A Comprehensive Urine Proteome Database Generated From Patients With Various Renal Conditions and Prostate Cancer

Adam C. Swensen[1], Jingtang He[1], Alexander C. Fang[1], Yinyin Ye[1], Carrie D. Nicora[1], Tujin Shi[1], Alvin Y. Liu[2], Tara K. Sigdel[3], Minnie M. Sarwal[3] and Wei-Jun Qian[1]*

[1] Integrative Omics, Pacific Northwest National Laboratory, Biological Sciences Division, Richland, WA, United States, [2] Department of Urology, University of Washington, Seattle, WA, United States, [3] Department of Surgery, University of California, San Francisco, San Francisco, CA, United States

Urine proteins can serve as viable biomarkers for diagnosing and monitoring various diseases. A comprehensive urine proteome database, generated from a variety of urine samples with different disease conditions, can serve as a reference resource for facilitating discovery of potential urine protein biomarkers. Herein, we present a urine proteome database generated from multiple datasets using 2D LC-MS/MS proteome profiling of urine samples from healthy individuals (HI), renal transplant patients with acute rejection (AR) and stable graft (STA), patients with non-specific proteinuria (NS), and patients with prostate cancer (PC). A total of ~28,000 unique peptides spanning ~2,200 unique proteins were identified with a false discovery rate of <0.5% at the protein level. Over one third of the annotated proteins were plasma membrane proteins and another one third were extracellular proteins according to gene ontology analysis. Ingenuity Pathway Analysis of these proteins revealed 349 potential biomarkers in the literature-curated database. Forty-three percentage of all known cluster of differentiation (CD) proteins were identified in the various human urine samples. Interestingly, following comparisons with five recently published urine proteome profiling studies, which applied similar approaches, there are still ~400 proteins which are unique to this current study. These may represent potential disease-associated proteins. Among them, several proteins such as serpin B3, renin receptor, and periostin have been reported as pathological markers for renal failure and prostate cancer, respectively. Taken together, our data should provide valuable information for future discovery and validation studies of urine protein biomarkers for various diseases.

Keywords: LC-MS/MS, urine proteome, proteomics, urinary biomarkers, prostate cancer, kidney disease

INTRODUCTION

The production and elimination of urine is essential for the removal of waste products generated by cellular metabolism and other processes. Kidneys use special structures, particularly glomeruli, to filter blood (1, 2). Important substances such as water, salts, glucose, other nutrients, and most proteins are reabsorbed by the kidneys. Only select proteins are removed for excretion in urine.

Therefore, urine protein excretion in healthy adults is usually limited to <150 mg/day (3). Urine protein excretion beyond this value is defined as proteinuria (4), which is often a sign of kidney damage. Proteins in urine can originate from the kidney, bladder, prostate gland, ureter, urethra, or even from distant organs, and tissues. Since urine can be collected in large quantities using non-invasive procedures, urine proteins are particularly suitable for use as biomarkers to diagnose and monitor dysfunction involving these organs. Some urine protein biomarkers are critical for diagnosing and monitoring diseases such as prostate cancer (5, 6) and kidney failure (7–9). To facilitate the discovery of novel urine protein biomarkers, it is necessary to generate a comprehensive urine protein database from samples collected from patients with various disease conditions and healthy patients.

Mass spectrometry (MS)-based proteomics provides a powerful analytical tool for large-scale identification of proteins in urine. There have been many urine proteome profiling studies using different separation approaches coupled with MS. For instance, Adachi et al. (10) and Kentsis et al. (11) employed combinations of SDS-PAGE, ultra-centrifugation, and reverse phase liquid chromatography (LC) for fractionation and identified a total of 1,543 and 2,362 proteins, respectively, from urine samples of healthy individuals *via* LC-MS/MS analysis. Using SDS-PAGE and lectin enrichment followed by LC-MS/MS, Marimuthu et al. (12) identified 1,823 proteins from healthy human urine. Gel-free methods have also been used for urine proteome profiling. For example, Li et al. (13) applied a multidimensional LC-MS/MS method and identified 1,310 urine proteins. Expanded coverage of 3,000–6,000 proteins from the human urine proteome have been recently reported by applying more complex ligand library bead-binding equalization techniques or multi-dimensional gel electrophoresis coupled with multi-dimensional LC-MS/MS approaches (14, 15). However, one of the limitations of these global urine proteome profiling studies was the focus mainly on healthy individuals such that many disease-associated proteins could be missed from these studies. Therefore, it would be valuable to have a comprehensive urine proteome database derived from both healthy and disease conditions as a reference resource for guiding urine protein biomarker discovery.

In our previous studies, we have performed comparative studies of urine of renal patients and healthy individuals with the purpose of identifying potential urinary protein biomarkers for acute renal transplant rejection (16, 17). In order to generate a urine proteome database originating from multiple disease conditions as a reference resource, we combined datasets from urine samples from patients suffering from prostate cancer, renal transplant, and non-specific proteinuria, as well as healthy individuals using a commonly applied 2D-LC-MS/MS

workflow. Urine proteins in each group of samples were digested into peptides which were pre-fractionated by either strong cation exchange or high-pH reversed-phase LC. Peptides in each fraction were analyzed by LC-MS/MS, resulting in the identification of a total of ∼28,000 unique peptides across ∼2,200 urinary proteins. The final database was annotated with observation counts at both protein and peptide levels from each biological condition as well as the annotation of presence or absence in five recent urine proteome profiling studies. Approximately 400 proteins were only observed in the current study, possibly suggesting the observation of disease-associated proteins. Since the database was generated from several disease conditions and annotated against other urine proteome databases from healthy individuals, our database could serve as a global reference for guiding future biomarker discovery studies using urine as the source sample.

MATERIALS AND METHODS

Urine Collection and Processing

A total of 45 urine samples from 10 renal transplant patients with proven acute rejection (AR), 10 renal transplant patients with stable graft (STA), 10 non-specific proteinuria patients (NS), 10 healthy individuals (HI), and 5 prostate cancer (PC) patients were utilized for global urine proteome profiling. The patient demographics of patients with renal conditions including healthy controls were the same as described previously with an age range of 3–21 (16). The PC urine samples were from pre-operation patients with an age range of 60–75. This research was approved by the Institutional Review boards at Stanford University, University of California San Francisco, University of Washington, and Pacific Northwest National Laboratory in accordance with federal regulations. ∼50 mL urine samples were collected from each patient in sterile containers. Samples were centrifuged at 2,000 × g for 20 min at room temperature within 1 h of collection. The supernatant was collected and stored at −80°C for further analysis.

Proteins in the urine supernatant were concentrated with 10-kDa Amicon Ultra-15 centrifugal filter units (Millipore). The final protein concentration was measured by bicinchoninic (BCA) assay (Pierce). After concentration, 45 μg of urine proteins were pulled from each sample and combined according to their clinical categories, namely, PC, AR, STA, NS, and HI. The pooled protein samples were denatured by 8 M urea, reduced by 10 mM dithiothreitol (DTT), alkylated by with 40 mM iodoacetamide, and digested by trypsin as previously described (16). The final peptide concentrations were measured using the BCA assay.

Peptides from pooled AR, pooled STA, pooled NS, and pooled HI samples were fractionated by strong cation exchange (SCX) chromatography as previously described (16), and peptides from pooled PC samples were fractionated by high-pH reversed-phase separation and concatenated into 24 fraction as previously described (18).

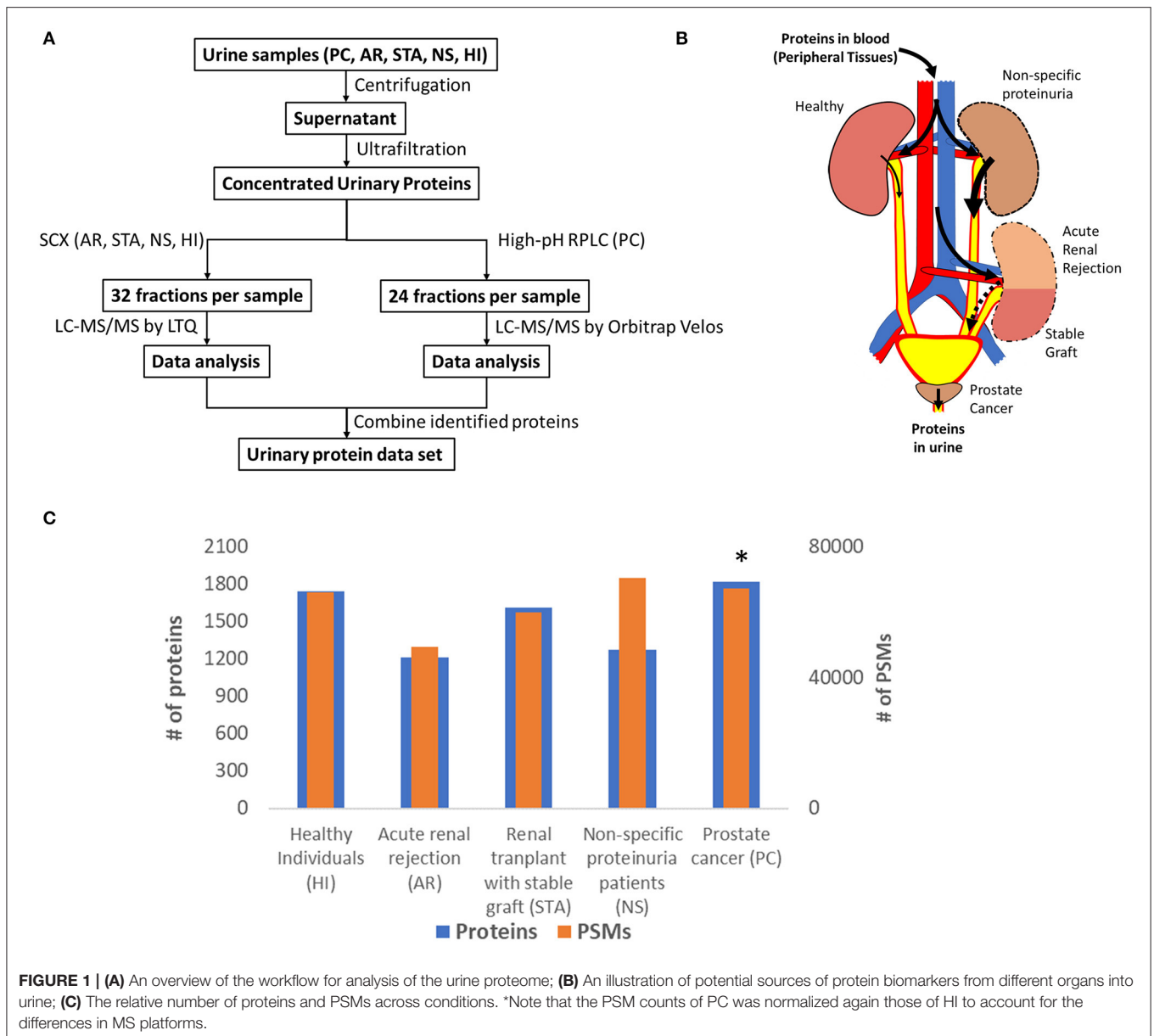**Abbreviations:** AR, acute rejection; CD, Cluster of Differentiation; DAVID, Database for Annotation, Visualization and Integrated Discovery; GO, Gene Ontology; HI, healthy individuals; IPA, Ingenuity Pathway Analysis; LC, Liquid Chromatography; MS, Mass Spectrometry; MSGF, MS Generating-Function; NS, non-specific proteinuria; STA, stable graft; PC, prostate cancer; PSM, Peptide-Spectrum Match.

LC-MS/MS Analyses

The peptide fractions were analyzed by LC-MS/MS. Specifically, the peptide fractions from SCX were analyzed by LTQ linear ion trap mass spectrometer (Thermo Fisher Scientific) coupled with a customized LC system as previously described (16). The peptide fractions from high pH reversed-phase LC fractionation were analyzed an LTQ Orbitrap Velos mass spectrometer (Thermo Fisher Scientific) coupled with a similar customized LC system. LC columns were prepared in-house by slurry packing 3-μm Jupiter C18 (Phenomenex, Torrence, CA) into 35-cm × 75 μm i.d fused silica (Polymicro Technologies Inc., Phoenix, AZ). A 100-min LC gradient with a 300 nL/min flowrate was applied for separations. The resolution of the MS scan was 120,000 with Top-20 data-dependent MS/MS acquisitions on CID mode.

Proteomics Data Analysis

All MS/MS spectra were searched against the UniProtKB/Swiss-Prot protein knowledgebase release 2013_09 using MSGF+ (Release 2019.07.03). The search parameters were as follows: (1) fixed modification, carbamidomethyl of C; (2) variable modification, oxidation of M; (3) allowing two missed cleavages; (4) parent ion mass tolerance: 1.0 Da for LTQ data and 20 ppm for Orbitrap Velos data; (5) fragment ion mass tolerance, 1.0 Da. MS Generating-Function (MSGF) scores were generated for all identified spectra by computing rigorous p-values (MSGF SpecEvalue) (19). The FDRs for final peptide and protein identifications were controlled to be <0.1 and <0.5%, respectively. Each identified peptide was assigned to the "first hit" protein from database searching output to avoid redundant

FIGURE 1 | **(A)** An overview of the workflow for analysis of the urine proteome; **(B)** An illustration of potential sources of protein biomarkers from different organs into urine; **(C)** The relative number of proteins and PSMs across conditions. *Note that the PSM counts of PC was normalized again those of HI to account for the differences in MS platforms.

A **Urine Proteomics**

	1D LC-MS	2D LC-MS (Includes MDC)	SDS-PAGE (Incudes SDS-PAGE + Lectin Enrichment)	GELFrEE and LP-IEF (Followed by 2D LC-MS)	Micro-vesicle and Exosome Enrichment	CPLL Bead Enrichment (including unbound)
Current Study		2193				
Adachi et al., 2006			1543			
Li et al., 2010		1310				
Marimuthu el al., 2011			1823			
Nagaraj et al., 2011	808					
Zerefos et al., 2012			588			
Santucci et al., 2014		1176			1615	1794
Zhao et al., 2017	808	3162		6085		

B **Urinary Proteomics in Disease**

	Acute Renal Rejection	Stable Graft	Chronic allograft nephropathy	BK virus nephritis	Acute Kidney Injury in preterm infants	Bladder cancer
Sigdel et al., 2016	958 proteins quantified; a panel of 35 peptide signatures for segregating patient groups.					
Chen et al., 2010						638 proteins; 55 candidates
Jung et al., 2020					1,810 proteins; 5 validated	

FIGURE 2 | (A) Comparison of methods used for urine proteome analysis and the number of proteins detected using the various methods **(B)** Highlights of several recent biomarker discovery and verification studies in several disease conditions.

assignment of peptides to multiple proteins (or protein group).

Gene Ontology (GO) annotation for cellular component and biological process of the identified urine proteins was performed by using the Database for Annotation, Visualization and Integrated Discovery (DAVID 6.8) bioinformatics resource (20, 21). Biomarkers were screened from the identified proteins using the Ingenuity Pathway Analysis biomarker filter module.

RESULTS AND DISCUSSION

Global Profiling of the Urine Proteome

Many urinary proteomics studies have been completed over many years with varying levels of protein coverage. Among the global deep proteomics studies with >2,000 unique protein IDs, most have reported on samples from healthy subjects. Our purpose here was to create a relatively comprehensive urine proteome database with urinary proteins that would be detectable in both diseased and healthy conditions. Importantly, we aimed to demonstrate what proteins could be detected using commonly applied standard LC-MS/MS techniques. The concept was to combine datasets from samples related to renal or other conditions relevant to the urinary tract from several independent global urine proteome profiling efforts from our laboratory as summarized in **Figure 1**, including an illustration of potential sources of urinary protein biomarkers from different disease

conditions. The first profiling efforts involved pooled urine samples from a renal rejection study with four different clinical conditions (AR, STA, NS, HI); where following protein digestion, peptides were fractionated into 32 fractions per sample by strong cation exchange chromatography (SCX) and analyzed by LC-MS/MS on a LTQ instrument. The second study involved pooled samples from prostate cancer patients; where peptides were fractionated by high-pH reversed-phase LC into 24 fractions and analyzed by LC-MS/MS on an Orbitrap Velos instrument. The combined dataset includes a total of ∼150 LC-MS/MS analyses to generate the final urine proteome database. Following database searching with the MSGF+ algorithm, we identified a total of ∼28,000 unique peptides (**Supplementary Table 1**) and ∼2,200 unique proteins (**Supplementary Table 2**) with FDR <0.5% at the protein level based on decoy database search and using stringent filtering criteria.
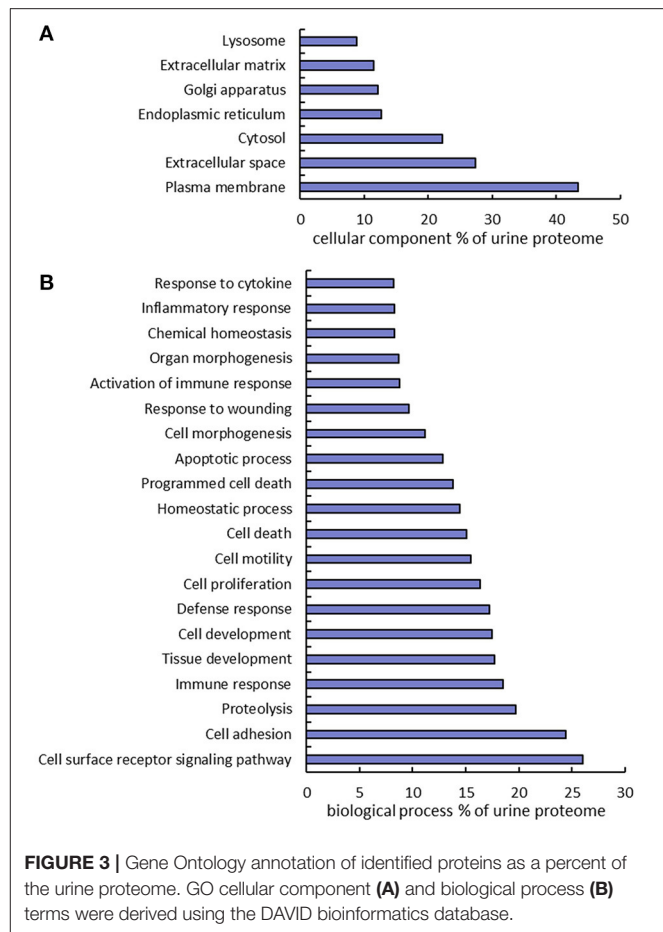
Figure 1C shows the number of total proteins and total PSMs from different conditions. We note that there are substantial differences in the age ranges of patient cohorts, fractionation methods, and instrumentation platforms being used for the two independent profiling studies. A global normalization of the PSM counts for PC against those from HI to account for the instrumentation platform differences between the two studies as outlined in **Figure 1A**. Despite these differences, the proteome coverage in terms of the number of proteins identified from

each condition was still relatively comparable. An interesting observation is the PSM counts for urine albumin, where 27,474 and 18,116 PSMs for serum albumin in the NS and AR conditions, respectively, but only 11,974 PSMs for albumin in the HI condition. Several other highly abundant serum proteins such as serotransferrin, retinol-binding protein 4, protein AMBP, and alpha-1 antitrypsin were also observed with high PSMs in disease conditions compared to HI. These observations are consistent with proteinuria as a known biomarker of kidney disease (22).

Comparison to Previous Urine Profiling or Biomarker Studies

Next, we compared our current set of urine proteins with results from several recent published global urine proteome profiling studies, which were mostly from healthy donors. Using a combination of 15 prior studies (including most listed in **Figure 2**), Farrah et al. (23) compiled a comprehensive list for the PeptideAtlas project urine section (UrinePA) and found 2,491 unique proteins confidently detected in those studies. The current most accessible proteome profiling method is the 2D-LC-MS/MS workflow, which consists of pre-fractionation with either high pH reverse phase LC or strong cation exchange (SCX) chromatography. Several groups have used this method to obtain relatively high urine proteome coverage (13, 24). Other specialized techniques have also been applied to expand the urine proteome coverage, including micro-vesicle and exosome enrichment prior to MS sample preparation (14), 2D SDS-PAGE separation and spot excision followed by LC-MS (10, 12, 25), multi-dimensional gel electrophoresis followed by multi-dimensional LC-MS (15), and using combinatorial peptide ligand library (CPLL) binding beads to "equalize" protein abundances (24). A urine proteome database with deeper coverage (∼6,000 proteins) was recently reported using several specialized methods including gel-free electrophoresis and isoelectric focusing (15). However, it is still unclear whether disease associated proteins would be missed from these efforts focusing on samples mainly from healthy individuals. Our urinary proteome database was generated using easily accessible techniques and incorporated multiple disease conditions in order to provide a useful baseline reference resource for guiding urine biomarker discovery. The comparison of the urine proteome coverage from different studies using various techniques is illustrated in **Figure 2**.
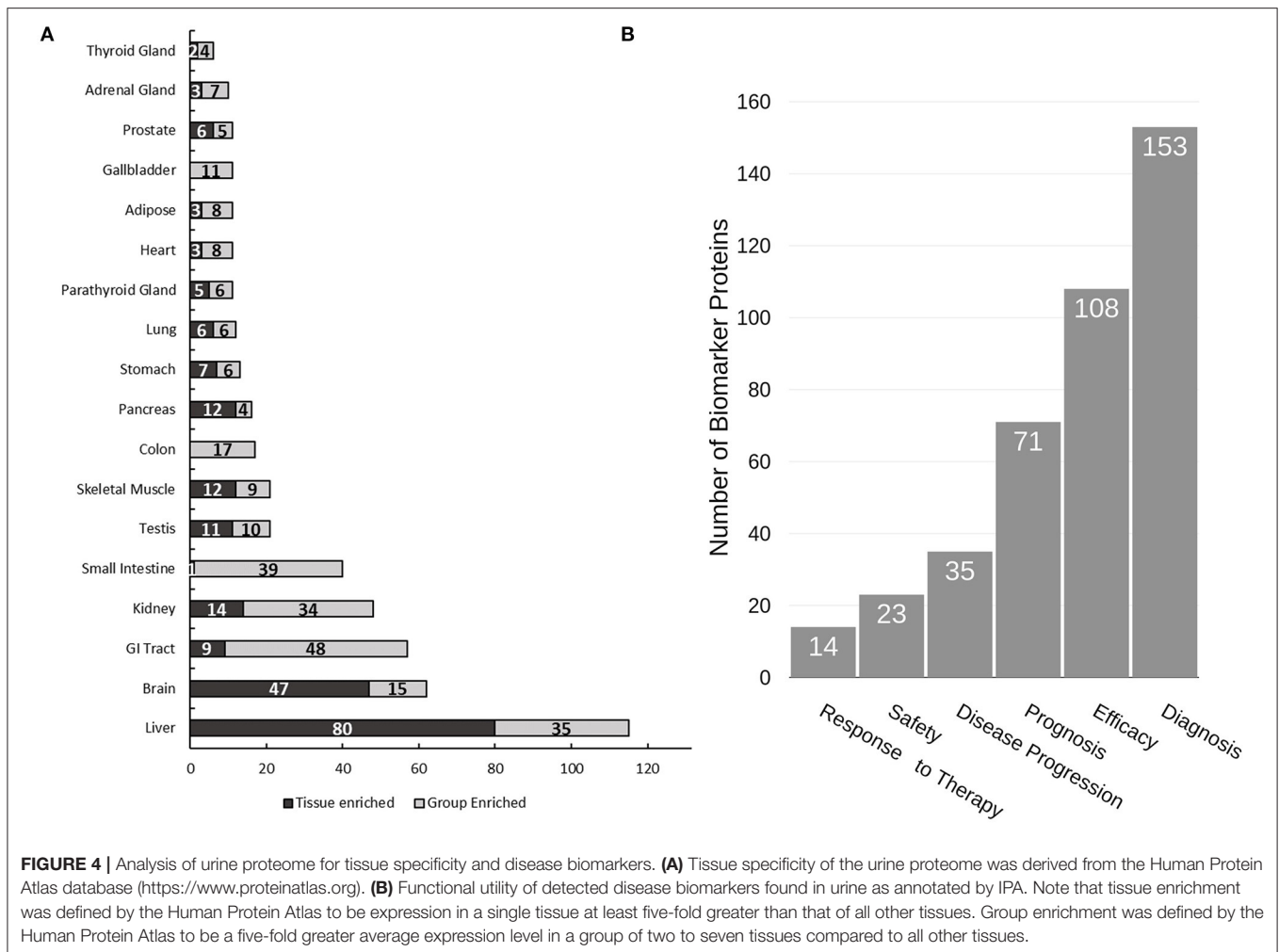
One of the primary interests of urine proteomics is the discovery of novel biomarkers of various diseases (26). While most urine biomarker studies were based on validating target panel of protein or peptide signatures in different disease conditions, there were several studies applying global discovery approach followed by targeted verification of candidate markers. We also highlighted several biomarker discovery and verification efforts in **Figure 2**, including the study by Sigdel et al. (17) for renal transplantation conditions, the bladder cancer study by Chen et al. (27), and the acute kidney injury study in preterm infants by Jung et al. (28). Importantly, all of the reported candidate biomarkers from these studies were identified in our current dataset (**Supplementary Table 2**), again

FIGURE 3 | Gene Ontology annotation of identified proteins as a percent of the urine proteome. GO cellular component **(A)** and biological process **(B)** terms were derived using the DAVID bioinformatics database.

suggesting the relative comprehensiveness of the current urine proteome database.

Gene Ontology Analysis of the Urine Proteome

The urine proteins identified in this work were classified based on Gene Ontology (GO) cellular component and biological process annotation terms using the Database for Annotation, Visualization, and Integrated Discovery (DAVID 6.8) bioinformatics resources. Of note, urine proteins were annotated as extracellular space and plasma membrane proteins at 27.4 and 43.3%, respectively (**Figure 3A**). Two earlier studies also found that plasma membrane proteins were enriched in urine samples, where ∼20% (10) and 31% (12) of urine proteins identified from healthy human urine were plasma membrane proteins. In terms of biological processes, 24.4% proteins were found to function in cell adhesion, which is consistent with the previously mentioned enrichment of extracellular space and plasma membrane proteins. It is well-known that the key protein components involved in various cell-adhesion structures (adherens junction, focal adhesion, desmosome, tight junction, and so on) are localized in the plasma membrane and extracellular matrix. Many examples of these proteins (cadherins, desmocollins, desmogleins, integrins, collagens, and

FIGURE 4 | Analysis of urine proteome for tissue specificity and disease biomarkers. **(A)** Tissue specificity of the urine proteome was derived from the Human Protein Atlas database (https://www.proteinatlas.org). **(B)** Functional utility of detected disease biomarkers found in urine as annotated by IPA. Note that tissue enrichment was defined by the Human Protein Atlas to be expression in a single tissue at least five-fold greater than that of all other tissues. Group enrichment was defined by the Human Protein Atlas to be a five-fold greater average expression level in a group of two to seven tissues compared to all other tissues.

fibronectins) were identified in this work. It is not surprising that cell-adhesion related proteins are enriched in urine because they are exposed on cell surfaces which increases the likelihood of release into urine. Besides cell adhesion, there are several other major biological processes that are enriched, including proteolysis (19.7%), immune response (18.5%), cell proliferation (16.4%), and response to wounding (9.7%) (**Figure 3B**).

Cluster of Differentiation Antigens

An important feature of our dataset is that many CD antigens were identified from human urine. CD antigens often play critical roles in cell signaling and cell adhesion. They are commonly used as markers for immunophenotyping and for diagnosis, monitoring, and treatment of diseases. Out of all 394 known human CD proteins in the Uniprot database, 178 (45%) were identified from the human urine samples in this study (**Supplementary Table 4**). Since CD proteins are cell-surface proteins, it is not a surprise that many CD molecules are released from cells into body fluids such as blood and urine. Also worthy of note, two extensively used prostate cancer stem cell markers CD133 (29) and CD44 (30) were identified. These CD proteins could be useful as diagnostic markers for other diseases.

Candidate Biomarkers Identified From the Urine Proteome

Due to the non-invasive nature of urine collection, urine proteins are ideal biomarkers for diagnosis of renal diseases and other diseases related to the urinary tract including prostate and bladder cancer. Indeed, several promising biomarkers have already been reported (31–36). Herein we performed a "biomarker filter" analysis of the identified urine proteins using the Ingenuity Pathway Analysis (IPA) software. The "biomarker filter" is an IPA module which allows identification of biomarker candidates based on prior curated literature data. Three hundred and forty-nine proteins were identified as candidate biomarkers (**Supplementary Table 5**) following biomarker filtering analysis. These biomarkers were categorized based on their applications. 153, 108, 71, and 35 proteins were relevant to diagnosis, efficacy, prognosis, and disease progression, respectively (**Figure 4B**). We also analyzed the tissue specificity of 349 tissue-enriched proteins based on human protein atlas database (**Figure 4A**). Of these, 230 were expressed in kidneys, 200 were expressed in the prostate gland, and 181 were expressed in bladder, among many other represented tissues (**Supplementary Table 6**). However, most of these proteins were shared by multiple tissues or organs. In the

TABLE 1 | Selected potential disease-associated proteins only detected in the current study.

Protein ID	Gene symbol	Protein name	PSMs in different conditions				
			HI	AR	STA	NS	PC
H31_HUMAN	HIST1H3A	Histone H3.1					**49**
H2B1B_HUMAN	HIST1H2BB	Histone H2B type 1-B	2		2	1	**24**
H33_HUMAN	H3F3A	Histone H3.3					**13**
H2A1A_HUMAN	HIST1H2AA	Histone H2A type 1-A	1	1		1	**10**
H31T_HUMAN	HIST3H3	Histone H3.1t					**10**
INS_HUMAN	INS	Insulin	1		2	1	**119**
SPB3_HUMAN	SERPINB3	Serpin B3	28	11	**104**	26	
RENR_HUMAN	ATP6AP2	Renin receptor	2		2	1	**120**
POSTN_HUMAN	POSTN	Periostin	2	1		1	**12**

Bold values indicate significantly elevated level compared to all other conditions.

study by Kentsis et al. (11) they presented a list of 403 biomarkers and their associations to 27 common and 500 rare diseases. We observed 187 of the biomarkers they detected in our data. Of the 349 biomarkers we detected as potential biomarkers based on IPA database, they only detected 70 of them, most likely due to the use of urine from healthy conditions.

Potential Disease-Associated Proteins

Our dataset could prove valuable for identifying potential disease-associated proteins by combining it with other available urine proteome databases. We were able to compare our data to six existing datasets where comparable profiling approaches were applied (10–15); each differing primarily by the extensiveness and complexity of pre-fractionation methods utilized. Indeed, ~400 unique proteins were only observed our dataset (**Supplementary Table 3**). We note that while some of the unique proteins may be due to protein accession ID discrepancies between different studies, the data suggest that many of these proteins are associated with the analyzed disease conditions.

We note that the primary purpose of dataset is to provide a qualitative database of detectable urine proteins in health and diseases. Due to the differences in age ranges of patient cohorts, sample collection and storage times, methods, and instrumentation platform being used, these data were not designed to be used for quantitative comparative analysis between disease conditions. Given these limitations, the differences in PSM counts for a given protein between disease conditions need to be interpreted with caution. In particular, due to the variable nature of proteins in urine, quantitative measurements must be well-controlled with good experimental design since it was reported that inter-patient variability may exceed 47% and intra-patient variability may exceed 45% if not carefully controlled (37). These limitations of the current study also explain sometime drastic difference being observed across disease conditions for some proteins. One example is the consistent detection of myoglobin in healthy individuals and the renal conditions, but the low PSM count in PC condition. Possible reasons include the independent patient cohorts with large differences in age ranges as well as the difference in sample collection and storage.

Nevertheless, we did observe several interesting proteins as potential disease-associated proteins (**Table 1**). For example, a number of histone proteins were predominantly detected in prostate cancer samples, supporting the recent report that extracellular or circulatory histones may reflect tissue injury or cell death (38). Indeed, due to the proximity of the prostate lumen to the urethra, PC can result in a variety of proteins shed into urine originating from injured epithelial cells, basal cells, cancer cells, blood, and immune cells (39). This is connected to the increase of nucleic acids (40) and nuclear proteins such as histones. Moreover, Serpin B3 was detected at highest level in the renal transplant patients with stable graft (STA), consistent with the report that Serpin B3 was an important healing biomarker. On the other hand, detection of high level of insulin in the prostate cancer samples is also worthy of noting since the older individuals are much more likely to have higher levels of circulating insulin due to insulin resistance (41). Renin receptor and Periostin were two markers detected primarily in prostate urine samples and both proteins have been reported as viable cancer biomarkers (42, 43).

Concluding Remarks

We have generated a comprehensive urine proteome database through LC-MS/MS profiling of urine samples from prostate cancer patients, renal transplant patients with acute rejection or stable graft, non-specific proteinuria patients, and healthy individuals. The overall analyses resulted in the identification of ~28,000 unique peptides and ~2,200 unique proteins. Over 40% of the identified proteins were annotated as plasma membrane proteins and over three-fourths were extracellular proteins. IPA biomarker filter analysis revealed that 349 proteins are potential candidate biomarkers relevant to diagnosis, efficacy, prognosis, and disease progression. Moreover, 45% (178) of all known CD proteins were identified in these human urine samples. Presumably due to the inclusion of several disease conditions, our study identified ~400 proteins that were not detected in previous profiling studies using similar approaches. Among them, several interesting disease-associated protein markers were identified. While the database will be useful resource for determining the detectability of proteins of interest, the dataset was not

intended for making quantitative comparisons between evaluated conditions due to the limitations of the study design. We also provide a list of all detected peptides, which is intended to be used to guide surrogate peptide selection in specific targeted mass spectrometry assays. Together, this comprehensive urine proteome dataset could serve as a valuable reference resource for future biomarker discovery efforts using urine as the source sample.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: Massive.ucsd.edu with Accession: MSV000086484.

ETHICS STATEMENT

This research was approved by the Institutional Review boards at Stanford University, University of California San Francisco, University of Washington, and Pacific Northwest National Laboratory in accordance with federal regulations.

AUTHOR CONTRIBUTIONS

AL, TKS, MS, and W-JQ designed the experiments. CN and TS performed the experiments. AS, JH, AF, and YY analyzed the data. AS, JH, AF, and W-JQ wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2021.548212/full#supplementary-material

REFERENCES

1. Haraldsson B, Sorensson J. Why do we not all have proteinuria? An update of our current understanding of the glomerular barrier. *News Physiol Sci.* (2004) 19:7–10. doi: 10.1152/nips.01461.2003

2. Thongboonkerd V, Malasit P. Renal and urinary proteomics: current applications and challenges. *Proteomics.* (2005) 5:1033–42. doi: 10.1002/pmic.200401012

3. Shaw AB, Risdon P, Lewis-Jackson JD. Protein creatinine index and Albustix in assessment of proteinuria. *Br Med J (Clin Res Ed).* (1983) 287:929–32. doi: 10.1136/bmj.287.6397.929

4. Bergstein JM. A practical approach to proteinuria. *Pediatr Nephrol.* (1999) 13:697–700. doi: 10.1007/s004670050684

5. Ploussard G, de la Taille A. Urine biomarkers in prostate cancer. *Nat Rev Urol.* (2010) 7:101–9. doi: 10.1038/nrurol.2009.261

6. Laxman B, Morris DS, Yu J, Siddiqui J, Cao J, Mehra R, et al. A first-generation multiplex biomarker analysis of urine for the early detection of prostate cancer. *Cancer Res.* (2008) 68:645–9. doi: 10.1158/0008-5472.CAN-07-3224

7. Han WK, Waikar SS, Johnson A, Betensky RA, Dent CL, Devarajan P, et al. Urinary biomarkers in the early diagnosis of acute kidney injury. *Kidney Int.* (2008) 73:863–9. doi: 10.1038/sj.ki.5002715

8. Nickolas TL, Barasch J, Devarajan P. Biomarkers in acute and chronic kidney disease. *Curr Opin Nephrol Hypertens.* (2008) 17:127–32. doi: 10.1097/MNH.0b013e3282f4e525

9. Devarajan P. Emerging biomarkers of acute kidney injury. *Contrib Nephrol.* (2007) 156:203–12. doi: 10.1159/000102085

10. Adachi J, Kumar C, Zhang Y, Olsen JV, Mann M. The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins. *Genome Biol.* (2006) 7:R80. doi: 10.1186/gb-2006-7-9-r80

11. Kentsis A, Monigatti F, Dorff K, Campagne F, Bachur R, Steen H. Urine proteomics for profiling of human disease using high accuracy mass spectrometry. *Proteomics Clin Appl.* (2009) 3:1052–61. doi: 10.1002/prca.200900008

12. Marimuthu A, O'Meally RN, Chaerkady R, Subbannayya Y, Nanjappa V, Kumar P, et al. A comprehensive map of the human urinary proteome. *J Proteome Res.* (2011) 10:2734–43. doi: 10.1021/pr2003038

13. Li QR, Fan KX, Li RX, Dai J, Wu CC, Zhao SL, et al. A comprehensive and non-prefractionation on the protein level approach for the human urinary proteome: touching phosphorylation in urine. *Rapid Commun Mass Spectrom.* (2010) 24:823–32. doi: 10.1002/rcm.4441

14. Santucci L, Candiano G, Petretto A, Bruschi M, Lavarello C, Inglese E, et al. From hundreds to thousands: widening the normal human Urinome. *Data Brief.* (2014) 1:25–8. doi: 10.1016/j.dib.2014.08.006

15. Zhao M, Li M, Yang Y, Guo Z, Sun Y, Shao C, et al. A comprehensive analysis and annotation of human normal urinary proteome. *Sci Rep.* (2017) 7:3024. doi: 10.1038/s41598-017-03226-6

16. Sigdel TK, Kaushal A, Gritsenko M, Norbeck AD, Qian WJ, Xiao W, et al. Shotgun proteomics identifies proteins specific for acute renal transplant rejection. *Proteomics Clin Appl.* (2010) 4:32–47. doi: 10.1002/prca.200900124

17. Sigdel TK, Gao Y, He J, Wang A, Nicora CD, Fillmore TL, et al. Mining the human urine proteome for monitoring renal transplant injury. *Kidney Int.* (2016) 89:1244–52. doi: 10.1016/j.kint.2015.12.049

18. Wang Y, Yang F, Gritsenko MA, Clauss T, Liu T, Shen Y, et al. Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells. *Proteomics.* (2011) 11:2019–26. doi: 10.1002/pmic.201000722

19. Kim S, Gupta N, Pevzner PA. Spectral probabilities and generating functions of tandem mass spectra: a strike against decoy databases. *J Proteome Res.* (2008) 7:3354–63. doi: 10.1021/pr8001244

20. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* (2009) 4:44–57. doi: 10.1038/nprot.2008.211

21. Huang da W, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, et al. DAVID bioinformatics resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res.* (2007) 35:W169–75. doi: 10.1093/nar/gkm415

22. Zandi-Nejad K, Eddy AA, Glassock RJ, Brenner BM. Why is proteinuria an ominous biomarker of progressive kidney disease? *Kidney Int Suppl.* (2004) 92:S76–89. doi: 10.1111/j.1523-1755.2004.09220.x

23. Farrah T, Deutsch EW, Omenn GS, Sun Z, Watts JD, Yamamoto T, et al. State of the human proteome in 2013 as viewed through PeptideAtlas: comparing the kidney, urine, and plasma proteomes for the biology- and disease-driven human proteome project. *J Proteome Res.* (2014) 13:60–75. doi: 10.1021/pr4010037

24. Santucci L, Candiano G, Petretto A, Bruschi M, Lavarello C, Inglese E, et al. From hundreds to thousands: widening the normal human urinome (1). *J Proteomics*. (2015) 112:53–62. doi: 10.1016/j.jprot.2014.07.021

25. Zerefos PG, Aivaliotis M, Baumann M, Vlahou A. Analysis of the urine proteome *via* a combination of multi-dimensional approaches. *Proteomics*. (2012) 12:391–400. doi: 10.1002/pmic.201100212

26. Thomas S, Hao L, Ricke WA, Li L. Biomarker discovery in mass spectrometry-based urinary proteomics. *Proteomics Clin Appl*. (2016) 10:358–70. doi: 10.1002/prca.201500102

27. Chen YT, Chen CL, Chen HW, Chung T, Wu CC, Chen CD, et al. Discovery of novel bladder cancer biomarkers by comparative urine proteomics using iTRAQ technology. *J Proteome Res*. (2010) 9:5803–15. doi: 10.1021/pr100576x

28. Jung YH, Han D, Shin SH, Kim EK, Kim HS. Proteomic identification of early urinary-biomarkers of acute kidney injury in preterm infants. *Sci Rep*. (2020) 10:4057. doi: 10.1038/s41598-020-60890-x

29. Richardson GD, Robson CN, Lang SH, Neal DE, Maitland NJ, Collins AT. CD133, a novel marker for human prostatic epithelial stem cells. *J Cell Sci*. (2004) 117:3539–45. doi: 10.1242/jcs.01222

30. Patrawala L, Calhoun T, Schneider-Broussard R, Li H, Bhatia B, Tang S, et al. Highly purified CD44+ prostate cancer cells from xenograft human tumors are enriched in tumorigenic and metastatic progenitor cells. *Oncogene*. (2006) 25:1696–708. doi: 10.1038/sj.onc.1209327

31. Mishra J, Dent C, Tarabishi R, Mitsnefes MM, Ma Q, Kelly C, et al. Neutrophil gelatinase-associated lipocalin (NGAL) as a biomarker for acute renal injury after cardiac surgery. *Lancet*. (2005) 365:1231–8. doi: 10.1016/S0140-6736(05)74811-X

32. Han WK, Bailly V, Abichandani R, Thadhani R, Bonventre JV. Kidney Injury Molecule-1 (KIM-1): a novel biomarker for human renal proximal tubule injury. *Kidney Int*. (2002) 62:237–44. doi: 10.1046/j.1523-1755.2002.00433.x

33. Parikh CR, Jani A, Melnikov VY, Faubel S, Edelstein CL. Urinary interleukin-18 is a marker of human acute tubular necrosis. *Am J Kidney Dis*. (2004) 43:405–14. doi: 10.1053/j.ajkd.2003.10.040

34. Bolduc S, Lacombe L, Naud A, Gregoire M, Fradet Y, Tremblay RR. Urinary PSA: a potential useful marker when serum PSA is between 2.5 ng/mL and 10 ng/mL. *Can Urol Assoc J*. (2007) 1:377–81. doi: 10.5489/cuaj.444

35. Irani J, Salomon L, Soulie M, Zlotta A, de la Taille A, Dore B, et al. Urinary/serum prostate-specific antigen ratio: comparison with free/total serum prostate-specific antigen ratio in improving prostate cancer detection. *Urology*. (2005) 65:533–7. doi: 10.1016/j.urology.2004.10.003

36. Schostak M, Schwall GP, Poznanovic S, Groebe K, Muller M, Messinger D, et al. Annexin A3 in urine: a highly specific noninvasive marker for prostate cancer early detection. *J Urol*. (2009) 181:343–53. doi: 10.1016/j.juro.2008.08.119

37. Nagaraj N, Mann M. Quantitative analysis of the intra- and inter-individual variability of the normal urinary proteome. *J Proteome Res*. (2011) 10:637–45. doi: 10.1021/pr100835s

38. Silk E, Zhao H, Weng H, Ma D. The role of extracellular histone in organ injury. *Cell Death Dis*. (2017) 8:e2812. doi: 10.1038/cddis.2017.52

39. Eskra JN, Rabizadeh D, Pavlovich CP, Catalona WJ, Luo J. Approaches to urinary detection of prostate cancer. *Prostate Cancer Prostatic Dis*. (2019) 22:362–81. doi: 10.1038/s41391-019-0127-4

40. Casadio V, Calistri D, Salvi S, Gunelli R, Carretta E, Amadori D, et al. Urine cell-free DNA integrity as a marker for early prostate cancer diagnosis: a pilot study. *Biomed Res Int*. (2013) 2013:270457. doi: 10.1155/2013/270457

41. Chang AM, Halter JB. Aging and insulin secretion. *Am J Physiol Endocrinol Metab*. (2003) 284:E7–12. doi: 10.1152/ajpendo.00366.2002

42. Tischler V, Fritzsche FR, Wild PJ, Stephan C, Seifert HH, Riener MO, et al. Periostin is up-regulated in high grade and high stage prostate cancer. *BMC Cancer*. (2010) 10:273. doi: 10.1186/1471-2407-10-273

43. Wang J, Nishiyama A, Matsuyama M, Wang Z, Yuan Y. The (pro)renin receptor: a novel biomarker and potential therapeutic target for various cancers. *Cell Commun Signal*. (2020) 18:39. doi: 10.1186/s12964-020-0531-3