Published in final edited form as:
Proteomics Clin Appl. 2012 June ; 6(0): . doi:10.1002/prca.201100108.

Application of systems biology principles to protein biomarker discovery: Urinary exosomal proteome in renal transplantation

Trairak Pisitkun#1, **Maria T. Gandolfo**#2, **Samarjit Das**2, **Mark A. Knepper**1, and **Serena M. Bagnasco**2

[1]Epithelial Systems Biology Laboratory, NHLBI, NIH, Bethesda, MD

[2]Department of Pathology, Johns Hopkins University, Baltimore, MD

# These authors contributed equally to this work.

Abstract

Purpose—In MS-based studies to discover urinary protein biomarkers, an important question is how to analyze the data to find the most promising potential biomarkers to be advanced to large-scale validation studies. Here, we describe a ‘systems biology-based’ approach to address this question.

Experimental design—We analyzed large-scale LC-MS/MS data of urinary exosomes from renal allograft recipients with biopsy-proven evidence of immunological rejection or tubular injury. We asked whether bioinformatic analysis of urinary exosomal proteins can identify protein groups that correlate with biopsy findings and whether the protein groups fit with general knowledge of the pathophysiological mechanisms involved.

Results—LC-MS/MS analysis of urinary exosomal proteomes identified more than 1000 proteins in each pathologic group. These protein lists were analyzed computationally to identify Biological Process and KEGG Pathway terms that are significantly associated with each pathological group. Among the most informative terms for each group were: “sodium ion transport” for tubular injury; “immune response” for all rejection; “epithelial cell differentiation” for cell-mediated rejection; and “acute inflammatory response” for antibody-mediated rejection. Based on these terms, candidate biomarkers were identified using a novel strategy to allow a dichotomous classification between different pathologic categories.

Conclusions and clinical relevance—The terms and candidate biomarkers identified make rational connections to pathophysiological mechanisms, suggesting that the described bioinformatic approach will be useful in advancing large-scale biomarker identification studies toward a validation phase.

Keywords

exosome; proteomics; rejection; transplant; urine

Corresponding author: Mark A. Knepper, MD, PhD, Epithelial Systems Biology Laboratory, Building 10, Room 6N260, National Institutes of Health, Bethesda, MD 20892-1603, Tel: 301-496-3064 Fax: 301-402-1443 knepperm@nhlbi.nih.gov.

Conflict of interest statement
The authors have no competing financial interests to disclose.

Introduction

The main objective of the present study is to develop a bioinformatic strategy that can reveal exploitable patterns in protein mass spectrometry data from urinary samples. The question being addressed is how to proceed from mass spectrometry-generated lists of proteins to a choice of candidate biomarkers for future testing in validation clinical trials. Here we describe and illustrate an approach derived from the field of systems biology that uses computational tools to discover biologically meaningful differences among experimentally derived protein lists [1, 2]. This approach is based on the idea that the objective of biomarker discovery studies should be to identify *biological processes* that are deranged rather than to look for specific proteins that may stand out. The rationale has been presented previously [3]. Briefly, it is based on the idea that protein biomarkers that "make sense" from the perspective of pathophysiology are more likely to succeed in the clinical setting than randomly discovered proteins that have uncertain connection to the relevant disease mechanisms (see *Discussion*). In this study, as a 'proof of principle', we apply such an approach to proteins detected by LC-MS/MS (liquid chromatography-tandem mass spectrometry) analysis of urinary exosome samples from three groups of renal transplant patients with different pathologic findings in the allograft biopsy: tubular injury, cell-mediated rejection, and antibody-mediated rejection. The bioinformatic approach outlined identified biological processes and candidate biomarkers associated with the different patient groups that fit with known pathophysiological mechanisms.

Materials and methods

Patients and samples

Urine samples were collected prospectively from adult renal transplant recipients followed at the Johns Hopkins Comprehensive Transplant Center. These patients underwent a kidney graft biopsy, either for cause (due to rise in baseline serum creatinine) or as protocol biopsy in patients with stable renal function. All patients received standard immunosuppressive treatment with prednisone, mycophenolate, and tacrolimus. Exclusion factors included combined transplantation of kidney and other organs, evidence of infection in the kidney (bacterial or viral such as BK/polyoma, CMV, or Herpes), primary glomerular disease (*de novo* or recurrent), post-transplant lymphoproliferative disorder (PTLD). Kidney graft biopsies were processed for routine studies as previously described including staining with hematoxilin-eosin (H&E), periodic acid-Schiff (PAS) methenamine silver, and Masson's trichrome for light microscopy examination [4]. C4d staining by immunofluorescence was routinely performed on frozen sections of transplant kidney biopsies (alternatively C4d immunoperoxidase-based stain was also available on sections from paraffin embedded tissue). Biopsies were graded for cellular- and antibody-mediated rejection according to the Banff score [5, 6]. Based on the graft biopsy findings, patients were assigned to four groups: *Non-Specific Findings* (N), mild to moderate *Tubular Injury* (TI), *Cell-Mediated Rejection* (CMR), or *Antibody-Mediated Rejection* (AMR). The study protocol was approved by the Johns Hopkins Medical Institutions Review Board.

Freshly voided urine (20-200 ml) was collected in the morning, before the biopsy (but was not the first morning urine sample), and was processed immediately to isolate urinary exosomes using the differential centrifugation procedure described by Gonzales et al [7]. Processed protein samples solubilized in Laemmli reagent were pooled for each pathological group for analysis by LC-MS/MS.

Sample processing and mass spectrometry analysis

Triplicate sets of 200 μg of urinary exosomal proteins pooled from each pathological group (7 TI, 6 CMR, 3 AMR, and 2 N, supporting Information Table S1) were separated by one dimensional SDS/PAGE electrophoresis using 10% polyacrylamide gels. After staining with Coomassie blue, the gels were de-stained, cut in multiple slices, dehydrated, reduced, alkylated, and subjected to trypsin digestion at 37°C overnight to obtain peptides, which were reconstituted in 0.1% formic acid for analysis as described [8]. LC-MS/MS was performed on triplicate aliquots of digested protein samples using LTQ Orbitrap XL (two runs) and LTQ Orbitrap Velos (one run) (Thermo Scientific, Waltham, MA). Reversed-phase C18 chromatographic separation of trapped peptides was carried out on a prepacked Beta Basic C18 PicoFrit column (75 μm i.d. × 10 cm length; New Objective) at 300 nL/min using the following gradient: 2–5% solvent B for 2 min; 5–45% solvent B for 45 min; 45–50% solvent B for 5 min; 50–95% solvent B for 5 min (solvent A: 0.1% formic acid in 98% water, 2% acetonitrile; solvent B: 0.1% formic acid in 100% acetonitrile). To identify peptide sequences, the MS data was searched against National Center for Biotechnology Information (NCBI) Reference Sequence human protein database (April 8, 2010, 38767 entries) which included a list of common contaminating proteins and scored using SEQUEST algorithm on Proteome Discoverer software (ver. 1.1, Thermo Scientific). Precursor ion tolerance was 50 ppm, while fragment ion tolerance was 1.0 Da. Two missed trypsin cleavage sites were allowed. Static modifications included carbamidomethylation of cysteine (+57.021 Da) and variable modifications included oxidation of methionine (+15.995 Da). The peptide false discovery rate was limited to < 2% for individual peptides using the target-decoy approach [9]. All MS spectral files associated with this manuscript may be downloaded from ProteomeCommons.org Tranche using the following hashes:

1. **LTQ Orbitrap XL Data 1st set**

   p7DyXnHc+EyANszEZXPsThX/EzRPhJR0YmfDYGZJoRREeKgRm00/
   jM47tDEH5Gh0wzO/Bk N5BcA0IQPCCN2QmJAWs4UAAAAAAAAfow==

2. **LTQ Orbitrap XL Data 2nd set**

   RD+8p8/vHKT7QhWVYe0hMsVKKEZitqPGBgv60DDvoGuOQ1ZaGz6Bpgmy
   +o/2FSOe+Y4v 1BfFtzlZcKEOJ43T7HwEVkUAAAAAAAAiHg==

3. **LTQ Orbitrap Velos Data**

   DlI2b2OfY46ui/
   Fln5JivUyHW03ZxU4Pv5xapK4xl9FB4wdSyW1Ffo4gQ38tjNwZo4SehiBbar
   RU0k8QiwrOiS5phrYAAAAAAAACyg==

For the evaluation of relative abundance of exosomal proteins among groups, the spectral counts were normalized by dividing by the number of theoretical tryptic peptides detectable by LC-MS/MS (criteria for predicting a detectable tryptic peptide are as followed: theoretical m/z = 300-2000 for charge states +1 to +3, peptide length = 6-35, possible missed cleavage sites < 3). These values were further normalized by dividing by the median value for each group to account for potential differences in sample size.

Bioinformatic analysis

Identified proteins were classified as ‘extrinsic to exosomes’ if they possessed the *Gene Ontology Cellular Component* term of “extracellular space”, but lacked the terms “anchor” and “cytosol”. The remaining proteins were placed in the ‘intrinsic to exosomes’ subgroup. Protein lists were analyzed using the DAVID bioinformatic tool (Database for Annotation, Visualization and Integrated Discovery, NIAID, Bethesda, MD http://david.abcc.niifcrf.gov/) [10] to determine what *Gene Ontology Biological Process* terms are over-represented relative to the concatenated lists from all four conditions. The DAVID tool was also applied to determine what KEGG pathways (*Kyoto Encyclopedia of Genes and Genomes*) [11] are over-represented in each pathological group.

Results

Proteomes were profiled by LC-MS/MS in urinary exosome samples from renal allograft recipients with acute decrease in renal function and distinct biopsy findings: tubular injury (TI), cell-mediated rejection (CMR), and antibody-mediated rejection (AMR) (supporting Information Table S1). We also examined urinary exosomes of two allograft recipients with stable renal function undergoing “protocol” kidney graft biopsy, with non-specific findings (N). We selected cases that were stereotypical of each category, recognizing that real-world cases often display less distinct histology. Exosome samples were pooled within each histology class (supporting Information Table S1) and analyzed by LC-MS/MS. 1989 urine exosomal proteins were identified in renal transplant patients (false discovery rate < 2%, target-decoy analysis). This total exceeds the 1160 proteins reported in the *Urinary Exosome Protein Database* (http://dir.nhlbi.nih.gov/papers/lkem/exosome/). A list of all identified proteins along with associated median-normalized spectral counts, a measure of relative abundance (*Materials and methods*), can be accessed online through http://helixweb.nih.gov/ESBL/Database/EXORT/ (username: clp, password: Esbl!@#$, login to be removed upon acceptance of this paper). Each group had a large subset of uniquely identified proteins, TI: 353 proteins; CMR: 322 proteins; and AMR 165 proteins, with 1073 proteins present in more than one group.

We classified the identified proteins into one of two subgroups (*Materials and methods*): ‘intrinsic to exosomes’ or ‘extrinsic to exosomes’ (see the online database). The ‘extrinsic’ proteins are presumably plasma proteins that have crossed the glomerular filter and are nonspecifically associated with exosomes. The ‘intrinsic’ proteins were used for bioinformatic analysis. Overall, 92% of the proteins identified were classified as intrinsic, suggesting that they were a component of exosomes (or similar membranous structures) released from kidney cells.

Interestingly, there seems to be a higher proportion of extrinsic proteins in the transplanted control group than in the non-transplanted group (chi-square, $p < 0.0001$) when comparing the most abundant proteins (based on spectral counts) found in the group of transplant recipients with non-specific findings versus the most abundant proteins found in a group of healthy, non-transplanted volunteers that we studied previously [7], (supporting Information Table S2). Although these results need to be confirmed on a larger number of observations in transplanted patients without kidney pathology, they suggest that transplanted kidneys excrete higher levels of plasma proteins than non-transplanted kidneys even with no obvious pathological changes suggestive of tubular injury or rejection, and are also consistent with previously noted differences in the proteomic profile of urine from patients with transplanted or native kidneys [12, 13].

Identification of biological processes over-represented in groups

We analyzed the intrinsic exosomal proteins in each group to determine what characteristics distinguish the groups based on the biological processes that they take part in. This analysis was conducted with an online bioinformatic analysis tool called “DAVID” (*Materials and methods*), using the *Gene Ontology (GO) Database*, which attaches hierarchical descriptors to all annotated proteins coded by the human genome. We applied the *Biological Process GO* terms to identify which *Biological Process* terms are represented statistically more frequently in a list of exosomal proteins from a given patient group than from the entire set of exosomal proteins.

We asked whether any *GO Biological Process* terms are over-represented in the urinary exosomal proteins from kidneys of each patient group with abnormal biopsy findings (compared with the list of all proteins identified in transplanted kidneys). Table 1 shows the *GO Biological Processes* identified in each group with $p < 0.05$ (Fisher Exact). A comparison of Table 1A, 1B, and 1C reveals that several terms were enriched in more than one group. For example the term "immune response" is found in both graft rejection groups. Thus, proteins populating the "immune response" list can be deemed potential discriminators between rejection and tubular injury, but not between the two forms of rejection. supporting Information Table S3 shows proteins in this "immune response" list (cell-mediated and antibody-mediated rejection combined). Each pathological group in Table 1 is associated with at least one unique term (represented in **Bold**). As discussed below, most of these appear to be related to the underlying pathophysiology.

For the *Tubular Injury* group, the *Gene Ontology Biological Process* term "sodium ion transport" was the only unique term (Table 1A). This protein list (Table 2) includes a large number of transporters, many of them are sodium-coupled solute transporters expressed in the apical brush border membrane of the proximal tubule. Based on the transcriptomic databases of proximal tubule, medullary thick ascending limb, and inner medullary collecting duct [14], [15], 25 out of all 29 proteins (86%) in Table 2 were expressed with relatively high signal values in proximal tubule. This finding is presumably indicative of proximal tubule damage and therefore makes sense from a pathophysiological perspective.

In the *Cell-Mediated Rejection* group (Table 1B), there are several unique terms, including three similar terms: "actin filament based process", "actin cytoskeleton organization", and "cytoskeleton organization". The proteins in the former group are shown in supporting Information Table S4. An additional unique term in the *Cell-Mediated Rejection* group is "epithelial cell differentiation" (Table 3). These are predominantly proteins involved in intermediate filament organization and include four different uroplakin isoforms. The latter are expressed in the renal ureter and pelvis, and point to the possibility that the transitional epithelium of the urinary drainage system is a target tissue in cell-mediated allograft rejection.

In the *Antibody-Mediated Rejection* group (Table 1C), several unique terms appear, including three similar terms relevant to protein trafficking. The proteins in the "protein localization" group are listed in supporting Information Table S5. Additional unique terms in the *Antibody-Mediated Rejection* group are "response to unfolded protein“ (supporting Information Table S6) and "acute inflammatory response" (Table 4). The latter category is dominated by components of the complement pathway. These categories refer to aspects of pathophysiology that appear logical in the context of current knowledge about antibody-mediated rejection (see *Discussion*).

Identification of KEGG biological pathways over-represented in pathological groups

We next analyzed the data to determine what characteristics distinguish the groups based on the KEGG biological pathways that they take part in, asking the question: "What KEGG pathways are represented statistically more frequently in a list of urinary exosomal proteins from a given patient group than in the list of all identified intrinsic exosomal proteins in all groups?"

The findings are summarized in Table 5. There were no unique KEGG pathways for exosomal proteins in the *Tubular Injury* group. One pathway was over-represented in the two Rejection groups, namely "endocytosis" (supporting Information Table S7 shows the "endocytosis" proteins combined from the two Rejection groups). Two KEGG pathways were uniquely over-represented in the list of exosomal proteins found in the *Cell-Mediated*

Rejection group, viz. "tight junction" (supporting Information Table S8) and "glutathione metabolism" (supporting Information Table S9). Interestingly the "tight junction" KEGG pathway with the *Cell-Mediated Rejection* group proteins include three proteins that are critically involved in epithelial polarization through PDZ interactions, viz. Crumbs3 (*CRB3*), PALS1 (*MPP5*), and PATJ (*INADL*) [16], and may be important in the recovery of tubular integrity damaged by "Tubulitis", a signature pathologic lesion in this type of rejection. Also as seen in Table 5, four KEGG pathways were uniquely over-represented in the *Antibody-Mediated Rejection* group, viz. "antigen processing and presentation" (Table 6), "neurotrophin signaling pathway", "pathways in cancer" and "pathogenic E. coli infection" (supporting Information Tables S10-S12). The "antigen processing and presentation" list is dominated by major histocompatibility complex proteins and heat shock proteins.

A strategy for identification of candidate biomarker proteins from *Biological Processprotein* lists

The approach described above selects groups of MS-identified proteins associated via *Gene Ontology Biological Process* terms or KEGG Pathways that relate to pathophysiological mechanisms in different patient groups. In full-scale biomarker studies, the next task would be to select candidate biomarkers from the protein groups that are predicted from the mass spectrometry data to have the potential to discriminate patient groups. In this study, we developed a strategy to identify candidate biomarkers from our 'proof-of-principle' data using the protein lists associated with the over-represented *Biological Processes* for each pathologic category (Tables 2-4 and supporting Information Tables S3-6). In this context, a first practical goal would be to identify pairs of proteins that will allow a dichotomous classification into *Rejection* vs. *Tubular Injury* categories (Figure 1). A second goal would be to choose protein pairs that allow samples to be categorized as *Antibody-Mediated Rejection* vs. *Cell-Mediated Rejection* (Figure 2). Selecting pairs of proteins that may have sufficient discriminating power between two different conditions, based on ratios of their abundance, would avoid the need of normalization factors, a main hurdle in optimizing assay conditions for urine biomarkers.

As rationalized in the *Discussion*, criteria for selection of candidate proteins from the *Gene Ontology Biological Process* lists obtained in the present study were: 1) maximization of the spectral count ratio between the two states being discriminated and 2) high absolute abundance based on spectral counts. Figure 1A shows a heat map representation of the dichotomous classification between *All Rejection* vs. *Tubular Injury*. The numbers for each pair of proteins are discrimination factors between the two conditions for a protein pair [Values are $\log_2(R_R/R_{TI})$, where R_R is the median-normalized spectral count ratio of the two proteins for *All Rejection* and R_{TI} is the ratio of the two proteins in *Tubular Injury*]. The best discrimination factors are the highest positive numbers. Figure 1B shows the same type of information for the comparison between *Antibody-Mediated Rejection* versus *Tubular Injury* and Figure 1C shows comparable information for *Cell-Mediated Rejection* versus *Tubular Injury*. Figure 2 shows the dichotomous classification between *Antibody-Mediated Rejection* and *Cell-Mediated Rejection* in which the median-normalized spectral count ratio between potential markers are represented in the same manner as in Figure 1. All proteins presented in Figures 1 and 2 were also selected from the *Biological Process* protein lists only if their absolute abundance levels in the pathologic category associated with a particular protein list were at least 2-fold above the median value for that pathologic category (median-normalized spectral count values > 2).

Discussion

Typically biomarker development studies involve at least three distinct stages: discovery, validation, and implementation [17]. Here, we describe a bioinformatic approach to the analysis of discovery data. We used state-of-the-art mass spectrometry techniques for proteomic analysis of urinary exosomes from renal transplant recipients with renal dysfunction and biopsy-proven tubular injury, cell-mediated rejection, or antibody-mediated rejection. Although this approach was developed in the context of urinary proteomics data, it theoretically could be applied to analysis of data from other body fluids.

A central question addressed in this study is: “how best to analyze large-scale mass spectrometry data from urinary biomarker discovery studies to find proteins that have a meaningful connection to the pathophysiology of the disease process under study?” Optimally, candidate protein biomarkers must be sufficiently abundant to allow measurement in whole urine and should ‘make sense’ in terms of the pathophysiological processes involved [3]. The latter requirement is based on three considerations: a) measurements of biomarkers consistent with current pathophysiological knowledge will more readily be integrated with other types of results in future clinical studies; b) such biomarkers will be more readily acceptable and useful to the clinical nephrology community than those with no known connection to pathophysiology; and c) biomarkers that fit with recognized mechanisms are less likely to be false positives.

Because of the considerations outlined in the foregoing *Discussion*, we have devised a ‘process-oriented’ approach to bioinformatic interpretation of the protein mass spectrometry discovery data. Thus, we are not looking for specific proteins that are enriched in individual patient groups, but rather we wish to identify classes of proteins (involved in specific biological processes) that can be shown statistically to be associated with individual patient groups. To implement this approach, we used the online computer utility, DAVID (see *Materials and methods*), to identify either *Gene Ontology Biological Process* (GO-BP) terms or KEGG Pathway terms that are selectively associated with lists of proteins identified in urinary exosomes in each renal transplant subgroup. Based on this approach, we can summarize the key observations in our ‘proof of principle’ samples as follows:

1. The only GO-BP term significantly associated with the *Tubular Injury* exosomal protein list was “sodium ion transport” (Table 1A). The list of proteins in this group (Table 2) consists chiefly of transporter proteins characteristic of the proximal tubule brush border but also includes proteins from the thick ascending limb of Henle (*SLC12A1*) and distal convoluted tubule (*SLC12A3*). These findings are consistent with the idea that tubular injury is seen mainly in proximal tubule, thick ascending limb of Henle, and distal convoluted tubule, where ATP demand is high because of the energy cost of transepithelial solute transport. This study did not single out any of the proteins previously proposed to be associated with ischemic tubule injury such as KIM-1 [18], Hsp72 [19], klotho [20], IL-6 [21], NGAL [22], L-FABP [23], netrin-1 [24], or fetuin-A [25], presumably because the pathophysiology is dependent on the context (transplanted kidney with immunosuppression) and because our bioinformatic analysis focused on functional classes of proteins.
2. Proteins classified with the GO-BP term “immune response” were enriched in the two rejection groups but not the *Tubular Injury* group (Table 1). This classification makes sense in the context of the immunological nature of rejection. The proteins in this group are dominated by proteins of the major histocompatibility complex and complement components (supporting Information Table S3).

3. The GO-BP term "epithelial cell differentiation" is uniquely associated with *Cell-Mediated Rejection* (Table 1B). Despite its name, this category is made up largely of proteins associated with stratified epithelia including the transitional epithelium of the pelvis and ureter (Table 3). For example, we identified four uroplakins, abundant proteins that form a large part of the surface area of the umbrella cells of transitional epithelia [26, 27]. This result is compatible with the hypothesis that cell-mediated rejection may affect multiple epithelial cells of the transplanted organ, including the pelvis and ureter that are not affected in the other two pathologies.
4. Several GO-BP terms are associated with proteins in the *Antibody-Mediated Rejection* group including "acute inflammatory response" (Table 1C). Several complement components were found in the "acute inflammatory response" category (Table 4), a finding that fits with the proposed roles for the complement pathway in the pathophysiology of antibody-mediated rejection, and with current efforts to devise therapies based on inhibition of terminal complement activation with targeted monoclonal antibodies [28].

In a full-blown discovery study, the ultimate objective would be the development and testing of immunoassays (or other assays) for selected biomarker proteins. A major problem in designing useful urinary assays is normalization. The need for normalization arises from the fact that water excretion is highly variable, depending on physiological factors, rendering measurement of the absolute concentrations of a given biomarker substance virtually useless for making practical comparisons. Much has been written about the use of various measures such as creatinine, osmolality, Tamm-Horsfall protein, etc., to normalize urinary measurements. Such normalizing factors add work and variability to the read-out. However, because many clinical problems (including the one described in this paper) involve the development of a dichotomous classification that distinguishes one state from another, the normalization problem may be obviated; it is theoretically sufficient to form ratios of measured biomarker abundances, thereby canceling out the normalization factor. In the example developed for this study, two dichotomies are seen: 1) between rejection and tubular injury and 2) between cell-mediated rejection and antibody-mediated rejection. Consequently, in such studies it is rational to identify pairs of biomarker proteins, which, when measured under arbitrary loading conditions, give a concentration ratio that correlates with disease state. If the assays are linear, the ratio should be independent of water excretion rate or sample size, and separate normalization should be unnecessary.

Given the above observations, we selected candidate biomarkers representative of the biological processes identified by bioinformatic analysis. Operationally, we identified pairs of candidate biomarker proteins that potentially could allow a dichotomous classification into *Rejection* vs. *Tubular Injury* (Figure 1) and, among those classified as Rejection, will allow samples to be categorized as *Antibody-Mediated* vs. *Cell-Mediated* (Figure 2). Among the proteins in Figures 1 and 2, *SERPING1* [29] and *S100A9* [30] have already been reported to be candidate biomarker proteins for acute rejection in renal allografts. Many membrane protein candidates were identified and have not been reported in the literatures associated with renal allograft rejection. This observation is likely to be due to the use of urinary exosomes in this study. Urinary exosomes are tiny (40–80 nm) membranous structures secreted by every renal tubule epithelial cell type, as well as podocytes and transitional epithelia from the urinary collecting system [31].

In summary, the object of this paper was to describe and illustrate an approach to bioinformatic analysis of biomarker discovery data in order to facilitate identification of candidate biomarkers for future testing in validation clinical trials. For this purpose, we prepared samples from a small number of actual renal transplant patients with different pathological changes in the allograft. The primary goal of this study was not to validate protein biomarkers for different processes that could cause an elevation of serum creatinine in renal transplant patients. Optimally, such validation studies would warrant the analysis of larger numbers of samples than used here. Nevertheless, data presented here, when combined with results from other biomarker studies [12] [13] [32] [33] [34] [35], could be useful in the design of validation studies for a limited number of protein biomarker candidates that make pathophysiological sense.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Parts of this study were funded with a grant from the National Kidney Foundation of Maryland (to SMB) and the operating budget of Division of Intramural Research, National Heart, Lung, and Blood Institute (Project ZO1-HL001285 to MAK). Mass spectrometry was conducted in the National Heart, Lung, and Blood Institute Proteomics Core Facility (director, Marjan Gucek). A portion of the data described in this study was presented in Abstract form at the American Society of Nephrology Renal Week 2009.

References

[1]. Khositseth S, Pisitkun T, Slentz DH, Wang G, et al. Quantitative protein and mRNA profiling shows selective post-transcriptional control of protein expression by vasopressin in kidney cells. Mol Cell Proteomics. 2011; 10:M110–004036. [PubMed: 20940332]

[2]. Gunaratne R, Braucht DW, Rinschen MM, Chou CL, et al. Quantitative phosphoproteomic analysis reveals cAMP/vasopressin-dependent signaling pathways in native renal thick ascending limb cells. Proc Natl Acad Sci U S A. 2010; 107:15653–15658. [PubMed: 20713729]

[3]. Knepper MA. Common sense approaches to urinary biomarker study design. J Am Soc Nephrol. 2009; 20:1175–1178. [PubMed: 19470673]

[4]. Bagnasco SM, Tsai W, Rahman MH, Kraus ES, et al. CD20-positive infiltrates in renal allograft biopsies with acute cellular rejection are not associated with worse graft survival. American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons. 2007; 7:1968–1973. [PubMed: 17617861]

[5]. Racusen LC, Colvin RB, Solez K, Mihatsch MJ, et al. Antibody-mediated rejection criteria - an addition to the Banff 97 classification of renal allograft rejection. American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons. 2003; 3:708–714. [PubMed: 12780562]

[6]. Solez K, Colvin RB, Racusen LC, Haas M, et al. Banff 07 classification of renal allograft pathology: updates and future directions. American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons. 2008; 8:753–760. [PubMed: 18294345]

[7]. Gonzales PA, Pisitkun T, Hoffert JD, Tchapyjnikov D, et al. Large-scale proteomics and phosphoproteomics of urinary exosomes. J Am Soc Nephrol. 2009; 20:363–379. [PubMed: 19056867]

[8]. Pisitkun T, Shen RF, Knepper MA. Identification and proteomic profiling of exosomes in human urine. Proc Natl Acad Sci U S A. 2004; 101:13368–13373. [PubMed: 15326289]

[9]. Elias JE, Gygi SP. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat Methods. 2007; 4:207–214. [PubMed: 17327847]

[10]. 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. [PubMed: 19131956]

[11]. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000; 28:27–30. [PubMed: 10592173]

[12]. Clarke W, Silverman BC, Zhang Z, Chan DW, et al. Characterization of renal allograft rejection by urinary proteomic analysis. Ann Surg. 2003; 237:660–664. discussion 664-665. [PubMed: 12724632]

[13]. O'Riordan E, Orlova TN, Podust VN, Chander PN, et al. Characterization of urinary peptide biomarkers of acute rejection in renal allografts. Am J Transplant. 2007; 7:930–940. [PubMed: 17331118]

[14]. Huling JC, Pisitkun T, Song JH, Yu MJ, et al. Gene expression databases for kidney epithelial cells. Am J Physiol Renal Physiol. 2012; 302:F401–407. [PubMed: 22114206]

[15]. Wright J, Morales MM, Sousa-Menzes J, Ornellas D, et al. Transcriptional adaptation to Clcn5 knockout in proximal tubules of mouse kidney. Physiol Genomics. 2008; 33:341–354. [PubMed: 18349385]

[16]. Wang Q, Margolis B. Apical junctional complexes and cell polarity. Kidney Int. 2007; 72:1448–1458. [PubMed: 17914350]

[17]. Granger CB, Van Eyk JE, Mockrin SC, Anderson NL. National Heart, Lung, And Blood Institute Clinical Proteomics Working Group report. Circulation. 2004; 109:1697–1703. [PubMed: 15078805]

[18. 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; 62:237–244. [PubMed: 12081583]

[19]. Barrera-Chimal J, Perez-Villalva R, Cortes-Gonzalez C, Ojeda-Cervantes M, et al. Hsp72 is an early and sensitive biomarker to detect acute kidney injury. EMBO Mol Med. 2011; 3:5–20. [PubMed: 21204265]

[20]. Aiello S, Noris M. Klotho in acute kidney injury: biomarker, therapy, or a bit of both? Kidney Int. 2010; 78:1208–1210. [PubMed: 21116272]

[21]. Dennen P, Altmann C, Kaufman J, Klein CL, et al. Urine interleukin-6 is an early biomarker of acute kidney injury in children undergoing cardiac surgery. Crit Care. 2010; 14:R181. [PubMed: 20942931]

[22]. Bolignano D, Donato V, Coppolino G, Campo S, et al. Neutrophil gelatinase-associated lipocalin (NGAL) as a marker of kidney damage. Am J Kidney Dis. 2008; 52:595–605. [PubMed: 18725016]

[23]. Matsui K, Kamijo-Ikemori A, Hara M, Sugaya T, et al. Clinical significance of tubular and podocyte biomarkers in acute kidney injury. Clin Exp Nephrol. 2011; 15:220–225. [PubMed: 21153750]

[24]. Ramesh G, Krawczeski CD, Woo JG, Wang Y, Devarajan P. Urinary netrin-1 is an early predictive biomarker of acute kidney injury after cardiac surgery. Clin J Am Soc Nephrol. 2010; 5:395–401. [PubMed: 20007677]

[25]. Zhou H, Pisitkun T, Aponte A, Yuen PS, et al. Exosomal Fetuin-A identified by roteomics: a novel urinary biomarker for detecting acute kidney injury. Kidney Int. 2006; 70:1847–1857. [PubMed: 17021608]

[26]. Riedel I, Liang FX, Deng FM, Tu L, et al. Urothelial umbrella cells of human ureter are heterogeneous with respect to their uroplakin composition: different degrees of urothelial maturity in ureter and bladder? Eur J Cell Biol. 2005; 84:393–405. [PubMed: 15819416]

[27]. Wu XR, Kong XP, Pellicer A, Kreibich G, Sun TT. Uroplakins in urothelial biology, function, and disease. Kidney Int. 2009; 75:1153–1165. [PubMed: 19340092]

[28]. Stegall MD, Gloor JM. Deciphering antibody-mediated rejection: new insights into mechanisms and treatment. Curr Opin Organ Transplant. 2010; 15:8–10. [PubMed: 19890210]

[29]. Sigdel TK, Kaushal A, Gritsenko M, Norbeck AD, et al. Shotgun proteomics identifies proteins specific for acute renal transplant rejection. Proteomics Clin Appl. 2010; 4:32–47. [PubMed: 20543976]

[30]. Eikmans M, Roos-van Groningen MC, Sijpkens YW, Ehrchen J, et al. Expression of surfactant protein-C, S100A8, S100A9, and B cell markers in renal allografts: investigation of the prognostic value. J Am Soc Nephrol. 2005; 16:3771–3786. [PubMed: 16251238]

[31]. Knepper MA, Pisitkun T. Exosomes in urine: who would have thought…? Kidney Int. 2007; 72:1043–1045. [PubMed: 17943150]

[32]. Quintana LF, Sole-Gonzalez A, Kalko SG, Banon-Maneus E, et al. Urine proteomics to detect biomarkers for chronic allograft dysfunction. J Am Soc Nephrol. 2009; 20:428–435. [PubMed: 19056874]

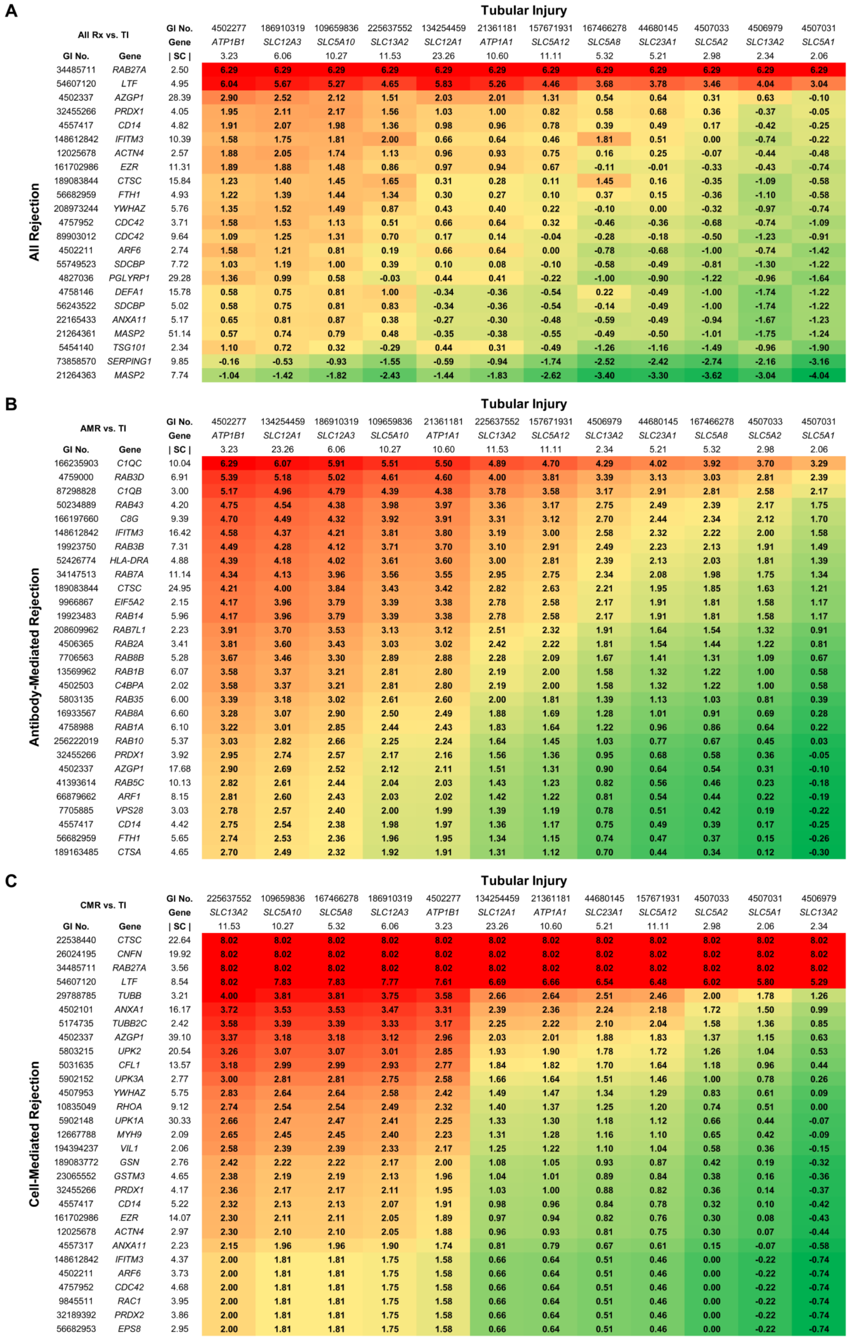
[33]. Ling XB, Sigdel TK, Lau K, Ying L, et al. Integrative urinary peptidomics in renal transplantation identifies biomarkers for acute rejection. J Am Soc Nephrol. 2010; 21:646–653. [PubMed: 20150539]

[34]. Srivastava M, Eidelman O, Torosyan Y, Jozwik C, et al. Elevated expression levels of ANXA11, integrins beta3 and alpha3, and TNF-alpha contribute to a candidate proteomic signature in urine for kidney allograft rejection. Proteomics Clin Appl. 2011; 5:311–321. [PubMed: 21591265]

[35]. Metzger J, Chatzikyrkou C, Broecker V, Schiffer E, et al. Diagnosis of subclinical and clinical acute T-cell-mediated rejection in renal transplant patients by urinary proteome analysis. Proteomics Clin Appl. 2011; 5:322–333. [PubMed: 21538920]

Statement of clinical relevance

A rising serum creatinine concentration in renal transplant patients raises a diagnostic dilemma: Is the rise in creatinine due to rejection or to tubular injury? If the former is true, what is the pathophysiological basis of the rejection? Presently these questions are best answered by examination of tissue from renal biopsy, an invasive procedure that requires hours to provide a diagnosis. The ultimate long-term objective of the present line of investigation is to devise immunological assays that can predict the biopsy results, possibly allowing earlier start of appropriate therapeutic intervention pending biopsy findings. The question addressed in this paper is: “how best to analyze large-scale mass spectrometry data from urinary biomarker discovery studies to find proteins that have a meaningful connection to the pathophysiology of the disease process under study?” To address this, we used state-of-the-art mass spectrometry techniques for analysis of urinary exosomes from patients with renal transplants, including those with biopsy evidence of tubular injury, cell-mediated rejection, and antibody-mediated rejection. We have developed a ‘process-oriented’ approach to bioinformatic interpretation of the protein mass spectrometry discovery data that facilitates identification of candidate biomarkers for future testing in validation clinical trials.



A

All Rejection

Tubular Injury

All Rx vs. TI		GI No.	4502277	186910319	109659836	225637552	134254459	21361181	157671931	167466278	44680145	4507033	4506979	4507031
		Gene	*ATP1B1*	*SLC12A3*	*SLC5A10*	*SLC13A2*	*SLC12A1*	*ATP1A1*	*SLC5A12*	*SLC5A8*	*SLC23A1*	*SLC5A2*	*SLC13A2*	*SLC5A1*
GI No.	**Gene**	**| SC |**	3.23	6.06	10.27	11.53	23.26	10.60	11.11	5.32	5.21	2.98	2.34	2.06
34485711	*RAB27A*	2.50	6.29	6.29	6.29	6.29	6.29	6.29	6.29	6.29	6.29	6.29	6.29	6.29
54607120	*LTF*	4.95	6.04	5.67	5.27	4.65	5.83	5.26	4.46	3.68	3.78	3.46	4.04	3.04
4502337	*AZGP1*	28.39	2.90	2.52	2.12	1.51	2.03	2.01	1.31	0.54	0.64	0.31	0.63	-0.10
32455266	*PRDX1*	4.05	1.95	2.11	2.17	1.56	1.03	1.00	0.82	0.58	0.68	0.36	-0.37	-0.05
4557417	*CD14*	4.82	1.91	2.07	1.98	1.36	0.98	0.96	0.78	0.39	0.49	0.17	-0.42	-0.25
148612842	*IFITM3*	10.39	1.58	1.75	1.81	2.00	0.66	0.64	0.46	1.81	0.51	0.00	-0.74	-0.22
12025678	*ACTN4*	2.57	1.88	2.05	1.74	1.13	0.96	0.93	0.75	0.16	0.25	-0.07	-0.44	-0.48
161702986	*EZR*	11.31	1.89	1.88	1.48	0.86	0.97	0.94	0.67	-0.11	-0.01	-0.33	-0.43	-0.74
189083844	*CTSC*	15.84	1.23	1.40	1.45	1.65	0.31	0.28	0.11	1.45	0.16	-0.35	-1.09	-0.58
56682959	*FTH1*	4.93	1.22	1.39	1.44	1.34	0.30	0.27	0.10	0.37	0.15	-0.36	-1.10	-0.58
208973244	*YWHAZ*	5.76	1.35	1.52	1.49	0.87	0.43	0.40	0.22	-0.10	0.00	-0.32	-0.97	-0.74
4757952	*CDC42*	3.71	1.58	1.53	1.13	0.51	0.66	0.64	0.32	-0.46	-0.36	-0.68	-0.74	-1.09
89903012	*CDC42*	9.64	1.09	1.25	1.31	0.70	0.17	0.14	-0.04	-0.28	-0.18	-0.50	-1.23	-0.91
4502211	*ARF6*	2.74	1.58	1.21	0.81	0.19	0.66	0.64	0.00	-0.78	-0.68	-1.00	-0.74	-1.42
55749523	*SDCBP*	7.72	1.03	1.19	1.00	0.39	0.10	0.08	-0.10	-0.58	-0.49	-0.81	-1.30	-1.22
4827036	*PGLYRP1*	29.28	1.36	0.99	0.58	-0.03	0.44	0.41	-0.22	-1.00	-0.90	-1.22	-0.96	-1.64
4758146	*DEFA1*	15.78	0.58	0.75	0.81	1.00	-0.34	-0.36	-0.54	0.22	-0.49	-1.00	-1.74	-1.22
56243522	*SDCBP*	5.02	0.58	0.75	0.81	0.83	-0.34	-0.36	-0.54	-0.14	-0.49	-1.00	-1.74	-1.22
22165433	*ANXA11*	5.17	0.65	0.81	0.87	0.38	-0.27	-0.30	-0.48	-0.59	-0.49	-0.94	-1.67	-1.23
21264361	*MASP2*	51.14	0.57	0.74	0.79	0.48	-0.35	-0.38	-0.55	-0.49	-0.50	-1.01	-1.75	-1.24
5454140	*TSG101*	2.34	1.10	0.72	0.32	-0.29	0.44	0.31	-0.49	-1.26	-1.16	-1.49	-0.96	-1.90
73858570	*SERPING1*	9.85	-0.16	-0.53	-0.93	-1.55	-0.59	-0.94	-1.74	-2.52	-2.42	-2.74	-2.16	-3.16
21264363	*MASP2*	7.74	-1.04	-1.42	-1.82	-2.43	-1.44	-1.83	-2.62	-3.40	-3.30	-3.62	-3.04	-4.04

B

Antibody-Mediated Rejection

Tubular Injury

AMR vs. TI		GI No.	4502277	134254459	186910319	109659836	21361181	225637552	157671931	4506979	44680145	167466278	4507033	4507031
		Gene	*ATP1B1*	*SLC12A1*	*SLC12A3*	*SLC5A10*	*ATP1A1*	*SLC13A2*	*SLC5A12*	*SLC13A2*	*SLC23A1*	*SLC5A8*	*SLC5A2*	*SLC5A1*
GI No.	**Gene**	**| SC |**	3.23	23.26	6.06	10.27	10.60	11.53	11.11	2.34	5.21	5.32	2.98	2.06
166235903	*C1QC*	10.04	6.29	6.07	5.91	5.51	5.50	4.89	4.70	4.29	4.02	3.92	3.70	3.29
4759000	*RAB3D*	6.91	5.39	5.18	5.02	4.61	4.60	4.00	3.81	3.39	3.13	3.03	2.81	2.39
87298828	*C1QB*	3.00	5.17	4.96	4.79	4.39	4.38	3.78	3.58	3.17	2.91	2.81	2.58	2.17
50234889	*RAB43*	4.20	4.75	4.54	4.38	3.98	3.97	3.36	3.17	2.75	2.49	2.39	2.17	1.75
166197660	*C8G*	9.39	4.70	4.49	4.32	3.92	3.91	3.31	3.12	2.70	2.44	2.34	2.12	1.70
148612842	*IFITM3*	16.42	4.58	4.37	4.21	3.81	3.80	3.19	3.00	2.58	2.32	2.22	2.00	1.58
19923750	*RAB3B*	7.31	4.49	4.28	4.12	3.71	3.70	3.10	2.91	2.49	2.23	2.13	1.91	1.49
52426774	*HLA-DRA*	4.88	4.39	4.18	4.02	3.61	3.60	3.00	2.81	2.39	2.13	2.03	1.81	1.39
34147513	*RAB7A*	11.14	4.34	4.13	3.96	3.56	3.55	2.95	2.75	2.34	2.08	1.98	1.75	1.34
189083844	*CTSC*	24.95	4.21	4.00	3.84	3.43	3.42	2.82	2.63	2.21	1.95	1.85	1.63	1.21
9966867	*EIF5A2*	2.15	4.17	3.96	3.79	3.39	3.38	2.78	2.58	2.17	1.91	1.81	1.58	1.17
19923483	*RAB14*	5.96	4.17	3.96	3.79	3.39	3.38	2.78	2.58	2.17	1.91	1.81	1.58	1.17
208609962	*RAB7L1*	2.23	3.91	3.70	3.53	3.13	3.12	2.51	2.32	1.91	1.64	1.54	1.32	0.91
4506365	*RAB2A*	3.41	3.81	3.60	3.43	3.03	3.02	2.42	2.22	1.81	1.54	1.44	1.22	0.81
7706563	*RAB8B*	5.28	3.67	3.46	3.30	2.89	2.88	2.28	2.09	1.67	1.41	1.31	1.09	0.67
13569962	*RAB1B*	6.07	3.58	3.37	3.21	2.81	2.80	2.19	2.00	1.58	1.32	1.22	1.00	0.58
4502503	*C4BPA*	2.02	3.58	3.37	3.21	2.81	2.80	2.19	2.00	1.58	1.32	1.22	1.00	0.58
5803135	*RAB35*	6.00	3.39	3.18	3.02	2.61	2.60	2.00	1.81	1.39	1.13	1.03	0.81	0.39
16933567	*RAB8A*	6.60	3.28	3.07	2.90	2.50	2.49	1.88	1.69	1.28	1.01	0.91	0.69	0.28
4758988	*RAB1A*	6.10	3.22	3.01	2.85	2.44	2.43	1.83	1.64	1.22	0.96	0.86	0.64	0.22
256222019	*RAB10*	5.37	3.03	2.82	2.66	2.25	2.24	1.64	1.45	1.03	0.77	0.67	0.45	0.03
32455266	*PRDX1*	3.92	2.95	2.74	2.57	2.17	2.16	1.56	1.36	0.95	0.68	0.58	0.36	-0.05
4502337	*AZGP1*	17.68	2.90	2.69	2.52	2.12	2.11	1.51	1.31	0.90	0.64	0.54	0.31	-0.10
41393614	*RAB5C*	10.13	2.82	2.61	2.44	2.04	2.03	1.43	1.23	0.82	0.56	0.46	0.23	-0.18
66879662	*ARF1*	8.15	2.81	2.60	2.43	2.03	2.02	1.42	1.22	0.81	0.54	0.44	0.22	-0.19
7705885	*VPS28*	3.03	2.78	2.57	2.40	2.00	1.99	1.39	1.19	0.78	0.51	0.42	0.19	-0.22
4557417	*CD14*	4.42	2.75	2.54	2.38	1.98	1.97	1.36	1.17	0.75	0.49	0.39	0.17	-0.25
56682959	*FTH1*	5.65	2.74	2.53	2.36	1.96	1.95	1.34	1.15	0.74	0.47	0.37	0.15	-0.26
189163485	*CTSA*	4.65	2.70	2.49	2.32	1.92	1.91	1.31	1.12	0.70	0.44	0.34	0.12	-0.30

C

Cell-Mediated Rejection

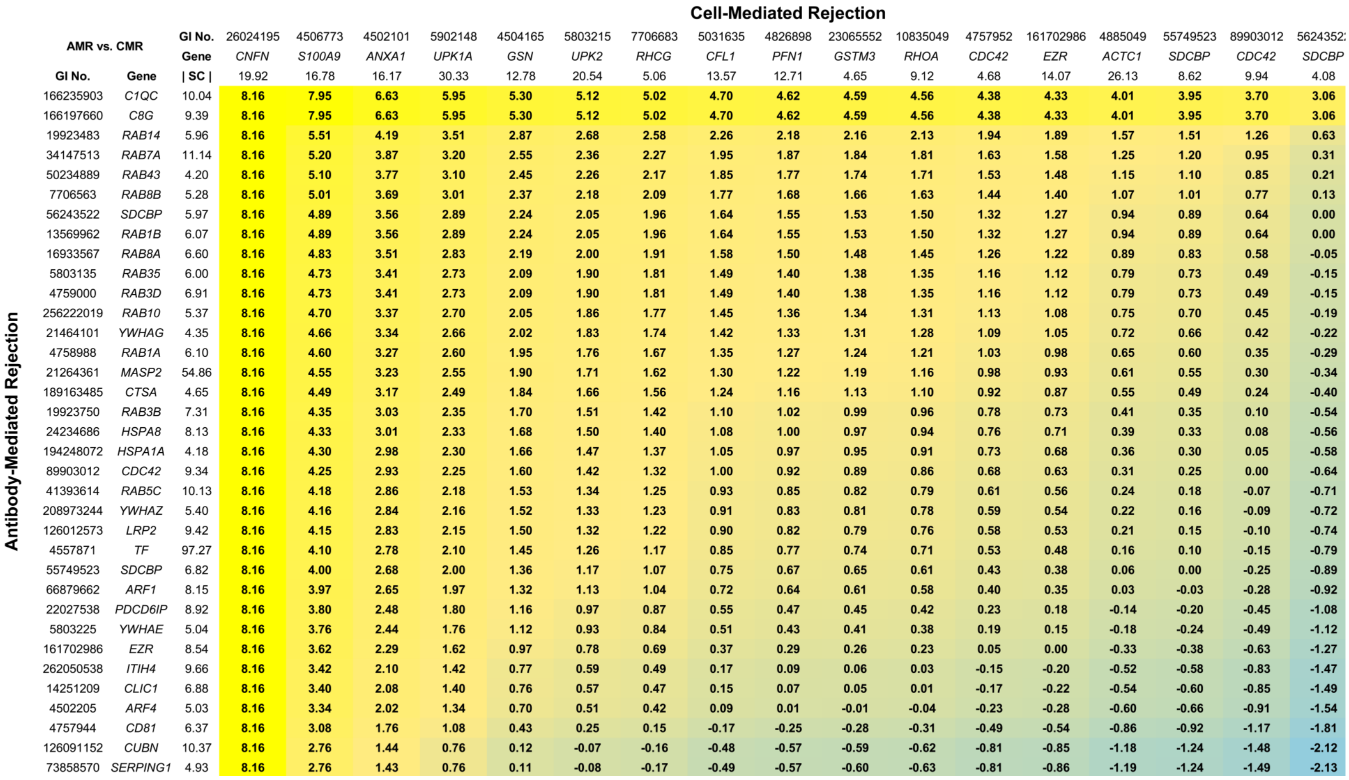
Tubular Injury

CMR vs. TI		GI No.	225637552	109659836	167466278	186910319	4502277	134254459	21361181	44680145	157671931	4507033	4507031	4506979
		Gene	*SLC13A2*	*SLC5A10*	*SLC5A8*	*SLC12A3*	*ATP1B1*	*SLC12A1*	*ATP1A1*	*SLC23A1*	*SLC5A12*	*SLC5A2*	*SLC5A1*	*SLC13A2*
GI No.	**Gene**	**| SC |**	11.53	10.27	5.32	6.06	3.23	23.26	10.60	5.21	11.11	2.98	2.06	2.34
22538440	*CTSC*	22.64	8.02	8.02	8.02	8.02	8.02	8.02	8.02	8.02	8.02	8.02	8.02	8.02
26024195	*CNFN*	19.92	8.02	8.02	8.02	8.02	8.02	8.02	8.02	8.02	8.02	8.02	8.02	8.02
34485711	*RAB27A*	3.56	8.02	8.02	8.02	8.02	8.02	8.02	8.02	8.02	8.02	8.02	8.02	8.02
54607120	*LTF*	8.54	8.02	7.83	7.83	7.77	7.61	6.69	6.66	6.54	6.48	6.02	5.80	5.29
29788785	*TUBB*	3.21	4.00	3.81	3.81	3.75	3.58	2.66	2.64	2.51	2.46	2.00	1.78	1.26
4502101	*ANXA1*	16.17	3.72	3.53	3.53	3.47	3.31	2.39	2.36	2.24	2.18	1.72	1.50	0.99
5174735	*TUBB2C*	2.42	3.58	3.39	3.39	3.33	3.17	2.25	2.22	2.10	2.04	1.58	1.36	0.85
4502337	*AZGP1*	39.10	3.37	3.18	3.18	3.12	2.96	2.03	2.01	1.88	1.83	1.37	1.15	0.63
5803215	*UPK2*	20.54	3.26	3.07	3.07	3.01	2.85	1.93	1.90	1.78	1.72	1.26	1.04	0.53
5031635	*CFL1*	13.57	3.18	2.99	2.99	2.93	2.77	1.84	1.82	1.70	1.64	1.18	0.96	0.44
5902152	*UPK3A*	2.77	3.00	2.81	2.81	2.75	2.58	1.66	1.64	1.51	1.46	1.00	0.78	0.26
4507953	*YWHAZ*	5.75	2.83	2.64	2.64	2.58	2.42	1.49	1.47	1.34	1.29	0.83	0.61	0.09
10835049	*RHOA*	9.12	2.74	2.54	2.54	2.49	2.32	1.40	1.37	1.25	1.20	0.74	0.51	0.00
5902148	*UPK1A*	30.33	2.66	2.47	2.47	2.41	2.25	1.33	1.30	1.18	1.12	0.66	0.44	-0.07
12667788	*MYH9*	2.09	2.65	2.45	2.45	2.40	2.23	1.31	1.28	1.16	1.10	0.65	0.42	-0.09
194394237	*VIL1*	2.06	2.58	2.39	2.39	2.33	2.17	1.25	1.22	1.10	1.04	0.58	0.36	-0.15
189083772	*GSN*	2.76	2.42	2.22	2.22	2.17	2.00	1.08	1.05	0.93	0.87	0.42	0.19	-0.32
23065552	*GSTM3*	4.65	2.38	2.19	2.19	2.13	1.96	1.04	1.01	0.89	0.84	0.38	0.16	-0.36
32455266	*PRDX1*	4.17	2.36	2.17	2.17	2.11	1.95	1.03	1.00	0.88	0.82	0.36	0.14	-0.37
4557417	*CD14*	5.22	2.32	2.13	2.13	2.07	1.91	0.98	0.96	0.84	0.78	0.32	0.10	-0.42
161702986	*EZR*	14.07	2.30	2.11	2.11	2.05	1.89	0.97	0.94	0.82	0.76	0.30	0.08	-0.43
12025678	*ACTN4*	2.97	2.30	2.10	2.10	2.05	1.88	0.96	0.93	0.81	0.75	0.30	0.07	-0.44
4557317	*ANXA11*	2.23	2.15	1.96	1.96	1.90	1.74	0.81	0.79	0.67	0.61	0.15	-0.07	-0.58
148612842	*IFITM3*	4.37	2.00	1.81	1.81	1.75	1.58	0.66	0.64	0.51	0.46	0.00	-0.22	-0.74
4502211	*ARF6*	3.73	2.00	1.81	1.81	1.75	1.58	0.66	0.64	0.51	0.46	0.00	-0.22	-0.74
4757952	*CDC42*	4.68	2.00	1.81	1.81	1.75	1.58	0.66	0.64	0.51	0.46	0.00	-0.22	-0.74
9845511	*RAC1*	3.95	2.00	1.81	1.81	1.75	1.58	0.66	0.64	0.51	0.46	0.00	-0.22	-0.74
32189392	*PRDX2*	3.86	2.00	1.81	1.81	1.75	1.58	0.66	0.64	0.51	0.46	0.00	-0.22	-0.74
56682953	*EPS8*	2.95	2.00	1.81	1.81	1.75	1.58	0.66	0.64	0.51	0.46	0.00	-0.22	-0.74

Figure 1. Heat map representation of discrimination factors for candidate protein biomarker pairs in discrimination of (A) *All Rejection* versus *Tubular Injury*; (B) *Antibody-Mediated Rejection* versus *Tubular Injury*; and (C) *Cell-Mediated Rejection* versus *Tubular Injury*

Proteins for *All Rejection* were selected from the “immune response” list (supporting Information Table S3) and common proteins enriched in both *Antibody-Mediated Rejection* and *Cell-Mediated Rejection*. Proteins for *Antibody-Mediated Rejection* were selected from the “acute inflammatory response” list (Table 4), the “protein localization” list (supporting Information Table S5), and the “response to unfolded proteinâ list (supporting Information Table S6). Proteins for *Cell-Mediated Rejection* were selected from the “epithelial cell differentiation” list (Table 3) and the “actin filament based process” list (supporting

Information Table S4). Proteins for *Tubular Injury* were selected from the “sodium ion transport” list (Table 2). Discrimination factors (colored boxes) are $\log_2(R_X/R_{TI})$, where R_X is the median-normalized spectral count ratio of the two proteins for indicated Rejection group and R_{TI} is the ratio of the two proteins in *Tubular Injury*. Colors stratify discrimination factors from high (red) to low (green). The spectral counts shown for specific proteins represent the median-normalized spectral count values (|SC|) obtained in tubular injury (across top) or in appropriate rejection state (left); values for ‘all rejection’ represent average of cell-mediated and antibody-mediated values. Proteins are listed by both the GI accession number and the official gene symbol. Full annotation of the candidate biomarkers listed can be obtained using the GI accession number at the *Entrez Protein* online site (http://www.ncbi.nlm.nih.gov/protein).



Cell-Mediated Rejection

Antibody-Mediated Rejection

AMR vs. CMR		**GI No.**	26024195	4506773	4502101	5902148	4504165	5803215	7706683	5031635	4826898	23065552	10835049	4757952	161702986	4885049	55749523	89903012	5624352
		Gene	*CNFN*	*S100A9*	*ANXA1*	*UPK1A*	*GSN*	*UPK2*	*RHCG*	*CFL1*	*PFN1*	*GSTM3*	*RHOA*	*CDC42*	*EZR*	*ACTC1*	*SDCBP*	*CDC42*	*SDCBP*
GI No.	**Gene**	**| SC |**	19.92	16.78	16.17	30.33	12.78	20.54	5.06	13.57	12.71	4.65	9.12	4.68	14.07	26.13	8.62	9.94	4.08
166235903	*C1QC*	10.04	**8.16**	**7.95**	**6.63**	**5.95**	**5.30**	**5.12**	**5.02**	**4.70**	**4.62**	**4.59**	**4.56**	**4.38**	**4.33**	**4.01**	**3.95**	**3.70**	**3.06**
166197660	*C8G*	9.39	**8.16**	**7.95**	**6.63**	**5.95**	**5.30**	**5.12**	**5.02**	**4.70**	**4.62**	**4.59**	**4.56**	**4.38**	**4.33**	**4.01**	**3.95**	**3.70**	**3.06**
19923483	*RAB14*	5.96	**8.16**	**5.51**	**4.19**	**3.51**	**2.87**	**2.68**	**2.58**	**2.26**	**2.18**	**2.16**	**2.13**	**1.94**	**1.89**	**1.57**	**1.51**	**1.26**	**0.63**
34147513	*RAB7A*	11.14	**8.16**	**5.20**	**3.87**	**3.20**	**2.55**	**2.36**	**2.27**	**1.95**	**1.87**	**1.84**	**1.81**	**1.63**	**1.58**	**1.25**	**1.20**	**0.95**	**0.31**
50234889	*RAB43*	4.20	**8.16**	**5.10**	**3.77**	**3.10**	**2.45**	**2.26**	**2.17**	**1.85**	**1.77**	**1.74**	**1.71**	**1.53**	**1.48**	**1.15**	**1.10**	**0.85**	**0.21**
7706563	*RAB8B*	5.28	**8.16**	**5.01**	**3.69**	**3.01**	**2.37**	**2.18**	**2.09**	**1.77**	**1.68**	**1.66**	**1.63**	**1.44**	**1.40**	**1.07**	**1.01**	**0.77**	**0.13**
56243522	*SDCBP*	5.97	**8.16**	**4.89**	**3.56**	**2.89**	**2.24**	**2.05**	**1.96**	**1.64**	**1.55**	**1.53**	**1.50**	**1.32**	**1.27**	**0.94**	**0.89**	**0.64**	**0.00**
13569962	*RAB1B*	6.07	**8.16**	**4.89**	**3.56**	**2.89**	**2.24**	**2.05**	**1.96**	**1.64**	**1.55**	**1.53**	**1.50**	**1.32**	**1.27**	**0.94**	**0.89**	**0.64**	**0.00**
16933567	*RAB8A*	6.60	**8.16**	**4.83**	**3.51**	**2.83**	**2.19**	**2.00**	**1.91**	**1.58**	**1.50**	**1.48**	**1.45**	**1.26**	**1.22**	**0.89**	**0.83**	**0.58**	**-0.05**
5803135	*RAB35*	6.00	**8.16**	**4.73**	**3.41**	**2.73**	**2.09**	**1.90**	**1.81**	**1.49**	**1.40**	**1.38**	**1.35**	**1.16**	**1.12**	**0.79**	**0.73**	**0.49**	**-0.15**
4759000	*RAB3D*	6.91	**8.16**	**4.73**	**3.41**	**2.73**	**2.09**	**1.90**	**1.81**	**1.49**	**1.40**	**1.38**	**1.35**	**1.16**	**1.12**	**0.79**	**0.73**	**0.49**	**-0.15**
256222019	*RAB10*	5.37	**8.16**	**4.70**	**3.37**	**2.70**	**2.05**	**1.86**	**1.77**	**1.45**	**1.36**	**1.34**	**1.31**	**1.13**	**1.08**	**0.75**	**0.70**	**0.45**	**-0.19**
21464101	*YWHAG*	4.35	**8.16**	**4.66**	**3.34**	**2.66**	**2.02**	**1.83**	**1.74**	**1.42**	**1.33**	**1.31**	**1.28**	**1.09**	**1.05**	**0.72**	**0.66**	**0.42**	**-0.22**
4758988	*RAB1A*	6.10	**8.16**	**4.60**	**3.27**	**2.60**	**1.95**	**1.76**	**1.67**	**1.35**	**1.27**	**1.24**	**1.21**	**1.03**	**0.98**	**0.65**	**0.60**	**0.35**	**-0.29**
21264361	*MASP2*	54.86	**8.16**	**4.55**	**3.23**	**2.55**	**1.90**	**1.71**	**1.62**	**1.30**	**1.22**	**1.19**	**1.16**	**0.98**	**0.93**	**0.61**	**0.55**	**0.30**	**-0.34**
189163485	*CTSA*	4.65	**8.16**	**4.49**	**3.17**	**2.49**	**1.84**	**1.66**	**1.56**	**1.24**	**1.16**	**1.13**	**1.10**	**0.92**	**0.87**	**0.55**	**0.49**	**0.24**	**-0.40**
19923750	*RAB3B*	7.31	**8.16**	**4.35**	**3.03**	**2.35**	**1.70**	**1.51**	**1.42**	**1.10**	**1.02**	**0.99**	**0.96**	**0.78**	**0.73**	**0.41**	**0.35**	**0.10**	**-0.54**
24234686	*HSPA8*	8.13	**8.16**	**4.33**	**3.01**	**2.33**	**1.68**	**1.50**	**1.40**	**1.08**	**1.00**	**0.97**	**0.94**	**0.76**	**0.71**	**0.39**	**0.33**	**0.08**	**-0.56**
194248072	*HSPA1A*	4.18	**8.16**	**4.30**	**2.98**	**2.30**	**1.66**	**1.47**	**1.37**	**1.05**	**0.97**	**0.95**	**0.91**	**0.73**	**0.68**	**0.36**	**0.30**	**0.05**	**-0.58**
89903012	*CDC42*	9.34	**8.16**	**4.25**	**2.93**	**2.25**	**1.60**	**1.42**	**1.32**	**1.00**	**0.92**	**0.89**	**0.86**	**0.68**	**0.63**	**0.31**	**0.25**	**0.00**	**-0.64**
41393614	*RAB5C*	10.13	**8.16**	**4.18**	**2.86**	**2.18**	**1.53**	**1.34**	**1.25**	**0.93**	**0.85**	**0.82**	**0.79**	**0.61**	**0.56**	**0.24**	**0.18**	**-0.07**	**-0.71**
208973244	*YWHAZ*	5.40	**8.16**	**4.16**	**2.84**	**2.16**	**1.52**	**1.33**	**1.23**	**0.91**	**0.83**	**0.81**	**0.78**	**0.59**	**0.54**	**0.22**	**0.16**	**-0.09**	**-0.72**
126012573	*LRP2*	9.42	**8.16**	**4.15**	**2.83**	**2.15**	**1.50**	**1.32**	**1.22**	**0.90**	**0.82**	**0.79**	**0.76**	**0.58**	**0.53**	**0.21**	**0.15**	**-0.10**	**-0.74**
4557871	*TF*	97.27	**8.16**	**4.10**	**2.78**	**2.10**	**1.45**	**1.26**	**1.17**	**0.85**	**0.77**	**0.74**	**0.71**	**0.53**	**0.48**	**0.16**	**0.10**	**-0.15**	**-0.79**
55749523	*SDCBP*	6.82	**8.16**	**4.00**	**2.68**	**2.00**	**1.36**	**1.17**	**1.07**	**0.75**	**0.67**	**0.65**	**0.61**	**0.43**	**0.38**	**0.06**	**0.00**	**-0.25**	**-0.89**
66879662	*ARF1*	8.15	**8.16**	**3.97**	**2.65**	**1.97**	**1.32**	**1.13**	**1.04**	**0.72**	**0.64**	**0.61**	**0.58**	**0.40**	**0.35**	**0.03**	**-0.03**	**-0.28**	**-0.92**
22027538	*PDCD6IP*	8.92	**8.16**	**3.80**	**2.48**	**1.80**	**1.16**	**0.97**	**0.87**	**0.55**	**0.47**	**0.45**	**0.42**	**0.23**	**0.18**	**-0.14**	**-0.20**	**-0.45**	**-1.08**
5803225	*YWHAE*	5.04	**8.16**	**3.76**	**2.44**	**1.76**	**1.12**	**0.93**	**0.84**	**0.51**	**0.43**	**0.41**	**0.38**	**0.19**	**0.15**	**-0.18**	**-0.24**	**-0.49**	**-1.12**
161702986	*EZR*	8.54	**8.16**	**3.62**	**2.29**	**1.62**	**0.97**	**0.78**	**0.69**	**0.37**	**0.29**	**0.26**	**0.23**	**0.05**	**0.00**	**-0.33**	**-0.38**	**-0.63**	**-1.27**
262050538	*ITIH4*	9.66	**8.16**	**3.42**	**2.10**	**1.42**	**0.77**	**0.59**	**0.49**	**0.17**	**0.09**	**0.06**	**0.03**	**-0.15**	**-0.20**	**-0.52**	**-0.58**	**-0.83**	**-1.47**
14251209	*CLIC1*	6.88	**8.16**	**3.40**	**2.08**	**1.40**	**0.76**	**0.57**	**0.47**	**0.15**	**0.07**	**0.05**	**0.01**	**-0.17**	**-0.22**	**-0.54**	**-0.60**	**-0.85**	**-1.49**
4502205	*ARF4*	5.03	**8.16**	**3.34**	**2.02**	**1.34**	**0.70**	**0.51**	**0.42**	**0.09**	**0.01**	**-0.01**	**-0.04**	**-0.23**	**-0.28**	**-0.60**	**-0.66**	**-0.91**	**-1.54**
4757944	*CD81*	6.37	**8.16**	**3.08**	**1.76**	**1.08**	**0.43**	**0.25**	**0.15**	**-0.17**	**-0.25**	**-0.28**	**-0.31**	**-0.49**	**-0.54**	**-0.86**	**-0.92**	**-1.17**	**-1.81**
126091152	*CUBN*	10.37	**8.16**	**2.76**	**1.44**	**0.76**	**0.12**	**-0.07**	**-0.16**	**-0.48**	**-0.57**	**-0.59**	**-0.62**	**-0.81**	**-0.85**	**-1.18**	**-1.24**	**-1.48**	**-2.12**
73858570	*SERPING1*	4.93	**8.16**	**2.76**	**1.43**	**0.76**	**0.11**	**-0.08**	**-0.17**	**-0.49**	**-0.57**	**-0.60**	**-0.63**	**-0.81**	**-0.86**	**-1.19**	**-1.24**	**-1.49**	**-2.13**

Figure 2. Heat map representation of discrimination factors for candidate protein biomarker pairs in discrimination of *Antibody-Mediated Rejection* versus *Cell-Mediated Rejection*

Proteins for *Antibody-Mediated Rejection* were selected from the "acute inflammatory response" list (Table 4), the "protein localization" list (supporting Information Table S5), and the "response to unfolded protein" list (supporting Information Table S6). Proteins for *Cell-Mediated Rejection* were selected from the "epithelial cell differentiation" list (Table 3) and the "actin filament based process" list (supporting Information Table S4). Discrimination factors (colored boxes) are $\log_2(R_{AMR}/R_{CMR})$, where R_{AMR} is the median-normalized spectral count ratio of the two proteins for the *Antibody-Mediated Rejection* group and R_{CMR} is the ratio of the two proteins in the *Cell-Mediated Rejection* group. Colors stratify discrimination factors from high (yellow) to low (blue). The spectral counts shown for specific proteins represent the median-normalized spectral count values (|SC|) obtained in the appropriate rejection states. Proteins are listed by both the GI accession number and the official gene symbol. Full annotation of the candidate biomarkers listed can be obtained using the GI accession number at the *Entrez Protein* online site (http://www.ncbi.nlm.nih.gov/protein).

Table 1

Gene Ontology Biological Process terms over-represented in lists of exosomal proteins associated with different biopsy findings. Categories unique to a given group are shown in **Bold.**

	No. Proteins	p-value	FDR (%)
A. Tubular Injury			
GO:0007264 Small GTPase mediated signal transduction	64	0.0036	6.2
G0:0006814 Sodium ion transport	26	0.0050	8.6
GO:0010033 Response to organic substance	73	0.0073	12.4
B. Cell-Mediated Rejection			
GO:0030029 Actin filament-based process	47	0.0005	0.8
GO:0030855 Epithelial cell differentiation	21	0.0018	3.2
GO:0007264 Small GTPase mediated signal transduction	62	0.0021	3.7
GO:0030036 Actin cytoskeleton organization	41	0.0026	4.5
GO:0007010 Cytoskeleton organization	63	0.0031	5.5
GO:0006955 Immune response	49	0.0118	19.1
GO:0043603 Cellular amide metabolic process	16	0.0118	19.2
C. Antibody-Mediated Rejection			
GO:0007264 Small GTPase mediated signal transduction	64	<0.0000	0.0
GO:0015031 Protein transport	97	<0.0000	0.1
GO:0045184 Establishment of protein localization	97	0.0001	0.1
GO:0008104 Protein localization	103	0.0003	0.5
GO:0010033 Response to organic substance	64	0.0018	3.1
GO:0007242 Intracellular signaling cascade	105	0.0034	6.0
GO:0044057 Regulation of system process	26	0.0039	6.7
GO:0006986 Response to unfolded protein	13	0.0063	10.6
GO:0002526 Acute inflammatory response	18	0.0066	11.1
GO:0010035 Response to inorganic substance	28	0.0083	13.7
GO:0051605 Protein maturation by peptide bond cleavage	15	0.0086	14.2
GO:0006955 Immune response	44	0.0094	15.5

Table 2

Protein list for Gene Ontology Biological Process ***“sodium ion transport”*** enriched in urinary exosomes from transplants with Tubular Injury compared with all transplants. Asterisks indicate proteins that were expressed with relatively high signal values in proximal tubule based on the transcriptomic databases of proximal tubule, medullary thick ascending limb, and inner medullary collecting duct [14], [15].

Protein Name	**GI Accession**	**Gene Symbol**
sodium/potassium-transporting ATPase subunit alpha-1 isoform a *	21361181	*ATP1A1*
sodium/potassium-transporting ATPase subunit beta-1 *	4502277	*ATP1B1*
sodium/potassium-transporting ATPase subunit gamma isoform 2 *	11125764	*FXYD2*
solute carrier family 12 member 3 isoform 1	186910315	*SLC12A3*
solute carrier family 12 member 3 isoform 3	186910319	*SLC12A3*
solute carrier family 12 member 1 isoform A	134254459	*SLC12A1*
solute carrier family 12 member 2 *	4506975	*SLC12A2*
solute carrier family 13 member 1 *	19923838	*SLC13A1*
solute carrier family 13 member 2 isoform b *	4506979	*SLC13A2*
solute carrier family 13 member 2 isoform c *	225637552	*SLC13A2*
solute carrier family 13 member 3 isoform a *	31377715	*SLC13A3*
solute carrier family 13 member 3 isoform b *	58761541	*SLC13A3*
sodium-dependent phosphate transport protein 1 *	154800443	*SLC17A1*
sodium-dependent phosphate transport protein 4 isoform a *	148743791	*SLC17A3*
solute carrier family 22 member 5 *	4507005	*SLC22A5*
solute carrier family 22 member 4 *	24497490	*SLC22A4*
solute carrier family 23 member 1 isoform a *	44680145	*SLC23A1*
sodium-dependent phosphate transport protein 2B isoform a	110611906	*SLC34A2*
sodium-coupled neutral amino acid transporter 2 *	21361602	*SLC38A2*
electrogenic sodium bicarbonate cotransporter 1 isoform 2 *	4507025	*SLC4A4*
sodium-coupled monocarboxylate transporter 1 *	167466278	*SLC5A8*
sodium/glucose cotransporter 1 *	4507031	*SLC5A1*
sodium/glucose cotransporter 5 isoform 2 *	109659836	*SLC5A10*
sodium/myo-inositol cotransporter 2 *	17941285	*SLC5A11*
sodium-coupled monocarboxylate transporter 2 *	157671931	*SLC5A12*
sodium/glucose cotransporter 2 *	4507033	*SLC5A2*
sodium/glucose cotransporter 4 isoform 2 *	206597483	*SLC5A9*
sodium- and chloride-dependent creatine transporter 1 isoform 3 *	218563758	*SLC6A8*
sodium/hydrogen exchanger 3 *	194239733	*SLC9A3*

Table 3

Protein list for Gene Ontology Biological Process ***"epithelial cell differentiation"*** enriched in urinary exosomes from transplants with Cell-Mediated Rejection compared with all transplants.

Protein Name	**GI Accession**	**Gene Symbol**
ammonium transporter Rh type C	7706683	*RHCG*
annexin A1	4502101	*ANXA1*
cornifelin	26024195	*CNFN*
desmoplakin isoform I	58530840	*DSP*
envoplakin	156104874	*EVPL*
ezrin	161702986	*EZR*
filaggrin	60097902	*FLG*
glutathione S-transferase Mu 3	23065552	*GSTM3*
hornerin	57864582	*HRNR*
periplakin	45439327	*PPL*
protein-glutamine gamma-glutamyltransferase K	4507475	*TGM1*
tumor susceptibility gene 101 protein	5454140	*TSG101*
uroplakin-1a	5902148	*UPK1A*
uroplakin-1b	49619237	*UPK1B*
uroplakin-2 precursor	5803215	*UPK2*
uroplakin-3a isoform 1	5902152	*UPK3A*

Table 4

Protein list for Gene Ontology Biological Process ***“acute inflammatory response”*** enriched in urinary exosomes from transplanted kidneys with Antibody-Mediated Rejection compared with all transplants.

Protein Name	GI Accession	Gene Symbol
14-3-3 protein zeta/delta	208973244	*YWHAZ*
alpha-1-antichymotrypsin precursor	50659080	*SERPINA3*
C4b-binding protein alpha chain precursor	4502503	*C4BPA*
complement C1q subcomponent subunit B precursor	87298828	*C1QB*
complement C1q subcomponent subunit C precursor	166235903	*C1QC*
complement component C7 precursor	45580688	*C7*
complement component C8 gamma chain precursor	166197660	*C8G*
complement component C9 precursor	4502511	*C9*
complement factor B preproprotein	67782358	*CFB*
complement receptor type 1 isoform F precursor	86793036	*CR1*
complement receptor type 1 isoform S precursor	86793109	*CR1*
inter-alpha-trypsin inhibitor heavy chain H4 isoform 1 precursor	31542984	*ITIH4*
inter-alpha-trypsin inhibitor heavy chain H4 isoform 2 precursor	262050538	*ITIH4*
mannan-binding lectin serine protease 2 isoform 1 precursor	21264363	*MASP2*
mannan-binding lectin serine protease 2 isoform 2 precursor	21264361	*MASP2*
peroxiredoxin-2 isoform a	32189392	*PRDX2*
plasma protease C1 inhibitor precursor	73858570	*SERPING1*
PREDICTED: complement C3-like, partial	169218213	*LOC100133511*
serotransferrin precursor	4557871	*TF*
tuberin isoform 4	116256350	*TSC2*

Table 5

KEGG pathways over-represented in lists of urinary exosomal proteins associated with different biopsy findings. Pathways unique to a histological category are indicated in **Bold.**

	No. Proteins	p-value	FDR (%)
A. Tubular Injury			
none			
B. Cell-Mediated Rejection			
hsa04530:Tight junction	28	0.01	16.62
hsa04144:Endocytosis	40	0.02	19.59
hsa00480:Glutathione metabolism	16	0.06	53.13
C. Antibody-Mediated Rejection			
hsa04612:Antigen processing and presentation	15	<0.01	4.20
hsa04722:Neurotrophin signaling pathway	22	0.01	6.73
hsa05200:Pathways in cancer	25	0.01	11.71
hsa05130:Pathogenic Escherichia coli infection	20	0.01	14.04
hsa04144:Endocytosis	36	0.01	16.30

Table 6

List of proteins identified in urinary exosomes of patients in Antibody-Mediated Rejection group that are present in KEGG pathway ***"antigen processing and presentation".***

Protein Name	GI Accession	Gene Symbol
calreticulin precursor	4757900	*CALR*
cathepsin S isoform 1 preproprotein	23110962	*CTSS*
HLA class I histocompatibility antigen, A-1 alpha chain precursor	24797067	*HLA-A*
major histocompatibility complex, class I, B precursor	17986001	*HLA-B*
major histocompatibility complex, class I, B precursor	17986001	*HLA-B*
HLA class II histocompatibility antigen, DM alpha chain precursor	18765715	*HLA-DMA*
HLA class II histocompatibility antigen, DM beta chain precursor	4504399	*HLA-DMB*
HLA class II histocompatibility antigen, DR alpha chain precursor	52426774	*HLA-DRA*
heat shock protein HSP 90-alpha isoform 2	154146191	*HSP90AA1*
heat shock protein HSP 90-alpha isoform 2	154146191	*HSP90AA1*
heat shock protein HSP 90-beta	20149594	*HSP90AB1*
heat shock 70 kDa protein 1A/1B	194248072	*HSPA1A*
heat shock 70 kDa protein 1A/1B	194248072	*HSPA1A*
heat shock 70 kDa protein 1-like	124256496	*HSPA1L*
heat shock-related 70 kDa protein 2	13676857	*HSPA2*
heat shock 70 kDa protein 6	34419635	*HSPA6*
heat shock 70 kDa protein 6	34419635	*HSPA6*
heat shock cognate 71 kDa protein isoform 2	24234686	*HSPA8*
proteasome activator complex subunit 1 isoform 1	5453990	*PSME1*