

# Urinary cell transcriptomics and acute rejection in human kidney allografts

Akanksha Verma,<sup>1,2,3,4</sup> Thangamani Muthukumar,<sup>5,6</sup> Hua Yang,<sup>5</sup> Michelle Lubetzky,<sup>5,6</sup> Michael F. Cassidy,<sup>5</sup> John R. Lee,<sup>5,6</sup> Darshana M. Dadhania,<sup>5,6</sup> Catherine Snopkowski,<sup>5</sup> Divya Shankaranarayanan,<sup>5,6</sup> Steven P. Salvatore,<sup>7</sup> Vijay K. Sharma,<sup>5</sup> Jenny Z. Xiang,<sup>8</sup> Iwijn De Vlaminck,<sup>9</sup> Surya V. Seshan,<sup>7</sup> Franco B. Mueller,<sup>5</sup> Karsten Suhre,<sup>10</sup> Olivier Elemento,<sup>1,2,3</sup> and Manikkam Suthanthiran<sup>5,6</sup>

<sup>1</sup>Department of Physiology and Biophysics, <sup>2</sup>Caryl and Israel Englander Institute for Precision Medicine, and <sup>3</sup>Institute for Computational Biomedicine, Weill Cornell Medical College, New York, New York, USA. <sup>4</sup>Tri-Institutional Training Program in Computational Biology and Medicine, Cornell University, Weill Cornell Medical College, and Sloan-Kettering Institute, New York, New York, USA. <sup>5</sup>Division of Nephrology and Hypertension, Department of Medicine, Weill Cornell Medical College, New York, New York, USA. <sup>6</sup>Department of Transplantation Medicine, NewYork-Presbyterian/Weill Cornell Medical Center, New York, New York, USA. <sup>7</sup>Division of Renal Pathology, Department of Pathology and Laboratory Medicine, and <sup>8</sup>Genomics Resources Core Facility, Department of Microbiology and Immunology, Weill Cornell Medical College, New York, New York, USA. <sup>9</sup>Meinig School of Biomedical Engineering, Cornell University, Ithaca, New York, USA. <sup>10</sup>Bioinformatics Core, Weill Cornell Medical College, Doha, Qatar.

**BACKGROUND.** RNA sequencing (RNA-Seq) is a molecular tool to analyze global transcriptional changes, deduce pathogenic mechanisms, and discover biomarkers. We performed RNA-Seq to investigate gene expression and biological pathways in urinary cells and kidney allograft biopsies during an acute rejection episode and to determine whether urinary cell gene expression patterns are enriched for biopsy transcriptional profiles.

**METHODS.** We performed RNA-Seq of 57 urine samples collected from 53 kidney allograft recipients (patients) with biopsies classified as acute T cell-mediated rejection (TCMR;  $n = 22$ ), antibody-mediated rejection (AMR;  $n = 8$ ), or normal/nonspecific changes (No Rejection;  $n = 27$ ). We also performed RNA-Seq of 49 kidney allograft biopsies from 49 recipients with biopsies classified as TCMR ( $n = 12$ ), AMR ( $n = 17$ ), or No Rejection ( $n = 20$ ). We analyzed RNA-Seq data for differential gene expression, biological pathways, and gene set enrichment across diagnoses and across biospecimens.

**RESULTS.** We identified unique and shared gene signatures associated with biological pathways during an episode of TCMR or AMR compared with No Rejection. Gene Set Enrichment Analysis demonstrated enrichment for TCMR biopsy signature and AMR biopsy signature in TCMR urine and AMR urine, irrespective of whether the biopsy and urine were from the same or different patients. Cell type enrichment analysis revealed a diverse cellular landscape with an enrichment of immune cell types in urinary cells compared with biopsies.

**CONCLUSIONS.** RNA-Seq of urinary cells and biopsies, in addition to identifying enriched gene signatures and pathways associated with TCMR or AMR, revealed genomic changes between TCMR and AMR, as well as between allograft biopsies and urinary cells.

**Authorship note:** AV, TM, and HY are co-first authors.

**Conflict of interest:** The authors have declared that no conflict of interest exists.

**Copyright:** © 2020, American Society for Clinical Investigation.

**Submitted:** July 8, 2019

**Accepted:** January 22, 2020

**Published:** February 27, 2020.

**Reference information:** *JCI Insight*. 2020;5(4):e131552.

<https://doi.org/10.1172/jci.insight.131552>

## Introduction

Although kidney transplantation is the treatment of choice for patients afflicted with end-stage kidney disease (1), acute rejection with a cumulative incidence of 10%–27% within 3 years of transplantation limits its benefits (2–5) and is an important contributor to allograft failure (6–9). Currently, 2 major categories of acute rejection, acute T cell-mediated rejection (TCMR) and acute antibody-mediated rejection (AMR) are diagnosed based on histological features displayed by the allograft (10, 11). Whether cellular and molecular mechanisms are unique to each type of acute rejection or shared in part between these 2 major subtypes of acute rejection remains mostly unresolved.

The invasive kidney allograft biopsy remains the standard-of-care approach for assessing kidney allograft status; however, the biopsy procedure is not without risks, such as bleeding, arteriovenous fistula formation, and — in rare cases — even death (12–14). Much effort has therefore been invested to develop noninvasive tools to ascertain kidney allograft status. In this regard, our laboratory developed preamplification enhanced real-time quantitative PCR (customized RT-qPCR) assays for the absolute quantification of mRNAs, and we reported that the urinary cell mRNA profiles are diagnostic and prognostic of TCMR in the kidney allograft (15–22). However, only a limited number of transcripts, albeit mechanistically sound, were evaluated in our studies, and — very importantly — the transcript levels were primarily associated with biopsy histology, leaving unanswered the mechanistic issue of whether the urinary cell mRNA expression patterns reflect accurately the gene expression patterns within an acutely rejecting kidney allograft.

To address the gaps in knowledge, we performed RNA sequencing (RNA-Seq) of urinary cells and kidney allograft biopsies and investigated whether the urinary cell gene expression patterns reflect TCMR or AMR biopsy gene signatures. We also examined whether TCMR and AMR are associated with a unique, as well as shared, gene expression pattern and biological pathway. Although much mechanistic information has been gained by gene expression profiling using the microarray platforms, we chose RNA-Seq to profile urinary cells and allograft biopsy specimens, since RNA-Seq has a number of advantages over hybridization-based microarrays, such as resolution at the single base level, lower background noise, and measurement of mRNA abundance over a broader dynamic range compared with microarrays (23–25).

RNA-Seq and bioinformatics of genome-wide expression profiles of urinary cells elucidated gene signatures distinguishing kidney graft recipients with TCMR or AMR from those without acute rejection in their biopsies. Intriguingly, we observed not only TCMR- or AMR-specific gene expression patterns and biological pathways, but also shared gene expression patterns and pathways. Our analysis demonstrated that urinary cell gene expression patterns associated with TCMR or AMR are enriched for genes expressed in TCMR biopsies or AMR biopsies. Gene expression-based cell type enrichment analysis revealed a diverse cellular landscape with an enrichment of immune cell types in urinary cells compared with biopsies.

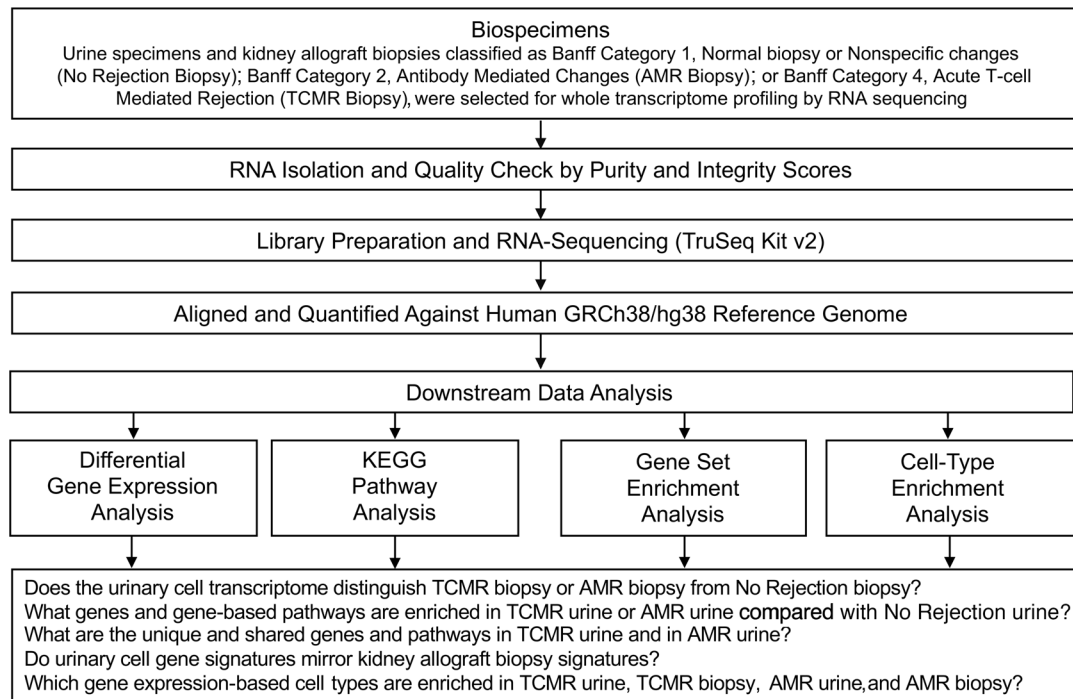
## Results

*Study design.* Figure 1 illustrates the overall study design and the key issues addressed by whole transcriptome profiling of urinary cells and kidney allograft biopsies with RNA-Seq. Urine samples and kidney allograft biopsies were selected from our biorepository to include 3 major Banff diagnostic categories (10): Banff Category 1, normal biopsy or nonspecific changes (designated in this study as No Rejection [NR]); Banff Category 2, antibody-mediated changes (designated in this study as AMR); or Banff Category 4, acute TCMR (designated in this study as TCMR).

*Characteristics of the urine RNA-Seq cohort.* The characteristics of the kidney allograft recipients (patients) whose urine samples were RNA sequenced and included in data analysis (urine RNA-Seq cohort) are summarized in Supplemental Table 1 (Supplemental material available online with this article; <https://doi.org/10.1172/jci.insight.131552DS1>). Among the 57 urine samples collected from 53 patients, 22 were from 20 patients with TCMR biopsies, 8 were from 8 patients with AMR biopsies, and 27 were from 25 patients with NR biopsies. In this investigation, the urine samples collected at the time of TCMR biopsy are designated as TCMR urine, the urine samples collected at the time of AMR biopsy are designated as AMR urine, and the urine samples collected at the time of NR biopsy are designated as NR urine.

All TCMR and AMR biopsies were for-cause biopsies and were performed to determine the basis for graft dysfunction, and 24 of 27 NR biopsies were surveillance biopsies (Supplemental Table 1). Recipient and donor information, induction and maintenance immunosuppression, time from transplant to biopsy, presence or absence of donor specific antibodies (DSA) before transplantation and at the time of allograft biopsy, and Banff acute and chronic scores of the biopsies are summarized in Supplemental Table 1. Among the NR biopsies, the acute Banff scores for tubulitis (t), interstitial inflammation (i), glomerulitis (g), and peritubular capillaritis (ptc) were all 0 in 23 of 27 biopsies. Among the remaining 4, the t score was 1, i score was 1, and ptc score was 1 in 1 biopsy; 1 biopsy had an i score of 1; and 2 biopsies had ptc score of 1. Thus, subclinical inflammation was nonexistent in 23 of 27 NR biopsies and minimal in others.

*Characteristics of kidney allograft biopsy RNA-Seq cohort.* The characteristics of the kidney allograft recipients from whom kidney allograft biopsies were obtained and RNA sequenced (kidney allograft biopsy cohort), summarized in Supplemental Table 2. Among the 49 biopsies from 49 kidney allograft recipients, 12 biopsies were classified as TCMR, 17 as AMR, and the remaining 20 as NR biopsies. All TCMR biopsies

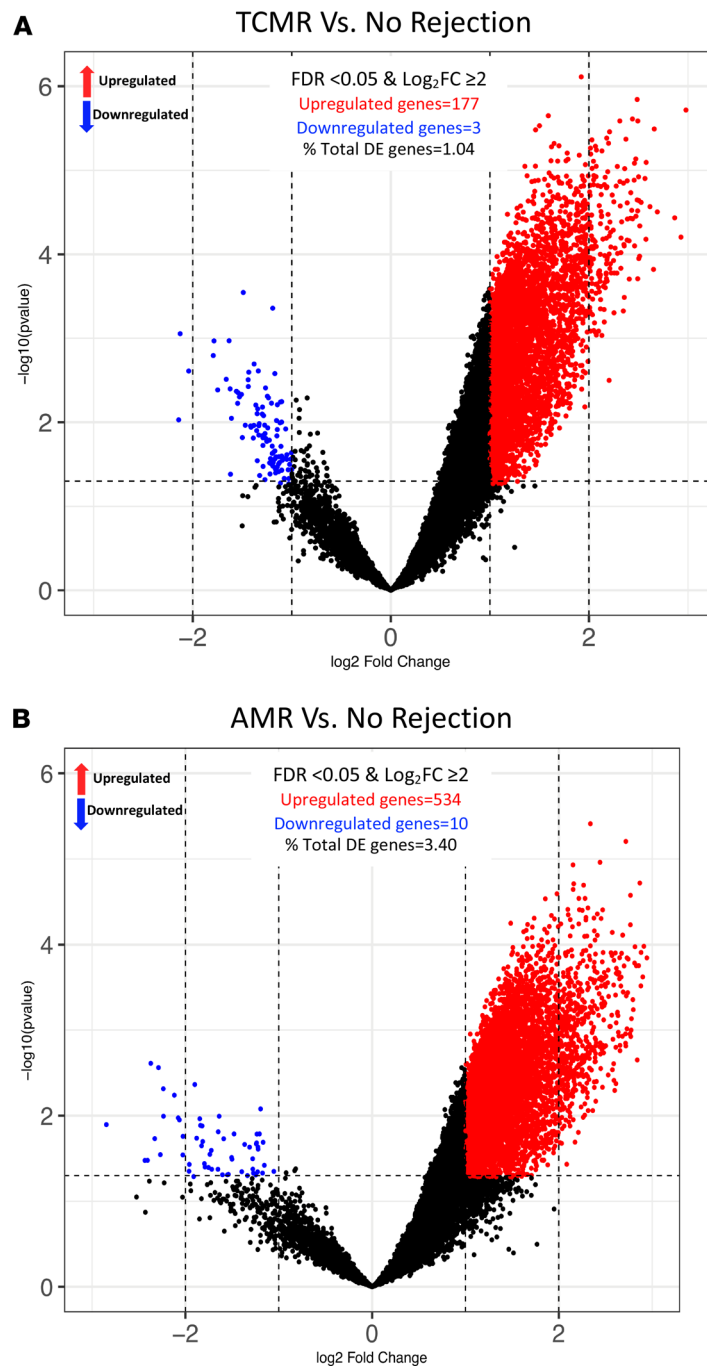


**Figure 1. Flowchart for RNA sequencing of urine samples and kidney allograft biopsy specimens.** Urine samples and kidney allograft biopsy specimens were selected from the Weill Cornell biorepository to include 3 major diagnostic categories: Banff Category 1, normal biopsy or nonspecific changes (designated in this report as No Rejection); Banff Category 2, antibody-mediated changes (AMR); or Banff Category 4, T cell-mediated rejection (TCMR). Total RNA was isolated from urinary cells and from the kidney allograft biopsies, and the quantity, purity, and integrity of the isolated RNA were determined. TrueSeq sample preparation kit v2 was used to prepare individual cDNA libraries, and RNA sequencing was performed using Illumina sequencer; the sequence read data were stored in FASTQ format. Sequenced reads were aligned to the human reference genome GRCh38/hg38 using STAR aligner. Aligned reads were quantified against the reference annotation to obtain fragments per kilobase per million (FPKM) and raw counts using CuffLinks (v2.2.1) and HTSeq, respectively. Among the 70 urine specimens selected for RNA sequencing, sequence reads from 13 were excluded for downstream analysis based on less than 20% alignment to the human reference genome GRCh38/hg38 and RNA integrity number (RIN) lower than 2 (RNA quality thresholds). Sequencing data from the remaining 57 urine specimens from 53 kidney allograft recipients (patients) with biopsies classified as No Rejection biopsy ( $n = 27$  biopsies from 25 patients), AMR ( $n = 8$  biopsies from 8 patients), or TCMR ( $n = 22$  biopsies from 20 patients) were included in downstream data analysis. RIN and sequence reads from all 49 kidney allograft biopsies (No Rejection biopsies,  $n = 20$  biopsies from 20 patients), AMR ( $n = 17$  biopsies from 17 patients), or TCMR ( $n = 12$  biopsies from 12 patients) from 49 kidney allograft recipients met RNA quality thresholds and were included in downstream data analysis. Among the urine and biopsy samples included in data analysis, 11 were paired samples (i.e., urine and biopsy were from the same kidney allograft recipient).

and all AMR biopsies were for-cause biopsies and were performed to determine the basis for graft dysfunction, and all NR biopsies were surveillance biopsies (Supplemental Table 2). Recipient and donor information, induction and maintenance immunosuppression, time from transplant to biopsy, presence or absence of DSA before transplantation and at the time of allograft biopsy, and Banff acute and chronic scores of the biopsies are provided in Supplemental Table 2. All 17 AMR biopsies were C4d<sup>+</sup>, and none of the TCMR or NR biopsies were C4d<sup>+</sup>. There were no borderline or mixed rejection biopsies included in this cohort. Among the 49 biopsies included in this study, 11 biopsies (4 TCMR, 3 AMR, and 4 NR biopsies) were from 11 patients who also provided urine specimens (paired biopsy and urine samples) for RNA-Seq.

**RNA-Seq of urinary cells.** Urine is rich in DNA and RNA hydrolyzing enzymes (26). Due to the expected fragmentation of RNA in urinary cells, we applied the following quality control (QC) metrics: RNA Integrity Number (RIN) of 2 or greater and a minimum of 20% unique reads aligned to a human reference genome before downstream data analysis. The positive correlation ( $r = 0.45$ ,  $P = 0.00043$ ) between RIN score and alignment percentage is shown in Supplemental Figure 1. Supplemental Figure 2 shows 3-dimensional plots of unsupervised principal component analysis (PCA) of the urinary cell RNA-Seq data, and the scree plots provide the eigenvalues (proportion of variance explained) and the number of principal components explaining the variance in gene expression.

**Differential gene expression analysis.** The volcano plots show differences in urinary cell gene expression between TCMR urine and NR urine (Figure 2A) and between AMR urine and NR urine (Figure 2B).



**Figure 2. Differential gene expression analysis.** Total RNA isolated from 57 urine specimens collected at the time of 57-kidney allograft biopsies from 53 kidney allograft recipients (22 urine specimens from 20 patients with 22 TCMR biopsies, 8 urine specimens from 8 patients with 8 AMR biopsies, and 27 No Rejection urine specimens from 25 patients with 27 No Rejection biopsies) were RNA sequenced. Differential gene expression analysis was performed using the limma package in R. (A and B) Volcano plots display differences in urinary cell gene expression between TCMR urine and No Rejection urine (A) and between AMR urine and No Rejection urine (B). The x axis depicts the log<sub>2</sub> fold change (FC) in gene expression and the y axis is the -log<sub>10</sub> P value. A positive FC (red dots) denotes increased expression in TCMR urine or AMR urine versus No Rejection urine. A negative FC (blue dots) denotes increased expression in No rejection urine versus TCMR urine or AMR urine. With log<sub>2</sub> FC ≥ 1 and FDR-adjusted P < 0.1 as the thresholds for differential expression, 4702 genes (27% of total genes) were differentially expressed in TCMR urine versus No Rejection urine, and 6516 genes (41%) were differentially expressed in AMR urine versus No Rejection urine. Supplemental Table 3 lists the differentially expressed genes. A total of 180 genes (1.04%) were differentially expressed in TCMR urine versus No Rejection urine, and 534 genes (3.40%) were differentially expressed in AMR urine versus No Rejection urine with log<sub>2</sub>FC ≥ 2 and FDR-adjusted P < 0.05 as the thresholds for differential expression.

We identified 4702 genes (27% of all genes) as differentially expressed in TCMR urine versus NR urine, and 6516 genes (41%) as differentially expressed in AMR urine versus NR urine with log<sub>2</sub> Fold Change [FC] > 1 and FDR < 0.1 as the thresholds for differential expression (Supplemental Table 3). On further stringent curation (log<sub>2</sub> FC > 2 and FDR < 0.05), 180 genes were differentially expressed (177 upregulated and 3 downregulated) between TCMR urine and NR urine, and 544 genes were differentially expressed (534 upregulated and 10 downregulated) between AMR urine and NR urine (Figure 2), suggesting a gene signature representative of the genome-wide transcriptional changes in the urinary cells of kidney allograft recipients with TCMR biopsies or AMR biopsies.

Figure 3 shows heatmaps based on 180 differentially expressed genes in TCMR urine versus NR urine (Figure 3A) and 544 differentially expressed genes in AMR urine versus NR urine (Figure 3B). Seven of the 22 TCMR urine samples clustered with NR urine and 2 of the 8 AMR urine samples clustered with NR urine samples. We investigated whether the presence of ptc or ptc and glomerulitis

(microvascular inflammation) in TCMR biopsies affected the clustering. Supplemental Figure 3 shows that neither the ptc score nor the presence or absence of microvascular inflammation in TCMR biopsies affected clustering of TCMR urine samples. Nine of 22 NR urine samples were associated with the presence of DSA. Hierarchical clustering analysis did not show clustering of NR urine samples based on the basis of presence or absence of DSA (Supplemental Figure 4).

We identified unique genes expressed in TCMR urine and in AMR urine and shared genes expressed in both TCMR urine and AMR urine. Among the 180 genes differentially expressed between TCMR urine versus NR urine and the 544 genes differentially expressed between AMR urine versus NR urine, 127 genes were shared between TCMR urine and AMR urine signatures (Figure 3C). Supplemental Table 4 lists the shared 127 genes, as well as the 50 genes unique to TCMR urine and the 407 genes unique to AMR urine.

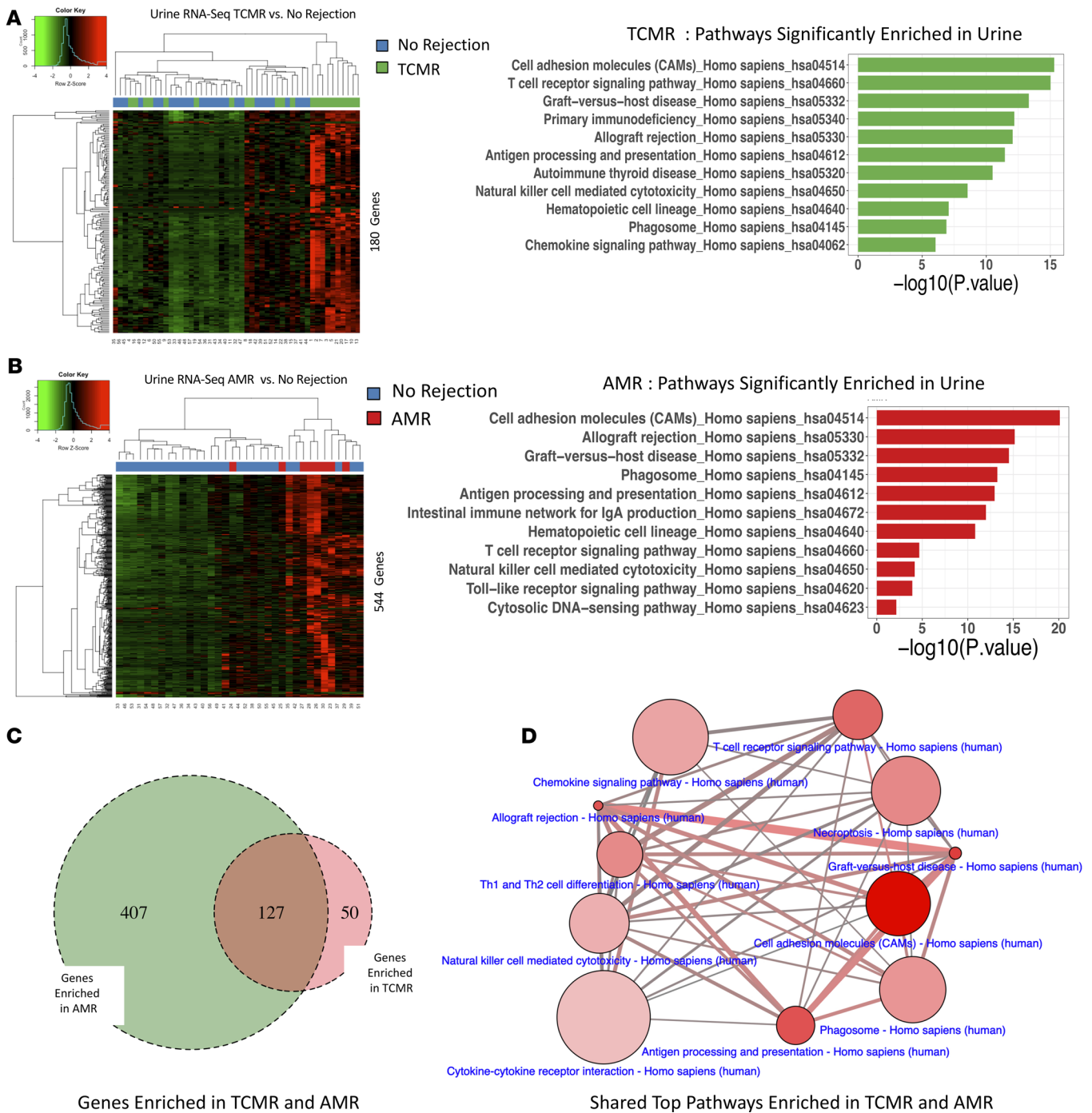
We performed a pathway analysis using the Kyoto Encyclopedia of Genes and Genomes (KEGG; <https://www.genome.jp/kegg/>) molecular pathways database to functionally annotate the observed gene expression changes. We identified key immune-related and allograft rejection pathways significantly enriched in TCMR urine versus NR urine (Figure 3A) and in AMR urine versus NR urine (Figure 3B), and for the pathways enriched and shared between TCMR urine and AMR urine (Figure 3D). Chemokine-signaling pathway and pathways for T cell receptor signaling, allograft rejection, graft-versus-host disease, Th1 and Th2 cell differentiation, necroptosis, NK cell-mediated cytotoxicity, cell adhesion molecules, cytokine-to-cytokine receptor interaction, phagosome, and antigen processing and presentation were significantly enriched and shared between TCMR urine and AMR urine. Primary immunodeficiency pathway was uniquely enriched in TCMR urine and TLR signaling pathway, and cytosolic DNA-signaling pathways were uniquely enriched in AMR urine.

*Expression of chemokines in urinary cells.* Chemokines have been reported to play a key role in the recruitment of immune cells to the allograft. We used a curated list of 45 chemokines to investigate which of the chemokine mRNAs are overexpressed in TCMR urine versus NR urine and in AMR urine versus NR urine. We overlapped the list of 45 chemokines with the 180 differentially expressed genes in TCMR urine and the 544 differentially expressed genes in AMR urine. Our analysis revealed that mRNAs for 4 chemokines — CXCL9, CXCL10, CXCL11, and CCL5 — are positively enriched in TCMR urine versus NR urine. Similarly, mRNAs for 6 chemokines — CXCL5, CXCL9, CXCL10, CCL2, CCL5, and CCL18 — were positively enriched in the AMR urine versus NR urine.

*Development of urinary cell diagnostic signatures.* To develop a urinary cell diagnostic signature based on RNA-Seq data, we performed penalized multinomial logistic lasso regression analysis to estimate generalized linear models for discriminating TCMR from NR. The normalized expression values in fragments per kilobase of transcript per million reads (FPKM) of the 180 genes differentially expressed between TCMR urine and NR urine were tested as predictors in the regression analysis. Our analysis identified a 13-gene urinary cell gene signature that distinguished TCMR from NR with a high degree of accuracy (Supplemental Figure 5). The area under the receiver-operating characteristic curve (AUC) of the urinary cell gene signature developed using RNA-Seq data was 0.92 (95% CI, 0.85–0.99). The 13 genes constituting the diagnostic signature, along with their regression coefficients, are listed in Supplemental Figure 5.

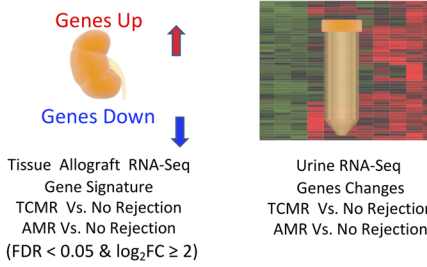
We next compared the diagnostic performance of 13-gene RNA-Seq signature with the diagnostic performance of RT-qPCR-based CTOT-04 (Clinical Trials in Organ Transplantation 04) urinary cell 3-gene TCMR diagnostic signature of 18S-normalized CD3E mRNA, 18S-normalized CXCL10/IP-10 mRNA, and 18S rRNA (21). Among the 49 urine samples that were RNA sequenced (22 TCMR urine and 27 NR urine) and included in the regression analysis, sufficient RNA from 46 urine samples (22 TCMR urine and 24 NR urine) was available for the measurement of CD3E mRNA, CXCL10/IP-10 mRNA, and 18S rRNA using customized RT-qPCR assays. Supplemental Table 5 shows the absolute copy numbers of CD3E mRNA, CXCL10/IP-10 mRNA, and 18S rRNA in TCMR urine and in NR urine. The AUC of urinary cell 3-gene signature was 0.73 (95% CI, 0.59–0.88) and was significantly lower than the AUC of the RNA-Seq-based urinary cell signature ( $P = 0.03$ , DeLong's test; Supplemental Figure 5).

We also examined whether TCMR could be distinguished from AMR using RNA-Seq data. In order to develop a prediction model discriminating TCMR from AMR, we performed regression analysis using, as predictors, the 50 uniquely overexpressed genes in TCMR urine versus NR urine and the 407 uniquely overexpressed genes in AMR urine versus NR urine. Penalized multinomial logistic lasso regression analysis was used to compute a signature for discriminating TCMR from AMR. This exploratory analysis identified a 13-gene set in urine that discriminated perfectly (AUC 1.0) AMR from TCMR. Supplemental Table 6 lists



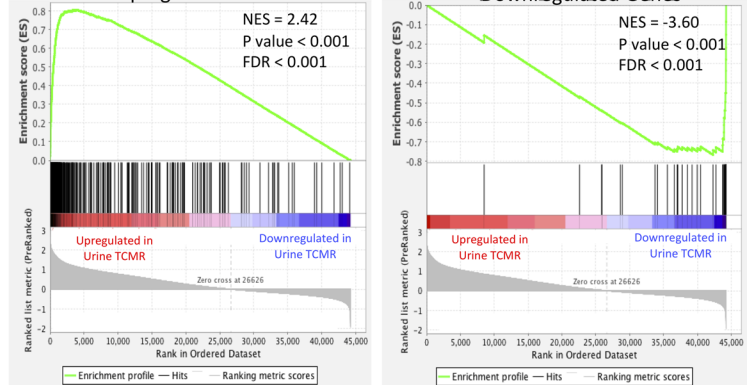
**Figure 3. Semisupervised heatmap and pathway analysis of differentially expressed genes in urinary cells.** (A and B) Semisupervised hierarchical clustering analysis and KEGG pathways based on the 180 genes that were differentially expressed in TCMR urine versus No Rejection urine (A) and the 544 genes that were differentially expressed in AMR urine versus No Rejection urine (B), with  $\log_2FC \geq 2$  and  $FDR < 0.05$  as the threshold for differential expression. Supervised gene name-based pathway analysis was performed using ENRICH. Pathways with  $FDR < 0.05$  were considered enriched. (C) Venn diagram shows unique and shared genes upregulated (enriched) in TCMR urine versus No Rejection urine and in AMR urine versus No Rejection urine. The abundance of 177 genes was higher in TCMR urine versus No Rejection urine, and the abundance of 534 genes was higher in AMR urine versus No Rejection urine. Among these genes, 127 formed an overlapping set of genes shared between the TCMR urine and AMR urine signatures. Supplemental Table 4 lists the shared 127 genes. (D) The shared top pathways enriched in TCMR urine and AMR urine based on the set of 127 shared genes is visualized using Consensus Pathway Database Networks.

**A** Kidney Allograft RNA-Seq Signature Vs. Urine RNA-Seq

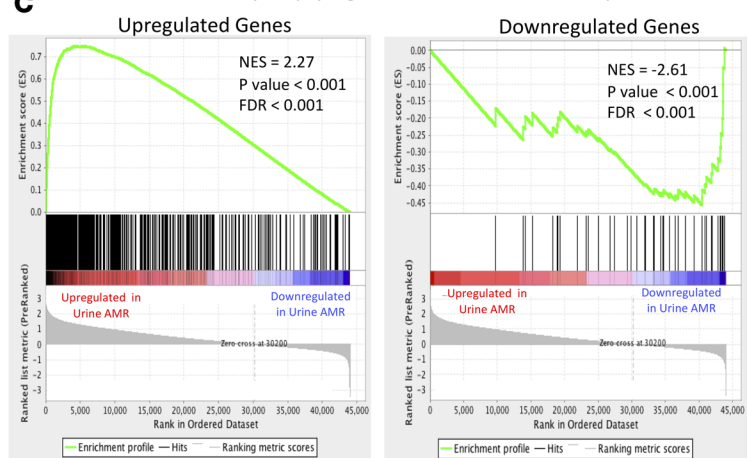


Gene Set Enrichment Analysis

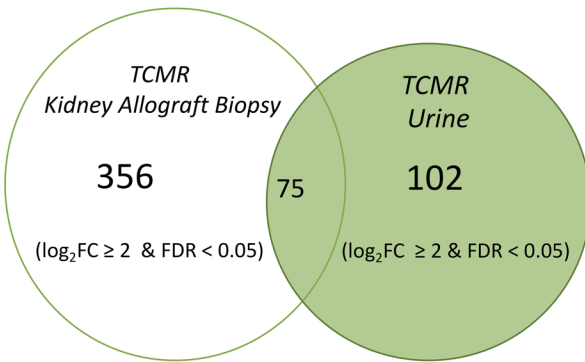
**B** TCMR: Kidney Biopsy Signature is enriched in Urinary Cells  
Upregulated Genes



**C** AMR: Kidney Biopsy Signature is enriched in Urinary Cells  
Upregulated Genes

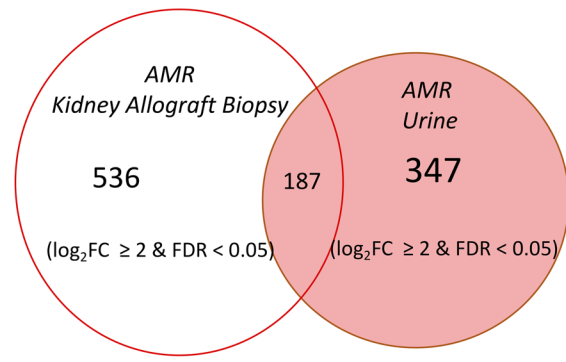


**D**



TCMR Shared Enriched Genes

**E**



AMR Shared Enriched Genes

**Figure 4. Kidney allograft biopsy gene signatures are enriched in urinary cells.** Kidney allograft biopsy signatures derived by RNA sequencing of 49 biopsies from 49 patients. **(A)** Gene Set Enrichment Analysis was performed to compare biopsy gene signatures with urinary cell gene expression patterns. **(B)** TCMR biopsy up- or downregulated signatures were enriched in urinary cells (both  $FDR < 0.001$ ). The ranked list of genes in the biopsy (x axis) and the enrichment score (ES) (y axis) are shown. A positive ES indicates that the top-ranked genes in the biopsy are enriched in urinary cells. The top portion of the plot shows the ES (green line). The ES for this gene set is the score at the peak of the plot. The middle portion shows where urinary cell genes appear in the ranked list of biopsy genes. The bottom portion shows the value of the ranking metric moving down the list of ranked genes and goes from positive (correlation with TCMR) to negative (correlation with No Rejection). The normalized enrichment score (NES) accounts for differences in gene set size and correlations between the urine and biopsy gene sets. The FDR is the estimated probability that a gene set with a given NES represents a false-positive finding. Enrichment of top differentially expressed upregulated (Positive NES) and downregulated (Negative NES) biopsy signatures in urine are shown. **(C)** AMR biopsy upregulated and downregulated signatures were enriched in urinary cells (both  $FDR < 0.001$ ). Positive NES and Negative NES biopsy signatures in urine are shown. **(D)** Venn diagram depicts top differentially expressed genes in TCMR biopsy versus No Rejection biopsy and TCMR urine versus No Rejection urine; 75 genes were shared between the TCMR biopsy and TCMR urine. **(E)** Venn diagram depicts top differentially expressed genes in AMR biopsy versus No Rejection biopsy and in AMR urine versus No Rejection urine; 187 genes were shared between the AMR biopsy and AMR urine.

the 13 genes, along with their regression coefficients. In view of the small sample size and the large number of variables (genes) included in the regression analysis, this prediction model and the model discriminating TCMR from NR should be validated using an independent and larger cohort of kidney allograft recipients.

*Comparison of transcriptional profiles between urinary cells and kidney allograft biopsies.* We determined whether the transcriptional changes in urinary cells reflect the global gene expression changes in kidney allograft biopsies. We used Gene Set Enrichment Analysis (GSEA) (27) to evaluate whether the gene signatures identified in the kidney allograft biopsies are enriched in urine (Figure 4A). Our analyses revealed that the TCMR biopsy versus NR biopsy upregulated gene signature is significantly upregulated in the TCMR urine versus NR urine (Figure 4B, normalized enrichment score [NES] = 2.42,  $P < 0.001$ , FDR < 0.001). Similarly, the upregulated gene signature in AMR biopsy versus NR biopsy was significantly upregulated in AMR urine versus NR urine (Figure 4C, NES = 2.27,  $P < 0.001$ , FDR < 0.001). The downregulated gene signature in TCMR biopsies versus NR biopsies was enriched in TCMR urine versus NR urine (Figure 4B, NES = -3.60,  $P < 0.001$ , FDR < 0.001), and the downregulated signature in AMR biopsies versus NR biopsies was enriched in AMR urine versus NR urine (Figure 4C, NES = -2.61,  $P < 0.001$ , FDR < 0.001).

We determined the number of genes that overlapped between urinary cells and kidney allografts. Among the 177 upregulated genes in TCMR urine versus NR urine, 75 (42%) genes overlapped between TCMR urine and TCMR biopsy (Figure 4D). Among the 534 genes upregulated in AMR urine versus NR urine, 187 (35%) genes overlapped between AMR urine and AMR biopsy (Figure 4E). Supplemental Table 7 lists the genes unique to kidney allograft biopsies, those unique to urine, and the genes shared between the biopsies and the urine samples.

Data illustrated in Figure 4 comparing transcriptional profiles of biopsies and urinary cells include biopsies and urine samples collected from the same patient (paired samples) and biopsies and urine samples obtained from different patients (unpaired samples). We determined whether the correlation exists not only when the analysis is restricted to paired samples, but also when the analysis is restricted to unpaired samples. Supplemental Figure 6 shows the significant correlation between paired TCMR biopsy and TCMR urine sample ( $P < 0.001$ , Pearson correlation,  $n = 4$  paired samples), paired AMR biopsy and AMR urine sample ( $P < 0.001$ ,  $n = 3$  paired samples), and paired NR biopsy and NR urine sample ( $P < 0.001$ ,  $n = 4$  paired samples). We then performed comparison of transcriptional profiles restricted to unpaired samples. Our analysis showed significant enrichment of the TCMR biopsy gene signature in TCMR urine and significant enrichment of the AMR biopsy gene signature in AMR urine (Supplemental Figure 7). Both upregulated and downregulated genes in TCMR biopsies versus NR biopsies showed significant enrichment in unpaired TCMR urine versus NR urine (NES = 3.12,  $P < 0.001$ , and FDR < 0.001 for the upregulated genes and NES = -4.01,  $P < 0.001$ , and FDR < 0.001 for the downregulated genes). In a similar fashion, both upregulated and downregulated genes in AMR biopsies versus NR biopsies showed significant enrichment in unpaired AMR urine versus NR urine (NES = 3.03,  $P < 0.001$ , and FDR < 0.001 for the upregulated genes and NES = -1.95,  $P < 0.001$ , and FDR < 0.001 for the downregulated genes). Thus, transcriptional changes in urinary cells reflected transcriptional changes in kidney allograft biopsies not only when urine and biopsies are from the same patient, but also when the biopsies and urine are from different patients.

*Shared transcriptional profiles in TCMR urine or AMR urine.* We performed GSEA to determine whether the observed TCMR or AMR gene enrichment in urinary cells is unique to TCMR or AMR biopsy diagnosis or whether they are shared. Supplemental Figure 8 shows that the TCMR biopsy gene signature is enriched in AMR urine (NES = 2.033,  $P < 0.001$ , FDR < 0.001 for the genes upregulated in TCMR biopsies and NES = -3.15,  $P < 0.001$ , FDR < 0.001 for the genes downregulated in TCMR biopsies) and the AMR biopsy gene signature is enriched in TCMR urine (NES = 2.68,  $P < 0.001$ , FDR < 0.001 for the genes upregulated in AMR biopsies and NES = -3.102,  $P < 0.001$ , FDR < 0.001 for the genes downregulated in AMR biopsies).

Prior microarray-based studies have identified molecular signatures of TCMR, AMR, IFN- $\gamma$ , endothelial associated transcripts (ENDAT), or DSA (28). We investigated whether the microarray-based gene signatures are enriched in our RNA-Seq data using GSEA. Our biopsy RNA-Seq data show that the previously reported TCMR, AMR, and IFN- $\gamma$  signatures are enriched in TCMR biopsy versus NR biopsy transcriptional profile and that these signatures and the ENDAT and DSA signatures are enriched in AMR biopsy versus NR biopsy transcriptional profile (Supplemental Figure 9).

Similar comparisons performed on urine RNA-Seq data show that the previously reported ENDAT, TCMR, AMR, IFN- $\gamma$ , and DSA signatures are enriched in TCMR urine versus NR urine transcriptional profiles and that ENDAT, TCMR, AMR, and IFN- $\gamma$  signatures — but not the DSA signature — are enriched in AMR urine versus NR urine transcriptional profiles (Supplemental Figure 10).



*Enrichment of immune cell types in urinary cell RNA-Seq profiles.* Because urinary cell RNA-Seq profile is likely a mixture of RNA from multiple cell types such as graft infiltrating cells and kidney parenchymal cells, we used bulk RNA cell type enrichment analysis using the gene expression data to recover the identity of the cell types found in TCMR urine and in AMR urine. Using a gene signature expression-based cell type enrichment tool xCell (29), cell type enrichment scores (ES) across 64 immune and stromal cell types were obtained for both urinary cells and kidney allograft specimens. Our data analysis demonstrated that there were 28 cell types differentially enriched in TCMR urine versus NR urine ( $FDR < 0.1$ ), with 24 of 28 cell types positively enriched and the remaining 4 cell types negatively enriched (Figure 5A). To further confirm the observed immune gene/cell type enrichment in TCMR urine versus NR urine, we performed customized RT-qPCR assays to validate a panel of mRNAs contributing to the immune score. Figure 5B shows that urinary cell levels of mRNA for CD3D ( $P < 0.01$ , Mann-Whitney  $U$  test), CD8B ( $P < 0.01$ ), CD27 ( $P < 0.001$ ), CXCL9 ( $P < 0.05$ ), CXCL13 ( $P < 0.05$ ), CXCR1 ( $P < 0.05$ ), IL-2RA ( $P < 0.05$ ), IFN- $\gamma$  ( $P < 0.01$ ), IL-10 ( $P < 0.01$ ), and FOXP3 ( $P < 0.05$ ) are all higher in TCMR urine versus NR urine.

Seventeen cell types were quantitated following analysis of AMR urine versus NR urine, but only 1 cell type — plasmacytoid DCs (pDCs) — was differentially enriched at  $FDR < 0.1$  in AMR urine versus NR urine (Supplemental Figure 11). Importantly, the overall immune score for TCMR urine and the immune score for AMR urine were significantly higher than the immune score for NR urine (Supplemental Figure 12).

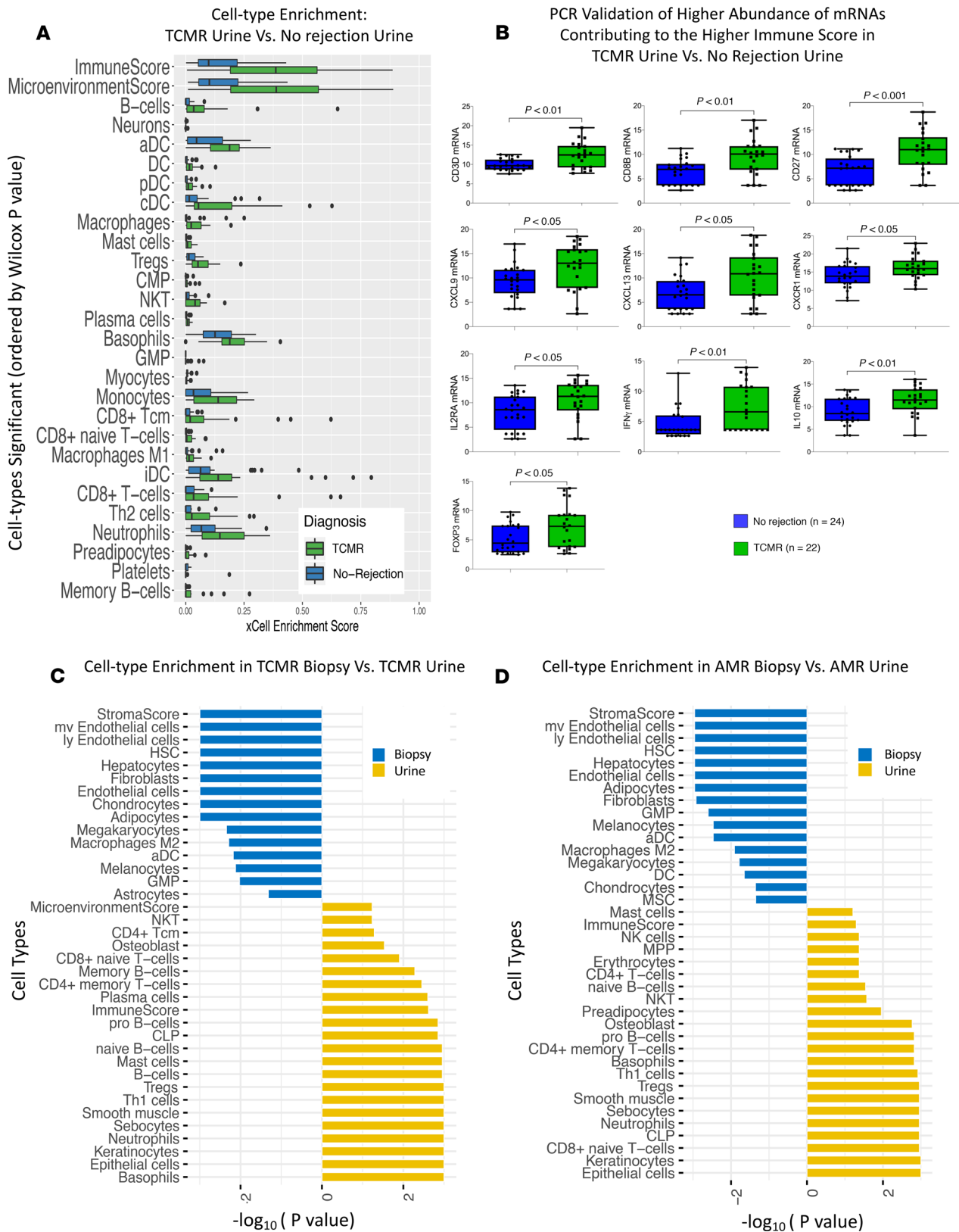
We compared cell type ES between TCMR urine and TCMR biopsy and between AMR urine and AMR biopsy. The rationale for these comparisons was to assess whether urinary cell populations inferred by mRNA expression patterns are reflective of kidney allograft cell populations. Our analyses identified that the urinary cells are enriched for immune cell types as compared with kidney allograft biopsies. Across both rejection phenotypes, the urine showed enhanced enrichment of various immune cell types, indicating a stronger and more diverse immune landscape signal in urinary cells compared with the kidney allograft (Figure 5, C and D). On the other hand, the kidney allografts showed enrichment of select stromal cell types, and the stromal cell score was higher in the biopsies compared with urinary cells.

## Discussion

Unbiased whole transcriptome profiling by RNA-Seq of urinary cells and bioinformatics enabled the discovery of a 180-gene set including multiple genes that have not previously been associated with TCMR. In a similar fashion, a 544-gene set including multiple genes that have not been associated with AMR has been identified by RNA-Seq of urinary cells. The enrichment of biopsy gene signatures in urinary cells and the enrichment of immune cell types in the urine compared with kidney allograft biopsies support urine as a suitable surrogate for the kidney allograft biopsy for the diagnosis of TCMR and AMR in human kidney allografts.

Allograft biopsies have thus far provided key insights regarding immune mechanisms of acute rejection; however, repetitive biopsies to capture the dynamics of the immune process is neither practical nor safe. Another concern is the extent to which the histological features in a biopsy core are representative of allograft pathology (30). In light of these challenges, our demonstration that urine gene expression patterns reflect and reinforce immune pathways previously identified using preclinical transplant models or by profiling the allograft suggests that urine profiling may help assess noninvasively the dynamics of antiallograft immune repertoire.

Key mechanistic insights regarding acute rejection emerged from whole transcriptome profiling of urinary cells. T cell receptor signaling pathway was significantly enriched during TCMR and AMR and suggested that not only TCMR, but also AMR involves T cell signaling. Enrichment of antigen processing and enrichment of TCR signaling pathways provided a biologic basis for allograft recognition, T cell stimulation, and immune rejection. Enrichment of chemokine signaling and cytokine-to-cytokine receptor pathways offered a basis for the homing and expansion of donor reactive and graft infiltrating cells. The finding that genes encoding cell adhesion proteins are enriched during an episode of acute rejection provides a mechanistic basis for interstitial infiltration and tubulitis, the histological hallmarks of TCMR, and for peritubular capillaritis and glomerulitis, the histological characteristics of AMR. Cytolytic mechanisms have been recognized to contribute to allograft damage, and our findings that genes encoding NK cell-mediated cytotoxicity are enriched suggest a mechanism for kidney allograft injury during an episode of TCMR. Phagosome maturation is an important component of immunity, and the enrichment of the phagosome pathway genes bring into focus the participation of not only the adaptive immune system, but also the innate immune system in the acute rejection process.



**Figure 5. Cell type enrichment analysis using xCell.** Cell type enrichment analysis of the RNA sequencing data determined using xCell, a bioinformatics tool that generates cell type enrichment scores based on gene expression data for 64 immune and stromal cell types. (A) The x axis depicts the xCell enrichment score, and the y axis lists 26 of the 64 cell types that were differentially enriched (FDR < 0.1, Wilcoxon test with Benjamini-Hochberg correction) in TCMR urine versus No Rejection urine. Twenty-two of the 26 cell types were positively enriched, and 4 were negatively enriched in TCMR urine versus No Rejection urine. Box plots of the immune score (composite score of immune cell types) and the microenvironment score (composite scores of immune cell types and stromal cell types) are also shown. (B) Box plots depict the RT-qPCR validation of differential abundance of several mRNAs contributing to the higher

immune score in TCMR urine versus No Rejection urine. mRNA abundance was quantified using customized RT-qPCR assays and shown as  $\log_2$  copies per  $\mu\text{g}$  of total RNA in the urinary cells on the y axis. Number of No Rejection urine samples and the number of TCMR urine samples are shown in parenthesis. *P* values of pairwise group comparisons are based on the Mann-Whitney *U* test. **(C)** Cell type enrichment scores of TCMR biopsy compared with the cell type enrichment scores of TCMR urine. **(D)** Cell type enrichment scores of AMR biopsy compared with the cell type enrichment scores of AMR urine. In **C** and **D**, the y axis lists only differentially enriched cell types (FDR-adjusted  $P < 0.1$ , Wilcoxon test with Benjamini-Hochberg correction), and the x axis shows the  $-\log_{10}$  *P* values. The blue bars represent cell types significantly enriched in the kidney allograft biopsies compared with urinary cells (higher score in the biopsies), and yellow bars represent cell types significantly enriched in the urinary cells compared with the kidney allograft (higher score in urinary cells). Stromal cell types were enriched in biopsy compared with urine, and immune cell types were enriched in urine compared with biopsy.

A major advantage of RNA-Seq is its ability to identify potentially novel biomarkers, expressed at very low or high levels, that otherwise may not be discovered using microarray platforms (31). Genes listed in Supplemental Tables 3 and 4 show that there are hundreds of potentially novel genes associated with the diagnosis of TCMR or AMR. In the multicenter CTOT-04 study, we discovered and validated that the 18S rRNA normalized urinary cell levels of mRNA for CD3E, CXCL10/IP-10, granzyme B, and perforin are diagnostic of TCMR (21). These mRNAs were indeed among the top 180 differently expressed genes in TCMR urine versus NR urine. Several additional mRNAs such as mRNA for CD2, CD8A, CCL5, GZMA, NKG7, and CTLA4 were found in this study to be diagnostic of TCMR, as well as AMR. CD2 is expressed on the T cell surface, and signaling via CD2 results in T cell activation, whereas CD2 blockade results in allograft tolerance (32–34). CD8 is a coreceptor for the T cell receptor and binds major histocompatibility complex class I proteins. CD8<sup>+</sup> T cells play a variety of roles in allograft rejection, and CD8 has not been evaluated as a urinary biomarker of acute rejection. CCL5 has been shown to play a role in acute rejection; however, its utility as a biomarker has not been investigated in an unbiased fashion afforded by RNA-Seq (35). Granzymes A and B are serine-proteinases stored in the granules of activated cytotoxic T lymphocytes, and RNA-Seq of urinary cells has validated their diagnostic utility (36). The granzymes may also have an additional role in triggering inflammation (37). NKG7 is a NK cell granule protein and is reported to be upregulated within the allografts in patients with acute rejection (38). CTLA4 has been studied in view of its T cell coinhibitory role and is the target of a clinically used immunosuppressant, belatacept.

Prioritization of genes identified by differential gene expression analysis of RNA-Seq data may lead to highly accurate diagnostic signatures. Indeed, we identified a 13-gene signature, a subset of the 180 differentially expressed genes in TCMR urine versus NR urine that outperformed the RT-qPCR-based CTOT-04 urinary cell TCMR signature in discriminating TCMR from NR. Differential gene expression resolved by RNA-Seq of urinary cells also enabled the development of a urinary cell gene signature that perfectly discriminated TCMR from AMR. These diagnostic models, however, may be overfitted in view of a relatively small sample size and the large number of variables (genes) used as predictors in the penalized multinomial regression analysis, which developed the signatures. The CTOT-04 urinary cell TCMR diagnostic signature has been validated using an independent cohort in kidney allograft recipients (21), and the prediction models developed in this study should be validated using an independent and larger cohort of kidney allograft recipients.

Our semisupervised heatmaps showing that acute rejection samples fail to form a single cluster molecularly are reminiscent of the findings of Sarwal et al. that gene expression patterns associated with histologically indistinguishable acute rejection biopsies cluster as 3 distinct subtypes at the molecular level (39). The observed heterogeneity may reflect hitherto unrecognized subtypes currently considered as a single category based on histological features. Whether the molecular heterogeneity helps personalize antirejection therapy or provides prognostic biomarkers remains to be investigated.

How suitable is the urine as a surrogate biospecimen for the standard-of-care kidney allograft biopsy? To address this issue, we investigated whether the gene expression patterns seen in the urine reflect gene expression patterns within the kidney allograft. We used GSEA to analyze genes and pathways shared between urine and biopsy specimens (27). Our analysis identified that 75 genes are shared between TCMR biopsies and TCMR urine, and 187 genes are shared between AMR biopsies and AMR urine. To our knowledge, this is the first demonstration that gene expression patterns of urinary cells reflect gene expression patterns of kidney allograft biopsies. Cell type enrichment analysis demonstrating enrichment of immune cell types in the urine compared with biopsies support the concept that the kidney allograft may function as *in vivo* flow cytometer and sort graft infiltrating cells in urine (40).

A limitation of our study is the relatively small size of the study cohorts. Thus, it is important to conduct additional RNA-Seq studies of urinary cells and kidney allograft biopsies to validate the transcriptional

profiles identified in the current investigation. An issue inherent to clinical studies is the heterogeneity of study participants and a lack of homogeneity of biospecimens. Both these factors may have contributed to the molecular heterogeneity within the same diagnostic category we observed. The lack of homogeneity, however, may improve the generalizability of our findings.

In sum, unbiased characterization of urinary cell transcriptomes by RNA-Seq and bioinformatics, in addition to elucidating the gene signatures and biological pathways associated with TCMR and AMR, helped decipher shared gene expression patterns between these 2 histologically distinct entities. Whole transcriptome profiling by RNA-Seq also demonstrated that TCMR and AMR biopsy gene signatures are enriched in urinary cells. Cell type enrichment analysis indicated a stronger and more diverse immune landscape signal in urinary cells compared with the kidney allograft biopsies.

## Methods

*Study cohort for urinary cell RNA-Seq and urine specimen processing.* We performed RNA-Seq of 70 urine samples from 66 kidney allograft recipients (patients) transplanted and followed at our center, NewYork-Presbyterian/Weill Cornell Medicine. The urine samples were from our biorepository and were selected to represent 3 major diagnostic categories: Banff Category 1, normal biopsy or nonspecific changes (NR); Banff Category 2, antibody-mediated changes (AMR); or Banff Category 4, TCMR. Among the 70 urine samples selected, 27 urine were from 25 patients with 27 TCMR biopsies, 10 were from 10 patients with AMR biopsies, and 33 urine samples were from 25 patients with 33 NR biopsies.

In this study of RNA-Seq of urinary cells, we aimed to investigate the 2 major types of acute rejection, TCMR and AMR, and compare each type to NR samples. We did not include other diagnostic categories such as polyoma virus-associated nephropathy. Data from the current study should serve as a useful precursor, as other complex kidney allograft diagnoses are investigated.

Approximately 50 mL of urine was obtained from each patient and centrifuged at 1250 g for 30 minutes at room temperature within 4 hours of collection. We isolated total RNA from urinary cells using the RNeasy mini kit (Qiagen) followed by DNase I (Qiagen) treatment. The quantity and purity of the RNA isolated from the urine cell pellet were measured using the NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific). We used Agilent 2100 Bioanalyzer (Agilent Technologies Inc.) to measure the integrity of RNA. Preparation of RNA sample library and RNA-Seq were performed by the Genomics Core Laboratory at Weill Cornell Medicine. rRNA was removed from total RNA using Illumina Ribo Zero Gold for human/mouse/rat kit. Using Illumina TruSeq RNA Sample Library Preparation v2 kit (Illumina), mRNA was fragmented into small pieces using divalent cations under elevated temperature. The cleaved RNA fragments were copied into first strand cDNA using reverse transcriptase and random primers. Second strand cDNA synthesis followed, using DNA Polymerase I and RNase H. The cDNA fragments then went through an end repair process, the addition of a single “A” base, and then ligation of the adapters. The products were then purified and enriched with PCR to create the final cDNA library. The normalized cDNA libraries were pooled and sequenced on Illumina HiSeq4000 sequencer with pair-end 100 cycles.

Urine is rich in DNA and RNA hydrolyzing enzymes (26). The inherent difficulty in isolating intact RNA from urinary cells is reflected in part by the lower RIN of the total RNA isolated from the urinary cells compared with the total RNA isolated from kidney allograft biopsies; the median and interquartile range (IQR) of RIN of RNA from the urine was 6.1 (IQR 4.7–7.1), and the RIN of the RNA isolated from the kidney allograft biopsies was 7.9 (IQR 7.4–8.7). Because of expected RNA degradation in urine, we chose 20% read alignment against human GRCh38/hg38 Reference genome and a RIN score of 2 as the minimum thresholds to pass onto downstream analysis. Based on these thresholds, we included 57 of the 70 urine samples sequenced consecutively for downstream data analysis and excluded 13 urine samples (5 from 5 patients with TCMR biopsies, 2 from 2 patients with AMR biopsies, and 6 from 6 patients with NR biopsies). Thus, 81% of urine samples yielded good-quality RNA for sequencing. Among the 57 urine samples with RNA-Seq data that met sequencing quality thresholds, 22 were from 20 recipients with biopsies showing TCMR, 8 were from 8 recipients with biopsies showing AMR, and 27 were from 25 recipients with biopsies classified as NR biopsies. None of the biopsies qualified for mixed rejection category (i.e., concurrent TCMR and AMR), and we did not include any other diagnostic categories in our RNA-Seq study. Supplemental Table 1 is a summary of characteristics of RNA-Seq cohort.

Among 30 urine specimens collected from patients with biopsies classified as TCMR or AMR, 28 of 30 were collected before the invasive biopsy procedure, and the remaining 2 specimens were collected 1 day

after the procedure. Among the 27 urine samples collected from patients with biopsies classified as NR, 26 of 27 were collected before the biopsy procedure. None of the patients received antirejection therapy before the collection of urine samples.

*Kidney allograft biopsies.* Kidney allograft biopsies were done under ultrasound guidance and were read and reported independently by 2 transplant pathologists at our center, based on the Banff 2017 update of the Banff 1997 classification of allograft pathology (10, 41). Biopsy tissue sections were stained with H&E, periodic acid–Schiff, and Masson's trichrome. Biopsies were also stained for polyomavirus and C4d.

Among the urinary cell RNA-Seq cohort, all 22 TCMR biopsies and all 8 AMR biopsies were for-cause biopsies and were performed to determine the cause of kidney allograft dysfunction. Among the 27 biopsies categorized as normal/nonspecific (NR), 24 biopsies were surveillance biopsies and were performed in patients without graft dysfunction, and the remaining 3 were for-cause biopsies performed to investigate the basis for graft dysfunction (Supplemental Table 1). Among the kidney allograft biopsy cohort, all 12 TCMR biopsies and all 17 AMR biopsies were for-cause biopsies and were performed to determine the cause of kidney allograft dysfunction, and all 20 NR biopsies were surveillance biopsies (Supplemental Table 2).

*Study cohort for kidney allograft biopsy RNA-Seq and biopsy specimen processing.* We analyzed 49 kidney allograft biopsy specimens from 49 adult kidney recipients transplanted and followed at our center. Supplemental Table 2 is a summary of kidney allograft biopsy cohort. Among the 49 patients, 11 provided urine specimens for RNA-Seq and are included among the 57 urine samples that were RNA sequenced. At the time of allograft biopsy, a portion of the biopsy tissue was immediately immersed in RNAlater RNA stabilization solution (Invitrogen) and stored in  $-80^{\circ}\text{C}$ . From the stored sample, we isolated total RNA using the miRNeasy mini kit (Qiagen). The quantity (absorbance at 260 nm) and purity (ratio of the absorbance at 260 nm and 280 nm) of the RNA isolated from the kidney tissue were measured using the NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific). We used Agilent 2100 Bioanalyzer (Agilent Technologies) to measure the integrity of RNA.

*RNA-Seq of kidney allograft biopsies.* Preparation of RNA sample library and RNA-seq were performed by the Genomics Core Laboratory at Weill Cornell Medicine. Using Illumina TruSeq RNA Sample Library Preparation v2 kit (Illumina), poly(A)-selected mRNA was fragmented into small pieces using divalent cations under elevated temperature. The cleaved RNA fragments were copied into first strand cDNA using reverse transcriptase and random primers. Second strand cDNA synthesis followed, using DNA Polymerase I and RNase H. The cDNA fragments then went through an end repair process, the addition of a single "A" base, and then ligation of the adapters. The products were then purified and enriched with PCR to create the final cDNA library. The normalized cDNA libraries were pooled and sequenced on Illumina HiSeq2500 or 4000 sequencer.

Urinary cell and kidney allograft biopsy RNA-Seq data have been deposited in NCBI's Gene Expression Omnibus, accession numbers GSE131179 and GSE142667.

*Urinary cell and kidney allograft biopsy sequencing data analysis and differential expression.* The raw sequencing data were stored in FASTQ format. Raw sequenced reads were aligned to the Human reference genome (Version hg38 from Genome Reference Consortium GRCh38) using STAR (Version 2.4.2) aligner (42). Aligned reads were quantified against the reference annotation to obtain FPKM and raw counts using CuffLinks (v 2.2.1) and HTSeq, respectively (43, 44). Genes expressed with counts per million (cpm)  $> 2$  in 2 or more samples were included for all downstream analysis. We used limma-voom, an R package, for the differential expression analysis of RNA-Seq data. This package applies a function that fits multiple linear models by weighted or generalized least squares and uses an empirical Bayes method to moderate the standard errors of the estimated log-fold changes.

Genes significantly differentially expressed in urine with  $\log_2 \text{FC} \geq 2$  and  $\text{FDR} < 0.1$  or  $< 0.05$  were considered for supervised downstream analysis. In order to classify the samples based on gene expression profiles, we first performed unsupervised PCA using  $\log_2$  transformed FPKM expression values in R statistical software (45). A supervised gene name-based pathway analysis to functionally annotate differentially expressed genes was performed using ENRICH (46) and the KEGG 2016 Human database (47). Pathways with  $\text{FDR} < 0.05$  were considered to be significantly enriched.

GSEA and pathway analysis using GSEA software from the Broad Institute (Massachusetts Institute of Technology, Cambridge, Massachusetts, USA) were used to compare kidney allograft biopsy signature against urinary cell gene signature in addition to identifying functions of differentially expressed genes (27). In order to curate a high confidence signature representative of the kidney tissue biopsy, stringent thresholds of  $\log_2 \text{FC} > 2$

and  $FDR < 0.05$  were established to signify up- or downregulated genes. Genes were ranked by the t-statistic value obtained from comparisons, and the preranked version of the tool was used to identify significantly enriched biological pathways. In the output, the ES reflects the degree to which a gene set is overrepresented at the top or bottom of a ranked list of genes. GSEA calculates the ES by walking down the ranked list of genes, increasing a running-sum statistic when a gene is in the gene set, and decreasing it when it is not. The magnitude of the increment depends on the correlation of the gene with the phenotype. The ES is the maximum deviation from zero encountered in walking the list. This computational method determines whether an a priori–defined set is overrepresented among the upregulated or downregulated spectrum of a ranked list of genes.

**Cell type enrichment.** Various cell type abundance estimation was performed using the bioinformatics tool xCell (29), using normalized bulk RNA-Seq expression data (FPKM) as the input. The relative cell type abundance was quantified and visualized across all samples. Abundance for each cell type across the diagnostic groups and between kidney allograft biopsy and urinary cells was compared using Wilcoxon rank-sum test in R. Cell types with  $FDR < 0.1$  were considered to be significantly differentially enriched.

**Customized RT-qPCR assays.** RNA for RT-qPCR assays were reverse transcribed to cDNA using TaqMan Reverse Transcription Reagents (Applied Biosystems) at a final concentration of 1.0  $\mu\text{g}$  of total RNA in 100  $\mu\text{L}$  volume. We designed gene-specific oligonucleotide primers and fluorogenic probes, using Primer Express software for the measurement of mRNAs and a housekeeping/reference gene, 18S rRNA. The probes were labeled with 6-carboxy-fluorescein (FAM) at the 5' end and 6-carboxy-tetramethylrodamine (TAMRA) or dihydrocyclopyrroloindole tripeptide minor groove binder (MGB) at the 3' end. FAM functioned as the reporter dye and TAMRA or MGB as the quencher dye. We did a 2-step customized RT-qPCR assay, a preamplification step followed by measurement of the absolute levels of mRNAs, using our previously described standard curve method in Applied Biosystems QuantStudio 6 Flex Real-Time PCR System (Thermo Fisher Scientific). The standard curve copy numbers in our PCR assays ranged from 25–2.5 million copies, and for data analysis, 18S rRNA  $\geq 5.0 \times 10^7$  copies/ $\mu\text{g}$  of total RNA and TGF- $\beta$ 1 mRNA  $\geq 1.0 \times 10^2$  copies/ $\mu\text{g}$  of total RNA were used as a measure of RNA adequacy in a specimen (21).

**Statistics.** For RNA-Seq data, differential expression analysis to compare the mRNA expression profiles of different diagnostic categories was performed on voom normalized raw counts using the *limma* package (<http://bioconductor.org/packages/release/bioc/html/limma.html>) in R (45). Probability values were adjusted for FDR by the Benjamini-Hochberg method. Penalized multinomial logistic lasso regression, as implemented in R (version 3.3.3) and the *glmnet* package (version 2.0-13), was used to estimate generalized linear models that discriminate TCMR from NR samples using differentially expressed genes as potential predictors and to discriminate TCMR from AMR using uniquely expressed genes as potential predictors (48). A tuning parameter ( $\lambda$ ) is used in penalized regression to specify the shrinkage coefficient. The  $\lambda$  that minimizes the cross-validation prediction error rate ( $\lambda_{\text{min}}$ ) was determined using the function *cv.glmnet*, and  $\lambda_{\text{min}}$  was used in the *glmnet* function. Receiver operating characteristic (ROC) curve construction, as well as calculation of AUC and comparison of AUC curves, was performed using R packages *pROC* v1.10.0 and *cvAUC* v1.1.0. Three-group comparisons of mRNA abundance were evaluated using the Kruskal-Wallis test. Pairwise group comparisons of mRNA abundance were evaluated using the Mann-Whitney *U* test;  $P < 0.05$  was considered statistically significant.

**Study approval.** Kidney transplant recipients reported herein provided written informed consent to participate in the study, and the informed consent was obtained before their inclusion in the study. The IRB at Weill Cornell Medicine approved the study. The clinical and research activities that we report here are consistent with the principles of the “Declaration of Istanbul on Organ Trafficking and Transplant Tourism” and the “World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects” (49, 50).

## Author contributions

TM and MS designed the research studies. HY, MFC, CS, and JZX conducted experiments. AV, TM, HY, ML, DMD, JRL, SPS, VKS, JZX, SVS, and DS acquired data. SPS and SVS classified kidney allograft biopsies based on the Banff 2017 update of the Banff 1997 classification of allograft pathology. AV, TM, HY, MC, ML, CS, JZX, ID, FBM, KS, OE, and MS analyzed data. JZX, OE, and MS provided reagents. AV, TM, ML, OE, and MS wrote the manuscript. AV, TM, and HY are co–first authors. AV performed the bioinformatic analysis and is listed as first among the 3 co–first authors. The other 2 co–first authors are listed alphabetically. All authors read and approved the contents of the manuscript.

## Acknowledgments

This work was supported, in part, by awards from the NIH to MS (NIH MERIT Award, R37AI051652), OE (UL1TR002384 and R01CA194547), TM (K08DK087824), and Weill Cornell Medical College (Clinical and Translational Science Center Award UL1TR000457). OE is supported by awards from the Leukemia and Lymphoma Society's Specialized Center of Research (180078-02 and 7021-20). An award from the Mendez National Institute of Transplantation Foundation (principal investigator, MS) supported, in part, RNA-Seq studies reported in this manuscript. The authors are thankful to Tijana Bzenic for her meticulous assistance in the preparation of this manuscript.

Address correspondence to: Manikkam Suthanthiran, Division of Nephrology and Hypertension, Department of Transplantation Medicine, 525 East 68th Street, Box 3, New York, New York 10065, USA. Phone: 212.746.4430; Email: msuthan@med.cornell.edu

1. Wolfe RA, et al. Comparison of mortality in all patients on dialysis, patients on dialysis awaiting transplantation, and recipients of a first cadaveric transplant. *N Engl J Med.* 1999;341(23):1725–1730.
2. Souillou JP, et al. Randomized controlled trial of a monoclonal antibody against the interleukin-2 receptor (33B3.1) as compared with rabbit antithymocyte globulin for prophylaxis against rejection of renal allografts. *N Engl J Med.* 1990;322(17):1175–1182.
3. Hanaway MJ, et al. Alemtuzumab induction in renal transplantation. *N Engl J Med.* 2011;364(20):1909–1919.
4. Vincenti F, et al. Belatacept and Long-Term Outcomes in Kidney Transplantation. *N Engl J Med.* 2016;374(4):333–343.
5. Hart A, et al. OPTN/SRTR 2017 Annual Data Report: Kidney. *Am J Transplant.* 2019;19 Suppl 2:19–123.
6. Ishikawa A, et al. Quantitative assessment of the first acute rejection as a predictor of renal transplant outcome. *Transplantation.* 1999;68(9):1318–1324.
7. El-Zoghby ZM, et al. Identifying specific causes of kidney allograft loss. *Am J Transplant.* 2009;9(3):527–535.
8. Alkadi MM, et al. Kidney allograft failure in the steroid-free immunosuppression era: A matched case-control study. *Clin Transplant.* 2017;31(11).
9. Burton H, et al. Causes of renal allograft failure in the UK: trends in UK Renal Registry and National Health Service Blood and Transplant data from 2000 to 2013. *Nephrol Dial Transplant.* 2019;34(2):355–364.
10. Haas M, et al. The Banff 2017 Kidney Meeting Report: Revised diagnostic criteria for chronic active T cell-mediated rejection, antibody-mediated rejection, and prospects for integrative endpoints for next-generation clinical trials. *Am J Transplant.* 2018;18(2):293–307.
11. Roufosse C, et al. A 2018 Reference Guide to the Banff Classification of Renal Allograft Pathology. *Transplantation.* 2018;102(11):1795–1814.
12. Morgan TA, Chandran S, Burger IM, Zhang CA, Goldstein RB. Complications of Ultrasound-Guided Renal Transplant Biopsies. *Am J Transplant.* 2016;16(4):1298–1305.
13. Redfield RR, et al. Nature, timing, and severity of complications from ultrasound-guided percutaneous renal transplant biopsy. *Transpl Int.* 2016;29(2):167–172.
14. Plattner BW, Chen P, Cross R, Leavitt MA, Killen PD, Heung M. Complications and adequacy of transplant kidney biopsies: A comparison of techniques. *J Vasc Access.* 2018;19(3):291–296.
15. Li B, 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(13):947–954.
16. Ding R, et al. CD103 mRNA levels in urinary cells predict acute rejection of renal allografts. *Transplantation.* 2003;75(8):1307–1312.
17. Muthukumar T, et al. Serine proteinase inhibitor-9, an endogenous blocker of granzyme B/perforin lytic pathway, is hyperexpressed during acute rejection of renal allografts. *Transplantation.* 2003;75(9):1565–1570.
18. Tatapudi RR, et al. Noninvasive detection of renal allograft inflammation by measurements of mRNA for IP-10 and CXCR3 in urine. *Kidney Int.* 2004;65(6):2390–2397.
19. Muthukumar T, et al. Messenger RNA for FOXP3 in the urine of renal-allograft recipients. *N Engl J Med.* 2005;353(22):2342–2351.
20. Afaneh C, et al. Urinary cell levels of mRNA for OX40, OX40L, PD-1, PD-L1, or PD-L2 and acute rejection of human renal allografts. *Transplantation.* 2010;90(12):1381–1387.
21. Suthanthiran M, et al. Urinary-cell mRNA profile and acute cellular rejection in kidney allografts. *N Engl J Med.* 2013;369(1):20–31.
22. Matignon M, et al. Urinary cell mRNA profiles and differential diagnosis of acute kidney graft dysfunction. *J Am Soc Nephrol.* 2014;25(7):1586–1597.
23. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 2009;10(1):57–63.
24. Zhao S, Fung-Leung WP, Bittner A, Ngo K, Liu X. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. *PLoS ONE.* 2014;9(1):e78644.
25. Zhang W, et al. Comparison of RNA-seq and microarray-based models for clinical endpoint prediction. *Genome Biol.* 2015;16:133.
26. Bryzgunova OE, Laktionov PP. Extracellular Nucleic Acids in Urine: Sources, Structure, Diagnostic Potential. *Acta Naturae.* 2015;7(3):48–54.
27. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA.* 2005;102(43):15545–15550.
28. Halloran PF, et al. Review: The transcripts associated with organ allograft rejection. *Am J Transplant.* 2018;18(4):785–795.
29. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 2017;18(1):220.

30. Sorof JM, Vartanian RK, Olson JL, Tomlanovich SJ, Vincenti FG, Amend WJ. Histopathological concordance of paired renal allograft biopsy cores. Effect on the diagnosis and management of acute rejection. *Transplantation*. 1995;60(11):1215–1219.
31. Wang Z, Liu P, Zeng G, Randhawa P. Application of RNA-seq Derived Diagnostic Algorithms of T-cell Mediated Kidney Rejection (TCMR) to Publicly Available DNA Microarray-based Gene Expression Datasets. Paper presented at: American Transplant Congress; June 2–6, 2018; Seattle, Washington, USA. <http://atcmeetingabstracts.com/abstract/application-of-rna-seq-derived-diagnostic-algorithms-of-t-cell-mediated-kidney-rejection-tcmr-to-dna-microarray-based-gene-expression-datasets>. Accessed February 5, 2020.
32. Sanders ME, et al. Human memory T lymphocytes express increased levels of three cell adhesion molecules (LFA-3, CD2, and LFA-1) and three other molecules (UCHL1, CDw29, and Pgp-1) and have enhanced IFN-gamma production. *J Immunol*. 1988;140(5):1401–1407.
33. Suthanthiran M. A novel model for antigen-dependent activation of normal human T cells. Transmembrane signaling by crosslinkage of the CD3/T cell receptor-alpha/beta complex with the cluster determinant 2 antigen. *J Exp Med*. 1990;171(6):1965–1979.
34. Kapur S, Khanna A, Sharma VK, Li B, Suthanthiran M. CD2 antigen targeting reduces intragraft expression of mRNA-encoding granzyme B and IL-10 and induces tolerance. *Transplantation*. 1996;62(2):249–255.
35. Ruster M, Sperschneider H, Fünfstück R, Stein G, Gröne HJ. Differential expression of beta-chemokines MCP-1 and RANTES and their receptors CCR1, CCR2, CCR5 in acute rejection and chronic allograft nephropathy of human renal allografts. *Clin Nephrol*. 2004;61(1):30–39.
36. van Ham SM, et al. Urinary granzyme A mRNA is a biomarker to diagnose subclinical and acute cellular rejection in kidney transplant recipients. *Kidney Int*. 2010;78(10):1033–1040.
37. Froelich CJ, Pardo J, Simon MM. Granule-associated serine proteases: granzymes might not just be killer proteases. *Trends Immunol*. 2009;30(3):117–123.
38. Khatri P, et al. A common rejection module (CRM) for acute rejection across multiple organs identifies novel therapeutics for organ transplantation. *J Exp Med*. 2013;210(11):2205–2221.
39. Sarwal M, et al. Molecular heterogeneity in acute renal allograft rejection identified by DNA microarray profiling. *N Engl J Med*. 2003;349(2):125–138.
40. Lee JR, et al. Urinary cell mRNA profiles predictive of human kidney allograft status. *Immunol Rev*. 2014;258(1):218–240.
41. Racusen LC, et al. The Banff 97 working classification of renal allograft pathology. *Kidney Int*. 1999;55(2):713–723.
42. Dobin A, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15–21.
43. Trapnell C, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010;28(5):511–515.
44. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166–169.
45. R Core Team. R Foundation for Statistical Computing. <http://www.R-project.org>. Accessed February 5, 2020.
46. Kuleshov MV, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. 2016;44(W1):W90–W97.
47. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res*. 2016;44(D1):D457–D462.
48. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw*. 2010;33(1):1–22.
49. International Summit on Transplant Tourism Organ Trafficking. The Declaration of Istanbul on Organ Trafficking and Transplant Tourism. *Clin J Am Soc Nephrol*. 2008;3(5):1227–1231.
50. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013;310(20):2191–2194.