Mapping the pharmacogenetic landscape in a Ugandan population: implications for personalized medicine in an underrepresented population

Samarasinghe, S.R.; Lee, S.B.; Corpas, M.; Fatumo, S.; Guchelaar, H.J.; Nagaraj, S.H.

Mapping the Pharmacogenetic Landscape in a Ugandan Population: Implications for Personalized Medicine in an Underrepresented Population

Sumudu Rangika Samarasinghe[1], Seung-been Lee[2], Manuel Corpas[3], Segun Fatumo[4,5], Henk-Jan Guchelaar[6] and Shivashankar H. Nagaraj[1,7,*]

Africans are extremely underrepresented in global genomic research. African populations face high burdens of communicable and non-communicable diseases and experience widespread polypharmacy. As population-specific genetic studies are crucial to understanding unique genetic profiles and optimizing treatments to reduce medication-related complications in this diverse population, the present study aims to characterize the pharmacogenomics profile of a rural Ugandan population. We analyzed low-pass whole genome sequencing data from 1998 Ugandans to investigate 18 clinically actionable pharmacogenes in this population. We utilized PyPGx to identify star alleles (haplotype patterns) and compared allele frequencies across populations using the Pharmacogenomics Knowledgebase PharmGKB. Clinical interpretations of the identified alleles were conducted following established dosing guidelines. Over 99% of participants displayed actionable phenotypes across the 18 pharmacogenes, averaging 3.5 actionable genotypes per individual. Several variant alleles known to affect drug metabolism (i.e., *CYP3A5*1*, *CYP2B6*9*, *CYP3A5*6*, *CYP2D6*17*, *CYP2D6*29*, and *TMPT*3C*)—which are generally more prevalent in African individuals—were notably enriched in the Ugandan cohort, beyond reported frequencies in other African peoples. More than half of the cohort exhibited a predicted impaired drug response associated with CFTR, IFNL3, CYP2B6, and CYP2C19, and approximately 31% predicted altered CYP2D6 metabolism. Potentially impaired CYP2C9, SLCO1B1, TPMT, and DPYD metabolic phenotypes were also enriched in Ugandans compared with other African populations. Ugandans exhibit distinct allele profiles that could impact drug efficacy and safety. Our findings have important implications for pharmacogenomics in Uganda, particularly with respect to the treatment of prevalent communicable and non-communicable diseases, and they emphasize the potential of pharmacogenomics-guided therapies to optimize healthcare outcomes and precision medicine in Uganda.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

☑ Several attempts to study certain pharmacogenes in Africans, including a few studies from different parts of Uganda, have been reported. However, due to the high ethnolinguistic diversity in Africa, it is unclear if these studies fully capture the pharmacogenomics (PGx) diversity in the region.

WHAT QUESTION DID THIS STUDY ADDRESS?

☑ This study addresses the lack of African genetic diversity in PGx research by investigating the unique PGx profile of a rural Ugandan population compared with global populations, including other African ancestries. It aims to assess the potential benefits of PGx testing in the context of the region's high disease burden.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

☑ Over 99% of participants predicted actionable phenotypes, averaging 3.5 per individual. We identified clinically important PGx alleles prevalent in Ugandans, surpassing global frequencies, including those previously reported in Africans. Additionally, we report novel PGx variants with potential actionable implications.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

☑ Our findings may aid in developing cost-effective PGx panels that capture population-specific drug response profiles. This enables tailored genotype-guided treatments for a broader African community, supporting robust precision public health strategies in high-risk groups for widespread benefits.

Received December 7, 2023; accepted April 27, 2024. doi:10.1002/cpt.3309

[1]Centre for Genomics and Personalised Health, Queensland University of Technology, Brisbane, Queensland, Australia; [2]Macrogen Inc, Seoul, Korea; [3]College of Liberal Arts and Sciences, University of Westminster, London, UK; [4]Department of Non-communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK; [5]Precision Healthcare University Research Institute, Queen Mary University of London, London, UK; [6]Department of Clinical Pharmacy and Toxicology, Leiden University Medical Center, Leiden, The Netherlands; [7]Translational Research Institute, Queensland University of Technology, Brisbane, Queensland, Australia. *Correspondence: Shivashankar H. Nagaraj (shiv.nagaraj@qut.edu.au)

Africa comprises a multitude of ethnically distinct populations and is home to the highest level of genetic diversity worldwide.[1] The African continent bears a substantial burden of disease, accounting for >20% of the global disease burden.[2] Infectious diseases, notably malaria, tuberculosis (TB), and HIV/AIDS, are the leading causes of mortality in the region.[2] Additionally, noncommunicable diseases, including cardiovascular disease (CVD), diabetes, cancer, and chronic respiratory diseases, are rapidly increasing in prevalence.[2,3] The complications arising from comorbidities and coinfections (i.e., HIV/TB and HIV/malaria), drug–drug interactions owing to polypharmacy, and emerging drug resistance (i.e., multidrug-resistant TB) also present considerable challenges for patient treatment and place additional strain on the African healthcare system.[4] Given the complex genetic diversity in Africa and the high burden of disease, it is crucial to conduct population-specific genetic studies in the context of precision medicine to advance clinical genomic research within the region.[4–6]

Over the past decade, pharmacogenomics (PGx)—the study of genetics in drug responses—has evolved as an effective tool to improve the efficacy and safety of treatments, and the clinical implementation of PGx has proven to play an important role in disease control and patient management in many developed countries.[7–9] However, the majority of PGx studies have focused on people of European ancestry, whereas African populations remain largely underrepresented.[6,10,11] As variability in drug response is largely influenced by geographical and ethnic backgrounds,[12,13] more genetic studies are essential to overcome the barriers currently faced by African healthcare systems.[14] There are substantial variations in the distribution of alleles known to affect drug metabolism in African populations, such as *G6PD* and several *CYP* alleles (e.g., *CYP2D6*17*, *CYP2D6*29*, *CYP2B6*6*, *CYP2B6*18*, and *CYP2C19*17*) (see Sitabule et al. for review),[5] highlighting the need to better characterize population-specific genetic profiles to accurately implement PGx interventions across diverse African populations. Despite this, PGx studies within African populations are limited.[15] Of >100 drugs with clinical PGx guidelines,[16] only approximately 15 have been reported in studies that explore the genetic basis of drug response in African populations,[5,14] resulting in a notable gap that limits the clinical benefits of the latest innovations in genomics and precision medicine for African populations.

Located in East Africa, Uganda is home to a population composed of diverse ethnolinguistic groups.[6,17] Modern Ugandans represent a blend of numerous structured populations that have been separated for many years. Additionally, Uganda's history is marked by a rich tapestry of migrations from neighboring regions spanning centuries, followed by substantial intra-regional movement and genetic mixing. The presence of haplotypes believed to have originated from out-of-Africa suggests Eurasian back-migration into Uganda, shaping its genetic complexity.[17] Like many other populations in Africa, Ugandans face a high burden of disease and substantial complications from coinfections that require the concurrent use of multiple drugs for prolonged treatment periods.[18–20] Uganda is one of the most heavily impacted countries in the world with respect to HIV/TB burden, with a 59.4% HIV prevalence rate among drug-resistant TB patients.[21,22] Similarly, persistently high HIV/malaria rates have been documented among Ugandans.[23] As a result, Ugandans face issues such as the high risk of reduced drug efficacy, treatment failures, and adverse drug reactions (ADRs) to commonly used antimalarials, antibiotics, antiretrovirals, and anti-TB drugs, as well as the persistent threat of drug-resistant microbial selection.[24–26] Despite the high burden of disease and polypharmacy, only 7% of PGx studies have been conducted in East Africans, while the vast majority of PGx studies (75%) include North Africans or African Americans.[14]

Efforts to identify genetic factors influencing the efficacy and safety of common drugs among Ugandans have been limited. Some studies have explored the impact of specific genetic variants on drug responses in this population. For instance, *CYP2B6* 516G>T and 983T>C on efavirenz pharmacokinetics in Ugandan and Zambian children[27] and variants in *NAT2*, *SLCO1B1*, and *PXR* on isoniazid exposure among HIV/TB-coinfected Ugandan patients.[28] More recently, research in Ugandan and South African populations has identified additional relevant genetic variants (i.e., *CYP2C9*8*, **9*, **11*, and *CYP2C* cluster SNP rs12777823) for warfarin response.[29] However, a comprehensive understanding of the full range of pharmacogenetic variants in this population is still lacking.

To address this gap, we used low-pass whole genome sequencing (WGS) data from 1998 individuals from a geographically defined rural population in southwest Uganda.[17] Our study, based on the largest WGS dataset from Africa to date,[17] represents the most extensive PGx characterization in an African population. It provides actionable insight across various genes for a genetically distinct Ugandan population, specifically focusing on common medications for prevalent diseases in Uganda. We anticipate that our discoveries will pave the way for robust precision public health strategies in this high-risk population.

MATERIALS AND METHODS

Study population

The Uganda Genome Resource (UGR) consists of genotype array data for 5,000 individuals (genotyped using Illumina HumanOmni2.5–8 chip) and WGS data for 2000 individuals (UG2G; sequenced using Illumina HiSeq 2000 technology with 75 bp paired-end reads at 4x average coverage per sample) from the Uganda General Population Cohort (GPC), which comprises members from a rural community in southwest Uganda[17] (**Text S1**). Alignment BAM/CRAM files (GRCh37) for WGS data for 1998 individuals and genotype array data for 4,778 individuals from the UGR[17] were included in this study.

Selection of pharmacologically relevant genes for analysis

We selected the following 18 clinically actionable pharmacogenes with associated clinical guidelines from the Clinical Pharmacogenetics Implementation Consortium (CPIC)[16] (https://cpicpgx.org) to characterize the actionable PGx profile of this rural Ugandan community: *CYP2D6*, *CYP2C9*, *CYP2C19*, *CYP3A5*, *CYP4F2*, *CYP2B6*, *VKORC1*, *IFNL3*, *UGT1A1*, *ABCG2*, *SLCO1B1*, *RYR1*, *CFTR*, *TPMT*, *DPYD*, *NUDT15*, *G6PD*, and *CACNA1S*. While there are no official guidelines, we included the *CYP2C8* and *NAT2* genes because of their relevance to certain frequently prescribed medications for African populations. Additionally, we included 67 Very Important Pharmacogenes (VIPs)—defined by PharmGKB[30] as being significantly associated with drug administration, distribution, metabolism, excretion (ADME) functions, and drug targets (**Text S1**)—to identify population-specific known and novel variants that are not included in current star allele definitions but are potentially actionable in this cohort.

Identification of star alleles/phased haplotype patterns

Initial alignment BAM/CRAM files were preprocessed using SAMtools v.1.9[31] and Genome Analysis Toolkit (GATK) v4.1.9.0[32] (**Text S1**).

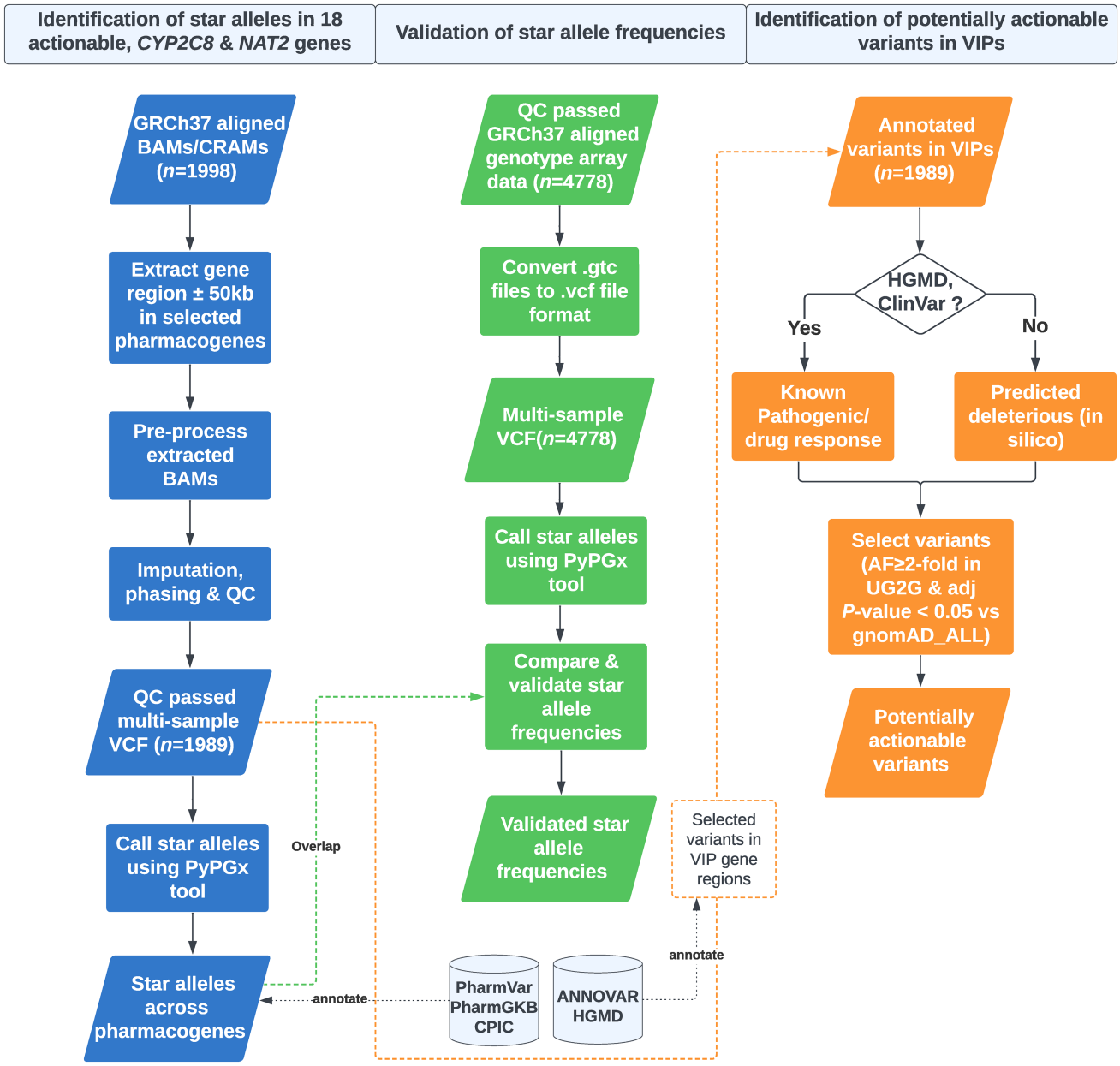


Figure 1 Workflow for identifying drug response-related variants in a Ugandan cohort. The key steps involved in the identification and validation of star alleles and the identification of potentially actionable variants across important pharmacogenes in the cohort are as follows: (i) Star alleles in 18 clinically actionable pharmacogenes, *CYP2C8* and *NAT2* genes were identified by analyzing WGS data from 1989 Ugandan individuals; (ii) Validation of the identified star alleles was performed using an orthogonal genotype array dataset consisting of 4,778 Ugandan individuals from the same region; (iii) Potentially actionable, known pathogenic, and predicted deleterious variants in VIPs—as defined by PharmGKB—were identified using WGS data. BAMs, binary alignment map files; CPIC, Clinical pharmacogenetics implementation consortium; CRAMs, compressed reference-oriented alignment map files; gtc, genotype call files; HGMD, Human gene mutation database; PharmGKB, Pharmacogenomics knowledge base; PharmVar, Pharmacogene variation consortium; VCF, variant call format; VIPs, very important pharmacogenes; WGS, whole genome sequencing.

Imputation, phasing, and QC steps were followed using BCFtools v.1.10.2,[31] GLIMPSE software suite v1.1.1,[33] and PLINK v1.9.[34] GLIMPSE takes advantage of reference panels to generate high-quality genotype calls and accurate phased haplotypes for low-pass sequenced datasets[33] (see **Text S1** for details). Highly related individuals (PLINK—king-cutoff of 0.177) were excluded. The filtered per-chromosome multisample VCF files were used as inputs in PyPGx v0.19.0[35,36] to identify star alleles across 18 clinically actionable pharmacogenes, *CYP2C8* and *NAT2* genes (**Figure 1**). PyPGx calls the two-star alleles and the corresponding genotype per gene per individual, and the metabolic phenotype (normal metabolizer (NM), intermediate metabolizer (IM), poor metabolizer (PM), rapid metabolizer (RM), ultrarapid metabolizer (UM), etc.) is assigned using PyPGx and CPIC based on the called genotype. Allele frequencies (AFs) for star allele haplotypes, diplotypes, and phenotypes for global populations (Sub-Saharan African (SSA), African American (AA), American, European, South Asian, East Asian, and Oceanian) were assigned from PharmGKB gene information tables,[30] except *NAT2* frequencies were adopted from the literature[37,38] (**Text S1**).

Validation of allele frequencies in an orthogonal genomic dataset

This analysis was conducted as an additional layer of evidence to substantiate the accuracy of our findings using low-pass WGS data. We called star alleles in the genotype array data corresponding to 4,778 Ugandan individuals in 18 actionable pharmacogenes, *CYP2C8* and *NAT2*, using PyPGx. Star alleles detected in both datasets (i.e., WGS and genotype array data) were selected for a frequency/proportion comparison. We conducted a two-sample z-test for equality of proportions with a continuity correction (**Text S1**). We ultimately selected 30 star alleles across 11 pharmacogenes for frequency comparison. Of these, for 26 alleles, there was insufficient evidence to support the claim that their frequencies/proportions were significantly different, and the observed concordance between the two datasets was 87%. Furthermore, when comparing frequencies at the variant level, we observed a concordance of 93.9% (**Table S1**; see **Text S1** for more details).

Analysis of potentially actionable variants in VIPs

The variants in the 67 VIPs were selected from phased and imputed variant files (**Figure 1**). The VCF was then annotated using the Human Gene Mutation Database (HGMD).[39] ANNOVAR was used to add ClinVar annotations, dbSNP150 rsIDs, and AFs for global populations (1,000 genomes and gnomAD).[40] Variants identified with pathogenic impact or drug response-related according to the corresponding HGMD and/or ClinVar annotations were selected as known pathogenic variants in the Ugandan cohort (**Text S1**). Variants without any functional definition in either HGMD or ClinVar were assessed using nine *in silico* function prediction algorithms. Variants predicted as potentially function-altering by at least two tools and/or variants with a CADD phred score of >20 were defined as "potentially deleterious" variants that may feasibly interfere with drug responses in this population (**Text S1**). Any variants included in the current star allele definitions were excluded. Variants with AF >2-fold in the UG2G vs. gnomAD_ALL population were selected, and the selected variant AFs were compared between the study cohort and gnomAD_ALL using Fisher's exact test with Benjamini-Hochberg correction for multiple testing. Variant frequency differences with an adjusted *P*-value of <0.05 were considered as statistically significant.

RESULTS

The spectrum of clinically actionable star alleles

We performed a comprehensive analysis of variant star alleles in a Ugandan cohort (QC passed $n = 1989$) using 18 actionable pharmacogenes known to be associated with important clinical interventions.[16] Nearly all subjects (99.6%) were found to carry at least one clinically actionable genotype, with a median of three and an average of 3.5 genotypes (**Figure 2a**). We identified a total of 43 known actionable star alleles with potential clinical implications across 13 pharmacogenes associated with the processing of >100 drugs with CPIC guidelines.[16] Of these, 28 star alleles were common within the Ugandan cohort (AF >0.01), of which 11 were observed with an AF of >0.1 (i.e., *CYP3A5*1*, *IFNL3* [rs12979860 T allele], *CYP2B6*9*, *CYP3A5*6*, *CYP2D6*17*, *CYP2C19*17*, *CYP3A5*3*, *CYP2D6*29*, *CYP2C19*2*, *G6PD* A-.202A.376G, and *TMPT*3C*) (**Table S2**). Of the 11 star alleles, *CYP3A5*1*, *CYP3A5*6*, *CYP2B6*9*, *CYP2D6*17*, *CYP2D6*29*, and *TMPT*3C* were enriched among Africans compared with global populations, as reported by PharmGKB population frequencies. Additionally, we noticed that AFs of these alleles were notably higher in the Ugandan cohort compared with reported frequencies in Africans (i.e., SSA and AA; **Table 1**, **Table S3**).

Of the three *CYP3A5* alleles detected in the Ugandan cohort, the *CYP3A5*1* normal and **6* no-function alleles were enriched with AFs of 0.58 and 0.23, respectively, and were generally more prevalent in Africans relative to other global populations (**Table 1**). In contrast, the frequency of the *CYP3A5*3* no-function allele was lowest in this cohort (AF = 0.195) and tended to be generally lower in Africans compared with other global populations (**Table 1**). The decreased function allele *CYP2B6*9* exhibited a notably high frequency within the cohort (AF = 0.4), surpassing both those of global populations and previously reported frequencies in AA (i.e., AA: AF = 0.05; others: AF = 0.01–0.1). In addition, two *CYP2C19* alleles were among the common alleles in the Ugandan cohort (AF >0.01), including *CYP2C19*17* (an increased function allele; AF = 0.195) and *CYP2C19*2* (no-function allele; AF = 0.15). However, these frequencies were not markedly different compared with other African or global population frequencies reported in PharmGKB (**Table 1**). A total of 15 *CYP2D6* star alleles were identified in the Ugandan cohort, of which the decreased function alleles *CYP2D6*17* and *CYP2D6*29* were the most prevalent, with AFs of 0.22 and 0.18, respectively. Both alleles were highly enriched in populations of African ancestry compared with other global populations (*CYP2D6*17*: African ancestry AF ~ 0.2, others: AF <0.03; and *CYP2D6*29*: African ancestry AF = 0.09–0.12, others: AF <0.015) (**Table 1**). Similarly, the prevalence of the no-function allele *TPMT*3C* was notably higher in the cohort (AF = 0.11), while the frequency of this allele was considerably lower in other African (SSA: AF = 0.053, AA: AF = 0.024) and global populations (AF <0.02) (**Table 1**). Additionally, the *IFNL3* rs12979860 T unfavorable response allele—the second most prevalent variant allele observed (AF = 0.57)—and the *G6PD* A-.202A.376G, which belongs to the *G6PD* class-III-deficient alleles, exhibited an AF of 0.12 in the Ugandan cohort (**Table 1**, **Table S3**).

We also predicted several other prevalent star alleles (AF >0.01) with an overall high AF in populations of African ancestry compared with global populations (i.e., decreased function alleles *CYP2C9*8*, *CYP2C9*11*, *CYP2C19*9*, and *DPYD* c.557 A>G; no-function allele *CYP2B6*18*). In contrast, some alleles detected in the Ugandan cohort exhibited generally lower AFs in African populations compared with global populations (i.e.,

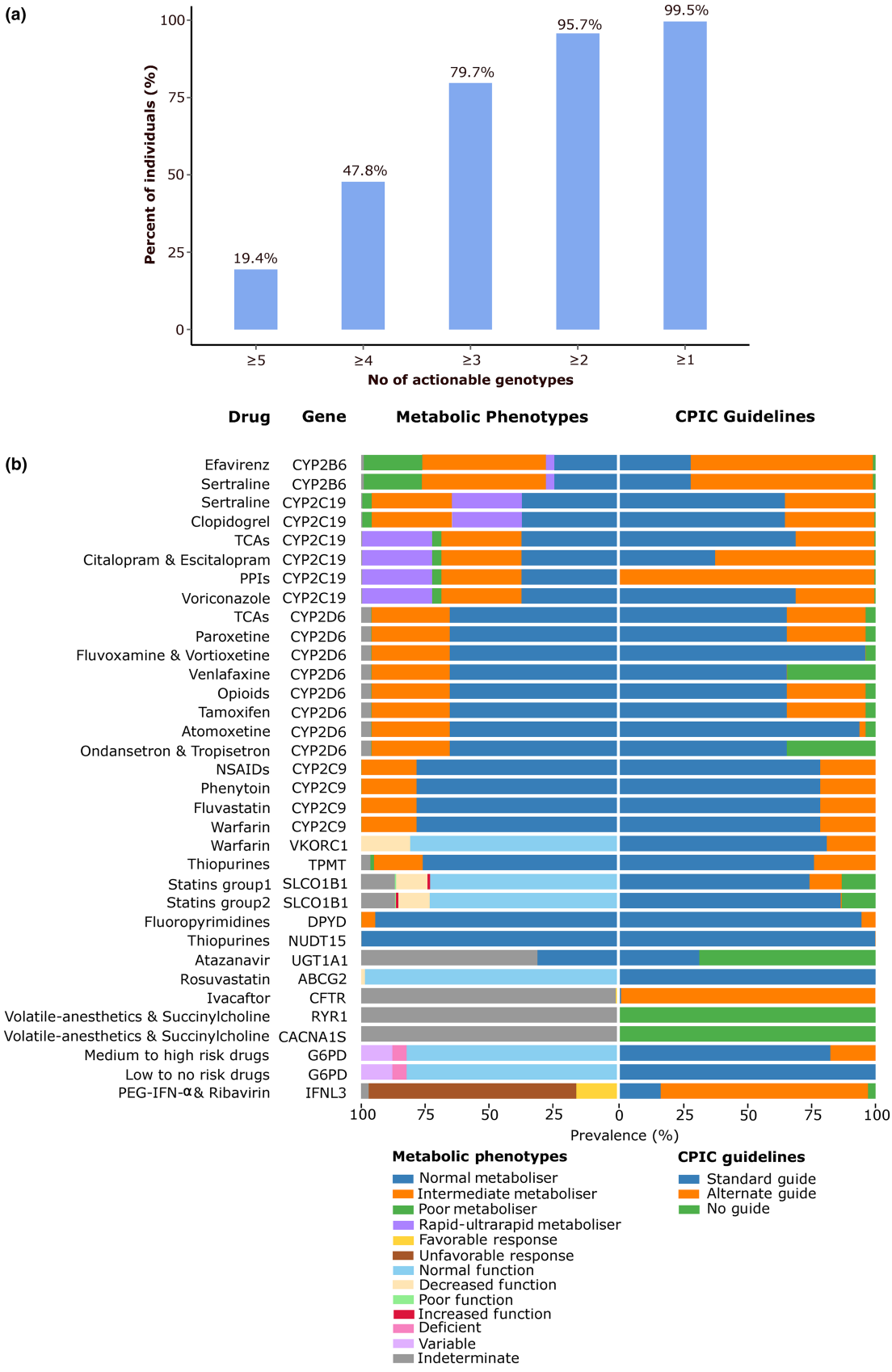


Figure 2 (Continues)

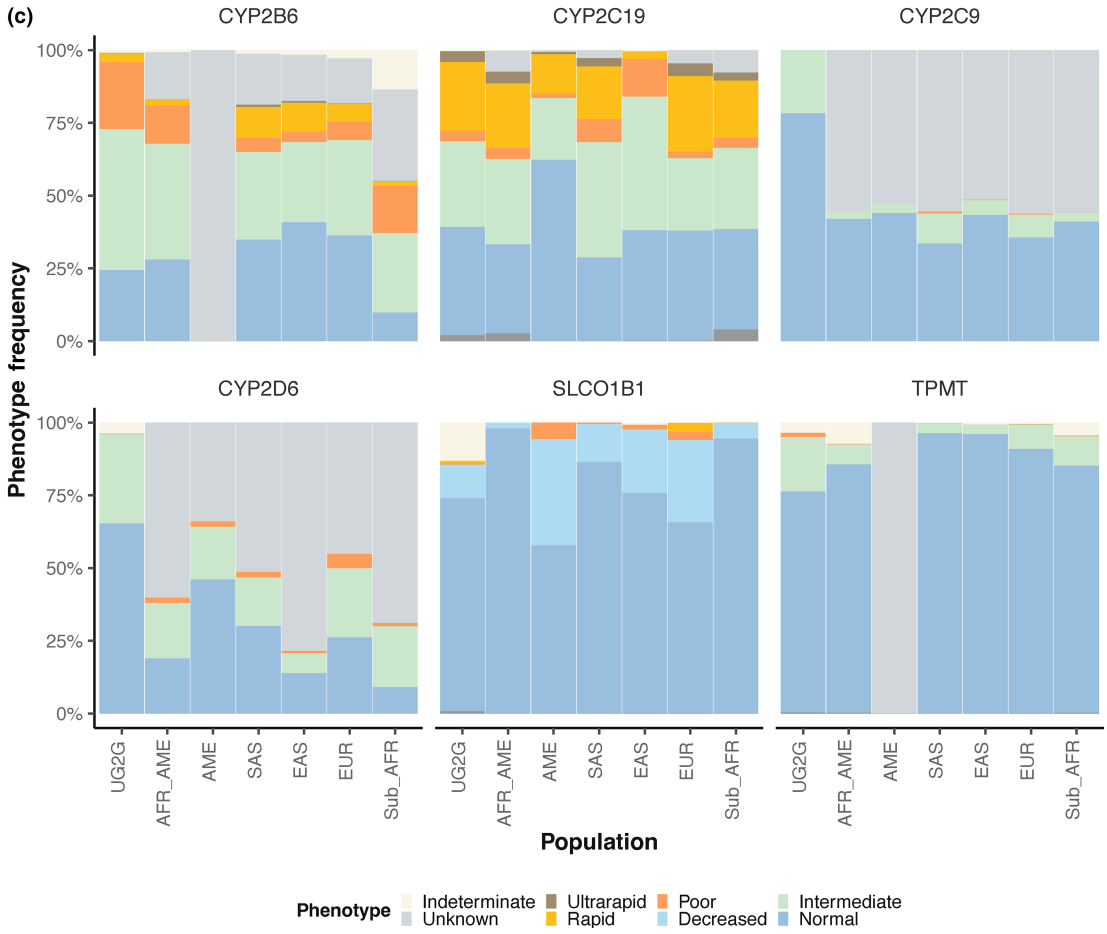


Figure 2 Actionable metabolic phenotypes in the Ugandan cohort. (**a**) Distribution of clinically actionable genotypes/phenotypes identified across 18 pharmacogenes in the Ugandan cohort. (**b**) An overview of the metabolic phenotypes identified within the Ugandan cohort and associated dosage recommendations based on the Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines (https://cpicpgx.org). The X-axis represents the prevalence of metabolic phenotypes and the associated CPIC dosage recommendation, while the Y-axis represents drug-gene combinations. *Standard Guide*: Standard drug dose is recommended without alterations; *Alternate Guide*: Recommended to avoid the standard drug dose/alternative dosage guidelines are available; *No Guide*: Dosage guidelines are not available in the CPIC. (**c**) Prevalence of metabolic phenotypes in the Ugandan cohort vs. global populations. Six out of the 18 clinically actionable genes analyzed, with allele frequency data available in the Pharmacogenomics Knowledgebase (PharmGKB) for population frequency comparisons, are depicted in the figure. AFR_AME, African American; EAS, East Asian; EUR, European; low-to-no-risk-drugs_G6PD, Primaquine single low dose (0.25 mg/kg) for *Plasmodium falciparum malaria* and Quinine; Medium-to-high-risk-drugs_G6PD, medium risk drugs – Primaquine medium dose (0.75 mg/kg or 45 mg once weekly for 8 weeks) for Plasmodium vivax malaria and Nitrofurantoin; high risk drugs – Primaquine standard dose (0.25–0.5 mg/kg daily for 14 days), Dapsone, and Methylene blue;[54] NSAIDs_CYP2C9, meloxicam, piroxicam, tenoxicam, celecoxib, flurbiprofen, ibuprofen, and lornoxicam; Opioids_CYP2D6, codeine and tramadol; PPIs_CYP2C19, omeprazole, lansoprazole, and dexlansoprazole; SAS, South Asian; Statins_group1_SLCO1B1, atorvastatin, lovastatin, pitavastatin, and simvastatin; Statins_group2_SLCO1B1, fluvastatin, pravastatin, and rosuvastatin; Sub_AFR, Sub-Saharan African; TCAs_CYP2C19, imipramine; TCAs_CYP2D6, amitriptyline, nortriptyline, clomipramine, doxepin, and imipramine; UG2G, Ugandan cohort.

decreased function alleles *CYP4F2*3* and *VKORC1* rs9923231 T; no-function alleles *SLCO1B1*15* and *CYP2D6*4*) (**Table 1**, **Table S3**).

Prevalence of actionable metabolic phenotypes and associated clinical interventions in Ugandans

CFTR, IFNL3, CYP2B6, and CYP2C19 demonstrated the highest prevalence of known actionable phenotypes in the Ugandan cohort. More than half of Ugandans in the cohort were predicted to have an impaired response to the associated specific drugs or drug classes. Also, certain metabolic phenotypes (i.e., IMs of CYP2B6, CYP2D6, CYP2C9, TPMT, DPYD, and PMs of TPMT) were more prevalent in the Ugandan cohort compared with both African and other global populations, of which TPMT IMs, TPMT PMs, and DPYD IMs were generally high among Africans compared with other populations as per PharmGKB (**Figure 2b,c**, **Tables S2, S3**).

Only eight individuals in the Ugandan cohort (0.4%) carried one of the 39 *CFTR* mutations (i.e., D1270N-*CFTR*), while 99.6% were non-carriers. The prevalence of the predicted IFNL3 actionable metabolic phenotypes in the cohort was 81.25%: C/T heterozygotes (49.27%) and T/T homozygotes (31.98%). Furthermore, the predicted prevalence of CYP2B6 actionable phenotypes was 71.4%, including 48.5% IMs and

Table 1 Common star alleles identified across 18 clinically actionable pharmacogenes within the Ugandan cohort

Gene	Allele	Core variants	Variant consequence	Allele function	Allele frequency					
					UG2G	Sub-AFR	AFR-AME	EUR	SAS	EAS
CYP2D6	*1	rs1135840 rs16947	No variant impact	Normal	0.267	0.078	0.201	0.283	0.287	0.245
	*2[a]	None	None	Normal	0.159	0.198	0.156	0.186	0.295	0.121
	*4[a]	rs3892097	Splice defect	No function	0.035	0.034	0.048	0.185	0.091	0.005
	*10[a]	rs16947 rs1065852	No variant impact Missense	Decreased	0.022	0.056	0.038	0.016	0.087	0.436
	*17[a]	rs28371706	Missense	Decreased	0.218	0.193	0.169	0.004	0.001	0.0001
	*29[a]	rs59421388 rs61736512	Missense	Decreased	0.180	0.121	0.088	0.001	0.003	0.0001
	*41[a]	rs28371725	Splice defect	Decreased	0.038	0.115	0.037	0.092	0.123	0.023
	*45	rs28371710	Missense	Normal	0.051	0.042	0.068	nr	nr	0
CYP2C9	*1	None	None	Normal	0.732	0.726	0.871	0.793	0.772	0.915
	*8[a]	rs7900194	Missense	Decreased	0.071	0.076	0.059	0.002	0.001	0.004
	*9	rs2256871	Missense	Normal	0.152	0.130	0	nr	0	0
	*11[a]	rs28371685	Missense	Decreased	0.014	0.026	0.014	0.002	0.001	0.0003
	*31	rs57505750	Missense	Decreased	0.013	nr	nr	nr	nr	0.002
CYP2C19	*1	None	None	Normal	0.566	0.552	0.547	0.625	0.544	0.596
	*2[a]	rs12769205 rs4244285	Splice defect	No function	0.155	0.157	0.181	0.147	0.270	0.284
	*9	rs17884712	Missense	Decreased	0.013	0.027	0.014	0.001	nr	0.0001
	*13	rs17879685	Missense	Normal	0.022	0	0.012	0.002	nr	0.0001
	*15	rs17882687	Missense	Normal	0.021	0.053	0.014	0.002	nr	0.001
	*17[a]	rs12248560	Expression	Increased	0.195	0.173	0.207	0.215	0.171	0.021
	*35	rs12769205	Splice defect	No function	0.02	0.032	0.016	0	nr	0
CYP2B6	*1	None	None	Normal	0.425	0.315	0.413	0.491	0.608	0.640
	*2	rs8192709	Missense	Normal	0.065	0.031	0.031	0.049	0.041	0.046
	*9	rs3745274	Splice defect	Decreased	0.402	nr	0.046	0.015	0.059	0.034
	*18	rs28399499	Missense	No function	0.07	0.058	0.033	0	nr	0
	*22	rs34223104	No variant impact	Increased	0.031	0.031	0.011	0.014	nr	nr
CYP3A5	*1	rs776746	No variant impact	Normal	0.578	0.479	0.453	0.074	0.327	0.254
	*3[a]	None	None	No function	0.195	0.241	0.316	0.924	0.673	0.746
	*6[a]	rs10264272 rs776746	Splice defect No variant impact	No function	0.228	0.193	0.111	0.002	0	0.0007
CYP4F2	*1	None	None	Normal	0.668	0.668	0.924	0.566	0.428	0.683
	*2	rs3093105	Missense	Unassigned	0.278	0.233	nr	0.160	0.169	0.084
	*3	rs2108622	Missense	Decreased	0.054	0.099	0.076	0.275	0.403	0.233
VKORC1	Reference (C)	None	None	Normal	0.901	0.892	0.899	0.587	0.844	0.133
	rs9923231(T)[a]	rs9923231	2 kb-upstream	Increased	0.099	0.108	0.101	0.413	0.156	0.866
IFNL3	Reference (C)	None	None	Favorable	0.396	nr	nr	nr	nr	nr
	rs12979860 (T)	rs12979860	Intron	Unfavorable	0.566	nr	nr	nr	nr	nr
	rs12980275 (G)	rs12980275	None	Unassigned	0.037	nr	nr	nr	nr	nr

(Continued)

Table 1 (Continued)

Gene	Allele	Core variants	Variant consequence	Allele function	Allele frequency					
					UG2G	Sub-AFR	AFR-AME	EUR	SAS	EAS
TPMT	*1	None	None	Normal	0.870	0.921	0.923	0.953	0.981	0.980
	*3C[a]	rs1142345	Missense	No function	0.110	0.053	0.024	0.005	0.011	0.016
	*8	rs56161402	Missense	Uncertain	0.015	0.024	0.006	0.0001	9.38E-05	0.00
UGT1A1	*1	None	None	Normal	0.55	0.493	0.031	0.361	0.541	0.706
	*80	rs887829	Intron	Unknown	0.45	nr	0.450	0.314	nr	nr
SLCO1B1	*1	None	None	Normal	0.132	0.171	0.230	0.403	0.470	0.255
	*14	rs2306283 rs11045819	Missense	Increased	0.02	0	nr	0.136	nr	nr
	*15	rs2306283 rs4149056	Missense	No function	0.031	0.028	0.01	0.150	0.065	0.125
	*20	rs2306283 rs34671512	Missense	Unassigned	0.071	nr	nr	0.037	nr	nr
	*27	rs2306283 rs59113707	Missense	Uncertain	0.011	nr	nr	nr	nr	nr
	*30	rs2306283 rs79135870	Missense	Uncertain	0.017	nr	nr	nr	nr	nr
	*31	rs2306283 rs59502379	Missense	Loss-of-function	0.033	nr	nr	nr	nr	nr
	*37	rs2306283	Missense	Normal	0.637	0.801	0.760	0.253	0.460	0.615
	*41	rs77271279	Splice defect	Uncertain	0.026	nr	nr	0.0001	nr	nr
	*43	rs2306283 rs11045819 rs11045852 rs74064213	Missense	Unknown	0.016	nr	nr	nr	nr	nr
DPYD	Reference	rs1801265	No variant impact	Normal	0.358	nr	nr	nr	nr	nr
	c.496A>G	rs2297595	Missense	Normal	0.026	0.031	0.032	0.110	0.085	0.016
	c.557A>G	rs115232898	Missense	Decreased	0.025	0.026	0.012	0.0001	nr	nr
	c.1218G>A	rs61622928	Missense	Normal	0.085	0.082	0.064	0.0003	0.0002	nr
	c.1371C>T	rs57918000	Synonymous	Normal	0.019	0.021	0.016	0.0002	0.0001	nr
	c.3067C>A	rs114096998	Missense	Normal	0.013	0.035	0.045	0.00001	nr	nr
	c.1349C>T	rs72975710	Missense	Normal	0.01	0.004	nr	nr	nr	nr
	*9A	None	None	Normal	0.233	0.446	0.427	0.227	0.255	0.076
	c.1627A>G (*5)	rs1801159	Missense	Normal	0.216	0.158	0.14	0.195	0.094	0.255
RYR1	Reference	None	None	Normal	1	nr	nr	nr	nr	nr
ABCG2	Reference	rs2231142 (G)	None	Normal	0.993	0.994	0.965	0.896	0.907	0.693
CFTR	Reference	None	None	Normal	0.998	nr	nr	nr	nr	nr
NUDT15	Reference	None	None	Normal	0.999	nr	nr	0.993	0.930	0.879
CACNA1S	Reference	None	None	Normal	1	nr	nr	nr	nr	nr
G6PD	*A	rs1050829	Missense	IV/Normal	0.195	nr	0.318	0.0005	0.001	0
	Reference (*B)	None	None	IV/Normal	0.688	nr	0.557	0.995	0.956	0.974
	*A-.202A.376G	rs1050829 rs1050828	Missense	III/Deficient	0.116	nr	nr	nr	nr	nr

Star alleles in 18 clinically actionable genes were identified using the PyPGx tool's GRCh37 assembly. Allele frequencies (AFs) in Uganda and global populations were compared as given in the PharmGKB gene information tables. Star alleles identified with AF>0.01 (i.e., >1%) were considered as common alleles within the cohort. Excluding CNVs/SVs, given the challenge of accurately detecting them with low-coverage data, may result in potentially inflated AFs, particularly for *CYP2D6* (*2), *CYP2B6* (*1), *CYP4F2* (*1), and *G6PD* (*B). Thus, interpretation should be cautious considering this limitation. Star allele definitions follow the GRCh37 reference version. AFR-AME, African American; CNVs, copy number variants; EUR, European; EAS, East Asian; SAS, South Asian; Sub-AFR, Sub-Saharan African; SVs, structural variants; nr, Not Reported in PharmGKB; UG2G, Ugandan cohort. [a]Association for Molecular Pathology (AMP) Tier 1 minimum sets of alleles for PGx testing (https://www.pharmgkb.org/ampAllelesToTest). AMP Tier 1 alleles not reported in **Table 1**: *CYP2D6*3*, *CYP2D6*5*, *CYP2D6*6*, *CYP2D6*9*, *CYP2C9*2*, *CYP2C9*3*, *CYP2C9*5*, *CYP2C9*6*, *CYP2C19*3*, *TPMT*2*, *TPMT*3A*, *TPMT*3B*, *NUDT15*2*, *NUDT15*3*, *CYP3A5*7* (see **Text S1** for details). Refer to **Table S3** for all the star alleles detected in this cohort.

22.9% PMs. CYP2B6 IMs have previously been reported with low prevalence among SSA (39.8%) compared with the present cohort and from 33% to 47% in global populations. Conversely, the prevalence of CYP2B6 PMs in SSA was close to that of the Ugandan cohort (23.7%) but was notably lower in other populations (AA: 15.7%, others: 4–13%), except for Oceanians (41%) (**Figure 2c**, **Table S3**). In the Ugandan cohort, 62.4% of subjects were predicted to exhibit impaired CYP2C19 metabolic activity. The prevalence of these metabolizers was comparable to that reported for African ancestry populations, while displaying substantial variation across other populations (**Figure 2c**, **Table S3**). Ugandan and European frequencies of CYP2D6 IMs

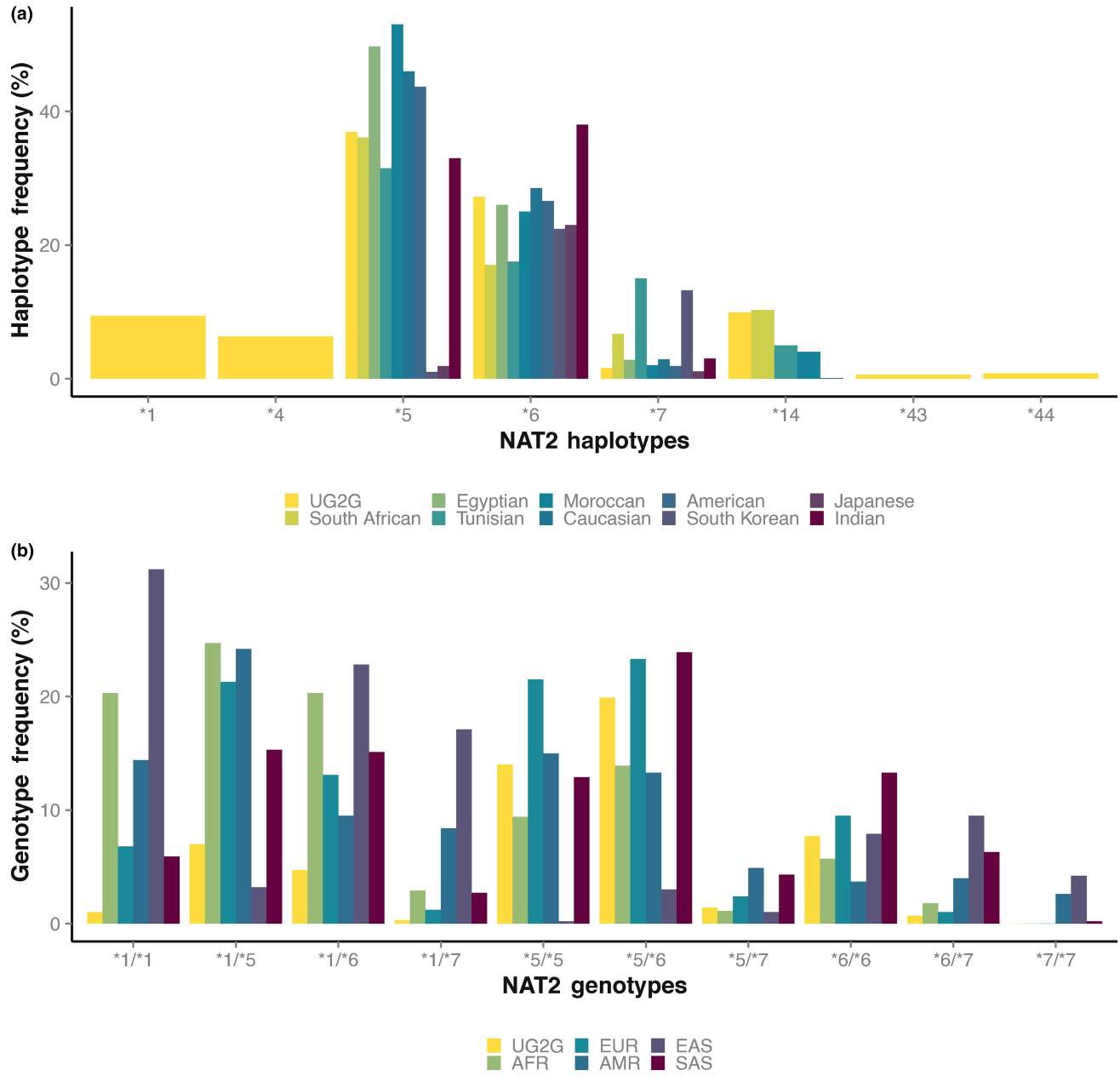


Figure 3 Distribution of *NAT2* haplotype and genotype frequencies in the Ugandan cohort vs. global populations. (**a**) *NAT2* haplotype frequencies in Ugandans vs. global populations. Haplotype frequencies for global populations were adopted from Guaoua et al.[37] These populations were chosen as representatives for AFR (South African, Egyptian, and Tunisians), EUR (Caucasians), AMR (Americans), EAS (South Koreans and Japanese), and SAS (Indians). (**b**) *NAT2* genotype frequencies in Ugandans vs. global populations. Genotype frequencies for global populations were reported by Fukunaga et al., calculated for 2,504 individuals from the 1,000 Genomes Project.[38] In **a** and **b**, frequency data for *NAT2* haplotypes and genotypes for some populations were not available in the literature for comparison. *NAT2*1* is the reference allele with no known variants (referred to as *NAT2*4* by Guaoua et al.[37] and Fukunaga et al.[38]) and the *NAT*4* allele is defined by rs1208 core variant. AFR, Africans; AMR, Americans; EAS, East Asians; EUR, Europeans; SAS, South Asians; UG2G, Ugandan cohort. **1*: Functional/fast allele (reference); **5*, **6*, **7*, **14*: Non-functional alleles; **4*, **43*, **44*: Function not defined. **1/*1*: Rapid acetylator phenotype; **1/*5*, **1/*6*, **1/*7*: Intermediate acetylator phenotype; **5/*5*, **5/*6*, **5/*7*, **6/*6*, **6/*7*, **7/*7*: Slow acetylator phenotype. **7/*7* slow acetylator phenotype was not detected in the Ugandan cohort and was not reported in Africans or Europeans in the 1,000 Genomes Project.[38]

were similar (31%); however, this prevalence was lower in other groups, including populations of African ancestry (SSA: 26.5%, AA: 23.5%, Others: 7–24%) (**Figure 2c**, **Table S3**).

Additionally, more than 10% of the Ugandan cohort was predicted to exhibit altered functionality for—CYP2C9, TPMT VKORC1, G6PD, and SLCO1B1—with important clinical implications for a wide range of drugs (**Figure 2b**). We observed a markedly higher prevalence of CYP2C9 IMs in the Ugandan cohort (21.7%) compared with the prevalence in Africans (5%) and other global populations (3% to 18%), according to PharmGKB (**Figure 2c**, **Table S3**). The predicted prevalence of TPMT IMs and PMs in the Ugandan cohort was 19.2% and 1.41%, respectively. Comparatively, the prevalence of TPMT IMs and PMs in other populations, including those of African descent, was notably lower, with frequencies below 10% and 0.3%, respectively (**Figure 2c**, **Table S3**). Within the Ugandan cohort, 5.6% of subjects were predicted to be G6PD deficient, while 12.1% were predicted to have a variable phenotype associated with a variable risk of hemolytic anemia (**Figure 2c**, **Table S3**). Additionally, 13.6% of the cohort exhibited altered SLCO1B1 metabolism, of which 11.5% was predicted to have decreased SLCO1B1 function. This prevalence is notably higher compared with the frequencies of other African populations (SSA: 5.43%, AA: 1.98%), but not of other global populations (13–36%). Although the prevalence of DPYD IMs in the cohort was 5%, this was notably high compared with <1% prevalence previously reported in all other populations, including African populations, according to PharmGKB (**Figure 2c**, **Table S3**).

Genetic variability of *NAT2* genotypes potentially influencing TB treatment with isoniazid in the Ugandan cohort

We observed *NAT2* alleles previously reported as non-functional (i.e., *NAT2*5* (rs1801280), *NAT2*6* (rs1799930), and *NAT2*7* (rs1799931)),[38,41] with AFs 0.37, 0.27, and 0.02, respectively (**Table S3**). The frequencies of *NAT2* haplotypes were variable across different African populations, as shown in **Figure 3a**. In our cohort, **5/*5*, **5/*6*, **5/*7*, **6/*6*, and **6/*7* genotypes (43.7%) were predicted to exhibit a slow acetylator phenotype to isoniazid. Additionally, **1/*5*, **1/*6*, and **1/*7* (12%) were predicted as intermediate, while **1/*1* (1%) were rapid acetylators (**Figure 3b**, **Table S3**).

Distributions of potentially actionable variants in important pharmacogenes

Across 67 VIPs, we identified 48 significantly enriched, potentially actionable variants with a pathogenic impact or drug response-related per HGMD or ClinVar, and of these, 28 were nonsynonymous variants (AF≥2-fold in UG2G and adjusted *P*-value of <0.05 vs. gnomAD_ALL) (**Figure 4**, **Table S4**). Of the 48 variants, 26 were significantly enriched in the Ugandan cohort compared with gnomAD_AFR (**Table S4**).

Using *in silico* algorithms to predict function, we identified 12 deleterious, potentially actionable variants that were significantly enriched in the Ugandan cohort compared with global populations (AF ≥2-fold in UG2G and adjusted *P*-value of <0.05 vs. gnomad_ALL) (**Figure 5**, **Table S4**). Of the 12 predicted deleterious variants, seven were significantly enriched in the cohort compared with gnomAD_AFR (**Table S4**, **Text S1**).

DISCUSSION

Our study, based on the largest WGS dataset from Africa to date, represents the most thorough PGx characterization in an African population, highlighting substantial genetic diversity among ethnolinguistic groups across Africa. Notably, six variant star alleles displayed higher frequencies in the Ugandan cohort (*CYP3A5*1*, *CYP3A5*6*, *CYP2B6*9*, *CYP2D6*17*, *CYP2D6*29*, and *TPMT*3C*) compared with global populations, including African population frequencies reported previously.[16] More than half of Ugandans in the present study predicted impaired drug responses associated with CFTR, IFNL3, CYP2B6, and CYP2C19, while 31% were predicted to have altered CYP2D6 metabolism. Actionable phenotypes such as CYP2B6, CYP2D6, CYP2C9, TPMT, DPYD IMs, TPMT PMs, and SLCO1B1 decreased metabolizers were markedly enriched within the Ugandan cohort compared with global populations, including SSA and AA. The prevalence of TPMT and DPYD actionable metabolic phenotypes was generally high among African ancestry, while SLCO1B1 decreased metabolizer prevalence was generally low among Africans.[30] Additionally, certain potentially actionable variants were significantly enriched among Ugandan VIPs, 48 of which were known pathogenic variants and 12 of which were predicted to be deleterious. Distinct frequency patterns observed in the region likely stem from gene flow from Eurasian and East African Nilo-Saharan populations during historical demographic events, contributing to varying degrees of admixture. Additionally, adaptive evolutionary events like selection or genetic drift, along with relatively low levels of linkage disequilibrium in their genetic profiles, may explain the differences in the prevalence of functional alleles and allelic diversity compared with other African populations.[17]

CFTR non-carriers (99.6%) lacked at least one of the 39 listed *CFTR* variants needed for ivacaftor therapy in cystic fibrosis diagnosis, making them unsuitable.[16] IFNL3 enzyme impairment in the Ugandan cohort suggests potential unfavorable responses to hepatitis C virus treatment.[42] Among 81.25% of unfavorable responders, nearly 50% (C/T heterozygotes) might be eligible for shorter therapy, while 32% (T/T homozygotes) predict a high risk of side effects due to prolonged drug exposure.[42] African-ancestry populations generally show lower response rates to these treatments compared with Caucasians and Asians.[42]

African populations exhibit higher frequencies of *CYP2B6*6* and *CYP2B6*18* alleles, impacting CYP2B6 function.[5] However, our cohort showed a notable prevalence of *CYP2B6*9* and **18*, lacking *CYP2B6*6* owing to the absence of the core variant rs2279343. Impaired CYP2B6 function elevates the risk of central nervous system toxicities with efavirenz, a primary antiretroviral for HIV/AIDS treatment,[43] and severe side effects with the antidepressant sertraline.[44] Dose adjustments or alternate drugs are recommended for high-risk individuals carrying actionable CYP2B6 phenotypes to avoid drug-induced complications.[43,44]

(a)

(b)

Variant_ID	Gene	Ref_Seq	Homo_alt	Het
rs143333036	*F5*	*NC_000001.10: g.169510483 T>C*	3	113
rs9332485	*F5*	*NC_000001.10: g.169555582 C>T*	8	193
rs5273	*PTGS2*	*NC_000001.10: g.186643768 A>G*	5	153
rs41313031	*SCN5A*	*NC_000003.11: g.38603947 G>A*	3	96
rs41313691	*SCN5A*	*NC_000003.11: g.38645522 G>T*	2	139
rs41313697	*SCN5A*	*NC_000003.11: g.38646357 A>C*	1	96
rs6791924	*SCN5A*	*NC_000003.11: g.38646357 A>C*	36	403
rs35761343	*NR1I2*	*NC_000003.11: g.119534626 G>A*	1	63
rs2308488	*HLA–B*	*NC_000006.11: g.31323184 C>T*	5	139
rs34447885	*SLC22A1*	*NC_000006.11: g.160543008 C>T*	3	152
rs35270274	*SLC22A1*	*NC_000006.11: g.160575907 G>T*	7	155
rs17290699	*EGFR*	*NC_000007.13:: g.55268897 A>C*	4	116
rs2229107	*ABCB1*	*NC_000007.13: g.87138659 A>T*	29	395
rs2228065	*ALOX5*	*NC_000010.10: g.45920506 G>A*	13	339
rs17861157	*CYP1A2*	*NC_000015.9: g.75043592 C>A*	4	262
rs2227945	*BRCA1*	*NC_000017.10: g.41244130 T>C*	9	206
rs56082113	*BRCA1*	*NC_000017.10: g.41245090 T>C*	5	162
rs55688530	*BRCA1*	*NC_000017.10: g.41249297 G>T*	0	61
rs4314	*ACE*	*NC_000017.10: g.61561304 C>T*	0	52
rs4318	*ACE*	*NC_000017.10: g.61562373 A>G*	123	711
rs34694816	*RYR1*	*NC_000019.9: g.38964275 A>G*	91	657
rs112105381	*RYR1*	*NC_000019.9: g.38964306 C>G*	0	59
rs143398211	*RYR1*	*NC_000019.9: g.38985195 G>A*	1	74
rs35180584	*RYR1*	*NC_000019.9: g.38995998 C>G*	2	123
rs148772854	*RYR1*	*NC_000019.9: g.39034444 C>T*	0	57
rs73933023	*RYR1*	*NC_000019.9: g.39057615 C>T*	0	149
rs28399454	*CYP2A6*	*NC_000019.9: g.41351267 C>T*	21	280
rs112337232	*CYP2A13*	*NC_000019.9: g.41596082 C>G*	18	271

Figure 4 Potentially actionable and known pathogenic variants in very important pharmacogenes (VIPs) within the Ugandan cohort. (**a**) The bubble plot shows 28 potentially actionable nonsynonymous variants (AF ≥2-fold in UG2G vs. gnomAD_ALL) with significantly different allele frequency distributions within the Ugandan cohort compared with the gnomAD_ALL population (adjusted *P*-value of <0.05). These variants have been identified as having pathogenic impact or being drug response-related per HGMD and/or ClinVar annotations. (**b**) The table shows the number of individuals carrying one of the three genotypes for the identified 28 nonsynonymous variants. HGMD, Human Gene Mutation Database; Hom_alt, homozygous alternate (1/1); Het, heterozygous (0/1); homozygous reference (0/0)=[total 1989 – (Hom_alt + Het)]; Ref_Seq, Reference Sequence; UG2G, Ugandan cohort; gnomAD_ALL, gnomAD all populations; gnomAD_AFR, gnomAD African; gnomAD_AMR, gnomAD American; gnomAD_EAS, gnomAD East Asian; 1KGP3_ALL, 1,000 Genomes all populations; 1KGP3_AFR, 1,000 Genomes African; 1KGP3_AMR, 1,000 Genomes admixed American; 1KGP3_EUR, 1,000 Genomes European; 1KGP3_EAS, 1,000 Genomes East Asian; 1KGP3_SAS, 1,000 Genomes South Asian.

The complications arising from medication exposure when treating CVD, cancer, and diabetes highlight the pressing need to understand drug-gene interactions in African populations.[2,4,45] In the present study, 24% of Ugandans may face issues with TPMT metabolism, impacting anti-cancer treatments and potentially leading to thiopurine-induced adverse effects.[16] TPMT deficiency is the primary genetic cause of thiopurine intolerance among Europeans and Africans.[16] Additionally, 5.5% of individuals classified as impaired DPYD metabolizers might face risks related to fluoropyrimidine toxicities.[16] We also observed a potentially actionable variant rs11572103 (NC_000010.10:g.96818106T>A) in *CYP2C8*, which could impact the metabolism of the chemotherapy drug paclitaxel.[39]

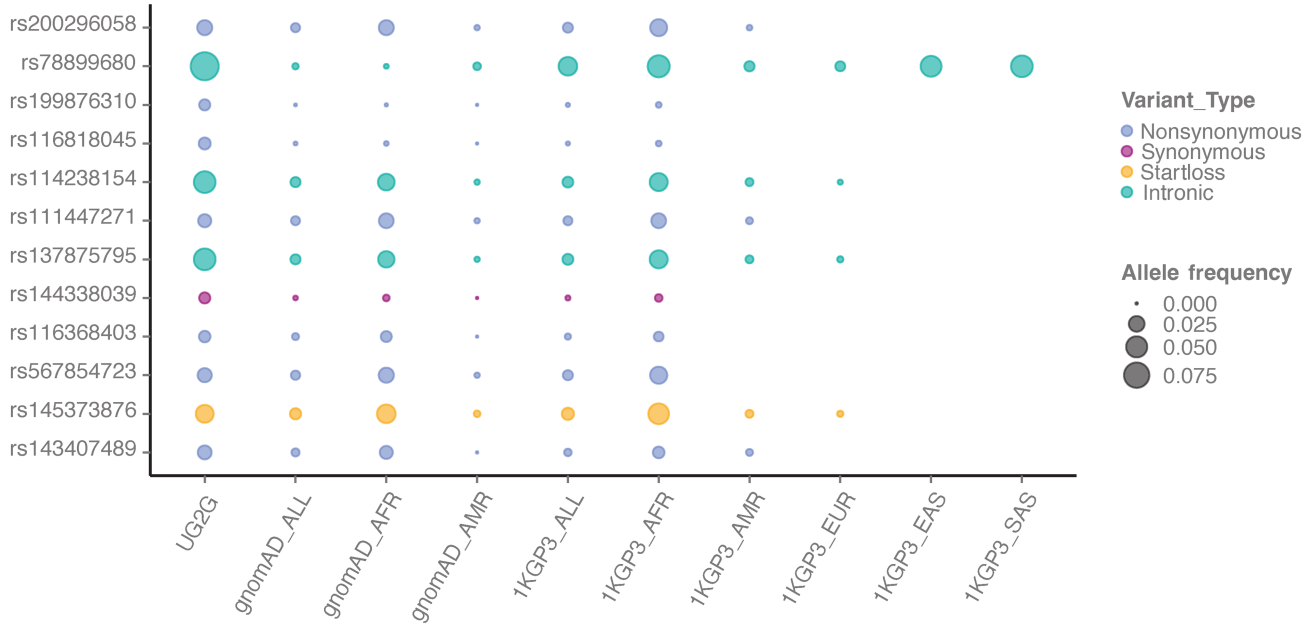
A high prevalence of *CYP2D6*3*, **17*, and **29* genotypes has previously been reported in African populations.[46] In the present study, *CYP2D6*17* and *CYP2D6*29* were highly prevalent (AF ~0.2), while the no-function *CYP2D6*4* allele and the decreased function *CYP2D6*10* and **41* alleles exhibited frequencies <0.04. Impaired CYP2D6 function affects the metabolism of a quarter of commonly used medications.[46] 31% of CYP2D6 IMs and PMs in the cohort have important clinical implications for the use of tamoxifen in breast cancer treatment and the use of 5HT3 receptor antagonists in managing side effects of chemotherapy or radiation.[16,47] However, alleles characterized by copy number or structural variants (CNVs/SVs) could potentially be missed within this cohort owing to the difficulties inherent in accurately identifying them with low-coverage data. For instance, the *CYP2D6*5* no-function allele might be interpreted as the *CYP2D6*2* reference allele with normal function. The presence of alleles having CNVs/SVs could elevate actionable phenotypes, consequently increasing the predicted clinical significance for the cohort. Furthermore, impaired CYP2D6 function affects the metabolism of commonly used antidepressants, antipsychotics, analgesics, antiarrhythmics, or β-blockers used for cardiovascular-related complications. These individuals face a high risk of treatment failure and ADRs, requiring alternative therapies.[16]

Our results show important implications for the use of anticoagulants, such as warfarin (associated with *CYP2C9* and *VKORC1* variants),[48] antiplatelets such as clopidogrel (associated with *CYP2C19* variants),[49] and statins used in the treatment of CVD (associated with *SLCO1B1* variants).[50] Moreover, *CYP2C9* is responsible for metabolizing several commonly used drugs (i.e., nonsteroidal anti-inflammatory drugs (NSAIDs): ibuprofen and celecoxib; antiepileptics: phenytoin; and statins: Fluvastatin).[16] Reduced enzymatic activity in CYP2C9 IMs and PMs leads to a higher risk of severe side effects, such as serious gastrointestinal, renal, and cardiovascular complications upon NSAID use[51] or bleeding upon treatment with warfarin.[48] Warfarin metabolism is primarily affected by the presence of the decreased function alleles *CYP2C9*2*, **3*, and the *VKORC1* [rs9923231] T allele, which lead to reduced drug metabolism. The decreased function alleles *CYP2C9*5*, **6*, **8*, and **11* are also crucial in determining warfarin dosage, particularly for individuals of African descent.[48] In the Ugandan cohort, except for *CYP2C9*6*, the other alleles were detected, with **8* being the most prevalent. Nearly 19% of the cohort carried at least one *VKORC1* [rs9923231] T allele.

In addition to clopidogrel, *CYP2C19* plays a crucial role in the response to various other drugs (i.e., tricyclic antidepressants: amitriptyline and imipramine; selective serotonin reuptake inhibitors: citalopram and sertraline; protein-pump inhibitors: omeprazole; and antifungals: voriconazole).[16] *CYP2C19*35*, with AF = 0.02 in the Ugandan cohort, has previously been reported only in SSA (AF = 0.03) and in AA (AF = 0.02). Altered enzyme activity predicted in >50% of the Ugandan cohort has important implications for CYP2C19 IMs and PMs, potentially leading to increased drug exposure and a higher risk of ADRs. In contrast, CYP2C19 RMs and UMs experience lower plasma drug concentrations, potentially contributing to a greater risk of treatment failure.[16] There are no guidelines currently available for gene-drug associations for diabetes treatment. However, we report here the potentially actionable nonsynonymous variant rs34447885 (NC_000006.11:g.160543008C>T) in *SLC22A1*, which is associated with reduced metformin uptake.[39]

The first-line treatment for TB—a combination of isoniazid, rifampicin, pyrazinamide, and ethambutol—has experienced complications in some patients.[41] The FDA drug label for isoniazid highlights the risk of drug-induced hepatotoxicity owing to slow acetylation, increasing drug concentrations in the blood.[41] While no established guidelines exist for TB drugs, utilizing *NAT2* genotyping to regulate isoniazid doses has shown promising results in optimizing treatment and reducing ADRs.[52] African populations have previously demonstrated substantial diversity within the *NAT2* gene, while other global populations have shown substantially lower genetic variation.[41,53] In our cohort, 43.7% and 12% of individuals were predicted to be slow and intermediate acetylators, respectively, while 1% were predicted to be rapid acetylators, requiring adjusted isoniazid doses to achieve therapeutic efficacy, reduce ADRs, and prevent drug resistance.[52] The cohort's *NAT2* actionability could rise substantially if all *NAT2* alleles were assigned functionality. With limited TB drug options, individualized treatment based on genetic variations is critical to optimize outcomes and curb drug-resistant TB strains.

(a)



(b)

Variant_ID	Gene	Ref_Seq	Homo_alt	Het
rs200296058	*VDR*	*NC_000012.11: g.48249448 C>A*	0	43
rs78899680	*SLC22A1*	*NC_000006.11: g.160577063 G>T*	1	89
rs199876310	*PTGIS*	*NC_000020.10: g.48129793 C>T*	0	52
rs116818045	*P2RY12*	*NC_000003.11: g.151056254 G>A*	0	47
rs114238154	*NQO1*	*NC_000016.9: g.69760341 A>G*	3	129
rs111447271	*ENOSF1*	*NC_000018.9: g.674325 G>A*	0	75
rs137875795	*CYP2E1*	*NC_000010.10: g.135341013 G>A*	0	40
rs144338039	*CYP2A13*	*NC_000019.9: g.41596396 A>T*	3	57
rs116368403	*CYP2A13*	*NC_000019.9: g.41600254 C>A*	1	75
rs567854723	*COL18A1*	*NC_000021.8: g.46918258 A>G*	20	335
rs145373876	*COL18A1*	*NC_000021.8: g.46918241 C>T*	6	200
rs143407489	*COL18A1*	*NC_000021.8: g.46918295 C>T*	6	201

Figure 5 Potentially actionable and predicted deleterious variants in very important pharmacogenes (VIPs) within the Ugandan cohort. (**a**) The bubble plot shows potentially actionable 12 deleterious variants (AF ≥2-fold in UG2G vs. gnomAD_ALL) with significantly different allele frequencies in the Ugandan cohort compared with the gnomAD_ALL population (adjusted *P*-value of <0.05). The variants were identified as deleterious if they were predicted as potentially deleterious by at least two *in silico* functionality prediction algorithms (SIFT "D", PolyPhen2 "D" or "P", LRT "D", MutationTaster "D" or "A", MutationAssessor "H" or "M", FATHMM "D", MetaSVM "D", MetaLR "D", or CADD score > 20). (**b**) The table shows the number of individuals carrying one of the three genotypes identified for the 12 deleterious variants. Hom_alt, homozygous alternate (1/1); Het, heterozygous (0/1), homozygous reference (0/0) = [total 1989 – (Hom_alt + Het)]; Ref_Seq, Reference Sequence; CADD score, combined annotation-dependent depletion score; UG2G, Ugandan cohort; gnomAD_ALL, gnomAD all populations; gnomAD_AFR, gnomAD African; gnomAD_AMR, gnomAD American; gnomAD_EAS, gnomAD East Asian; 1KGP3_ALL, 1,000 Genomes all populations; 1KGP3_AFR, 1,000 Genomes African; 1KGP3_AMR, 1,000 Genomes admixed American; 1KGP3_EUR, 1,000 Genomes European; 1KGP3_EAS, 1,000 Genomes East Asian; 1KGP3_SAS, 1,000 Genomes South Asian.

Although G6PD deficiency is prevalent in equatorial African populations and offers a degree of protection against severe malaria, it also poses a significant risk for ADRs such as acute hemolytic anemia when patients are treated with the antimalarial drug primaquine.[54] The A-.202A.376G allele is associated with G6PD deficiency and is commonly reported in African populations.[54] In our cohort, the A-.202A.376G allele was the most prevalent *G6PD*-deficient allele, resulting in 5.6% of the cohort being predicted as

G6PD-deficient. CPIC guidelines recommend avoiding standard primaquine doses in *G6PD*-deficient patients.[54] Similarly, drugs in the high-risk category such as rasburicase—used for prophylaxis and to treat hyperuricemia in patients undergoing chemotherapeutic treatment for malignancies—should be avoided in G6PD-deficient patients.[54] Given the high prevalence of malaria and G6PD deficiency in the region, genotyping for *G6PD* and enzyme activity testing in variable respondents (predicted to be 12.1% of the cohort) is necessary to understand drug response variability and optimize treatment while preventing resistance from subtherapeutic drug exposure. Similarly, the *CYP2C8* gene is known for its involvement in the metabolism of antimalarials, anti-cancer drugs, and anti-TB drugs (see Rajman et al. for review),[46] although there is a lack of sufficient evidence pertaining to this gene to establish prescription guidelines. *CYP2C8*2* decreased function allele reported variable frequencies, ranging from 10–22% across African populations, including Uganda, while being notably less common (0–1.6%) in Caucasians and Asians.[46] In the present study, the prevalence of *CYP2C8*2* was 18.9%, with 34.2% of the cohort predicted to carry at least one *CYP2C8*2* allele, potentially leading to reduced clearance of the antimalarial drug amodiaquine.[55]

There are some limitations to this study. CNVs/SVs present in some pharmacogenes (i.e., *CYP2D6, CYP2B6*) were excluded from our analysis owing to difficulties in accurately identifying these variants using low-pass WGS data. Likewise, our results do not include star alleles associated with these variants, potentially changing the predicted metabolic status of those genes in some individuals **Text S1**.

In summary, our analysis emphasizes the necessity of acknowledging the diversity within African populations, particularly in the context of PGx. We stress that a singular African population cannot adequately represent the entire PGx landscape of a diverse country or continent. Therefore, there is a crucial need for population-specific PGx profiling across Africa. Our research, conducted in a rural Ugandan community, identifies clinically significant variants, urging their inclusion in preemptive PGx testing to enhance medication efficacy and safety. We recommend future PGx studies utilizing advanced sequencing in various Ugandan ethnic groups to uncover common and population-specific rare variants. This approach can inform the development of cost-effective, tailored PGx panels, facilitating genotype-guided treatments and advancing precision public health strategies for the broader population.

SUPPORTING INFORMATION

Supplementary information accompanies this paper on the *Clinical Pharmacology & Therapeutics* website (www.cpt-journal.com).

ACKNOWLEDGMENTS

This study makes use of the Uganda Genome Resource (UGR) which constitutes whole genome sequencing data for 2000 individuals and genotype arrya data for 5000 individuals from the Uganda General Population Cohort (GPC) - a population-based open cohort study established in 1989 by the Medical Research Council (MRC), UK in collaboration with the Uganda Virus Research Institute (UVRI) and LSHTM (Uganda Research Unit).[17] We thank all participants who contributed to the UGR. The Uganda GPC was supported by the MRC, UK and the UK Department for International Development (DFID) under the MRC/DFID Concordant agreement, through core funding to the MRC/UVRI and LSHTM Uganda Research Unit. S.F. was supported by the Wellcome Trust grant number 220740/Z/20/Z to Segun Fatumo. S.R.S PhD is funded by the QUT Postgraduate Research Award (QUTPRA) (International) Scholarship from the Queensland University of Technology. The authors would like to acknowledge the following people: Sudhir Jadhao and Stacie O'Brien for providing custom-written scripts for downloading the initial BAM/CRAM files for the Ugandan cohort from EGA; Anoop Joseph for assisting with the statistical comparison to validate star allele frequencies between the genotype array and WGS data; and Simon Lee for assisting with gtf_to_vcf tool installation. Open access publishing facilitated by Queensland University of Technology, as part of the Wiley - Queensland University of Technology agreement via the Council of Australian University Librarians.

FUNDING

No funding was received for this work.

CONFLICT OF INTEREST

The authors declared no competing interests for this work.

AUTHOR CONTRIBUTIONS

S.R.S., S.H.N., M.C., S.F., and H-J.G. wrote the manuscript. S.H.N. and S.R.S. designed the research. S.R.S. performed the research. S.R.S and S.L. analyzed the data.

1. Pereira, L., Mutesa, L., Tindana, P. & Ramsay, M. African genetic diversity and adaptation inform a precision medicine agenda. *Nat. Rev. Genet.* **22**, 284–306 (2021).
2. Niohuru, I. Disease Burden and Mortality. In *Healthcare and Disease Burden in Africa* 35–85 (Springer, Cham, Switzerland, 2023).
3. Gouda, H.N. *et al*. Burden of non-communicable diseases in sub-Saharan Africa, 1990-2017: results from the global burden of disease study 2017. *Lancet Glob. Health* **7**, e1375–e1387 (2019).
4. Mpye, K.L., Matimba, A., Dzobo, K., Chirikure, S., Wonkam, A. & Dandara, C. Disease burden and the role of pharmacogenomics in African populations. *Global Health Epidemiol Genomics* **2**, e1 (2017).
5. Sitabule, B.R., Othman, H., Choudhury, A., Twesigomwe, D. & Hanchard, N.A. Promoting pharmacogenomics in Africa: perspectives from variation in G6PD and other Pharmacogenes. *Clin. Pharmacol. Ther.* **113**, 476–479 (2023).
6. Gurdasani, D. *et al*. The African genome variation project shapes medical genetics in Africa. *Nature* **517**, 327–332 (2015).
7. Swen JJ, et al. A 12-gene pharmacogenetic panel to prevent adverse drug reactions: an open-label, multicentre, controlled, cluster-randomised crossover implementation study. *Lancet*. 2023;**401**:347–356.
8. Krebs K, Milani L. Translating pharmacogenomics into clinical decisions: do not let the perfect be the enemy of the good. *Hum. Genomics* 2019;**13**, 39.
9. Relling, M.V. & Evans, W.E. Pharmacogenomics in the clinic. *Nature* **526**, 343–350 (2015).
10. Magavern, E.F., Gurdasani, D., Ng, F.L. & Lee, S.S. Health equality, race and pharmacogenomics. *Br. J. Clin. Pharmacol.* **88**, 27–33 (2022).
11. Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature* **538**, 161–164 (2016). https://doi.org/10.1038/538161a

12. Eastwood, S.V., Mathur, R., Sattar, N., Smeeth, L., Bhaskaran, K. & Chaturvedi, N. Ethnic differences in guideline-indicated statin initiation for people with type 2 diabetes in UK primary care, 2006–2019: a cohort study. *PLoS Med.* **18**, e1003672 (2021).
13. Samarasinghe, S.R., Hoy, W., McMorran, B.J., Guchelaar, H.-J. & Nagaraj, S.H. The pharmacogenomic landscape of an indigenous Australian population. *Front. Pharmacol.* **14**, 1180640 (2023). https://doi.org/10.3389/fphar.2023.1180640.
14. Radouani, F. *et al*. A review of clinical pharmacogenetics studies in African populations. *Perinat. Med.* **17**, 155–170 (2020).
15. Soko ND, Muyambo S, Dandara MTL, Kampira E, Blom D, Jones ESW, Rayner B, Shamley D, Sinxadi P, Dandara C Towards evidence-based implementation of pharmacogenomics in southern Africa: comorbidities and polypharmacy profiles across diseases. *J Personalized Med.* **13**(8), 1185 (2023). https://doi.org/10.3390/jpm13081185
16. Relling, M.V. & Klein, T.E. CPIC: clinical pharmacogenetics implementation consortium of the pharmacogenomics research network. *Clin. Pharmacol. Ther.* **89**, 464–467 (2011).
17. Gurdasani D, Carstensen T, Fatumo S, Chen G, Franklin CS, Prado-Martinez J, Bouman H, Abascal F, Haber M, Tachmazidou I, Mathieson I, Ekoru K, DeGorter MK, Nsubuga RN, Finan C, Wheeler E, Chen L, Cooper DN, Schiffels S, Chen Y, Ritchie GRS, Pollard MO, Fortune MD, Mentzer AJ, Garrison E, Bergström A, Hatzikotoulas K, Adeyemo A, Doumatey A, Elding H, Wain LV, Ehret G, Auer PL, Kooperberg CL, Reiner AP, Franceschini N, Maher D, Montgomery SB, Kadie C, Widmer C, Xue Y, Seeley J, Asiki G, Kamali A, Young EH, Pomilla C, Soranzo N, Zeggini E, Pirie F, Morris AP, Heckerman D, Tyler-Smith C, Motala AA, Rotimi C, Kaleebu P, Barroso I, Sandhu MS Uganda genome resource enables insights into population history and genomic discovery in Africa. *Cell* 2019;**179**:984–1002.e36.
18. World Health Organization. *Country Disease Outlook Uganda* 139–140 (WHO Regional Office for Africa, Republic of Congo, 2023) <https://www.afro.who.int/sites/default/files/2023-08/Uganda.pdf>.
19. Nasasira, M., Kalyango, J.N., Mupere, E. & Baluku, J.B. Incidence and predictors of adverse drug events among people receiving drug resistant tuberculosis treatment in Uganda: 8-year retrospective cohort study. *Ther. Clin. Risk Manag.* **18**, 1117–1127 (2022).
20. Dwyer-Lindgren, L. *et al*. Mapping HIV prevalence in sub-Saharan Africa between 2000 and 2017. *Nature* **570**, 189–193 (2019).
21. Baluku, J.B. *et al*. Treatment outcomes of drug resistant tuberculosis patients with multiple poor prognostic indicators in Uganda: a countrywide 5-year retrospective study. *J Clin Tuber Oth Myc* **23**, 100221 (2021).
22. Aturinde A, Farnaghi M, Pilesjö P, Mansourian A. Spatial analysis of HIV-TB co-clustering in Uganda. *BMC Infect. Dis.* 2019;**19**:612.
23. Orishaba, P. *et al*. Increased malaria parasitaemia among adults living with HIV who have discontinued cotrimoxazole prophylaxis in Kitgum district, Uganda. *PLoS One* **15**(11), e0240838 (2020).
24. Ndagije, H.B. *et al*. The burden of adverse drug reactions due to artemisinin-based antimalarial treatment in selected Ugandan health facilities: an active follow-up study. *Drug Saf.* **41**, 753–765 (2018).
25. Kiguba, R., Karamagi, C. & Bird, S.M. Antibiotic-associated suspected adverse drug reactions among hospitalized patients in Uganda: a prospective cohort study. *Pharmacol. Res. Perspect* **5**, e00298 (2017).
26. Micheni, L.N., Kassaza, K., Kinyi, H., Ntulume, I. & Bazira, J. Rifampicin and isoniazid drug resistance among patients diagnosed with pulmonary tuberculosis in southwestern Uganda. *PLoS One*, **16**(10), e0259221 (2021). https://doi.org/10.1371/journal.pone.0259221.
27. Bienczak, A. *et al*. The impact of genetic polymorphisms on the pharmacokinetics of efavirenz in African children. *Br. J. Clin. Pharmacol.* **82**, 185–198 (2016).
28. Calcagno, A. *et al*. The influence of pharmacogenetic variants in HIV/tuberculosis coinfected patients in Uganda in the SOUTH study. *Clin. Pharmacol. Ther.* **106**, 450–457 (2019).
29. Asiimwe, I.G. *et al*. A genome-wide association study of plasma concentrations of warfarin enantiomers and metabolites in sub-Saharan black-African patients. *Front. Pharmacol.* **13**, 967082 (2022).
30. Whirl-Carrillo, M. *et al*. An evidence-based framework for evaluating pharmacogenomics knowledge for personalized medicine. *Clin. Pharmacol. Ther.* **110**, 563–572 (2021).
31. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H Twelve years of SAMtools and BCFtools. *Gigascience*. **10**(2), giab008 (2021). https://doi.org/10.1093/gigascience/giab008
32. McKenna, A. *et al*. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
33. Rubinacci S, Ribeiro DM, Hofmeister RJ, Delaneau O. Efficient phasing and imputation of low-coverage sequencing data using large reference panels. *Nat. Genet.* 2021;**53**, 120–126.
34. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M. & Lee, J.J. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
35. Lee, S.B., Shin, J.Y., Kwon, N.J., Kim, C. & Seo, J.S. ClinPharmSeq: a targeted sequencing panel for clinical pharmacogenetics implementation. *PLoS One* **17**, e0272129 (2022).
36. Lee SB, Wheeler MM, Patterson K, McGee S, Dalton R, Woodahl EL, Gaedigk A, Thummel KE, Nickerson DA. Stargazer: a software tool for calling star alleles from next-generation sequencing data using CYP2D6 as a model. *Genet Med.* **21**(2), 361–372 (2019). https://doi.org/10.1038/s41436-018-0054-0
37. Guaoua, S. *et al*. Distribution of allelic and genotypic frequencies of NAT2 and CYP2E1 variants in Moroccan population. *BMC Genet.* **15**, 156 (2014).
38. Fukunaga, K. *et al*. Functional characterization of the effects of N-acetyltransferase 2 alleles on N-acetylation of eight drugs and worldwide distribution of substrate-specific diversity. *Front. Genet.* **12**, 652704 (2021).
39. Stenson, P.D. *et al*. The human gene mutation database (HGMD((R))): optimizing its use in a clinical diagnostic or research setting. *Hum. Genet.* **139**, 1197–1207 (2020).
40. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
41. Oelofse, C., Sima, C.A.A.N., Moller, M. & Uren, C. Pharmacogenetics as part of recommended precision medicine for tuberculosis treatment in African populations: could it be a reality? *Clin. Transl. Sci.* **16**, 1101–1112 (2023).
42. Muir, A.J. *et al*. Clinical pharmacogenetics implementation consortium (CPIC) guidelines for IFNL3 (IL28B) genotype and PEG interferon-alpha-based regimens. *Clin. Pharmacol. Ther.* **95**, 141–146 (2014).
43. Desta, Z. *et al*. Clinical pharmacogenetics implementation consortium (CPIC) guideline for CYP2B6 and efavirenz-containing antiretroviral therapy. *Clin. Pharmacol. Ther.* **106**, 726–733 (2019).
44. Bousman, C.A. *et al*. Clinical pharmacogenetics implementation consortium (CPIC) guideline for CYP2D6, CYP2C19, CYP2B6, SLC6A4, and HTR2A genotypes and serotonin reuptake inhibitor antidepressants. *Clin. Pharmacol. Ther.* **114**, 51–68 (2023).
45. Damanhouri, Z.A., Alkreathy, H.M., Alharbi, F.A., Abualhamail, H. & Ahmad, M.S. A review of the impact of pharmacogenetics and metabolomics on the efficacy of metformin in type 2 diabetes. *Int. J. Med. Sci.* **20**, 142–150 (2023).
46. Rajman, I., Knapp, L., Morgan, T. & Masimirembwa, C. African genetic diversity: implications for cytochrome P450-mediated drug metabolism and drug development. *EBioMedicine* **17**, 67–74 (2017).
47. Twesigomwe, D. *et al*. Characterization of CYP2D6 pharmacogenetic variation in sub-Saharan African populations. *Clin. Pharmacol. Ther.* **113**, 643–659 (2022).
48. Johnson, J.A. *et al*. Clinical pharmacogenetics implementation consortium (CPIC) guideline for pharmacogenetics-guided warfarin dosing: 2017 update. *Clin. Pharmacol. Ther.* **102**, 397–404 (2017).
49. Lee CR, Luzum JA, Sangkuhl K, Gammal RS, Sabatine MS, Stein CM, Kisor DF, Limdi NA, Lee YM, Scott SA, Hulot JS, Roden DM, Gaedigk A, Caudle KE, Klein TE, Johnson JA, Shuldiner AR.

Clinical pharmacogenetics implementation consortium guideline for CYP2C19 genotype and clopidogrel therapy: 2022 update. *Clin. Pharmacol. Ther.* 2022;**112**:959–967.
50. Cooper-DeHoff, R.M. *et al*. The clinical pharmacogenetics implementation consortium guideline for SLCO1B1, ABCG2, and CYP2C9 genotypes and statin-associated musculoskeletal symptoms. *Clin. Pharmacol. Ther.* **111**, 1007–1021 (2022).
51. Theken, K.N. *et al*. Clinical pharmacogenetics implementation consortium guideline (CPIC) for CYP2C9 and nonsteroidal anti-inflammatory drugs. *Clin. Pharmacol. Ther.* **108**, 191–200 (2020).
52. Masiphephethu, M.V. *et al*. Pharmacogenetic testing for NAT2 genotypes in a Tanzanian population across the lifespan to guide future personalized isoniazid dosing. *Tuberculosis (Edinb.)* **136**, 102246 (2022).
53. Sabbagh, A., Langaney, A., Darlu, P., Gerard, N., Krishnamoorthy, R. & Poloni, E.S. Worldwide distribution of NAT2 diversity: implications for NAT2 evolutionary history. *BMC Genet.* **9**, 21 (2008).
54. Gammal, R.S. *et al*. Expanded clinical pharmacogenetics implementation consortium guideline for medication use in the context of G6PD genotype. *Clin. Pharmacol. Ther.* **113**, 973–985 (2023).
55. Parikh, S., Ouedraogo, J.B., Goldstein, J.A., Rosenthal, P.J. & Kroetz, D.L. Amodiaquine metabolism is impaired by common polymorphisms inCYP2C8: implications for malaria treatment in Africa. *Clin. Pharmacol. Therap.* **82**, 197–203 (2007).