Specialty Journal of Pharmacognosy, Phytochemistry, and Biotechnology
ISSN: 3062-441X
2021, Volume 1, Page No: 112-122

Available online at: www.galaxypub.co/page/journals

Galaxy Publication

Galaxy Publication

Pharmacogenomics and Ancestral Genetics in Colombia: An Analysis of All PharmGKB Drug Variant Annotations

David Wilson[1], Andrew Foster[1]*, Kevin Brooks[1], Daniel Hart[1]

[1]Department of Natural Products Chemistry, School of Pharmacy, University of Mississippi, Oxford, USA.

***E-mail** ✉ andrew.foster.np@yahoo.com

Received: 05 May 2021; **Revised:** 16 August 2021; **Accepted:** 16 August 2021

ABSTRACT

This study aimed to map Colombia's pharmacogenomic (PGx) landscape with respect to ancestry by combining all PharmGKB variant-drug annotations with local allele-frequency data, in order to measure clinically relevant differences among ancestral groups and identify gaps that limit fair access to precision medicine. We analyzed 4,462 PharmGKB variant annotations spanning 1994–2024, narrowing the dataset to 1,216 significant single-nucleotide polymorphisms (SNPs) reported in 552 studies. Allele frequencies were compiled for five Colombian populations: two largely African-descended (Palenque [PLQ], Chocó [CHG]) and three predominantly European-descended (ATQCES, ATQPGC, CLM) using the CÓDIGO database. Spearman correlation coefficients assessed similarities between population-specific PGx profiles, and SNPs with frequency differences exceeding 25 percentage points were cataloged. The global PGx literature is heavily skewed toward European ancestry, representing 51.5% of 651,532 participants, while African ancestry accounted for only 0.46% (n = 3,031). European-leaning Antioquian populations showed strong internal correlations ($r^2 \geq 0.90$), whereas PLQ displayed weak or inverse correlations with these groups ($r^2 = -0.20$ to -0.02) and only minimal similarity with CHG ($r^2 = 0.12$). Among the SNPs, 28 were highly frequent in PLQ (>75%) but uncommon in Europeans (<50%), while 44 exhibited the reverse pattern. Notable examples include CYP3A4 rs3735451-C (rivaroxaban; 87.1% vs. 23.2%), CYP3A5 rs776746-T (tacrolimus; 85% vs. 23.5%), and rs55881666-C (duloxetine; 15% vs. 84%). Globally, 71.5% of PGx studies originated in high-income nations. The substantial allele-frequency disparities and pronounced research biases underscore the need for ancestry-informed PGx testing and locally tailored dosing strategies in Colombia, while the study's analytic framework and variant catalog provide a foundation for implementing precision pharmacotherapy in Latin-American admixed populations.

Keywords: Colombia, Pharmacogenomic variants, genomics, Precision medicine, Pharmacogenetics

How to Cite This Article: Wilson D, Foster A, Brooks K, Hart D. Pharmacogenomics and Ancestral Genetics in Colombia: An Analysis of All PharmGKB Drug Variant Annotations. Spec J Pharmacogn Phytochem Biotechnol. 2021;1:112-22. https://doi.org/10.51847/3wHgDGEBl1

Introduction

The concept of precision pharmacotherapy—matching medications and dosing to an individual's genetic profile—has shifted from theoretical promise to an essential component of clinical care, especially as non-communicable diseases continue to rise globally [1]. Central to this approach, pharmacogenomics (PGx) uncovers how inherited genetic variation influences drug absorption, metabolism, efficacy, and adverse reactions [2]. However, the practical application of PGx findings remains uneven because the majority of variant–drug associations were derived from populations of primarily European descent, even though allele frequencies and linkage patterns differ considerably across ancestral backgrounds [3]. These differences have meaningful clinical implications: even minor shifts in variant prevalence can affect the cost-effectiveness of pharmacogenomic testing and impact overall population health outcomes [4].

Regions such as Latin America, sub-Saharan Africa, and parts of South-East Asia, which carry high disease burdens, are particularly underrepresented in PGx research, perpetuating therapeutic disparities and reducing the relevance of dosing guidelines, decision-support systems, and risk stratification tools for these populations [5].

Recognizing this inequity, global initiatives have emphasized the importance of diversity in genomics research. The Global Alliance for Genomics and Health, the WHO Genomic Medicine Implementation framework, and the U.S. Precision Medicine Initiative all highlight the need to extend pharmacogenomic discovery and clinical translation to low- and middle-income countries (LMICs) [6-8].

Specialist organizations—including CPIC [9], PharmVar [10], and PharmGKB [11]—have similarly called for ancestry-aware pharmacogenomic recommendations, noting that the lack of local allele-frequency data compromises the reliability and safety of current guidelines [9-11]. Evidence further shows that ancestry bias hinders optimization of multi-drug therapies, confounds comparative-effectiveness research, and complicates the estimation of variant penetrance in admixed populations, underscoring the need for population-specific PGx data from under-studied regions [9-11].

Colombia serves as a compelling case study. Centuries of admixture among Indigenous, European, and African populations have produced complex genomic ancestries that vary across small geographic regions, from the predominantly African-descended population of San Basilio de Palenque to the largely European-descended communities of Antioquia [12]. Emerging national resources, such as the CÓDIGO database, provide allele-frequency data reflecting this diversity [13]. Yet, these data have not been systematically linked to curated variant–drug annotations. Consequently, clinicians often rely on extrapolated frequency data from foreign populations, risking inaccurate classification of metabolizer status and inappropriate drug dosing. These limitations hinder patient safety and reduce health-system efficiency, particularly as electronic health records increasingly integrate pre-emptive pharmacogenomic testing [14].

A comprehensive, ancestry-stratified PGx map for Colombia could deliver immediate clinical and public-health benefits. Locally calibrated screening panels could anticipate adverse drug reactions in high-burden areas such as psychiatry and oncology, while genetic counseling for monogenic pharmacogenetic disorders (e.g., dihydropyrimidine dehydrogenase deficiency) and cascade testing could be more effectively guided [15]. At the population level, ancestry-specific allele frequencies could inform essential-medicine formularies, target cost-effective pharmacogenomic interventions, improve trial design, support pharmacovigilance, and align national policy with global precision-medicine standards [16].

This study integrates all PharmGKB variant–drug annotations with allele-frequency data from five Colombian subpopulations to (i) construct a detailed pharmacogenomic landscape by ancestry, (ii) identify variants with the largest inter-ancestry frequency differences and direct clinical relevance, and (iii) contextualize global research output through bibliometric analysis. By filling a key knowledge gap in an admixed LMIC setting, this work aims to support the equitable implementation of pharmacogenomics and provide a framework applicable to other underrepresented populations.

Materials and Methods

Study design

This research employed a cross-sectional analysis of pharmacogenomic variant annotations.

Data collection

All available variant–drug associations were retrieved from PharmGKB on March 5, 2025. Two datasets were downloaded: one containing all variant–drug associations, including effects on dosing, drug response, and metabolism, and another with study-level metadata, such as cohort sizes and biogeographical classifications. After merging these datasets, the initial compilation included 4,462 variant annotations from 1,225 studies.

Variants were filtered according to three criteria: (i) single-nucleotide polymorphisms (SNPs) with a single reported allele, excluding genotypes and complex variants; (ii) statistically significant associations ($p < 0.05$); and (iii) availability of allele-frequency data. Allele frequencies for five Colombian populations were obtained from the CÓDIGO database using automated Python queries [13]. These populations included two with predominantly African ancestry—Palenque (PLQ, n = 34, 84% African ancestry) and Chocó (CHG, n = 96, 76% African ancestry)—and three predominantly European populations from Antioquia—ATQCES (n = 404, 50.5% European), ATQPGC (n = 624, 55% European), and CLM (n = 96, 62.9% European). Ancestry proportions were calculated using ADMIXTURE, with full admixture profiles available in the CÓDIGO database.

Allele frequencies for PLQ and ATQPGC were determined via whole-genome sequencing, while the remaining populations were analyzed through whole-exome sequencing or genome-wide genotyping. The CÓDIGO

database (release 1.0) contained 1,441 Colombian samples across 14 populations, encompassing 95,254,482 non-redundant variants from eight independent datasets. Following filtering, the final dataset included 1,216 variant annotations reported across 552 studies, which formed the foundation for all subsequent analyses.

Statistical analysis

We analyzed the curated set of 1,216 pharmacogenomic variant annotations by computing Spearman correlation coefficients across populations using reported allele frequencies. Variants lacking frequency information were omitted from the relevant pairwise comparisons. To assess ancestry-specific patterns, we calculated the mean allele frequency for each ancestry group: the PLQ population represented African ancestry, while the European ancestry estimate was derived from the average of ATQCES, ATQPGC, and CLM populations. Additionally, participants from the 1,225 source studies were classified as either cases or controls based on the biogeographical ancestry provided in the original publications, with individuals assigned to European, Asian, American, African, Other, or Unknown categories.

Bibliometric analysis

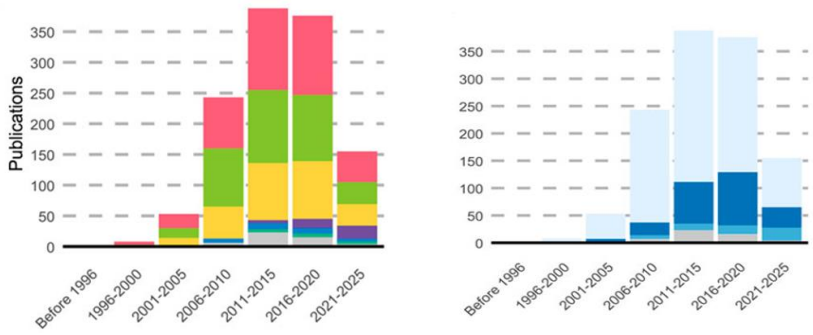
To contextualize population-level allele-frequency differences within the global pharmacogenomic research landscape, metadata for all 1,225 studies was retrieved using PubMed IDs through the NCBI Entrez API, Unpaywall API, and Crossref API. Information collected included article title, publication date, number of authors, first author's country, citation counts, open-access status, and journal-specific metrics (H-index and quartile) extracted from the SCImago Journal Rank database corresponding to the year of publication (1999–2024). First author affiliations were further mapped to WHO regions [17] and World Bank income classifications [18] to allow descriptive analyses of study origin and resource context.

By integrating allele-frequency data with bibliometric metadata, we were able to relate observed genetic variability in Colombian populations to the broader evidence base, highlighting the alignment between local frequency contrasts and the clinical significance of variant–drug associations. All analyses were executed in R (v4.4.0), and the code, datasets, and detailed annotations are publicly available at https://doi.org/10.5281/zenodo.15361131.

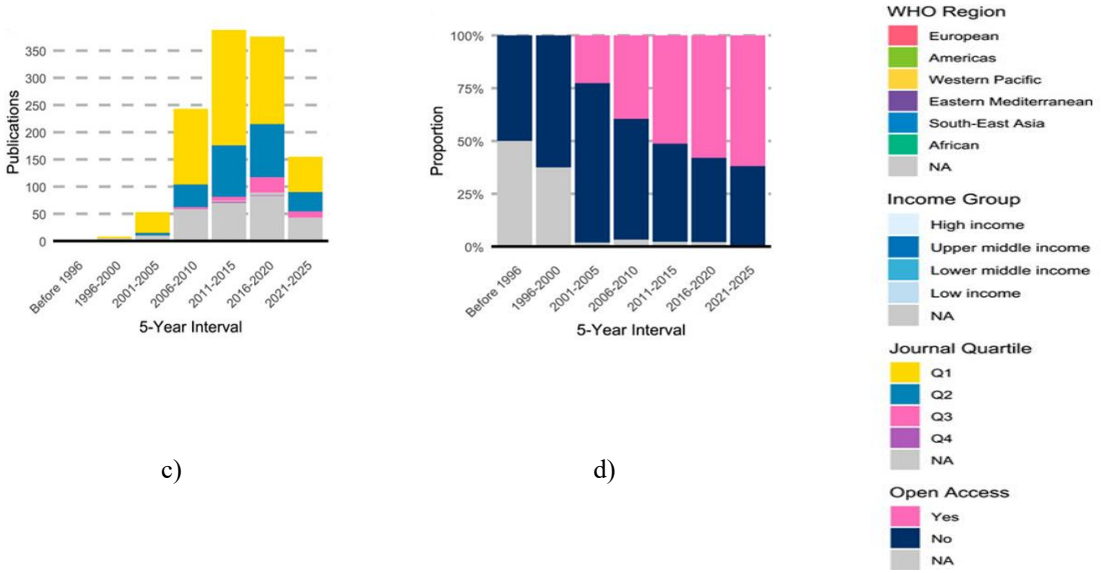
Results and Discussion

Overview of study characteristics

Across the initial dataset, 4,462 pharmacogenomic variants were reported in 1,225 studies, representing the full spectrum of PharmGKB variant–drug annotations. The earliest study, published in 1994 in Pharmacogenetics and Genomics, documented that the T allele was associated with reduced warfarin metabolism relative to the C allele [19]. Geographically, studies were primarily concentrated in Europe (34.5%) and the Americas (30.7%) **(Figure 1a; Table 1)**, and most were conducted in high-income countries (71.5%; **Figure 1b; Table 2**). Citation patterns reflected this distribution, with 27,351 citations (38.8%) from the Americas, 25,501 (36.2%) from Europe, and 61,432 (87.2%) from high-income nations. Open-access prevalence was highest in the Americas (ratio 2.34) and among high-income countries (1.19) **(Tables 1 and 2)**. Regarding journal quality, 65.3% of publications appeared in Q1-ranked journals **(Figure 1c)**, yet more than half (50.7%) of all articles were not freely accessible **(Figure 1d).**



a) b)



c) d)

Figure 1. Temporal patterns and attributes of PharmGKB-indexed pharmacogenomics studies (n = 1,225). (a) Number of publications stratified by WHO region in 5-year intervals. (b) Breakdown of studies according to World Bank income categories. (c) Distribution of articles across journal quartiles (Q1–Q4). (d) Proportion of studies available as open access over time. NA: not available.

Table 1. Baseline characteristics of publications on pharmacogenomics by region (N = 1,225).

Variable	Americas	Europe	Western Pacific	South-East Asia	Eastern Mediterranean	Africa	Not retrieved
Publications (%)	376 (30.7)	424 (34.6)	289 (23.9)	31 (2.53)	41 (3.35)	16 (1.31)	48 (3.92)
Total citations (mean per paper)	27,351 (72.7)	25,501 (60.1)	13,334 (46.1)	669 (21.6)	337 (8.22)	394 (24.6)	2,855 (59.5)
Median H-index (IQR)	154 (145)	154 (114)	118 (80)	126 (78.5)	103 (81.2)	154 (61.8)	152 (70.5)
Open access/No open access (ratio)	262/112 (2.34)	190/ 223 (0.85)	105/173 (0.61)	10/20 (0.5)	20/17 (1.18)	7/9 (0.78)	27/19 (1.42)
Journal quartile (%) (n = 946)							
Q1	205 (71.18)	209 (67.42)	139 (57.92)	12 (52.17)	10 (34.48)	10 (71.43)	33 (78.57)
Q2	75 (26.04)	86 (27.74)	86 (35.83)	10 (43.48)	13 (44.83)	4 (28.57)	2 (4.76)
Q3	8 (2.78)	14 (4.52)	14 (5.83)	0	6 (20.69)	0	6 (14.29)
Q4	0	1 (0.32)	1 (0.42)	1 (4.35)	0	0	1 (2.38)

Table 2. Baseline characteristics of publications on pharmacogenomics by income group (N = 1,225).

Variable	High income	Upper-middle income	Lower-middle income	Low- income	Not retrieved
Publications (%)	876 (71.5)	239 (19.5)	59 (4.82)	3 (0.25)	48 (3.9)
Total citations (mean per paper)	61,432 (87.5)	5,067 (7.19)	938 (1.33)	149 (0.21)	2,855 (4.05)
Median H-index (IQR)	154 (114)	106 (69.8)	126 (80.5)	140 (14)	125 (70.5)
Open access/No open access (ratio)	468/394 (1.19)	100/126 (0.79)	24/33 (0.73)	2/ 1 (2)	27/19 (1.42)
Journal quartile (%) (n = 946)					
Q1	460 (70.4)	104 (50.2)	19 (45.2)	2 (100)	33 (78.6)
Q2	175 (26.8)	83 (40.1)	16 (38.1)	0	2 (4.8)

Q3	18 (2.8)	18 (8.7)	6 (14.3)	0	6 (14.3)
Q4	0	2 (1)	1 (2.4)	0	1 (2.4)

Influence of genetic ancestry on pharmacogenomic variation in colombia

Analysis of the 1,216 filtered variant annotations **(Figure 2a)** revealed clear ancestry-dependent patterns in Colombian populations. Spearman correlation coefficients were very high among populations with predominantly European ancestry (ATQCES, ATQPGC, and CLM), all exceeding 0.9, indicating highly concordant allele-frequency profiles across these groups **(Figure 2b).** By contrast, the population with the largest African ancestry fraction, San Basilio de Palenque, displayed strikingly divergent pharmacogenomic profiles compared with the European-leaning populations (ATQCES: $r^2 = -0.20$; ATQPGC: $r^2 = -0.06$; CLM: $r^2 = -0.02$), each of which contains less than 13% African ancestry. San Basilio de Palenque also showed minimal similarity to Chocó ($r^2 = 0.12$), a population with predominantly African ancestry but higher European admixture (11.5%). The Chocó population, in turn, exhibited moderate positive correlations with the European-ancestry groups (ATQCES: $r^2 = 0.79$; ATQPGC: $r^2 = 0.75$; CLM: $r^2 = 0.80$), reflecting partial concordance in pharmacogenomic profiles.

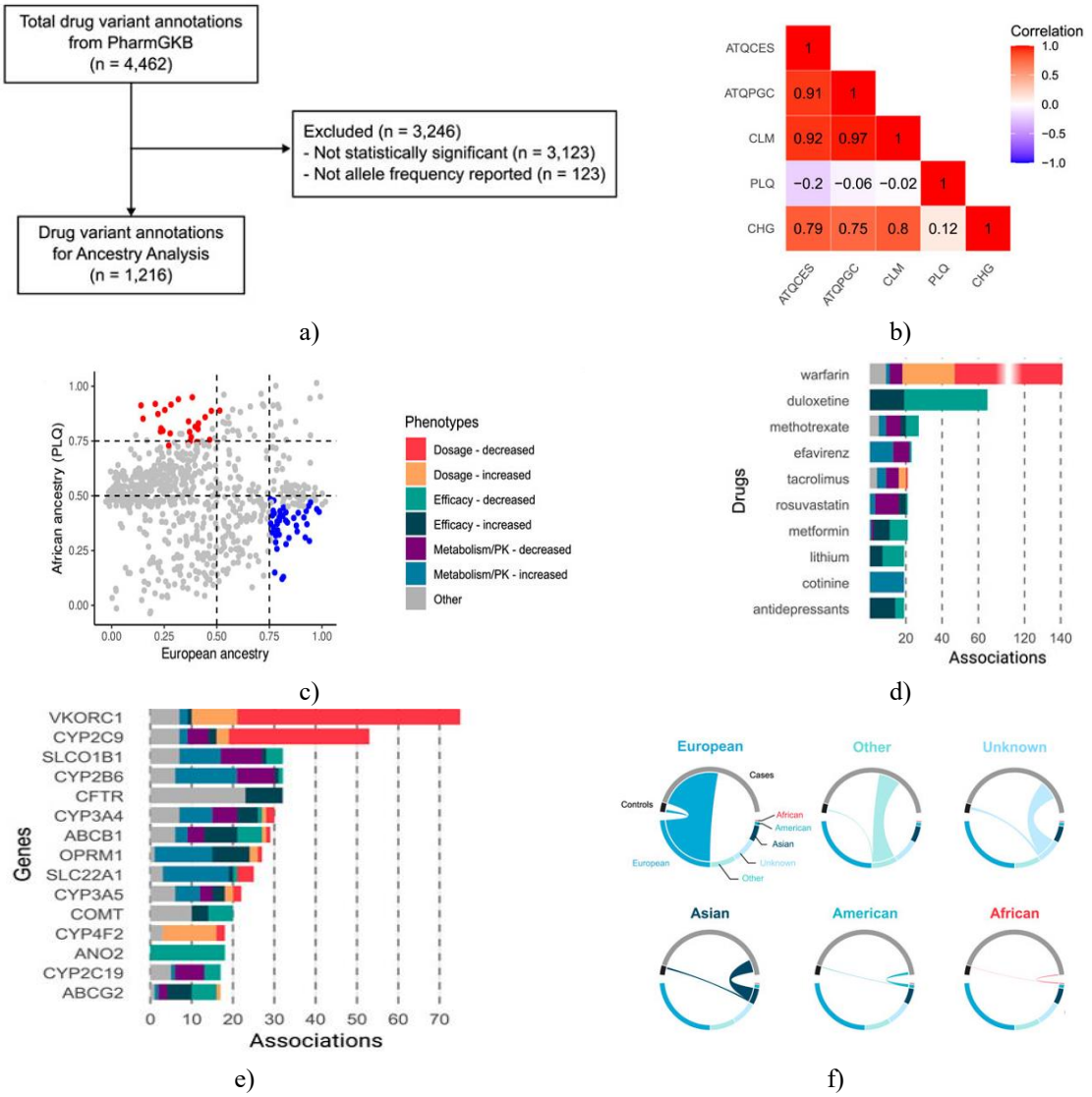


Figure 2. Patterns of ancestry, variant selection, and distribution of drug- and gene-level associations in Colombian populations. (a) Schematic of the filtering workflow applied to 4,462 PharmGKB drug variant annotations, yielding 1,216 variants for analysis. (b) Heatmap illustrating Spearman correlation values for allele frequencies across five Colombian populations: ATQCES, ATQPGC, CLM, PLQ (San Basilio de Palenque), and CHG (Chocó). (c) Comparison of allele frequencies in PLQ (y-axis, African ancestry) versus

the averaged allele frequency of ATQCES, ATQPGC, and CLM (x-axis, European ancestry) for all significant SNPs. (d) Ten drugs with the largest number of reported variant associations, classified by phenotype. (e) Fifteen genes with the highest number of documented associations, color-coded according to phenotype category. (f) Chord diagrams depicting the ancestry composition of participants in the 1,225 pharmacogenomic studies, separated by ancestry category and study role (case vs. control).

Due to the distinct pharmacogenomic landscape of San Basilio de Palenque, we focused on identifying alleles that were especially frequent in this population relative to predominantly European-ancestry groups. To facilitate this comparison, we calculated the mean allele frequency across ATQCES, ATQPGC, and CLM—populations chosen for their strong pairwise similarity—and plotted these values against the frequencies observed in PLQ for all significant SNPs **(Figure 2c).** This approach revealed 28 SNPs that were highly prevalent in PLQ (allele frequency >75%) but rare in European populations (mean frequency <50%), while 44 SNPs displayed the reverse trend, being common in Europeans (mean frequency >75%) and less frequent in PLQ (allele frequency <50%). The majority of these variants were associated with drug efficacy or pharmacokinetic/metabolism traits. **Table 3** highlights a subset of SNPs with the largest inter-population frequency differences.

Table 3. Pharmacogenomic variants with differential allele frequencies by ancestry and their associated phenotype and drug effects.

Gene	SNP	Drug	Phenotype and direction of effect	PLQ (AF%)	European (AF%)	Reference
SNPs most predominant in San Basilio de Palenque						
CYP3A4	rs3735451-C	Rivaroxaban	Metabolism/PK increased	87.1	23.23	Li *et al*. (2024)[20]
CYP3A5	rs776746-T	Tacrolimus	Dosage, Metabolism/PK increased	85	23.5	Kim *et al*. (2012)[21]
CDKAL1	rs7754840-C	DPP-4 inhibitors	Efficacy increased	76.1	28.4	Osada *et al*. (2016)[22]
SNPs most predominant in European populations						
Not retrieved	rs55881666-C	Duloxetine	Efficacy decreased	15	84	Maciukiewicz *et al*. (2018)[23]
PTPRC	rs10919563-G	Adalimumab, etanercept, infliximab	Efficacy increased	28	75.6	Cui *et al*. (2010)[24]
ABCB1	rs7787082-G	Clozapine	Efficacy decreased	29	79.3	Lee *et al*. (2012) [25]

AF, allele frequency; PLQ, san basilio de palenque; SNPs, Single Nucleotide Polymorphisms.

In the PLQ population, certain SNPs showed markedly higher frequencies compared to European populations. For instance, rs3735451-C, which is linked to elevated rivaroxaban levels in atrial fibrillation patients [20], was observed in 87.1% of PLQ individuals, whereas only 23.2% of Europeans carried this allele. Similarly, rs776746-T, associated with accelerated tacrolimus metabolism in kidney transplant recipients [21], was found in 85% of PLQ participants but just 23.5% in Europeans. Rs7754840-C, connected to enhanced response to DPP-4 inhibitors in diabetes [22], occurred in 76.1% of PLQ individuals, contrasting with 28.4% in European populations **(Table 3).**

Conversely, several variants were more frequent in European populations. Rs55881666-C, which correlates with reduced duloxetine efficacy in major depressive disorder [23], was present in 84% of Europeans but only 15% of PLQ participants. Rs10919563-G, linked to improved responses to biologics such as adalimumab, etanercept, or infliximab in rheumatoid arthritis [24], was observed in 75.6% of Europeans versus 28% of PLQ individuals. Rs7787082-G, associated with lower clozapine response in schizophrenia [25], appeared in 79.3% of Europeans compared to 29% in PLQ **(Table 3).**

Global focus in pharmacogenomics

We further explored which drugs and genes dominate the pharmacogenomic literature. The ten drugs with the largest number of reported significant associations **(Figure 2d)** accounted for 30.9% of all curated links. Warfarin

was the most frequently studied, representing 11.5% of associations, followed by duloxetine (5.3%) and methotrexate (2.2%). Most commonly reported drug-related phenotypes involved reduced dosing requirements (25.2%) and diminished efficacy (21.8%). Among genes, 36.7% of all associations involved the top 15 most-studied genes **(Figure 2e),** with VKORC1 leading at 6.1%, followed by CYP2C9 (4.3%) and SLCO1B1, CYP2B6, and CFTR (each 2.6%). Importantly, several highly represented genes—CYP3A4, CYP3A5, VKORC1, and SLCO1B1—also displayed major allele-frequency differences between Colombian populations, reinforcing their clinical significance for locally tailored pharmacogenomics.

Predominance of European ancestry in PGx research

Across the 1,225 studies analyzed, a total of 651,532 participants were included, comprising 609,280 cases (93.6%) and 42,252 controls (6.4%). Individuals of European ancestry constituted the largest proportion (n = 336,073; 51.5%) **(Figure 2f).** Participants of Asian and Native/admixed American ancestry made up 72,359 (11.1%) and 14,470 (2.22%), respectively, while only 3,031 participants (0.46%) were of African ancestry. The remaining individuals either had unreported ancestry (n = 110,073; 16.8%) or were classified as other (n = 115,526; 17.7%).

This work presents the first detailed, ancestry-stratified pharmacogenomic map at the variant level for Colombia, offering one of the clearest demonstrations of how global PGx discovery often fails to translate effectively to local, highly admixed populations. By merging 1,216 statistically significant PharmGKB annotations with allele-frequency data from five Colombian sub-populations, we examined three intersecting layers of evidence—bibliometric trends, ancestry representation, and population genetics—to assess whether current global knowledge adequately supports clinical decision-making in a tri-hybrid Latin American context [12].

The distribution of PGx research remains strongly skewed toward wealthier nations. Among the 1,225 publications underpinning all PharmGKB variant–drug associations, 71.5% originated from high-income countries, and 65.3% appeared in top-quartile journals, yet only half were openly accessible, with the lowest availability seen in lower-middle-income settings. Even though the Americas contributed nearly one-third of studies, the locations of first authors and citation impact continue to reflect a traditional North–South gradient [26]. While this pattern aligns with previous bibliometric analyses showing dominance by resource-rich regions, the magnitude of the disparity emphasizes a persistent implementation gap: variants most likely to inform clinical decisions in LMICs remain the least explored or validated locally [4, 5]. A 2.34 open-access ratio in the Americas suggests regional efforts to reduce paywall barriers, yet restricted access continues to limit national pharmacovigilance and formulary decision-making [27].

Ancestry-related bias continues to be a critical obstacle in translating PGx research into practice [3]. Across 651,532 participant records, 51.5% were of European ancestry, compared with only 0.46% of African and 2.22% of Native/admixed American ancestry. For a country like Colombia, shaped by centuries of tri-hybrid admixture [12], this imbalance has direct clinical implications: applying dosing guidance derived from European-centric cohorts carries unmeasured risk. Additionally, the fact that 17% of individuals had unreported ancestry represents a preventable source of misclassification that could propagate through meta-analyses and guideline development. Although the direction of bias is predictable, its persistence into 2025, despite calls for global equity in research [6-8], indicates that conventional funding and peer-review incentives alone are insufficient to correct disparities.

Fine-scale correlation analyses reveal a bifurcated Colombian PGx landscape. The three Antioquian populations (ATQCES, ATQPGC, CLM) showed near-identical allele-frequency profiles ($r2 \geq 0.90$), consistent with the expectation that shared founder histories produce highly similar PGx patterns. In contrast, the San Basilio de Palenque (PLQ) population, which has the highest African ancestry, exhibited weakly negative or negligible correlations with European-leaning groups ($r2 = -0.20$ to -0.02) and minimal concordance with Chocó, another Afro-descendant population ($r2 = 0.12$).

The low correlation between PLQ and Chocó was somewhat unexpected given their shared African heritage, suggesting that founder effects, differential admixture, and localized selection pressures have shaped distinct allele distributions. Clinically, this highlights the risk of assuming homogeneity among "African-ancestry" populations; intra-continental variation can rival inter-continental differences at pharmacologically actionable loci [28].

These population-specific differences have tangible clinical implications. We identified 28 SNPs common in PLQ but rare among European Colombians, and 44 with the opposite pattern. For example, rs3735451-C in CYP3A4, linked to increased rivaroxaban exposure, was present in 87.1% of PLQ residents versus 23.2% of European Colombians, potentially affecting bleeding risk [29]. Similarly, rs776746-T in CYP3A5, associated with faster

tacrolimus metabolism, occurred in 85% of PLQ individuals compared to 23.5% in Europeans, reinforcing the need for dose adjustments in Afro-descendant transplant recipients [30]. Conversely, rs55881666-C, which reduces duloxetine efficacy, was four times more common in Europeans (84% vs. 15%), highlighting potential risks of applying European-derived dosing strategies to PLQ patients [31].

The divergence between global research priorities and local clinical needs is also evident. Drugs such as warfarin, duloxetine, and methotrexate dominate the PGx literature, together comprising nearly one-fifth of all significant associations, while VKORC1 and CYP2C9 are the most frequently studied genes. Although warfarin remains relevant, newer medications like rivaroxaban now dominate anticoagulant therapy in Colombia, and duloxetine does not represent the bulk of antidepressant use [32]. This mismatch reinforces critiques that pharmacogenomics has largely focused on "variant discovery" rather than aligning research with pressing local clinical demands ("clinical-need hunting") [33].

Even variants showing partial correlation between Chocó and European populations ($r2 = 0.75–0.80$) are informative, revealing intermediate allele sharing due to historical admixture. Such patterns caution against oversimplified ancestry classifications and argue for incorporating continuous local ancestry or probabilistic genotype imputation in decision-support tools. Additionally, two-thirds of PharmGKB variants identified in Colombia did not meet the $p < 0.05$ threshold, emphasizing that many associations remain population-specific or underpowered, and underscoring the need for replication studies [34].

Collectively, these findings reinforce the importance of ancestry-informed PGx panels to enable equitable precision medicine [35]. For health-technology assessments, the allele-frequency contrasts documented here can refine cost-effectiveness calculations, optimizing the number-needed-to-genotype and number-needed-to-treat [36]. For regulatory agencies, the overrepresentation of European participants highlights the need for post-marketing surveillance in admixed populations, especially for drugs with narrow therapeutic indices. Finally, national research bodies may leverage bibliometric insights to negotiate open-access policies, facilitating the dissemination of locally relevant pharmacogenomic knowledge.

Research gaps and opportunities

The underrepresentation of African-ancestry populations, even within Latin America, signals a clear need for collaborative research with Afro-Colombian communities, focusing on study designs co-developed to overcome historical mistrust. Pharmacogenomic trials that account for detailed local ancestry could provide models for other countries with highly mixed populations. Linking electronic health records to prospective biobanks in Colombia would allow adaptive healthcare systems where allele-frequency data continuously informs clinical decision-making [36].

To speed up translation into practice, national health authorities and research institutions could develop ancestry-informed dosing guidelines, implement PGx alerts in digital health systems, and fund studies targeting high-risk groups such as Afro-Colombians. Engaging underrepresented communities and co-designing protocols is essential for ethical and sustainable integration.

Advancing precision medicine

By comparing global pharmacogenomic evidence with Colombia's unique genetic landscape, this work highlights critical data gaps and presents an analytic approach that other LMICs could adopt. The study provides actionable variant lists ready for testing, informs guideline committees about ancestry-specific effect sizes, and offers epidemiologic data for payers to support reimbursement decisions [37-41]. Overall, this approach facilitates the transition from standardized treatment strategies toward ancestry-aware therapies, improving efficacy, safety, and cost-efficiency across the region [26, 42].

Limitations and next steps

This study was designed as a meta-research exercise rather than a reevaluation of pharmacogenomic effect sizes. Only variants from peer-reviewed studies reporting clearly defined clinical outcomes were included. By integrating these findings with allele-frequency data from the CÓDIGO database, ancestry-specific variants with potential clinical relevance were highlighted. The results should be interpreted as hypothesis-generating, and prospective validation using individual-level genomic and longitudinal clinical data in Colombian cohorts is necessary to confirm clinical impact.

Conclusion

This work delivers the first ancestry-resolved pharmacogenomic map of Colombia, merging PharmGKB variant–drug annotations with allele-frequency data from five distinct sub-populations. It shows that global PGx evidence is predominantly European, leaving key loci under-characterized in admixed and African-ancestry Colombians. Clinically relevant allele-frequency differences were quantified—for instance, the CYP3A5*3 enhancer rs776746-T occurs four times more often in San Basilio de Palenque than in European-leaning Antioquia—indicating a clear need for ancestry-informed dosing of medications such as tacrolimus, rivaroxaban, and duloxetine. Bibliometric analyses reveal persistent barriers to access and geographic inequities that limit LMIC researchers' ability to replicate and implement PGx findings.

These insights equip policymakers and payers with the necessary information—allele frequencies, effect sizes, and knowledge gaps—to evaluate the cost-effectiveness of pre-emptive genotyping and prioritize high-impact tests. Regulators can use the variant catalogue to refine pharmacovigilance and enforce post-marketing monitoring in high-risk populations. For researchers, underrepresented variants and populations are highlighted as priorities for replication studies, multicenter trials, and guideline development. This framework is readily transferable to other admixed LMICs, supporting a more equitable global precision medicine agenda.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. Hicks JK, Aquilante CL, Dunnenberger HM, Gammal RS, Funk RS, Aitken SL, et al. Precision pharmacotherapy: integrating pharmacogenomics into clinical pharmacy practice. J Am Coll Clin Pharm. 2019;2(3):303–13. doi:10.1002/jac5.1118
2. Cecchin E, Stocco G. Pharmacogenomics and personalized medicine. Genes (Basel). 2020;11(6):679. doi:10.3390/genes11060679
3. Corpas M, Siddiqui MK, Soremekun O, Mathur R, Gill D, Fatumo S. Addressing ancestry and sex bias in pharmacogenomics. Annu Rev Pharmacol Toxicol. 2024;64:53–64. doi:10.1146/annurev-pharmtox-030823-111731
4. Karamperis K, Katz S, Melograna F, Ganau FP, Van Steen K, Patrinos GP, et al. Genetic ancestry in population pharmacogenomics unravels distinct geographical patterns related to drug toxicity. iScience. 2024;27(10):110916. doi:10.1016/j.isci.2024.110916
5. Zhou Y, Park Y, Camara MD, Lauschke VM. Opportunities and challenges of population pharmacogenomics. Ann Hum Genet. 2025;89:384–97. doi:10.1111/ahg.12596
6. Ambrosino E, Abou Tayoun AN, Abramowicz M, Zilfalil BA, Boughtwood T, Hamdi Y, et al. The WHO genomics program of work for equitable implementation of human genomics for global health. Nat Med. 2024;30(10):2711–3. doi:10.1038/s41591-024-03225-x
7. National health institutes. the precision medicine initiative. Available from: https://www.nih.gov/sites/default/files/research-training/initiatives/pmi/pmi-infographic.pdf (Accessed May 11, 2025)
8. The global alliance for genomics and health. Strategic road map. Available from: https://www.ga4gh.org/about-us/strategic-road-map/ (Accessed May 11, 2025)
9. Relling MV, Klein TE. CPIC: clinical pharmacogenetics implementation consortium of the pharmacogenomics research network. Clin Pharmacol Ther. 2011;89(3):464–7. doi:10.1038/clpt.2010.279
10. Gaedigk A, Casey ST, Whirl-Carrillo M, Miller NA, Klein TE. Pharmacogene variation consortium: a global resource and repository for pharmacogene variation. Clin Pharmacol Ther. 2021;110(3):542–5. doi:10.1002/cpt.2321

11. Thorn CF, Klein TE, Altman RB. PharmGKB: the pharmacogenomics knowledge base. Methods Mol Biol. 2013;1015:311–20. doi:10.1007/978-1-62703-435-7_20
12. Nagar SD, Moreno AM, Norris ET, Rishishwar L, Conley AB, O'Neal KL, et al. Population pharmacogenomics for precision public health in Colombia. Front Genet. 2019;10:241. doi:10.3389/fgene.2019.00241
13. Mariño-Ramírez L, Sharma S, Hamilton JM, Nguyen TL, Gupta S, Natarajan AV, et al. The consortium for genomic diversity, ancestry, and health in Colombia (CÓDIGO): building local capacity in genomics and bioinformatics. Commun Biol. 2025;8(1):1–13. doi:10.1038/s42003-025-08496-9
14. Chande AT, Rishishwar L, Ban D, Nagar SD, Conley AB, Rowell J, et al. The phenotypic consequences of genetic divergence between admixed Latin American populations: antioquia and chocó, Colombia. Genome Biol Evol. 2020;12(9):1516–27. doi:10.1093/gbe/evaa154
15. Riess O, Sturm M, Menden B, Liebmann A, Demidov G, Witt D, et al. Genomes in clinical care. NPJ Genom Med. 2024;9(1):20. doi:10.1038/s41525-024-00402-2
16. Gemmati D, Varani K, Bramanti B, Piva R, Bonaccorsi G, Trentini A, et al. Bridging the gap everything that could have been avoided if we had applied gender medicine, pharmacogenetics and personalized medicine in the gender-omics and sex-omics era. Int J Mol Sci. 2019;21(1):296. doi:10.3390/ijms21010296
17. World Health Organization. Countries/areas by WHO region. Available from: https://apps.who.int/violence-info/Countries%20and%20areas%20by%20WHO%20region%20-%2012bfe12.pdf (Accessed May 11, 2025)
18. World Bank. World Bank Country and Lending Groups Washington D.C. Available from: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups (Accessed May 11, 2025)
19. Rettie AE, Wienkers LC, Gonzalez FJ, Trager WF, Korzekwa KR. Impaired (S)-warfarin metabolism catalysed by the R144C allelic variant of CYP2C9. Pharmacogenet Genomics. 1994;4(1):39–42. doi:10.1097/00008571-199402000-00005
20. Li X, Gu Z, Wang Z, Xu Q, Ma C, Lv Q. Mutant CYP3A4/5 correlated with clinical outcomes by affecting rivaroxaban pharmacokinetics and pharmacodynamics in patients with atrial fibrillation. Cardiovasc Drugs Ther. 2024;38(6):1315–25. doi:10.1007/s10557-023-07495-4
21. Kim I, Noh H, Ji E, Han N, Hong SH, Ha J, et al. Identification of factors affecting tacrolimus level and 5-year clinical outcome in kidney transplant patients. Basic Clin Pharmacol Toxicol. 2012;111(4):217–23. doi:10.1111/j.1742-7843.2012.00892.x
22. Osada UN, Sunagawa H, Terauchi Y, Ueda S. A common susceptibility gene for type 2 diabetes is associated with drug response to a DPP-4 inhibitor: pharmacogenomic cohort in Okinawa, Japan. PLOS ONE. 2016;11(5):e0154821. doi:10.1371/journal.pone.0154821
23. Maciukiewicz M, Marshe VS, Hauschild AC, Foster JA, Rotzinger S, Kennedy JL, et al. GWAS-based machine learning approach to predict duloxetine response in major depressive disorder. J Psychiatr Res. 2018;99:62–8. doi:10.1016/j.jpsychires.2017.12.009
24. Cui J, Saevarsdottir S, Thomson B, Padyukov L, van der Helm-van Mil AHM, Nititham J, et al. Rheumatoid arthritis risk allele PTPRC is also associated with response to anti-tumor necrosis factor alpha therapy. Arthritis Rheumatol. 2010;62(7):1849–61. doi:10.1002/art.27457
25. Lee ST, Ryu S, Kim SR, Kim MJ, Kim S, Kim JW, et al. Association study of 27 annotated genes for clozapine pharmacogenetics: validation of preexisting studies and identification of a new candidate gene, ABCB1, for treatment response. J Clin Psychopharmacol. 2012;32(4):441–8. doi:10.1097/JCP.0b013e31825ac35c
26. Lozada-Martinez ID, Lozada-Martinez LM, Cabarcas-Martinez A, Ruiz-Gutierrez FK, Aristizabal Vanegas JG, Amorocho Lozada KJ, et al. Historical evolution of cancer genomics research in Latin America: a comprehensive visual and bibliometric analysis until 2023. Front Genet. 2024a;15:1327243. doi:10.3389/fgene.2024.1327243
27. Else H. Open access is working - but researchers in lower-income countries enjoy fewer benefits. Nature. 2024. doi:10.1038/d41586-024-01748-4
28. Munung NS. Science and society: pathways to equitable access and delivery of genomics medicine in Africa. Curr Genet Med Rep. 2025;13(1):1. doi:10.1007/s40142-024-00211-0

29. Baturina O, Chashkina M, Andreev D, Mirzaev K, Bykova A, Suvorov A, et al. Pharmacokinetic and pharmacogenetic predictors of major bleeding events in patients with an acute coronary syndrome and atrial fibrillation receiving combined antithrombotic therapy. J Pers Med. 2023;13(9):1371. doi:10.3390/jpm13091371
30. Genvigir FDV, Campos-Salazar AB, Felipe CR, Tedesco-Silva H Jr, Medina-Pestana JO, Doi SQ, et al. CYP3A53 and CYP2C83 variants influence exposure and clinical outcomes of tacrolimus-based therapy. Pharmacogenomics. 2020;21(1):7–21. doi:10.2217/pgs-2019-0120
31. Maciukiewicz M, Marshe VS, Tiwari AK, Fonseka TM, Freeman N, Rotzinger S, et al. Genetic variation in IL-1β, IL-2, IL-6, TSPO, and BDNF and response to duloxetine or placebo treatment in major depressive disorder. Pharmacogenomics. 2015;16(17):1919–29. doi:10.2217/pgs.15.136
32. Machado-Duque ME, Machado-Alba JE, Gaviria-Mendoza A, Valladales-Restrepo LF, Parrado-Fajardo IY, Ospina-Castellanos M, et al. Identification of medication errors through a monitoring and minimization program in outpatients in Colombia, 2018-2019. Biomedica. 2021;41(1):79–86. doi:10.7705/biomedica.5544
33. El-Gowilly SM, Metwaly HA, Makhlouf D, Elmansoury N, Abuiessa SA, Sorour AA, et al. Analysis of the current situation of pharmacogenomics in terms of educational and healthcare needs in Egypt and Lebanon. Pharmacogenomics. 2024;25(10–11):429–40. doi:10.1080/14622416.2024.2403967
34. Feres M, Feres MFN. Absence of evidence is not evidence of absence. J Appl Oral Sci. 2023;31:ed001. doi:10.1590/1678-7757-2023-ed001
35. Nagar SD, Conley AB, Jordan IK. Population structure and pharmacogenomic risk stratification in the United States. BMC Biol. 2020;18(1):140. doi:10.1186/s12915-020-00875-4
36. Picón-Jaimes YA. Innovation and digital transformation in health education: opportunities to drive technological development in the training of future professionals. Inge CUC. 2024m;20(2):99–105. doi:10.17981/ingecuc.20.2.2024.10
37. Tamraz B, Shin J, Khanna R, Van Ziffle J, Knowles S, Stregowski S, et al. Clinical implementation of preemptive pharmacogenomics testing for personalized medicine at an academic medical center. J Am Med Inform Assoc. 2025;32(3):566–71. doi:10.1093/jamia/ocae293
38. Lozada-Martinez ID, Lozada-Martinez LM, Fiorillo-Moreno O. Leiden manifesto and evidence-based research: are the appropriate standards being used for the correct evaluation of pluralism, gaps and relevance in medical research? J R Coll Physicians Edinb. 2024b;54(1):4–6. doi:10.1177/14782715241227991
39. Lozada-Martinez ID, Neira-Rodado D, Martinez-Guevara D, Cruz-Soto HS, Sanchez-Echeverry MP, Liscano Y. Why is it important to implement meta-research in universities and institutes with medical research activities? Front Res Metr Anal. 2025a;10:1497280. doi:10.3389/frma.2025.1497280
40. Lozada-Martinez ID, Hernandez-Paez DA, Fiorillo-Moreno O, Picón-Jaimes YA, Bermúdez V. Meta-research in biomedical investigation: gaps and opportunities based on meta-research publications and global indicators in health. Sci Hum Dev Publ. 2025b;13(1):7. doi:10.3390/publications13010007
41. Lozada-Martinez ID, Hernandez-Paez D, Jiménez Zárate YE, Delgado P. Scientometrics and meta-research in medical research: approaches required to ensure scientific rigor in an era of massive low-quality research. Rev Assoc Med Bras. 2025c;71(4):e20241612. doi:10.1590/1806-9282.20241612
42. Pérez-Fontalvo NM, De Arco-Aragón MA, Jimenez-García JDC, Lozada-Martinez ID. Molecular and computational research in low- and middle-income countries: development is close at hand. J Taibah Univ Med Sci. 2021;16(6):948–9. doi:10.1016/j.jtumed.2021.06.010