Published in final edited form as:

Transplantation. 2010 December 27; 90(12): 1401–1408. doi:10.1097/TP.0b013e3182000085.

Single Nucleotide Polymorphisms, Acute Rejection and Severity of Tubulitis in Kidney Transplantation, Accounting for Center-to-Center Variation

A Israni, MD, MS,

Department of Nephrology, Hennepin County Medical Center, Department of Epidemiology & Community Health, University of Minnesota, Minneapolis, MN

Robert Leduc, PhD*,

Division of Biostatistics, University of Minnesota, Minneapolis, MN

John Holmes*,

Division of Epidemiology, Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, P.A

Pamala A. Jacobson, PharmD,

Department of Experimental and Clinical Pharmacology, University of Minnesota, Minneapolis, MN

Vishal Lamba, PhD,

Department of Genetics, Cell Biology and Development, University of Minnesota, Minneapolis, MN

Weihua Guan, PhD,

Division of Biostatistics, University of Minnesota, Minneapolis, MN

David Schladt, MS,

Division of Biostatistics, University of Minnesota, Minneapolis, MN

Jinbo Chen, PhD,

Division of Biostatistics, Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, P.A

Arthur J. Matas, MD, and

Department of Surgery, University of Minnesota, Minneapolis, MN

William S. Oetting, PhD

Corresponding Author: Ajay K. Israni, M.D., M.S., Department of Medicine, 701 Park Avenue, Minneapolis, MN 55415-1829.

Telephone # 612-347-5871, Fax # 612-347-2003, isran001@umn.edu.

* Authors contributed equally

Conflict of Interest

None to report

Author Contributions:

A Israni, MD, MS, - Research design, writing of paper, performance of research, contributed to analytical tools.

*Robert Leduc, PhD - Research design, writing of paper, performance of research, contributed to analytical tools, data analysis.

*John Holmes - Research design, writing of paper, performance of research, contributed to analytical tools, data analysis.

Pamala A. Jacobson, PharmD - Research design, writing of paper, performance of research.

Vishal Lamba, PhD, - Writing of paper, performance of research, contributed to analytical tools.

William S. Oetting, PhD, - Research design, writing of paper, performance of research.

Weihua Guan, PhD - Research design, writing of paper, contributed to analytical tools, data analysis.

David Schladt, MS, - Research design, writing of paper, performance of research, contributed to analytical tools, data analysis.

Jinbo Chen, PhD, - Research design, contributed to analytical tools, data analysis.

Arthur J. Matas, MD, - Research design, writing of paper, performance of research, contributed to analytical tools.

Department of Experimental and Clinical Pharmacology and Institute of Human Genetics,
University of Minnesota, Minneapolis, MN

For the DeKAF investigators

Abstract

Background—Acute rejection (AR) is associated with worse renal allograft outcomes.

Therefore, this study investigated single nucleotide polymorphisms (SNPs) to identify genetic variants associated with AR, accounting for center variation, in a multi-center, prospective, observation study.

Methods—We enrolled patients from 6 transplant centers, 5 in the U.S. and one in Canada. A total of 2,724 SNPs were genotyped. We accounted for center variation in AR rates by stratifying by transplant center and using novel knowledge discovery methods.

Results—There was significant center variation in AR rates across the six transplant sites. ($p < 0.0001$) Accounting for this difference and clinical factors independently associated with AR, we identified 15 novel SNPs associated with AR with stratification by transplant center ($p < 0.05$). We also identified 15 novel SNPs associated with severity of tubulitis scores, after adjusting for transplant center and other clinical factors independently associated with severity of tubulitis. ($p < 0.05$) There was some overlap with one SNP associated with AR and also associated with severity of tubulitis, among the top 15 SNPs.

Conclusion—Center-to-center variation is a major challenge to genomic studies focused on AR. The SNPs associated with AR and severity of tubulitis in this study, will need to be validated in independent cohort of kidney transplant recipients.

Keywords

single nucleotide polymorphisms; acute rejection; tubulitis

Introduction

Acute rejection (AR) is associated with worse long-term allograft survival. AR is also a risk factor for reduced renal allograft function post-transplant - both because treatment of AR does not always lead to a return to baseline kidney function and because there may be ongoing subclinical damage. AR is also a risk factor for cardiovascular disease whereas reduced renal function has been shown to be a risk for cardiovascular and overall mortality. (1) Despite a consensus that immunosuppressive medication regimens should be designed to prevent AR and its consequences, controversy remains over the optimal regimen. Recently it has been proposed that immunosuppression be individualized in relation to the patient's immunological risk.(2)

Using a multi-center cohort of kidney and simultaneous kidney-pancreas (SPK) transplant recipients, we studied the association of single nucleotide polymorphisms (SNPs) with AR. One of our initial observations was that the incidence of AR varied significantly by center. Such variation can result in associations at one center that cannot be validated at other sites. We therefore did our analyses using standard statistical methods and novel knowledge discovery methods that accounted for this center-to-center variation. We found a number of SNPs were associated with AR. Defining SNPs associated with increased rejection risk will potentially allow for individualizing immunosuppression and to gain better insight into mechanisms of AR.

Results

Baseline Characteristics and Outcomes

990 kidney transplant recipients were enrolled from 6 transplant centers. The demographic and transplant related characteristics for these recipients are listed in Table 1. In the first year post-transplant, 181 (18 %) recipients experienced AR and 177 (98 %) were biopsy confirmed. However, the rate of AR varied across the 6 study centers: 12%, 30%, 15%, 14%, 0% and 9% for center 1–6, respectively ($p < 0.0001$). The median time to AR diagnosis was 33 days (range 7–364 days). The recipients' pathologists reported the following types of AR: 70% cellular, 18% antibody mediated, 10% with features of both and 2% unknown. Treatment for AR was as follows: 52% steroids alone, 4% antibodies alone, 25% steroid and antibodies together, 12% steroids followed by antibodies and 7% other.

SNPs Associated with AR

In a Cox proportional hazards model adjusted by recipient race only and stratified by transplant center, the top 15 SNPs that were associated with AR, all with $p < 0.05$, are shown in Table 2. (All SNPs genotyped are in Supplemental Table 1) These simple models for the entire list of SNPs tested for association with AR are shown in Supplemental Table 2. The SNPs are listed by increasing p-values. Many of the top 15 SNPs such as rs2227931 in *ATR*, rs2267130 in *CHEK2*, rs3088142 in *DUSP13* represent signaling pathways which play an important role in activation of T cells. The SNP in *VANGL1* (rs4839469) was also associated with AR. In the multivariate model, similar SNPs remained as top SNPs associated with AR. (Table 3A) In a multivariate model, the factors that were independently associated with AR were: recipient factors (race, age, gender, weight), PRA presence, number of HLA mismatches, T or B-cell cross-match positive, antibody induction, type of calcineurin inhibitor used, steroid use at day 14 post-transplant, simultaneous kidney-pancreas transplant (versus kidney transplant alone), cause of ESRD, living donor (versus deceased donor), donor age and transplant center. Given that transplant center was independently associated with AR even after accounting for these clinical factors, we analyzed the SNPs associated with AR, stratifying by transplant center. The multivariate models for the entire list of SNPs tested for association with AR are shown in Supplemental Table 3. None of the SNPs were statistically significant, after accounting for FDR of 10%. In the analysis limited to acute cellular rejection only, similar SNPs were seen among the top 15 SNPs. (Footnote, Table 3A)

SNPs Associated with AR Using Three Knowledge Discovery Methods

We determine combination of SNPs associated with AR. The 14 SNPs associated with AR as determined by all three methods are highlighted in Table 3B. None of these SNPs were ranked among the top 15 SNPs in the multivariate analysis as shown in Table 3A.

SNPs Associated with Severity of AR

Given the predominance of cellular rejection in our study population, we divided our cohort into 3 groups, no biopsy ($n = 609$), t_1 ($n = 271$) and t_2 ($n = 83$). There were no t-scores available for 27 patients from the local pathologist reading of the biopsy confirmed AR. These 27 subjects were excluded from this analysis. The distribution of the t-scores on the remaining 154 patients' AR biopsies were as follows: 0 (17%), 1 (37 %), 2 (32 %) and 3 (14 %). There were 200 patients with no AR, but had a kidney allograft biopsy, in a similar time-frame post-transplant (no AR group with biopsy at median of 32 days and range 3 to 366 days versus AR group with median 30 days and range of 7 to 364 days). The distribution of pathology t-scores in these 200 patients were as follows: 0 (89%), 1 (9%), 2 (2 %) and 3 (<1%). In both groups, the t-scores were highly correlated with the i-scores

($p < 0.001$, Fisher's exact test), therefore the analysis was not repeated for the i-scores. The factors independently associated with severity of acute rejection were transplant center (Supplemental Table 4), recipient race, recipient age, recipient gender, PRA, number of HLA mismatches, donor age, type of calcineurin inhibitor use at baseline, recipient weight at time of transplant, antibody induction, and smoking status. The SNPs associated with having more severe tubulitis scores ($t = 2$ versus $t = 1$), are shown in Table 4. One of these top 15 SNPs associated with severity of AR did overlap with the SNPs associated with AR namely rs6163 in *CYP17A1* ($p = 3.30 \times 10^{-3}$). The two SNPs most strongly associated with t-scores in this analysis were: rs2228059 in *IL15RA* ($p = 4.8 \times 10^{-4}$) and rs811925 in *PRDM1* ($p = 2.3 \times 10^{-4}$). The entire list of SNPs tested for association with severity of AR shown in Supplemental Table 5. None of the SNPs were statistically significant, after accounting for FDR of 10%.

Discussion

The goal of this study was to identify genetic variants associated with AR. However, we found significant center-to-center variation in occurrence of AR among the study centers. Such center-to-center variation could explain why some SNPs associated with AR at one center, may be difficult to validate at other centers. We also identified SNPs associated with AR using several methods including traditional statistical methods with stratification by transplant center. We also used novel knowledge discovery methods to identify potential SNP-SNP interactions after accounting for the center-to-center variation in occurrence of AR. These knowledge discovery methods allowed us to analyze each transplant center separately. The most strongly associated SNPs with AR were involved in signaling pathways genes such as rs2227931 in *ATR*, and rs3088142 in *DUSP13*. Such genes may play a role in T cell activation. We also determined that different novel SNPs were associated with severity of t-scores on biopsy such as rs811925 in *PRDM1* and rs2228059 in *IL15RA*. *PRDM1* is involved in repressing the beta-interferon gene.(3) whereas *IL15RA* is a high affinity receptor for the inflammatory cytokine interleukin 15.(4)

Our study is the first to account for significant center-to-center variation in the occurrence of AR. Prior multi-center SNP studies have not accounted for center-to-center variation in AR. (5, 6) The most significant SNPs associated with AR in the present study, ranked by p-value, were predominantly in intracellular pathways that could activate T cells. These include SNPs in the rs2227931 in *ATR*, rs2267130 in *CHEK2*, and rs3088142 in *DUSP13* genes. *ATR* is a critical checkpoint kinase that leads to phosphorylation of several proteins.(7) It is reasonable to assume that greater *ATR* activity could lead to cell activation during an immune response. Thus, the C allele of the coding synonymous SNP rs2227931 in *ATR* is associated with increased risk of AR in this study, possibly by increasing *ATR* activity through splice variation. The *CHEK2* is a critical checkpoint regulator in cell cycle replication. *CHEK2* is activated by phosphorylation in response to DNA damage.(8) Thus, the C allele of the intronic SNP rs2267130 in *CHEK2* is associated with decreased risk of AR, possibly by reducing the activity of *CHEK2*. *DUSP13* is a member of a protein-tyrosine phosphatase superfamily that cooperates with protein kinases to regulate cell proliferation and differentiation.(9) The T allele of missense SNP rs3088142 in *DUSP13* gene is associated with an increased risk of AR and probably leads to increased activity of *DUSP13*. The A-allele of missense SNP rs4839469 in *VANGL1* is associated with AR. *VANGL1* is a novel protein since its mutations in this gene have been associated with neural tube defects(10) and *VANGL1* is involved in tumor cell migration.(11) Its role in AR is not known. It is possible that other studies could find similar SNPs associated with AR, if they genotyped them in their cohort.

Previous studies have found other SNPs associated with AR namely *ABCB1* gene haplotype(12), *ABCB1* 2677T allele, *IMPDH2* 3757C allele and *IL-10*-592A,(5) and *IMPDH1* (rs2278293 and rs2278294)(13). Other studies have also genotyped SNPs in 20 other genes that have provided conflicting associations with transplant outcomes.(6) These genes had SNPs that were genotyped in the present study but not found to be significantly associated with AR. It is possible that these SNPs are not associated with AR after accounting for the center-to-center variation or center specific practices.

This study is the first to determine SNPs associated with severity of AR as determined by t-scores. It is well known that there is variation in the severity of acute rejection, (14) thus it is not surprising that not all the same SNPs were associated with risk of AR and severity of tubulitis. This study describes the clinical and genetic factors that are associated with this variation in severity of AR by studying the severity of tubulitis on allograft biopsy. The *PRDM1* gene encodes a protein that acts as a repressor of beta-interferon gene expression. The protein binds specifically to the *PRDI* (positive regulatory domain I element) of the beta-IFN gene promoter.(15) The C allele of the missense SNP rs811925 was associated with increased odds of a tubulitis score greater than or equal to two compared to t-score less than or equal to 1. (Table 4) Similarly, a missense SNP (rs2228059) in *IL15RA*, which is a high affinity cytokine receptor for the inflammatory cytokine interleukin 15, was seen to be associated with severity of tubulitis. This receptor is structurally related to *IL2R* alpha, the high affinity receptor for *IL2*.(16) In murine models, *IL15* intra-articular injections induce a local tissue inflammatory infiltrate composed to predominantly T cells.(4) (Table 4)

It is possible that the association of SNPs in some of the candidate genes was not seen in our study after accounting for multiple testing because of small effect sizes of these SNPs and limitations of sample size. We are planning on validating the SNPs shown in Table 3A, 3B and 4, in the subsequent 2,000 kidney transplant recipients that are being enrolled by this study. We plan to genotype only the top SNPs from the present study in this subsequent validation cohort. The number of SNPs that will be genotyped in the larger validation cohort will be determined by the number of AR events in the validation cohort. Since the validation cohort will need to account for multiple testing using a rigorous test, only a few top SNPs shown in Table 3A, 3B and 4 will be eligible for genotyping in the validation cohort. The genes of those SNPs association that have an association in the validation cohort will need to be studied for gene expression in a future study. The current study does not collect any samples for gene expression. Another potential limitation is that our t and i-scores were not read by a central pathologist blinded to the clinical information. However, the central pathologist scores and local pathologist had similar rates of AR using the biopsy scores of $t > 1$ and $i > 1$ (McNemar's Test, $p = 0.60$, data not shown).(17)

In summary, this study has described significant center-to-center variation in AR which is a challenge especially for all multicenter, genomics study. This variation can explain the difficulty in validating SNPs or other genomics markers found in one center, in another center. We have analyzed SNPs associated with AR, accounting for this center-to-center variation. SNPs potentially involved in T cell activation such as rs2227931 in *ATR*, *CHEK2* (rs2267130) and *DUSP13* (rs3088142) are novel findings in this study. This study is also unique for determining SNPs in *PRDM1* (rs811925) and *IL15RA* (rs2228059), associated with severity of AR using biopsy scores. Our study will validate the top SNPs (Table 3A, 3B and 4) in an ongoing, larger cohort of 2,000 kidney transplant recipients, after accounting for multi-testing. In the future, with a larger cohort of recipients, this study may have enough power to conduct a genome-wide association study.

Methods

Clinical Data

Between, 10/03/2005 and 4/10/2008, 990 kidney and SPK recipients at 6 transplant centers (Table 1) were enrolled in a prospective study of kidney allograft function. Patients were consented for participation at the time of or soon after transplantation. All kidney transplant recipients undergoing a kidney or SPK transplant were eligible. Patients with non-kidney solid organ transplants were not eligible. The Institutional Review Boards at each of the study sites approved this study.

Immunosuppression and AR treatment was center-specific. Clinical data were collected at the time of transplant and regularly until allograft failure and maintained in a central database. Delayed graft function was defined as need for dialysis in the first week post-transplantation. AR was defined by the treating physician. Local pathologist reading of biopsies was collected for all biopsies done for cause.

Genotyping

SNP genotyping was done primarily using a customized Affymetrix GeneChip having a total of 3,404 SNP assays(18) on a Affymetrix GeneChip Scanner 3000 Targeted Genotyping System (Affymetrix, Santa Clara, CA).(19) Additional variants were genotyped using the SNPlex (Applied Biosystems Inc, Foster City, California) and Sequenom (Sequenom, Inc, San Diego, CA) platforms, as per manufacturer's recommendation. We selected SNPs for genotyping if they were known or thought to be functional within biologically relevant genes to transplantation and included genes in pathways associated with immunity, cell signaling, cell growth and proliferation, and drug absorption, disposition, metabolism and excretion. In the absence of functional variants, intragenic tagging variants were selected. The design of this custom SNP chip has been described previously.(18)

Genotyping Quality Control—We assessed data quality using negative controls (water) and duplicate samples (3% on Affymetrix, 7% on SNPlex, and 1% on Sequenom). On the Affymetrix platform, we genotyped 31 individuals in duplicated with >99% concordance. For all platforms, we eliminated SNPs with concordance rates <90% and with call rates <60%. We genotyped 20 SNPs on multiple platforms and had a concordance rate of >97% and SNP call rates were present for greater than 82% of the samples. SNPs that deviated from Hardy Weinberg equilibrium ($p\text{-value} < 1 \times 10^{-6}$) were removed. SNPs were excluded from further analysis if the minor allele frequency (MAF) was <5% in both the African American and non-African American population. In the final analysis, we included 2,552 SNPs from the Affymetrix chip platform, 165 from the SNPlex platform and 7 from the Sequenom platform (Supplemental Table 1) with a total of 2,724 SNPs.

Data Analyses

We studied the association of SNPs with AR. Because we found a statistically different incidence of AR by site, the standard statistical methods included stratification by transplant center. The stratified analysis excludes the center with no AR event. For the final models, the p-values for SNP association, below the FDR set at 10% were considered statistically significant after accounting for multiple testing. Since standard statistical methods would not allow for analyzing SNPs at each center separately due to limited power, we also utilized several knowledge discovery methods to determine SNPs at each transplant center individually.

Standard Statistical Analysis of Single SNPs

Analysis was conducted utilizing SAS v9.1 (The SAS Institute, <http://www.sas.com>) using an additive genetic model. Separate Cox proportional hazards models were used to investigate the association of each SNP genotype with time to AR, adjusting for recipient race and stratifying by transplant center.(20). This analysis stratified by transplant center, automatically excludes any center that does not have an episodes of AR. Model selection was performed twice: once stratifying by transplant center and another time without.(21) Retained covariates from each model were combined to make a baseline model. Each SNP was added into the baseline model individually and the model was stratified by the transplant center.

Knowledge Discovery Methods

In order to analyze SNPs at each center separately, we utilized knowledge discovery methods. We utilized three correlational-based feature selection in the Weka software suite of programs(22) to select SNPs associated with AR at each of the transplant centers which had cases of AR. The correlation-based feature selections included a best-first search procedure (23) which was used to determine the SNPs eligible for the other two methods namely, a rule discovery tool named “JRip”(24) and decision trees were constructed with J48.(25) The SNPs in rules that classified for presence of AR were extracted into the final feature set. The SNPs (nodes) of decision trees that resulted in positive outcome (presence of AR) were extracted to the final feature set. Each of the 3 knowledge discovery methods were conducted with 10-fold cross-validation with each fold preserving the distribution of the outcome of AR.

A final SNP set was created combining the SNPs found by the three knowledge discovery methods. Duplicates were removed from the final set. Then using a Cox proportional hazards model, we determined the association with AR for these SNPs, across all the transplant centers in the study.

Statistical Analysis of Biopsy tubulitis (t)-scores

In order to assess severity of AR, the study population was divided into groups: no biopsy, biopsy with t-score 1 and t-score 2. We then conducted a multinomial logistic regression analysis adjusted by clinical center.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We acknowledge the dedication and hard work of our coordinators at each of the six clinical sites: University of Alberta, Nicoleta Boboccea, Tina Wong, Adrian Geambasu and Alyssa Sader; University of Manitoba, Myrna Ross and Kathy Peters; University of Minnesota, Mandi DeGrote and Jill Nagorski; Hennepin County Medical Center, Lisa Berndt; Mayo Clinic, Tom DeLeeuw; University of Iowa; Wendy Wallace and Tammy Lowe; University of Alabama, Catherine Barker. We also acknowledge the dedicated work of our research scientists: Marcia Brott, Becky Willaert, Jennifer Vigliaturo and Winston Wildebush.

Funding Source

This work was supported by the National Institutes of Health NIAID Genomics of Transplantation (5U19-AI070119) and DeKAF (5U01-AI058013)

References

1. Israni A, Snyder JJ, Skeans MA, et al. Predicting coronary heart disease after kidney transplantation: Patient Outcomes in Renal Transplantation (PORT) Study. *American Journal of Transplantation*. 2010; 10:338. [PubMed: 20415903]
2. KDIGO. KDIGO Clinical Practice Guideline for the care of kidney transplant recipients. *American Journal of Transplantation*. 2009; 9(Supplement 3)
3. Tam W, Gomez M, Chadburn A, Lee J, Chan W, Knowles D. Mutational analysis of PRDM1 indicates a tumor-suppressor role in diffuse large B-cell lymphomas. *Blood*. 2006; 107 (10):4090. [PubMed: 16424392]
4. Waldmann TADS, Tagaya Y. Contrasting roles of IL-2 and IL-15 in the life and death of lymphocytes: implications for immunotherapy. *Immunity*. 2001; 14 (2):105. [PubMed: 11239443]
5. Grinyo J, Vanrenterghem Y, Nashan B, et al. Association of four DNA polymorphisms with acute rejection after kidney transplantation. *Transplant International*. 2008; 21 (9):879. [PubMed: 18444945]
6. Marder B, Schroppel B, Murphy B, Marder B, Schroppel B, Murphy B. Genetic variability and transplantation. *Current Opinion in Urology*. 2003; 13 (2):81. [PubMed: 12584466]
7. Lindsey-Boltz L, Serçin O, Choi J, Sancar A. Reconstitution of human claspin-mediated phosphorylation of Chk1 by the ATR (ataxia telangiectasia-mutated and rad3-related) checkpoint kinase. *Journal of Biological Chemistry*. 2009; 284 (48):33107. [PubMed: 19828454]
8. Stracker TH, Usui T, Petrini JH. Taking the time to make important decisions: the checkpoint effector kinases Chk1 and Chk2 and the DNA damage response. *DNA Repair*. 2009; 8 (9):1047. [PubMed: 19473886]
9. Chen HH, Luche R, Wei B, NKT. Characterization of two distinct dual specificity phosphatases encoded in alternative open reading frames of a single gene located on human chromosome 10q22.2. *J Biol Chem*. 2004; 279 (27):41404. [PubMed: 15252030]
10. Kibar Z, Bosoi CM, Kooistra M, et al. Novel mutations in VANGL1 in neural tube defects. *Hum Mutation*. 2009; 30 (7):E706.
11. Lee JK, Bae JA, Sun EG, et al. KITENIN increases invasion and migration of mouse squamous cancer cells and promotes pulmonary metastasis in a mouse squamous tumor model. *FEBS Letters*. 2009; 583 (4):711. [PubMed: 19166844]
12. Bandur S, Petrasek J, Hribova P, et al. Haplotypic structure of ABCB1/MDR1 gene modifies the risk of the acute allograft rejection in renal transplant recipients. *Transplantation*. 2008; 86 (9):1206. [PubMed: 19005401]
13. Wang J, Yang JW, Zeevi A, et al. IMPDH1 gene polymorphisms and association with acute rejection in renal transplant patients. *Clinical Pharmacology & Therapeutics*. 2008; 83 (5):711. [PubMed: 17851563]
14. Shinn C, Malhotra D, Chan L, Cosby RL, Shapiro JJ. Time course of response to pulse methylprednisolone therapy in renal transplant recipients with acute allograft rejection. *American Journal of Kidney Disease*. 1999; 34 (2):304.
15. Barnes NA, Stephenson SJ, Tooze RM, GMD. Amino acid deprivation links BLIMP-1 to the immunomodulatory enzyme indoleamine 2,3-dioxygenase. *J Immunol*. 2009; 183 (9):5768. [PubMed: 19828629]
16. Lorenzen IDA, Jacques Y, Grötzinger J. The structure of the interleukin-15 alpha receptor and its implications for ligand binding. *J Biol Chem*. 2006; 281 (10):6642. [PubMed: 16377614]
17. Grande J, Matas A, Kasiske B, et al. Reproducibility of the Banff Classification Scheme in Patients with Chronic Allograft Dysfunction. *American Society of Nephrology*; Philadelphia: 2008.
18. Van Ness B, Ramos C, Haznadar M, et al. Genomic variation in myeloma: design, content, and initial application of the Bank On A Cure SNP Panel to detect associations with progression-free survival. *BMC Med*. 2008; 6(26)
19. Hardenbol PBJ, Jain M, Nilsson M, Namsaraev EA, Karlin-Neumann GA, et al. Multiplexed genotyping with sequence-tagged molecular inversion probes. *Nat Biotechnol*. 2003; 21 (6):673. [PubMed: 12730666]

20. Grambsch P, Therneau T. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*. 1994; 81:515.
21. Cox DR. Regression models and life-tables. *Journal of the Royal Statistics Society*. 1972; 34:187.
22. Hall MFE, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA Data Mining Software: An Update. *SIGKDD Explorations*. 2009; 11 (1):10.
23. Hall, M. Correlation-based feature selection for discrete and numeric class machine learning. *Proceedings of the 17th International Conference on Machine Learning*; San Francisco: Morgan Kaufman; 2000. p. 359
24. Cohen, WW. Fast effective rule induction. *Proceedings of the 12th International Conference on Machine Learning*; Tahoe City, CA: Morgan Kaufmann; 1995. p. 115
25. Breiman, LFJ.; Olshen, RA.; Stone, CJ. *Classification and regression trees*. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software; 1984.
26. Quinlan, R. *C4.5: Programs for Machine Learning*. San Mateo: Morgan Kaufmann Publishers; 1993.

DeKAF Investigators

J. Michael Cecka, M.D., UCLA Immunogenetics Center, Los Angeles, CA 90095, Email: mcecka@ucla.edu

John Connett, Ph.D., Division of Biostatistics. University of Minnesota, Minneapolis, MN 55455, Email: john-c@biostat.umn.edu

Fernando G. Cosio, M.D., Division of Nephrology, Mayo Clinic, Rochester, MN 55905, Email: Cosio.Fernando@mayo.edu

Robert Gaston, M.D., University of Alabama, Division of Nephrology, Birmingham, AL 35294-0006, Email: rgaston@uab.edu

Sita Gourishankar M.D., Division of Nephrology and Immunology, University of Alberta, Edmonton, Alberta, Canada, Email: sitag@ualberta.ca

Joseph P. Grande, M.D., Ph.D., Mayo Clinic College of Medicine, Rochester MN 55905, Email: Grande.Joseph@mayo.edu

Lawrence Hunsicker, M.D., Nephrology Division, Iowa City, IA 52242-1082, Email: lawrence-hunsicker@uiowa.edu

Bertram Kasiske, M.D., Department of Medicine, Hennepin County Medical Center and the University of Minnesota, Minneapolis, MN 55415, Email: kasis001@umn.edu

Rosalyn Mannon, University of Alabama, Division of Nephrology, Birmingham, AL 35294-0006, Email: rmannon@uab.edu

David Rush, M.D., Health Sciences Center, Winnipeg MB, Canada, Email: drush@exchange.hsc.mb.ca

Table 1

Characteristics of kidney transplant recipients at each of the six transplant centers (n=990). These six transplant centers were University of Minnesota, Hennepin County Medical Center (Minneapolis, Minnesota), Mayo Clinic (Rochester, Minnesota), University of Iowa, University of Alabama and University of Alberta.

Characteristic	n= 990	p-value
Ethnicity: White	77 %	<0.0001
Black	17 %	
Asian	3 %	
Other	2 %	
Not Known	<1 %	
Hispanic	2 %	0.35
Male	62 (%)	0.86
Mean age at enrollment	49 ±14 yrs	0.001
Cause of End Stage Renal		<0.0001
Disease:		
Diabetes	31 %	
Glomerular disease	20 %	
HTN	12 %	
Polycystic kidney dz	12 %	
Other	21 %	
Unknown	4%	
Living donor transplant	59 %	<0.0001
Living Donor: Black *	4 %	0.01
Mean donor age in years	40 ±14	<0.0001
Male donor *	46 %	0.13
Cold Ischemia time >24 h *	7 %	<0.0001
Prior kidney Transplant	14 %	0.0003
Allograft function in first week post transplant *		<0.0001
Delayed graft function	8 %	
Slow graft function	13%	
No Slow/delayed function	79%	
Final PRA present *	36%	<0.0001
T or B Crossmatch positive *	5 %	<0.0001
Plasmapheresis prior to tx	3 %	<0.0001
Zero HLA mismatches	12 %	0.04
Antibody Induction:		<0.0001
IL-2 blockers		
Monoclonal		
None		
Other		
Polyclonal		

Characteristic	n= 990	p-value
Smoking status: *		0.0017
Never	60%	
Past	31%	
Current	9%	
Pre-emptive transplant	31 %	<0.0001
Steroid withdrawal by day 14 post-transplant	48%	<0.0001
CNI type:		<0.0001
Cyclosporine	35%	
Tacrolimus	62%	
None	3%	
SPK	6%	<0.0001
Prior Non-kidney Tx *	11%	<0.0001
CMV Recipient/Donor Status *		0.01
Recipient (-)/ Donor (-)	22%	
Recipient (+)/ Donor (+)	62%	
Recipient (-)/ Donor (+)	16%	

* Missing data: Living donor race status missing in 176 subjects, Living donor gender missing in 2 subjects, Delayed/slow allograft function missing in 92 subjects, Cold Ischemia time missing in 111 subjects, Final PRA missing in 5 subjects, B/T cell crossmatch missing in 37 subjects, Smoking status missing in 1 subject, Prior non-kidney transplant missing in 44 subjects, CMV recipient/ donor status missing in 44 subjects.

Table 2

Simple model results with top 15 SNPs (ranked by p-value) associated with time to acute rejection, adjusted by recipient race and stratified by transplant center. All variant data from dbSNP Build 131

SNP	Gene	Variant	H.R.#	95 % C.I.	p-value	Minor Allele in AA *	MAF ⁺ in AA	Minor Allele in non-AA	MAF ⁺ in non-AA
Rs2227931	ATR	coding-synon	1.45	(1.17–1.80)	7.43E-04	C	0.19	C	0.38
Rs2267130	CHEK2	Intron	0.69	(0.56–0.86)	8.96E-04	C	0.09	C	0.46
Rs3088142	DUSP13	missense,	1.41	(1.15–1.73)	9.46E-04	C	0.29	T	0.44
Rs2017662	TNFRSF17	coding-synon	1.77	(1.26–2.48)	9.76E-04	T	0.24	T	0.05
Rs3743591	TNFRSF17	untranslated-5	1.76	(1.25–2.47)	1.09E-03	G	0.24	G	0.05
Rs4839469	VANGL1	Missense	1.53	(1.17–1.99)	1.84E-03	A	0.07	A	0.14
Rs1800457	CYB5R3	Missense	2.79	(1.42–5.45)	2.78E-03	G	0.27	G	0
Rs2229032	ATR	Missense	0.61	(0.43–0.85)	3.41E-03	A	0.09	A	0.16
Rs196912	ERN1	coding-synon	1.4	(1.11–1.76)	3.99E-03	C	0.34	T	0.23
Rs524	PPP1R15A	coding-synon	0.69	(0.53–0.89)	4.10E-03	G	0.47	A	0.25
Rs3093816	CCNH	Intron	1.35	(1.10–1.65)	4.45E-03	C	0.25	C	0.41
Rs2227929	ATR	coding-synon	1.37	(1.10–1.70)	4.97E-03	C	0.19	C	0.38
Rs500079	PPP1R15A	Missense	0.67	(0.51–0.89)	5.22E-03	T	0.45	C	0.26
Rs1042858	RRM1	coding-synon	1.53	(1.13–2.06)	5.36E-03	G	0.14	G	0.08
Rs4253199	ERCC6	Missense	3.08	(1.39–6.84)	5.68E-03	T	0.06	0	0

(Abbreviations: AA= African Americans, MAF= Minor Allele Frequency, synon=synonymous, untranslated-5= variant in the untranslated 5' end of gene)

Analysis conducted using a Cox proportional hazards model

** There was linkage disequilibrium (LD) between the two SNPs in ATR namely rs2227931 and rs 2227929 with $r^2 > 0.95$ in both AA and non-AAs. Similarly, there was LD between the two SNPs in TNFRSF17 of >0.95 in both AAs and non-AA. There was LD between the two SNPs in PPP1R15A, with $r^2 = 0.93$ in AA and $r^2 = 0.97$ in non-AA. The remaining SNPs in the table had an $r^2 < 0.2$.

Table 3A

Multivariate model results with top 15 SNPs (ranked by p-value) associated with time to acute rejection, stratified by transplant center. Model also adjusted for recipient factors (race, age, gender, weight), PRA presence, number of HLA mismatches, T or B-cell cross-match positive, antibody induction, type of calcineurin inhibitor used, steroid use at day 14 post-transplant, simultaneous kidney-pancreas transplant (versus kidney transplant alone), cause of ESRD, living donor (versus deceased donor) and donor age. All variant data from dbSNP Build 131

SNP	Gene	Variant	HR [#]	95% C.I.	p-value	Minor Allele in AA [*]	MAF ⁺ in AA	Minor Allele in Non-AA	MAF ⁺ in non-AA
rs2227931	ATR	coding-synon	1.51	[1.20–1.90]	4.09E-04	C	0.19	C	0.38
rs2267130 [*]	CHEK2	Intron	0.67	[0.53–0.84]	5.84E-04	C	0.09	C	0.46
rs4839469 [*]	VANGU1	Missense	1.62	[1.23–2.13]	6.10E-04	A	0.07	A	0.14
rs3088142	DUSP13	Missense	1.43	[1.15–1.78]	1.17E-03	C	0.29	T	0.44
rs2229032	ATR	Missense	0.55	[0.38–0.80]	1.75E-03	A	0.09	A	0.16
rs2228224 [*]	GLI1	Missense	0.69	[0.54–0.88]	2.45E-03	A	0.23	G	0.41
rs2227929 [*]	ATR	coding-synon	1.43	[1.13–1.80]	2.74E-03	C	0.19	C	0.38
rs3783408	MAP4K5	nearGene-5	1.47	[1.13–1.90]	3.49E-03	A	0.08	A	0.26
rs4348159 [*]	UGT2B7	coding-synon	0.47	[0.28–0.78]	3.62E-03	T	0.28	T	0.09
rs4253199	ERCC6	Missense	3.77	[1.54–9.26]	3.76E-03	T	0.06	0	0
rs6163	CYP17A1	coding-synon	0.72	[0.57–0.90]	3.76E-03	A	0.38	A	0.42
rs743572	CYP17A1	coding-synon	0.72	[0.57–0.90]	3.80E-03	G	0.38	G	0.42
rs2072651	TPP1	Intron	1.49	[1.13–1.96]	4.50E-03	T	0.2	T	0.17
rs3125001	NOTCH1	Intron	0.65	[0.48–0.88]	4.75E-03	C	0.33	T	0.39
rs3093816	CCNH	Intron	1.38	[1.10–1.72]	4.84E-03	C	0.25	C	0.41

(Abbreviations: AA= African Americans, MAF= Minor Allele Frequency, synon=synonymous, untranslated-5= variant in the untranslated 5' end of gene, near-gene-5= variation within 2000 bases of 5' end of gene)

[#] Analysis conducted using a Cox proportional hazards model.

^{*} Same SNPs seen in the top 15 SNPs associated with cellular acute rejection.

^{**} There was some linkage disequilibrium (LD) between the two SNPs in CYP17A1, with $r^2 = 1.00$ in non-African Americans and African Americans, separately. There was also some LD between the two SNPs in ATR, with $r^2 > 0.95$. The remaining SNPs in the table had an $r^2 < 0.2$.

Table 3B

SNPs identified using knowledge discovery methods. The SNPs identified by three correlational-based feature selection methods, namely, best-first search, JRIP and decision trees* and a 10-fold cross-validation#. The SNPs are shown with their specific ranking, from Table 3A and Supplemental Table 3, in the multivariate model.

Rank	SNP	Gene	Variant	HR	95% CI	P-Value
60	rs6271	DBH	Missense	1.67	[1.07–2.61]	2.41E-02
199	rs3093106	CYP4F2	coding-synon	1.26	[0.97–1.63]	8.15E-02
378	rs2231924	FLJ10213	Missense	0.84	[0.67–1.06]	1.48E-01
405	rs7588635	MERTK	Missense	1.61	[0.83–3.14]	1.62E-01
667	rs2295155	CARD10	Intron	1.21	[0.86–1.72]	2.70E-01
964	rs1822017	DCBLD2	Missense	1.15	[0.84–1.58]	3.89E-01
1073	rs1863703	STK36	Missense	1.16	[0.80–1.70]	4.28E-01
1116	rs2066471	MTHFR	Intron	0.90	[0.68–1.18]	4.46E-01
1259	rs1801018	BCL2	coding-synon	0.92	[0.73–1.17]	4.95E-01
1475	rs162555	CYP1B1	Intron	1.09	[0.82–1.45]	5.69E-01
1773	rs1887994	ESR2	Intron	1.09	[0.74–1.60]	6.68E-01
2012	rs2072052	CDK4	nearGene-5	0.96	[0.75–1.23]	7.51E-01
2554	rs1800471	TGFB1	Missense	1.02	[0.65–1.60]	9.41E-01
2656	rs1043424	PINK1	Missense	1.00	[0.78–1.29]	9.77E-01

(Abbreviations: synon=synonymous, near-gene-5= variation within 2000 bases of 5' end of gene)

* = A final SNP set was created combining the SNPs found by the three knowledge discovery methods. These three methods included the best-first search procedure.(23) A second knowledge discovery method utilized a rule discovery tool named “JRip” which is based on a propositional rule learner, Repeated Incremental Pruning to Produce Error Reduction, or RIPPER.(24) The SNPs in rules that classified for presence of AR were extracted into the final feature set. For the third knowledge discovery method, decision trees were constructed with J48, a tool that is analogous to C4.5.(26) and classification and regression tree analysis.(25) The SNPs (nodes) of pathways that resulted in positive outcome (presence of AR) were extracted to the final feature set. Duplicates were removed from the final set. Then using a Cox proportional hazards model, we determined the association with AR for these SNPs, across all the transplant centers in the study and in both African Americans and non-African Americans.

= Each of the 3 knowledge discovery methods were conducted with 10-fold cross-validation. In the 10-fold cross-validation, the dataset was divided into 10 slices, folds, each of which preserves the distribution of the outcome of AR. There were a total of 10 runs, in which a fold was removed from the dataset and reserved for testing, and the remaining nine folds are used for training, or discovering hypotheses (rules, trees, or features) that may explain AR. After the system was optimized on training, the resulting hypotheses were tested against the data in the testing set. The results (sensitivity, specificity, predictive values) were calculated for each fold, and then averaged over the 10 runs. The sensitivity across the runs ranged between 85 and 95%. We evaluated the resulting sets of features, rules, or trees on separate validation sets and found minimal evidence of overfitting, in that the sensitivity remained high on validation.

Table 4

SNPs associated with severity of AR using a multinomial logistic regression with 3 outcome groups, no biopsy (n= 610), t₁ (n= 271) and t₂ (n = 83) groups. The top 15 SNPs are ranked by p-values for the t₂ versus t₁ groups. The model was adjusted by transplant center, recipient race, recipient age, recipient gender, PRA (positive or negative), number of HLA mismatches, donor age, type of calcineurin inhibitor use at baseline, recipient weight at time of transplant, antibody induction, and smoking status. All variant data from dbSNP Build 131

SNP	Gene	Variant	Allele	t ₂ vs. t ₁	
				OR	P-value
rs811925	PRDM1	Missense	C	2.41 [1.51,3.84]	2.27E-04
rs2228059	IL15RA	Missense	A	2.00 [1.36,2.93]	4.14E-04
rs2283512	ABCC1	Intron	T	1.91 [1.29,2.82]	1.26E-03
rs619824	CYP17A1	Intron	T	0.54 [0.37,0.79]	1.50E-03
rs743572	CYP17A1	untranslated-5	G	0.55 [0.38,0.81]	2.56E-03
rs6162	CYP17A1	coding-synon	A	0.56 [0.39,0.82]	2.81E-03
rs6163	CYP17A1	coding-synon	A	0.56 [0.38,0.82]	3.07E-03
rs2664538	MMP9	Missense	G	0.54 [0.36,0.81]	3.10E-03
rs880324	NFATC2	Intron	A	0.48 [0.29,0.79]	3.49E-03
rs3088440	C9orf53	untranslated-3	A	2.21 [1.30,3.77]	3.59E-03
rs275652	AGTR1	near-gene-5	C	1.97 [1.25,3.12]	3.75E-03
rs2193587	DGKG	Missense	G	1.83 [1.19,2.82]	5.66E-03
rs12720356	TYK2	Missense	G	2.23 [1.25,3.96]	6.47E-03
rs2267668	PPARD	Intron	G	0.41 [0.22,0.78]	6.58E-03
rs1137282	KRAS	untranslated-3	C	1.79 [1.17,2.72]	7.16E-03

(Abbreviations: synon=synonymous, untranslated-5= variant in the untranslated 5' end of gene, untranslated-3= variant in the untranslated 3' end of gene, near-gene-5= variation within 2000 bases of 5' end of gene)

= Analysis conducted using a multinomial logistic regression model.

*** There was some linkage disequilibrium between the four SNPs in CYP17A1, with $r^2 = 0.74-1.00$ in non-African Americans and $0.30-1.00$ in African Americans. The remaining SNPs in the table had an $r^2 < 0.2$.