Interference of globin genes with biomarker discovery for allograft rejection in peripheral blood samples

Li Li,1 Lihua Ying,1 Maarten Naesens,1,4 Wenzhong Xiao,2 Tara Sigdel,1 Sue Hsieh,1 Jon Martin,1 Rong Chen,3 Kang Liu,1 Michael Mindrinos,2 Ron Davis,2 and Minnie Sarwal1

1Pediatrics Department, Stanford University, Stanford; 2Stanford Genome Technology Center, Stanford University, Palo Alto; 3Stanford Medical Informatics, Stanford University, Stanford, California; and 4Department of Nephrology and Renal Transplantation, University Hospitals Leuven, Leuven, Belgium

Submitted 17 September 2007; accepted in final form 26 October 2007

WHOLE GENOME GENE EXPRESSION profiling using microarray is a powerful technology for studying molecular biology and leads to new insights into the pathogenesis, classification, evolution, and prognosis of a wide range of disease processes (19). In kidney transplantation, the systematic study of the gene expression patterns of biopsy samples from normal and dysfunctional renal allografts has revealed consistent differences in gene expression patterns associated with biopsy-proven acute graft rejection (AR), nephrotoxic effects of drugs, chronic allograft nephropathy (CAN), and biopsies from normal engrafted kidneys (21).

Next to its value in the identification and elucidation of biological processes underlying human disease, microarray technology is a powerful tool in the discovery of new biomarkers for disease (19). Biomarkers detectable in peripheral blood would be ideal for the detection or staging of a disease, as peripheral blood is readily available from the patient and accessible without invasive procedures. In the case of transplantation, where the detection of rejection is usually made on invasive biopsies, it could be hypothesized that noninvasive transcriptional profiling of peripheral blood will reveal rejection-specific expression patterns from circulating immune cells. However, in kidney transplant rejection, the analysis of gene expression data in whole blood has proven difficult for detecting significant genes specific for acute graft rejection. Previous studies have demonstrated that the abundance of globin genes in whole blood may mask the underlying biological differences between whole blood samples. In the present study, we compared the gene expression profiles of peripheral blood of nine stable renal allograft recipients with seven matched patients having an ongoing acute renal transplant rejection, using four different protocols of preparation, amplification, and synthesis of cRNA or cDNA and hybridization on the Affymetrix platform. We demonstrated that the globin reduction method is not sufficient to mask clinically relevant rejection-specific transcriptome profiles in whole blood. Applying an additional mathematical depletion of the globin genes improves the efficacy of globin reduction but cannot remove the confounding influence of globin gene hybridization. Sampling of peripheral blood leukocytes alone, without the confounding influence of globin mRNA, provides sensitive and specific peripheral signatures for graft rejection, with many of these signals overlapping with rejection-driven tissue (kidney)-specific signatures from matched biopsies. Similar applications may exist for array-based biomarker discovery for other diseases associated with changes in leukocyte trafficking, activation, or function.

kidney transplantation; microarray analysis; graft rejection; peripheral blood leukocytes

Article published online before print. See web site for date of publication (http://physiolgenomics.physiology.org).

Address for reprint requests and other correspondence: M. Sarwal, Dept. of Pediatrics, G320, 300 Pasteur Dr., Stanford, CA 94304 (e-mail: msarwal@stanford.edu).
**METHODS**

**Patients and Samples**

Sixteen whole blood samples from sixteen pediatric recipients of a kidney allograft were included. Seven samples (AR group) were obtained at the time of a biopsy-proven AR episode according to the Banff classification (IA, IIA, IB) (20), and nine samples were collected from patients with stable graft function during regular follow-up after transplantation (stable group). Except for the presence of rejection, both groups were matched for recipient gender, recipient age, immunosuppression protocol time after transplantation, donor gender, type of donor source, and donor age (Table 1). Sixteen additional blood samples [7 from patients experiencing a biopsy-proven AR (Banff grade: IA, IIA, and IB) and 9 samples from stable patients] were used for the analysis of gene expression profiles in proven AR (Banff grade: IA, IIA, and IB) and 9 samples from stable patients were also matched for peripheral blood leukocyte samples (Table 1). AR or stable patients in whole blood samples were also matched with peripheral blood leukocyte samples (Table 1). The blood samples were collected between May 2002 and February 2006. Written, informed consent was obtained from all the subjects, and the study was approved by the institutional review board of Stanford University.

**Whole Blood Samples**

**Sample collection and RNA extraction.** Whole blood was collected in PAXgene Blood RNA Tubes (PreAnalytiX, Qiagen), and total RNA was extracted using the PAXgene Blood RNA Kit (PreAnalytiX, Qiagen). Total RNA concentration was measured by NanoDrop ND-1000 (NanoDrop Technologies, Wilmington, DE), and the integrity of extracted total RNA was assessed with the Agilent 2100 Bioanalyzer using RNA Nano Chips (Agilent Technologies, Santa Clara, CA). Total RNA was stored at −80°C until sample preparation for the microarray experiments.

**Microarray target preparation, hybridization, and gene expression profiling.** Hybridization was performed on GeneChip Human Genome U133 Plus 2.0 Arrays (Affymetrix, Santa Clara, CA) after three different methods of preparation, amplification, and synthesis of cRNA or cDNA were used for each sample (Fig. 1). PAXGENE T7 PROTOCOL. The PAXgene T7 protocol (T7) employed one-cycle cDNA synthesis with the use of T7-Oligo(dT) Primer. Two micrograms of total RNA were used for the synthesis of cRNA as the final product. The manufacturer’s protocol (GeneChip Expression Analysis Technical Manual, Affymetrix) was followed to synthesize biotin-labeled cRNA.

**GLOBIN REDUCTION BASED ON PAXGENE T7 PROTOCOL.** For the globin reduction based on the PAXgene T7 protocol (T7+GR), to reduce the amount of cRNA generated from globin mRNA, 8 μg of total RNA were subjected to a globin depletion step using GeneChip Globin-Reduction Kit (Affymetrix/PreAnalytiX). After the globin reduction step for the cRNA synthesis, the manufacturer’s protocol (Expression Analysis Technical Manual, Affymetrix) was followed, as mentioned above.

**NUGEN PROTOCOL.** cDNA was synthesized from total RNA using the Ovation Biotin RNA Amplification and Labeling System. The Ovation Biotin RNA Amplification and Labeling System (NuGEN Technologies, San Carlos, CA) was used to synthesize biotin-labeled cDNA from 100 ng of total RNA as starting material. The synthesized cDNA was hybridized onto GeneChip Human Genome U133 Plus 2.0 Array. The chips were washed and scanned as recommended by the Ovation Biotin RNA Amplification and Labeling System User Guide (version 1.0).

**Peripheral Blood Leukocyte Protocol**

Whole blood (5–10 ml) was collected in sodium heparin tubes, and the initial separation of the white blood cells involved a centrifugation step to recover the plasma fraction, followed by bicarbonate lysis of the erythrocytes. Following lysis, the remaining leukocytes (peripheral blood leukocytes; PBL) were recovered by centrifugation. From the resulting leukocytes of each sample, a total of 2 μg of RNA was extracted using the RNeasy Midi Kit (Qiagen, Valencia, CA). Sample preparation was done using the T7 protocol, with hybridization on GeneChip Human Genome U133 Plus 2.0 Arrays (Affymetrix).

**Microarray Data Analysis**

DChip (13, 14) software was used to calculate probe level intensities and quality measures including median intensity, percentage of probe set outliers, and percentage of single probe outliers for each chip. Quartile-quartile normalization was performed, and gene expression values were transformed to log₂ for analysis. Two-class unpaired significance analysis of microarray (SAM) (24) was used to determine significant differential gene expression between the AR and stable patient groups for each sample preparation and hybridization method separately. A false discovery rate (FDR; q < 5%) was used to determine the differential genes between the two classes. GeneSpring

---

**Table 1. Patient demographic information for 16 samples for T7, T7+GR, and NuGEN and PBL protocols**

<table>
<thead>
<tr>
<th>Protocols</th>
<th>AR (n = 7)</th>
<th>STA (n = 9)</th>
<th>P Value</th>
<th>AR (n = 7)</th>
<th>STA (n = 9)</th>
<th>P Value</th>
<th>P of AR (WB vs. PBL)</th>
<th>P of STA (WB vs. PBL)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recipients</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender, %females</td>
<td>29%</td>
<td>22%</td>
<td>1.0</td>
<td>14%</td>
<td>33%</td>
<td>0.4</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Mean age, yr</td>
<td>11.5±5.2</td>
<td>11.8±6.7</td>
<td>0.9</td>
<td>12.5±5.8</td>
<td>12.5±6.1</td>
<td>1.0</td>
<td>0.7</td>
<td>0.8</td>
</tr>
<tr>
<td>Age range, yr</td>
<td>1–18</td>
<td>2–18</td>
<td>NA</td>
<td>1–18</td>
<td>3–17</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Males</td>
<td>13.4±3.2</td>
<td>13.3±6.8</td>
<td>1.0</td>
<td>14.1±4.4</td>
<td>15.9±3.0</td>
<td>0.4</td>
<td>0.8</td>
<td>0.4</td>
</tr>
<tr>
<td>Females</td>
<td>6.5±7.4</td>
<td>6.5±6.1</td>
<td>1.0</td>
<td>2.8*</td>
<td>5.5±4.0</td>
<td>NA</td>
<td>NA</td>
<td>0.8</td>
</tr>
<tr>
<td>Immunosuppression, %SF</td>
<td>57%</td>
<td>56%</td>
<td>1.0</td>
<td>71%</td>
<td>67%</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Sample after post-txp, yr</td>
<td>2.2±2.1</td>
<td>1.7±1.1</td>
<td>0.6</td>
<td>3.2±2.9</td>
<td>2.2±1.5</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Time range, yr</td>
<td>0.5–5.7</td>
<td>0.4–3.5</td>
<td>NA</td>
<td>0.5–8.9</td>
<td>0.5–4.9</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Donors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type, %LRD</td>
<td>86%</td>
<td>78%</td>
<td>1.0</td>
<td>57%</td>
<td>89%</td>
<td>0.3</td>
<td>0.6</td>
<td>1.0</td>
</tr>
<tr>
<td>Gender, %females</td>
<td>57%</td>
<td>78%</td>
<td>0.6</td>
<td>67%</td>
<td>56%</td>
<td>1.0</td>
<td>1.0</td>
<td>0.6</td>
</tr>
<tr>
<td>Mean age, yr</td>
<td>36.9±5.4</td>
<td>33.8±6.0</td>
<td>0.3</td>
<td>32.3±17.2</td>
<td>40.4±7.5</td>
<td>0.3</td>
<td>0.5</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Values are means ± SD. T7, T7 PAXgene; GR, globin reduction; PBL, peripheral blood leukocytes; AR, acute rejection; STA, stable; WB, whole blood; SF, steroid free; txp, transplant; LRD, living remaining donors; NA, not applicable. *n = 1 for recipient female in stable samples.
GX (Agilent, Palo Alto, CA) was used to analyze the significant genes on Gene Ontology (GO) annotation. Prediction analysis of microarray (PAM) (23) was performed for the classification of the two phenotypes based on the leave-one-out cross validation. Singular value decomposition (SVD) (9), which is a data-driven mathematical framework in the field of pattern recognition that determines unique orthogonal gene and corresponding array expression patterns in the data set (1, 16, 22), was applied to the data set to determine the eigengenes as principle components across all samples (12). SVD allowed the identification of patterns present in the data from the correlations between the array expression patterns and independent biological processes.

RESULTS
Comparison of Gene Expression Signals on Three Different Protocols in Whole Blood Samples

SAM was not able to detect significant gene expression differences between the AR phenotype and stable patients after T7 and NuGEN amplification protocols, either with an FDR <5% or with an FDR <10%. Only two genes, C1QA and C1QB, which are the A and B chains, respectively, of the early complement cascade, were detected in the T7+GR protocol with FDR = 0%. Aside from the two complement genes, the next most significant expressed gene could only be found with the minimum FDR of 23% in the T7+GR samples, and no significant genes could be found within T7 and NuGEN protocol samples (FDR ≤ 40% for both sample sets; see Table 3, top).

To analyze whether this low discovery rate related to quality control issues in the hybridization of the samples by the three protocols, the DChip program was used to perform quartile-quartile normalization with perfect-matched modeling. In addition, DChip reported prenormalization intensity of the probes and detection calls (Table 2A). The NuGEN protocol samples had the lowest median signal intensity; however, these samples also had the lowest standard deviation and the highest mean percent present call (55.8 ± 3.9%) among these three protocols.

Table 2.

<table>
<thead>
<tr>
<th></th>
<th>T7</th>
<th>T7+GR</th>
<th>NuGEN</th>
<th>PBL</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: median intensity and present call</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median intensity</td>
<td>92.3±19.1</td>
<td>88.8±17.7</td>
<td>62.8±7.4</td>
<td>93.25±45.45</td>
</tr>
<tr>
<td>Present call, %</td>
<td>45.4±5.2</td>
<td>50.3±3.5</td>
<td>55.8±3.9</td>
<td>60.05±43.45</td>
</tr>
<tr>
<td>B: hemoglobin gene proportions and hypergeometric P values</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemoglobin, %</td>
<td>12/21 (57%)</td>
<td>9/21 (43%)</td>
<td>9/21 (43%)</td>
<td></td>
</tr>
<tr>
<td>P value</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
</tbody>
</table>

A: microarray experimental quality of median intensity and present call for T7, T7+GR, NuGEN, and PBL protocols. Values are means ± SD. B: hemoglobin gene proportions and hypergeometric P values for 3 WB-based protocols in eigengene1 by singular value decomposition (SVD).
cols, which indicates that cDNA is more specific than cRNA binding. The T7 protocol had the lowest mean percent present call (45%). Nevertheless, no outlier was detected on the basis of the DChip algorithm for the three protocols used on peripheral whole blood.

In addition to sample quality issues, to estimate any influence of inter- and intragroup sample expression variation, Spearman correlation coefficients based on the nonnormal distribution of 54,000 genes across all samples were calculated among the same sample phenotype to determine outliers (Fig. 2, A–C) for the three protocols. The assumption is that samples within the same condition (rejection or stable) should be most similar at a global level. Stable samples, as expected, significantly correlated with each other and clustered together in the box-and-whisker plot (Fig. 2, A–C) in T7 ($r_{\text{median}} = 0.95–0.96; r_{\text{range}} = 0.93–0.97$), T7+GR ($r_{\text{median}} = 0.95–0.97; r_{\text{range}} = 0.94–0.98$), and NuGEN protocols ($r_{\text{median}} = 0.90–0.95; r_{\text{range}} = 0.88–0.97$). Spearman correlation values were also highly significant in AR samples; however, the ranges were broader than those in stable samples in the box-and-whisker plot in T7 ($r_{\text{median}} = 0.90–0.96; r_{\text{range}} = 0.88–0.97$), T7+GR ($r_{\text{median}} = 0.90–0.97; r_{\text{range}} = 0.87–0.97$), and NuGEN protocols ($r_{\text{median}} = 0.93–0.97; r_{\text{range}} = 0.93–0.97$). Thus, by both DChip and Spearman correlation analysis, all samples passed the quality control screening.

Impact of Globin Gene Reduction and SVD in Whole Blood Samples

We previously reported specific and highly detectable transcriptional profiles for acute graft rejection in tissue biopsies (21), with 1,340 genes discovered between the two classes (FDR <5%). Given that the signal for AR in peripheral whole blood was very difficult to detect, even after globin reduction, it is likely that biological differences for AR are either minimal to absent in peripheral blood, or, indeed, the confounding effect of globin genes persists despite the globin reduction protocol, masking discovery for rejection biomarkers in blood. To test the latter hypothesis, the impact of globin genes was assessed for the three protocols.

On the Affymetrix GeneChip Human Genome U133 Plus 2.0 platform, 21 probes are annotated as hemoglobin genes; these are HBA, HBB, HBG, HBM, HBE, HBZ, HBQ, and HBD. With the use of SVD, patterns of gene expression were observed to have similar trends, as shown in Fig. 3A for all three protocols. SVD decomposes the patterns into 16 eigengenes and ranks them by contribution, which was calculated from the SVD equation. The contribution of the first eigengene was almost twice that of the second eigengene and was strongly associated with globin genes, as shown in Table 2B; 57, 43, and 43% hemoglobin genes were ranked as top gene expression values for T7, T7+GR, and NuGEN, respectively, all with $P < 0.0001$ which were calculated by hypergeometric statistics. Figure 3B indicates that, despite attempts at globin reduction by the T7+GR protocol, expression of globin gene values ranked highest across the entire 54,000-probe set for all three protocols.

After the identification of globin genes as the highest ranking eigengene, we removed this first eigengene, with the hypothesis that this predominant globin-driven eigengene may mask the rejection signature in whole blood. Removing this eigengene is mathematically equivalent to setting the expression level of this eigengene to zero. Removing this eigengene retains all the genes and arrays in the original microarray data set but changes the expression matrix. The three matrices (the eigernarrays matrix, the eigengene expression matrix, and the eigengene matrix) were then multiplied to reconstruct a microarray data set after removal of the globin eigengene.

AR signature improved after removal of the globin eigengene, especially in the protocol employing a first step of globin depletion (T7+GR) and in the more stringent NuGEN protocol. However, even in these protocols, the signal differences between AR and stable samples remained week and only yielded seven, one, and one probe, respectively, for the T7+GR, T7, and NuGEN protocols, using a FDR <5% (Table 3, bottom). This suggests that either the overall biological expression signature for AR is very weak in peripheral blood, or else the confounding effect of globin genes, despite globin reduction and SVD, continues to cause sufficient interference in the hybridization of whole blood samples and thus still masks the signature of rejection-regulated transcription in peripheral blood. To differentiate between these causes, we performed array studies on isolated PBL from a matched group of AR and stable peripheral blood samples. Our hypothesis was that, if AR-related peripheral blood gene expression is weak overall, biomarker discovery will continue to be a problem regardless of sampling whole blood or PBL.

Improvement of Signal Detection for AR-Specific Genes Using PBL

Sixteen matched PBL samples (7 AR and 9 stable) were run on the same platform using the PBL protocol (Table 1). As expected, there was no signal for globin gene mRNA in the PBL samples. DChip was used to perform the quantile-quantile normalization, and Spearman correlation coefficients were calculated to assess outliers among the 16 samples. Intragroup correlations were $r_{\text{median}} = 0.94–0.97$ with $r_{\text{range}} = 0.94–0.97$ ($P < 0.0001$) in stable samples, and $r_{\text{median}} = 0.93–0.97$ with $r_{\text{range}} = 0.92–0.98$ ($P < 0.0001$) in AR samples (Fig. 2D). The percent present call was 60.05 ± 43.45%, indicating that >60% of probes could be detected, with a median intensity of 93.25 ± 45.45 across all arrays with excellent experimental quality (Table 2A).

Unlike the results from whole blood, where almost no AR-specific signal could be detected, two-class unpaired SAM analysis comparing AR ($n = 7$) and stable ($n = 9$) PBL samples could identify ~8,000 differentially expressed transcripts, representing ~6,000 annotated genes with a cutoff FDR <5% (~2,400 upregulated and ~5,600 downregulated transcripts in AR). PAM perfectly segregated the samples obtained in stable patients from the samples obtained in patients experiencing AR, with 100% sensitivity, 100% specificity, 100% positive predictive value, and 100% negative predictive value (Fig. 4A). The significant genes identified in PBL for rejection were interrogated in rejecting biopsy tissue from our ongoing analysis of biopsy rejection data from matched samples on the same Affymetrix platform. Approximately 2,000 genes that were the most significant (FDR <5% by SAM for both lists) overlapped between PBL and biopsy signatures. The fold change of these genes showed a similar trend in both biopsy and blood rejection ($R^2 = 0.88$; Fig. 4B), but the fold
Fig. 2. Box plot of sample correlations for T7 protocol (A), T7+GR protocol (B), NuGEN protocol (C), and PBL protocol (D). Box plots are shown for calculated Spearman correlation coefficients among the same phenotype samples. Correlations of stable samples are shown at left for T7, T7+GR, NuGEN, and PBL protocols with $r = (0.88 - 0.98)$. Spearman correlation in acute rejection (AR) samples was $r = (0.87 - 0.97)$. 
change for this gene set was significantly higher (P = 0.02) in biopsy than PBL (Fig. 4B). With the use of GO and hypergeometric analysis, the biological processes, molecular functions, and cellular components of the differentially expressed genes were identified. The expression pattern of these genes bore biological relevance, as many of the genes were significantly associated with immune responses (383 genes, P = 2.06e-13), response to stimuli (333 genes, P = 1.25e-9), cell death (256 genes, P = 0.048), lymphocyte proliferation (11 genes, P = 0.04), and cell cycling (362 genes, P = 0.001). Functional characterization of the genes associated with immune response showed that the immune response-specific changes largely related to changes in inflammatory response (113 genes, P = 2.73e-10), humoral immune response (74 genes, P = 0.00234), innate immune response (29 genes, P = 0.0186), acute-phase response (15 genes, P = 0.013), immune regulation (36 genes, P = 0.00782), and cellular defense response (44 genes, P = 0.06). Many genes were also associated with responses to stress (498 genes, P = 7.89e-11) and biotic stimuli (432 genes, P = 9.35e-13).

DISCUSSION
A major potential of gene expression analyses in whole blood samples for immune-mediated diseases, such as infection, autoimmune diseases and transplantation, lies in the assessment of genes regulating or mediating the functions of circulating leukocytes. In organ transplant rejection, immune cells trafficking into and affecting the graft also circulate in the periphery and may reflect or allow for monitoring of the ongoing process in the graft.

Expression profiling using whole blood RNA is a minimally invasive method for patient monitoring and thus holds promise for its application in molecular diagnostics and clinical medicine. Nevertheless, high levels of globin mRNA may confound biomarker discovery, which has promoted the development of globin reduction methods. As shown here and in other published studies (7, 15), the globin reduction method improves the detection sensitivity of Affymetrix Human Genome U133 Plus 2.0 arrays by increasing call rates from higher signal intensities and lower signal-to-noise ratios and permitting detection of additional “masked” genes. Nevertheless, the discovery of candidate biomarker genes for acute rejection diagnosis in whole blood remains difficult. This is likely due to several reasons.
gene expression from biopsy tissue with acute rejection are much lower compared with acute rejection blood RNA isolated using PAXgene collection tubes (as in this cell line with or without spiked globin mRNA and human (15) comparing these two globin reduction methods in a Jurkat bin-Reduction Kit (Affymetrix) are available. A recent study GLOBINclear (Ambion, Foster City, CA) and GeneChip Glo- 
is [equivalently to that of an identical RNA sample without globin mRNA excess. Therefore, none of these protocols can completely remove globin gene mRNA. Despite the T7+GR protocol, the remaining globin genes (9/21 still detectable after globin depletion) in our whole blood sample still continue to drive the main eigenvector in the data set.

Third, the confounding influence of globin gene expression becomes pronounced, as there is significant enrichment (P < 0.001; Table 2B) for globin genes on the Affymetrix Human Genome U133 Plus 2.0 platform. We show that application of an additional mathematical depletion of globin genes using SVD, shown to be similarly useful in other studies (2, 18), improves the efficacy of globin reduction but does not fully unmask gene expression differences for acute rejection that can be identified with confidence when the PBL analysis is done after an erythrocyte depletion protocol. Thus, despite the application of globin reduction protocols, overrepresentation of globin genes on the Affymetrix platform can be a confounder for biomarker discovery, particularly in diseases with subtle gene expression signatures.

Fourth, because molecular heterogeneity exists in transcriptional analysis of tissue renal allograft rejection, as previously shown by our group (21), a unifying peripheral inflammatory signature across all rejections may not always be detectable. Last, but not the least, the question of interobserver variability in histological end point definition for the purpose of functional genomic marker identification is a critical issue. In recently published studies (5, 17), only 77% concordance existed between experienced heart transplant pathologists in grade 3A/2R rejection and 44% in grade 2 rejection. Intra- and interobserver variability in renal pathology readouts have also been shown to be substantial (8). In our study, we have addressed this by having a single pathologist blindly review and score all biop-
sies as rejection or stable on a single day to minimize variability in pathology diagnoses.

Transcriptional analysis of PBL enhanced rejection-associated gene expression discovery in this study compared with PAXgene whole blood mRNA expression analysis (6). Previous studies have shown that neither method could recover a profile equivalent to that of an identical RNA sample without globin mRNA excess. Therefore, none of these protocols can completely remove globin gene mRNA. Despite the T7+GR protocol, the remaining globin genes (9/21 still detectable after globin depletion) in our whole blood sample still continue to drive the main eigenvector in the data set.

Acknowledgments

We are grateful to Dr. Atul Butte for critical discussions and valuable comments and to Dr. Neeraja Kambham for the blinded pathology readouts for
all biopsies. We also are grateful to our pediatric transplant nephrology team for helping with clinical sample collection.

GRANT

The study was funded by National Institute of Allergy and Infectious Diseases Grant RO1-AI-061739, awarded to M. Sarwal.

REFERENCES


