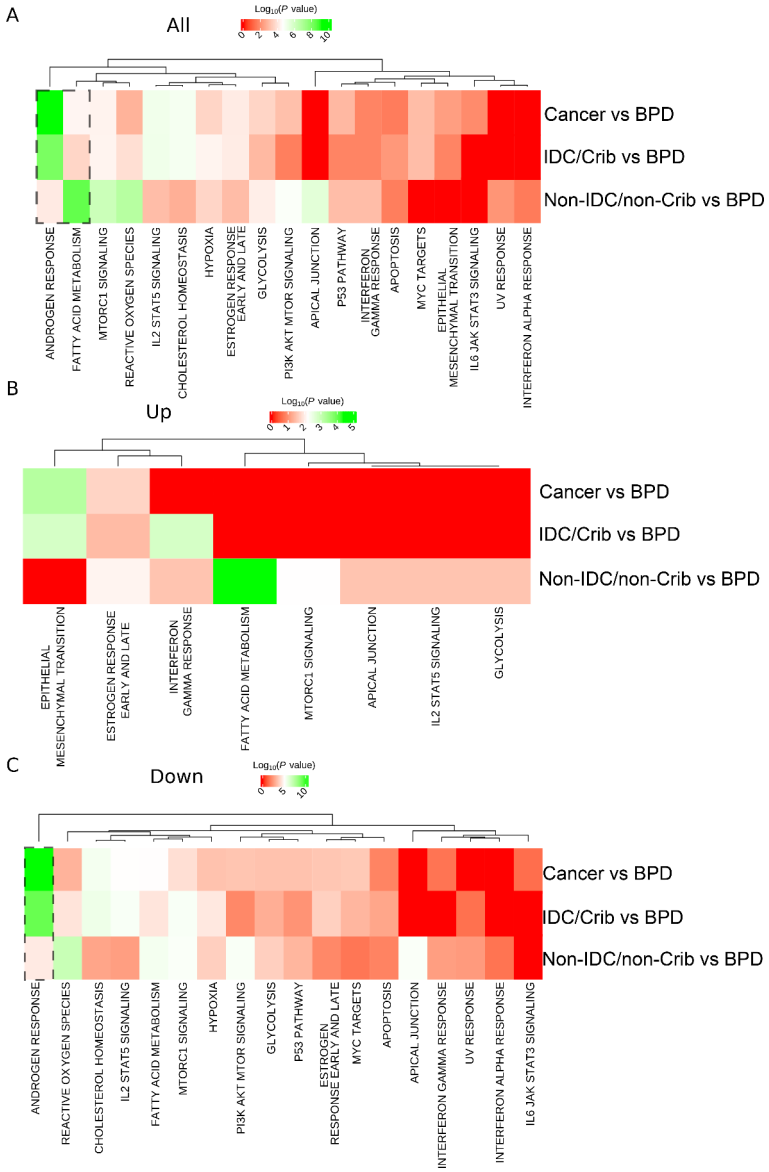


**Fig. 6**. Functional enrichment analysis of cancer hallmark proteins based on significant regulated proteins for pairwise comparisons to control: Heatmap summarizing the functional enrichment of cancer hallmark proteins for each of the three pairwise comparisons to control based on (**A**) all regulated proteins, (**B**) up-regulated proteins and (**C**) down-regulated proteins. . Functional analysis based on significant regulated proteins for pairwise comparisons of different histological subgroups to non-IDC/non-Crib and cribriform: Functional enrichment against cancer hallmarks was performed for (**D**) all regulated proteins ($p < 0.05$), (**E**) all up-regulated proteins, and (**F**) all down-regulated proteins for each of the pairwise comparisons.

Finally, G protein subunit gamma 5 in hepatocellular carcinoma is a prognostic biomarker and is correlated with immune infiltrates[30].

Gleason grade and IDC/Crib are correlated (Table 1). However, adjusting for ISUP in the limma regression models still resulted in the proteins S100A10, also known as p11, and Prostaglandin E synthase enzyme3 (PTGES3) significantly regulated after correction for multiple testing (Table S7). S100A10 is considered a significant cancer promotor[31]. Although, in urinary sEVs from IDC/Crib S100A10 was significantly down regulated in urinary sEVs whereas in advanced tumor stages it is reported up regulated. PTGES3 is a well-studied oncogene, also named p23. PTGES3 has been suggested to be overexpressed in multiple cancers, including breast cancer, colorectal cancer, cervical cancer and lung[32]. PTGES3 also correlates with poor prognosis. PTGES3 is required for proper functioning of the glucocorticoid and other steroid receptors. Again, PTGES3 exhibited a significant downregulation in urinary small extracellular vesicles and particles (sEVs), contrasting with the observed reverse dysregulation in tissue. Perhaps diminished extracellular vesicle and particle (EV) secretion on the tissue level of these oncoproteins may contribute to the elevated tissue levels.

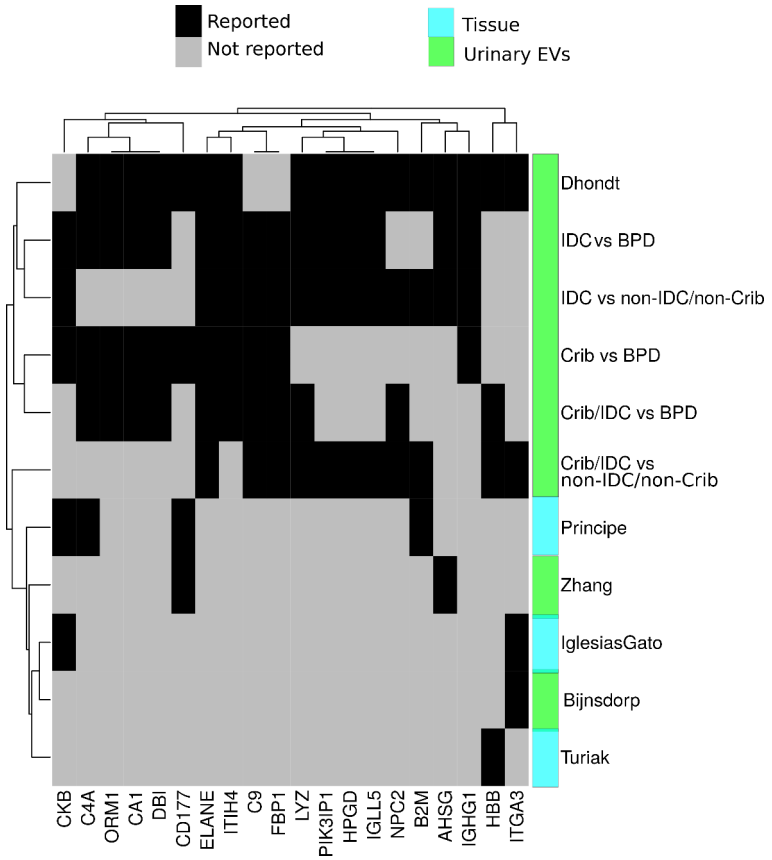


**Fig. 7**. Heatmap overview of the 20 most frequent reported differential regulated proteins from current study compared with previous studies.

| References | Year | Target sample |
|---|---|---|
| Dhondt et al[20] | 2020 | Urinary EVs |
| Sequeiros et al[23] | 2017 | Urinary EVs |
| Fujita et al[21] | 2017 | Urinary EVs |
| Zhang et al[22] | 2020 | Human Seminal Plasma |
| Principe et al[33] | 2012 | Prostatic secretions |
| Kawahara et al[34] | 2019 | Prostate tissue |
| Turiak et al[35] | 2019 | Prostate tissue |
| Iglesias-Gato[18] | 2016 | Prostate tissue |
| Bijnsdorp et al[36] | 2013 | Cell line |

**Table 2**. Overview of past prostate cancer MS-based proteomics studies.

Significantly regulated proteins identified in our study based on pairwise comparisons were compared with previous proposed biomarkers based on either prostate tissue-specific[33], prostate tissue marker[18,34,35] or urinary and cell line EVs marker[20–23,36] (Fig. 7, Table 2). Note there are no previous studies on IDC and Crib targeting urinary EV proteome.

Strikingly, our study did not show any particular overlap with the previous studies by Fujita et al. [21] and Kawahara et al. [34]. Most overlapping proteins are well described in association with cancer. For example, Fetuin-A (AHSG) is described as driving pancreatic, prostate, and glioblastoma tumors[37]. Circulating blood and urine B2M is a well-established marker of cancer[38]. ELANE and CD177 are neutrophil markers with previous association with prostate cancer.

On the other hand, the peripheral zone constitutes the most common site of origin of neoplasms in the aged prostate and its stroma contains, among other cell types, fibroblasts. These, by inducing epithelial transformation (EMT) and stimulating survival signaling, contribute to an increase in cancer cells invasion and metastisation.

A previous study on PCa[39] identified HIST1H2BE as predictor of Gleason grade, which we identified as correlating with Gleason score (Fig. 5). According to research conducted on a cohort of acute lymphoblastic

leukemia (ALL), in deceased patients, the overexpression of JCHAIN was presumably connected to tumor aggression[40]. JCHAIN encodes the immunoglobulin J chain and joins the monomer units of IgA and IgM. HPX has long been regarded as the ultimate scavenger of labile heme and the plasma protein with the highest affinity for heme. A previous study observed low levels of HPX in prostate tumors and in the plasma of prostate cancer patients[41]. Perhaps increased urinary secretion of HPX occurs in cancer. Lung cancer, gastric cancer, and colorectal cancer were demonstrated to have altered invasive and metastatic capacities in response to the serine protease inhibitor serpinA1[42]. serpinA1 was also identified when comparing PCa to BPD (Fig. 3D) and when correlating protein expression with prePSA.

Androgen response (AR) refers to the ability of PCa cells to respond to androgens. Indeed, AR plays a crucial role in both the onset and spread of PCa[43]. The majority of androgen-independent or hormone-refractory PCa express AR, and AR expression is sustained throughout disease progression. AR transcriptional activation in reaction to antiandrogens or other endogenous hormones, mutations of the androgen receptor, particularly mutations that result in a relaxation of AR ligand specificity, may contribute to the progression of prostate cancer and the failure of endocrine therapy. There is evidence to suggest that androgen response can decline in advanced or aggressive PCa[44]. Interestingly, we also observed IL2 Stat5 pathway dysregulation, which might be related to the fact that, in PCa cells, transcription factor Stat5 synergizes with the androgen receptor[45].

Overall, significantly regulated proteins in urinary EVs mostly resemble the profiles characterized in previous studies on PCa as well as other cancers.

This study has several limitations. The sample size was estimated to obtain proof of concept as a pilot study and is single centered. It's essential to acknowledge that while our findings are promising, they may not be directly applicable to all patient populations or settings. The study was conducted in a specific context with a particular patient cohort, and external validation is necessary to ensure generalizability to other groups. Global label free quantitation was performed to target many proteins in an unbiased manner. More precise targeted protein quantitation can be performed in follow up studies to measure concentration of biomarkers. Participants were randomly assigned into two experimental batches except from non-IDC/non-Crib and potential batch effect was corrected for in the analysis. All patients fulfilling the eligibility criteria were enrolled thereby minimizing selection bias. Reporting bias was minimized by only addressing pairwise comparisons related to the objectives in the study design. Future studies must improve the clinical study design to improve the disentanglement of the association between Gleason grade and IDC/Crib. The project funding was for one year limiting the whole study to one and half year.

## Conclusion

Based on our analysis, we conclude that the proteome of small urinary EVs holds diagnostic potential and reflects the proteome of prostate tissue. Urinary EV proteome highlighted proteins with a role in androgen response, epithelial mesenchymal transition, fatty acid metabolism and reactive oxygen species as prognostic factors for prostate cancer.

## Online methods

### Patients

Patients were selected for EV isolation followed by EV proteome profiling based on criteria defined below. Information was collected prospectively in a single time for each patient before any treatment (treatment naïve). Patient enrollments started in April 2020 and ended in May 2022. All enrollments were from a single center, Centro Hospitalar e Universitário Lisboa Central, Lisbon, Portugal. Enrollment was continued until the number of samples were in accordance with the study protocol. The final cohort was composed of BPD (N = 24) patients without Crib and IDC histological prostate patterns (non-IDC/non-Crib, N = 21), and patients with Crib and/or IDC histological prostate patterns (IDC/Crib, N = 55) at biopsy.

*Inclusion Criteria*

1. Men over 18 years old.
2. No previous history of PCa treatments.

*Exclusion Criteria*

1. History of other forms of focal treatment of PCa
2. Radical surgery performed in the context of "salvage" strategy, due to recurrence or local persistence.
3. Neoadjuvant and/or adjuvant treatment (includes any type of hormonotherapy as LHRH agonists/antagonists)
4. History of urothelial cancer (bladder or upper urinary tract)

Mid-stream urine (30–120 mL) from PCa suspects were collected, and immediately frozen at − 80° C and stored upon collection until EV isolation. The experimental protocols were approved by the medical agencies and ethics committees of NOVA Medical School (82/2020/CEFCM). All patients signed an informed consent before trial participation.

The collected clinical data included prePSA (ELISA), histological type, Gleason grade/score (World Health Organization (WHO) and the College of American Pathologists (CAP)), number of positive cores, histological patterns Crib and IDC-p at biopsy. Samples were organized into two experimental batches for urinary EV isolation followed by LC–MS analysis with random batch allocation of the sample groups BPD and IDC/Crib. Subsequent analysis adjusted for batch effects in the limma regression models to minimize confounding batch effects. All

clinical measurements were obtained in a blinded manner without knowledge of the final clinical outcome. This approach ensured that the assessors performing the measurements remained unbiased and uninfluenced by the eventual results, minimizing potential researcher bias. The numbers of samples to collect were estimated based on standard error obtained on LC–MS from previous studies performed in our group and estimated with the R function pwr assuming paired testing. The number of samples collected and the single center collection were considered appropriate for a pilot study. Strobe check list is presented in supplementary data.

### Isolation of extracellular vesicles and particles from urine

Frozen urine specimens were thawed and centrifuged at 3000×*g* for 20 min at 4 °C and then at 12,000×*g* for 60 min at 4 °C. Clarified urine was ultracentrifuged in an Optima TM L-80XP ultracentrifuge (Beckman Coulter, Brea, CA, USA) at 170,000×*g* at 4 °C for 120 min with a Type 32 Ti rotor to pellet EVs. The supernatant was carefully removed, and crude EV-containing pellets were resuspended in ice-cold PBS. This study has been registered with the EV-TRACK knowledgebase (EV-TRACK ID: EV240158)[46]. We have followed the MISEV2018 guidelines[10] for the characterization and reporting of extracellular vesicles, ensuring transparency and reproducibility of our methods and results. Details of our experimental protocols, including EV isolation, characterization techniques, and relevant data, are available on the EV-TRACK platform for public reference.

### Protein measurements

Following manufacturer's instructions, a bicinchoninic acid (BCA) protein assay kit (Pierce Biotechnology, Rockford, IL, USA) was used to measure the protein concentrations in isolated exosome fractions.

### Western blotting

For western blotting (WB) assay, 5 µg sEV protein was mixed with Laemmli sample buffer (BioRad) boiled for 10 min at 100 °C. Then, the samples were resolved by SDS-PAGE followed by transfer onto nitrocellulose membranes (Cytiva). Blocking was performed during 1 h with 5% skim milk in TBST 0,1% or PBST 0,1% Tween or 5% BSA in TBST 0,1% Tween. Primary antibodies (CD63, SICGEN (AB0047); Alix, SICGEN (AB0327), TOM20, BD Biosciences (612,278), GRP75, Cell Signaling Technology (2816S)) were incubated overnight at 4 °C and secondary antibodies (HRP-AffiniPure Donkey Anti-Goat IgG (H + L), HRP-AffiniPure Goat Anti-Mouse IgG (H + L), HRP- Affini Pure Goa Anti-Rabbit IgG (H + L), Jackson Immuno-Research) during 1 h at room temperature (RT). Development was performed using ECL™ prime Western blotting detection reagent (Cytiva) and the Chemidoc Touch Imager (BioRad).

### Nanoparticle tracking EV measurements

A NanoSight NS300 instrument (Malvern Panalytical, Malvern, UK) was used to determine the concentrations and sizes of the EVs in the samples. Samples were diluted in PBS to a final volume of 1 ml to reach the ideal particle concentration of $1\times10^8$–$2\times10^9$ particles/mL. The samples were loaded to the sample chamber in a continuous flow by a syringe pump. The instrument was equipped with a 488 nm laser and a sCMOS camera. The focus for each sample was manually adjusted to achieve optimal visualization of particles and for each measurement five videos of 60 s were captured. For all experiments the following settings were used: temperature: 25 °C; Syringe speed: 20; Viscosity: 0.9 cP; camera level setting ranged from 13–14 in light scatter mode (LSM). After capture, the videos have been analysed by the in-build NanoSight Software NTA 3.4 Build 3.4.4 with a detection threshold of 5. To minimize variability, all camera and detection threshold settings were kept the same and all particles over 300 nm of diameter were excluded from the analysis.

### *Electron* microscopy

5 µL of each sample was incubated on glow-discharged (0.5 min) formvar-carbon coated copper mesh grids (Electron Microscopy Sciences) for 2 min, before washing 10 times with dH2O. Samples were negatively stained with 2% uranyl acetate in dH2O for 2 min, before blotting dry and imaging with a Hitachi H-7650 TEM equipped with an AMT XR41 M digital camera.

### Peptide sample preparation

Samples containing a minimum of 20 µg of total EV proteins were further processed by the filter-aided sample preparation (FASP) method. In short, protein solutions containing SDS and DTT were loaded onto filtering columns (Millipore, Billerica, MA, USA) and washed exhaustively with 8 M urea (GE, Healthcare, Marlborough, MA, USA) in HEPES buffer (Sigma-Aldrich, Saint Louis, MO, USA) as previously described[47,48]. Proteins were equilibrated with ammonium bicarbonate solution prior to trypsin digestion overnight at 37 °C (Sigma-Aldrich, Saint Louis, MO, USA). Overnight cleavage of proteins was carried out using sequencing-grade trypsin (Promega, Madison, WI, USA).

### Mass spectrometry analysis

As previously described[14], samples were analyzed by mass spectrometry-based proteomics using nano-LC-MSMS equipment (Dionex RSLCnano 3000) coupled to an Exploris 480 Orbitrap mass spectrometer (Thermo Scientific, Hemel Hempstead, UK). In brief, samples were loaded onto a custom-made fused capillary pre-column (2 cm length, 360 µm OD, 75 µm ID, flowrate 5 µL per minute for 6 min) packed with ReproSil Pur C18 5.0 µm resin (Dr. Maisch, Ammerbuch-Entringen, Germany), and separated using a capillary column (25 cm length, 360 µm outer diameter, 75 µm inner diameter) packed with ReproSil Pur C18 1.9-µm resin (Dr. Maisch, Ammerbuch-Entringen, Germany) at a flow of 250 nL per minute. A 56 min linear gradient from 89% A (0.1% formic acid) to 32% B (0.1% formic acid in 80% acetonitrile) was applied. Mass spectra were acquired in positive ion mode in a data-dependent manner by switching between one Orbitrap survey MS scan (mass

range m/z 350–m/z 1200) followed by the sequential isolation and higher-energy collision dissociation (HCD) fragmentation and Orbitrap detection of fragment ions of the most intense ions with a cycle time of 2 s between each MS scan. MS and MSMS.

settings: maximum injection times were set to "Auto", normalized collision energy was 30%, ion selection threshold for MSMS analysis was 10,000 counts, and dynamic exclusion of sequenced ions was set to 30 s.

## Database search

The data obtained from the 200 LC–MS runs of urine EV samples from 24 controls and 76 PCa cases, characterized following radical prostatectomy (55 with and 21 without Cribriform pattern and/or IDC) each run as technical duplicates were analyzed. The LC–MS data were searched using VEMS[49] and MaxQuant[50] (Version 2.1.0.0). The MSMS spectra were searched against a standard human proteome database from UniProt (3AUP000005640).

Permuted protein sequences, where arginine and lysine were not permuted, were included in the database for VEMS and FDR in MaxQuant version 2.1.0.0 were based on reversed sequences. 1% FDR threshold was applied for peptide and protein identifications. Trypsin cleavage allowing a maximum of four missed cleavages was used. Carbamidomethyl cysteine was included as fixed modification. Methionine oxidation, lysine and N-terminal protein acetylation, were included as variable modifications. No restriction was applied for minimal peptide length for VEMS search. All other search parameters were default values. The downstream analysis presented is based on the MaxQuant results.

## Estimation of analytical variability

In our comprehensive proteomic investigation, we meticulously evaluated the precision of our experimental raw measurements (prior to quality filtering or normalization), as evidenced by a calculated average coefficient of variation (CV) of 34.1%. The mean CV was estimated to 13.1% after normalization. This average CV is based on all measurements on all proteins in the technical replicas. This indicative measure underscores the reliability and consistency of protein abundance quantification across technical replicates, affirming the robustness of our proteomic profiling methodology.

## Statistical analysis

Statistical analysis of identified proteins was performed in R statistical programming language. Quantitative data from MaxQuant and VEMS were analyzed in R statistical programming language version 4.04 (The R Foundation, Vienna, Austria). Protein label free quantitation (iBAQ) and protein spectral counts from the two programs were preprocessed by removing common MS contaminants, followed by a $\log_2(x+1)$ transformation and removing common MS contaminants. iBAQ values from the duplicated measurements were averaged. No imputation of missing or zero value protein quantitation values were performed in the analysis. Information on sample grouping based on histological patterns which were used for pairwise comparisons were complete for all samples. Protein iBAQ values were subjected to statistical analysis utilizing the R package limma[19], where the contrast for different pairwise comparisons was specified for the main clinical groups BPD, non-IDC/non-Crib and IDC/Crib (Tables S1–S3). Samples were processed in two large batches to minimize experimental bias. For sensitivity analysis, various linear regression models including terms to correct for batch effect and PSA were tested and these models displayed minimal effect on the number significantly regulated proteins called after correction for multiple testing. For example, batch effect had no effect for the comparison IDC/Crib versus BPD and cancer versus BPD. For non-IDC/non-Crib versus BPD only a difference of two more significantly regulated proteins were observed.

Correction for multiple testing was applied using the method of Benjamini & Hochberg[51]. Volcano plots were constructed with ggplot software (The R Foundation, Vienna, Austria). To test for increasing trend in iBAQ values as Gleason grade increase, the Jonckheere's test were calculated using the R package clinfun[52]. It examined whether there is a significant trend in iBAQ values across the increasing levels of Gleason grade. A low p-value indicates strong evidence against the null hypothesis of no trend, suggesting a significant increasing pattern. For correlation analysis a few missing values were present for same patients and cases with missing values for correlation analysis were excluded. Sensitivity of the analysis was assessed by comparing protein markers obtained by correlating to clinical parameters that are known to correlate with sample grouping.

## Functional enrichment analysis

Functional enrichment based on the hypergeometric probability test was performed as described previously in R[53,54]. Functional enrichment was based on extracting all functional categories for which at least one of the samples showed a significant enrichment based on the hypergeometric probability test[53,54]. For these functional categories, the matching proteins' gene names and numbers of proteins matching the functional categories were extracted, and the estimated p values were $-\log_{10}$ transformed and plotted as heatmaps. Functional enrichment was performed for all identified proteins in each sample group and for deregulated proteins when comparing sample groups. Cellular component (CC), biological process (BP), molecular function (MF), KEGG and cancer hallmark functional annotations were considered in the analysis.

## Data availability

The mass spectrometry proteomics data that support the findings of this study have been deposited in ProteomeXchange Consortium[55] via the PRIDE[56] partner with the PXD043874 accession code.

Received: 11 May 2024; Accepted: 3 October 2024
Published online: 23 October 2024

## References


1. Sung, H. et al. Global cancer statistics 2020: GLOBOCAN Estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.***71**, 209–249. https://doi.org/10.3322/caac.21660 (2021).
2. Bernardino, R. M. et al. Prostate cancer with cribriform pattern: Exclusion criterion for active surveillance?. *Arch. Ital. Urol. Androl. Organo Ufficiale [di] Soc. Ital. Ecogr. Urol. Nefrol.*https://doi.org/10.4081/aiua.2020.3.235 (2020).
3. Masoomian, M. et al. Concordance of biopsy and prostatectomy diagnosis of intraductal and cribriform carcinoma in a prospectively collected data set. *Histopathology***74**, 474–482. https://doi.org/10.1111/his.13747 (2019).
4. Truong, M. et al. Impact of gleason subtype on prostate cancer detection using multiparametric magnetic resonance imaging: Correlation with final histopathology. *J. Urol.***198**, 316–321. https://doi.org/10.1016/j.juro.2017.01.077 (2017).
5. Bernardino, R. & Fleshner, N. Re: Sensitivity of multiparametric MRI and targeted biopsy for detection of adverse pathologies (cribriform gleason pattern 4 and intraductal carcinoma): Correlation of detected and missed prostate cancer foci with whole mount histopathology. *Eur. Urol.*https://doi.org/10.1016/j.eururo.2023.01.028 (2023).
6. Ericson, K. J. et al. Diagnostic accuracy of prostate biopsy for detecting cribriform gleason pattern 4 carcinoma and intraductal carcinoma in paired radical prostatectomy specimens: Implications for active surveillance. *The J. Urol.***203**, 311–319. https://doi.org/10.1097/JU.0000000000000526 (2020).
7. Yanez-Mo, M. et al. Biological properties of extracellular vesicles and their physiological functions. *J. Extracell. Vesicles***4**, 27066. https://doi.org/10.3402/jev.v4.27066 (2015).
8. Conde-Vancells, J. et al. Characterization and comprehensive proteome profiling of exosomes secreted by hepatocytes. *J. Proteome Res.***7**, 5157–5166. https://doi.org/10.1021/pr8004887 (2008).
9. Bernardino, R. M. M. et al. Extracellular vesicle proteome in prostate cancer: A comparative analysis of mass spectrometry studies. *Int. J. Mol. Sci.*https://doi.org/10.3390/ijms222413605 (2021).
10. Thery, C. et al. Minimal information for studies of extracellular vesicles 2018 (MISEV2018): A position statement of the International society for extracellular vesicles and update of the MISEV2014 GUIDELINES. *J. Extracell. Vesicles***7**, 1535750. https://doi.org/10.1080/20013078.2018.1535750 (2018).
11. Webber, J. & Clayton, A. How pure are your vesicles?. *J. Extracell. Vesicles*https://doi.org/10.3402/jev.v2i0.19861 (2013).
12. Hurwitz, S. N. et al. Proteomic profiling of NCI-60 extracellular vesicles uncovers common protein cargo and cancer type-specific biomarkers. *Oncotarget***7**, 86999–87015. https://doi.org/10.18632/oncotarget.13569 (2016).
13. Hoshino, A. et al. Extracellular vesicle and particle biomarkers define multiple human cancers. *Cell***182**, 1044–1061. https://doi.org/10.1016/j.cell.2020.07.009 (2020).
14. Carvalho, A. S. et al. Proteomic landscape of extracellular vesicles for diffuse large B-cell lymphoma subtyping. *Int. J. Mol. Sci.*https://doi.org/10.3390/ijms222011004 (2021).
15. Carvalho, A. S. et al. Is the proteome of bronchoalveolar lavage extracellular vesicles a marker of advanced lung cancer?. *Cancers*https://doi.org/10.3390/cancers12113450 (2020).
16. Matthiesen, R. et al. Extracellular vesicles in diffuse large b cell lymphoma: Characterization and diagnostic potential. *Int. J. Mol. Sci.*https://doi.org/10.3390/ijms232113327 (2022).
17. Prakash, A. et al. Integrated view of baseline protein expression in human tissues. *J. Proteome Res.***22**, 729–742. https://doi.org/10.1021/acs.jproteome.2c00406 (2023).
18. Iglesias-Gato, D. et al. The proteome of primary prostate cancer. *Eur. Urol.***69**, 942–952. https://doi.org/10.1016/j.eururo.2015.10.053 (2016).
19. Smyth, G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.*https://doi.org/10.2202/1544-6115.1027 (2004).
20. Dhondt, B. et al. Unravelling the proteomic landscape of extracellular vesicles in prostate cancer by density-based fractionation of urine. *J. Extracell. Vesicles***9**, 1736935. https://doi.org/10.1080/20013078.2020.1736935 (2020).
21. Fujita, K. et al. Proteomic analysis of urinary extracellular vesicles from high Gleason score prostate cancer. *Sci. Rep.***7**, 42961. https://doi.org/10.1038/srep42961 (2017).
22. Zhang, X., Vos, H. R., Tao, W. & Stoorvogel, W. Proteomic profiling of two distinct populations of extracellular vesicles isolated from human seminal plasma. *Int. J. Mol. Sci.*https://doi.org/10.3390/ijms21217957 (2020).
23. Sequeiros, T. et al. Targeted proteomics in urinary extracellular vesicles identifies biomarkers for diagnosis and prognosis of prostate cancer. *Oncotarget***8**, 4960–4976. https://doi.org/10.18632/oncotarget.13634 (2017).
24. Kalantari, S., Jafari, A., Moradpoor, R., Ghasemi, E. & Khalkhal, E. Human urine proteomics: Analytical techniques and clinical applications in renal diseases. *Int. J. Proteom.***2015**, 782798. https://doi.org/10.1155/2015/782798 (2015).
25. Grzmil, M. et al. Bax inhibitor-1 is overexpressed in prostate cancer and its specific down-regulation by RNA interference leads to cell death in human prostate carcinoma cells. *Am. J. Pathol.***163**, 543–552. https://doi.org/10.1016/S0002-9440(10)63682-6 (2003).
26. Cai, J. et al. TMBIM1 promotes proliferation and attenuates apoptosis in glioblastoma cells by targeting the p38 MAPK signalling pathway. *Translat. Oncol.***19**, 101391. https://doi.org/10.1016/j.tranon.2022.101391 (2022).
27. Zhang, V. X. et al. Antioxidant supplements promote tumor formation and growth and confer drug resistance in hepatocellular carcinoma by reducing intracellular ROS and induction of TMBIM1. *Cell Biosci.***11**, 217. https://doi.org/10.1186/s13578-021-00731-0 (2021).
28. Malik, A., Srinivasan, S. & Batra, J. A new era of prostate cancer precision medicine. *Front. Oncol.***9**, 1263. https://doi.org/10.3389/fonc.2019.01263 (2019).
29. Zhang, W. et al. GNG5 is a novel oncogene associated with cell migration, proliferation, and poor prognosis in glioma. *Cancer Cell Int.***21**, 297. https://doi.org/10.1186/s12935-021-01935-7 (2021).
30. Wang, H., Yu, L., Cui, Y. & Huang, J. G Protein subunit gamma 5 Is a prognostic biomarker and correlated with immune infiltrates in hepatocellular carcinoma. *Dis. Markers***2022**, 1313359. https://doi.org/10.1155/2022/1313359 (2022).
31. Saiki, Y. & Horii, A. Multiple functions of S100A10, an important cancer promoter. *Pathol. Int.***69**, 629–636. https://doi.org/10.1111/pin.12861 (2019).
32. Gao, P. et al. High expression of PTGES3 is an independent predictive poor prognostic biomarker and correlates with immune infiltrates in lung adenocarcinoma. *Int. Immunopharm.***110**, 108954. https://doi.org/10.1016/j.intimp.2022.108954 (2022).
33. Principe, S. et al. Identification of prostate-enriched proteins by in-depth proteomic analyses of expressed prostatic secretions in urine. *J. Proteome Res.***11**, 2386–2396. https://doi.org/10.1021/pr2011236 (2012).
34. Kawahara, R. et al. Tissue proteome signatures associated with five grades of prostate cancer and benign prostatic hyperplasia. *Proteomics***19**, e1900174. https://doi.org/10.1002/pmic.201900174 (2019).
35. Turiak, L. et al. High sensitivity proteomics of prostate cancer tissue microarrays to discriminate between healthy and cancerous tissue. *J. Proteom.***197**, 82–91. https://doi.org/10.1016/j.jprot.2018.11.009 (2019).
36. Bijnsdorp, I. V. et al. Exosomal ITGA3 interferes with non-cancerous prostate cell functions and is increased in urine exosomes of metastatic prostate cancer patients. *J. Extracell. Vesicles*https://doi.org/10.3402/jev.v2i0.22097 (2013).
37. Ochieng, J. et al. Impact of fetuin-A (AHSG) on tumor progression and type 2 diabetes. *Int. J. Mol. Sci.*https://doi.org/10.3390/ijms19082211 (2018).
38. Prizment, A. E. et al. Circulating beta-2 microglobulin and risk of cancer: The atherosclerosis risk in communities study (ARIC). *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.***25**, 657–664. https://doi.org/10.1158/1055-9965.EPI-15-0849 (2016).

39. Chen, X. et al. An accurate prostate cancer prognosticator using a seven-gene signature plus Gleason score and taking cell type heterogeneity into account. *PloS one***7**, e45178. https://doi.org/10.1371/journal.pone.0045178 (2012).
40. Tomar, A. K., Agarwal, R. & Kundu, B. Most variable genes and transcription factors in acute lymphoblastic leukemia patients. *Interdiscip. Sci. Comput. Life Sci.***11**, 668–678. https://doi.org/10.1007/s12539-019-00325-y (2019).
41. Canesin, G. et al. Scavenging of labile heme by hemopexin is a key checkpoint in cancer growth and metastases. *Cell Rep.***32**, 108181. https://doi.org/10.1016/j.celrep.2020.108181 (2020).
42. Kwon, C. H. et al. Snail and serpinA1 promote tumor progression and predict prognosis in colorectal cancer. *Oncotarget***6**, 20312–20326. https://doi.org/10.18632/oncotarget.3964 (2015).
43. Culig, Z. & Santer, F. R. Androgen receptor signaling in prostate cancer. *Cancer Metastasis Rev.***33**, 413–427. https://doi.org/10.1007/s10555-013-9474-0 (2014).
44. Stelloo, S. et al. Androgen receptor profiling predicts prostate cancer outcome. *EMBO Mol. Med.***7**, 1450–1464. https://doi.org/10.15252/emmm.201505424 (2015).
45. Tan, S. H. et al. Transcription factor Stat5 synergizes with androgen receptor in prostate cancer cells. *Cancer Res.***68**, 236–248. https://doi.org/10.1158/0008-5472.CAN-07-2972 (2008).
46. Van Deun, J. et al. EV-TRACK: Transparent reporting and centralizing knowledge in extracellular vesicle research. *Nat. Methods***14**, 228–232. https://doi.org/10.1038/nmeth.4185 (2017).
47. Wisniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods***6**, 359–362. https://doi.org/10.1038/nmeth.1322 (2009).
48. Ferreira, I. et al. Messages from the small intestine carried by extracellular vesicles in prediabetes: A proteomic portrait. *J. Proteome Res.***21**, 910–920. https://doi.org/10.1021/acs.jproteome.1c00353 (2022).
49. Carvalho, A. S. et al. Global mass spectrometry and transcriptomics array based drug profiling provides novel insight into glucosamine induced endoplasmic reticulum stress. *Mol. Cell. Proteom. MCP***13**, 3294–3307. https://doi.org/10.1074/mcp.M113.034363 (2014).
50. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.***26**, 1367–1372. https://doi.org/10.1038/nbt.1511 (2008).
51. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B***57**, 289–300 (1995).
52. clinfun: Clinical trial design and data analysis functions v. R package version 1.1.1 (2023).
53. Hackenberg, M. & Matthiesen, R. Annotation-Modules: A tool for finding significant combinations of multisource annotations for gene lists. *Bioinformatics***24**, 1386–1393. https://doi.org/10.1093/bioinformatics/btn178 (2008).
54. Carvalho, A. S., Molina, H. & Matthiesen, R. New insights into functional regulation in MS-based drug profiling. *Sci. Rep.***6**, 18826. https://doi.org/10.1038/srep18826 (2016).
55. Deutsch, E. W. et al. The ProteomeXchange consortium in 2020: Enabling 'big data' approaches in proteomics. *Nucleic Acids Res.***48**, D1145–D1152. https://doi.org/10.1093/nar/gkz984 (2020).
56. Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.***47**, D442–D450. https://doi.org/10.1093/nar/gky1106 (2019).


## Acknowledgements

Acknowledgement: We thank Instituto Gulbenkian de Ciência for use of transmission electron microscopy.

## Author contributions

Conceptualization, R.B., R.L. R.M.; methodology, R.M., H.C.B. A.S.C.; formal analysis, RB, A.S.C., R.M.; investigation R.B., H.C.B., M.H., L.A., A.S.C., R.M.; resources, R.B., R.M.; writing—original draft preparation, R.B., N.F. and R.M.; writing—review and editing, RB, ASC, RL, RS, HP, JP, HCB, LCP, RH, NF, RM; project administration, R.B., A.S.C. and R.M.; funding acquisition, R.B.. All authors have read and agreed to the published version of the manuscript.

## Funding

The project is funded by Liga Portuguesa Contra o Cancro – Terry Fox Grant. R.M. is supported by Fundação para a Ciência e a Tecnologia (CEEC position, DOI: https://doi.org/10.54499/CEECIND/03906/2017/CP1421/CT0004). A.S.C. is supported by Fundação para a Ciência e a Tecnologia (DOI: https://doi.org/10.54499/DL57/2016/CP1457/CT0013). R.B is supported by FCT (Grant number 2022.13386.BD). R.M. and A.S.C. receive funding by programme and National Funds through FCT—Portuguese Foundation for Science and Technology under the projects number PTDC/BTM-TEC/1746/2021 and European Union to advance EV research (Horizon2020 GA n° 101079264, EVCA). We acknowledge the COST Action CA20113388 "PROTEOCURE" supported by COST (European Cooperation in Science and Technology). This article is a result of the projects (iNOVA4Health – UIDB/04462/2020 and UIDP/04462/2020, and by the Associated Laboratory LS4FUTURE (LA/P/0087/2020), two programs financially supported by Fundação para a Ciência e Tecnologia / Ministério da Ciência, Tecnologia e Ensino Superior.

## Declarations

## Competing interests

The authors of this research paper declare the existence of a potential conflict of interest due to the submission of a provisional patent application related to the published data.

## Additional information

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-75272-w.

**Correspondence** and requests for materials should be addressed to R.B., A.S.C. or R.M.