

## Applications of urinary proteomics in biomarker discovery

SHAO Chen<sup>\*</sup>, WANG Yan & GAO YouHe

*National Key Laboratory of Medical Molecular Biology, Department of Physiology and Pathophysiology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences/School of Basic Medicine, Peking Union Medical College, Beijing 100005, China*

Received May 14, 2010; accepted August 8, 2010; published online March 31, 2011

Urine is an important source of biomarkers. This article reviews current advances, major challenges, and future prospects in the field of urinary proteomics. Because the practical clinical problem is to distinguish diseases with similar symptoms, merely comparing samples from patients of a particular disease to those of healthy individuals is inadequate for finding biomarkers with sufficient diagnostic power. In addition, the variation of expression levels of urinary proteins among healthy individuals and individuals under different physiological conditions adds to the difficulty in identifying biomarkers. We propose that establishing the natural variation in urinary protein expression among a healthy population can serve as a reference to help identify protein abundance changes that are caused by disease, not by individual variations or physiological changes. We also discuss that comparing protein expression levels between urine and plasma may reveal the physiological function of the kidney and that may facilitate biomarker discovery. Finally, we propose that establishing a data-sharing platform for data collection and integrating results from all urinary biomarker studies will help promote the development of urinary proteomics.

**urine, proteomics, biomarker**

**Citation:** Shao C, Wang Y, Gao Y H. Applications of urinary proteomics in biomarker discovery. *Sci China Life Sci*, 2011, 54: 409–417, doi: 10.1007/s11427-011-4162-1

Biofluid proteomics is becoming an important research field along with the development of proteomics technologies. Urine is an ideal source for clinical research as it can be obtained non-invasively in large quantities. Proteins in the urine are mainly composed of plasma proteins that pass through the glomerular filtration barrier as well as proteins secreted from the kidney and urinary tract. Studies on the diagnosis and therapy of kidney and other urological diseases have progressed slowly, primarily a result of disease diagnosis still relying on subtle histopathological changes in the majority of cases. Early diagnosis, prognosis and evaluation of therapy progression are still very difficult in a clinical setting. Urinary proteomics provides a new direction for disease diagnosis, treatment, monitoring and prognosis.

Considerable achievements have recently been made in

urinary proteomics. Still, the number of urinary proteomics studies published is still far fewer than those published for plasma proteomics. PubMed searches using “(urine OR urinary) AND (proteome OR proteomics)” as keywords returned 837 articles (May 2010), whereas using “(blood OR plasma OR serum) AND (proteome OR proteomics)” as keywords returned 5446 articles. However, the PubMed results should not lead one to conclude that urinary proteomics is less important. Compared with plasma, urine has unique advantages that make it a suitable source for both physiological research and disease biomarker discovery. First, urine can be collected continually and non-invasively. Second, changes in the urinary proteome directly reflect changes of the urinary system, and third, because the urinary proteome contains a number of plasma proteins, some changes of plasma proteome can also be found in urine. Urine therefore is not only a good source from which to study urological diseases, but it can also reflect the condi-

<sup>\*</sup>Corresponding author (email: scshaochen@126.com)

tions of the whole body. Moreover, some proteins/peptides may be excreted into urine immediately after being released into plasma, which may not be detected in plasma but in urine. Finally, the complexity of urinary proteome is lower than plasma proteome, making it easier to detect changes in low-abundance proteins.

The main drawback of using urine proteomics in clinical applications is that the urinary proteome changes greatly among different individuals, and among the same individual under different physiological conditions. In this review, we introduce the current status of urinary proteomics studies, including the profiling of the normal human urinary proteome, the determination of individual differences and the dynamic variation of normal human urinary proteome, and the applications of urinary proteomics to biomarker discovery. Finally, major challenges in studying urinary proteomics and potential future developments are discussed.

## 1 Profiling of normal human urinary proteome

As early as the 1990s, people started to identify proteins in urine. In 2004, Pieper *et al.* [1] separated nearly 1400 protein spots using two-dimensional gel electrophoresis (2-DE) and identified 150 distinct proteins from 420 spots by mass spectrometry. Sun *et al.* [2] employed liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) to profile normal human urinary proteome. They identified 226 proteins by three different separation methods (1D gel followed by 1DLC, 1DLC and 2DLC). Using high-resolution LTQ-FT and LTQ-Orbitrap mass spectrometers, Adachi *et al.* [3] identified 1534 proteins in normal human urine. Recently, Li *et al.* [4] developed two robust approaches, integrated multidimensional liquid chromatography (IMD-LC) and Yin-yang multidimensional liquid chromatography (Yin-yang MDLC) tandem mass spectrometry, to identify 1310 proteins. To date, over 2000 proteins have been identified in normal human urine.

It was noted that some proteins had greater experimental molecular weights than the theoretically possible from amino acid sequence alone, indicating the presence of post-translational modifications [5]. Wang *et al.* [6] identified 225 N-glycoproteins in Con-A enriched normal human urinary proteome, whereas Li *et al.* [4] identified 31 phosphoproteins.

Normal human urine contains large numbers of small vesicles with diameters less than 100 nm, named exosomes, which are secreted from renal epithelial cells. Exosomes are important because they contain a number of disease-related proteins. In 2009, Gonzales *et al.* [7] identified 1132 proteins in urinary exosomes, including 14 phosphoproteins.

## 2 Individual differences and dynamic variation of normal human urinary proteome

It is known that the urinary proteome differs between healthy individuals, particularly between men and women. In addition to the inter-individual differences, the urinary proteome from the same individual varies at different time points due to the effect of exercise, diet, lifestyle and other factors.

Khan and Packer [8] studied the urinary proteome variation of a healthy male volunteer. They examined urinary profiles at three time points in one day, and on different days over a six-week period using 2-DE. Results showed that the urinary proteome changed significantly over time. Urine collected in the morning contained more proteins than in the afternoon and evening. In addition, the inter-day proteome variation was observed to be greater than the variation at different time points in the same day.

Thongboonkerd *et al.* [9] compared urine samples from four men and four women by 2-DE. They also observed that urine samples collected in the morning had the greatest amount of proteins. The existence of inter-individual variation was also confirmed. Albumin and transferrin were reported to have the greatest coefficients of variation.

Similar conclusions were reached in urinary peptidomics studies conducted by Jürgens *et al.* [10] who identified 12 sex-specific peptides from 10 male and 10 female volunteers by liquid chromatography and matrix assisted laser desorption ionization time-of-flight mass spectrometry (LC-MALDI-TOF). In another study, Weissinger *et al.* [11] profiled urinary peptides of 57 healthy individuals and several patients with kidney diseases by capillary electrophoresis mass spectrometry (CE-MS). Approximately 900–2500 peaks could be detected in each sample. However, only 173 and 690 peaks were detected consistently in more than 90% and 50% of samples, respectively.

In 2009, Sun *et al.* [12] compared urine samples from three healthy male and three healthy female volunteers. Five kinds of urine samples (first morning void, second morning void, excessive water-drinking void, random void, and 24 h void) were collected from each volunteer to study inter-day, inter-individual, and inter-gender variations. Urinary proteins were divided into two groups: proteins expressed consistently in all samples with relatively stable abundances, and proteins whose expression levels varied. The authors proposed that any qualitative or quantitative expression changes of the stably expressed proteins might serve as biomarkers.

## 3 Applications of urinary proteomics in biomarker discovery

Biomarker discovery is a very active area in urinary pro-

teomics, which is moving out of the discovery phase and into an era of validation of candidate biomarkers by large-scale prospective analysis.

### 3.1 Cancers

#### 3.1.1 Bladder cancer

Bladder cancer is one of the most common cancers of the urinary system. Urine cytology, the most commonly used non-invasive test for bladder cancer, suffers from low sensitivity and is not effective for early diagnosis [13]. Several protein biomarkers have already been discovered, including nuclear matrix protein 22, cytokeratin 8 and 18 fragments (UBC) and bladder tumor antigen (BTA). These biomarkers are generally more sensitive than urine cytology, but have lower specificity. Therefore, urine cytology is still the gold standard for bladder cancer detection and cannot yet be substituted by biomarkers. The prevailing opinion is that a single biomarker will have limited diagnostic value, and that using a panel of biomarkers will achieve better diagnostic performance [14]. Proteomics technologies have been used to help researchers detect changes in expression levels of multiple proteins simultaneously.

Irmak *et al.* [15] used 2-DE to compare urine samples from bladder cancer patients at different stages, patients on follow-up and normal controls. In their study, orosomucoid and zinc- $\alpha$ 2-glycoprotein were identified as potential biomarkers. Saito *et al.* [16] identified various amounts of matrix metalloproteinase 2 and 9 (MMP-2 and 9), fibronectin, and their fragments as potential urinary biomarkers by gelatin-affinity purification and 2-DE separation. They also observed that the 2-DE patterns of these proteins correlated well with the levels of tumor invasiveness that was identified by histopathological examination. Tan *et al.* [17] employed LC-MS/MS to identify a novel candidate biomarker PLK2, while Tsui *et al.* [18] found that bikunin might serve as a potential biomarker. All of the above biomarker candidates were not only identified by proteomics approaches, but were verified by other methods (such as Western blotting) on independent sample sets. However, further investigations using larger sample sets are still needed before using these biomarkers in a clinical setting.

In addition to screening one or a few proteins as biomarkers, some studies focused on discovering signature proteome/peptidome patterns for diseases using CE-MS or surface enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF-MS). Mueller *et al.* [19] used SELDI-TOF-MS to search for a discriminating expression pattern from a training sample set of 30 bladder cancer patients and 30 healthy subjects. Urine samples were analyzed on weak cationic exchanger (CM10) and a strong anionic exchanger (SAX2) chips. Sensitivity of 80% and specificity of 90%–97% were achieved. The authors then applied the pattern to a test set of 42 urine samples (21 pa-

tients and 21 controls), sensitivities decreased 20%–30% and specificities decreased more than 30% on both chips. In another CE-MS study, Theodorescu *et al.* [20] used a panel of 22 peptides to classify patients with urothelial carcinoma (100% sensitivity and 100% specificity). The model also performed well when distinguishing patients with urothelial carcinoma from patients with other malignant and non-malignant genitourinary diseases. One peptide was identified as fibrinopeptide A, a known biomarker for ovarian and gastric cancer.

#### 3.1.2 Prostate cancer

Prostate cancer is a common cancer of the male reproductive system. Prostate specific antigen (PSA) is the most widely used prostate cancer marker yet does not have satisfactory detection power due to the lack of specificity. The expression level of PSA changes in various prostate diseases such as benign prostatic hyperplasia and prostatitis [21]. Efforts to discover specific biomarkers for prostate cancer have been undertaken using proteomics methods. Using 2-DE, Rehman *et al.* [22] compared the urinary proteome after prostatic massage of patients with prostate cancer with benign prostatic hyperplasia patients. With this approach, Calgranulin B/MRP-14 was identified as a potential biomarker for prostate cancer. Theodorescu *et al.* [23] used CE-MS to identify a panel of 9 urinary peptides that were differentially expressed between prostate cancer patients and controls. Sensitivity and specificity were 92% and 96%, respectively. M'Koma *et al.* [24] profiled the urinary proteome of 407 patients by MALDI-TOF and were able to classify patients with prostate cancer, benign prostatic hyperplasia and high-grade prostatic intraepithelial neoplasia, with a specificity of 71.2%–80.8% and sensitivity of 67.4%–81.0%.

#### 3.1.3 Renal cell carcinoma

Renal cell carcinoma (RCC) is the most common primary renal malignant neoplasm. No promising protein biomarkers have been used in clinical diagnosis until just recently. In 2004, Rogers *et al.* [25] employed a neural network approach to detect signature patterns from SELDI-TOF-MS results. Urine samples were collected from healthy volunteers, and patients with RCC or other urological diseases. Sensitivities and specificities of 81.8%–83.3% were achieved in an initial blind test. However, when the model was retested 10 months later, the sensitivities and specificities declined markedly, ranging from 41.0% to 76.6%. The authors demonstrated that factors contributing to the change included sample stability and instrument variability. In 2008, Wu *et al.* [26] employed the same SELDI-TOF-MS technology and found a group panel of peptides with masses of 4020, 4637, 5070 and 5500 kD as potential biomarkers. A sensitivity of 67.8% and a specificity of 81.4% were achieved in a blind test. On the other hand, Bosso *et al.* [27] used another set of three peptides for RCC diagnosis and

obtained a sensitivity of 100% and a specificity of 95%. One of the peptides was a fragment of uromodulin.

### 3.1.4 Other cancers

In addition to cancers in the urinary system, urinary proteomics has been applied to several non-urological cancers. Ward *et al.* [28] used SELDI-TOF-MS and MALDI-TOF-MS to study changes in urinary protein abundance caused by colon cancer, and identified 19 peptides that changed significantly. A logistic regression classification model based on peptide peak intensity was set up and achieved 78% sensitivity and 87% specificity. Three of the peaks were identified as fragments of fibrinogen, hepcidin-20 and  $\beta$ 2-microglobulin, respectively.

In studies of ovarian cancer, Ye *et al.* [29] reported that glycosylated eosinophil-derived neurotoxin and C-terminal osteopontin fragments could be used together for early diagnosis, with 93% specificity and 72% sensitivity. Petri *et al.* [30] adopted an equalizer beads approach and SELDI-TOF-MS to analyze urine samples from patients who had benign gynecological tumors, borderline tumors and malignant epithelial ovarian cancer. The three peaks that changed most significantly were fibrinogen alpha fragment, collagen alpha 1 (III) fragment, and fibrinogen beta NT fragment.

Tantipaiboonwong *et al.* [31] studied urinary proteome changes between lung cancer patients and healthy volunteers by 2-DE. Several differentially expressed proteins were identified as potential biomarkers, including CD59 glycoprotein, transthyretin, G(M2) activator protein, and immunoglobulin light chain.

## 3.2 Kidney transplantation

One of the most serious problems in kidney transplantation is acute rejection, which can lead to fatal consequences if not diagnosed in time. Renal biopsy is the gold standard to detect acute rejection presently. However, there are risks in performing renal biopsy on critically ill patients. Moreover, it is not acceptable to perform renal biopsy many times within a short time period. Therefore, there is an urgent need to find a non-invasive detection method to diagnose acute rejection early.

SELDI-TOF-MS analyses were performed by three independent groups to characterize urinary proteome signatures of renal allograft rejection. Each group identified a unique cluster of proteins as potential biomarkers. Clarke *et al.* [32] found a panel of peptides with masses of 6.5, 6.6, 6.7, 7.1 and 13.4 kD had the highest diagnostic power, as measured by the area under the ROC curve. In another analysis using a CART (Classification and Regression Tree) algorithm, the combination of two novel biomarker candidates with masses of 10.0 and 3.4 kD achieved a sensitivity of 83% and specificity of 100%. Schaub *et al.* [33,34] compared the urinary proteome of four patient groups following kidney transplantation: stable, acute rejection, acute tubular

necrosis, and recurrent (or *de novo*) glomerulopathy, as well as healthy controls. Peak clusters within the *m/z* regions of 5270–5550, 7050–7360 and 10530–11100 were identified and sequenced. Sequencing revealed the peaks as  $\beta$ 2-microglobulin fragments that were cleaved by aspartic protease at a urine pH<6. O’Riordan [35] reported that peaks at 2003.0, 2802.6, 4756.3, 5872.4, 6990.6, 19018.8 and 25665.7 were essential in the distinction of acute rejection and stable transplant groups. Two of these peaks were identified as  $\beta$ -defensin-1 (4.7 kD) and  $\alpha$ -1-antichymotrypsin (4.4 kD) [36].

Except for acute rejection, Quintana *et al.* [37] profiled the urinary proteome of 39 kidney chronic allograft dysfunction (CAD) patients and 32 controls by LC-MS/MS. Specific peptides derived from uromodulin and kininogen increased significantly in the control group. Low expression of uromodulin fragment at *m/z* 638.03 coupled with high expression at *m/z* 642.61 diagnosed CAD in virtually all cases. In addition, peptides at *m/z* 645.59 and 642.61 were able to classify patients with different forms of CAD, with specificities and sensitivities over 90% in the training set, and of approximately 70% in an independent test set.

## 3.3 Acute kidney injury

Acute kidney injury (AKI) or acute renal failure (ARF) has high morbidity and mortality rates. Serum creatinine is used routinely in clinical practice. However, it is neither specific nor sensitive. Moreover, types of AKI cannot be determined by serum creatine measurements. Due to the lack of early diagnostic biomarkers, patients are often diagnosed as AKI when the optimal time of treatment has already been missed.

Nguyen *et al.* [38] investigated early biomarkers for acute ischemic renal injury. Urine samples from patients undergoing cardiopulmonary bypass (CPB) were collected and analyzed by SELDI-TOF-MS. Proteins with masses of 28.5, 43 and 66 kD were used to predict ARF at 2 h following CPB with 100% sensitivity and specificity. Ho *et al.* [39] also used SELDI-TOF-MS to study acute ischemic renal injury.  $\beta$ 2-microglobulin and other known tubular injury biomarkers were identified. Hepcidin-25, a peptide hormone produced by the liver that regulates iron homeostasis, was also reported to be a novel biomarker.

## 3.4 Diabetic nephropathy

Diabetic nephropathy is a serious chronic complication of diabetes. Microalbuminuria is a predictor for diabetic nephropathy, but does not have satisfactory specificity and sensitivity.

To identify better biomarkers, Mischak *et al.* [40] profiled the urinary peptidome of healthy individuals: patients with type II diabetes and patients with diabetic nephropathy by CE-MS. Three peptides were sequenced and found to be

insulin-like peptide 3, uromodulin and an albumin fragment. On the other hand, Rao *et al.* [41] used 2-DE to analyze the urinary proteome of type II diabetes patients with normo-, micro-, or macroalbuminuria. They identified seven proteins to be progressively upregulated with increasing albuminuria as well as four proteins that showed progressive downregulation. The upregulated proteins were  $\alpha$ -1B glycoprotein, zinc- $\alpha$ <sub>2</sub>-glycoprotein,  $\alpha$ <sub>2</sub>-HS-glycoprotein, vitamin D-binding protein, calgranulin B,  $\alpha$ <sub>1</sub>-antitrypsin, and hemopexin, and the downregulated proteins were transthyretin, apolipoprotein A-I,  $\alpha$ <sub>1</sub>-microglobulin/bikunin precursor (AMBP), and plasma retinol-binding protein. The majority of these proteins were plasma glycoproteins. In another study, Jiang *et al.* [42] discovered and validated the 80 kD soluble fragment of E-cadherin, a novel candidate biomarker for diabetic nephropathy, by 2-DE, Western blotting and ELISA. Sensitivity and specificity of this biomarker were 78.8% and 80%. Immunohistochemical staining showed that E-cadherin expression decreased significantly in renal tubular epithelial cells of patients with diabetic nephropathy versus healthy controls.

Other studies compared diabetic nephropathy with other kidney diseases. Rossing *et al.* [43] demonstrated the importance of employing proper control groups in biomarker discovery. To obtain the diabetic nephropathy patterns, the urinary proteome of patients with diabetic nephropathy was first compared to patients with diabetes and normoalbuminuria. Patterns were set up based on 65 CE-MS peaks (many of those were peptide fragments of uromodulin and collagen) by two different mathematical models. This cluster of biomarkers was able to predict whether patients with diabetes and microalbuminuria would progress toward overt diabetic nephropathy. The authors then found that the same biomarkers also identified other chronic renal diseases (IgA nephropathy, focal segmental glomerulosclerosis, FSGS; membranous glomerulonephritis, MGN; and minimal change disease, MCD) as diabetic nephropathy 71% of the time, indicating that the 'diabetic nephropathy pattern' was in fact the chronic renal disease pattern. Seventeen complementary CE-MS peaks were then used to distinguish diabetic nephropathy from other chronic renal diseases, resulting in sensitivity and specificity of 81% and 91%, respectively.

In a SELDI-TOF-MS study, Dihazi *et al.* [44] compared urine samples from patients with type II diabetes, diabetic nephropathy, non-diabetic renal diseases and normal controls. They found that the process of diabetic nephropathy correlated with ubiquitination. Release of high amounts of ubiquitin ribosomal fusion protein UbA52 as well as downregulation of a processed form of ubiquitin at  $m/z$  6188 in urine samples could potentially be used to diagnose diabetic nephropathy. In a similar study, Lapolla *et al.* [45] adopted MALDI-TOF technology and identified three potential peptide biomarkers to differentiate diabetes, diabetic nephropathy and non-diabetic renal diseases. Lower expression of

uromodulin fragment ( $m/z$  1912) and upregulation of a peptide of collagen alpha-5 (IV) chain precursor ( $m/z$  1219) happened in general cases of nephropathy, whereas a peptide of collagen alpha-1(I) chain precursor ( $m/z$  2049) showed a slight increase of expression level specifically in the case of diabetic nephropathy.

### 3.5 Fanconi syndrome

Fanconi syndrome is a disorder due to complex defects in renal proximal tubular function. Proteins filtered by the glomerulus cannot be reabsorbed by tubulus in Fanconi syndrome patients. Cutillas *et al.* [46] compared urine samples from patients with Dent's disease, a variant of Fanconi syndrome, to urine samples from healthy individuals using three different proteomics approaches. Vitamin and prosthetic group carrier proteins, complement components, apolipoproteins, and several cytokines were found in higher amounts in the urinary proteome of Dent's disease patients, whereas kidney originated proteins were in smaller amounts. This study helps reveal the function of the proximal tubulus and its reabsorption mechanisms, and may provide the basis for disease treatment.

Drube *et al.* [47] applied CE-MS to study renal Fanconi syndrome in children. Urine samples were collected from seven pediatric patients with cystinosis, and six patients with ifosfamide-induced Fanconi syndrome, 54 healthy volunteers, 45 patients with other renal diseases (MCD, MGN, FSGS, and lupus nephritis). They identified osteopontin, uromodulin and collagen alpha-1 were as potential biomarkers, with 89% specificity and 82% sensitivity.

### 3.6 Distinguishing between multiple glomerular diseases

Identifying biomarkers for various glomerular diseases is becoming one of the most valuable and productive subfields in proteomics studies of kidney diseases. Merely comparing the urinary proteome from certain patients to that from healthy individuals cannot meet the clinical requirement for distinguishing between different glomerular diseases with similar symptoms.

Mischak *et al.* [11,40,48] have worked years on identifying signature urinary proteome patterns with CE-MS technology and bioinformatics for a series of glomerular diseases that may lead to renal failure. They compared urine samples from patients with MCD, FSGS, MGN and healthy controls to identify peptide expression patterns. The classification accuracy was 71.4% for MCD and FSGS and 92.9% for MGN [11]. In a similar study, Chalmers *et al.* [49] also used CE-MS to classify MCD, MGN, FSGS, IgA nephropathy and diabetic nephropathy by their urinary peptide patterns. Some of the candidate biomarkers were identified by Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) as albumin fragments.

Varghese *et al.* [50] profiled the urinary proteome of FSGS, lupus nephritis, MGN and diabetic nephropathy by 2-DE and employed a neural network algorithm with a training set of 16 samples. Another 16 samples served as test set. The above diseases were identified with sensitivities of 75%–86% and specificities of 67%–92%.

IgA nephropathy is the most common chronic glomerulonephritis in adults. Besides renal biopsy, there are no reliable methods for diagnosis. Haubitz *et al.* [48] examined the urinary peptidome from patients with IgA nephropathy, MGN and normal volunteers, and identified a unique signature expression pattern for IgA nephropathy. Using normal subjects and MGN patients as controls respectively, sensitivities were 100% and 77% and specificities were 90% and 100%, respectively. When comparing this pattern to their previously identified MCD, FSGS and diabetic nephropathy patterns, specificities of 100% and sensitivities of 100% were achieved.

### 3.7 Non-urological diseases

Urinary proteomics has also been applied to non-urological diseases. Kaiser *et al.* [51] and Weissinger *et al.* [52] both investigated potential biomarkers for graft-versus-host disease after allogeneic hematopoietic stem cell transplantation. Kaiser *et al.* identified 16 peptides from 40 disease samples and achieved a sensitivity and specificity of 100% and 82%, respectively. Weissinger *et al.* set up a model of 31 peptides and achieved a sensitivity of 83.1% and specificity of 75.6% in a blind test based on 599 samples.

Zimmerli *et al.* [53] studied urinary proteome changes caused by coronary artery disease using CE-MS. They identified 15 differentially expressed peptides as potential biomarkers, and achieved 98% sensitivity and 83% specificity in a blind test based on 59 samples.

Buhimschi *et al.* [54] identified a urinary proteome pattern of pre-eclampsia using SELDI-TOF-MS. This pattern was able to distinguish pre-eclampsia from other hypertensive or proteinuric disorders during pregnancy. The 21 amino acid C-terminal fragment of SERPINA1 and a cleavage product of albumin were identified as potential biomarkers. SERPINA1 was also correlated with a severe form of the disease.

Taneja *et al.* [55] studied both the urinary and plasma proteome of acute hepatitis E by 2-DE. Plasma transthyretin and urinary  $\alpha$ 1-microglobulin were identified as potential diagnosis biomarkers.

### 3.8 Animal models

Partially due to lack of biomarkers, early stage patients of many diseases are hard to find in clinical practice. To discover biomarkers for early diagnosis, animal models are sometimes very helpful. Other advantages of the animal model are that animal samples exhibit less inter-individual

variations than humans [56] and can be used to study interventions that cannot be tested on humans. More animal models will be developed and applied to biomarker discovery as more animal genomes are sequenced.

Holly *et al.* [57] studied sepsis-induced acute renal failure in a rat model. A series of proteins, including albumin, meprin-1-alpha and serine protease inhibitors, were identified as potential biomarkers. Meprin-1-alpha was also reported to be a potential drug target by the authors.

To distinguish between two glomerular diseases, Wang *et al.* [58] established two rat models: adriamycin nephropathy to mimic human FSGS, and Thy1.1 nephritis to mimic human mesangial proliferative glomerulonephritis (MsPGN). They profiled a ConA-enriched urinary glycoproteome by reverse phase liquid chromatography tandem mass spectrometry (RPLC-MS/MS) to identify proteins whose expression levels changed in the early stage of diseases. Among the identified proteins, 39 showed different directions of abundance changes, which could potentially be used to distinguish between the two glomerular diseases, whereas seven with similar direction of change may indicate general renal damages.

Large quantities of disease-related proteins are contained in urinary exosomes. In a rat model of cisplatin-induced acute kidney injury (AKI), Zhou *et al.* [59] detected upregulation of Fetuin-A in urinary exosomes in AKI rats using 2-DE. It is interesting to note, a similar change of Fetuin-A was observed in 3 ICU patients with AKI. Animal research also has limitations. For example, when urine samples are collected from rats, fecal contamination is inevitable. Animal studies tend to be successful only when the models can closely mimic human diseases.

## 4 Comments and prospects

### 4.1 Discovering diseases-specific biomarkers

In the current era, a major clinical problem is to distinguish between different diseases with similar symptoms to help doctors choose appropriate treatments. Simply comparing the urinary proteome of patients with a certain disease with healthy individuals is insufficient to solve this problem. Rigorous experimental design, especially comparing multiple diseases simultaneously, is the key to the success of biomarker studies.

For example, there is an emergent need for earlier diagnosis of acute renal allograft rejection. However, experiments designed to compare acute rejection and stable transplant groups are not able to find biomarkers that have sufficient specificity. The clinical problem in renal transplantation is that existing laboratory examinations, as well as early symptoms such as fever and oliguria, cannot distinguish acute rejection from other complications such as nephrotoxicity induced by immune inhibitors, acute tubular necro-

sis and urinary tract infection. Thus, appropriate treatments cannot be determined. Therefore, samples from all possible complications of renal allograft should serve as controls in studies of acute rejection.

As described in section 3.4, studies of diabetic nephropathy by Rossing *et al.* [43] illustrated the importance of comparing various diseases in the discovery of disease-specific biomarkers. When non-diabetic renal diseases were not included in the study, only biomarkers describing general renal damages were identified. Specific biomarkers were identified only when all diseases that might be confused with the target disease were included as control groups.

Another issue is the selection of healthy controls. For example, since cancers as well as chronic kidney disease usually occur in middle-aged and elderly people, urine samples from healthy young volunteers are not suitable controls to study these diseases. Zurbig *et al.* [60] observed similarity between biomarkers for kidney aging and for chronic kidney disease, indicating that age is an essential factor that must be accounted for when designing biomarker studies.

#### 4.2 Overlap between biomarkers identified by different proteomics methods

Three clusters of potential biomarkers were identified by three independent research groups, in a set of SELDI-TOF-MS studies on acute renal allograft rejection [32,33,35]. The poor overlap of these results using even the same type of mass spectrometers is not surprising, considering that different methods of sample preparation, laser sources, chips and data analysis algorithms were used. In addition, different immunosuppressive therapies result in different urinary protein expression patterns. The diverse biomarker candidates identified by different strategies may be understood in this way: They do not contradict, but complement each other, so that researchers can get more valuable information by integrating the knowledge obtained by different proteomics methods.

On the other hand, Rogers *et al.* [25] reported that sensitivity and specificity of the same cluster of biomarkers declined significantly when retested 10 months after the original research. Sample stability and instrument variability might contribute to that change. Because protein identification by mass spectrometry can be influenced by many factors, proteomics technologies (except multiple reaction monitoring) may not be suitable methods for biomarker validation. Methods with high reproducibility should be chosen to validate proteomics results.

#### 4.3 Influence of individual difference and dynamic variation in urinary proteomics studies

As mentioned in section 2, the urinary proteome can differ significantly among healthy individuals and even with time of

collection within the same individual due to exercise, diet, lifestyle and other factors. This physiological variation makes it more troublesome to discover biomarkers from human urine samples. Candidate biomarkers identified by proteomics approaches using small sample sets should be verified with larger sample sets in order to distinguish changes caused by diseases from those resulting from normal physiological variations.

Except for enlarging the size of validation sets, setting up a project to determine the variation range of each urinary protein under different physiological conditions in a normal population may be another way to solve the problem of physiological variation. Despite the large numbers of samples needed to obtain statistically significant results, once completed, the project would benefit many researchers. Results of all the urinary proteomics studies could be compared to identify proteins with expression changes that differ from normal expression variations in a healthy population, and therefore serve as potential biomarkers.

#### 4.4 Comparing urinary proteome with plasma proteome

About 30% of the urinary proteome comes from plasma [1,61]. Plasma proteins become part of the urinary proteome after filtration through the glomerulus and reabsorption into the tubulus. Therefore it is reasonable to assume that comparing the plasma proteome to the urinary proteome may lead to a better understanding of kidney physiology. Jia *et al.* [62] proposed the “black box theory”, in which a kidney was treated as a stand-alone black box, with plasma proteins and urine proteins as its input and output. Using bioinformatics methods, changes of proteins after kidney handling were identified to reveal some general rules of kidney function of protein handling. This work was only a pilot study and lacked quantitative analysis. Refined quantitative studies are therefore needed to calculate the filtration rate for each protein. The filtration rate changes of some proteins are potentially related to renal damage at a certain site. It is possible that it is not the absolute changes of proteins in a single biofluid, but their changes in plasma/urine abundance ratios that can serve as markers of disease in clinical practice.

#### 4.5 Establishment of web-based data-sharing platforms

Due to limited experimental resources, a single research group can handle only modest numbers of samples and diseases, making it difficult to identify disease-specific biomarkers with high confidence. The establishment of a data-sharing platform to collect all existing urinary protein/peptide expression profiles in normal and disease conditions could integrate worldwide urinary proteomics studies.

By comparing studies of the same diseases, researchers

can establish whether other groups have also identified a protein/peptide identified in their study. This would allow investigators to have higher confidence in identifying potential biomarkers. Researchers can also easily get information of biomarker candidates identified by other groups but not in their study to improve their understanding of disease mechanisms. More importantly, the disease-specificity of biomarker candidates could be better evaluated by comparing studies across various diseases. We believe that there are clear benefits to data-sharing platforms that can promote biomarker discovery.

Coon *et al.* [63] have established the Human Urinary Database ([http://mosaiques-diagnostics.de/diapatpcms/mosaiquescms/front\\_content.php?idcat=257](http://mosaiques-diagnostics.de/diapatpcms/mosaiquescms/front_content.php?idcat=257)) to collect 5010 polypeptides (very few of them are full-length proteins) identified by CE-MS from 3926 urine samples. Urinary peptide profiles of 13 different kidney diseases, nine cancers, four transplant-related diseases, and eight non-urological diseases were deposited in the database. Of those polypeptides deposited, 444 have sequence information.

Our group has carried out part of the data collection and website construction. In addition, a literature-based knowledge-sharing website is planned in the near future to assist all urinary proteomics researchers.

*We thank the following students in our lab for the collection and curation of literature references: Jia LuLu, Wei LiLong, Zhang Dan, Mu Yi, Guo ZhengGuang, Wang JiangNing, Zhu LiSi, Yang Fan, Li XunDou and Li MengLin.*

- 1 Pieper R, Gatlin C L, McGrath A M, *et al.* Characterization of the human urinary proteome: A method for high-resolution display of urinary proteins on two-dimensional electrophoresis gels with a yield of nearly 1400 distinct protein spots. *Proteomics*, 2004, 4: 1159–1174
- 2 Sun W, Li F, Wu S, *et al.* Human urine proteome analysis by three separation approaches. *Proteomics*, 2005, 5: 4994–5001
- 3 Adachi J, Kumar C, Zhang Y, *et al.* The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins. *Genome Biol*, 2006, 7: R80
- 4 Li Q R, Fan K X, Li R X, *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*, 24: 823–832
- 5 Lafitte D, Dussol B, Andersen S, *et al.* Optimized preparation of urine samples for two-dimensional electrophoresis and initial application to patient samples. *Clin Biochem*, 2002, 35: 581–589
- 6 Wang L, Li F, Sun W, *et al.* Concanavalin A-captured glycoproteins in healthy human urine. *Mol Cell Proteomics*, 2006, 5: 560–562
- 7 Gonzales P A, Pisitkun T, Hoffert J D, *et al.* Large-scale proteomics and phosphoproteomics of urinary exosomes. *J Am Soc Nephrol*, 2009, 20: 363–379
- 8 Khan A, Packer N H. Simple urinary sample preparation for proteomic analysis. *J Proteome Res*, 2006, 5: 2824–2838
- 9 Thongboonkerd V, Chutipongtanate S, Kanlaya R. Systematic evaluation of sample preparation methods for gel-based human urinary proteomics: Quantity, quality, and variability. *J Proteome Res*, 2006, 5: 183–191
- 10 Jürgens M, Appel A, Heine G, *et al.* Towards characterization of the human urinary peptidome. *Comb Chem High Throughput Screen*, 2005, 8: 757–765
- 11 Weissinger E M, Wittke S, Kaiser T, *et al.* Proteomic patterns established with capillary electrophoresis and mass spectrometry for diagnostic purposes. *Kidney Int*, 2004, 65: 2426–2434
- 12 Sun W, Chen Y, Li F, *et al.* Dynamic urinary proteomic analysis reveals stable proteins to be potential biomarkers. *Proteomics Clin Appl*, 2009, 3: 370–382
- 13 Vrooman O P, Witjes J A. Urinary markers in bladder cancer. *Eur Urol*, 2008, 53: 909–916
- 14 Goodison S, Rosser C J, Urquidí V. Urinary proteomic profiling for diagnostic bladder cancer biomarkers. *Expert Rev Proteomics*, 2009, 6: 507–514
- 15 Irmak S, Tilki D, Heukeshoven J, *et al.* Stage-dependent increase of orosomucoid and zinc-alpha2-glycoprotein in urinary bladder cancer. *Proteomics*, 2005, 5: 4296–4304
- 16 Saito M, Kimoto M, Araki T, *et al.* Proteome analysis of gelatin-bound urinary proteins from patients with bladder cancers. *Eur Urol*, 2005, 48: 865–871
- 17 Tan L B, Chen K T, Yuan Y C, *et al.* Identification of urine PLK2 as a marker of bladder tumors by proteomic analysis. *World J Urol*, 2010, 28: 117–122
- 18 Tsui K H, Tang P, Lin C Y, *et al.* Bikunin loss in urine as useful marker for bladder carcinoma. *J Urol*, 2010, 183: 339–344
- 19 Mueller J, von Eggeling F, Driesch D, *et al.* ProteinChip technology reveals distinctive protein expression profiles in the urine of bladder cancer patients. *Eur Urol*, 2005, 47: 885–893
- 20 Theodorescu D, Wittke S, Ross M M, *et al.* Discovery and validation of new protein biomarkers for urothelial cancer: A prospective analysis. *Lancet Oncol*, 2006, 7: 230–240
- 21 Goo Y A, Goodlett D R. Advances in proteomic prostate cancer biomarker discovery. *J Proteomics* advance online publication 14 April 2010
- 22 Rehman I, Azzouzi A R, Catto J W, *et al.* Proteomic analysis of voided urine after prostatic massage from patients with prostate cancer: A pilot study. *Urology*, 2004, 64: 1238–1243
- 23 Theodorescu D, Fliser D, Wittke S, *et al.* Pilot study of capillary electrophoresis coupled to mass spectrometry as a tool to define potential prostate cancer biomarkers in urine. *Electrophoresis*, 2005, 26: 2797–2808
- 24 M'Koma A E, Blum D L, Norris J L, *et al.* Detection of pre-neoplastic and neoplastic prostate disease by MALDI profiling of urine. *Biochem Biophys Res Commun*, 2007, 353: 829–834
- 25 Rogers M A, Clarke P, Noble J, *et al.* Proteomic profiling of urinary proteins in renal cancer by surface enhanced laser desorption ionization and neural-network analysis: Identification of key issues affecting potential clinical utility. *Cancer Res*, 2003, 63: 6971–6983
- 26 Wu D L, Zhang W H, Wang W J, *et al.* Proteomic evaluation of urine from renal cell carcinoma using SELDI-TOF-MS and tree analysis pattern. *Technol Cancer Res Treat*, 2008, 7: 155–160
- 27 Bosso N, Chinello C, Picozzi S, *et al.* Human urine biomarkers of renal cell carcinoma evaluated by ClinProt. *Proteomics Clin Appl*, 2008, 2: 1036–1046
- 28 Ward D G, Nyangoma S, Joy H, *et al.* Proteomic profiling of urine for the detection of colon cancer. *Proteome Sci*, 2008, 6: 19
- 29 Ye B, Skates S, Mok S C, *et al.* Proteomic-based discovery and characterization of glycosylated eosinophil-derived neurotoxin and COOH-terminal osteopontin fragments for ovarian cancer in urine. *Clin Cancer Res*, 2006, 12: 432–441
- 30 Petri A L, Simonsen A H, Yip T T, *et al.* Three new potential ovarian cancer biomarkers detected in human urine with equalizer bead technology. *Acta Obstet Gynecol Scand*, 2009, 88: 18–26
- 31 Tantipaiboonwong P, Sinchaikul S, Sriyam S, *et al.* Different techniques for urinary protein analysis of normal and lung cancer patients. *Proteomics*, 2005, 5: 1140–1149
- 32 Clarke W, Silverman B C, Zhang Z, *et al.* Characterization of renal allograft rejection by urinary proteomic analysis. *Ann Surg*, 2003, 237: 660–664
- 33 Schaub S, Wilkins J, Weiler T, *et al.* Urine protein profiling with surface-enhanced laser-desorption/ionization time-of-flight mass spectrometry. *Kidney Int*, 2004, 65: 323–332
- 34 Schaub S, Wilkins J A, Antonovici M, *et al.* Proteomic-based identi-



- fication of cleaved urinary beta2-microglobulin as a potential marker for acute tubular injury in renal allografts. *Am J Transplant*, 2005, 5: 729–738
- 35 O’Riordan E, Orlova T N, Mei J J, *et al.* Bioinformatic analysis of the urine proteome of acute allograft rejection. *J Am Soc Nephrol*, 2004, 15: 3240–3248
- 36 O’Riordan E, Orlova T N, Podust V N, *et al.* Characterization of urinary peptide biomarkers of acute rejection in renal allografts. *Am J Transplant*, 2007, 7: 930–940
- 37 Quintana L F, Campistol J M, Alcolea M P, *et al.* Application of label-free quantitative peptidomics for the identification of urinary biomarkers of kidney chronic allograft dysfunction. *Mol Cell Proteomics*, 2009, 8: 1658–1673
- 38 Nguyen M T, Ross G F, Dent C L, *et al.* Early prediction of acute renal injury using urinary proteomics. *Am J Nephrol*, 2005, 25: 318–326
- 39 Ho J, Lucy M, Krokhn O, *et al.* Mass spectrometry-based proteomic analysis of urine in acute kidney injury following cardiopulmonary bypass: a nested case-control study. *Am J Kidney Dis*, 2009, 53: 584–595
- 40 Mischak H, Kaiser T, Walden M, *et al.* Proteomic analysis for the assessment of diabetic renal damage in humans. *Clin Sci (Lond)*, 2004, 107: 485–495
- 41 Rao P V, Lu X, Standley M, *et al.* Proteomic identification of urinary biomarkers of diabetic nephropathy. *Diabetes Care*, 2007, 30: 629–637
- 42 Jiang H, Guan G, Zhang R, *et al.* Identification of urinary soluble E-cadherin as a novel biomarker for diabetic nephropathy. *Diabetes Metab Res Rev*, 2009, 25: 232–241
- 43 Rossing K, Mischak H, Dakna M, *et al.* Urinary proteomics in diabetes and CKD. *J Am Soc Nephrol*, 2008, 19: 1283–1290
- 44 Dihazi H, Muller G A, Lindner S, *et al.* Characterization of diabetic nephropathy by urinary proteomic analysis: Identification of a processed ubiquitin form as a differentially excreted protein in diabetic nephropathy patients. *Clin Chem*, 2007, 53: 1636–1645
- 45 Lapolla A, Seraglia R, Molin L, *et al.* Low molecular weight proteins in urines from healthy subjects as well as diabetic, nephropathic and diabetic-nephropathic patients: A MALDI study. *J Mass Spectrom*, 2009, 44: 419–425
- 46 Cutillas P R, Chalkley R J, Hansen K C, *et al.* The urinary proteome in Fanconi syndrome implies specificity in the reabsorption of proteins by renal proximal tubule cells. *Am J Physiol Renal Physiol*, 2004, 287: F353–364
- 47 Drube J, Schiffer E, Mischak H, *et al.* Urinary proteome pattern in children with renal Fanconi syndrome. *Nephrol Dial Transplant*, 2009, 24: 2161–2169
- 48 Haubitz M, Wittke S, Weissinger E M, *et al.* Urine protein patterns can serve as diagnostic tools in patients with IgA nephropathy. *Kidney Int*, 2005, 67: 2313–2320
- 49 Chalmers M J, Mackay C L, Hendrickson C L, *et al.* Combined top-down and bottom-up mass spectrometric approach to characterization of biomarkers for renal disease. *Anal Chem*, 2005, 77: 7163–7171
- 50 Varghese S A, Powell T B, Budisavljevic M N, *et al.* Urine biomarkers predict the cause of glomerular disease. *J Am Soc Nephrol*, 2007, 18: 913–922
- 51 Kaiser T, Kamal H, Rank A, *et al.* Proteomics applied to the clinical follow-up of patients after allogeneic hematopoietic stem cell transplantation. *Blood*, 2004, 104: 340–349
- 52 Weissinger E M, Schiffer E, Hertenstein B, *et al.* Proteomic patterns predict acute graft-versus-host disease after allogeneic hematopoietic stem cell transplantation. *Blood*, 2007, 109: 5511–5519
- 53 Zimmerli L U, Schiffer E, Zurbig P, *et al.* Urinary proteomic biomarkers in coronary artery disease. *Mol Cell Proteomics*, 2008, 7: 290–298
- 54 Buhimschi I A, Zhao G, Funai E F, *et al.* Proteomic profiling of urine identifies specific fragments of SERPINA1 and albumin as biomarkers of preeclampsia. *Am J Obstet Gynecol*, 2008, 199: 551.e1–16
- 55 Taneja S, Sen S, Gupta V K, *et al.* Plasma and urine biomarkers in acute viral hepatitis E. *Proteome Sci*, 2009, 7: 39
- 56 Thongboonkerd V, Klein J B, Arthur J M. Proteomic identification of a large complement of rat urinary proteins. *Nephron Exp Nephrol*, 2003, 95: e69–78
- 57 Holly M K, Dear J W, Hu X, *et al.* Biomarker and drug-target discovery using proteomics in a new rat model of sepsis-induced acute renal failure. *Kidney Int*, 2006, 70: 496–506
- 58 Wang Y, Chen Y, Zhang Y, *et al.* Differential ConA-enriched urinary proteome in rat experimental glomerular diseases. *Biochem Biophys Res Commun*, 2008, 371: 385–390
- 59 Zhou H, Pisitkun T, Aponte A, *et al.* Exosomal Fetuin-A identified by proteomics: A novel urinary biomarker for detecting acute kidney injury. *Kidney Int*, 2006, 70: 1847–1857
- 60 Zurbig P, Decramer S, Dakna M, *et al.* The human urinary proteome reveals high similarity between kidney aging and chronic kidney disease. *Proteomics*, 2009, 9: 2108–2117
- 61 Marshall T, Williams K. Two-dimensional electrophoresis of human urinary proteins following concentration by dye precipitation. *Electrophoresis*, 1996, 17: 1265–1272
- 62 Jia L, Zhang L, Shao C, *et al.* An attempt to understand kidney’s protein handling function by comparing plasma and urine proteomes. *PLoS One*, 2009, 4: e5146
- 63 Coon J, Zurbig P, Dakna M, *et al.* CE-MS analysis of the human urinary proteome for biomarker discovery and disease diagnostics. *Proteomics Clin Appl*, 2008, 2: 964–973

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.