Comprehensive profiling of ADME genes identified rare variants associated with tacrolimus pharmacokinetic variability

Seung Hwan Song

Department of Medicine

The Graduate School, Yonsei University

Comprehensive profiling of ADME genes identified rare variants associated with tacrolimus pharmacokinetic variability

Directed by Professor Myoung Soo Kim

Doctoral Dissertation
submitted to the Department of Medicine
the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree
of Doctor of Philosophy

Seung Hwan Song

December 2020

This certifies that the Doctoral Dissertation of Seung Hwan Song is approved.

Thesis Supervisor : Myoung Soo Kim

Thesis Committee Member#1 : Beom Seok Kim

Thesis Committee Member#2 : Hyung Joon Ahn

Thesis Committee Member#3: Min Goo Lee

Thesis Committee Member#4: Woong Kyu Han

The Graduate School
Yonsei University

December 2020

ACKNOWLEDGEMENTS

First and foremost, I would thank God for being able to complete this project with success.

I would like to express my deep and sincere gratitude to my thesis supervisor, Professor Myoung Soo Kim, for giving me the opportunity to do research and providing invaluable guidance throughout my Ph. D study and research. I could not have imagined having a better advisor and mentor for my research.

I am extremely thankful and pay my gratitude to my thesis committee: Min Goo Lee, Hyung Joon Ahn, Woong Kyu Han, and Beom Seok Kim for their valuable guidance and support for completion of this projection.

I extend my gratitude to The Research institute for Transplantation, Yonsei University College of Medicine and Dr. Jihoon G. Yoon for giving me this opportunity.

I also acknowledge with a deep sense of reverence, my gratitude towards my family who has always supported me.

At last but not least gratitude goes to all of my friends who directly or indirectly helped me to complete this project report.

Any omission in this brief acknowledgement does not mean lack of gratitude.

Thanking You
Seung Hwan Song

<TABLE OF CONTENTS>

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Comprehensive profiling of ADME genes identified rare variants associated with tacrolimus pharmacokinetic variability

Seung Hwan Song

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Myoung Soo Kim)

Tacrolimus (TAC) is an immunosuppressant that is widely prescribed following allogenic organ transplant. Due to the narrow therapeutic window and large inter-individual pharmacokinetic (PK) variability, optimizing TAC dosing based on genetic markers is necessary. We enrolled 1,133 participants from four different cohorts receiving TAC as the major immunosuppressant drug, consisting of three cohorts with kidney transplant recipients and one cohort with healthy volunteers from clinical trials. We estimated the effect of clinical factors to appropriately control for confounding variables. A single and multiple variant association analysis was conducted using the detected variants with the Korea Biobank Array including 833,535 markers or targeted sequencing for 114 pharmacogenes. This large-scale multi-cohort analysis indicated that clinical variables influence a large portion of the TAC PK variability; therefore, appropriately controlling for confounding variables is essential for a robust pharmacogenetic study using the continuous trait. *CYP3A5**3

was the only common polymorphism associated with TAC PK variability, whereas the effect of *POR**28 was negligible. Multiple variant association analysis using the 61 absorption, distribution, metabolism, and excretion (ADME) genes revealed that *CYP3A4* and *CYP1A1* rare variants are associated with additional TAC PK variability in a context-dependent manner. *CYP3A4* rare variant carriers among intermediate metabolizers displayed higher TAC trough levels, whereas *CYP1A1* rare variant carriers among poor metabolizers had lower TAC trough levels than non-carrier controls. Our study suggested that rare variants of *CYP3A4* and *CYP1A1* genes are associated with TAC PK variability. Comprehensive profiling of ADME genes may thus guide optimal, personalized TAC-dosing regimens.

Key words : tacrolimus; pharmacokinetics; pharmacogenomics; ADME; kidney transplantation; *CYP3A5*; *CYP3A4*; *CYP1A1*; genome-wide association study

Comprehensive profiling of ADME genes identified rare variants associated with tacrolimus pharmacokinetic variability

Seung Hwan Song

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Myoung Soo Kim)

I. INTRODUCTION

Tacrolimus (TAC), also known as FK-506, is an immunosuppressant widely used after allogenic transplants, such as liver, heart, kidney, and hematopoietic stem cell transplantation, to prevent allograft rejection. TAC and cyclosporine A (CsA) are both classified as calcineurin inhibitors, which blocks the activation, proliferation, and differentiation of T lymphocytes, thus eliciting immune suppression.[1] Although CsA was discovered before TAC in the 1970s and started to be used as an immunosuppressive drug in the early 1980s, TAC has gradually replaced CsA as the primary immunosuppressant of choice for allogenic transplants, owing to its higher potency and better clinical outcomes.[2-6] However, several adverse events have been reported with TAC use, including post-transplant nephrotoxicity, post-transplant diabetes, hypertension, and dyslipidemia.[7-9]

Due to the narrow therapeutic window and large inter-individual pharmacokinetic (PK) variability, therapeutic drug monitoring (TDM) is recommended for TAC use to minimize adverse drug events, mainly for nephrotoxicity and acute rejections. TAC TDM is usually performed by

measuring the trough drug concentration (C_0). The target range of C_0 is 4–12 ng/mL in renal transplant patients with some adjustments depending on the post-transplant period, induction or combination therapy, and recipient immune status.[10,11] However, TDM-guided dosing is time-consuming and imperfect. Thus, developing a personalized regimen is still required for optimal immune suppression without adverse drug events.

Recently, the Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines for TAC suggested incorporating genetic factors affecting the TAC PK metabolism.[12] The current CPIC guideline only includes the *CYP3A5* genotype and related metabolizer phenotype. Accordingly, a 1.5 to 2-fold increased starting dose of TAC is recommended for *CYP3A5* extensive (EM) or intermediate (IM) metabolizers compared with poor metabolizers (PM). Although clinical factors and the *CYP3A5* genotype explain up to 50% of inter-individual PK variability, further genetic factors may need to be incorporated for more precise dosing.[13] Therefore, recent studies have focused on identifying other genetic loci or variants, particularly in absorption, distribution, metabolism, and excretion (ADME) genes, that may additionally account for TAC PK variability using genomic tools, such as single nucleotide polymorphism (SNP) arrays or massively parallel sequencing.[13-16]

The major canonical pathway involved in TAC metabolism is phase 1 drug metabolism by two cytochrome P450 3A enzymes (*CYP3A5* and *CYP3A4*). *CYP3A5* plays a dominant role in the oxidative metabolism of TAC, such as via its 13-O-demethylation and hydroxylation.[17,18] Furthermore, frequent non-expression polymorphisms of *CYP3A5*, including *3, *5, *6, and *7 alleles found in different ethnicities, emphasize the role of *CYP3A4* in TAC metabolism.[19] Other molecular markers associated with TAC PKs include variants in *POR*, *ABCB1*, *CYP3A7*, *PPARA*, *NR1I2*, *COMT*, *IL-10*, and *CYB5R2* genes.[2,14,16,20,21] However, due to poor correlation of PK parameters (C_{max}, C_0,

and AUC), limited numbers of study subjects, drug-drug interactions, and other confounding factors, the clinical evidence of these putative genetic markers have been weak or even conflicting.[2,10,22]

To overcome these limitations, we enrolled multiple large cohorts with kidney transplantation recipients and healthy volunteers and used genetic investigation tools, including the SNP array (the Korean Biobank Array or KoreanChip) and targeted sequencing for the 114 ADME/pharmacogenomics (ADME/PGx) genes.[15,23] Our study suggested that comprehensive profiling of rare variants in ADME genes may help establish an efficient-personalized TAC dosing strategy.

II. MATERIALS AND METHODS

Study cohort

This study was conducted using a retrospective cross-sectional design and consists of four different cohorts from single- or multiple-centers in the Republic of Korea (**Appendix Figure A1**). We enrolled a total of 1,133 participants from four different cohorts receiving TAC as the major immunosuppressant consist of three cohorts (cohort 1, n = 346; cohort 2, n = 592; cohort 3, n = 105) with kidney transplant recipients and one cohort (cohort 4, n = 90) with healthy volunteers from clinical trials. Genomic DNA (gDNA) was obtained from the peripheral blood of all participants and isolated using a QIAamp DNA Mini Kit (Qiagen; Hilden, Germany) as per the manufacturer's protocol. GWAS and rare variant analysis was conducted using the Korea Biobank Array to investigate the contribution of common polymorphisms and Korean-specific rare variants in cohort 1 (**Appendix Table A1**).[23] A total of 195 participants in cohorts 3 and 4 were sequenced for 114 ADME/PGx genes and

rare variant analysis was conducted. This study was approved by each center's local ethics committee, i.e., YUHS (IRB No. 1-2015-0056), SNUH (IRB No. H-1604-124-756), and SMC (IRB No. 2018-09-071), and was performed in adherence to the Guidelines for Good Clinical Practices and the Declaration of Helsinki.

Target PK parameter for analysis

Due to the dynamic changes in drug metabolism and concentration, spot sampling has limitations that do not fully reflect PK characteristics. Compared with the clinical setting, healthy male volunteers enrolled from the clinical trial in cohort 4 had relatively stable dynamics and the PK parameters were more accurately measured using 18 samplings within 96 h. Due to limited sampling numbers and heterogeneous host factors, the correlation between PK parameters in cohort 3 was relatively lower than in cohort 4. In cohorts 2 and 3, PK parameter could be estimated only for TAC C_0. For these reasons, we used a natural log (ln)-transformed dose-adjusted TAC C_0 (ln C_0/D) for cohorts 1–3 and ln C_{max}/D for cohort 4 as target phenotype as described in the previous study.[13]

Genome-wide association study using the Korea Biobank Array

For the participants in cohort 1, SNP genotyping was conducted using the Korea Biobank Array (or KoreanChip) v1.1, which includes a total of 833,535 markers and 1,881 pharmacogenetics/ADME-related sites. GWAS was performed for 596,102 common single variants (minor allele frequency, MAF > 1%) based on an additive inheritance and linear regression model adjusting

for clinical variables. Genome-wide significance level was considered *P* value $< 5\times10^{-8}$. Additionally, we attempted to validate 44 reported markers in Korean population that were previously assessed in the European-American and African-American populations.[24] Among these markers, 24 SNPs were included in the KoreanChip and detected in the Korean population.

Targeted sequencing for 114 pharmacogenes

Using the 195 gDNA samples collected from cohorts 3 and 4, targeted sequencing for 114 ADME/PGx genes based on the hybrid capture method was conducted using Illumina HiSeq 2500 System. Bioinformatics analysis was performed using our in-house pipeline as previously described.[15] For variant QC, genotype refinement workflow and variant filtering were applied to the variant call format (VCF) output as follows: variants with genotype quality (GQ) < 25, depth of coverage (DP) < 20, missing genotypes > 10%, and violating the Hardy-Weinberg equilibrium ($P < 10^{-6}$) were excluded. Haplotype and star nomenclature data were obtained from the PharmVar (https://www.pharmvar.org/).[25]

Rare variant association analysis

We selected 61 ADME genes for rare variants analysis that were represented in the 114 ADME/PGx panel and overlapped with the 298 known ADME genes (Figure 3A, http://pharmaadme.org/).[15,26] Functional missense, nonsense, frameshift, and splicing variants (±2bp) were eligible for the analysis. We performed gene-level aggregation of the multiple rare variants to constitute gene sets and evaluated their association with TAC PK variability in the two

independent cohorts (cohort 3 and 4) by Sequence Kernel Association Test (SKAT), Optimized SKAT (SKAT-O), and Burden Test using the R package 'SKAT' (Figure 3B, Supplemental Table 2).[27,28] ADME genes with *P* values < 0.05, calculated by any of the three tests, were reviewed to determine whether the rare variants and genes were associated with TAC PK variability in a relevant manner.

SNP genotyping using TaqMan assay

For the participants (n = 592) in cohort 2, two representative SNPs including rs776746 ([T/C], CYP3A5*3) and rs1057868 ([C/T], POR*28) were genotyped using Taqman assay with ABI PRISM 7900HT real-time PCR system (Applied Biosystems; Foster City, CA, USA).

Structural analysis of *CYP1A1*

The crystal structure of the *CYP1A1* was obtained from the Research Collaboratory for Structural Bioinformatics (RCBB) Protein Data Bank (www.rcsb.org, PDB Entry: 4I8V). The detected rare variants were computationally rendered using the program PyMOL v2.3 (DeLano Scientific, South San Francisco, CA, USA).

Statistical analysis

The normal distribution of log-transformed PK parameters was evaluated by the Shapiro-Wilk test. Comparisons of the mean values of two groups were carried out using the unpaired *t*-test with Welch's correction for continuous variables. Multiple R^2 values were used to measure the explanatory power of the

multiple regression models. A significance level of $P < 0.05$ was used throughout the analyses.

III. RESULTS

Clinical factors attributed to TAC PK variability

Several studies have been conducted to investigate the genetic markers associated with inter-individual TAC PK variabilities. However, many genetic markers that were suggested to be associated with TAC PK were not validated in different clinical settings or displayed inconsistent results. Therefore, we enrolled four different cohorts from different clinical settings and first analyzed the clinical variables and how they confound the PK parameter (**Table 1, Appendix Figure A1**). The first cohort included 346 kidney transplant recipients enrolled from a single center. Living and deceased donors comprised 73.1% and 26.9% of the transplantations, respectively. Most of the recipients (93.6%) received antibody induction therapies, and the interleukin 2 receptor antagonist (IL2RA) therapy was the most frequent induction modality (84.7%) followed by the anti-thymocyte globulin (ATG) infusion. Some recipients received combination therapy or a triple regimen, wherein other immunosuppressants, including Mycophenolate mofetil (MMF; 26.0%), Myfortic (10.7%), and Mizoribine (1.2 %), were co-administered. TAC C_0 levels were periodically monitored after transplant, and we selected three time points (at post-op 1 week, 1 month, and 3 months) for the analysis. Two patients (0.6%) changed the main immunosuppressant to CsA between post-op 1month and 3months; therefore, TAC C_0 at post-op 3 months was not available. Notably, the median TAC C_0 level (4.0, 4.9, and 5.4 ng/mL at post-op 1 week, 1month, 3months) was relatively lower in cohort 1 than in the other cohorts, and the

variances of the target PK parameter (ln C_0/D) were time-dependent and stabilized after post-op 1 month (**Table 2**). The coefficient of variation (CV) of the target PK parameter (ln C_0/D) was 25.4% greater at post-op 1 week than at post-op 3 months, suggesting that patient status, drug-drug interaction, days post-operation, and a plethora of other clinical factors contributed to the TAC PK variability, particularly in the initial period (**Table 2**).

The *CYP3A5**3 allele is a well-established marker with a clear association to TAC PK variability, whereas *POR**28 allele showed inconsistent results with conflicting interpretations. To precisely estimate the true effect of the *POR**28 genotype on TAC PK variability, we evaluated the effect size of clinical factors and representative genotypes (*CYP3A5**3, *POR**28) among the three different

Table 1. Characteristics of the study cohorts (*n* = 1,133)

Cohort	Cohort 1: YUHS_KT		Cohort 2: KOTRY		Cohort 3: SNU_KT		Cohort 4: Tac_CT	
Subject characteristic	Kidney transplant recipient		Kidney transplant recipient		Kidney transplant recipient		Healthy male volunteer	
No. of subjects	346		592		105		90	
Investigation tool	SNP array (the Korea Biobank Array)		SNP genotyping (*CYP3A5**3, *POR**28)		Targeted sequencing for 114 ADME/PGx genes		Targeted sequencing for 114 ADME/PGx genes	
Analysis	Single variant association test		*CYP3A5**3, *POR**28 effect size estimation		Single and multiple variant association test		Single and multiple variant association test	
Age, years, median [SD]	49	[11.5]	51	[12.0]	50	[13.0]	25	[6.4]
Age group, years, n (%)								
18-34	49	14.2%	90	15.2%	23	21.9%	78	86.7%
35-64	285	82.4%	470	79.4%	75	71.4%	12	13.3%
65-84	12	3.5%	33	5.6%	7	6.7%	0	0.0%
Sex								
Male	186	53.8%	357	60.3%	67	63.8%	90	100.0%
Female	160	46.2%	235	39.7%	38	36.2%	0	0.0%

BMI, median [SD]		22.4	[3.3]	21.2	[3.3]	22.4	[3.5]	22.9	[2.1]
*CYP3A5**3 allele (rs776746)									
	EM (*1/*1)	14	4.0%	34	5.7%	8	7.6%	8	8.9%
	IM (*1/*3)	133	38.4%	197	33.3%	30	28.6%	32	35.6%
	PM (*3/*3)	199	57.5%	361	61.0%	67	63.8%	50	55.6%
*POR**28 allele (rs1057868)									
	POR *1/*1	105	30.3%	196	33.1%	40	38.1%	29	32.2%
	POR *1/*28	171	49.4%	278	47.0%	44	41.9%	47	52.2%
	POR *28/*28	70	20.2%	118	19.9%	21	20.0%	14	15.6%
Donor									
	Living-related or unrelated	253	73.1%	399	67.4%	50	47.6%		
	Deceased	93	26.9%	193	32.6%	55	52.4%		
Surgery	Initial transplant	317	91.6%	556	93.9%	NA	NA		
	Retransplant	29	8.4%	36	6.1%	NA	NA		
Antibody induction therapy									
	Monoclonal (IL2RA)	293	84.7%	427	72.1%	NA	NA		
	Polyclonal (ATG)	25	7.2%	133	22.5%	NA	NA		
	Combination	5	1.4%	27	4.6%	NA	NA		
	None	23	6.6%	5	0.8%	NA	NA		

Co-administered immunosuppressant							
MMF	90	26.0%	525	88.7%	NA	NA	
Myfortic	37	10.7%	NA	NA	NA	NA	
Mizoribine	4	1.2%	23	3.9%	NA	NA	
Pharmacokinetic parameters	C_0 at post-op 1wk, 1m, and 3m		C_0 at post-op 2wk		C_0, CL, V_d from TDM at post-op 1wk		C_{max}, AUC, CL, $t_{1/2}$, Vd from constituted samplings

Abbreviations: EM, Extensive metabolizer; IM, Intermediate metabolizer; PM, Poor metabolizer; IL2RA, interleukin 2 receptor antagonist; ATG, anti-thymocyte globulin; MMF, Mycophenolate mofetil; C_0, trough drug concentration; CL, clearance; AUC, area under the curve; Vd, volume of distribution; NA, Not available;

Table 2. Multiple regression model of the target pharmacokinetic parameters in kidney transplant recipients

Cohort	**Cohort 1: YUHS_KT**			**Cohort 2: KOTRY**	**Cohort 3: SNU_KT**
Sampling	post-op 1 week	post-op 1 month	post-op 3 months	post-op 2 weeks	post-op 1 week
Tacrolimus trough conc. (C_0; ng/mL), median [SD]	4.0 [2.2]	4.9 [2.6]	5.4 [2.4]	8.8 [3.7]	9.5 [2.7]
Daily tacrolimus dose (mg), median [SD]	2.5 [1.9]	4.0 [2.0]	3.5 [1.9]	6.0 [4.2]	9.0 [4.5]
Daily dose normalized tacrolimus trough conc. (ng/mL/mg), median [SD]	1.4 [1.5]	1.3 [1.0]	1.6 [1.1]	1.4 [1.5]	1.5 [1.0]
Log-transformed dose normalized TAC C_0 (ln C_0/D), mean [SD]	0.95 [0.41]	0.915 [0.34]	0.98 [0.34]	0.92 [0.39]	0.84 [0.35]
Coefficient of variation (CV) of ln C_0/D	43.0%	36.8%	34.3%	42.4%	41.4%
Relative CV (%)	125.4	107.3	100.0	123.6	120.7
Model	Variation explained by the model (Multiple R^2, %)				
M1: Clinical variables	15.07	13.10	14.16	24.2	12.95
M2: Clinical variables + *CYP3A5**3	30.17	43.67	48.07	40.0	26.72
M3: Clinical variables + *CYP3A5**3 + *POR**28	30.22	45.12	49.93	40.3	28.37

Significance					
M1 vs. M2	1.87×10^{-14}	2.2×10^{-16}	2.2×10^{-16}	2.2×10^{-16}	3.3×10^{-4}
M2 vs. M3	n.s.	4.8×10^{-2}	2.0×10^{-2}	n.s.	n.s.
Effect size of the genotype attributed to the variance of target PK parameter, %					
*CYP3A5**3	15.10	30.57	33.91	15.80	13.77
*POR**28	0.05	1.45	1.86	0.30	1.65

cohorts by building multivariate regression models. Based on these models, clinical variables explained 12.95 to 24.20% of the PK variability depending on the clinical settings or the sampling time (**Table 2**). 'Age', 'BSA', and '*CYP3A5* genotype' were significant factors in cohort 1, regardless of the sampling time, whereas 'Sex' was only significant at post-op 3 months (**Supplementary Table S1**). Even after accounting for these confounding variables, only the *CYP3A5* genotype significantly affected PK variability commonly in the three cohorts. The *CYP3A5**3 genotype explained 12.43 to 33.91% of the PK variability depending on the cohorts, whereas the *POR**28 genotype only provided 0.05 to 2.90% of additive variability with marginal significance. Moreover, the effect size of the *POR**28 genotype was far lower than that of the *CYP3A5**3 genotype, with no significance in the multiple regression model. These results suggested that the effect of the *POR*28* genotype to TAC PK variability is negligible when compared with the other clinical variables or the *CYP3A5**3 genotype. Taken together, our results clearly showed that the contribution of clinical factors to the TAC PK variability is larger than the effect of common polymorphisms, except *CYP3A5**3, particularly in the initial post-op phase, and suggested that the confounding factors should be appropriately corrected for to estimate the true effect of genetic markers on TAC PK variability.

Single variant association analysis

By adjusting confounding factors based on multiple linear regression models for the target PK parameter (ln C_0/D), genome-wide association study (GWAS) was conducted in cohort 1. Regardless of the sampling time, no other markers except the *CYP3A5**3 allele (rs776746) exhibited significant results at a genome-wide significant level ($P = 5 \times 10^{-8}$) for the three different time points,

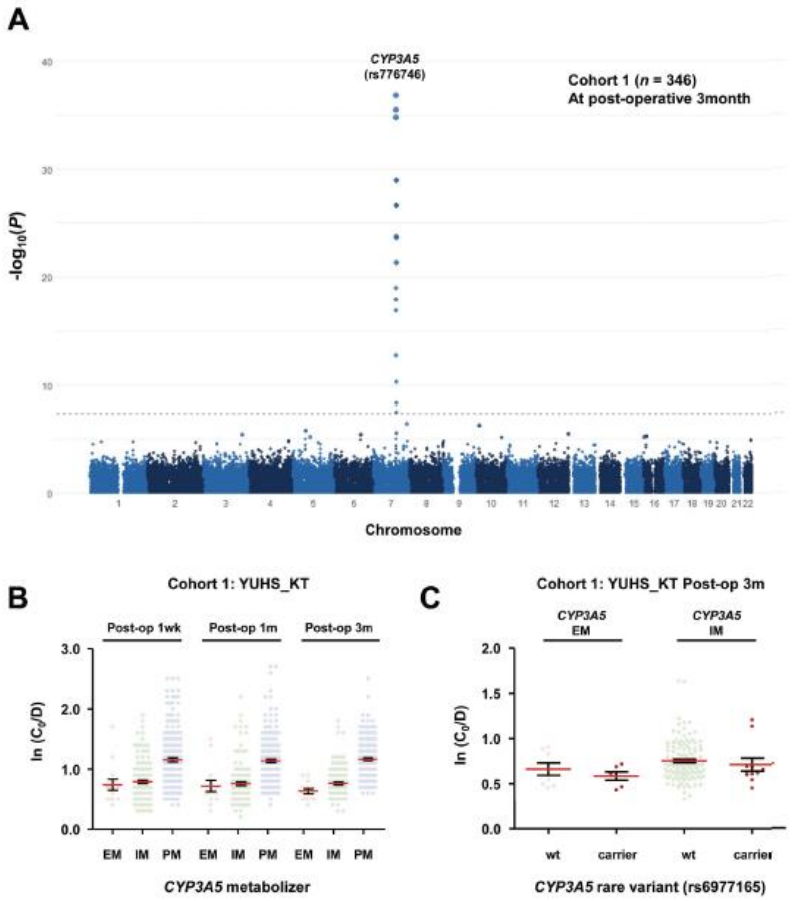


Figure 1. Genome-wide association study identified only a single SNP (rs776746) as a genetic determinant for tacrolimus pharmacokinetic variability. (A) Manhattan plot using the natural log-transformed dose-adjusted tacrolimus trough level (ln C_0/D) measured at post-operative 3 months in cohort 1. (B) Distribution of the target PK parameter (ln C_0/D) by the *CYP3A5* metabolizer phenotype and sampling time. (C) *CYP3A5* rare variant marker (rs6977165) was detected in 17 patients (4.9%) but needs to be further evaluated. Red bar: mean, black error bar: standard error mean.

based on an additive inheritance mode (**Figure 1A**). The *P* values of the *CYP3A5**3 allele were 4.8×10^{-16}, 9.6×10^{-31}, and 1.6×10^{-37} at post-op 1 week, 1 month, and 3 months, respectively (**Appendix Figure A2**). The target PK parameter (ln C_0/D) exhibited wider distribution in the IM and PM groups than the EM groups, and the median value of ln C_0/D gradually increased according to the numbers of *CYP3A5**3 allele (**Figure 1B**). We conditioned the SNP (rs776746) and performed GWAS to find other genetic determinants of TAC PK variability, however, no other common polymorphism showed significant association with TAC PK variability. Additionally, we investigated rare variant markers of the *CYP3A5* gene included in the KoreanChip (**Appendix Table A1**). Only one marker (rs6977165) was detected in 17 patients (4.9%; n = 6 in the EM, n = 11 in the IM), and no patients having *CYP3A5* *3/*3 genotype (PM) were found to harbor this rare variant, indicating that this variant is located in the *CYP3A5**1 haplotype block (**Appendix Figure A3**). This variant was expected to be a stop-loss variant with 12 amino acid extensions (p.X141W*12), however, this marker exhibited no remarkable effects on the PK variability between carrier and non-carriers (**Figure 1C**).

Rare variants associated with TAC PK variability

Next, we investigated rare variants detected by targeted sequencing in cohorts 3 and 4. We first estimated the effect of rare variants detected in known actionable pharmacogenes, including *CYP3A5* and *CYP3A4* (**Table 3**). Most of the detected variants were predicted to be functionally deleterious (splicing defects or frameshift variants) or to be variants of uncertain significance (missense variants). Rare variants in the *CYP3A5* gene were mostly negligible due to the co-occurrence of a functionally non-expressing genotype (*3/*3) in the PM

groups (T007, T076, and T078 in cohort 3; B1228 and B1251 in cohort 4). Due to technical limitations in short-read sequencing, cis- or trans- positions between the *3 allele and splicing variants observed in the IM group (T034 and T070 in cohort 3) were undetermined. This limitation of haplotyping impaired the discrimination of the *CYP3A5* metabolizer phenotype between IM and PM in these patients. Although the *CYP3A5* PM group (*3/*3) is prevalent, most of rare variants in the *CYP3A4* gene were observed in the EM or IM groups. Due to the central role of *CYP3A4* and *CYP3A5* in xenobiotic metabolism, selective pressure may account for extremely rare cases in displaying both the *CYP3A5* non-expressing genotype and *CYP3A4* loss-of-function (LOF) variant.[29] After conditioning to the IM group, the recipients harboring rare variants in the *CYP3A4* gene (T020, T063, T065, and T089) have higher TAC C_0 levels (P = 0.0269) and lower clearance (P = 0.0348) compared to the wild-type groups in cohort 3 (**Figure 2A, B**). Similarly, participants (B1236, B1288, R018, and R025) having rare variants of the *CYP3A4* gene in cohort 4 tended to have relatively higher TAC C_{max} levels and lower clearance; however, statistical analysis was unavailable due to limited case numbers (**Figure 2C, D**). Taken together, these findings have raised issues related to precise haplotyping, functional characterization of unknown rare variants, and the needs for sufficient case numbers of rare variants for statistical analysis.

Beyond the known actionable pharmacogenes, we performed multiple variants association test in the 61 selected ADME genes (**Figure 3A**). Rare variants were aggregated using gene-level collapsing approach and their association was analyzed using SKAT, SKAT-O, and Burden tests (**Figure 3B**). Most of rare variants had MAF < 0.001, particularly for LOF variants. Although each variant had very low allele frequency, proportions of individuals having rare variants were not negligible after gene-level aggregation, even up to 23.6% in *CES1* gene,

Table 3. Rare variants of *CYP3A5*, *CYP3A4*, and *CYP1A1* genes detected by the sequencing study

Gene (transcript)	Sequence change	Amino acid change	Variant classification	Star allele*	*in vitro* function	Cohort 3: SNU_KT (*n* = 105)	Cohort 4: Tac_CT (*n* = 90)	Minor allele frequency (*n* = 195)
CYP3A5 (NM_000777)	c.218+1G>T	splice defect	Splicing	-	LOF	T034 (IM)**	-	0.26%
	c.277G>T	p.V93L	Missense	-	?	T078 (PM)	-	0.26%
	c.432+2T>C	splice defect	Splicing	*3G	LOF	-	B1228 (PM), B1251 (PM)	0.51%
	c.432+1G>T	splice defect	Splicing	-	LOF	T070 (IM)**	-	0.26%
	c.1463T>C	p.I488T	Missense	*3F	?	T076 (PM)	-	0.26%
	c.1493C>A	p.T498N	Missense	-	?	T007 (PM)	-	0.26%
CYP3A4 (NM_001202855)	c.653C>G	p.P218R	Missense	*5	?	-	B1236 (PM)	0.26%
	c.827dupA	p.D276Efs*9	Frameshift	*6	LOF	T089 (IM)	-	0.26%
	c.875T>C	p.L292P	Missense	*18	?	T020 (IM), T042 (EM), T063 (IM), T065 (IM)	B1288 (EM), R018 (IM), R025 (EM)	1.79%
	c.1496dupT	p.V499fs	Frameshift	-	LOF	T063 (IM)	-	0.26%

	c.184G>C	p.A62P	Missense	*12	Decreased	T008 (PM), T031 (IM), T033 (PM), T042 (EM), T044 (PM), T071 (EM), T087 (PM), T098 (EM)	R045 (IM)	2.30%
CYP1A1 (NM_000499)	c.229C>T	p.R77X	Nonsense	-	LOF	-	R028 (PM)	0.26%
	c.403C>T	p.R135W	Missense	-	?	-	B1289 (PM)	0.26%
	c.518C>G	p.T173R	Missense	-	?	T059 (PM)	B1298 (IM)	0.51%
	c.1129C>T	p.R377X	Nonsense	-	LOF	T053 (IM)	-	0.26%
	c.1214G>A	p.R405H	Missense	-	?	T050 (IM)	B1278 (PM), R028 (PM)	0.77%
	c.1235G>A	p.W412X	Nonsense	-	LOF	T060 (IM)	-	0.26%
	c.1275dupT	p.E426X	Nonsense	-	LOF	T002 (PM)	-	0.26%
	c.1430G>A	p.R477Q	Missense	-	?	T081 (PM)	-	0.26%
	c.1475C>G	p.P492R	Missense	*11	?	T092 (PM)	R026 (PM)	0.51%

Abbreviations: EM, Extensive metabolizer; IM, Intermediate metabolizer; PM, Poor metabolizer; LOF, loss-of-function

* Star alleles were obtained from the PharmVar database (https://www.pharmvar.org/);

** cis- or trans- positions are undetermined for these patients, thus *CYP3A5* metabolizer phenotype is based on *3 allele

indicating that rare variants of ADME genes may contribute to unexplained portion of PK variability (**Figure 3C, D**). Because the PK parameter was significantly influenced by several clinical factors described previously, we used cohort 4 as the exploratory dataset which had the precisely measured PK parameters and the fewest confounding variables, and cohort 3 was used as the replication dataset.

Among the 61 selected ADME genes, *CYP1A1* exhibited the lowest *P* values by SKAT and SKAT-O in the exploratory cohort and also by SKAT-O and Burden test in the replication cohort (**Appendix Table A3**). Therefore, we further investigated the rare variants in the *CYP1A1* gene and their association with TAC PK parameters in rare variant carriers versus non-carriers.

A total of 10 rare variants in the *CYP1A1* gene were detected in 21 participants (11.8%) from cohorts 3 (n = 105) and 4 (n = 90): 15 patients (14.2%) in cohort 3 and 6 participants (6.7 %) in cohort 4 (**Figure 3D, Table 3**). All these rare variants were located in exon 2 or 6 with missense or nonsense variants, implying mostly LOF variants (**Appendix Figure A4**). *CYP1A1* rare variant carriers only existed in the IM and PM groups in cohort 4. Interestingly, *CYP1A1* rare variant carriers in the PM groups had relatively lower TAC C_{max} levels than non-carriers in the exploratory dataset (P = 0.0461; **Figure 4A**). This finding was also replicated in cohort 3 and *CYP1A1* rare variant carriers in the PM groups had relatively lower TAC C_0 levels than the non-carriers (P = 0.0159; **Figure 4B**). Although all detected *CYP1A1* rare variant markers were not included in the KoreanChip, we tested whether this tendency also held in cohort 1. KoreanChip detected two rare variant markers of *CYP1A1* in cohort 1 (**Table 3**, **Appendix Table A2**). These include the A62P variant (*12; rs143070677), which was the most frequent rare variant detected in cohorts 3 and 4 (allele frequency: 2.3%), and the P492R

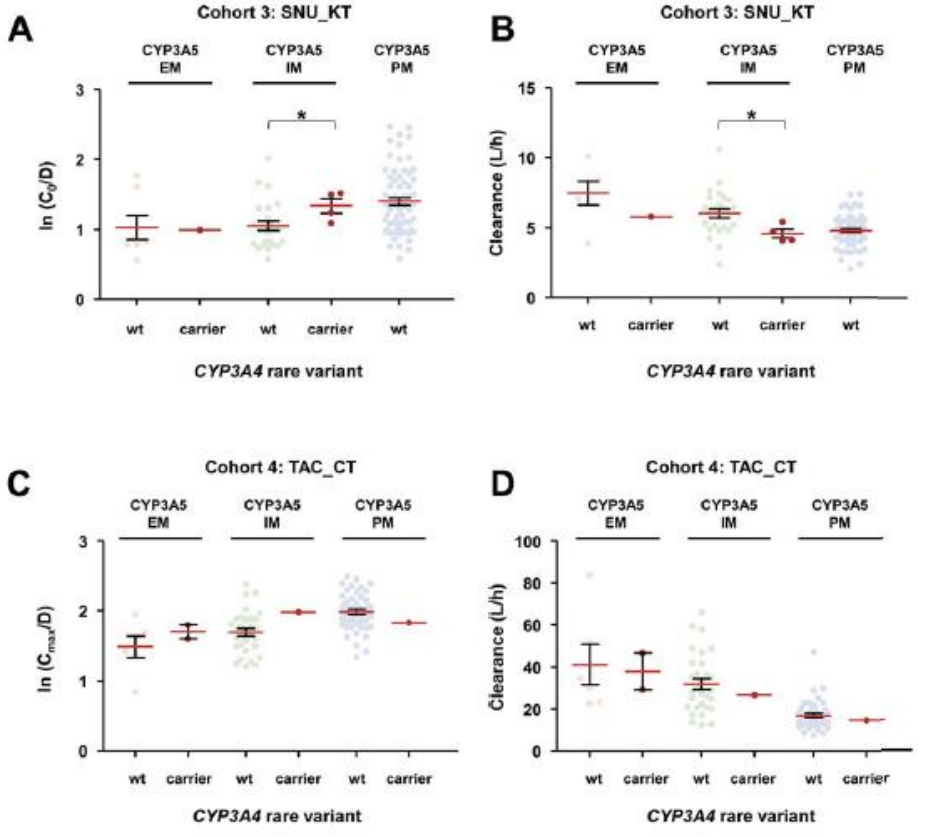


Figure 2. Rare variant profiling of actionable pharmacogenes identified several candidate variants in *CYP3A4* associated with tacrolimus pharmacokinetic variability. (A, B) *CYP3A4* rare variant carriers in the intermediate metabolizer (IM) groups showed relatively higher dose-adjusted TAC C_0 level and lower clearance in the renal transplant recipients (cohort 3). (C, D) Although not remarkable due to the limited number of cases, a similar trend was observed in the healthy volunteers (cohort 4). Red bar: mean, black error bar: standard error mean. *, *P* value < 0.05

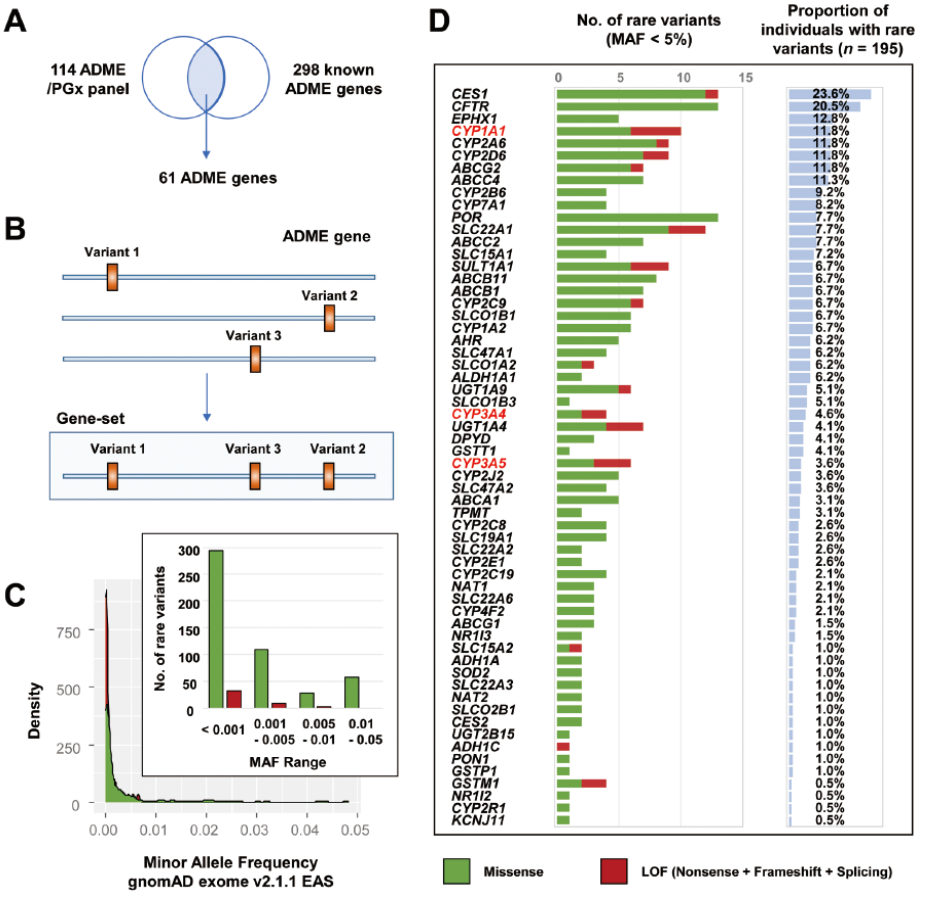


Figure 3. Rare variant association analysis in 61 ADME genes using rare variants (MAF < 5%) detected in cohorts 3 and 4. (A) Selection of 61 target ADME genes included in the 114 ADME/PGx panel and overlapped with the 298 known ADME genes (http://pharmaadme.org/). (B) Gene-level aggregation of detected multiple rare variants to constitute collapsed gene sets. (C) Distribution of rare variants according to minor allele frequency (MAF) obtained from gnomAD exome v2.1.1 EAS database. (D) Numbers of detected rare variants and proportion of individuals harboring rare variants in 61 ADME genes.

variant (*11; rs28399430) located in highly conserved regions. *CYP1A1* A62P (*12) and P492R (*11) variants were detected in 13 patients (3.76%) each and one patient had both variants. All carriers (n = 25) had at least one *3 allele in *CYP3A5* and were stratified into the IM or PM groups. Interestingly, the finding that TAC C_0 levels were lower in the PM groups with *CYP1A1* rare variant carriers than non-carriers was also replicated in cohort 1. Among the 199 patients in the PM group, 13 patients (6.5%) were *CYP1A1* A62P or P492R variant carriers. These patients had relatively lower TAC C_0 levels than the non-carriers particularly in the initial post-op period (**Figure 4C**).

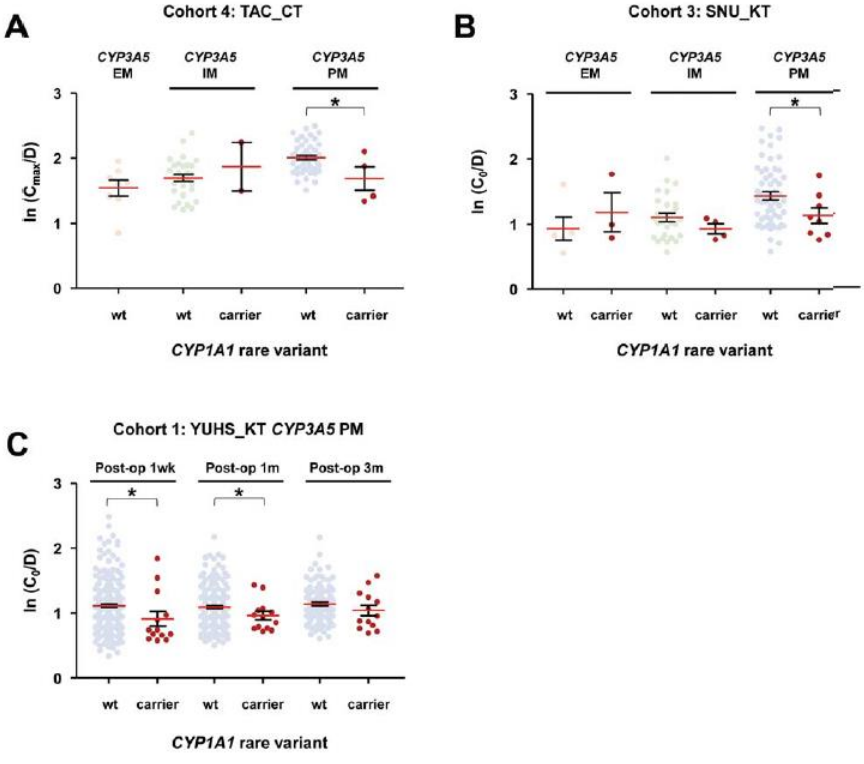


Figure 4. *CYP1A1* rare variant carriers in *CYP3A5* non-expressing groups showed lower tacrolimus trough levels than non-carriers in the three independent cohorts. (A, B) Rare variant association test identified that *CYP1A1* rare variant carriers in the poor metabolizer (PM) group had relatively lower dose-adjusted TAC C_0 level than non-carriers in the healthy volunteers (exploratory dataset: cohort 4) and kidney transplant recipients (replication dataset: cohort 3). (C) Similar tendency are held at post 1 week and 1 month in the PM group of cohort 1 using two *CYP1A1* rare variants (A62P, *12; P492R, *11) included in the KoreanChip. Red bar: mean, black error bar: standard error mean. *, *P* value < 0.05

IV. DISCUSSION

Although TDM is employed to achieve the optimal therapeutic levels and to prevent adverse drug events, a more finely tuned initial dosing is required for TAC administration in clinical practice. Several studies have tried to overcome this unmet need by incorporating genetic information and clinical data using data-driven computerized modeling.[30,31] On the other hand, pharmacogenetic studies using SNP arrays or massively parallel sequencing techniques have focused on uncovering genetic markers associated with TAC PK variability. Although these efforts have identified various candidates, such as variants in *POR*, *ABCB1*, *CYP3A7*, *PPARA*, *NR1I2*, *COMT*, *IL-10*, and *CYB5R2* genes, the *CYP3A5* metabolizer phenotype is the only established and reproducible marker thus far.[2,12]

Due to conflicting reports on the association between candidate markers and TAC PK, such as *POR**28,[24,32,33] we believed that not only the size of the clinical cohorts but the confounding factors to phenotype need to be corrected for a robust pharmacogenetic study. Our analysis performed on multiple cohorts suggested that clinical variables, such as the clinical setting and the sampling time, contributed to 12.95 to 24.20% of PK variability. We estimated the effect size of *CYP3A5**3 and *POR**28 using multiple regression models, and our results clearly showed that the effect size of *CYP3A5**3 is far greater than that of *POR**28 and that the *POR**28 genotype has a negligible effect (**Table 2**). This result implies that the effect size of a novel genetic marker should be greater than the putative uncontrolled confounding variables to be uncovered. Our GWAS result and the previous studies indicated that not common variants but rare variants with a greater effect size may significantly contribute to the remaining portion of TAC PK variability (**Figure 1**).[34]

Consequently, we first analyzed the rare variants in the known actionable pharmacogenes (*CYP3A5*, *CYP3A4*). Because the PM group (59.8%) is the major portion of the *CYP3A5* metabolizer phenotype among Koreans, most of the rare variants located in the *3 haplotype block are expected to have no additional effects. Particularly, the technical limitation related to short-read sequencing illustrates the significance of appropriate phasing and haplotyping for the prediction of the functional impact of the detected rare variants. In contrast, the variation of the TAC PK parameter was more remarkable in the PM group than the IM or EM groups (**Figure 1B**). Although one rare variant (rs6977165) linked to the *1 haplotype was observed, the contribution of this variant to PK variability was uncertain (**Figure 1C**).

For the next step, we performed gene-level aggregation of rare variants and tested the association to the PK parameters in 61 ADME genes (**Figure 3**). We used cohort 4 (n = 90) as an exploratory dataset for its high-quality PK profile and fewer confounding factors, and cohort 3 (n = 105) from the clinical setting as a validation dataset. Interestingly, *CYP1A1* was the only gene having SKAT-O P values < 0.05 in both the exploratory and the validation datasets. A major proportion of the rare variant carriers were predicted to have functionally deleterious variants using multiple *in-silico* prediction tools (**Appendix Table A4**). However, there have been no reports that *CYP1A1* directly metabolizes TAC but *CYP1A1* rare variant carriers exhibited lower TAC C_0 levels in the PM groups indicating gain-of-function effects than a LOF effect. Hence, we postulated that this association is not caused by *CYP1A1* being directly involved in TAC metabolism, but rather by indirect effects such as induction of *CYP3A4* expression at a transcriptional or post-transcriptional level. Since EM and IM groups express *CYP3A5*, induction of *CYP3A4* in TAC metabolism may not be remarkable when compared to PM group. However, in *CYP3A5* non-expressing

group (PM), *CYP3A4* may be the major enzyme responsible for TAC metabolism. A recent study measuring mRNA level induced by TAC reported that the expression profile of six cytochrome P450 enzymes including *CYP1A1*, *CYP2B6*, *CYP3A5*, *CYP4A11*, *CYP19A1*, and *CYP17A1* is different between groups with high and low TAC serum concentration in liver transplant recipients.[35] Moreover, a study using humanized *CYP3A4* mice reported that hepatic *CYP3A4* expression is suppressed by polycyclic aromatic hydrocarbons, which is mainly exemplified by *CYP1A1*.[36] Because *CYP1A1* and *CYP3A4* widely participate in the metabolism of many xenobiotics,[37] the *CYP1A1* LOF variant carrier may display different compensatory changes and transcriptional regulation, in particular the relief of *CYP1A1*-mediated suppression of *CYP3A4* expression, resulting in higher *CYP3A4* induction and lower TAC C_0 levels than non-carriers.[38] Interestingly, a recent study quantifying the relative abundance of *CYP* isoforms in human hepatic tissue reported that the relative abundance of *CYP1A1* and *CYP3A4* among recruited samples was widely distributed, indicating a high level of inter-individual variability.[39]

V. CONCLUSION

To our knowledge, this is the first study reporting the association between rare variants in the *CYP1A1* gene and TAC PK variability. Although this finding was replicated in independent cohorts, the true physiology underlying this association is elusive. Because of the known allelic imbalance of the *CYP3A4* expression pattern, epigenetic regulation or post-transcriptional regulation of *CYP3A4* may be the significant factor underlying wide inter-individual variability of TAC PK.[40] Future studies with genome sequencing for regulatory variants located in non-

coding regions and transcriptome analysis may explain this unknown portion of the observed variability. Additionally, functional studies for the *CYP1A1* rare variants that explain the effect of amino acid substitutions may be necessary, although *CYP1A1* A62P (*12) has been reported to be deleterious.[41] Furthermore, our study emphasized the importance of clinical variables, which could be potentially confounding to the PK phenotype, thus inhibiting a robust pharmacogenetic study. Advanced technical method allowing precise haplotyping and *in vitro* measurement of functional variants may be preceded to fully adopt pharmacogenomics data for clinical decision support.[42] Together, our study provides evidence that comprehensive profiling of rare variants, particularly *CYP3A4* and *CYP1A1*, may help achieve a more finely tailored and personalized TAC dosing regimen.

REFERENCES

1. Azzi JR, Sayegh MH, Mallat SG. Calcineurin inhibitors: 40 years later, can't live without. J Immunol 2013;191:5785-91.

2. Barbarino JM, Staatz CE, Venkataramanan R, Klein TE, Altman RB. PharmGKB summary: cyclosporine and tacrolimus pathways. Pharmacogenetics and genomics 2013;23:563-85.

3. Pirsch JD, Miller J, Deierhoi MH, Vincenti F, Filo RS. A comparison of tacrolimus (FK506) and cyclosporine for immunosuppression after cadaveric renal transplantation. FK506 Kidney Transplant Study Group. Transplantation 1997;63:977-83.

4. Sonoda T, Takahara S, Takahashi K, et al. Outcome of 3 years of immunosuppression with tacrolimus in more than 1,000 renal transplant recipients in japan. Transplantation 2003;75:199-204.

5. Webster AC, Woodroffe RC, Taylor RS, Chapman JR, Craig JC. Tacrolimus versus ciclosporin as primary immunosuppression for kidney transplant recipients: meta-analysis and meta-regression of randomised trial data. BMJ (Clinical research ed) 2005;331:810.

6. Grimm M, Rinaldi M, Yonan NA, et al. Superior prevention of acute rejection by tacrolimus vs. cyclosporine in heart transplant recipients--a large European trial. American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons 2006;6:1387-97.

7. Mihatsch MJ, Kyo M, Morozumi K, Yamaguchi Y, Nickeleit V, Ryffel B. The side-effects of ciclosporine-A and tacrolimus. Clin Nephrol

1998;49:356-63.

8. van Hooff JP, Christiaans MH, van Duijnhoven EM. Tacrolimus and posttransplant diabetes mellitus in renal transplantation. Transplantation 2005;79:1465-9.

9. Naesens M, Kuypers DR, Sarwal M. Calcineurin inhibitor nephrotoxicity. Clinical journal of the American Society of Nephrology : CJASN 2009;4:481-508.

10. Schiff J, Cole E, Cantarovich M. Therapeutic monitoring of calcineurin inhibitors for the nephrologist. Clinical journal of the American Society of Nephrology : CJASN 2007;2:374-84.

11. Brunet M, van Gelder T, Asberg A, et al. Therapeutic Drug Monitoring of Tacrolimus-Personalized Therapy: Second Consensus Report. Ther Drug Monit 2019;41:261-307.

12. Birdwell KA, Decker B, Barbarino JM, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guidelines for CYP3A5 Genotype and Tacrolimus Dosing. Clinical pharmacology and therapeutics 2015;98:19-24.

13. Oetting WS, Schladt DP, Guan W, et al. Genomewide Association Study of Tacrolimus Concentrations in African American Kidney Transplant Recipients Identifies Multiple CYP3A5 Alleles. American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons 2016;16:574-82.

14. 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.

15. Han SM, Park J, Lee JH, et al. Targeted Next-Generation Sequencing for Comprehensive Genetic Profiling of Pharmacogenes. Clinical pharmacology and therapeutics 2017;101:396-405.

16. Dorr CR, Wu B, Remmel RP, et al. Identification of genetic variants associated with tacrolimus metabolism in kidney transplant recipients by extreme phenotype sampling and next generation sequencing. The pharmacogenomics journal 2019;19:375-89.

17. Kamdem LK, Streit F, Zanger UM, et al. Contribution of CYP3A5 to the in vitro hepatic clearance of tacrolimus. Clin Chem 2005;51:1374-81.

18. Dai Y, Hebert MF, Isoherranen N, et al. Effect of CYP3A5 polymorphism on tacrolimus metabolic clearance in vitro. Drug metabolism and disposition: the biological fate of chemicals 2006;34:836-47.

19. Kuehl P, Zhang J, Lin Y, et al. Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. Nat Genet 2001;27:383-91.

20. Li CJ, Li L, Lin L, et al. Impact of the CYP3A5, CYP3A4, COMT, IL-10 and POR genetic polymorphisms on tacrolimus metabolism in Chinese renal transplant recipients. PLoS One 2014;9:e86206.

21. Tang JT, Andrews LM, van Gelder T, et al. Pharmacogenetic aspects of the use of tacrolimus in renal transplantation: recent developments and ethnic considerations. Expert Opin Drug Met 2016;12:555-65.

22. Andrews LM, Li Y, De Winter BCM, et al. Pharmacokinetic considerations related to therapeutic drug monitoring of tacrolimus in kidney transplant patients. Expert Opin Drug Met 2017;13:1225-36.

23. Moon S, Kim YJ, Han S, et al. The Korea Biobank Array: Design and Identification of Coding Variants Associated with Blood Biochemical Traits. Sci Rep 2019;9:1382.

24. 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;19:175-84.

25. Gaedigk A, Ingelman-Sundberg M, Miller NA, Leeder JS, Whirl-Carrillo M, Klein TE. The Pharmacogene Variation (PharmVar) Consortium: Incorporation of the Human Cytochrome P450 (CYP) Allele Nomenclature Database. Clinical pharmacology and therapeutics 2018;103:399-401.

26. Hovelson DH, Xue Z, Zawistowski M, et al. Characterization of ADME gene variation in 21 populations by exome sequencing. Pharmacogenetics and genomics 2017;27:89-100.

27. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. American journal of human genetics 2011;89:82-93.

28. Lee S, Emond MJ, Bamshad MJ, et al. Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. American journal of human genetics 2012;91:224-37.

29. van Herwaarden AE, Wagenaar E, van der Kruijssen CM, et al.

Knockout of cytochrome P450 3A yields new mouse models for understanding xenobiotic metabolism. The Journal of clinical investigation 2007;117:3583-92.

30. Storset E, Asberg A, Skauby M, et al. Improved Tacrolimus Target Concentration Achievement Using Computerized Dosing in Renal Transplant Recipients--A Prospective, Randomized Study. Transplantation 2015;99:2158-66.

31. Tang J, Liu R, Zhang YL, et al. Application of Machine-Learning Models to Predict Tacrolimus Stable Dose in Renal Transplant Recipients. Sci Rep 2017;7:42192.

32. de Jonge H, Metalidis C, Naesens M, Lambrechts D, Kuypers DR. The P450 oxidoreductase *28 SNP is associated with low initial tacrolimus exposure and increased dose requirements in CYP3A5-expressing renal recipients. Pharmacogenomics 2011;12:1281-91.

33. Jannot AS, Vuillemin X, Etienne I, et al. A Lack of Significant Effect of POR*28 Allelic Variant on Tacrolimus Exposure in Kidney Transplant Recipients. Ther Drug Monit 2016;38:223-9.

34. 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. The pharmacogenomics journal 2018;18:501-5.

35. Wang J, Li K, Zhang X, et al. The correlation between the expression of genes involved in drug metabolism and the blood level of tacrolimus in liver transplant receipts. Sci Rep 2017;7:3429.

36. Crosby M, Riddick DS. Suppression of Hepatic CYP3A4 Expression and Activity by 3-Methylcholanthrene in Humanized PXR-CAR-CYP3A4/3A7 Mice. Drug metabolism and disposition: the biological fate of chemicals 2019;47:279-82.

37. Pavek P, Dvorak Z. Xenobiotic-induced transcriptional regulation of xenobiotic metabolizing enzymes of the cytochrome P450 superfamily in human extrahepatic tissues. Curr Drug Metab 2008;9:129-43.

38. Kumar R, Mota LC, Litoff EJ, et al. Compensatory changes in CYP expression in three different toxicology mouse models: CAR-null, Cyp3a-null, and Cyp2b9/10/13-null mice. PLoS One 2017;12:e0174355.

39. Lang D, Radtke M, Bairlein M. Highly Variable Expression of CYP1A1 in Human Liver and Impact on Pharmacokinetics of Riociguat and Granisetron in Humans. Chem Res Toxicol 2019;32:1115-22.

40. Hirota T, Ieiri I, Takane H, et al. Allelic expression imbalance of the human CYP3A4 gene and individual phenotypic status. Hum Mol Genet 2004;13:2959-69.

41. Lee SH, Kang S, Lee S, Ryu D-Y. Functional characterization of cytochrome P450 1A1 Ala62Pro variant by recombinant protein expression (615.3). The FASEB Journal 2014;28:615.3.

42. Sangkuhl K, Whirl-Carrillo M, Whaley RM, et al. Pharmacogenomics Clinical Annotation Tool (PharmCAT). Clinical pharmacology and therapeutics 2020;107:203-10.

APPENDICES

Appendix Figures

Appendix Tables

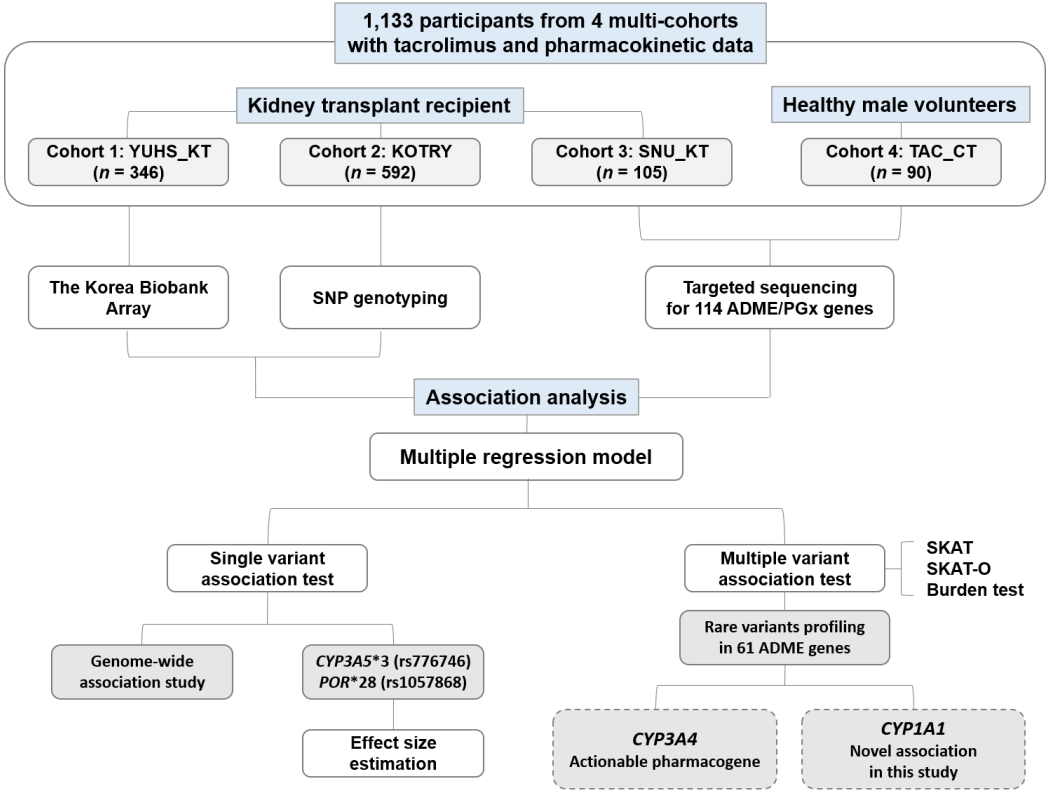


Figure A1. Schematics of enrolled cohorts and association analysis strategy.

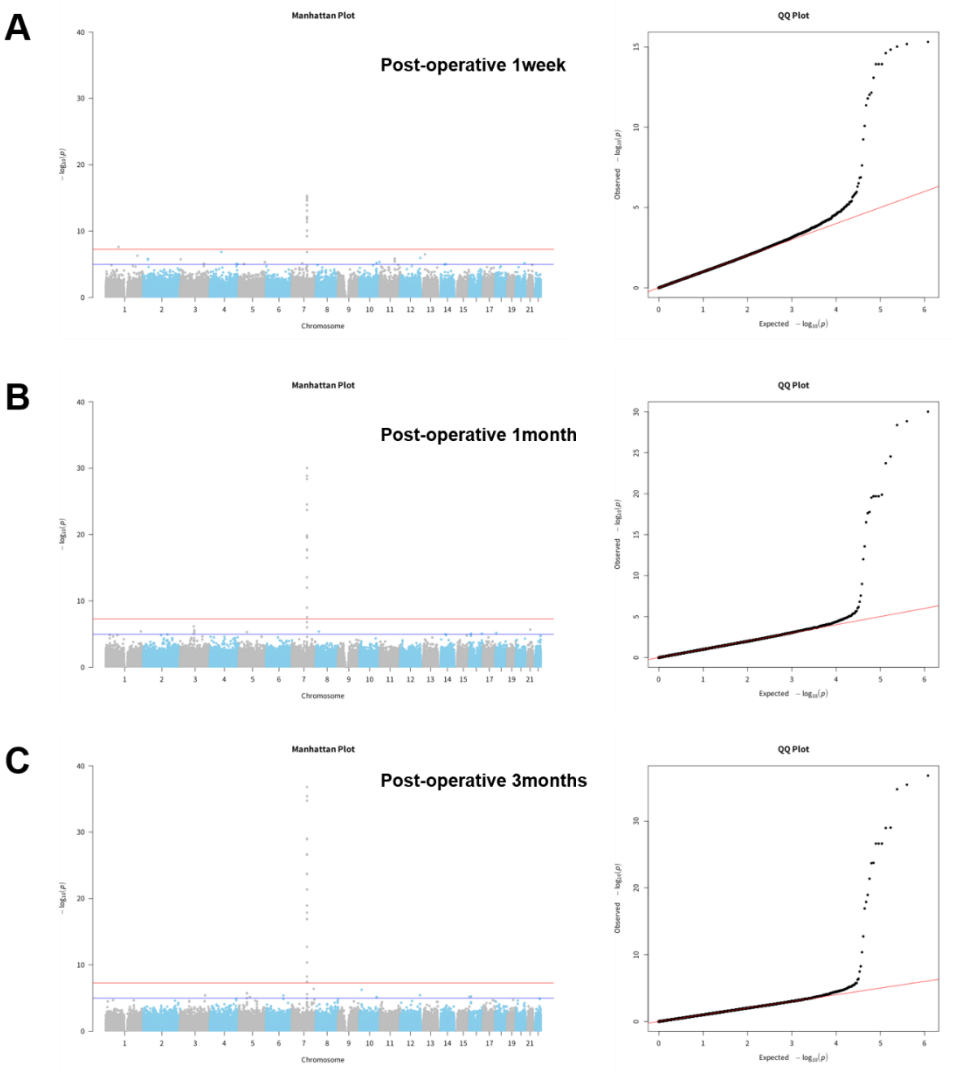
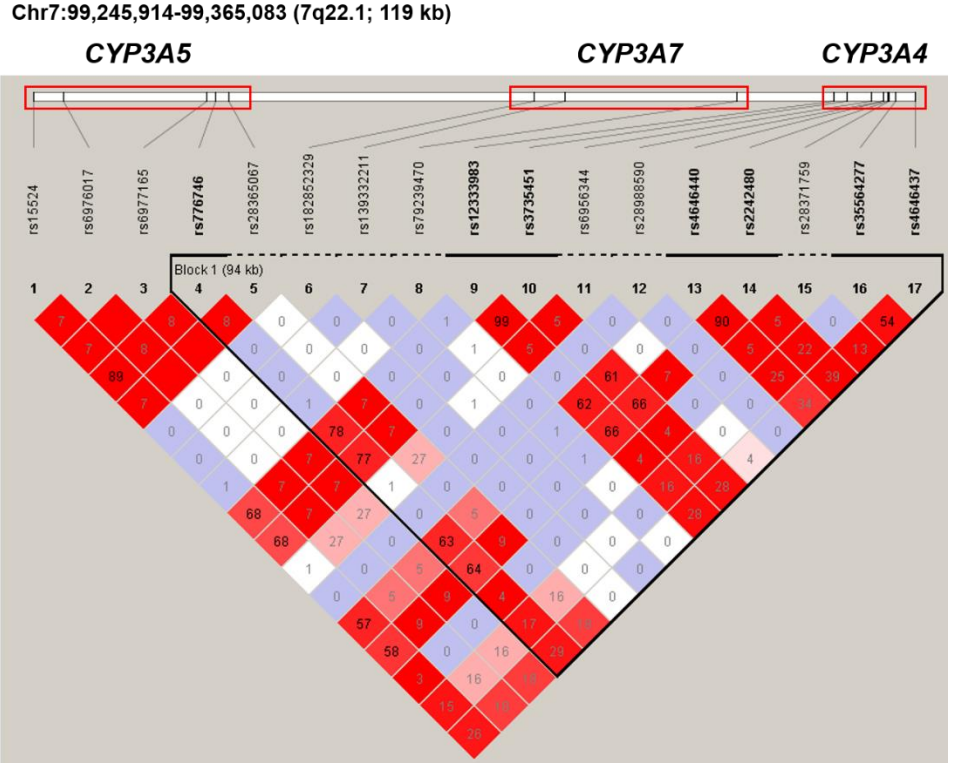


Figure A2. Manhattan plot and QQ plot of genome-wide association study performed in the cohort 1. Genome-wide association study identified only a single SNP (rs776746) as a genetic marker for tacrolimus pharmacokinetic variability at (A) post-operative 1week ($P = 4.8\times10^{-16}$), (B) post-operative 1month ($P = 9.6\times10^{-31}$), (C) post-operative 3months ($P = 1.6\times10^{-37}$) in cohort 1 ($n = 346$)



SNP	rs776746	rs12333983	rs3735451	rs4646440	rs2242480	rs35564277	rs4646437	Frequency (n = 346)
Haplotype	C/T	T/A	T/C	G/A	C/T	T/C	G/A	
H1	C	T	T	G	C	T	G	72.6%
H2	T	A	C	A	T	T	G	9.3%
H3	T	A	C	A	T	C	A	5.2%
H4	T	A	C	G	C	T	G	3.5%
H5	T	A	C	A	T	T	A	2.8%
H6	C	A	C	G	C	T	G	2.1%
H7	T	T	T	G	C	T	G	1.3%

Figure A3. Linkage disequilibrium values (R^2) and haplotype blocks in *CYP3A* region (7q22.1). Haplotype view of *CYP3A* family (*CYP3A5*, *CYP3A7*, and

CYP3A4) region observed in cohort 1 (*n* = 346) exhibited 7 representative haplotypes constituted of 7 SNPs (rs776746, rs12333983, rs3735451, rs4646440, rs2242480, rs35564277, rs4646437) included in the KoreanChip.

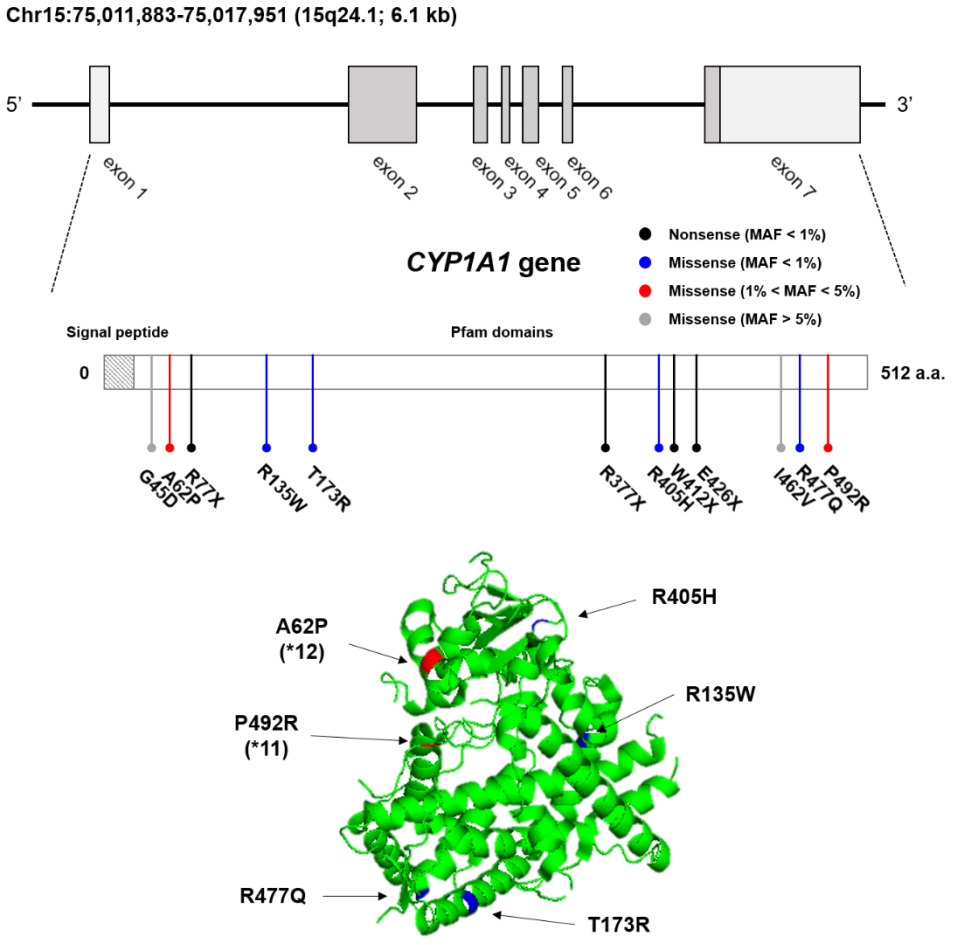


Figure A4. Functional variants of *CYP1A1* gene detected by targeted sequencing. Schematic representation of *CYP1A1* gene coding sequence variants and their minor allele frequency detected by targeted sequencing in the cohort 3 and 4 (*n* = 195)

Table A1. Rare variant markers of *CYP3A5*, *CYP3A4*, and *CYP1A1* genes included in the KoreanChip

Marker ID	rsID	Chr	Position	Reference allele	Alternative allele	Gene	Star allele*	Sequence change	Functional impact	gnomAD EAS AF
AX-83336237	rs41303343	7	99250393	-	A	*CYP3A5*	*7	p.T346Yfs*3	LOF	0.0000544
AX-42283597	rs10264272	7	99262835	C	T		*6	p.K208=	LOF, splicing defect	0
AX-83484666	rs55965422	7	99264573	A	G		*5	c.432+2T>C	LOF, splicing defect	0.00631
AX-86705637	rs759812876	7	99264574	C	A		-	c.432+1C>A	LOF, splicing defect	0.000435
AX-11714998	rs56244447	7	99270276	A	C		*3D	p.L82R	LOF	0
AX-83471153	rs28383468	7	99273815	G	A		*3B	p.H30Y	LOF	0
AX-95861086	rs55817950	7	99273821	G	A		*8	p.R28C	Decreased	0.0000544
AX-15792293	rs6977165	7	99269397	T	C		-	p.X141Wext*12	?	0.0304
AX-59024238	rs4986910	7	99358524	A	G	*CYP3A4*	*3	p.M445T	?	0
AX-106717568	rs4986909	7	99359670	G	A		*13	p.P416L	Decreased	0
AX-82883060	rs12721629	7	99359800	G	A		*12	p.L373F	?	0
AX-83380668	rs67784355	7	99359829	G	A		*11	p.T363M	Decreased	0.000163

AX-59024272	rs28371759	7	99361626	A	G		*18	p.L293P	Decreased	0.0205
AX-83021231	rs4646438	7	99364034	-	T		*6	p.D277Efs*9	LOF	0.00239
AX-112930804	rs145582851	7	99364062	C	T		-	p.R268Q	?	0
AX-36380417	rs12721627	7	99366093	G	C		*16	p.T185S	Decreased	0.00103
AX-15792439	rs35599367	7	99366316	G	A		*22	c.522-191C>T	Decreased	0
AX-83018455	rs4986908	7	99367392	C	G		*10	p.D174H	splicing site	0.0000544
AX-106718954	rs4986908	7	99367392	C	T		-	p.D174N	splicing site	0
AX-83202102	rs72552799	7	99367788	C	T		*8	p.R130Q	Decreased	0
AX-83474939	rs55951658	7	99367825	T	C		*4	p.I118V	Decreased	0.00517
AX-112972020	rs142296281	7	99370218	G	A		-	p.R105W	splicing site	0
AX-83159691	rs56343424	15	75012837	C	T		-	p.R511H	?	0.0002
AX-56429218	rs28399430	15	75012894	G	C		*11	p.P492R	?	0.0054
AX-113526318	rs140680363	15	75012939	C	T	*CYP1A1*	-	p.R477Q	?	0.0003
AX-83313569	rs41279188	15	75012979	G	T		*5	p.R464S	?	0
AX-50946413	rs2278970	15	75012981	G	C		-	p.A463G	?	0

AX-83569597	rs180744198	15	75013005	C	T	-	p.R455Q	?	0.0054
AX-94298666	rs377421606	15	75013327	A	G	-	p.W412R	?	0.0003
AX-83151369	rs56127934	15	75013347	C	T	-	p.R405H	?	0.001
AX-106714988	rs2856833	15	75013563	G	T	-	p.F381L	?	0
AX-83250267	rs4987133	15	75014027	A	G	-	p.I286T	?	0.0002
AX-83537132	rs34260157	15	75014049	G	A	-	p.R279W	?	0
AX-113043569	rs200600221	15	75014823	G	A	-	p.R206C	?	0.0001
AX-82919218	rs28399427	15	75014921	G	C	-	p.T173R	?	0.0093
AX-82893753	rs45442501	15	75015036	G	A	-	p.R135W	?	0.0008
AX-106718917	rs2229150	15	75015162	G	A	-	p.R93W	?	0.00005798
AX-86668287	rs371190271	15	75015210	G	A	-	p.R77X	LOF	0.0003
AX-85004407	rs554072118	15	75015215	T	G	-	p.Q75P	?	0.0019
AX-94294022	rs143070677	15	75015255	C	G	*12	p.A62P	Decreased	0.0045
AX-113986033	rs761140013	15	75015260	T	G	-	p.H60P	?	0.00005814

*Star alleles were obtained from PharmVar (https://www.pharmvar.org/

Table A2. Estimated effect size of clinical variables associated with the target pharmacokinetic parameters in the multiple regression model

Cohort	Cohort 1: YUHS_KT			Cohort 2: KOTRY	Cohort 3: SNU_KT
Sampling	post-op 1 week	post-op 1 month	post-op 3 months	post-op 2 weeks	post-op 1 week
Sex, M vs. F	0.077 [-0.041,0.194]	0.07 [-0.017,0.157]	0.188 [0.105,0.272] ***	0.016 [-0.068,0.1]	0.217 [-0.003,0.438]
Age	0.008 [0.004,0.011] ***	0.005 [0.002,0.007] ***	0.004 [0.001,0.007] **	0.005 [0.003,0.007] ***	0.005 [-0.001,0.01]
Height	0.075 [0.012,0.139] *	0.086 [0.04,0.133] ***	0.028 [-0.017,0.073]	0.014 [-0.024,0.051]	0.051 [-0.068,0.17]
Body weight	0.073 [-0.007,0.153]	0.036 [-0.023,0.095]	0.048 [-0.008,0.105]	0.014 [-0.039,0.068]	0.084 [-0.013,0.181]
BMI	0.132 [-0.005,0.269]	0.184 [0.082,0.285] ***	0.034 [-0.064,0.132]	0.017 [-0.062,0.096]	0.073 [-0.16,0.305]
BSA	-8.927 [-15.734,-2.121] **	-7.585 [-12.616,-2.554] **	-4.591 [-9.39,0.209]	-1.848 [-6.246,2.55]	-8.176 [-19.573,3.221]
CYP3A5, PM vs. IM	-0.328 [-0.411,-0.244] ***	-0.384 [-0.445,-0.322] ***	-0.396 [-0.455,-0.337] ***	-0.262 [-0.32,-0.203] ***	-0.238 [-0.386,-0.09] **
CYP3A5, PM vs. EM	-0.386 [-0.589,-0.184] ***	-0.501 [-0.651,-0.351] ***	-0.527 [-0.67,-0.384] ***	-0.45 [-0.568,-0.331] ***	-0.424 [-0.693,-0.156] **
POR, *1/*1 vs. *1/*28	0.014 [-0.076,0.104]	-0.034 [-0.101,0.032]	-0.031 [-0.094,0.033]	-0.057 [-0.119,0.005]	-0.033 [-0.173,0.107]
POR, *1/*1 vs. *28/*28	-0.006 [-0.118,0.107]	0.016 [-0.067,0.099]	-0.076 [-0.156,0.003]	0.019 [-0.058,0.096]	0.095 [-0.084,0.274]
IL2RA, No vs. Yes	0.047 [-0.098,0.192]	-0.058 [-0.165,0.05]	-0.01 [-0.114,0.094]	-0.053 [-0.193,0.087]	NA

ATG, No vs. Yes	0.055 [-0.124,0.234]	-0.073 [-0.206,0.059]	-0.059 [-0.186,0.068]	0.001 [-0.12,0.121]	NA
Medical center, A vs. B	NA	NA	NA	-0.259 [-0.362,-0.156] ***	NA
Medical center, A vs. C	NA	NA	NA	-0.292 [-0.396,-0.188] ***	NA

The equation (M3) was used for the calculation of the estimated effect size (β) of the clinical variables [95% confidence interval].

Significant factors were marked as follows: *, *P* value < 0.05; **, *P* value < 0.01; ***, *P* value < 0.001;

NA: not available

Table A3. Rare variant association tests performed in 61 ADME genes using SKAT, SKAT-O, and Burden test

Cohort	Cohort 4: Tac_CT				Cohort 3: SNU_KT			
Dataset	Exploratory dataset				Validation dataset			
Genes	No. of variants	*P* value			No. of variants	*P* value		
		SKAT	SKAT-O	Burden		SKAT	SKAT-O	Burden
CYP1A1	6	0.007	0.012	0.063	8	0.527	0.029	0.015
CYP7A1	2	0.024	0.015	0.010	4	0.461	0.634	0.734
EPHX1	5	0.026	0.049	0.302	4	0.314	0.467	0.451
CYP2J2	4	0.054	0.035	0.020	1	0.096	0.096	0.096
SLC22A2	1	0.071	0.071	0.071	2	0.742	0.853	0.806
TPMT	2	0.102	0.045	0.029	1	0.764	0.764	0.764
CYP2B6	3	0.150	0.229	0.179	4	0.105	0.177	0.180
UGT1A4	5	0.153	0.240	0.836	4	0.137	0.122	0.080
CFTR	9	0.155	0.216	0.144	8	0.379	0.287	0.170
SLC47A2	3	0.160	0.247	0.933	2	0.895	0.786	0.656
SLC22A6	2	0.160	0.224	0.176	2	0.874	0.862	0.761
SULT1A1	4	0.180	0.085	0.049	6	0.653	0.206	0.122
ADH1A	1	0.190	0.190	0.190	1	0.127	0.127	0.127
AHR	4	0.207	0.197	0.127	2	0.128	0.085	0.058
SLCO1B1	3	0.225	0.340	0.615	5	0.072	0.128	0.922
NAT1	1	0.229	0.229	0.229	2	0.795	1.000	0.933
SLC15A2	1	0.259	0.259	0.259	1	0.406	0.406	0.406

CYP2C9	4	0.264	0.402	0.841	4	0.233	0.360	0.979
CES2	2	0.282	0.382	0.664	0	NA	NA	NA
SLC47A1	3	0.295	0.425	0.333	4	0.916	0.868	0.705
ABCB1	2	0.301	0.265	0.189	6	0.981	0.918	0.763
ABCG2	5	0.305	0.325	0.212	6	0.626	0.815	0.963
CYP3A5	1	0.308	0.308	0.308	5	0.051	0.081	0.077
ADH1C	2	0.308	0.182	0.124	0	NA	NA	NA
SLC15A1	2	0.351	0.201	0.136	3	0.561	0.705	0.584
POR	7	0.352	0.532	0.947	7	0.045	0.082	0.818
CYP3A4	2	0.366	0.480	0.385	3	0.518	0.573	0.434
UGT1A9	5	0.373	0.264	0.162	4	0.636	0.569	0.398
DPYD	3	0.380	0.495	0.650	1	0.447	0.447	0.447
SLC22A1	6	0.416	0.224	0.133	8	0.175	0.287	0.883
ABCA1	3	0.432	0.470	0.361	4	0.626	0.802	0.689
CYP2E1	1	0.450	0.450	0.450	1	0.349	0.349	0.349
CYP2D6	3	0.503	0.459	0.317	9	0.285	0.450	0.377
ABCC2	4	0.503	0.351	0.227	5	0.117	0.195	0.980
SLCO1B3	4	0.507	0.684	0.513	4	0.153	0.246	0.652
CES1	9	0.539	0.674	0.810	11	0.831	0.632	0.438
ALDH1A1	2	0.546	0.683	0.738	1	0.590	0.590	0.590
CYP1A2	4	0.581	0.463	0.311	4	0.863	0.853	0.685
CYP4F2	1	0.648	0.648	0.648	2	0.365	0.240	0.167
CYP2A6	7	0.715	0.888	0.729	4	0.360	0.516	0.444

SLCO1A2	3	0.738	0.859	0.713	2	0.135	0.192	0.305
NR1I3	1	0.756	0.756	0.756	1	0.411	0.411	0.411
GSTT1	1	0.766	0.766	0.766	1	0.553	0.553	0.553
ABCB11	6	0.769	0.311	0.190	3	0.379	0.147	0.092
ABCG1	1	0.777	0.777	0.777	2	0.410	0.266	0.187
NAT2	2	0.793	0.739	0.605	0	NA	NA	NA
CYP2C19	1	0.810	0.810	0.810	3	0.205	0.170	0.119
KCNJ11	1	0.832	0.832	0.832	0	NA	NA	NA
CYP2C8	2	0.857	0.794	0.654	3	0.038	0.064	0.888
GSTP1	1	0.878	0.878	0.878	1	0.798	0.798	0.798
ABCC4	5	0.904	1.000	0.782	7	0.073	0.120	0.106
UGT2B15	1	0.909	0.909	0.909	1	0.527	0.527	0.527
SLC19A1	2	0.926	1.000	0.947	2	0.622	0.437	0.324
CYP2R1	0	NA	NA	NA	1	0.522	0.522	0.522
GSTM1	0	NA	NA	NA	4	0.402	0.402	0.402
NR1I2	0	NA	NA	NA	1	0.669	0.669	0.669
PON1	0	NA	NA	NA	1	0.891	0.891	0.891
SLC22A3	0	NA	NA	NA	2	0.761	0.588	0.455
SLCO2B1	0	NA	NA	NA	2	0.184	0.259	0.343
SOD2	0	NA	NA	NA	2	0.781	1.000	0.944

NA, Not available: no detected rare variant (MAF < 5%)

Table A4. Functional prediction of missense variants in *CYP3A5*, *CYP3A4*, and *CYP1A1* genes

Gene (Transcript)	dbSNP ID	Sequence change	Amino acid change	Star allele*	Minor allele frequency (n = 195)	gnomAD exome ALL	*in vitro* function	No. of algorithms predicting as deleterious**
CYP3A5 (NM_000777)	rs759099183	c.277G>T	p.V93L	-	2.56E-03	2.84E-05	?	6/10
	rs28365085	c.1463T>C	p.I488T	*3F	2.56E-03	4.00E-04	?	6/10
	rs749945666	c.1493C>A	p.T498N	-	2.56E-03	2.03E-05	?	1/10
CYP3A4 (NM_001202855)	rs55901263	c.653C>G	p.P218R	*5	2.56E-03	1.00E-04	?	7/10
	rs28371759	c.875T>C	p.L292P	*18	1.79E-02	1.50E-03	Decreased	1/10
CYP1A1 (NM_000499)	rs143070677	c.184G>C	p.A62P	*12	2.31E-02	3.00E-04	Decreased	7/10
	rs45442501	c.403C>T	p.R135W	-	2.56E-03	3.00E-04	?	10/10
	rs28399427	c.518C>G	p.T173R	-	5.13E-03	7.00E-04	?	1/10
	rs56127934	c.1214G>A	p.R405H	-	7.69E-03	4.00E-04	?	0/10
	rs140680363	c.1430G>A	p.R477Q	-	2.56E-03	4.47E-05	?	0/10
	rs28399430	c.1475C>G	p.P492R	*11	5.13E-03	4.00E-04	?	7/10

*Star alleles were obtained from PharmVar (https://www.pharmvar.org/)

** PolyPhen-2, SIFT, LRT, MutationTaster, FATHMM, PROVEAN, MetaLR, REVEL, CADD, and DANN algorithms were used for the *in silico* functional prediction. Deleterious prediction was made for the categories of 'damaging', 'probably damaging' , 'disease causing' or 'deleterious' by each prediction tools or by scaled rank score cut-off values > 0.5, respectively

ABSTRACT(IN KOREAN)

흡수, 분포, 대사, 배설 관련 유전자의 프로파일을 통한 면역억제제 Tacrolimus의 약동학에 기여하는 희귀변이 발굴

<지도교수 김 명 수 >

연세대학교 대학원 의학과

송 승 환

타크롤리무스는 장기 이식 수술후 주요 면역억제제로 널리 사용되고 있다. 하지만 타크롤리무스의 좁은 치료 범위와 개인간 약동학의 차이로 인해 타크롤리무스를 개인마다 최적화하는 과정이 필요하다. 본 연구는 타크롤리무스를 주요 면역억제제로서 투여받은 4개의 다른 연구군 (독립된 기관에서 신장이식을 받은 3개의 연구군, 신장 이식을 받지 않은 건강한 성인 남성 지원자로 이루어진 연구군), 총 1,133명의 연구자들을 분석하였다. 각각의 그룹은 833,535 개 유전적 마커를 포함하는 한국인칩 과 114 개 약물 유전자에 대한 표적 염기 서열 분석을 통해 단일 및 다중 변이체 연관성 분석을 시행하였다. 본 연구는 대규모의 다양한 연구군을 분석함에 있어서 교란 변수들을 제어하기 위해 약동학에 영향을 줄수 있는 임상적인 변수를 제어하였다. 건강한 성인 남성 지원자로 이루어진 연구에서 발견된 2가지의 유전적 변이들을 신장이식을 받은 다른 연구군(연구군 1,2,3)에서 재현하였을 때, CYP3A5 * 3은 타크롤리무스의 역동학적 가변성에 영향을 주는 것으로 확인되었지만 다른 유전적 변이인 POR * 28의 효과는

유의미한 차이를 보여주지 못했다. 61 개의 흡수, 분포, 대사 및 배설과 관련된 유전자를 분석한 연구(연구군 3, 4)에서는 CYP3A4 및 CYP1A1 희귀 변이가 상황에 따라 타크롤리무스의 약동학의 변동성에 영향을 주는 것을 확인하였다. CYP3A5의 효소 활성도가 약간 있는 유전적 변이를 지닌 연구군에서 CYP3A4 희귀 변이가 동반될 경우 타크롤리무스의 혈중 농도가 높은 반면, CYP3A5의 효소 활성도가 거의 없는 유전적 변이가 있는 연구군에서 CYP1A1 희귀 변이가 동반되면, CYP1A1 희귀 변이가 없는 연구군에 비해 타크롤리무스의 혈중 농도가 낮아지는 것이 관찰되었다. 본 연구는 CYP3A4 및 CYP1A1 유전자의 희귀 변이가 타크롤리무스의 약동학에 관련이 있음을 보여준다. 이러한 흡수, 분포, 대사, 배설 관련 유전자의 전반적인 프로파일은 개인에게 최적화된 면역억제제 투여 요법을 설립할 수 있을 것으로 기대된다.

핵심되는 말 : 타크롤리무스; 약동학; 약물유전체학; 신장이식; 전장 유전체 연관분석