PLOS ONE

RESEARCH ARTICLE

Genotypic and phenotypic landscapes of 51 pharmacogenes derived from whole-genome sequencing in a Thai population

Natnicha Wankaew[1], Pajaree Chariyavilaskul[2,3], Monpat Chamnanphon[2,4], Adjima Assawapitaksakul[5,6], Wanna Chetruengchai[5,6], Monnat Pongpanich[7,8]*, Vorasuk Shotelersuk[5,6]

1 Program in Bioinformatics and Computational Biology, Graduate School, Chulalongkorn University, Bangkok, Thailand, **2** Clinical Pharmacokinetics and Pharmacogenomics Research Unit, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand, **3** Department of Pharmacology, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand, **4** Department of Pathology, Faculty of Medicine, Srinakharinwirot University, Nakornnayok, Thailand, **5** Center of Excellence for Medical Genomics, Medical Genomics Cluster, Department of Pediatrics, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand, **6** Excellence Center for Genomics and Precision Medicine, King Chulalongkorn Memorial Hospital, the Thai Red Cross Society, Bangkok, Thailand, **7** Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, Bangkok, Thailand, **8** Age-related Inflammation and Degeneration Research Unit, Chulalongkorn University, Bangkok, Thailand

* monnat.p@chula.ac.th

Citation: Wankaew N, Chariyavilaskul P, Chamnanphon M, Assawapitaksakul A, Chetruengchai W, Pongpanich M, et al. (2022) Genotypic and phenotypic landscapes of 51 pharmacogenes derived from whole-genome sequencing in a Thai population. PLoS ONE 17(2): e0263621. https://doi.org/10.1371/journal.pone.0263621

Editor: Laith Al-Eitan, Jordan University of Science and Technology, JORDAN

Received: August 24, 2021

Accepted: January 22, 2022

Published: February 17, 2022

Data Availability Statement: Data cannot be shared publicly because of Center of Excellence for Medical Genomics, Faculty of medicine, Chulalongkorn University privacy terms. Data are available from the Center of Excellence for Medical Genomics Data Access / Ethics Committee for researchers who meet the criteria for access to confidential data. Point of contact for Data Access / Ethics Committee of Center of Excellence for Medical Genomics, Faculty of medicine,

Abstract

Differences in drug responses in individuals are partly due to genetic variations in pharmacogenes, which differ among populations. Here, genome sequencing of 171 unrelated Thai individuals from all regions of Thailand was used to call star alleles of 51 pharmacogenes by Stargazer, determine allele and genotype frequencies, predict phenotype and compare high-impact variant frequencies between Thai and other populations. Three control genes, *EGFR*, *VDR*, and *RYR1*, were used, giving consistent results. Every individual had at least three genes with variant or altered phenotype. Forty of the 51 pharmacogenes had at least one individual with variant or altered phenotype. Moreover, thirteen genes had at least 25% of individuals with variant or altered phenotype including *SLCO1B3* (97.08%), *CYP3A5* (88.3%), *CYP2C19* (60.82%), *CYP2A6* (60.2%), *SULT1A1* (56.14%), *G6PD* (54.39%), *CYP4B1* (50.00%), *CYP2D6* (48.65%), *CYP2F1* (46.41%), *NAT2* (40.35%), *SLCO2B1* (28.95%), *UGT1A1* (28.07%), and *SLCO1B1* (26.79%). Allele frequencies of high impact variants from our samples were most similar to East Asian. Remarkably, we identified twenty predicted high impact variants which have not previously been reported. Our results provide information that contributes to the implementation of pharmacogenetic testing in Thailand and other Southeast Asian countries, bringing a step closer to personalized medicine.

Introduction

Genetic variations play an important role in personalized drug response and constitute pharmacogenetics biomarkers for drug dosing, effectiveness, and toxicity [1]. Variant frequencies

Chulalongkorn University is Dr. Chureerat Phokaew (email: chureerat.p@chula.ac.th).

Funding: This work was supported by Chulalongkorn University Graduate Scholarship to Commemorate the 72nd Anniversary of His Majesty King Bhumibol Adulyadej to NW, Ratchadapiseksompotch Fund, Chulalongkorn University [764002-HE01], Thailand Science Research and Innovation Fund [CU_FRB640001_01_30_10], Thailand Research Fund [DPG6180001], and Health Systems Research Institute [64-132]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

in pharmacogenes differ significantly among various populations [2–4]. However, these data in Asian populations, especially Thai, are limited. A meta-analysis of 56,945 individuals studying allele frequencies of 12 *CYP* genes in several ethnicities, but only 7.6% were data of East Asians [2]. A study in the United States provides allele frequencies in *CYP2D6* of 104,509 individuals where only 0.2% were Asians [5].

High-throughput DNA sequencing technologies with decreasing costs lead genome and exome sequencing to be increasingly used in clinical medicine. They provide the molecular diagnosis for the primary disease and offer a determination of genetic variations in pharmacogenes. To the best of our knowledge, there has been only one study in the Thai population using genome sequencing to investigate pharmacogenes [6]. The participants, however, were affected with Brugada syndrome, which is more common in the Northeastern part of Thailand; therefore, the results may not represent the general Thai population. In addition, the study determined variants in only 25 pharmacogenes. Variants of other 26 genes associated with the pharmacokinetics of drugs treating the top 10 diseases causing the highest mortality for Thais (https://www.cdc.gov/globalhealth/countries/thailand/) including cancer, non-communicable diseases, and infections have not been determined.

In an effort to determine the star allele profile for a more comprehensive list of relevant and meaningful pharmacogenes in a general Thai population, this study explored allele frequencies, genotype frequencies, phenotype prediction together with deleterious variants in 51 pharmacogenes using WGS data of 171 unrelated healthy Thai individuals from all geographical regions of Thailand.

Materials and methods

Study participants

The genome sequencing data of 171 healthy unrelated Thais parents of children with various rare diseases visiting the Genetics Clinic of King Chulalongkorn Memorial Hospital were enrolled in the study.

Ethics statement

The study was approved by the Institutional Review Board of the Faculty of Medicine, Chulalongkorn University (IRB number 264/62). All participants provided their written informed consent.

The margin of error calculation

The margin of error was calculated based on Cochran's sample size formula [7] at a 95% confidence interval, with a sample size of 171, the Thai population size of approximately 67 million people, and a population proportion having an attribute, e.g., normal phenotype at 0.5, which would yield a maximum margin of error.

Star allele analysis

We first assessed the quality of our data using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Next, we used BWA [8] version 0.7.17 (BWA-MEM algorithm) to align reads in Fastq files to the reference genome hg19. To speed up the computation, we extracted reads from genome location of 51 pharmacogenes including:- *CACNA1S*, *CFTR*, *CYP1A1*, *CYP1A2*, *CYP1B1*, *CYP2A6*, *CYP2A13*, *CYP2B6*, *CYP2C8*, *CYP2C9*, *CYP2C19*, *CYP2D6*, *CYP2E1*, *CYP2F1*, *CYP2J2*, *CYP2R1*, *CYP2S1*, *CYP2W1*, *CYP3A4*, *CYP3A5*, *CYP3A7*, *CYP3A43*, *CYP4B1*, *CYP26A1*, *CYP4F2*, *CYP19A1*, *DPYD*, *G6PD*, *GSTM1*, *GSTP1*, *GSTT1*, *IFNL3*, *NAT1*,

NAT2, *NUDT15*, *POR*, *RYR1*, *SLC15A2*, *SLC22A2*, *SLCO1B1*, *SLCO1B3*, *SLCO2B1*, *SULT1A1*, *TBXAS1*, *TPMT*, *UGT1A1*, *UGT1A4*, *UGT2B7*, *UGT2B15*, *UGT2B17*, and *VKORC1*.

Once the bam files were obtained, we used GATK version 4.1.3.0 [9, 10] and followed GATK best practice workflow to obtain called variants in a variant call format (VCF) file. Namely, we identified duplicate reads with the MarkDuplicates tool then performed base quality score recalibration (BQSR). To call variants, HaplotypeCaller per-sample, which generated GVCF file for each sample, was applied, then all GVCFs files were combined, and joint calling using GenotypeGVCFs was performed. Then, we filtered variants with the following criteria. Single nucleotide variants (SNVs) that did not match any of these conditions—QualByDepth (QD) <2.0, FisherStrand (FS) >60.0, RMSMappingQuality (MQ) <40, MappingQualityRankSumTest (MQRankSum) <12.5, and ReadPosRankSumTest (ReadPosRankSum) <-8.0—were passed. Indels that did not match any of these conditions—QD <2.0, FS >200.0, and ReadPosRankSum <-20.0—were passed.

Afterwards, we obtained read depth in the region of all 51 genes for all samples using GATK-DepthOFCoverage (version 3.8.1) which calculated depth from bam files. For compatibility with Stargazer (https://stargazer.gs.washington.edu/stargazerweb), we referred to this read depth file as a target GDF (GATK-DepthOfCoverage Format) file.

To call star alleles and obtain predicted phenotypes, we used Stargazer [11, 12] version 1.0.8, where *EGFR*, *VDR*, and *RYR1* were used as a control gene one at a time. Inputs that Stargazer requires are VCF and GDF files, which were obtained as described previously. Stargazer internally used Beagles [13, 14] to phase haplotype with the 1000 genomes project (1KGP) as a reference panel. The reference panel was changed to only East and Southeast Asian samples.

Pathogenic pharmacogenetic variants analysis

The VCF file was used to calculate allele frequency by plink (version 1.7). Single nucleotide variants and small insertion and deletions were identified as known pharmacogenetic variants by available information from dbSNP (https://www.ncbi.nlm.nih.gov/snp) and PharmGKB (https://www.pharmgkb.org) databases. All variants were annotated with 1000 genomes global and continental minor allele frequencies and variant consequences and its impact on protein function using Ensemble Variant Effect Predictor web tools (http://grch37.ensembl.org/Homosapiens/Tools).

The distribution of high impact variants was compared among Thais (THA) and the five populations from 1000 genomes project (https://www.internationalgenome.org/data), which are Africans (AFR), Americans (AMR), East Asians (EAS), European (EUR), and South Asians (SAS) using Chi-square test. The statistical analyzes were completed by the standard statistical package (R version 4.0.3), and a p-value of <0.01 was used as a significant level after controlling for the critical false discovery rate of 0.05 by the Benjamini-Hochberg procedure [15].

Results

Population

Data from 171 individuals (male/female = 91/80) were analyzed. Geographical areas were available from 65.5% of participants and showed that they were from all regions of Thailand (S1 Table in S2 File). With 171 individuals, the maximum margin of error for the proportion of population carrying certain star alleles is approximately 7.5% at a 95% confidence level.

Sequence quality and read depth

Quality scores across all bases of all samples were >30. All genes except *G6PD*, *GSTM1*, *GSTT1*, *UGT2B15*, and *UGT2B17* had an average read depth >40x (Fig 1). Among three

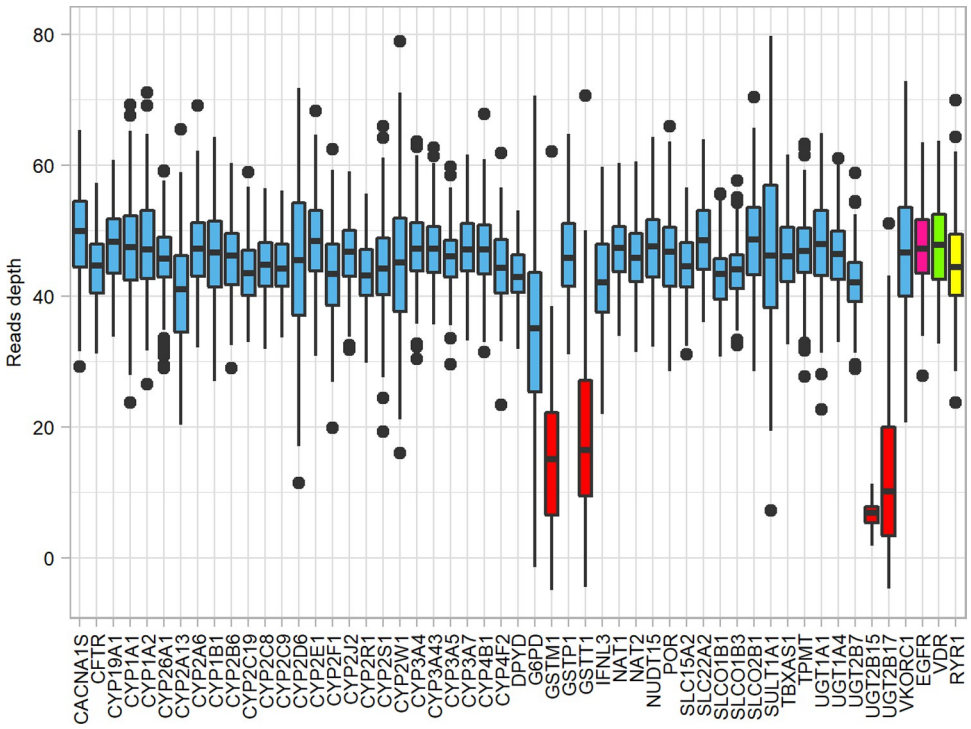

Fig 1. Average read depth. The overview of the average read depths of the 51 pharmacogenes (blue and red) and three control genes, *EGFR* (pink), *VDR* (green), and *RYR1* (yellow). Four red boxes represent genes with low depth coverage, which are *GSTM1*, *GSTT1*, *UGT2B15*, and *UGT2B17*.

https://doi.org/10.1371/journal.pone.0263621.g001

control genes, *EGFR* had the highest average read depth with a mean ± standard deviation (SD) of 43.91 ± 7.45 followed by *VDR* (43.30 ± 7.76) and *RYR1* (40.39 ± 7.82).

Control genes

The three available control genes in Stargazer yielded 99.08% identical results where *EGFR vs VDR*, *EGFR vs RYR1*, and *VDR vs RYR1* had 99.75%, 99.09%, and 99.31% identical results, respectively.

Using *EGFR* as the control gene would be unable to predict genotypes of two genes, *UGT2B15* (n = 99) and *CYP2D6* (n = 7), followed by *VDR* with three genes, *UGT2B15* (n = 92), *CYP2D6* (n = 5), and *SLC22A2* (n = 1), and *RYR1* with four genes, *UGT2B15* (n = 83), *CYP2D6* (n = 8), *SLC22A2* (n = 2), and *CYP2E1* (n = 1) (S1A Fig in S1 File).

UGT2B15 showed the highest incidence of unpredictable genotype, followed by *CYP2D6*, *SLC22A2*, and *CYP2E1*. Stargazer was able to call star alleles in 164, 165, and 162 individuals using *EGFR*, *VDR*, and *RYR1* as the control genes, respectively (S1B Fig in S1 File). The results presented here were based on *EGFR* as the control gene.

Allele and genotype frequencies and phenotype prediction

The thirteen genes that more than 25% of our samples had variant or altered phenotype were *SLCO1B3*, *CYP3A5*, *CYP2C19*, *CYP2A6*, *SULT1A1*, *G6PD*, *CYP4B1*, *CYP2D6*, *CYP2F1*, *NAT2*, *SLCO2B1*, *UGT1A1*, *SLCO1B1* (Fig 2A, S2-S5 Figs in S1 File). Overall, Stargazer detected 196 star alleles in 51 genes (S2 Table in S2 File). Stargazer was unable to call star allele for some genes or predict phenotype (unknown phenotype) in specific individuals. Therefore, the

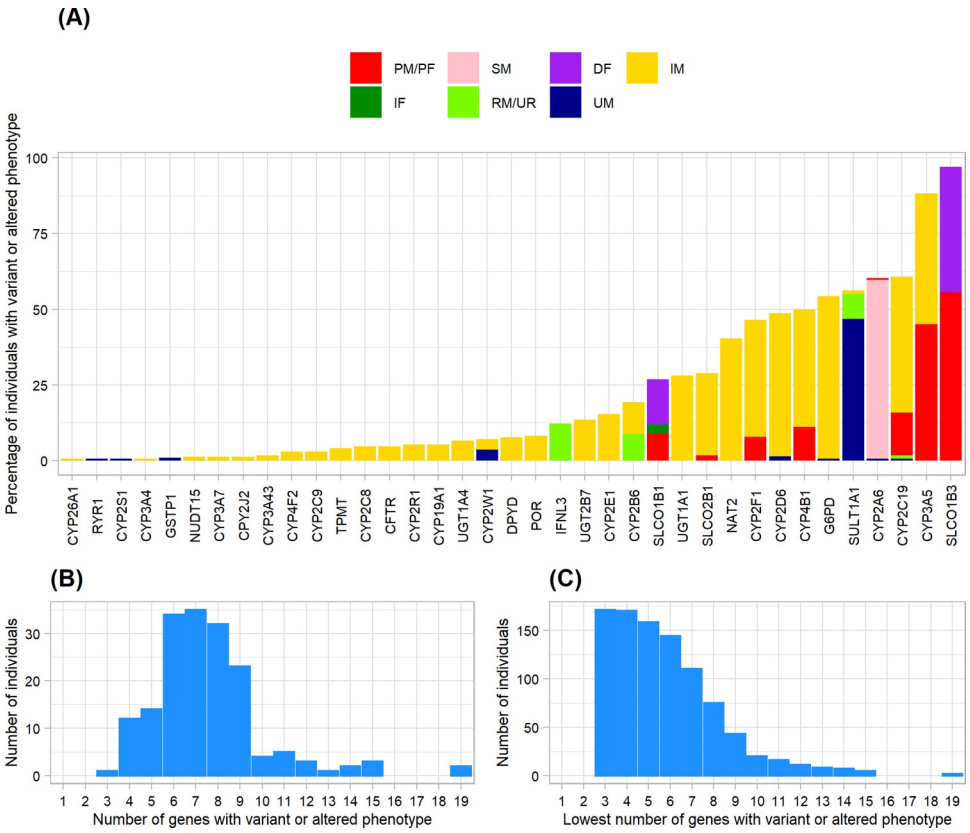

Fig 2. Variant or altered phenotype: Genes and number of individuals. (A) Percentage of individuals with variant or altered phenotype. Poor metabolizer (PM; red) or poor function (PF; red) for *SLCO1B1*, *SLCO1B3*, and *SLCO2B1*; slow metabolizer (SM; pink); decreased function (DF; purple); intermediate metabolizer (IM; yellow); increased function (IF; dark green); rapid metabolizer (RM; green) or unfavorable response (UR; green) for *IFNL3*; ultra-rapid metabolizer (UM; dark blue). (B) The number of individuals with a particular number of genes where individuals had variant or altered phenotype. (C) The number of individuals with at least a particular number of genes where individuals had variant or altered phenotype.

https://doi.org/10.1371/journal.pone.0263621.g002

percentage calculated in Fig 2A was based on the number of informative phenotypes, i.e., excluding individuals with uncalled and unknown phenotypes.

Only 5 individuals (2.9%) had *SLCO1B3*1/*1* normal function phenotype. *SLCO1B3*S1*, a decreased-function allele, is the most common allele present in 166 individuals (97.1%). Of these, 95 individuals (55.6%) had poor-function phenotype (*SLCO1B3*S1/*S1* or *SLCO1B3*S1/*DEL*) and 71 individuals (41.5%) had decreased-function phenotype (*SLCO1B3*1/*S1*) (Fig 2A, S4 Fig in S1 File).

*CYP3A5*3*, a no-function allele, was observed in 74 individuals (43.3%) with the intermediate metabolizer phenotype (*CYP3A5*1/*3*) (S2C Fig in S1 File) and 77 individuals (45.0%) with the poor metabolizer phenotype (*CYP3A5*3/*3*) (Fig 2A, S2 Fig in S1 File).

In *CYP2C19*, 77 individuals (45.0%) were classified as intermediate metabolizers carrying a gene deletion allele (**DEL*) or one no-function **2*, **3*, or **5* allele (*CYP2C19*1/*2*, **1/*3*, **1/*5*, and **1/*DEL*). 24 individuals (14%) were predicted to be poor metabolizers carrying two no-function **2* and **3* alleles (*CYP2C19*2/*2*, **2/*3*, and **3/*3*). Increased-function *CYP2C19*17* allele was found in two individuals with the rapid metabolizer phenotype (*CYP2C19*1/*17*) and one individual with the ultra-rapid metabolizer phenotype (*CYP2C19*17*/17*) (Fig 2A, S2 Fig in S1 File).

In *CYP2A6*, 98 individuals (59%) were predicted to be slow metabolizers carrying one or more decreased-function alleles (*CYP2A6*7*, **9*, **10*, **11*, **12*, **19*, **21*, and **35)* or no-function allele (*CYP2A6*4*). Diplotypes were *CYP2A6*1/*4*, **1/*7*, **1/*9*, **1/*10*, **1/*12*, **1/*19*, **1/*35*, **4/*7*, **4/*9*, **4/*11*, **4/*35*, **7/*12*, **7/*21*, **10/*11*, and **10/*35*. One individual was classified as poor metabolizer (*CYP2A6*4/*4*) and one was ultra-rapid metabolizer (*CYP2A6*1/*1x2*) (Fig 2A, S2 Fig in S1 File).

In *SULT1A1*, several diplotypes carrying ≥2 copies of a gene (*SULT1A1*1/*1x2*, **1/*1x3*, **1/*1x4*, **1/*1x5*, **1x2/*1x2*, **1x2/*1x3*, **1x2/*1x6*, **1x2/*2x2*, **1x2/*2x3*, **1x3/*1x6*, and **1x3/*2x3*) were predicted as ultra-rapid metabolizers (n = 80; 46.8%). Fourteen individuals (8.2%) carried *SULT1A1*1x2/*2* diplotypes and were predicted to be rapid metabolizers. In addition, two individuals predicted to be intermediate metabolizers (*SULT1A1*1/*S1* and *SULT1A1*2/*2*) as they carried *SULT1A1*S1* no-function allele and *SULT1A1*2* decreased-function allele (Fig 2A, S3 Fig in S1 File).

The allele frequencies of *G6PD*1* (normal-function) and *G6PD*DEL* (no-function) alleles were the two most common *G6PD* alleles occurring at 66.1% and 26.6%, respectively. *G6PD*DEL* allele along with other deficiency alleles (*G6PD*8*, **21*, **28*, **31*, **50*, and **51*) comprised diplotypes (*G6PD*1/*DEL*, **21/*DEL*, **28/*DEL*, **51/*51*, **51/*DEL*, and **8/*DEL*) which had predicted phenotype as intermediate metabolizer in 92 individuals (53.8%). Out of 14 diplotypes, *G6PD*1/*DEL* had the highest frequency (49.1%) followed by *G6PD*1/*1* (36.8%). In addition, one individual had *G6PD*1/*1x2* diplotype and was predicted to be an ultra-rapid metabolizer. A high frequency of a deletion observed corresponded to male subjects (Fig 2A, S5 Fig in S1 File).

Thirty-five individuals (38.9%) carrying *CYP4B1*1/*2* and ten individuals (11.1%) carrying *CYP4B1*2/*2* or *CYP4B1*S1/*2* were predicted to be intermediate and poor metabolizers, respectively (Fig 2A, S2 Fig in S1 File).

Nineteen distinct *CYP2D6* alleles were detected in our subjects (S2D Fig in S1 File). These included gene duplication (*CYP2D6*1x2*, **2x2*, **2x3*, **10x2*, and **39x2*), gene deletion (*CYP2D6*5*) and gene rearrangement (*CYP2D6*36+*10* and **36x3+*10*). *CYP2D6*36+*10* and *CYP2D6*10* decreased-function alleles were the most common allele found and contributed to predicted intermediate metabolizer phenotype (*CYP2D6*10/*10*, **10/*36+*10*, **10/*36x3+*10*, **10/*41*, **36+*10/*36+*10*, **36+*10/*41*, **4/*10*, **4/*36+*10*, **5/*10*, and **5/*36+*10*) in 62 individuals. Moreover, intermediate metabolizer phenotype (*CYP2D6*1/*4*, **1/*5*, **2/*5*, **4/*41*, **5/*39*, and **5/*41*) was resulted from other decreased-function (*CYP2D6*41*) and no-function alleles (*CYP2D6*4* and *CYP2D6*5*) in 7 individuals. In addition, two subjects were classified as ultra-rapid metabolizer (*CYP2D6*1/*2x2* and *CYP2D6*1/*2x3*) (Fig 2A, S2 Fig in S1 File).

*CYP2F1*2* (no-function allele) contributed to intermediate metabolizer phenotype (*CYP2F1*1/*2*) in 59 individuals (38.6%) and poor metabolizer phenotype (*CYP2F1*2/*2*) in 12 individuals (7.8%) (Fig 2A, S2 Fig in S1 File).

*NAT2*5*, **6*, **7* decreased-function alleles and *NAT2*DEL* no-function allele comprised intermediate metabolizer diplotypes *(NAT2*1/*DEL*, **5/*6*, **5/*7*, **5/*DEL*, **6/*6*, **6/*7*, **6/*DEL*, **7/*7*, and **7/*DEL*) in 69 individuals (40.4%) (Fig 2A, S3 Fig in S1 File).

A large number of the subjects (71.1%) had *SLCO2B1*1/*1* normal function phenotype. However, 57 individuals had unknown phenotype (*SLCO2B1*1/*S464F*, **S1/*S464F*, and **S464F/*S464F*) due to the unknown function of **S464F* allele. A no-function *SLCO2B1*S1* allele contributed to intermediate function phenotype (*SLCO2B1*1/*S1*) in 31 individuals (27.2%) and poor-function phenotype (*SLCO2B1*S1/*S1*) in 2 individuals (1.8%) (Fig 2A, S4 Fig in S1 File).

*UGT1A1*6*, *7, *27, and *60 decreased-function alleles comprised intermediate metabolizer diplotypes (*UGT1A1*27/*27, *27/*60, *6/*27, *6/*6, *6/*60, *6/*7, and *60/*60) in 48 individuals (28%) (Fig 2A, S3 Fig in S1 File).

Poor-, decreased-, normal- and increased-function phenotypes were observed for *SLCO1B1*. *SLCO1B1*15* and *SLCO1B1*17* decreased-function allele and *SLCO1B1*DEL* no-function allele comprised diplotypes (*SLCO1B1*15/*17, *15/*DEL, *17/*DEL, and *1B/*DEL) that were predicted to be poor-function phenotype in 15 individuals (8.9%). Twenty-five individuals (14.3%) had decreased-function phenotype carrying *SLCO1B1*1/*15, *1/*17, *1B/*15, *1B/*17, or *35/*DEL where *1B was a normal-function allele and *35 was an increased-function allele. Five individuals (3%) carrying *SLCO1B1*1/*35, *14/*35, *1B/*35 were predicted with increased-function phenotype where *14 was an increased-function allele (Fig 2A, S4 Fig in S1 File).

A high frequency of a no-function allele (deletion) was observed in *GSTM1*, *GSTT1*, *UGT2B17*, and *UGT2B15* and this corresponded with poor or intermediate metabolizer phenotype in the first three genes and unknown or unpredictable phenotype in *UGT2B15* (S3 Fig in S1 File). Nevertheless, results in these four genes should be interpreted with caution due to a very low read depth, which might explain a high frequency of a deletion observed (S3 Fig in S1 File).

All individuals had normal phenotype for *CACNA1S* and had unknown phenotypes for *VKROC1*. In *CYP1A1*, *CYP1A2*, *CYP1B1*, *CYP2A13*, *NAT1*, *SLC15A2*, *SLC22A2*, and *TBXAS1*, the predicted phenotypes were either normal metabolizers or unknown.

Individuals mostly have variant or altered phenotypes in 6–8 genes (Fig 2B). Every individual has variant or altered phenotype in at least three genes but no more than 19 genes. One hundred and fifty-eight individuals had variant or altered phenotypes in at least five genes (Fig 2C).

Pharmacogenetic variants frequencies comparison

We identified 28,061 variants within 51 pharmacogenes in the 171 individuals, where the majority were intron variants (S3 Table in S2 File). Variants were classified according to their functional consequences, for example, frameshift, missense or synonymous. There were 322 novel variants with high, moderate or low impacts (S3 Table in S2 File). Of these, 21 (Novel1 to Novel21) were high-impact variants, which had not been reported in 1000 Genomes Project Phase 3 (Table 1; Fig 3). However, one of them, Novel6, which corresponds to rs370320936, has been reported in Singaporean and Malaysian (https://www.ncbi.nlm.nih.gov/projects/SNP/snp_retrieve.cgi?subsnp_id=657184749). Therefore, twenty variants were novel. Allele frequencies of 6 variants representing *DPYD* (rs189768576), *CYP3A5* (rs373134805, rs55965422), *SLCO1B1* (rs200994482, rs183624077), and *CYP2D6* (rs147960066) were not significantly different among the six populations (Fig 3B). The remaining eight variants were significantly different between Thais and one or more populations. Three of the eight variants (rs4986893, rs79527462, and rs3892097) were not different between Thais and East Asians but different to the other four populations.

Another variant (rs71358943) was not different among Asians (Thais, East Asian, and South Asian) but different from African, admixed American, and European. Two variants (rs61469810 and rs3833221) had the same allele frequencies among populations except for African, which had a much higher frequency. For rs3215983 and rs11935951, frequencies varied across populations where Thais were different from African, South Asian, and European at rs3215983 and different from African and admixed American at rs11935951.

Discussion

This study reported the star allele profile, diplotypes, and the corresponding predicted phenotypes of 51 pharmacogenes using WGS of 171 unrelated healthy Thais. Compared to the

Table 1. The novel variants (which were not reported in 1000 Genomes Project Phase 3) with high impacts identified in the 171 individuals.

Novel	Genes	Chr	Positions	Reference Allele	Alternate Allele	Allele frequency
1	*CACNA1S*	1	201016674	G	A	0.0058
2	*CFTR*	7	117230480	G	T	0.0029
3	*CYP1A1*	15	75013810	G	GATGGCGACGTACAATCTTTTCAGAAACATCATTAAGAACATCATCAAATGTAAAATAACCATCCTCTGAAGCGAGTTGA	0.0029
4	*CYP2A13*	19	41600328	C	CCTCCCTAA	0.0029
5	*CYP2B6*	19	41497355	AG	A	0.0029
6*	*CYP2C19*	10	96602646	C	A	0.0029
7	*CYP2D6*	22	42523521	C	CCCAAA	0.0029
8	*CYP2D7*	22	42537234	C	CCCAAA	0.0058
9	*CYP2D7*	22	42538514	T	TC	0.4152
10	*CYP2F1*	19	41627939	G	GA	0.0029
11	*CYP2W1*	7	1024201	C	T	0.0029
12	*CYP2W1*	7	1024653	CCT	C	0.0029
13	*CYP3A5*	7	99250232	A	C	0.0029
14	*CYP4F2*	19	15997044	GCCCTCA	G	0.0205
15	*SLC22A2*	6	160671579	C	T	0.0029
16	*SLCO1B1*	12	21294593	G	A	0.0029
17	*SLCO1B3*	12	21015371	CAT	C	0.0029
18	*SLCO1B3*	12	21036518	TTATC	T	0.0058
19	*SLCO1B3*	12	21054403	T	C	0.0029
20	*UGT1A4*	2	234627639	CG	C	0.0088
21	*UGT2B7*	4	69962949	TG	T	0.0029

*Novel6 was identified in Singaporean and Malaysian (https://www.ncbi.nlm.nih.gov/projects/SNP/snp_retrieve.cgi?subsnp_id=657184749) and was annotated as rs370320936

https://doi.org/10.1371/journal.pone.0263621.t001

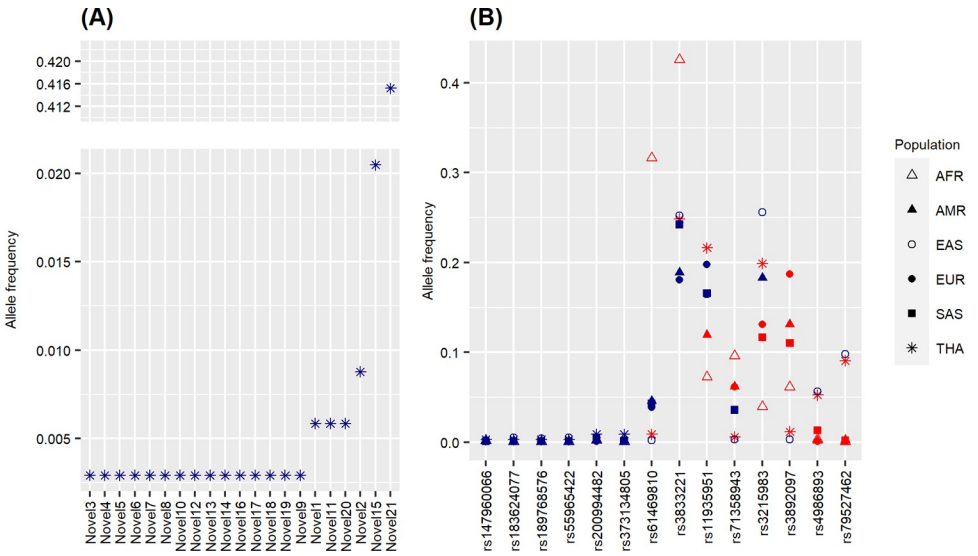

Fig 3. Allele frequency comparisons of the 35 high impact variants among Thais (THA), African (AFR), admixed American (AMR), East Asian (EAS), European (EUR), and South Asian (SAS) were obtained from the 1000 genome project database. (A) Variants whose allele frequencies are not available on the 1000 Genome Project database but found in Thai are denoted with Novel1-Novel21. (B) Δ, ▲, ○, ●, ■, * denote allele frequency of AFR, AMR, EAS, EUR, SAS, and THA, respectively. Allele frequencies that were significantly different ($p < 0.05$) between THA and other populations were shown in red, and the non-statistically significant variants were shown in blue.

https://doi.org/10.1371/journal.pone.0263621.g003

previous report of pharmacogenes in a Thai population recruiting participants with Brugada syndrome endemic in Thailand's Northeastern region [6], our study enrolled participants from all regions of Thailand (S1 Table in S2 File). However, participants were recruited from one hospital, King Chulalongkorn Memorial Hospital, in Bangkok. This mainly results from the fact that all 20 clinical geneticists except one are in Bangkok [16, 17]; therefore, participants enrolled from the Genetics Clinic of the Hospital, a tertiary referral center were from throughout the country. In addition, 26 more genes were added to the 25 genes in the previous study [6]. These newly studied genes are associated with drugs approved by the Thai Food and Drug Administration, Ministry of Public Health, Thailand (https://porta.fda.moph.go.th) and are used as the treatment for the top 10 diseases in Thai people (S4 Table in S2 File).

Of the 51 studied genes, 40 had at least one individual with variant or altered phenotype. This highlights the need for dose alteration of a prescribed drug or an alternative treatment. Pharmacogenetic testing would benefit everyone, as our data showed that no one had a normal phenotype in all 51 genes. In addition, the high impact SNPs allele frequencies were compared to the other five populations. Twenty variants identified have not been reported elsewhere. Therefore, this study provides genes that should be included in a preemptive pharmacogenetic panel for the Thai population.

The thirteen genes where more than 25% of our samples had variant or altered phenotype is involved in the absorption, distribution, metabolism or excretion of various drugs. Therefore, the need for a change in clinical management depends on the genes and drugs. On the contrary, for some genes, e.g., *CYP26A1*, *CACNA1S*, or *TBXAS1*, the majority of the Thai cohort has a normal phenotype.

One hundred sixty-six individuals had either poor- or decreased-function phenotype in *SLCO1B3* due to **S1*. The core SNPs of *SLCO1B3***S1* is rs7311358 (c.699G>A). One study analyzed these SNPs in combination with rs4149117 (c.334T>G) and found that rs7311358 AA +AG genotypes and the rs4149117 GG+GT genotypes were associated with a higher

probability of not responding to the standard dose of imatinib as a treatment in chronic myeloid leukaemia [18]. In addition, carriers of rs4149117(G)-rs7311358(A) haplotype had a significantly lower uptake of mycophenolic acid glucuronide into the cells than the reference haplotype [19].

CYP3A, one of the major CYP families, encodes for various catalyzing enzymes involved in many drug metabolism including nifedipine, cyclosporine, tacrolimus, erythromycin, midazolam, alprazolam, and triazolam [20, 21]. For *CYP3A5*, the *CYP3A5*3* allele results in the absence of protein expression due to an improper mRNA splicing [22]. *CYP3A5*3* (rs776746) was highly present in Caucasians (85–95%), Mexicans (75%), and Asians (65–85%) and to a lower degree in African Americans (27–55%) [23]. The number of Asians agrees with our results (66.7%). Therefore, to achieve therapeutic drug concentrations, an intermediate metabolizer phenotype such as *CYP3A5*1/*3* was recommended to receive an increased dose of tacrolimus (1.5–2 times higher than standard dosing). In contrast, poor metabolizer phenotype such as *CYP3A5*3/*3* should receive the standard dosing of medication based on the tacrolimus package insert [24].

CYP2C19 is well known for the metabolism of particular substances such as omeprazole, imipramine, and diazepam. *CYP2C19*17* is mainly found in the European, African, and admixed American, while *CYP2C19*2* is commonly found in Asia [2]. In our study, *CYP2C19*2* frequency was also high. The homozygous poor metabolizer diplotype (including *CYP2C19*2/*2*) patients showed an efficient eradication of *Heterobacter pylori* when treated with omeprazole and amoxicillin since the drug remains in the bloodstream longer than normal phenotype and they can avoid drug resistance during therapy [25, 26]. For citalopram and escitalopram, intermediate metabolizer phenotype such as *CYP2C19*1/*2* was recommended to initiate therapy with recommended starting dose while poor metabolizer phenotype such as *CYP2C19*2/*2* should receive a 50% dose reduction or select alternative drug not predominantly metabolized by CYP2C19 [27].

CYP2A6 is involved in the metabolism of a large number of xenobiotic and is responsible for approximately 3% of the drugs metabolized by CYP enzymes. *CYP2A6*4*, **7*, **10*, **11*, **19*, and **35* are highly distributed over East Asians, while *CYP2A6*14*, **28*, and **34* have higher frequencies in Caucasian populations [28]. In our data, 3–12% of the population carried decreased-function alleles (*CYP2A6*4*, **7*, **9*, and **10*) where their homozygous and heterozygous diplotypes were predicted as the slow metabolizer. Individuals with *CYP2A6*4*, *7*, **9*, *and *10* were associated with a reduction in nicotine metabolism [29–31] and reduced metabolism of tegafur compared to *CYP2A6* wild type when treated with S-1 (an oral anticancer agent) [32, 33]. In addition, healthy postmenopausal women with *CYP2A6*4*, **7*, **9* had a lower clearance of letrozole than *CYP2A6*1* [34].

SULT1A1 activity plays a crucial role in the metabolism, bioactivation, detoxification of procarcinogens, deactivating catecholamines, and the sulfate conjugation of steroid hormones, including estrogens [35]. *SULT1A1* is in a region with high repetitive sequences and segmental duplication [36], resulting in a highly polymorphic copy number variant (0–6 copies) [37, 38]. A study in Caucasians reported that 64% of the population had two copies, and 32% had $\geq$ 3 copies [37]. The increasing copies of *SULT1A1* were negatively correlated with a conversion of estrone-sulfate to estrone ratio in men [37]. *SULT1A1*2* was common in Caucasians and African Americans but not in Chinese (allele frequencies of 0.332, 0.294, and 0.08, respectively) [39]. Our data agree with those studies (the frequency of 0.07 and ~23% with two copies).

For the X-linked gene, *G6PD*, men have one copy number of the gene while women have two. Therefore, men will be *G6PD* deficient if they inherit only one mutant gene, but women usually need to inherit two abnormal genes [40]. G6PD deficiency has been recognized as a common inherited enzymopathy where G6PD Viangchan (**51*) and G6PD Mahidol variant

(*31) is highly prevalent in Thais [25]. Other G6PD variants such as G6PD Canton (*8), G6PD Kaiping (*28), G6PD Union (*50), and G6PD Gaohe (*153) were found in Chinese, Indian and Southeast Asia populations [41]. Most men carried the normal allele in our cohort except eight individuals carried *G6PD*8*, *21, *28, or *51. The core SNP of *G6PD*8* is rs72554665 (g.153760484C>A). A report showed that after ingesting fava beans, a 26-month-old Chinese-Japanese boy carrying rs72554665(A) allele had severe hemolytic anemia [42]. Moreover, rasburicase is contraindicated for males carrying class I, II or III allele e.g., Canton allele due to the risk of acute hemolytic anemia [43].

CYP4B1 is an interface between the metabolism of xenobiotics such as 2-aminofluorene and endobiotic, including ligands where P450 acts on them to modify endogenous processes [44]. It is predominantly expressed in human lungs and might contribute to carcinogenesis [45]. About 20% of our samples carried *CYP4B1*2*. A study in Japanese found that individuals with *CYP4B1*1/*2 or *2/*2 genotypes had a 1.75-fold increased risk of bladder cancer [46]. The frequencies of the *CYP4B1*2* was 0.328 in Japanese and 0.147 in French Caucasians [47].

CYP2D6 and its pseudogene, *CYP2D7*, are responsible for metabolizing and eliminating >20% of drugs with variability among different ethnicity [48]. *CYP2D6*2* and *CYP2D6*4* are common in Europeans, Africans, South Asians, and admixed Americans, while *CYP2D6*5* (gene deletion) and *CYP2D6*10* are dominantly observed in East Asian [2]. In addition, *CYP2D6*41* were remarkably found in South Asian [2]. Here, we found that *CYP2D6*36+*10* had the highest frequency followed by *CYP2D6*10* in our cohort while *CYP2D6*2*, *4, *5, and *41 were found in a low proportion. For intermediate metabolizers with *CYP2D6*10*, e.g., *CYP2D6*10/*10* and *CYP2D6*10/*41*, a guideline recommends initiating atomoxetine at 40 mg/day and increasing to 80 mg/day after two weeks if no clinical response and in the absence of adverse events [49].

The expression of human *CYP2F1* is found restrictively in lung tissues with the role of metabolizing inhaled compounds [50]. The no-function allele, *CYP2F1*2*, caused by silence mutation was reported at a high frequency, while other variants were less common (0.6–7.2%) in French [50]. The *CYP2F1*2* was also common in our cohort.

Individuals with *NAT2* slow acetylator genotypes (homozygotes or compound heterozygotes for *NAT2*5, *6, or *7) are associated with an increased risk of anti-tuberculosis drug-induced liver injury [51], risk of cotrimoxazole adverse events in patients with systemic lupus erythematosus [52], and risk of sulfasalazine-induced toxicity [53]. However, almost half of our samples were intermediate acetylator.

SLCO2B1 is expressed in the luminal membrane of small-intestinal enterocytes and might play a role in the uptake of drugs from the intestinal lumen [54]. Intermediate and poor function phenotype in our samples was due to carrying *SLCO2B1*S1*, an in-frame deletion variant. The *SLCO2B1*S1* was a new star allele reported by Lee et al. (2019) and was found in only East Asian samples [12].

Genetic polymorphism of *UGT1A1* is associated with diseases such as Gilbert syndrome (*UGT1A1*6/*6*) [55], head and neck cancers [56], colorectal cancer [57], and coronary artery disease [58]. While *UGT1A1*60* was observed in Asians (Korean, Chinese, and Japanese), African Americans, and European Americans [59], *UGT1A1*6* is a common allele in only Asians [60]. Individuals with 1 or 2 alleles of *UGT1A1*6* have a higher risk of irinotecan-induced neutropenia [61, 62]. In addition, the *UGT1A1*6* allele is associated with an increased risk of hyperbilirubinemia in Thai HIV-infected patients treated with indinavir [63]. Besides *UGT1A1*1*, *UGT1A1*60* had the highest allele frequency in our cohort, followed by *UGT1A1*6*.

SLCO1B1 gene encodes for a membrane-bound sodium-independent organic anion transporter protein (OATP1B1) that plays a role in the hepatic uptake of many endogenous and

xenobiotic compounds [64]. Our samples' poor and decreased function phenotypes were primarily due to carrying *SLCO1B1*15* or *SLCO1B1*17*. A guideline for administrating simvastatin for heterozygous (one decrease function allele), e.g., *SLCO1B1*1b/*15*, **1b/*17* or homozygous variant (two decreased-function allele), e.g., *SLCO1B1*15/*15*, **15/*17*, **17/*17* stated that a lower dose or alternative statin should be prescribed because it implies intermediate/high myopathy risk for heterozygous or homozygous respectively [65]. In addition, *SLCO1B1*15* was found to be associated with an increased relative bioavailability of pravastatin [66], exposure to pitavastatin [67], concentrations of repaglinide [68], risk of drug-induced liver injury due to rifampin [69] and associated with decreased clearance of olmesartan [70], metabolism of rosuvastatin [71], transport of atrasentan [72].

Our data confirm similar star allele distribution in both frequency and predicted phenotype to a previous report in the Thai population [6]. The result of *CACNA1S* was the same. Nineteen genes (*CFTR*, *CYP2B6*, *CYP2C8*, *CYP2C9*, *CYP2C19*, *CYP2D6*, *CYP3A4*, *CYP3A5*, *CYP4F2*, *GSTP1*, *NAT1*, *NAT2*, *NUDT15*, *RYR1*, *SLCO1B1*, *TPMT*, *UGT1A1*, *UGT1A4*, and *VKROC1*) had a very similar pattern for both allele frequency and phenotype. However, three genes (*DPYD*, *G6PD*, and *IFNL3*) showed different findings. *DPYD*S12* had the highest allele frequency in our study resulting in intermediate metabolizer phenotype, while *DPYD*1* had the highest allele frequency, and *DPYD*S12* was not found in the previous study [6], resulting in normal phenotype shown. We detected many star alleles in *G6PD*, i.e., *G6PD*1x2*, **21*, **28*, **31*, **50*, **51*, **8*, and **DEL* resulting in intermediate metabolizer phenotype while the previous study [6] detected only *G6PD*1*. For *IFNL3*, both studies detected *IFNL3*1* and *IFNL3*S3*. However, we found *IFNL3*DEL* in one haplotype and the predicted phenotype was different. We reported unfavorable responses while the previous study reported rapid metabolizer phenotype. Since our results for *GSTM1* and *UGT2B15* might not be reliable, we did not compare the results for these two genes.

The three control genes provide consistent results in our case due to their similar read depths. The important thing to note is that the read depth of a control gene and a calling gene should be comparable; otherwise, a duplication/deletion might be reported if a control gene has a lower/higher read depth. In our study, when results disagree between the three control genes, *RYR1* reported a duplication, whereas the other two control genes reported a normal copy. When we inspected read depth, we found that *RYR1* had a lower read depth than others.

The 51 pharmacogenes we studied were all the genes that Stargazer version 1.0.8 can call. This is the highest number of genes able to be called at the present, compared with other tools, e.g., Aldy [73], StellarPGx [74] and Astrolabe [75] which support 35, 13, and 9 genes, respectively. Sixteen genes had PharmGKB level 1A clinical annotation for specific variant-drug pairs (S2 Table in S2 File). Genes with levels 3 or 4 clinical annotations might be nontrivial as some variants were more frequent in our population than other populations; therefore, the variant-drug pairs might be understudied in our context.

Our data is from a short-read sequencing technique whose accuracy for calling structural variants, e.g., copy number variation, might be suboptimal. This might affect the results of some genes, e.g., *CYP2D6*.

Our study recruited healthy parents of children with rare diseases; thus, more prevalent pathogenic variants for rare diseases are expected, as observed in the Thai Reference Exome (T-REx) Database with similar inclusion criteria [76]. T-REx recruited unaffected parents of children with rare diseases and found a more enriched pathogenic variants, compared with gnomAD. However, this study did not recruit any patients with diseases related to pharmacogenes such as G6PD deficiency. We intentionally recruited only parents of patients with 1) *de novo* mutations such as achondroplasia or 2) diseases unrelated to these 51 pharmacogenes such as methylmalonic academia and Duchenne muscular dystrophy. Therefore, the increased

prevalence of the pathogenic variants in these healthy parents or even the patients themselves should not affect this study conclusion.

In conclusion, our study provides pharmacogenomics implications for drug prescription and guidance of population-specific genotyping policy to improve drug response or prevent adverse drug reactions for Thais. In addition, the study provides genes with a high proportion of the population expressing variant or altered phenotypes, which should be included in a pre-emptive pharmacogenomics panel.

Supporting information

S1 File. This file contains S1-S5 Figs.
(DOCX)

S2 File. This file contains S1-S4 Tables.
(DOCX)

Acknowledgments

We gratefully thank Dr Chureerat Phokaew for facilitating the HPC.

Author Contributions

Conceptualization: Pajaree Chariyavilaskul, Monpat Chamnanphon, Monnat Pongpanich, Vorasuk Shotelersuk.

Data curation: Adjima Assawapitaksakul.

Formal analysis: Natnicha Wankaew.

Funding acquisition: Vorasuk Shotelersuk.

Resources: Wanna Chetruengchai, Vorasuk Shotelersuk.

Supervision: Pajaree Chariyavilaskul, Monnat Pongpanich, Vorasuk Shotelersuk.

Writing – original draft: Natnicha Wankaew, Monnat Pongpanich.

Writing – review & editing: Pajaree Chariyavilaskul, Monnat Pongpanich, Vorasuk Shotelersuk.

References

1. Oates JT, Lopez D. Pharmacogenetics: An Important Part of Drug Development with A Focus on Its Application. Int J Biomed Investig. 2018; 1(2):111. Epub 2018/05/27. https://doi.org/10.31531/2581-4745.1000111 PMID: 32467882.
2. Zhou Y, Ingelman-Sundberg M, Lauschke VM. Worldwide Distribution of Cytochrome P450 Alleles: A Meta-analysis of Population-scale Sequencing Projects. Clin Pharmacol Ther 2017; 102(4):688–700. Epub 2017/05/26. https://doi.org/10.1002/cpt.690 PMID: 28378927.
3. Adrián LL, Naranjo M, Rodrigues-Soares F, Penas LE, Fariñas H, Tarazona-Santos E. Interethnic variability of CYP2D6 alleles and of predicted and measured metabolic phenotypes across world populations. Expert Opin Drug Metab Toxicol. 2014; 10(11):1569–1583. Epub 2014/10/16. https://doi.org/10.1517/17425255.2014.964204 PMID: 25316321.
4. Zhang H, De T, Zhong Y, Perera MA. The Advantages and Challenges of Diversity in Pharmacogenomics: Can Minority Populations Bring Us Closer to Implementation? Clin Pharmacol Ther Clinical pharmacology and therapeutics. 2019; 106(2):338–349. https://doi.org/10.1002/cpt.1491 PMID: 31038731.
5. Del Tredici AL, Malhotra A, Dedek M, Espin F, Roach D, Zhu G-D, et al. Frequency of CYP2D6 Alleles Including Structural Variants in the United States. Front Pharmacol. 2018; 9:305–305. https://doi.org/10.3389/fphar.2018.00305 PMID: 29674966.

6. Mauleekoonphairoj J, Chamnanphon M, Khongphatthanayothin A, Sutjaporn B, Wandee P, Poovorawan Y, et al. Phenotype prediction and characterization of 25 pharmacogenes in Thais from whole genome sequencing for clinical implementation. Sci Rep. 2020; 10(1):18969. https://doi.org/10.1038/s41598-020-76085-3 PMID: 33144648

7. Cochran WG. Sampling techniques. 3rd ed. New York: John Wiley & Sons; 1977.

8. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 2009; 25(14):1754–1760. https://doi.org/10.1093/bioinformatics/btp324 PMID: 19451168

9. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics. 2013; 43(1110):11.10.11–11.10.33. Epub 2014/11/29. https://doi.org/10.1002/0471250953.bi1110s43 PMID: 25431634; PubMed Central PMCID: PMC4243306.

10. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011; 43(5):491–498. Epub 2011/04/10. https://doi.org/10.1038/ng.806 PMID: 21478889.

11. Lee SB, Wheeler MM, Patterson K, McGee S, Dalton R, Woodahl EL, et al. Stargazer: a software tool for calling star alleles from next-generation sequencing data using CYP2D6 as a model. Genet Med. 2019; 21(2):361–372. Epub 2018/06/08. https://doi.org/10.1038/s41436-018-0054-0 PMID: 29875422; PubMed Central PMCID: PMC6281872.

12. Lee SB, Wheeler MM, Thummel KE, Nickerson DA. Calling Star Alleles With Stargazer in 28 Pharmacogenes With Whole Genome Sequences. Clin Pharmacol Ther Clinical pharmacology and therapeutics. 2019; 106(6):1328–1337. Epub 2019/07/26. https://doi.org/10.1002/cpt.1552 PMID: 31206625.

13. Browning BL, Zhou Y, Browning SR. A One-Penny Imputed Genome from Next-Generation Reference Panels. Am J Hum Genet. 2018; 103(3):338–348. https://doi.org/10.1016/j.ajhg.2018.07.015 PMID: 30100085

14. Browning S, Browning B. Rapid and Accurate Haplotype Phasing and Missing-Data Inference for Whole-Genome Association Studies By Use of Localized Haplotype Clustering. Am J Hum Genet. 2007; 81(5):1084–1097. https://doi.org/10.1086/521987 PMID: 17924348

15. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. Behav Brain Res. 2001; 125(1–2):279–284. Epub 2001/10/30. https://doi.org/10.1016/s0166-4328(01)00297-2 PMID: 11682119.

16. Shotelersuk V, Tongsima S, Pithukpakorn M, Eu-Ahsunthornwattana J, Mahasirimongkol S. Precision medicine in Thailand. American journal of medical genetics Part C, Seminars in medical genetics. 2019; 181(2):245–253. Epub 2019/03/20. https://doi.org/10.1002/ajmg.c.31694 PMID: 30888117.

17. Shotelersuk V, Limwongse C, Mahasirimongkol S. Genetics and genomics in Thailand: challenges and opportunities. Molecular genetics & genomic medicine. 2014; 2(3):210–216. Epub 2014/06/18. https://doi.org/10.1002/mgg3.83 PMID: 24936510; PubMed Central PMCID: PMC4049361.

18. De Lima LT, Bueno CT, Vivona D, Hirata RD, Hirata MH, Hungria VT, et al. Relationship between SLCO1B3 and ABCA3 polymorphisms and imatinib response in chronic myeloid leukemia patients. Hematology (Amsterdam, Netherlands). 2015; 20(3):137–142. Epub 2014/07/25. https://doi.org/10.1179/1607845414Y.0000000181 PMID: 25056761.

19. Picard N, Yee SW, Woillard JB, Lebranchu Y, Le Meur Y, Giacomini KM, et al. The role of organic anion-transporting polypeptides and their common genetic variants in mycophenolic acid pharmacokinetics. Clin Pharmacol Ther Clinical pharmacology and therapeutics. 2010; 87(1):100–108. Epub 2009/11/04. https://doi.org/10.1038/clpt.2009.205 PMID: 19890249.

20. Thervet E, Anglicheau D, King B, Schlageter MH, Cassinat B, Beaune P, et al. Impact of cytochrome p450 3A5 genetic polymorphism on tacrolimus doses and concentration-to-dose ratio in renal transplant recipients. Transplantation. 2003; 76(8):1233–1235. Epub 2003/10/28. https://doi.org/10.1097/01.TP.0000090753.99170.89 PMID: 14578760.

21. Bjornsson T, Wagner J, Donahue S, Harper D, Karim A, Khouri M, et al. A review and assessment of potential sources of ethnic differences in drug responsiveness. Journal of clinical pharmacology. 2003; 43(9):943–967. Epub 2003/09/16. https://doi.org/10.1177/0091270003256065 PMID: 12971027.

22. Liu H, Xu Q, Huang W, Zhao Q, Jiang Z, Kuang X, et al. CYP3A5 and CYP3A7 genetic polymorphisms affect tacrolimus concentration in pediatric patients with nephrotic range proteinuria. European journal of clinical pharmacology. 2019; 75(11):1533–1540. https://doi.org/10.1007/s00228-019-02726-w PMID: 31401678

23. Zhang JJ, Zhang H, Ding XL, Ma S, Miao LY. Effect of the P450 oxidoreductase 28 polymorphism on the pharmacokinetics of tacrolimus in Chinese healthy male volunteers. European journal of clinical pharmacology. 2013; 69(4):807–812. Epub 2012/10/26. https://doi.org/10.1007/s00228-012-1432-1 PMID: 23097010.

24. Birdwell K, Decker B, Barbarino J, Peterson J, Stein C, Sadee W, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guidelines for CYP3A5 Genotype and Tacrolimus Dosing. Clin Pharmacol Ther Clinical pharmacology and therapeutics. 2015; 98(1):19–24. Epub 2015/03/25. https://doi.org/10.1002/cpt.113 PMID: 25801146; PubMed Central PMCID: PMC4481158.

25. Tanigawara Y, Aoyama N, Kita T, Shirakawa K, Komada F, Kasuga M, et al. CYP2C19 genotype-related efficacy of omeprazole for the treatment of infection caused by Helicobacter pylori. Clin Pharmacol Ther Clinical pharmacology and therapeutics. 1999; 66(5):528–534. Epub 1999/12/01. https://doi.org/10.1016/S0009-9236(99)70017-2 PMID: 10579481.

26. Sagar M, Bertilsson L, Stridsberg M, Kjellin A, Mårdh S, Seensalu R. Omeprazole and CYP2C19 polymorphism: effects of long-term treatment on gastrin, pepsinogen I, and chromogranin A in patients with acid related disorders. Aliment Pharmacol Ther. 2000; 14(11):1495–1502. Epub 2000/11/09. https://doi.org/10.1046/j.1365-2036.2000.00835.x PMID: 11069321.

27. Hicks J, Bishop J, Sangkuhl K, Müller D, Ji Y, Leckband S, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2D6 and CYP2C19 Genotypes and Dosing of Selective Serotonin Reuptake Inhibitors. Clin Pharmacol Ther Clinical pharmacology and therapeutics. 2015; 98 (2):127–134. Epub 2015/05/15. https://doi.org/10.1002/cpt.147 PMID: 25974703; PubMed Central PMCID: PMC4512908.

28. López-Flores LA, Pérez-Rubio G, Falfán-Valencia R. Distribution of polymorphic variants of CYP2A6 and their involvement in nicotine addiction. EXCLI J. 2017; 16:174–196. Epub 2017/05/17. https://doi.org/10.17179/excli2016-847 PMID: 28507465; PubMed Central PMCID: PMC5427481.

29. Nakajima M, Fukami T, Yamanaka H, Higashi E, Sakai H, Yoshida R, et al. Comprehensive evaluation of variability in nicotine metabolism and CYP2A6 polymorphic alleles in four ethnic populations. Clin Pharmacol Ther Clinical pharmacology and therapeutics. 2006; 80(3):282–297. Epub 2006/09/06. https://doi.org/10.1016/j.clpt.2006.05.012 PMID: 16952495.

30. Peamkrasatam S, Sriwatanakul K, Kiyotani K, Fujieda M, Yamazaki H, Kamataki T, et al. In vivo evaluation of coumarin and nicotine as probe drugs to predict the metabolic capacity of CYP2A6 due to genetic polymorphism in Thais. Drug Metab Pharmacokinet. 2006; 21(6):475–484. Epub 2007/01/16. https://doi.org/10.2133/dmpk.21.475 PMID: 17220563.

31. Yoshida R, Nakajima M, Watanabe Y, Kwon JT, Yokoi T. Genetic polymorphisms in human CYP2A6 gene causing impaired nicotine metabolism. Br J Clin Pharmacol. 2002; 54(5):511–517. Epub 2002/11/26. https://doi.org/10.1046/j.1365-2125.2002.01667.x PMID: 12445030; PubMed Central PMCID: PMC1874463.

32. Fujita K, Yamamoto W, Endo S, Endo H, Nagashima F, Ichikawa W, et al. CYP2A6 and the plasma level of 5-chloro-2, 4-dihydroxypyridine are determinants of the pharmacokinetic variability of tegafur and 5-fluorouracil, respectively, in Japanese patients with cancer given S-1. Cancer science. 2008; 99 (5):1049–1054. Epub 2008/04/03. https://doi.org/10.1111/j.1349-7006.2008.00773.x PMID: 18380793.

33. Kim KP, Jang G, Hong YS, Lim HS, Bae KS, Kim HS, et al. Phase II study of S-1 combined with oxaliplatin as therapy for patients with metastatic biliary tract cancer: influence of the CYP2A6 polymorphism on pharmacokinetics and clinical activity. Br J Cancer. 2011; 104(4):605–612. Epub 2011/02/18. https://doi.org/10.1038/bjc.2011.17 PMID: 21326246; PubMed Central PMCID: PMC3049596.

34. Tanii H, Shitara Y, Horie T. Population pharmacokinetic analysis of letrozole in Japanese postmenopausal women. European journal of clinical pharmacology. 2011; 67(10):1017–1025. Epub 2011/04/16. https://doi.org/10.1007/s00228-011-1042-3 PMID: 21494765.

35. Hildebrandt M, Adjei A, Weinshilboum R, Johnson JA, Berlin DS, Klein TE, et al. Very important pharmacogene summary: sulfotransferase 1A1. Pharmacogenet Genomics. 2009; 19(6):404–406. Epub 2009/05/20. https://doi.org/10.1097/FPC.0b013e32832e042e PMID: 19451861.

36. Martin J, Han C, Gordon LA, Terry A, Prabhakar S, She X, et al. The sequence and analysis of duplication-rich human chromosome 16. Nature. 2004; 432(7020):988–994. Epub 2004/12/24. https://doi.org/10.1038/nature03187 PMID: 15616553.

37. Liu J, Zhao R, Ye Z, Frey AJ, Schriver ER, Snyder NW, et al. Relationship of SULT1A1 copy number variation with estrogen metabolism and human health. J Steroid Biochem Mol Biol. 2017; 174:169–175. Epub 2017/09/05. https://doi.org/10.1016/j.jsbmb.2017.08.017 PMID: 28867356; PubMed Central PMCID: PMC5675753.

38. Hebbring SJ, Adjei AA, Baer JL, Jenkins GD, Zhang J, Cunningham JM, et al. Human SULT1A1 gene: copy number differences and functional implications. Hum Mol Genet. 2007; 16(5):463–470. Epub 2006/12/26. https://doi.org/10.1093/hmg/ddl468 PMID: 17189289.

39. Carlini EJ, Raftogianis RB, Wood TC, Jin F, Zheng W, Rebbeck TR, et al. Sulfation pharmacogenetics: SULT1A1 and SULT1A2 allele frequencies in Caucasian, Chinese and African-American subjects. Pharmacogenetics. 2001; 11(1):57–68. Epub 2001/02/24. https://doi.org/10.1097/00008571-200102000-00007 PMID: 11207031.

40. Guindo A, Fairhurst RM, Doumbo OK, Wellems TE, Diallo DA. X-linked G6PD deficiency protects hemizygous males but not heterozygous females against severe malaria. PLoS Med. 2007; 4(3):e66. Epub 2007/03/16. https://doi.org/10.1371/journal.pmed.0040066 PMID: 17355169; PubMed Central PMCID: PMC1820604.

41. Jitueakul S, Buncherd H, Thawornpan P, Tun A, Thanapongpichat S. Characterization of G6PD genotypes in G6PD deficiency patients from Suratthani Hospital, Thailand. J Assoc Med Sci. 2018; 51.

42. Shibuya A, Hirono A, Ishii S, Fujii H, Miwa S. Hemolytic crisis after excessive ingestion of fava beans in a male infant with G6PD Canton. Int J Hematol. 1999; 70(4):233–235. Epub 2000/01/22. PMID: 10643148.

43. Relling M, McDonagh E, Chang T, Caudle K, McLeod H, Haidar C, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guidelines for Rasburicase Therapy in the Context of G6PD Deficiency Genotype. Clin Pharmacol Ther Clinical pharmacology and therapeutics. 2014; 96(2):169–174. Epub 2014/05/03. https://doi.org/10.1038/clpt.2014.97 PMID: 24787449; PubMed Central PMCID: PMC4111801.

44. Baer BR, Rettie AE. CYP4B1: An Enigmatic P450 at the Interface between Xenobiotic and Endobiotic Metabolism. Drug Metab Rev. 2006; 38(3):451–476. https://doi.org/10.1080/03602530600688503 PMID: 16877261.

45. Lim S, Alshagga M, Ong C, Chieng J, Pan Y. Cytochrome P450 4B1 (CYP4B1) as a target in cancer treatment. Human & experimental toxicology. 2020; 39(6):785–796. Epub 2020/02/15. https://doi.org/10.1177/0960327120905959 PMID: 32054340.

46. Sasaki T, Horikawa M, Orikasa K, Sato M, Arai Y, Mitachi Y, et al. Possible Relationship Between the Risk of Japanese Bladder Cancer Cases and the CYP4B1 Genotype. Jpn J Clin Oncol. 2008; 38 (9):634–640. https://doi.org/10.1093/jjco/hyn081 PMID: 18713828

47. Hiratsuka M, Nozawa H, Konno Y, Saito T, Konno S, Mizugaki M. Human CYP4B1 Gene in the Japanese Population Analyzed by Denaturing HPLC. Drug Metab Pharmacokinet. 2004; 19(2):114–119. https://doi.org/10.2133/dmpk.19.114 PMID: 15499177

48. Leitão LPC, Souza TP, Rodrigues JCG, Fernandes MR, Santos S, Santos NPC. The Metabolization Profile of the CYP2D6 Gene in Amerindian Populations: A Review. Genes (Basel). 2020; 11(3). Epub 2020/03/04. https://doi.org/10.3390/genes11030262 PMID: 32121156; PubMed Central PMCID: PMC7140882.

49. Brown J, Bishop J, Sangkuhl K, Nurmi E, Mueller D, Dinh J, et al. Clinical Pharmacogenetics Implementation Consortium Guideline for Cytochrome P450 (CYP)2D6 Genotype and Atomoxetine Therapy. Clin Pharmacol Ther Clinical pharmacology and therapeutics. 2019; 106(1):94–102. Epub 2019/02/26. https://doi.org/10.1002/cpt.1409 PMID: 30801677; PubMed Central PMCID: PMC6612570.

50. Tournel G, Cauffiez C, Billaut-Laden I, Allorge D, Chevalier D, Bonnifet F, et al. Molecular analysis of the CYP2F1 gene: identification of a frequent non-functional allelic variant. Mutat Res. 2007; 617(1–2):79–89. Epub 2007/03/01. https://doi.org/10.1016/j.mrfmmm.2007.01.007 PMID: 17327131.

51. Suvichapanich S, Fukunaga K, Zahroh H, Mushiroda T, Mahasirimongkol S, Toyo-Oka L, et al. NAT2 ultra-slow acetylator and risk of anti-tuberculosis drug-induced liver injury: a genotype-based meta-analysis. Pharmacogenet Genomics. 2018; 28(7):167–176. Epub 2018/05/22. https://doi.org/10.1097/FPC.0000000000000339 PMID: 29781872.

52. Soejima M, Sugiura T, Kawaguchi Y, Kawamoto M, Katsumata Y, Takagi K, et al. Association of the diplotype configuration at the N-acetyltransferase 2 gene with adverse events with co-trimoxazole in Japanese patients with systemic lupus erythematosus. Arthritis Res Ther. 2007; 9(2):R23. Epub 2007/03/06. https://doi.org/10.1186/ar2134 PMID: 17335581; PubMed Central PMCID: PMC1906798.

53. Yee J, Kim SM, Han JM, Lee N, Yoon HY, Gwak HS. The association between NAT2 acetylator status and adverse drug reactions of sulfasalazine: a systematic review and meta-analysis. Sci Rep. 2020; 10 (1):3658. https://doi.org/10.1038/s41598-020-60467-8 PMID: 32107440

54. Shitara Y, Maeda K, Ikejiri K, Yoshida K, Horie T, Sugiyama Y. Clinical significance of organic anion transporting polypeptides (OATPs) in drug disposition: their roles in hepatic clearance and intestinal absorption. Int J Hematol. 2013; 34(1):45–78. https://doi.org/10.1002/bdd.1823 PMID: 23115084.

55. Yamamoto K, Sato H, Fujiyama Y, Doida Y, Bamba T. Contribution of two missense mutations (G71R and Y486D) of the bilirubin UDP glycosyltransferase (UGT1A1) gene to phenotypes of Gilbert's syndrome and Crigler-Najjar syndrome type II. Biochim Biophys Acta. 1998; 1406(3):267–273. Epub 1998/06/19. https://doi.org/10.1016/s0925-4439(98)00013-1 PMID: 9630669.

56. Lacko M, Roelofs HM, Te Morsche RH, Voogd AC, Ophuis MB, Peters WH, et al. Genetic polymorphism in the conjugating enzyme UGT1A1 and the risk of head and neck cancer. Int J Cancer. 2010; 127(12):2815–2821. Epub 2011/02/26. https://doi.org/10.1002/ijc.25296 PMID: 21351260.

57. Tang KS, Chiu HF, Chen HH, Eng HL, Tsai CJ, Teng HC, et al. Link between colorectal cancer and polymorphisms in the uridine-diphosphoglucuronosyltransferase 1A7 and 1A1 genes. World J

Gastroenterol. 2005; 11(21):3250–3254. Epub 2005/06/02. https://doi.org/10.3748/wjg.v11.i21.3250 PMID: 15929176; PubMed Central PMCID: PMC4316057.

58. Lingenhel A, Kollerits B, Schwaiger JP, Hunt SC, Gress R, Hopkins PN, et al. Serum bilirubin levels, UGT1A1 polymorphisms and risk for coronary artery disease. Exp Gerontol. 2008; 43(12):1102–1107. Epub 2008/09/16. https://doi.org/10.1016/j.exger.2008.08.047 PMID: 18790042; PubMed Central PMCID: PMC2648823.

59. Kim JY, Cheong HS, Park BL, Kim LH, Namgoong S, Kim JO, et al. Comprehensive variant screening of the UGT gene family. Yonsei Med J. 2014; 55(1):232–239. Epub 2013/12/18. https://doi.org/10.3349/ymj.2014.55.1.232 PMID: 24339312; PubMed Central PMCID: PMC3874916.

60. Akaba K, Kimura T, Sasaki A, Tanabe S, Wakabayashi T, Hiroi M, et al. Neonatal hyperbilirubinemia and a common mutation of the bilirubin uridine diphosphate-glucuronosyltransferase gene in Japanese. J Hum Genet. 1999; 44(1):22–25. Epub 1999/02/04. https://doi.org/10.1007/s100380050100 PMID: 9929972

61. Onoue M, Terada T, Kobayashi M, Katsura T, Matsumoto S, Yanagihara K, et al. UGT1A1*6 polymorphism is most predictive of severe neutropenia induced by irinotecan in Japanese cancer patients. International journal of clinical oncology. 2009; 14(2):136–142. Epub 2009/04/25. https://doi.org/10.1007/s10147-008-0821-z PMID: 19390945.

62. Jada SR, Lim R, Wong CI, Shu X, Lee SC, Zhou Q, et al. Role of UGT1A1*6, UGT1A1*28 and ABCG2 c.421C>A polymorphisms in irinotecan-induced neutropenia in Asian cancer patients. Cancer science. 2007; 98(9):1461–1467. Epub 2007/07/14. https://doi.org/10.1111/j.1349-7006.2007.00541.x PMID: 17627617.

63. Boyd M, Srasuebkul P, Ruxrungtham K, Mackenzie P, Uchaipichat V, Stek M Jr., et al. Relationship between hyperbilirubinaemia and UDP-glucuronosyltransferase 1A1 (UGT1A1) polymorphism in adult HIV-infected Thai patients treated with indinavir. Pharmacogenet Genomics. 2006; 16(5):321–329. Epub 2006/04/13. https://doi.org/10.1097/01.fpc.0000197465.14340.d4 PMID: 16609363.

64. Niemi M, Pasanen MK, Neuvonen PJ. Organic Anion Transporting Polypeptide 1B1: a Genetically Polymorphic Transporter of Major Importance for Hepatic Drug Uptake. Pharmacological reviews. 2011; 63 (1):157–181. Epub 2011/01/20. https://doi.org/10.1124/pr.110.002857 PMID: 21245207.

65. Ramsey LB, Johnson SG, Caudle KE, Haidar CE, Voora D, Wilke RA, et al. The Clinical Pharmacogenetics Implementation Consortium Guideline for SLCO1B1 and Simvastatin-Induced Myopathy: 2014 Update. Clin Pharmacol Ther Clinical pharmacology and therapeutics. 2014; 96(4):423–428. Epub 2014/06/12. https://doi.org/10.1038/clpt.2014.125 PMID: 24918167; PubMed Central PMCID: PMC4169720.

66. Ide T, Sasaki T, Maeda K, Higuchi S, Sugiyama Y, Ieiri I. Quantitative Population Pharmacokinetic Analysis of Pravastatin Using an Enterohepatic Circulation Model Combined With Pharmacogenomic Information on SLCO1B1 and ABCC2 Polymorphisms. Journal of clinical pharmacology. 2009; 49 (11):1309–1317. Epub 2009/09/25. https://doi.org/10.1177/0091270009341960 PMID: 19776292.

67. Oh ES, Kim CO, Cho SK, Park MS, Chung JY. Impact of ABCC2, ABCG2 and SLCO1B1 polymorphisms on the pharmacokinetics of pitavastatin in humans. Drug Metab Pharmacokinet. 2013; 28 (3):196–202. Epub 2012/09/26. https://doi.org/10.2133/dmpk.dmpk-12-rg-068 PMID: 23007012.

68. Pei Q, Liu JY, Yin JY, Yang GP, Liu SK, Zheng Y, et al. Repaglinide-irbesartan drug interaction: effects of SLCO1B1 polymorphism on repaglinide pharmacokinetics and pharmacodynamics in Chinese population. European journal of clinical pharmacology. 2018; 74(8):1021–1028. Epub 2018/05/12. https://doi.org/10.1007/s00228-018-2477-6 PMID: 29748863.

69. Li LM, Chen L, Deng GH, Tan WT, Dan YJ, Wang RQ, et al. SLCO1B1 *15 haplotype is associated with rifampin-induced liver injury. Molecular medicine reports. 2012; 6(1):75–82. Epub 2012/05/09. https://doi.org/10.3892/mmr.2012.900 PMID: 22562052.

70. Suwannakul S, Ieiri I, Kimura M, Kawabata K, Kusuhara H, Hirota T, et al. Pharmacokinetic interaction between pravastatin and olmesartan in relation to SLCO1B1 polymorphism. J Hum Genet. 2008; 53 (10):899–904. https://doi.org/10.1007/s10038-008-0324-9 PMID: 18641915

71. Choi J, Lee M, Cho J-Y, Lee J-E, Kim K, Park K. Influence of OATP1B1 Genotype on the Pharmacokinetics of Rosuvastatin in Koreans. Clin Pharmacol Ther Clinical pharmacology and therapeutics. 2008; 83(2):251–257. Epub 2007/06/15. https://doi.org/10.1038/sj.clpt.6100267 PMID: 17568401

72. Katz DA, Carr R, Grimm DR, Xiong H, Holley-Shanks R, Mueller T, et al. Organic anion transporting polypeptide 1B1 activity classified by SLCO1B1 genotype influences atrasentan pharmacokinetics. Clin Pharmacol Ther Clinical pharmacology and therapeutics. 2006; 79(3):186–196. Epub 2006/03/04. https://doi.org/10.1016/j.clpt.2005.11.003 PMID: 16513443

73. Numanagić I, Malikić S, Ford M, Qin X, Toji L, Radovich M, et al. Allelic decomposition and exact genotyping of highly polymorphic and structurally variant genes. Nature Communications. 2018; 9(1):828. https://doi.org/10.1038/s41467-018-03273-1 PMID: 29483503

74. Twesigomwe D, Drögemöller BI, Wright GEB, Siddiqui A, da Rocha J, Lombard Z, et al. StellarPGx: A Nextflow Pipeline for Calling Star Alleles in Cytochrome P450 Genes. Clinical Pharmacology & Therapeutics. 2021; 110(3):741–749. https://doi.org/10.1002/cpt.2173 PMID: 33492672

75. Twist GP, Gaedigk A, Miller NA, Farrow EG, Willig LK, Dinwiddie DL, et al. Constellation: a tool for rapid, automated phenotype assignment of a highly polymorphic pharmacogene, CYP2D6, from whole-genome sequences. npj Genomic Medicine. 2016; 1(1):15007. https://doi.org/10.1038/npjgenmed.2015.7 PMID: 29263805

76. Shotelersuk V, Wichadakul D, Ngamphiw C, Srichomthong C, Phokaew C, Wilantho A, et al. The Thai reference exome (T-REx) variant database. Clin Genet. 2021; 100(6):703–712. Epub 20210922. https://doi.org/10.1111/cge.14060 PMID: 34496037.