METHODS

Patient information

Sixty-seven allograft biopsies from 50 children and young adults (15 months - 22 years in age) who received kidney transplants at Stanford between 1995-2001 were analyzed. The study was governed by informed consent and Stanford Institutional Review Board approval. Immunosuppression consisted of treatment with glucocorticoids, a calcineurin inhibitor (tacrolimus or cyclosporine), an antimetabolite (azathioprine or mycophenolate mofetil) and daclizumab induction. Graft losses occurred at a mean time of 10 months post-biopsy and range of 18 months to 8 years post-transplantation.

Biopsy samples

Biopsies were obtained during acute (>10% rise in serum creatinine from baseline) and chronic (Schwartz glomerular filtration rate\(^{14} \leq 50 \text{ ml/min/1.73m}^2\)) allograft dysfunction (n=52), at engraftment (n=8) and at protocol time points with stable graft function (n=7; Schwartz glomerular filtration rate > 80 ml/min/1.73 m\(^2\)). All but 5 biopsies were obtained prior to treatment intensification. Pathological analyses were done blinded to knowledge of clinical outcomes. Half of each biopsy sample was frozen in liquid nitrogen and the remainder was fixed with 10% buffered formalin for light microscopy. Eight 3 \(\mu\)m thick sections were obtained
from paraffin embedded tissues and stained with hematoxylin and eosin (H&E) and Periodic–
Acid Schiff (PAS); all biopsies were “adequate”, as defined by Banff 1997 criteria\(^5\) and had \(\geq 10\) glomeruli and two arteries.

Evidence of AR,\(^5\) chronic allograft nephropathy (CAN),\(^15\) drug toxicity (DT),\(^16\) recurrent disease and infection (IN) were documented (Table 1). No biopsy had features of post-transplant lymphoproliferative disorder or viral inclusions. Immunohistochemical staining for CD20, CD4, CD8 and PCNA (see Supplemental Information for methods) were done on untreated AR samples. Additionally, an independent set of 31 archived AR biopsy samples were analyzed by immunohistochemical staining with antibodies against CD20.

**Microarray Hybridization**

Sixty-seven biopsies were analyzed by hybridization to 67 DNA microarrays. The microarrays each contained 28,032 DNA spots (24,497 unique cDNA clones), representing \(~12,440\) human genes (i.e., unique UniGene clusters). Forty-three percent of the genes were represented at least twice, most of these by independent cDNA clones, thus providing internal controls for the reproducibility of gene expression quantification.
Total RNA was isolated (mean 3.4 µg) from frozen biopsies using Tri Reagent (MRC Inc., Cincinnati, OH) with the GIBCOBRL RNA extraction protocol. A “common reference” RNA pool\(^1\) was used as an internal standard in all microarray hybridizations to allow expression levels of individual genes to be measured on a common, internally standardized scale in every patient sample. Sample or reference RNA samples were subjected to two successive rounds of amplification, using a protocol modified from Wang et al\(^1\) (mean sample amplified RNA = 15 µg). After amplification, the hybridization and image analysis protocols were as described at http://www.microarrays.org.

**Data Analysis**

Raw data were collected and analyzed using GenePix Pro 3.0 (Axon Instruments). After flagging obvious blemishes, raw data was deposited into the Stanford Microarray Database (http://genome-www4.stanford.edu/MicroArray/SMD/index.html). Non-flagged spots with fluorescence intensity < 2.5 fold over background and genes with technically adequate measurements in < 75% of all samples were excluded. A group of 1340 cDNAs with mRNA levels differing from the median by at least 2.9 fold in at least 6 samples were selected for
further analysis. A hierarchical clustering algorithm was applied to both genes and samples using the Pearson correlation coefficient as a measure of similarity in the pattern of variation in gene expression between pairs of genes or pairs of samples. The results were visualized and analyzed with TreeView (M. Eisen; http://rana.lbl.gov).

To estimate the ability of individual genes to distinguish between two sample groups, SAM analysis (http://www-stat-class.stanford.edu/SAM/SAMServlet) was used. Enrichment of specific functional groups of genes was assessed by two-tailed p-values using the hypergeometric distribution as described. A group of 86 “T-cell specific” genes were identified as those highly expressed in thymus, whole blood, and the MOLT-4 T-cell line, relative to 44 other unique human tissues and cell lines, of which 46 were present in our dataset (see Supplemental Table for list). A group of 2610 transcripts were identified as “T-cell inducible” in primary human T lymphocytes, of which 145 were present in our dataset (Diehn et al, PNAS in press, see Supplemental Table for list). Kaplan-Meier survival analyses were assessed using the log-rank method of Cox (significance value at p < 0.05), used to determine the relationship between graft survival or recovery of graft function (defined as return of the serum creatinine to baseline one month after treatment of AR) and CD20 density.
Immunohistochemistry

Immunohistochemical staining was performed using CD20, CD8, PCNA (from Dako, Carpinteria, CA) and CD4 (Novacstra, Vector Laboratories, Burlinglame, CA) antibodies on all AR samples and selected samples with infection and drug toxicity. In addition, 31 archived AR biopsies were also stained with CD20. Sections were stained with monoclonal antibodies to human CD20 (dilution 1:400), CD8 (dilution 1:25), CD4 (dilution 1:10) and PCNA (dilution 1:100), using Biogenex I 1000 automated antigen retrieval system and Biogenex I 6000 automated stainer. The cores were scanned in a blinded fashion for CD20, CD4 and CD8 cell density. Each core biopsy was evaluated by a single pathologist for staining by each antibody across the entire specimen, with all represented 0.69mm high-power fields (hpfs) enumerated for the respective antibody. Cells were counted per hpf and the number of hpf counted were documented in each core. For each specimen, the single hpf with the highest CD20 count was then identified, and arbitrary threshold cell counts of > 275 and <100 in the selected hpf were used to assign "CD20 positive" or "CD20 negative" status. The CD20 density counts on the web site are the maximal CD20 densities over any single hpf after counts were completed over the entire core.
Supplemental Information

Additional information on methods, immunohistochemistry images and analyses are available at our Web supplement (http://genome-www.stanford.edu/rejection/).