Introduction

Human urine is mainly composed of shed cells, debris, and secreted components from the urinary tract as well as blood components that have passed through glomerular filtration and renal tubule reabsorption. Therefore, urine contains useful information not only regarding the kidney and urinary tract, but also about more distant organs. Analysis of the urinary proteome could aid the discovery of biomarkers for both urogenital and systemic diseases. Moreover, compared to serum, human urine is relatively simple and easy to collect, which makes urinary proteome analysis an attractive approach in clinical proteomics research.

Because inherent and environmental factors may influence the components of the urinary proteome, the biological and technical variations are important issues for urinary proteome research. Many groups [1-6] have contributed data regarding this issue and found that (1) a considerable degree of variation can be found in intra-day (collection from one volunteer at different daily time points), intra-individual (collection from one volunteer on different days), and inter-individual (collection from different volunteers) samples; (2) the variation of five intra-day samples (first morning, second morning, 24 h, random, and water loading void) was similar; (3) the variation of intra-individual samples was less than that of inter-individual samples; and (4) technical variation was less than biological variation. Although great variations have been found in different urinary samples, a number of urinary proteins were demonstrated to be consistently present in urine samples collected at different time points and from different individuals [4]. Moreover, Nagaraj *et al.* used healthy volunteers to construct a common dataset of 500 urinary proteins [5]. Taken together, the findings to date indicate that the urinary proteome is relatively stable and a good source for disease biomarkers.

Since the first urinary proteome analysis was published in 2001 [7], many clinical urinary proteome differential analyses have been reported, including analyses of samples from urogenital diseases (kidney transplantation [8], diabetic nephropathy [9], obstructive nephropathy [10], bladder cancer [11], prostate cancer [12], and others) and non-urogenital diseases (hematopoietic stem cell transplantation [13], coronary artery disease [14], and others).

A urinary proteome database from normal human samples plays an important role in biomarker discovery. In the discovery stage, a database could be used as a control for a disease group. In the validation stage, the candidate proteins/peptides could be selected from the database for MS-based or immuno-based validation. Many groups have analyzed the normal human urinary proteome using various approaches and have identified more than 2500 urinary proteins to date [15]. Analyses of the normal urinary proteome have usually analyzed pooled or individual samples from several volunteers [15-17]. However, due to the variations in the urinary proteome, it is still unknown whether these data represent the true pattern of the normal urinary proteome. If the sample number was less than the minimal number required for a representative database in a group, then an analysis may only represent the pattern of selected individuals and not the entire group, which would be misleading for subsequent studies. Therefore, to obtain a representative urinary proteome, it is necessary to define the minimal urinary sample number needed. To the best of our knowledge, such an analysis has not been conducted to date.

In this study, inter-individual and inter-gender variations were taken into consideration to achieve a representative urinary proteome. An individual urinary proteome analysis of 10 male and 10 female normal overnight samples from healthy volunteers was used to define the minimal samples number required. Because the data-dependent acquisition mode in LC/MS/MS analysis is biased against low abundance proteins [18], replicate experimental strategies are often used to obtain a comprehensive analysis [18-20], and therefore this strategy was also adopted for this study. To determine how many runs are necessary to obtain a comprehensive result for one urine sample by 1DLC/MS/MS, a pooled sample from ten male samples was analyzed with forty runs. Based on these calculations, 10 male and 10 female urinary samples were then analyzed by replicate 1DLC/MS/MS. For qualitative analysis by intra-gender and inter-gender analysis, the minimal sample number for male, female, and normal groups was estimated. For quantitative analysis, the variation of protein abundance was defined by spectrum count and western blotting methods. And then the minimal sample number for quantitative proteomic analysis was estimated. The overall workflow is shown in Figure 1.

Materials and methods

Apparatus

An LTQ XL mass spectrometer was purchased from Thermo-Fisher (San Jose, CA). A 1200 nano-HPLC system was obtained from Agilent (Foster, CA). An ADVANCE CaptiveSpray source for Thermo and C18 reverse phase capillary column was purchased from Michrom Bioresources (Auburn, CA). The microwave oven used in this study was solid-state Whirlpool model VIP271 (Shanghai, China), and the maximum output power was 850 W.

Reagents

Deionized water from a MilliQ RG ultrapure water system (Bedford, MA) was used at all times. HPLC grade

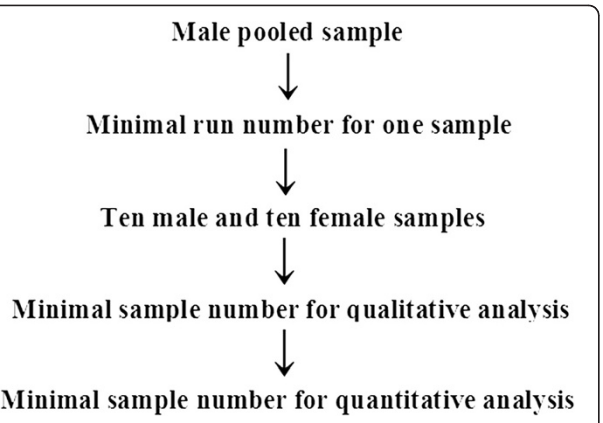

Figure 1 The overall workflow of this study. Pooled male samples were analyzed in replicate to define the minimal 1DLC/MS/MS number for a comprehensive analysis. After validating the analytical completeness in 10 males and 10 females, the minimal sample size in a male/female group was defined. Within a group, the minimal sample size and the ratio of males and females was further investigated.

acetonitrile (ACN) and formic acid, ammonium bicarbonate, iodoacetamide, dithiothreitol (DTT), sequencing grade modified trypsin, and protease inhibitor phenylmethylsulfonyl fluoride (PMSF) were purchased from Sigma-Aldrich (St. Louis, MO).

Urine collection

Overnight urine samples were collected from twenty consenting individuals, including 10 males and 10 females (average age 28 and 31 y, respectively). The donors had no acute or chronic illnesses and were not taking any prescription or over-the-counter medications. No female was menstruating at the time of urine collection. Each specimen was collected in 250 mL conical tubes. The samples were immediately acidified to pH 2.7 with hydrochloric acid and then cooled to 4 °C to prevent bacterial growth and proteolysis.

Acetone precipitation

All procedures were performed at 4 °C. Urine samples were centrifuged at 5,000 x g for 30 min and the pellets were removed. The supernatants were then precipitated with 75% v/v acetone for 16 h followed by centrifugation at 12,000 x g for 30 min. The pellets were resuspended in lysis buffer (7 M urea, 2 M thiourea, 50 mM Tris, and 50 mM dithioerythritol) and subjected to protein quantitation by the Bradford method. One pooled male sample was mixed using ten male samples with an "equal amounts of protein" criterion.

Protein digestion

Each sample was digested with trypsin as previously described [21]. Briefly, each sample was reduced with DTT by heating at 100 °C for 5 min and then alkykated with iodoacetamide at room temperature in the dark for 45 min. The samples were then digested with trypsin (1:50) for 1 min under microwave irradiation at 850 W using the following method: samples were placed into 1.5 mL polypropylene vials and a container with 1,000 mL of water was placed beside the sample vials to absorb the extra microwave energy. The microwave oven was turned on for 1 min. After microwave irradiation, the vials were removed from the microwave oven and lyophilized to near dryness.

1DLC/MS/MS

All lyophilized samples were redissolved in 0.1% formic acid (buffer A) at a concentration of 5 mg/mL before MS analysis. All peptide mixtures were analyzed on a reverse phase C18 capillary LC column from Michrom Bioresources (100 μm x 150 mm, 3 μm, 0.5 μL/min). The elution gradient was 5-30% buffer B (0.1% formic acid, 99.9% ACN; flow rate, 0.5 μL/min) for 100 min. Eluted peptides were analyzed by an LTQ XL electrospray ion trap mass spectrometer. Ions were detected in a survey scan from 400 to 2000 amu followed by 10 data-dependent MS/MS scans (1 μscan each, isolation width 2 amu, 35% normalized collision energy, dynamic exclusion for 90 s) in a completely automated fashion.

Western blot analysis

Western blots were performed for three proteins: alpha 1 antitrypsin, ceruloplasmin, and beta-2-microglobulin to confirm the variation in urine. For each protein, 16 normal human overnight urine samples (10 females and 6 males) were used. A urine sample from a stage IV diabetic nephropathy patient was used as a control. Thirty micrograms urine protein from each urine sample after acetone precipitation was separated on a 4-12% NuPAGE gel (Invitrogen). Proteins were then transferred to polyvinylidene difluoride (PVDF) membranes (Millipore, USA). The membranes were blocked for 1 h at room temperature in Tris-buffered saline (TBS) with 5% skim milk, and then incubated overnight at 4 °C in a 1% milk solution containing mouse monoclonal anti-alpha 1 antitrypsin (SERPINA1) antibody (1:1000, ab9400, Abcam), mouse monoclonal anti-ceruloplasmin antibody (1:1000, ab51083, Abcam), and rabbit monoclonal anti-beta-2-microglobulin antibody (1:1000, ab15976, Abcam), respectively. The membranes were washed three times for 5 min with TBST (Tris-buffered saline with 0.05% Tween-20), and then incubated with horseradish peroxidase-labeled goat anti-mouse or mouse anti-rabbit IgG secondary antibodies (diluted 1:5000, Abcam) at room temperature for 3 h. After washing three times for 5 min each in TBST, the membrane was visualized with an ECL detection kit (Millipore, Bedford,