

Population-scale catalogue of clinically actionable pharmacogenomic variation in a large Middle Eastern Levantine cohort

Nagham Nafiz Hendi

Applied Science Private University (ASU)

Laith M. Alawneh

Aberdeen Royal Infirmary

Noureddin Ismail

Oxford Brookes University

Rania Abdellatif

Minia University

Nathir M. Obeidat

University of Jordan

Basel Abu-Jamous

Pan.bio, Inc.

Malek Zihlif

University of Jordan

Nancy Hakooz

University of Jordan

Said I. Ismail

sismail@hbku.edu.qa

Hamad bin Khalifa University

Article

Keywords: pharmacogenomics, clinically actionable variation, precision prescribing, Levantines, Middle Eastern genomics

Posted Date: October 9th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11089585/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Population-scale catalogue of clinically actionable pharmacogenomic variation in a large Middle Eastern Levantine cohort

Nagham Nafiz Hendi¹, Laith M. Alawneh², Nouredin Ismail³, Rania Abdellatif^{4,5}, Nathir M. Obeidat⁶, Basel Abu-Jamous⁷, Malek Zihlif⁸, Nancy Hakooz^{9,†}, Said I. Ismail^{10,*}

¹ Department of Clinical Pharmacy & Therapeutics, Faculty of Pharmacy, Applied Science Private University (ASU), Amman, Jordan; n_hendi@asu.edu.jo; ORCID: 0000-0003-2737-0565

² Department of Trauma and Orthopaedics, Aberdeen Royal Infirmary, Aberdeen, United Kingdom; laith.alawneh@nhs.scot

³ School of Biological and Medical Sciences, Faculty of Health, Science, and Technology (HST), Oxford Brookes University, Oxford, United Kingdom; 19250829@brookes.ac.uk

⁴ Qatar Genome Program (QGP), Qatar Precision Health Institute (QPHI), Qatar Foundation, Doha, Qatar; rabdellatif@qf.org.qa

⁵ Department of Pharmacology and Toxicology, Faculty of Pharmacy, Minia University, Minia, Egypt; rania.abdellatief@mu.edu.eg

⁶ Department of Internal Medicine, School of Medicine, University of Jordan, Amman, Jordan; nobeidat@ju.edu.jo

⁷ Pan.bio, Inc., Boston, United State; basel@pan.bio

⁸ Department of Biological Sciences, Faculty of Science, University of Jordan, Amman, Jordan; m.zihlif@ju.edu.jo

⁹ Department of Biopharmaceuticals and Clinical Pharmacy, School of Pharmacy, University of Jordan, Amman, Jordan; nhakooz@ju.edu.jo

¹⁰ College of Health and Life Sciences (CHLS), Hamad Bin Khalifa University (HBKU), Ar-Rayyan, Qatar; sismail@hbku.edu.qa; ORCID: 0000-0002-8425-4010

* Corresponding author: Said I. Ismail; sismail@hbku.edu.qa; ORCID: 0000-0002-8425-4010.

† Nancy Hakooz and Said I. Ismail jointly supervised this study and are co-senior authors.

Abstract

Genotype-guided prescribing depends on reliable population-specific evidence, yet pharmacogenomic (PGx) diversity in the Levant remains largely uncharacterized. To our knowledge, this is the largest population-scale PGx study conducted in this ancient region, characterizing PGx variation and its potential clinical actionability in 4,322 unrelated individual living in Jordan. Array-imputed genotypes were annotated using PharmGKB, while diplotypes, predicted phenotypes, and medication-specific prescribing recommendations were generated using PharmCAT. Several clinically relevant allele frequencies differed substantially from those reported in the 1000 Genomes Project (1KGP) populations. The *CYP3A5* rs776746-C (89.1% versus 62.2%; adjusted $P < 1.0 \times 10^{-300}$) and *VKORC1* rs9923231-T (49.6% versus 35.0%; $P = 9.88 \times 10^{-62}$) were enriched, whereas *CYP2C19* rs4244285-A was depleted (10.6% versus 22.1%; $P = 3.79 \times 10^{-75}$). Overall, 99.3% of participants had at least one medication-specific PGx result classified as potentially

actionable, involving a median of six implicated medications per participant. Warfarin-related dosing recommendations applied to 89.9% of participants, while predicted *CYP3A5* poor metabolism, reduced-function *UGT1A1*, and reduced-metabolism *CYP2C19* phenotypes were identified in 84.2%, 58.0%, and 25.6%, respectively. Jordanian allele frequencies showed substantially greater concordance with the gnomAD Middle Eastern reference population than with broader global populations, demonstrating marked regional specificity. These findings establish a population-scale PGx reference for Jordan and provide a foundation for prioritizing locally relevant gene-drug pairs, targeted testing, prospective clinical validation, and future precision-prescribing implementation in Jordan and across the Levant.

Keywords: pharmacogenomics; clinically actionable variation; precision prescribing; Levantines; Middle Eastern genomics.

Introduction

Interindividual variability in drug response remains a major clinical challenge, contributing to differences in therapeutic efficacy, toxicity, and adverse drug reactions. Although clinical and environmental factors influence treatment outcomes, inherited genetic variation is a major determinant of response to many commonly prescribed medications (Nemer & Hendi, 2023). Variation within pharmacogenes may alter drug pharmacokinetics, pharmacodynamics, and immune-mediated hypersensitivity, providing the biological basis for pharmacogenomics (PGx) and its application to genetically informed prescribing (Evans & Relling, 1999).

The clinical translation of PGx has been enabled by the development of curated knowledge resources and evidence-based prescribing guidelines. International initiatives, including the Pharmacogenomics Knowledgebase (PharmGKB)(Whirl-Carrillo et al., 2012), Pharmacogene Variation Consortium (PharmVar)(Gaedigk et al., 2018; Gaedigk et al., 2019), Clinical Pharmacogenetics Implementation Consortium (CPIC)(Relling & Klein, 2011), Dutch Pharmacogenetics Working Group (DPWG)(Swen et al., 2011), together with PGx information in regulatory drug labels, support the standardized interpretation for an expanding number of clinically actionable gene–drug pairs. Correspondingly, population-based studies indicate that most individuals carry at least one clinically actionable genotype and diplotype, often with prescribing implications for multiple medications (Jithesh et al., 2022; Van Driest et al., 2014).

A major challenge for the equitable implementation of PGx is ensuring that available evidence is applicable across diverse populations. The frequencies of clinically actionable alleles, diplotypes, and predicted phenotypes vary substantially because of demographic history, ancestry, natural selection, and genetic drift, resulting in population-specific differences in the prevalence and clinical impact of PGx variation (Genomes Project et al., 2015; Karczewski et al., 2021). Current evidence derived predominantly from European-enriched or broadly aggregated reference datasets may therefore inadequately represent PGx variation in underrepresented populations, including those from the Middle East and North Africa (MENA) (Dunnenberger et al., 2015; van der Wouden et al., 2017). Population-specific genomic resources are therefore essential for accurately estimating the clinical yield of PGx testing and guiding implementation strategies.

The Levant has served as a major crossroads of human migration and population exchange, resulting in a complex demographic history and fine-scale population structure (Haber et al., 2013). Despite this, Levantine populations remain largely underrepresented in PGx research. Although recent studies have expanded knowledge of clinically actionable

PGx variation in Qatar, the United Arab Emirates, and other Middle Eastern populations mainly from the Arabian Peninsula (Al-Mahayri et al., 2022; Al-Mahayri et al., 2020; El Rouby et al., 2022; Eljilany et al., 2023; Goljan et al., 2022; Jithesh et al., 2022; Velayutham et al., 2025), comprehensive population-scale data from Jordan and the Levant remain limited. This gap limits accurate assessment of clinically actionable PGx variation and hinders evaluation of how accurately existing reference populations reflect prescribing-relevant genetic diversity in the region.

Here, we characterize the PGx landscape of 4,322 unrelated individuals living in Jordan recruited through multiple healthcare institutions across several Jordanian governorates. Situated within the Levant at the crossroads of Africa, Asia, and Europe, Jordan has been shaped by repeated episodes of human migration and admixture over millennia, shaping the genetic architecture of contemporary Levantine populations (Haber et al., 2017). This geographically distributed cohort enables the characterization of clinically actionable PGx variation in an understudied population and provides an important reference for regional precision medicine. We estimated the prevalence of actionable variants and compared allele frequencies with regional and global reference populations. Our study establishes the first comprehensive population-scale PGx reference for Jordanians. It provides a foundation for population-informed implementation of precision prescription while contributing relevant evidence to the wider Levant and MENA region.

Results

Population differentiation of the Jordanian pharmacogenome

PGx analysis was completed for 4,322 unrelated Jordanian individuals retained after genotype quality control, following the workflow outlined in Fig. 1a. Principal-component analysis (PCA) placed Jordanians within the West Eurasian genetic continuum, nearest to European and South Asian populations, while revealing a distinct and internally heterogeneous distribution (Fig. 1b; Supplementary Figs. S1a–c). Among the variance represented by the first 20 PCs, PC1 and PC2 accounted for 44.1% and 21.9% of respectively.

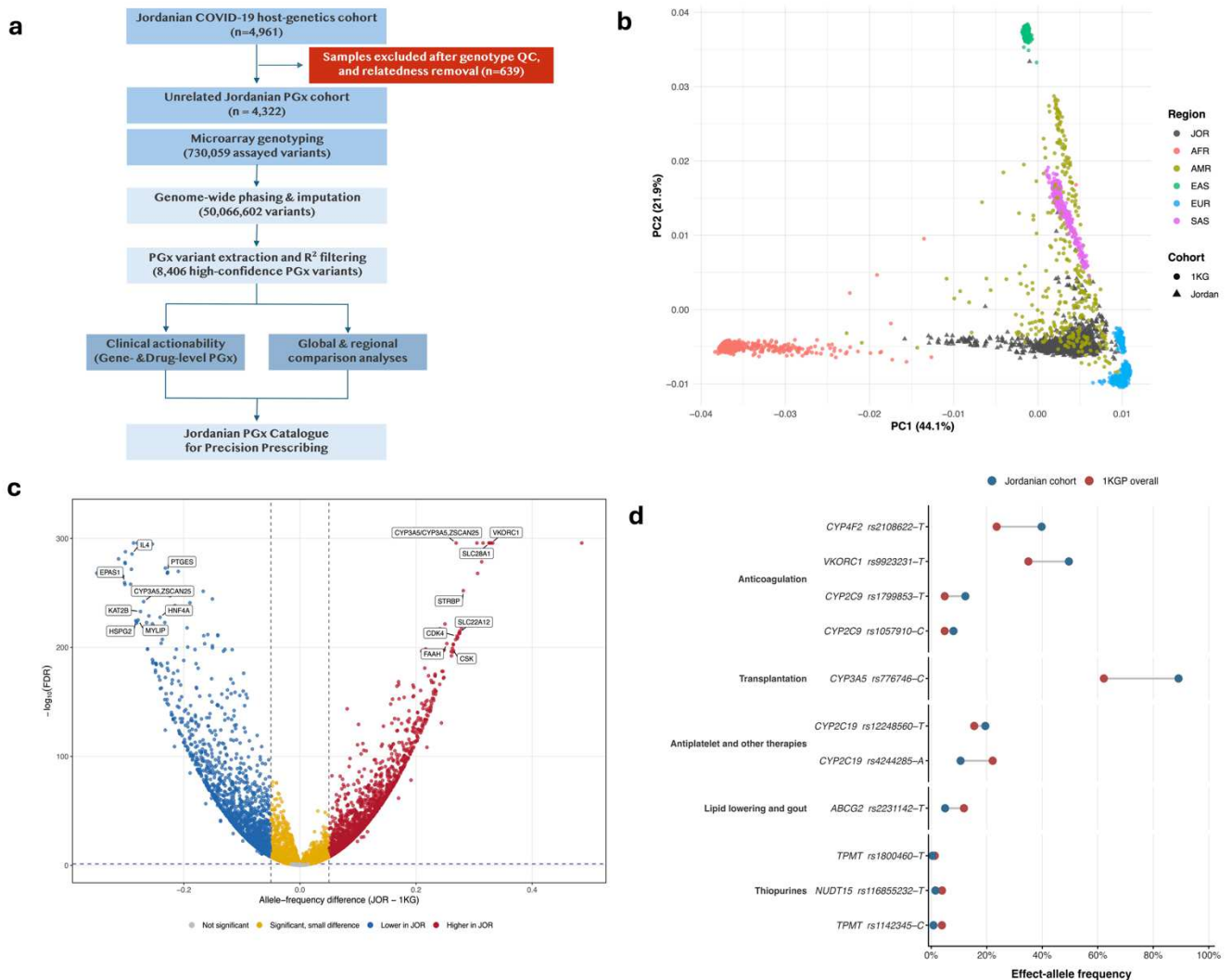


Fig. 1: Population differentiation of the Jordanian pharmacogenome.

a, Study design and pharmacogenomic analysis workflow. Quality-controlled (QC) microarray genotypes from 4,322 unrelated Jordanians were phased, imputed, and filtered ($R^2 \geq 0.80$) to generate a high-confidence pharmacogenomic (PGx) dataset for PharmCAT-based clinical interpretation and population-level comparison with the 1000 Genomes Project (1KGP), gnomAD, and published QGP data. **b**, Principal component (PC) analysis of 4,322 Jordanian and 2,574 1KGP participants. Axes show PC1 and PC2; points represent individuals, colours identify populations (AFR, African; AMR, Admixed American; EAS, East Asian; EUR, European; JOR, Jordanian; SAS, South Asian) and shapes distinguish Jordanian from 1KGP participants. **c**, Volcano plot of PharmGKB allele-frequency differentiation between Jordanians and 1KGP overall. Axes show the allele-frequency difference (JOR - 1KGP) and $-\log_{10}$ false discovery rate (FDR). Dashed vertical lines denote frequency differences of ± 0.05 and the horizontal line FDR = 0.05; colours indicate statistical significance with an absolute difference < 0.05 , effect magnitude, and direction. Top differentiated genes are labelled. **d**, Effect-allele frequencies of clinically relevant PharmGKB variants selected to represent major therapeutic areas and contrasting

patterns of population differentiation. Points show cohort-specific frequencies, connecting lines indicate between-cohort differences, and variants are grouped by therapeutic context.

Among 6,334 allele-compatible variants compared with 2,574 unrelated individuals from the 1000 Genomes Project (1KGP), allele frequencies differed significantly for 5,409 variants after false-discovery-rate correction ($FDR < 0.05$), of which 4,359 variants remained significant after Bonferroni correction. Of the FDR-significant variants, 2,546 were more frequent, and 2,863 were less frequent in Jordanians. PharmGKB annotations associated these variants with 1,810 unique gene annotations and 918 unique drugs or drug classes (Fig. 1c; Supplementary Fig. S1g; Supplementary Table S1).

Consistent with the PCA findings, the Jordanians PGx allele-frequency was most similar to that of Europeans, with the smallest median absolute allele-frequency differences (0.06; 4,047 FDR-significant variants), followed by South Asians (0.08; 4,924 variants;) and Admixed Americans (0.08; 4,403 variants). Differentiation was greatest relative to East Asians (5,648; 0.13) and Africans (5,643; 0.14) (Supplementary Figs. S1d–f and Supplementary Table S1).

Several clinically relevant PGx variants contributed prominently to this differentiation (Fig. 1d and Table 1). The *CYP3A5* loss-of-function allele rs776746-C was markedly enriched in Jordanians relative to 1KGP overall (89.1% versus 62.2%; odds ratio (OR) = 4.97, 95% CI: 4.55–5.42; Bonferroni-adjusted $P_{adj} < 1.00 \times 10^{-300}$), consistent with a high prevalence of the predicted *CYP3A5* poor-metabolizer (nonexpresser) phenotype. Enrichment was also observed for *VKORC1* rs9923231-T allele (49.6% versus 35.0%; OR = 1.83, 95% CI: 1.70–1.96; $P_{adj} = 9.68 \times 10^{-59}$) and *CYP2C9* rs1799853-T (12.3% versus 4.9%; OR = 2.74, 95% CI: 2.37–3.16; adjusted $P_{adj} = 3.19 \times 10^{-43}$), both are relevant to genotype-informed warfarin dosing.

Conversely, the *CYP2C19* rs4244285-A was less frequent in Jordanians than in 1KGP overall (10.6% versus 22.1%; OR = 0.42, 95% CI: 0.38–0.46; $P_{adj} = 2.70 \times 10^{-72}$) as was *ABCG2* rs2231142-T (5.0% versus 11.8%; OR = 0.39, 95% CI: 0.35–0.45; $P_{adj} = 8.32 \times 10^{-45}$). Lower frequencies were also observed for the thiopurine-associated *TPMT* rs1142345-C allele (0.8% versus 3.9%; OR = 0.20, 95% CI: 0.15–0.26; $P_{adj} = 2.08 \times 10^{-32}$) and *NUDT15* rs116855232-T (1.5% versus 3.9%; OR = 0.38, 95% CI: 0.30–0.47; $P_{adj} = 5.54 \times 10^{-15}$). In contrast, the frequency of *IFNL3* rs12979860-T showed no significant difference between Jordanians and 1KGP (Table 1).

Clinical actionability of the Jordanian pharmacogenome

PharmCAT generated gene-level findings across multiple pharmacogenes, although variation in the completeness of allele-defining positions resulted in gene-specific differences in diplotype resolution and phenotype assignment. Phenotype frequencies were therefore calculated using gene-specific denominators restricted to participants with an assigned phenotype, while ambiguous, indeterminate, and no-call results were reported separately (Fig. 2 and Table 2; Supplementary Fig. S2a). Gene-level results were obtained for all participants for *ABCG2*, *CYP3A5*, *DPYD*, *NUDT15*, and *VKORC1*. Phenotype-assigned results were also high for *UGT1A1* (99.6%), *TPMT* (99.4%), and *G6PD* (96.8%), but were lower for *CYP2C19* (64.5%), *CYP2C9* (63.6%), *SLCO1B1* (47.3%), and *CYP2B6* (32.8%). Gene-specific distributions of resolved diplotypes, assigned phenotypes, and unresolved results are provided in Supplementary Tables S2–S4, respectively.

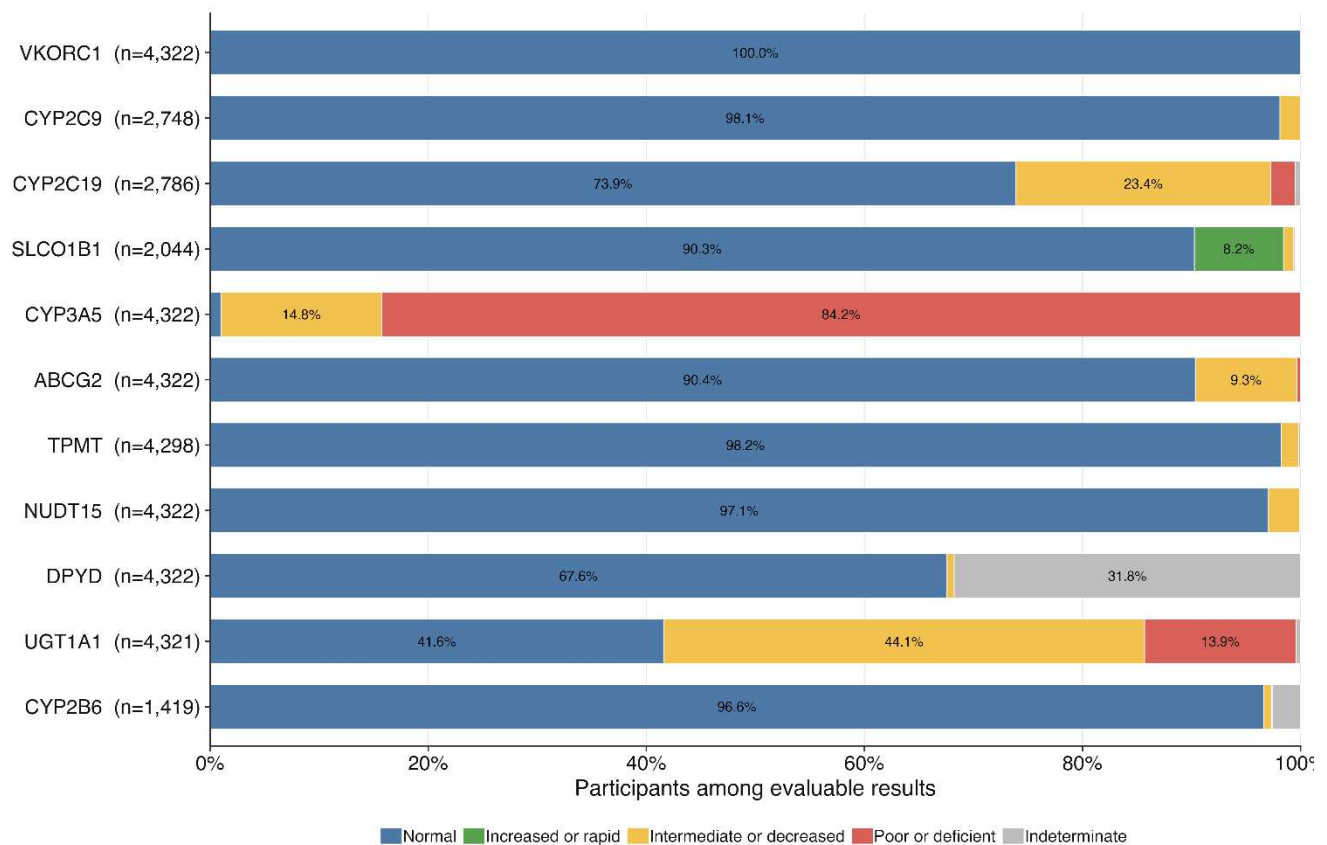


Fig. 2: Predicted pharmacogenomic phenotype distributions in Jordanians.

Horizontal stacked bars show the proportions of participants classified as normal, increased or rapid, intermediate or decreased, poor or deficient, or indeterminate for each pharmacogene. Percentages were calculated using gene-specific denominators comprising participants with a retained gene-level PharmCAT result, as indicated beside each pharmacogene. Ambiguous diplotypes and no-call results were excluded and summarized separately.

Table 2. Major clinically actionable pharmacogenomic findings in Jordanians.

Pharmacogene	Evaluable participants, n/N (%)	Representative medications	Major actionable genotype(s)/diplotype(s), n (% of cohort)	Predicted functional phenotype	Principal prescribing action	Participants with any guideline-linked actionable result, n (% of cohort)	Evidence source(s)
<i>VKORC1</i>	4,322/4,322 (100.0%)	Warfarin; acenocoumarol; phenprocoumon	–1639 AG, 2,109 (48.8%); –1639 AA, 1,089 (25.2%)	Increased warfarin sensitivity	Genotype-informed dose selection and INR monitoring	3,865 (89.4%)	DPWG; FDA association; FDA label
<i>CYP2C19</i>	2,801/4,322 (64.8%)	Clopidogrel; voriconazole; proton-pump inhibitors; antidepressants	*1/*2, 540 (12.5%); *2/*2, 55 (1.3%); *1/*3, 12 (0.3%); additional diplotypes in Supplementary Table 3	Rapid, ultrarapid, intermediate or poor metabolizer	Drug-specific alternative therapy, dose modification or enhanced monitoring	2,772 (64.1%)	CPIC; DPWG; FDA association; FDA label
<i>CYP2C9</i>	2,751/4,322 (63.7%)	Warfarin; phenytoin; non-steroidal anti-inflammatory drugs	*1/*11, 20 (0.5%); *1/*8, 19 (0.4%); *1/*6, 9 (0.2%)	Intermediate or poor metabolizer	Lower starting dose, alternative therapy or enhanced monitoring	2,701 (62.5%)	CPIC; DPWG; FDA association; FDA label
<i>UGT1A1</i>	4,322/4,322 (100.0%)	Irinotecan; atazanavir; belinostat	*1/*80+*28, 1,874 (43.4%); *80+*28/*80+*28, 582 (13.5%)	Intermediate or poor metabolizer	Drug-specific dose reduction, alternative therapy or toxicity monitoring	2,507 (58.0%)	CPIC; DPWG; FDA association; FDA label
<i>CYP3A5</i>	4,322/4,322 (100.0%)	Tacrolimus	*1/*3, 609 (14.1%); *1/*1, 43 (1.0%); *1/*6, 24 (0.6%); *1/*7, 5 (0.1%)	Tacrolimus expresser	Higher initial dose followed by therapeutic drug monitoring	681 (15.8%)	CPIC; DPWG; FDA association
<i>ABCG2</i>	4,322/4,322 (100.0%)	Rosuvastatin; allopurinol	rs2231142 G/T, 403 (9.3%); T/T, 14 (0.3%)	Decreased or poor function	Drug-specific dose adjustment or alternative therapy	419 (9.7%)	CPIC; DPWG
<i>CYP2B6</i>	1,437/4,322 (33.2%)	Efavirenz; sertraline	*1/*18, 10 (0.2%); additional diplotypes in Supplementary Table 3	Intermediate or poor metabolizer	Lower dose, slower titration or therapeutic drug monitoring	247 (5.7%)	CPIC; DPWG
<i>NUDT15</i>	4,322/4,322 (100.0%)	Azathioprine; mercaptopurine; thioguanine	*1/*3, 124 (2.9%); other actionable diplotypes in Supplementary Table 3	Intermediate or poor metabolizer	Reduced starting dose or alternative non-thiopurine therapy	194 (4.5%)	CPIC; DPWG; FDA association; FDA label
<i>TPMT</i>	4,298/4,322 (99.4%)	Azathioprine; mercaptopurine; thioguanine	*1/*3C, 34 (0.8%); *1/*3A, 33 (0.8%)	Intermediate or poor metabolizer	Reduced starting dose or alternative non-thiopurine therapy	191 (4.4%)	CPIC; DPWG; FDA association; FDA label
<i>DPYD</i>	4,322/4,322 (100.0%)	Capecitabine; fluorouracil; tegafur	Multiple actionable variant combinations; complete distribution in Supplementary Table 3	Intermediate metabolizer or decreased function	Reduced starting dose, alternative therapy or toxicity monitoring	150 (3.5%)	CPIC; DPWG; FDA label
<i>SLCO1B1</i>	2,068/4,322 (47.8%)	Atorvastatin; simvastatin; rosuvastatin; other statins	Multiple actionable diplotypes; complete distribution in Supplementary Table 3	Decreased or poor function	Lower statin dose or selection of an alternative statin	47 (1.1%)	CPIC; DPWG
<i>G6PD</i>	4,183/4,322 (96.8%)	Rasburicase; primaquine; dapsone; methylene blue	Phenotype-based classification; diplotype not applicable	Deficient or variable activity	Drug avoidance, alternative therapy or enhanced monitoring	14 (0.3%)	CPIC; FDA label

Evaluable participants had an unambiguous PharmCAT result, including usable calls with missing-position warnings. Representative genotypes/diplotypes occurring in at least five participants are shown. Percentages use the full cohort denominator (n = 4,322). Complete distributions and medication-specific recommendations are provided in Supplementary Tables 3 and 4, respectively. **Abbreviations:** CPIC, Clinical Pharmacogenetics Implementation Consortium; DPWG, Dutch Pharmacogenetics Working Group; FDA, U.S. Food and Drug Administration; INR, international normalized ratio.

Overall, 4,293 participants (99.3%) carried at least one medication-specific actionable PGx result, involving a median of six medications per participant (interquartile range, 4–8). All participants had PGx guidance for at least one medication (Supplementary Fig. S2c and Supplementary Table S5). Actionable results supported by CPIC and DPWG recommendations were identified in 99.0% and 63.7% of participants, respectively. The U.S. Food and Drug Administration (FDA)-recognized PGx associations were identified in 92.7% of participants, including results supported by FDA-approved drug labelling in 69.9% (Supplementary Fig. S2d).

Warfarin-related prescribing recommendations and the predicted *CYP3A5* nonexpresser phenotype were among the most prevalent findings. The *VKORC1* rs9923231-T allele was carried by 74.0% of participants (heterozygous AG, 49.0%; homozygous AA, 25.2%). Integrated *VKORC1*–*CYP2C9* interpretation generated warfarin-related prescribing recommendations for 3,882 participants (89.9 %). All 4,322 participants received an assigned *CYP3A5* phenotype, 84.2% were predicted poor metabolizers (nonexpressers), whereas 15.8% were expressers comprising 14.8% intermediate and 1.0% normal metabolizers. Reduced *CYP2C19* metabolism was identified in 713 of 2,786 participants with an assigned phenotype (25.6%). For *CYP2C9*, 51 of 2,748 phenotype-assigned participants (1.9%) were intermediate metabolizers and one (<0.1%) was a poor metabolizer (Supplementary Fig. S2b and Supplementary Table S4).

Variation in drug transport and glucuronidation was strongly gene-dependent. For *ABCG2*, decreased- and poor-function phenotypes were identified in 403 (9.3%) and 14 (0.3%), respectively. Reduced *SLCO1B1* function, comprising decreased- and poor-function phenotypes, was observed in 21 of 2,044 participants (1.0%) with an assigned phenotype. Reduced *UGT1A1* activity was more frequent, affecting 2,507 of 4,321 participants (58.0%), including 1,906 (44.1%) intermediate and 601 (13.9%) poor metabolizers. The predominant *UGT1A1* diplotypes were *1/*80+*28, identified in 1,874 of 4,322 participants (43.4%), and *80+*28/*80+*28, identified in 582 (13.5%) (Supplementary Fig. S2b and Supplementary Tables S3 and S4).

Predicted phenotypes associated with severe drug toxicity were comparatively infrequent (Supplementary Table S4). Among 4,322 participants with *NUDT15* results, 124 (2.9%) were intermediate metabolizers and three (0.07%) were poor metabolizers. Among 4,298 participants with *TPMT* results, 69 (1.6%) were intermediate metabolizers and one (<0.1%) was a poor metabolizer; seven additional results were indeterminate. For *DPYD*, 28 of 4,322 participants (0.6%) were intermediate metabolizers, whereas 1,373 (31.8%) had indeterminate classifications. Among 4,183 participants with *G6PD* results, nine (0.2%)

were classified as deficient and five (0.1%) as variable, corresponding to a combined non-normal phenotype frequency of 14/4,183 (0.3%).

Global and regional differentiation of clinically relevant pharmacogenomic variation

Actionable phenotype frequencies were compared across populations using harmonized gene-specific definitions and matched phenotype-assigned denominators (Fig. 3 and Supplementary Table S6). Compared with 1KGP overall, Jordanians had lower frequencies of actionable results involving *ABCG2* (9.7% versus 22.6%; Bonferroni- $P_{adj}=1.06\times10^{-47}$, *CYP3A5* (15.8% versus 47.6%; $P_{adj}=1.21\times10^{-178}$), *NUDT15* (4.5% versus 16.3%; $P_{adj}=2.50\times10^{-61}$), *TPMT* (4.4% versus 16.3%; $P_{adj}=2.58\times10^{-62}$), and *G6PD* (0.3% versus 1.3%; $P_{adj}=2.05\times10^{-05}$) (Fig. 3a). Relative to Europeans, Jordanians also had lower actionable frequencies for *ABCG2* (9.6% versus 19.6%; $P_{adj}=6.02\times10^{-11}$) and *TPMT* (4.4% versus 8.2%; $P_{adj}=1.97\times10^{-03}$), but a higher frequency for *CYP3A5* (15.8% versus 10.1%; $P_{adj}=8.90\times10^{-03}$). Differences for *NUDT15* and *G6PD* were not significant after Bonferroni correction (Fig. 3b).

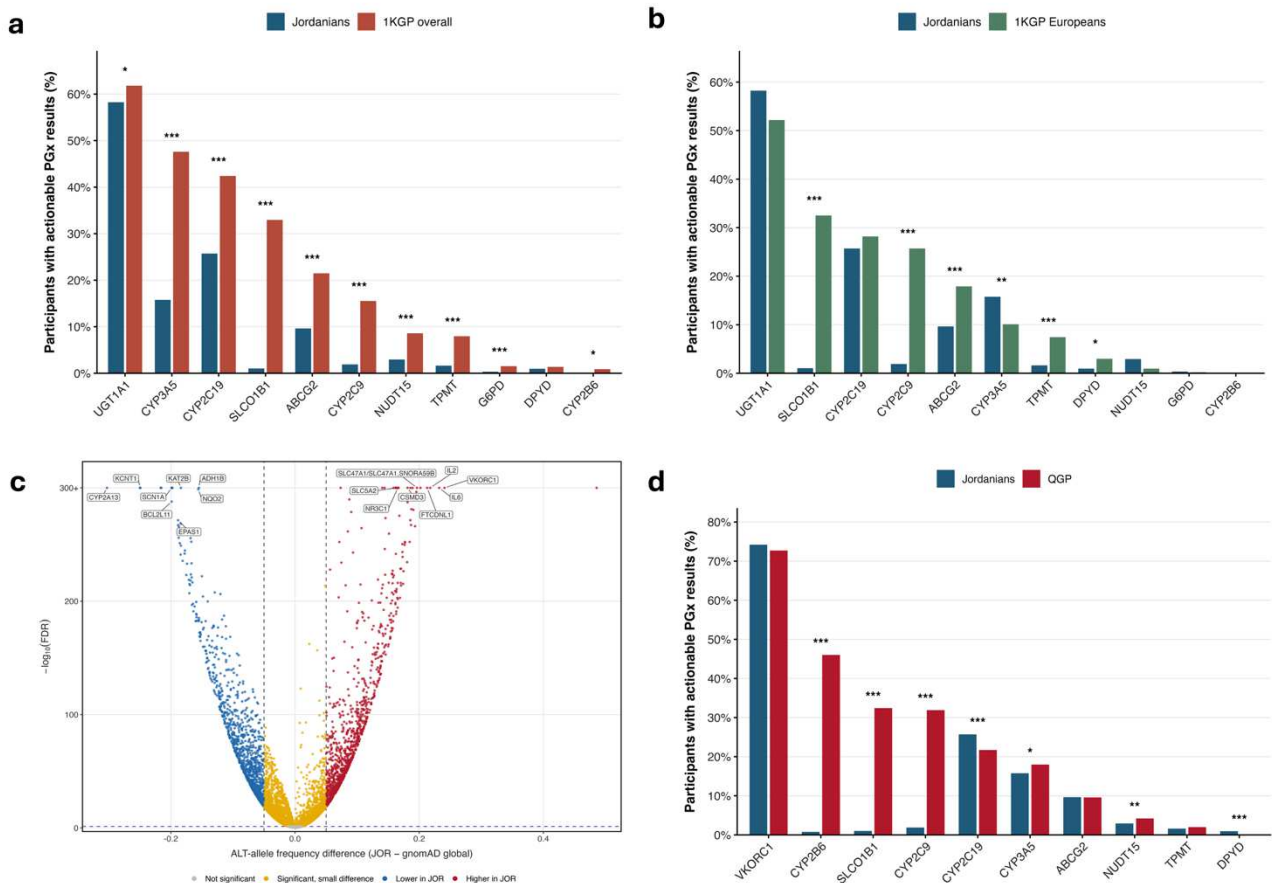


Fig. 3: Global and regional differentiation of clinically relevant pharmacogenomic variation in Jordanians.

a, b, Frequencies of harmonized gene-specific actionable phenotypes in Jordanians (JOR; total cohort, n = 4,322) compared with **(a)** 1KGP overall (n = 2,574) and **(b)** the 1KGP European superpopulation (EUR; n = 525). Percentages were calculated using gene-specific endpoint-assessable denominators, identical definitions applied to both populations. **c**, Volcano plot of exact-allele frequency differences between Jordanians and Genome Aggregation Database (gnomAD) global genomes. Axes represent the alternative-allele-frequency difference (JOR-gnomAD global) and $-\log_{10}$ false discovery rate (FDR). Dashed vertical lines indicate differences of ± 0.05 and the horizontal line indicates FDR = 0.05. Values exceeding the plotted y-axis range are displayed at 300+, and selected differentiated genes are labelled. **d**, Harmonized gene-specific PGx endpoints in Jordanians and the Qatar Genome Program (QGP; n = 6,045), based on matched phenotype- or genotype-level definitions (Supplementary Table S8). QGP counts were reconstructed from published frequencies. For **a, b** and **d**, two-sided tests of proportions or Fisher's exact tests for sparse data were applied, with Bonferroni correction: $*P_{adj} < 0.05$, $**P_{adj} < 1 \times 10^{-4}$, and $***P_{adj} \leq 1 \times 10^{-6}$.

At the variant level, 5,291 of 6,154 PharmGKB-annotated alleles (86.0%) differed significantly between Jordanians and the gnomAD global population after FDR correction, including 4,415 that remained significant after Bonferroni correction. Among the FDR-significant alleles, 2,514 were more frequent, and 2,777 were less frequent in Jordanians. Population-specific comparisons revealed substantially fewer differences relative to the gnomAD Middle Eastern population than to non-Finnish Europeans (94 versus 5,379 FDR-significant variants, respectively (Fig. 3c; Supplementary Table S7).

Several clinically actionable variants differed markedly in frequency between Jordanians and the gnomAD global population (Table 1; Supplementary Fig. S3). The largest differences involved genes relevant to anticoagulation and transplantation. The *VKORC1* rs9923231-T (49.6% versus 31.9%; $P_{adj}=2.0 \times 10^{-252}$), *CYP4F2* rs2108622-T (39.8% versus 23.0%; $P_{adj}=4.2 \times 10^{-276}$), and *CYP3A5* rs776746-C (89.1% versus 72.9%; $P_{adj}=2.0 \times 10^{-242}$) alleles were all enriched in Jordanians. The decreased-function *CYP2C9* alleles rs1799853-T (12.3% versus 8.8%; $P_{adj}=6.5 \times 10^{-28}$) and rs1057910-C (8.0% versus 4.9%; $P_{adj}=5.5 \times 10^{-36}$) were also more frequent. Conversely, the *CYP2C19* rs4244285-A (10.6% versus 16.8%) and *ABCG2* rs2231142-T (5.0% versus 9.3%) alleles were less frequent in Jordanians, as were the thiopurine-associated *TPMT* variants rs1142345-C (0.8% versus 4.3%) and rs1800460-T (0.4% versus 2.6%) (all $P_{adj} < 2 \times 10^{-60}$). Frequencies of the principal *VKORC1*, *CYP4F2*, *CYP3A5*, *CYP2C19*, and *ABCG2* alleles in the gnomAD Middle Eastern population more closely approximated those observed in Jordanians than did the corresponding global or non-Finnish European populations.

Table 1. Population frequencies of selected clinically relevant pharmacogenomic alleles.

Pharmacogene	<i>VKORC1</i>	<i>CYP4F2</i>	<i>CYP3A5</i>	<i>CYP2C9</i>		<i>CYP2C19</i>	<i>ABCG2</i>	<i>TPMT</i>		<i>NUDT15</i>	<i>UGT1A1</i>	<i>IFNL3</i>
Variant–effect allele	rs9923231–T	rs2108622–T	rs776746–C	rs1799853–T	rs1057910–C	rs4244285–A	rs2231142–T	rs1142345–C	rs1800460–T	rs116855232–T	rs887829–T	rs12979860–T
Functional consequence	Reduced VKORC1 expression	Reduced enzymatic activity	Loss of function; non–expresser	Decreased function	Decreased function	No function	Decreased function	No function	No function	Decreased function	Reduced expression	Altered treatment response
Representative medications	Warfarin	Warfarin	Tacrolimus	Warfarin; phenytoin; NSAIDs	Warfarin; phenytoin; NSAIDs	Clopidogrel; voriconazole; proton–pump inhibitors	Rosuvastatin; allopurinol	Thiopurines	Thiopurines	Thiopurines	Irinotecan; atazanavir	Peginterferon; ribavirin
Jordanian cohort	49.6%	39.8%	89.1%	12.3%	8.0%	10.6%	5.0%	0.8%	0.4%	1.5%	35.6%	35.6%
Jordanian cohort compared with the 1000 Genomes Project												
1KGP Europeans	38.4%	28.2%	94.5%	12.6%	7.2%	14.4%	9.5%	2.9%	2.8%	0.2%	29.7%	30.9%
1KGP overall	35.0%	23.6%	62.2%	4.9%	4.8%	22.1%	11.8%	3.8%	1.3%	3.9%	35.4%	35.6%
Difference (Jordan – 1KGP overall)	14.6%	16.2%	26.9%	7.4%	3.2%	–11.6%	–6.8%	–3.1%	–0.9%	–2.4%	0.2%	0.0%
OR (95% CI)	1.83 (1.70–1.96)	2.14 (1.98–2.31)	4.97 (4.55–5.42)	2.74 (2.38–3.16)	1.71 (1.47–1.98)	0.42 (0.38–0.46)	0.39 (0.35–0.45)	0.20 (0.15–0.27)	0.32 (0.21–0.47)	0.38 (0.30–0.47)	1.01 (0.94–1.09)	1.00 (0.93–1.08)
<i>P</i> _{adj}	9.88×10^{–62}	1.80×10^{–83}	<1×10^{–300}	2.29×10^{–46}	1.75×10^{–12}	3.79×10^{–75}	6.25×10^{–48}	1.14×10^{–35}	8.41×10^{–9}	1.94×10^{–18}	0.797	0.97
Jordanian cohort compared with gnomAD genomes												
gnomAD Middle Eastern	53.4%	39.1%	89.5%	13.9%	6.5%	8.6%	5.4%	1.7%	1.4%	0.3%	33.0%	38.1%
gnomAD non–Finnish European	37.8%	28.6%	93.1%	12.7%	6.6%	14.8%	10.9%	4.2%	3.8%	0.3%	32.1%	31.5%
gnomAD global	31.9%	23.0%	72.9%	8.8%	4.9%	16.8%	9.3%	4.3%	2.6%	1.3%	36.1%	38.6%
Difference (Jordan – gnomAD global)	17.7%	16.8%	16.2%	3.5%	3.1%	–6.2%	–4.3%	–3.5%	–2.2%	0.2%	–0.5%	–3.0%
OR (95% CI)	2.10 (2.01–2.19)	2.21 (2.12–2.32)	3.05 (2.84–3.26)	1.45 (1.36–1.55)	1.68 (1.55–1.82)	0.59 (0.55–0.63)	0.51 (0.46–0.56)	0.18 (0.14–0.23)	0.15 (0.11–0.21)	1.20 (1.00–1.43)	0.98 (0.94–1.02)	0.88 (0.84–0.92)
<i>P</i> _{adj}	1.99×10^{–252}	4.18×10^{–276}	2.00×10^{–242}	6.50×10^{–28}	5.46×10^{–36}	2.40×10^{–51}	1.90×10^{–41}	5.85×10^{–57}	1.84×10^{–36}	0.053	0.384	3.59×10^{–8}

Variants were selected for clinical relevance, therapeutic breadth, and contrasting population-frequency patterns. Frequencies refer to the stated effect allele. Differences, odds ratios (ORs), and false-discovery-rate (FDR)-adjusted P values (*P*_{adj}) compare Jordanians with with 1KGP overall or gnomAD global genomes, as indicated. gnomAD comparisons used exact GRCh38 chromosome-position-reference-alternate matches. Bold values indicate FDR < 0.05. Abbreviations: CI, confidence interval; 1KGP, 1000 Genomes Project; gnomAD, Genome Aggregation Database; NSAIDs, non-steroidal anti-inflammatory drugs.

To assess regional variation, we compared matched gene-specific PGx endpoints between the Jordanian cohort and published aggregate data from Qatar Genome Program (QGP) (Fig. 3d and Supplementary Table S8). The largest differences were observed for *CYP2B6* intermediate or poor metabolism (11/1,382, 0.80%, in Jordanians versus approximately 2,781/6,045, 46.00%, in QGP; $P_{adj}=4.10 \times 10^{-214}$), decreased or poor *SLCO1B1* function (21/2,033, 1.03% versus approximately 1,959/6,045, 32.41%; $P_{adj}=5.06 \times 10^{-177}$), and *CYP2C9* intermediate or poor metabolism (52/2,748, 1.89% versus approximately 1,928/6,045, 31.89%; $P_{adj}=5.86 \times 10^{-213}$), all of which were less frequent in Jordanians. Reduced *CYP2C19* metabolism was modestly more frequent in Jordanians than in QGP (713/2,772, 25.72% versus 1,312/6,045, 21.70%; $P_{adj}=3.13 \times 10^{-4}$). Smaller differences were observed for *CYP3A5*, *NUDT15*, and *DPYD* (all $P_{adj} < 0.05$), whereas *VKORC1*, *TPMT*, and *ABCG2* did not differ significantly. Sensitivity analyses accounting for rounding of the published QGP frequencies did not alter these findings.

Coverage of complex pharmacogenomic loci

The array-imputed dataset included 186 variants across four pharmacogenes with complex genomic architecture, comprising 103 variants in *CYP2D6*, 40 in *HLA-A*, 11 in *HLA-B*, and 32 in *MT-RNR1*. Although the *CYP2D6* marker rs3892097-T and selected *HLA*-region proxy variants were represented, the available array data did not support validated *CYP2D6* structural-variant, hybrid-allele, diplotype calling, or classical *HLA-A* and *HLA-B* allele assignment. In addition, the clinically actionable *MT-RNR1* variants m.1095T>C, m.1494C>T, and m.1555A>G could not be reliably resolved. Consequently, these loci were excluded from definitive clinical phenotype assignment and actionable-burden analyses (Supplementary Tables S9 and S10).

Discussion

This study provides the first population-scale characterization of clinically actionable PGx variation in a Levantine population. Analysis of 4,322 unrelated Jordanians extends the available national evidence beyond individual candidate genes and clinically selected cohorts by integrating allele frequencies, predicted diplotypes and phenotypes, and medication-specific prescribing recommendations. 99.3% of participants had at least one medication-specific PGx result classified as potentially actionable, with a median of six implicated medications per individual. These findings indicate substantial potential reach for PGx-guided prescribing. Notably, the marked differences from global reference populations demonstrate that the expected clinical yield of PGx testing may be inaccurately estimated when inferred from ancestrally mismatched datasets.

Previous PGx research in Jordan has largely focused on specific gene–drug associations. Over 30 studies have been published, primarily involving disease-specific or case-based cohorts genotyped using limited variant panels. Most were conducted at single centres or within geographically restricted populations, and only five included more than 500 participants. Although clinically relevant associations were reported for genes such as *VKORC1*, *CYP2C9*, and *CYP2C19*, the broader population distribution and potential prescribing relevance of PGx variation remained largely undefined (Yehya & Altaany, 2021). More recently, whole-genome sequencing of 76 Jordanian patients with genetic disorders identified PGx biomarkers relevant to warfarin, clopidogrel, antidepressants, and anticancer therapy (Jarrar et al., 2025).

The present study extends this evidence provides the first large-scale characterization of potentially actionable PGx variation in Jordan and translates thousands of variants into clinically interpretable phenotypes and medication-specific prescribing recommendations. By leveraging a multi-centre cohort, recruitment through healthcare institutions across several Jordanian governorates provides broader geographic representation than previous single-centre studies. Thus, this study advances the national PGx evidence beyond individual variant discovery towards population-level potential of Gx-guided prescription.

Patterns of PGx variation broadly reflected the position of Jordanians within the West Eurasian continuum and the demographic history of the Levant. Differences were smallest relative to Europeans, followed by South Asians and Admixed Americans, and greatest relative to East Asians and Africans. Broad continental comparisons, however, did not fully capture regional population structure (Gouveia et al., 2023). Jordanian allele frequencies showed substantially greater concordance with the gnomAD Middle Eastern reference, consistent with shared regional ancestry. Residual differentiation may reflect the fine-scale genetic structure and admixture patterns of populations in the Levant and Arabian Peninsula (Elliott et al., 2022; Haber et al., 2013; Rodriguez-Flores et al., 2016). These findings support the use of regional reference datasets to provide greater resolution of PGx variation, while population-specific data remain essential for accurately estimating the prevalence of clinically relevant alleles and informing PGx clinical implementation.

The high prevalence of potentially actionable PGx variation shifts the implementation question from whether PGx testing is relevant to which gene–drug pairs offer the greatest clinical value. A pre-emptive panel provides reusable information across multiple prescriptions. However, the presence of genotype-linked prescribing guidance alone does not establish clinical utility (Keeling et al., 2019). Prioritization should encompass common phenotypes affecting widely prescribed medicines and less frequent variants associated

with severe, preventable adverse outcomes. Evidence that panel-guided prescribing can reduce adverse drug reactions supports this risk-and evidence-based approach (Swen et al., 2023). Integration of the population-level PGx findings with national prescribing patterns, disease burden, and clinical outcomes is therefore the critical next step towards a clinically effective and economically sustainable PGx programme in Jordan.

Anticoagulation emerged as a strong candidate for prospective implementation. Warfarin-related recommendations were identified for 89.9% of participants, driven by the high prevalence of functional variation in *VKORC1* rs9923231-T, *CYP4F2* rs2108622-T, and *CYP2C9* alleles rs1799853-T and rs1057910-C. This population-level pattern reinforces earlier evidence linking *VKORC1* and *CYP2C9* variation to warfarin sensitivity and dose requirements in Jordanian cardiovascular patients (Al-Eitan et al., 2019). These pharmacogenes are central components of CPIC genotype-guided warfarin dosing recommendations (Nemer & Hendi, 2023; Relling & Klein, 2011). These findings therefore support prospective evaluation of a locally, genotype-guided dosing approach that integrate PGx result with established clinical factors and continued INR monitoring. The Clinical utility should be assessed through anticoagulation control, reductions in bleeding and thromboembolic outcomes, and health care utilization compared to standard care.

Transplantation represents another relevant pathway for targeted PGx implementation. Although most participants were predicted *CYP3A5* poor metabolizers, 15.8% were classified as expressers. For patients receiving tacrolimus, *CYP3A5* expressers generally require a higher genotype-guided starting dose than nonexpressers; starting doses are generally recommended, followed by therapeutic drug monitoring to achieve and maintain the target concentration (Birdwell et al., 2015; Chen & Prasad, 2018). Jordanian *CYP3A5*-associated actionability was substantially lower than that in 1KGP overall but higher than that in Europeans, illustrating how reference-population selection alters the anticipated yield of PGx testing. A similar expresser frequency of approximately 18% was reported in Qatar (Jithesh et al., 2022), suggesting partial regional similarity. These findings support evaluation of pre-emptive *CYP3A5* testing within Jordanian transplant programmes, where genotype-guided initial dosing can be combined with therapeutic drug monitoring and established clinical factors.

Substantial clinical actionability was also observed across cardiovascular medicine, oncology, infectious diseases, gastroenterology, and psychiatry. Reduced *CYP2C19* metabolism was identified in 713 of 2,786 participants with an assigned phenotype (25.6%). However, its clinical implications vary by medication and indication. For clopidogrel, these phenotypes reduce active-metabolite formation and platelet inhibition, supporting genotype-guided treatment in defined cardiovascular settings (Dean & Kane,

2012). Reduced CYP2C19 metabolism may increase exposure to proton-pump inhibitors, potentially enhancing acid suppression and treatment efficacy (Lima et al., 2021). For voriconazole, reduced metabolism may increase drug exposure and toxicity risk, supporting genotype-informed dose selection alongside therapeutic drug monitoring (Moriyama et al., 2017).

Reduced *UGT1A1* activity was identified in 58.0% of participants, indicating a potentially broad impact on medications, such as irinotecan and atazanavir. However, the strength of prescribing recommendations depends on the phenotype, medication, dose, indication, and applicable guideline. Consequently, a high phenotype prevalence alone does not justify population-wide testing (Johnson et al., 2022). The priority of *UGT1A1*, *CYP2C19*, and other gene–drug pairs should be determined by local medication use, clinical outcomes, disease burden, and the feasibility of integrating results into prescribing workflows.

Conversely, toxicity-associated phenotypes, involving *NUDT15*, *TPMT*, *DPYD*, and *G6PD*, were less frequent but clinically potentially consequential. Their relatively low observed frequencies should not be interpreted as limited clinical value because the associated outcomes—including thiopurine-induced myelosuppression, fluoropyrimidine toxicity, and drug-induced haemolysis—can be serious and potentially preventable. For thiopurines, CPIC recommends genotype-guided dose modification based on both *TPMT* and *NUDT15*, whose relative contributions to toxicity risk vary across ancestry groups (Maillard et al., 2026; Relling & Klein, 2011). Because reduced function in either gene can increase the risk of severe thiopurine-related myelosuppression, testing both genes before treatment is more appropriate than selecting a single gene based on presumed ancestry. More broadly, implementation priorities should consider phenotype frequency alongside the severity and preventability of drug-related harm, local medication exposure, and the availability of effective risk-mitigation strategies.

Regional evidence increasingly indicates that clinically relevant PGx variation is widespread but heterogeneous across Middle Eastern populations (Goljan et al., 2022; Jithesh et al., 2022). Our harmonized comparison with QGP reinforces this heterogeneity, suggesting differences in *CYP2B6*, *SLCO1B1*, and *CYP2C9*, and more modest differences in *CYP2C19*, *CYP3A5*, *NUDT15*, and *DPYD*. In contrast, *VKORC1*, *TPMT*, and *ABCG2* showed broadly comparable frequencies. Restricting CYP2C19 to reduced-metabolism phenotypes changed the magnitude and direction of the observed difference, underscoring the importance of harmonized phenotype definitions in population comparisons.

These findings suggest that PGx profiles may vary within the region; however, the magnitude of individual gene-level differences should be interpreted cautiously. These differences may reflect population structure and analytical heterogeneity, including sequencing platforms, variant coverage, callability, and phenotype-translation methods. Whole-genome sequencing provides greater resolution of *CYP2D6*, classical *HLA*, and mitochondrial variation than array-based imputation. Apparent regional differences may therefore reflect demographic history and methodological variation (Ismail et al., 2026). Standardized sequencing, allele definition phenotype translation, and matched comparison criteria will be important for establishing a robust Middle Eastern PGx reference framework.

Several limitations should frame interpretation. Array-based imputation captured a broad range of common PGx variation but could not comprehensively resolve structurally complex loci, including *CYP2D6*, classical *HLA* alleles, and clinically important mitochondrial variants. Therefore, the detectable scope of PGx variation was incomplete. Diplotype resolution and phenotype assignment varied across pharmacogenes, particularly *CYP2B6*, *SLCO1B1*, *CYP2C9*, and *CYP2C19*, necessitating gene-specific denominators and limiting direct comparisons across genes. The high proportion of indeterminate *DPYD* classifications further limits interpretation of its apparently low actionable-phenotype frequency. PharmCAT-derived diplotypes and phenotypes were computational predictions and were not validated against pharmacokinetics, pharmacodynamics, or adverse events. Medication exposure and longitudinal clinical data were unavailable, preventing estimation of recommendation uptake or clinical benefit. External comparisons may also be influenced by differences in sequencing technology, variant ascertainment, allele definitions, callability thresholds, phenotype-translation methods, ancestry composition, and sample size. Although the cohort was large, further sampling is required to establish its representation of the complete geographic, demographic, and ancestral diversity of Jordan.

Future efforts should combine population-scale whole-genome sequencing with dedicated *CYP2D6*, *HLA*, and mitochondrial analysis, supported by orthogonal validation of clinically important calls. Integration of genomic profiles with prescriptions, therapeutic drug monitoring, and longitudinal outcomes will be essential to determine clinical utility. Initial prospective studies should prioritize well-supported pathways with plausible local impact, including warfarin dosing, tacrolimus therapy, antiplatelet therapy, thiopurines, and fluoropyrimidines. Transparent and locally validated clinical decision support systems could integrate genomic, clinical, and medication data for real-time precision prescribing. Regional Levant studies should then assess clinical and cost-effectiveness, health-system integration, and the distinction between shared and population-specific PGx variation.

In conclusion, we provided the first population-scale characterization of clinically actionable pharmacogenomic variation in Jordanians and established a population-specific evidence base for clinical translation. These findings support locally prioritized testing and prospective validation, while providing an essential foundation for precision prescribing in Jordan and across the Levant.

Methods

Study population

The study included genotype data from 4,961 Jordanian individuals originally recruited through the Jordan coronavirus disease 2019 (COVID-19) Host Genomics Initiative between December 2020 and early 2022 as part of international efforts to investigate host genetic determinants of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) susceptibility and COVID-19 severity (Initiative, 2021, 2023). Participants were recruited through governmental, university, military, and private healthcare institutions across northern, central, and southern Jordan (Supplementary Table S11).

Following genotype quality control and relatedness filtering, 4,322 unrelated Jordanian individuals were retained for PGx analyses, including 2,053 females (47.5%) and 2,269 males (52.5%), with a mean age of 56.0 years (standard deviation (SD): ± 17.5 years).

Ethics approval

The original Jordanian COVID-19 host-genomics studies were approved by the Scientific Research Ethics Committee of the Jordanian Ministry of Health on 9 December 2020 (approval no. MoH/REC/2020/211) and conducted in accordance with the Declaration of Helsinki and applicable national ethical and regulatory requirements. Informed consent was obtained from all participants under the approved study protocols. On 5 May 2024, the Jordanian Ministry of Health granted authorization for secondary use of the data across several genomic investigations, including the characterization of clinically actionable pharmacogenomic genotypes. The present study involved a secondary analysis of de-identified genotype data, with individual-level demographic information limited to age and sex.

Microarray genotyping

DNA samples were genotyped using the Infinium Global Screening Array-24 v3.0 BeadChip platform (Illumina, San Diego, CA, USA). Raw genotype intensity data were processed, and genotype clustering was performed using GenomeStudio software (Illumina). The initial

SNP-array dataset contained 730,059 assayed variants, with an overall genotyping rate of 94.5%. Variant genomic coordinates were assigned according to the GRCh38/hg38 Illumina GSA v3.0 manifest.

Quality control was performed using PLINK v1.90 and PLINK v2.0 (Purcell et al., 2007). Sample-level quality control excluded individuals with genotype missingness > 5% ($n = 123$), discordance between reported and genetically inferred sex ($n = 2$), or autosomal heterozygosity more than four standard deviations from the cohort mean ($n = 10$). Pairwise relatedness was assessed using identity-by-descent and KING-robust kinship estimates (Manichaikul et al., 2010), and one individual from each related pair was excluded at a kinship threshold of 0.0884, corresponding approximately to second-degree relatedness ($n = 490$).

Variant-level quality control excluded markers with a minor allele frequency (MAF) < 1%, high genotype missingness > 1%, or Hardy–Weinberg equilibrium $P < 1 \times 10^{-6}$ (Wigginton et al., 2005). Variants within low-complexity genomic regions were also excluded to reduce potential genotyping and alignment artefacts. The final quality-controlled dataset comprised 529,319 autosomal and non-autosomal variants in 4,329 participants, with an overall genotyping rate of 99.99%.

Phasing and genotype imputation

Quality-controlled SNP-array genotypes were phased and imputed using the Michigan Imputation Server v2.0.9 (<https://imputationserver.sph.umich.edu>). Phasing was performed with Eagle2 v2.4 (Loh et al., 2016), and genotype imputation was conducted against the 1KGP Phase 3 high-coverage (30×) reference panel on GRCh38 coordinates (Byrska-Bishop et al., 2022; Genomes Project et al., 2015). Following post-imputation relatedness and heterozygosity filtering, the final phased imputed dataset contained 50,066,602 variants in 4,322 unrelated individuals and was used for downstream PGx analysis.

Pharmacogenomic variant selection

PharmGKB clinical annotations were used to characterize PGx variation beyond the variants required for star-alleles (haplotypes) assignment (<https://www.pharmgkb.org/downloads/>, accessed June 2026). Variant-level analyses were restricted to simple variants with dbSNP reference SNP (rs) identifiers. Of 7,301 unique rsIDs identified, 7,290 were mapped to GRCh38 coordinates using Ensembl annotation resources (Dyer et al., 2025). After normalization, 6,353 variants (87.0%) were

matched to the Jordanian dataset by chromosome, position, reference allele, and alternate allele. Allele counts and frequencies were calculated using PLINK v2.0.

Variants required for pharmacogene haplotype and diplotype inference were selected using the genomic target regions and allele definition implemented in PharmCAT v3.2.0, with allele and phenotype definitions released on 9 February 2026 (Klein & Ritchie, 2018; Li et al., 2023; Sangkuhl et al., 2020). Of 26,589 extracted variants, 607 directly genotyped variants were retained without applying imputation-quality thresholds, whereas 7,799 imputed variants met the prespecified imputation-quality threshold of $R^2 \geq 0.80$. The remaining 18,183 variants (68.4%) were excluded because of low or missing imputation quality. The final PharmCAT input contained 8,406 high-quality variants in 4,322 participants, with imputation-derived phase information preserved for haplotype and diplotype inference.

Star alleles were assigned only when the phased genotype matched a PharmCAT allele definition across the required defining positions. Accordingly, detection of an individual defining variant in the variant-level dataset was not considered sufficient evidence for assignment of the corresponding star allele. To investigate the apparent absence of *CYP2C9**2 (rs1799853), *CYP2C9**3 (rs1057910), and *CYP2C19**17 (rs12248560) were specifically audited across the quality-filtered VCF, preprocessed PharmCAT input, allele-matching warnings and individual reports. Defining variants that were present but could not be incorporated into a unique phased haplotype were retained in variant-level analyses but were not assigned as star alleles or interpreted as reference alleles. Unresolved assignments were reported separately and excluded from diplotype- and phenotype-frequency calculations.

Pharmacogenomic analysis

Pharmacogene haplotypes, diplotypes, functional phenotypes and prescribing recommendations were determined using PharmCAT v3.2.0, Named Allele Matcher v2.0.0, and the allele and phenotype definitions released on 9 February 2026. The phased, quality-filtered VCF was processed using the PharmCAT preprocessor, Named Allele Matcher, Phenotyper, and Reporter modules. Observed haplotypes were matched to the implemented allele definitions and resolved diplotypes were translated into functional phenotypes. Genotype-specific prescribing recommendations were obtained from CPIC guidelines, PharmGKB-curated DPWG guidelines, and pharmacogenomic information contained in US FDA-approved drug labels. Individual reports were generated for all 4,322 participants.

Variant-completeness and allele-matching warnings produced by PharmCAT were retained for quality assessment. A resolved diplotype was defined as a single, unambiguous diplotype assignment, including warning-flagged assignments retained despite incomplete defining-position coverage. A phenotype-assigned result was defined as a resolved gene-level result translated into an interpretable phenotype. Indeterminate phenotypes, ambiguous diplotypes and no-call results were summarized separately and were not treated as negative findings. Diplotype and phenotype frequencies were calculated using gene-specific resolved-diplotype and phenotype-assigned denominators, respectively.

A result was classified as actionable when the inferred diplotype or phenotype was linked to an explicit recommendation to alter drug selection, dose, treatment initiation, toxicity risk, titration, or clinical monitoring. Drug–gene associations or toxicity statements without a corresponding prescribing recommendation were classified as PGx guidance but were not considered actionable in the primary analysis. Recommendations were summarized separately by source to distinguish CPIC-, DPWG-, and FDA-supported information. Major actionable diplotypes were defined as those observed in at least five participants. Actionable-result frequencies were summarized by gene, diplotype, phenotype, drug, and guideline source. Individual PGx burden was summarized as the number of unique drugs associated with an actionable recommendation and the number of drugs with any PGx guidance for each participant.

For population comparisons, phased high-coverage 1KGP GRCh38 data were initially available for 3,202 individuals; analyses were restricted to 2,574 unrelated participants with complete population assignments and harmonized variant data. Variant-level allele frequencies were compared with the combined 1KGP cohort and its five superpopulations. Gene-level comparisons used PharmCAT v3.2.0 and its bundled allele-definition release, applying consistent calling and warning-handling criteria to both cohorts. Published aggregate PGx frequencies from 6,045 QGP whole-genomes were used as regional comparisons (Jithesh et al., 2022). Comparisons were restricted to gene-specific phenotype- or genotype-level endpoints that could be matched across studies. Jordanian frequencies were calculated using endpoint-assessable denominators, and QGP counts were reconstructed from published percentages when counts were unavailable. Differences were assessed using two-sided pooled two-proportion tests, with Fisher’s exact test applied to sparse categories and Bonferroni correction across the ten endpoints compared. Sensitivity analyses examined the effect of rounding the published QGP percentages.

A complementary variant-level analysis characterized PharmGKB-annotated variation beyond PharmCAT diplotype calls. Allele frequencies in Jordanians were compared with those in the overall 1KGP cohort and its superpopulations. For comparison with gnomAD v4.1 genomes, PASS variants were harmonized to the Jordanian dataset by exact agreement of GRCh38 chromosome, position, reference allele, and alternate allele (<https://gnomad.broadinstitute.org/>; accessed June 2026) (Karczewski et al., 2021). Comparisons were conducted separately against gnomAD global, Middle Eastern, and non-Finnish European frequencies. All exactly matched alleles were included in the analysis, with representative clinically relevant variants selected for presentation based on PharmGKB evidence, effect magnitude, and therapeutic relevance.

The complex pharmacogenes *CYP2D6*, *HLA-A*, *HLA-B*, and *MT-RNR1* were assessed separately. The corresponding genomic regions contained 103, 40, 11, and 32 variants, respectively. However, the microarray-imputed data did not support reliable *CYP2D6* structural-allele inference, classical *HLA* genotyping, or mitochondrial heteroplasmy assessment. These genes were therefore included in the coverage analysis but were not assigned definitive clinical diplotypes or phenotypes.

Population structure analysis

Population structure was evaluated by jointly analysing the 4,322 unrelated Jordanian participants with 2,574 unrelated individuals from the 1KGP, comprising 681 African (AFR), 351 Admixed American (AMR), 506 East Asian (EAS), 525 European (EUR), and 511 South Asian (SAS) individuals. Variants were matched by chromosome and position, and reference and alternate alleles were aligned before merging. The datasets were restricted to shared autosomal biallelic variants on GRCh38 with concordant genomic coordinates and alleles. Duplicate and allele-incompatible sites were excluded before merging in PLINK v1.9 using `--bmerge`. Linkage disequilibrium (LD) pruning was performed in PLINK v1.9 using a 200-variant sliding window advanced in five-variant increments and an r^2 threshold of 0.05 (`--indep-pairwise 200 5 0.05`). Pruning retained 191,785 variants; after automatic exclusion of 6,097 non-autosomal variants, 185,688 autosomal variants were used to calculate the genetic relationship matrix and the first 20 principal components (PCs).

Statistical analysis

Differences in allele frequencies and actionable-result proportions between the Jordanian cohort and the 1KGP or gnomAD reference populations were assessed using two-sided pooled two-proportion z-tests. Effect sizes were reported as absolute allele-frequency differences and odds ratios with 95% CIs. Variant-wide comparisons were adjusted using

the Benjamini–Hochberg FDR procedure, whereas comparisons of predefined clinically actionable pharmacogenes were corrected using the Bonferroni method. An FDR- or Bonferroni-adjusted P -value < 0.05 , as applicable, was considered statistically significant.

Diplotype, phenotype, and actionable-result frequencies were calculated using gene-specific denominators, comprising participants with resolved diplotypes, assigned interpretable phenotypes, and reportable medication-specific recommendations, respectively. Harmonized population comparisons used gene-specific endpoint-assessable denominators. Results are presented as counts, percentages, and 95% Wilson CIs. Individual PGx burden was summarized using the mean, standard deviation, median, interquartile range, and observed range. All statistical tests were two-sided. Statistical analyses and graphical presentation were performed using R v4.4.1.

Data availability

All study-generated aggregate data supporting the findings are provided in the Article and Supplementary Information. Individual-level genotype and phenotype data are not publicly available because participant consent and applicable ethical approvals do not permit unrestricted data sharing. Access requests may be directed to the corresponding author and will be considered by the Jordan COVID-19 Host Genomics Initiative (Initiative, 2021, 2023), subject to institutional and ethical approval and execution of an appropriate data-use agreement. The 1KGP data are available through the International Genome Sample Resource (<https://www.internationalgenome.org>), gnomAD v4.1 data through the Genome Aggregation Database (<https://gnomad.broadinstitute.org>), and PharmGKB annotations through PharmGKB (<https://www.pharmgkb.org>). Published aggregate Qatar Genome Program data were obtained from (Jithesh et al., 2022).

Code availability

The analyses were conducted using publicly available software, including PLINK, BCFtools/HTSlib, PharmCAT and R. Custom analysis scripts are available from the corresponding author upon reasonable request.

Acknowledgements

We thank all participants and the healthcare institutions involved in the Jordan COVID-19 Host Genomics Initiative, particularly the Jordanian Ministry of Health. We also acknowledge the Institute for Molecular Medicine Finland (FIMM), University of Helsinki, for performing the genotyping of the original cohort.

Funding

Genotyping of the original cohort was supported by the Jordanian Ministry of Health and the Institute for Molecular Medicine Finland, University of Helsinki. The present secondary pharmacogenomic analysis received no dedicated research funding. Open-access publication charges will be supported by Qatar National Library.

Author contributions

N.N.H. conceived and designed the study, conducted the investigation, developed the methodology and software, performed the data curation and formal analysis, prepared the visualizations, and wrote the original manuscript draft. **L.M.A.** contributed to investigation. **N.I.** contributed to data curation. **R.A.** contributed to validation and manuscript review and editing. **N.M.O.** contributed resources. **N.H.** contributed to the supervision and manuscript review. **S.I.I.** conceived and supervised the study, contributed to project administration and resources, secured funding, and reviewed and edited the manuscript. All authors reviewed and approved the final manuscript.

Corresponding author

Correspondence to Said I. Ismail; e-mail: sismail@hbku.edu.qa.

Ethics declarations

Competing interests

The authors declare no competing interests.

Supplementary information

Supplementary Figs. S1–3 and Supplementary Tables S1–11.

Figure legends

Fig. 1: Population differentiation of the Jordanian pharmacogenome.

a, Study design and pharmacogenomic analysis workflow. Quality-controlled (QC) microarray genotypes from 4,322 unrelated Jordanians were phased, imputed, and filtered ($R^2 \geq 0.80$) to generate a high-confidence pharmacogenomic (PGx) dataset for PharmCAT-based clinical interpretation and population-level comparison with the 1000 Genomes Project (1KGP), gnomAD, and published QGP data. **b**, Principal component (PC) analysis of 4,322 Jordanian and 2,574 1KGP participants. Axes show PC1 and PC2; points represent

individuals, colours identify populations (AFR, African; AMR, Admixed American; EAS, East Asian; EUR, European; JOR, Jordanian; SAS, South Asian) and shapes distinguish Jordanian from 1KGP participants. **c**, Volcano plot of PharmGKB allele-frequency differentiation between Jordanians and 1KGP overall. Axes show the allele-frequency difference (JOR – 1KGP) and $-\log_{10}$ false discovery rate (FDR). Dashed vertical lines denote frequency differences of ± 0.05 and the horizontal line FDR = 0.05; colours indicate statistical significance with an absolute difference < 0.05 , effect magnitude, and direction. Top differentiated genes are labelled. **d**, Effect-allele frequencies of clinically relevant PharmGKB variants selected to represent major therapeutic areas and contrasting patterns of population differentiation. Points show cohort-specific frequencies, connecting lines indicate between-cohort differences, and variants are grouped by therapeutic context.

Fig. 2: Predicted pharmacogenomic phenotype distributions in Jordanians.

Horizontal stacked bars show the proportions of participants classified as normal, increased or rapid, intermediate or decreased, poor or deficient, or indeterminate for each pharmacogene. Percentages were calculated using gene-specific denominators comprising participants with a retained gene-level PharmCAT result, as indicated beside each pharmacogene. Ambiguous diplotypes and no-call results were excluded and summarized separately.

Fig. 3: Global and regional differentiation of clinically relevant pharmacogenomic variation in Jordanians.

a, b, Frequencies of harmonized gene-specific actionable phenotypes in Jordanians (JOR; total cohort, $n = 4,322$) compared with **(a)** 1KGP overall ($n = 2,574$) and **(b)** the 1KGP European superpopulation (EUR; $n = 525$). Percentages were calculated using gene-specific endpoint-assessable denominators, identical definitions applied to both populations. **c**, Volcano plot of exact-allele frequency differences between Jordanians and Genome Aggregation Database (gnomAD) global genomes. Axes represent the alternative-allele-frequency difference (JOR–gnomAD global) and $-\log_{10}$ false discovery rate (FDR). Dashed vertical lines indicate differences of ± 0.05 and the horizontal line indicates FDR = 0.05. Values exceeding the plotted y-axis range are displayed at 300+, and selected differentiated genes are labelled. **d**, Harmonized gene-specific PGx endpoints in Jordanians and the Qatar Genome Program (QGP; $n = 6,045$), based on matched phenotype- or genotype-level definitions (Supplementary Table S8). QGP counts were reconstructed from published frequencies. For **a, b** and **d**, two-sided tests of proportions or Fisher’s exact tests for sparse data were applied, with Bonferroni correction: $*P_{adj} < 0.05$, $**P_{adj} < 1 \times 10^{-4}$, and $***P_{adj} \leq 1 \times 10^{-6}$.

References

- Al-Eitan, L. N., Almasri, A. Y., & Khasawneh, R. H. (2019, May). Effects of CYP2C9 and VKORC1 polymorphisms on warfarin sensitivity and responsiveness during the stabilization phase of therapy. *Saudi Pharm J*, 27(4), 484-490.
<https://doi.org/10.1016/j.jsps.2019.01.011>
- Al-Mahayri, Z. N., Khasawneh, L. Q., Alqasrawi, M. N., Altoum, S. M., Jamil, G., Badawi, S., Hamza, D., George, L., AlZaabi, A., Ouda, H., Al-Maskari, F., AlKaabi, J., Patrinos, G. P., & Ali, B. R. (2022, Sep 25). Pharmacogenomics implementation in cardiovascular disease in a highly diverse population: initial findings and lessons learned from a pilot study in United Arab Emirates. *Hum Genomics*, 16(1), 42.
<https://doi.org/10.1186/s40246-022-00417-9>
- Al-Mahayri, Z. N., Patrinos, G. P., Wattanapokayakit, S., Iemwimangsa, N., Fukunaga, K., Mushiroda, T., Chantratita, W., & Ali, B. R. (2020, Dec 4). Variation in 100 relevant pharmacogenes among emiratis with insights from understudied populations. *Sci Rep*, 10(1), 21310. <https://doi.org/10.1038/s41598-020-78231-3>
- Birdwell, K. A., Decker, B., Barbarino, J. M., Peterson, J. F., Stein, C. M., Sadee, W., Wang, D., Vinks, A. A., He, Y., Swen, J. J., Leeder, J. S., van Schaik, R., Thummel, K. E., Klein, T. E., Caudle, K. E., & MacPhee, I. A. (2015, Jul). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guidelines for CYP3A5 Genotype and Tacrolimus Dosing. *Clin Pharmacol Ther*, 98(1), 19-24.
<https://doi.org/10.1002/cpt.113>
- Byrska-Bishop, M., Evani, U. S., Zhao, X., Basile, A. O., Abel, H. J., Regier, A. A., Corvelo, A., Clarke, W. E., Musunuri, R., Nagulapalli, K., Fairley, S., Runnels, A., Winterkorn, L., Lowy, E., Human Genome Structural Variation, C., Paul, F., Germer, S., Brand, H., Hall, I. M., Talkowski, M. E., Narzisi, G., & Zody, M. C. (2022, Sep 1). High-coverage whole-genome sequencing of the expanded 1000 Genomes Project cohort including 602 trios. *Cell*, 185(18), 3426-3440 e3419.
<https://doi.org/10.1016/j.cell.2022.08.004>
- Chen, L., & Prasad, G. V. R. (2018). CYP3A5 polymorphisms in renal transplant recipients: influence on tacrolimus treatment. *Pharmacogenomics Pers Med*, 11, 23-33.
<https://doi.org/10.2147/PGPM.S107710>

- Dean, L., & Kane, M. (2012). Clopidogrel Therapy and CYP2C19 Genotype. In V. M. Pratt, S. A. Scott, M. Pirmohamed, B. Esquivel, B. L. Kattman, & A. J. Malheiro (Eds.), *MINI Medical Genetics Summaries*. <https://www.ncbi.nlm.nih.gov/pubmed/28520346>
- Dunnenberger, H. M., Crews, K. R., Hoffman, J. M., Caudle, K. E., Broeckel, U., Howard, S. C., Hunkler, R. J., Klein, T. E., Evans, W. E., & Relling, M. V. (2015). Preemptive clinical pharmacogenetics implementation: current programs in five US medical centers. *Annu Rev Pharmacol Toxicol*, 55, 89-106. <https://doi.org/10.1146/annurev-pharmtox-010814-124835>
- Dyer, S. C., Austine-Orimoloye, O., Azov, A. G., Barba, M., Barnes, I., Barrera-Enriquez, V. P., Becker, A., Bennett, R., Beracochea, M., Berry, A., Bhai, J., Bhurji, S. K., Boddu, S., Branco Lins, P. R., Brooks, L., Ramaraju, S. B., Campbell, L. I., Martinez, M. C., Charkhchi, M., Cortes, L. A., Davidson, C., Denni, S., Dodiya, K., Donaldson, S., El Houdaigui, B., El Naboulsi, T., Falola, O., Fatima, R., Genez, T., Martinez, J. G., Gurbich, T., Hardy, M., Hollis, Z., Hunt, T., Kay, M., Kaykala, V., Lemos, D., Lodha, D., Mathlouthi, N., Merino, G. A., Merritt, R., Mirabueno, L. P., Mushtaq, A., Hossain, S. N., Perez-Silva, J. G., Perry, M., Pilizota, I., Poppleton, D., Prosovetskaia, I., Raj, S., Salam, A. I. A., Saraf, S., Saraiva-Agostinho, N., Sinha, S., Sipos, B., Sitnik, V., Steed, E., Suner, M. M., Surapaneni, L., Sutinen, K., Tricomi, F. F., Tsang, I., Urbina-Gomez, D., Veidenberg, A., Walsh, T. A., Willhoft, N. L., Allen, J., Alvarez-Jarreta, J., Chakiachvili, M., Cheema, J., da Rocha, J. B., De Silva, N. H., Giorgetti, S., Haggerty, L., Ilesley, G. R., Keatley, J., Loveland, J. E., Moore, B., Mudge, J. M., Naamati, G., Tate, J., Trevanion, S. J., Winterbottom, A., Flint, B., Frankish, A., Hunt, S. E., Finn, R. D., Freeberg, M. A., Harrison, P. W., Martin, F. J., & Yates, A. D. (2025, Jan 6). Ensembl 2025. *Nucleic Acids Res*, 53(D1), D948-D957. <https://doi.org/10.1093/nar/gkae1071>
- El Rouby, N., Shahin, M. H., Bader, L., Khalifa, S. I., & Elewa, H. (2022, Feb). Genomewide association analysis of warfarin dose requirements in Middle Eastern and North African populations. *Clin Transl Sci*, 15(2), 558-566. <https://doi.org/10.1111/cts.13176>
- Eljilany, I., Elewa, H., & Al-Badriyeh, D. (2023, Jun). Cost-Benefit Analysis of Genotype-Guided Interruption Days in Warfarin Pre-Procedural Management. *Curr Probl Cardiol*, 48(6), 101128. <https://doi.org/10.1016/j.cpcardiol.2022.101128>
- Elliott, K. S., Haber, M., Daggag, H., Busby, G. B., Sarwar, R., Kennet, D., Petraglia, M., Petherbridge, L. J., Yavari, P., Heard-Bey, F. U., Shobi, B., Ghulam, T., Haj, D., Al Tikriti, A., Mohammad, A., Antony, S., Alyileili, M., Alaydaroos, S., Lau, E., Butler, M., Yavari, A., Knight, J. C., Ashrafian, H., & Barakat, M. T. (2022, Mar 2). Fine-Scale Genetic Structure in the United Arab Emirates Reflects Endogamous and

- Consanguineous Culture, Population History, and Geography. *Mol Biol Evol*, 39(3). <https://doi.org/10.1093/molbev/msac039>
- Evans, W. E., & Relling, M. V. (1999, Oct 15). Pharmacogenomics: translating functional genomics into rational therapeutics. *Science*, 286(5439), 487-491. <https://doi.org/10.1126/science.286.5439.487>
- Gaedigk, A., Ingelman-Sundberg, M., Miller, N. A., Leeder, J. S., Whirl-Carrillo, M., Klein, T. E., & PharmVar Steering, C. (2018, Mar). The Pharmacogene Variation (PharmVar) Consortium: Incorporation of the Human Cytochrome P450 (CYP) Allele Nomenclature Database. *Clin Pharmacol Ther*, 103(3), 399-401. <https://doi.org/10.1002/cpt.910>
- Gaedigk, A., Sangkuhl, K., Whirl-Carrillo, M., Twist, G. P., Klein, T. E., Miller, N. A., & PharmVar Steering, C. (2019, Jan). The Evolution of PharmVar. *Clin Pharmacol Ther*, 105(1), 29-32. <https://doi.org/10.1002/cpt.1275>
- Genomes Project, C., Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., Korbel, J. O., Marchini, J. L., McCarthy, S., McVean, G. A., & Abecasis, G. R. (2015, Oct 1). A global reference for human genetic variation. *Nature*, 526(7571), 68-74. <https://doi.org/10.1038/nature15393>
- Goljan, E., Abouelhoda, M., ElKalioby, M. M., Jabaan, A., Alghithi, N., Meyer, B. F., & Monies, D. (2022). Identification of pharmacogenetic variants from large scale next generation sequencing data in the Saudi population. *PLoS One*, 17(1), e0263137. <https://doi.org/10.1371/journal.pone.0263137>
- Gouveia, M. H., Bentley, A. R., Leal, T. P., Tarazona-Santos, E., Bustamante, C. D., Adeyemo, A. A., Rotimi, C. N., & Shriner, D. (2023, Nov 7). Unappreciated subcontinental admixture in Europeans and European Americans and implications for genetic epidemiology studies. *Nat Commun*, 14(1), 6802. <https://doi.org/10.1038/s41467-023-42491-0>
- Haber, M., Doumet-Serhal, C., Scheib, C., Xue, Y., Danecek, P., Mezzavilla, M., Youhanna, S., Martiniano, R., Prado-Martinez, J., Szpak, M., Matisoo-Smith, E., Schutkowski, H., Mikulski, R., Zalloua, P., Kivisild, T., & Tyler-Smith, C. (2017, Aug 3). Continuity and Admixture in the Last Five Millennia of Levantine History from Ancient Canaanite and Present-Day Lebanese Genome Sequences. *Am J Hum Genet*, 101(2), 274-282. <https://doi.org/10.1016/j.ajhg.2017.06.013>

Haber, M., Gauguier, D., Youhanna, S., Patterson, N., Moorjani, P., Botigue, L. R., Platt, D. E., Matisoo-Smith, E., Soria-Hernanz, D. F., Wells, R. S., Bertranpetit, J., Tyler-Smith, C., Comas, D., & Zalloua, P. A. (2013). Genome-wide diversity in the levant reveals recent structuring by culture. *PLoS Genet*, 9(2), e1003316. <https://doi.org/10.1371/journal.pgen.1003316>

Initiative, C.-H. G. (2021, Dec). Mapping the human genetic architecture of COVID-19. *Nature*, 600(7889), 472-477. <https://doi.org/10.1038/s41586-021-03767-x>

Initiative, C.-H. G. (2023, Sep). A second update on mapping the human genetic architecture of COVID-19. *Nature*, 621(7977), E7-E26. <https://doi.org/10.1038/s41586-023-06355-3>

Ismail, S. I., Saad, C., Elrayess, M. A., Khleifat, A. A. L., Hendi, N. N., AlSarraj, Y., Messai Badji, R., & Muftah, W. (2026, Jul 7). Mapping the path to clinical implementation of multi-omics. *Nat Genet*. <https://doi.org/10.1038/s41588-026-02663-2>

Jarrar, Y., Hakooz, N., Hashem, D. A., Homsy, M., Ajeen, S. A., Ghedan, M. A., Almasri, R., Bdier, A., & Alraqad, M. (2025, Oct 18). Genome-wide screening of pharmacogenomic biomarkers in jordanian patients with genetic disorders. *Pharmacogenomics J*, 25(6), 28. <https://doi.org/10.1038/s41397-025-00387-w>

Jithesh, P. V., Abuhaliqa, M., Syed, N., Ahmed, I., El Anbari, M., Bastaki, K., Sherif, S., Umlai, U. K., Jan, Z., Gandhi, G., Manickam, C., Selvaraj, S., George, C., Bangarusamy, D., Abdel-Latif, R., Al-Shafai, M., Tatari-Calderone, Z., Estivill, X., Pirmohamed, M., & Qatar Genome Program Research, C. (2022, Feb 15). A population study of clinically actionable genetic variation affecting drug response from the Middle East. *NPJ Genom Med*, 7(1), 10. <https://doi.org/10.1038/s41525-022-00281-5>

Johnson, B., Ouyang, K., Frank, L., Truty, R., Rojahn, S., Morales, A., Aradhya, S., & Nykamp, K. (2022, Sep). Systematic use of phenotype evidence in clinical genetic testing reduces the frequency of variants of uncertain significance. *Am J Med Genet A*, 188(9), 2642-2651. <https://doi.org/10.1002/ajmg.a.62779>

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alfoldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., England, E. M., Seaby, E. G., Kosmicki, J. A., Walters, R. K., Tashman, K., Farjoun, Y., Banks, E., Poterba, T., Wang, A., Seed, C., Whiffin, N., Chong, J. X., Samocha, K. E., Pierce-Hoffman, E.,

- Zappala, Z., O'Donnell-Luria, A. H., Minikel, E. V., Weisburd, B., Lek, M., Ware, J. S., Vittal, C., Armean, I. M., Bergelson, L., Cibulskis, K., Connolly, K. M., Covarrubias, M., Donnelly, S., Ferreira, S., Gabriel, S., Gentry, J., Gupta, N., Jeandet, T., Kaplan, D., Llanwarne, C., Munshi, R., Novod, S., Petrillo, N., Roazen, D., Ruano-Rubio, V., Saltzman, A., Schleicher, M., Soto, J., Tibbetts, K., Tolonen, C., Wade, G., Talkowski, M. E., Genome Aggregation Database, C., Neale, B. M., Daly, M. J., & MacArthur, D. G. (2021, Feb). Author Correction: The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, 590(7846), E53.
<https://doi.org/10.1038/s41586-020-03174-8>
- Keeling, N. J., Rosenthal, M. M., West-Strum, D., Patel, A. S., Haidar, C. E., & Hoffman, J. M. (2019, May). Preemptive pharmacogenetic testing: exploring the knowledge and perspectives of US payers. *Genet Med*, 21(5), 1224-1232.
<https://doi.org/10.1038/gim.2017.181>
- Klein, T. E., & Ritchie, M. D. (2018, Jul). PharmCAT: A Pharmacogenomics Clinical Annotation Tool. *Clin Pharmacol Ther*, 104(1), 19-22.
<https://doi.org/10.1002/cpt.928>
- Li, B., Sangkuhl, K., Keat, K., Whaley, R. M., Woon, M., Verma, S., Dudek, S., Tuteja, S., Verma, A., Whirl-Carrillo, M., Ritchie, M. D., & Klein, T. E. (2023, May). How to Run the Pharmacogenomics Clinical Annotation Tool (PharmCAT). *Clin Pharmacol Ther*, 113(5), 1036-1047. <https://doi.org/10.1002/cpt.2790>
- Lima, J. J., Thomas, C. D., Barbarino, J., Desta, Z., Van Driest, S. L., El Rouby, N., Johnson, J. A., Cavallari, L. H., Shakhnovich, V., Thacker, D. L., Scott, S. A., Schwab, M., Uppugunduri, C. R. S., Formea, C. M., Franciosi, J. P., Sangkuhl, K., Gaedigk, A., Klein, T. E., Gammal, R. S., & Furuta, T. (2021, Jun). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2C19 and Proton Pump Inhibitor Dosing. *Clin Pharmacol Ther*, 109(6), 1417-1423.
<https://doi.org/10.1002/cpt.2015>
- Loh, P. R., Danecek, P., Palamara, P. F., Fuchsberger, C., Y, A. R., H, K. F., Schoenherr, S., Forer, L., McCarthy, S., Abecasis, G. R., Durbin, R., & A, L. P. (2016, Nov). Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet*, 48(11), 1443-1448. <https://doi.org/10.1038/ng.3679>
- Maillard, M., Schwab, M., Whirl-Carrillo, M., Moyer, A. M., Suarez-Kurtz, G., Pui, C. H., Stein, C. M., Klein, T. E., Spahn, C., Kwon, S., Hartono, J. L., de Boer, N. K., Ahmad, T., Antillon-Klussmann, F. G., Caudle, K. E., Kato, M., Yeoh, A. E. J., Schmiegelow, K., &

- Yang, J. J. (2026, Apr). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for Thiopurine Dosing Based on TPMT and NUDT15 Genotypes: 2025 Update. *Clin Pharmacol Ther*, 119(4), 916-927. <https://doi.org/10.1002/cpt.70209>
- Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., & Chen, W. M. (2010, Nov 15). Robust relationship inference in genome-wide association studies. *Bioinformatics*, 26(22), 2867-2873. <https://doi.org/10.1093/bioinformatics/btq559>
- Moriyama, B., Obeng, A. O., Barbarino, J., Penzak, S. R., Henning, S. A., Scott, S. A., Agundez, J., Wingard, J. R., McLeod, H. L., Klein, T. E., Cross, S. J., Caudle, K. E., & Walsh, T. J. (2017, Jul). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guidelines for CYP2C19 and Voriconazole Therapy. *Clin Pharmacol Ther*, 102(1), 45-51. <https://doi.org/10.1002/cpt.583>
- Nemer, G., & Hendi, N. N. (2023). Pharmacogenomics of Cardiovascular Diseases: The Path to Precision Therapy. In M. Khullar, A. Mittal, & A. Patil (Eds.), *Pharmacogenomics and Pharmacogenetics in Drug Therapy*. IntechOpen. <https://doi.org/10.5772/intechopen.113236>
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I., Daly, M. J., & Sham, P. C. (2007, Sep). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*, 81(3), 559-575. <https://doi.org/10.1086/519795>
- Relling, M. V., & Klein, T. E. (2011, Mar). CPIC: Clinical Pharmacogenetics Implementation Consortium of the Pharmacogenomics Research Network. *Clin Pharmacol Ther*, 89(3), 464-467. <https://doi.org/10.1038/clpt.2010.279>
- Rodriguez-Flores, J. L., Fakhro, K., Agosto-Perez, F., Ramstetter, M. D., Arbiza, L., Vincent, T. L., Robay, A., Malek, J. A., Suhre, K., Chouchane, L., Badii, R., Al-Nabet Al-Marri, A., Abi Khalil, C., Zirie, M., Jayyousi, A., Salit, J., Keinan, A., Clark, A. G., Crystal, R. G., & Mezey, J. G. (2016, Feb). Indigenous Arabs are descendants of the earliest split from ancient Eurasian populations. *Genome Res*, 26(2), 151-162. <https://doi.org/10.1101/gr.191478.115>
- Sanguhl, K., Whirl-Carrillo, M., Whaley, R. M., Woon, M., Lavertu, A., Altman, R. B., Carter, L., Verma, A., Ritchie, M. D., & Klein, T. E. (2020, Jan). Pharmacogenomics Clinical Annotation Tool (PharmCAT). *Clin Pharmacol Ther*, 107(1), 203-210. <https://doi.org/10.1002/cpt.1568>

Swen, J. J., Nijenhuis, M., de Boer, A., Grandia, L., Maitland-van der Zee, A. H., Mulder, H., Rongen, G. A., van Schaik, R. H., Schalekamp, T., Touw, D. J., van der Weide, J., Wilffert, B., Deneer, V. H., & Guchelaar, H. J. (2011, May). Pharmacogenetics: from bench to byte--an update of guidelines. *Clin Pharmacol Ther*, 89(5), 662-673. <https://doi.org/10.1038/clpt.2011.34>

Swen, J. J., van der Wouden, C. H., Manson, L. E., Abdullah-Koolmees, H., Blagec, K., Blagus, T., Bohringer, S., Cambon-Thomsen, A., Cecchin, E., Cheung, K. C., Deneer, V. H., Dupui, M., Ingelman-Sundberg, M., Jonsson, S., Joefield-Roka, C., Just, K. S., Karlsson, M. O., Konta, L., Koopmann, R., Kriek, M., Lehr, T., Mitropoulou, C., Rial-Sebbag, E., Rollinson, V., Roncato, R., Samwald, M., Schaeffeler, E., Skokou, M., Schwab, M., Steinberger, D., Stingl, J. C., Tremmel, R., Turner, R. M., van Rhenen, M. H., Davila Fajardo, C. L., Dolzan, V., Patrinos, G. P., Pirmohamed, M., Sunder-Plassmann, G., Toffoli, G., Guchelaar, H. J., & Ubiquitous Pharmacogenomics, C. (2023, Feb 4). A 12-gene pharmacogenetic panel to prevent adverse drug reactions: an open-label, multicentre, controlled, cluster-randomised crossover implementation study. *Lancet*, 401(10374), 347-356. [https://doi.org/10.1016/S0140-6736\(22\)01841-4](https://doi.org/10.1016/S0140-6736(22)01841-4)

van der Wouden, C. H., Cambon-Thomsen, A., Cecchin, E., Cheung, K. C., Davila-Fajardo, C. L., Deneer, V. H., Dolzan, V., Ingelman-Sundberg, M., Jonsson, S., Karlsson, M. O., Kriek, M., Mitropoulou, C., Patrinos, G. P., Pirmohamed, M., Samwald, M., Schaeffeler, E., Schwab, M., Steinberger, D., Stingl, J., Sunder-Plassmann, G., Toffoli, G., Turner, R. M., van Rhenen, M. H., Swen, J. J., Guchelaar, H. J., & Ubiquitous Pharmacogenomics, C. (2017, Jul). CORRIGENDUM: Implementing Pharmacogenomics in Europe: Design and Implementation Strategy of the Ubiquitous Pharmacogenomics Consortium. *Clin Pharmacol Ther*, 102(1), 152. <https://doi.org/10.1002/cpt.725>

Van Driest, S. L., Shi, Y., Bowton, E. A., Schildcrout, J. S., Peterson, J. F., Pulley, J., Denny, J. C., & Roden, D. M. (2014, Apr). Clinically actionable genotypes among 10,000 patients with preemptive pharmacogenomic testing. *Clin Pharmacol Ther*, 95(4), 423-431. <https://doi.org/10.1038/clpt.2013.229>

Velayutham, D., Bastaki, K., Irfan, A., Abuhaliqa, M., AlMulla, A., Qatar Genome Program Research, C., Ghuloum, S., Azeem, M. W., Pirmohamed, M., & Jithesh, P. V. (2025). Clinically actionable pharmacogenomic landscape of antidepressants and antipsychotics in Qatar: a population-based cohort study. *Camb Prism Precis Med*, 3, e4. <https://doi.org/10.1017/pcm.2025.2>

Whirl-Carrillo, M., McDonagh, E. M., Hebert, J. M., Gong, L., Sangkuhl, K., Thorn, C. F., Altman, R. B., & Klein, T. E. (2012, Oct). Pharmacogenomics knowledge for personalized medicine. *Clin Pharmacol Ther*, 92(4), 414-417.
<https://doi.org/10.1038/clpt.2012.96>

Wigginton, J. E., Cutler, D. J., & Abecasis, G. R. (2005, May). A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet*, 76(5), 887-893.
<https://doi.org/10.1086/429864>

Yehya, A., & Altaany, Z. (2021, Oct). A Decade of Pharmacogenetic Studies in Jordan: A Systemic Review. *Pharmacogenomics J*, 21(5), 543-550.
<https://doi.org/10.1038/s41397-021-00236-6>

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTables.xlsx](#)
- [SupplementaryFigures.docx](#)