RESEARCH ARTICLE

Identification of a protein signature in renal allograft rejection

Diane M. Cibrik[1], Roscoe L. Warner[2], Mallika Kommareddi[1], Peter Song[3], Fu L. Luan[1] and Kent J. Johnson[2]

[1] Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA
[2] Department of Pathology, University of Michigan, Ann Arbor, MI, USA
[3] Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA

Purpose: Serum creatinine functions as a poor surrogate marker of renal allograft dysfunction and long-term graft survival. By measuring multiple proteins simultaneously in the serum of transplant patients, we can identify unique protein signatures of graft dysfunction.

Experimental design: We utilized training and validation cohorts composed of healthy and volunteer subjects, stable renal transplant patients, and renal transplant patients experiencing acute allograft rejection. Utilizing our antibody microarray, we measured 108 proteins simultaneously in these groups.

Results: Using Mann–Whitney tests with Bonferroni correction, we identified ten serum proteins from 19 renal transplant patients with stable renal function, which are differentially expressed, compared to healthy control subjects. In addition, we identified 17 proteins that differentiate rejecting renal transplant recipients from stable renal transplant. Validation cohorts substantiated these findings.

Conclusion and clinical relevance: Our preliminary results support that a specific pattern of protein expression or "protein signature" may be able to differentiate between stable transplant patients from those with rejection. Future studies will focus on other etiologies of renal allograft dysfunction and the effect of treatment on protein expression and long-term outcome.

Correspondence: Dr. Diane M. Cibrik, 3914 Taubman Health Center, 1500 E. Medical Center Dr., Ann Arbor, MI 48109, USA
E-mail: dcibrik@umich.edu
Fax: 734-232-4160

Abbreviations: ATN, acute tubular necrosis; **BKV**, BK virus; **CAN**, chronic allograft nephropathy; **CNI**, calcineurin inhibitor; **EGF**, epidermal growth factor; **GM-CSF**, granulocyte-macrophage colony-stimulating factor; **GROα**, growth regulated oncogene-alpha; **HSP27**, heat shock protein 27; **IL**, interleukin; **IL-1 R1**, IL-1 receptor 1; **KIM-1**, kidney injury molecule 1; **LARS**, least angle regression analysis; **LDA**, linear discriminant analysis; **MCP-1**, monocyte chemotactic protein 1; **MFI**, median fluorescent intensity; **MIF**, macrophage migration inhibitory factor; **MIP-3α**, macrophage inflammatory protein 3-alpha; **ROC**, receiver-operator curve; **RT**, room temperature; **Scr**, serum creatinine; **TGF**, transforming growth factor; **TIMP-4**, tissue inhibitor matrix metalloproteinase 4; **TNF-RII**, tumor necrosis factor receptor II; **VEGF-R2**, vascular endothelial growth factor receptor 2

1 Introduction

Although one year kidney graft survival is greater than 90% due to improved immunosuppression, long-term kidney allograft outcome has not changed dramatically over the last decade [1]. Several factors that influence long-term kidney transplant function include immunologic injury from either acute cellular or antibody-mediated rejection, calcineurin inhibitor (CNI) nephrotoxicity, BK virus (BKV) associated nephropathy, and metabolic derangements, all of which result in chronic allograft injury and subsequent loss of the transplanted graft [2]. The long-term success of kidney transplantation traditionally relies on serum creatinine (Scr) levels as the surrogate marker to follow renal allograft function and graft survival [3]. The early and prompt detection of renal allograft dysfunction and understanding of its associated mechanisms may improve the long-term outcome. Renal allograft biopsy, though imperfect and invasive, is the current gold standard to achieve such objective.

For years, investigators have been searching for alternative, noninvasive, and easy to use methods to detect

renal allograft dysfunction. Initial studies included blood and urine flow cytometry [4, 5], and over the years, enzyme-linked immunoassays [6–9], NMR [10], mass spectroscopy [11, 12], PCR [13–17], and gene microarrays [18, 19] of both urine and blood have been used in an attempt to diagnose the etiology of renal allograft dysfunction [20]. Although each of these methods has strengths and weaknesses, these techniques have led to identification of specific genes, transcripts, and proteins important in renal transplant injury. Yet, widespread clinical application of each of those methods has not been accomplished.

We have developed an antibody microarray that can measure multiple proteins, including those involved in inflammation, rejection, and renal tubular injury, simultaneously in biological samples such as serum from transplant patients. Our array can measure many proteins of interest including those involved in inflammation, rejection, and renal tubular injury and has been previously been used in our lab to identify biomarkers in vasculitis [21]. Our current antibody array measures 108 proteins simultaneously and is limited only by the number of antibody pairs commercially available for human protein detection. In this study, we have identified a set of proteins that have differential expression among healthy controls, stable renal transplant patients, and renal transplant patients with rejection.

2   Materials and methods

2.1   Study population

We included a training cohort and a validation cohort. Each cohort is composed of three groups: healthy and volunteer subjects, stable renal transplant patients, and renal transplant patients experiencing acute rejection of renal allograft.

Healthy and volunteer subjects came from two sources: those enrolled as healthy control in the Pfizer CRC clinic (Protocol #394-0) and those employees in our institution. All subjects met the following criteria: (i) normal renal function, as judged by normal Scr levels; (ii) no over the counter pain medications for 1 wk, (iii) no steroidal anti-inflammatory drugs for at least 6 months, and (iv) no overt signs of illness. In addition, other serum chemistries and cell blood counts were normal.

Stable renal transplant patients were defined as having normal and stable Scr levels (Scr < 1.5 mg/dL) and without a clinical indication for a renal allograft biopsy (a Scr increase ≥ 25% over the baseline values). The "to be discarded serum samples" on their routine clinic visit day were collected as well as other pertinent laboratory values and clinical information.

Renal transplant patients experiencing acute rejection underwent renal allograft biopsy to confirm the clinical suspicion of rejection. The Banff 07 classification was used to grade the rejection. The discarded serum samples on the day of renal allograft biopsy were collected. The clinical information

and other pertinent laboratory values were collected as well on the day of biopsy.

Samples collection for training cohort occurred from October 2007 to February 2010 and for validation cohort from March 2010 through February 2011.

This study was approved by the institutional review board at the University of Michigan (Protocol #IRB 2000-0005)

2.2   Immunosuppression

Triple drug regimen consisting of a CNI, mycophenolic acid, and prednisone was used for all but five renal transplant patients. Four patients were on experimental immunosuppression at the time of serum collection and one patient was on no immunosuppression (self-discontinuation due to lack of insurance; Scr 0.9; no proteinuria).

2.3   Antibody array

Antibody microarrays were generated as previously described [22] using commercially available antibody pairs and standard proteins (R&D Systems, Minneapolis, MN). The 108 antibodies used for this analysis were run as three separate MicroArrays. This method was done based on problems with nonspecific binding, which frequently occur when polyclonal antibodies are utilized in this technology. Antibody arrays were spotted using a Piezorray™ (PerkinElmer, Boston, MA) targeted to 16 specific subarrays, designed to work with the ProPlate™ system (Grace Bio-labs/Molecular Probes, Eugene, OR). Capture antibodies were diluted in spotting buffer and spotted in five replicate spots onto EpoxyES microarray slides (Erie Scientific, Portsmouth, NH; 39). As a control for nonspecific binding, isotype control and nonspecific IgG antibodies (rabbit, mouse, and goat) were spotted in each array. For optimization, array components were tested over a range of concentrations; capture antibodies (25–500 µg/mL), standard curves of commercial proteins (1.22–5000 pg/mL), and secondary antibody (diluted 1:200–1:1000). Antibody microarrays were hybridized, scanned, and quantified as outlined below. Standard curves for the arrays were considered acceptable when they achieved an r^2 of 0.95 or greater for quantification. An optimized antibody pair was defined as providing a median fluorescent intensity (MFI) that responded in a dose-dependent manner and did not saturate the MFI. Once an optimal array was achieved, percent recovery and cross-reactivity of standards (1000 pg/mL) spiked into normal human serum was determined and any antibody that showed nonspecific cross-reactivity was excluded from the array.

In addition, we performed conventional ELISAs on several of the current patient samples and found a 0.94 correlation with the MicroArray values. As the MicroArray is an ELISA-based system, we utilized the same antibody pairs (R&D DuoSets) for the conventional ELISA and produced similar standard curves and values as in the MicroArray.

2.4 Hybridization

Standards were diluted into Tris-buffered saline pH 7.4 (Rockland, Gilbertsville, PA) with 0.1% Tween-20 (TBS-t) and 5% BSA (Sigma, St. Louis, MO) and serum samples (80 μL of serum) were diluted fourfold in TBS-t. The EpoxyES-coated slides were placed in the ProPlate™ system and washed five times with TBS-t and blocked for 30 min at room temperature (RT) with 0.5% dry milk, 0.5% BSA (Sigma) in TBS-t on a rotating platform at RT. Chips were washed thrice with TBS-t, standards and diluted serum samples added to the chip (100 μL per well). They were covered and placed on a rotating platform at 4°C overnight. Arrays were washed five times with TBS-t, secondary antibodies diluted in blocking solution were added and incubated for 2 h on a rotating platform at RT. Arrays were washed five times with TBS-t, 100 μL of Alexa Fluor 647-strepavidin (Molecular Probes) diluted 1:2500 in blocking solution was added and incubated for 30 min at RT on a rotating platform in the dark. Arrays were washed three times in TBS-t, removed from the ProPlate™ system, dipped in TBS, then into double distilled H_2O, and dried using compressed air. Slides were stored in the dark at RT until scanned.

2.5 Quantification

Slides were scanned on an Axon 4000B using GenePix 6.0 (Axon Instruments/Molecular Devices, Sunnyvale, CA). A single laser was used and laser intensity was adjusted so that the highest standards were not saturated and the same intensity was used for all slides processed using the same experiment. MFI data minus the background was used to generate the standard curve. Unknowns were calculated using the curve fitting function and quantifiable analytes were plotted based on concentration. Unknowns that were not within the standard linear range or had a high nonspecific background in the isotype antibody controls were excluded from analysis.

2.6 Statistics

Of the original 108 biomarkers (see Appendix), 24 biomarkers were excluded from the healthy subjects versus stable renal transplant analysis and 31 biomarkers were excluded from the stable versus rejecting renal transplant analysis because either the standard deviation was zero or the results did not fall on the standard curve. For the remaining biomarkers, the nonparametric Mann–Whitney rank-based test was used to assess the difference between the medians for each biomarker in each cohort since the majority of the biomarkers did not follow a normal distribution and the Mann–Whitney test does not require normality and is robust to outliers. Interestingly, the two-sample *t*-test gave almost the same results for the majority of the proteins analyzed even though the *t*-test is suitable for samples with a normal distribution (data not shown). In addition to the univariate tests, (i) a least angle regres-

sion analysis (LARS) was used to select a set of simultaneous markers that best predict group membership in a multivariate logistic regression model [23] and (ii) a multivariate analysis was performed that was adjusted for the recipient's race, age, and gender; Scr, donor source, and time post-transplantation. The results were identical to the original, unadjusted analysis (data not shown). Unlike traditional regression methods, LARS can be used when the number of predictors is larger than the sample size without over fitting the model [23, 24]. The key feature of the LARS approach is that regression coefficients are penalized toward zero with the least predictive markers being excluded by estimating their coefficients exactly zero. The advantage of the LARS method is to adjust the effects of other biomarkers so that a relatively weak biomarker can stand out to a greater extent. This procedure yields a prediction model that is both parsimonious and interpretable. Interactions between selected markers were also tested. We ran the LASSO with all biomarkers included initially in order to select important markers in a multivariate setting. Those markers selected by the LASSO were then examined for their interaction terms because it is difficult to interpret an interaction term when one of the involved factors is not important in its main effect. Two-way interaction terms were created between each pair of these markers and included in the LASSO along with the main effects, which were forced into the final model. No interaction term was selected by the LASSO for inclusion in the final model.

To examine the prediction capability, receiver-operator curves (ROCs) were generated on the biomarkers that were found to be significant by the Mann–Whitney test. *t*-Tests, Mann–Whitney tests, and ROC were repeated on the validation cohort.

To assess whether actual transplant patients could be predicted by the model, that is the set of ten biomarkers, a linear discriminant analysis (LDA) was used for the ten healthy control subjects and the 19 stable renal transplant recipients. The same model was then used to predict group membership in the validation cohort.

The analysis was completed with and without one outlier (sample #22) and the results compared. To account for multiple testing using a Bonferroni correction, significance of the Mann–Whitney test was achieved at a level of $\alpha = 5.952 \times 10^{-4}$ when comparing healthy subjects to stable renal transplant patients and at a level of 6.494×10^{-4} when comparing stable to rejecting renal transplant patients. The Mann–Whitney test, *t*-tests, LDA, and ROC were performed in SAS 9.1 (SAS Institute, Cary, NC) and the LARS procedure in the GLM path package in R 2.9.1 [24].

3 Results

Demographic and pertinent clinical information for the training and validation cohorts are shown in Tables 1 and 2. In the training cohort (Table 1), healthy subjects and rejecting transplant patients were younger and predominantly male as

Table 1. Demographic variables of the training cohorts

	Healthy controls ($n = 10$)	Stable patients ($n = 19$)	Rejecting patients ($n = 15$)
Mean age (years)	41.70 ± 15.45	59.9 ± 13.0	45 ± 11.1
R. race (Caucasian)	N/A	19 (100%)	12 (80%)
R. gender (n,% male)	6 (60%)	5 (26%)	9 (60%)
ESRD (diabetes)	N/A	6	4
Months post-transplant	N/A	50.4 ± 31.0	13.1 ± 18.9
Transplant type (n,% living donor [LD])	N/A	11 (58%)	8 (53%)
Donor age (years)[a]	N/A	37.2 ± 13.4	34.9 ± 16.9
Donor gender (n,% male)[a]	N/A	6 (33%)	8 (61%)
CMV status (D+/R−)	N/A	2	4
Mean scr (mg/dL)	0.95 ± 0.25	1.1 ± 0.23	2.6 ± 0.86
CNI (CSA/TAC)[b]	N/A	10/8	12/2

a) Donor age and gender are missing for some patients.
b) One patient was not on any CNI and a few patients were on study drug.
ESRD, end-stage renal disease; R, recipient; D, donor; CMV, cytomegalovirus; CSA, cyclosporine; TAC, tacrolimus.

compared to stable renal transplant recipients. Patients with rejection had shorter post-transplant follow-up (13.1 ± 18.9 vs. 50.4 ± 31.0 months) and had higher Scr levels (2.6 ± 0.9 vs. 1.1 ± 0.2 mg/dL) as compared to stable renal transplant patients. Similar patterns were observed among subjects within the validation cohort (Table 2). While cyclosporine-based immunosuppression was more often used among transplant patients within the training cohort, the patients within the validation cohort received more frequently Tac-based immunosuppression.

Of the 108 proteins that were assayed (Appendix for complete list of proteins on array), ten protein biomarkers were identified that were significantly different between healthy subjects and stable renal transplant patients (Table 3). Although the renal biomarkers, cystatin C and kidney injury molecule-1 (KIM-1), differ significantly between healthy subjects and stable renal transplant patients as would be expected, levels of cystatin C were significantly increased (188 005 vs. 84 023 pg/mL; $p = 3.63 \times 10^{-5}$) and levels of KIM-1 were significantly decreased (872 vs. 3620 pg/mL; $p = 6.70 \times$

10^{-6}) as compared to the healthy controls. In addition, monocyte chemotactic protein-1 and -3 (MCP-1; MCP-3) and tissue inhibitor matrix metalloproteinase-4 (TIMP-4) were significantly higher in stable renal transplant recipients. In contrast, five markers were significantly lower in stable renal transplant patients as compared to healthy subjects: epidermal growth factor (EGF), granulocyte-macrophage colony-stimulating factor (GM-CSF), interleukin-5 (IL-5), IL-1 receptor 1, and macrophage migration inhibitory factor (MIF).

Table 4 shows the protein signature differentiating stable renal transplant recipients from those renal transplant patients with rejection. Seventeen proteins were found to be significantly different between these two groups. Compared to stable transplant recipients, patients with rejection had lower amounts of E-cadherin, EGF, erythropoietin receptor (EPO-R), growth regulated oncogene-alpha (GROα), interleukin 6 (IL-6), MCP-1, macrophage inflammatory protein 3-alpha (MIP-3α), transforming growth factor (TGF)β1, and TGFβ2 and higher amounts of GM-CSF, interleukin 1 receptor 1 (IL-1R1), interleukin 12 p70 (IL-12p70), KIM-1, MIF,

Table 2. Demographic variables of the validation cohorts

	Healthy controls ($n = 15$)	Stable patients I ($n = 15$)[a]	Stable patients II[b] ($n = 14$)[a]	Rejecting patients ($n = 13$)
Mean age (years)	30.26 ± 8.45	49.9 ± 12.6	53.5 ± 15.96	38.77 ± 14.86
R. race (Caucasian)	7 (47%)	13 (86%)	12 (86%)	6 (46%)
R. gender (n,% male)	10 (67%)	8 (53%)	6 (43%)	4 (31%)
ESRD (diabetes)	N/A	6	5	4
Months post-transplant	N/A	$56.1 + 50.9$	54.4 ± 45.5	14.0 ± 17.24
Transplant type (n,% LD)	N/A	5 (33%)	5 (36%)	6 (46%)
Donor age (years)[c]	N/A	30.1 ± 14.1	33.7 ± 17.1	39.7 ± 13.0
Donor gender (n,% male)[c]	N/A	5 (42%)	6 (43%)	6 (46%)
CMV status (D+/R−)	N/A	4	5	3
Mean scr (mg/dL)	N/A	1.0 ± 0.15	1.05 ± 0.13	2.29 ± 0.85
CNI (CSA/TAC)[a]	N/A	6/8	6/7	3/10

a) One patient was not on any CNI and a few patients were on study drug.
b) A new stable transplant validation cohort was used in the rejection analysis.
c) Donor age and gender are missing for some patients.
ESRD, end-stage renal disease; R, recipient; D, donor; CMV, cytomegalovirus; CSA, cyclosporine; TAC, tacrolimus.

Table 3. Protein signature of stable renal transplant recipients versus healthy controls

Marker	Healthy controls ($N = 10$)	Stable transplants ($N = 19$)	*p*-Value
Cystatin-C	84 023 (38 433, 123 381)	188 005 (69 878, 400 000)	3.630×10^{-5}
EGF	1756 (1078, 2259)	419 (151, 2840)	2.690×10^{-5}
GM-CSF	2275 (967, 4000)	412 (149, 4000)	4.680×10^{-5}
IL-1 R1	15 970 (7284, 20 419)	2667 (2405, 30 299)	1.930×10^{-5}
IL-5	409 (294, 751)	82 (22, 648)	4.643×10^{-4}
KIM-1	3620 (1378, 10 872)	872 (450, 4102)	6.700×10^{-6}
MCP-1	418 (275, 644)	841 (471, 1499)	3.000×10^{-6}
MCP-3[a)]	358 (249, 673)	1943 (222, 2000)	1.044×10^{-4}
MIF	6163 (2665, 7498)	2541 (1150, 6797)	4.880×10^{-5}
TIMP-4[a)]	6667 (5437, 7843)	8000 (6521, 8000)	1.300×10^{-6}

a) *p*-Value is conservative due to truncation of data in the stable transplants.
Median values are in picograms/milliliter with the range in parentheses.
p-Value $\leq 5.952 \times 10^{-4}$ was consider statistically significant.

osteopontin, tumor necrosis factor receptor II (TNF-RII), and vascular endothelial growth factor receptor 2 (VEGF-R2). Furthermore, six of these 17 proteins, namely EGF, GM-CSF, IL-1R1, KIM-1, MCP-1, and MIF, were also differentially expressed between healthy controls and stable renal transplant patients. In addition, a multivariate analysis was performed that was adjusted for the recipient's race, age, and gender; Scr, donor source, and time post-transplantation. The results were identical to the original, unadjusted analysis (data not shown).

The protein signatures for those renal transplant recipients with biopsy-proven chronic allograft nephropathy (CAN), acute tubular necrosis (ATN), CNI nephrotoxicity, and BKV nephropathy are shown in Table 5. Significant differences in the levels of several biomarkers can be seen when each condition was compared to stable renal transplant recipients. For example, in those patients with CAN, significant differences were observed in connective tissue growth factor, E-cadherin, and KIM-1 as compared to stable renal transplant recipients. Since tubular injury and dropout are common features of CAN, it is not unexpected to see elevations in such biomarkers as KIM-1. In contrast, HSP27 and nephrin were significantly decreased in patients with BKV and a significant difference in HSP27 expression was only seen in those patients with BKV nephropathy (HSP27 expressions was numerically higher in all other groups as compared to stable patients). Interestingly, HSP27 has been shown to be involved in viral infection, replication, and downstream signaling events in virus-infected cells [25, 26]. As for renal transplant patients with ATN or CNI nephrotoxicity, other biomarkers were significantly

Table 4. Protein signature of stable patients and rejecting patients

Biomarkers	Stable transplants ($n = 19$)	Rejecting transplants ($n = 15$)	Mann–Whitney *p*-value
TGF-β2	2320 (2044, 9628)	980 (85, 1234)	5.388×10^{-10}
E-cadherin	8488 (1892, 48 000)	170 (3, 360)	1.078×10^{-9}
GROα	5281 (2344, 8000)	358 (124, 976)	1.078×10^{-9}
TGF-β1	256 (157, 1111)	89 (45, 121)	1.078×10^{-9}
IL-6	575 (183, 4000)	73 (4, 231)	4.310×10^{-9}
IL-1 R1[a)]	**2667 (2405, 30 299)**	**32 000 (8530, 32 000)**[b)]	**2.802×10^{-8}**
EGF[a)]	**418 (151, 2840)**	**114 (8, 208)**	**3.233×10^{-8}**
MIP-3α	375 (164, 1167)	82 (18, 304)	4.849×10^{-8}
TNF-RII	2138 (1146, 4177)	5994 (565, 6400)	7.423×10^{-6}
KIM-1[a)]	**872 (450, 4102)**	**2364 (999, 5706)**	**9.663×10^{-6}**
Osteopontin	24 145 (11 911, 51 053)	66 802 (27 754, 89 010)	1.211×10^{-5}
VEGF-R2	12 (12, 8000)	8000 (18, 8000)[b)]	1.491×10^{-5}
Epo-R	57 387 (10 926, 80 000)	22 201 (13 635, 47 932)	8.929×10^{-5}
MIF[a)]	**2541 (1150, 6797)**	**3874 (2800, 6214)**	**1.096×10^{-4}**
IL-12p70	2900 (88, 8000)	8000 (6900, 8000)[b)]	1.168×10^{-4}
MCP-1[a)]	**841 (471, 1499)**	**122 (27, 4000)**	**1.300×10^{-4}**
GM-CSF[a)]	**412 (149, 4000)**	**1100 (210, 2305)**	**3.494×10^{-4}**

a) Statistically different between healthy controls and stable patients.
b) *p*-Value is conservative due to truncation of the data in rejecting patients.
Bold values are protein biomarkers common to both stable and rejecting patients.
Median values are in picograms/milliliter with the range in parentheses.
p-Value $\leq 6.494 \times 10^{-4}$ was consider statistically significant.

Clinical Relevance

While the 1 year kidney graft survival is greater than 90% due to improved immunosuppression, long-term outcome has not changed dramatically over the last decade. Factors that influence long-term kidney transplant function include immunologic injury from acute cellular or antibody-mediated rejection, CNI nephrotoxicity, and BKV-associated nephropathy, all of which result in chronic allograft injury and subsequent loss of the transplanted graft. The long-term success of kidney transplantation traditionally depends on Scr levels as the surrogate marker of renal allograft function. Although invasive, renal allograft biopsy is currently the accepted approach used to detect renal allograft dysfunction. Using the protein array, early and prompt detection of renal allograft dysfunction may help improve long-term outcomes. Our preliminary results support that a specific pattern of protein expression or "protein signature" may be able to differentiate between stable transplant patients from those with rejection.

different when compared to stable renal transplant recipients (Table 5). Across all conditions, transforming growth factor beta 1 (TGF-β1) and vascular endothelial growth factor were significantly lower than in stable renal transplant patients (data not shown).

Tables 6 and 7 display the LDA models for the biomarkers of both the training and validation cohorts. Table 6 displays the LDA model to determine the accuracy of these ten biomarkers to predict whether a protein signature defined a healthy subject or a stable renal transplant patient. For the training cohort, no control subjects and three transplant patients were misclassified by the LDA model. For the validation cohort, no healthy control subjects and no stable renal transplant patients were misclassified as having a stable transplant patient and healthy subject profile, respectively. Nonetheless, the p-values for the table and whether the AUC = 0.50 were statistically significant with an AUC of 1.0 for both the training and the validation cohorts, demonstrating that these ten biomarkers could predict accurately the protein signature of a healthy subject versus a renal transplant recipient. Table 7 displays the LDA model to determine the accuracy of these 17 protein biomarkers to predict whether a protein signature defined a stable renal transplant patient or a rejecting renal transplant patient.

For both the training and validation cohorts, no stable renal transplant patients and no transplant patients with rejection were misclassified by the LDA model yielding an AUC of 1.0 for both the training and validation cohorts.

The ROCs analysis for the biomarkers identified in the training cohort by the Mann–Whitney test demonstrated the AUC for the ten biomarkers using the validation cohort is 1.0, suggesting that these ten biomarkers can differentiate healthy subjects from stable renal transplant patients (Fig. 1A). Similarly, the AUC for the 17 biomarkers differentiating stable from rejecting renal transplant patients in the validation set is 1.0 (data not shown). In addition, the ROC of the individual biomarker MCP-1 (Fig. 1B), using the validation cohort, has an AUC that approaches 1.0 (0.960). The ROCs for the other biomarkers were similar to the ROCs generated for MCP-1 and MIF (data not shown). The training set ROC (AUC = 1.0) was similar to the validation set ROC (data not shown).

4   Discussion

For many years, studies have attempted to link a specific gene, transcript, or protein or a pattern of genes or transcripts with renal allograft dysfunction [4–19]. To date, no specific gene,

Table 5. Differences in protein expression of the seven protein biomarkers common to kidney transplant patients diagnosed with CAN, ATN, CNI, or BKV injury

Biomarkers	Stable transplants[a] (n = 12)	CAN[a] (n = 12)	ATN[a] (n = 13)	CNI[a] (n = 6)	BKV[a] (n = 12)
Clusterin (ng/mL)	202 778 ± 48 446	205 208 ± 22 014	102 486 ± 12 449[b]	115 138 ± 17 805[b]	136 762 ± 35 181
CTGF (pg/mL)	968 ± 681	315 ± 62[b]	702 ± 511	327 ± 476	969 ± 898
E-cadherin (pg/mL)	4914 ± 2971	825 ± 192[b]	104 ± 39[b]	1412 ± 1692	5617 ± 2633
Fibronectin (ng/mL)	509 838 ± 86 752	480 208 ± 33 037	184 879 ± 37 565[b]	252 520 ± 47 431[b]	347 560 ± 184 975
HSP27 (pg/mL)	3703 ± 2414	6174 ± 1409	4124 ± 1800	15 263 ± 8916	268 ± 124[b]
KIM-1 (pg/mL)	744 ± 565	2994 ± 1459[b]	453 ± 346	640 ± 847	1172 ± 2290
Nephrin (ng/mL)	67 ± 28	146 ± 104	4.2 ± 11[b]	2.6 ± 1.1[b]	8.3 ± 12.2[b]

a) Mean values are either in picograms/milliliter or nanograms/milliliter along with the SD.
b) Proteins significantly different from stable patients: p-value ≤ 0.001136.
Each cohort in this table was compared to stable renal transplant recipients.
CTGF, connective tissue growth factor.

Table 6. Linear discriminant analysis model for the ten biomarkers of the training and validation cohorts

Cross-validation	Control$_{Actual}$	Txp.$_{Actual}$	
Using markers significant in Mann–Whitney tests for the training cohort			
			Apparent error rate = 3/30 or 10%
Control$_{Predicted}$	10	3	
Transplant$_{Predicted}$	0	16	AUC = 1.0 ($p = 1.51 \times 10^{-6}$)
Making predictions using the validation cohort			
			Apparent error rate = 0
Controls$_{Predicted}$	15	0	
Transplant$_{Predicted}$	0	15	AUC = 1.0 ($p = 1.29 \times 10^{-8}$)

Linear discriminant analysis of ten healthy subjects versus 19 stable renal transplant patients in the training cohort to assess whether transplant patients could be predicted by the model. New healthy subjects and stable transplant patients were used for the validation analysis. Txp, stable renal transplant patients; AUC, area under concentration curve.

protein, or pattern has been identified that is specific for renal allograft dysfunction. The problems with these previous studies are twofold. First, a priori assumptions have been made to link a specific transcript or protein with a particular condition. Second, upregulated genes and transcripts may not be translated into proteins so that differences observed among genes or transcripts may not correlate with renal pathology. Most likely, a pattern of proteins and not a single protein will define the complex, pathological process leading to renal allograft dysfunction. Thus, identification of a protein signature that is specific for the various etiologies of renal allograft dysfunction is paramount.

Our study demonstrates that protein array can be used to identify a specific protein signature for stable renal allograft function from those with impaired function from rejection. To avoid a priori assumptions, the 108 proteins on the array represent a large variety of cellular proteins including, but not limited to, cytokine/chemokines, their receptors and antagonists, cell adhesion molecules, growth and angiogenesis factors, proteases and their inhibitors, matrix metalloproteinases (MMPs) and their inhibitors, HSPs, complement factors, and anticoagulant factors. By assaying a wide variety of proteins, a priori assumptions are limited only by the number of commercially available antibody pairs for plating.

Our preliminary results support that a specific pattern of protein expression or "protein signature" may be able to differentiate between stable transplant patients from those with rejection. First, we identified a protein signature that distinguishes healthy subjects from renal transplant patients with good allograft function and no previous history of rejection. Ten proteins reliably differentiate stable renal transplant recipients from healthy subjects in both the training and validation cohorts. Of these ten proteins, six proteins have lower expression and four proteins (cystatin C, MCP-1, MCP-3, TIMP-4) have higher expression in the stable renal transplant cohort as compared to healthy controls. Second, we identified a protein signature that distinguishes stable renal transplant patients from those with rejection. Seventeen proteins differentiate rejecting renal transplant recipients from stable renal transplant patients in both training and validation cohorts. Of these 17 proteins, nine have lower levels expression and eight have higher levels of expression in the rejecting cohort as compared to the stable renal transplant cohort. Interestingly, six proteins are common to both the protein signature of the stable and the rejecting patients but vary significantly in their levels of expression. For example, EGF expression is highest in the healthy controls and decreases in expression in the stable and rejecting renal transplant patients. In contrast, the other five proteins (GMCSF, IL-1 R1, KIM-1, MCP-1, MIF) are depressed in stable renal transplant patients and are at higher levels in the healthy controls and in the rejecting renal transplant patients suggesting similar

Table 7. Linear discriminant analysis model for the 17 biomarkers of the training and validation cohorts

Cross-validation in training set	Stable$_{Actual}$	Rejecting$_{Actual}$	
Using markers significant in Mann–Whitney tests for the training cohort			
			Apparent error rate = 0
Stable$_{Predicted}$	19	0	
Rejecting$_{Predicted}$	0	15	AUC = 1 ($p = 1.30 \times 10^{-7}$)
Making predictions using the validation cohort			
			Apparent error rate = 0
Stable$_{Predicted}$	14	0	
Rejecting$_{Predicted}$	0	13	AUC = 1 ($p = 2.72 \times 10^{-6}$)

Linear discriminant analysis of 19 stable renal transplant patients versus 15 renal transplant patients with rejection in the training cohort to assess whether rejecting transplant patients could be predicted by the model. New stable and rejecting transplant patients were used for the validation analysis; AUC, area under concentration curve.

A

ROC Curve for 10 Markers in the Validation Set
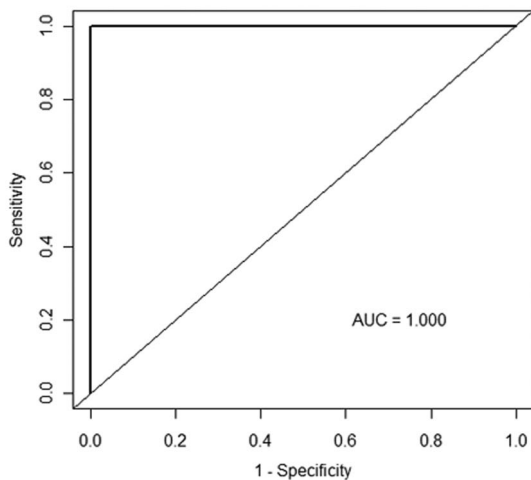
B

ROC Curve for MCP-1 in the Validation Set
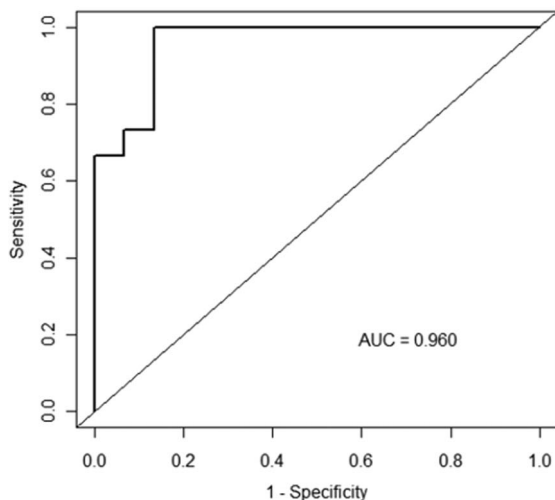
Figure 1. ROCs for the protein biomarkers for the validation cohort and one representative proteins. (A) ROC using all ten significant biomarkers that differentiate stable renal transplant patients from healthy control subjects. (B) ROC for MCP-1 in the validation cohort.

immunoreactivity between the healthy controls and rejecting renal transplant patients. Given the various levels of protein expression among the cohorts, our protein array reflects the spectrum and complexity of the disease process.

In a previous study, KIM-1, a protein marker present in urine of patients with renal disease, was undetectable in healthy controls [27]. In our study, KIM-1 was detected in healthy controls but the detection limit of our array is 1000 times more sensitive than the ELISA used in their study and the biological fluid of choice in our study was serum, not urine, due to the stability of proteins at a neutral pH. In a contrasting study using LC and mass spectroscopy, Ling and colleagues identified a urine peptidome and correlated it

with renal transplant biopsy transcriptome for rejection [28]. They identified a six gene biomarker panel for acute rejection (COL1A2, COL3A1, UMOD, MMP-7, SERPING1, TIMP-1). Interestingly, our protein array identified a TIMP-4 that differentiated healthy subjects from stable renal transplant patients and that was also present in patients with rejection but only trended toward statistical significance. As these authors found, we believe that collagen biomarkers are also important determinants of renal allograft dysfunction and are currently being incorporated into our protein array.

Our protein array has several advantages. First, it allows for the measurement of over a 100 proteins simultaneously and identifies a pattern of protein biomarkers specific for renal allograft dysfunction. Second, since the protein array is noninvasive, it could be used to follow patients longitudinally and to detect subclinical renal allograft dysfunction before it would be detected by conventional methods (i.e. Scr). By detecting renal allograft dysfunction earlier, long-term graft outcomes could be positively impacted. Third, since the array allows for detection of protein differences in serum, an acidic pH as in urine is not a concern. Protein array chips can also be tailored for specific pathological conditions so that 100 proteins do not need to be assayed every time. Lastly, by identifying specific proteins involved in the rejection process, new targets for therapeutic agents could be developed.

There are several limitations to our study. Since the array can only be constructed from antibody pairs that are commercially available, this limitation restricts the number of proteins that can be assayed in the serum. As more antibody pairs become available, the numbers of proteins that can be assayed on the array will increase to where 150 proteins per chip and several chips at a given time can be assayed simultaneously. Another major limitation is the lack of protocol renal transplant biopsies in the stable renal transplant cohort to validate their Scr levels. Since renal transplant protocol biopsies have now become standard at our institution, future studies will no longer have this limitation. Due to the nature of this work being a pilot study, other protein patterns of renal allograft dysfunction such as BK nephropathy have yet to be elucidated, even though preliminary data from our protein array can differentiate between ATN and CNI toxicity. The sample size in our study limits the ability to perform interaction models among the proteins of interest and the ROCs. For example, the AUC from the ROC for the validation cohort is 1.0, which most likely reflects the small sample size. In the Ling et al. paper [28], the AUCs approached 1.0 as well, most likely due to the small sample size in their study. Although our array can measure proteins in the urine (unpublished data), we chose to use serum samples due to the neutral pH and stability of proteins in serum to protect protein integrity. The majority patients in our cohorts were Caucasian (only ~20% of patients in our institution are African American). In contrast, the validation cohort did include Asians and one African American and yet the results were similar to the training transplant cohort. Lastly, a control cohort of patients

with other inflammatory conditions such as lupus nephritis or rheumatoid arthritis would have reinforced of our results.

A statistical concern is that biomarkers with nonsignificant p-values may strongly discriminate between two cohorts when brought together in a protein signature. In the current literature, the univariate analysis approach based on p-values is popular (perhaps the first choice of screening analysis) because of its computational convenience. However, it has two major weaknesses: (i) biomarkers that are not associated with the disease but strongly associated with a significant disease predictor can be selected as important biomarkers [27]; and (ii) a biomarker that is marginally uncorrelated but jointly associated with the disease outcome cannot be identified by such a simple univariate analysis method based on p-values [28]. These shortcomings can be overcome by joint screening methods such as the LASSO used in our manuscript.

In summary, using protein array methodology, we were able to differentiate stable renal transplant patients from those patients with rejection. In addition, our results support that a group of proteins are involved in the disease process and the pattern of protein expression or a "protein signature" is necessary to differentiate between various disease processes. Given the promising results of our current study, future directions for the protein array will include determination of the protein signature for other etiologies of renal allograft dysfunction such as BK nephropathy, chronic allograft injury, antibody-mediated rejection, and CNI nephrotoxicity. In addition, patient samples will be collected prospectively and longitudinally to determine the effect of treatment on the protein signature and to correlate it with patient outcomes.

5 References

[1] Meier-Kriesche, H. U., Schold, J. D., Srinivas, T. R., Kaplan, B., Lack of improvement in renal allograft survival despite a marked decrease in acute rejection rates over the most recent era. *Am. J. Transplant* 2004, *4*, 378–383.

[2] Vadivel, N., Tullius, S. G., Chandraker, A., Chronic allograft nephropathy. *Semin. Nephrol.* 2007, *27*, 414–429.

[3] Kaplan, B., Schold, J., Meier-Kriesche, H. U., Poor predictive value of serum creatinine for renal allograft loss. *Am. J. Transplant* 2003, *3*, 1560–1565.

[4] Magee, C. C., Denton, M. D., Womer, K. L., Khoury, S. J., Sayegh, M. H., Assessment by flow cytometry of intracellular cytokine production in the peripheral blood cells of renal transplant recipients. *Clin. Transplant* 2004, *18*, 395–401.

[5] Roberti, I., Reisman, L., Serial evaluation of cell surface markers for immune activation after acute renal allograft rejection by urine flow cytometry—correlation with clinical outcome. *Transplantation* 2001, *71*, 1317–1320.

[6] Ting, Y. T., Coates, P. T., Marti, H. P., Dunn, A. C. et al., Urinary soluble HLA-DR is a potential biomarker for acute renal transplant rejection. *Transplantation* 2010, *89*, 1071–1078.

[7] Sutherland, S. M., Li, L., Sigdel, T. K., Wadia, P. P. et al., Protein microarrays identify antibodies to protein kinase Czeta that are associated with a greater risk of allograft loss in pediatric renal transplant recipients. *Kidney Int.* 2009, *76*, 1277–1283.

[8] Platt, R. E., Wu, K. S., Poole, K., Newstead, C. G., Clark, B., Soluble CD30 as a prognostic factor for outcome following renal transplantation. *J. Clin. Pathol.* 2009, *62*, 662–663.

[9] Schaub, S., Nickerson, P., Rush, D., Mayr, M. et al., Urinary CXCL9 and CXCL10 levels correlate with the extent of subclinical tubulitis. *Am. J. Transplant* 2009, *9*, 1347–1353.

[10] Foxall, P. J., Mellotte, G. J., Bending, M. R., Lindon, J. C., Nicholson, J. K., NMR spectroscopy as a novel approach to the monitoring of renal transplant function. *Kidney Int.* 1993, *43*, 234–245.

[11] Quintana, L. F., Solé-Gonzalez, A., Kalko, S. G., Bañon-Maneus, E. et al., Urine proteomics to detect biomarkers for chronic allograft dysfunction. *J. Am. Soc. Nephrol.* 2009, *20*, 428–435.

[12] Schaub, S., Mayr, M., Hönger, G., Bestland, J. et al., Detection of subclinical tubular injury after renal transplantation: comparison of urine protein analysis with allograft histopathology. *Transplantation* 2007, *84*, 104–112.

[13] Dadhania, D., Snopkowski, C., Ding, R., Muthukumar, T. et al., Validation of noninvasive diagnosis of BK virus nephropathy and identification of prognostic biomarkers. *Transplantation* 2010, *90*, 189–197.

[14] Cheng, O., Thuillier, R., Sampson, E., Schultz, G. et al., Connective tissue growth factor is a biomarker and mediator of kidney allograft fibrosis. *Am. J. Transplant* 2006, *6*, 2292–2306.

[15] Tatapudi, R. R., Muthukumar, T., Dadhania, D., Ding, R. et al., Noninvasive detection of renal allograft inflammation by measurements of mRNA for IP-10 and CXCR3 in urine. *Kidney Int.* 2004, *65*, 2390–2397.

[16] Muthukumar, T., Dadhania, D., Ding, R., Snopkowski, C. et al., Messenger RNA for FOXP3 in the urine of renal-allograft recipients. *N. Engl. J. Med.* 2005, *353*, 2342–2351.

[17] Li, B., Hartono, C., Ding, R., Sharma, V. K. et al., Noninvasive diagnosis of renal-allograft rejection by measurement of messenger RNA for perforin and granzyme B in urine. *N. Engl. J. Med.* 2001, *344*, 947–954.

[18] Sarwal, M., Chua, M. S., Kambham, N., Hsieh, S. C. et al., Molecular heterogeneity in acute renal allograft rejection identified by DNA microarray profiling. *N. Engl. J. Med.* 2003, *349*, 125–138.

[19] Mas, V., Maluf, D., Archer, K., Yanek, K. et al., Establishing the molecular pathways involved in chronic allograft nephropathy for testing new noninvasive diagnostic markers. *Transplantation* 2007, *83*, 448–457.

[20] Matas, A. J., Gillingham, K. J., Payne, W. D., Najarian, J. S., The impact of an acute rejection episode on long-term renal allograft survival. *Transplantation* 1994, *57*, 857–859.

[21] Olle, E. W., Deogracias, M. P., Messamore, J. E., McClintock, S. D. et al., Screening of serum samples from Wegener's granulomatosis patients using antibody arrays. *Proteomic Clin. Appl.* 2007, *1*, 1212–1220.

[22] Olle, E. W., Sreekumar, A., Warner, R. L., McClintock, S. D. et al., Development of an internally controlled antibody microarray. *Mol. Cell. Proteomics* 2005, *4*, 1664–1672.

[23] White, A. M., Daly, D. S., Varnum, S. M., Anderson, K. K. et al., ProMAT: Protein Microarray Analysis Tool. *Bioinformatics* 2006, *22*, 1278–1279.

[24] Park, M. Y., Hastie, T., L1 regularization path algorithm for generalized linear models. *J. R. Stat. Soc.* 2007, *69*, 659–677.

[25] Mathew, S. S., Della Sleva, M. P., Burch, A. D., Modification and reorganization of the cytoprotective cellular chaperone HSP27 during herpes simplex virus type 1 infection. *J. Virol*. 2009, *83*, 9304–9312.

[26] Rajaiya, J., Yousuf, M. A., Singh, G., Stanish, H., Chodosh, J., Heat shock protein 27 mediated signaling in viral infection. *Biochemistry* 2012, *51*, 5695–5702.

[27] Van Timmeren, M. M., van den Heuvel, M. C., Bailly, V., Bakker, S. J. L. et al., Tubular kidney injury molecule-1 (KIM-1) in human renal disease. *J. Pathol.* 2007, *212*, 209–217.

[28] Ling, X. B., Sigdel, T. K., 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.

Appendix

Table A1 Human antibody array performance

Proteins	Range (pg/mL)	Curve fit r^2	Proteins	Range (pg/mL)	Curve fit r^2
ACE	31–8000	0.992	IL-12p70	750–2000	0.991
Activin A	333–8000	0.989	IL-13	6–6000	0.965
ADAM8	354–4000	0.994	IL-15	4–1000	0.996
Adiponectin	1000–4000	0.981	IL-18 BPα	1875–6000	0.997
ALCAM	7.81–8000	0.976	IP-10	125–2000	0.991
BCAM	208–4000	0.992	Leptin	8–2000	0.961
BLC/BCA-1	83–1000	0.982	L-selectin	5–5000	0.982
β-NGF	105–2000	0.999	MCP-1	42–1000	0.976
Cathepsin S	58–1000	0.984	MCP-2	300–4000	0.997
Complement D	8–2000	0.987	MCP-3	23–500	0.994
CTGF	4–1000	0.996	MCP-4	2–500	0.992
Dectin-1	2500–10 000	0.991	MIF	250–2000	0.993
E-cadherin	12–12 000	0.999	MIP-1α	26–1000	0.983
EGF	104–1000	0.994	MIP-1β	90–1000	0.991
Endostatin	16–4000	0.999	MIP-1δ	4–1000	0.971
Eotaxin	104–1000	0.994	MIP-3α	90–1000	0.992
Epo-R	10 000–20 000	0.999	MIP-3β	83–1000	0.999
E-selectin	6–6000	0.987	MMP-3	2–2000	0.952
Fas ligand	63–2000	0.986	MMP-9	8–2000	0.997
Fcg RIIB/C	542–3000	0.985	MMP-13	1480–10 000	0.982
FGF basic	4–1000	0.974	OSM	500–2000	0.998
FGF-4	375–1000	0.999	P-cadherin	1667–4000	0.987
FGF-6	260–5000	0.971	PDGF-AB	73–1000	0.995
Fractalkine	5000–20000	0.982	PDGF-BB	183–2000	0.992
G-CSF	125–2000	0.983	Pro-cathepsin S	167–2000	0.992
GM-CSF	250–1000	0.998	P-selectin	8–8000	0.999
GROα	375–2000	0.989	RANTES	16–1000	0.997
HGF	375–8000	0.985	Resistin	8–2000	0.996
HSP27 (IC)	31–2000	0.998	sICAM-1	4–1000	0.976
HSP70 (IC)	1667–10 000	0.997	sTNF RI	233–1000	0.995
ICAM-3	500–2000	0.997	Survivin (IC)	667–4000	0.992
IFNγ	42–1000	0.997	TARC	96–500	0.995
IGFBP-2	4–4000	0.987	TGF-α	47–1000	0.997
IGFBP-3	8–8000	0.976	TGF-β1	500–4000	0.997
IL-1 RI	917–8000	0.997	TGF-β2	1250–4000	0.997
IL-1 sRII	500–2000	0.960	TGF-β3	500–2000	0.990
IL-1α	208–1000	0.998	TIMP-1	500–2000	0.960
IL-1β	42–1000	0.998	TIMP-2	8–2000	0.974
IL-2	94–1000	0.982	TIMP-4	750–2000	0.988
IL-3	333–1000	0.987	TNFα	333–1000	0.997
IL-4	333–2000	0.991	uPAR	8–2000	0.975
IL-5	39–1500	0.996	VCAM-1	4–1000	0.976
IL-6	52–1000	0.982	VEGF	83–2000	0.984
IL-6 sR	8–1000	0.997	VEGF-D	375–2000	0.992
IL-8	73–2000	0.993	VEGF R2(IC)	8–2000	0.983
IL-10	4–4000	0.962	PAI-I	39–160 000	0.999
Fetuin-A/AHSG	195–800 000	0.997	IL-18	1–4000	0.999
IL-2 sRα	6–24 000	0.994	KIM-1	4–16 000	0.999
IL-2 Rβ	6–24 000	0.995	DAF	195–800 000	0.995
CLUSTERIN	156–640 000	0.998	Lactoferrin	4–14 000	0.999
Osteopontin	49–200 000	0.998	Annexin V	98–400 000	0.996
NGAL	10–40 000	0.995	Cystatin C	98–400 000	0.999
TNF RII	2–6400	0.998	Lymphotoxin β R	1–4000	0.997
Fibronectin	4–80 000	0.976	Lymphotoxin α//β	195–800 000	0.996

Human antibody chip range of detection (pg/mL) and correlation of the standard curves.