

## Longitudinal Analysis of Whole Blood Transcriptomes to Explore Molecular Signatures Associated With Acute Renal Allograft Rejection

Heesun Shin<sup>1,2,7</sup>, Oliver Günther<sup>1</sup>, Zsuzsanna Hollander<sup>1,3,7</sup>, Janet E. Wilson-McManus<sup>1</sup>, Raymond T. Ng<sup>1,4</sup>, Robert Balshaw<sup>1,5</sup>, Paul A. Keown<sup>1,2</sup>, Robert McMaster<sup>1,6</sup>, Bruce M. McManus<sup>1,3,7</sup>, Nicole M. Isbel<sup>8</sup>, Greg Knoll<sup>9</sup> and Scott J. Tebbutt<sup>1,2,7</sup>; for the PROOF Centre of Excellence Team

<sup>1</sup>NCE CECR PROOF Centre of Excellence, Vancouver, BC. <sup>2</sup>University of British Columbia (UBC) Department of Medicine, Vancouver, BC. <sup>3</sup>UBC Department of Pathology and Laboratory Medicine, Vancouver, BC. <sup>4</sup>UBC Department of Computer Science, Vancouver, BC. <sup>5</sup>UBC Department of Statistics, Vancouver, BC. <sup>6</sup>UBC Department of Medical Genetics, Vancouver, BC. <sup>7</sup>Institute for HEART + LUNG Health, Vancouver, BC. <sup>8</sup>Department of Nephrology, Princess Alexandra Hospital, and University of Queensland, Brisbane Australia. <sup>9</sup>The Ottawa Hospital Research Institute, Ottawa, ON, Canada.

**ABSTRACT:** In this study, we explored a time course of peripheral whole blood transcriptomes from kidney transplantation patients who either experienced an acute rejection episode or did not in order to better delineate the immunological and biological processes measurable in blood leukocytes that are associated with acute renal allograft rejection. Using microarrays, we generated gene expression data from 24 acute rejectors and 24 nonrejectors. We filtered the data to obtain the most unambiguous and robustly expressing probe sets and selected a subset of patients with the clearest phenotype. We then performed a data-driven exploratory analysis using data reduction and differential gene expression analysis tools in order to reveal gene expression signatures associated with acute allograft rejection. Using a template-matching algorithm, we then expanded our analysis to include time course data, identifying genes whose expression is modulated leading up to acute rejection. We have identified molecular phenotypes associated with acute renal allograft rejection, including a significantly upregulated signature of neutrophil activation and accumulation following transplant surgery that is common to both acute rejectors and nonrejectors. Our analysis shows that this expression signature appears to stabilize over time in nonrejectors but persists in patients who go on to reject the transplanted organ. In addition, we describe an expression signature characteristic of lymphocyte activity and proliferation. This lymphocyte signature is significantly downregulated in both acute rejectors and nonrejectors following surgery; however, patients who go on to reject the organ show a persistent downregulation of this signature relative to the neutrophil signature.

**KEYWORDS:** blood transcriptomics, microarray, kidney transplant rejection, peripheral whole blood, neutrophil to lymphocyte ratio

**CITATION:** Shin et al. Longitudinal Analysis of Whole Blood Transcriptomes to Explore Molecular Signatures Associated With Acute Renal Allograft Rejection. *Bioinformatics and Biology Insights* 2014:8 17–33 doi: 10.4137/BBI.S13376.

**RECEIVED:** October 8, 2013. **RESUBMITTED:** November 17, 2013. **ACCEPTED FOR PUBLICATION:** November 17, 2013.

**ACADEMIC EDITOR:** J.T. Efrid, Associate Editor

**TYPE:** Original Research

**FUNDING:** This research was supported by Genome BC ([www.genomebc.ca](http://www.genomebc.ca)) and Genome Canada ([www.genomecanada.ca](http://www.genomecanada.ca)). HS was a recipient of MITACS-Accelerate research internship (<http://www.mitacs.ca/accelerate>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**COMPETING INTERESTS:** NMI has received a grant from Novartis and advisory board fees paid to her institution from Pfizer, outside the published work. Other authors disclose no competing interests.

**COPYRIGHT:** © the authors, publisher and licensee Libertas Academica Limited. This is an open-access article distributed under the terms of the Creative Commons CC-BY-NC 3.0 License.

**CORRESPONDENCE:** Scott.Tebbutt@hli.ubc.ca

### Background

The success of kidney transplantation has increased dramatically over the past several decades largely due to a better understanding of the immune biology and the molecular processes underlying allograft assimilation and rejection in addition to improved management of immunosuppression regimens.<sup>1–3</sup>

However, acute rejection, caused by an immune response in the recipient against alloantigens of the donor graft, remains a major complication and prevents the long-term assimilation of the allograft. Management of the immune response associated with rejection is, therefore, essential for the prevention of irreversible damage to the graft.



T lymphocytes are known to be the principal mediators of acute allograft rejection, infiltrating graft tissues and interceding cell-mediated cytotoxicity reactions.<sup>4,5</sup> Therefore, the majority of studies examining gene expression changes associated with transplant rejection have used lymphocyte-biased samples or platforms such as peripheral blood mononuclear cells (PBMCs)<sup>6</sup> and Lymphochip cDNA microarrays.<sup>7,8</sup> More recent evidence has revealed an important role for the innate immune system in allograft rejection.<sup>9</sup> It has been demonstrated that the almost complete depletion of T cells is not sufficient to prevent allograft rejection. Furthermore, it has been suggested that rejection might be associated with innate immune activities related to ischemia-reperfusion injury.<sup>10,11</sup> A more complete understanding of activities of the immune system during transplantation and allograft rejection is, therefore, needed to facilitate the discovery and development of more targeted and successful tolerance strategies.<sup>12</sup>

Molecular profiling of circulating whole blood is a viable approach for monitoring the status of immune system activities associated with kidney transplantation and allograft rejection.<sup>13</sup> The identification of specific molecular signatures associated with allograft rejection has potential utility for the development of surveillance tools for real-time tracking of the rejection status of transplant patients. The feasibility of this type of approach is based on the utility of peripheral whole blood as a surrogate for biopsy tissue, as the blood carries educated immune cells circulating from the site of injury to lymphoid organs as it flows throughout the body. Blood cell molecular signatures are known to reflect system-wide pathological changes and physiological homeostasis in the body,<sup>14</sup> which can reveal a variety of immune responses. Profiling of blood transcriptomes by using genomic technologies, such as microarray, provides a snapshot of transcriptional activities in peripheral whole blood that can be used to characterize molecular signatures present in the sample. These techniques have, therefore, surfaced as a powerful approach that can be used to exploit disease-associated molecular phenotypes in blood.<sup>15,16</sup>

In this study, we use stringently filtered microarray data to identify and characterize transcriptional changes that are associated with kidney transplantation and acute allograft rejection. Using a data-driven exploratory analysis approach to eliminate potential confounding factors within the data, we identify biologically relevant genes representing physiological homeostasis and immunological activities that are associated with kidney transplantation and acute allograft rejection. This analysis encompasses data representing various time points from pretransplant and posttransplant, allograft rejection, and postrejection. This work provides potential biological insight into immune system activities throughout the kidney transplantation process including allograft acceptance or rejection.

## Methods

**Ethics statement.** This study was conducted in collaboration with the University of British Columbia investigators

and was approved by the Providence Health Care Research Ethics Board. All patients in the study gave informed written consent.

**Patient group and sample selection.** As described in our previous work,<sup>16</sup> all subjects receiving a renal transplant in the University of British Columbia renal transplant program from January 1, 2005, through December 31, 2009, were invited to participate in this study. Those who agreed and signed consent were enrolled and were followed routinely by the transplant program team throughout their course.<sup>16</sup> Patients received a standardized treatment protocol including basiliximab 20 mg intravenously (IV) on days 0 and 4, methylprednisolone 125 mg IV on the day of transplantation tapering to zero by day 3 posttransplant, tacrolimus 0.075 mg/kg twice a day, and mycophenolate 1000 mg twice a day. Tacrolimus concentrations were measured by tandem mass spectrometry, and the dose was adjusted to achieve 12-hour trough levels of 8–12 ng/mL in month 1, 6–9 ng/mL in month 2, and 4–8 ng/mL thereafter. Allograft rejection was diagnosed by normal clinical and laboratory parameters, confirmed by biopsy, and graded according to the Banff 97 working classification of renal allograft pathology.<sup>17</sup> Banff categories 2 and 4 (antibody-mediated or acute/active cellular rejection) were considered significant. Subjects with borderline changes (category 3) were analyzed separately. All demographic, clinical, diagnostic, and therapeutic data were recorded longitudinally in an electronic database, and there was no patient loss to follow-up during the study.

A total of 252 blood samples were obtained during the first year in PAXgene<sup>TM</sup> tubes (Qiagen, Valencia, CA). Samples were obtained immediately prior to transplantation, at 0.5, 1, 2, 3, 4, 8, 12, 26, and 52 weeks posttransplant and at the time of suspected rejection. However, a complete collection of time point samples was not attained in every case. Additionally, samples were not uniformly collected for every patient for every time point. Graft tissue was obtained pretransplant and at the time of all biopsies performed for clinical purposes posttransplant. All samples were stored in a biobank until required for analysis. Blood samples from 20 healthy volunteers collected at a single time point, treated identically, served as controls.<sup>16</sup> The study employed a case-control design<sup>18</sup> to compare differential gene expression in subjects with or without acute rejection (AR). Patients were considered eligible for acceptance into the program and subsequent analysis if they were less than 75 years of age; did not have pretransplant immunosuppression or immunological desensitization; received an AHG-CDC crossmatch negative kidney transplant from a deceased or non-HLA-identical living donor; did not receive depleting antibody induction therapy; were able to receive oral immunosuppression; and had no evidence of infection, disease recurrence, or other major comorbid events. Cases with AR diagnosed during the first 12 months posttransplant were matched as closely as possible for age, sex, degree of sensitization, organ source, and



date of transplantation with nonrejection controls (NR), who had no evidence of acute allograft rejection during the period of follow-up. Complete blood counts with differentials were obtained for some samples.

**RNA extraction and microarray data generation.** Total RNA was extracted using PAXgene™ Blood RNA kits (Qiagen, Valencia, CA) and integrity and concentration determined using the Agilent 2100 BioAnalyzer (Agilent Technologies, Palo Alto, CA). Gene expression was analyzed at the CAP/CLIA certified Microarray Core Laboratory, Genome Core at the Children's Hospital, Los Angeles, CA, using Affymetrix Human Genome U133 Plus 2.0 arrays. Quality of the samples, hybridization, chips, and scanning were reviewed using the tools in the Bioconductor packages affy version 1.16.0 and affyPLM version 1.14.0.

**Microarray data processing and filtering.** GCRMA<sup>19,20</sup> was used for normalization and background correction using R statistical language. Expression values were analyzed on the log-base 2 scale. We included not only the 252 samples from renal transplant recipients and 20 healthy controls, but also 136 samples from cardiac transplant recipients for normalization and background correction to increase the sample size for more robust data normalization. PANP (Presence-Absence Calls From Negative Strand Matching Probesets)<sup>21</sup> bioconductor R package was used for microarray filtering of background signals from unspecific hybridization (lower  $P$  value = 0.02 and upper  $P$  value = 0.01 for absent/marginal/present calls). On average, the unspecific binding signal was around a  $\log_2$  signal intensity of 6 across all the samples at a  $P$  value of 0.01. Probe sets that were "present" in at least 80% of the samples and "marginal" in the remaining 20% of the samples were selected for further analysis. PLANdbAffy<sup>22</sup> database was used to remove ambiguous mapping of probe sets to multiple genes (ie, probe sets with at least 1 "green probe," which are uniquely mapping probes, were used). Only probe sets with a minimum 9 out of 11 probes that mapped consistently to the same target and probes sets that directly aligned with known transcripts (ie, grades A and B) based on Affymetrix annotations (NetAffx™ release 31) were used. Of the 54,000 probe sets present on the arrays, 5619 passed our rigorous filtering pipeline.

**Data exploration and analysis.** MultiExperiment Viewer (MeV)<sup>23</sup> was used for the microarray data analysis. Data reduction tools including principal component analysis (PCA)<sup>24</sup> in

MeV were used (standard default inputs) in order to reveal internal structures or hidden associations in the gene expression data with regards to clinical background, patient demographics, and transplantation outcome. In all cases, the first 2 principal components accounted for more than approximately 70% of the variance. A pattern-matching algorithm, Pavlidis Template Matching (PTM)<sup>25</sup> and Significance Analysis of Microarrays (SAM)<sup>25</sup> in MeV were used to determine transcripts with differential expression between acute rejectors (AR) and nonrejectors (NR) that changed over time approaching rejection (adjusted  $P$  value false discovery rate [FDR] < 0.05) (0.05 FDR is a commonly used false discovery rate cut-off). We used bioinformatic analyses, including InnateDB<sup>27</sup> and GOstat,<sup>28</sup> for the identification of biological pathways and processes overrepresented in the various lists of differentially expressed genes. oPOSSUM<sup>29</sup> was used for detection of overrepresented transcription factor binding sites in the promoters of differentially expressed genes.

## Results

### Patient group description and confounder analysis.

The samples collected for this study represent 48 transplant patients from different demographic and clinical backgrounds and include various time points from end-stage kidney failure through the posttransplantation period, including rejection and postrejection time points. Twenty-four patients experienced an acute kidney rejection episode, while 24 did not reject the transplanted organ. Table 1 summarizes the demographics of the 48 patients (Additional file 1, detailed 48 patient demographic and clinical data). In addition, single time point samples from 20 healthy subjects were used as a reference for several analyses, including as controls for immunotherapy.

We performed a confounder analysis of patient demographics and clinical data in order to ensure balanced sample matching of acute rejectors (AR) with nonrejectors (NR). Pearson's chi-squared test and analysis of variance (ANOVA) were used for categorical and continuous variables, respectively. As shown in Table 2, potential confounders, including patient demographics and clinical variables such as treatment regimen, did not show any statistically significant difference between ARs and NRs. However, as expected, renal blood urea nitrogen (BUN) and creatinine levels were significantly different between the 2 patient groups (adjusted  $P$  value = 3.65E-03 and 4.36E-06, respectively) indicating declining kidney function

**Table 1.** Patient demographics.

CLASS	NUMBER OF SAMPLES	GENDER	AGE (MEAN, MEDIAN)	AGE (MIN, MAX)	ETHNICITY
ARs	24	17 Male and 7 Female	46.8, 50.0	20.3, 71.8	22 Caucasian, 2 Asian Indian
NRs	24	14 Male and 10 female	46.8, 49.9	20.7, 58.6	19 Caucasian, 1 Asian Indian, 2 Asian Oriental, 2 North American Indian

**Table 2.** Confounder analysis for 24 ARs and 24 NRs.

VARIABLE NAME	SAMPLES AVAILABLE	ANALYSIS	ADJ P-VALUE
Age	48	ANOVA	8.82E-01
Ethnicity	48	chisq	3.53E-01
Gender	48	chisq	6.63E-01
Blood type	48	chisq	7.39E-01
Primary disease	48	chisq	6.99E-01
Sample collection days post transplant	48	ANOVA	3.23E-01
RENAL_BUN	48	ANOVA	3.65E-03
RENAL_CREATININE	48	ANOVA	4.36E-06
HEMA_HGB	48	ANOVA	1.76E-01
HEMA_Neutrophils	47	ANOVA	1.02E-01
HEMA_PLT	48	ANOVA	4.22E-01
HEMA_WBC	48	ANOVA	4.30E-01
IMMS_TACROLIMUS	42	ANOVA	9.04E-01
IMM_TACROLIMUS_WB	44	ANOVA	6.75E-01
IMMS_MMF	42	ANOVA	6.63E-01
ACT_CMV	48	chisq	6.47E-01
ACT_EBV	48	chisq	7.48E-01
ACT_HBCAB	48	chisq	7.39E-01
ACT_HBSAB	48	chisq	6.47E-01
ACT_HEPATITIS_B	48	chisq	7.39E-01
ACT_HEPATITIS_C	48	chisq	7.39E-01
ACT_HIV	48	chisq	6.47E-01
DONOR_AGE	48	ANOVA	1.90E-01
DONOR_BLOOD_TYPE	48	chisq	7.82E-01
DONOR_CMV	48	chisq	6.47E-01
DONOR_ETHNICITY	48	chisq	6.47E-01
DONOR_GENDER	48	chisq	7.39E-01

in the AR patients. In addition, although statistically not significant, neutrophil count (adjusted  $P$  value = 1.02E-01) and donor age (adjusted  $P$  value = 1.90E-01) were slightly higher in AR samples as compared with NR samples.

**Microarray data filtering.** We generated transcriptome data for 272 peripheral whole blood samples from the 48 kidney transplantation patients as well as 20 healthy volunteers. We subjected the transcriptome data to a rigorous filtering process to remove obsolete, erroneously mapped, or poorly performing probe sets (see Methods). This approach has been shown to increase the reliability and consistency of gene expression analyses and improve the correlation of microarray and RT-qPCR results.<sup>30,31</sup> Of the 54,000 probe sets present on the arrays, 5619 passed our filtering pipeline, representing roughly 4000 unique genes that are equivalent to approximately 25% of the human blood transcriptome (Additional file 2).<sup>14</sup> Given that, on average, only 30% to 40% of expressed genes in any individual tissue are detected by microarray technologies,<sup>32</sup>

our transcriptome complement was, therefore, a reasonable representation of robustly detectable and actively transcribed genes within the technical limitations of this methodology.

**An initial comparison of nonrejection and acute rejection patient samples did not reveal a gene expression signature specific to acute allograft rejection.** In an attempt to uncover differential gene expression signatures that correlated with acute allograft kidney rejection, we compared transcriptome data from 24 AR samples at the time of rejection with their matched 24 NR samples. To do this, we first simplified the analysis by performing data reduction using PCA (principal component analysis),<sup>24</sup> a method that creates a visualization of different clusters in two- or three-dimensional space, allowing the highest variations to be determined. PCA analysis revealed no significant variance (in the 2 principal components) that separated AR samples from NR samples (Fig. 1A). However, we found that the most significant of the criteria that separated the sample data was time posttransplantation, regardless of rejection status. Specifically, samples from early time points posttransplant separated from samples taken in weeks 3 and beyond (Fig. 1B). This result highlighted the importance of precise sample matching between the AR and NR patients groups with respect to sample collection time posttransplantation. A closer examination of the initial sample matching between AR and NR patients revealed an imbalance with respect to collection time in the first week post transplant. For example, 7 AR samples collected at day 3 or 4 were originally matched with 7 NR samples from day 6 or 7. Given the potential for the presence of dynamic and extensive transcriptional changes post surgery, the imprecise time matching of the samples may have introduced a confounding factor to the analysis, potentially masking rejection-associated expression changes.

**Gene expression changes common to transplant patients are time-dependent and include known targets of immunosuppression and processes involved in inflammation and innate immunity.** In order to further investigate the time-dependent gene expression changes that were separating the patient data, we compared the 24 AR and 24 NR patient samples using 20 healthy samples as a reference. PCA analysis of patient and control samples revealed that the highest variance present in these data was a molecular signature comprised of genes that showed an overall lower expression level among transplant patient samples when compared with control samples (Additional file 3, A and B). Furthermore, a high variability in gene expression levels was observed within the transplant patient data. PCA analysis using only the genes present in the differential gene expression signature revealed that patient samples collected at earlier time points (ie, those taken from a few days posttransplant) clustered more distantly from the control subject samples, regardless of their rejection status (Additional file 3, A). These genes were more severely downregulated in both early AR and NR time point samples. A clear separation was observed between samples collected

at days 2–4 and days 6–10 posttransplant (Fig. 1C). Using InnateDB to identify specific biological processes associated with the highly downregulated genes revealed an overrepresentation of pathways related to T cell and B cell proliferation and activation such as TCR, downstream signaling in naïve CD8+ T cells, IL-2, IL-12 pathways, and BCR (Table 3). Having identified a time dependent downregulated gene expression signature that was common to transplant patients, we hypothesized that an expression signature representing upregulated genes would also be present. Pavlidis template matching (PTM)<sup>25</sup> is an algorithm that can be tailored to specify a template expression profile, or use a specific gene as a template, to search for matches to the template within an entire dataset. The match is based on the Pearson correlation between the template and the genes in the data set, and the threshold criterion for matching can be either the magnitude of the correlation coefficient or the significance (*P* value) of the correlation coefficient. We used PTM to define a gene expression pattern that was highest at the earliest time point and lower at the later time point in order to detect genes with elevated expression in the early time point samples (Additional file 3, C). Examination of biological pathways overrepresented in this cluster of genes revealed major mediators in the early innate immune response (Table 4), including TLR signaling and proinflammatory IL-1 (IL1B) pathways. Together, these data suggested that the differential gene expression signatures in early time point samples might arise from the contribution

of strong immunosuppressive treatment immediately following transplantation.

**A comparison of closely matched acute rejector and nonrejector samples from a relatively late time point (week 3 postsurgery and beyond) revealed acute allograft rejection-associated differential gene expression signatures.** We found that dynamic expression changes common to both AR and NR patients at early time points, combined with an imbalance in sample matching in this period, was potentially confounding the identification of gene expression signatures associated specifically with transplant rejection. Therefore, to better reveal such signatures, we focused our analysis on samples taken from a relatively late rejection patient group (ie, minimum 15 days posttransplant) (Additional file 4, demographics and clinical data and the confounding analysis results for late 8 ARs and matched 8 NRs). This limited our analysis to a group of 8 AR patients who had a confirmed rejection episode between weeks 3 and 12 postsurgery and their matched NR patient samples. This highly matched patient case-control group did not show any significant confounding factors arising from demographics or clinical background except for the previously stated and expected creatinine and BUN readings.

**Table 3.** InnateDB analysis of biological processes associated with the genes downregulated in both AR and NR transplant patient samples (compared with healthy controls) taken after 2 or 3 days following transplant.

PATHWAY NAME	PATHWAY ID	SOURCE NAME	P-VALUE (CORRECTED)
Downstream signaling in naïve CD8+ T cells	9376	PID NCI	4.21E-08
IL12-mediated signaling events	9386	PID NCI	4.50E-07
T cell receptor signaling pathway	563	KEGG	1.85E-06
TCR signaling in naïve CD8+ T cells	9330	PID NCI	1.83E-06
TCR signaling in naïve CD4+ T cells	9420	PID NCI	5.30E-06
Lck and fyn tyrosine kinases in initiation of tcr activation	4116	PID BIOCARTA	1.51E-05
IL2 signaling events mediated by STAT5	9363	PID NCI	2.36E-05
BCR	3931	NETPATH	7.43E-05
Role of mef2d in t-cell apoptosis	4085	PID BIOCARTA	4.01E-04
Primary immunodeficiency	2815	KEGG	8.24E-04

**Table 4.** InnateDB analysis of biological processes associated with the genes upregulated in both AR and NR transplant patient samples (compared to healthy controls) taken after 2 or 3 days following transplant.

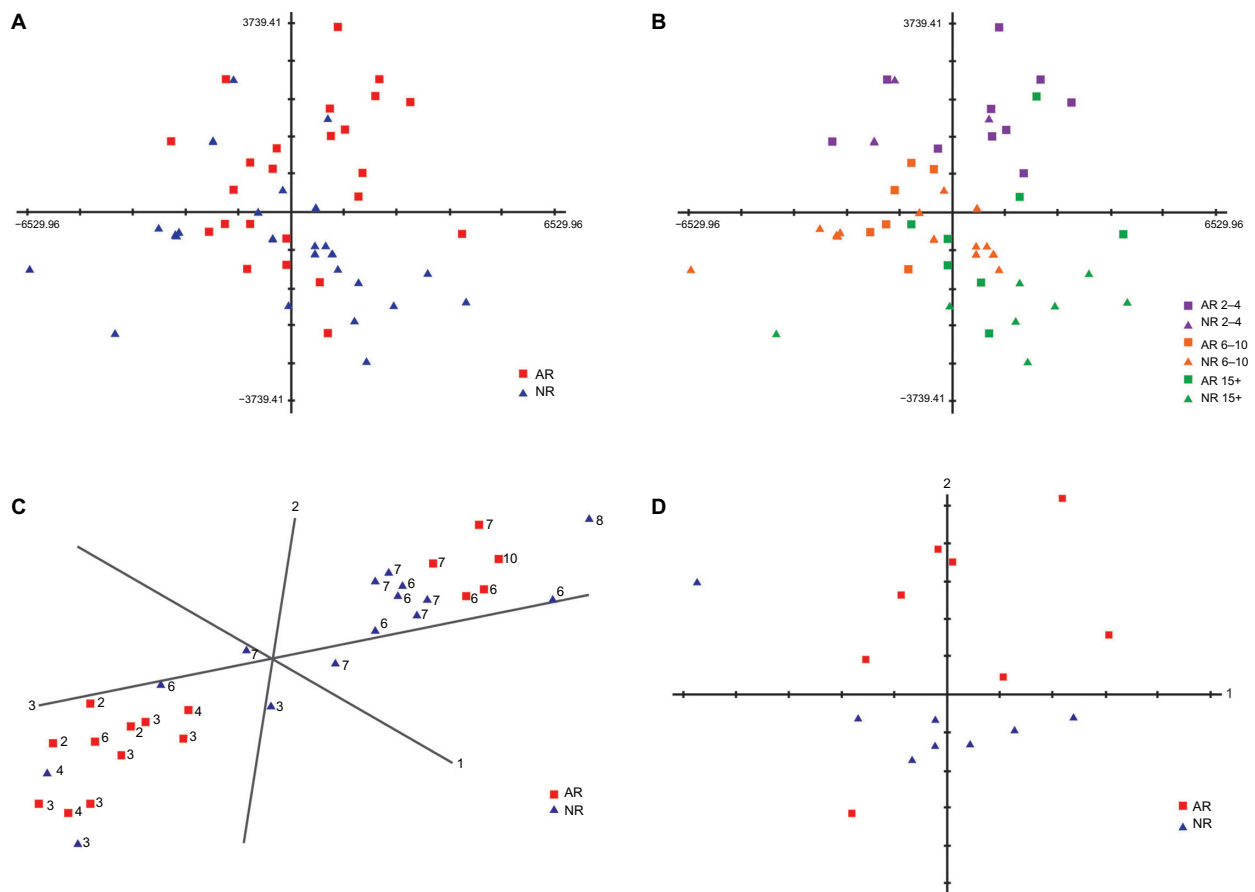
PATHWAY NAME	PATHWAY ID	SOURCE NAME	P-VALUE (CORRECTED)
Toll-like receptor signaling pathway	564	KEGG	1.90E-02
Endogenous TLR signaling	10525	PID NCI	1.33E-02
Complement and coagulation cascades	456	KEGG	1.47E-02
Osteoclast differentiation	10367	KEGG	1.31E-02
Hematopoietic cell lineage	415	KEGG	1.85E-02
Cytokine receptor degradation signaling (JAK-STAT pathway and regulation pathway Diagram)	9939	INOH	1.67E-02
Glycogen breakdown (glycogenolysis)	12814	REACTOME	1.86E-02
Regulation of IFNA signaling	12587	REACTOME	1.86E-02
Negative feedback regulation of JAK STAT pathway by (cytokine receptor degradation signaling)	9645	INOH	1.79E-02
NOD-like receptor signaling pathway	8112	KEGG	3.61E-02
IL1	10429	NETPATH	4.41E-02

However, the age of the donor was found to be significantly different between the AR and NR patient groups (adjusted  $P$  value =  $2.78E-03$ ). In general, the transplant donors of AR patients were older than for NR patients. It has previously been reported that increased donor age might have a significant and negative effect on transplant outcome.<sup>33,34</sup>

We generated a PCA plot comparing the 8 AR samples at the time of rejection against their 8 time point matched NR samples. This plot demonstrated a much better separation between the AR and NR samples (Fig. 1D), revealing a significant differential expression signature specific to the AR patient group. Using SAM, we identified 120 genes that displayed significantly lower expression and 11 genes that were significantly more highly expressed in the AR samples ( $FDR < 0.05$ ) (Additional files 5 and 6). Using InnateDB, we found that the downregulated expression signature was enriched for genes involved in lymphocyte activities

including *TCR*, *CD28*, *JAK-STAT*, *EPO*, *IL-2*, *IL-10* anti-inflammatory, and *IL-7* signaling pathways (Additional file 7). Conversely, genes that were upregulated were indicative of innate immune responses and included neutrophil related genes such as the granule protein genes, *ANXA3*, *MMP25*, and *RAB31*.

**Differential gene expression analysis in time course samples from 15 days posttransplant up to rejection further expands rejection-associated gene expression signatures.** Having identified significantly differentially expressed genes at the point that the rejection state was confirmed, we wished to identify additional genes that were more highly or lowly expressed in patients as they approached rejection. To do this, we expanded our analysis to include all available time course samples from the late AR patient group from day 15 up to the rejection time point (19 AR and 16 NR samples in total) and analyzed them using SAM. This approach uncovered



**Figure 1.** Comparison of 24 ARs at rejection and 24 matched NRs. **A.** PCA plot of 24 AR patient samples at the time of rejection and their 24 matched NR patient samples. AR and NR samples do not separate clearly. **B.** The same PCA plot of 24 AR and 24 NR patient samples as in A. Samples are highlighted by the time (days) of rejection (ie, time of sample collection since transplant). Sample separation based on the time of collection can be seen indicating the presence of a time-dependent gene expression signature. **C.** PCA plot of samples from days 2 to 10 posttransplant generated using the time-dependent gene expression signature. Overall, a strong separation of samples from days 2 to 4 and days 6 to 10 posttransplant is observed (the number shown next to each sample point is the number of days posttransplant for that sample). **D.** PCA plot of 8 late ARs (week 3 and beyond, posttransplant) and the matched NRs using all the filtered probe set data. These late ARs separate from NRs more clearly compared to the early ARs and NRs.



additional genes that were significantly upregulated or downregulated specifically in the AR samples (Additional file 8). Biological activities overrepresented by this expanded list of downregulated genes also included lymphocyte activities such as TCR, IL12, IL7, and IL2 pathways (Table 5). Biological activities overrepresented by the extended highly expressed gene set included known markers of systemic inflammation, including IL-1, IL-6, TLR, TNF alpha, NFκB, and Pentose phosphate pathways (Table 6). Additionally, we found that the most significant differentially expressed genes from this cluster overlapped with 286 neutrophil granule protein genes identified from a proteomic analysis of the azurophil, specific, and gelatinase granules from human neutrophils.<sup>35</sup>

Commonly identified genes included *ANXA3*, *LCN2*, *LTF*, *MMP9*, *PGLYRP1*, *MGAM*, *PAD14*, *CR1*, *DYSF*, *SLPI*, *SIGLEC5*, *LYZ*, *SLC2A3*, *RAB31*, *ITGAM*, *HCK*, *QPCT*, *FLOT1*, *MMP25*, *FLOT2*, *PTPRJ*, *GSN*, *ADAM8*, *MVP*, *RAB27A*, *TLN1*, *LILRB2*, *IQGAP1*, *FPR1*, and *GNB2*. Overall, approximately 14% of the AR associated upregulated genes included neutrophil granule protein genes.

**Upregulated genes associated with acute allograft rejection are highly enriched in neutrophils; genes that are downregulated are enriched in lymphocytes.** To further investigate the biological functions of the differentially expressed genes associated with rejection, we used gene enrichment profiles created by Benita et al<sup>36</sup> to examine their

**Table 5.** Overrepresented pathways defined by the cluster of genes significantly downregulated in ARs compared to NRs identified from the InnateDB analysis using time course samples collected minimum 15 days post-TX.

PATHWAY NAME	PATHWAY ID	SOURCE NAME	PATHWAY P-VALUE (CORRECTED)
TCR	10409	NETPATH	1.88E-06
IL12-mediated signaling events	10615	PID NCI	3.81E-06
TCR signaling in naïve CD8+ T cells	10538	PID NCI	6.03E-06
The co-stimulatory signal during t-cell activation	4018	PID BIOCARTA	1.92E-05
Downstream signaling in naïve CD8+ T cells	10485	PID NCI	2.14E-05
Lck and fyn tyrosine kinases in initiation of tcr activation	4116	PID BIOCARTA	2.19E-05
Il-7 signal transduction	4180	PID BIOCARTA	1.16E-04
Generation of second messenger molecules	12530	REACTOME	1.36E-04
Role of mef2d in t-cell apoptosis	4085	PID BIOCARTA	1.52E-04
T cell receptor signaling pathway	563	KEGG	2.81E-04
Primary immunodeficiency	2815	KEGG	4.11E-04
TCR signaling in naïve CD4+ T cells	10446	PID NCI	9.56E-04
IL12 signaling mediated by STAT4	10626	PID NCI	9.58E-04
Translocation of ZAP-70 to Immunological synapse	12531	REACTOME	9.99E-04
IL-7	10408	NETPATH	1.24E-03
Phosphorylation of CD3 and TCR zeta chains	12532	REACTOME	1.58E-03
Nef and signal transduction	12866	REACTOME	2.87E-03
PD-1 signaling	12525	REACTOME	2.94E-03
T cell receptor signaling pathway	4156	PID BIOCARTA	2.97E-03
Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell	12533	REACTOME	3.33E-03
Downstream TCR signaling	12529	REACTOME	3.66E-03
Activation of csk by camp-dependent protein kinase inhibits signaling through the t cell receptor	4160	PID BIOCARTA	4.52E-03
Hematopoietic cell lineage	415	KEGG	4.66E-03
Cell surface interactions at the vascular wall	12632	REACTOME	5.12E-03
CD4 T cell receptor signaling (through Vav, Rac and JNK cascade ( CD4 T cell receptor signaling (JNK cascade) )	10184	INOH	5.12E-03
Cytosolic tRNA aminoacylation	12381	REACTOME	5.47E-03
Interleukin-7 signaling	12386	REACTOME	5.79E-03
Valine, leucine and isoleucine biosynthesis	408	KEGG	5.79E-03
IL2 signaling events mediated by STAT5	10632	PID NCI	9.08E-03

**Table 6.** Overrepresented pathways defined by the cluster of genes significantly upregulated in ARs compared to NRs identified from the InnateDB analysis using time course samples collected minimum 15 days post-TX.

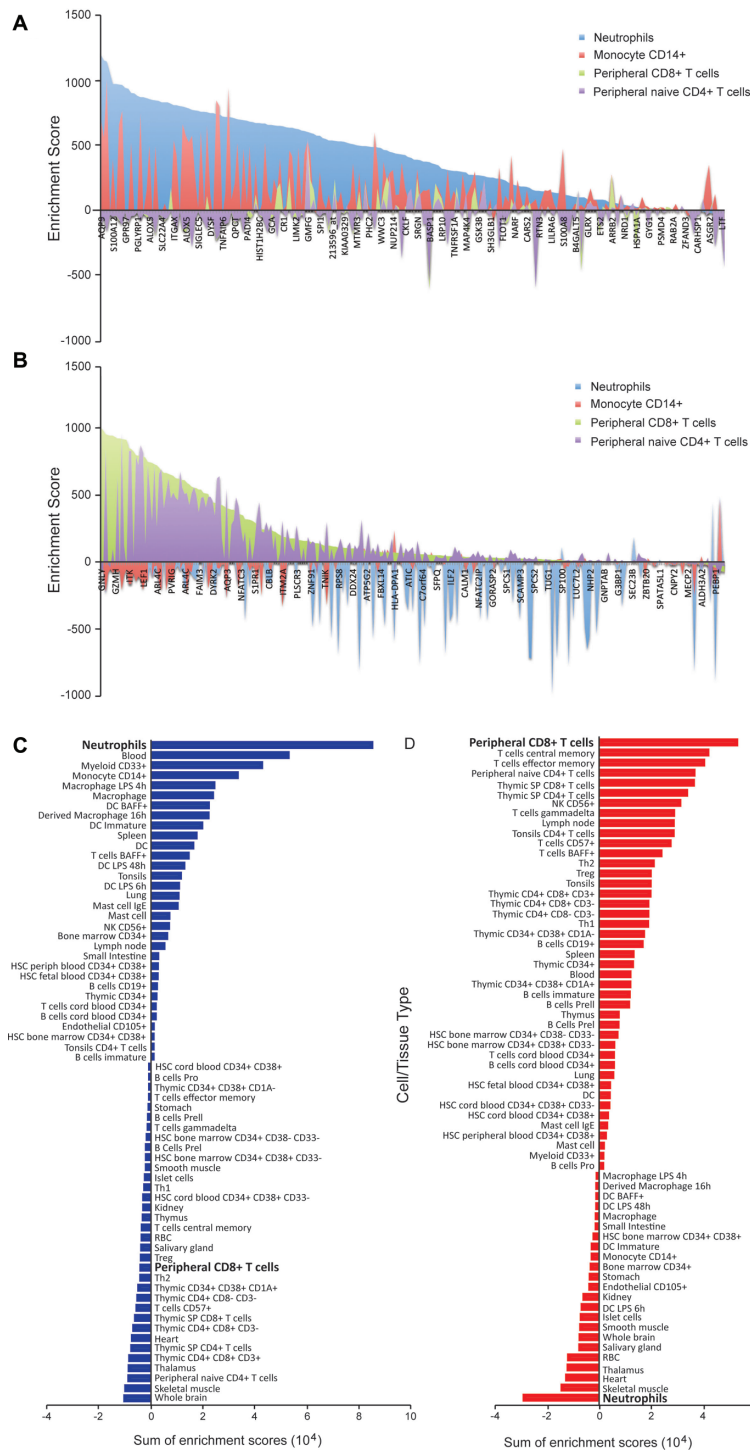
PATHWAY NAME	PATHWAY ID	SOURCE NAME	P-VALUE (CORRECTED)
IL1	10429	NETPATH	1.19E-02
Role of Calcineurin-dependent NFAT signaling in lymphocytes	10555	PID NCI	1.53E-02
Pentose phosphate pathway (hexose monophosphate shunt)	12813	REACTOME	2.27E-02
MAPK signaling pathway	487	KEGG	3.29E-02
Lissencephaly gene (LIS1) in neuronal migration and development	10548	PID NCI	3.37E-02
BMP receptor signaling	10455	PID NCI	3.40E-02
TNFalpha	10418	NETPATH	3.42E-02
IL6	10415	NETPATH	3.46E-02
NfκB activation by nontypeable hemophilus influenzae	4159	PID BIOCARTA	3.57E-02
Reelin signaling pathway	10605	PID NCI	3.57E-02
Signal transduction through il1r	4064	PID BIOCARTA	3.78E-02
TWEAK	10426	NETPATH	3.78E-02
Pentose phosphate pathway	470	KEGG	3.86E-02
Endogenous TLR signaling	10525	PID NCI	3.90E-02
Trk receptor signaling mediated by the MAPK pathway	10503	PID NCI	3.90E-02
Toll-like receptor pathway	3951	PID BIOCARTA	3.92E-02
Map kinase inactivation of smrt corepressor	4027	PID BIOCARTA	3.98E-02
IL3	10406	NETPATH	3.99E-02
Mapkinase signaling pathway	4137	PID BIOCARTA	4.15E-02
Pentose phosphate cycle (Pentose phosphate cycle )	9670	INOH	4.34E-02
Retinoic acid receptors-mediated signaling	10644	PID NCI	4.39E-02
Tnf/stress related signaling	4103	PID BIOCARTA	4.55E-02
TGF beta Receptor	10428	NETPATH	4.67E-02

significance across various cell and tissue types. As expected, neutrophil activity-associated genes that were upregulated showed significant enrichment in neutrophils (Fig. 2A). Conversely, the downregulated genes were found to be highly enriched in CD8+ T cells and mostly absent from neutrophils (Fig. 2B). This analysis provided evidence that the gene expression signatures might be cell specific and representative of both an increase in neutrophil activity and reduction of lymphocyte development and differentiation. Furthermore, the sum of enrichment scores for all genes from each signature across various cell and tissue types indicated that the AR associated upregulated genes are most highly enriched in the myeloid lineage while the downregulated genes are most highly enriched in T lymphocytes (Fig. 2C and 2D).

**Rejection-associated gene expression signatures are modulated in a manner associated with rejection treatments.** We reasoned that if the differential gene expression signatures we have described were specifically associated with acute allograft rejection, these signatures might be modulated in response to rejection therapies applied after primary confirmation of the rejection episode. To test this hypothesis, we interrogated patient transcriptome data collected after the time that rejection was confirmed and when rejection treatments were initiated. A PCA plot generated using the most significantly differentially expressed genes from both neutrophil and lymphocyte molecular signatures revealed a clear separation of the 8 AR patients and their matched 8 NR counterparts (Fig. 3A). Additionally, the expression levels of these genes were altered in AR patients undergoing treatment for rejection (ie, treated with solumedrol, prednisone, or antithymocyte globulin [ATG]). Specifically, genes upregulated in the neutrophil signature showed a decreased expression after rejection treatment (Fig. 3B). Conversely, downregulated genes associated with the lymphocyte gene signature exhibited a general increase in expression post rejection treatment (Fig. 3C).

**Analysis of external case-control patient data demonstrates concordance with molecular expression signatures that are associated with acute allograft rejection.** To independently assess the reproducibility of the rejection-associated signatures, we obtained additional data from an independent multicentre international cohort of 15 AR and 22 NR kidney transplant patients. Data from this external patient group were generated using Affymetrix Human GeneST 1.1 arrays (Additional file 9, A and B), which differed from our microarray data, which were generated using Affymetrix U133 plus 2.0 arrays. To determine if the differential gene expression signatures were reproducible across both platforms, we generated transcriptome data using the Human GeneST 1.1 arrays for the rejection samples from our 8 AR patients described above and their matched NR samples. A PCA plot comparison of the neutrophil and lymphocyte signature genes generated from both platforms showed a similar separation. This indicated that these rejection-associated gene expression signatures were consistent across the 2 microarray platforms (Additional file 9, C and D). Subsequent SAM analysis

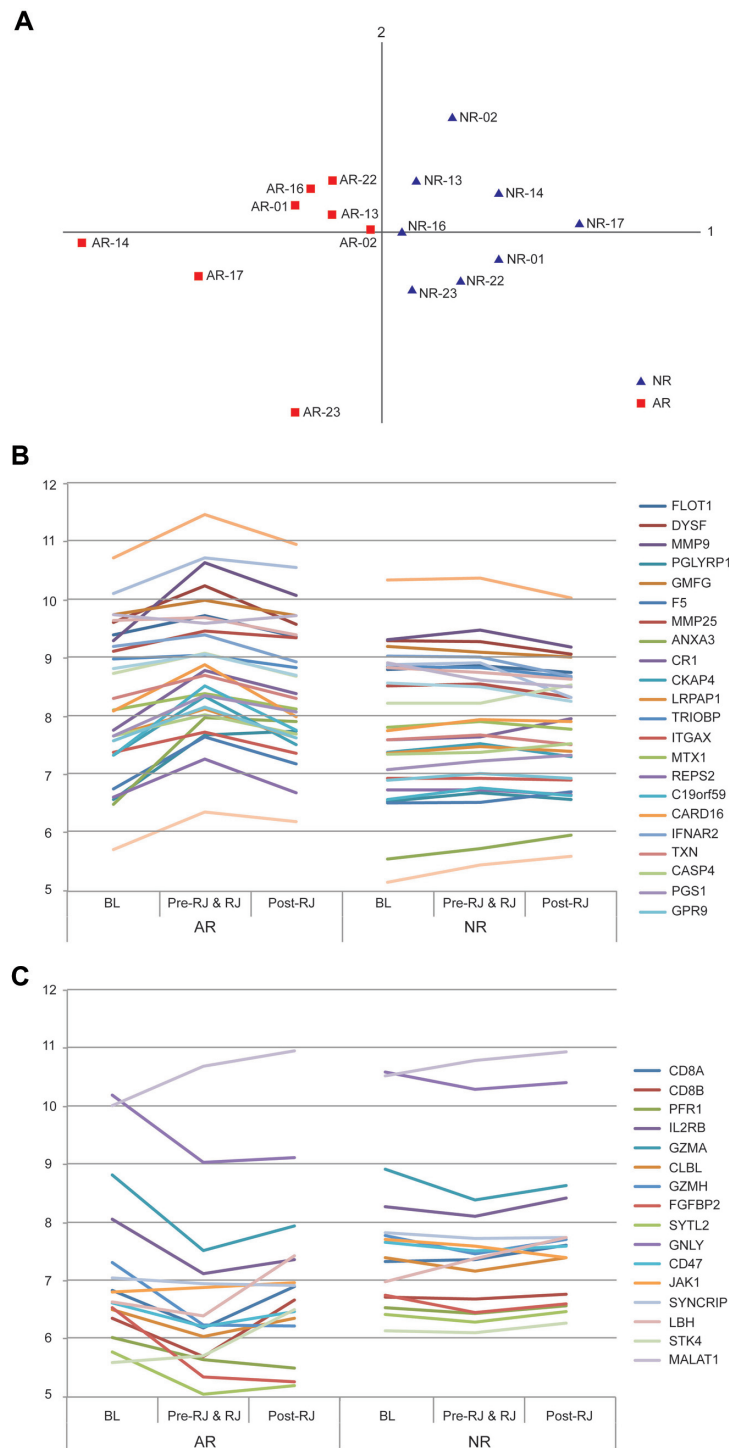




**Figure 2.** Genes upregulated approaching rejection are highly neutrophil enriched while genes downregulated are highly lymphocyte enriched. **A.** Comparison of gene enrichment scores (Benita et al<sup>36</sup>) for genes that are upregulated approaching rejection in neutrophils and monocyte CD14+ versus peripheral CD4+ and CD8+ T cells. These genes are highly enriched in neutrophils and monocytes compared to peripheral T cells. A positive score indicates enrichment and negative score indicates absence. **B.** Comparison of gene enrichment scores for genes that are downregulated approaching rejection in neutrophils and monocytes versus peripheral CD4+ and CD8+ T cells indicates these genes are highly enriched in T lymphocytes. **C** and **D.** Sum of enrichment scores for various blood cell and different tissue types shows that upregulated genes represent a neutrophil signature (C) and that downregulated genes represent a T lymphocyte signature (D).

of statistically significant differential gene expression signatures between the 15 AR and 22 NR patients in the external validation patient group demonstrated that neutrophil genes showed an overall upregulation in AR samples in comparison with NR

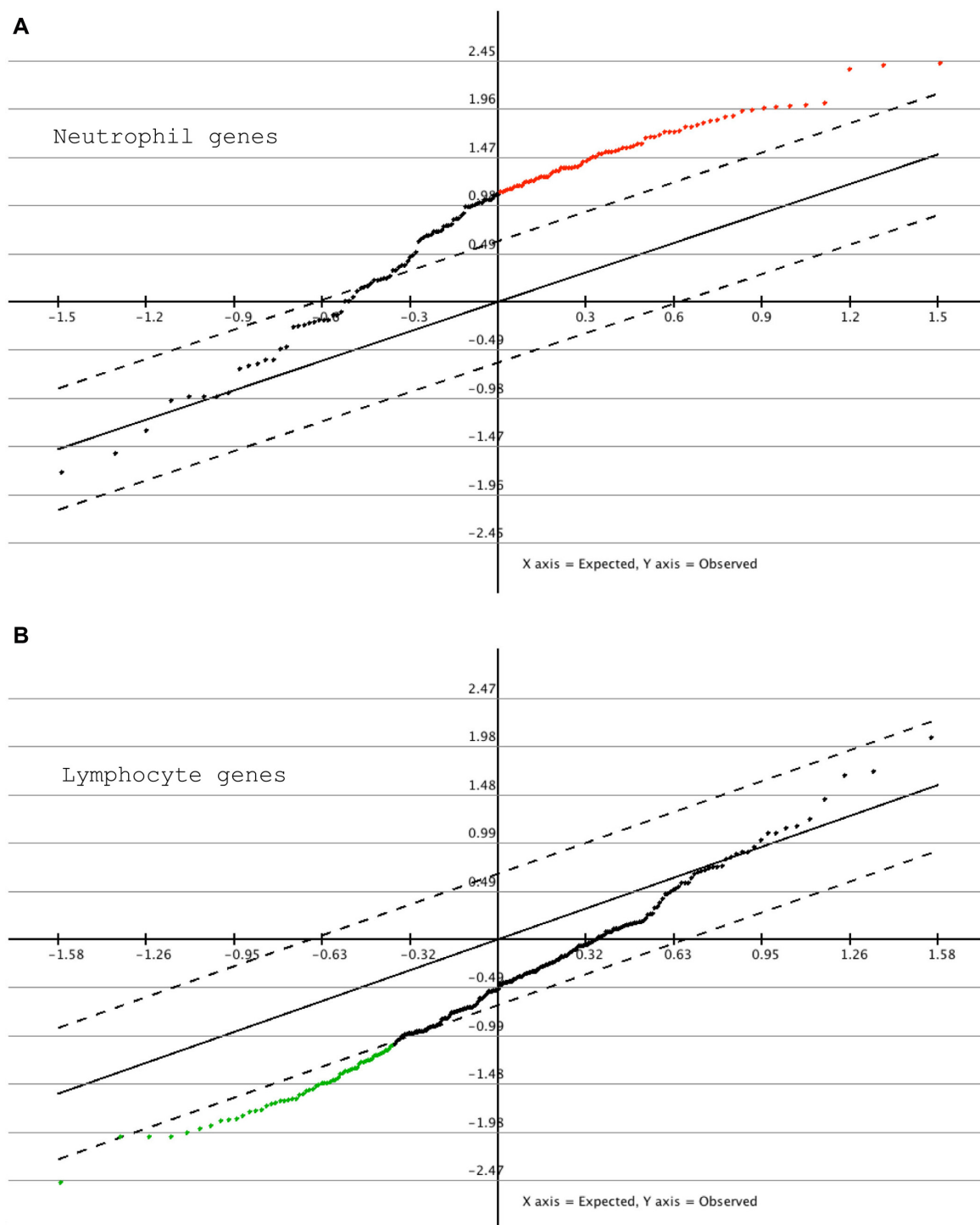
samples. Approximately half of the signature genes showed statistically significant differential expression in this analysis (FDR < 0.05) (Fig. 4A). Conversely, lymphocyte genes were in general downregulated in ARs in this patient group, with



**Figure 3. A.** PCA plot generated by the representative set of the most significant genes from both neutrophil and lymphocyte signatures shows clear separation of ARs from NRs. **B.** Average gene expression levels in 3 time periods: pretransplant (BL), posttransplant up to rejection (Pre-RJ and RJ), and postrejection (Post-RJ). A representative set of the neutrophil signature genes show upregulation from BL towards rejection and downregulation after rejection treatment in AR samples; these same genes show a more stable level of expression in NR samples. **C.** A representative set of the lymphocyte signature genes show downregulation from BL towards rejection and upregulation after rejection treatment in AR samples; these same genes show more stable expression in NR samples.

approximately a third of the signature genes showing a statistically significant differential expression ( $FDR < 0.05$ ) (Fig. 4B). Given that the transcriptome samples from the external validation cohort were not precisely time-matched, the level of validation described in this analysis might represent an overestimate or underestimate

of the actual reproducibility of the rejection signatures. Overall, however, this analysis demonstrated that the gene expression signatures we have identified as associated with acute allograft rejection replicate to a reasonably significant degree in an independent group of kidney transplant patients.



**Figure 4.** Validation of the rejection-associated signatures using external kidney transplant study data. Validation study microarray data were generated using the GeneST 1.1 platform on 15 AR and 22 NR sample sets (BiT2). **A.** SAM analysis of 15 AR samples when compared with the 22 NR samples indicates a general upregulation of neutrophil specific genes, of which approximately half are statistically significant (red data points). **B.** A general downregulation of most lymphocyte specific genes is observed, of which approximately a third are statistically significantly downregulated (green data points).

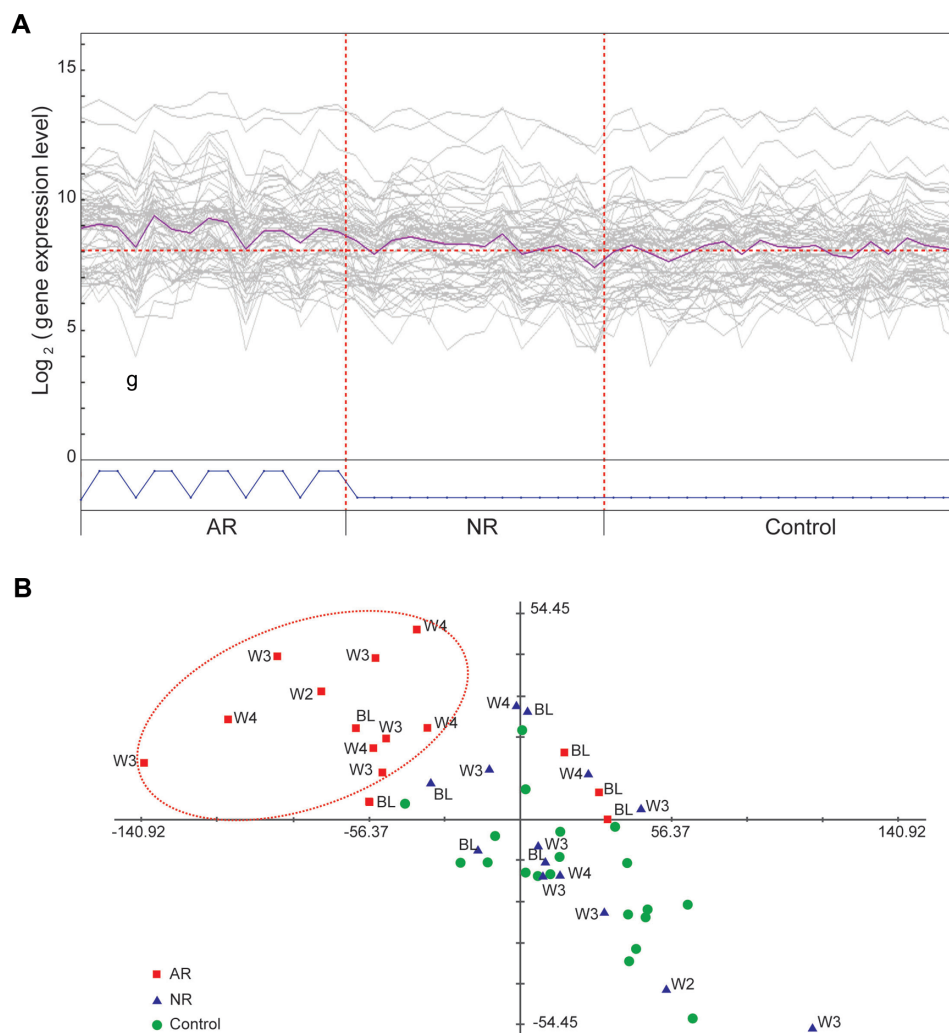
**Biological processes that are representative of the neutrophil gene expression signature are increasingly upregulated posttransplantation and approaching rejection.** We have described differential gene expression signatures in peripheral blood leukocytes that are associated with acute

allograft rejection. In particular, the increasing gene expression changes described by the neutrophil associated signature have potential utility for predicting transplant patient outcome. To refine this expression signature, we interrogated patient time course data for genes whose expression was specifically

upregulated in AR patients as they approached rejection using an artificially defined pattern of increasing expression. For this analysis, we selected time point data that encompassed samples collected from pretransplant up to and including confirmation of rejection. For consistency, we selected precisely matched time points, which limited this analysis to a small cohort of 5 AR patients and their matched NR samples. Three time points—at pretransplant baseline (BL), at 7 days prior to rejection (-7), and at the time of rejection (RJ) (ie, biopsy confirmation)—were selected within this patient group. We then performed pattern matching analysis using PTM to identify genes that were upregulated specifically in AR patients from pretransplant levels and maintained at a high expression level as they approached the rejection time point (Fig. 5A). Twenty-nine genes followed this pattern with an  $R$  greater than 0.5 (Table 7). A PCA plot using these data confirmed that the

AR patients were increasingly separating from their matched NR patients and control patients as they approached rejection (Fig. 5B). Consistent with the findings described above, the genes that most closely followed this pattern of expression were related to activated neutrophils and known neutrophil granule protein genes.<sup>37–40</sup>

**Functional and transcriptional coherency genes that covary with *DYSF* include functionally related neutrophil protein genes and potential novel members of a biological network.** Having identified genes that were specifically upregulated in AR patients as they approached rejection, we wished to further investigate the potential for identifying tightly coregulated genes within this pattern. To do this, we took the expression pattern of one of the most significantly upregulated genes in rejection patients, *DYSF* ( $P$  value = 2.79E-06) and identified additional genes that covaried stringently with its expression



**Figure 5.** Pattern matching of gene expression changes over time reveals a molecular signature associated with AR. **A.** PTM analysis showing genes matching the expression signature of upregulation from pretransplant to posttransplant in AR patients only. Five AR patients and 5 matched NR patients with the precisely matched time points were used for the analysis. The 3 time points used (plotted consecutively for each patient) are pretransplant (BL), 7 days prerejection (-7), and at rejection (RJ). **B.** PCA plot visualizing the expression signature described in A. The AR patients (red squares, with week [W] posttransplant data indicated) are moving away from their pretransplant time point (BL) as well as from control patients (green circles) and NR patients (blue triangles) as they are approaching rejection.

**Table 7.** Twenty-nine genes upregulated from pretransplant levels and maintained at a high expression level as they approach the rejection time point identified from the PTM (Pavlidis Template Matching) analysis.

PROBESET	GENE SYM	ENTREZ	R	P-VALUE
204220_at	GMFG	9535	0.653009	3.67E-07
207384_at	PGLYRP1	8993	0.619712	2.05E-06
208864_s_at	TXN	7295	0.617343	2.30E-06
218660_at	DYSF	8291	0.613392	2.79E-06
221058_s_at	CKLF	51192	0.612538	2.90E-06
201186_at	LRPAP1	4043	0.599121	5.41E-06
209369_at	ANXA3	306	0.594741	6.60E-06
228758_at	BCL6	604	0.594195	6.76E-06
1552701_a_at	CARD16	114769	0.587988	8.89E-06
200999_s_at	CKAP4	10970	0.587098	9.25E-06
202795_x_at	TRIOBP	11078	0.585577	9.88E-06
204714_s_at	F5	2153	0.56244	2.60E-05
220404_at	GPR97	222487	0.561969	2.65E-05
210386_s_at	MTX1	4580	0.554397	3.58E-05
210142_x_at	FLOT1	10211	0.546598	4.84E-05
205645_at	REPS2	9185	0.540405	6.12E-05
224637_at	OST4	1.00E+08	0.538789	6.50E-05
223451_s_at	CKLF	51192	0.534389	7.65E-05
235568_at	C19orf59	199675	0.527377	9.88E-05
209310_s_at	CASP4	837	0.523987	1.12E-04
204786_s_at	IFNAR2	3455	0.520773	1.25E-04
202121_s_at	CHMP2 A	27243	0.519799	1.29E-04
204971_at	CSTA	1475	0.513932	1.59E-04
229597_s_at	WDFY4	57705	0.51145	1.73E-04
208749_x_at	FLOT1	10211	0.511329	1.74E-04
201463_s_at	TALDO1	6888	0.510643	1.78E-04
212531_at	LCN2	3934	0.508234	1.93E-04
203814_s_at	NQO2	4835	0.504014	2.22E-04
229967_at	CMTM2	146225	0.502453	2.34E-04

pattern over the entire time course from the 48 patients. From this approach, we identified 369 genes that covaried with *DYSF* ( $R > 0.6$ ) (Additional file 10). This group comprised the majority of genes upregulated approaching rejection (~75%) and included many previously identified neutrophil granule proteins (~21%).<sup>35</sup> The genes whose expression most tightly correlated with *DYSF* included other neutrophil granule membrane protein genes such as *NCF4* and *FLOT1* ( $R > 0.85$ ) as well as the neutrophil granule membrane marker protein CD63. Gene Ontology (GO) analysis revealed that top hits ( $R > 0.75$ ) were significantly enriched with genes encoding integral membrane proteins (Additional file 11). From this analysis we observed functional and transcriptional coherency within our datasets and identified novel genes potentially functionally related to genes that are operating in activated neutrophils.

**Identification of highly represented transcription factor binding sites within the differentially expressed gene signatures reveal potential transcriptional factors regulating the biological response.** Finally, our finding that the acute allograft associated differential gene expression patterns displayed a high degree of transcriptional coherency suggested the potential for coregulation of these molecular signatures by specific transcription factors. The identification of such transcriptional hubs might provide suitable targets for the development of therapeutic strategies directed at modulating these expression signatures. In order to gain insight into differential transcriptional regulation in these cell specific signatures, we examined overrepresented transcription factor binding sites among the neutrophil and lymphocyte enriched genes by using oPOSSUM.<sup>29</sup> From this analysis, we identified the most highly represented site in the genes of the T lymphocyte signature was specific to the ETS family transcription factor, GABP ( $z$  score = 15.23, Fisher score = 3.85E-04). Furthermore, analysis of transcription factor binding sites using the genes described in the neutrophil signature identified NF- $\kappa$ B/RelA as the most highly represented transcription binding site in this gene signature ( $z$  score = 8.10, Fisher score = 2.38E-02) (Additional file 12).

## Discussion

In this study, we have used whole blood transcriptomics to investigate biological processes associated with kidney transplantation and acute allograft rejection. We have been able to identify molecular signatures that correlate with the underlying biology associated with both kidney transplantation and subsequent rejection. These signatures are found to be leukocyte cell type-specific, which potentially reflects systemic inflammation and immune dysfunction in patients following surgery and approaching rejection. Questions as to whether such activities are causal for allograft rejection are beyond the scope of this study, though investigating the primary kidney tissues (not available to us in this instance) might be a useful next step.

While transcriptional activities in peripheral whole blood can provide a system-wide picture of complex immune activities and networks, challenges and limitations to this type of approach exist. Interpretation of the vast amount of multi-dimensional genomic data can be complex and problematic. Limited patient selection and variation in patient treatment can also introduce potentially confounding factors into the data. Transcript abundance in highly complex tissues, such as blood, may be convoluted because transcript expression levels originate not only from differential regulation of gene transcription but also from changes in cell type population composition.<sup>41,42</sup> Previously, Shen-Orr et al have developed a deconvolution method (csSAM), which takes advantage of complete blood cell count and differential (CBC/diff) data to statistically estimate differential gene expression changes within each cell type.<sup>43</sup> Using this approach, Shen-Orr et al<sup>43</sup>



showed that groups of significantly upregulated genes within neutrophils and monocytes and significant downregulated genes within lymphocytes were associated with acute rejection in pediatric kidney transplant patients. While we have not undertaken csSAM analysis on our “late” adult kidney rejection samples (due to small sample size and lack of comprehensive CBC/diff data for all patients), our results are consistent with those of Shen-Orr et al<sup>43</sup> and build on their findings, providing further support for the molecular changes in these cell types associated with acute allograft rejection. Additionally, our analysis of time course data provides a biological context with which to interpret the observed expression changes motivated by a desire for biological understanding of the pattern of gene expression changes seen in the transplant recipient over time leading up to rejection. We have also collected whole blood samples that provide an unbiased representation of the transcriptomes from all blood cell types, including neutrophils. This is in contrast to most previously published studies, which are based on lymphocyte biased data that underrepresents or excludes neutrophils (such as peripheral blood mononuclear cell [PBMC] samples),<sup>6</sup> or the use of a lymphocyte biased platform (ie, lymphochip cDNA platform).<sup>7,8</sup>

From the initial analysis of all 24 AR and 24 NR patient samples, we characterized a differential gene expression signature that was associated with early time points posttransplant (ie, first 4 days posttransplant) and common to both patient groups regardless of their eventual rejection status. This signature was comprised of genes known to be involved in immunosuppression, such as the *IL-2* and *TCR* pathway. That such a signature existed in all patient samples might not be surprising given that immunosuppressive treatments are a standard regimen for transplant patients. Furthermore, samples collected within the first 4 days since transplant surgery in both patient groups showed the most significant differential gene expression, potentially reflecting the heavier immunosuppressive induction therapy provided to patients immediately after surgery. Additionally, using a pattern matching approach we were able to define a gene signature representing upregulated genes in early samples following surgery that included major mediators in the early innate immune response, including TLR signaling and proinflammatory IL-1 (IL1B) pathways. However, the presence of these expression signatures was potentially problematic for elucidating rejection-associated events and highlighted the need to more precisely match AR and NR patients based on sample collection time. To reduce these potential confounding factors we focused our analysis on samples from AR and NR patient groups that were more precisely matched and collected at least 15 days following transplant surgery.

The analysis of time-matched postsurgery transcriptome samples revealed gene expression signatures that were significantly different between the AR and NR patient groups. The most significant differences were correlated with cell type-specific processes and adaptive and innate immune responses

that can be categorized into 2 general signatures, T lymphocyte activity and neutrophil activity. We found that genes significantly downregulated in transplant patients approaching rejection were enriched in T lymphocytes and involved in processes of T lymphocyte proliferation and activation. Conversely, we found that gene transcripts significantly elevated in the acute rejection patient group as they approached rejection were highly enriched in neutrophils and correlated with neutrophil activities. These molecular signatures appeared to be specific to the acute rejection of the allograft as we observed a general stabilization of the molecular signature in patient samples taken postrejection and after immunosuppressive treatment had been initiated to alleviate the rejection. Furthermore, we found that both the neutrophil and the lymphocyte gene expression signatures that we had described were reasonably conserved in an independent group of kidney transplant patients from a multicentre international study.

We found that lymphocyte genes were in general downregulated in all patient samples. One possible explanation for this is the immunosuppression regime the patients are exposed to as described above. However, persistent suppression of lymphocyte genes was a highly significant feature unique to AR patients approaching rejection when compared with recovering NR patients. Since depression of lymphocyte count and elevation of neutrophil count is known to occur in response to infections or glucocorticoid administration, it is possible that the AR associated signature is due to the presence of an infection or due to steroid treatment. However, at the time of sample collection, patients were not undergoing treatment with glucocorticoids. Furthermore, the treatment regimen did not differ between AR and NR groups. Additionally, the selection of the patient groups and matching of AR and NR patients was carefully designed to avoid potential confounding factors such as demographics and treatment regimen and included clinical records of patient infection status, as described in the confounder analysis.

The suppression of lymphocyte-associated genes in the AR group approaching rejection may represent a state of immune dysfunction and might potentially be associated with uremic conditions.<sup>44,45</sup> This notion is supported by the significantly higher levels of creatinine in AR as well as the observation that a large number of chemokines as well as interleukins and their receptors, which promote the development and differentiation of T, B, and hematopoietic cells, are downregulated in AR patients. For example, the IL-7 (IL7R) pathway was one of the most highly downregulated interleukin pathways and a number of studies have demonstrated the role of IL-7 and its signal transducer, STAT5, in the development, differentiation, and survival of T cells.<sup>46–50</sup> The JAK-STAT pathway, a pleiotropic cascade used to transduce a multitude of signals for development and homeostasis in animals, is also downregulated in our data, and it has been shown that dysregulation of JAK-STAT signaling can result in immune deficiency conditions.<sup>51</sup> Overall, these molecular changes indicate

that, in patients approaching rejection, pathways that function in lymphocyte activities are downregulated (at least in relation to neutrophil activities [see below]), potentially representing immune dysfunction and deficiency.

The second category of transcriptional changes that we identified in AR patients was an increase in genes related to neutrophil activities. As shown in the gene enrichment analysis, most genes in this signature are highly enriched in neutrophils and indeed many are almost completely specific to neutrophils. A large proportion of neutrophil granule proteins, which represent neutrophil activation, chemotaxis, and degranulation, are included in this signature. Neutrophils are known to be key mediators in innate immunity and inflammation. The initial inflammatory response, represented by the upregulation of neutrophil enriched genes in our data, is present in all patients immediately after (first 4 days) transplant surgery. However, this signature remains high or is possibly reactivated in patients approaching rejection while appearing to be returning to a more normal range in nonrejection patients.

Neutrophils are known to be critical mediators of ischemic reperfusion injury after organ transplantation.<sup>52</sup> For example, syngeneic lung transplantations can stimulate the expansion of neutrophil progenitors, leading to the accumulation of neutrophils within the peripheral blood and graft tissues.<sup>53</sup> Furthermore, ischemia reperfusion injury prevents immunosuppression-mediated acceptance of mouse lung allografts unless granulopoiesis is inhibited, linking neutrophil accumulation and activities directly with tissue rejection.<sup>54</sup> Neutrophils mediate tissue damage by cytotoxic and proinflammatory cytokine production and can infiltrate the organ within hours after surgical trauma or ischemia has been established.<sup>55</sup> This early inflammation is due to the innate response to tissue injury independent of the adaptive immune systems and occurs before the T cell response.<sup>10,11,56</sup> In fact, it has previously been shown that innate immune cells are able to respond to the allograft even in the absence of T lymphocytes.<sup>10,57,58</sup> Based on these findings and our own observations, we hypothesize that inflammation arising due to surgery or ischemia reperfusion injury immediately following transplant surgery is returning to a normal range in stabilizing patients, whilst remaining high in patients who go on to reject the organ. These persistent or reactivated innate immune responses might be factors that act to prevent immunosuppression mediated graft acceptance, or alternatively are a consequence of other factors that initiate the rejection.

We have described an overall concordance of the gene expression signatures. The concerted response of these genes suggested a potential for the existence of expression control hubs mediated by specific transcription factors. Using over-represented transcription factor binding site analysis we found GABP, an ETS family transcription factor known as an essential regulator of IL7-IL7R $\alpha$  signals that is central to

T cell proliferation and development,<sup>59</sup> as the top candidate transcription factor that might regulate transcription of the downregulated lymphocyte genes. Similarly, transcription binding site analysis using the neutrophil signature genes indicated RelA/NF- $\kappa$ B as a transcription factor likely to contribute to the regulation of the expression of genes involved in inflammation.<sup>60,61</sup> In addition to well-known functions of toll-like receptors (TLRs) in the induction of the immune response under pathological conditions, modulated by transcription factors such as nuclear factor NF- $\kappa$ B leading to the production of proinflammatory cytokines, TLR activation is found to play a role in differentiation of hematopoietic stem cells as well as prevention of apoptosis in neutrophils.<sup>62</sup>

Finally, a high neutrophil to lymphocyte ratio (NLR) has been shown to be a potential predictor of poor outcomes of transplantation, as well as in various diseases.<sup>63–65</sup> In our data, NLR was overall moderately increased in patients approaching rejection, when compared with nonrejectors. Our transcriptionally defined rejection-associated molecular signatures correlate with the observed NLR. The presence of an upregulated neutrophil molecular signature and downregulated lymphocyte molecular signature in patients approaching rejection could potentially represent a “molecular NLR” that might have potential for use as a monitoring tool for posttransplant patients, as the molecular changes preceding and modulating such global changes in cell type populations in blood, as well as cell type-specific gene regulation, may be more sensitive and specific than a direct count of cell populations.

## Conclusions

In this study, we describe the use of whole blood transcriptomes for the identification and characterization of molecular signatures associated with kidney transplantation and acute allograft rejection. From this exploratory approach, we have identified an increased gene expression signature comprising genes involved in neutrophil activation that is indicative of systemic inflammation. We find that this signature persists in patients approaching rejection. Additionally, we have described a gene expression signature indicative of immune dysfunction, potentially arising due to conditions associated with the kidney malfunction or uremic conditions. Furthermore, we have demonstrated that circulating whole blood transcriptome profiling is representative of immune responses associated with transplant rejection and that this information might have potential utility for the identification of indicators that can predict acute kidney transplant rejection. The relative ease of obtaining circulating whole blood for the purpose of monitoring rejection makes this approach fundamentally appealing. Finally, characterizing the molecular basis of gene expression signatures associated with the biological and immunological processes that take place during kidney transplant allograft rejection and recovery may provide insight into the improvement of post-transplant maintenance regimens and targeted treatment options.



## List of abbreviations

NR: Non-Rejector, AR: Acute Rejector, NLR: Neutrophil to Lymphocyte count Ratio, NSMPs: Negative Strand Mapping Probes, PCA: Principal Component Analysis, MeV: MultiExperiment Viewer, PTM: Pavlidis Template Matching, SAM: Significance Analysis of Microarrays.

## Acknowledgements

The authors thank all the patients and healthy donors that provided the samples for this research. We acknowledge contributions of clinical research coordinators, nurses, and physicians who made this work possible, including Dr. David Landsberg at St. Paul's Hospital, Vancouver. We also appreciate Dr. Alice Mui, UBC for helpful discussions.

## Author contributions

HS carried out the data analysis and drafted the manuscript. OG and ZH participated in data analysis. JEW participated in the design of the study and coordination. RTN, RB, RM, and BMM conceived of the study and participated in its design. PAK and SJT conceived of the study, participated in the design of the study and in drafting the manuscript. NMI, GK provided reagents and materials for the study and performed experiments. All authors read and approved the final manuscript.

## DISCLOSURES AND ETHICS

As a requirement of publication the authors have provided signed confirmation of their compliance with ethical and legal obligations including but not limited to compliance with ICMJE authorship and competing interests guidelines, that the article is neither under consideration for publication nor published elsewhere, of their compliance with legal and ethical guidelines concerning human and animal research participants (if applicable), and that permission has been obtained for reproduction of any copyrighted material. This article was subject to blind, independent, expert peer review. The reviewers reported no competing interests.

## Supplementary Data

The following additional data are available with the online version of this paper.

**Additional file 1.** Patient demographic and clinical data for 24 ARs and matched 24 NRs.

**Additional file 2.** The 5619 probe sets, which survived filtering process.

**Additional file 3.** Comparison of 24 ARs at rejection and 24 matched NRs against healthy controls. **A.** PCA plot of 24 AR patient samples at the time of rejection, their 24 matched NR patient samples, and 20 healthy control subject samples. Control samples show tight clustering while patient samples are more dispersed. Limited overlap is observed between normal samples and both AR and NR samples. **B.** An expression plot of the genes that differentiates control samples from all NR and AR patient samples. These genes display an overall lower level of expression in both AR and NR patients and an increase in the variability of expression between patients when compared with control samples. **C.** PTM analysis using

a defined gene expression pattern that is the highest at the earliest time points posttransplant and lower at the later time points in both AR and NR. Genes following this pattern potentially represent a signature that is associated with the transplant surgery and ischemic injury.

**Additional file 4. A.** Patient demographic and clinical data for 8 late rejection ARs and their matched 8 NRs. **B.** Confounder analysis for 8 ARs and 8 NRs.

**Additional file 5.** SAM analysis using late rejector samples only—Genes significantly downregulated in ARs versus NRs at the time of rejection.

**Additional file 6.** SAM analysis using late rejector samples only—Genes significantly upregulated in ARs versus NRs at the time of rejection.

**Additional file 7.** InnateDB analysis using late rejector samples only (day 15 and beyond posttransplant)—overrepresented pathways in the downregulated genes in acute rejectors

**Additional file 8.** SAM analysis using time course samples collected minimum 15 days post-TX—Cluster of genes significantly upregulated and downregulated in ARs compared to NRs.

**Additional file 9.** Cell type-specific gene expression signatures replicate across different microarray platforms. **A-D.** PCA plots of neutrophil and lymphocyte signature genes (**A** and **B**) and 8 late AR and matched NR samples (**C** and **D**). Gene expression signatures (gene PCA plots **A** and **B**) show significant separation of AR and NR samples consistently over the 2 different platforms (sample PCA plots **C** and **D**).

**Additional file 10.** The 369 genes that covary with *DYSF* ( $R > 0.6$ ).

**Additional file 11.** Gene Ontology (GO) analysis. Overrepresented cellular component GO terms by the genes whose expression most tightly correlates with *DYSF*.

**Additional file 12.** Overrepresented transcription factor binding site analysis.

## REFERENCES

- Cohen DJ, St Martin L, Christensen LL, Bloom RD, Sung RS. Kidney and pancreas transplantation in the United States, 1995–2004. *Am J Transplant.* 2006;6:1153–69.
- Morris PJ. Transplantation—a medical miracle of the 20th century. *N Engl J Med.* 2004;351:2678–80.
- Sayegh MH, Carpenter CB. Transplantation 50 years later—progress, challenges, and promises. *N Engl J Med.* 2004;351:2761–6.
- Hall BM. Cells mediating allograft rejection. *Transplantation.* 1991;51:1141–51.
- Sayegh MH, Watschinger B, Carpenter CB. Mechanisms of T cell recognition of alloantigen. The role of peptides. *Transplantation.* 1994;57:1295–302.
- Deng MC, Eisen HJ, Mehra MR, et al. Noninvasive discrimination of rejection in cardiac allograft recipients using gene expression profiling. *Am J Transplant.* 2006;6:150–60.
- Li L, Khatri P, Sigdel TK, et al. A peripheral blood diagnostic test for acute rejection in renal transplantation. *Am J Transplant.* 2012;12:2710–8.
- Sarwal M, Chua MS, Kambham N, et al. Molecular heterogeneity in acute renal allograft rejection identified by DNA microarray profiling. *N Engl J Med.* 2003;349:125–38.
- Kitchens WH, Uehara S, Chase CM, Colvin RB, Russell PS, Madsen JC. The changing role of natural killer cells in solid organ rejection and tolerance. *Transplantation.* 2006;81:811–7.
- He H, Stone JR, Perkins DL. Analysis of robust innate immune response after transplantation in the absence of adaptive immunity. *Transplantation.* 2002;73:853–61.





11. He H, Stone JR, Perkins DL. Analysis of differential immune responses induced by innate and adaptive immunity following transplantation. *Immunology*. 2003;109:185–96.
12. Perico N, Cattaneo D, Sayegh MH, Remuzzi G. Delayed graft function in kidney transplantation. *Lancet*. 2004;364:1814–27.
13. Sarwal MM. Deconvoluting the ‘omics’ for organ transplantation. *Curr Opin Organ Transplant*. 2009;14:544–51.
14. Liew CC, Ma J, Tang HC, Zheng R, Dempsey AA. The peripheral blood transcriptome dynamically reflects system wide biology: a potential diagnostic tool. *J Lab Clin Med*. 2006;147:126–32.
15. Kurian S, Grigoryev Y, Head S, Campbell D, Mondala T, Salomon DR. Applying genomics to organ transplantation medicine in both discovery and validation of biomarkers. *Int Immunopharmacol*. 2007;7:1948–60.
16. Gunther OP, Balshaw RF, Scherer A, et al. Functional genomic analysis of peripheral blood during early acute renal allograft rejection. *Transplantation*. 2009;88:942–51.
17. Racusen LC, Solez K, Colvin RB, et al. The Banff 97 working classification of renal allograft pathology. *Kidney Int*. 1999;55:713–23.
18. Etminan M, Samii A. Pharmacoepidemiology I: a review of pharmacoepidemiology study designs. *Pharmacotherapy*. 2004;24:964–9.
19. Vardhanabhuti S, Blakemore SJ, Clark SM, Ghosh S, Stephens RJ, Rajagopalan D. A comparison of statistical tests for detecting differential expression using Affymetrix oligonucleotide microarrays. *OMICS*. 2006;10:555–66.
20. Wu Z, Irizarry RA, Gentleman R, Murillo FM, Spencer F. A Model-Based Background Adjustment for Oligonucleotide Expression Arrays. *J Amer Stat Assoc*. 2004;99:909–17.
21. Warren PTD, Martini PGV, Jackson J, Bienkowska J. PANP—a new method of gene detection on oligonucleotide expression arrays. In: Proceedings from the IEEE 7th International Symposium on Bioinformatics and Bioengineering; October 14–17, 2007; Boston, MA; 108–15.
22. Nurtdinov RN, Vasiliev MO, Ershova AS, Lossev IS, Karyagina AS. PLANDbAffy: probe-level annotation database for Affymetrix expression microarrays. *Nucleic Acids Res*. 2010;38:D726–30.
23. Saeed AI, Bhagabati NK, Braisted JC, et al. TM4 microarray software suite. *Methods Enzymol*. 2006;411:134–93.
24. Wold S, Esbensen K, Geladi P. Principal component analysis. *Chemometr Intell Lab Syst*. 1987;2:37–52.
25. Pavlidis P, Noble WS. Analysis of strain and regional variation in gene expression in mouse brain. *Genome Biol*. 2001;2(10):RESEARCH0042.
26. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A*. 2001;98:5116–21.
27. Lynn DJ, Winsor GL, Chan C, et al. InnateDB: facilitating systems-level analyses of the mammalian innate immune response. *Mol Syst Biol*. 2008;4:218.
28. Beissbarth T, Speed TP. GOstat: find statistically overrepresented Gene Ontologies within a group of genes. *Bioinformatics*. 2004;20:1464–5.
29. Ho Sui SJ, Mortimer JR, Arenillas DJ, et al. oPOSSUM: identification of over-represented transcription factor binding sites in co-expressed genes. *Nucleic Acids Res*. 2005;33:3154–64.
30. Mieczkowski J, Tyburczy ME, Dabrowski M, Pokarowski P. Probe set filtering increases correlation between Affymetrix GeneChip and qRT-PCR expression measurements. *BMC Bioinformatics*. 2010;11:104.
31. Dai M, Wang P, Boyd AD, et al. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res*. 2005;33:e175.
32. Su AI, Cooke MP, Ching KA, et al. Large-scale analysis of the human and mouse transcriptomes. *Proc Natl Acad Sci U S A*. 2002;99:4465–70.
33. de Fijter JW. The impact of age on rejection in kidney transplantation. *Drugs Aging*. 2005;22:433–49.
34. Emiroglu R, Yagmurdu MC, Karakayali F, et al. Role of donor age and acute rejection episodes on long-term graft survival in cadaveric kidney transplantations. *Transplant Proc*. 2005;37:2954–6.
35. Lominadze G, Powell DW, Luerman GC, Link AJ, Ward RA, McLeish KR. Proteomic analysis of human neutrophil granules. *Mol Cell Proteomics*. 2005;4:1503–21.
36. Benita Y, Cao Z, Giallourakis C, Li C, Gardet A, Xavier RJ. Gene enrichment profiles reveal T-cell development, differentiation, and lineage-specific transcription factors including ZBTB25 as a novel NF-AT repressor. *Blood*. 2010;115:5376–84.
37. Kupiec-Weglinski JW, Busuttill RW. Ischemia and reperfusion injury in liver transplantation. *Transplant Proc*. 2005;37:1653–6.
38. Zhang X, Kluger Y, Nakayama Y, et al. Gene expression in mature neutrophils: early responses to inflammatory stimuli. *J Leukoc Biol*. 2004;75:358–72.
39. Borregaard N, Sorensen OE, Theilgaard-Monch K. Neutrophil granules: a library of innate immunity proteins. *Trends Immunol*. 2007;28:340–5.
40. Faurschou M, Borregaard N. Neutrophil granules and secretory vesicles in inflammation. *Microbes Infect*. 2003;5:1317–27.
41. Shannon CP, Hollander Z, Wilson-McManus J, et al. White blood cell differentials enrich whole blood expression data in the context of acute cardiac allograft rejection. *Bioinform Biol Insights*. 2012;6:49–61.
42. Grigoryev YA, Kurian SM, Avnur Z, et al. Deconvoluting post-transplant immunity: cell subset-specific mapping reveals pathways for activation and expansion of memory T, monocytes and B cells. *PLoS One*. 2010;5:e13358.
43. Shen-Orr SS, Tibshirani R, Khatri P, et al. Cell type-specific gene expression differences in complex tissues. *Nat Methods*. 2010;7:287–9.
44. Hauser AB, Stinghen AE, Kato S, et al. Characteristics and causes of immune dysfunction related to uremia and dialysis. *Perit Dial Int*. 2008;28(suppl 3):S183–7.
45. Kato S, Chmielewski M, Honda H, et al. Aspects of immune dysfunction in end-stage renal disease. *Clin J Am Soc Nephrol*. 2008;3:1526–33.
46. Yao Z, Cui Y, Watford WT, et al. Stat5a/b are essential for normal lymphoid development and differentiation. *Proc Natl Acad Sci U S A*. 2006;103:1000–5.
47. Tripathi P, Kurtulus S, Wojciechowski S, et al. STAT5 is critical to maintain effector CD8+ T cell responses. *J Immunol*. 2010;185:2116–24.
48. Rochman Y, Spolski R, Leonard WJ. New insights into the regulation of T cells by gamma(c) family cytokines. *Nat Rev Immunol*. 2009;9:480–90.
49. Palmer MJ, Mahajan VS, Trajman LC, et al. Interleukin-7 receptor signaling network: an integrated systems perspective. *Cell Mol Immunol*. 2008;5:79–89.
50. Hand TW, Cui W, Jung YW, et al. Differential effects of STAT5 and PI3K/AKT signaling on effector and memory CD8 T-cell survival. *Proc Natl Acad Sci U S A*. 2010;107:16601–6.
51. Aaronson DS, Horvath CM. A road map for those who don't know JAK-STAT. *Science*. 2002;296:1653–5.
52. de Perrot M, Liu M, Waddell TK, Keshavjee S. Ischemia-reperfusion-induced lung injury. *Am J Respir Crit Care Med*. 2003;167:490–511.
53. Kreisel D, Sugimoto S, Tietjens J, et al. Bcl3 prevents acute inflammatory lung injury in mice by restraining emergency granulopoiesis. *J Clin Invest*. 2011;121:265–76.
54. Kreisel D, Sugimoto S, Zhu J, et al. Emergency granulopoiesis promotes neutrophil-dendritic cell encounters that prevent mouse lung allograft acceptance. *Blood*. 2011;118:6172–82.
55. Jaeschke H, Farhood A, Smith CW. Neutrophils contribute to ischemia/reperfusion injury in rat liver in vivo. *FASEB J*. 1990;4:3355–9.
56. Christopher K, Mueller TF, Ma C, Liang Y, Perkins DL. Analysis of the innate and adaptive phases of allograft rejection by cluster analysis of transcriptional profiles. *J Immunol*. 2002;169:522–30.
57. Kirk AD, Hale DA, Mannon RB, et al. Results from a human renal allograft tolerance trial evaluating the humanized CD52-specific monoclonal antibody alemtuzumab (CAMPATH-1H). *Transplantation*. 2003;76:120–9.
58. Wu T, Bond G, Martin D, Nalesnik MA, Demetris AJ, Abu-Elmagd K. Histopathologic characteristics of human intestine allograft acute rejection in patients pretreated with thymoglobulin or alemtuzumab. *Am J Gastroenterol*. 2006;101:1617–24.
59. Xue HH, Bollenbacher J, Rovella V, et al. GA binding protein regulates interleukin 7 receptor alpha-chain gene expression in T cells. *Nat Immunol*. 2004;5:1036–44.
60. McDonald PP, Cassatella MA. Activation of transcription factor NF-kappa B by phagocytic stimuli in human neutrophils. *FEBS Lett*. 1997;412:583–6.
61. Miskolci V, Rollins J, Vu HY, Ghosh CC, Davidson D, Vancurova I. NFkappaB is persistently activated in continuously stimulated human neutrophils. *Mol Med*. 2007;13:134–42.
62. McGettrick AF, O'Neill LA. Toll-like receptors: key activators of leucocytes and regulator of haematopoiesis. *Br J Haematol*. 2007;139:185–93.
63. Imtiaz F, Shafique K, Mirza SS, Ayoob Z, Vart P, Rao S. Neutrophil lymphocyte ratio as a measure of systemic inflammation in prevalent chronic diseases in Asian population. *Int Arch Med*. 2012;5:2.
64. Azab B, Daoud J, Naeem FB, et al. Neutrophil-to-lymphocyte ratio as a predictor of worsening renal function in diabetic patients (3-year follow-up study). *Ren Fail*. 2012;34:571–6.
65. Zahorec R. Ratio of neutrophil to lymphocyte counts—rapid and simple parameter of systemic inflammation and stress in critically ill. *Bratisl Lek Listy*. 2001;102:5–14.