

## RESEARCH ARTICLE

## Actionable pharmacogenetic variants in Hong Kong Chinese exome sequencing data and projected prescription impact in the Hong Kong population

Mullin Ho Chung Yu<sup>1</sup> <sup>✉</sup>, Marcus Chun Yin Chan<sup>1</sup> <sup>✉</sup>, Claudia Ching Yan Chung<sup>1</sup>, Andrew Wang Tat Li<sup>2</sup>, Chara Yin Wa Yip<sup>2</sup>, Christopher Chun Yu Mak<sup>1</sup> <sup>✉</sup>, Jeffrey Fong Ting Chau<sup>1</sup> <sup>✉</sup>, Mianne Lee<sup>1</sup> <sup>✉</sup>, Jasmine Lee Fong Fung<sup>1</sup> <sup>✉</sup>, Mandy Ho Yin Tsang<sup>1</sup>, Joshua Chun Ki Chan<sup>1</sup> <sup>✉</sup>, Wilfred Hing Sang Wong<sup>1</sup> <sup>✉</sup>, Jing Yang<sup>1</sup> <sup>✉</sup>, William Chun Ming Chui<sup>2</sup>, Patrick Ho Yu Chung<sup>3</sup>, Wanling Yang<sup>1</sup> <sup>✉</sup>, So Lun Lee<sup>4,5</sup> <sup>✉</sup>, Godfrey Chi Fung Chan<sup>1,5,6</sup>, Paul Kwong Hang Tam<sup>3,7</sup> <sup>✉</sup>, Yu Lung Lau<sup>1,5,6</sup> <sup>✉</sup>, Clara Sze Man Tang<sup>3,7\*</sup> <sup>✉</sup>, Kit San Yeung<sup>1\*</sup> <sup>✉</sup>, Brian Hon Yin Chung<sup>1,4,5,6\*</sup> <sup>✉</sup>

**1** Department of Paediatrics and Adolescent Medicine, LKS Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong SAR, China, **2** Department of Pharmacy, Queen Mary Hospital, Pokfulam, Hong Kong SAR, China, **3** Department of Surgery, LKS Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong SAR, China, **4** Department of Paediatrics and Adolescent Medicine, Duchess of Kent Children's Hospital, Pokfulam, Hong Kong SAR, China, **5** Department of Paediatrics and Adolescent Medicine, Queen Mary Hospital, Pokfulam, Hong Kong SAR, China, **6** Department of Paediatrics and Adolescent Medicine, The Hong Kong Children's Hospital, Kowloon Bay, Hong Kong SAR, China, **7** Dr Li Dak-Sum Research Centre, The University of Hong Kong–Karolinska Institutet Collaboration in Regenerative Medicine, Pokfulam, Hong Kong SAR, China

<sup>✉</sup> These authors contributed equally to this work.

\* [claratang@hku.hk](mailto:claratang@hku.hk) (CSMT); [jsks@connect.hku.hk](mailto:jsks@connect.hku.hk) (KSY); [bhychung@hku.hk](mailto:bhychung@hku.hk) (BHYC)

## Abstract

Preemptive pharmacogenetic testing has the potential to improve drug dosing by providing point-of-care patient genotype information. Nonetheless, its implementation in the Chinese population is limited by the lack of population-wide data. In this study, secondary analysis of exome sequencing data was conducted to study pharmacogenomics in 1116 Hong Kong Chinese. We aimed to identify the spectrum of actionable pharmacogenetic variants and rare, predicted deleterious variants that are potentially actionable in Hong Kong Chinese, and to estimate the proportion of dispensed drugs that may potentially benefit from genotype-guided prescription. The projected preemptive pharmacogenetic testing prescription impact was evaluated based on the patient prescription data of the public healthcare system in 2019, serving 7.5 million people. Twenty-nine actionable pharmacogenetic variants/alleles were identified in our cohort. Nearly all (99.6%) subjects carried at least one actionable pharmacogenetic variant, whereas 93.5% of subjects harbored at least one rare deleterious pharmacogenetic variant. Based on the prescription data in 2019, 13.4% of the Hong Kong population was prescribed with drugs with pharmacogenetic clinical practice guideline recommendations. The total expenditure on actionable drugs was 33,520,000 USD, and it was estimated that 8,219,000 USD (24.5%) worth of drugs were prescribed to patients with an implicated actionable phenotype. Secondary use of exome sequencing data for



## OPEN ACCESS

**Citation:** Yu MHC, Chan MCY, Chung CCY, Li AWT, Yip CYW, Mak CCY, et al. (2021) Actionable pharmacogenetic variants in Hong Kong Chinese exome sequencing data and projected prescription impact in the Hong Kong population. *PLoS Genet* 17(2): e1009323. <https://doi.org/10.1371/journal.pgen.1009323>

**Editor:** Nita Limdi, University of Alabama at Birmingham, UNITED STATES

**Received:** August 12, 2020

**Accepted:** December 30, 2020

**Published:** February 18, 2021

**Copyright:** © 2021 Yu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Genotype data is available in the [Supporting Information](#) files in statistical summary. Additional data cannot be shared publicly due to ethical concerns regarding patient data. Full data are available from the the HKU/HA HK West Institutional Review Board (UW12-211, UW12-383, UW 05-282 T/945, UW 12-382, UW 12-469)(contact via Dr Brian Chung) for researchers who meet the criteria for access to confidential data.

**Funding:** This study was supported by the Society for the Relief of Disabled Children (BHYC), the Health and Medical Research Fund (BHYC), Li Ka Shing Donation Account: Enhanced New Staff Start-up Packages (PHYC), the Children's Heart Foundation (BHYC), and the Edward and Yolanda Wong Fund (BHYC). 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.

pharmacogenetic analysis is feasible, and preemptive pharmacogenetic testing has the potential to support prescription decisions in the Hong Kong Chinese population.

## Author summary

Pharmacogenetic testing provides relevant drug phenotype information to guide personalized drug prescription, which potentially improves drug efficacy and prevent adverse drug reactions. However, its implementation in the Chinese population is limited by the lack of Chinese-specific pharmacogenetics data. In this study, we studied the spectrum of 133 actionable pharmacogenetic variants and rare deleterious variants in 108 pharmacogenes using an exome sequencing consisting of 1116 Hong Kong Chinese subjects. It was found that nearly all individuals carried at least one actionable pharmacogenetic variant and one rare, predicted deleterious pharmacogenetic variant. In addition, we projected the potential prescription impact of actionable pharmacogenetic variants using prescription data of the Hong Kong's public healthcare system. We estimated that one-seventh of the Hong Kong population received at least one of the 36 drugs with clinical pharmacogenetics guideline recommendations. The findings demonstrated the potential of pharmacogenetic testing in improving patient care and resource allocation in Chinese. The cohort dataset also supports clinical implementation of pharmacogenetics in the Chinese population.

## Introduction

Pharmacogenetics is the study of variability in drug response caused by genetic variations [1]. It is estimated that over two million hospitalized patients develop severe adverse drug reactions in the United States annually, incurring a direct medical cost of 200 billion US dollars (USD) [2]. This indicates a huge potential in reducing healthcare costs even if only a small proportion of adverse drug reactions are preventable by genotype-guided prescription. The current clinical applications of pharmacogenetics are mostly limited to reactive pharmacogenetic testing, in which investigations are ordered only when certain high-risk medications are prescribed, or after an adverse drug reaction has occurred. In contrast, preemptive pharmacogenetic testing allows the optimization of dosing based on genotype information at the time of prescription, minimizing the risk of undesired outcomes [3].

The pharmacogenetics community has established open-access platforms to facilitate the implementation of clinical pharmacogenetics. The Clinical Pharmacogenetics Implementation Consortium (CPIC) provides peer-reviewed, evidence-based guidelines for individual gene-drug combinations to assist clinicians to incorporate pharmacogenetics into daily practice [4]. The Pharmacogenomics Knowledgebase (PharmGKB) provides a multitude of pharmacogenetic information and assigns a level of evidence on variant-medication pairs based on the quantity and quality of available evidence [5,6]. PharmGKB level 1A/B represents the highest level of evidence, which is either supported by large-scale studies or is widely accepted by the pharmacogenetics community. Multiple large population pharmacogenetics studies revealed that nearly all subjects (96.2–97.0%) had at least one actionable pharmacogenetic variant, with a median of two to three [7,8]. PharmGKB level 2A/B annotations are assigned to variant-drug combinations with moderate evidence of association, whereas levels 3 and 4 annotations are assigned to variants with low and preliminary levels of evidence, respectively [6]. The study of pharmacogenetic variants is often limited by sample size. Many actionable pharmacogenetic variants are known because they are relatively common in the population, or their associated

adverse reactions are severe. In contrast, the effects of rare variants are largely unknown, but they should not be neglected since they are consistently found in the population and have been predicted to account for nearly all inter-individual variability in more than half of the known pharmacogenes [7,9–12]. Therefore, the study of rare pharmacogenetic variants is important as it can potentially improve the prediction of drug responses.

It is known that pharmacogenetic variations exist across different ethnicities. Currently, Chinese pharmacogenetic data is limited. For example, Asians account for 2% of the eMERGE-PGRNseq cohort [7]. The first large-scale analysis of actionable pharmacogenetic variants in Chinese was published only recently, yet the study did not examine rare variants and its prescription pattern analysis was performed based on data from a children hospital [13]. Besides, the Southern Chinese sub-population accounted for 22.4% of the study subjects and was underrepresented [13]. To address these issues, we examined the spectrum of 133 actionable pharmacogenetic variants and rare deleterious variants in 108 pharmacogenes using an exome sequencing cohort consisting of 1116 Hong Kong (HK) Chinese subjects, who are representative of the Southern Chinese subpopulation. In addition, the potential prescription impact of actionable pharmacogenetic variants was projected on the HK population.

## Materials and methods

### Ethics statement

Written informed consent was obtained for each participant, and this study was approved by the HKU/HA HK West Institutional Review Board (UW12-211, UW12-383, UW 05-282 T/945, UW 12-382, UW 12-469).

### Subjects and exome sequencing (S1 Fig)

A total of 1,141 unrelated, self-reported Chinese were enrolled for exome sequencing for rare disease diagnosis or complex disease research from 2012 to 2019. Exome sequencing was performed on genomic DNA derived from peripheral blood or buccal mucosa by Illumina sequencing platforms, and different exome capture kits were used (S1 Table). The processing of raw exome sequencing data is described in detail in the Supplementary Methods (S1 Text). Briefly, variant calling was performed using a pipeline based on the Genome Analysis Toolkit (GATK), and human leukocyte antigen (HLA) typing was performed using HLA typing from High-quality Dictionary (HLA-HD) [14,15]. The exome sequencing dataset was subjected to stringent quality control (QC) procedures at the sample, variant, and genotype levels and the output data were annotated using wANNOVAR [16]. To avoid over-representation of disease-associated variants, the samples collected from subjects with respiratory diseases and neuromuscular disorders were removed for *CFTR* and *RYR1* analysis, respectively. In this study, a rare variant was defined as a variant having a Genome Aggregation Database (gnomAD) global allele frequency (AF) <1%. A missense variant was considered deleterious when it possessed a Phred-scaled Combined Annotation Dependent Depletion (CADD) score  $\geq 20$  [17], or Rare Exome Variant Ensemble Learner (REVEL) score  $\geq 0.7$ , or PREDICT score  $\geq 0.6$ ; whereas a loss-of-function (LoF) variant was considered deleterious when it possessed a Phred-scaled CADD score  $\geq 20$  or a Loss-Of-Function Transcript Effect Estimator (LOFTEE) of “high-confidence” [18–20].

### Selection of actionable pharmacogenetic variants and high-confidence pharmacogenes

To examine the spectrum of known actionable pharmacogenetic variants, those with clinical annotations of PharmGKB level 1A or 1B were selected since they represent variants with the

highest level of evidence for clinical actionability of gene-drug associations [5,6,21]. In addition, the literature was reviewed to identify Chinese-specific pharmacogenetic variants that were actionable based on the CPIC guideline [22,23]. This review resulted in a list of 133 actionable pharmacogenetic variants and HLA alleles in 19 genes (S2 Table). To analyze rare pharmacogenetic variants that are potentially actionable, genes with at least one PharmGKB level 1 or 2 clinical annotation were selected, resulting in a list of 108 genes that were considered “high confidence pharmacogenes” (S3 Table) [5,6].

## Projected prescription impact of actionable pharmacogenetic variants in Hong Kong

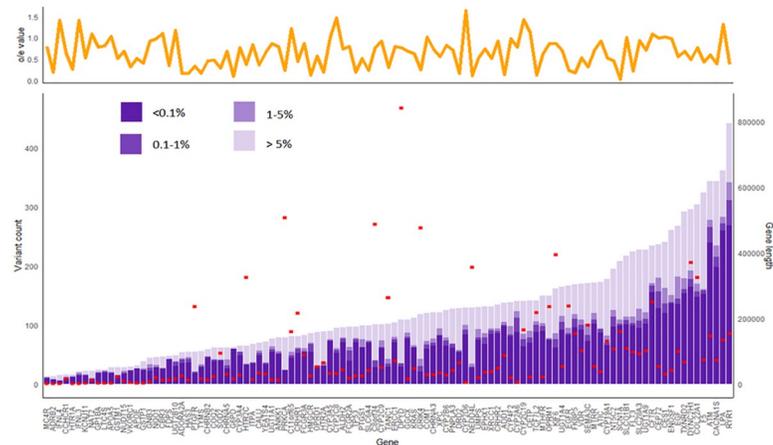
HK is a city in Southern China with a population of 7.5 million, of which 92% are ethnic Han Chinese [24]. In HK, over 90% of all secondary and tertiary health services are provided by the public healthcare system under the management of the Hospital Authority (HA) [25,26]. All public healthcare service data, including drug prescriptions in the inpatient and outpatient settings, are available in the Clinical Data Analysis and Reporting System (CDARS) database in an unlisted and anonymous manner.

Based on CPIC gene-drug annotations, 44 drugs were matched with known actionable pharmacogenetic variants observed in the exome dataset. Under the public healthcare system in 2019, 36 of these drugs were dispensed by pharmacies and had available data, and eight drugs were not prescribed. All patients in HK who had a prescription record of one or more of the 36 drugs at one or more of the hospitals under the HA between January 1, 2019 and December 31, 2019 were identified from the CDARS database. The number of subjects, drug quantity, and unit cost of each drug were retrieved for the estimation of prescription impact and drug expenditure. Duplications from multiple prescriptions within the same year, different routes of administration and doses for each drug were removed to identify the unique number of subjects. Pharmacophenotypes were assigned based on the genotype definition stated in the CPIC guideline. Phenotypes with at least one prescription recommendation based on the CPIC guideline were considered as actionable. The frequency of each actionable phenotype was derived from the allele frequency of the Hong Kong Chinese exome dataset based on the Hardy-Weinberg equation, except for *CYP2D6*. The frequency of actionable phenotype of *CYP2D6* was retrieved from a published Hong Kong study, since exome sequencing cannot detect haplotype and copy number variations accurately [27]. The projected prescription impact for each drug was estimated by multiplying the frequency of actionable phenotype by the total number of subjects for each individual drug retrieved from CDARS. Drug expenditures were converted from HK dollars (HKD) to USD based on the conversion rate of 7.8 HKD = 1.0 USD.

## Results

### Exome sequencing data characteristics

A total of 1,116 samples, including 622 males and 494 females, passed the sample-level QC procedures. Among the 108 high confidence pharmacogenes, 104 genes had at least 8X mean coverage in >75% of the samples (S2 Fig). The exceptions were *CCHCR1*, *TNF*, *IFNL4*, and *GSTM1*. A total of 13,165 variants were identified in the 108 pharmacogenes, among which 11,415 were non-coding, 1,719 were exonic, and 31 were canonical splice site variants (S3 Fig). Of all variants identified, 3,501 (26.6%) have never been reported in public databases including gnomAD, dbSNP, and ClinVar (S4 Table). A significant linear relationship between gene transcript length and total number of variants in each gene ( $p = 0.0073$ ) was observed, with an increase of 0.17 variants per kilobase of gene length (Fig 1).



**Fig 1. Allele frequency of variants within the 108 pharmacogenes in the dataset.** In the upper panel, the yellow line graph shows the gnomAD loss-of-function constraint metric (*o/e* score) of the respective genes. In the lower panel, the purple bars denote the variant counts in the 108 high-confidence pharmacogenes, while the red rectangles indicate respective gene transcript lengths. Consistent across genes, most variants belong to the very rare category (AF <0.1%). The relationship between variant count, gene transcript length, and constraint (*o/e* score reported in gnomAD) was analyzed using multiple linear regression analysis. There was significant association between gene transcript length and total variant count ( $P = 0.0073$ ). In general, the number of variants increased by 0.17 for every kilobase increase in gene length, although outliers existed. In the highly polymorphic gene *CYP2D6*, 29.5 variants were observed for every kilobase of gene length.

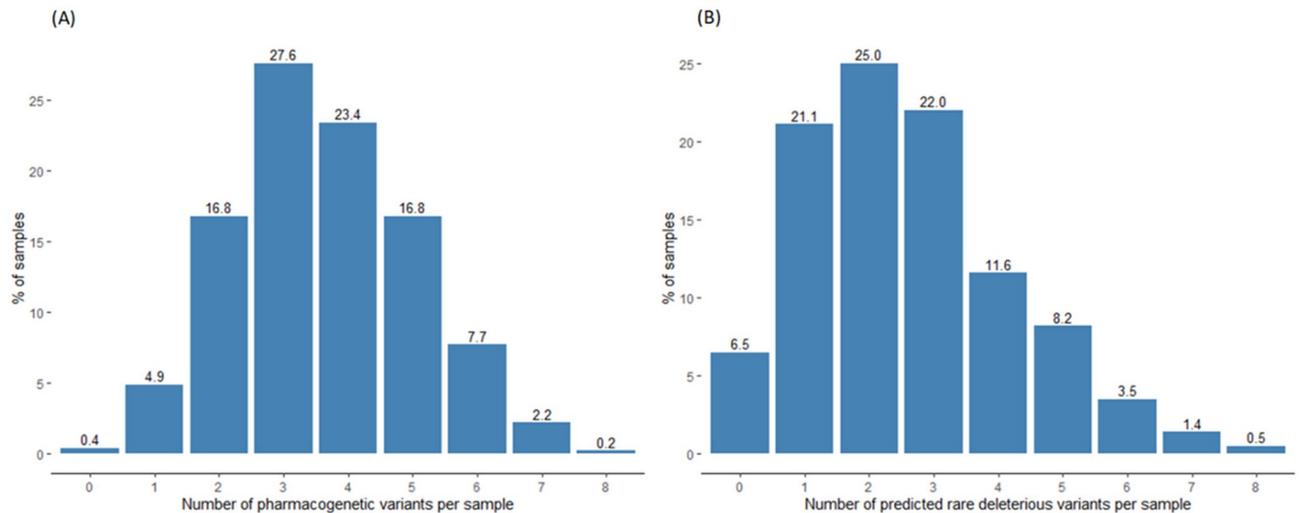
<https://doi.org/10.1371/journal.pgen.1009323.g001>

### Spectrum of known actionable pharmacogenetic variants

The majority of the 129 known actionable pharmacogenetic variants and four HLA alleles were well-covered in the exome sequencing data, except for four variants which could not be detected by exome sequencing because they are located in non-coding regions (S2 Table). For more than 90% of the samples, depths of >8X and >30X were achieved in 121 (93.8%) and 62 (48.1%) variants, respectively (S4 Fig). In our cohort, 25 known actionable variants and all four HLA alleles were identified, accounting for 15 genes and 44 implicated drugs (S5 Table). 104 actionable variants are absent in the HK Chinese population (AF = 0). The most prevalent variant in our cohort was rs1065852 in *CYP2D6* (AF = 60.95%), a marker single nucleotide polymorphism (SNP) of a markedly reduced or null allele, while the most prevalent HLA risk allele was *HLA-B\*15:02* (AF = 9.68%; S6 Table). Analyzing using a per-sample approach, 1,111 (99.6%) individuals harbored at least one actionable variant, with a median of four (Fig 2A). At the gene level, *CYP2C19* (57.21%), *CYP3A5* (43.38%), and *CYP2B6* (40.51%) were the genes with the highest frequency of actionable phenotypes (Table 1). In terms of individual drugs, the antiplatelet drug clopidogrel (57.21%), immunosuppressant tacrolimus (43.38%), and anti-coagulant warfarin (43.13%) had the highest frequency of actionable phenotypes (S7 Table).

### Projected prescription impact of actionable pharmacophenotypes in HK

Based on the available prescription data in CDARS in 2019, a total of 1,006,046 HK Chinese patients had received at least one of the 36 drugs with CPIC guideline recommendations, representing 13.4% of the HK population. In general, the percentage of people prescribed with drugs with CPIC guideline recommendations increased with ages (S8 Table). For elderly patients (aged >60), 616,553 received one of the 36 drugs, accounting for 31.9% of the total elderly population. Preemptive pharmacogenetic testing is projected to have the greatest impact for simvastatin (146,167 patients, frequency: 25.81%), clopidogrel (26,304 patients, frequency: 57.21%), and ibuprofen (12,000 patients, frequency: 5.39%; Fig 3). In terms of drug



**Fig 2. Number of pharmacogenetic variants identified in exome sequencing per sample.** (A) Nearly all (99.6%) Hong Kong Chinese carried at least one known actionable pharmacogenetic variant, with a median of four variants. The highest number of actionable variants in a single sample was eight, occurring in 0.2% of samples. (B) Among 1,116 Hong Kong Chinese, 1,043 (93.5%) carried at least one rare deleterious variant in the 108 high-confidence pharmacogenes, with a median of two variants per individual. The highest number of variants observed was eight, occurring in 0.5% of the population.

<https://doi.org/10.1371/journal.pgen.1009323.g002>

costs, the 36 drugs had a total expenditure of 33,520,000 USD, which accounted for 3.6% of the annual drug expenditure of HA in the 2018–2019 fiscal year. It was estimated that 8,219,000 USD (i.e., 24.5% of drug expenditures among the 36 implicated drugs) worth of drugs were prescribed to patients with an implicated actionable phenotype. Remarkably, tacrolimus (4,301,000 USD), escitalopram (777,000 USD), and simvastatin (710,000 USD) accounted for 70.4% of the total expenditure of drugs that were prescribed to patients with an implicated actionable phenotype (Fig 3). The use of genotype-guided prescriptions was estimated to have the greatest impact for ibuprofen (1,417 patients) in the pediatric population (S5 Fig).

### Spectrum of rare, predicted deleterious variants that are potentially actionable

A total of 13,165 variants were identified in the 108 pharmacogenes and 3,586 of them (27.2%) were rare variants (gnomAD AF < 1%; S4 Table). Five hundred and thirty-one rare variants were predicted to be deleterious, including 475 (89.5%) missense variants and 56 (10.5%) LoF variants (S6 Fig and S9 Table). Among the 531 rare deleterious variants, 96 (18.1%) have never been reported in public databases including dbSNP, gnomAD, and ClinVar.

Out of the 108 pharmacogenes, 97 (89.8%) had at least one deleterious variant. Genes with the highest number of rare deleterious variants included *CACNA1S* (n = 26), *CFTR* (n = 26), and *LPA* (n = 24; S7 Fig). Analyzing using a per-sample approach, 1,043 (93.5%) individuals harbored at least one rare deleterious pharmacogenetic variant, and the median was two (Fig 2B). Considering only the ten cytochrome P450 (CYP) genes, 435 (39.0%) individuals carried at least one rare deleterious variant in a CYP gene.

### Discussion

As exome sequencing becomes increasingly available, it has the potential to facilitate personalized medicine without significant additional costs. This study demonstrated the secondary use of exome data for pharmacogenetic analysis, and we found that the majority of known

**Table 1. Frequency of actionable pharmacophenotypes in Hong Kong Chinese.**

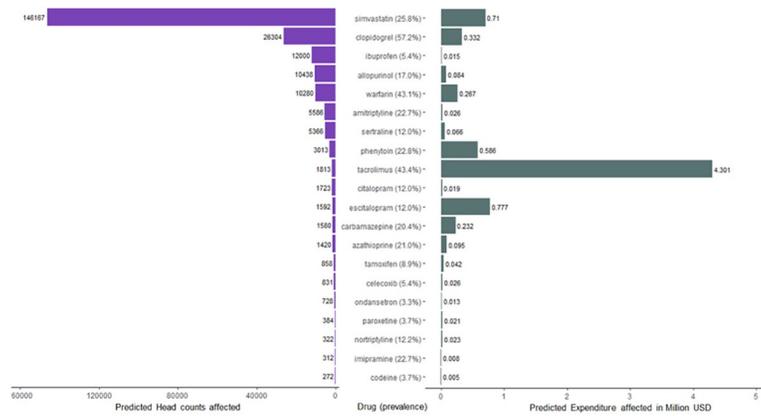
Gene	Allele	Actionable phenotype	Genotype definition	Frequency (%)	
				by phenotype	by gene
CYP2C19	No function: *2, *3, *6	Intermediate metabolizer	One normal function allele and one no function allele	45.25	57.21
		Poor metabolizer	Two no function alleles	11.96	
CYP3A5	Functional allele: *1	Extensive metabolizer	Two functional alleles	6.13	43.38
		Intermediate metabolizer	One functional allele and one nonfunctional allele	37.25	
CYP2B6	Decreased function: *6	Intermediate metabolizer	One normal function allele and one decreased function allele	35.28	40.51
		Poor metabolizer	Two decreased function alleles	5.23	
CYP4F2	rs2108622	Carrier of decreased function allele	Carrier of rs2108622 T allele	39.89	39.89
HLA-B	*15:02, *57:01, *58:01	*15:02 positive	Heterozygous or homozygous for HLA-B*15:02	18.42	33.88
		*57:01 positive	Heterozygous or homozygous for HLA-B*57:01	0.18	
		*58:01 positive	Heterozygous or homozygous for HLA-B*58:01	17.04	
SLCO1B1	rs4149056	Intermediate function	TC genotype at rs4149056	23.89	25.81
		Low function	CC genotype at rs4149056	1.92	
NUDT15	No function: *2, *3	Intermediate metabolizer	One normal function allele and one no function allele	17.63	18.58
		Poor metabolizer	Two no function alleles	0.95	
IFNL3	rs12979860	Unfavorable response genotype	Carrier of rs12979860 T allele	13.57	13.57
CYP2D6	Duplication of functional allele: *1x2, *1x3, *2x2, *2x3	Ultrarapid metabolizer	Carrier of duplications of functional alleles	3.3	12.24
				Reduced function: *10, *10x2, *14B, *36-*10, (*36-*10)x2, *41	
	No function: *4, *5, *6, *14A, *36, *36x2	Intermediate metabolizer	One decreased function and one no function allele	8.54	
		Poor metabolizer	Two no function alleles	0.4	
CYP2C9	No function: *3	Intermediate metabolizer	One normal function allele and one no function allele	5.32	5.39
		Poor metabolizer	Two no function alleles	0.07	
UGT1A1	Decreased function: *6, *28, *37	Poor metabolizer	Two decreased function alleles	5.08	5.08
TPMT	No function: *3A, *3C	Intermediate metabolizer	One normal function allele and one no function allele	2.92	2.94
		Poor metabolizer	Two no function alleles	0.02	
HLA-A	*31:01	*31:01 positive	Heterozygous or homozygous for HLA-A*31:01	2.48	2.48
G6PD	WHO class II-III: KaiPing, Canton, Gaohe, Chinese-5, Hechi, Maewo, Quing Yan	Deficient (male)	One deficient (class II–III) allele	3.68	3.68
		Deficient (female)	Two deficient (class II–III) alleles	0.14	0.14
CFTR	rs115545701, rs78655421	Favorable response genotype	Homozygous or heterozygous for CFTR variants listed in the FDA-approved drug label	0.09	0.09

In this study, 15 pharmacogenes were found to have actionable pharmacogenetic variants. The frequency of actionable phenotypes was derived based on the Hardy-Weinberg equation, except for *CYP2D6* where the frequency was retrieved from a published Hong Kong study by Chan et al.<sup>25</sup>

N/A, not applicable.

<https://doi.org/10.1371/journal.pgen.1009323.t001>

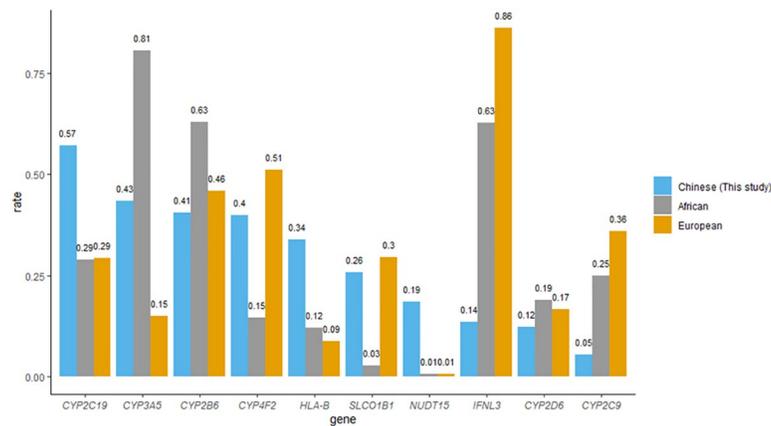
actionable pharmacogenetic variants and the coding regions of most pharmacogenes were well-covered. To our knowledge, this is the second study to investigate actionable pharmacogenetic variations in the Chinese population and it contained the largest sample size of Chinese subjects to date. Our data are representative of Southern Chinese, a Chinese subpopulation



**Fig 3. Top 20 drugs with the highest estimated prescription impact on headcount.** This figure illustrates the top 20 drugs with the highest estimated prescription impact on headcount (left panel) and their respective predicted expenditure (right panel). Overall, most of the projected prescription impact was concentrated in a few drugs. The top three drugs projected to affect the greatest number of patients were the lipid-lowering drug simvastatin (146,167 patients, frequency: 25.81%), clopidogrel (26,304 patients, frequency: 57.21%), and anti-inflammatory drug ibuprofen (12,000 patients, frequency: 5.39%). It was estimated that 8,219,000 USD worth of drugs were prescribed to patients with an implicated actionable phenotype, where tacrolimus (4,301,000 USD), escitalopram (777,000 USD), and simvastatin (710,000 USD) accounted for 70.4% of the total expenditure. CPIC only recommends genotype-guided prescription of clopidogrel in patients with acute coronary syndrome receiving percutaneous coronary intervention. Indications for prescription is however not included in the CDARS system.

<https://doi.org/10.1371/journal.pgen.1009323.g003>

that is underrepresented in the literature. To the best of our knowledge, we, for the first time, have provided information on rare, predicted deleterious pharmacogenetic variants in Chinese. Additionally, the potential prescription impact of actionable pharmacogenetic variants was projected in the HK population of 7.5 million.



**Fig 4. Frequency of the top ten actionable phenotypes in Hong Kong Chinese compared to that of Africans and Europeans.** This figure compares the top ten actionable phenotypes in Hong Kong Chinese with that of Africans and Europeans. The actionable phenotype frequency of Africans and Europeans was retrieved from data from Chanfreau-Coffinier et al., Walker et al. and supplementary information in the CPIC guideline [30,36]. Actionable phenotypes include *CYP2C19* IM and PM; *CYP3A5* EM and IM; *CYP2B6* IM and PM; carrier of *CYP4F2* decreased function allele; carrier of *HLA-B*\*15:02; \*57:01 and \*58:01; *SLCO1B1* intermediate and low-function diplotypes; *NUDT15* IM and PM; *IFNL3* unfavorable response genotype; *CYP2D6* UM, IM, and PM; and *CYP2C9* IM and PM. *CYP2C19*\*17 is not readily detected in exome sequencing and therefore *CYP2C19* RM and UM were not included. The frequency of actionable phenotypes in *CYP2C19*, *HLA-B*, and *NUDT15* was found to be higher in Hong Kong Chinese than in Europeans and Africans. In contrast, actionable genotypes in *CYP3A5*, *CYP2B6*, and *CYP2D6* were more prevalent in Africans, whereas actionable phenotypes in *CYP4F2*, *SLCO1B1*, *IFNL3*, and *CYP2C9* were more prevalent in Europeans; however, all of these genes were within the top ten actionable phenotypes in Hong Kong. IM, intermediate metabolizer; PM, poor metabolizer; EM, extensive metabolizer; UM, ultrarapid metabolizer; RM, rapid metabolizer.

<https://doi.org/10.1371/journal.pgen.1009323.g004>

## Burden and projected prescription impact of known actionable pharmacogenetic variations

Consistent with other populations, nearly all HK Chinese (99.6%) harbored at least one pharmacogenetic variant, with a median of four variants [7,8]. Nonetheless, the spectrum of actionable genotypes was different compared to that of African and European populations (Fig 4). The highest actionable phenotype in Europeans and Africans was *IFNL3* and *CYP3A5*, respectively, which had a frequency of more than 80% in their respective populations; however, less than half of the HK Chinese carried an actionable phenotype in these genes. In contrast, while *NUDT15* ranked seventh among the drugs with highest actionable phenotypes in HK Chinese (frequency: 18.58%), only 1% of Europeans and Africans carried actionable phenotypes in this gene. It is therefore more important to consider defective *NUDT15* alleles when azathioprine is prescribed in the Chinese population. In contrast to *NUDT15*, which is the key determinant of azathioprine-induced myelosuppression in Asians, defective *TPMT* alleles should be considered in Europeans instead [28,29]. This suggests that even when the same drug is prescribed, different pharmacogenes should be considered in each population. Furthermore, considering the same pharmacogene, the alleles contributing to actionable phenotypes could be different across different populations. For example, while *G6PD* deficiency is common in Africans and Chinese, the A allele alone explains nearly all *G6PD* deficiencies in Africans, whereas seven *G6PD* alleles contribute to the deficiency in HK Chinese [30].

Based on prescription data of the HK public healthcare system, 13.4% of the HK population (1,006,046 patients) received at least one of the 36 drugs with CPIC guideline recommendations. The total expenditure on CPIC actionable drugs in 2019 was 33,520,000 USD, and it was estimated that 8,219,000 USD (24.5%) worth of drugs were prescribed to subjects with an implicated actionable phenotype. Pharmacogenetic results would improve patient care and allocation of resources. For example, in order to reduce the risk of thiopurine-related myelosuppression, 1,354 (17.63%) patients taking azathioprine would require a starting dose reduction of 30–80%, and 66 (0.95%) patients would require alternative medications. Another example is the lipid-lowering drug simvastatin, which was found to have the greatest prescription impact in terms of headcount in our study. A previous study on *SLCO1B1* reported an 18% cumulative risk of simvastatin-induced myopathy for the CC genotype (frequency in HK: 1.92%) and a 3% cumulative risk for the CT genotype (frequency in HK: 23.89%) [31]. Projecting from the frequency of actionable phenotype derived from our dataset onto the CDARS prescription data, it was estimated that 146,167 patients prescribed with simvastatin had an actionable phenotype in 2019. If all of these patients prescribed with a lower dose of simvastatin or another statin, it was estimated that 6,019 cases of simvastatin-induced myopathy can be prevented. Dosing can also be increased based on genotype-guided prescription. Tacrolimus, an immunosuppressive drug used by organ transplant recipients to lower the risk of organ rejection, was predicted to affect 1,813 (43.38%) people, accounting for the highest expenditure (4,301,000 USD) among all the CPIC actionable drugs in HK. Based on the pharmacogenetic results, these 1,813 people required a 1.5 to 2 times increase in starting dose, further increasing the prescription cost. Nevertheless, genotype-guided prescription in tacrolimus is known to achieve therapeutic concentrations earlier and had fewer out-of-range concentrations when compared to standard dosing [32]. With expansion of the CPIC guidelines to include other commonly prescribed drugs such as tramadol and proton pump inhibitors, the prescription impact of pharmacogenetic testing is likely to increase in the near future.

## Spectrum of rare, deleterious pharmacogenetic variants

The effect of rare pharmacogenetic variants on drug responses should not be neglected, since they may account for nearly all inter-individual variabilities in more than half of the

pharmacogenes [11]. At present, pharmacogenetic testing is performed by SNP arrays targeting specific alleles and hence, the detection of rare variants is impossible. In contrast, exome or genome sequencing would result in an abundance of rare variants, but determining the effect of rare variants on drug responses is difficult without functional data [19]. In the present study, we aimed to maximize the discovery of potentially deleterious pharmacogenetic variants. We found that 93.5% of subjects carried at least one rare, deleterious pharmacogenetic variant, with a median of two variants. This suggests that despite being individually rare, pharmacogenetic variants with AF < 1% are collectively common. The rare, deleterious pharmacogenetic variants reported in this study can be prioritized for functional studies, especially for variants with consensus deleterious effects predicted across multiple bioinformatics tools and with known gene mechanisms. An example would be the *CYP2C9* splice variant c.1291+1G>T, which was predicted to be deleterious by both CADD and LOFTEE. Although this variant has not been reported in PharmVar, other LoF variants in *CYP2C9* have been reported as “no function” [26]. In the future, saturation mutagenesis studies will likely aid in determining how deleterious rare pharmacogenetic variants are [33].

### Study limitations

First, due to the technical limitations of exome sequencing, non-coding regions, copy number variations, structural variations, and loci with high genomic complexity either were poorly/not covered with exome sequencing or were difficult to detect with bioinformatics. For example, four actionable variants located in non-coding regions were not sequenced, and four variants, *UGT1A1*\*28, *IFNL3* rs12979860, *CYP2C19* rs4244285 and *CYP3A5* rs776746, were sequenced in less than 70% of the samples (S2 Table). The reported AF of these variants should therefore be interpreted with caution. Advancements in bioinformatics may have overcome some technical difficulties, such as using HLA-HD in this study for HLA typing. The AF of HLA identified in this study agreed well with that of the HK Bone Marrow Donor Registry (S6 Table), suggesting the accuracy of HLA-HD [34]. In the future, copy number variations and structural variations could be more reliably identified from exome data with bioinformatics. The second limitation is related to the projected impact of preemptive pharmacogenetic testing. The prescription data of private medical practitioners and over-the-counter medications were not available in this study. Therefore, our analysis was limited to prescription data in the public healthcare setting. Although our data were limited to public hospitals, the public healthcare system accounted for approximately 90% of all secondary and tertiary health services in HK. Lastly, indications for drug prescription could not be retrieved from the CDARS database. The CPIC guideline only recommended *CYP2C19* genotype-guided prescription of clopidogrel in patients with acute coronary syndrome receiving anti-platelet therapy for percutaneous coronary intervention, therefore the projected impact of clopidogrel could be over-estimated [35].

In conclusion, nearly all individuals carried at least one actionable pharmacogenetic variant and one rare, deleterious pharmacogenetic variant in our cohort. It was estimated that one-seventh of the HK population received at least one of the 36 drugs with CPIC guideline recommendations, and 8,219,000 USD worth of drugs were prescribed to patients with an implicated actionable phenotype. This indicates that preemptive pharmacogenetic testing has the potential to improve patient care and allocation of resources significantly.

### Supporting information

**S1 Text. Supplementary Methods.**  
(DOCX)

**S1 Table. Sequencing platform and capture kit used in the Hong Kong Chinese exome sequencing data.**

(XLSX)

**S2 Table. List of the 133 actionable pharmacogenetic variants/alleles.** <sup>#</sup> Call rate refers to the % of samples having coverage  $\geq 8X$  in the exome sequencing cohort. ADR, adverse drug reaction; PK, pharmacokinetics.

(XLSX)

**S3 Table. List of the 108 high confidence pharmacogenes.**

(XLSX)

**S4 Table. Spectrum of variants identified in the 108 high-confidence pharmacogenes.**

(XLSX)

**S5 Table. Actionable pharmacogenetic variants identified in the Hong Kong Chinese exome sequencing cohort.** <sup>#</sup> Call rate refers to the % of samples having coverage  $\geq 8X$  in the exome sequencing cohort. <sup>^</sup> \*4, \*10, \*36, \*47, \*49, \*54, \*56, \*57, \*69, \*72, \*99, \*100, \*101, \*114. AF<sub>HK</sub>, allele frequency in Hong Kong; LoF, loss-of-function; NSAID, non-steroidal anti-inflammatory drug (includes celecoxib, ibuprofen, and piroxicam); SSRI, selective serotonin reuptake inhibitor (includes citalopram, escitalopram, and sertraline); TCA, tricyclic antidepressant (includes amitriptyline, clomipramine, doxepin, imipramine, and trimipramine).

(XLSX)

**S6 Table. Table of actionable HLA risk alleles identified in the Hong Kong Chinese exome sequencing cohort.** AF<sub>HK</sub>, allele frequency in Hong Kong. <sup>a</sup>Data from Hong Kong Bone Marrow Donor Registry (HKBM DR)[34]; <sup>b</sup>Data from Chinese Bone Marrow Donor Program (CBMD)[37].

(XLSX)

**S7 Table. Frequency of actionable phenotype(s) and projected impact for each drug.** <sup>a</sup>Carrier rate of *G6PD* deficiency projected from male-female ratio of 45.6:54.4 in Hong Kong.<sup>b</sup>Variants include p.(Arg74Trp) and p.(Arg117His). Cystic fibrosis is uncommon in Chinese, hence Ivacaftor is not prescribed in public healthcare systems. UM, ultrarapid metabolizer; IM, intermediate metabolizer; EM, extensive metabolizer; PM, poor metabolizer.

(XLSX)

**S8 Table. The number and percentage of people prescribed with one of the 36 drugs with CPIC guideline recommendation in different age group.** Based on the available prescription data in CDARS in 2019, a total of 1,006,046 Hong Kong Chinese patients had received at least one of the 36 drugs with CPIC guideline recommendations, representing 13.4% of the Hong Kong population. The percentage of people prescribed with one of the 36 drugs with CPIC guideline recommendations increased with age.

(XLSX)

**S9 Table. List of 531 rare, deleterious variants identified in this study.**

(XLSX)

**S1 Fig. Analysis flowchart of this study.** Concatenated exome sequencing data was first run through sample-, variant-, and genotype-level quality control (QC) procedures. In the analysis of known actionable pharmacogenetic variants, we first extracted data based on a curated list of 129 variants and four HLA alleles, and subsequently projected the prescription impact in the Hong Kong public healthcare system. We further processed the dataset for analysis of rare

variants of the 108 high-confidence pharmacogenes. The final list of rare, predicted deleterious variants included only missense and loss-of-function (LoF) variants with gnomAD allele frequency (AF) <1% and at least one deleterious prediction by bioinformatics algorithm.

(TIF)

**S2 Fig. Coverage of the coding regions of the 108 high-confidence pharmacogenes.** In general, exome sequencing covered the 108 high-confidence pharmacogenes well, with 104 of them having a mean coverage of at least 8X in over 75% of the samples. Genes that did not have a mean coverage of 8X included *CCHCR1*, *TNF*, *IFNL4*, and *GSTM1*. Mean coverages of 30X and 50X were achieved in over 75% of samples for 90 and 31 genes, respectively.

(TIF)

**S3 Fig. Spectrum and functional consequence of variants identified in the 108 high-confidence pharmacogenes.** A total of 13,165 variants were identified in the 108 high-confidence pharmacogenes, with 10,192 (77.4%) being intronic variants. Coding variants accounted for 1,719 of the variants, with the majority (58.6%) being nonsynonymous variants and 2.5% being loss-of-function (frameshift, stop-gain, and start-loss) variants. SNV, single-nucleotide variant; UTR3, 3' untranslated region; UTR5, 5' untranslated region.

(TIF)

**S4 Fig. Cumulative known actionable variant count by proportion of samples with specified coverage.** In our cohort, >8X read depth and >30X read depth were achieved in >90% of samples in 121/129 (93.8%) and 62/129 (48.1%) known actionable variants, respectively.

(TIF)

**S5 Fig. Top 20 drugs with the highest estimated prescription impact on headcount in the pediatric population (age < 19).** The top three drugs with the highest pharmacogenetic impact in the pediatric population (age <19) based on headcount were ibuprofen (1417 patients, frequency:5.39%), atomoxetine (235 patients, frequency:12.24%), and sertraline (156 patients, frequency:11.96%). The top three drugs with highest pharmacogenetic impact based on expenditures were tacrolimus (238,000 USD), atomoxetine (48,000 USD), and escitalopram (32,000 USD).

(TIF)

**S6 Fig. Venn diagrams showing the overlapping deleterious predictions using different bioinformatics tools.** (A) Among the 829 rare (gnomAD global AF <1%) missense variants in the 108 high-confidence pharmacogenes, 475 variants were predicted to be deleterious by at least one of the three bioinformatics tools (CADD, REVEL, and PREDICT), and 89 variants had consensus deleterious predictions. There were 354 rare missense variants that were not predicted as deleterious by all of the bioinformatics tools. (B) Among the 63 rare LoF variants, 56 were predicted to be deleterious by either CADD or LOFTEE, and 43 variants had consensus deleterious predictions. There were 7 rare LoF variants that were not predicted as deleterious by both of the bioinformatics tools. AF, allele frequency; LoF, loss-of-function.

(TIF)

**S7 Fig. Manhattan plot of gene distribution among rare, predicted deleterious variants.** The Manhattan plot summarizes the gene distribution of 531 rare, predicted deleterious variants identified and their chromosome number (chr) in the human genome. *CACNA1S* (n = 26), *CFTR* (n = 26), and *LPA* (n = 24) had the highest number of rare deleterious variants among the 108 high-confidence pharmacogenes. LoF, loss-of-function.

(TIF)

## Author Contributions

**Conceptualization:** Mullin Ho Chung Yu, Claudia Ching Yan Chung, Brian Hon Yin Chung.

**Data curation:** Mullin Ho Chung Yu, Marcus Chun Yin Chan, Claudia Ching Yan Chung, Andrew Wang Tat Li.

**Formal analysis:** Mullin Ho Chung Yu, Marcus Chun Yin Chan, Claudia Ching Yan Chung, Wilfred Hing Sang Wong, Kit San Yeung.

**Funding acquisition:** Patrick Ho Yu Chung, Brian Hon Yin Chung.

**Investigation:** Mullin Ho Chung Yu, Marcus Chun Yin Chan, Claudia Ching Yan Chung.

**Methodology:** Mullin Ho Chung Yu, Marcus Chun Yin Chan, Claudia Ching Yan Chung, Christopher Chun Yu Mak, Jeffrey Fong Ting Chau, Mianne Lee, Jasmine Lee Fong Fung, Mandy Ho Yin Tsang, Joshua Chun Ki Chan, Clara Sze Man Tang.

**Project administration:** Clara Sze Man Tang, Kit San Yeung, Brian Hon Yin Chung.

**Resources:** Mullin Ho Chung Yu, Marcus Chun Yin Chan, Andrew Wang Tat Li, Chara Yin Wa Yip, Wilfred Hing Sang Wong, Jing Yang, William Chun Ming Chui, Patrick Ho Yu Chung, Wanling Yang, So Lun Lee, Godfrey Chi Fung Chan, Paul Kwong Hang Tam, Yu Lung Lau, Clara Sze Man Tang.

**Software:** Mullin Ho Chung Yu, Wilfred Hing Sang Wong, Clara Sze Man Tang.

**Supervision:** Kit San Yeung, Brian Hon Yin Chung.

**Validation:** Mullin Ho Chung Yu, Marcus Chun Yin Chan, Clara Sze Man Tang, Kit San Yeung.

**Visualization:** Mullin Ho Chung Yu, Marcus Chun Yin Chan, Claudia Ching Yan Chung.

**Writing – original draft:** Mullin Ho Chung Yu, Marcus Chun Yin Chan, Claudia Ching Yan Chung.

**Writing – review & editing:** Clara Sze Man Tang, Kit San Yeung, Brian Hon Yin Chung.

## References

1. International Conference on Harmonisation; Guidance on E15 Pharmacogenomics Definitions and Sample Coding; Availability. Notice. Fed Regist. 2008; 73(68):19074–6. Epub 2008/08/06. PMID: [18677821](https://pubmed.ncbi.nlm.nih.gov/18677821/).
2. Doble B, Schofield DJ, Roscioli T, Mattick JS. Prioritising the application of genomic medicine. NPJ Genom Med. 2017; 2:35. Epub 2017/12/22. <https://doi.org/10.1038/s41525-017-0037-0> PMID: [29263844](https://pubmed.ncbi.nlm.nih.gov/29263844/); PubMed Central PMCID: PMC5698310.
3. Roden DM, Van Driest SL, Mosley JD, Wells QS, Robinson JR, Denny JC, et al. Benefit of Preemptive Pharmacogenetic Information on Clinical Outcome. Clin Pharmacol Ther. 2018; 103(5):787–94. Epub 2018/01/30. <https://doi.org/10.1002/cpt.1035> PMID: [29377064](https://pubmed.ncbi.nlm.nih.gov/29377064/); PubMed Central PMCID: PMC6134843.
4. Relling MV, Klein TE. CPIC: Clinical Pharmacogenetics Implementation Consortium of the Pharmacogenomics Research Network. Clin Pharmacol Ther. 2011; 89(3):464–7. Epub 2011/01/29. <https://doi.org/10.1038/clpt.2010.279> PMID: [21270786](https://pubmed.ncbi.nlm.nih.gov/21270786/); PubMed Central PMCID: PMC3098762.
5. Barbarino JM, Whirl-Carrillo M, Altman RB, Klein TE. PharmGKB: A worldwide resource for pharmacogenomic information. Wiley Interdiscip Rev Syst Biol Med. 2018; 10(4):e1417. Epub 2018/02/24. <https://doi.org/10.1002/wsbm.1417> PMID: [29474005](https://pubmed.ncbi.nlm.nih.gov/29474005/); PubMed Central PMCID: PMC6002921.
6. McDonagh EM, Whirl-Carrillo M, Garten Y, Altman RB, Klein TE. From pharmacogenomic knowledge acquisition to clinical applications: the PharmGKB as a clinical pharmacogenomic biomarker resource. Biomark Med. 2011; 5(6):795–806. Epub 2011/11/23. <https://doi.org/10.2217/bmm.11.94> PMID: [22103613](https://pubmed.ncbi.nlm.nih.gov/22103613/); PubMed Central PMCID: PMC3339046.

7. Bush WS, Crosslin DR, Owusu-Obeng A, Wallace J, Almoquera B, Basford MA, et al. Genetic variation among 82 pharmacogenes: The PGRNseq data from the eMERGE network. *Clin Pharmacol Ther.* 2016; 100(2):160–9. Epub 2016/02/10. <https://doi.org/10.1002/cpt.350> PMID: 26857349; PubMed Central PMCID: PMC5010878.
8. Wright GEB, Carleton B, Hayden MR, Ross CJD. The global spectrum of protein-coding pharmacogenomic diversity. *Pharmacogenomics J.* 2018; 18(1):187–95. Epub 2016/10/26. <https://doi.org/10.1038/tj.2016.77> PMID: 27779249; PubMed Central PMCID: PMC5817389.
9. Gordon AS, Tabor HK, Johnson AD, Snively BM, Assimes TL, Auer PL, et al. Quantifying rare, deleterious variation in 12 human cytochrome P450 drug-metabolism genes in a large-scale exome dataset. *Hum Mol Genet.* 2014; 23(8):1957–63. Epub 2013/11/28. <https://doi.org/10.1093/hmg/ddt588> PMID: 24282029; PubMed Central PMCID: PMC3959810.
10. Fujikura K, Ingelman-Sundberg M, Lauschke VM. Genetic variation in the human cytochrome P450 supergene family. *Pharmacogenet Genomics.* 2015; 25(12):584–94. Epub 2015/09/05. <https://doi.org/10.1097/FPC.000000000000172> PMID: 26340336.
11. Ingelman-Sundberg M, Mkrтчian S, Zhou Y, Lauschke VM. Integrating rare genetic variants into pharmacogenetic drug response predictions. *Hum Genomics.* 2018; 12(1):26. Epub 2018/05/26. <https://doi.org/10.1186/s40246-018-0157-3> PMID: 29793534; PubMed Central PMCID: PMC5968569.
12. 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 2017/04/06. <https://doi.org/10.1002/cpt.690> PMID: 28378927; PubMed Central PMCID: PMC5600063.
13. Qin W, Du Z, Xiao J, Duan H, Shu Q, Li H. Evaluation of clinical impact of pharmacogenomics knowledge involved in CPIC guidelines on Chinese pediatric patients. *Pharmacogenomics.* 2020; 21(3):209–19. Epub 2020/01/23. <https://doi.org/10.2217/pgs-2019-0153> PMID: 31967514.
14. Kawaguchi S, Higasa K, Shimizu M, Yamada R, Matsuda F. HLA-HD: An accurate HLA typing algorithm for next-generation sequencing data. *Hum Mutat.* 2017; 38(7):788–97. Epub 2017/04/19. <https://doi.org/10.1002/humu.23230> PMID: 28419628.
15. Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, et al. Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv.* 2018:201178. <https://doi.org/10.1101/201178>
16. Yang H, Wang K. Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nat Protoc.* 2015; 10(10):1556–66. Epub 2015/09/18. <https://doi.org/10.1038/nprot.2015.105> PMID: 26379229; PubMed Central PMCID: PMC4718734.
17. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* 2019; 47(D1):D886–d94. Epub 2018/10/30. <https://doi.org/10.1093/nar/gky1016> PMID: 30371827; PubMed Central PMCID: PMC6323892.
18. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet.* 2016; 99(4):877–85. Epub 2016/09/27. <https://doi.org/10.1016/j.ajhg.2016.08.016> PMID: 27666373; PubMed Central PMCID: PMC5065685.
19. Zhou Y, Mkrтчian S, Kumondai M, Hiratsuka M, Lauschke VM. An optimized prediction framework to assess the functional impact of pharmacogenetic variants. *Pharmacogenomics J.* 2019; 19(2):115–26. Epub 2018/09/13. <https://doi.org/10.1038/s41397-018-0044-2> PMID: 30206299; PubMed Central PMCID: PMC6462826 that they have no conflict of interest.
20. Karczewski K. LOFTEE (Loss-Of-Function Transcript Effect Estimator) [cited 2019 4 April]. Available from: <https://github.com/konradjk/loftee>.
21. Gaedigk A, Sangkuhl K, Whirl-Carrillo M, Twist GP, Klein TE, Miller NA. The Evolution of PharmVar. *Clin Pharmacol Ther.* 2019; 105(1):29–32. Epub 2018/12/12. <https://doi.org/10.1002/cpt.1275> PMID: 30536702; PubMed Central PMCID: PMC6312487.
22. Zhang X, Yin JF, Zhang J, Kong SJ, Zhang HY, Chen XM. UGT1A1\*6 polymorphisms are correlated with irinotecan-induced neutropenia: a systematic review and meta-analysis. *Cancer Chemother Pharmacol.* 2017; 80(1):135–49. Epub 2017/06/07. <https://doi.org/10.1007/s00280-017-3344-3> PMID: 28585035.
23. SK M. Recent perspectives in glucose-6-phosphate dehydrogenase (G6PD) deficiency. Topical Update—The Hong Kong College of Pathologists 2007 [cited 2020 23 January]. Available from: <http://www.hkcpath.org/article/recent-perspectives-glucose-6-phosphate-dehydrogenase-g6pd-deficiency>.
24. Census and Statistics Department TGoHK SAR. Women and men in Hong Kong—Key statistics, 2019 Edition 2019 [cited 2020 23 January]. Available from: <https://www.censtatd.gov.hk/hkstat/sub/sp180.jsp?productCode=B1130303>.

25. Hong Kong Hospital Authority. Hospital authority statistical report 2015–2016 2016 [cited 2020 24 March]. Available from: <https://www3.ha.org.hk/data/HAStatistics/StatisticalReport/>.
26. Hospital Authority. Hospital Authority Annual Report 2018–2019 [cited 2020 24 March]. Available from: <https://www.legco.gov.hk/yr19-20/chinese/counmtg/papers/cm20191218-sp077-ec.pdf>
27. Chan W, Li MS, Sundaram SK, Tomlinson B, Cheung PY, Tzang CH. CYP2D6 allele frequencies, copy number variants, and tandems in the population of Hong Kong. *J Clin Lab Anal.* 2019; 33(1):e22634. Epub 2018/08/03. <https://doi.org/10.1002/jcla.22634> PMID: 30069923; PubMed Central PMCID: PMC6430334.
28. Moriyama T, Nishii R, Perez-Andreu V, Yang W, Klussmann FA, Zhao X, et al. NUDT15 polymorphisms alter thiopurine metabolism and hematopoietic toxicity. *Nat Genet.* 2016; 48(4):367–73. Epub 2016/02/16. <https://doi.org/10.1038/ng.3508> PMID: 26878724; PubMed Central PMCID: PMC5029084.
29. Ford LT, Berg JD. Thiopurine S-methyltransferase (TPMT) assessment prior to starting thiopurine drug treatment; a pharmacogenomic test whose time has come. *J Clin Pathol.* 2010; 63(4):288–95. Epub 2010/04/01. <https://doi.org/10.1136/jcp.2009.069252> PMID: 20354201.
30. Chanfreau-Coffinier C, Hull LE, Lynch JA, DuVall SL, Damrauer SM, Cunningham FE, et al. Projected Prevalence of Actionable Pharmacogenetic Variants and Level A Drugs Prescribed Among US Veterans Health Administration Pharmacy Users. *JAMA Netw Open.* 2019; 2(6):e195345. Epub 2019/06/08. <https://doi.org/10.1001/jamanetworkopen.2019.5345> PMID: 31173123; PubMed Central PMCID: PMC6563578.
31. Link E, Parish S, Armitage J, Bowman L, Heath S, Matsuda F, et al. SLCO1B1 variants and statin-induced myopathy—a genomewide study. *N Engl J Med.* 2008; 359(8):789–99. Epub 2008/07/25. <https://doi.org/10.1056/NEJMoa0801936> PMID: 18650507.
32. Min S, Papaz T, Lafreniere-Roula M, Nalli N, Grasemann H, Schwartz SM, et al. A randomized clinical trial of age and genotype-guided tacrolimus dosing after pediatric solid organ transplantation. *Pediatr Transplant.* 2018; 22(7):e13285. Epub 2018/09/05. <https://doi.org/10.1111/ptr.13285> PMID: 30178515.
33. Matreyek KA, Starita LM, Stephany JJ, Martin B, Chiasson MA, Gray VE, et al. Multiplex assessment of protein variant abundance by massively parallel sequencing. *Nat Genet.* 2018; 50(6):874–82. Epub 2018/05/23. <https://doi.org/10.1038/s41588-018-0122-z> PMID: 29785012; PubMed Central PMCID: PMC5980760.
34. Kwok J, Guo M, Yang W, Lee CK, Ho J, Tang WH, et al. HLA-A, -B, -C, and -DRB1 genotyping and haplotype frequencies for a Hong Kong Chinese population of 7595 individuals. *Hum Immunol.* 2016; 77(12):1111–2. Epub 2016/10/25. <https://doi.org/10.1016/j.humimm.2016.10.005> PMID: 27769748.
35. Scott SA, Sangkuhl K, Gardner EE, Stein CM, Hulot JS, Johnson JA, et al. Clinical Pharmacogenetics Implementation Consortium guidelines for cytochrome P450-2C19 (CYP2C19) genotype and clopidogrel therapy. *Clin Pharmacol Ther.* 2011; 90(2):328–32. Epub 2011/07/01. <https://doi.org/10.1038/clpt.2011.132> PMID: 21716271; PubMed Central PMCID: PMC3234301.
36. Walker GJ, Harrison JW, Heap GA, Voskuil MD, Andersen V, Anderson CA, et al. Association of Genetic Variants in NUDT15 With Thiopurine-Induced Myelosuppression in Patients With Inflammatory Bowel Disease. *Jama.* 2019; 321(8):773–85. Epub 2019/02/27. <https://doi.org/10.1001/jama.2019.0709> PMID: 30806694; PubMed Central PMCID: PMC6439872.
37. He Y, Li J, Mao W, Zhang D, Liu M, Shan X, et al. HLA common and well-documented alleles in China. *Hla.* 2018; 92(4):199–205. Epub 2018/08/04. <https://doi.org/10.1111/tan.13358> PMID: 30073798.