Establishing long-term allograft acceptance without the requirement for continuous immunosuppression, a condition known as allograft tolerance, is a highly desirable therapeutic goal in solid organ transplantation. Determining which recipients would benefit from withdrawal or minimization of immunosuppression would be greatly facilitated by biomarkers predictive of tolerance. In this study, we identified the largest reported cohort to our knowledge of tolerant renal transplant recipients, as defined by stable graft function and receiving no immunosuppression for more than 1 year, and compared their gene expression profiles and peripheral blood lymphocyte subsets with those of subjects with stable graft function who are receiving immunosuppressive drugs as well as healthy controls. In addition to being associated with clinical and phenotypic parameters, renal allograft tolerance was strongly associated with a B cell signature using several assays. Tolerant subjects showed increased expression of multiple B cell differentiation genes, and a set of just 3 of these genes distinguished tolerant from nontolerant recipients in a unique test set of samples. This B cell signature was associated with upregulation of \textit{CD20} mRNA in urine sediment cells and elevated numbers of peripheral blood naive and transitional B cells in tolerant participants compared with those receiving immunosuppression. These results point to a critical role for B cells in regulating alloimmunity and provide a candidate set of genes for wider-scale screening […]
Identification of a B cell signature associated with renal transplant tolerance in humans

Kenneth A. Newell,1 Adam Asare,2,3 Allan D. Kirk,1 Trang D. Gisler,2,3 Kasia Bourcier,2,3 Manikkam Suthanthiran,4 William J. Burlingham,5 William H. Marks,6 Ignacio Sanz,7 Robert I. Lechler,8,9 Maria P. Hernandez-Fuentes,8,9 Laurence A. Turka,3,10 and Vicki L. Seyfert-Margolis,3,11 for the Immune Tolerance Network ST507 Study Group

1Emory University, Atlanta, Georgia, USA. 2University of California, San Francisco, California, USA. 3Immune Tolerance Network, Bethesda, Maryland, USA (www.immunetolerance.org). 4Cornell University Medical Center, New York, New York, USA. 5University of Wisconsin, Madison, Wisconsin, USA. 6Swedish Medical Center, Seattle, Washington, USA. 7University of Rochester, Rochester, New York, USA. 8MRC Centre for Transplantation, King’s College, London, United Kingdom. 9Indices of Tolerance EU consortium (www.transplant-tolerance.org.uk). 10Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts, USA. 11Food and Drug Administration, Silver Spring, Maryland, USA.

Establishing long-term allograft acceptance without the requirement for continuous immunosuppression, a condition known as allograft tolerance, is a highly desirable therapeutic goal in solid organ transplantation. Determining which recipients would benefit from withdrawal or minimization of immunosuppression would be greatly facilitated by biomarkers predictive of tolerance. In this study, we identified the largest reported cohort to our knowledge of tolerant renal transplant recipients, as defined by stable graft function and receiving no immunosuppression for more than 1 year, and compared their gene expression profiles and peripheral blood lymphocyte subsets with those of subjects with stable graft function who are receiving immunosuppressive drugs as well as healthy controls. In addition to being associated with clinical and phenotypic parameters, renal allograft tolerance was strongly associated with a B cell signature using several assays. Tolerant subjects showed increased expression of multiple B cell differentiation genes, and a set of just 3 of these genes distinguished tolerant from nontolerant recipients in a unique test set of samples. This B cell signature was associated with upregulation of CD20 mRNA in urine sediment cells and elevated numbers of peripheral blood naive and transitional B cells in tolerant participants compared with those receiving immunosuppression. These results point to a critical role for B cells in regulating alloimmunity and provide a candidate set of genes for wider-scale screening of renal transplant recipients.

Introduction

Advances in immunosuppression over 2 decades have led to vast improvements in both control of acute rejection and short-term graft survival in renal transplantation. However, similar improvements in long-term outcomes have not yet been achieved, and concerns over the morbidity of lifelong regimens of immunosuppressive drugs remain (1, 2). Establishing long-term allograft acceptance without the requirement for continuous immunosuppression, a condition known as allograft tolerance, is therefore a highly desirable therapeutic goal in kidney transplantation (3–5).

Unlike liver transplantation, where it is estimated that up to 20% of recipients may be withdrawn from all immunosuppression (6–12), tolerance to renal allografts appears to be far less frequent (13–15). Although tolerance has been achieved in numerous animal models of renal transplantation (16), attempts to induce long-term allograft tolerance in humans have been far less successful, and may be complicated by the potential loss of the engrafted kidney during immunosuppression minimization or withdrawal.

Several recent studies have attempted to identify biomarkers of tolerance in liver and kidney transplantation (15, 17–20). In liver transplantation, the proportion of γδ T cells (specifically TCR γδ cells; ref. 19) and the ratio of plasmacytoid to myeloid dendritic cells (20) and B cells (18) were shown to be increased in tolerant liver transplant recipients relative to those stable on immunosuppression (9, 19, 21). Additionally, specific patterns of expressed genes were shown to be associated with tolerant participants compared with those stable on immunosuppression and with healthy controls (19, 21). Similarly, 2 studies have shown that tolerant kidney transplant recipients have distinct patterns of expressed genes and T cell receptor gene use (13, 15).

In this study, we recruited the largest reported cohort to our knowledge of tolerant kidney transplant recipients (n = 25), 20 of whom ceased taking immunosuppression as a result of medication nonadherence. We sought to identify immune parameters that would discriminate tolerant individuals from subjects with stable allograft function while on immunosuppression, as well as healthy (nontransplanted) controls. We found that tolerant patients exhibited increased numbers of total and naïve B cells and had enhanced expression of B cell differentiation and activation genes compared with subjects receiving immunosuppression. Most notably, the tolerant cohort differentially expressed 3 B cell genes that were highly predictive of tolerance in a new test set of patients. These markers are strong candidates for clinical testing as a means to predict kidney transplant recipients who may benefit from minimization or withdrawal of immunosuppression, and for monitoring their status during immunosuppression withdrawal.

Results

Study population clinical characteristics. We enrolled 3 groups of participants into the study: those operationally tolerant, who had stable graft function despite receiving no immunosuppression for
at least 1 year (TOL, \(n = 25\)); those with stable graft function while on immunosuppressive medications (SI, \(n = 33\)); and healthy (nontransplanted) control subjects (HC, \(n = 42\)). Participants in the TOL and SI groups had excellent renal function, even though TOL participants had ceased taking immunosuppressive medications for at least 1 year. Ages, genders, and primary diseases leading to renal failure were similar between the TOL and SI groups (Table 1). There was no statistical difference in the frequency of humoral sensitization in these 2 groups, as 8 of 25 TOL and 15 of 33 SI patients had detectable antibodies directed against HLA molecules, whereas 4 of the 31 evaluable patients in the SI group had donor-directed anti-HLA antibodies. The differences between the groups were not statistically significant.

Participants in the TOL group were more closely matched for HLA than the SI group (\(P < 0.05\); Table 1). Whereas the majority of our TOL group was HLA matched, 5 participants were not matched for any HLA antigens. Time from transplant to enrollment was found to be significantly higher in the TOL group (\(P < 0.05\)), while the number of patients with cadaveric donors was significantly greater in the SI group (\(P < 0.05\)). There were no other significant differences in clinical assessments between these groups in this study. Several assays were performed on peripheral blood, including immunophenotyping and gene expression profiling by microarray and multiplex real-time PCR (Supplemental Table 1; supplemental material available online with this article; doi:10.1172/JCI39933DS1).

B cell gene signatures distinguish TOL from SI participants. For the purposes of predictive modeling using gene expression studies, patients in the TOL and SI groups were subdivided, based on study enrollment, into training sets (TOL-TRN, \(n = 19\); SI-TRN, \(n = 27\)) and test sets (TOL-TST, \(n = 6\); SI-TST, \(n = 6\)). All test set patients were identified, and samples collected, subsequent to the identification and study of the training set patients and therefore are an independent cohort.

Microarrays were used to detect expressed gene profiles of whole-blood total RNA from subjects in the TOL-TRN and SI-TRN groups. RNA samples were selected, and cRNA was made and hybridized in 5 batches. The strategy used for normalizing batch effects involved building batch as a covariant into the statistical models used in our analyses. Statistical analysis for differentially expressed genes between these 2 groups was performed, and the top differentially expressed genes were ranked in order, based on their mean fold change differences.

<table>
<thead>
<tr>
<th>Donor age, yr</th>
<th>TOL-TRN ((n = 19))</th>
<th>51 (27–77)</th>
<th>TOL-TST ((n = 6))</th>
<th>56.5 (47–66)</th>
<th>SI ((n = 33))</th>
<th>52.5 (31–69)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Donor type, (n)</td>
<td>Cadaveric</td>
<td>4</td>
<td>8</td>
<td>14</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>Living-related</td>
<td>14</td>
<td>3</td>
<td>1</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living-unrelated</td>
<td>–</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data missing</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender, (n)</td>
<td>Female</td>
<td>7</td>
<td>3</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>12</td>
<td>3</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race, (n)</td>
<td>Asian</td>
<td>1</td>
<td>–</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>–</td>
<td>–</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>18</td>
<td>6</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recipient age, yr</td>
<td>51 (27–77)</td>
<td>51 (42–63)</td>
<td>44 (26–75)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary cause of renal failure, % ((n))</td>
<td>Cystic/polycystic kidney disease</td>
<td>5.3% (1)</td>
<td>16.7% (1)</td>
<td>9.1% (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>5.3% (1)</td>
<td>16.7% (1)</td>
<td>3% (1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glomerulonephritis</td>
<td>36.8% (7)</td>
<td>33% (2)</td>
<td>24.2% (8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypertension</td>
<td>–</td>
<td>–</td>
<td>12.1% (4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IgA nephropathy</td>
<td>5.3% (1)</td>
<td>16.7% (1)</td>
<td>12.1% (4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obstructive/reflux nephropathy</td>
<td>15.8% (3)</td>
<td>–</td>
<td>6.1% (2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poststreptococcal glomerular nephritis</td>
<td>10.5% (2)</td>
<td>–</td>
<td>6.1% (2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyelonephritis/interstitial nephritis</td>
<td>5.3% (1)</td>
<td>–</td>
<td>6.1% (1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systemic lupus erythematosus</td>
<td>–</td>
<td>–</td>
<td>9.1% (3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data missing</td>
<td>15.8% (3)</td>
<td>16.7% (1)</td>
<td>15.2% (5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creatinine level, mg/dl</td>
<td>1 (0.6–9.8)</td>
<td>0.95 (0.7–1.5)</td>
<td>1.4 (0.7–2.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interval between transplant and enrollment, yr</td>
<td>20.42 (7–40.25)</td>
<td>11.79 (5–19.67)</td>
<td>5.5 (1–41.17)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*HLA mismatch</td>
<td>0.83 ± 1.59 of 6</td>
<td>0.67 ± 1.15 of 6</td>
<td>3.60 ± 1.82 of 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time of immunosuppression cessation, yr</td>
<td>13 (1–52)</td>
<td>5.5 (2–22)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reason for cessation, % ((n))</td>
<td>Medical condition</td>
<td>10.5% (2)</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noncompliance</td>
<td>73.7% (14)</td>
<td>100% (6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data missing</td>
<td>15.8% (3)</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We found that 5 unique genes — \(\text{TUBB2A, TCL1A, BRDG1, HTPAP, and PPPAPDC1B}\) — reached statistical significance after a false discovery rate (FDR) correction was applied to the data. Of the 30 genes found to have a 2-fold increase in expression in the TOL-TRN versus SI-TRN group (Figure 1), 22 were B cell specific (as defined in the OMIM database; http://www.ncbi.nlm.nih.gov/omim). Many of these genes are involved in B cell activation and differentiation, including genes encoding Ig heavy, light, and joining chains and HLA (22, 23). In contrast, few genes were observed to be more highly expressed in the SI-TRN group. One of these, \(\text{TUBB2A}\), was highly differentially expressed (7-fold difference). Overexpression of this gene may be due to prolonged use of calcineurin inhibitors, which have previously been shown to induce tubulin expression (24). Comparison of whole-blood total RNA from TOL-TRN and HC groups revealed no significant differences in expression levels. The similarity in profiles between TOL-TRN and HC groups was illustrated using hierarchical clustering of the 30 genes that were more than 2-fold differentially expressed between TOL-TRN and SI-TRN (Figure 1).

Further evidence for B cell involvement in tolerance was provided by gene expression profiles of urinary cell sediments. Of
the 18 genes tested by quantitative real-time PCR, only the CD20 transcript was significantly higher in TOL-TRN than SI-TRN participants (Figure 2). We performed the same analyses on urine sedimentary cells from HC and found significantly lower expression of CD20 and FOXP3 in HC relative to TOL-TRN (Figure 2, A and B). We also found significantly lower expression of CD3 and Perforin in HC compared with TOL-TRN (Figure 2, C and D). Collectively, these findings are consistent with healthy subjects having few or no lymphocytes in their urine; indeed, urine from 7 of the 25 HC tested did not have sufficient RNA because of their low or no numbers of urine sedimentary cells. These samples were excluded from our analysis.

Multiplex real-time PCR confirms a B cell gene signature and identifies 3 genes that predict tolerance in a test set of patients. Multiplex real-time PCR was performed on all participants to develop a more quantitative approach for defining tolerance-specific expressed gene profiles and to support our microarray findings. Probe-primer sets for 228 genes were made for the multiplex real-time PCR, which was run on all participants in a single batch (Supplemental Table 2). These genes were selected from the following lists: (a) the top 50 differentially expressed genes from our microarray analyses; (b) genes defined as tolerant specific in liver transplant recipients, as reported by Martinez-Llordella et al. (21); (c) genes found to be family members of our differentially expressed genes identified by microarray; and (d) genes believed to play a role in immunological tolerance as reported in the literature, such as CD40.

Statistical analysis of the multiplex real-time PCR data revealed 31 unique genes that were different between the TOL-TRN and SI-TRN groups (P < 0.05; Figure 3), whereas no genes were significantly different between TOL and HC (Supplemental Table 3). Of those found to be differentially expressed in TOL-TRN versus SI-TRN, 17 were originally detected by microarray. None of the genes found to be involved in liver transplant tolerance by Martinez-Llordella et al. (21) were differentially expressed between our TOL and SI cohorts. Again, most of the genes overexpressed in the TOL cohort were B cell specific (26 of 30), with many encoding κ/λ light chains of Ig.
cross-validation (LOOCV), the LDA model found 3 genes to be most predictive, giving a positive predictive value (PPV) of 83% and a negative predictive value (NPV) of 84% (Table 2). To directly test the predictive value of these 3 genes, we used a separate test set cohort of 6 TOL and 6 SI patients (receiving standard triple-drug immunosuppression), termed TOL-TST and SI-TST, respectively. All 6 TOL-TST participants were correctly identified as TOL, and 5 of 6 SI-TST participants were correctly identified as SI, yielding PPV and NPV of 86% and 100%, respectively (Table 3).

The genes used in this predictive model are IGKV4-1, IGLLI, and IGKVID-15; their expression levels for each patient are shown in Figure 4. These genes clearly separated TOL from SI patients for the majority of participants in the training set and for all but 1 SI participant in the test set (Figure 4). Box plots of mRNA copy numbers for each of these genes are shown in Figure 5. All 3 of these genes encode κ or λ light chains, which are upregulated during the transition from pre- to mature B cells and during class switch and receptor editing that occurs after stimulation of mature B cells with antigen.

Flow cytometry of whole blood revealed increased numbers of naive B cells in TOL participants. We next performed immunophenotyping on whole blood samples shipped overnight from the collection site to a centralized flow cytometry facility. Subsets for analysis were selected from the 32 possible combinations of each 5-color panel (i.e., 2⁵). Each unique combination based on the selected region combinations used (approximately 15 per panel) was treated as independent variables for the purposes of statistical analysis of the flow data in total. A multiple test correction was applied to the data, and those populations determined to be significantly different were evident between TOL and HC and between TOL and SI were defined. A significant difference was found between the TOL and SI groups in the number of total B cells (CD19⁺) and naive B cells (CD19⁺CD27⁺IgM⁺IgD⁺; Figure 6, A and B). The mean B cell number of the TOL cohort was significantly greater than in the SI group (P < 0.01), but did not reach significance compared with the HC cohort (P = 0.10). However, there were significant differences between the TOL cohort and both the SI and HC group with respect to naive B cell numbers (P ≤ 0.05, Figure 6B). In addition, CD86⁺CD19⁺ B cells and memory B cells (CD19⁺CD27⁺IgM⁺IgD⁺) were significantly different between the TOL and HC cohorts (P < 0.01 and P = 0.03, respectively, Figure 6, C and D).

Many subpopulations of lymphocytes were measured in our analyses (Supplemental Table 4), and we did find other lymphocyte subpopulations that were significantly different between the TOL and SI or TOL and HC groups. These included HLA-DR⁺CD4⁺ T cells and NK cells, among others (data not shown). Of those that were significantly different, we highlight here the B cells and B cell subsets, as these differences correlated with our findings using microarrays and PCR.

Statistical analyses of total WBC count means for each group showed no differences, indicating that the observed group-specific differences were not caused by overall changes in the number of total WBCs. Mean WBC counts for the TOL, SI, and HC groups were 6.7 × 10⁹, 8.1 × 10⁹, and 5.8 × 10⁹ cells/μL, respectively. Rather, there appeared to be a redistribution of B cell subsets in the TOL group relative to the others, in which the IgM⁺IgD⁻ cells made up a greater proportion of the B cells in TOL participants than in the other groups. Furthermore, the observation that selected populations of B cells distinguished TOL kidney transplant recipients from HC subjects suggests that the B cell signature associated with tolerance is not simply a consequence of immunosuppression.

As the flow cytometric studies above analyzed all of our patients, we sought to validate the phenotypic B cell signature in an independent set of kidney transplant recipient patients. For this purpose, we reexamined our subjects’ frozen peripheral blood mononuclear cells (PBMCs) and compared them with frozen PBMCs collected from participants enrolled in the Indices of Tolerance (IOT) study of tolerant renal transplant recipients (25). These studies were performed with additional surface markers to be able to differentiate between naive and transitional cells and discriminate transitional subsets (26). Frozen PBMCs from our patients and...
3 of the IOT cohorts (TOL; SI, specifically calcineurin inhibitors; and HC) were studied by flow cytometry. The percentage of total B cells and naive B cells in the lymphocyte gate was higher in each TOL group compared with the respective SI group in both our Immune Tolerance Network (ITN) samples and the IOT samples (Figure 7, A and B). We noted lower numbers of naive B cells in all groups of samples provided by the IOT collaborative relative to our samples. Differences in freezing methods may have caused these shifts. Importantly, the patients in the IOT cohort were not as well HLA-matched with their donors as were those in our TOL cohort, which suggests that the increased B cell numbers were not a direct result of HLA matching.

Next we examined transitional B cells (CD19+CD38+CD24+IgD+) in these samples. These cells were of interest because of our gene signature and the regulatory role that has been proposed in murine models for cells with a similar immature phenotype (27). In both IOT and ITN samples, there were increased numbers of transitional B cells in the TOL versus SI group comparisons (Figure 7C).

Additionally, predictive modeling of the frozen flow assays was performed, in which we treated the ITN frozen flow results as a training set and the IOT flow results as an independent test set. In these analyses we also showed transitional B cells to be the most predictive, with a PPV of 85% and NPV of 96% in training using Table 2.

<table>
<thead>
<tr>
<th>Multiplex RT-PCR, ITN training set</th>
<th>Actual TOL</th>
<th>Actual SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted TOL</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>Predicted SI</td>
<td>4</td>
<td>21</td>
</tr>
</tbody>
</table>

Confusion matrix and classification summary showing 3-gene prediction model of TOL-TRN samples using LDA (see Methods). LDA identified 3 genes that predict tolerance (IGKV1D-13, IGKV4-1, and IGLL1) using the TOL-TRN versus SI-TRN multiplex RT-PCR data in the model (PPV, 83%; NPV, 84%).
our samples and a PPV of 63% and NPV of 90% (sensitivity, 83%; specificity, 75%) in the European samples that served as the test set (Tables 4 and 5).

Elevated expression of intracellular IL-10 in TOL patients. Finally, because we observed increases in B cell numbers and B cell subpopulations in TOL and HC patients, we sought to determine whether these cells produce immunomodulatory cytokines, such as IL-10 or TGFβ, that might be implicated in tolerance (27, 28). Thus, intracellular cytokine staining was performed on frozen PBMCs from 21 TOL, 32 SI, and 13 HC patients. The cells were studied both unstimulated and cultured in the presence of PMA plus ionomycin for 5 hours. Total B cells and 6 subsets, including T1 and T2 transitional B cells (CD38–CD24+), switched memory (CD27+IgD–), CD27+ memory (CD27+IgD–), naive and T3 transitional (CD27–IgD–), unswitched memory (CD27–IgD–), and plasmablasts (CD27–CD38–IgD–), were analyzed for IL-10 and TGFβ intracellular staining.

We observed a statistically significant increase in T1 plus T2 transitional B cells expressing IL-10 in the TOL group relative to the SI or HC groups (Figure 8A). Such a result must be interpreted with caution, however, as the overall percentages were extremely low, with many samples containing no IL-10–producing cells, and with a large overlap in the range of the groups. This may be the result of using stimulation conditions that were limited in length and intensity in order to maintain the cell surface phenotype and thus determine IL-10 production within different B cell subpopulations. No differences in the number of TGFβ–expressing B cells or subsets between cells from TOL or any other group were found (Figure 8B).

Discussion

This report, along with the companion study (25), is thus far the largest study to our knowledge of tolerant renal transplant recipients, both in terms of the number of participants studied and the variety of biological parameters analyzed. Most significantly, not only did our study identify what we believe to be new biomarkers of tolerance, it confirmed these markers across 3 independent cohorts of kidney transplant recipients, with varying degrees of HLA matching and differing patient demographics and geography. That said, many clinical factors associated with tolerance in this study are also known to identify low-risk kidney transplant recipient populations. Clinically, our TOL cohort received well-matched kidneys from living donors, and only 1 of the evaluable patients had humoral sensitization to the donor; conversely, 4 participants of our SI group had donor-specific antibodies.

This observation contrasts with that of Roussey-Kesler et al. (14), for whom the majority of tolerant kidney transplant recipients received allografts from deceased donors that were not well matched. Of their 10 participants, 5 experienced at least 1 episode of acute rejection, and 3 developed anti-donor HLA antibodies. Whereas the main reason for discontinuing immunosuppression in that study was life-threatening infections or malignancies, in the present study, discontinuation of immunosuppression in 20 of 25 participants was due to medication nonadherence, with the other 5 ceasing medication under medical supervision for complications associated with immunosuppression. Therefore, the results reported here may be more relevant to the goal of identification of clinically stable participants that may benefit from immunosuppressive drug withdrawal.

The expressed gene signature identified in this study distinguished TOL from SI participants and primarily consisted of B cell–specific genes, in particular genes involved in B cell differentiation. The association of B cell biomarkers and the tolerant phenotype was confirmed by flow cytometric analyses, which distinguished TOL not only from SI, but also from HC. These results suggest that this fingerprint may not be simply a consequence of differences induced by immunosuppression.

Table 3
Multiplex RT-PCR, ITN test set

<table>
<thead>
<tr>
<th>Actual TOL</th>
<th>Actual SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted TOL</td>
<td>6</td>
</tr>
<tr>
<td>Predicted SI</td>
<td>0</td>
</tr>
</tbody>
</table>

Confusion matrix and classification summary showing 3-gene prediction model of TOL-TST samples using LDA (see Methods). LDA identified 3 genes that predict tolerance (IGKV1D-13, IGKV4-1, and IGLL1) using the TOL-TST versus SI-TST multiplex RT-PCR data in the model (PPV, 86%; NPV, 100%).

Figure 4
Transcripts that best distinguish TOL from SI. Multiplex RT-PCR gene expression levels of IGKV1D-13, IGKV4-1, and IGLL1 for the 19 TOL-TRN and 24 SI-TRN samples, and for the 6 TOL-TST and 6 SI-TST samples, in log2 normalized number of molecules.
We were able to narrow this signature down to 3 genes that predicted tolerance with 100% accuracy in our test set of patients. Because we found that this signature of 3 genes was highly predictive for tolerance, a simple PCR assay may prove to be an easy test for screening kidney transplant patients that may benefit from weaning immunosuppression. Moreover, these gene expression analyses, combined with flow cytometry, may help to identify drug treatment regimens that promote tolerance based on an appearance of this protolerance signature. Importantly, these 3 genes are all expressed during the differentiation of B cells from pre- to mature B cells or during B cell activation-induced transition. This suggests that transitioning or maturing B cells are involved in tolerance induction and/or maintenance; alternatively, these cells may be suppressed by immunosuppression. Studies have shown that mature B cells in transition after antigen stimulation undergo receptor editing as means of promoting tolerance to self antigens (29, 30). Although receptor editing is believed to take place centrally in the bone marrow, it is possible that this mechanism is used peripherally in transplant recipients, as also shown in autoimmune models (31).

Our B cell gene signature correlates in part with findings of Brouard et al. (15), who, in cDNA microarray studies, found that transcripts of CD20, a trans-membrane protein expressed on pre- and mature B cells, were differentially expressed in tolerant

Figure 5
Box plots of log2 normalized mRNA copy numbers for the 3 genes found to be the best classifiers among the 31 identified as differentially expressed. (A) IGKV1D-13. (B) IGKV4-1. (C) IGLL1. Genes were derived from LDA, where they were found to have the best predictive value for TOL (Tables 2 and 3). Boxes depict IQR; whiskers denote 1.5 x IQR; values beyond this range are shown as outliers. P values are shown for statistically significant differences.

Figure 6
5-color flow cytometry of whole blood samples shows different numbers of B cell subsets. (A) Total B cells, as defined by CD19+ cells. The total B cell counts for TOL, SI, and HC cohorts were 287, 120, and 176 cells/μl, respectively. (B) Naive B cells, as defined by CD19+CD27-IgM+IgD+ cells. The naive B cell counts for TOL, SI, and HC cohorts were 190, 61, and 90 cells/μl, respectively. (C) Memory B cells, as defined by CD27-IgM+IgD+. The mean numbers for the memory B cell subset were 54.2 and 20.8 cells/μl for TOL and HC, respectively. (D) CD86+ B cells, as defined by CD86+CD19+. Mean numbers of CD86+CD19+ B cells for TOL and HC cohorts were 22 and 4.5 cells/μl, respectively. All values are shown as log2 absolute numbers, a calculation of the percent of lymphocyte gate multiplied by the total wbc count obtained from the same sample on a Coulter Counter. Boxes depict IQR; whiskers denote 1.5 x IQR; values beyond this range are shown as outliers. P values are shown for statistically significant differences.
kidney transplant recipients compared with other cohorts. When we mapped genes associated with tolerance from their study to the microarray and assessed their expression in our subjects, only CD20 was found to be increased in our TOL group. This was confirmed in our multiplex real-time PCR assay. None of the other tolerance-associated genes from that study were differentially expressed in our TOL participants compared with other groups in our study. However, it should be noted that the study by Brouard et al. compared the tolerant patient group to a group described as nontolerant. This group was composed of both patients doing well on immunosuppression and those with established chronic allograft nephropathy and/or interstitial fibrosis and tubular atrophy without evidence of any specific etiology (IFTA) and thus is not directly comparable to our comparison group composed only of SI patients.

Our choice to limit the comparison group to stable patients doing well without chronic allograft nephropathy/IFTA was based on the concept that this population is the most clinically relevant, as these individuals would be considered candidates for immunosuppressive minimization, whereas those with IFTA would not. The genes found in our signature also differed from those reported by Martinez-Llordella et al. (21); however, it should be noted that they examined liver transplant recipients, a group in which tolerance is recognized as being more common and for which the mechanisms of tolerance induction and maintenance are likely to be different.

Comparisons of microarray data from our study and the IOT study profile (25) revealed increased CD20 and TCL1A expression in the TOL groups of both studies. TCL1A differences are noteworthy because expression of this oncoprotein is highest in immature cells but is low or absent in mature B cells; therefore, its overexpression in the TOL group appears to reflect the abundance of transitional cells (32). Similarly, the TOL group displayed increased expression of VH4-34, which is also preferentially expressed in transitional and naive cells compared with memory cells (33). Our 3 most predictive genes, IGKV4-1, IGLLA, and IGKV1D-13, were not part of the IOT or Brouard gene list (15), most likely because their analyses were performed on Agilent and cDNA microarrays, which have a much more restricted set of genes, and because of differences in RNA purification methods, which have been shown to affect expressed gene profiles (34).

Our observation of increased total B cell numbers and naive B cells in the peripheral blood of TOL suggests that these cells may be important regulators of the antidonor immune response. Naive B cells are the major population of cells in the peripheral B cell pool and are known to be poor antigen-presenting cells for naive T cells (35, 36). Antigens presented by naive B cells have been shown to render animals tolerant (37) and were recently shown to stimulate naive T cells toward development into regulatory as opposed to effector T cells (38). Another B cell subset — unswitched memory cells — that we found to be elevated in TOL patients has also been reported to promote the production of the regulatory cytokine IL-10 (39).

However, functional studies performed on these B cell subsets did not unequivocally demonstrate a cytokine-mediated regulatory mechanism. We found no evidence of increased TGFβ-expressing B cells or B cell subpopulations in the TOL group relative to the SI or HC groups. In the case of IL-10, as noted previously, the results, while statistically significant, must be interpreted cautiously. Further studies will be necessary to determine whether these cells are truly associated with tolerance after renal transplantation, and, if so, whether they play a causative role and by what mechanism they do so. A potential role for B cells in transplantation tolerance is

Figure 7

9-color flow cytometry of frozen PBMCs from both ITN and European IOT samples. (A) Total B cells, expressed as the percent CD19+ cells in the lymphocyte gate. (B) Naive B cells, defined as CD19+CD27-IgD+, shown as the percent of the total CD19+ gate. (C) Transitional B cells, defined as CD19+CD38+CD24-IgD+, shown as the percent of the total CD19+ gate. Boxes depict IQR; whiskers denote 1.5 × IQR; values beyond this range are shown as outliers. P values are shown for statistically significant differences.
Renal allograft recipients were enrolled into 2 groups: TOL (n = 25), defined as individuals who, for at least 1 year prior to enrollment, had not taken immunosuppressive medications and had stable renal function and serum creatinine within 25% of baseline (as evaluated by 3 experienced transplant physicians); and SI (n = 33), with clinically stable renal function (using the same criteria as TOL) while on a maintenance triple-drug immunosuppressive regimen (including a calcineurin or mammalian target of rapamycin inhibitor, an antiproliferative agent, and corticosteroids). An additional group of normal HC participants (n = 42) with no known history of renal disease/dysfunction or evidence of acute medical illness was enrolled. Breakdown of samples used in each assay are shown in Supplemental Table 1.

For the purposes of predictive modeling and verification of our initial findings, we divided the TOL participants into a training set (TOL-TRN; n = 19) and a test set (TOL-TST; n = 6) based on enrollment prior to or after November 2007, respectively. The SI group was also divided into a training set (SI-TRN; n = 27) and a test set (SI-TST; n = 6) for modeling and verification, also based on enrollment timing.

Finally, an independent set of kidney transplant recipient patients was collected in Europe (25), and frozen PBMCs from this collection were used for flow cytometry analysis. Frozen PBMCs from our patients and 3 of the European cohorts (1 TOL, 1 HC, and 1 SI, specifically calcineurin inhibitors), were studied by flow cytometry at the same time in an independent laboratory (see below).

Table 4

<table>
<thead>
<tr>
<th>Actual TOL</th>
<th>Actual SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted TOL</td>
<td>22</td>
</tr>
<tr>
<td>Predicted SI</td>
<td>1</td>
</tr>
</tbody>
</table>

Confusion matrix and classification summary by SVM (see Methods) showing transitional B cells as the most predictive population when frozen PBMCs from ITN samples were analyzed (PPV, 85%; NPV, 96%).

Table 5

<table>
<thead>
<tr>
<th>Actual TOL</th>
<th>Actual SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted TOL</td>
<td>5</td>
</tr>
<tr>
<td>Predicted SI</td>
<td>1</td>
</tr>
</tbody>
</table>

Confusion matrix and classification summary by SVM (see Methods) showing transitional B cells as the most predictive population when frozen PBMCs from an independent set of European samples were analyzed (PPV, 63%; NPV, 90%).
stain-lyse method, cells from blinded samples were stained with 5-color monoclonal antibody panels conjugated to FITC, PE, PERCP, PECY7, or APC (Becton Dickinson; ref. 49). Marker/fluorochrome combinations used are described in Supplemental Table 4.

For flow cytometry on frozen cells, PBMCs were stained in PBS plus 2 mM EDTA, 0.5% BSA, 5% normal mouse serum, and 5% normal rat serum on ice for 30 minutes with fluorochrome-conjugated mouse anti-human monoclonal antibodies (see Supplemental Table 4). After washing with PBS plus 2 mM EDTA and 0.5% BSA, cells were stained with LIVE/DEAD aqua–fluorescent reactive dye (Invitrogen) in PBS on ice for 30 minutes, then fixed with 0.5% formaldehyde. Samples were blinded and run on a LSRII flow cytometer (BD Biosciences) at the University of Rochester.

For intracellular cytokine (IL-10 and TGFβ) evaluations, cells were divided into 2 samples and cultured in complete media (RPMI supplemented with 20% BSA) with brefeldin (1 μg/ml) and monensin (2 μM) in the presence or absence of stimulation (500 ng/ml PMA and 500 ng/ml ionomycin) for 5 hours. After culture, cells were washed with FACS Buffer (PBS plus 0.5% BSA) 2x and then surface stained with extracellular antibody cocktail for 30 minutes at 4°C. Cells were then washed 2x with PBS and stained with LIVE/DEAD aqua–fluorescent reactive dye (Invitrogen) for 30 minutes at 4°C. Cells were then washed 2x with FACS buffer and then fixed in 4% paraformaldehyde (Sigma-Aldrich) in PBS for 30 minutes at 4°C. Cells were subsequently permeabilized with Perm/Wash Buffer (catalog no. L34957; Invitrogen) at 1:1,000 in PBS for 30 minutes at 4°C. Cells were then washed 2x with FACS buffer and then fixed in 4% paraformaldehyde (Sigma-Aldrich) in PBS for 30 minutes at 4°C. Cells were then intracellularly stained in Perm/Wash buffer with the ICS antibody cocktail for 1 hour at 4°C. Finally, cells were washed 2x with Perm/Wash buffer and resuspended in 4% paraformaldehyde in PBS until analysis on the LSR II Flow Cytometer (BD) and with Flow Jo Software (Tree Star Inc.)

MassARRAY quantitative gene expression. Multiplexed primer and competitive template designs were created using the MassARRAY quantitative gene expression (MassARRAY QGE) Assay Design software (version 1.0; Sequenom) for random hexamer priming, such that at least 1 PCR primer spanned an exonic boundary per transcript assayed. The 228 genes tested were assayed in a series of 20-plex reactions on RNA from blinded samples and are shown in Supplemental Table 2. Copy gene number determination was conducted as described previously (34).

Statistics. Microarray data background adjustment, normalization, and summarization were performed using the robust multichip average method. Microarray quality assurance was carried out by detecting outlier arrays based on standard posthybridization quality metrics (51). A linear mixed effect model was used for correcting potential processing batch effect as well as for estimating clinical group effect. Pairwise comparisons were performed with Tukey adjustment for group-level multiple comparisons, to identify differentially expressed genes between different clinical groups. The Benjamini and Hochberg approach for control of FDR was adopted for genewise multiple testing adjustments (52). Statistical criteria for identification of differentially expressed genes were FDR-adjusted P < 0.05 and fold change greater than 1.5.

Urine RT-PCR data was normalized against 18S-rRNA; peripheral blood MassARRAY QGE was normalized against a set of 5 stable housekeeping genes. For urine RT-PCR, peripheral blood MassARRAY QGE, and flow cytometry data, a Shapiro-Wilk test was adopted to check whether data
came from a normally distributed population. Although a log2 data transformation substantially reduced extent of deviation from the normal distribution, logarithm-transformed data of a substantial number of genes or cell populations still deviated from the normal distribution. Hence, a nonparametric Wilcoxon rank-sum test was conducted on a per-gene basis for pairwise comparisons between the clinical groups. Gene- or cell population-wise multiple testing adjustments were performed using the FDR method of Benjamini and Hochberg (52). A statistical criterion for identification of differentially expressed genes or cell populations was FDR-adjusted P < 0.05. Hierarchical clustering images for microarray and MassARRAY QGE data were generated based on Pearson correlation using GeneSpring GX 7.3.1 (Agilent). Box plot images were generated using S-Plus (TIBCO). Classification approaches were applied to MassARRAY QGE and flow cytometry on frozen cells using log2 transformed data. In all instances, variable selection was performed using 1-way ANOVA testing TOL versus SI with multiple groups of variables ranging in size from 1 to 10 with a step of 1. LDA with equal prior probabilities using LOOCV (54). The model was constructed using a training set of 19 TOL and 24 SI samples and then applied to an ITN test set of 6 TOL and 6 SI samples collected and processed after November 2007.

The normalized, log2-transformed expression levels for the 3 genes IGKV1D-13, IGKV4-1, and JGLL1 were used to generate a probability score between 0 and 1 for each participant’s membership in the TOL group. LOOCV was applied to determine how accurately the learning algorithm predicted data that it was not trained on. In using this approach, the LDA model was trained multiple times, using all but 1 of the training set data points. The sample that was removed was restated iteratively to generate the best PPV and NPV during training. Feature selection was embedded within the LOOCV process.

Support vector machine (SVM) was the most effective for flow cytometry data in determining 1 population to have the best PPV. The SVM model was configured with cost-based shrinking from 1 to 1,001 with step 100 and tolerance from 0.001. The kernel used a radial basis function (RBF) set at 1/no. variables and LOOCV (54). The model was constructed using 23 ITN TOL and 31 ITN SI samples and applied to the test set of European IOT group samples consisting of 6 TOL and 12 SI samples. Partek Genomics Suite (version 6.4; Partek) was used for prediction approaches that allowed us to test over 1,000 models using various feature combinations and SVM permutations.

\[
P_{\text{test}} = \frac{\exp \left( \beta_1 + \beta_2 G_1 + \beta_3 G_2 + \beta_4 G_3 \right)}{1 + \exp \left( \beta_1 + \beta_2 G_1 + \beta_3 G_2 + \beta_4 G_3 \right)}
\]

(Equation 1)

Classification was based on the posterior probabilities with a 0.5 cutoff; thus, participants with a score of 0.5 or higher would be assigned to the TOL group. LOOCV was applied to determine how accurately the learning algorithm predicted data that it was not trained on. In using this approach, the LDA model was trained multiple times, using all but 1 of the training set data points. The sample that was removed was restated iteratively to generate the best PPV and NPV during training. Feature selection was embedded within the LOOCV process.

Support vector machine (SVM) was the most effective for flow cytometry data in determining 1 population to have the best PPV. The SVM model was configured with cost-based shrinking from 1 to 1,001 with step 100 and tolerance from 0.001. The kernel used a radial basis function (RBF) set at 1/no. variables and LOOCV (54). The model was constructed using 23 ITN TOL and 31 ITN SI samples and applied to the test set of European IOT group samples consisting of 6 TOL and 12 SI samples. Partek Genomics Suite (version 6.4; Partek) was used for prediction approaches that allowed us to test over 1,000 models using various feature combinations and SVM permutations.


