



## OPEN ACCESS

EDITED BY  
Michał Korostynski,  
Institute of Pharmacology PAS, Poland

REVIEWED BY  
Katarzyna Tomala,  
Jagiellonian University, Poland  
Yitian Zhou,  
Karolinska Institutet (KI), Sweden

\*CORRESPONDENCE  
Jan-Gowth Chang,  
d6781@mail.cmu.org.tw  
Fuu-Jen Tsai,  
d0704@mail.cnuh.org.tw

†These authors share senior authorship

SPECIALTY SECTION  
This article was submitted to  
Pharmacogenetics and  
Pharmacogenomics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 20 May 2022  
ACCEPTED 06 July 2022  
PUBLISHED 25 August 2022

CITATION  
Lu H-F, Liu T-Y, Chou Y-P, Chang S-S,  
Hsieh Y-W, Chang J-G and Tsai F-J  
(2022), Comprehensive  
characterization of pharmacogenes in a  
Taiwanese Han population.  
*Front. Genet.* 13:948616.  
doi: 10.3389/fgene.2022.948616

COPYRIGHT  
© 2022 Lu, Liu, Chou, Chang, Hsieh,  
Chang and Tsai. This is an open-access  
article distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does  
not comply with these terms.

# Comprehensive characterization of pharmacogenes in a Taiwanese Han population

Hsing-Fang Lu<sup>1</sup>, Ting-Yuan Liu<sup>2</sup>, Yu-Pao Chou<sup>2,3</sup>,  
Shih-Sheng Chang<sup>4,5</sup>, Yow-Wen Hsieh<sup>6,7</sup>,  
Jan-Gowth Chang<sup>2,8\*†</sup> and Fuu-Jen Tsai<sup>1,9,10,11\*†</sup>

<sup>1</sup>Department of Medical Research, China Medical University Hospital, Taichung, Taiwan, <sup>2</sup>Center for Precision Medicine, China Medical University Hospital, Taichung, Taiwan, <sup>3</sup>Department of Laboratory Medicine, China Medical University Hospital, Taichung, Taiwan, <sup>4</sup>Division of Cardiovascular Medicine, China Medical University Hospital, Taichung, Taiwan, <sup>5</sup>School of Medicine, College of Medicine, China Medical University, Taichung, Taiwan, <sup>6</sup>Department of Pharmacy, China Medical University Hospital, Taichung, Taiwan, <sup>7</sup>School of Pharmacy, College of Pharmacy, China Medical University, Taichung, Taiwan, <sup>8</sup>Epigenome Research Center, China Medical University Hospital, Taichung, Taiwan, <sup>9</sup>School of Chinese Medicine, China Medical University, Taichung, Taiwan, <sup>10</sup>Division of Medical Genetics, Children's Hospital of China Medical University, Taichung, Taiwan, <sup>11</sup>Department of Biotechnology and Bioinformatics, Asia University, Taichung, Taiwan

Pharmacogenetic (PGx) testing has not been well adopted in current clinical practice. The phenotypic distribution of clinically relevant pharmacogenes remains to be fully characterized in large population cohorts. In addition, no study has explored actionable PGx alleles in the East Asian population at a large scale. This study comprehensively analyzed 14 actionable pharmacogene diplotypes and phenotypes in 172,854 Taiwanese Han individuals by using their genotype data. Furthermore, we analyzed data from electronic medical records to investigate the effect of the actionable phenotypes on the individuals. The PGx phenotype frequencies were comparable between our cohort and the East Asian population. Overall, 99.9% of the individuals harbored at least one actionable PGx phenotype, and 29% of them have been prescribed a drug to which they may exhibit an atypical response. Our findings can facilitate the clinical application of PGx testing and the optimization of treatment and dosage individually.

## KEYWORDS

pharmacogenetics, CYP, HLA, population genetics, SNP microarray

## Introduction

Pharmacogenetics (PGx) is the study of how individuals respond differently to drugs based on their genes (Lavertu et al., 2018). Genetic guidance for selecting suitable treatment drugs is a key components of precision medicine, especially for identifying responders and nonresponders to a certain treatment, selecting the appropriate dosage, and decreasing the risk of adverse events. To reduce the barrier to implementing PGx in clinical practice, the Clinical Pharmacogenetics Implementation Consortium (CPIC) has provided updated, evidence-based, and peer-reviewed guidelines on PGx based on

published clinical annotation (Relling and Klein, 2011). The US Food and Drug Administration (FDA) has evaluated gene–drug interactions and has presented adequate scientific evidence in the form of three tables of PGx associations. These tables, based on genetic markers, guides health-care providers to implement appropriate measures to ensure the safety and efficacy of prescribed medications (Kim et al., 2021).

To standardize the interpretation of gene–drug associations, methods used to identify pharmacogene haplotypes must be consistent among different laboratories to enable the clinical application of the drug (Tayeh et al., 2022). The “star-allele nomenclature” is used to predict the activity or function of pharmacogenes (Robarge et al., 2007). Several pharmacogene naming tools are used to construct the PGx phenotype by using genome sequencing data according to the standardized nomenclature (Twist et al., 2016; Numanagić et al., 2018; Lee et al., 2019). However, because of the small sample size of sequencing studies, determining the PGx star allele frequency distribution in the general population would be difficult (Mauleekoonphairoj et al., 2020). Compared with whole-genome sequencing (WGS) data, the use of single-nucleotide variation (SNV; formerly single-nucleotide polymorphism) array data is more affordable and suitable for biobank studies. The Pharmacogenomics Clinical Annotation Tool (PharmCAT) and the Population-scale Pharmacogenetic Allele and Phenotype Caller (PGxPOP) can be used to construct the PGx gene star allele and determine the phenotype annotation by incorporating definitions from the CPIC database and using genotyping array data as input (Sanguhl et al., 2020; McInnes et al., 2021). This method generates consolidated and reproducible results, which can be easily compared to interpret the PGx star allele results.

In 2021, the UK Biobank released the PGx allele and phenotype frequencies for clinically relevant genes in 487,409 individuals (McInnes et al., 2021). Most of the participants harbored active PGx variants that may affect treatment outcomes. Besides, nearly a quarter of individuals have been prescribed a medication when they possessed actionable pharmacogene variants toward the medication. Therefore, characterizing PGx phenotypes in larger populations is crucial to determine treatment outcomes. Moreover, to identify the effect of genetic variations on drug responses, the phenotypic constitution of pharmacogenes must be elucidated.

Wang et al. (2021) reported clinical PGx testing results for a Chinese population and demonstrated the benefit of PGx-guided treatment among certain drugs. The Japan PGx Data Science Consortium (JPDSC) collected the data of approximately 3,000 Japanese individuals and built a genotype database of 2.5 million SNPs (Kamitsuji et al., 2015). However, these studies did not apply entire variants to define a PGx haplotype and construct the comprehensive phenotype frequency of the population. To date, pharmacogenes have not been well characterized in the East Asian population, and systemic research on the PGx allele frequency in Han Chinese populations is still lacking. Therefore, we constructed

14 clinically relevant pharmacogenes from nearly 150,000 Taiwanese individuals. Moreover, we evaluated medication prescription records for each participant to identify any unmet need in PGx testing.

## Materials and methods

### Sample collection

We used participant data collected from the China Medical University Hospital Precision Medicine Project (CMUH, Taiwan) and Taiwan Biobank in this study. The CMUH Biobank contains electronic medical records (EMRs) and blood samples of nearly 175,000 individuals. The genotyping method and imputation procedure were described in detail previously (Liu et al., 2021). Clinical information and baseline characteristics of each individual were determined from the EMRs. Regarding the medication records, we mapped the generic names of 66 gene-related drugs to Anatomical Therapeutic Chemical (ATC) codes. We searched the EMR for the ATC codes of prescriptions used in general practice for the participants and found the prescription records of 39 drugs. For each medication, only the first prescription was considered. The prospective Taiwan Biobank contains the clinical information and genetic data of the general Taiwanese population collected from medical institutions and community health centers in Taiwan, including the WGS and genotyping data (Wei et al., 2021). This study was approved by the China Medical University and Hospital Research Ethics Committee (CMUH110-REC3-225).

### Quality control

PLINK1.9 (Purcell et al., 2007) was used for quality control. We excluded SNVs with a call rate of  $<0.98$ , a minor allele frequency (MAF) of  $<0.01$ , and a Hardy-Weinberg equilibrium (HWE) value of  $<1 \times 10^{-6}$ . At the individual level, SNVs with a call rate of  $<0.98$ , a heterozygosity of  $\pm 5$  standard deviations (SDs) from the mean, or a kinship of  $>0.354$  were excluded. To determine the population structure, we combined the data of 1,000 genome East Asian population (EAS) with our data and conducted principal component analysis by using PLINK2 (Chang et al., 2015). Principal components that did not fall within the three interquartile ranges of 1,000 genome EAS were defined as outliers and excluded from the subsequent analysis.

### Pharmacogene allele construction

After imputation, the variants with an INFO  $R^2$  of  $\geq 0.3$ , a MAF of  $>0.0001$ , and a  $p$  value of HWE  $>1 \times 10^{-7}$  in the variant call format

**TABLE 1** Concordance evaluation of diplotypes and phenotypes between the TWB whole-genome sequencing and TWB imputed array data.

	PharmCAT		PGxPOP	
	Diplotype	Phenotype	Diplotype	Phenotype
<i>CYP2B6</i>	99.30%	99.70%	99.52%	99.52%
<i>CYP2C19</i>	100.00%	100.00%	100.00%	100.00%
<i>CYP2C9</i>	99.80%	99.80%	99.73%	99.73%
<i>CYP3A5</i>	99.80%	99.80%	99.73%	99.73%
<i>CYP4F2</i>	99.90%	100.00%	99.93%	99.93%
<i>NUDT15</i>	99.50%	99.50%	98.49%	99.18%
<i>SLCO1B1</i>	88.60%	99.60%	99.66%	99.86%
<i>TPMT</i>	99.90%	99.90%	99.93%	99.93%
<i>UGT1A1</i>	99.70%	99.80%	99.86%	99.93%
<i>VKORC1</i>	99.90%	99.90%	99.93%	99.93%

TWB, Taiwan Biobank.

(VCF) were included in the analysis. The default settings of PharmCAT v1.0.0 (Sangkuhl et al., 2020) and PGxPOP (McInnes et al., 2021) were used to generate PGx phenotypes from the imputed VCF file. GRCh38 was applied as reference. These tools extract gene variants and infer star allele haplotypes for each individual by using the CPIC allele definition. To enhance the computational efficiency, we extracted the specific chromosomes of the PharmCAT target region before applying PharmCAT. A total of 10 genes (*CYP2B6*, *CYP2C19*, *CYP2C9*, *CYP3A5*, *CYP4F2*, *DPYD*, *NUDT15*, *SLCO1B1*, *TPMT*, and *VKORC1*) were included in the analysis. If multiple PGx phenotypes were inferred by PharmCAT for an individual, we recorded the phenotype as “NA.” To compare the PGx gene allele and phenotype results between the SNP array and WGS data from the Taiwan Biobank, we inferred the star allele of the phased WGS data ( $n = 1,462$ ).

## HLA imputation

We used the HLA genotype imputation with attribute bagging (HIBAG) (Zheng et al., 2014) to impute HLA alleles with four-digit resolution. The default settings of HIBAG were used during the analysis. Four hypersensitivity-related HLA genes—*HLA-A*, *HLA-B*, *HLA-DQA1*, and *HLA-DRB1*, were imputed in this study. An imputation posterior probability of >0.5 was considered reliable.

## Allele frequency data from pharmacogenomics knowledge base

The Pharmacogenomics Knowledge Base (PharmGKB) (Whirl-Carrillo et al., 2021) curates information from latest

pharmacogenomic studies and annotates the gene–drug associations of variants. Moreover, PharmGKB provides a summary of the allele frequencies of pharmacogenes from previous published studies. We downloaded the diplotype and phenotype frequencies in the East Asian population from the PharmGKB website (<https://www.pharmgkb.org/page/pgxGeneRef>).

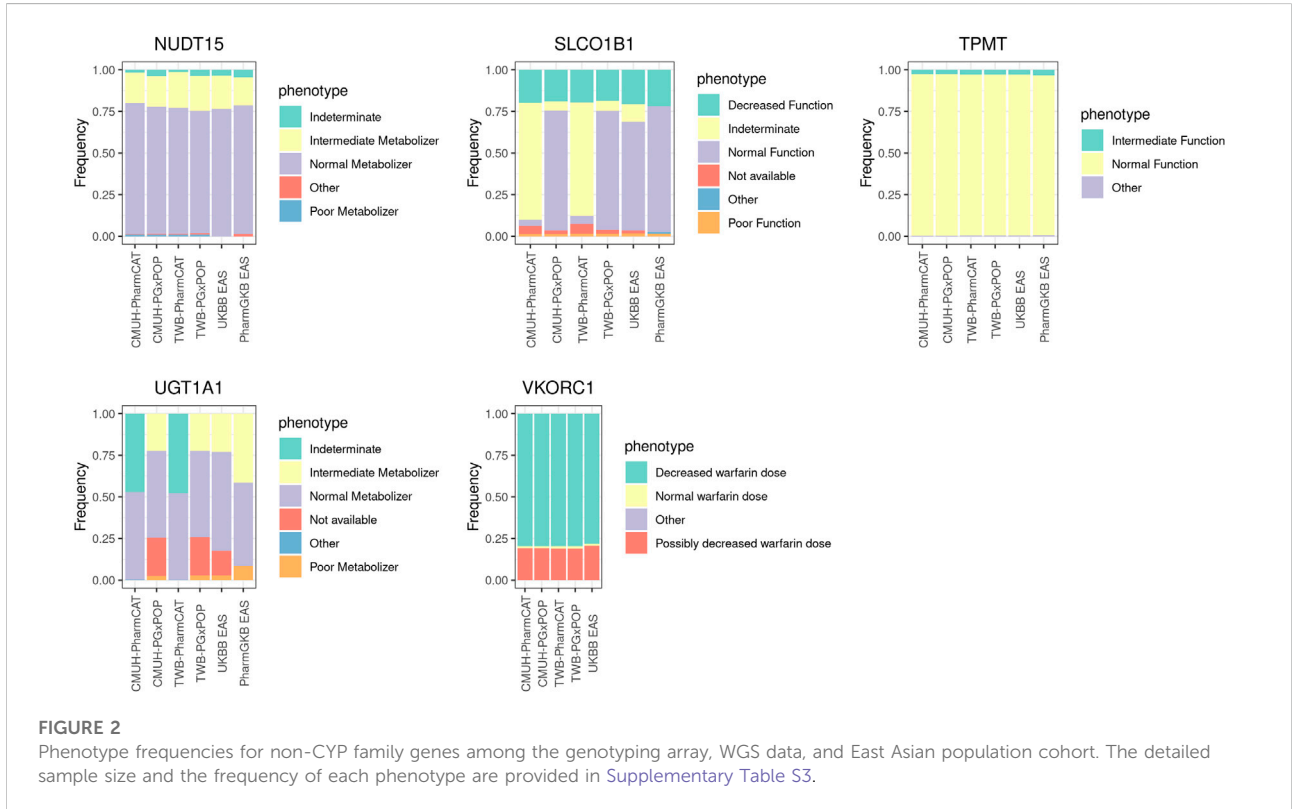
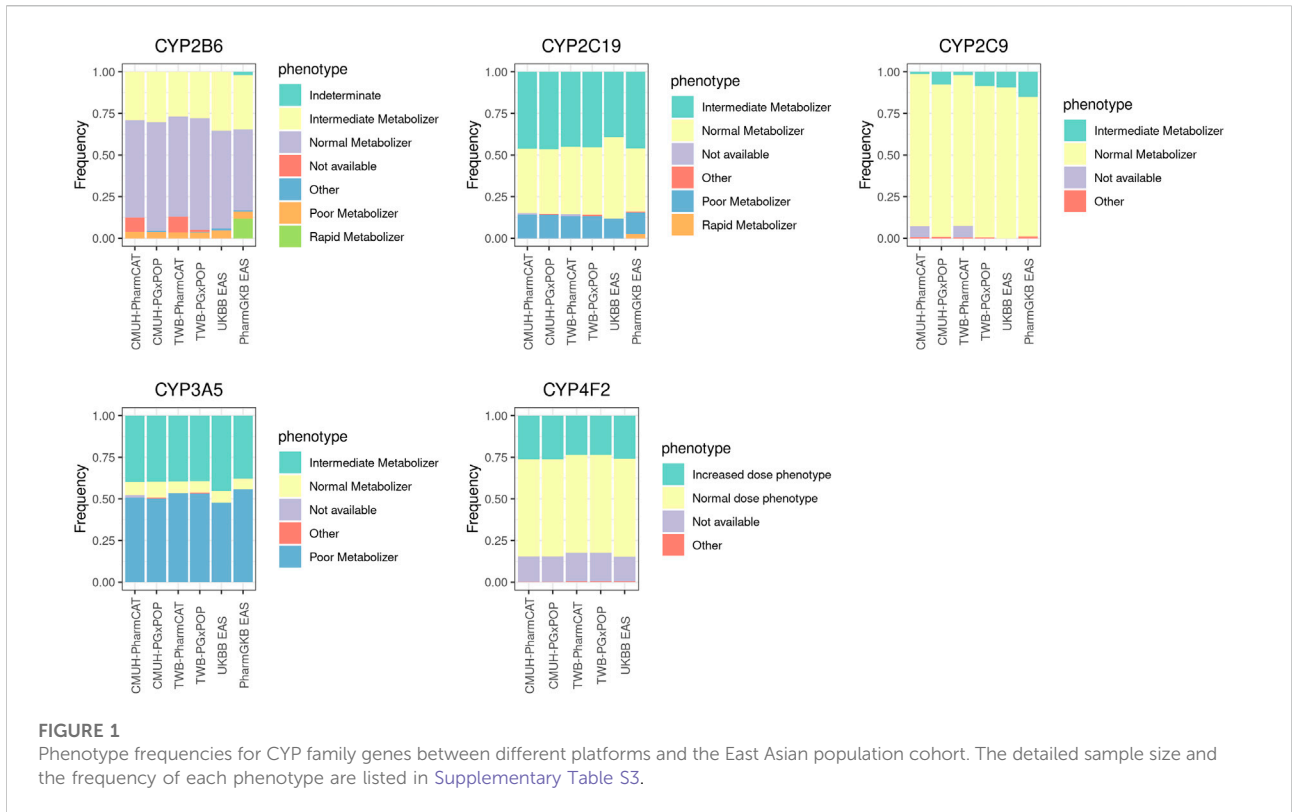
## US Food and Drug Administration guidance

We applied the output from PharmCAT and HIBAG and intersected the data with gene drug lists provided by FDA. A total of 14 genes (*CYP2B6*, *CYP2C19*, *CYP2C9*, *CYP3A5*, *CYP4F2*, *DPYD*, *NUDT15*, *SLCO1B1*, *TPMT*, *VKORC1*, *HLA-A*, *HLA-B*, *HLA-DQA1*, and *HLA-DRB1*) remained. We conducted the PGx study in accordance with the recommendation of the US FDA. We divided PGx associations into high, moderate, and low on the basis of the solidity of evidence from FDA (<https://www.fda.gov/drugs/science-and-research-drugs/table-pharmacogenomic-biomarkers-drug-labeling>).

## Results

### Population and concordance check

The Taiwanese Han population mainly comprises East Asians. After the exclusion of outliers from the EAS population, the genotyping array information of 172,854 individuals was included in this study. The average age of the cohort was  $49.46 \pm 20.96$  years, and 54% were women. The Taiwan Biobank contains WGS and SNP array data. We used two pharmacogene phenotype callers named PharmCAT and PGxPOP to estimate the phenotype of each pharmacogene. Among most of the genes, the phenotype was consistent between the two tools, with a concordance of >90% (Supplementary Table S1). The detailed frequencies of the haplotype and phenotype of each gene are presented in Supplementary Tables S2, S3. Compared with the output from PharmCAT, the phenotype frequencies generated by PGxPOP were similar to the results from UKBB EAS. Furthermore, we conducted a concordance check between the WGS and SNP array data for alleles and phenotypes among 1,462 individuals. The concordance rate of most of the genes was approximately 0.9; only the concordance rate of the *SLCO1B1* diplotype from PharmCAT was 0.89 (Table 1). All phenotypes were highly consistent between the SNP array and WGS results.



## Cytochrome p450 star allele analysis

The phenotypic distributions of *CYP2B6*, *CYP2C19*, *CYP2C9*, *CYP3A5*, and *CYP4F2* are presented in Figure 1, and the frequency of these genes was compared with that in the East Asian population reported by UKBB and PharmGKB. In general, the phenotype distribution of the four CYP family genes (*CYP2B6*, *CYP2C19*, *CYP2C9*, *CYP3A5*, and *CYP4F2*) was similar between our results and those obtained for the East Asian population. Currently, PharmGKB provides only the allele frequency of *CYP4F2*. The *CYP4F2*\*3 allele frequency was 0.17 in our study and 0.22 in the East Asian population cohort. However, a few individuals in our cohort harbored the *CYP4F2*\*2 allele (0.001).

## Noncytochrome p450 star allele analysis

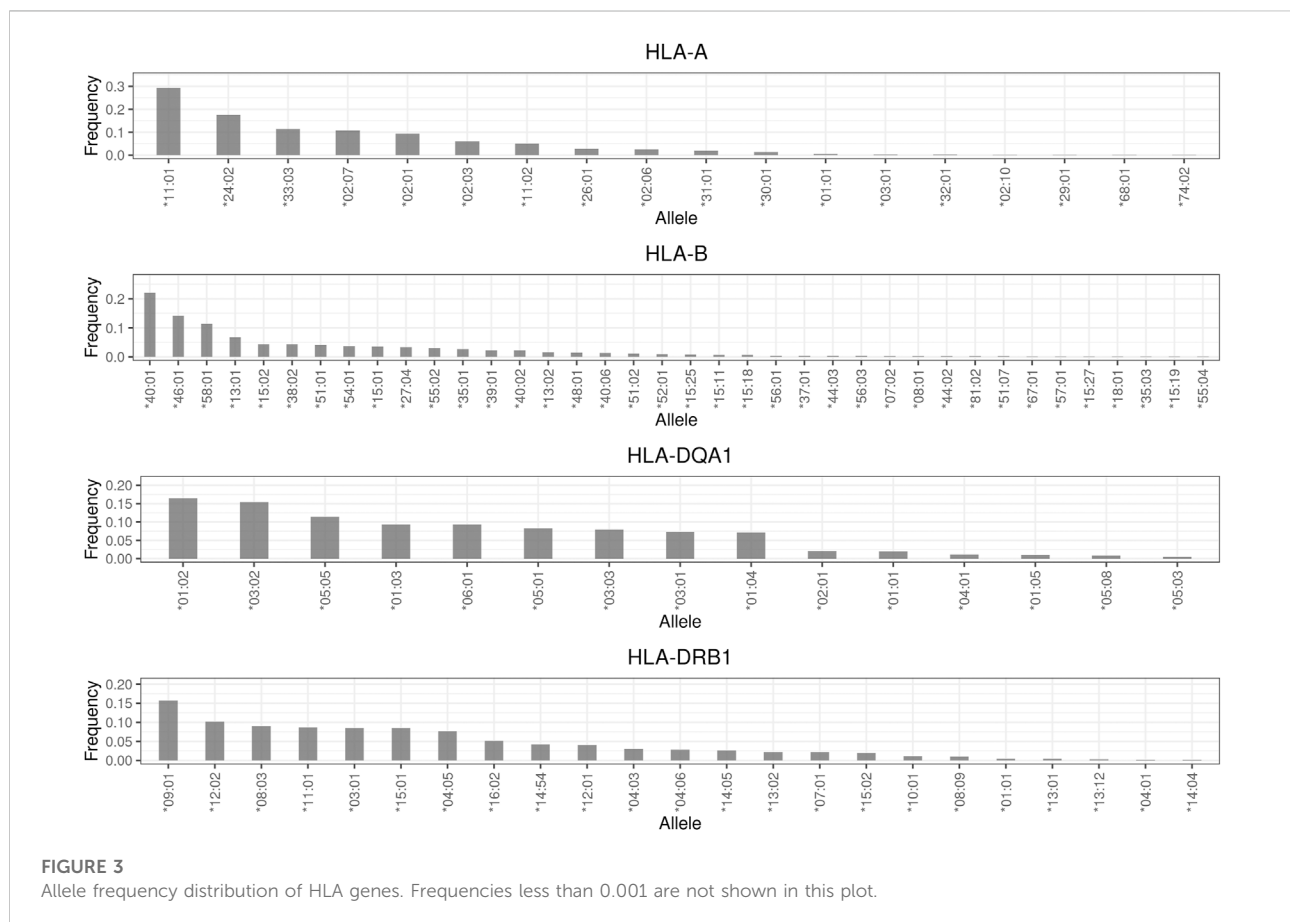
Five PGx genes were included (*UGT1A1*, *TPMT*, *NUDT15*, *SLCO1B1*, and *VKORC1*) in our study (Figure 2). In general, the phenotype frequency distributions of *TPMT*, *NUDT15*, and *VKORC1* in our cohort were similar to those in the East Asian

population. Some differences were noted among different phenotype callers. For example, the most common phenotype of *SLCO1B1* was indeterminate (70%) from PharmCAT, whereas normal function was the dominant phenotype in PGxPOP and UKBB EAS populations. The most common star allele of *SLCO1B1* was \*37 (61%) in the PharmGKB EAS population, while \*1B appeared in approximately 60% of the CMUH array and UKBB EAS samples. Moreover, we evaluated the *UGT1A1* haplotype distribution and observed that the \*60 allele frequency was 23% in our cohort but the allele did not appear in the PharmGKB EAS (Supplementary Table S2).

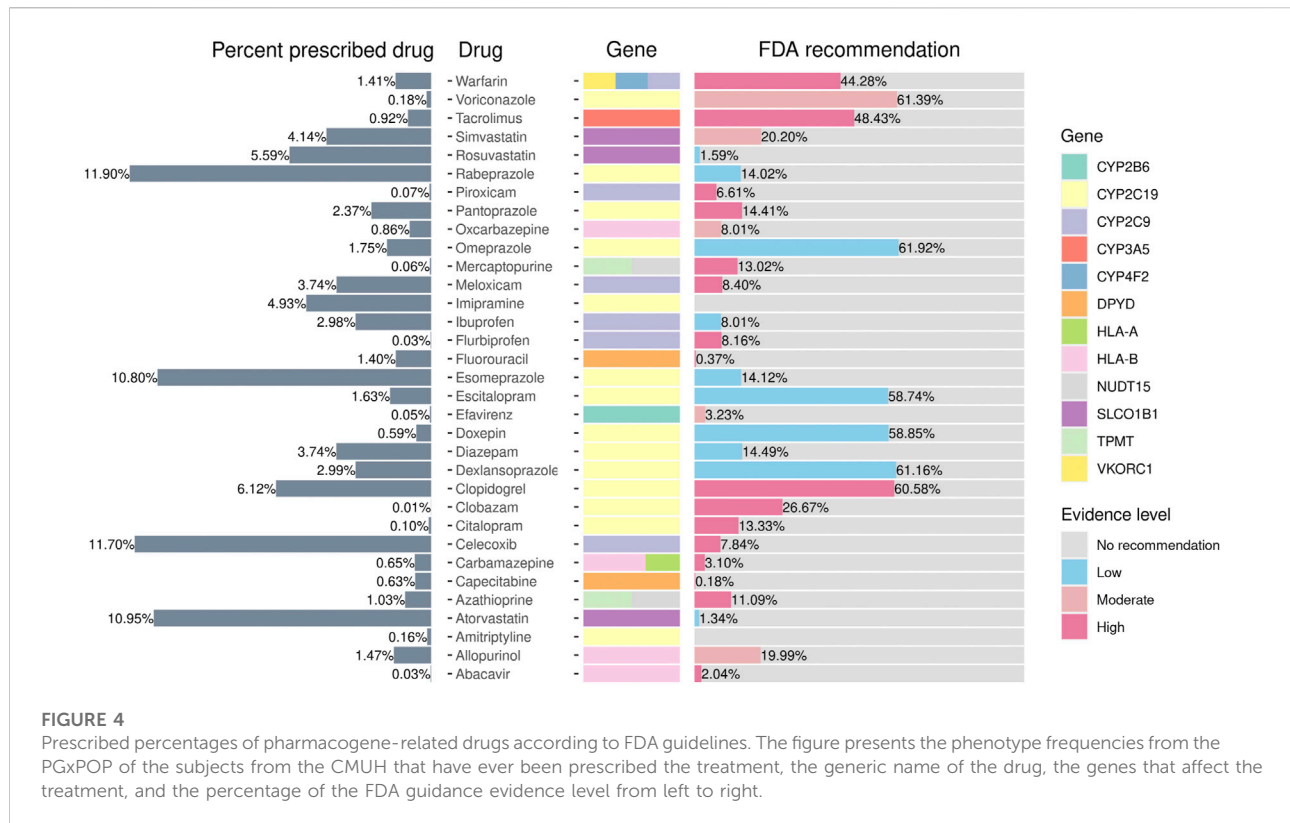
Our results revealed that nearly 80% of the *VKORC1* rs9923231 diplotype is TT, which may increase the sensitivity of an individual to warfarin. The allele frequency of the rs9923231 variant T was 0.89 in the CMUH genotyping array and 0.87 in the East Asian population cohort.

## HLA results

The detailed frequency distributions of the drug responses of four HLA genes, namely *HLA-A*, *HLA-B*, *HLA-DRB1*, and *HLA-DQA1*, are described in Supplementary Tables S3–S6. In brief, the most







common alleles of these four genes were *HLA-A\*11:01* (29.3%), *HLA-B\*40:01* (22%), *HLA-DQA1\*01:02* (16.5%), and *HLA-DRB1\*09:01* (15.7%), respectively (Figure 3). Among the drug hypersensitivity-related alleles, *HLA-B\*58:01* (11.4%) was the most common, followed by *HLA-B\*15:02* (4.4%). The frequency of the other three alleles (*HLA-A\*31:01*, *HLA-DQA1\*02:01*, and *HLA-DRB1\*07:01*) was approximately 2%. The least common allele was *HLA-B\*57:01*, with a frequency of 0.2%. In general, the frequencies of the four *HLA* alleles were consistent with those reported in a previous study (Huang et al., 2020).

## Medication use

We extracted data on the prescription rates of 12 medications relevant to PGx-related treatment from the CMUH EMR. We observed that 99.9% of individuals may atypically respond to at least one medicine. Furthermore, 29% of the individuals harbored the actionable variants for the prescribed drug. As presented in Figure 4, the top three drugs prescribed were the protein pump inhibitor rabeprazole (11.9%), NSAID celecoxib (11.70%), and HMG-CoA reductase inhibitor atorvastatin (10.95%). Notably, 6,414 (61%) clopidogrel recipients belonged to the phenotype that warrants a change in treatment according to the US FDA's recommendation. Among them, 46% were the intermediate metabolizers and 14% were the poor metabolizers of *CYP2C19*. The anticoagulant warfarin is

another drug that merits investigation. Although only approximately 1.41% of the individuals received warfarin in the present cohort, 44% of them may require the adjustment of the treatment dosage. Moreover, the dosage may need to be modified in 48% of tacrolimus users ( $n = 772$ ) to prevent rejection risk.

## Discussion

To date, this is the largest PGx study to report the distribution of diplotypes and phenotypes in the Taiwanese population. This study comprehensively investigated the haplotype frequency of 14 treatment response-related genes. Moreover, through the incorporation of prescription records in the analysis, the potentially affected population of the atypical PGx star allele on drug responses was investigated.

Although the call rates of the alleles were not identical between the genotyping array and WGS data, most of the identified phenotypes were consistent. Because one phenotype may contain different alleles, the response to a medication is mainly driven by the phenotype rather than the allele. Moreover, the phenotypic distribution of most pharmacogenes is consistent with that of the East Asian population. These results indicated that imputed genotyping data may generate robust phenotypes for each

individual, which may help in treatment selection in clinical settings.

The antiplatelet agent clopidogrel is widely used to reduce the risk of cardiovascular events in high-risk patients (Yusuf et al., 2001). *CYP2C19* plays a crucial role in the bioactivation of clopidogrel (Sangkuhl et al., 2010). Exposure to the active clopidogrel metabolite was significantly lower in the *CYP2C19* intermediate metabolizers and poor metabolizers, thereby diminishing the antiplatelet effect (Brandt et al., 2007; Varenhorst et al., 2009; Duarte and Cavallari, 2021). Moreover, real-world evidence indicates that the atherothrombotic risk was lower in the carriers of the *CYP2C19* loss-of-function variant treated with alternative drugs than in those receiving clopidogrel (Beitelshees et al., 2022). Our study demonstrated that 61% of the clopidogrel recipients exhibited the phenotype that may not have had sufficient exposure to the active clopidogrel metabolite. These results indicate that phenotype-guided treatment is crucial in clinical practice and help in prescribing suitable drugs to patients.

The haplotype and phenotype frequencies of *UGT1A1* and *SLCO1B1* were not consistent with those in the East Asian population from PharmGKB, which can be explained by several plausible reasons. First, rare variants and structural variations play a crucial role in star allele haplotyping. However, the SNV microarray may not fully capture these variations. Second, frequency distributions in the East Asian population of the PharmGKB were determined in the basis of previous studies. Because these studies did not test for all known variant alleles (Premawardhena et al., 2003; Teh et al., 2012), once the alleles were negative for sequenced variation, the star allele may default assign to \*1. Moreover, these studies had small sample sizes (Maeda et al., 2014) and focused on specific phenotypes; therefore, the results may not be representative of the general population. Third, a previous study demonstrated that haplotype call rates are affected by the design of pharmacogenomic assays, and assays that examine various variants generate more accurate results compared with those that examine a relatively limited number of variants (Samwald et al., 2015). Therefore, the discrepancies in the phenotype frequency may be caused by the differing variant coverage and design of each platform (Kalman et al., 2016). Additional large-scale population-based studies should be conducted to comprehensively determine the frequency of pharmacogenes through the incorporation of structural variation information.

*CYP2D6* is an important pharmacogene involved in the metabolism of nearly 20% of widely used medications (Taylor et al., 2020). In addition to SNVs, the precise haplotype definition of *CYP2D6* depends on copy number variants and structural alterations (Beoris et al., 2016). A previous study indicated that genotyping arrays cannot capture variations in *CYP2D6* comprehensively due to the highly polymorphic nature of this gene (Twesigomwe et al., 2020). The phenotype of *CYP2D6* called by the genotyping array may be biased consequently (McInnes et al., 2021). Therefore, we did not include the phenotype frequency of *CYP2D6* in this study.

In conclusion, our study comprehensively evaluated the allele and phenotypic distributions of clinically crucial PGx genes in a Taiwanese Han population through systemic imputation and annotation. By analyzing the prescription information from the EMRs, we identified the effect of each drug on a specific population and highlighted the importance of PGx in clinical settings. The study results provide insights into the phenotypic frequencies of clinically relevant pharmacogenes and may be used as reference for designing future PGx studies.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Materials, further inquiries can be directed to the corresponding authors.

## Ethics statement

The studies involving human participants were reviewed and approved by the China Medical University and Hospital Research Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

H-FL, J-GC, and F-JT contributed to the conception and design of the study. T-YL, S-SC, and Y-WH organized the database. H-FL and Y-PC performed the statistical analysis. H-FL conducted data curation and wrote the manuscript.

## Funding

H-FL was supported by China Medical University Hospital, Taiwan (DMR-111-190). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their

affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Beitelshees, A. L., Thomas, C. D., Empey, P. E., Stouffer, G. A., Angiolillo, D. J., Franchi, F., et al. (2022). CYP2C19 genotype-guided antiplatelet therapy after percutaneous coronary intervention in diverse clinical settings. *J. Am. Heart Assoc.* 11, e024159. doi:10.1161/JAHA.121.024159
- Beoris, M., Amos Wilson, J., Garces, J. A., and Lukowiak, A. A. (2016). CYP2D6 copy number distribution in the US population. *Pharmacogenet. Genomics* 26, 96–99. doi:10.1097/FPC.0000000000000188
- Brandt, J. T., Close, S. L., Iturria, S. J., Payne, C. D., Farid, N. A., Ernest, C. S., 2nd, et al. (2007). Common polymorphisms of CYP2C19 and CYP2C9 affect the pharmacokinetic and pharmacodynamic response to clopidogrel but not prasugrel. *J. Thromb. Haemost.* 5, 2429–2436. doi:10.1111/j.1538-7836.2007.02775.x
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., Lee, J. J., et al. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 4, 7. doi:10.1186/s13742-015-0047-8
- Duarte, J. D., and Cavallari, L. H. (2021). Pharmacogenetics to guide cardiovascular drug therapy. *Nat. Rev. Cardiol.* 18, 649–665. doi:10.1038/s41569-021-00549-w
- Huang, Y. H., Khor, S. S., Zheng, X., Chen, H. Y., Chang, Y. H., Chu, H. W., et al. (2020). A high-resolution HLA imputation system for the Taiwanese population: a study of the taiwan biobank. *Pharmacogenomics* J. 20, 695–704. doi:10.1038/s41397-020-0156-3
- Kalman, L. V., Agúndez, J., Appell, M. L., Black, J. L., Bell, G. C., Boukouvala, S., et al. (2016). Pharmacogenetic allele nomenclature: international workgroup recommendations for test result reporting. *Clin. Pharmacol. Ther.* 99, 172–185. doi:10.1002/cpt.280
- Kamitsuji, S., Matsuda, T., Nishimura, K., Endo, S., Wada, C., Watanabe, K., et al. (2015). Japan PGx data science consortium database: SNPs and HLA genotype data from 2994 japanese healthy individuals for pharmacogenomics studies. *J. Hum. Genet.* 60, 319–326. doi:10.1038/jhg.2015.23
- Kim, J. A., Ceccarelli, R., and Lu, C. Y. (2021). Pharmacogenomic biomarkers in US FDA-approved Drug labels (2000–2020). *J. Pers. Med.* 11, 179. doi:10.3390/jpm11030179
- Lavertu, A., McInnes, G., Daneshjou, R., Whirl-Carrillo, M., Klein, T. E., Altman, R. B., et al. (2018). Pharmacogenomics and big genomic data: from lab to clinic and back again. *Hum. Mol. Genet.* 27, R72–R78. doi:10.1093/hmg/ddy116
- Lee, S.-B., Wheeler, M. M., Patterson, K., Mcgee, S., Dalton, R., Woodahl, E. L., et al. (2019). Stargazer: a software tool for calling star alleles from next-generation sequencing data using CYP2D6 as a model. *Genet. Med.* 21, 361–372. doi:10.1038/s41436-018-0054-0
- Liu, T. Y., Lin, C. F., Wu, H. T., Wu, Y. L., Chen, Y. C., Liao, C. C., et al. (2021). Comparison of multiple imputation algorithms and verification using whole-genome sequencing in the CMUH genetic biobank. *Biomed. (Taipei)* 11, 57–65. doi:10.37796/2211-8039.1302
- Maeda, H., Hazama, S., Shavkat, A., Okamoto, K., Oba, K., Sakamoto, J., et al. (2014). Differences in UGT1A1, UGT1A7, and UGT1A9 polymorphisms between Uzbek and Japanese populations. *Mol. Diagn. Ther.* 18, 333–342. doi:10.1007/s40291-014-0083-6
- Mauleekoonphairoj, J., Chamnanphon, M., Khongphatthanayothin, A., Sutjaporn, B., Wandee, P., Poovorawan, Y., et al. (2020). Phenotype prediction and characterization of 25 pharmacogenes in Thais from whole genome sequencing for clinical implementation. *Sci. Rep.* 10, 18969. doi:10.1038/s41598-020-76085-3
- McInnes, G., Lavertu, A., Sangkuhl, K., Klein, T. E., Whirl-Carrillo, M., Altman, R. B., et al. (2021). Pharmacogenetics at scale: An analysis of the UK biobank. *Clin. Pharmacol. Ther.* 109, 1528–1537. doi:10.1002/cpt.2122
- Numanagić, I., Malikić, S., Ford, M., Qin, X., Toji, L., Radovich, M., et al. (2018). Allelic decomposition and exact genotyping of highly polymorphic and structurally variant genes. *Nat. Commun.* 9, 828. doi:10.1038/s41467-018-03273-1
- Premawardhana, A., Fisher, C. A., Liu, Y. T., Verma, I. C., De Silva, S., Arambepola, M., et al. (2003). The global distribution of length polymorphisms of the promoters of the glucuronosyltransferase 1 gene (UGT1A1): hematologic and evolutionary implications. *Blood Cells Mol. Dis.* 31, 98–101. doi:10.1016/s1079-9796(03)00071-8
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. a. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi:10.1086/519795
- Relling, M. V., and Klein, T. E. (2011). CPIC: clinical pharmacogenetics implementation consortium of the pharmacogenomics research network. *Clin. Pharmacol. Ther.* 89, 464–467. doi:10.1038/clpt.2010.279
- Robarge, J. D., Li, L., Desta, Z., Nguyen, A., and Flockhart, D. A. (2007). The star-allele nomenclature: retooling for translational genomics. *Clin. Pharmacol. Ther.* 82, 244–248. doi:10.1038/sj.clpt.6100284
- Samwald, M., Blagec, K., Hofer, S., and Freimuth, R. R. (2015). Analyzing the potential for incorrect haplotype calls with different pharmacogenomic assays in different populations: a simulation based on 1000 genomes data. *Pharmacogenomics* 16, 1713–1721. doi:10.2217/pgs.15.108
- Sangkuhl, K., Klein, T. E., and Altman, R. B. (2010). Clopidogrel pathway. *Pharmacogenet. Genomics* 20, 463–465. doi:10.1097/FPC.0b013e3283385420
- Sangkuhl, K., Whirl-Carrillo, M., Whaley, R. M., Woon, M., Lavertu, A., Altman, R. B., et al. (2020). Pharmacogenomics clinical annotation tool (PharmCAT). *Clin. Pharmacol. Ther.* 107, 203–210. doi:10.1002/cpt.1568
- Tayeh, M. K., Gaedigk, A., Goetz, M. P., Klein, T. E., Lyon, E., Mcmillin, G. A., et al. (2022). Clinical pharmacogenomic testing and reporting: a technical standard of the American college of medical genetics and genomics (ACMG). *Genet. Med.* 24, 759–768. doi:10.1016/j.gim.2021.12.009
- Taylor, C., Crosby, I., Yip, V., Maguire, P., Pirmohamed, M., Turner, R. M., et al. (2020). A review of the important role of CYP2D6 in pharmacogenomics. *Genes (Basel)* 11, E1295. doi:10.3390/genes11111295
- Teh, L. K., Hashim, H., Zakaria, Z. A., and Salleh, M. Z. (2012). Polymorphisms of UGT1A1\*6, UGT1A1\*27 & UGT1A1\*28 in three major ethnic groups from Malaysia. *Indian J. Med. Res.* 136, 249–259.
- Twesigomwe, D., Wright, G. E. B., Drögemöller, B. I., Da Rocha, J., Lombard, Z., Hazelhurst, S., et al. (2020). A systematic comparison of pharmacogene star allele calling bioinformatics algorithms: a focus on CYP2D6 genotyping. *NPJ Genom. Med.* 5, 30. doi:10.1038/s41525-020-0135-2
- Twist, G. P., Gaedigk, A., Miller, N. A., Farrow, E. G., Willig, L. K., Dinwiddie, D. L., et al. (2016). Constellation: a tool for rapid, automated phenotype assignment of a highly polymorphic pharmacogene, CYP2D6, from whole-genome sequences. *NPJ Genom. Med.* 1, 15007. doi:10.1038/npjgenmed.2015.7
- Varenhorst, C., James, S., Erlinge, D., Brandt, J. T., Braun, O. O., Man, M., et al. (2009). Genetic variation of CYP2C19 affects both pharmacokinetic and pharmacodynamic responses to clopidogrel but not prasugrel in aspirin-treated patients with coronary artery disease. *Eur. Heart J.* 30, 1744–1752. doi:10.1093/eurheartj/ehp157
- Wang, Y., Xiao, F., Chen, Y., Xiao, L. D., Wang, L. Y., Zhan, Y., et al. (2021). Analytics of the clinical implementation of pharmacogenomics testing in 12 758 individuals. *Clin. Transl. Med.* 11, e586. doi:10.1002/ctm.2586
- Wei, C.-Y., Yang, J.-H., Yeh, E.-C., Tsai, M.-F., Kao, H.-J., Lo, C.-Z., et al. (2021). Genetic profiles of 103, 106 individuals in the Taiwan Biobank provide insights into the health and history of Han Chinese. *NPJ Genom. Med.* 6, 10. doi:10.1038/s41525-021-00178-9
- Whirl-Carrillo, M., Huddart, R., Gong, L., Sangkuhl, K., Thorn, C. F., Whaley, R., et al. (2021). An evidence-based framework for evaluating pharmacogenomics knowledge for personalized medicine. *Clin. Pharmacol. Ther.* 110, 563–572. doi:10.1002/cpt.2350
- Yusuf, S., Zhao, F., Mehta, S. R., Chrolavicius, S., Tognoni, G., and Fox, K. K. (2001). Effects of clopidogrel in addition to aspirin in patients with acute coronary syndromes without ST-segment elevation. *N. Engl. J. Med.* 345, 494–502. doi:10.1056/NEJMoa010746
- Zheng, X., Shen, J., Cox, C., Wakefield, J. C., Ehm, M. G., Nelson, M. R., et al. (2014). HIBAG-HLA genotype imputation with attribute bagging. *Pharmacogenomics* J. 14, 192–200. doi:10.1038/tj.2013.18

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.948616/full#supplementary-material>