

OPEN

# An Integrated Clinical and Genetic Prediction Model for Tacrolimus Levels in Pediatric Solid Organ Transplant Recipients

Sandar Min, MBBS, MPH,<sup>1</sup> Tanya Papaz, HBA,<sup>1</sup> A. Nicole Lambert, MD,<sup>2</sup> Upton Allen, MBBS,<sup>3,4</sup> Patricia Birk, MD,<sup>5</sup> Tom Blydt-Hansen, MD,<sup>6</sup> Bethany J. Foster, MD,<sup>7</sup> Hartmut Grasmann, MD,<sup>3,4</sup> Lorraine Hamiwka, MD,<sup>8</sup> Catherine Litalien, MD,<sup>9</sup> Vicky Ng, MD,<sup>3,4</sup> Nouredine Berka, PhD,<sup>10</sup> Patricia Campbell, MD,<sup>11</sup> Claude Daniel, PhD,<sup>12</sup> Chee Loong Saw, PhD,<sup>13</sup> Kathryn Tinkam, MD,<sup>14</sup> Simon Urschel, MD,<sup>15</sup> Sara L. Van Driest, MD, PhD,<sup>2</sup> Rulan Parekh, MD, MS,<sup>3,4,16</sup> and Seema Mital, MD<sup>1,3,4</sup>

**Background.** There are challenges in achieving and maintaining therapeutic tacrolimus levels after solid organ transplantation (SOT). The purpose of this genome-wide association study was to generate an integrated clinical and genetic prediction model for tacrolimus levels in pediatric SOT. **Methods.** In a multicenter prospective observational cohort study (2015–2018), children <18 years old at their first SOT receiving tacrolimus as maintenance immunosuppression were included (455 as discovery cohort; 322 as validation cohort). Genotyping was performed using a genome-wide single nucleotide polymorphism (SNP) array and analyzed for association with tacrolimus trough levels during 1-y follow-up. **Results.** Genome-wide association study adjusted for clinical factors identified 25 SNPs associated with tacrolimus levels; 8 were significant at a genome-wide level ( $P < 1.025 \times 10^{-7}$ ). Nineteen SNPs were replicated in the validation cohort. After removing SNPs in strong linkage disequilibrium, 14 SNPs remained independently associated with tacrolimus levels. Both traditional and machine learning approaches selected organ type, age at transplant, rs776746, rs12333983, and rs12957142 SNPs as the top predictor variables for dose-adjusted 36- to 48-h posttacrolimus initiation (T1) levels. There was a significant interaction between age and organ type with rs776476\*1 SNP ( $P < 0.05$ ). The combined clinical and genetic model had lower prediction error and explained 30% of the variation in dose-adjusted T1 levels compared with 18% by the clinical and 12% by the genetic only model. **Conclusions.** Our study highlights the importance of incorporating age, organ type, and genotype in predicting tacrolimus levels and lays the groundwork for developing an individualized age and organ-specific genotype-guided tacrolimus dosing algorithm.

(*Transplantation* 2022;106: 597–606).

Received 21 August 2020. Revision received 1 December 2020.

Accepted 2 January 2021.

<sup>1</sup> Genetics and Genome Biology Program, Hospital for Sick Children, Toronto, ON, Canada.

<sup>2</sup> Department of Pediatrics, Vanderbilt University Medical Center, Nashville, TN.

<sup>3</sup> Department of Pediatrics, Hospital for Sick Children, University of Toronto, Toronto, ON, Canada.

<sup>4</sup> Department of Pediatrics, Transplant and Regenerative Medicine Centre, Hospital for Sick Children, Toronto, ON, Canada.

<sup>5</sup> Department of Pediatrics and Child Health, Health Sciences Centre Winnipeg, Winnipeg, MB, Canada.

<sup>6</sup> Division of Pediatric Nephrology, University of British Columbia, Vancouver, BC, Canada.

<sup>7</sup> Division of Nephrology, Montreal Children's Hospital, McGill University Health Centre, Montreal, QC, Canada.

<sup>8</sup> Division of Pediatric Nephrology, Alberta Children's Hospital, University of Calgary, Calgary, AB, Canada.

<sup>9</sup> Division of General Pediatrics, Department of Pediatrics, Centre Hospitalier Universitaire Sainte-Justine, Montreal, QC, Canada.

<sup>10</sup> Department of Pathology & Laboratory Medicine, University of Calgary, Calgary, AB, Canada.

<sup>11</sup> Department of Laboratory Medicine and Pathology, University of Alberta, Edmonton, AB, Canada.

<sup>12</sup> INRS - Centre Armand-Frappier Santé Biotechnologie, Laval, QC, Canada.

<sup>13</sup> Division of Hematology, McGill University Health Centre, Montreal, QC, Canada.

<sup>14</sup> Departments of Medicine and Laboratory Medicine & Pathobiology, University of Toronto, University Health Network, Toronto, ON, Canada.

<sup>15</sup> Division of Pediatric Cardiology, University of Alberta, Edmonton, AB, Canada.

<sup>16</sup> Division of Nephrology, Department of Medicine, University Health Network and University of Toronto, Toronto, ON, Canada.

S.Mital, B.J.F., U.A., S.U., C.L., T.B.-H., P.B., L.H., R.P., H.G., and V.N. participated in the conception or design of the work. S.Min, T.P., A.N.L., S.L.V.D., C.L.S., C.D., K.T., N.B., P.C. did the data collection. S.Min and S.Mital did data analysis, interpretation, and drafting of the article. All authors did critical review and final approval of the article.

The authors declare no conflicts of interest.

This study is part of the Canadian National Transplant Research Program (CNTRP) funded by the Canadian Institutes for Health Research (CIHR) and partners (grant number TFU 127880); Astellas Pharma Canada, Inc.; and SickKids Transplant and Regenerative Medicine Centre. S.Mital is supported by the Heart and Stroke Foundation of Canada Chair. R.P. is supported by the Canada Research Chairs program. S.L.V.D. is supported by IRSA 1015006 from the Burroughs Wellcome Fund. A.N.L. is supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health (award number T32HL105334). The Vanderbilt University Medical Center's BioVU sample and data repository is supported by Vanderbilt University Medical Center, the NIH funded Shared Instrumentation Grant S10RR025141, and Clinical and Translational Science Awards grants UL1TR002243, UL1TR000445, and UL1RR024975.

Supplemental Visual Abstract; <http://links.lww.com/TP/C153>.

Supplemental digital content (SDC) is available for this article. Direct URL citations appear in the printed text, and links to the digital files are provided in the HTML text of this article on the journal's Web site ([www.transplantjournal.com](http://www.transplantjournal.com)).

## INTRODUCTION

Tacrolimus, a calcineurin inhibitor, is one of the most commonly used immunosuppressive medications after solid organ transplantation. Therapeutic drug monitoring is required to maintain blood levels within a target range. There are significant challenges in achieving and maintaining therapeutic drug concentrations in pediatric transplant recipients, related in part to developmental changes in drug metabolism.<sup>1</sup>

Tacrolimus is almost completely metabolized through the cytochrome P450 enzymes, *CYP3A5* and *CYP3A4*, in the liver and to a lesser extent in enterocytes.<sup>2</sup> First steady state concentration is usually achieved at 36–48 hours after initiation of tacrolimus.<sup>3</sup> Delay in achieving stable therapeutic levels in the early transplant period is associated with increased risk of developing donor-specific antibodies, rejection, and graft loss.<sup>4,5</sup> Single nucleotide polymorphisms (SNPs) in *CYP3A5* and *CYP3A4* genes are important contributors to the variation in tacrolimus levels and dosage requirements.<sup>6,7</sup> The Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines<sup>8</sup> recommend adjusting tacrolimus starting dosage based on recipient *CYP3A5* \*3, \*6, and \*7 genotype status but fail to take into account a potential contribution of other SNPs in the recipient (and donor), or of clinical factors like age, sex, organ type, concomitant interacting medications like *CYP3A4* inducers/inhibitors which also contribute to variability in tacrolimus levels and dosing requirements.<sup>9,10</sup> Additionally, most studies are limited to adult kidney transplant recipients and do not explore differences by organ type.<sup>11,12</sup>

The purpose of this study was to identify recipient and donor genotypes associated with tacrolimus levels across all organs and develop a prediction model for dose-adjusted tacrolimus levels that incorporate both genetic and clinical factors in a large cohort of pediatric transplant recipients.

## MATERIALS AND METHODS

### Study Cohort

Children <18 years old undergoing their first heart, liver, kidney, or lung transplant were recruited from 2015 to 2018 as part of a multicenter observational cohort study involving 7 Canadian pediatric transplant centers (POSITIVE cohort). Details of the study design have been previously published.<sup>13</sup> DNA samples from eligible pediatric transplant recipients in BioVU, a biorepository of deidentified patient samples at the Vanderbilt University Medical Center, were used for external validation.<sup>14</sup> Patients who received tacrolimus immunosuppression after transplant were eligible for this analysis. Retransplants and multiple

organ transplants were excluded. Written informed consent was obtained from participants and their parents or legal guardians. Living donors were consented for blood or saliva for DNA. For deceased donors, anonymized archived DNA samples were obtained from the local HLA laboratory after verification of donor consent for research by the local organ donor procurement organization. The study was approved by the Institutional Research Ethics Board at each center.

### Data Collection

Pretransplant data, including demographics, dates of listing and transplant, primary diagnosis, height, weight, blood type, and donor information, were captured from medical records and an enrollment questionnaire. Posttransplant data included laboratory, medication and outcomes at 6 time points: 36–48 hours posttacrolimus initiation, 7±3 days, 14±3 days, 30±3 days, 3±1 months, and 12±3 months posttransplant. Tacrolimus trough levels were measured using the liquid chromatography-tandem mass spectrometry (LC-MS-MS) assay in accredited institutional Therapeutic Drug Monitoring laboratories except for Montreal Children's Hospital, which used the Abbott ARCHITECT assay. Target therapeutic ranges for tacrolimus levels by organ are described in Table S1 (SDC, <http://links.lww.com/TP/C152>). All patients received maintenance immunosuppression that included mycophenolate mofetil in addition to tacrolimus as per standard-of-care. Induction immunosuppression varied by organ type and center and included a combination of steroids, antithymocyte globulin (ATG) in heart recipients, and basiliximab in kidney recipients (ATG in selected cases). Only selected liver recipients (ATG or basiliximab) and selected infant lung recipients (ATG) received induction immunosuppression at transplant.

### Genotyping

DNA from recipients and donors was genotyped using the Axiom Transplant Genotyping Array (ThermoFisher, Carlsbad, CA) developed by the iGeneTRAiN international consortium.<sup>15</sup> This array has ~782 000 markers enriched for transplant-relevant SNPs. IMPUTE2 software<sup>16</sup> with 1000 genomes as its reference was used to impute missing genotypes for 752 011 SNPs on the array. Four lakh eighty-seven thousand seven hundred ninety-five autosomal variants meeting SNP call rate >90%, minor allele frequency (MAF) >0.01 and Hardy-Weinberg exact test  $P > 10^{-6}$  were included in the analysis.

### Statistical Analysis

Continuous variables were reported as mean (SD) or median (interquartile range), and categorical variables were reported as frequencies and proportions. Means were compared using Student *t*-test, medians were compared using Wilcoxon rank-sum test and proportions were compared with Pearson's  $\chi^2$  or Fisher's exact test. All statistical tests were performed using R software version 3.4 ([www.r-project.org](http://www.r-project.org)) and Stata software version 16.

### GWAS for SNPs Associated With Tacrolimus Levels

Clinical factors associated with tacrolimus levels in the discovery cohort were first identified using a multivariable

Correspondence: Seema Mital, MD, Hospital for Sick Children, 555 University Ave, Toronto, ON M5G 1X8, Canada. ([seema.mital@sickkids.ca](mailto:seema.mital@sickkids.ca)).

Copyright © 2021 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

ISSN: 0041-1337/20/1063-597

DOI: 10.1097/TP.00000000000003700

linear-mixed effects model with backward deletion with retention  $P < 0.1$ . Next, to adjust for population stratification based on ethnicity, a Principal Component Analysis of ancestry markers mapped to the HapMap3 population was performed in the discovery cohort. First and second principal components (PC1 and PC2) explaining the highest genetic variation due to ancestry were included as covariates in the statistical models. To identify SNPs associated with tacrolimus levels, genome-wide analysis was conducted with adjustment for clinical factors and PC1 and PC2 in the discovery cohort with genome-wide significance defined as a  $P < 1.025 \times 10^{-7}$  based on Bonferroni correction for multiple testing.<sup>17</sup> For initial discovery, SNPs that met a more liberal  $P < 1 \times 10^{-5}$  were included for downstream analyses.

### GWAS Power Calculations

Kelly, Stallard, and Whittaker method<sup>18</sup> was used to calculate the sample size required to achieve a prespecified power for the genome-wide association study (GWAS) given a type I error rate. Based on MAF range for significant SNPs of 0.07–0.35, a sample size of 270 was sufficient to fulfill 80% of power with 0.05 type I error rate to observe an additive SNP effect with  $\beta$ -coefficient 1.74.

### Redundancy Analysis

Redundant SNPs in strong linkage disequilibrium (LD) were excluded by constructing an LD plot that included the significant SNPs from GWAS as well as 3 SNPs deemed clinically relevant from previous studies (*CYP3A4*\*22, *CYP3A5*\*6, and *CYP3A5*\*7).<sup>6,7</sup> SNPs that were highly correlated ( $r^2 > 0.75$ ) were grouped into haplotype blocks. Additional redundancy analysis was conducted to remove highly correlated SNPs within each block.<sup>19</sup> *Subgroup analyses* were performed using linear-mixed models for association of independent SNPs with tacrolimus levels stratified by organ type.

### Interaction Between Age and Organ Type With rs776746\*1 SNP on Dose-adjusted T1 Levels

To determine if the association of genotype with tacrolimus levels was influenced by age and organ type, linear regression analysis was performed in the combined discovery and validation cohorts to determine interaction effect between (i) age and (ii) organ type, with the top-ranked SNP (rs776746) for log-transformed dose-adjusted first tacrolimus trough levels after drug initiation (T1), that is, first trough level 36–48 hours after tacrolimus initiation divided by total daily starting dose. Carriers of rs776746\*1 SNP were defined as *CYP3A5* expressors and noncarriers were defined as *CYP3A5* nonexpressors. Patients were categorized into 3 age groups—infant (<2 y), child (2–10 y), and adolescent ( $\geq 11$  y) since infancy and adolescence are the periods characterized by marked physiological changes secondary to developmental maturation, growth, and hormonal influences that can independently influence drug disposition. The interaction analysis was adjusted for organ type. Similarly, to test interaction between organ type and SNP, the analysis was adjusted for age. Lung transplants were excluded from the interaction analysis due to small numbers. For patients in whom both recipient and donor genotypes were available, the association of donor genotype with tacrolimus levels was analyzed.

### Prediction Model for Dose-adjusted Tacrolimus Levels

The final set of independent (nonredundant) SNPs was used for model development. Given the importance of achieving therapeutic tacrolimus levels early after transplant when variability is highest, we built a prediction model for dose-adjusted T1 levels. Since a machine learning approach may provide better prediction accuracy than traditional approaches,<sup>20</sup> we first used least absolute shrinkage and selection operator (lasso) model to select variables that predicted log-transformed dose-adjusted T1 levels (level/dose). The tuning parameter  $\lambda$ ) was selected using 10-fold cross-validation. The model with minimum out of sample mean squared error was selected. Next, we used linear regression with stepwise backward deletion to determine the significant predictors of log-transformed dose-adjusted T1 levels in the discovery cohort and compared the performance of linear and lasso models in the cohort. Three linear regression models were developed using (1) clinical predictors only, (2) SNPs only, and (3) both clinical predictors and SNPs. All models were tested in the validation cohort, and prediction accuracy of each model was compared using the mean squared error and coefficient of determination,  $r^2$ . Finally, we combined the discovery and validation cohorts to compare model performance across different organs using linear regression analysis.

## RESULTS

Overall, 897 transplant recipients in the discovery and validation cohorts were genotyped; 88 patients with incomplete data and 32 recipients where genotyping did not meet quality control criteria were excluded. The first 455 eligible transplant recipients (2375 levels) with at least 2 available tacrolimus levels during follow-up recruited into POSITIVE were used for GWAS discovery and prediction model development, 322 recipients (213 from POSITIVE, 109 from BioVU) (1274 levels) were used for model validation, and 146 matched recipient-donor pairs were used for donor genotype association analysis. The characteristics of the discovery and validation cohorts are shown in Table 1. There were only minor differences in ethnicity and organ transplant distribution between the validation and discovery cohorts with more kidney transplants in validation compared with discovery cohort. Figure 1A shows the variability in tacrolimus trough levels during 1-y posttransplant follow-up in the discovery cohort ( $n = 2375$  tacrolimus levels) at 6 posttransplant time points: 36–48 hours post-Tac initiation ( $n = 398$ ), 7 days ( $n = 410$ ), 14 days ( $n = 409$ ), 30 days ( $n = 403$ ), 3 months ( $n = 398$ ), and 12 months posttransplant ( $n = 357$ ). Overall, 13% of levels were missing. Forty-four percent of tacrolimus levels were in the presence of *CYP3A4* inhibitors with the 2 most frequent *CYP3A4* inhibitors being amlodipine in 50% and fluconazole in 7% patients (Figure S1, SDC, <http://links.lww.com/TP/C227>). A multivariable mixed effects model identified older age, organ type (higher levels in heart, liver, lung versus kidney), tacrolimus dosage, and concomitant *CYP3A4* inhibitor use as factors associated with higher tacrolimus levels (Figure 1B). Higher levels in heart recipients were seen despite similar target concentrations as kidney recipients; higher levels in liver recipients occurred despite similar target concentrations as kidney recipients after 3 months posttransplant.

**TABLE 1.**  
Characteristics of transplant recipients

	Discovery (N = 455)	Validation (N = 322)	P
	n (%)	n (%)	
Median age at transplant (IQR) (y)	4.5 (1.0–11.7)	6.1 (2.0–13.1)	0.003
Sex: male	247 (54%)	176 (55%)	0.918
Self-reported race			
White	287 (63%)	249 (77%)	<0.001
Black	18 (4%)	19 (6%)	
Aboriginal	14 (3%)	12 (4%)	
Asian	89 (20%)	30 (9%)	
Mixed	40 (9%)	11 (3%)	
Unknown	7 (2%)	1 (0.3%)	
Organ type			
Kidney	133 (29%)	151 (47%)	<0.001
Heart	151 (33%)	80 (25%)	
Liver	161 (35%)	85 (26%)	
Lung	10 (2%)	6 (2%)	
Donor type			
Deceased	301 (66%)	209 (65%)	0.718

IQR, interquartile range.

### GWAS SNPs Associated With Tacrolimus Levels

Figure S2 (SDC, <http://links.lww.com/TP/C228>) shows the Principal Component Analysis plot of ancestry markers in the discovery cohort mapped to the HapMap3 population. After adjustment for age, gender, PC1, PC2, organ type, and concomitant CYP3A4 inhibitor use, 25 SNPs were significantly associated with tacrolimus trough levels ( $P < 10^{-3}$ ) in the discovery cohort, of which 8 were significant at a genome-wide level ( $P < 1.025 \times 10^{-7}$ ) (Table S2, SDC, <http://links.lww.com/TP/C152>). The MAF of significant SNPs ranged from 0.07 to 0.35. Figure 2A shows the GWAS Manhattan plot with 21 of the 25 significant SNPs residing on chromosome 7; 14 mapped to pharmacogenes—CYP3A4, CYP3A5, CYP3A7, and CYP3A7-CYP3A1. Figure 2B shows a quantile-quantile (Q-Q) plot of the observed versus expected  $P$  for SNP association. The most strongly associated SNP was the previously known rs776746 (CYP3A5) ( $P = 9.71 \times 10^{-13}$ ). Nineteen of these 25 SNPs from the discovery cohort were also significant in the POSITIVE validation cohort (Table S2, SDC, <http://links.lww.com/TP/C152>).

We constructed an LD plot of all 25 SNPs, and, using a correlation coefficient threshold of  $r^2 > 0.75$ , we identified 4 haplotype blocks (Figure S3, SDC, <http://links.lww.com/TP/C228>). After excluding redundant SNPs (ie, SNPs that were in linear combination) in each haplotype group, 14 independent SNPs were retained for prediction modeling (Figure 2C). Of these 14 SNPs, rs72816873, rs12957142, and rs35154575 were associated with higher tacrolimus levels and the remainder with lower levels. Seven SNPs mapped to pharmacogenes—CYP3A4, CYP3A5, CYP3A4-CYP3A7, CYP3A7-CYP3A1, and CYP3A5-ZKSCAN5. The others mapped to ARPC1B-ARPC1A, VPS35-ORC6, MN1-LINC01422, KC6-LOC101927900, ARPC1A, ZC3H13-SIAH3, and MYH16 genes.

### Subgroup Analyses

Subgroup analysis of the 14 independent SNPs in the discovery cohort by organ type showed that all 14 SNPs were significantly associated with tacrolimus levels in heart, 11 SNPs were significant in kidney, and 6 were significant in liver recipients ( $P < 0.05$ ) (Table S2, SDC, <http://links.lww.com/TP/C152>). Because of the small number, we did not perform subgroup analysis in lung transplant recipients ( $n = 16$ ). Subgroup analysis of 146 recipients with available donor genotypes (63 kidney, 67 heart, 10 liver, 6 lung) showed no association of donor genotype for the 14 independent SNPs with tacrolimus levels in the liver and nonliver recipients.

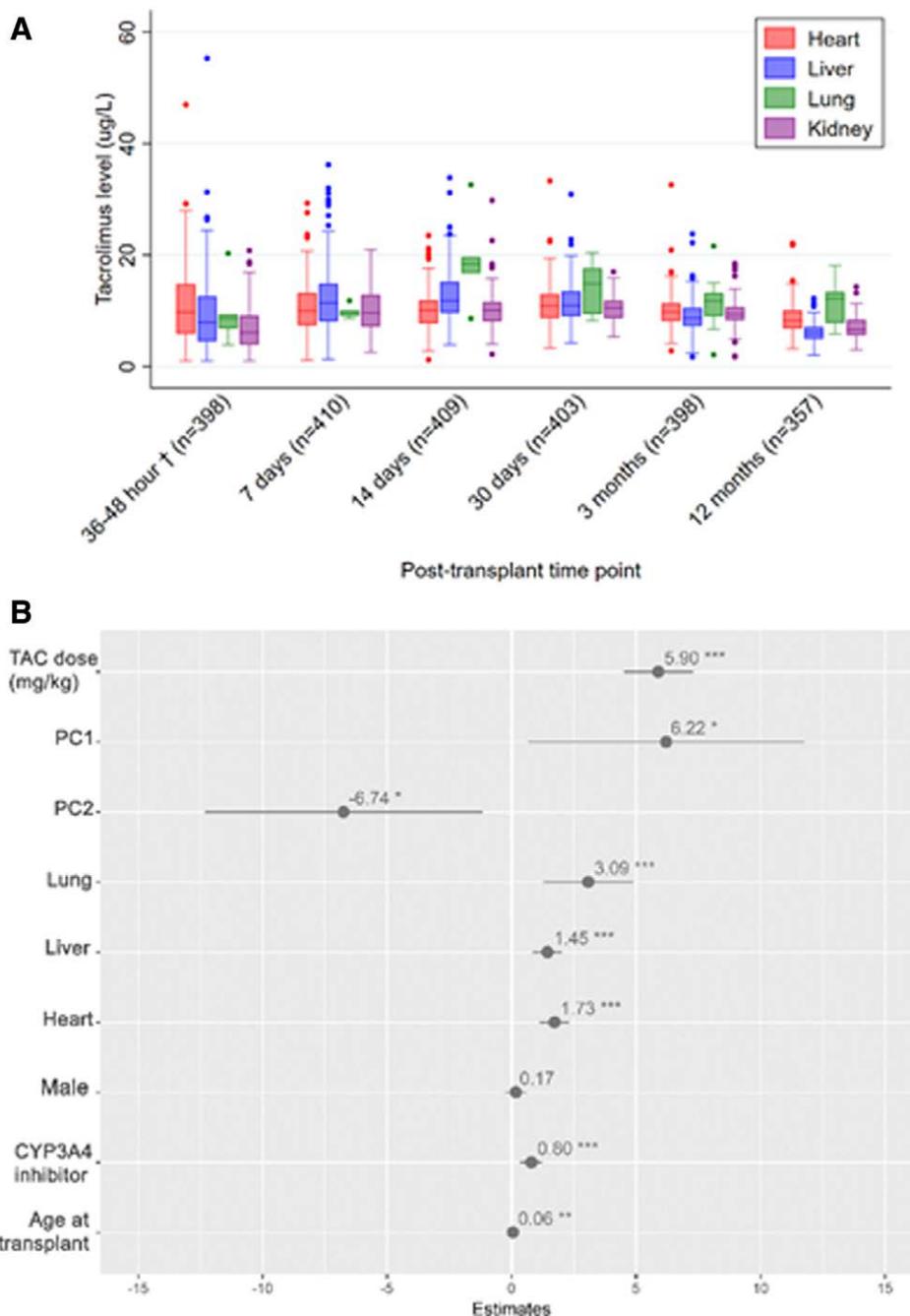
### Interaction Between Age and Organ Type With rs776746\*1 SNP on Dose-adjusted T1 Levels

There was a significant interaction between age and rs776746\*1 SNP on T1 levels after adjusting for organ type ( $P = 0.032$ ). Although the levels were higher in CYP3A5 nonexpressors compared with expressors across all 3 age groups ( $P < 0.01$ ), the difference in levels was greater in infants (effect size 0.73) and adolescents (effect size 0.67) compared with children (effect size 0.34) ( $P < 0.05$  for children versus other age groups) (Figure 3A). There was also a significant interaction between organ type and rs776746\*1 SNP on association with log-transformed dose-adjusted T1 levels after adjusting for age ( $P < 0.01$ ). Although the levels were higher in CYP3A5 nonexpressors compared with expressors across all 3 organ groups ( $P < 0.001$ ), the difference in levels was greater in heart (effect size 0.92) than in kidney (effect size 0.42) and liver transplant recipients (effect size 0.46) ( $P < 0.01$  for heart versus other organs) (Figure 3B).

### Prediction Model for T1 Tacrolimus Levels

We first used a machine learning approach with the lasso model, which selected 15 variables as predictors of T1 levels. The most important predictors were organ type, age at transplant, rs776746, rs12333983, and rs12957142 SNPs. Linear regression with stepwise deletion also selected the same variables as significant predictors of T1 levels. The strength and direction of effect of clinical variables and SNPs selected from the lasso model is described in Figure 4A and B.

We used the selected variables to build 3 prediction models for dose-adjusted T1 levels using the discovery cohort—clinical only, SNP only, and clinical+SNP. Model performance was assessed in the validation cohort. The combined model had a lower prediction error than the clinical only or SNP only model (Table 2). The relationship between predicted and observed levels is illustrated in Figure 4C. The combined model explained 30% of the variation in T1 levels as compared with 18% by the clinical only and 12% by the SNP only model. Although we selected 3 SNPs, rs776746 was the most important SNP as it explained 10% of the variation in dose-adjusted T1 levels in the overall cohort compared with 4% by rs12333983 and 1% by rs12957142. When analyzed across organ groups, age and the 3 top SNPs explained 23% of the variability in heart but only 14% and 9% variability in dose-adjusted T1 levels in liver and kidney, respectively (Figure 4D).

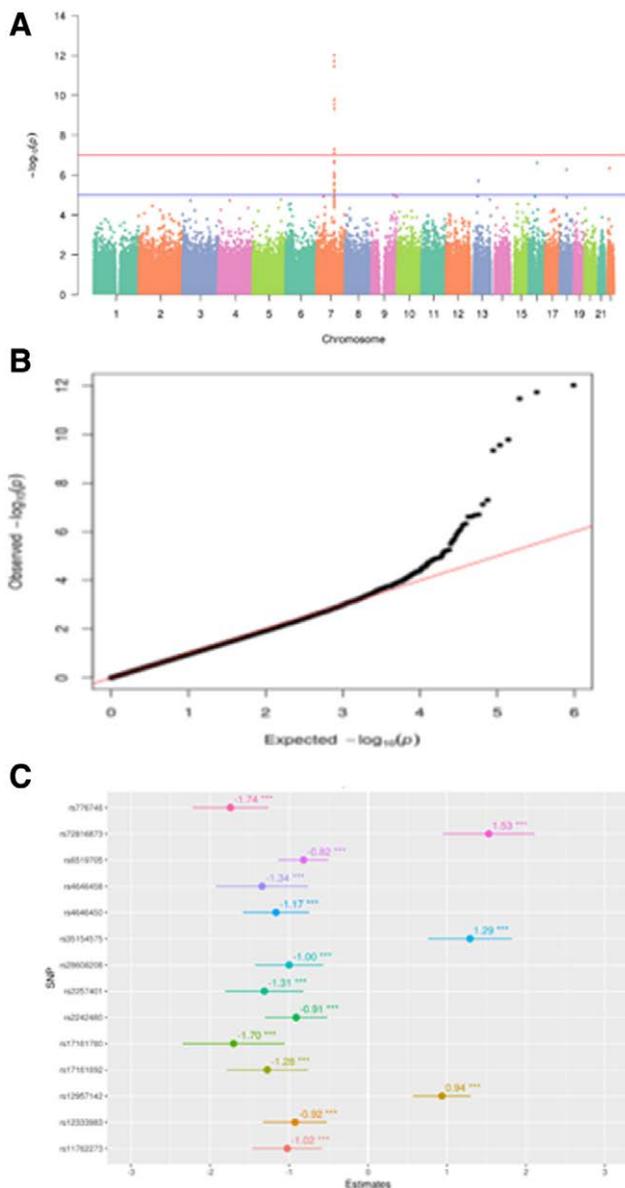


**FIGURE 1.** A, Box plot of tacrolimus levels by time after transplant (n=455 recipients; 2375 tacrolimus levels). The x-axis shows tacrolimus levels at 6 posttransplant time points (36–48 h post-Tac initiation (n=398), 7 d (n=410), 14 d (n=409), 30 d (n=403), 3 mo (n=398), and 12 mo posttransplant (n=357) in kidney (purple), heart (red), liver (blue), and lung (green) transplant recipients. The boxes represent medians and interquartile ranges, the whiskers represent minimum and maximum values, and the dots represent outliers. B, Forest plot of variables associated with tacrolimus trough levels during 1-y follow-up posttransplant using linear-mixed effects model. Tacrolimus levels were higher in liver, heart, and lung recipients compared with kidney recipients. Tacrolimus levels were higher with increasing age, concomitant CYP3A4 inhibitor use, and higher tacrolimus dosage. Dots represent parameter estimate for tacrolimus levels, and bars represent 95% confidence intervals. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ . †Posttacrolimus initiation. PC1, first principal component; PC2, second principal component; TAC, tacrolimus.

**DISCUSSION**

Current practice guidelines for genotype-guided tacrolimus starting dose after organ transplant do not take into consideration interacting clinical factors that can influence tacrolimus levels. Our study used an unbiased approach to identify and validate known and novel SNPs that were independently associated with tacrolimus levels

posttransplant after adjusting for clinical variables. The top-ranked clinical and genetic variables explained 30% of the variability in tacrolimus starting levels (T1). A combined model incorporating clinical and genetic variables had lower prediction error for dose-adjusted T1 levels compared with the clinical or SNP only models, and there were performance differences by organ subtype.



**FIGURE 2.** GWAS SNPs associated with tacrolimus trough levels. A, Manhattan plot showing SNPs associated with tacrolimus levels on GWAS analysis with majority of the significantly associated SNPs residing on chromosome 7. The red line represents genome-wide significance at  $P < 1.025 \times 10^{-7}$ , and the blue line represents  $P < 1 \times 10^{-5}$ . B, Q-Q plot of the observed vs expected  $P$  for SNP association with tacrolimus levels. C, Forest plot of SNPs independently associated with tacrolimus levels during 1-y follow-up posttransplant. Colored dots represent parameter estimate for tacrolimus levels, and bars represent 95% confidence intervals. GWAS, genome-wide association study; SNP, single nucleotide polymorphism.

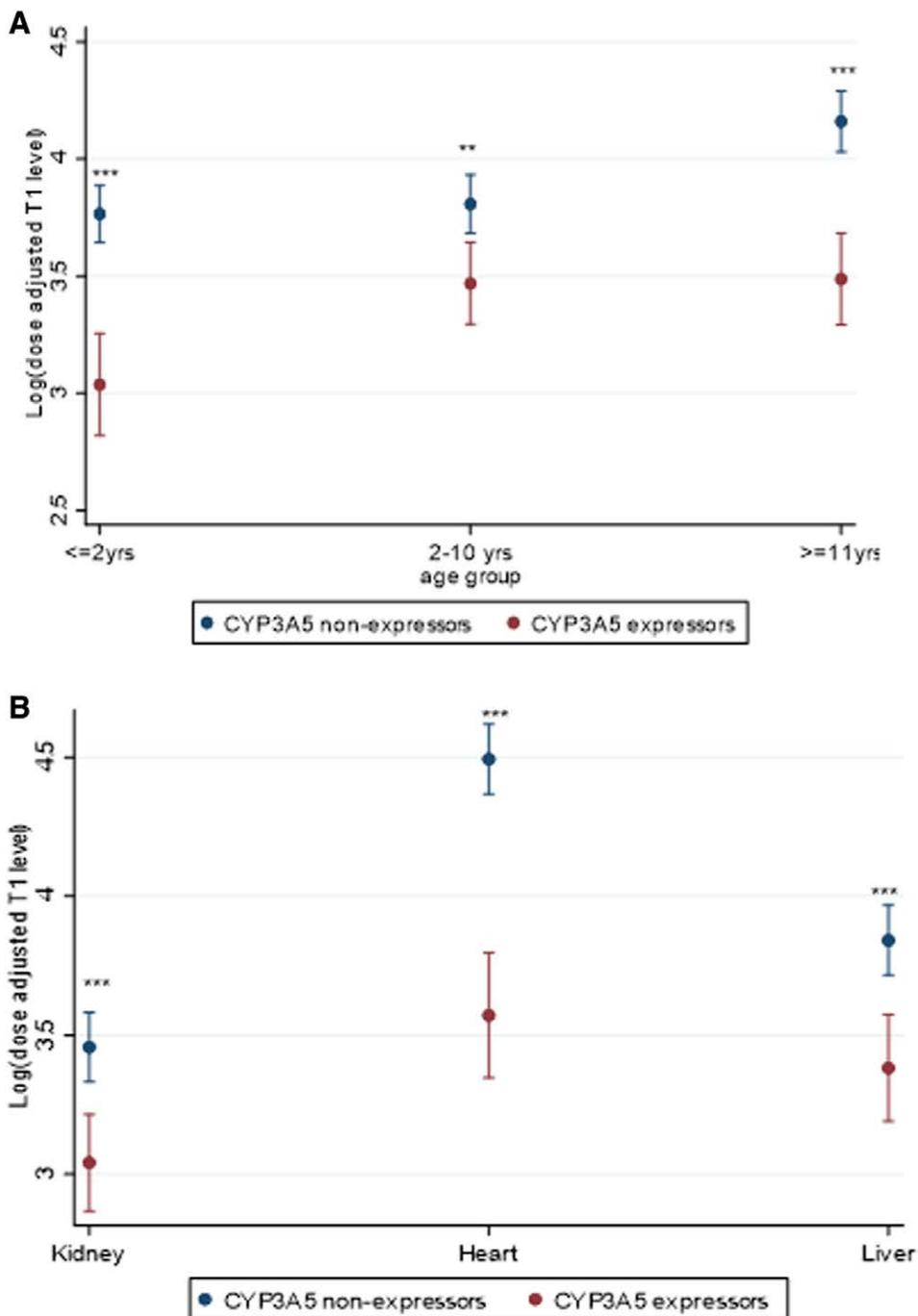
Overall, our study has several important findings. Unlike previous studies that were mostly in adults and in kidney recipients, our study used a GWAS approach to pharmacogenetic discovery that included all pediatric solid organ transplant recipients. This allowed the identification of additional SNPs on chromosome 7 (on *CYP3A4*, *CYP3A5*, *CYP3A7*, and *CYP3A7-CYP3A1* genes) beyond rs776746 (*CYP3A5*) and also allowed the development of an integrated prediction model that included organ-specific differences. It highlights the importance of moving beyond only a genotype-guided dosing to using an

integrated clinical and genetic prediction model to inform individualized tacrolimus dosing.

Besides confirming an association of previously reported factors with tacrolimus levels including organ type, age, tacrolimus dosage, and concomitant use of a CYP3A4 inhibitor drug,<sup>1,3,21,22</sup> a GWAS adjusted for these factors identified known and new variants associated with tacrolimus levels with a cluster located on chromosome 7 that mapped to the *CYP3A* family of pharmacogenes.<sup>7,23</sup> The most significant SNP was rs776746, an established pharmacogenetic SNP associated with tacrolimus levels and for which CPIC guidelines recommend a higher starting dose in CYP3A5 expressors.<sup>8</sup> Of the remaining independent SNPs, 5 SNPs (rs2257401, rs2242480, rs12333983, rs4646450, rs4646458) mapped to pharmacogenes *CYP3A4*, *CYP3A5*, and *CYP3A7*.<sup>24-27</sup> The relatively high proportion of Asians (20%) in our study cohort may explain the significant finding with rs2257401, a SNP in LD with rs776746, that has been associated with tacrolimus levels in a Korean kidney transplant population.<sup>24,28</sup> There is limited data on the SNP, rs12333983 (3'-UTR 27674A>T), known to be associated with hepatic *CYP3A4* expression,<sup>29</sup> rs17161780 that maps to *CYP3A5-ZKSCAN5*, and rs12957142 an intergenic SNP. Two SNPs, *CYP3A5*\*6 and *CYP3A5*\*7, for which CPIC guidelines exist, did not reach significance at a genome-wide level in our study, possibly because these SNPs are associated with tacrolimus levels mainly in an African-American population,<sup>16</sup> an ethnic group that was under-represented in our cohort. *CYP3A4*\*22 (rs35599367) loss of function variant, that has been associated with dosage-adjusted tacrolimus levels in other studies,<sup>23,30</sup> was also not significant in our study likely because only 1 patient was homozygous for this variant. The *ABCB1* 3435C>T SNP (rs1045642)<sup>31</sup> previously reported to be associated with tacrolimus metabolism did not reach significance in our study probably because the effect of *ABCB1* polymorphism on tacrolimus pharmacokinetics is small.<sup>32</sup> Six SNPs did not map to pharmacogenetics.

We used a machine learning approach to build a prediction model for dose-adjusted tacrolimus T1 levels that combined clinical and genetic factors. We focused primarily on T1 levels since the ability to predict accurate starting dose of tacrolimus can improve the achievement and maintenance of on-target tacrolimus concentrations.<sup>10</sup> Both linear regression and lasso models selected organ type, age at transplant, rs776746, rs12333983, and rs12957142 as important predictors of dose-adjusted T1 levels. The machine learning model had a lower prediction error than the linear regression model in the validation cohort, which is consistent with other studies that have reported greater prediction accuracy with machine learning compared with a logistic regression approach.<sup>20,33</sup> The performance of the combined clinical and SNP model was superior to that of the clinical or SNP only models. The combined model explained 30% of the variability in tacrolimus dose-adjusted T1 levels across the cohort. The variability in dose-adjusted T1 levels explained by the model was highest in heart and lower in liver and kidney recipients.

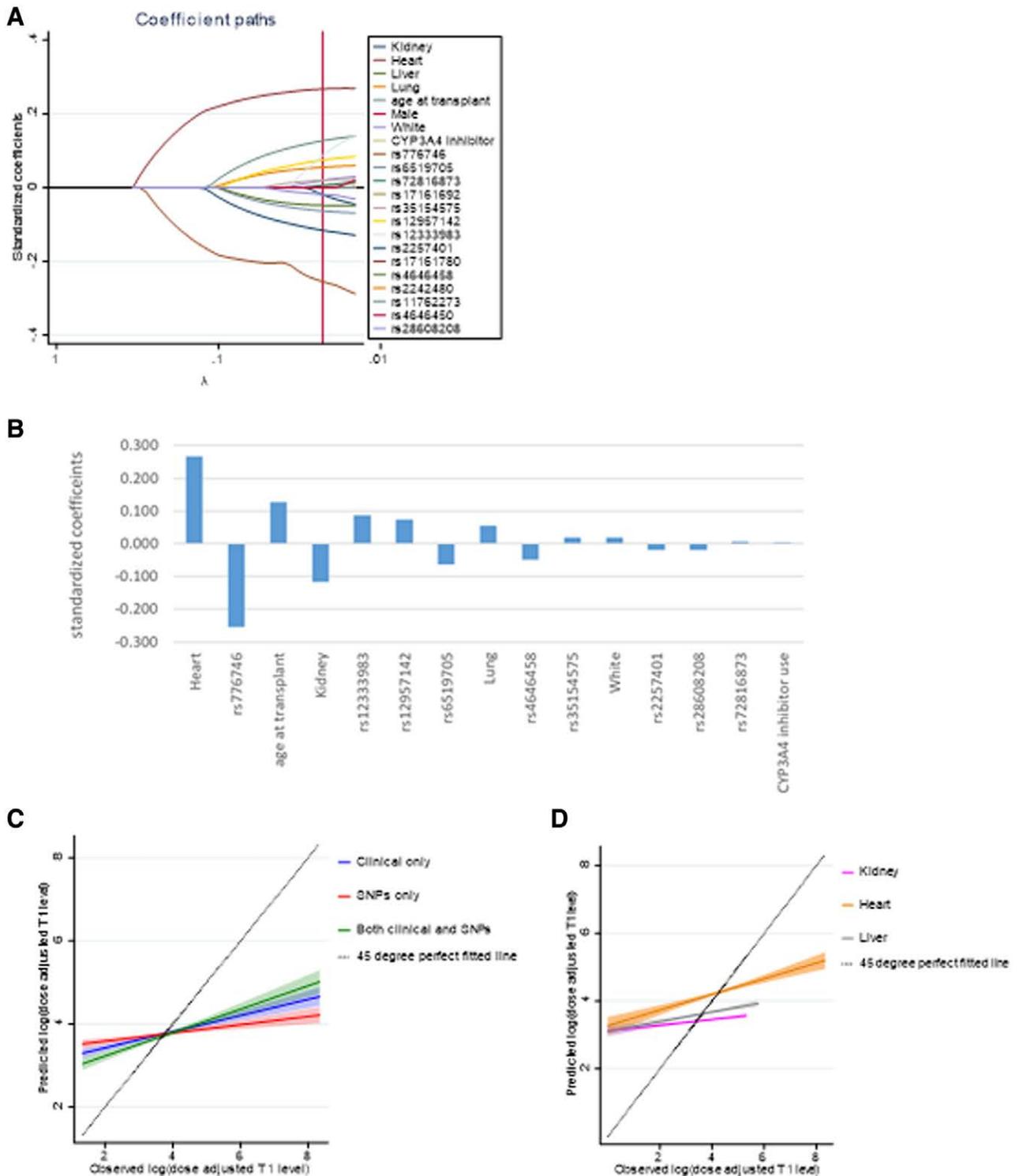
To further define how age and organ type influenced the association of genotype with T1 levels, we performed an interaction analysis using the top-ranked SNP, rs776746\*1, and log-transformed dose-adjusted T1 levels and found a



**FIGURE 3.** Interaction of age and organ type with association of rs776746\*1 SNP with T1 levels. A, Interaction plot of age and rs776746\*1 SNP (adjusted for organ type). Blue dots represent log-transformed dose-adjusted T1 levels in *CYP3A5* nonexpressors, red dots represent log-transformed dose-adjusted T1 levels in *CYP3A5* expressors, and bars represent 95% CI in 3 age groups. The difference in log-transformed dose-adjusted T1 levels between *CYP3A5* nonexpressors and expressors was higher in infants ( $n=238$ ; effect size 0.73;  $P<0.001$ ) and adolescents ( $n=219$ ; effect size 0.67;  $P<0.001$ ) and lower in children ( $n=234$ ; effect size 0.34;  $P<0.01$ ) ( $P<0.05$  between children and other age groups). B, Interaction plot of organ type and rs776746\*1 SNP (adjusted for age). Blue dots represent log-transformed dose-adjusted T1 levels in *CYP3A5* nonexpressors, red dots represent log-transformed dose-adjusted T1 levels in *CYP3A5* expressors, and bars represent 95% CI in 3 different organ types. The difference in log-transformed dose-adjusted T1 levels between *CYP3A5* nonexpressors and *CYP3A5* expressors was higher in heart (effect size 0.92;  $P<0.001$ ) than in kidney (effect size 0.42;  $P<0.001$ ) or liver recipients (effect size 0.46;  $P<0.001$ ) ( $P<0.01$  between heart and other organ recipients). \*\*\* $P<0.001$ , and \*\* $P<0.01$ . SNP, single nucleotide polymorphism.

significant interaction of age and organ type with this association. The difference in T1 levels between nonexpressors and expressors was highest in infant and adolescent age groups and in heart recipients. While *CYP3A5* expression appears to be independent of age,<sup>34,35</sup> tacrolimus

levels are influenced by several other factors that are age-dependent. For example, infants have lower dose-adjusted T1 levels likely due to a relatively large liver size with high plasma clearance of the drug.<sup>36</sup> Also, age influences the expression of other enzymes like *CYP3A4* (active in



**FIGURE 4.** Model performance for dose-adjusted T1 level prediction. A, Coefficient path of lasso model with a red vertical line indicating selected  $\lambda=0.23$ , which has smallest out of sample MSE. The y-axis indicates standardized coefficient. Each colored line represents the independent variable and its coefficient. The variable with the largest standardized coefficient has highest impact on dose-adjusted T1 levels. As variables enter the model from left to right, the variable which enters the model first (ie, diverges from the 0 line) is the most important predictor and the variable which enters last is the least important variable. The x-axis represents the tuning parameter ( $\lambda$ ) of lasso model. B, Standardized coefficient plot of the lasso model showing the relationship between the variables and dose-adjusted T1 levels with the variables above the zero line being positively related while those below the line being negatively associated with dose-adjusted T1 levels. C, Plot of observed vs dose-adjusted T1 levels predicted by clinical only model (blue), SNP only model (red), and combined clinical and SNP model (green) in the validation cohort. The dotted line represents a 45-degree perfect fitted line. (D) Plot of observed vs dose-adjusted T1 levels predicted by the combined clinical and SNP model in heart (orange), liver (gray), and kidney (magenta) recipients. The dotted line represents a 45-degree perfect fitted line. MSE, mean squared error; SNP, single nucleotide polymorphism.

**TABLE 2.****Prediction model performance in discovery and validation cohorts**

Model performance metrics	Discovery cohort (n = 455)		Validation cohort (n = 322)	
	MSE	r <sup>2</sup>	MSE	r <sup>2</sup>
SNP only	0.76	0.12	0.76	0.05
Clinical only	0.71	0.18	0.65	0.18
Clinical+SNP (Linear regression)	0.61	0.30	0.63	0.21
Clinical+SNP (Lasso)	0.60	0.30	0.61	0.23

MSE, mean squared error; r<sup>2</sup>, coefficient of determination; SNP, single nucleotide polymorphism.

adults) and CYP3A7 (active in infants) that CYP3A5 non-expressors have to rely on for tacrolimus clearance.<sup>35,37</sup> Finally, changes in growth hormone and sex hormone levels in children between 5 and 15 years may influence hepatic CYP3A4 expression as these hormones enhance CYP3A4 expression.<sup>38</sup> All these factors likely explain the age-related variability in the association of CYP3A5 genotype with dose-adjusted T1 levels. Further studies are needed to explore the basis of higher levels and greater SNP effect in heart transplants compared with other organ types. Regardless, these findings reinforce the importance of including both clinical and genetic factors when developing tacrolimus dosing algorithms and that an age and organ-specific approach will be needed to optimize prediction models in different organ types.

### Clinical Significance

Several trials have evaluated CYP3A5-guided tacrolimus starting dose with mixed results.<sup>39,40</sup> The DeKAF study group incorporated 4 clinical factors (time post-transplant, age, steroid use, and calcium channel blocker use) in addition to CYP3A5 genotype,<sup>41</sup> but the ability to predict tacrolimus clearance could not be replicated in an independent validation cohort<sup>42</sup> highlighting the challenges with modeling a complex phenotype and possibly an unaccounted effect of other genetic variants.<sup>43</sup> A pediatric trial of solid organ transplants at our center showed the importance of incorporating age into CYP3A5-guided dosing to improve on-target tacrolimus concentrations.<sup>10</sup> The findings of the present study underscore the importance of incorporating not just age and CYP3A5 genotype but also additional SNPs and clinical predictors into individualized dosing. Achieving on-target tacrolimus concentrations through individualized dosing has the potential to reduce the need for therapeutic drug monitoring, reduce costs and hospital length of stay, as well as reduce complications related to tacrolimus over- or under-dosing.<sup>44</sup>

### Limitations

Although there were some differences between the discovery and validation cohort characteristics, the genotype findings were independent of clinical confounders and therefore remained significant in the validation cohorts. The study was underpowered for analysis of donor genotype, especially in liver recipients, for analysis of other factors, for example, hemoglobin, albumin, liver function as potential covariates of tacrolimus levels, and for association with clinical outcomes.<sup>3,45,46</sup> The study was underpowered to

detect an association of 2 SNPs that are included in CPIC guidelines due to underrepresentation of African-American population in our cohort. Our study was also underpowered to detect an association between donor genotype and tacrolimus levels in liver transplants because of a small number of paired liver donor-recipient genotypes. Future work is needed to include a larger sample size with a more diverse population to address these limitations.

In summary, using GWAS, we identified pharmacogenetic SNPs beyond previously known SNPs that were associated with tacrolimus levels independent of age, organ type, ethnicity and concomitant medications in a pediatric transplant cohort. Using machine learning, we demonstrated the superiority of a combined genetic and clinical prediction model for starting tacrolimus levels compared with a clinical or genetic only model, highlighting the importance of a precision medicine approach to tailored medical therapy in this population. While further refinement of the model is needed, the findings of our study pave the way for future development of individualized tacrolimus dosing.

### ACKNOWLEDGMENTS

We acknowledge The Centre for Applied Genomics at the Hospital for Sick Children for genotyping of the study cohort. We thank our colleague, Sangsoon Woo, from Axio Research, who provided statistical data analysis for the GWAS.

### REFERENCES

- Kearns GL, Abdel-Rahman SM, Alander SW, et al. Developmental pharmacology—drug disposition, action, and therapy in infants and children. *N Engl J Med*. 2003;349:1157–1167.
- Christians U, Jacobsen W, Benet LZ, et al. Mechanisms of clinically relevant drug interactions associated with tacrolimus. *Clin Pharmacokinet*. 2002;41:813–851.
- Sikma MA, van Maarseveen EM, van de Graaf EA, et al. Pharmacokinetics and toxicity of tacrolimus early after heart and Lung transplantation. *Am J Transplant*. 2015;15:2301–2313.
- Gueta I, Markovits N, Yarden-Bilavsky H, et al. High tacrolimus trough level variability is associated with rejections after heart transplant. *Am J Transplant*. 2018;18:2571–2578.
- Davis S, Gralla J, Klem P, et al. Tacrolimus inpatient variability, time in therapeutic range, and risk of De Novo Donor-specific antibodies. *Transplantation*. 2020;104:881–887.
- Elens L, Capron A, van Schaik RH, et al. Impact of CYP3A4\*22 allele on tacrolimus pharmacokinetics in early period after renal transplantation: toward updated genotype-based dosage guidelines. *Ther Drug Monit*. 2013;35:608–616.
- Oetting WS, Schladt DP, Guan W, et al; DeKAF Investigators. Genome-wide association study of tacrolimus concentrations in African American kidney transplant recipients identifies multiple CYP3A5 alleles. *Am J Transplant*. 2016;16:574–582.
- Birdwell KA, Decker B, Barbarino JM, et al. Clinical pharmacogenetics implementation consortium (CPIC) guidelines for CYP3A5 genotype and tacrolimus dosing. *Clin Pharmacol Ther*. 2015;98:19–24.
- Gijsen V, Mital S, van Schaik RH, et al. Age and CYP3A5 genotype affect tacrolimus dosing requirements after transplant in pediatric heart recipients. *J Heart Lung Transplant*. 2011;30:1352–1359.
- Min S, Papaz T, Lafreniere-Roula M, et al. A randomized clinical trial of age and genotype-guided tacrolimus dosing after pediatric solid organ transplantation. *Pediatr Transplant*. 2018;22:e13285.
- Rojas L, Neumann I, Herrero MJ, et al. Effect of CYP3A5\*3 on kidney transplant recipients treated with tacrolimus: a systematic review and meta-analysis of observational studies. *Pharmacogenomics J*. 2015;15:38–48.
- Haufroid V, Wallemacq P, Van Kerckhove V, et al. CYP3A5 and ABCB1 polymorphisms and tacrolimus pharmacokinetics in renal transplant

- candidates: guidelines from an experimental study. *Am J Transplant*. 2006;6:2706–2713.
13. Papaz T, Allen U, Blydt-Hansen T, et al. Pediatric outcomes in transplant: PersOnalSing immunosuppression to ImproVe efficacy (POSITIVE study): the collaboration and design of a National transplant precision medicine program. *Transplant Direct*. 2018;4:e410.
  14. Birdwell KA, Grady B, Choi L, et al. The use of a DNA biobank linked to electronic medical records to characterize pharmacogenomic predictors of tacrolimus dose requirement in kidney transplant recipients. *Pharmacogenet Genomics*. 2012;22:32–42.
  15. Keating BJ, Van Setten J, Jacobson P, et al. Design and implementation of the international genetics and translational research in transplantation network. *Transplantation*. 2015;99:2401–2412.
  16. Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. *G3 (Bethesda)*. 2011;1:457–470.
  17. Bonferroni C. *Teoria Statistica Delle Classi e Calcolo Delle Probabilità*. Vol 8. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze; 1936.
  18. Kelly PJ, Stallard N, Whittaker JC. Statistical design and analysis of pharmacogenetic trials. *Stat Med*. 2005;24:1495–1508.
  19. DeSarbo WS, Hwang H, Jedidi K. *Redundancy Analysis*. Wiley StatsRef: Statistics Reference Online. 2016:1–18.
  20. Churpek MM, Yuen TC, Winslow C, et al. Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. *Crit Care Med*. 2016;44:368–374.
  21. Montini G, Ujka F, Varagnolo C, et al. The pharmacokinetics and immunosuppressive response of tacrolimus in paediatric renal transplant recipients. *Pediatr Nephrol*. 2006;21:719–724.
  22. Wallemacq P, Armstrong WW, Brunet M, et al. Opportunities to optimize tacrolimus therapy in solid organ transplantation: report of the European consensus conference. *Ther Drug Monit*. 2009;31:139–152.
  23. Oetting WS, Wu B, Schladt DP, et al. Genome-wide association study identifies the common variants in CYP3A4 and CYP3A5 responsible for variation in tacrolimus trough concentration in Caucasian kidney transplant recipients. *Pharmacogenomics J*. 2018;18:501–505.
  24. Sohn M, Kim MG, Han N, et al. Whole exome sequencing for the identification of CYP3A7 variants associated with tacrolimus concentrations in kidney transplant patients. *Sci Rep*. 2018;8:18064.
  25. Liu J, Ouyang Y, Chen D, et al. Donor and recipient P450 gene polymorphisms influence individual pharmacological effects of tacrolimus in Chinese liver transplantation patients. *Int Immunopharmacol*. 2018;57:18–24.
  26. Liu MZ, He HY, Zhang YL, et al. IL-3 and CTLA4 gene polymorphisms may influence the tacrolimus dose requirement in Chinese kidney transplant recipients. *Acta Pharmacol Sin*. 2017;38:415–423.
  27. Wang P, Mao Y, Razo J, et al. Using genetic and clinical factors to predict tacrolimus dose in renal transplant recipients. *Pharmacogenomics*. 2010;11:1389–1402.
  28. Oetting WS, Wu B, Schladt DP, et al. Attempted validation of 44 reported SNPs associated with tacrolimus troughs in a cohort of kidney allograft recipients. *Pharmacogenomics*. 2018;19:175–184.
  29. Lamba V, Panetta JC, Strom S, et al. Genetic predictors of interindividual variability in hepatic CYP3A4 expression. *J Pharmacol Exp Ther*. 2010;332:1088–1099.
  30. Elens L, van Schaik RH, Panin N, et al. Effect of a new functional CYP3A4 polymorphism on calcineurin inhibitors' dose requirements and trough blood levels in stable renal transplant patients. *Pharmacogenomics*. 2011;12:1383–1396.
  31. Vafadari R, Bouamar R, Hesselink DA, et al. Genetic polymorphisms in ABCB1 influence the pharmacodynamics of tacrolimus. *Ther Drug Monit*. 2013;35:459–465.
  32. Staatz CE, Goodman LK, Tett SE. Effect of CYP3A and ABCB1 single nucleotide polymorphisms on the pharmacokinetics and pharmacodynamics of calcineurin inhibitors: part I. *Clin Pharmacokinet*. 2010;49:141–175.
  33. Merali ZG, Witwi CD, Badhiwala JH, et al. Using a machine learning approach to predict outcome after surgery for degenerative cervical myelopathy. *PLoS One*. 2019;14:e0215133.
  34. Lacroix D, Sonnier M, Moncion A, et al. Expression of CYP3A in the human liver—evidence that the shift between CYP3A7 and CYP3A4 occurs immediately after birth. *Eur J Biochem*. 1997;247:625–634.
  35. Stevens JC, Hines RN, Gu C, et al. Developmental expression of the major human hepatic CYP3A enzymes. *J Pharmacol Exp Ther*. 2003;307:573–582.
  36. Murry DJ, Crom WR, Reddick WE, et al. Liver volume as a determinant of drug clearance in children and adolescents. *Drug Metab Dispos*. 1995;23:1110–1116.
  37. Strougo A, Yassen A, Monneau C, et al. Predicting the “First dose in children” of CYP3A-metabolized drugs: evaluation of scaling approaches and insights into the CYP3A7-CYP3A4 switch at young ages. *J Clin Pharmacol*. 2014;54:1006–1015.
  38. Liddle C, Goodwin BJ, George J, et al. Separate and interactive regulation of cytochrome P450 3A4 by triiodothyronine, dexamethasone, and growth hormone in cultured hepatocytes. *J Clin Endocrinol Metab*. 1998;83:2411–2416.
  39. Thervet E, Lorient MA, Barbier S, et al. Optimization of initial tacrolimus dose using pharmacogenetic testing. *Clin Pharmacol Ther*. 2010;87:721–726.
  40. Shuker N, Bouamar R, van Schaik RH, et al. A randomized controlled trial comparing the efficacy of Cyp3a5 genotype-based with body-weight-based tacrolimus dosing after living donor kidney transplantation. *Am J Transplant*. 2016;16:2085–2096.
  41. Passey C, Birnbaum AK, Brundage RC, et al. Dosing equation for tacrolimus using genetic variants and clinical factors. *Br J Clin Pharmacol*. 2011;72:948–957.
  42. Boughton O, Borgulya G, Cecconi M, et al. A published pharmacogenetic algorithm was poorly predictive of tacrolimus clearance in an independent cohort of renal transplant recipients. *Br J Clin Pharmacol*. 2013;76:425–431.
  43. Elens L, Hesselink DA, van Schaik RH, et al. The CYP3A4\*22 allele affects the predictive value of a pharmacogenetic algorithm predicting tacrolimus predose concentrations. *Br J Clin Pharmacol*. 2013;75:1545–1547.
  44. Levy AR, Sobolev B, James D, et al. The costs of change: direct medical costs of solid organ transplantation in British Columbia, Canada, 1995–2003. *Value Health*. 2009;12:282–292.
  45. Brown NW, Gonde CE, Adams JE, et al. Low hematocrit and serum albumin concentrations underlie the overestimation of tacrolimus concentrations by microparticle enzyme immunoassay versus liquid chromatography-tandem mass spectrometry. *Clin Chem*. 2005;51:586–592.
  46. Fukudo M, Yano I, Masuda S, et al. Population pharmacokinetic and pharmacogenomic analysis of tacrolimus in pediatric living-donor liver transplant recipients. *Clin Pharmacol Ther*. 2006;80:331–345.