Urine Proteomics to Detect Biomarkers for Chronic Allograft Dysfunction

Luís F. Quintana,*† Amanda Solé-Gonzalez,† Susana G. Kalko,‡ Elisenda Bañon-Maneus,† Manel Solé,§ Fritz Diekmann,†‡ Alex Gutierrez-Dalmau,† Joaquín Abian,¶ and Josep M. Campistol*†

*Servicio de Nefrología y Trasplante renal, †Laboratorio Experimental de Nefrología y Trasplante renal, ‡Unidad de Bioinformática, §Servicio Anatomía Patológica, Hospital Clinic, Institut d’Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Universitat de Barcelona, and ¶Laboratorio de Proteómica Consejo Superior de Investigaciones Científicas (CSIC)/Universidad Autónoma, Instituto de Investigacions Biomèdiques de Barcelona/CSIC-Institut d’Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain; and Department of Nephrology, Charité Campus Mitte, Berlin, Germany

ABSTRACT

Despite optimal immunosuppressive therapy, more than 50% of kidney transplants fail because of chronic allograft dysfunction. A noninvasive means to diagnose chronic allograft dysfunction may allow earlier interventions that could improve graft half-life. In this proof-of-concept study, we used mass spectrometry to analyze differences in the urinary polypeptide patterns of 32 patients with chronic allograft dysfunction (14 with pure interstitial fibrosis and tubular atrophy and 18 with chronic active antibody-mediated rejection) and 18 control subjects (eight stable recipients and 10 healthy control subjects). Unsupervised hierarchical clustering showed good segregation of samples in groups corresponding mainly to the four biomedical conditions. Moreover, the composition of the proteome of the pure interstitial fibrosis and tubular atrophy group differed from that of the chronic active antibody-mediated rejection group, and an independent validation set confirmed these results. The 14 protein ions that best discriminated between these two groups correctly identified 100% of the patients with pure interstitial fibrosis and tubular atrophy and 100% of the patients with chronic active antibody-mediated rejection. In summary, this study establishes a pattern for two histologic lesions associated with distinct graft outcomes and constitutes a first step to designing a specific, noninvasive diagnostic tool for chronic allograft dysfunction.


During the past three decades, the incidence and prevalence of ESRD has increased each year all over the world.1 Kidney transplantation is the treatment of choice for ESRD because it prolongs survival,2 improves quality of life, and is less costly than dialysis;3 however, despite these improvements, a substantial proportion of grafts develop progressive dysfunction and fail within a decade, even with the use of appropriate dosages of immunosuppressive drugs to prevent acute rejection.4 Chronic allograft dysfunction (CAD) causes more than 50% of graft losses.5,6 Although patients can return to dialysis after transplant failure, loss of a functioning graft is associated with a three-fold increase in the risk for death,7,8 a substantial decrease in quality of life in survivors, and a four-fold increase in cost.1,3

The decline in function, often associated with hypertension and proteinuria, constitutes a clinical
syndrome that has been called chronic allograft nephropathy (CAN). The histopathologic hallmarks of these patients are chronic interstitial fibrosis, tubular atrophy, vascular occlusive changes, and glomerulosclerosis, usually evaluated by the Banff working classification.10 Major outcomes discussed at the last Banff Conference included the elimination of the non-specific term CAN and recognition of the entity “chronic active antibody-mediated rejection” (CAAR).11 The rationale for this update was the improper use of “CAN” as a generic term for all causes of chronic renal allograft dysfunction with interstitial fibrosis and tubular atrophy (IF/TA), which hampers accurate diagnosis and appropriate therapy, and increasing recognition of the role of alloantibody in chronic renal allograft deterioration and the corresponding histologic changes, making the identification of an antibody-mediated component of chronic rejection feasible.11

Effective strategies to prevent renal function deterioration should focus on the early detection and treatment of patients who develop CAD. In addition to elevated serum creatinine, usually associated with proteinuria and arterial hypertension, more specific and sensitive markers are needed to identify high-risk patients or initial lesions without any changes in serum creatinine or proteinuria.1,11

New analytic tools that allow rapid screening and accurate protein identification in body fluids are now emerging within the field of proteomic science. High-throughput mass spectrometry (MS) methods allow simultaneous detection of a large number of proteins in a large set of biologic tissues or samples. Protein fingerprinting MS methods using modern matrix-assisted laser desorption/ionization-time of-flight MS (MALDI-MS) instrumentation can detect hundreds of peak signals that, as a whole, could be considered a reflex of the body’s physiologic status.12 To date, MALDI-MS has been successfully used to detect patterns of substantial overexpression of proteins in cancer cells.13–15 Urine seems to be an ideal source of potential biomarkers, and urine proteomic approaches have been used in numerous attempts to define biomarkers for a variety of nephro-urologic disorders.16–18 The aim of this study was to evaluate whether chromatography by solid-phase extraction coupled to MS would differentiate urinary polypeptide patterns in patients with pure IF/TA, patients with CAAR, and two control groups: Healthy individuals and stable renal transplant recipients.

**RESULTS**

**Clinical and Histologic Characteristics of Patients with CAD**

The analysis included 50 individuals: 32 patients with CAD (eight in training set and six in validation set with IF/TA with no other cause and 10 in training set and eight in validation set with CAAR) and 18 control subjects (10 healthy individuals and eight stable renal transplant recipients). Table 1 shows the baseline characteristic of patients with CAD and control subjects.

There was no evidence of CAAR or transplant glomerulopathy (TG), and C4d was negative in all patients with pure IF/TA (G1). Mean glomerular double contour (CG) score was 1.89, and C4d was positive in all patients in the CAAR group (G2). Evidence of chronic active T cell-mediated rejection was excluded in all samples from this group. Table 2 summarizes Banff scores in the IF/TA and CAAR groups. Table 3 shows the HLA of both recipients and donors and the status of circulating donor-specific anti-HLA antibodies in patients with CAAR and with the HLA analysis in the prebiopsy period. All patients had available HLA matching data; none was HLA identical. The mean HLA-A-B-DR mismatches was 4.16 (minimum 2, maximum 6), and 100% of patients with CAAR had donor-specific antibody (DSA) by FLOWPRA Specific Antibody Detection Test (One Lambda, Canoga Park, CA).

**Urinary Proteome of Healthy Individuals Compared with Those of Stable Transplant Recipients and Patients with CAD**

We found some differences between control subjects and patients with CAD. Patients with CAD (IF/TA and CAAR) always had peak clusters in three regions corresponding to mass/charge (m/z) values of 2628 to 2922, 4307 to 4799, and 8303 to 8850, whereas the healthy urine protein profile had no peak clusters in these m/z regions. All urine samples from healthy individuals showed a peak m/z of 9754. Subsequently, an informatics analysis was performed in a blind and systematic manner as cases were added. (All data containing the peak intensities and molecular weights in all patients are provided online as supplemental information.)

Unsupervised hierarchical clustering using the 2000 proteins with the largest variability in the whole set of spectrums in

---

**Table 1.** Clinical characteristic of study cohorts and controls groups

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>IF/TA Group (n = 14)</th>
<th>CAAR Group (n = 18)</th>
<th>Stable Renal Transplant Recipients (n = 8)</th>
<th>Healthy Control Subjects (n = 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training Set (n = 8)</td>
<td>Validation Set (n = 6)</td>
<td>Training Set (n = 10)</td>
<td>Validation Set (n = 8)</td>
</tr>
<tr>
<td>Age (yr)</td>
<td>51.00 ± 10.69</td>
<td>49.90 ± 12.09</td>
<td>47.22 ± 17.07b</td>
<td>51.00 ± 15.20</td>
</tr>
<tr>
<td>Creatinine (mg/dl)</td>
<td>3.20 ± 1.68</td>
<td>3.10 ± 1.48</td>
<td>2.98 ± 1.64b</td>
<td>2.80 ± 0.82</td>
</tr>
<tr>
<td>Proteinuria (g/24 h)</td>
<td>2.67 ± 2.90</td>
<td>2.63 ± 2.70</td>
<td>3.11 ± 3.33b</td>
<td>3.29 ± 2.96</td>
</tr>
<tr>
<td>GFR (ml/min per 1.73 m²)</td>
<td>28.88 ± 17.65</td>
<td>29.78 ± 15.60</td>
<td>33.44 ± 12.05b</td>
<td>27.90 ± 10.10</td>
</tr>
</tbody>
</table>

aData are means ± SD.

b T test NS between IF/TA and CAAR groups.
Diffusional Expression Analysis of Patients with CAD

In the unsupervised cluster analysis, patients with CAAR were grouped together, separately from patients with IF/TA. These results suggest that the composition of the urine proteome of pure IF/TA is subtly different from that of the CAAR urine proteome. Subsequently, we tried to determine the differences in the urine proteomic profile among patients with pure IF/TA and those with CAAR.

Differential expression analysis in a first set of samples (training set) showed statistically significant differences between the IF/TA and CAAR groups; these differences identified possible biomarkers that might confidently be used to distinguish between these two entities. Using the significance analysis of microarrays (SAM) module in MeV, we reported 366 m/z values with fold change (CAAR and IF/TA) < 0.5 and a significance of q < 0.01 (q = false discovery rate). Mass proteins with fold change < 0.5 in the m/z range 2850 to 3050 are shown in Figure 2.

To validate protein signature obtained in the training set, we considered an independent set of samples (validation set) from IF/TA and CAAR groups, equivalent to the training set. Unsupervised clustering of the validation set using the 366 m/z observed in the previous SAM analysis on the training set shows very good segregation of this new set in the two groups (Figure 3). Furthermore, we analyzed the 366 selected protein ions using discriminant analysis. The protein ions that best discriminated between the two groups were 1529.22, 1529.45, 1539.8, 1540.03, 1541.87, 1542.1, 1574.79, 1575.02, 1575.25, 1575.48, 1575.72, 1587.86, 1657.4, and 6146.92, resulting in correct identification of 100% (eight of eight) of the pure IF/TA group and 100% (10 of 10) of the CAAR group. Logistic regression (LR) analysis resulted in selection of 1539.8, 1540.03, 1542.1, 1575.48, 1587.86, and 1657.4. A receiver operating characteristic curve constructed from the logistic regression scores gave a high area under the curve value of 1.0 (Figure 4). This distinct protein signature is a suitable model for further identification of biomarker candidates.

**DISCUSSION**

Like surface enhanced laser desorption/ionization technology, direct polypeptide profiling by solid-phase extraction and
MALDI-MS reduces the complexity of a sample by selective adsorption of proteins to an active surface and is a relatively simple proteomic approach that allows rapid differential diagnosis of patients and information transfer between the laboratory and the clinical context. The processes that result in CAD are poorly understood. Although determination of organ failure often relies on measurable physiologic parameters, the early stages of CAD are difficult to diagnose. Detection requires invasive procedures to obtain graft biopsies for histologic evaluation. To our knowledge, this is the first preliminary report on this broad-based, unbiased, noninvasive proteomic and bioinformatics approach as a means to design a noninvasive diagnostic tool for CAD.

The large quantity of protein data generated requires complementary bioinformatics tools to identify differential patterns of value, and a mathematical model was used to distinguish between the overlapping sets of results. To avoid biases and misinterpretation between the CAAR and IF/TA groups, the samples analyzed were taken from patients with similar follow-up after transplantation, the same immunosuppressive treatment, similar renal function, and similar Banff score on IF/TA. The only difference between the two groups was the presence of histologic markers of CAAR and HLA DSA in the CAAR group. Allergic drug reactions, systemic or intragraft infections, and other inflammatory processes were reasonably excluded.

Initially, comparative bioinformatics urine proteomic analysis detected differences among healthy individuals, stable transplant recipients, and patients with IF/TA and CAAR, showing an excellent clinical correlation. Importantly, the presence of IF/TA is another criterion required for a diagnosis of CAAR, and the absence of differences in the IF/TA scores of the two groups is a plausible explanation for the finding that both groups (two CAAR samples were interchangeable with two IF/TA samples) share some mass values. The same phenomenon occurred in two stable renal transplant samples and two samples from healthy individuals, indicating that these groups have some urine proteins in common.

When we analyzed our two groups of patients with CAD in greater detail, the differential expression analysis confirmed differences between patients with pure IF/TA and patients with IF/TA plus CAAR. The mass range analysis highlighted these differences, whereas comparison with traditional markers (creatinine and proteinuria) of CAD demonstrated that our markers had discriminatory power. Moreover, the independent and equivalent validation set confirms the results from the training set.

The presence of IF/TA in protocol renal allograft biopsies is an independent predictor of graft survival. Recently, IF/TA associated with transplant vasculopathy, subclinical rejection, and TG was shown to be associated with a poorer outcome and prognosis than IF/TA without additional lesions (pure IF/TA). Although histologic lesions predict graft survival, their accuracy has been poorly characterized. Preliminary data suggest that their predictive value in graft survival is at least not inferior to that of acute rejection or renal function. In addition, quantitative evaluation of histologic lesions has been performed in an attempt to provide a better surrogate marker for graft half-life. Both morphometry and image analysis have been used to quantify interstitial fibrosis and transplant vasculopathy, but neither of these parameters has been

Figure 1. (A) Detail of unsupervised hierarchical clustering using the 2000 m/z features with the largest variability in the entire set of samples (HCL* module of MeV package). (B) Matrix representation of the distances (inverse of the similarity) between samples, using the same features as those used in A (GDM module of MeV package).
properly validated. An alternative strategy to increase the predictive value of protocol biopsies in graft outcome was validation with composite end points containing histologic, clinical, biologic, and analytical parameters.28

This study reveals a differential link between specific clinical factors, histologic findings, and urine proteome among patients with CAAR and those with IF/TA without additional lesions. Despite the limited number of patients analyzed, the unsupervised cluster analysis differentiated between patients with pure IF/TA and those with CAAR (IF/TA + CAAR), supporting the importance of the recent update of the Banff working classification of CAD.

In view of the complex relationship among clinical and biologic CAD factors and among early histopathologic changes and graft outcome, the search for a urine biomarker panel to identify the various causes of CAD seems absolutely necessary to improve long-term graft survival. Moreover, the different urine protein excretion patterns in patients with pure IF/TA versus those with CAAR detected in this study will support the identification of different urine biomarkers for these two histologic lesions with distinct outcomes; notably, the CAD pattern had peak clusters in three regions with high linkage, which could be used in investigation of urinary markers of graft fibrosis. The combination of protocol graft biopsies with simultaneous proteomic analyses at different times after renal transplantation could provide more accurate information on the exact mechanisms involved in the development of CAD. This information could be very helpful not only for an early diagnosis of CAD but also in the treatment and prevention of the leading cause of graft loss.

Additional studies including a larger and more diverse transplant recipient population are required to confirm our data. Reproduction of these patterns in future studies would represent a step forward in identifying each protein and in translating this relatively intricate procedure into a biochemical test of clinical utility.

CONCISE METHODS

Study Cohorts
Fifty individuals were included in the study: 32 patients with clinical and histopathologic characterization of CAD and 18 control subjects. The patients were divided in two groups: (1) 14 patients (nine men and five women) with IF/TA and no evidence of any specific cause (IF/TA group) and (2) 18 patients (12 men and six women) with CAAR defined by morphologic features including TG and IF/TA with or without peritubular capillary loss, and fibrous intimal thickening in arteries without duplication of the internal elastica, diffuse C4d
deposition in peritubular capillaries (PTC), and the presence of DSA (CAAR group).

All transplant recipients received immunosuppressive treatment with a calcineurin inhibitor, mycophenolate mofetil, and prednisone. There were no significant differences between the IF/TA and CAAR groups with respect to age, gender, diabetes duration, arterial BP, body mass index, and GFR (Table 1).

The control subjects were divided in two groups: (1) Stable renal transplant recipients, who were eight live-donor recipients of a first renal graft at the end of the first month after surgery with normal pretransplantation biopsy and after immunosuppressive treatment with tacrolimus, mycophenolate mofetil, and prednisone (mean age 36.2 ± 8.0, mean serum creatinine 1.08 ± 0.30 mg/dl, and mean proteinuria 205.2 ± 50.0 mg/d) and (2) healthy control subjects, who were 10 volunteers with normal BP (systolic BP <130 mmHg and diastolic BP <80 mmHg) and no history of diabetes, ischemic heart disease, stroke, or peripheral vascular disease (mean age 43 ± 10 yr; seven men and three women, mean creatinine 0.91 ± 0.30 mg/dl, GFR 110 ± 10 ml/min, mean microalbuminuria/creatinuria 4 ± 1 mg/g, mean proteinuria/creatinuria 39 ± 6 mg/g). The study was approved by the institutional review board at the Hospital Clinic in Barcelona, and both patients and control subjects gave informed consent for the collection and analysis of their urine.

Histopathology
Transplant biopsies consisted of two cores obtained with 18-gauge needles using ultrasound guidance because of clinical indication. Paraffin sections were prepared and stained with hematoxylin-eosin, trichrome, periodic acid-Schiff, and periodic acid-Schiff–methenamine silver.

The biopsies were analyzed and were scored according to the Banff classification by a pathologist who was blinded to the results of molecular studies. TG was diagnosed by light microscopy on the basis of double contours of glomerular basement membranes and was analyzed on a BD FACSCalibur cytometer. Antibody specificities to class I and class II antibodies by FlowPRA. Beads were specificities to FlowPRA Single Antigen I and II beads (One Lambda). Beads were specificities determined by FlowPRA Specific class I and or II and class I and class II antibodies by FlowPRA. When positive, antibody Anti-HLA Antibody Analysis
Recipient sera taken in the prebiopsy period were screened for HLA class I and class II antibodies by FlowPRA. When positive, antibody specificities were determined by FlowPRA Specific class I and or II and or FlowPRA Single Antigen I and II beads (One Lambda). Beads were analyzed on a BD FACSCalibur cytometer. Antibody specificities to HLA-A, B, DRb1, DRb3, DRb4, DRb5, DQB1 were evaluated.

Sample Preparation and Purification
Fifty milliliters of early morning urine was collected immediately before renal biopsy, adding protease inhibitor cocktail (Complete Mini; Roche, Mannheim, Germany); rapidly frozen in dry ice; and stored at −80°C until analyzed. Urine samples were concentrated and separated from organic salts by solid-phase extraction using a reverse-phase HLB Oasis 94226 (Waters, Milford, MA) as the stationary phase and according to the following protocol: Conditioning in 10 ml of 100% acetonitrile (ACN), equilibration in 10 ml of 0.1% trifluoroacetic acid (TFA)/5% ACN, loading of sample (acidified with 0.1% TFA at pH 3 and final concentration of 5% ACN), washing with 10 ml 0.1% TFA/5% CAN, and elution with 2 ml of 0.1% TFA/60% ACN.

Each sample was applied and dried on an uncoated MALDI target plate using the sandwich technique (1 μl of sample plus 2 μl of matrix). A saturated solution of 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid; Sigma, St. Louis, MO) in ACN/H2O/TFA 50/50/0.3 was used as matrix. Ten spots were prepared and analyzed per sample.

Mass Spectrometry Data Acquisition
Mass spectra were acquired with a Voyager-DE PRO TM Biospectrometry Workstation (Applied Biosystems, Foster City, CA), and data were obtained by using the linear acquisition mode under delayed extraction conditions. Instrument settings were an accelerating voltage of 25 kV, 91% grid voltage, 0.05% guidewire voltage, delay time of 220 ns, and bin size of 2 ns. Three spectra (three acquisitions of 150 shots each) were acquired for each replicate. Internal calibration standards were insulin bovine 0.5 pmol/μl, thioredoxin from Escherichia coli 2.75 pmol/μl, and apomyoglobin from horse 4 pmol/μl (standard mixture 3 from Applied Biosystems). Deposition of the samples on the MALDI target and mass spectra acquisition were performed randomly.

Data Analysis
Baseline Correction.
The chemical noise present in each spectrum was estimated by determining the minimal measured intensity value in successive 100 m/z windows. A function fit to these minima by least squares was defined as baseline and subtracted from the spectrum.

Binning of Peaks.
Data were analyzed for the molecular weight range of 1.0 to 20 kD. The three spectra from each replicate were averaged arithmetically, and automated peak detection was carried out on these averaged spectra in Data Explorer 4.3 software, which determines the centroid mass, height for the 50th percentile of the lower and upper boundary of the peak, and area and signal-to-noise ratio for each peak (signal-to-noise threshold of 2, a 5-kD to define peak width).

Preprocessing of MS Data
For the comparison of the peaks included in different spectra, peak alignment was performed with an agglomerative hierarchical clustering algorithm. The algorithm initially creates as many clusters as detected peaks in all spectra and successively merges the closest clusters if all of the peaks of the new cluster belong to different spectra and
the maximal mass distance between two peaks of the new cluster is \(<0.2\%\) of their mean. This tolerance corresponds to a measurement error of 0.1\%.\(^{22}\)

**Bioinformatics and Statistical Analysis**

Unsupervised bidimensional cluster analysis of features showing wide variability was performed and displayed using the MeV package in the TM4 software suite.\(^{20}\)

Statistical differential analysis was performed for the signal intensities of the peaks detected. The SAM method was used to identify significant differentially expressed features while controlling for the false discovery rate.\(^{19}\)

In high-throughput experiments, large number of features may display change with ordinary levels of significance \((i.e., \, P < 0.05)\) even when no biologic differences exist. One approach to surmount this problem in multiple comparisons is to identify the false discovery rate \((q\) value). A valuable feature of SAM analysis is that it gives estimates of the false discovery rate, which is the proportion of features likely to have been identified by chance as being significant.

The protein ions were then analyzed using discriminant analysis \((\text{Systat} 10.2; \text{Richmond, CA})\) to identify combinations of proteins that best discriminate between disease states. A logistic regression model was also built using the same protein ions to calculate prediction scores for each sample, allowing us to construct a receiver operating characteristic curve based on these values.\(^{31}\)

**ACKNOWLEDGMENTS**

This study was supported by grants from the *Fondo de Investigaciones Sanitarias*, FIS 03/0557 and FIS 06/1222, and an Emili Letang Fellowship-Sanitarias

**DISCLOSURES**

None.

**REFERENCES**


Supplemental information for this article is available online at http://www.jasn.org/.