UNIVERSITY OF KWAZULU-NATAL ™

INYUVESI YAKWAZULU-NATALI

Influence of Drug Transporter Gene Polymorphisms on Pharmacokinetics and Treatment Outcomes in African *Mycobacterium tuberculosis* Cohorts

By

Sahil Tulsi

Student No: 200309440

B.Sc. (Microbiology and Physiology), B.MedSci (Hons) (Medical Science), M.Med.Sc. (Medical Science)

Submitted in fulfilment for the degree of Doctor of Philosophy (Health Sciences – Virology),

School of Medicine, College of Health Sciences, University of KwaZulu-Natal

Supervisor
Professor Veron Ramsuran

December 2025

Declaration

I, Sahil Tulsi, declare that:

i. The research reported in this thesis, except where otherwise indicated, is my own original work.
ii. This thesis has not been submitted for any degree or examination at another university.
iii. This thesis does not contain other persons' data, pictures, graphs, or other information, unless specifically acknowledged as being sourced from other persons.
iv. This thesis does not contain other person's writing, unless specifically acknowledged as being sourced from other researchers.

Where other written sources have been quoted, then:
a) their words have been re-written, but the general information sourced has been referenced to the authors.
b) where their exact words have been used, their writing has been placed within quotation marks and referenced.

v. Where I have reproduced a publication of which I am an author, co-author, or editor, I have indicated in detail which part of the publication was written by myself alone and have fully referenced such publications.

vi. This thesis does not contain text, graphs, or tables copied and pasted from the internet, unless specifically referenced or acknowledged, and the source being detailed in the thesis and references section.

Signature:  Date: 1st December 2025

Mr. Sahil Tulsi

Dedication

This work in this PhD thesis is purely dedicated to the countless individuals across the African continent who continue to face the burden of tuberculosis.

It is my sincere hope that this research represents a small but meaningful contribution to the broader scientific effort aimed at understanding the genetic diversity of African populations and its implications for more effective, personalized treatment strategies.

By adding one piece to this extremely complex puzzle, I aim for this research and findings to aid in the support of the development of therapies that are not only scientifically sound but also equitable and responsive to the unique needs of African patients in need of unique treatment.

Lastly to all those striving to make science more inclusive and impactful for African patients, we share the same passion and end goal.

Acknowledgements

Mrs. Nazley Tulsi

You have been my pillar of strength until the end. Thank you for your unwavering support throughout this academic journey. I appreciate every sacrifice you have made so that I can make my contribution to science. During the most challenging moments your support has been a source of strength and motivation which brings me to this memorable time of completion. This achievement would not have been possible without your faith in me. I am truly grateful, with all my love and appreciation.

Professor Veron Ramsuran

You have been an inspiring leader, your constant drive pushes a student to deliver their best in every aspect and reflect their true potential. I would like to express my sincere gratitude for your exceptional leadership and unwavering support throughout the course of my degree. Your guidance, insight, and dedication have been instrumental to my academic success and personal growth, and it indeed has been a privilege to work under your supervision.
Thank you once again for your invaluable mentorship.

My Family & Friends

Thank you for your guidance throughout the years, your motivation and positive advice has been pivotal to reaching this point of submission.

The Ramsuran Lab

Thank you to a group of perfectly well-balanced team players who constantly support one another. I would like to extend my heartfelt thanks to all of you for your support, collaboration, and camaraderie throughout my doctoral journey. I am truly grateful to have been a part of such a dedicated and supportive team. Wishing you all continued success in your work and academic futures.

Funding

I would like to sincerely thank Novartis and GSK as well as Prof Veron Ramsuran
for providing funding throughout the duration of this study.

Publications

1. **Tulsi, S.,** Arumugam, T., Pramchand, I., Ramsuran, V,. 2025 Genetic diversity of polymorphisms within specific drug transporter genes in African individuals infected with Mycobacterium tuberculosis.
   In submission in : *Journal of Clinical Tuberculosis and Other Mycobacterial Diseases*
   Manuscript ID: JCTUBE-D-25-00260

2. **Tulsi, S.,** Gokul, A., Naidoo, A., Naidoo K., Singh, R., Swe Swe Han, K., Dlamini, N., Mahabeer, Y., Ramsuran, V., 2025 Potential implications of genetic variants in *ABCB1* and *SLC01B1* on moxifloxacin pharmacokinetics in South African patients infected with tuberculosis.
   In submission: *British Journal of Clinical Pharmacology*
   Manuscript ID: ecda1f83-b823-4073-b393-70ed6e64b7d3

3. **Tulsi, S.,** Sharma, PP., Gokul, A., Naidoo, A., Naidoo, K., Ramsuran, V., 2025 Associations of understudied variants in *Mycobacterium tuberculosis* infected African patients with HIV-1 infection, TB recurrence and Drug PK levels receiving anti-TB therapy by genome-wide association study.
   In submission in: *Clinical Genetics*
   Manuscript ID: 5becc035-8a98-4134-b08f-e8f176ef89b6

4. **Tulsi, S**., Sharma PP., Gokul A., Naidoo A., Naidoo K., Ramsuran V., 2025
   Associations of *ABC & SLC* gene family drug transporter variants and understudied SNPs with moxifloxacin drug PK levels in *Mycobacterium tuberculosis* infected African patients by genome-wide association study
   In submission in: *Pharmacogenetics and Genomics*
   Manuscript ID: PGEN-2025-096

5. Arumugam T., Naidoo A., **Tulsi S**., Ramsuran V., 2025
   Exploring the Human Leukocyte Antigen (HLA) Landscape in Tuberculosis-Affected South African Populations: Genetic Associations with TB Infection, HIV Co-infection, and Recurrence. *To be Submitted*

6. Sharaf A, Nesengani LT, Hayah I, Kuja JO, Mdyogolo S, Omotoriogun TC, Odogwu BA, Beedessee G, Smith RM, Barakat A, Moila AM, El Hamouchi A, Benkahla A, Boukteb A, Elmouhtadi A, Mafwila AL, Abushady AM, Elsherif AK, Ahmed B, Wairuri C, Ndiribe CC, Ebuzome C, Kinnear CJ, Ndlovu DF, Iraqi D, El Fahime E, Assefa E, Ouardi F, Belharfi FZ, Tmimi FZ, Markey FB, Radouani F, Zeukeng F, Mvumbi GL, Ganesan H, Hanachi M, Nigussie H, Charoute H, Benamri I, Mkedder I, Haddadi I, Meftah-Kadmiri I, Mubiru JF, Domelevo Entfellner JK, Rokani JB, Ogwang J, Daiga JB, Omumbo J, Ideozu JE, Errafii K, Labuschagne K, Komi KK, Tonfack LB, Hadjeras L, Ramantswana M, Chaisi M, Botes MW, Kilian M, Kvas M, Melloul M, Chaouch M, Khyatti M, Abdo M, Phasha-Muchemenye M, Hijri M, Mediouni MR, Hassan MA, Piro M, Mwale M, Maaloum M, Mavhunga M, Olivier NA, Aminou O, Arbani O, Souiai O, Djocgoue PF, Mentag R, Zipfel RD, Tata RB, Megnekou R, Muzemil S, Paez S, Salifu SP, Kagame SP, Selka S, Edwards S, Gaouar SBS, Reda SRA, Fellahi S, Khayi S, Ayed S, Madisha T, **Sahil T**, Udensi OU, Ras V, Ezebuiro V, Duru VC, David X, Geberemichael Y, Tchiechoua YH, Mungloo-Dilmohamud Z, Chen Z, Happi C, Kariuki T, Ziyomo C, Djikeng A, Badaoui B, Mapholi N, Muigai A, Osuji JO, Ebenezer TE. Establishing African genomics and bioinformatics programs through annual regional workshops. **Nat Genet. 2024 Aug;56(8):1556-1565. doi: 10.1038/s41588-024-01807-6. Epub 2024 Jul 8. PMID: 38977855.**

Table of Contents

Declaration ii

Dedication iii

Acknowledgements iv

Publications v

List of Figures viii

List of Tables x

Abbreviations/ Acronyms xii

Abstract xiv

Chapter 1 1

Introduction 1

Chapter 2 8

Genetic diversity of polymorphisms within specific drug transporter genes in African individuals infected with Mycobacterium tuberculosis 9

Chapter 3 25

Potential implications of genetic variants in ABCB1 and SLC01B1 on moxifloxacin pharmacokinetics in South African patients infected with tuberculosis. 26

Chapter 4 50

Associations of understudied variants in Mycobacterium tuberculosis infected African patients with HIV-1 predisposition, TB infection and recurrence by genome-wide association study. 51

Chapter 5 80

Associations of ABC & SLC gene family drug transporter variants and understudied SNPs with moxifloxacin drug PK levels in Mycobacterium tuberculosis infected African patients by genome-wide association study. 81

Chapter 6 114

6.1 Synthesis and conclusion 114

6.2 Limitations and Recommendations 118

Chapter 7 119

7.1 References: 119

Appendix A 122

Appendix B 144

Addendum A: Ethical approval 156

Addendum B: Guidelines for Thesis submission 157

Addendum C: Turnitin Digital Receipt 164

List of Figures

Chapter 1

Figure 1: Illustration representing localization of ABC and SLC transporters involved in pharmacokinetics.	**3**
Figure 2: Illustration showing typical plasma concentration–time curve illustrating the area under the curve (AUC) in pharmacodynamics.	**5**

Chapter 2

Figure 1: Schematic depiction of the spatial distribution of ABC and SLC transporters relevant to drug pharmacokinetics.	**11**

Chapter 3

Figure 1: Genotype frequencies of rs10261685, rs11045819 & rs1045642 in TRUTH patients grouped into recurrence compared to no recurrence subsets with no significant differences observed.	**35**
Figure 2: Allele frequencies of rs10261685, rs1045642, and rs11045819 in TRUTH patients, stratified into recurrence and no-recurrence subsets with no significant differences observed.	**36**
Figure 3: Allele frequencies of the rs10261685, rs1045642, and rs11045819 variants in TRUTH patients, comparing individuals who were co-infected with HIV to those who did not seroconvert.	**37**
Figure 4: Drug pharmacokinetic (PK) analysis of single nucleotide polymorphism (rs10261685, rs1045642, and rs11045819) showing the respective genotypes and corresponding moxifloxacin plasma levels measured at two timepoints [Timepoint 1 (T1) and Timepoint 2 (T2)] with a significant difference observed for rs11045819 at T2 between genotypes CA/CC ($p= 0.0232$, $R^2= 0.1117$).	**38**

Chapter 4

Figure 1: Manhattan plot generated from the genome-wide association study (GWAS) of TB-positive patient samples (n = 459) evaluated for genetic variants associated with HIV-1 infection status by Fisher's exact test for genome-wide significance.	**58**
Figure 2: Manhattan plot generated from the genome-wide association study (GWAS) of TB-positive patient samples (n= 427) assessing genetic variants associated with susceptibility to TB recurrence by Fisher's exact test.	**60**
Figure 3: Manhattan plot illustrating the genome-wide distribution of association signals derived from the GWAS comparing TB-positive patients (n= 985) with TB-negative individuals by Fisher's exact test.	**62**

Chapter 5

Figure 1: Manhattan plot representing genome-wide association results for moxifloxacin AUC levels that illustrates the distribution of association p-values for all genome-wide single nucleotide polymorphisms (SNPs) tested for their relationship with moxifloxacin area-under-the-curve (AUC) pharmacokinetic levels at timepoint 1 (T1).	**87**
Figure 2: Manhattan plot displaying genome-wide association results with distribution of the association p-values with moxifloxacin area-under-the-curve (AUC) pharmacokinetic levels at timepoint 2 (T2).	**92**

List of Tables

Chapter 2

Table 1: Genetic diversity & Minor Allele frequency of common ABC SNPs between Asian, African and European populations.	**14**

Chapter 3

Table 1: Demographics of patients summarized from the CAPRISA clinical trials (TRUTH, IMPRESS and InDEX).	**30**
Table 2: Minor allele frequencies (MAFs) for the genetic variants rs10261685, rs1045642, and rs11045819 are summarized across major global population groups, including Asian and European cohorts, as well as African and African American (Afro-American) populations.	**33**
Table 3: Genotype frequencies for the *ABCB1* gene variants rs10261685 and rs1045642, as well as the SLCO1B1 gene variant rs11045819, obtained from patients enrolled in the TRUTH study.	**34**

Chapter 4

Table 1: Summary of patient samples from the UKZN, TRUTH, and Insight studies included in the genome-wide association (GWAS) analysis.	**53**
Table 2: Summarized tabular representation of the top 10 single nucleotide polymorphisms (SNPs) most significantly associated with HIV-1 infection susceptibility by Fisher's exact test, as identified through genome-wide association study (GWAS) analysis.	**57**
Table 3: Expanded tabular representation of the top 10 single nucleotide polymorphisms (SNPs) most significantly associated with TB recurrence by Fisher's exact test as identified through genome-wide association study (GWAS) analysis.	**59**
Table 4. Comprehensive tabular summary of the top 10 single nucleotide polymorphisms (SNPs) demonstrating the strongest statistical association with	**61**

tuberculosis (TB) infection as identified through genome-wide association study (GWAS) analysis by Fisher's exact test.	

Chapter 5

Table 1: Table Summary presenting an overview of the top 10 single nucleotide polymorphisms (SNPs) within the ATP-binding cassette (ABC) gene family that demonstrated the strongest statistical associations with moxifloxacin area-under-the-curve (AUC) pharmacokinetic levels at timepoint 1 (T1).	**88**
Table 2: Summary of the top ten single nucleotide polymorphisms (SNPs) within the solute carrier (SLC) gene family that demonstrated the strongest associations with moxifloxacin area-under-the-curve (AUC) pharmacokinetic levels at timepoint 1 (T1).	**89**
Table 3: Summary of the top 10 SNPs within ABC gene family that showed the strongest associations with moxifloxacin area-under-the-curve (AUC) levels at T2.	**90**
Table 4: Summary of the top 10 SLC gene family SNPs showing the strongest associations with moxifloxacin area-under-the-curve (AUC) levels at T2.	**93**
Table 5: Common ABC and SLC gene family SNPs that showed statistically significant associations with moxifloxacin area-under-the-curve (AUC) levels at both T1 and T2.	**94**

Abbreviations/ Acronyms

ABC	Adenosine Trisphosphate Binding Cassette
ABCC	ATP-binding cassette subfamily C
ABCG	ATP-binding cassette, subfamily G
ADME	Absorption, Distribution, Metabolism, and Excretion
AFR	African
AfricaBP	Africa Bio-genome Project
AIDS	Acquired immunodeficiency syndrome
ALFA	Allele Frequency Aggregator
AP3B1	Adaptor protein complex-3
ART	Anti-retroviral therapy
ARV	Anti-retroviral
AS	Asian
ATP	Adenosine Triphosphate
AUC	Area under the curve
BRCA	Breast Cancer Gene
CAPRISA	Center for the AIDS Programme of Research in South Africa
CCR5	Chemokine Receptor 5
CD206	C-Lectin 206
CF	Cystic Fibrosis
DNA	Deoxyribonucleic Acid
DMET	Drug Metabolizing Enzymes and Transporters
EAS	East Asian
ETH	Ethambutol
EUR	European
EWAS	Epigenome-Wide Association Studies
FH	Familial Hypercholesterolaemia
GI	Gastrointestinal
GT	Genotype
GWAS	Genome-Wide Association Studies
HAART	Highly Active Antiretroviral Therapy
HIV	Human Immunodeficiency Virus
HLA	Human Leukocyte Antigen
IMPRESS	Improving Retreatment Success
INDEX	Individualized M(X) drug-resistant TB Treatment Strategy Study
INH	Isoniazid
INSTIs	Integrase Strand Transfer Inhibitors
INSIGHT	INSTIs for the management of HIV-associated TB
LNGI	Large National Genomic Initiative
LTBI	Latent TB Infection
LD	Linkage Disequilibrium
MA	Micro Array
MAF	Minor Allele Frequency
MBL	Mannose-Binding Lectin
MDR	Multi-Drug Resistance
MM	Multiple Myeloma

Mtb	*Mycobacterium tuberculosis*
MOXI	Moxifloxacin
NEO1	Neogenin 1
NGS	Next-Generation Sequencing
PBMC	Peripheral Blood Mononuclear Cells
PCR	Polymerase Chain Reaction
P-gp	P-glycoprotein
PGx	Pharmacogenomics
PMDA	Precision Medicine Diversity Array
PONV	Postoperative Nausea and Vomiting
PTPRA	Protein tyrosine phosphatase receptor type A
RIF	Rifampicin
rsID	Reference SNP Identification
RT	Real-Time
RT	Reverse Transcriptase
SA	South Africa
SC	Solute Carrier
SDG	Sustainable Development Goals
SLC	Solute carrier gene
SLC1OA1	Solute Carrier Family 10 Member 1
SNP	Single Nucleotide Polymorphism
SNV	Single Nucleotide Variant
SLOC1B1	Solute Carrier Organic Anion Transporter Family Member 1B1
TB	Tuberculosis
TDM	Therapeutic Drug Monitoring
UN	United Nations
US	United States
VL	Viral Load
VUS	Variant of Uncertain Significance
WES	Whole-exome sequencing
WHO	World Health Organization
WGS	Whole-genome sequencing
XDR	Extensively Drug-Resistant

Abstract

Background: Tuberculosis (TB) remains a major health challenge in Africa, and achieving the United Nations Sustainable Development Goal of ending TB by 2030 may be hindered by limited understanding of how African populations respond to current therapies. Emerging evidence indicates that African patients metabolize anti-TB drugs differently due to unique genetic variants, particularly in drug-transporter genes. Many variants known to affect drug response in Asian and European populations occur at lower frequencies or are absent in African groups, while numerous African-specific or understudied variants remain unexplored. This study examined the role of drug-transporter genes and genome-wide variants with pharmacodynamics, HIV and TB-associated biomarkers among South African patients receiving anti-TB therapy or HAART. Specifically, this study: (1) reviewed the genetic diversity of transporter-gene variants in African individuals with TB; (2) assessed the effects of specific *ABCB1* and *SLCO1B1* mutations on moxifloxacin pharmacokinetics (PK); (3) investigated genome-wide markers associated with HIV/TB co-infection and TB recurrence; and (4) identified ATP-Binding Cassette (ABC), Solute Carriers (SC) and genome-wide variants influencing moxifloxacin pharmacokinetics (PK).

Materials and Methods: A total of 1,407 black race South African participants from various clinical trials from rural and urban KwaZulu-Natal were included for this study across chapters, namely the TB Recurrence upon treatment with HAART study (TRUTH) n= 427, Improving Retreatment Success (IMPRESS) n=157, University of KwaZulu-Natal cohort (UKZN) n=778, Integrase Strand Trasnfer Inhibitors (INSTIs) for The Management of HIV-associated TB study (INSIGHT) n=32 and Individualized M(X) drug-resistant TB Treatment Strategy Study (InDEX), n= 13. TB diagnoses were confirmed through microscopy, PCR, and culture methods. Moxifloxacin pharmacokinetic measurements were obtained approximately 24 hours post-dose using mass spectrometry. *ABCB1* and *SLCO1B1* variants were genotyped by real-time PCR, and genome-wide Single Nucleotide Polymorphisms (SNPs) were genotyped by microarray technology.

Results and Conclusions: Three initial variants rs10261685, rs1045642 (*ABCB1*) and rs11045819 (*SLC01B1*) showed modest associations with moxifloxacin concentrations. The CC genotype of rs11045819 was linked to significantly reduced drug levels (p= 0.0232, R^2=

0.1117) compared to CA genotyped patients, suggesting rapid metabolizer status. These 3 SNPs however showed no significant associations with HIV acquisition, TB infection or recurrence, indicating the need to investigate broader genomic contributions. We then performed a genome-wide association study that identified several variants strongly associated with HIV and HIV/TB co-infected patients. For the purposes of this study we determined Bonferroni correction to account and adjust for multiple testing comparisons and determine genome-wide significance threshold at $p \leq 5.89 \times 10^{-8}$. Variant rs121908780 showed strong genome-wide significance ($p= 8.56 \times 10^{-8}$) and has been associated with cystic fibrosis and rs80357798 ($p= 2.70 \times 10^{-34}$), rs587778855 ($p= 2.98 \times 10^{-34}$) and rs886040490 ($p= 4.39 \times 10^{-34}$) associated with cancer risk in previous studies. Additional variants were associated with TB recurrence with allele frequencies that differed notably from those in non-African populations. For the drug PK analysis several ABC and SC family gene variants showed suggestive associations with moxifloxacin levels. Two SLC intronic variants, rs4414423 ($p= 9.50 \times 10^{-05}$ at T1, $p= 0.019$ at T2) and rs7805940 ($p= 0.0003$ at T1, p= 0.003 at T2), were nominally significant at both time points and associated with elevated drug concentrations. These findings highlight population-specific genetic influences on drug response and underscore the need for larger studies to validate the clinical significance of African-specific variants, particularly solute-carrier intronic mutations affecting anti-TB pharmacokinetics. In order to ensure optimal treatment outcomes in African patients , these future studies can aid in the development of therapeutics that are just as unique as the patients who will receive them.

Chapter 1

Introduction

Background

Tuberculosis (TB) remains one of the leading global health problems with approximately 5–10% latent tuberculosis infections (LTBI) individuals and of which 90–95% individuals will eventually develop active tuberculosis infection (1). The African continent, home to 17% of the world's population accounts for over 25% of Mycobacterium tuberculosis (TB) mortalities globally (2). Active TB infections is the cause of death in approximately 50% of individuals who are untreated, emphasizing the importance of early diagnosis and the administration of effective TB treatment, in addition to prevention in high-risk populations (3). Even with advanced research and development carried out over the past 50 years, TB standard treatment regimens and drugs administered for drug susceptible Tuberculosis (DS-TB) remain unchanged (4). Only recently, new drugs and regimens have been approved for both DS-TB and drug-resistant TB. Rifampicin, isoniazid, and pyrazinamide are the current DS-TB drugs that are used for first-line treatment regimens. Studies have been performed to investigate optimal TB doses in order to determine effective TB drug concentrations at sites of action (2, 5, 6). Several studies have shown first line rifampicin, isoniazid and pyrazinamide drug concentrations are observed at much lower levels in individuals of African descent, which therefore results in compromised treatment outcomes (2, 6, 7). Compromised bactericidal/bacteriostatic activity along with acquired drug non-susceptibility are initially driven by minimal drug concentrations, which often products of pharmacokinetic (PK) variability and are independent of adherence to treatment (8).

Apart from adherence to TB treatment, it is widely understood that patient TB treatment outcomes are often influenced by specific genetic polymorphisms (2). Genes that code for drug-metabolizing and transporter enzymes often harbour single nucleotide polymorphisms (SNPs) along with certain genetic polymorphisms which have been shown to significantly correlate with variable drug concentrations and pharmacokinetic parameters (9). It has been observed in a number of cases that the frequency of SNP alleles showing strong associations with various clinical conditions in European or Asian individuals are not present within individuals of African descent. An example demonstrating this being allele A of rs4148323,

known to be associated with hyperbilirubinemia in Asian populations occurring at 19% minor allele frequency (29), however not present within African individuals as reported in ALFA databases). Studies have shown that the greatest genetic variability across the world has been found in individuals from Africa (10). Genomic databases over-represent European data, whereas data on African genomes account for only 2% of the genomes analysed (11). Considering the previous statement, Africa is a region that is under studied.

The African continent has faced the burden of disease for many years with several outbreaks leading to the widespread of antimicrobial exposure (33). Such an environment can shape the African resistome which represents a dynamic reservoir of resistance genes that potentially exerts selective pressure either on a host or microbial genomes. This environment can potentially influence genetic variants in human drug transporter genes, such as those encoding ATP-binding cassette (ABC) and solute carrier (SLC) proteins, through adaptation to xenobiotic stress by evolution. In the context of anti-tuberculosis therapy, such genetic variability may affect intracellular drug concentrations, particularly for agents like rifampicin, moxifloxacin and isoniazid, leading to disparities in therapeutic efficacy and toxicity profiles. Consequently, the association between the African resistome and host genomic diversity is thus a critical factor underlying observed variability in drug PK among patients, with crucial implications for precision medicine and treatment optimization.

Drug transporter genes are of utmost importance across numerous areas of therapeutics, due to their vital role in processes that regulate both PK properties of drugs (absorption, distribution and elimination) and the development of drug resistance via decreased drug uptake on increase in efflux (9). Soluble carriers and ATP-binding cassette transporters are two of the most commonly studied membrane transporters (9). There are almost 400 individual proteins between these two transporters which have been identified to date. The proteins are widely distributed throughout the human body. Depending on the particular drug transporter, their role in both absorption and efflux of endogenous compounds or xenobiotics provides understanding that they could influence PK of therapeutic drugs used for clinical purposes (Fig. 1). Tuberculosis treatment drugs such as isoniazid is metabolized within the human body by N-acetyltransferase enzymes that are known to be coded by NAT2 genes, which often determine the metabolizer status of an individual (12). The rate of isoniazid clearance is often associated by NAT2 gene metabolic activity (13). In a recent study that reported the prevalence of genotypes (ABCB1, SLCO1B1 and NAT2) and their effect on PK parameters of isoniazid and

rifampicin, it was found that rapid and intermediate metabolizers showed 2.3- and 1.6-times faster clearance of isoniazid in comparison to slow metabolizers (2).

The ATP-binding family consists of 48 genes, with most research focusing on ABCB1 (P-glycoprotein) and ABCG2 (BCRP, MXR and ABCP) which encode transmembrane proteins that are understood to bind and hydrolyze ATP, using this energy to allow various molecules to be transported across the membranes (14-16). The transporter proteins from the ABC family are known to be associated with detoxification of the host and provide protection against xenobiotics (17).

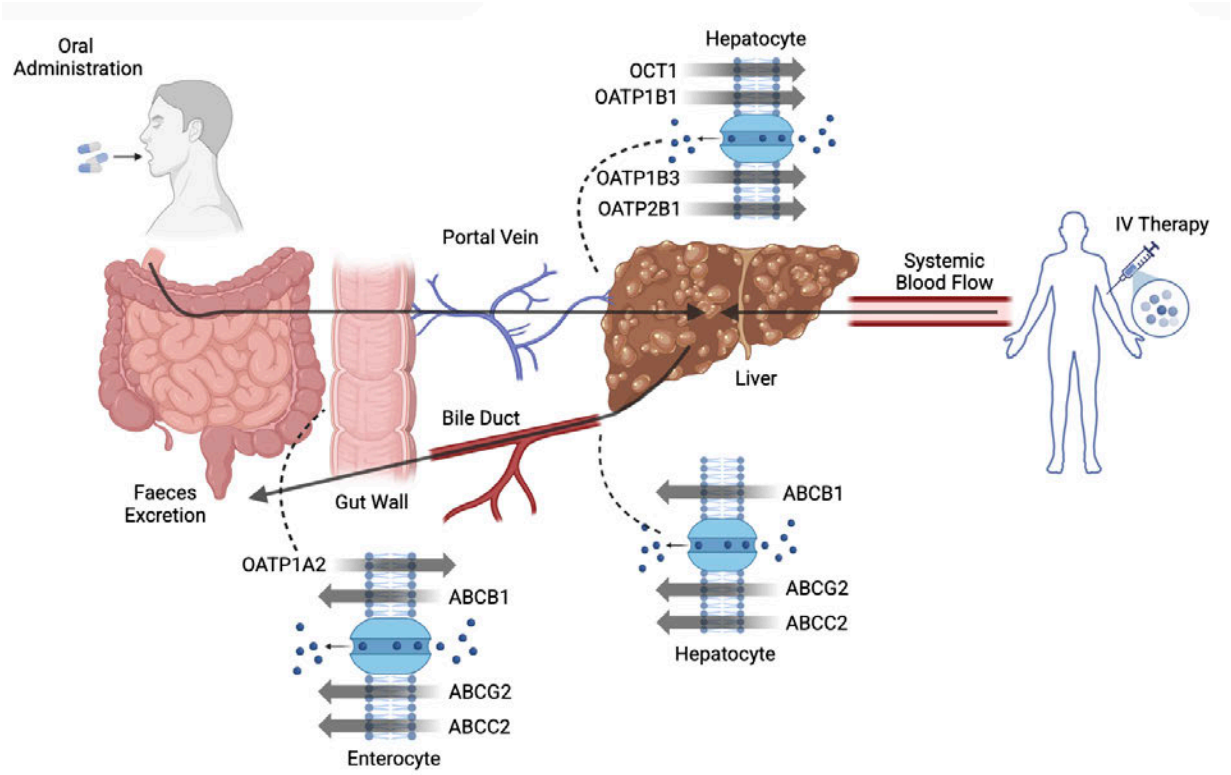


Figure 1: **Illustration representing localization of ABC and SLC transporters involved in pharmacokinetics.** Therapeutic drugs are orally administered which passes through the lumen of the intestine and gut wall and then up taken by the liver via the portal vein and elimination via the bile through the gut wall to eventually be excreted. Alternatively therapeutic doses are administered intravenously (IV) and transported to the liver via systematic circulation which is then eliminated via the bile through the gut wall and then excreted. Drug transporter proteins present in enterocytes (OATP1A2, ABCB1, ABCG2 & ABCC2) and hepatocytes (OCT1, OATP1B1, OATP1B3, OATP2B1, ABCB1, ABCG2 & ABCC2) are responsible for the absorption, metabolizing and release of therapeutics into the bloodstream that contribute to drug PK levels. Schematic illustration was created with Biorender Software (app.biorender.com).

Previous studies based on ABC transporter genes have shown alterations in intestinal drug absorption observed in mouse models (18, 19). It has been shown in studies that re-sequenced

the human ABC gene, several variants were identified which have been associated with functional activity of the encoded protein *in-vivo* (20-22). These mutations within the transporter gene may modulate the phenotypes of the transporter which will then have an impact on drug treatment, predisposes the patient to disease and affect toxicity levels (9). Variants in ABCB1 rs2032582 SNP have been previously investigated in a South African TB cohort and was found to be modestly associated with the interindividual variability reported in moxifloxacin pharmacokinetics and exposure (6). Further investigation is required to understand the effects of ABCB1 variants on drug transporter membrane and treatment outcomes in an African population.

Whilst most researchers have focused on ABC transporters for several years, only recently has the impact of genetic variation in solute carriers (SLCs) been investigated. In humans more than 300 individual proteins exist which are organized into 47 families of SLCs (23). Membrane proteins from the SLC families have been identified as passive transporters and exchangers. Organic anion transporting polypeptides (OATPs) are responsible for sodium-independent transport of a wide range of compounds such as: xenobiotics, steroid conjugates, thyroid hormones and bile salts (24). OATP1B1 (OATP2, OATP-C, and LST-1) are primarily found in the basolateral membrane of human liver hepatocytes (9). Due to its localization and it being an uptake transporter, its primary function involves transporting various substrates into the liver from the bloodstream (25). Variants such as: 521T>C, 388 A>G and 288 A>G among others within OATP1B1 that contribute to functional decrease in substrate uptake have been found in non-African populations (9).

Previous studies have shown that genetic variability observed within the African population, especially within the genes that code for drug-metabolizing and transporter enzymes, showed to have contributed significantly to variable drug concentrations and pharmacokinetics parameters (2, 5, 6, 26-28). In order to investigate the impact that specific SNPs impose on drug levels, analysis focussing on the area under the curve (AUC) enables us to understand the total exposure of a drug within the time that the host is exposed to it. The AUC plot as illustrated in Figure 2. entails drug concentration recorded from plasma samples versus time following oral administration and advises on the time the respective drug remains in systemic circulation. In the context of pharmacodynamics, the AUC often determines the therapeutic level or toxic effects clarifying the understanding of the relationship between treatment and patient outcomes. Plasma levels with high AUC values are generally associated with greater overall drug exposure, which at times can influence both the efficacy of a drug leading to potential

side effects or adverse events. Area under the curve data is regarded as a key parameter and datapoint for optimizing therapeutic regimens and exploring the possibility of combining various drug formulations.

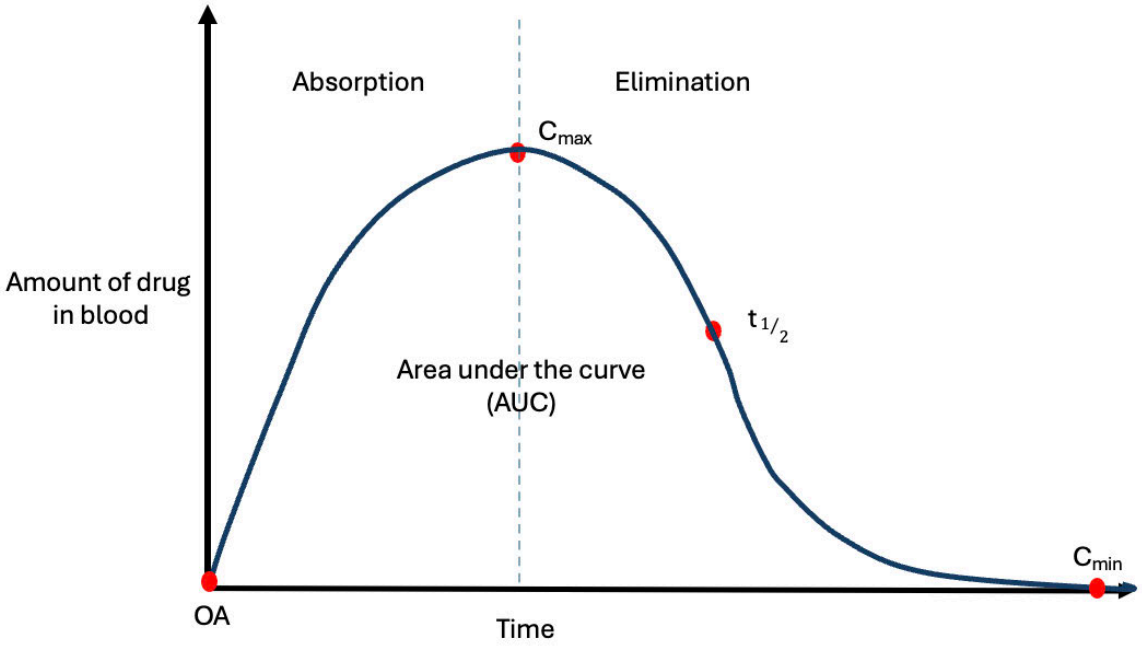


Figure 2: Illustration showing typical plasma concentration–time curve illustrating the area under the curve (AUC) in pharmacodynamics. The curve highlights the change in plasma drug concentration following oral administration (OA), beginning with the absorption phase, peaking at the maximum concentration (C_{max}), followed by drug half-life ($t_{1/2}$) and then a decline during the elimination phase until the minimum concentration (C_{min}). The area beneath the curve quantifies total drug exposure over time.

A recent study in South Africa reported on the prevalence of genetic variants on PK parameters of isoniazid and rifampicin, however no association was observed between rifampicin pharmacokinetics and SLCO1B1 and ABCB1 (2). Further studies are required to identify the effect of the genetic variation on TB drug therapy and their clinical relevance within African populations. Genome-wide association studies are pivotal in gaining better insights on understanding the diverse genomes of African population. Whilst there are several commercially available microarrays that can be utilized to address the lack of genetic information from African cohorts, the challenge posed is that there are several markers yet to be discovered which may potentially play crucial roles in treatment diversity among African populations.

Significance of the study:

This study will serve to provide researchers and drug developers with a greater understanding of how individuals of African descent metabolize anti-TB drugs based on their genetic mutations or combination of specific mutations (haplotype). Due to the great level of genetic diversity in African individuals this study in addition to focusing only on specific drug transporter investigation the study also takes an unbiased approach to screen for potential novel and understudied markers that may be associated with HIV-1 infection, TB recurrence and predisposition. This analysis will also help to modestly suggest/ strengthen future treatment guidelines for pharmacists and clinicians to administer TB drugs for patients in need as well as suggest understudied variants that could be of potential concern to virologists and microbiologist in the fields of research and diagnostics.

Study Aims and Objectives:

Study Aims:

The aim of this study is to investigate the role of genetic polymorphisms, found within the drug transporter genes that influence PK levels and TB outcomes within individuals of African descent, which will help to identify biomarkers for drug metabolism of these individuals that lead to adequate or poor treatment outcomes based on their respective drug levels.

Objectives:

Objective 1: To review and highlight the level of genetic diversity observed in African populations with additional focus on drug metabolism in the context of specific variants in drug transporter genes within ABC and SC families.

Objective 2: To Determine the potential role and implications of specific known genetic variants in *ABCB1* and *SLCO1B1* on moxifloxacin pharmacokinetics in South African patients infected with tuberculosis.

Objective 3: To investigate the role of both known and understudied SNPs in association with HIV-1 acquisition, TB infection and recurrence using an unbiased approach focusing on genome-wide markers using a genotyping by microarray approach using various cohorts.

Objective 4: To explore additional variants in drug transporter genes *ABCB1* and *SLC01B1* together with genome-wide variants that are most highly significant and associated with moxifloxacin drug PK levels in patients receiving anti-TB therapy. These SNPs could serve as biomarkers of TB drug metabolism in African descent.

Hypothesis:

Drug transporter genes play an integral in the pharmacokinetics of anti- tuberculosis drugs and may influence treatment outcomes. Single nucleotide polymorphisms existing in these genes have been identified from European and Caucasian cohorts previously. It can therefore be hypothesized that novel/ understudied SNPs found in genes that code for drug metabolism and transporter enzymes i.e. *ABCB1, ABCC2, ABCG2, SLCO1B1* & *SLC22A1* are also associated with drug metabolism and treatment outcomes in individuals of African descent. We further hypothesize that due to the genetic diversity in African genomes, we believe that there are several genome-wide markers that are associated with HIV-1 risk, TB recurrence and predisposition that may be present at lower frequencies or unique to African populations. Further studies are needed to identify these novel SNPs which could serve as biomarkers in TB treatment outcomes and drug metabolism.

Expedited ethical clearance for this study using human specimens was obtained from the University of KwaZulu-Natal Biomedical Research Ethics Committee (Ethical approval reference number: **(BREC/00004110/2022) (Addendum C).**

As per guidelines of the University of KwaZulu-Natal's for a Doctor of Philosophy (PhD) degree, a thesis by publication states that the thesis may comprise of at least three published papers or in press in accredited journals, of which the student (Mr Sahil Tulsi) needs to be the prime author for all publications. Within this thesis, a total of four manuscripts have been generated for the specific research question stipulated (one review paper and three experimental research papers with myself, Sahil Tulsi as the prime author). In addition two more manuscripts have been included in the addendum which are related to the research topic presented herein.

Chapter 2

Preface:

This chapter reviews and highlights the role of genetic polymorphisms in drug transporter genes, specifically the ABC and SLC families in influencing individual responses to anti-tuberculosis (TB) therapy with a focus on African populations. Populations of African descent exhibit greater genetic diversity than most other global populations. To date the genetic data captured in global databases represent African populations with approximately ~2-3% (53) . It has been observed that genetic diversity within these specific drug transporter genes vary significantly between African, Asian, and European populations, potentially affecting drug efficacy and safety. The review chapter emphasizes the limited data and lack of genetic information available on African populations and calls for more population-specific research to identify both known and novel variants that could impact drug pharmacokinetics and pharmacodynamics, ultimately improving TB treatment outcomes in Africa. It is likely that additional variants remain undiscovered within these genes in African populations. In light of both the fact that African genomes are severely understudied and the great level of diversity, it would be extremely beneficial for additional research studies to be conducted that can aid in identifying the role of these known and novel variants that are yet to be discovered in African populations. It is also possible that there are additional genome-wide variants that may potentially play differential roles among populations. Exploring the roles of these mutations specific to various populations can provide clarity on differential treatment outcomes across populations and possibly suggest dosing therapies for optimal management of tuberculosis at a level unique to the individual's needs.

Genetic diversity of polymorphisms within specific drug transporter genes in African individuals infected with Mycobacterium tuberculosis

Sahil Tulsi [1], Thilona Arumugam [1], Ishq Pramchand [1] and Veron Ramsuran [1,2,3]

[1]School of Laboratory Medicine, College of Health Sciences, University of KwaZulu-Natal, Durban, South Africa, [2] Centre for the AIDS Programme of Research in South Africa (CAPRISA), University of KwaZulu-Natal, Durban, South Africa, [3] International Foundation for Integrative Medical Research, Cape Town, South Africa

*Authors to whom correspondence should be addressed:
sahiltulsi@gmail.com; ramsuranv@ukzn.ac.za

Conflict of interest: None declared.
In submission in : *Journal of Clinical Tuberculosis and Other Mycobacterial Diseases*
Manuscript ID: JCTUBE-D-25-00260

Abstract

Genetic polymorphisms located within the human drug transporters, mainly adenosine triphosphate-binding cassette (ABC) and solute carriers (SLC) may play an influential role in fluctuating drug responses to Asian and European populations in comparison with African individuals undergoing anti-TB therapy. This review focuses on the importance of genetic polymorphisms in drug transporters and their potential contribution to treatment diversity in patients undergoing anti-TB therapy, with a specific focus on an African population. Genetic diversity among African individuals is observed at variable levels compared to other populations across the globe. Studies have shown trends that are observed in drug transporter genetic diversity among these respective populations and may result in enhanced or reduced transporter activity. Our review specifically focuses on the ABC and SLC gene families, of which limited data is available for the African population for other drug transporter genes. In addition, we highlight the need for emphasis on population specific studies in response to anti-TB treatment. Common genetic variants within the ABC and SLC gene families linked to drug levels among European populations are found at substantially different levels in Africans.

Additional unknown variants might be present within these genes that could contribute to diversity in drug levels among African populations. Therefore, additional research needs to be conducted to help further identify the role of these known variants in African populations, along with the identification of novel single nucleotide polymorphisms that could potentially influence drug pharmacokinetic and pharmacodynamic levels to optimize and improve treatment outcomes.

Introduction:

The African continent, home to 17% of the world's population accounts for over 25% of Mycobacterium tuberculosis (TB) mortalities globally (1). Despite advanced ongoing research and development carried out over the past 50 years, TB standard treatment regimens and drugs administered for drug-susceptible (DS)TB remain unchanged (2). Only over the recent years have new drugs and regimens been approved for both DS-TB and drug-resistant TB. Rifampicin, isoniazid, ethambutol and pyrazinamide are the current DS-TB drugs that are used for first-line treatment regimens. Several studies have been performed to investigate optimal dosing strategies to determine effective TB drug concentrations at sites of action (3-5). Recent studies have shown that first-line drugs namely isoniazid, pyrazinamide and rifampicin drug concentrations are observed at much lower levels in individuals of African descent, which therefore results in compromised treatment outcomes or treatment failure (4-6). Adding to the significant burden of TB related healthcare are both multidrug-resistance TB (MDR-TB) and extensively drug-resistant (XDR-TB) which have been the contributing to increased prevalence in Sub-Saharan Africa (7, 8). In 2022 it has been reported that 15% of global MDR-TB cases are represented by patients of African descent (9). Closely investigating genetic diversity in African populations can help researchers better understand this high representation of African TB cases in global statistics of MDR-TB.

Apart from adherence to TB treatment, it is widely understood that patient TB treatment outcomes are often influenced by specific genetic polymorphisms (4). Genetics and polymorphisms are a vital factor in the advancement of precision medicine along with a patient's lifestyle and environmental factors which helps to determine a treatment that would work best for them. Genes that code for drug-metabolizing and transporter enzymes often contain single nucleotide polymorphisms (SNPs) which have been shown to significantly correlate with variable drug concentrations and pharmacokinetic parameters (10). The extensive level of genetic variation observed across African populations plays a critical role in shaping tuberculosis TB pharmacogenetics. Genetic variability influences key determinants of

drug metabolism, treatment efficacy, and raises the risk of adverse drug reactions. As the complex interactions between TB, antimicrobial resistance, and the distinct challenges within African settings continue to unfold, there is growing recognition of the potential of precision medicine's ability to enhance treatment outcomes (54). However, the effective application of such approaches requires robust research efforts and thorough validation in diverse African cohorts. A comprehensive understanding of the interactions among host genetics, comorbid conditions, and drug resistance mechanisms is essential for improving TB management and clinical outcomes in this high-burden context.

Drug transporter genes are of utmost importance across multiple areas of therapeutics due to their vital role in processes that regulate both pharmacokinetic (PK) properties of drugs (absorption, distribution, and elimination) and the development of drug resistance via decreased drug uptake on increase in efflux (10).

Figure 1 represents the localization of Adenosine tri-phosphate binding cassette (ABC) and solute carriers (SLC) drug transporters in humans, with focus on the respective genes involved in absorption, distribution , and elimination.

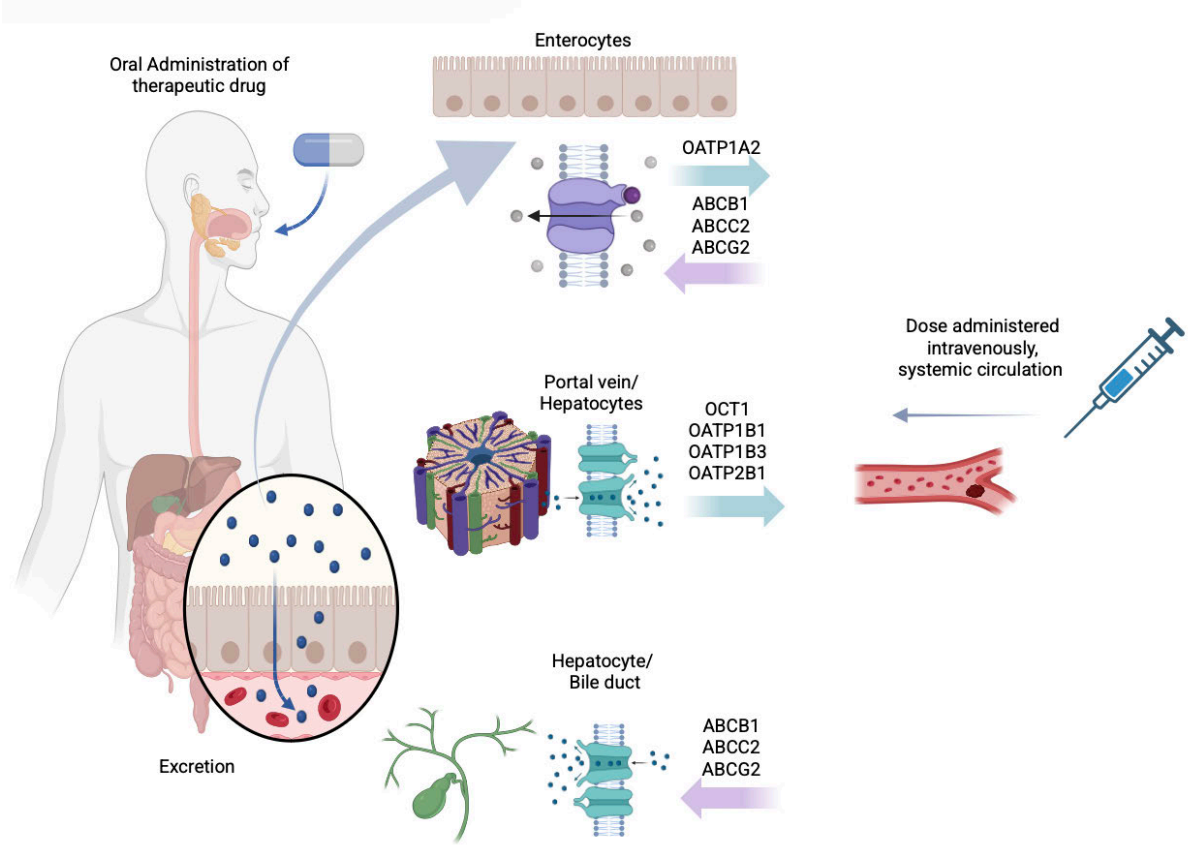


Figure 1: Schematic depiction of the spatial distribution of ABC and SLC transporters relevant to drug pharmacokinetics. Following oral administration, therapeutic agents transit

through the intestinal lumen and epithelial barrier before entering the liver via the portal circulation, after which they may be eliminated into the bile, transported back through the intestinal wall, and ultimately excreted. Alternatively, intravenously (IV) administered drugs reach the liver through the systemic circulation and undergo similar biliary elimination and excretion pathways. Transporter proteins expressed in enterocytes (OATP1A2, ABCB1, ABCG2, and ABCC2) and hepatocytes (OCT1, OATP1B1, OATP1B3, OATP2B1, ABCB1, ABCG2, and ABCC2) mediate the uptake, disposition, and efflux of therapeutic compounds, thereby shaping systemic pharmacokinetic profiles. Blue arrows facing right are indicative of drug transporter genes involved in uptake, with transporter genes that release drugs represented by left facing purple arrows. Illustration was created with Biorender software (app.biorender.com)

Both soluble carriers and ATP-binding cassette transporters are two of the most commonly studied membrane transporters (10). There are almost 400 individual proteins between these two families of drug transporters which have been identified to date. Mutations and changes in expression within these proteins can have direct implications to health outcomes and are crucial for treatment success in MDR-TB and XDR-TB patients. These proteins are widely distributed throughout the human body and closely studied to better understand their role for drug delivery in addition to their contribution of genetic variants in these transport proteins that lead to variable patient treatment diversity in the efficacy and safety for pharmacotherapy (11).

ABC Transporters: Localization, roles & mechanisms

In humans there are 49 genes that code for ABC transporters which include proteins that have shown to elicit great clinical impact such as the translocation of various metabolites and drugs (12). The transporter proteins from the ABC family are known to be associated with detoxification of the host and provide protection against xenobiotics (13). ABC family consist of seven subfamilies: ATP-binding cassette (ABC), Multidrug-resistance like protein (MRP), alanine dehydrogenase (ALD), organic anion binding protein (OABP), general control non-repressible 20 (GCN20), white mutation (White) and multidrug resistance/ transporter associated with antigen processing (MDR.TAP). The ABC family is also a member of the MDR/TAP which has been observed to be involved in multidrug resistance (14). Recent studies based on ABC transporter genes have shown alterations in intestinal drug absorption observed in mouse models (15, 16). It has been shown in studies that re-sequenced genes from the ABC family, several variants were identified which have been associated with the functional activity of the encoded protein *in vivo* (17-19). These mutations within the transporter gene may

modulate the phenotypes of the transporter which then have an impact on drug treatment, predisposes the patient to disease, and affect toxicity levels (10).

ABCB1, ABCC1, ABCC2 & ABCG2

From the ABC gene family, variants in four genes of interest namely *ABCB1, ABCC1, ABCC2 and ABCG2* have been investigated to better understand their role in drug response and pharmacokinetics (20). Drug transporter gene *ABCB1* (commonly known as P-glycoprotein (P-gp), *CD243*, and Multidrug resistance protein 1 (MDR1) have been identified as a membrane-associated protein from a superfamily of ATP-Binding cassette transporters. To date the most common SNPs identified in the ABCB1 gene are rs1128503, rs2032582, and rs1045642 (21). These trans-membrane efflux pumps facilitate the transportation of various molecules, exogenous compounds, and therapeutic agents across intra-and extra-cellular membranes. ABCB1 has been previously shown to influence the response to rifampicin (22). Chigutsa *et al* (2011) reported minor allele frequencies of *ABCB1* variants (rs1045642, rs2032582, rs1128503, and rs3842) among a South African cohort as 0.26, 0.19, 0.26 and 0.21 (3). The participants within this study were from mixed ancestry, i.e.40% were of black African ethnicity (mainly Xhosa), with the rest having mixed ancestry (mainly European and African origins (3).

The rs1128503 polymorphism, primarily known for its importance in impacting co-translational folding of nearby amino acids known to be essential for ATP binding and hydrolysis, is more prevalent in Asian and European populations (60-72% and 34-42%), respectively; however, observed at a frequency of only 26% in African populations (19, 23). Genetic diversity as such among different ethnicities across multiple geographies may suggest that there are several novel variants contributing to treatment diversity which are yet to be identified and characterized.

Other variants in *ABCB1* genes such as rs2032582 SNP have been previously investigated in a South African TB cohort and were found to be modestly associated with the interindividual variability reported in moxifloxacin pharmacokinetics and exposure (5). Table 1 below summarizes the diversity of polymorphism frequencies in genes from the ABC class among Asians, Africans & Europeans with differences in minor allele frequency greater than arbitrary values of ≥ 0.2 highlighted in yellow. Genetic diversity as such is a crucial factor that can

possibly define population specific treatment and the expectation of differential treatment outcomes and adverse events.

Table 1: Genetic diversity & Minor Allele frequency of common ABC SNPs between Asian, African and European populations as per ALFA project.

Gene	ABC specific	Minor Allele Frequency (MAF)				
	SNPs (mutation)	Asian	Difference	African	Difference	European
ABCB1	**rs1128503**	0.373	0.167	0.206	0.222	0.428
	rs1045642	0.382	0.161	0.221	0.259	0.48
	rs2032582	0.433	0.314	0.119	0.321	0.44
ABCG2	**rs2231137**	0.316	0.257	0.059	0	0.059
	rs72552713	0.007		0		0
	rs2231142	0.296	0.271	0.025	0.077	0.102

Minor allele frequencies highlighted in yellow fields are indicative of a difference with ≥ 0.2 between respective populations for each SNP. The value of 0.2 was an arbitrary value used for the purposes of this review.

Based on ABC genotypes prevalent in Asian, African, and European populations it has been observed that the frequency of the minor allele is lower in individuals of African descent as compared to the other two populations. Observations in genetic variance as such could be owing to the fact that African genomes are far more diverse than other populations' genetic profiles (24). Some of the functional actions of these prevalent ABC SNPs highlighted surface the importance and role of these drug transporters in a patient's response to therapy.

With reference to the variant in *ABCB1* rs1128503 which has been observed to be prevalent in Asians and Europeans (19, 23), responses to tramadol are usually clinically associated with rs1128503, however no significant associations were reported (25). African populations have a much lower minor allele frequency (0.206) of these *ABCB1* rs1128503 SNPs which contributes to the fact that reduced levels of TB drugs have been observed in individuals of African descent followed by compromised treatment outcomes (4-6). Additionally, in the *ABCG2* gene loss of transport activity has been observed in Asian populations with the presence of rs72552713 (Q126stop) polymorphism (26), whilst African and Europeans did not

have these SNPs present. Along with conventional TB treatment therapies, rifamycin's have also been used in some studies as a novel class of antibiotics that have been clinically approved (27). A study in 2019 investigated *ABCB1* polymorphisms in Mexican TB patients also known to be diabetics reported rs1045642 genotypes (CC or CT) to be associated with lower rifamycin levels compared to genotype TT (28), whilst other studies showed no significant effect on PK levels of rifamycin with rs1045642 in diabetics infected with TB (29). With reference to West-African populations the variant rs1045642 was observed at a minor allele frequency of 0.34 (30) and observed to be lower, around 0.07 to 0.12 in South African populations (31).

In addition to diabetic treatment serving their primary purpose, further studies may be required to better understand if diabetic drugs have a significant impact on rifamycin levels. With a great deal of invariance of these ABC SNPs that are dispersed among these populations, there is much room for debate on the need to further investigate the genetic profiles of African populations to strengthen the understanding patient therapy.

Solute carriers

Whilst most researchers have focused on ABC transporters for several years, only recently has the impact of genetic variation in SLCs been investigated in more detail. In humans, more than 300 individual proteins exist which are organized into 47 families of SLCs (32). Membrane proteins from the SLC families have been identified as passive transporters and exchangers. Organic anion-transporting polypeptides (OATPs) are responsible for the sodium-independent transport of a wide range of compounds such as xenobiotics, steroid conjugates, thyroid hormones, and bile salts (33). OATP1B1 (OATP2, OATP-C, and LST-1) are primarily found in the basolateral membrane of human liver hepatocytes (10). Due to its localization and being an uptake transporter (illustrated in **Figure 1)**, its primary function involves transporting various substrates into the liver from the bloodstream (34). Specific to anti-TB therapy, SLCs play a vital role in substrate transportation and their role in patients undergoing TB therapy.

SLCO1B1, SLCO1A2, SLCO2B1, SLC22A1* and *SLC22A6

Various genes that are members of the SLCO and SLC22 subfamilies are crucial to drug transport in TB patients. Rifampicin is a common and widely used anti-TB therapy drug that has been observed to be one of the multiple substrates transported by these solute carriers among several other anti-TB drugs part of various treatment regimens (35). Variants in SLCO1B1 genes have the potential to impact pharmacokinetics of anti-TB drugs such as rifampicin and isoniazid (36), eventually leading to altered drug levels, efficacy and possible

adverse drug reactions in TB patients . A recent genome-wide association study (GWAS) by Sanni et al., 2022 reported a polymorphism in the *SLC01B1* gene (rs11792875) with strong genome-wide significance (p = 2.44e-07) associated with rifampicin clearance in Sub-Saharan TB patients receiving rifampicin and isoniazid treatment (37). Microarray technology has become a valuable tool in screening for variants of significance in Genome-wide association studies (GWAS) however commercially available arrays would need more updated content with the use of data obtained from whole-genome sequencing (WGS) to identify more influential variants that are either underexplored or novel.

OATP1A2, OATP1B1, OATP1B3, OATP2B1* and *OCT1

Organic Anion Transporting Polypeptide (OATP) is a family of genes within the solute carrier drug transporters that encode membrane transport proteins responsible for the uptake of a wide range of endogenous and exogenous compounds including drugs into cells (38). As drug transporters, OATPs influence the pharmacokinetics of many medications, affecting their bioavailability, therapeutic efficacy, and potential for drug interactions (39). Genetic variants in the OATP1B1 transporter have been suggested to be partially responsible for the great large interindividual variation in exposure to rifampicin (40). Organic Cation Transporter 1 (OCT1) is also a drug transporter of interest encoding a membrane transporter protein responsible for the cellular uptake of drugs and metabolites. Studies have shown that OCT1 variants lead to differential drug responses among individuals leading to diverse treatment outcomes (41).

The SLC transporters referenced above are mainly localized in the human intestine, and liver, and those that are ubiquitously expressed all act as substrate transporters and inhibitors of rifampicin and other TB drugs with very similar mechanisms in common (42). A study has shown that variants of interest in African populations with functional consequences of decreased transporter activity in common are rs2306283 (N130D) and rs1363826500 (E667G) are present at much higher allele frequency levels (74% and 34% respectively) compared to Asian (54% and 0% respectively) and European populations(30% and 2% respectively) (43). Other variants such as rs4149056, rs2306283, and A288G among others within *OATP1B1* that contribute to a functional decrease in substrate uptake have been found in non-African populations (10). As seen in the data from several studies show that *SLCO1B1* variants are mainly known to contribute strongly toward decreased transport activity which could be indicative that one of the reasons African populations result in compromised treatment

outcomes is due to a higher frequency of these known SNPs. Multiple polymorphisms such as rs2306283 and rs11045819 have been investigated in studies showing significantly reduced levels of rifampicin levels in patients from Africa and America (44, 45).

Overall, for solute carriers, previous studies have shown that genetic variability observed within the African population, especially within the genes that code for drug-metabolizing and transporter enzymes, showed to have contributed significantly to variable drug concentrations and pharmacokinetic parameters (3-5, 44-46). Some notable *SLCO1B1* variants (c.521T>C, c.388A>G and c.467C>T) may influence the pharmacokinetics of anti-TB drugs, specifically statins, and possibly other substrates altering liver transporter organic anion-transporting polypeptide 1B1 (OATP1B1) which are more common in African populations in contrast to other non-African populations (47). A recent study in South Africa reported on the prevalence of genetic variants on PK parameters of isoniazid and rifampicin, however no association was observed between rifampicin pharmacokinetics and *SLCO1B1* and *ABCB1* (4). Further studies to understand these solute carrier variants and clinical associations are required to identify the effect of the genetic variation on TB drug therapy and their clinical relevance within African populations.

General observations in context and suggestions for future research

African populations have a greater range of genetic diversity in comparison to other populations which has been observed in whole-genome analysis (48). In light of this the difference in genetic profiles of African individuals affects the process of drugs and substrates being transported by ABC and SLC genes which leads to differential treatment options and outcomes compared to other populations. This review brings to the surface the fact that there is indeed a great need for additional future research to actively include more African populations in study cohorts for a greater understanding of their response to TB therapy. In addition to this, reasoning arises from the fact that in some cases research performed on drug transporters have reported on data that may contradict observations and findings from other studies, hence contributing more to the need for further investigation of genetic profiles in African populations.

Variant rs1128503 in *ABCB1* genes are mutations play a crucial role in ATP binding of substrates and hydrolysis, which was reported to have high frequencies in Asians and Europeans but less frequent in African populations (19, 23). Polymorphisms as such could

prove to be beneficial to Asian and European populations in terms of more effective substrate-binding ABC drug transporters with effective treatment outcomes as a result. African populations who lack the same level of allele frequency as Asians and Europeans would mean more individuals undergoing TB therapy in Africa are at a disadvantage due to the nature of their genetic diversity. Another two observations in African populations having lower allele frequencies than Asian and European populations were for variants rs1045642 (silent) and *ABCB1**13 (RS1128503/RS2032582/RS1045642 haplotype), which are associated with altering substrate specificity and inhibition by a small subset of modulators (23, 49, 50). Since genetic databases mainly represent European and Asian genetic profiles there is a lot of focus on known SNPs and their effect on therapy however very little research has contributed towards the identification of novel SNPs in African populations that may have a positive impact on drug transport.

The level of genetic diversity between Asian, European and African individuals may possibly also be observed from additional analysis such as linkage disequilibrium of *ABCB1* and *ABCC2* genes, that are likely to suggest further exploration of variants both known and novel in these genes of African descent will reveal more useful genetic information and their clinical relevance to drug transportation.

Frequencies of *SLCO1B1* variants rs2306283 (N130D) and rs1363826500 (E667G), both polymorphisms contributing to a decrease in transport activity were much higher in African populations (74% and 34% respectively) compared to non-African populations. Another study with 113 Ghanaian adolescents by Dompreh *et al* (44) reported an association between the presence of variant rs2306283 and rifampicin C_{max} (P=0.052), apparent oral clearance time per hour (P=0.093) and area under the time-concentration curve from 0 to 8 hours after dosing (P=0.085), (p values here >0,05: not statistically significant) however no impact on PK drug levels were observed in similar studies concerning South African and Ugandan populations (45, 46, 51). Nevertheless, apart from data that may contradict studies and variability in frequency among African and non-African populations, with high frequencies of detrimental polymorphisms such as rs2306283 (N130D) within African populations, it is unquestionable that further studies need to be performed in order to optimize drug therapy and dosing strategies for patients undergoing anti-TB therapy since the African genetic profiles are not an ideal fit for generalized treatment strategies among worldwide populations.

African Specific polymorphisms associating with TB

PK drug variability and response to TB treatment is a result partially driven by specific genetic polymorphisms in drug transporter genes. With African genomes only representing a minute portion of genomic databases, there are limited known polymorphisms specific to African populations that drive these responses to TB drugs. One example in solute carriers, rs1363826500 (E667G) is a polymorphism that is highly prevalent (34% frequency) in African populations and is associated with a reduction in transport activity (22, 43). Additional genotyping by sequencing or micro array studies are required to screen more African individuals to gain statistical power and impute for these known variants. While it is of utmost importance to further explore the role known SNPs play on TB drug transport in African individuals, the identification of novels SNPs that may play a role is equally as valuable to determine more effective therapy for TB patients. Several pan-African efforts on the continent to further develop genomics and build infrastructure in Africa are in favour of supporting African genome and exome sequencing in the best interest of patient healthcare for the African population. In addition to novel SNPs, it is a unfortunate that African genetic profiles only make up approximately 2% of genomes examined in the Human health and hereditary Africa (H3Africa) genome-wide association study (GWAS) . Under represented genetic profiles of African populations are at a disadvantage compared to other highly examined populations that may benefit from findings in genomic research which aid in early disease detection and drug design for precision medicine. A genetics study by Yu *et al* (52) that reported on patterns in non-coding regions in 50 DNA segments showed that the average nucleotide diversity (pi) was only 0.061% +/- 0.010% among Asians and 0.064% +/- 0.011% among Europeans but nearly twice as high (0.115% +/- 0.016%) among individuals of African descent. The African diversity estimate is seen to be much higher than that between Africans and Eurasians (0.096% +/- 0.012%) (52). This review supports any research that includes African cohorts in additional studies since these individuals respond differently compared to non-African populations. It is noticeable that for some studies that have already taken place, data was limited by sample size and the number of known polymorphisms which have an impact on PK drug levels in African populations. Pharmacogenomic studies face several challenges when assessing the impact of polymorphisms in drug transporter genes. The majority of pharmacogenomic studies include

mostly Asian and European populations, with occasional extrapolation of African-American data to African populations, even so, this is not an ideal representation of the genetic profiles and diversity observed in African populations.

Conclusions

The genetic diversity of ABC and SLC transporters has been reported to be much higher in African populations when compared to Asian and European populations. The impact of this difference in diversity along with findings resulting from treatment outcomes influenced by novel SNPs has crucial implications in the development of precision medicine leading to a more effective therapy, in particular for tuberculosis patients. Dosing therapies could be further optimized and advanced with more PK and PD studies that include African populations in the future. Future studies may also assess confounding factors such as drug-drug interactions, chronic liver disease and kidney disease to better understand and measure treatment outcomes. Another benefit would be the addition of genetic profile data for African populations which will aid the current situation being the lack of African descendant profiles in genetic databases. Population genetic studies focusing on African descent will address this lack of data thereof and is a key element to defining success for the future of patient specific treatment needs for African individuals. Nevertheless, data as such would supplement and contribute to a platform for further genetic studies in various research areas of interest.

Funding statement

Research reported in this review article was supported by Grants, Innovation and Product Development unit of the South African Medical Research Council with funds received from Novartis and GSK Global Health for Project Africa GRADIENT (Grant No. GSKNVS2/202101/005).

References

1. Organization WH. WHO Global tuberculosis report. 2018.
2. Health SADo. South African National Tuberculosis Management Guidelines. 2014.
3. Chigutsa E, Visser ME, Swart EC, Denti P, Pushpakom S, Egan D, et al. The SLCO1B1 rs4149032 polymorphism is highly prevalent in South Africans and is associated with reduced rifampin concentrations: dosing implications. Antimicrob Agents Chemother. 2011;55(9):4122-7.

4. Naidoo A, Chirehwa M, Ramsuran V, McIlleron H, Naidoo K, Yende-Zuma N, et al. Effects of genetic variability on rifampicin and isoniazid pharmacokinetics in South African patients with recurrent tuberculosis. Pharmacogenomics. 2019;20(4):225-40.
5. Naidoo A, Ramsuran V, Chirehwa M, Denti P, McIlleron H, Naidoo K, et al. Effect of genetic variation in UGT1A and ABCB1 on moxifloxacin pharmacokinetics in South African patients with tuberculosis. Pharmacogenomics. 2018;19(1):17-29.
6. McIlleron H, Wash P, Burger A, Norman J, Folb PI, Smith P. Determinants of rifampin, isoniazid, pyrazinamide, and ethambutol pharmacokinetics in a cohort of tuberculosis patients. Antimicrob Agents Chemother. 2006;50(4):1170-7.
7. Gandhi NR, Nunn P, Dheda K, Schaaf HS, Zignol M, van Soolingen D, et al. Multidrug-resistant and extensively drug-resistant tuberculosis: a threat to global control of tuberculosis. Lancet. 2010;375(9728):1830-43.
8. (WHO) WHO. Global Tuberculosis Report 2023. Geneva: WHO. 2023.
9. Naidoo K, Perumal R, Cox H, Mathema B, Loveday M, Ismail N, et al. The epidemiology, transmission, diagnosis, and management of drug-resistant tuberculosis—lessons from the South African experience. The Lancet Infectious Diseases. 2024;24(9):e559-e75.
10. Franke RM, Gardner ER, Sparreboom A. Pharmacogenetics of drug transporters. Curr Pharm Des. 2010;16(2):220-30.
11. Kerb R. Implications of genetic polymorphisms in drug transporters for pharmacotherapy. Cancer Lett. 2006;234(1):4-33.
12. Glavinas H, Krajcsi P, Cserepes J, Sarkadi B. The role of ABC transporters in drug resistance, metabolism and toxicity. Curr Drug Deliv. 2004;1(1):27-42.
13. Lin JH, Yamazaki M. Role of P-glycoprotein in pharmacokinetics: clinical implications. Clin Pharmacokinet. 2003;42(1):59-98.
14. Borst P, Evers R, Kool M, Wijnholds J. A family of drug transporters: the multidrug resistance-associated proteins. J Natl Cancer Inst. 2000;92(16):1295-302.
15. Jonker JW, Buitelaar M, Wagenaar E, Van Der Valk MA, Scheffer GL, Scheper RJ, et al. The breast cancer resistance protein protects against a major chlorophyll-derived dietary phototoxin and protoporphyria. Proc Natl Acad Sci U S A. 2002;99(24):15649-54.
16. Sparreboom A, van Asperen J, Mayer U, Schinkel AH, Smit JW, Meijer DK, et al. Limited oral bioavailability and active epithelial excretion of paclitaxel (Taxol) caused by P-glycoprotein in the intestine. Proc Natl Acad Sci U S A. 1997;94(5):2031-5.
17. Fromm MF. The influence of MDR1 polymorphisms on P-glycoprotein expression and function in humans. Adv Drug Deliv Rev. 2002;54(10):1295-310.
18. Lockhart AC, Tirona RG, Kim RB. Pharmacogenetics of ATP-binding cassette transporters in cancer and chemotherapy. Mol Cancer Ther. 2003;2(7):685-98.
19. Marzolini C, Paus E, Buclin T, Kim RB. Polymorphisms in human MDR1 (P-glycoprotein): recent advances and clinical relevance. Clin Pharmacol Ther. 2004;75(1):13-33.
20. Bruckmueller H, Cascorbi I. ABCB1, ABCG2, ABCC1, ABCC2, and ABCC3 drug transporter polymorphisms and their impact on drug bioavailability: what is our current understanding? Expert Opin Drug Metab Toxicol. 2021;17(4):369-96.
21. Khan A, Abbas M, Verma S, Verma S, Rizvi AA, Haider F, et al. Genetic Variants and Drug Efficacy in Tuberculosis: A Step toward Personalized Therapy. Glob Med Genet. 2022;9(2):90-6.

22. Ahmed S, Zhou Z, Zhou J, Chen SQ. Pharmacogenomics of Drug Metabolizing Enzymes and Transporters: Relevance to Precision Medicine. Genomics Proteomics Bioinformatics. 2016;14(5):298-313.
23. Fung KL, Gottesman MM. A synonymous polymorphism in a common MDR1 (ABCB1) haplotype shapes protein function. Biochim Biophys Acta. 2009;1794(5):860-71.
24. Gomez F, Hirbo J, Tishkoff SA. Genetic variation and adaptation in Africa: implications for human evolution and disease. Cold Spring Harb Perspect Biol. 2014;6(7):a008524.
25. Saiz-Rodriguez M, Valdez-Acosta S, Borobia AM, Burgueno M, Galvez-Mugica MA, Acero J, et al. Influence of Genetic Polymorphisms on the Response to Tramadol, Ibuprofen, and the Combination in Patients With Moderate to Severe Pain After Dental Surgery. Clin Ther. 2021;43(5):e86-e102.
26. Mizuarai S, Aozasa N, Kotani H. Single nucleotide polymorphisms result in impaired membrane localization and reduced atpase activity in multidrug transporter ABCG2. Int J Cancer. 2004;109(2):238-46.
27. Sileshi T MG, Makonnen E, Aklillu E. Effect of Genetic Variations in Drug-Metabolizing Enzymes and Drug Transporters on the Pharmacokinetics of Rifamycins: A Systematic Review. Pharmgenomics Pers Med 2022;15:561-571. 2022.
28. Huerta-Garcia AP, Medellin-Garibay SE, Salazar-Gonzalez RA, Ortiz-Alvarez A, Magana-Aquino M, Rodriguez-Pinal CJ, et al. Anthropometric and Genetic Factors Associated With the Exposure of Rifampicin and Isoniazid in Mexican Patients With Tuberculosis. Ther Drug Monit. 2019;41(5):648-56.
29. Ramesh K, Hemanth Kumar AK, Kannan T, Vijayalakshmi R, Sudha V, Manohar Nesakumar S, et al. SLCO1B1 gene polymorphisms do not influence plasma rifampicin concentrations in a South Indian population. Int J Tuberc Lung Dis. 2016;20(9):1231-5.
30. Ammour A, Aberkane M, Boudjema A, Boughrara W, Benchekor SM. The Frequencies distribution of CYP3A5 rs776746 and ABCB1 rs1045642 polymorphisms in the west Algerian population and relationships with pharmacogenetics. Afr Health Sci. 2024;24(1):307-12.
31. Swart M, Ren Y, Smith P, Dandara C. ABCB1 4036A>G and 1236C>T Polymorphisms Affect Plasma Efavirenz Levels in South African HIV/AIDS Patients. Front Genet. 2012;3:236.
32. Hediger MA, Romero MF, Peng JB, Rolfs A, Takanaga H, Bruford EA. The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteinsIntroduction. Pflugers Arch. 2004;447(5):465-8.
33. Hagenbuch B, Meier PJ. The superfamily of organic anion transporting polypeptides. Biochim Biophys Acta. 2003;1609(1):1-18.
34. Hsiang B, Zhu Y, Wang Z, Wu Y, Sasseville V, Yang WP, et al. A novel human hepatic organic anion transporting polypeptide (OATP2). Identification of a liver-specific human organic anion transporting polypeptide and identification of rat and human hydroxymethylglutaryl-CoA reductase inhibitor transporters. J Biol Chem. 1999;274(52):37161-8.
35. Mohzari YA, Musawa MA, Alrashed A, Zaeri B, Damfu N, Cahusac P, et al. The clinical impact of rifampicin-based anti-TB therapy and tenofovir alafenamide-containing ARV regimen drug Interaction in people living with HIV: Case series report. J Infect Public Health. 2023;16(12):2026-30.

36. Acocella G. Clinical pharmacokinetics of rifampicin. Clin Pharmacokinet. 1978;3(2):108-27.
37. Sanni S, Wachinou A, Merle C, Baba-Moussa L, Affolabi D. SLCO1B1 , NAT2 Polymorphisms and Pharmacokinetic Variability of Rifampicin and Isoniazid in Tuberculosis Patients from Sub-Saharan Africa. Journal of Tuberculosis Research. 2022;10:111-23.
38. Stieger B, Hagenbuch B. Organic anion-transporting polypeptides. Curr Top Membr. 2014;73:205-32.
39. Kalliokoski A, Niemi M. Impact of OATP transporters on pharmacokinetics. Br J Pharmacol. 2009;158(3):693-705.
40. Litjens CHC, van den Heuvel J, Russel FGM, Aarnoutse RE, Te Brake LHM, Koenderink JB. Rifampicin Transport by OATP1B1 Variants. Antimicrob Agents Chemother. 2020;64(10).
41. Shu Y, Brown C, Castro RA, Shi RJ, Lin ET, Owen RP, et al. Effect of genetic variation in the organic cation transporter 1, OCT1, on metformin pharmacokinetics. Clin Pharmacol Ther. 2008;83(2):273-80.
42. Lin L, Yee SW, Kim RB, Giacomini KM. SLC transporters as therapeutic targets: emerging opportunities. Nat Rev Drug Discov. 2015;14(8):543-60.
43. Konig J, Seithel A, Gradhand U, Fromm MF. Pharmacogenomics of human OATP transporters. Naunyn Schmiedebergs Arch Pharmacol. 2006;372(6):432-43.
44. Dompreh A, Tang X, Zhou J, Yang H, Topletz A, Adu Ahwireng E, et al. Effect of Genetic Variation of NAT2 on Isoniazid and SLCO1B1 and CES2 on Rifampin Pharmacokinetics in Ghanaian Children with Tuberculosis. Antimicrob Agents Chemother. 2018;62(3).
45. Weiner M, Peloquin C, Burman W, Luo CC, Engle M, Prihoda TJ, et al. Effects of tuberculosis, race, and human gene SLCO1B1 polymorphisms on rifampin concentrations. Antimicrob Agents Chemother. 2010;54(10):4192-200.
46. Hennig S, Naiker S, Reddy T, Egan D, Kellerman T, Wiesner L, et al. Effect of SLCO1B1 Polymorphisms on Rifabutin Pharmacokinetics in African HIV-Infected Patients with Tuberculosis. Antimicrob Agents Chemother. 2016;60(1):617-20.
47. Medwid S, Deckert R, Gryn SE, Kim RB. SLCO1B1 variants in a patient of African ancestry presenting with rosuvastatin-induced rhabdomyolysis: A case report. Br J Clin Pharmacol. 2025;91(1):232-5.
48. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68-74.
49. Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV, et al. A "silent" polymorphism in the MDR1 gene changes substrate specificity. Science. 2007;315(5811):525-8.
50. Woodahl EL, Hingorani SR, Wang J, Guthrie KA, McDonald GB, Batchelder A, et al. Pharmacogenomic associations in *ABCB1* and CYP3A5 with acute kidney injury and chronic kidney disease after myeloablative hematopoietic cell transplantation. Pharmacogenomics J. 2008;8(4):248-55.
51. Mpeta B, Kampira E, Castel S, Mpye KL, Soko ND, Wiesner L, et al. Differences in genetic variants in lopinavir disposition among HIV-infected Bantu Africans. Pharmacogenomics. 2016;17(7):679-90.

52. Yu N, Chen FC, Ota S, Jorde LB, Pamilo P, Patthy L, et al. Larger genetic differences within africans than between Africans and Eurasians. Genetics. 2002;161(1):269-74.
53. Omotoso OE, Teibo JO, Atiba FA, Oladimeji T, Adebesin AO, Babalghith AO. Bridging the genomic data gap in Africa: implications for global disease burdens. Global Health. 2022 Dec 9;18(1):103. doi: 10.1186/s12992-022-00898-2. PMID: 36494695; PMCID: PMC9733397
54. Wells, G., Hkimi, C., Zass, L., Kumuthini, J. (2026). Precision Medicine of Tuberculosis in Africa. In: Radouani, F., Zass, L., Fadlelmola, F.M. (eds) Precision Medicine in Africa: Promises and Opportunities. Springer, Singapore. https://doi.org/10.1007/978-981-96-8959-0_12

Chapter 3

Preface:

In chapter 2 the review concluded that common genetic variants within the *ABC* and *SLC* gene families found to influence drug levels in European populations are present at significantly different frequencies in African populations. Furthermore, the presence of additional, as-yet unidentified or understudied variants in these genes may contribute to the variability in drug response observed among African individuals. There is a concerning need for further research focused on characterizing both known and novel single nucleotide polymorphisms in African populations, to better understand their impact on drug pharmacokinetics and pharmacodynamics ultimately to enhance the effectiveness of anti-TB treatment outcomes.

In this chapter, we examined how three understudied variants (rs10261685, rs1045642, and rs11045819) in the drug-transporter genes *ABCB1* and *SLCO1B1* influence moxifloxacin pharmacokinetics in South African patients receiving anti-TB therapy, and their potential associations with TB recurrence, HIV-1 acquisition, and variability in drug concentrations. These 3 variants were not available on commercial genotyping microarrays hence our focus on individually investigating these SNPs as potential variants of concern within South African patients. Fortunately we were able to obtain access to patient samples from the CAPRISA study cohorts (TRUTH, IMPRESS & InDEX) which provided an ideal opportunity to explore the genetic diversity of these underrepresented samples with phenotypes of concern arising from KwaZulu-Natal which is known to be the epicentre of both HIV and TB epidemics in South Africa. The aim was to understand how these genetic factors may contribute to disease susceptibility, TB recurrence and association with moxifloxacin drug PK AUC levels at two timepoints that were a month apart. Although no significant associations were found between the three studied SNPs and TB recurrence or HIV-1 acquisition, the data suggests that other genetic polymorphisms across the genome may influence these outcomes. Some SNPs showed modest associations with variability in drug pharmacokinetics, metabolizer status, and potential treatment differences among African individuals. To better understand the clinical relevance

of these findings, larger studies such as Genome-Wide Association Studies (GWAS) with updated pharmacogenomic marker content tailored to African populations are needed.

Potential implications of genetic variants in *ABCB1* and *SLC01B1* on moxifloxacin pharmacokinetics in South African patients infected with tuberculosis.

By Sahil Tulsi [1], Anmol Gokul [1], Anushka Naidoo [2], Kogieleum Naidoo [2], Ravesh Singh [1], Khine Swe Swe Han [1], Nomonde Dlamini [1], Yesholata Mahabeer [1], Tilahun Ferede [1] and Veron Ramsuran [1,2,3]

[1] School of Medicine, College of Health Sciences, University of KwaZulu-Natal, Durban, South Africa

[2] Centre for the AIDS Programme of Research in South Africa (CAPRISA), University of KwaZulu-Natal, Durban, South Africa

[3] International Foundation for Integrative Medical Research, Cape Town, South Africa

*Authors to whom correspondence should be addressed:
sahiltulsi@gmail.com; ramsuranv@ukzn.ac.za
Conflict of interest: None declared.
In submission: *British Journal of Clinical Pharmacology*
Manuscript ID: ecda1f83-b823-4073-b393-70ed6e64b7d3

Abstract

Treatment outcomes are often influenced by several genetic factors in drug transporter genes. In this study we investigated the potential effects of genetic variants in drug transporter genes in *ABCB1 and SLC01B1* on moxifloxacin pharmacokinetics from South African patients receiving anti-TB therapy. We investigated understudied polymorphisms (rs10261685, rs1045642 & rs11045819) that we hypothesized were associated with tuberculosis recurrence, HIV-1 acquisition and impact on drug PK plasma concentration levels. Specific single nucleotide polymorphisms genotypes for *ABCB1& SLC01B1* variants were genotyped using real-time PCR based genotyping SNP assays for rs10261685, rs1045642 & rs11045819. A combination of high-performance liquid chromatography and mass spectrometry was utilized to measure and quantify drug PK levels of moxifloxacin. Genotypes AA/AC of rs10261685 &

AG/GG of rs1045642 were modestly associated with variance in levels of moxifloxacin at specific time points. Genotype CC of rs11045819 had significantly lower drug PK levels (P= 0.0232, R^2= 0.1117) when compared to patients with CA genotype (T2) suggesting individuals genotyped as CC could potentially identify as a rapid metabolizer status group for moxifloxacin. Despite no significant associations with TB recurrence & HIV acquisition and the 3 SNPs were observed in our population subset, the data from this study suggests that there are several other genome-wide polymorphisms that could be influencing recurrence of TB and predisposition to HIV-1 as well as have differential roles in pharmacodynamics compared to other more characterized cohorts. In addition to this it was observed that specific SNPs which were modestly associated with drug PK variability, metabolizer status and possible treatment diversity for some African individuals. Further clinical relevance of the effects of genetic variation on moxifloxacin pharmacokinetic properties need to be better understood in larger studies such as Genome Wide Association studies with updated commercial microarray PGx marker content for African populations.

Introduction

Genetic diversity in African individuals has been observed at a magnitude much larger compared to any other population worldwide (1). Variants unique to African genomes contribute significantly to several factors such as disease predisposition, disease onset and progression, treatment selection and differential response to therapy. Populations such as Europeans and Asians have been well characterised to date, whilst African genetic data is severely underrepresented in genomic databases, contributing approximately less than 2% of global genetic data (2). Tuberculosis and HIV remain among the most common health burdens to patients of African descent. Variants within drug transporter genes from European and Asian cohorts have been found to play crucial roles in HIV pathogenesis and treatment response as well as tuberculosis predisposition and metabolizer status of patients receiving anti-TB therapy (3, 4). Drug transporter gene mutations in African populations need to be further investigated in order to better understand their contribution to infectious disease predisposition, recurrence and treat diversity and patient outcomes. To date several studies have investigated the role of single nucleotide polymorphisms (SNP) that occur in drug transporter genes such as ATP-binding cassettes (ABC) and solute carriers (SC) focusing on their role in TB drug pharmacokinetics across several populations globally. It has been widely observed that single nucleotide mutations occurring in drug transporter genes can impact gene expression leading to the instability of messenger RNA (mRNA), metabolizing of drugs and kinetics of transport

which directly influence drug levels (5). There are several microarray based assays available from global commercial suppliers that investigate these pharmacogenomic (PGx) specific markers in addition to a genome-wide backbone enabling in-depth understanding as to how and why patients respond to treatments based on their genetic make-up. Variants rs10261685, rs1045642 & rs11045819 are not included in ThermoFisher Scientific and Illumina PGx focused commercial microarrays and may be associated with drug metabolism in African populations.

Some studies have particularly investigated the role of *ABCB1* genetic variants such as rs1128503 and rs1045642 in Indonesian populations and reported no significant correlation with TB therapy (6). Similar observations have also been noted in European populations for these *ABCB1* variants. Despite rs1045642 (3435T>C) understood as a silent mutation with variable allele frequencies across global populations (34%-90%), (23, 24,25) and associations with altered function in P-glycoprotein having low expression (p-gp) for the 3435 TT genotype and increased plasma levels of digoxin relative to the 3435 CC genotype have been observed across populations (7) with limited data for African patients warrants the need to further characterize this variant and understand if there is a potential role in pharmacogenetics. Differences in minor allele frequency across global population often poses the need to curiously expand on this research with specific focus on sampling additional African cohorts. With the limited data that represents the lack of clinical associations with TB for these *ABCB1* SNPs in Asian & European populations, our study focused on exploring the possibility of these variants having impact on understudied African populations to determine if they play a role in treatment diversity and outcome. A polymorphism that has been understood to influence drug metabolism and toxicity is rs10261685 however there is limited data for its role in patients of African descent (8). For an understudied SNP such as rs10261685 being a marker that is not commonly included in commercial PGx microarrays warrants further investigation in African populations due to its potential impact on drug levels. Another variant of concern for our study was rs11045819 found in the *SLCO1B1* gene that has been found to influence variable rifampicin PK levels including reduced drug concentrations in patients receiving anti-TB therapy in African populations (9-11). Continued focus on this variant rs11045819 can not only aid in validation of its impact on rifampicin treatment in African patients but help us better understand its impact on other anti-TB drugs such as moxifloxacin.

Over several years moxifloxacin has been administered as a common antibacterial and in some cases ideal for patients who are unable to withstand standardized regimens or at risk of

developing multidrug resistance to anti-TB drugs (12). Moxifloxacin has also been recommended by the WHO for the treating multidrug-resistant tuberculosis (MDR-TB) and has been useful for patients who may develop a level of toxicity towards isoniazid or standard first-line drugs drug-susceptible TB if toxicity develops to standard first-line drugs (12, 13). Moxifloxacin is gaining importance as a central component in the creation of new, shorter rifampicin-sparing treatment regimens for both drug-susceptible and multidrug-resistant tuberculosis (MDR-TB). It is currently being evaluated in multiple ongoing Phase III clinical trials (14, 15).

Whilst moxifloxacin has been administered in several anti-TB treatment regimens due to its highly potent bactericidal activity, studies have suggested that a standardized dose of approx. 400mg may not suffice as an optimal treatment for all patients leading to compromised treatment outcomes at risk of acquiring anti-TB drug resistance (16). Moxifloxacin has been observed to elicit differential interpatient variability in pharmacokinetic (PK) responses across TB patients and healthy individuals (17). Apart from genetic factors that contribute to variance in pharmacokinetics between patients, additional considerations that play influential roles in fluctuating responses are gender, weight, age and co-infection with HIV and antiretroviral drug interactions (18).

In this study we aim to investigate understudied PGx markers (rs10261685, rs1045642 & rs11045819) in African populations with patient samples from three independent clinical trials. The study initially focuses on understanding the roles of these 3 variants in the context of TB recurrence, and then furthering investigation to determine if the variants potentially influence HIV-1 acquisition to any extent. With one of the trials (CAPRISA IMPRESS) based on retreatment success outcomes from moxifloxacin we then interrogated the impact of the three SNPs to determine if they play a role in differential drug PK levels.

Methods

Study design:

For this retrospective study we investigated various sample types (buffy coats from whole blood and sputum) from 571 patients who were enrolled in three independent clinical trials [Improving Retreatment Success (IMPRESS), TB recurrence upon treatment with HAART (TRUTH) & The individualized multi-drug resistant TB treatment strategy (InDEX) conducted by The Centre for the Aids Programme of Research in South Africa (CAPRISA). Table 1 in the supplementary information provided in conjunction with this manuscript summarizes the

objectives of all 3 clinical trials that investigated individual objectives. Reference to study size, aims of each study, research objectives and study outcomes together with their ClinicalTrials.gov ID for reference can be accessed in the supplementary information. Full details on study objectives, inclusion/ exclusion criteria and design are available at ClinicalTrials.gov (NCT01539005, NCT0211468 & NCT03237182).

In parallel to CAPRISA IMPRESS study, a prospective pharmacokinetic (PK) sub study was initiated where blood samples were collected at enrolment and follow up visits from consenting study participants in the PK study. All study participants above 18 years of age with a history of confirmed TB within the last 3 years. Table 1 represents patient demographics from data that was available from each respective clinical trial. A Summary of the respective CAPRISA studies highlighting study size, aims, objectives and outcomes together with ClinicalTrials.gov ID has been presented in Supplementary Table 1.

Table 1: Demographics of patients summarized from the CAPRISA clinical trials (TRUTH, IMPRESS and InDEX).

Study	**TRUTH, n= 401**	**IMPRESS**	**InDEX**
Sample size	**n= 401**	**n= 157**	**n= 13**
Variable			
Age, years, median (IQR)	34 (29-39.5)	34 (24-44)	24 (18-59)
Sex, n(%) male/female	185 (46.13)/ 216(53.86)	103 (65.6)/ 54 (34.3)	9 (64.2)/ 5 (35.7)
African Descent, n(%)	401 (100)	157 (100)	13 (100)
HIV VL Detected, n(%)	48 (11.97)	97 (61.78)	3 (23.07)
Patients on ART, n(%)	44 (10.97)	94 (59.87)	9 (69.23)
TB recurrence, n(%)	48 (11.97)	-	-
Anti TB therapy, n(%)	401 (100)	157 (100)	13 (100)
Hypertension history, n(%)	20 (4.98)	-	-
Diabetes history, n(%)	3 (0.75)	-	-
Cancer history, (n%)	3 (0.75)	-	-

The above listed table represents the participants that were included in this retrospective study with a total of 401, 157 and 13 individuals respectively. Additional variables such as age, sex, race, HIV-1 viral load (HIV VL), number of individuals on antiretroviral treatment (ART), number of patients who presented with clinical TB recurrence, number of TB patients receiving anti-TB therapy, history of hypertension, history diabetes and history of cancer were also data captured and summarized for the purposes of this study.

Sample collection & storage:

TRUTH Study – Sputum samples were induced by each study participant and collected at each visit, transported on ice to the CAPRISA laboratory and stored at -80 degrees Celsius for optimal sample preservation. Venous whole blood specimens were collected in 4ml Ethylenediaminetetraacetic acid (EDTA) tubes

IMPRESS Study - For recruitment of study participants, whole blood samples were collected (24 hours post oral administration) at predefined timepoints (T1 and T2, approximately one month apart) for PK analysis who were receiving moxifloxacin in the intervention arm of the study receiving moxifloxacin as anti-TB therapy. Consenting patients in support of their samples being stored were collected as whole blood samples for downstream genomic testing.

InDEX Study - Sputum samples were induced by each study participant and collected at baseline and each follow up visit, transported on ice to the CAPRISA laboratory and stored at -80 degrees Celsius for optimal sample preservation. Venous whole blood specimens were collected in 4ml Ethylenediaminetetraacetic acid (EDTA) tubes in preparation of peripheral blood mononuclear cells (PBMC) processing.

Laboratory sample processing:

Pharmacokinetic testing – Plasma was isolated from whole blood samples that were collected from patients who were receiving moxifloxacin (24 hours post oral administration) at two predefined timepoints (namely T1 and T2 in this study that were taken approximately one month apart) for PK analysis. Following centrifugation of the EDTA tubes at a centrifugal force of 3000 rpm and cooled on ice prior to being shipped to the CAPRISA laboratory in order to be frozen for storage at -80°C. All plasma samples were shipped within one hour of collection from the clinic to the laboratory. Drug concentrations of Moxifloxacin were measured and quantified in all plasma samples at the pharmacology lab at the KwaZulu-Natal Research institute for tuberculosis and HIV (K-RITH) with the use of high-performance liquid chromatography (HPLC) followed by mass spectrometry (28). United States Food and Drug Administration (US FDA) guidelines (2011) (19) were carefully followed to develop and validate the bioanalytical method applied. Protein precipitation with acetonitrile followed by

dilution with water was carried out for sample preparation. Analytes were separated by chromatography using a Zorbax C18, 3.5 µm, 50 mm × 2.1 mm column. Detection was carried out with the ABI Sciex 5500 (AB Sciex LLC, MA, USA). QTrap mass spectrometer (AB Sciex LLC) was operated in positive ionization mode. The following transitions were used: precursor ion → product ion (all in units of m/z): moxifloxacin: 402.1 → 358.2 and 402.1 → 364.1. Ciprofloxacin was used as an internal standard: 331.6 → 231.0 and 331.6 → 288.1. Samples for quality control measures met the required criteria of FDA bioanalytical validation with precision <15%. Moxifloxacin drug levels were determined with a 22% acetonitrile/water/0.1% formic acid mobile phase. Drug pharmacokinetic (PK) analysis parameters were determined by calculating the area under the concentration–time curve (AUC), together with the maximum observed concentration (Cmax) and trough concentration (Cmin), ensuring the use of standard non-compartmental analysis methods. The volume of injection used was 2µl with an analytical run time of 5 minutes in total. Validation of this method performed over the concentration range of 50–5000 ng/ml. Assay precision was confirmed based on quality control samples that were evaluated at low, medium and high concentrations throughout the validation and analysis of samples ranged from 8.4 to 19.4% and accuracy ranged from 101.9 to 105%. The calculated carry over specifically at the lower limit of quantification was 5.4%. The liquid chromatography–tandem mass spectrometry system was integrated with a PC with Windows 7 operating system interfacing version 1.6.2 of AnalystR software, for chromatographic acquiring of data, integration of peaks and analyte quantification.

SNP Genotyping - Genomic deoxynucleic acid (DNA) was isolated from sputum samples and whole blood samples as per manufacturer's specifications using the QIAmp DNA blood kits (Qiagen, Hilden ,Germany). Isolated DNA samples were standardized to 20ng/µl for downstream processing. Custom Taqman assays were designed for 3 SNPs (rs10261685, rs1045642 & rs11045819), assay IDs referenced in Supplementary table 2 were genotyped using TaqMan SNP Genotyping mastermix (ThermoFisher Scientific, USA) on the QuantStudio 5 Real-time PCR System (Applied Biosystems, USA). Secondary data analysis and genotype calling was determined using the TaqMan Genotyper version 1.7.1 software (ThermoFisher Scientific, USA). Real-time PCR reactions were optimized with the following cycling conditions for 5µl reactions: Pre read of 60°C for 30 seconds, initial denaturation at

95°C for 10 minutes, with 40 cycles of 95°C for 15 seconds, annealing and extension at 60°C for 1 minute concluding with a post-read step that entailed 60°C for 30 seconds.

Ethical considerations

The Biomedical Research Ethics Committee of the University of KwaZulu-Natal granted expedited ethical approval for this study (BREC/00004110/2022)

Statistical Analysis

Fisher's exact tests were performed to determine the association between SNPs and TB recurrence using allele frequencies. Chi-Square Tests were utilized to compare the allele or genotype frequencies of the same 3 SNPs between two different groups (TB+HIV+ vs. TB+HIV-). Comparison of Drug PK data (AUC values) for each SNP at two different time points we used the Wilcoxon Signed Rank Test. All figures represented in this research was generated using GraphPad prism 9.3.2.

Results

Baseline data:

In this retrospective study we included a total of 571 TB positive patients of Black South African ancestry, 297 (52.8%) were male and 302 (52.7%) patients had detectable HIV-1 viral loads. Additional comorbidities such as history of diabetes (1.04%), hypertension (3.49%) and cancer (0.5%) were recorded at patient enrolment for the TRUTH study.

Genotype & Allele frequencies:

Genotype statistics of the SNPs for this study were very much aligned with the minor allele frequencies captured by the Allele Frequency Aggregator (ALFA) genomic database and summarised in Table 2. For all three SNPs (rs10261685, rs1045642 & rs11045819), samples that were successfully genotyped with their respective genotype frequencies are reported in Table 4. for all patients from the TRUTH study where we thereafter investigated the association of these frequencies with TB recurrence, co-infection with HIV-1 and PK drug variability. Patient samples that yielded undetermined genotyped were omitted from data analysis.

Table 2: Minor allele frequencies (MAFs) for the genetic variants rs10261685, rs1045642, and rs11045819 are summarised across major global population groups, including Asian and European cohorts, as well as African and African American (Afro-American) populations.

rsID	Minor Allele Frequency (MAF)			
	Asians	Europeans	African	African American
rs10261685, A>C	C (0.073)	C (0.026)	C (0.083)	C (0.082)
rs1045642, A>C/G/T	A (0.391)	G (0.478)	G (0.214)	G (0.217)
c, C>A/G/T	A (0.142)	A (0.158)	A (0.074)	A (0.075)

Frequency estimates were obtained from the Allele Frequency Aggregator (ALFA) database, which compiles high-quality allele frequency data from multiple large-scale genomic studies. This table provides a comparative overview of population-specific variation for each single nucleotide polymorphism (SNP), allowing assessment of inter-population differences that may be relevant to the genetic associations and pharmacogenetic in this analyses in this study.

Table 3: Genotype frequencies for the ABCB1 gene variants rs10261685 and rs1045642, as well as the SLCO1B1 gene variant rs11045819, obtained from patients enrolled in the TRUTH study.

Gene	SNP, Mutation	Sample no, n=	Genotype	Genotype Frequency
ABCB1	rs10261685, A>C	378	AA	0.84
			AC	0.16
			CC	0
	rs1045642, 3435C>T	390	AG	0.18
			GG	0.82
			AA	0
SLC01B1	rs11045819, C>A/G/T	399	AA	0
			CA	0.12
			CC	0.85
			CT	0.03

The data presented in this table highlights the distribution of homozygous reference, heterozygous, and homozygous variant genotypes for each single nucleotide polymorphism (SNP) within the study cohort. These frequencies reflect the genetic variation present in the sampled patient population. The data provide a basis for evaluating potential associations between these variants with potential pharmacogenomic relevance and clinical outcomes assessed within the TRUTH study.

To investigate the association of these 3 variants with TB recurrence from the CAPRISA TRUTH study, 401 study participants (214 female, 53%) were initially included in this retrospective study analysis to determine the link between drug transporter gene *ABCB1* & *SLC01B1* (rs10261685, rs1045642 & rs11045819) and the likelihood of increased susceptibility to TB recurrence downstream. Single nucleotide polymorphism referenced as rs11045819r is a generic rsID used for the purpose of this study since rs11045819 (often represented as an average) is a tri-allelic assay (C>A/G/T) which was split over 2 individual SNP assays (rs11045819 & rs11045819r) during genotyping by real-time PCR. Forty-eight individuals (24 male, 50%) were found to have recurrent TB at a later stage during their study participation (time point 2).

Genotype frequencies are presented in Table 4. For each SNP all genotype frequencies (Figure 1.) were in Hardy-Weinberg equilibrium, however no significant associations were observed with a particular SNP and TB recurrence after performing Chi-Square tests on GraphPad Prism software v9.2.0. Allele frequencies were also investigated and represented in Figure 2., although there were no associations observed that predisposed the patient to recurring TB at a later stage. To determine the impact of allele distribution and if there is a link to predicting predisposition to HIV acquisition, allele frequencies are presented in Figure 3. Chi-Square tests were performed however no significant associations of alleles present in rs10261685, rs1045642 & rs11045819 variants were observed.

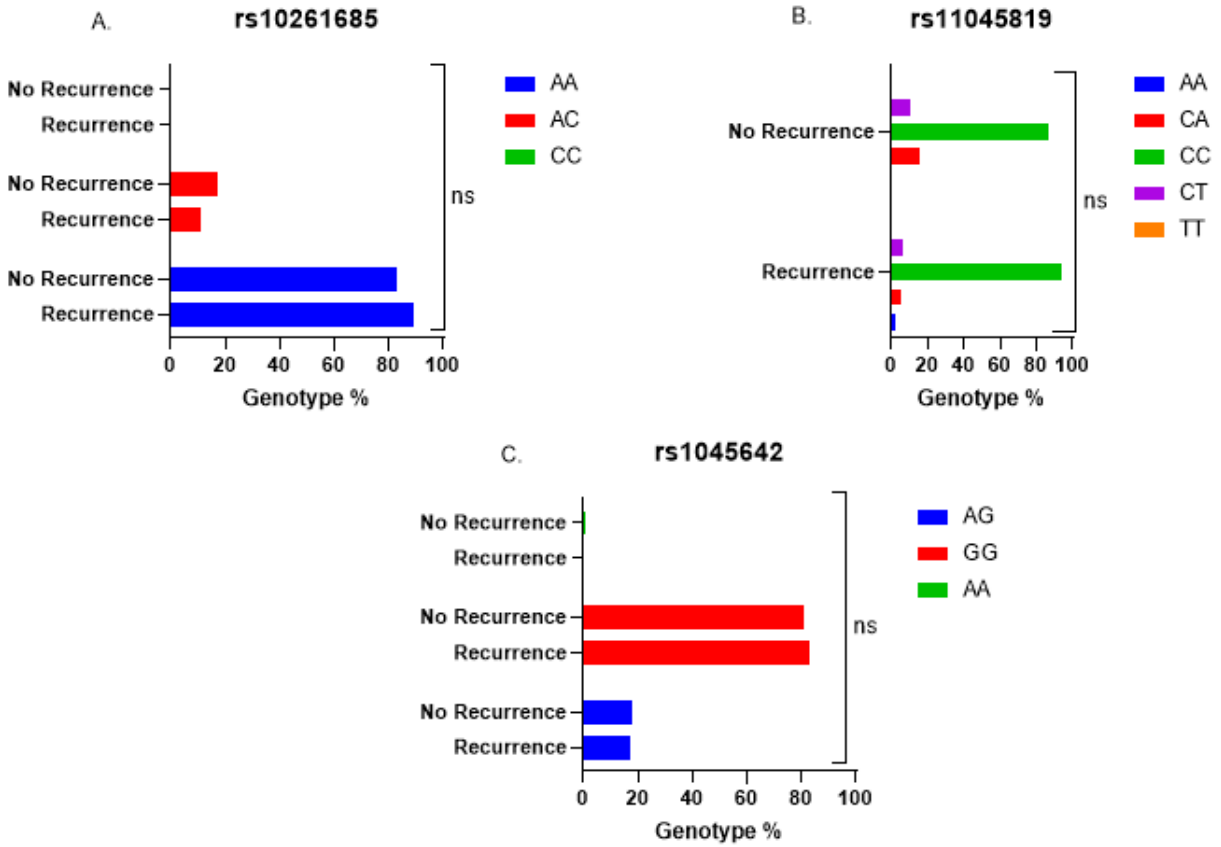


Figure 1. Genotype frequencies of rs10261685, rs11045819 & rs1045642 in TRUTH patients grouped into recurrence compared to no recurrence subsets with no significant differences observed. Genotype distributions were compared between the two clinical groups using chi-squared tests. A) AA/AC/CC Genotype percentages for rs10261685 observed for recurrence [AA, n= 41; AC, n= 5; CC, n= 0] & no-recurrence groups [AA, n= 276; AC, n= 56; CC, n= 0]. B) AA/CA/CC/TT genotype frequencies for rs11045819 observed for both recurrence [AA, n= 1; CA, n= 2; CC, n= 39; TT, n= 0] & no-recurrence groups [AA, n=0; CA, n= 45; CC, n= 229; TT, n= 0] C) AG/GG/AA genotype frequencies for rs1045642 observed for recurrence [AG, n= 8; GG, n=38; AA, n=0] & no-recurrence data groups [AG, n= 61; GG, n= 280; AA, n=3].

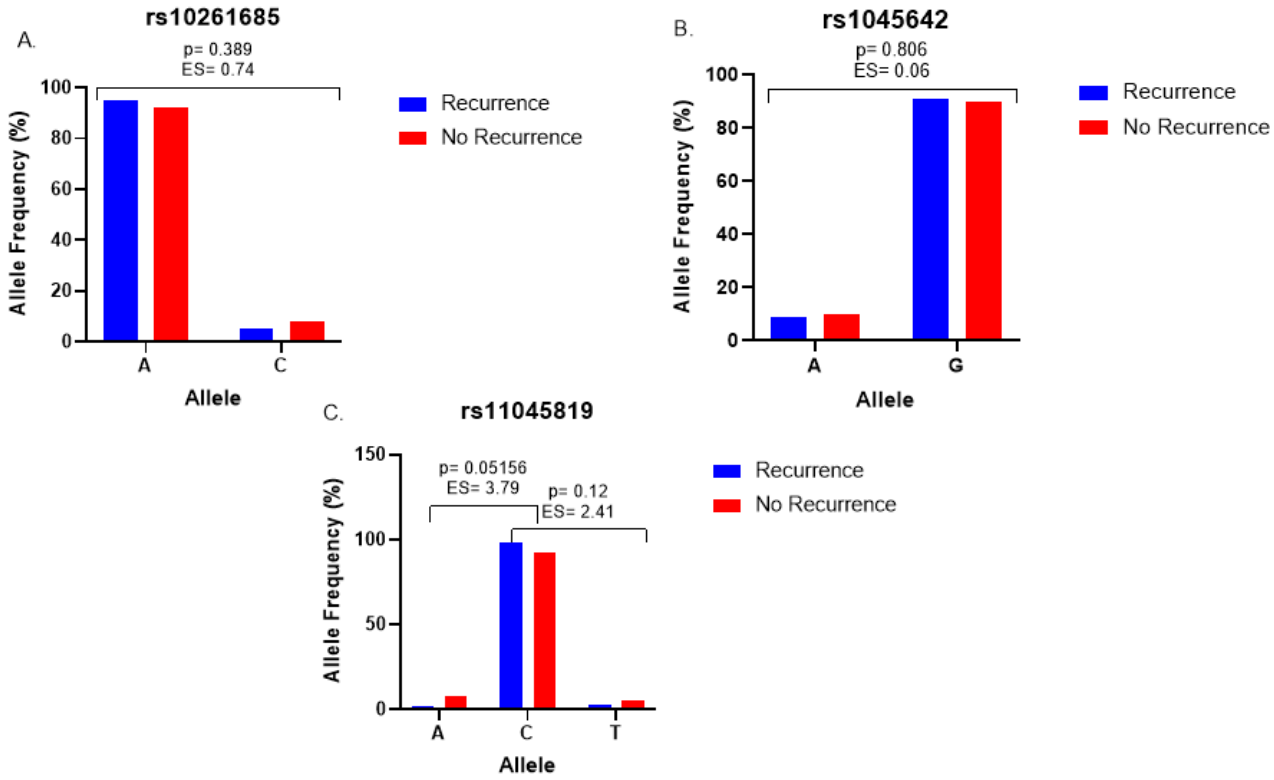


Figure 2. Allele frequencies of rs10261685, rs1045642, and rs11045819 in TRUTH patients, stratified into recurrence and no-recurrence subsets with no significant differences observed. The figure illustrates the distribution of allelic variants across these clinical groups, allowing visualization of potential differences (p= p value, ES= Effect Size) associated with recurrence status. Statistical comparisons of allele frequencies between the two subsets were performed using chi-squared tests to evaluate whether observed variations deviated significantly from expected distributions. A) Chi-square test results for A/C allele frequencies of rs10261685 based on recurrence & no-recurrence groups. B) Chi-square test results for A/G allele frequencies of rs1045642 based on recurrence & no-recurrence groups. C) Chi-square test results for A/C/T allele frequencies of rs11045819 based on recurrence & no-recurrence groups.

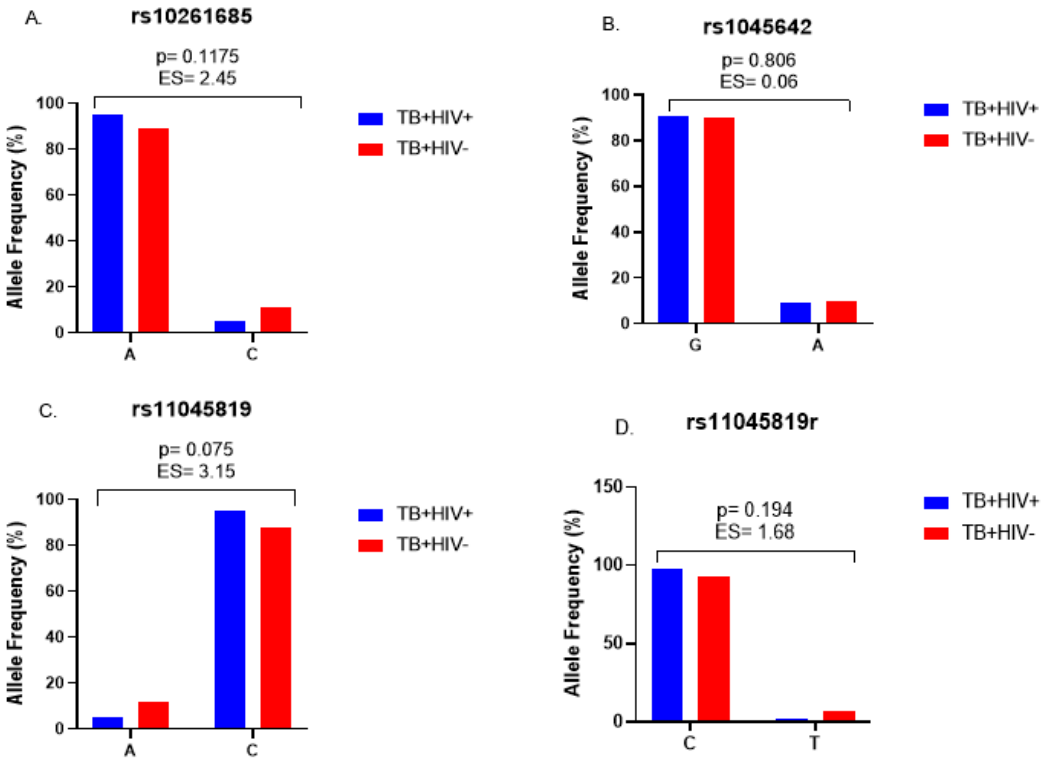


Figure 3. Allele frequencies of the rs10261685, rs1045642, and rs11045819 variants in TRUTH patients, comparing individuals who were co-infected with HIV to those who did not seroconvert. The figure displays the distribution of allelic variants across these two groups to highlight potential differences associated with HIV co-infection status. Statistical comparisons of allele frequencies were conducted using chi-squared tests to assess whether observed variations differed significantly from expected distributions, (p= p value, ES= Effect Size) A) Chi-square results for allele frequencies of A/C for rs10261685. B) Chi-Square results for allele frequencies G/A of rs1045642. C) Chi-square results of allele frequencies A/C for rs11045819. D) Chi-square results of allele frequencies C/T for rs11045819r

To understand the potential influence of SNPs rs10261685, rs1045642 & rs11045819 on drug PK levels were investigated from 46 study participants from the CAPRISA IMPRESS study with drug PK data available from two time points with samples collected on average one month apart. We then performed individual genotype analysis for moxifloxacin drug PK levels were then performed for all SNPs of interest (rs10261685, rs1045642 & rs11045819). Figure 4 represents moxifloxacin levels measured for genotypes of each respective SNP at the two time points (T1 & T2) where drug PK levels were recorded. Both paired and unpaired t-tests were performed to compare the impact each genotype had on moxifloxacin levels.

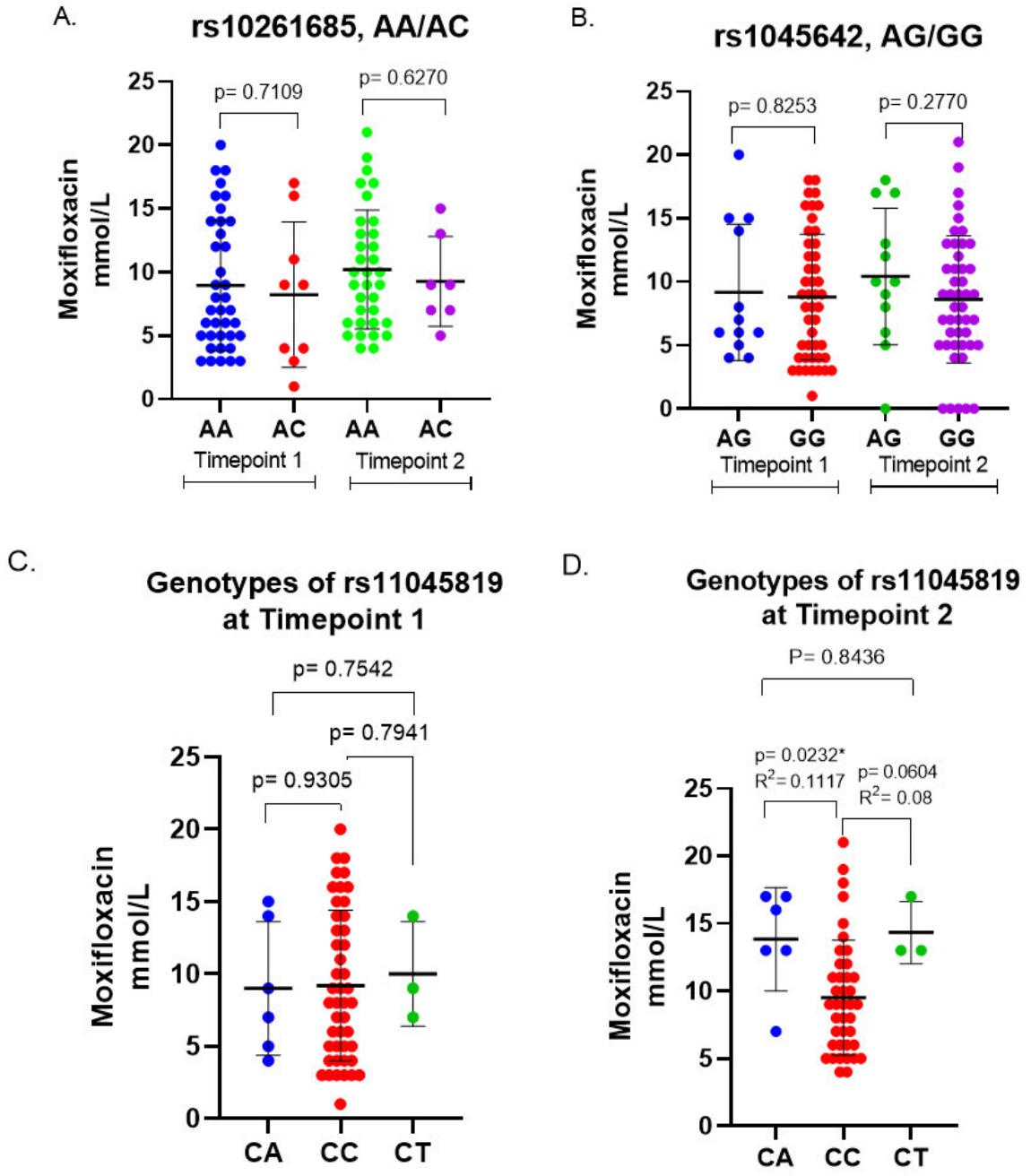


Figure 4. Drug pharmacokinetic (PK) analysis of single nucleotide polymorphism (rs10261685, rs1045642, and rs11045819) showing the respective genotypes and corresponding moxifloxacin plasma levels of patients (n=58) measured at two timepoints [Timepoint 1 (T1) and Timepoint 2 (T2)] with a significant difference observed for rs11045819 at T2 between genotypes CA/CC (p= 0.0232, R^2= 0.1117). This figure illustrates potential differences in drug exposure associated with each genotype. Statistical comparisons between groups were performed, with p-values (p) indicating the observed differences. A) Moxifloxacin drug PK levels for genotypes AA/AC of rs10261685 between T1 & T2. B) Moxifloxacin drug PK levels for genotypes AG/GG of rs1045642 between T1 & T2. C) Moxifloxacin drug PK levels for genotypes CA/CC/TT of rs11045819 at T1. D) Moxifloxacin drug PK levels for genotypes CA/CC/CT of rs11045819 observed at timepoint 2.

We then analyzed the influence on moxifloxacin of each genotype and timepoint of each respective SNPs. A combination of paired and unpaired t-tests were once again performed for rs10261685, rs1045642 and rs11045819 and its 3 genotypes. For SNP rs10261685 no significant differences were observed for genotypes AA vs AC at T1 (P= 0.7109) and T2 (0.6270), however patients with rs10261685 genotype AA at both T1 and T2 had slightly higher moxifloxacin levels compared to patients with the genotype AC at both time points. To date no clinical associations have been published in ClinVar suggesting that genotype AA of rs10261685 may possibly be a variant of uncertain significance (VUS) . A similar trend was observed for rs1045642, genotypes AG vs GG showed no significant differences at T1 (P= 0.8253) and T2 (0.2770), with patients having genotype AG resulting in slightly higher moxifloxacin levels at both time points as compared to the GG genotype patients . Variant rs11045819 represented in Figure 5.C focuses on genotypes (CA, CC, CT). Individual unpaired T-tests were performed to observe the genotype driving the resulted p-values. Between genotype CA vs CC at T1 no significant differences were observed (p= 0.9305) for variant rs11045819, the same were observed for genotype CC vs CT (p= 0.7941) and genotypes CA vs CT. (p= 0.7542) at T1. However time point 2 analysis of rs11045819 presented in Figure 5.D revealed that patients with genotype CC has significantly lower moxifloxacin levels compared to patients with genotype CA yielding a p-value = 0.0232, R^2= 0.1117. A very similar trend was observed with patient genotypes CC vs CT at T2 for rs11045819 with noticeably lower moxifloxacin levels observed in CC genotype patients compared to genotype CT patients however the difference noted in this comparison was not significant (p= 0.0604, R^2= 0.08). Genotypes CA vs CT for rs11045819 were also compared at T2 with no significant differences observed (p= 0.8436).

Discussion

African genomes are far more diverse than any other population (20), to an extent that standardized treatment regimens and dosing therapies may lead to suboptimal patient outcomes for tuberculosis. Despite treatment regimens not being changed for the last 50 years (21), there has been very little focus on gaining a better understanding of what makes patients of African descent unique in a sense of treatment diversity. In this study we investigated the possibility of determining the contributing factors for TB recurrence with respect to the effect of genetic polymorphisms in ABC and SC drug transporter genes. Genetic diversity in patients of African descent based on literature published could imply that further whole genome sequencing of African genomes is warranted in the context of better understanding disease predisposition and recurrence in African populations. Africa as a diseased burdened continent can only benefit from future genetic disease research in light of these research questions to the best interest of African patients response to treatment. Shifting focus from biomarkers representing association with recurrence and predisposition, interestingly the data generated in this study showed modest association with the SNPs of interest and pharmacokinetic drug variability.

With initial analysis of comparing moxifloxacin drug PK levels with median AUC values, we noticed that T2 values were only slightly higher than T1 data points. This minor difference ($p= 0.4359$, $R^2= 0.007$) represented in supplementary figure 1 suggests that the majority of patients had adhered to treatment and possibly most drug transporter genes have been functioning as expected to, however leaves room for investigation that there were some patients with decreased drug PK levels at T2. Assuming consistent adherence of anti-TB therapy, patients with lower levels of moxifloxacin at the second timepoint but higher concentrations at the initial timepoint could possibly be regarded as fast metabolizers and would need to explored in a future study. For patients who could possibly be fast metabolizers, there may be a need for them to receive higher dosages of anti-TB therapy in order for them to ensure that the treatment is optimal. These results possibly suggest that variance in genotypes for each SNP, or other novel SNPs may be contributing to some patients having lower levels after a month of initial moxifloxacin administration along with other epigenetic factors that could possibly impact this observation.

Further analysis of each genotype for the respective SNPs (rs10261685, rs1045642 & rs11045819) was performed to determine if some genotypes may play an influential role and impact the drug PK levels of moxifloxacin. For rs10261685 we observed that patients with

genotype AA at T2 had slightly elevated levels of moxifloxacin compared to patients with genotype AC at T2. Whilst the difference is minimal it is possible that AA genotyped patients are slow metabolizers compared to genotype AC patients. In a larger subset of samples it would be more likely to observe and note a significant difference for further investigation. To date there is limited research citing rs10261685 or clinical observations observed in African populations posing this SNP as a potential variant of uncertain significance (VUS) that should be explored in further studies. A similar observation was identified for rs1045642, with genotype AG being associated with higher moxifloxacin levels compared to GG genotypes at T2. Interestingly variant rs1045642 was reported to play a differential role in a study performed in South Korea where patients with the AA Genotype showed a decreased likelihood of postoperative nausea and vomiting (PONV) when treated with ondansetron as compared to genotypes AG + GG (26). In this study genotype AA for rs1045642 was only observed at a frequency of 1% in patients who did not present themselves with clinical TB recurrence (Figure 1.C) once again reiterating the diversity of African populations that requires further investigation whilst these SNPs are present at larger genotypic frequencies in other populations with defined associations.

With regards to analysis of variant rs11045819 we chose to rather utilize individual T-tests to highlight the respective comparisons driving significant differences among the 3 possible genotypes (CA/CC/CT). From the analysis presented in Figure 4.D the data suggests that patients with a CC genotype had significantly lower drug PK levels of moxifloxacin ($p= 0.0232$, $R^2= 0.1117$) as compared to patients who genotyped as CA. Even though a similar trend was observed when CC genotyped patients moxifloxacin levels were compared to patients with a CT genotype, the difference was not significant but possibly due to the small sample number in the CT genotyped group. Nevertheless this data finding suggests that patients with CA and CT genotypes can be regarded as slow metabolizers when compared to patients with a CC genotype for rs11045819 who are more likely to be classified as rapid metabolizers at T2 . If we relate to the observed moxifloxacin drug levels of rs11045819, genotype CC at T1 which was recorded approximately 1 month before T2 PK data was collected, it provides insight that within the space of a month these CC genotyped patients rapidly metabolized the moxifloxacin drug which explains the lower PK levels observed at T2 for CC genotyped patients as opposed to CA and CT genotyped individuals with higher moxifloxacin levels. In light of this finding our data suggests that patients with CC genotype in rs11045819 can be considered as rapid metabolizers for moxifloxacin and an exploratory marker with potential relevance pending confirmation in larger studies with greater statistical

significance for receiving a larger dose to ensure optimal treatment outcomes. In contrast a study conducted in the (United States) US reported that Caucasian and Afro-American patients who genotyped as AC were observed to have associations with an increased clearance of rifampin in healthy individuals as compared to patients that genotyped as CC for rs11045819 (27), which was the opposite effect observed in our moxifloxacin drug PK focused study for CC genotyped patients suggesting that these SNPs and genotypes play differential roles in patients of purely African descent.

Data as such for variants investigated in this study (rs10261685, rs1045642 & rs11045819) may be posed as potential PGx markers and should be considered for addition to future revisions of commercial arrays to benefit African populations in terms of understanding treatment diversity. It is becoming increasingly common for researchers who choose to perform genotyping by microarray as an alternative to costly next-generation sequencing, the addition of these markers can only benefit African populations once more samples are screened to gain statistical power and address the lack of genomic data in international databases.

Moxifloxacin is considered a drug that is rapidly absorbed when administered orally with an average half-life of 8.2-15 hours for patients who are considered healthy (22). Metabolizer status is of huge concern in determining patient outcomes in the context of optimal treatment. Data as such is supportive that these genetic variants in drug transporter genes could be playing a crucial role in the treatment of patients as well as the rate of antibacterial activity in combating *Mycobacterium tuberculosis* from a hosts perspective. For the remainder of genotypes for SNPs rs10261685 and rs1045642 and their impact on differential yet minimal drug PK levels these findings may indicate that these SNPs may not be playing as much of a pharmacokinetic role in African populations as compared to other global populations, leaving more room for future studies to further explore other novel or understudied variants relative to African patients and their impact on anti-TB therapy.

With consideration of study limitations which may require a larger study population for screening to confirm such findings, there may be a possibility that there are other genes which are yet to be investigated that could be playing a crucial role in recurrence. Without ruling out other potential influence of antiretrovirals, patient age, gender among several confounding factors could supplement this argument. Human immunodeficiency virus (HIV) is known to more likely infect individuals who exposed and are immunocompromised with comorbidities such as tuberculosis. Participants within the CAPRISA TRUTH study were also analyzed in

the context of predisposition to HIV-1. With a hypothesis specific to this study that drug transporter gene variants may potentially increase a patients risk in acquiring HIV-1 whilst co-infected with tuberculosis, we also interrogated the genotype dataset in search of understanding the possible link of these variants to HIV-1 predisposition. Despite no significant findings leading to concrete evidence that these SNPs play a role in HIV-1 susceptibility, this leaves much room for debate on the possibility of alternate genetic factors in other drug transporter genes that are associated with HIV-1 infection.

Conclusion

Genetic variants specific to patients of African descent are known to be far more diverse compared to other global populations. This study has observed the prevalence of polymorphisms in drug transporter genes *ABCB1* and *SLC01B1* from three independent clinical trials in South Africa. Whilst the investigated SNPs have been noted for associations in disease trends, patient treatment specific outcomes and drug transporter functions in other populations, none were observed in the context of the South African study participants in association with TB recurrence and susceptibility to HIV-1 co-infection. However in contrast we observed modest significant differences in moxifloxacin drug concentrations and PK parameters for genotypes. Larger studies need to further investigate the clinical relevance of these observed changes in moxifloxacin drug exposure for TB treatment outcomes. Moxifloxacin has been noted for its dose dependant potency and anti-bacterial properties, therefore additional research can only strengthen the suggestion that any genetic variation is probable to have effects on drug efficacy.

Ratios of PK–PD data points such AUC & MIC of values greater than 53 up to 100 have been previously reported to be associated with successful outcomes and minimizing the onset of anti-TB drug resistance. On the contrary in this retrospective study moxifloxacin drug concentrations observed from South African trial participants were observed at much lower levels. In summary these findings suggest that drug transporter gene variants associated with fluctuating drug levels in SA patients may potentially modulate the patients ability to maintain average expected AUC and MIC data values. Africa is diverse from a genomic perspective, this study suggests that there are several SNPs both known and novel that are yet to be identified as key PGx markers with potential to be implemented into future microarray-based designs to further our understanding of pharmacogenomics in the context of African populations.

Funding statement

Research reported in this review article was supported by Grants, Innovation and Product Development unit of the South African Medical Research Council with funds received from Novartis and GSK Global Health for Project Africa GRADIENT (Grant No. GSKNVS2/202101/005).

References

1. Gomez F, Hirbo J, Tishkoff SA. Genetic variation and adaptation in Africa: implications for human evolution and disease. Cold Spring Harb Perspect Biol. 2014;6(7):a008524.
2. Fatumo S, Yakubu A, Oyedele O, Popoola J, Attipoe DA, Eze-Echesi G, et al. Promoting the genomic revolution in Africa through the Nigerian 100K Genome Project. Nature Genetics. 2022;54(5):531-6.
3. Ma MK, Woo MH, McLeod HL. Genetic basis of drug metabolism. Am J Health Syst Pharm. 2002;59(21):2061-9.
4. Wilson JF, Weale ME, Smith AC, Gratrix F, Fletcher B, Thomas MG, et al. Population genetic structure of variable drug response. Nat Genet. 2001;29(3):265-9.
5. Sissung TM, Troutman SM, Campbell TJ, Pressler HM, Sung H, Bates SE, et al. Transporter pharmacogenetics: transporter polymorphisms affect normal physiology, diseases, and pharmacotherapy. Discov Med. 2012;13(68):19-34.
6. Barliana MI, Kusuma ASW, Insani WN, Alfian SD, Diantini A, Mutakin M, et al. Genetic variation of ABCB1 (rs1128503, rs1045642) and CYP2E1 rs3813867 with the duration of tuberculosis therapy: a pilot study among tuberculosis patients in Indonesia. BMC Res Notes. 2021;14(1):295.
7. Hodges LM, Markova SM, Chinn LW, Gow JM, Kroetz DL, Klein TE, et al. Very important pharmacogene summary: ABCB1 (MDR1, P-glycoprotein). Pharmacogenet Genomics. 2011;21(3):152-61.
8. Richardson M, Kirkham J, Dwan K, Sloan DJ, Davies G, Jorgensen AL. Association of variants within the GST and other genes with anti-tubercular agents related toxicity: a systematic review and meta-analysis. bioRxiv. 2019:515817.
9. Chigutsa E, Visser ME, Swart EC, Denti P, Pushpakom S, Egan D, et al. The SLCO1B1 rs4149032 polymorphism is highly prevalent in South Africans and is associated with reduced rifampin concentrations: dosing implications. Antimicrob Agents Chemother. 2011;55(9):4122-7.
10. Gengiah TN, Botha JH, Soowamber D, Naidoo K, Abdool Karim SS. Low rifampicin concentrations in tuberculosis patients with HIV infection. J Infect Dev Ctries. 2014;8(8):987-93.

11. Weiner M, Peloquin C, Burman W, Luo CC, Engle M, Prihoda TJ, et al. Effects of tuberculosis, race, and human gene SLCO1B1 polymorphisms on rifampin concentrations. Antimicrob Agents Chemother. 2010;54(10):4192-200.
12. Gillespie SH. The role of moxifloxacin in tuberculosis therapy. Eur Respir Rev. 2016;25(139):19-28.
13. WHO. Global Tuberculosis Report 2016. 2016.
14. Dawson R, Diacon AH, Everitt D, van Niekerk C, Donald PR, Burger DA, et al. Efficiency and safety of the combination of moxifloxacin, pretomanid (PA-824), and pyrazinamide during the first 8 weeks of antituberculosis treatment: a phase 2b, open-label, partly randomised trial in patients with drug-susceptible or drug-resistant pulmonary tuberculosis. Lancet. 2015;385(9979):1738-47.
15. Moodley R, Godec TR, Team ST. Short-course treatment for multidrug-resistant tuberculosis: the STREAM trials. Eur Respir Rev. 2016;25(139):29-35.
16. Gumbo T, Louie A, Deziel MR, Parsons LM, Salfinger M, Drusano GL. Selection of a moxifloxacin dose that suppresses drug resistance in Mycobacterium tuberculosis, by use of an in vitro pharmacodynamic infection model and mathematical modeling. J Infect Dis. 2004;190(9):1642-51.
17. Weiner M, Gelfond J, Johnson-Pais TL, Engle M, Peloquin CA, Johnson JL, et al. Elevated Plasma Moxifloxacin Concentrations and SLCO1B1 g.−11187G>A Polymorphism in Adults with Pulmonary Tuberculosis. Antimicrobial Agents and Chemotherapy. 2018;62(5):10.1128/aac.01802-17.
18. Naidoo A, Naidoo K, McIlleron H, Essack S, Padayatchi N. A Review of Moxifloxacin for the Treatment of Drug-Susceptible Tuberculosis. J Clin Pharmacol. 2017;57(11):1369-86.
19. FDA. Guidance for industry bioanalytical method validation. Department of Health and Human Services, Food and Drug Administration. 2011.
20. Sibomana O. Genetic Diversity Landscape in African Population: A Review of Implications for Personalized and Precision Medicine. Pharmgenomics Pers Med. 2024;17:487-96.
21. Health SADo. South African National Tuberculosis Management Guidelines. 2014.
22. Moise PA, Schentag JJ. Pharmacokinetic and pharmacodynamic modelling of antibiotic therapy. Curr Opin Infect Dis. 1998;11(6):673-80.
23. Schwab M, Eichelbaum M, Fromm MF. Genetic polymorphisms of the human MDR1 drug transporter. Annu Rev Pharmacol Toxicol. 2003;43:285-307. doi: 10.1146/annurev.pharmtox.43.100901.140233. Epub 2002 Jan 10. PMID: 12359865.
24. Dey S. Single nucleotide polymorphisms in human P-glycoprotein: its impact on drug delivery and disposition. Expert Opin Drug Deliv. 2006 Jan;3(1):23-35. doi: 10.1517/17425247.3.1.23. PMID: 16370938.
25. Sakaeda T, Nakamura T, Okumura K. Pharmacogenetics of MDR1 and its impact on the pharmacokinetics and pharmacodynamics of drugs. Pharmacogenomics. 2003 Jul;4(4):397-410. doi: 10.1517/phgs.4.4.397.22747. PMID: 12831320.
26. Choi EM, Lee MG, Lee SH, Choi KW, Choi SH. Association of ABCB1 polymorphisms with the efficacy of ondansetron for postoperative nausea and vomiting. Anaesthesia. 2010 Oct;65(10):996-1000. doi: 10.1111/j.1365-2044.2010.06476.x. PMID: 20707787.
27. Kwara A, Cao L, Yang H, Poethke P, Kurpewski J, Tashima KT, Mahjoub BD, Court MH, Peloquin CA. Factors associated with variability in rifampin plasma pharmacokinetics and the relationship between rifampin concentrations and induction of efavirenz

clearance. Pharmacotherapy. 2014 Mar;34(3):265-71. doi: 10.1002/phar.1388. Epub 2014 Jan 13. PMID: 24420746; PMCID: PMC3947488.

28. Upendra P. Dahal, Jeffrey P. Jones, John A. Davis, Dan A. Rock,Small Molecule Quantification by Liquid Chromatography-Mass Spectrometry for Metabolites of Drugs and Drug Candidates,Drug Metabolism and Disposition,Volume 39, Issue 12,2011,Pages 2355-2360, ISSN 0090-9556.

Supplementary information

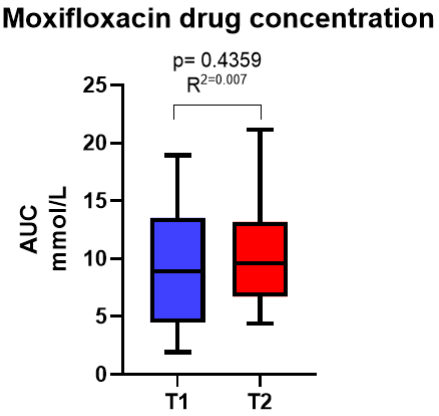
Supplementary Table 1: Summary of respective CAPRISA studies highlighting study size, aims, objectives and outcomes together with ClinicalTrials.gov ID for reference

Clinical Trial	Year study commenced	Cohort Size at enrollment	Study aim/ objective	Outcomes	ClinicalTrials.gov ID
TB Recurrence upon Treatment with HAART (TRUTH)	2009	550	To determine the extent of and reasons for relapse and re-infection in incident cases of tuberculosis (TB) in Human Immunodeficiency Virus (HIV)-infected patients on Highly Active Antiretroviral Therapy (HAART)	The primary endpoint of this study will be the development of recurrent TB. TB recurrence as a result of relapse will be defined as isolates of *M. tuberculosis* from the first and second episodes of TB which cluster in Restriction Fragment Length Polymerization (RFLP) analysis. TB recurrence as a result of re-infection will be defined as isolates from the first and second episode of TB which differ on RFLP analysis. Each study participant who has recurrent TB will be assessed by Interferon (IFN) gamma ELISPOT assay to compare cytotoxic lymphocyte (CTL) immune responses in TB relapse and re-infection.	NCT01539005
IMPROVING RETREATMENT SUCCESS (IMPRESS)	2015	330	The primary objective is to determine if a Moxifloxacin-containing regimen, [Isoniazid (H), Rifampicin (R), Pyrazinamide (Z), Moxifloxacin (M)], substituting Moxifloxacin for Ethambutol, of 24 weeks duration is superior to a control regimen [Isoniazid (H), Rifampicin(R), Pyrazinamide (Z), Ethambutol (E)] of 24 weeks duration in improving treatment outcomes in patients with recurrent TB.	The IMPRESS study concluded that replacing ethambutol with moxifloxacin in the treatment regimen for recurrent, drug-sensitive pulmonary TB did not significantly improve culture conversion rates or overall treatment success. Additionally, the substitution was associated with a higher incidence of adverse events. These findings suggest that moxifloxacin may not be a suitable replacement for ethambutol in this context.	NCT0211468
The Individualized M(X) drug-resistant TB Treatment Strategy Study, InDEX Study	2016	300	The overall aim of this study is to evaluate treatment success in patients with drug-resistant TB using a gene-derived individualized regimen versus the SOC.	the combination of these interventions did indeed result in a significant reduction in HIV incidence. Participants who received the interventions had a lower risk of acquiring HIV compared to those who did not, highlighting the potential of integrated approaches in HIV prevention.	NCT03237182

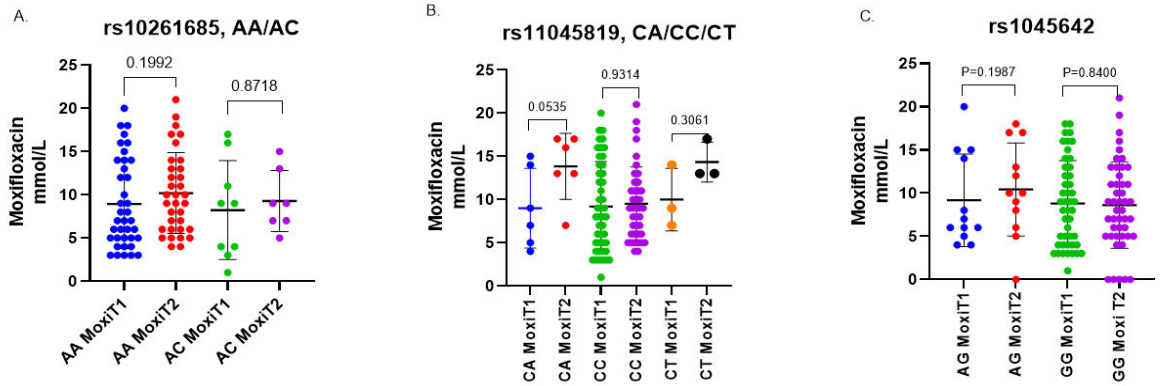
Supplementary Table 2: SNP Assay IDs and Context sequences used for this study

Gene	SNP	Assay ID/ Custom ID	Context Sequence [VIC/FAM]
ABCB1	rs10261685 [A/C]	C__33986526_10	TCACTGTGAGCAAAAAGTTTGACTT[A/C]AGTAAAAGGCAGCAAATTCAATGCA
ABCB1	rs1045642 [A/G]	C___7586657_20	TGTTGGCCTCCTTTGCTGCCCTCAC[A/G]ATCTCTTCCTGTGACACCACCCGGC
SLCO1B1	rs11045819[C/A]	AN7D7VU	-
SLCO1B1	rs11045819r[C/T]	AN9H2FR	-

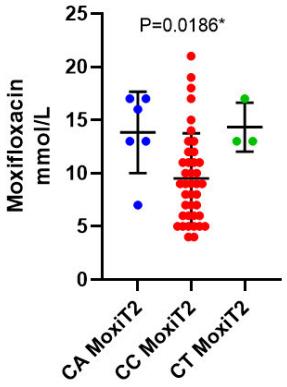
Supplementary Figure 1: Plasma moxifloxacin concentrations (mmol/L) measured at two distinct time points (T1 and T2), n=58. Comparison of drug levels between the two time points indicates consistent exposure, supporting evidence of routine adherence to the prescribed treatment regimen. Statistical analysis of the paired measurements was performed with unpaired t-tests.



Supplementary Figure 2: Drug PK Analysis of single nucleotide polymorphisms (rs10261685, rs1045642 & rs11045819) with respective genotypes (n=58) and moxifloxacin levels between timepoint 1 and timepoint 2. A) Moxifloxacin drug PK levels for genotypes AA/AC of rs10261685 between T1 & T2. B) Moxifloxacin drug PK levels for genotypes CA/CC/TT of rs11045819 between T1 & T2. C) Moxifloxacin drug PK levels for genotypes AG/GG of rs1045642 between T1 & T2. D) Moxifloxacin drug PK levels for genotypes CA/CC/CT of rs11045819 observed at timepoint 2.



D. **Genotypes of rs11045819 at T2**



Chapter 4

Preface:

In Chapter 3 investigated single nucleotide polymorphisms (rs10261685, rs1045642, and rs11045819) which modestly played a role in variable drug PK levels in European and Asian populations which was also not represented very well in commercially available microarrays. As some of the participants were infected with both TB and HIV-1, we then chose to investigate further using an unbiased approach to explore the contribution of genome-wide variants and their association with TB infection, recurrence and HIV-1 predisposition in study participants from three CAPRISA clinical trials. We were uniquely positioned to conduct this research owing to our access to well-characterized and highly valuable samples enrolled in the CAPRISA clinical trials. These samples encompassed the phenotypes of interest, providing an exceptional and timely platform upon which to perform the study. Leveraging this resource enabled us to robustly investigate and highlight the genetic diversity present within the South African population. Furthermore, the use of a genome-wide association approach—an unbiased, hypothesis-free method interrogating all genomic loci—allowed for a comprehensive assessment of genetic contributors underlying the traits of interest. The rationale behind this approach is due to the study in chapter 3 having found no significant associations between the three drug transporter gene SNPs with TB recurrence or HIV acquisition/ predisposition. Certain SNPs however showed modest associations with drug PK levels, metabolizer status, and treatment variability in some patients, indicating potential clinical relevance. However, the data suggests that other genome wide polymorphisms may influential roles in disease onset and progression which can be explored by genome-wide association studies.

Associations of understudied variants in *Mycobacterium tuberculosis* infected African patients with HIV-1 predisposition, TB infection and recurrence by genome-wide association study.

By Sahil Tulsi [1], Prem Prakash Sharma [1], Anmol Gokul [1], Anushka Naidoo [2], Kogieleum Naidoo [2] & Veron Ramsuran [1,2,3]

[1]School of Medicine, College of Health Sciences, University of KwaZulu-Natal, Durban, South Africa,

[2] Centre for the AIDS Programme of Research in South Africa (CAPRISA), University of KwaZulu-Natal, Durban, South Africa,

[3] International Foundation for Integrative Medical Research, Cape Town, South Africa

*Authors to whom correspondence should be addressed:

sahiltulsi@gmail.com; ramsuranv@ukzn.ac.za

Conflict of interest: None declared.

In submission in: *Clinical Genetics*

Manuscript ID: 5becc035-8a98-4134-b08f-e8f176ef89b6

Abstract

Genetic diversity in African patients contributes to a various factors influencing disease predisposition and treatment diversity in TB regimen patients is of huge concern among African populations since outcomes vary based on several genetic factors in drug transporter genes. In this genome-wide associations study we investigated the potential effects of underexplored genetic variants in various genes on HIV-1 acquisition, TB infection and recurrence. Several single nucleotide polymorphisms across the human genome were genotyped by microarray technology with Genome-wide association studies (GWAS) analysis performed post genotyping. Among several understudied mutations identified with strong significance, variant rs121908780, risk classified genotype homozygous reference allele yielded a strong association with HIV-1 infection observed ($p= 8.56\times10^{-35}$, OR >99) which has been previously reported as a pathogenic variant clinically associated with the development of cystic fibrosis. Another SNP with strong association with TB infection identified was rs886040602 (p=

9.33×10^{-87}, OR>99), risk categorized genotype homozygous reference allele previously reported as a pathogenic germline variant linked to the development of inherited breast/ ovarian cancer. The data in this study suggest that potential understudied biomarkers markers linked to disease onset and development in African patients need to be further explored and better understood in larger studies such as additional genome wide association studies that highlight the genetic diversity in African populations.

Introduction

African populations are extremely diverse from a genetic perspective, representing only a minute percentage in global databases to date of less than 3% (24). The addressable lack in representation of genetic information highlights that African genomes are severely understudied (53). Genome-wide association studies provide us with a panoramic view of a patients genetic landscape, providing a deeper understanding of patient biology and disease progression diversity in affected individuals. Genotyping by microarray has proven to be a more cost-effective approach and often preferred alternative to genotyping by sequencing (54) to screen for key known variants and star alleles in millions of markers across the human genome for patients at scale. Commercially available arrays provide lots of value by allowing for the screening of markers that can help to confirm patient ancestry, disease associated clinical variants and pharmacogenomic markers of interest among other clinical aspects to help better understand how and why patients respond to treatment the way they do.

To date several studies have focused on very specific and well documented single nucleotide polymorphisms (SNPs) occurring in host genes such as human leukocyte antigen (HLA), mannose-binding lectin 2 (*MBL2)*, C-lectin (*CD209)*, cytokines as well as chemokine receptor genes that have been observed to be associated with the development of human immunodeficiency virus (HIV) and tuberculosis across global populations (55-58). Additional genome-wide studies can broaden the scope of these host genes associated with the burden of disease and enable researchers to look at several more variants that can possibly influence viral or bacterial acquisition. Legacy technologies such as real-time PCR enabled researchers to investigate SNPs based on low throughput assays, limiting screening volumes and lacking in statistical power. Genotyping by microarray has evolved the method of screening, by investigating millions of genome-wide markers at once in a single assay making it easier to increase statical power in a timely fashion and enable easier imputation to a reference genome

for variant calling. In several large national genomics initiatives (LNGIs) such as the All of US project, UK Biobank, Emirati Genome project and Singapore Precise had undergone population genomic screening in order to compile a reference genome that truly reflects their respective populations genetic make-up. Even though the cost of whole genome sequencing has reduced drastically over the past two decades as next-generation sequencing technologies have advanced (59), it would not prove sustainable to genotype an entire population of millions with sequencing, hence the use of microarray to screen for known population specific variants once secondary and tertiary analysis has been completed population genomic programmes.

Genetic variants identified in human genes play various roles in differential populations, some SNPs often contribute to disease predisposition whilst other variants may contribute to resistance, an example being the variant CCR5-Δ32 in the *CCR5* gene has been significantly associated with HIV-1 resistance and observed mainly in European and Central Asian populations whilst not occurring in African populations (60). Variants linked to HIV-1 sporadic resistance are yet to be identified, along with novel variants associated with acquisition that remain undiscovered. Some studies have identified variants that are unique to African populations (Botswana study population) such as genes *AP3B1, PTPRA, and* NEO1 which are associated with HIV-1 subtype C acquisition (61). Further characterization of African genomes is warranted to unlock new findings that will shed light on disease predisposition. There is a growing need to further explore genetic variants that are associated with HIV and TB co-infection and TB recurrence in African patients. With Africa as a diverse population, there may be known mutations which play a significant role in other populations but are not necessarily impactful in the African ancestry which leave room for investigation of looking at markers across the genome that could potentially play a role disease development and progression.

With the African genome being severely understudied, this study focuses on both known and underexplored markers in the context of African diversity and their association with HIV-1 infection and tuberculosis. Understanding and defining key markers that may or may not be unique to African tuberculosis (TB) patients can help identify biomarkers occurring within specific genes that are often associated with predisposition and infection of HIV and TB. The use of commercial precision medicine focused arrays can provide a significant amount of value by aiding in the identification of population specific variants and their clinical association with the burden of HIV and TB.

Methods

Study Population & sample collection

This research study investigated the genome wide association of various factors contributing to TB infection or recurrence, HIV-1 acquisition and predisposition. A total of 1237 retrospectively stored samples from various South African study cohorts [University of KwaZulu-Natal (UKZN BREC/00005530/2023), Integrase Strand Transfer Inhibitors Insight for the management of HIV associated TB (Insight) and TB Recurrence upon treatment with HAART (TRUTH) were comprised of either peripheral mononuclear blood cells (PBMCs) or buffy coats or peripheral mononuclear cells (PBMCs)] were genotyped by microarray. Additional details on study participants, inclusion/ exclusion criteria and study designs are available at ClinicalTrials.gov [(TRUTH, NCT01539005) & NCT03237182). All individuals were adults 18 years and older, from KwaZulu-Natal, South Africa and of African descent. Table 1 summarizes the total number of samples included in the GWAS analysis and the respective groups for HIV-1 infection and TB infection/ recurrence (TB+HIV+, TB+HIV-, TB+Rec+ & TB+Rec-).

Table 1: Summary of patient samples from the UKZN, TRUTH, and Insight studies included in the genome-wide association (GWAS) analysis.

	No. of patient samples	**TB+HIV+, n=**	**TB+HIV-, n=**
Insight/ Truth Total co-infected Patients	459	417	42
TRUTH Cohort Study Patients	**No. of patient samples**	**TB Rec+, n=**	**TB Rec-, n=**
	427	58	369
UKZN Cohort Study Patients	**No. of patient samples**	**TB+, n=**	**TB-, n=**
	778	558	220

This table provides an overview of the distribution of participants across the different study cohorts and outlines the clinical categories used for comparison, including individuals who

were TB-positive and HIV-positive versus those who were TB-positive and HIV-negative, as well as patients classified as TB-recurrent versus TB non-recurrent. Notably, the cohort of TB/HIV co-infected patients was drawn specifically from both the UKZN and TRUTH studies, reflecting the combined contribution of these two datasets to the co-infection group. The table therefore contextualizes the sample composition and highlights the multi-study integration that underpins the subsequent GWAS findings.

Processing of laboratory samples

Genomic DNA isolation and Genotyping by MicroArray

For the genotyping by microarray approach we selected the Axiom Precision Medicine Diversity Array that investigates >850 000 genome-wide markers that in addition to the SNP backbone consist of approximately 5000 variants in over 1100 genes with known PGx value (ThermoFisher Scientific, USA). Genomic DNA (gDNA) was manually isolated from all sample types that included human specimens consisting of whole blood, sputum & peripheral blood mononuclear cells using standard commercial kits optimized for highest purity, yields and integrity. The quality and quantity of DNA was observed via absorbance ratios (OD_260/OD_280: 1.8–2.0; OD_260/OD_230 > 1.5) and nucleic acid integrity was confirmed by 1% agarose gel electrophoresis. To confirm DNA input requirements of 100ng as per Axiom PMD Array (ThermoFisher Scientific, USA), we used a fluorometric based assay (Qubit, ThermoFisher Scientific).

The template gDNA was amplified using whole-genome amplification (WGA) without being biased to any genes as per manufactures instructions. Post amplification we utilized enzymatic fragmentation to segment the amplicons into smaller fragments. Fragmented DNA was then hybridized to the microarray overnight in a hybridization oven. Hybridized DNA was then exposes to a series of enzymatically staining with fluorescent dyes and washing. To conclude the workflow the microarrays were scanned on the GeneTitan platform (ThermoFisher Scientific, USA) which captures and acquires the fluorescent intensities emitted by the dyes. Genotyping raw data was collected and processed by Axiom Analysis suite (ThermoFisher Scientific).

Data normalization was applied to adjust for variance in probes from batch/ lot effects from manufacturing. Calling of genotypes employed clustering algorithms with the use of curated cluster files. Low quality calls or poor performing SNPs with low call rates or high error rates were omitted as part of quality control metrics based on software defined thresholds (acceptable call rates $\geq 98\%$). Additional quality control checks were performed post-processing to filter for minor allele frequencies (MAF) and confirming Hardy-Weinberg equilibrium. Output

genotyping data was the input for genome-wide association study (GWAS) and imputed into reference panels (1000 Genomes) (24). Data visualization of genome-wide results were observed through Manhattan plots.

Statistical Analysis

Data Quality control (QC) analysis was performed with the use of the Axiom Array Analysis Suite, version 4.0.3.3 (ThermoFisher Scientific, Massachusetts, USA), with best practice workflow. The QC analysis workflow was performed for samples and only auto calls genotypes the samples that have passed the manufacturers defined QC thresholds, followed by classifying the probe sets used to identify the genotypes which are recommended for statistical tests in downstream study.

Genome-wide association analyses was performed to identify specific known genetic variants associated with traits of interest. Quality control (QC) of genotype data was conducted using PLINK v1.9, omitting data from individuals with high missingness (>5%), gender discrepancies, excess in heterozygosity, or cryptic relatedness (identity-by-descent > 0.125). SNP-level filtering excluded variants with a call rate <95%, minor allele frequency (MAF) >0.01 was included in analysis as the defined threshold, and only Variants showing deviation from Hardy–Weinberg equilibrium ($p > 1 \times 10^{-6}$ in controls or in the full cohort, depending on the study design) were included based on the predefined significance threshold.

Association testing was conducted using logistic regression (for binary traits) or linear regression (for continuous traits) under an additive genetic model, adjusting for relevant covariates including age, sex, and the top principal components (PCs) to control for population stratification. Principal component analysis (PCA) was performed using EIGENSOFT v7.2.1 to identify and adjust for ancestral structure with logistic regression analysis. Quantile-Quantile plots based on genome-wide association analysis comparing the observed distribution of $-\log_{10}(P)$ values to the expected distribution under the null hypothesis of no association were generated using R software and included in the Supplementary data, figures 1-6. Genomic inflation factor (λ), which quantifies the extent of systematic inflation in genome-wide association test statistics has also been reported alon with the logistic regression QQ plots in the Supplementary Figures: 2, 4 & 6.

To account and adjust for multiple testing comparisons we performed Bonferroni correction to determine genome-wide significance was set at $p \leq 5.89 \times 10^{-8}$ for the HIV-1 GWAS analysis

and $p \leq 5.79\times 10^{-8}$ for the TB recurrence GWAS analysis. Suggestive associations were considered at $p < 1 \times 10^{-5}$. Manhattan plots were generated to visualize the distribution of association statistics and identify genomic loci of interest highlighting SNPs with highest level of impact and association on the respective phenotypes.

Ethical considerations

This study was granted expedited ethical approval by the Biomedical Research Ethics Committee of the University of KwaZulu-Natal (BREC/00004110/2022).

Results

To investigate the association on genome-wide variants associated with HIV-1 infection among TB positive individuals we genotyped a total of 459 patients, 417 (91%) individuals were both TB and HIV positive and 42 (9%) individuals were TB positive but HIV negative as referenced in Table 1. Single nucleotide polymorphisms where either all HIV negative frequencies are missing or all HIV positive frequencies are missing were filtered out for analysis. From 864, 313 SNPs, a total of 16, 343 were omitted allowing 847, 970 SNPs to be included for Fisher statistical tests.

Table 2: Summarized tabular representation of the top 10 single nucleotide polymorphisms (SNPs) most significantly associated with HIV-1 infection susceptibility by Fisher's exact test, as identified through genome-wide association study (GWAS) analysis.

GWAS Analysis and association with HIV-1 Infection													
Chr	Gene	Position	SNP	Ref/Alt	p-Value	OR Ref/Ref	OR Ref/Alt	OR Alt/Alt	Class Ref/Ref	Class Ref/Alt	Class Alt/Alt	Clinical Association	Reference/ ClinVar Accession
7	CTFR	117592140	rs121908780	GAAATTCAATCCT/AGAAA	8.56E-35*	>99*	<0.001	NA	Risk	Protective	NA	Pathogenic Variant for Cystic fibrosis	RCV000046511.25
17	LASP1	38888179	rs145691953	G/A	1.45E-34*	>99*	<0.001	NA	Risk	Protective	NA	-	-
14	Undefined	48040161	rs76466646	C/T	2.06E-34*	>99*	0.003	0	Risk	Protective	Protective	-	-
13	BRCA2	32337189	rs397509342	AA/A	2.46E-34*	>99*	<0.001	NA	Risk	Protective	NA	-	-
2	MSH2	47478300	rs1553369652	ATA/A	2.58E-34*	>99*	<0.001	NA	Risk	Protective	NA	-	-
17	BRCA1	43091753	rs80357798	AA	2.70E-34*	>99*	<0.001	NA	Risk	Protective	NA	Pathogenic Germiline Inherited variant Breast Ovarian	RCV000112182.13, RCV000574583.11
13	RB1	48367575	rs587778855	A/T	2.98E-34*	>99*	<0.001	NA	Risk	Protective	NA	Germline pathogenic variant linked to retinoblastoma	RCV000114717.2
13	BRCA2	32325128	rs886040490	AAT/TA	4.39E-34*	>99*	<0.001	NA	Risk	Protective	NA	Pathogenic variant for Breast Ovarian, familial cancer	RCV000257355.5
12	TSPAN8	71181113	rs118011769	C/T	5.48E-34*	>99*	<0.001	NA	Risk	Protective	NA	-	-
9	ASTN2	116761901	rs75670657	G/A	6.03E-34*	>99*	0.003	0	Risk	Protective	Protective	-	-

For each SNP, the corresponding gene name, chromosomal location, rsID, observed genotype, and GWAS p-value derived from comparisons between TB+HIV+ and TB+HIV– patient groups are reported. This table summarizes the variants that exceeded genome-wide significance thresholds and highlights loci potentially involved in biological pathways influencing HIV-1 susceptibility. Genotypes with associated with HIV-1 risk are indicated by an asterisk (*). Notably, the SNP displaying the strongest association with HIV status—characterized by a p-value of $8.56 \times 10\ E^{-35}$ and prominently depicted in Figure 1 is included among the listed variants.

A total of 27, 967 SNP assays had P values of ≤ 0.05. Bonferroni correction was determined at P values ≤ 5.89×10^{-08}. From the 27, 967 SNPs we filtered out the top 10 most significant SNPs that were associated with HIV infection and referenced in the Manhattan plot shown in Figure 1. Variant rs121908780 in the CTFR gene on chromosome 7 was the SNP with the highest association with HIV status as represented in Figure 1, p = 8.56×10^{-35} with greatest significance over Bonferroni corrected P value of 5.89 E-08. Three additional SNPs [rs80357798 (p= 2.71E-34), rs587778855 (p= 2.98E-34), rs886040490, (p= 4.40E-34)] were highly significant and reported in ClinVar for various clinical associations referenced with accession numbers in Table 2. Six out of the 10 most significant SNPs reported in Table 2 are currently not published in ClinVar but still highly significant after comparison with Bonferroni corrected P value of 5.89 E-08 [rs145691953 (p= 1.46E-34), rs76466646 (p= 2.07E-34), rs397509342 (p= 2.46E-34), CAT>C 47478300:47478301 (p= 2.58E-34), rs118011769 (p= 5.49E-34) & rs75670657, (p= 6.04E-34)]. Data on genotypic frequencies for the top 10 associated SNPs with HIV-1 infection are represented in Supplementary Table 1.

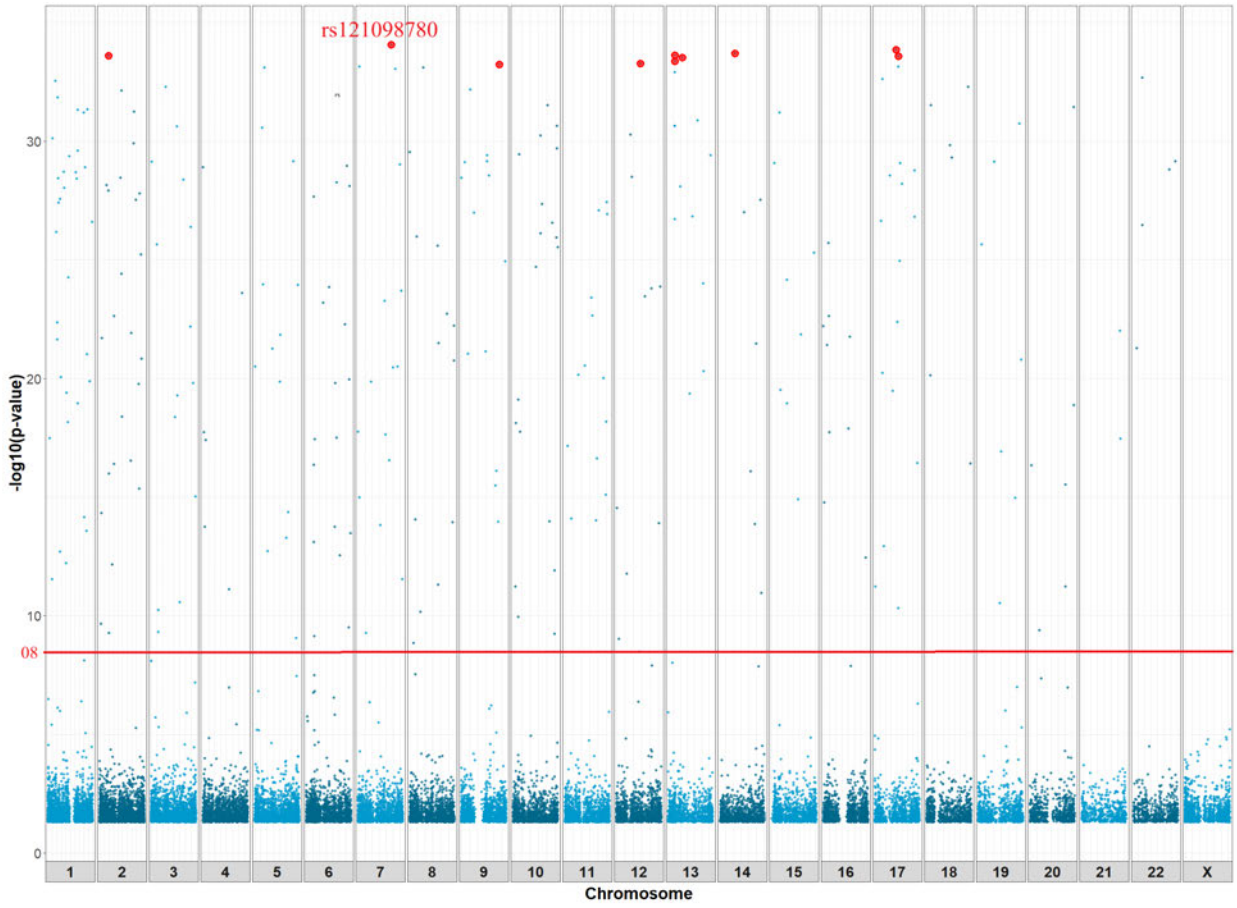


Figure 1: Manhattan plot generated from the genome-wide association study (GWAS) of TB-positive patient samples (n = 459) evaluated for genetic variants associated with HIV-1 infection status by Fisher's exact test for genome-wide significance. Each point on the plot represents an individual SNP mapped to its chromosomal position, with the $-\log_{10}$(p-value) displayed on the y-axis to illustrate the strength of association. The red horizontal line across the Manhattan plot indicates the Bonferroni-corrected genome-wide significance threshold ($p \leq 5.89 \times 10^{-8}$), providing a stringent cutoff for identifying variants with statistically meaningful associations at a genome level. Variant rs121098780, which exhibits the strongest association with HIV-1 infection ($p = 8.6 \times 10^{-35}$), surpasses this threshold by several orders of magnitude and is highlighted in red text to emphasize its exceptional significance within the dataset.

To advance our understanding of significant SNPs that may potentially influence the risk of TB recurrence we genotyped a total of 427 patient samples, with 58 (14%) who had recurrent TB infection and 369 (86%) individuals who showed no symptoms or presented themselves with clinical evidence of recurrent TB as summarized in Table 1. Of the 864, 313 markers, 1122 SNPs were omitted from data analysis due to missing frequencies for TB Rec+ and TB Rec- study groups. After omitting the 1122 SNPs, the remaining 863, 191 SNPs were analyzed with Fishers statistical test. A total of 29, 686 SNPs showed P values that were ≤ 0.05. Following Bonferroni correction to adjust for multiple comparisons P = 5.79 $\times 10^{-08}$ was determined. Figure 2 represents the 10 most significant SNPs [rs17075658 (p= 2.69E^{-06}), rs1987349241

(p= 7.05E^{-06}), rs4299914 (p= 7.98E^{-06}), rs7913065 (p= 9.38E^{-06}), rs4950973 (p= 9.89E^{-06}), rs3786544 (p= 1.30E^{-05}), rs17272023 (p= 1.52E^{-05}), rs116389662 (p= 1.64E^{-05}), rs79644900 (p= 1.77E^{-05}), rs77246883 (p= 1.79E^{-05})] genotyped and summarized in Table 3 with all presented P values greater than Bonferroni corrected P value ≤ 5.79 E^{-08} and were not considered significant, neither were the variants published in ClinVar for clinical associations. Genotype frequencies of the top 10 associated SNPs with TB recurrence infection are represented in Supplementary Table 2.

Table 3: Expanded tabular representation of the top 10 single nucleotide polymorphisms (SNPs) most significantly associated with TB recurrence by Fisher's exact test as identified through genome-wide association study (GWAS) analysis.

GWAS Analysis and association with TB recurrence											
Chr	Gene	Position	SNP	Ref/Alt	p-Value	OR Ref/Ref	OR Ref/Alt	OR Alt/Alt	Class Ref/Ref	Class Ref/Alt	Class Alt/Alt
13	NFASC	20054788	rs17075658	C/T	2.69E-06	0.03405995	29.36*	NA	Protective	Risk	NA
20	SLCO4A1	62672463	rs1987349241	T/C	7.05E-06	0.28638132	2.5142857*	5.96703297	Protective	Risk	Risk
7	SLC29A4	5301939	rs4299914	A/G	7.98E-06	4.1068580*	0.32926829	0.94693878	Risk	Protective	Neutral
10	GRM3-AS1	1960295	rs7913065	G/A	9.38E-06	0.26266094	3.7548701*	1.06481481	Protective	Risk	Neutral
1	NFASC	204919626	rs4950973	G/T	9.89E-06	0.23374676	4.173758*	3.2105263*	Protective	Risk	Risk
19	ZMYM2	55276501	rs3786544	C/T	1.30E-05	0.25735294	2.4712430*	3.12*	Protective	Risk	Risk
19	IRAIN	17356405	rs17272023	G/A	1.52E-05	0.15117371	5.9346504*	>99*	Protective	Risk	Risk
15	HSPBP1	98641883	rs116389662	G/A	1.64E-05	0.22553897	4.7390625*	<0.001	Protective	Risk	Protective
7	PLVAP	86796009	rs79644900	G/A	1.77E-05	0.03970417	21.173076*	>99*	Protective	Risk	Risk
1	SLCO4A1	60637838	rs77246883	C/T	1.79E-05	0.20723684	5.2419047*	0	Protective	Risk	Protective

For each SNP, the corresponding gene name, chromosomal location, rsID, observed genotype, and GWAS-derived p-value from comparisons between TB-recurrent and TB non-recurrent patient groups are provided. Risk genotypes are indicated with an asterisk (*), denoting allelic configurations that appear to confer an increased risk and likelihood of recurrence based on the observed associations. Although the SNPs listed do not reach genome-wide significance after applying Bonferroni correction, they are nonetheless considered suggestive or nominally significant signals, as all exhibit p-values < 5.0×10^{-6}. These variants may therefore represent biologically relevant loci warranting further investigation in larger or independent cohorts.

Between TB positive and TB negative individuals we investigated significant SNPs that may play an influential role in TB acquisition by genotyping a total of 778 individuals of which 558 (72%) were positive for TB and 220 (28%) were TB negative as summarized in Table 1. A total of 839, 167 SNPs were investigated using Fisher statistical tests after omitting 1075 markers from the initial 840, 242 markers due to missing frequencies for the TB positive and TB negative study groups. There were 31, 762 SNPs that had P values ≤ 0.05. After Bonferroni correction to adjust for multiple test comparisons, p=5.95×10^{-08} was determined. Figure 3 represents the Manhattan plot with reference to the top 10 SNPs with the highest significance [rs886040602 (p= 9.33E$^{-87)}$, rs12199036002 (p= 9.25E^{-86}), rs112194146 (p= 1.03E^{-85}), rs1057517572 (2.40E^{-84}), rs1064792936 (5.17E^{-84}), rs11587334 (p= 5.74E^{-84}), rs78396100 (p=

2.33E^{-83}), rs1555578543 (p= 5.13E^{-83}), rs1555594837 (p= 5.13E^{-83}), rs886055595 (p= 5.13E^{-83})] genotyped and also presented in Table 4 together with P values significantly lower when compared to Bonferonni corrected value of P≤ 5.95E^{-08}. We have summarized the genotype frequencies of the top 10 variants with TB infection and are represented in Supplementary Table 3.

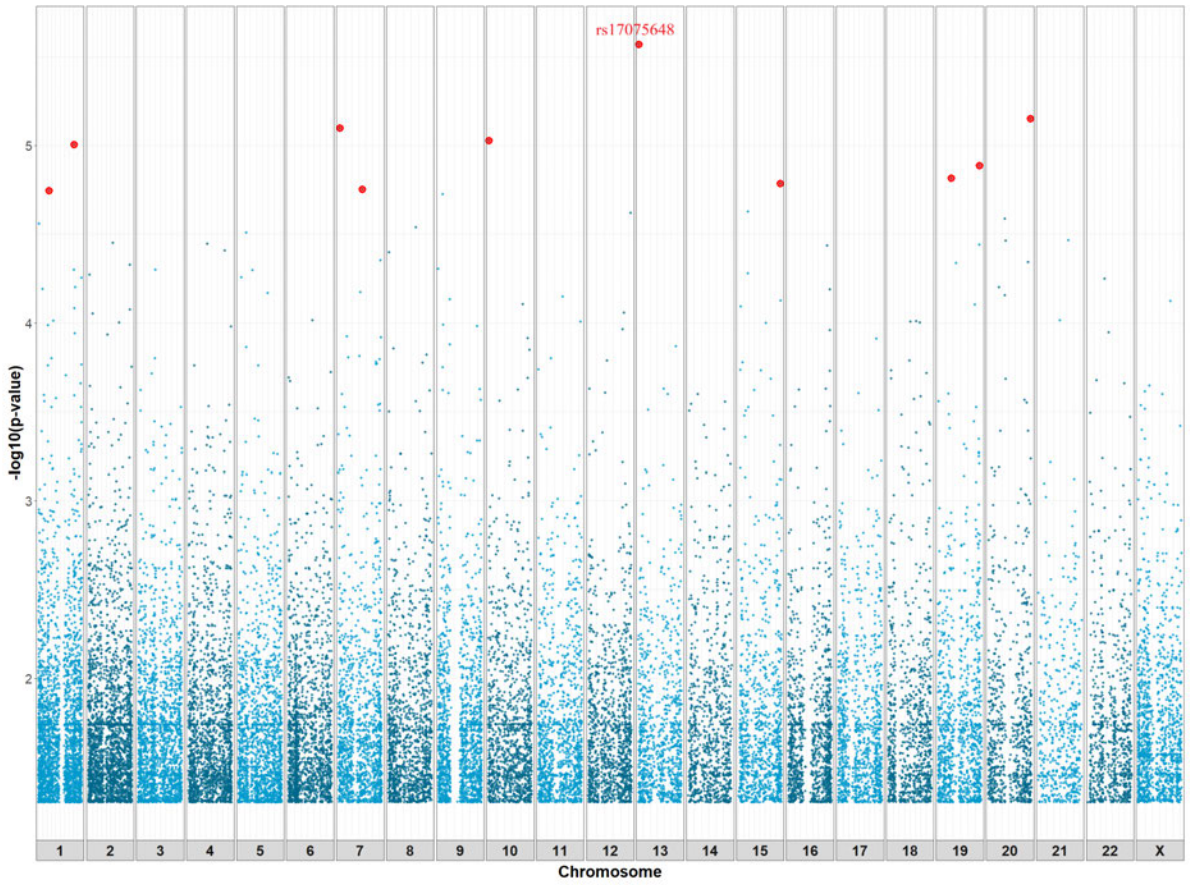


Figure 2: Manhattan plot generated from the genome-wide association study (GWAS) of TB-positive patient samples (n= 427) assessing genetic variants associated with susceptibility to TB recurrence by Fisher's exact test. Each plotted point corresponds to an individual SNP positioned according to its chromosomal location along the x-axis, with the –log$_{10}$ (p-value) displayed on the y-axis to indicate the strength of association with recurrent TB. The top 10 SNPs featured in this analysis are of nominal significance, reflecting p-values below 5.0 × 10E^{-05} but not surpassing the stringent threshold required for genome-wide significance. As these variants did not meet Bonferroni-corrected significance criteria, no dotted genome-wide threshold line is shown on the plot. The most notable association is observed for variant rs17075648, located on chromosome 13, which demonstrates the strongest signal within the dataset ($p = 2.7 \times 10^{-6}$) and is highlighted as the leading candidate SNP related to recurrent TB infection.

Table 4. Comprehensive tabular summary of the top 10 single nucleotide polymorphisms (SNPs) demonstrating the strongest statistical association with tuberculosis (TB) infection as identified through genome-wide association study (GWAS) analysis by Fisher's exact test.

GWAS Analysis and association with TB Infection													
Chr	Gene	Position	SNP	Ref/Alt	P Value	OR Ref/Ref	OR Ref/Alt	OR Alt/Alt	Class Ref/Ref	Class Ref/Alt	Class Alt/Alt	Clinical Associations	Reference/ Accession
13	BRCA2	32340008	rs886040602	T/TT	9.33E-87*	>99*	<0.001	NA	Risk	Protective	NA	Pathogenic Germiline Inherited variant Breast Ovarian	RCV000256755.13
2	MSH2	47410262	rs12199036002	CC/C	9.25E-86*	>99*	<0.001	NA	Risk	Protective	NA	-	-
5	Undefined	31180270	rs112194146	T/C	1.02E-85*	>99*	<0.001	NA	Risk	Protective	NA	-	-
13	BRCA2	32319317	rs1057517572	T/TT	2.40E-84*	>99*	<0.001	NA	Risk	Protective	NA	Pathogenic variant for hereditary breast ovarian cancer	VCV000371869.11
11	MYBPC3	47351272	rs1064792936	TGG/CCTCC	5.16E-84*	>99*	<0.001	NA	Risk	Protective	NA	-	-
1	FLG_AS1	1.52E+08	rs11587334	G/A	5.73E-84*	>99*	<0.001	NA	Risk	Protective	NA	-	-
13	ATP8A2	25788311	rs78396100	C/T	2.32E-83*	>99*	<0.001	NA	Risk	Protective	NA	-	-
17	BRCA1	43063878	rs1555578543	A/AG	5.12E-83*	>99*	<0.001	NA	Risk	Protective	NA	-	-
17	BRCA1	43099797	rs1555594837	CTTTTG/C	5.12E-83*	>99*	<0.001	NA	Risk	Protective	NA	-	-
2	APOB	21038000	rs886055595	C	5.12E-83*	>99*	<0.001	NA	Risk	Protective	NA	Likely Pathogenic variant for Hypercholesterolemia/hypobetalipoproteinemia	VCV000334183.5

Each SNP is presented together with its corresponding gene annotation, chromosomal position, reference SNP identifier (rsID), observed genotype, and the calculated p-value reflecting the strength of association between TB-infected and uninfected study participants. Notably, the SNP rs886040602, which exhibited the highest level of statistical significance ($p = 9.33 \times 10^{-87}$), is included among these top 10 variants. Genotypes associated with increased risk of TB infection are marked with an asterisk (*) to clearly distinguish risk-enhancing allelic combinations from non-risk genotypes. This table serves to highlight the most influential candidate loci that may contribute to genetic susceptibility to TB infection.

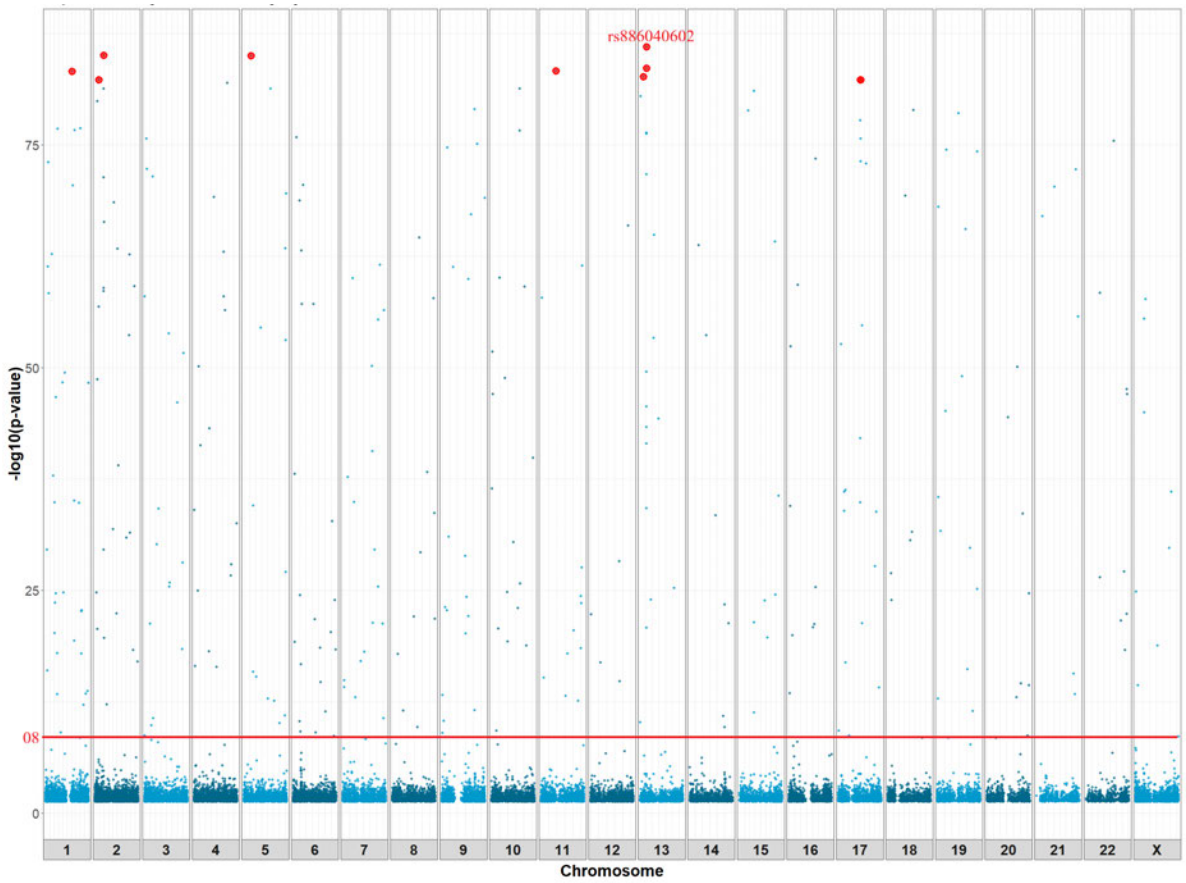


Figure 3: **Manhattan plot illustrating the genome-wide distribution of association signals derived from the GWAS comparing TB-positive patients (n= 985) with TB-negative individuals by Fisher's exact test.** Each point on the plot represents a single nucleotide polymorphism (SNP), plotted according to its chromosomal position on the x-axis and the $\log_{10}$ (p-value) of its association with TB infection on the y-axis. Statistical significance was assessed using Fisher's exact test**,** enabling detection of genotypic differences between infected and uninfected groups. The horizontal line spanning the plot denotes the Bonferroni-corrected significance threshold**,** used to account for multiple testing across the genome and to identify truly robust association signals. The SNP exhibiting the strongest association with TB infection, rs886040602 ($p= 9.33 \times 10E^{-87}$)**,** is highlighted as the most significant variant identified in this analysis, marking a prominent peak that surpasses the corrected threshold.

Discussion

Observed associations with HIV-1 infection

Various studies have shown that several genetic mutations are strongly associated with an increased risk of HIV-1 acquisition and infection (60-62). A South African study concluded that variants that confer resistance to HIV-1 infection are very rare in black individuals of South African descent (63), whilst HIV-1 resistance to infection due to mutations such as CCR5-Δ32 has been found in European populations with an allele frequency of at least 10% (64), It is clearly evident that African genomes are among the most diverse globally (65) , to an extent that patient outcomes in the context of HIV infection, TB recurrence or drug metabolism is subjective compared to populations that are far more well characterized from a genetic perspective.

In our study we looked at the diversity of genome-wide variants with an unbiased approach in association with HIV-1 infection, TB recurrence & infection as well as additional focus on variance in Drug PK levels for patients on anti-TB therapy. From our genome-wide analysis of 459 patients, 417 (91%) on individuals were co-infected with TB and HIV-1with 42 (9%) were TB positive but HIV-1 negative. Our findings reported that more than 27, 000 variants has p-values ≤ 0.05, however for the purposes of this study we chose to focus on the 10 most significant SNPs in association with HIV-1 infection between the two study groups. All 10 of the most SNPs were significantly associated with HIV-1 infection even after comparison to Bonferroni corrected P values of $5.89 \times 10E^{-08}$.

The variant with the strongest association seen in our study population was rs121908780, which has been characterized as a pathogenic frameshift mutation reported to have shown strong associations with cystic fibrosis (CF) in American and European populations (66, 67) and has been published in ClinVar (accession references : RCV000046511.22, RCV001004280.9). Tuberculosis is rarely seen to co-exist in patients with CF however could be regarded as a potential pathogen (68). Our clinical metadata had no record of patients with CF. The Fisher test analysis for rs121908780 yielded a P value = $8.56E^{-35}$ with an odds ratio (OR) >99 for homozygous reference allele (GAAATTCAATCCT) which we have classed as risk associated. These findings suggest that it is possible the variant rs121908780 is playing a differential role in South African individuals and would need to be explored in more detail with

additional research studies. Data as such again highlights the diversity observed in African populations which need to be further understood at a genome level. Genetic mutations rs397509342 and rs80357798 are variants found in *BRCA2* and *BRCA1* genes respectively which are known to be oncogenes investigated for both germline and somatic breast cancer development (69). Variant rs80357798 in *BRCA1* with a P value $2.70E^{-34}$ [OR >99 (Homozygous for Ref AA)] was classified as high odds ratio for HIV-1 risk in this study has also been published in ClinVar as a pathogenic variant and driver mutation for familial breast ovarian cancer (accession references: RCV000112182.13, RCV000574583.11). Another *BRCA2* mutation identified was rs886040490 with a P value = $4.39E^{-34}$ and OR > 99 for homozygous allele AAT considered as a high-odds risk SNP for HIV infection which has been published in ClinVar (accession reference : RCV000257355.5) as a pathogenic variant for breast and ovarian cancer. In contrast rs397509342 with a P value of $2.46E^{-34}$ and OR >99 for homozygous allele AA found in *BRCA2* has not yet been published in ClinVar as a pathogenic variant but was classified as a high-odds risk SNP in association with HIV-1 infection in this study. The remainder of the 6 SNPs rs145691953, rs76466646, rs397509342, rs1553369652, rs118011769 and rs75670657 with p values and odds rations that are strongly significant when compared to Bonferroni corrected P values of 5.89 E^{-08} as represented in Table 2 are not yet published in ClinVar. These understudied variants need to be further validated with additional studies focusing on their potential role in HIV acquisition specifically in African cohorts.

Observed associations with TB recurrence

Recurrence of tuberculosis infection is observed at various rates in several populations across the world (70). A Vietnamese study by Vree *et al*. showed that 8.6% of TB patients presented themselves with recurrent TB between 1-2 years post successful treatment (71). In contrast an Ethiopian study showed approximately half the recurrence rate compared to the Vree *et al* with 4.1% of patients that experiences recurrent TB infection (72). It is possible that genetic diversity of the Ethiopian individuals could have potentially contributed to the disparity in recurrence rate between the two populations. In our TB recurrence aspect of this study we observed a 13.58% rate of recurrence with 58 of 427 patients who presented themselves with recurrent clinical TB infections after having completed their TB regimen successfully as part of the TRUTH study (*clinical trials.gov ID NCT01539005*). Our GWAS analysis revealed that 29, 686 SNP assays from the entire 863, 191 had P values that were ≤ 0.05. All 10 variants reported in Table 3 were not significant when compared to Bonferroni corrected P value of 5.79

E^{-08}. From the 10 most significant SNPs that are reported in Table 4, the SNP assay with greatest significance was *NFASC* mutation rs17075658 ($p= 2.68E^{-06}$, OR = 29.36 for allele C/T) which was classified as a risk associated genotype with TB recurrence. Despite the P value not being considered as significant when compared with Bonferroni corrected P value of $P = 5.79 \times 10^{-08}$, the odds ratio of 29.36 is well above the threshold of 1.0 to be associated as a risk heterozygous genotype. Although the variant rs17075658 has not been published to date in ClinVar, we confirmed the MAF of the heterozygous genotype [C/T] reported as 0.000053 in European populations, 0.0 in Asian populations whilst occurring at 0.050171 as per ALFA database. With a higher MAF of rs17075658 the heterozygous genotype [C/T] can be considered as a TB recurrence risk-associated genotype in African populations, this data suggests once again that diversity in African populations is observed at far higher levels than compared to any other population. The third most significant SNP with a p value = 7.98×10^{-06} and OR = 4.10 for homozygous genotype [A/A] classified as a risk associated genotype was rs4299914 found in solute carrier 29 gene *SLC29A4.* Minor allele frequency of the homozygous allele [A/A] has been reported in Asian (0.121) whilst both African and European populations are reported to have higher frequencies recorded (0.19 & 0.21) respectively .Whilst no clinical significance of rs4299914 has been reported to date in ClinVar and not considered significant in this study after Bonferroni adjustment it is suggested that the homozygous genotype [A/A] may potentially pose as a risk factor (OR = 4.10) for TB recurrence in African populations and would need to be looked at in greater detail for future studies. The remainder of the SNPs from the top 10 significant variants are also not published in ClinVar to date but the study specific risk genotypes calculated suggest that the variants indeed need additional investigation for further understanding their modest role in TB recurrence.

Associations observed between TB positive and TB negative individuals

Some studies have found that there are genetic markers that can be unique to TB positive patients, a study in Taiwan concluded that mutations found in the *SP110* gene are markers associated with predisposition to both latent and active infection of TB (73). Another Southern African study in Mozambique showed that rs1800629 (TNF) -308G>A, a cytokine specific variant that was uniquely reported in Sub-Saharan patients with pulmonary TB (74). Although rs1800629 was not a marker present on our microarray bead chip used for this study therefore not investigated in this study it is evident that different variants, in various chromosomes are linked to TB susceptibility that varies between different populations. We are yet to further

characterize more variants linked to TB susceptibility that may occur in various regions of the human genome. The Fisher statistical analysis from this GWAS reported more than 31, 000 SNPs with P values ≤ 0.05, with a Bonferroni corrected P value ≤ $P=5.95E^{-08}$ determined. Table 4 summarizes the 10 most highly significant SNPs even after Bonferroni comparison. The highest significant marker observed was *BRCA2* variant rs886040602, $p= 9.33E^{-87}$ with OR >99 for homozygous genotype [T/T] has been identified as a high-risk marker for TB infection in this study. Clinical significance of the pathogenic variant rs886040602 associated with germline breast/ ovarian cancer has been reported in ClinVar (accession reference: RCV000256755.13). Similarly, the same trend has been observed with TB associated high-risk genotype [T/T] of rs1057517572 (*BRCA2*), $p= 2.40E^{-84}$ and OR >99 in this GWAS, and reported as a pathogenic variant for germline risk of inherited breast/ ovarian malignancies in ClinVar (accession reference: VCV000371869.11). Another interesting finding was the strong significance of rs886055595, $p= 5.13\times10^{-83}$, OR > 99 classified as high risk for TB infection with homozygous [C/C] genotype found in *APOB* gene. Other studies have reported rs886055595 has been classified as a likely pathogenic variant associated with Familial Hypercholesterolemia (FH) and reported in ClinVar (accession reference: VCV000334183.5). The remaining 7 of the top 10 variants presented in Table 4, (rs12199036002, rs112194146, rs1064792936, rs11587334, rs78396100, rs1555578543, rs1555594837) however were not reported yet in ClinVar for any confirmed clinical associations. All 7 of the above-mentioned variants yielded strong significant P values from the Fishers test compared to Bonferroni corrected P value ($P=5.95\times10^{-08}$), with homozygous reference genotypes identified as risk markers as shown in Table 4. These variants need to be investigated in more detail to uncover their contribution to TB infection or predisposition in African individuals.

Conclusion

This study underscores the critical importance of accounting for the extensive genetic diversity within African populations in the context of disease predisposition and development. Such diversity is a key factor in better understanding genetic aspects of HIV-1 infection, tuberculosis infection and recurrence as well as anti-TB drug metabolism that contributes to the variable treatment outcomes observed among patients undergoing therapy. Our findings highlight the presence of several genome-wide mutations in significant genes that may influence acquisition and predisposition to TB or HIV-1 infection, many of which remain underexplored. It is clear that many SNPs which play various roles in other global populations such as driver mutations

for breast, prostate or ovarian cancer however are strongly associated with the increased risk and burden of disease in African populations. To fully understand the underexplored genetic drivers of differential predisposition and disease development, large-scale investigations are needed to characterize these variants more comprehensively. Additionally, expanding whole genome sequencing (WGS) efforts across diverse African cohorts will be essential to uncover novel variants that could inform personalized TB and HIV-1 protective, prevention and treatment strategies. Advancing on suggestions from the data in this study will be pivotal in defining key biomarkers of disease prevention and progression for individuals of African descent, ultimately improving patient outcomes and equity in healthcare.

Funding statement

Research performed in this article was generously supported by Grants, Innovation and Product Development unit of the South African Medical Research Council (SAMRC) with funds received from both Novartis and GSK Global Health for Project Africa GRADIENT (Grant No. GSKNVS2/202101/005).

References

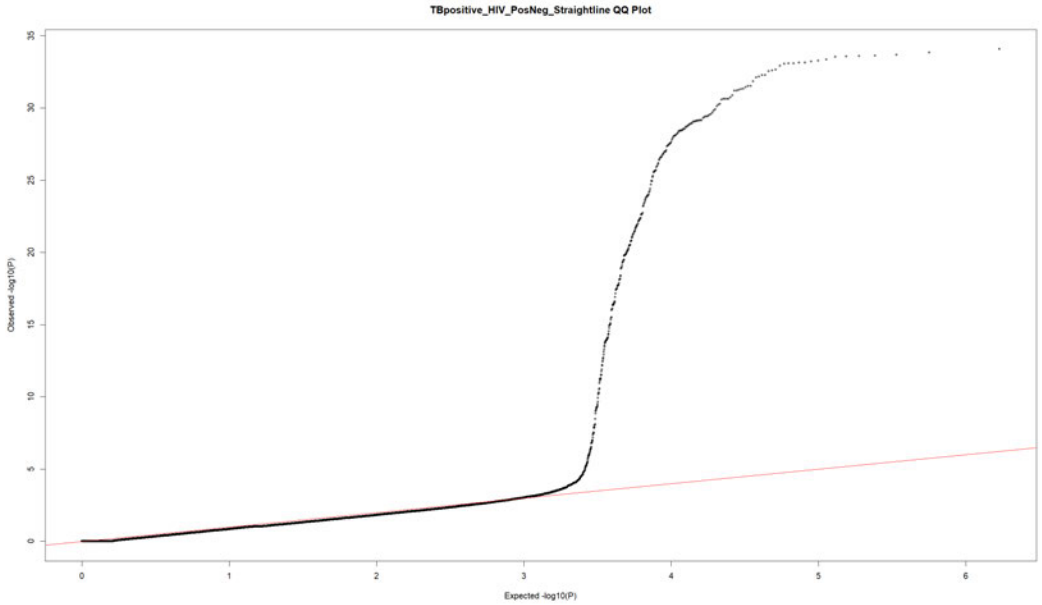
1. Gomez F, Hirbo J, Tishkoff SA. Genetic variation and adaptation in Africa: implications for human evolution and disease. Cold Spring Harb Perspect Biol. 2014;6(7):a008524.
2. Wonkam A, Adeyemo A. Leveraging our common African origins to understand human evolution and health. Cell Genom. 2023;3(3):100278.
3. Zhang S, Xiao K, He X, Hao M, Ma Y, Tian Y, et al. A novel efficient algorithm for common variants genotyping from low-coverage sequencing data. Computational and Structural Biotechnology Reports. 2025;2:100046.
4. Alagarasu K, Selvaraj P, Swaminathan S, Raghavan S, Narendran G, Narayanan PR. Mannose binding lectin gene variants and susceptibility to tuberculosis in HIV-1 infected patients of South India. Tuberculosis (Edinb). 2007;87(6):535-43.
5. Selvaraj P, Alagarasu K, Swaminathan S, Harishankar M, Narendran G. CD209 gene polymorphisms in South Indian HIV and HIV-TB patients. Infect Genet Evol. 2009;9(2):256-62.
6. Indrati AR, Sumarpo A, Haryanto J, Rosmiati NMD, Munaya S, Turbawaty DK, et al. Identification of cytokine signatures in HIV-infected individuals with and without Mycobacterium tuberculosis co-infection. Biomed Rep. 2024;21(3):131.
7. Tamene W, Marconi VC, Abebe M, Wassie L, Belay Y, Kebede A, et al. Differential expression of chemokine receptors on monocytes in TB and HIV S. Heliyon. 2023;9(6):e17202.
8. McCombie WR, McPherson JD. Future Promises and Concerns of Ubiquitous Next-Generation Sequencing. Cold Spring Harb Perspect Med. 2019;9(9).
9. Petrovski S, Fellay J, Shianna KV, Carpenetti N, Kumwenda J, Kamanga G, et al. Common human genetic variants and HIV-1 susceptibility: a genome-wide survey in a homogeneous African population. Aids. 2011;25(4):513-8.
10. Shevchenko AK, Zhernakova DV, Malov SV, Komissarov A, Kolchanova SM, Tamazian G, et al. Genome-wide association study reveals genetic variants associated with HIV-1C infection in a Botswana study population. Proceedings of the National Academy of Sciences. 2021;118(47):e2107830118.
11. An P, Winkler CA. Host genes associated with HIV/AIDS: advances in gene discovery. Trends in Genetics. 2010;26(3):119-31.
12. Williamson C, Loubser SA, Brice B, Joubert G, Smit T, Thomas R, et al. Allelic frequencies of host genetic variants influencing susceptibility to HIV-1 infection and disease in South African populations. Aids. 2000;14(4):449-51.
13. Galvani AP, Novembre J. The evolutionary history of the CCR5-Δ32 HIV-resistance mutation. Microbes and Infection. 2005;7(2):302-9.
14. Sibomana O. Genetic Diversity Landscape in African Population: A Review of Implications for Personalized and Precision Medicine. Pharmgenomics Pers Med. 2024;17:487-96.
15. Sosnay PR, Siklosi KR, Van Goor F, Kaniecki K, Yu H, Sharma N, et al. Defining the disease liability of variants in the cystic fibrosis transmembrane conductance regulator gene. Nat Genet. 2013;45(10):1160-7.

16. Castellani C, Cuppens H, Macek M, Jr., Cassiman JJ, Kerem E, Durie P, et al. Consensus on the use and interpretation of cystic fibrosis mutation analysis in clinical practice. J Cyst Fibros. 2008;7(3):179-96.
17. Marco A, Montales MT, Mittadodla P, Mukasa L, Bhaskar N, Bates J, et al. Tuberculosis reinfection in a pregnant cystic fibrosis patient. Respiratory Medicine Case Reports. 2015;16:57-9.
18. Mehrgou A, Akouchekian M. The importance of BRCA1 and BRCA2 genes mutations in breast cancer development. Med J Islam Repub Iran. 2016;30:369.
19. Moosazadeh M, Bahrampour A, Nasehi M, Khanjani N. The incidence of recurrence of tuberculosis and its related factors in smear-positive pulmonary tuberculosis patients in Iran: A retrospective cohort study. Lung India. 2015;32(6):557-60.
20. Vree M, Huong NT, Duong BD, Sy DN, Van LN, Hung NV, et al. Survival and relapse rate of tuberculosis patients who successfully completed treatment in Vietnam. Int J Tuberc Lung Dis. 2007;11(4):392-7.
21. Datiko DG, Lindtjørn B. Tuberculosis recurrence in smear-positive patients cured under DOTS in southern Ethiopia: retrospective cohort study. BMC Public Health. 2009;9:348.
22. Chang SY, Chen ML, Lee MR, Liang YC, Lu TP, Wang JY, et al. SP110 Polymorphisms Are Genetic Markers for Vulnerability to Latent and Active Tuberculosis Infection in Taiwan. Dis Markers. 2018;2018:4687380.
23. Mabunda N, Alvarado-Arnez LE, Vubil A, Mariamo A, Pacheco AG, Jani IV, et al. Gene polymorphisms in patients with pulmonary tuberculosis from Mozambique. Mol Biol Rep. 2015;42(1):71-6.
24. The 1000 Genomes Project Consortium. A global reference for human genetic variation.*Nature* **526**, 68–74 (2015). https://doi.org/10.1038/nature15393

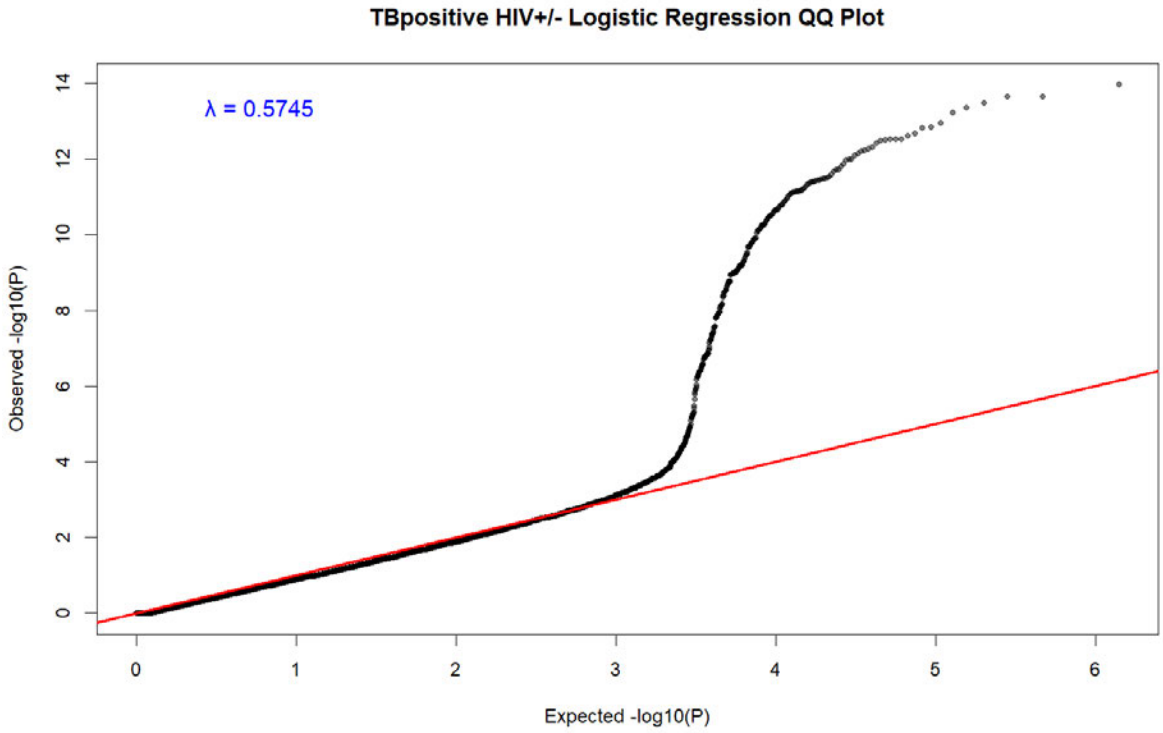
Supplementary Table 1: Top 10 Significant SNPs with GT Frequencies of TB+HIV- and TB+HIV+ Individuals										
SNP	**TB+HIV-**	**Omitted**	**Ref/Ref, n (%)**	**Ref/Alt, n (%)**	**Alt/Alt, n (%)**	**TB+HIV+**	**Omitted**	**Ref/Ref, n (%)**	**Ref/Alt, n (%)**	**Alt/Alt, n (%)**
rs121908780	42	-	0	42(100)	0 (0)	417	2	368 (88.7)	47 (11.3)	0 (0)
rs145691953	42	-	0	42(100)	0 (0)	417	1	368 (88.5)	48 (11.5)	0 (0)
rs76466646	42	-	0	41 (97.6)	1 (2.4)	417	0	367 (88)	50 (12)	0 (0)
rs397509342	42	-	0	42(100)	0	417	0	368 (88.3)	49 (11.8)	0 (0)
rs1553369652	42	1	0	41 (100)	0	417	3	368 (88.9)	46 (11.1)	0 (0)
rs80357798	42	-	0	42(100)	0	417	1	367 (88.2)	49 (11.8)	0 (0)
rs587778855	42	-	0	42(100)	0	417	2	366 (88.2)	49 (11.8)	0 (0)
rs886040490	42	1	0	41 (100)	0	417	2	368 (88.7)	47 (11.3)	0 (0)
rs118011769	42	-	0	42(100)	0	417	2	365 (88)	50 (12.1)	0 (0)
rs75670657	42	-	0	41 (97.6)	1 (2.4)	417	5	361 (87.6)	51 (12.4)	0 (0)

Supplementary Table 2: Top 10 Significant SNPs with GT Frequencies of No TB Recurrence vs TB Recurrence Individuals										
SNP	**No TB Recurrence**	**Omitted**	**Ref/Ref, n (%)**	**Ref/Alt, n (%)**	**Alt/Alt, n (%)**	**TB Recurrence**	**Omitted**	**Ref/Ref, n (%)**	**Ref/Alt, n (%)**	**Alt/Alt, n (%)**
rs17075658	369	0	367 (99)	2 (1)	0 (0)	58	0	50 (86.2)	8 (13.8)	0 (0)
rs1987349241	369	0	257 (70)	105 (28)	7 (2)	58	0	23 (39.7)	29 (50)	6 (10.3)
rs4299914	369	3	57 (15.5)	204 (55.7)	105 (28.6)	58	0	25 (43.1)	17 (29.3)	16 (27.6)
rs7913065	369	0	233 (63.1)	112 (30.3)	24 (6.5)	58	0	18 (31)	36 (62.1)	4 (6.9)
rs4950973	369	1	319 (86.6)	47 (12.7)	2 (0.5)	58	0	35 (60.3)	22 (37.9)	1 (1.7)
rs3786544	369	0	204 (55.2)	147 (39.8)	18 (4.8)	58	0	14 (24.1)	36 (62.1)	8 (13.8)
rs17272023	369	0	355 (96.2)	14 (3.7)	0 (0)	58	0	46 (79.3)	11 (19)	1 (1.7)
rs116389662	369	0	335 (90.7)	32 (8.6)	2 (0.5)	58	0	40 (69)	18 (31)	0 (0)
rs79644900	369	0	367 (99.4)	2 (0.5)	0 (0)	58	0	51 (87.9)	6 (10.3)	1 (1.7)
rs77246883	369	0	342 (92.6)	25 (6.7)	2 (0.5)	58	0	42 (72.4)	16 (27.6)	0 (0)

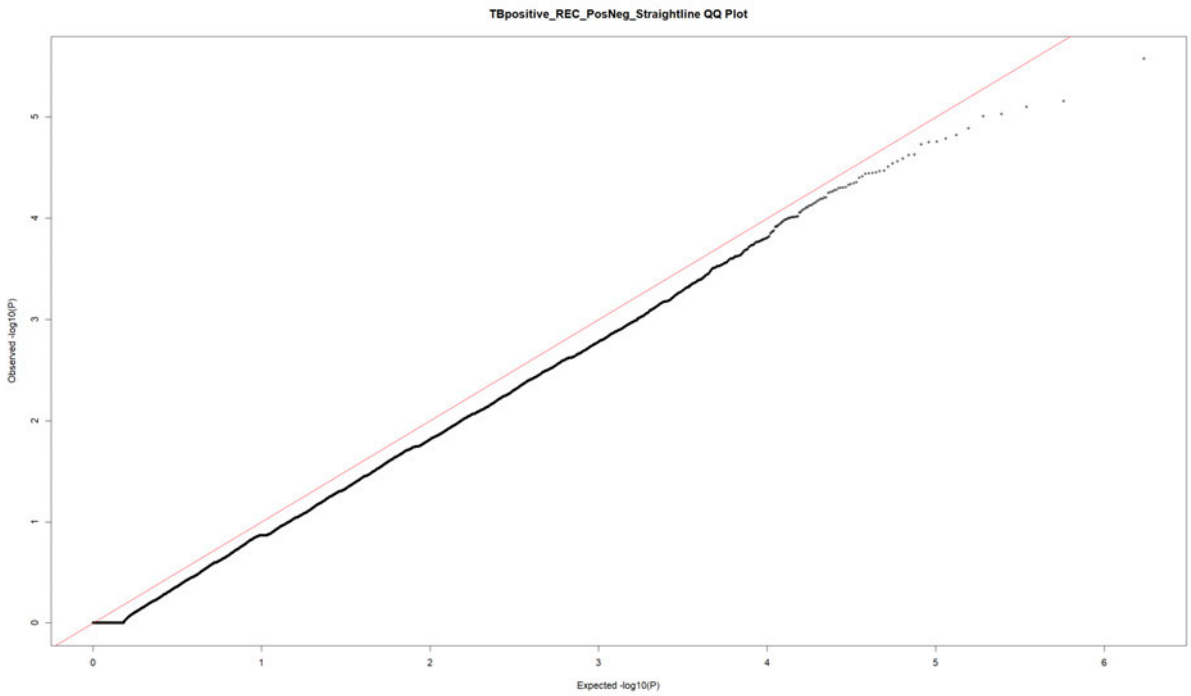
Supplementary Table 3: Top 10 Significant SNPs with GT Frequencies of TB Positive vs TB Negative Individuals										
SNP	**TB -ve**	**Omitted**	**Ref/Ref, n (%)**	**Ref/Alt, n (%)**	**Alt/Alt, n (%)**	**TB +ve**	**Omitted**	**Ref/Ref, n (%)**	**Ref/Alt, n (%)**	**Alt/Alt, n (%)**
rs886040602	215	153	0 (0)	62 (100)	0 (0)	553	0	553 (100)	0 (0)	0 (0)
rs12199036002	215	154	0 (0)	61 (100)	0 (0)	553	0	553 (100)	0 (0)	0 (0)
rs112194146	215	154	0 (0)	61 (100)	0 (0)	553	1	552 (100)	0 (0)	0 (0)
rs1057517572	215	153	0 (0)	62 (100)	0 (0)	553	50	503 (100)	0 (0)	0 (0)
rs1064792936	215	153	1 (1.6)	61 (98.4)	0 (0)	553	0	553 (100)	0 (0)	0 (0)
rs11587334	215	154	0 (0)	61 (100)	0 (0)	553	0	552 (99.8)	1 (0.2)	0 (0)
rs78396100	215	153	0 (0)	62 (100)	0 (0)	553	2	549 (99.6)	2 (0.4)	0 (0)
rs1555578543	215	154	1 (1.6)	60 (98.4)	0 (0)	553	0	553 (100)	0 (0)	0 (0)
rs1555594837	215	154	1 (1.6)	60 (98.4)	0 (0)	553	0	553 (100)	0 (0)	0 (0)
rs886055595	215	154	1 (1.6)	60 (98.4)	0 (0)	553	0	553 (100)	0 (0)	0 (0)



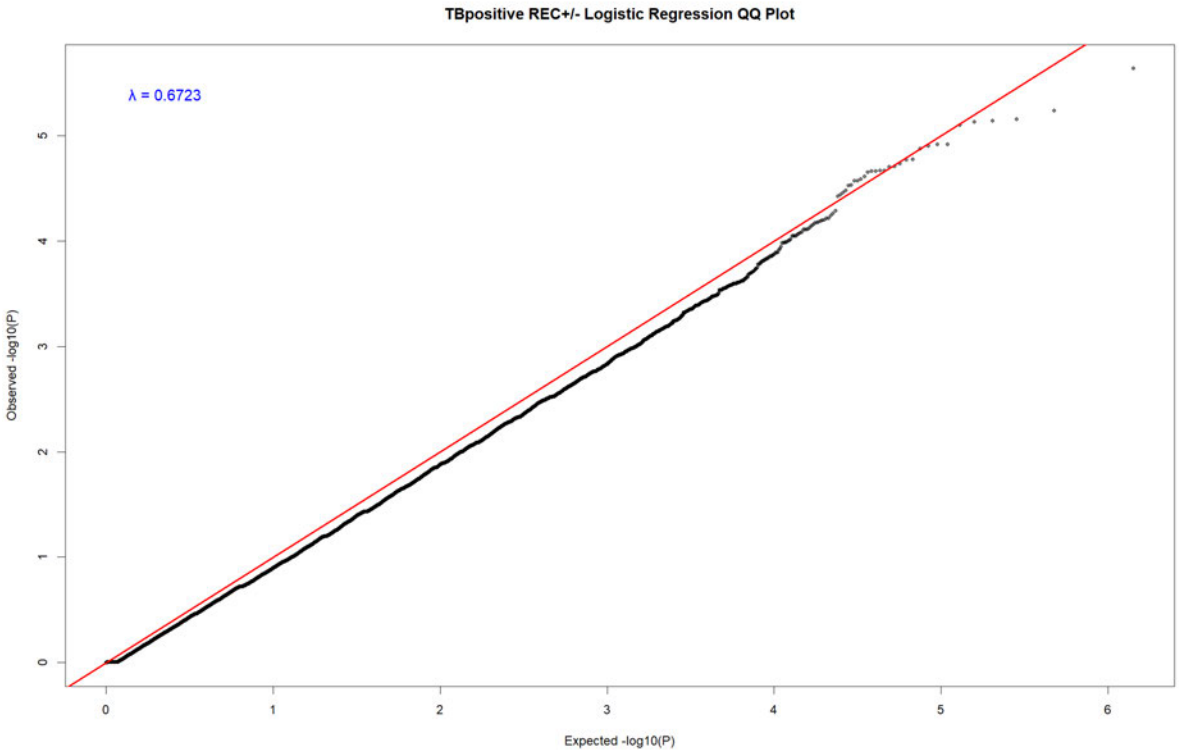
Supplementary Figure 1: Quantile–quantile (Q–Q) plot of genome-wide association results for HIV-1 susceptibility in tuberculosis (TB)-positive individuals (n=459). The represented Q–Q plot compares the observed distribution of $-\log_{10}(P)$ values from the genome-wide association analysis to the expected distribution under the null hypothesis of no association. Each point represents a single nucleotide polymorphism (SNP) tested for association with HIV-1 susceptibility among TB-positive patients. The solid diagonal line indicates the expected distribution of P-values under the null. Deviation of observed values above the expected line at the upper tail suggests the presence of true genetic associations with HIV-1 susceptibility. Early deviation from the null may indicate population stratification, cryptic relatedness, or other confounding factors. The genomic inflation factor (λ) is provided to assess the extent of test statistic inflation.



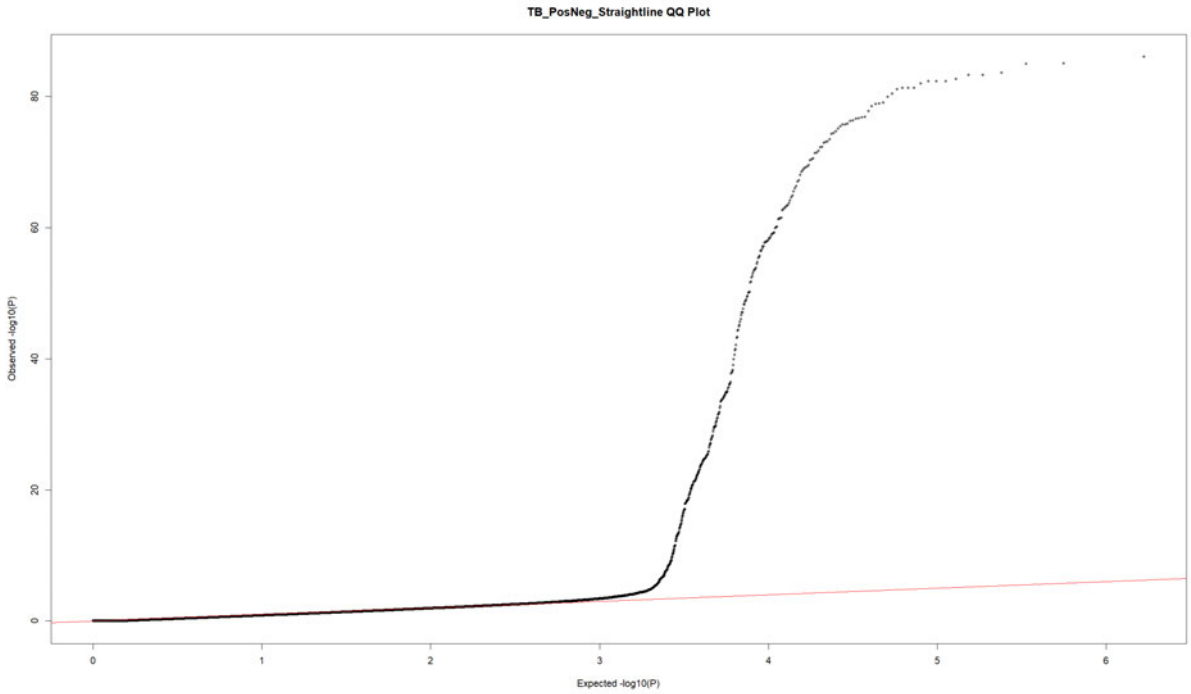
Supplementary Figure 2: Quantile–quantile (Q–Q) plot of logistic regression genome-wide association results for HIV-1 susceptibility among tuberculosis (TB)-positive individuals (n=459). The Q–Q plot displays the observed versus expected distribution of $-\log_{10}(P)$ values derived from logistic regression analysis testing the association between single nucleotide polymorphisms (SNPs) and HIV-1 susceptibility in TB-positive patients. The expected P-values under the null hypothesis of no genetic association are shown by the diagonal reference line. Each point corresponds to an individual SNP included in the genome-wide analysis. Concordance with the null distribution across most of the range indicates adequate control of population stratification and other confounders, while deviation of observed values above the expected line in the upper tail suggests potential true genetic associations. Systematic early deviation may reflect residual population structure, cryptic relatedness, or model misspecification. The genomic inflation factor (λ) is reported to quantify any inflation (λ= 0.5745) in the test statistics.



Supplementary Figure 3: Quantile–quantile (Q–Q) plot of genome-wide association results for single nucleotide polymorphism (SNP) associations with tuberculosis (TB) recurrence and those without TB recurrence (n=427). The Q–Q plot illustrates the relationship between the observed and expected distributions of $-\log_{10}(P)$ values obtained from the genome-wide association analysis of TB recurrence. Each point represents an individual SNP tested for association with the risk of recurrent TB. The diagonal reference line indicates the expected distribution of P-values under the null hypothesis of no association. Alignment of observed values with the expected distribution across the majority of SNPs suggests appropriate control for population structure and other confounding factors. Upward deviation of observed values from the expected line in the tail of the distribution indicates potential true genetic associations with TB recurrence.

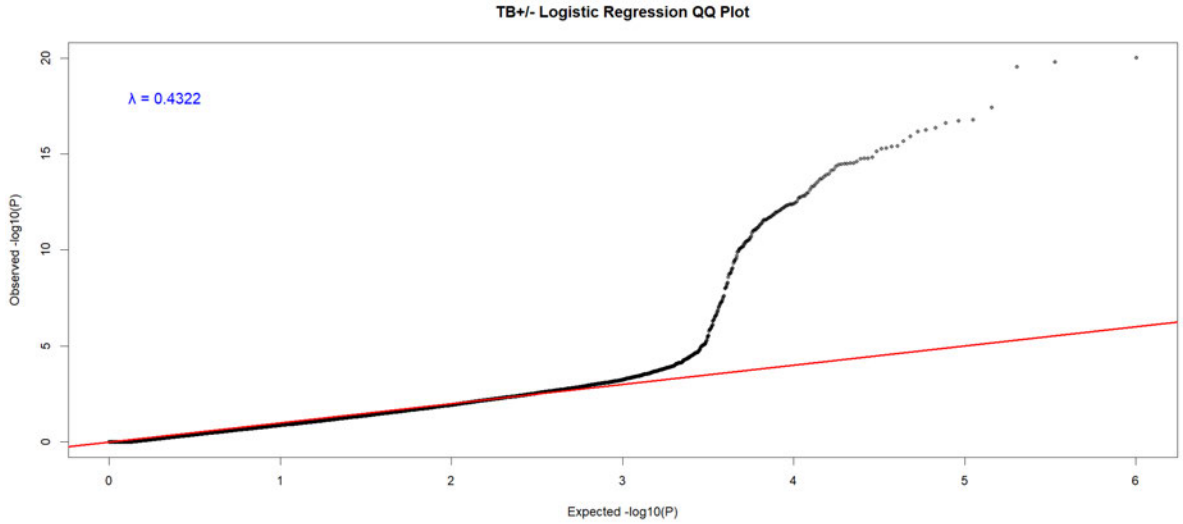


Supplementary Figure 4: Quantile–quantile (Q–Q) plot of logistic regression genome-wide association results for tuberculosis (TB) recurrence. The Q–Q plot shows the observed versus expected distribution of $-\log_{10}(P)$ values derived from logistic regression analysis testing the association between single nucleotide polymorphisms (SNPs) and TB recurrence status, comparing patients with recurrent TB to those without recurrence (n=427). Each point represents an individual SNP included in the genome-wide analysis. The diagonal reference line indicates the expected distribution of P-values under the null hypothesis of no association. Agreement between observed and expected values across most of the distribution suggests appropriate control of population stratification and other confounding factors. Upward deviation of observed values from the expected line in the upper tail indicates potential true genetic associations with TB recurrence. Systematic early deviation from the null may reflect residual confounding, cryptic relatedness, or model misspecification. The genomic inflation factor ,λ= 0.6723.



Supplementary Figure 5: Quantile–quantile (Q–Q) plot of genome-wide association results for single nucleotide polymorphism (SNP) associations with tuberculosis (TB) infection (n= 985).

The Q–Q plot compares the observed distribution of $-\log_{10}(P)$ values to the expected distribution under the null hypothesis of no association, based on genome-wide association analysis contrasting individuals with TB infection and TB-negative controls. Each point represents a SNP tested for association with TB infection status. The diagonal reference line indicates the expected null distribution of P-values. Concordance between observed and expected values across the majority of SNPs suggests adequate control of population stratification and other potential confounders. Deviation of observed values above the expected line in the upper tail indicates potential true genetic associations with susceptibility to TB infection. Early or systematic deviation from the null distribution may reflect residual population structure, cryptic relatedness, or technical artifacts.



Supplementary Figure 6: Quantile–quantile (Q–Q) plot of logistic regression genome-wide association results for single nucleotide polymorphism (SNP) associations with tuberculosis

(TB) infection (n= 985). The Q–Q plot presents the observed versus expected distribution of $-\log_{10}(P)$ values obtained from logistic regression analysis assessing the association between SNPs and TB infection status, comparing individuals with TB infection to TB-negative controls. Each point represents an individual SNP included in the genome-wide analysis. The diagonal reference line denotes the expected distribution of P-values under the null hypothesis of no association. Concordance between observed and expected values across most of the distribution indicates appropriate control for population stratification and other confounding factors. Deviation of observed values above the expected line in the upper tail suggests the presence of potential true genetic associations with susceptibility to TB infection. Early or systematic deviation from the null may indicate residual confounding, cryptic relatedness, or model misspecification. The genomic inflation factor, $\lambda = 0.4322$

Chapter 5

Preface:

In chapter 4 we investigated genetic variation with an unbiased approach focusing on genome-wide markers and their association with HIV predisposition, TB infection and recurrence. We reported on the top 10 significant markers associated with each of these phenotypic groups of which the majority are either to date understudied or have also been known to be associated with other clinical implications. Our final investigation in this research study involves determining the extent to which *ABC and SLC* variants as well as genome-wide markers are associated with variable drug PK levels in TB patients.We were particularly well positioned to undertake this investigation owing to our access to rigorously phenotyped samples derived from participants enrolled in CAPRISA IMPRESS trial who had undergone moxifloxacin treatment. This cohort provided a uniquely advantageous opportunity to focus specifically on the contribution of ABC and SLC drug transporter family of genes within the South African population. The availability of these high-quality samples created an optimal platform for interrogating the genetic determinants of treatment response, while our genome-wide, unbiased analytical framework ensured comprehensive coverage across all loci relevant to transporter-mediated drug disposition. Our main aim in this study was to define key drug transporter variants specific to African population and highlight the understudied variants in addition to reporting any other genome-wide variants with strong influence on drug PK levels.

Associations of *ABC* & *SLC* gene family drug transporter variants and understudied SNPs with moxifloxacin drug PK levels in *Mycobacterium tuberculosis* infected African patients by genome-wide association study.

By Sahil Tulsi [1], Prem Prakash Sharma [1], Anmol Gokul [1], Anushka Naidoo [2], Kogieleum Naidoo [2], [1] & Veron Ramsuran [1,2,3]

[1]School of Medicine, College of Health Sciences, University of KwaZulu-Natal, Durban, South Africa, [2] Centre for the AIDS Programme of Research in South Africa (CAPRISA), University of KwaZulu-Natal, Durban, South Africa, [3] International Foundation for Integrative Medical Research, Cape Town, South Africa

*Authors to whom correspondence should be addressed:
sahiltulsi@gmail.com; ramsuranv@ukzn.ac.za
Conflict of interest: None declared.
In submission in: *Pharmacogenetics and Genomics*
Manuscript ID: PGEN-2025-096

Abstract

Individuals of African descent harbour a great deal of genome-wide diversity that contribute to variance in pharmacodynamics and differential drug levels. In this genome-wide associations study we investigated the potential effects of mainly underexplored genetic variants in various *ABC* and *SLC* family genes in association with moxifloxacin drug levels at two time points from South African patients receiving undergoing TB therapy. Genome wide variants were genotyped from patient samples by microarray technology with GWAS analysis performed post genotyping. Various genome-wide markers and *ABC/ SLC* variants were identified with suggestive significance ($p \leq 1.0E^{-05}$) that are yet to be further investigated and clinically reported on with varying allele frequencies compared to European and Asian populations. Solute carrier intronic variants (rs4414423 & rs7805940) were notably reported as nominal significant SNPs at both time points with homozygous reference genotypes A/A [$p= 9.50e^{-05*}$, OR= 0.18 (0.04-0.71)] and homozygous reference T/T genotypes [$p= 7.21e^{-05*}$, OR= 0.15 (0.04-0.65)] respectively which were associated with higher median drug levels of moxifloxacin compared to other genotypes specific to each variant. Additionally several genome-wide variants were reported to have known clinical associations that differed from the

phenotypes that this study focussed on in the context of drug PK levels. Evidently genome-wide markers both known and unreported are playing differential roles with various associations across global populations. Further studies that include larger cohorts with specific focus on the effects of genetic variation and their association with varying moxifloxacin pharmacokinetic drug levels need to be better understood in larger studies that include intronic region variants in solute carriers particularly.

Introduction

Pharmacogenetics and genetic diversity particularly focuses on variants across the human genome and their impact and influence on how a patient metabolizes and responds to treatment (1). African populations are very unique due to their diversity from a genetic perspective, representing only a very small portion of global databases to date of less than 3% (2). The underrepresentation of genetic information from African populations reveals how little these genomes have been explored (3). Over time several studies have investigated drug transporter genes from the *ABC* gene family (4). A total of 49 genes have been identified within the *ABC* family that consists of transporters genes which include proteins known to elicit great clinical impact such as the transportation of various drugs and metabolites (5). Another gene family of recent significant interest has been solute carriers (SLC) with over 300 individual proteins that have been categorized into 47 families of *SLCs* that have been identified as passive transporters and exchangers. (6). Several genes that are members of the SLCO and *SLC22* subfamilies are crucial to drug transport in tuberculosis patients receiving anti-TB therapies (7).

Genome-wide association studies are crucial to determine the effect of both drug transporter gene variants in addition to the many more genome-wide biomarkers that are key in understanding treatment diversity in patients of underrepresented populations. Genotyping by microarray technology is known for its value in enabling the ability to screen for millions of known markers across the human genome at a reduced costing compared to genotyping be sequencing methods (8). The market for arrays available commercially is growing rapidly as they bring additional insights on markers that can help to confirm clinical associations, ancestry of individuals and pharmacogenomic (PGx) markers of interest among other clinical annotations to help better understand that patients metabolize treatments uniquely. Identification of underexplored mutations in poorly represented populations provide useful

clinical insights of biomarkers linked to various phenotypes from a pharmacokinetic and pharmacodynamic perspective.

The magnitude of several studies focusing on known single nucleotide polymorphisms (SNPs) that have been known to be associated with variance in drug PK levels has increased over the recent years of research (9) . Treatment diversity and variance in treatment outcomes in patients of African descent needs to be better characterized as the possibility of novel variants may play significant roles which we are yet to discover. Technology has evolved from traditional methods such as real-time PCR due to the limitations of single SNP investigation and low sample volumes making it more challenging to timeously obtain statistical power. As the goal of sequencing African populations at greater depth will be a large focus for several initiatives across the continent, array content is most likely to present a more accurate build of an African genome.

It is imperative that more detailed characterization of African genomes will surface new insights and findings that will help clarify and improve our understanding of treatment metabolism in African patients. Regimens for treating tuberculosis (TB) have barely been optimized for almost 5 decades leading to sub-optimal patient outcomes in African patients (10). A wealth of studies have investigated the effect of drug transporter genes that have been associated with drug level fluctuations in American, European and Asian populations but may not necessarily play the same role in African populations (11). Sub-optimal treatment outcomes in patients of African descent are becoming increasingly common, warranting further studies to investigate the molecular mechanisms that contribute to disparities in patient responses. Investigation of mutations in single drug transporter genes are limiting and often fail in providing a full understanding of patients drug metabolism and treatment success. With Africa as a diverse population, there may be known mutations which play a significant role in other populations but not necessarily impactful in the African ancestry which leave room for investigation of looking at marker across the genome that could potentially play a role in drug pharmacokinetics.

Treatment success is not only defined and dictated by a patients adherence to treatment, but also linked to several factors such age, metabolism, genetics, epigenetic factors among other variables (12). With the African genome being severely understudied, this study focuses on both known pharmacogenomic (PGx) markers and novel PGx markers in the context of African diversity. Understanding and defining key markers that may or may not be unique to African tuberculosis (TB) patients can help identify star alleles which represent standardized names for specific sets of genetic variations (regarded as haplotypes) occurring within specific genes that

are often associated with drug metabolism and patient response. The use of commercial PGx focused arrays provide a significant amount of value when analyzed together with Drug PK levels and patient outcomes.

Methods

Study Population & sample collection

The research performed in this study investigated the genome wide association of various genetic factors contributing to variable drug PK levels in patients on moxifloxacin treatment. A total of 58 retrospectively stored samples from the Centre for the AIDS programme of research in South Africa (CAPRISA) study - Improving retreatment success (IMPRESS) were comprised of either peripheral mononuclear blood cells (PBMCs) or buffy coats or peripheral mononuclear cells (PBMCs)] were genotyped by microarray. Additional details on study participants, inclusion/ exclusion criteria and study designs are available at ClinicalTrials.gov (IMPRESS, NCT0211468)). All individuals included in this study were adults above 18 years of age, from KwaZulu-Natal, South Africa and of African descent.

Processing of laboratory samples

Genomic DNA isolation and Genotyping by MicroArray

Genomic DNA (gDNA) was manually extracted from all sample types that included human specimens consisting of whole blood, sputum & peripheral blood mononuclear cells using standard commercial kits optimized for highest purity, yields and integrity. The quality and quantity of DNA was observed via absorbance ratios (OD_260/OD_280: 1.8–2.0; OD_260/OD_230 > 1.5) and nucleic acid integrity was confirmed by 1% agarose gel electrophoresis. To confirm DNA input requirements of 100ng as per Axiom PMD Array (ThermoFisher Scientific, USA), we used a fluorometric based assay (Qubit, ThermoFisher Scientific). For the genotyping by microarray approach we utilized the Axiom Precision Medicine Diversity Array that investigates >850 000 genome-wide markers that in addition to the SNP backbone consist of approximately 5000 variants in over 1100 genes with known PGx value (ThermoFisher Scientific, USA).

The template gDNA was amplified using whole-genome amplification (WGA) without being biased to any genes as per manufactures instructions. Post amplification we utilized enzymatic

fragmentation to segment the amplicons into smaller fragments. Fragmented DNA was then hybridized to the microarray overnight in a hybridization oven. Hybridized DNA was then exposes to a series of enzymatically staining with fluorescent dyes and washing. To conclude the workflow the microarrays were scanned on the GeneTitan platform (ThermoFisher Scientific, USA) which captures and acquires the fluorescent intensities emitted by the dyes. Genotyping raw data was collected and processed by Axiom Analysis suite (ThermoFisher Scientific).

Data normalization was applied to adjust for variance in probes from batch/ lot effects from manufacturing. Calling of genotypes employed clustering algorithms with the use of curated cluster files. Low quality calls or poor performing SNPs with low call rates or high error rates were omitted as part of quality control metrics based on software defined thresholds (acceptable call rates $\geq 98\%$). Additional quality control checks were performed post-processing to filter for minor allele frequencies (MAF) and confirming Hardy-Weinberg equilibrium. Output genotyping data was the input for genome-wide association study (GWAS) and imputed into reference panels (1000 genomes) (. Data visualization of genome-wide results were observed and presented as Manhattan plots.

Pharmacokinetic Analysis – Plasma samples sub-aliquoted after centrifuging EDTA-containing tubes at 3000 rpm. The samples were preserved on ice before being transported to the CAPRISA laboratory, where they were stored at -80°C. All samples were delivered from the clinic to the laboratory within one hour of collection. Moxifloxacin concentrations in plasma were analyzed at the pharmacology laboratory of the KwaZulu-Natal Research Institute for Tuberculosis and HIV (K-RITH), using high-performance liquid chromatography (HPLC) followed by mass spectrometry with two time points that were one month apart (T1 and T2). The bioanalytical method used was developed and validated according to the United States Food and Drug Administration (US FDA) 2011 guidelines (1). For sample preparation, proteins were precipitated with acetonitrile, then diluted with water. Chromatographic separation was performed on a Zorbax C18 column (3.5 μm, 50 mm × 2.1 mm). Detection was conducted using an ABI Sciex 5500 QTrap mass spectrometer (AB Sciex LLC, MA, USA), operating in positive ion mode. The monitored ion transitions were (in m/z): moxifloxacin – 402.1 → 358.2 and 402.1 → 364.1; ciprofloxacin (internal standard) – 331.6 → 231.0 and 331.6 → 288.1. Moxifloxacin levels were quantified using isocratic elution with a mobile phase composed of 22% acetonitrile in water containing 0.1% formic acid. Each sample injection was 2 μl, and the total run time per analysis was 5 minutes. Method validation covered a concentration range

from 50 to 5000 ng/ml. Precision was confirmed using quality control samples at low, medium, and high levels, with results ranging from 8.4% to 19.4%, and accuracy between 101.9% and 105%. Carryover at the lower limit of quantification was calculated at 5.4%. Data acquisition, peak integration, and analyte quantification were performed using Analyst® software version 1.6.2 on a Windows 7 PC integrated with the LC-MS/MS system.

Statistical Analysis

Post Genotyping by MicroArray - Data Quality control (QC) analysis was performed with the use of the Axiom Array Analysis Suite, version 4.0.3.3 (ThermoFisher Scientific, Massachusetts, USA), with best practice workflow. The QC analysis workflow was performed for samples and only auto calls genotypes the samples that have passed the manufacturers defined QC thresholds, followed by classifying the probe sets used to identify the genotypes which are recommended for statistical tests in downstream study.

Genome-wide association analyses was performed to identify specific known genetic variants associated with traits of interest. Quality control (QC) of genotype data was conducted using PLINK v1.9, omitting data from individuals with high missingness (>5%), gender discrepancies, excess in heterozygosity, or cryptic relatedness (identity-by-descent > 0.125). SNP-level filtering excluded variants with a call rate <95%, minor allele frequency (MAF) <1%, or deviation from Hardy-Weinberg equilibrium ($p < 1 \times 10^{-6}$ in controls or the entire cohort, depending on study design). Association testing was conducted with the use of Wilcoxon rank sum tests and adjusting for multiple testing comparisons we performed Bonferroni correction to determine genome-wide significance was set at $p \leq 6.0 \times 10^{-8}$ for the moxifloxacin drug PK AUC analysis. Suggestive associations were only considered at $p \leq 1.0 \times 10^{-5}$ and indicated with a "*" for ease of reference. Manhattan plots were generated (Figure 1 & Figure 2) to visualize the distribution of association statistics and identify the specific genomic loci of interest highlighting SNPs with highest impact and association on the respective phenotypes. Quantile-Quantile plots based on the genome-wide association analysis that compared the observed distribution of $-\log_{10}(P)$ values to the expected distribution under the null hypothesis of no association were generated using R software and included in the Supplementary data, Figures: 1-2.

Ethical considerations

This study and research performed was granted expedited ethical approval by the Biomedical Research Ethics Committee of the University of KwaZulu-Natal (BREC/00004110/2022).

Results

Drug PK levels were recorded from all 46 study participants at both time points T1 and T2 (on average one month apart for each patient) from plasma samples obtained from participants whole-blood taken approximately 24 hours after oral administration of moxifloxacin. For timepoint 1, from a total of 58 patient samples only 54 samples were included in drug PK analysis due to samples being below quality control (QC) criteria and 49 samples included in analysis for timepoint 2. From a total of 864, 313 SNPS, we omitted 6, 714 SNP assays due to performance below acceptable frequencies for the assay. We used a Wilcoxon rank-sum test to determine significance of all 857, 599 SNPs in the microarray. Manhattan plots representing genome-wide association results for both timepoints are shown in Figure 1. and 2. respectively. We adjusted for multiple comparisons and confirmed the Bonferroni corrected p-value of 6.0×10^{-08}. Genotypes of individuals were observed and reported in two groups (Group 1 and Group 2) for each SNP and listed in each table in descending order of P value significance. In some cases the third genotype group was listed as a separate data field for the same SNP.

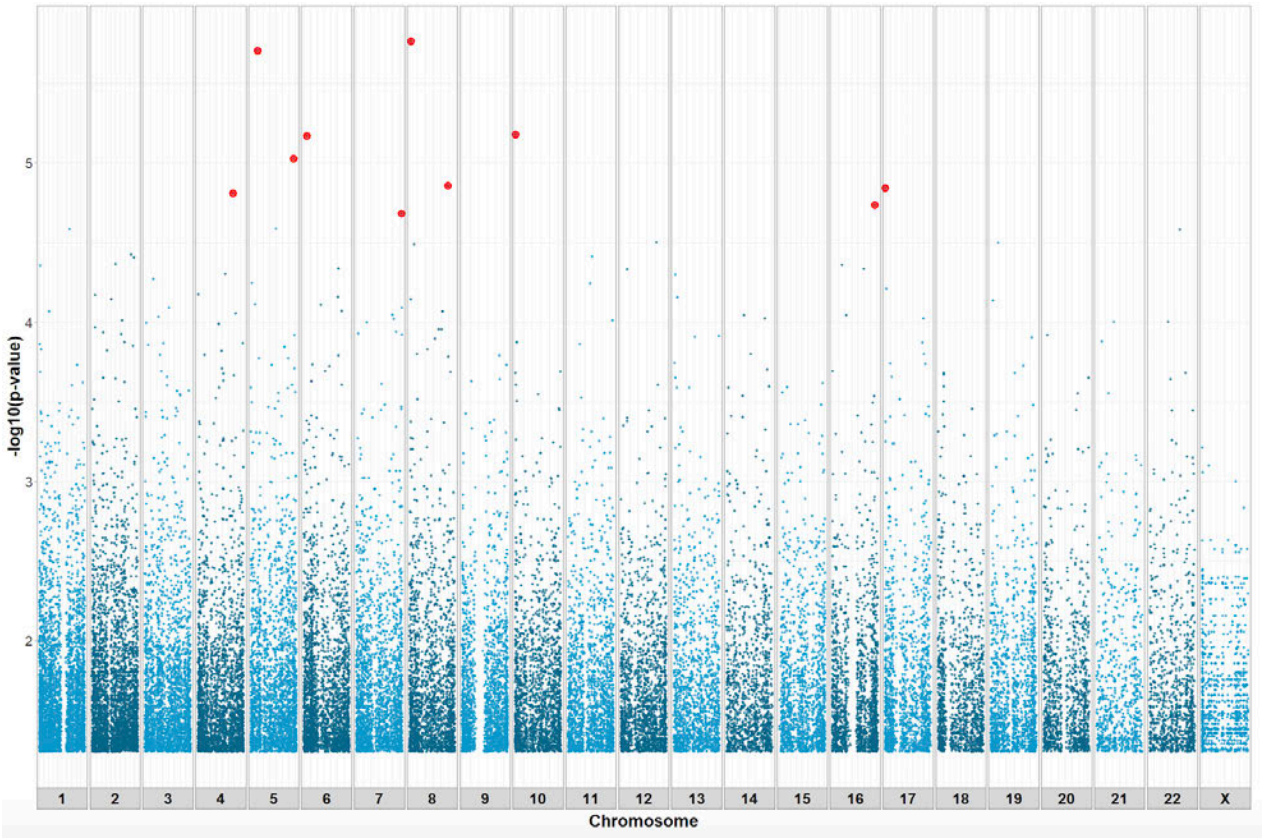


Figure 1: Manhattan plot representing genome-wide association results for moxifloxacin AUC levels (n=58) that illustrates the distribution of association p-values for all genome-wide single nucleotide polymorphisms (SNPs) tested for their relationship with moxifloxacin area-under-the-curve (AUC) pharmacokinetic levels at timepoint 1 (T1). The x-axis displays SNPs ordered by their chromosomal position from chromosome 1 through chromosome 22, thereby providing a continuous representation of genomic location across the entire genome. The y-axis represents the strength of statistical association, plotted as the negative logarithm of the p-value ($-\log_{10} p$), with higher points indicating stronger evidence of SNP–AUC association. The top 10 SNPs with the lowest p-values (greatest statistical significance) are highlighted as red data points, marking genomic regions that may be of particular biological or pharmacogenomic interest. Although none of these SNPs reach the stringent genome-wide significance threshold after applying Bonferroni correction, some SNPS indicated by red dots exceed the most commonly accepted threshold for suggestive significance ($p < 1.0\times10^{-05}$) and therefore pose as potentially meaningful candidates for further replication and functional follow-up studies.

A total of 41, 603 genome-wide SNPs yielded P values that were ≤ 0.05 for T1 and 43, 114 SNPs for timepoint 2. Variant rs6982690 (uncharacterized gene) was observed at the SNP with the highest level of suggested significance (p= 1.715×10-06), OR = 0.04, 95% CI (0.00-0.80) however not considered genome-wide significant when compared to Bonferroni corrected p value. The top 10 SNPs with the highest associations with moxifloxacin drug PK levels at T1 are represented in supplementary table A. Wilcoxon rank-sum analysis for T2 reported

rs9865956 (uncharacterized gene) as the variant with the highest suggestive significance (p= $2.11E^{-06}$), OR= 63, 95% CI (3.22-1231.34) even though not regarded as genome-wide significant after Bonferroni comparison. Supplementary table B summarizes the top 10 genome-wide SNPs from the GWAS analysis that was associated with drug PK levels at timepoint 2.

We then narrowed down the reporting and filtered the genotyping data for *ABC* and *SLC* family genes to investigate the SNPs that may yield p-values ≤ 0.05 in association with moxifloxacin drug PK levels to a total of 76 SNPs in 24 *ABC* family genes were identified at T1. The summarized SNPs in Supplementary Table C all yielded p values ≤ 0.01, with the top 10 suggested significant *ABC* gene family SNPs at T1 summarized and represented below in Table 1.

Table 1. Table Summary presenting an overview of the top 10 single nucleotide polymorphisms (SNPs) within the ATP-binding cassette (ABC) gene family that demonstrated the strongest statistical associations with moxifloxacin area-under-the-curve (AUC) pharmacokinetic levels at timepoint 1 (T1).

Chr	Gene Name	Pos	Ref/Alt	SNP	Group 1	Group 2	Group 1 n (%)	Group 2 n (%)	Group 1 Median	Group 2 Median	p-Value	OR, 95% CI
7	ABCB5	20678540	T/C	rs7805940	T/T	C/C	38 (70.4)	3 (5.6)	6.71	18.93	4E-04	0.084 (0.00-1.75)
21	ABCG1	42286991	C/G	rs425215	G/G	C/C	18 (33.3)	15 (27.8)	14.17	5.76	7E-04	13.75 (2.54-74.3)
21	ABCG1	42286991	C/G	rs425215	C/G	G/G	21 (38.9)	18 (33.3)	6.60	14.17	0.001	0.12 (0.02-0.56)
21	ABCG1	42304243	G/A	rs15661	G/A	G/G	23 (42.6)	7 (13.0)	6.60	15.08	0.001	0.10 (0.01-1.04)
11	ABCC8	17423592	A/G	rs929234	G/G	A/A	14 (25.9)	9 (16.7)	10.34	5.65	0.002	29.3 (2.55-336)
7	ABCB5	20661377	G/A	rs73684681	G/G	G/A	32 (59.3)	18 (33.3)	5.99	12.58	0.002	0.12 (0.03-0.49)
17	ABCA5	69253458	C/T	rs75169077	C/C	C/T	49 (90.7)	3 (5.6)	8.80	3.37	0.003	7.89 (0.38-160)
7	ABCF2-H2BK1	151225377	A/T	rs188409705	A/A	A/T	49 (90.7)	5 (9.3)	7.74	16.54	0.006	0.22 (0.02-2.12)
7	ABCB5	20678540	T/C	rs7805940	T/T	T/C	38 (70.4)	13 (24.1)	6.71	12.60	0.006	0.175 (0.04-0.74)
17	ABCA6	69129236	G/A	rs139091810	G/G	G/A	47 (87.0)	7 (13.0)	7.58	15.08	0.007	0.049 (0.00-0.92)

For each SNP, identified by its respective rsID**,** the table reports the number of participants stratified by genotype, along with the corresponding median AUC drug concentrations (mmol/L) observed within each genotype group. Additionally, the table includes the p-values**,** odds ratios (ORs), and 95% confidence intervals (CIs) associated with each SNP, providing a comprehensive summary of both effect size and statistical uncertainty. Despite these SNPs not meeting the threshold for suggestive statistical significance within the current study, it is important to note that the limited sample size may reduce the power to detect true genetic effects. As such, these variants may exhibit stronger or more definitive associations in future investigations employing larger, well-powered cohorts**,** where increased sample size

could enhance the ability to detect modest genetic contributions to variability in moxifloxacin pharmacokinetics. Thus, the SNPs highlighted here remain promising candidates for follow-up analyses and validation in expanded studies.

Variant rs7805940 found in gene *ABCB5* yielded the greatest level of significance (p = 0.0004, OR=0.0084) for genotype CC with higher AUC medians compared to patients who genotyped as T/T. Another variant of significance reported in *ABCG1* is rs425215 where patients who genotyped as G/G had higher median levels compared to C/C genotyped patients (p= 0.0007, OR= 13.75) suggesting an approximate 13-fold increase in associated moxifloxacin drug levels. Median AUC values of genotype G/G in rs425215 were also significantly higher compared to C/G genotypes (p= 0.0012, OR=0.12). In gene *ABCC8* we observed rs929234 with higher AUC median levels in G/G genotypes compared to A/A genotypes with a p value = 0.002, OR= 29.3 suggestive of a 29-fold increase in associated drug levels of moxifloxacin. Another variant with strong suggestive significance is rs75169077 in *ABCA5* where genotype C/C had significantly higher levels of moxifloxacin compared to C/T genotyped individuals (p= 0.0028, OR= 7.89).

For *SLC* family genes at T1 we identified a total of 468 SNPs with p-values ≤ 0.05 from 168 genes regarded as variants with suggested significance. In supplementary table D only SNPs with p values ≤ 0.01 have been summarized for reference. The top 10 SNPs with suggested significance are reported in Table 2 for *SLC* family gene variants associated with moxifloxacin drug levels at T1. Variant rs4414423 in *SLC24A4* yielded the strongest p value = 9.5×10^{-05} (*), OR= 0.18 for genotype A/A compared to genotype A/G suggesting an 82% increase in median moxifloxacin drug levels and has been regarded as a suggestive SNP in this study.

Table 2. Summary of the top ten single nucleotide polymorphisms (SNPs) within the solute carrier (SLC) gene family that demonstrated the strongest associations with moxifloxacin area-under-the-curve (AUC) pharmacokinetic levels at timepoint 1 (T1).

Chr	Gene Name	POS	Ref/ Alt	SNP	Group 1	Group 2	Group 1 n (%)	Group 2 n (%)	Group 1 Median	Group 2 Median	p-Value	OR, 95% CI
14	SLC24A4	92423645	A/G	rs4414423	A/G	A/A	31 (57.4)	16 (29.6)	6.11868	15.189	9.50e-05*	0.18 (0.04-0.71)
8	SLCO5A1	69694927	A/G	rs10504458	A/A	A/G	32 (59.3)	17 (31.5)	11.6722	4.4892	0.0004	10.2 (2.39-43.93)
7	SLC29A4	5303497	C/T	rs71529388	C/C	C/T	48 (88.9)	5 (9.3)	9.3542	3.6338	0.00054	12.9 (0.67-247.3)
3	SLC7A14-AS1	170654068	G/A	rs7611820	G/A	G/G	19 (35.2)	11 (20.4)	9.91348	4.4567	0.00081	9.75 (1.59-59.6)
15	SLC12A6	34269767	T/C	rs74010046	T/T	T/C	48 (88.9)	6 (11.1)	9.3542	4.0255	0.00089	16.6 (0.88-311.8)
8	SLCO5A1	69713707	A/G	rs6472480	A/A	A/G	23 (42.6)	21 (38.9)	5.64536	9.9135	0.00092	0.13 (0.03-0.53)
12	SLC15A4	128793867	G/C	rs76392376	G/G	G/C	45 (83.3)	8 (14.8)	9.03361	3.8915	0.00097	8.75 (0.99-77.11)
14	SLC25A21	37118109	C/T	rs66862591	C/C	C/T	17 (31.5)	28 (51.9)	5.64536	10.415	0.00099	0.14 (0.03-0.57)
17	SLC26A11	80252305	G/A	rs73432173	G/A	A/A	21 (38.9)	3 (5.6)	9.67478	3.1767	0.00099	7.66 (0.35-166.6)
5	SLCO6A1	102386596	G/T	rs78967500	G/G	G/T	41 (75.9)	12 (22.2)	6.59816	14.748	0.00126	0.128 (0.02-0.66)

For each SNP, the table reports the number of patients categorized by genotype, along with the corresponding median AUC drug concentrations (mmol/L) within each genotype-defined group. Additionally, p-values**,** odds ratios (ORs)**,** and 95% confidence intervals (CIs) are provided to describe both the magnitude of association and the precision of the estimated effect sizes. Among the SNPs evaluated, rs4414423 emerged as the top-ranking variant**,** showing the lowest p-value **(**$p = 9.50 \times 10^{-5}$**).**

SLCO5A1 variant rs10504458 showed higher AUC median levels of moxifloxacin in genotype A/A compared to A/G genotypes (p= 0.0004, OR= 10.2) suggesting more than a 10-fold increase observed. A similar observation was noted with rs71529388 of gene *SLC29A4* showing higher median drug levels observed in genotypes C/C compared to C/T (p= 0.00054, OR= 12.9). A variant of significance also observed was rs74010046 in *SLC12A6* showing higher AUC median levels (p= 0.00089, OR= 16.6, 95% CI 0.88-311.8) for T/T genotype compared to T/C genotype. For rs76392376 of *SLC15A4* patients who genotyped as G/G had higher median AUC values (p=0.00097, OR= 8.75, 95% CI 0.99-77.11) compared to G/G genotypes indicative of a 8-fold difference in associated drug levels. Genotype G/T of rs78967500 in *SLC06A1* is associated with an 87% increase in median drug level (p= 0.00126, OR= 0.128) compared to genotype G/G.

Table 3. Summary of the top 10 SNPs within ABC gene family that showed the strongest associations with moxifloxacin area-under-the-curve (AUC) levels at T2.

Chr	Gene Name	POS	Ref/ Alt	SNP	Group 1	Group 2	Group 1 n (%)	Group 2 n (%)	Group 1 Median	Group 2 Median	p-Value	OR, 95% CI
9	ABCA1	104818199	C/T	rs58866705	C/T	C/C	8 (16.3)	40 (81.6)	5.52	10.26	0.0002	0.12 (0.01-1.04)
21	ABCG1	42304243	G/A	rs15661	G/A	G/G	21 (42.9)	6 (12.2)	9.63	15.81	0.0015	0.07 (0.00-1.40)
21	ABCG1	42304243	G/A	rs15661	A/A	G/G	22 (44.9)	6 (12.2)	8.07	15.81	0.0021	0.05 (0.00-0.90)
1	ABCA4	94086591	A/G	rs138785943	A/A	A/G	45 (91.8)	4 (8.2)	10.18	5.37	0.0023	10.26 (0.52-201.64)
1	ABCD3	94502789	C/T	rs12130330	C/C	C/T	31 (63.3)	15 (30.6)	8.02	13.07	0.0025	0.23 (0.059-0.89)
2	ABCB11	168969927	A/G	rs80049571	A/A	G/G	34 (69.4)	2 (4.1)	8.27	20.40	0.0032	0.13 (0.02-2.82)
1	ABCB10	229524044	T/C	rs10916508	T/T	C/C	25 (51.0)	6 (12.2)	10.33	5.44	0.0033	5.42 (0.55-53.27)
16	ABCA17P	2404539	A/G	rs4261541	G/G	A/G	45 (91.8)	3 (6.1)	9.57	17.57	0.0036	0.13 (0.01-2.57)
7	ABCB5	20678540	T/C	rs7805940	T/T	T/C	34 (69.4)	12 (24.5)	8.71	13.02	0.0039	0.21 (0.05-0.91)
7	ABCA13	48638485	G/A	rs1918583	A/A	G/A	43 (87.8)	5 (10.2)	9.09	17.94	0.0041	0.07 (0.00-1.39)

For each respective SNP, the number of patients grouped by genotype is provided, along with the corresponding median AUC concentrations (mmol/L)**,** p-values**,** odds ratios (ORs), and 95% confidence intervals (CIs)**.** Although none of these SNPs reach statistical significance after correction for multiple testing, they represent the most notable ABC variants at this timepoint and may still warrant follow-up evaluation in larger, better-powered studies to clarify their potential influence on moxifloxacin pharmacokinetics.

The subsequent analysis of timepoint 2 was then performed to investigate the top 10 identified *ABC* and *SLC* gene family SNPs. Supplementary table E summarizes *ABC* Gene family variants with p-values ≤ 0.01 in association with moxifloxacin drug PK AUC values at T2.

All of the top 10 hits with highest level of significance presented in Table 3 yielded P values less than $10E^{-05}$ indicating that none of the *ABC* gene family SNPs were regarded as suggestive SNPs in this study.

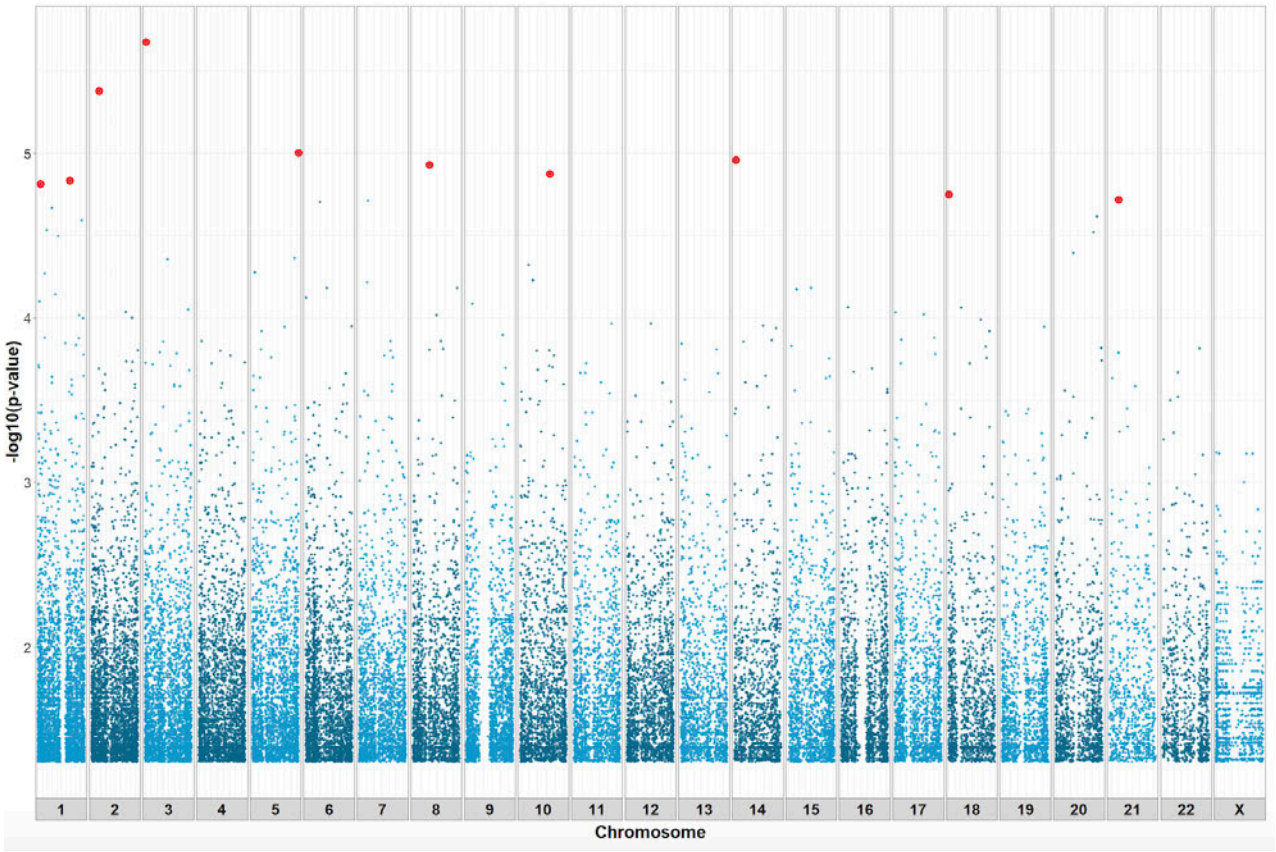


Figure 2. Manhattan plot displaying genome-wide association results with distribution of the association p-values with moxifloxacin area-under-the-curve (AUC) pharmacokinetic levels (n=58) at timepoint 2 (T2). The x-axis representing SNPs arranged by genomic position across all chromosomes, while the y-axis shows the strength of association expressed as $-\log_{10}(p)$, with higher values indicating stronger statistical evidence. The ten SNPs with the lowest p-values are highlighted as red data points**,** marking the variants with the most notable signals in this analysis. Whilst these SNPs represent the strongest associations detected at T2, they do not surpass the threshold for genome-wide significance after applying Bonferroni correction for multiple testing. Nonetheless, they are potential variants of concern for further exploration larger studies where increased statistical power may clarify their potential relevance to moxifloxacin pharmacokinetic variability.

Despite no SNPs with suggested significance were observed, variant rs58866705 of *ABCA1* showed that patients who genotyped as C/C were associated with an 82% increase in median drug levels compared to C/T genotyped individuals [p= 0.0002, OR= 0.12 and 95% CI (0.01-1.04)]. Another variant observed with a high odds ratio was rs138785943 of gene *ABCA4* with patient genotypes A/A had higher AUC median drug levels compared to A/A genotyped patients (p= 0.0023, OR=0.26, 95% CI-0.52-201.64).

Table 4. Summary of the top 10 SLC gene family SNPs showing the strongest associations with moxifloxacin area-under-the-curve (AUC) levels at T2.

Chr	Gene Name	POS	Ref/ Alt	SNP	Group 1	Group 2	Group 1 n (%)	Group 2 n (%)	Group 1 Median	Group 2 Median	p-Value	OR, 95% CI
1	SLC44A3-AS1	94651182	T/C	rs75457665	T/C	T/T	15 (30.6)	34 (69.4)	5.90	11.44	7.21e-05*	0.15 (0.04-0.65)
14	SLC25A21	36900653	A/G	rs17177739	A/G	A/A	7 (14.3)	42 (85.7)	5.15	10.33	0.000140139	0.05 (0.00-0.94)
20	SLC17A9	62956015	C/T	rs2427456	C/C	T/T	31 (63.3)	4 (8.2)	8.32	18.05	0.000152788	0.09 (0.00-1.86)
7	SLC26A4	107666761	C/T	rs2701688	C/T	T/T	22 (44.9)	6 (12.2)	11.58	6.51	0.000159261	26.86 (1.33-542.72)
11	SLC1A2	35341405	C/T	rs7942395	T/T	C/C	4 (8.2)	20 (40.8)	5.15	10.25	0.000376435	0.07 (0.00-1.59)
20	SLC24A3	19516069	G/A	rs16980713	G/G	G/A	44 (89.8)	5 (10.2)	9.33	17.94	0.000427923	0.06 (0.00-1.33)
15	SLC28A2-AS1	45262069	C/A	rs1060896	A/A	C/C	3 (6.1)	36 (73.5)	4.78	9.77	0.000437685	0.14 (0.00-2.96)
16	SLC38A8	84023351	C/T	rs9783761	C/T	C/C	28 (57.1)	10 (20.4)	8.27	14.10	0.000485224	0.02 (0.00-0.04)
1	SLC35D1	67021693	C/A	rs61780038	C/C	C/A	22 (44.9)	25 (51.0)	7.50	12.70	0.00048842	0.14 (0.04-0.53)
2	SLC25A12	171855597	T/TG	rs3836018	T/TG	TG/TG	25 (51.0)	19 (38.8)	7.35	12.70	0.000550753	0.08 (0.02-0.35)

For all variants in this table, the number of patients in each genotype group is shown alongside the corresponding median AUC values (mmol/L), p-values**,** odds ratios (ORs), and 95% confidence intervals (CIs).The SNP with the highest and suggested significance at T2 is one again an SLC family variant (rs75457665, p= 7.21×10^{-05}), reinforcing the potential relevance of solute carrier transporters in moxifloxacin pharmacokinetics. Although none of remaining SNPs reach statistical significance after Bonferroni correction**,** they remain the most notable signals within the SLC gene family and may merit further investigation in larger studies.

All *SLC* Gene family variants with P values ≤ 0.01 in association with moxifloxacin drug PK AUC values at T2 are summarized in supplementary table F. From the top 10 SLC SNPs, solute carrier variant rs75457665 of *SLC44A3-AS1* gene was observed as a suggestive SNP as represented in Table 4 with a p value of $7.21E^{-05}$ (*) for T/T genotyped individuals associated with 85% increase in median drug levels compared to T/C genotyped patients with an OR= 0.15, 95% CI (0.04-0.65). The remainder of the 9 SNPs in Table 4 were not considered as suggestive SNPs in this study with p values > 1.0×10^{-05}. Individuals who genotyped as T/T of rs2427456 showed the high median drug levels compared to C/C genotyped patients with an odds ratio of 0.09, CI 95% (0.00-1.86) and a p value = 0.00015. Variant rs2701688 of *SLC26A4* however showed a high odds ratio of 26.86, 95% CI (1.33-542.72) for C/T genotyped individuals having higher median drug levels compared to patients who genotyped as T/T (p= 0.00016).

Table 5. Common ABC and SLC gene family SNPs that showed statistically significant associations with moxifloxacin area-under-the-curve (AUC) levels at both T1 and T2.

SNP	Gene	T1, P Value	T1 OR, 95% CI	T2, p-Value	T2 OR, 95% CI
rs4414423	*SLC24A4*	9.50e-05*	0.18 (0.04-0.71)	0.0190338	0.21 (0.05-0.84)
rs7805940	*ABCB5*	0.000375235	0.084 (0.00-1.75)	0.0038	0.21 (0.05-0.91)
rs74010046	*SLC12A6*	0.00089	16.6 (0.88-311.8)	0.02468	13.15 (0.69-252.17)
rs66862591	*SLC25A21*	0.000987635	0.14 (0.03-0.57)	0.00468	0.14 (0.02-0.65)
rs15661	*ABCG1*	0.001451027	0.10 (0.01-1.04)	0.001533	0.07 (0.00-1.40)
rs73684681	*ABCB5*	0.0021	7.89 (0.38-160)	0.00373	0.30 (0.09-1.07)
rs75169077	*ABCA5*	0.0028	7.89 (0.38-160)	0.044	62 (0.37-155.88)
rs139091810	*ABCA6*	0.0071	0.049 (0.00-0.92)	0.03629	0.05 (0.00-0.85)

For each SNP, the p-values**,** odds ratios (ORs)**,** and 95% confidence intervals (CIs) are provided, allowing comparison of the magnitude and consistency of the effect across the two timepoints. Noting the study's sample size may have limited statistical power**,** potentially reducing the ability to detect additional true associations.

Among the top 10 *ABC* and *SLC* variants from each timepoint of the study, Table 5 summarizes the common SNPs observed at both timepoints with p values ≤ 0.05 . Eight variants (rs4414423, rs7805940, rs74010046, rs66862591, rs15661, rs73684681, rs75169077, rs139091810) were identified with P values that differed between T1 and T2. Variant rs4414423 which has been considered a SNP with suggestive significance at T1 ($p= 0.50E^{-05}$), had lower significance at T2 (p= 0.0190338) with an odds ratio of 0.18 initially at T1 followed by 0.019 at T2. Both SNPs rs66862591 & rs15661 each differed in p values between T1 and T2, however the odds ratio remained similar for both SNPs between the two time points (0.14 and 0.14) for rs66862591 and 0.10 and 0.07 for rs15661. A variant with a large difference in odds ratio between timepoints was rs75169077 which was observed as 7.89 (0.38-160) at T1 and 62 (0.37-155.88).

Discussion

Genome-wide Associations in association with moxifloxacin drug levels at both time points

Genetic studies have shown that African populations exhibit the greatest level of genetic diversity compared to any other population (2). Despite African genomes harbouring the greatest level of genetic diversity with less than 3% representation in global genetic databases (13), there are substantial gaps in understanding the genetic variation that need to be addressed over the coming years. This study embraces the reality of diversity in African genomes with understudied variants observed to have suggested significance with variable moxifloxacin drug levels.

From our genome-wide observations at timepoint 1 for the top 10 SNPs, we noted that specific variants rs6982690, rs2319809, rs145236962 to date have not been assigned to gene or published in ClinVar which are regarded as variants of uncertain significance (VUS). Although in this study the above mentioned genotype rs6982690 is not considered genome-wide significant after Bonferroni correction, the data suggests that genotype A/A is which was associated with higher median drug levels in this study compared to G/A genotypes potentially may be a marker that needs further investigation in African cohorts further warranted by the low allele frequencies observed within patients of African descent. It is highly likely that a similar trend can be observed with several understudied variants across the genome in African populations. The remainder of the variants in supplementary table 1 to date have not been linked to clinical associations or published which strengthens the suggestion that these underexplored variants may be of potential interest for further investigation in the context of pharmacogenetics and pharmacokinetics.

Supplementary table B represents all top 10 genome-wide variants observed at T2 with suggested p value significance $\leq 1.0\times10^{-05}$. Variant rs9865956 was observed as the SNP with the strongest association with high median drug levels observed in our study for rs9865956 may potentially indicate that patients with a A/G genotype are slow metabolizers and African populations may have a higher number of individuals with a similar phenotype with regards to drug metabolism compared to Asian and European populations with lower frequencies of this heterozygous genotype as per ALFA database. Patients regarded as slow metabolizers may require a lower dosage of anti-TB drugs due to slower clearance of the drugs compared to

Asian and European populations with lower frequencies of this genotype A/G for rs9865956. Additional studies need to focus more on genome-wide significant variants in association with drug pharmacokinetics as there are several markers with differential roles across various populations.

Observations of ABC & SLC family variants in association with moxifloxacin drug levels at both time points.

ABC gene family variants reported in Table 1 and 3 summarize the top 10 SNPs at both time points with p values ≤ 0.05 however none of the reported SNPs were of suggested significance. Genotype G/G of rs425215 in this study was associated with higher median drug levels (p= 0.0007) compared to C/C genotypes with an odds ratio of 13.75 (2.54-74.3) and is present at higher frequencies in African populations (0.45) compared to Europeans (0.39) and Asian populations (0.11). Interestingly a microarray based study that validated drug metabolizing enzymes and transporters (DMET) reported genotype G/G of rs425215 in *ABCG1* to be associated with high irinotecan-associated gastrointestinal (GI) toxicity along with other *ABCC5* variants in colorectal cancer patients (14). These findings modestly suggest that these known variants clearly play differential roles in various populations across the globe in addition to unreported variants that are yet to be understood. Evidently a study by *Fan et al*., 2024 reported that following whole-genome sequencing (WGS) from African individuals, their study analysis identified millions of unreported variants across the entire human genome that are predicted to be functionally important (15). Our analysis in this study focused only on the top 10 variants for both *ABC* and *SLC* family genes, however the remainder of the unreported variants may overlap with those identified in the study *by Fan et al*., 2024. From data as such from various studies it is supportive of the fact that integrating data from WGS with GWAS data is essential for comprehensively elucidating the specific and functional roles of drug transporter gene variants to accurately characterize their population-specific purpose and associations across global populations.

From the solute carrier gene family GWAS analysis, a variant observed that stood out at timepoint 1 analysis showed suggested significance from rs4414423 in *SLC24A4 (*p= 9.50e-^{05*}), OR= 0.18, 95% CI (0.04-0.71) with A/A genotypes having higher median drug levels compared to A/G genotypes. While this p-value does not meet the stringent criteria for genome-wide or suggestive significance under multiple-testing correction, it falls within the range typically considered nominally significant**,** indicating a potentially meaningful association with moxifloxacin pharmacokinetics. This suggests that rs4414423 may influence interindividual variability in drug exposure and warrants prioritization for further investigation. Even though the remaining 9 SNPs identified in this analysis achieve corrected statistical significance, it is important to acknowledge that the current study's sample size may limit statistical power. Nevertheless these variants and particularly the nominal top hit rs4414423 may demonstrate stronger or more definitive associations in future studies with larger, better-powered cohorts, where subtle genetic effects on moxifloxacin disposition could be more readily detected. This *SLC* variant rs4414423 to date has not been reported to have any clinical associations however may have potential to be investigated as a genome-wide SNP of significance in a larger sample sized study.

The remaining 9 of top 10 *SLC* variants (rs10504458, rs71529388, rs7611820, rs74010046, rs6472480, rs76392376, rs66862591, rs73432173 & rs78967500) at T1 however have not been reported with clinical associations to date and is worth exploring further in larger studies despite not being suggested as significant variants in this study. Timepoint 2 analysis revealed another unreported solute carrier intronic variant rs75457665 of *SLC44A3-AS1* showing suggested significance (p= 7.21e-^{05*}) with higher moxifloxacin drug levels in individuals who genotyped as T/T compared to T/C genotypes. Another variant observed within this study is rs2701688 of *SLC26A4* that showed to have high moxifloxacin drug levels in C/T genotypes compared to T/T genotyped individuals (p= 0.000159), OR= 26.86 95% CI (1.33-542.72). This intronic variant has been cited once in a study that investigated various drug transporter variants and their associated effect on aldosterone response to antihypertensive drug therapy (16). Another interesting variant identified in our study was rs1060896 found in gene *SLC28A2-AS1* which has also been reported as one of the SNPs containing risk alleles for the onset and development of multiple myeloma (MM) in absorption, distribution, metabolism, and excretion (ADME) genes (17). Between both time points in this study there were eight variants that had p values ≤ 0.05 presented in Table 5 with varying p-values however the majority of the variants

showed similar odds ratio values at both T1 and T2 indicating a nearly constant effect size in this study. While it is assumed that patient adherence to moxifloxacin dosing remained relatively constant between both time points, other biological and environmental factors such as epigenetic modifications, variable gene expression, and unmeasured covariates may contribute to the observed differences in pharmacokinetic responses between timepoints. Consequently, these findings highlight SNPs that are consistently associated across both timepoints, but future replication in larger, well-powered cohorts is necessary to confirm their functional relevance. The findings of this data modestly captures the variations of concern as a piloted subset of candidate variants for future analyses and follow up in larger cohort studies.

As observed in this study, it is evident that various drug transporter gene variants are observed with none to minimal clinical significance reported with allele frequencies that differ across mainly European, Asian and African populations. Additionally it has been observed that genome-wide variants play differential roles and are linked to phenotypes that are uncommon between populations. Even though several of the variants identified in this study are understudied intronic variants, they are potentially often missed within whole exome sequencing (WES) studies that exclude intron targets and should not be overlooked in future research due to their potential ability to affect RNA splicing leading to potential loss of gene function in some cases (18). From an alternate perspective, Appendix A highlights the importance of Human Leukocyte Antigen (HLA) and how its diversity may be associated with HIV/TB co-infection and the recurrence of TB in South African populations. Lastly we cannot rule out the possibility of simultaneous multi-omic changes (epigenetic, proteomic & metabolic) that may occur and differ between individuals whilst we only focus and analyze from a genetic perspective.

Conclusion

Genome-wide studies are pivotal in defining the significance and consequences of the extensive genetic diversity within African populations in the context of precision medicine. In conclusion, the findings from this research study provide compelling evidence suggesting that genome-wide variants require further investigation within African cohorts, as underexplored genetic variants may play distinct and differential roles across diverse populations. The results observed in this study shed light on the critical gaps in current genomic research, particularly

regarding the drug and metabolizing transporters in the *ABC* and *SLC* gene families, which remain severely understudied and underreported to date. Comprehensive exploration of these genetic variants is essential to advance our understanding of pharmacogenetic and pharmacodynamic variability in African populations. Given the high level of genetic diversity that has been stressed across African ancestries, prioritizing such research is vital to improving the accuracy, equity, and efficacy of precision medicine initiatives on a global scale.

Study limitations

Genome-wide significance ($p \leq 1.0E^{-08}$) would have been easier to obtain in larger study sample numbers for the drug PK analysis aspect of the study which would enable easier comparison with stringent Bonferroni corrected P values for significant differences at a genome perspective. Quantile-quantile plots represented in supplementary figures: 1 and 2 confirm the lack in statisitical power of the GWAS findings due to a limited sample size of n=58. Nevertheless the suggested or nominal significant SNPs with P values $\leq 1.0E^{-05}$ identified that we did not investigate further in this study provide context and rationale for larger studies to focus on *ABC*, *SLC* and other genome-wide variants in African populations with greater detail. Another limitation to consider is that it is possible that there is a level of variance in treatment adherence among study participants and time of sample which to an extend is beyond the control of this research study.

Funding statement

Research performed in this article was generously supported by Grants, Innovation and Product Development unit of the South African Medical Research Council (SAMRC) with funds received from both Novartis and GSK Global Health for Project Africa GRADIENT (Grant No. GSKNVS2/202101/005).

References

1. Mroziewicz M, Tyndale RF. Pharmacogenetics: a tool for identifying genetic factors in drug dependence and response to treatment. Addict Sci Clin Pract. 2010;5(2):17-29.
2. Gomez F, Hirbo J, Tishkoff SA. Genetic variation and adaptation in Africa: implications for human evolution and disease. Cold Spring Harb Perspect Biol. 2014;6(7):a008524.
3. Wonkam A, Adeyemo A. Leveraging our common African origins to understand human evolution and health. Cell Genom. 2023;3(3):100278.
4. Bruckmueller H, Cascorbi I. ABCB1, ABCG2, ABCC1, ABCC2, and ABCC3 drug transporter polymorphisms and their impact on drug bioavailability: what is our current understanding? Expert Opin Drug Metab Toxicol. 2021;17(4):369-96.
5. Glavinas H, Krajcsi P, Cserepes J, Sarkadi B. The role of ABC transporters in drug resistance, metabolism and toxicity. Curr Drug Deliv. 2004;1(1):27-42.
6. Hediger MA, Romero MF, Peng JB, Rolfs A, Takanaga H, Bruford EA. The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteinsIntroduction. Pflugers Arch. 2004;447(5):465-8.
7. Parvez MM, Kaisar N, Shin HJ, Lee YJ, Shin JG. Comprehensive Substrate Characterization of 22 Antituberculosis Drugs for Multiple Solute Carrier (SLC) Uptake Transporters In Vitro. Antimicrob Agents Chemother. 2018;62(9).
8. Zhang S, Xiao K, He X, Hao M, Ma Y, Tian Y, et al. A novel efficient algorithm for common variants genotyping from low-coverage sequencing data. Computational and Structural Biotechnology Reports. 2025;2:100046.
9. Alwi ZB. The Use of SNPs in Pharmacogenomics Studies. Malays J Med Sci. 2005;12(2):4-12.
10. Tiberi S, Munoz-Torrico M, Duarte R, Dalcolmo M, D'Ambrosio L, Migliori GB. New drugs and perspectives for new anti-tuberculosis regimens. Pulmonology. 2018;24(2):86-98.
11. da Rocha JEB, Othman H, Botha G, Cottino L, Twesigomwe D, Ahmed S, et al. The Extent and Impact of Variation in ADME Genes in Sub-Saharan African Populations. Front Pharmacol. 2021;12:634016.
12. Berry KM, Rodriguez CA, Berhanu RH, Ismail N, Mvusi L, Long L, et al. Treatment outcomes among children, adolescents, and adults on treatment for tuberculosis in two metropolitan municipalities in Gauteng Province, South Africa. BMC Public Health. 2019;19(1):973.
13. Omotoso OE, Teibo JO, Atiba FA, Oladimeji T, Adebesin AO, Babalghith AO. Bridging the genomic data gap in Africa: implications for global disease burdens. Global Health. 2022;18(1):103.
14. Di Martino MT, Arbitrio M, Leone E, Guzzi PH, Rotundo MS, Ciliberto D, et al. Single nucleotide polymorphisms of ABCC5 and ABCG1 transporter genes correlate to irinotecan-associated gastrointestinal toxicity in colorectal cancer patients: a DMET microarray profiling study. Cancer Biol Ther. 2011;12(9):780-7.
15. Fan S, Spence JP, Feng Y, Hansen MEB, Terhorst J, Beltrame MH, et al. Whole-genome sequencing reveals a complex African population demographic history and signatures of local adaptation. Cell. 2023;186(5):923-39.e14.

16. Barbalic M, Schwartz GL, Chapman AB, Turner ST, Boerwinkle E. Kininogen gene (KNG) variation has a consistent effect on aldosterone response to antihypertensive drug therapy: the GERA study. Physiol Genomics. 2009;39(1):56-60.
17. Scionti F, Agapito G, Caracciolo D, Riillo C, Grillone K, Cannataro M, et al. Risk Alleles for Multiple Myeloma Susceptibility in ADME Genes. Cells. 2022;11(2).
18. Hertzog A, Selvanathan A, Farnsworth E, Tchan M, Adams L, Lewis K, et al. Intronic variants in inborn errors of metabolism: Beyond the exome. Front Genet. 2022;13:1031495.

Supplementary Table A: Top 10 genome-wide significant SNPs associated with moxifloxacin AUC drug levels and T1.

Chr	Gene Name	POS	Ref/Alt	SNP	Group 1	Group 2	Group 1 (n)	Group 2 (n)	Group 1 Median	Group 2 Median	P Value	OR, 95% CI
8	-	5819814	G/A	rs6982690	G/A	A/A	25	9	6.21	16.88	1.715e-06	0.04 (0.00-0.79)
5	-	26365484	G/A	rs2319809	G/G	G/A	46	8	7.20	17.19	1.970e-06	0.04 (0.00-0.76)
10	ADARB2	1708820	T/A	rs10047298	T/T	T/A	43	8	9.91	3.95	6.592e-06	23.43 (1.27-432.02)
6	HIVEP1	12093098	T/C	rs6910367	C/C	T/C	18	29	14.75	5.76	6.739e-06	44.62 (5.06-392.80)
5	SLIT3	169258089	C/T	rs61174378	C/C	C/T	41	10	6.60	15.92	9.405e-06	0.06 (0.00-0.55)
8	HAS2-AS1	121853401	G/A	rs73328832	G/A	G/G	20	33	14.33	6.21	1.388e-05	13.03 (3.10-54.70)
17	RPH3AL	301039	A/G	rs72806046	A/G	A/A	19	33	5.77	12.60	1.434e-05	0.03 (0.00-0.19)
4	C4orf51	145692921	C/T	rs2132778	C/T	C/C	26	6	5.94	15.90	1.544e-05	0.04 (0.00-0.92)
16	CDH13	83747110	T/C	rs56338482	T/T	T/C	49	5	7.74	17.84	1.833e-05	0.074 (0.00-1.41)
7	-	155582418	C/T	rs145236962	C/C	C/T	47	7	7.58	16.88	2.079e-05	0.04 (0.00-0.92)

Supplementary Table B: Top 10 significant SNPs associated with moxifloxacin AUC drug levels and T2.

Chr	Gene Name	POS	Ref/ Alt	SNP	Group 1	Group 2	Group 1 (n)	Group 2 (n)	Group 1 Median	Group 2 Median	P Value	OR, 95% CI
3	-	3277703	A/G	rs9865956	A/G	G/G	25	10	11.707	5.6394	2.11e-06	63 (3.22-1231.34)
2	-	38291559	GA/G	rs199538426	GA/G	GA/GA	5	44	5.0955	10.255	4.19e-06	0.07 (0.00-1.45)
5	-	1.81E+08	A/G	rs615560	A/G	G/G	24	9	13.19	5.8994	9.95e-06	19.42 (2.03-185.73)
14	OXA1L-DT	22749545	C/T	rs8014504	C/T	C/C	11	36	5.3794	10.753	1.10e-05	0.06 (0.00-0.55)
8	-	51289810	G/A	rs112384918	G/A	G/G	7	41	5.3794	10.326	1.17e-05	0.05 (0.00-0.98)
10	GRID1	85904089	T/A	rs28376465	T/A	T/T	8	39	5.2838	10.334	1.33e-05	0.04 (0.00-0.77)
1	ATP1B1	1.69E+08	G/A	rs114970959	G/G	G/A	40	9	8.2689	15.628	1.46e-05	0.03 (0.00-0.59)
1	KAZN	14899559	T/C	rs2807555	T/C	T/T	18	8	7.9233	16.48	1.53e-05	0.02 (0.00-0.49)
18	-	833496	G/A	rs77849147	G/G	G/A	43	6	9.0905	18.053	1.78e-05	0.05 (0.00-1.05)
21	-	19796780	C/T	rs6517902	C/C	T/T	17	9	7.6346	17.332	1.92e-05	0.017 (0.00-0.37)

Supplementary Table C: Gene family *ABC* with respective variants with P values ≤ 0.01 in association with moxifloxacin drug PK AUC values at T1.

Chr	Gene Name	Pos	Ref/Alt	SNP	Group 1	Group 2	Group 1 (n)	Group 2 (n)	Group 1 Median	Group 2 Median	P Value	OR, 95% CI
7	ABCB5	20678540	T/C	rs7805940	T/T	C/C	38	3	6.71	18.93	0.0004	0.084 (0.00-1.75)
21	ABCG1	42286991	C/G	rs425215	G/G	C/C	18	15	14.17	5.76	0.0007	13.75 (2.54-74.3)
21	ABCG1	42286991	C/G	rs425215	C/G	G/G	21	18	6.60	14.17	0.0012	0.12 (0.02-0.56)
21	ABCG1	42304243	G/A	rs15661	G/A	G/G	23	7	6.60	15.08	0.0015	0.10 (0.01-1.04)
11	ABCC8	17423592	A/G	rs929234	G/G	A/A	14	9	10.34	5.65	0.002	29.3 (2.55-336)
7	ABCB5	20661377	G/A	rs73684681	G/G	G/A	32	18	5.99	12.58	0.0021	0.12 (0.03-0.49)
17	ABCA5	69253458	C/T	rs75169077	C/C	C/T	49	3	8.80	3.37	0.0028	7.89 (0.38-160)
7	ABCF2-H2BK	151225377	A/T	rs188409705	A/A	A/T	49	5	7.74	16.54	0.0056	0.22 (0.02-2.12)
7	ABCB5	20678540	T/C	rs7805940	T/T	T/C	38	13	6.71	12.60	0.0056	0.175 (0.04-0.74)
17	ABCA6	69129236	G/A	rs139091810	G/G	G/A	47	7	7.58	15.08	0.0071	0.049 (0.00-0.92)
9	ABCA1	104791243	A/G	rs10121901	G/G	A/A	30	5	9.47	4.46	0.0083	16.28 (0.82-321)
17	ABCA6	69103517	A/G	rs4141177	A/A	A/G	38	16	9.79	4.66	0.0093	4.6 (1.24-16.96)
21	ABCG1	42281867	A/G	rs492338	G/G	A/G	21	24	10.77	5.98	0.0102	4.85 (1.37-17.19)
7	ABCB5	20747650	C/T	rs6461517	C/T	C/C	25	14	12.57	6.40	0.011	5.31 (1.26-22.2)
17	ABCA6	69101071	T/C	rs1468514	C/C	T/T	12	16	5.76	10.83	0.0113	0.09 (0.014-0.57)
1	ABCA4	94110804	A/G	rs3789435	A/A	A/G	15	31	10.89	6.21	0.0118	3.16 (0.86-11.55)
4	ABCG2	88192345	T/C	rs13152371	T/T	C/C	21	9	5.76	13.46	0.0125	0.17 (0.02-1.06)
16	ABCC1	16120015	G/A	rs2238476	G/G	G/A	50	4	7.77	16.19	0.0126	0.09 (0.00-1.85)
7	ABCB5	20714159	T/C	rs73684821	T/T	C/C	34	5	8.71	4.41	0.0141	4.5 (0.45-44.5)
9	ABCA1	104884939	C/T	rs4149268	C/T	C/C	16	3	10.40	3.56	0.0144	8.86 (0.39-199)
1	ABCA4	93993879	A/G	rs12239339	A/A	G/G	21	3	9.03	3.63	0.0158	7.66 (0.35-166.6)
13	ABCC4	95179748	T/G	rs4148493	G/G	T/G	9	26	5.65	8.92	0.0163	0.24 (0.04-1.4)
17	ABCA10	69204660	A/G	rs9898868	G/G	A/G	35	18	9.67	4.66	0.0193	3.9 (1.13-13.38)
7	ABCA13	48582270	C/T	rs6583453	C/T	T/T	27	2	7.58	18.19	0.0197	0.13 (0.00-3.18)

Supplementary Table D: *SLC* Gene family variants with P values ≤ 0.01 in association with moxifloxacin drug PK AUC values at T1.

Chr	Gene Name	POS	Ref/ Alt	SNP	Group 1	Group 2	Group 1 (n)	Group 2 (n)	Group 1 Median	Group 2 Median	P Value	OR, 95% CI
14	SLC24A4	92423645	A/G	rs4414423	A/G	A/A	31	16	6.11868	15.189	9.50e-05*	0.18 (0.04-0.71)
8	SLCO5A1	69694927	A/G	rs10504458	A/A	A/G	32	17	11.6722	4.4892	0.0004	10.2 (2.39-43.93)
7	SLC29A4	5303497	C/T	rs71529388	C/C	C/T	48	5	9.3542	3.6338	0.00054	12.9 (0.67-247.3)
3	SLC7A14-AS1	170654068	G/A	rs7611820	G/A	G/G	19	11	9.91348	4.4567	0.00081	9.75 (1.59-59.6)
15	SLC12A6	34269767	T/C	rs74010046	T/T	T/C	48	6	9.3542	4.0255	0.00089	16.6 (0.88-311.8)
8	SLCO5A1	69713707	A/G	rs6472480	A/A	A/G	23	21	5.64536	9.9135	0.00092	0.13 (0.03-0.53)
12	SLC15A4	128793867	G/C	rs76392376	G/G	G/C	45	8	9.03361	3.8915	0.00097	8.75 (0.99-77.11)
14	SLC25A21	37118109	C/T	rs66862591	C/C	C/T	17	28	5.64536	10.415	0.00099	0.14 (0.03-0.57)
17	SLC26A11	80252305	G/A	rs73432173	G/A	A/A	21	3	9.67478	3.1767	0.00099	7.66 (0.35-166.6)
5	SLCO6A1	102386596	G/T	rs78967500	G/G	G/T	41	12	6.59816	14.748	0.00126	0.128 (0.02-0.66)
11	SLC35F2	107802697	C	rs12577389	C/T	C/C	13	40	4.46001	10.224	0.00142	0.12 (0.02-0.62)
18	SLC14A2	45564677	T	rs13382004	T/C	C/C	27	10	11.1306	4.0092	0.00159	18 (1.96-164.99)
17	SLC47A2	19710749	G	rs113690031	G/G	G/A	46	8	9.79413	4.6451	0.00174	23.97 (1.30-440.376)
1	SLC6A17	110181792	AGT	rs34512655	A/A	AGT/A	42	12	9.79413	5.6423	0.00192	7.35 (1.42-27.84)
15	SLC28A1	84904404	A	rs2290271	A/C	C/C	21	7	5.65758	13.756	0.00199	0.05 (0.00-0.54)
4	SLC2A9-AS1	10057994	G	rs714873	A/A	G/A	29	22	6.11868	13.176	0.00209	0.24 (0.07-0.79)
8	SLCO5A1	69672393	T	rs1138541	T/T	C/C	21	10	13.7557	5.1231	0.00223	8 (1.32-48.18)
6	SLC17A3	25873015	T	rs648433	T/T	T/C	50	4	8.91755	3.247	0.00225	10.53 (0.53-205.94)
6	SLC17A1	25820200	T	rs6913879	C/C	T/C	13	18	6.82375	14.752	0.00233	0.05 (0.00-0.33)
6	SLC22A16	110456759	A	rs6907567	A/G	A/A	30	17	6.35842	14.924	0.00249	0.20 (0.05-0.75)
6	SLC22A16	110456925	T	rs714368	T/C	T/T	30	17	6.35842	14.924	0.00249	0.20 (0.05-0.75)
17	SLC26A11	80252305	G	rs73432173	G/G	A/A	30	3	8.59106	3.1767	0.00257	7.96 (0.37-167.53)
5	SLCO6A1	102420358	G	rs6596486	G/G	A/A	29	4	6.11868	17.907	0.00259	0.05 (0.00-1.22)
5	SLC12A7	1054716	T	rs4975569	T/C	C/C	18	35	13.917	6.5982	0.00308	5.92 (1.60-21.85)
14	SLC7A8	23171585	A	rs2268873	G/G	A/G	8	20	5.23464	11.851	0.00317	0.04 (0.00-0.48)
8	SLCO5A1	69713707	A	rs6472480	A/A	G/G	23	10	5.64536	13.769	0.00327	0.06 (0.01-0.43)
15	SLCO3A1	91931055	T	rs6496873	G/G	T/T	23	6	8.1676	14.776	0.00336	0.07 (0.00-1.40)
7	SLC29A4	5279756	G	rs10279978	G/G	G/A	9	8	13.4603	5.0527	0.0037	24.5 (1,78-336.24)
15	SLCO3A1	91992789	G	rs114014316	G/G	G/A	50	4	7.76689	17.167	0.00385	0.09 (0.00-1.85)
3	SLC6A6	14458757	C	rs17039625	C/T	C/C	28	16	10.8341	5.7073	0.00435	4 (1.02-15.53)
17	SLC39A11	72690302	C	rs4076588	A/A	C/A	17	25	5.65758	10.774	0.00436	0.08 (0.01-0.38)
3	SLC6A6	14424348	C	rs77301812	C/T	C/C	15	38	14.9238	6.711	0.00437	4.21 (1.13-15.72)
6	SLC17A1	25825891	C	rs6456703	C/T	C/C	19	25	14.5801	6.1187	0.00438	4.77 (1.22-18.5)
1	SLC44A5	75410352	C	rs4998167	C/C	C/T	39	15	9.93593	5.6576	0.00457	6.4 (1.54-26.4)
8	SLCO5A1	69793915	G	rs147059714	G/G	G/A	42	11	6.4039	13.756	0.00475	0.28 (0.06-1.21)
2	SLC16A14	230055802	G	rs4973253	G/A	A/A	26	19	10.8341	5.757	0.00481	2.74 (0.80-9.30)
6	SLC17A1	25825891	C	rs6456703	C/T	T/T	19	10	14.5801	7.1995	0.00482	33.75 (3.24-351.07)
2	SLC4A3	219639240	G	rs78104328	G/G	G/A	45	8	9.67478	4.6597	0.00486	9.57 (1.08-84.50)
2	SLC5A6	27202638	C	rs145748011	C/C	C/A	47	7	7.7399	15.077	0.00491	0.13 (0.02-1.21)
8	SLCO5A1	69730611	T	rs61342871	T/T	T/C	38	14	6.16416	12.325	0.00495	0.26 (0.07-0.99)
3	SLC12A8	125110376	C	rs72963853	C/C	T/T	38	2	6.71096	18.389	0.00513	0.13 (0.01-2.94)
1	SLC44A3	94892540	A	rs859063	A/A	A/G	31	17	9.93593	5.6454	0.00533	5.14 (1.36-19.52)
10	SLC16A12	89471239	C	rs11203141	C/C	C/T	46	8	7.65759	16.288	0.00542	0.10 (0.01-0.97)
11	SLC1A2-AS1	35273849	A	rs73439110	A/A	A/G	48	5	7.76689	16.036	0.00545	0.07 (0.00-1.36)
11	SLC29A2	66366172	G	rs8187661	G/G	G/A	42	11	9.79413	4.2705	0.00553	6 (1.15-31.23)
6	SLC22A23	3440080	G	rs9328172	G/A	G/G	25	19	13.7557	7.7939	0.00557	1.66 (0.50-5.56)
3	SLC4A7	27391865	A	rs17019727	A/A	A/G	49	5	7.7399	15.808	0.00559	0.07 (0.00-1.42)
5	SLC6A7	150213844	C	rs115236806	C/T	C/C	5	49	3.82331	9.0336	0.00559	0.07 (0.00 - 1.42)
8	SLCO5A1	69709973	T	rs2933045	T/T	C/C	2	25	18.7559	8.1676	0.0057	5.4 (0.24-123.81)
14	SLC25A21	36776774	A	rs11846596	A/G	A/A	30	7	6.4039	17.844	0.00579	0.23 (0.04-1.40)
1	SLC44A3	94892540	A	rs859063	G/G	A/G	6	17	11.1309	5.6454	0.0059	6.5 (0.85-49.69)
1	SLC6A17	110155174	A	rs76482195	A/C	A/A	10	42	13.608	6.711	0.00594	5.88 (1.11-31.17)
3	SLC4A7	27395380	T	rs73149794	C/C	T/T	2	24	18.211	8.0873	0.00615	5 (0.22-115.06)
8	SLC39A14	22391152	C	rs57210246	C/C	C/T	44	10	9.9247	5.6981	0.00638	33 (1.86-599.57)
9	SLC1A1	4542890	A	rs928209	A/A	A/G	42	10	9.47354	4.0453	0.00643	5.33 (1.01-28.21)

10	SLC29A3	71336036	G	rs16928705	G/A	G/G	32	16	6.5167	13.324	0.00643	0.13 (0.03-0.59)
16	SLC38A8	84011039	A	rs113857440	A/A	A/G	47	7	7.7399	16.036	0.00652	0.13 (0.01-1.21)
5	SLC6A7	150222404	A	rs112896997	A/A	A/G	26	19	10.9522	6.5982	0.00653	4.87 (1.36-17.47)
8	SLCO5A1	69683309	A	rs77341699	A/A	A/G	45	9	9.67478	4.4892	0.00655	10.94 (1.26-95.06)
1	SLC30A10	219903146	G	rs113489650	G/G	G/A	45	9	6.82375	13.46	0.00655	0.03 (0.00-0.65)
6	SLC22A23	3438831	C	rs13213207	C/C	C/T	49	4	7.79388	17.167	0.00665	0.09 (0.00-1.93)
1	SLC30A10	219947627	C	rs2789801	T/T	C/C	30	4	7.98074	3.5033	0.00668	7.90 (0.39-159.78)
22	SLC5A4-AS1	32244362	G	rs2294211	G/A	G/G	23	21	5.76756	10.894	0.00702	0.21 (0.06-0.78)
11	SLC22A8	63010425	T	rs55880001	T/T	T/C	35	15	6.20964	13.756	0.00712	0.13 (0.03-0.55)
6	SLC17A1	25820200	T	rs6913879	T/T	T/C	23	18	6.59816	14.752	0.00727	0.26 (0.07-1.04)
14	SLC25A21	37165634	G	rs10134334	A/A	G/A	6	25	5.29883	11.131	0.0073	0.09 (0.00 -0.94)
1	SLC1A7	53139045	C	rs58722687	C/C	C/T	37	14	6.20964	13.015	0.00751	0.166 (0.03-0.70)
1	SLC9A1	27148998	T	rs11247612	T/T	C/C	31	2	7.79388	18.826	0.00758	0.165 (0.00-3.73)
5	SLC1A3	36651262	T	rs56357960	T/T	T/C	23	19	14.0784	6.8238	0.00759	6.4 (1.65-24.77)
20	SLCO4A1	62671558	CTT	rs3060740	CTT/CT	CTT/C	7	26	4.4567	9.9247	0.00768	0.34 (0.05-2.10)
12	SLC2A13	39928044	T	rs6581439	T/C	C/C	12	39	15.1886	7.5753	0.00782	3.88 (0.91-16.58)
3	SLC9A9	143304755	C	rs7428843	T/T	C/T	26	21	5.64838	10.894	0.00787	0.177 (0.05-0.63)
17	SLC39A11	72987719	C	rs61172383	C/C	C/T	31	20	10.7738	5.0642	0.00808	4.24 (1.27-14.18)
3	SLC7A14-AS1	170670518	A	rs16847978	A/A	A/G	26	23	5.9886	12.597	0.00809	0.16 (0.05-0.56)
19	SLC8A2	47443142	A	rs57112810	A/A	A/G	35	18	6.11868	10.355	0.00838	0.37 (0.11-1.23)
19	SLC8A2	47454274	C	rs73568302	C/C	C/T	35	18	6.11868	10.355	0.00838	0.37 (0.11-1.23)
10	SLC35G1	93921322	G	rs1223470	G/A	A/A	28	12	9.48477	5.0339	0.00844	4.63 (1.02-21.00)
5	SLC25A48	135511781	G	rs10213718	G/G	G/A	42	11	7.19951	14.58	0.00857	0.30 (0.07-1.33)
2	SLC8A1	40409100	A	rs6755144	A/G	G/G	29	6	9.03361	4.0109	0.00905	6.15 (0.64-59.47)
17	SLC6A4	30216371	C	rs140700	C/C	C/T	51	3	7.79388	17.844	0.00917	0.12 (0.00-2.59)
6	SLC22A23	3413063	T	rs6926754	T/C	C/C	31	4	10.7738	3.8915	0.00921	14.04 (0.69-283.94)
17	SLC39A11	72688232	A	rs55688472	A/G	A/A	25	26	5.76756	10.834	0.00927	0.35 (0.11-1.10)
11	SLC29A2	66365865	A	rs2279861	A/A	A/G	32	19	10.3436	5.6454	0.00948	2.20 (0.69-7.06)
17	SLC47A2	19703667	G	rs111301691	G/G	G/A	50	4	8.91755	3.8183	0.00952	10.53 (0.54-205.95)
14	SLC24A4	92412450	G	rs4904903	G/G	G/A	48	5	9.3542	3.8233	0.00975	14.06 (0.74-268.70)
3	SLC6A6	14451736	C	rs116344061	C/C	C/T	49	5	7.7399	15.077	0.00986	0.07 (0.00-1.42)
1	SLC30A10	219947627	C	rs2789801	C/T	C/C	20	4	9.9247	3.5033	0.00998	16.2 (0.76-343.82)
4	SLC4A4	71322934	G	rs146763256	G/G	G/A	47	7	7.7399	16.879	0.01016	0.13 (0.01-1.21)
7	SLC26A3	107791267	T	rs41280242	T/T	C/C	33	5	9.91348	4.2705	0.01037	5.42 (0.55-54.01)
1	SLC35F3	233994762	G	rs34117807	G/G	A/A	30	4	9.9247	4.1959	0.01039	4.5 (0.42-48.53)
15	SLCO3A1	92143026	A	rs74028820	A/A	G/G	25	6	10.7738	5.0673	0.01041	26.76 (1.34-532.73)
14	SLC8A3	70054853	A	rs2839867	C/C	A/A	22	10	7.68458	14.938	0.01062	0.17 (0.03-1.01)
2	SLC7A15P	20394353	C	rs11096662	C/T	C/C	11	41	5.76756	9.9135	0.01062	0.06 (0.01-0.55)
14	SLC35F4	57885537	C	rs73298617	C/T	C/C	13	39	4.83011	9.6748	0.01065	0.23 (0.06-0.98)
20	SLC24A3	19290861	T	rs58483859	T/T	T/C	37	16	7.57527	14.34	0.01105	0.22 (0.06-0.84)
14	SLC35F4	57848562	A	rs74565776	A/A	A/G	45	9	9.67478	4.4567	0.01121	4.37 (0.81-23.43)
17	SLC39A11	72719238	A	rs2567505	A/G	A/A	20	8	12.117	5.7012	0.01137	7 (1.08-45.16)
14	SLC25A21	37021604	G	rs17105971	G/G	G/A	44	9	6.71096	12.597	0.01149	0.21 (0.04-1.17)
5	SLC12A2-DT	127959270	C	rs148631990	C/C	C/T	49	4	7.7399	16.826	0.01149	0.09 (0.00-1.78)
14	SLC7A8	23126512	A	rs2236135	A/G	G/G	22	8	9.80536	5.0673	0.01158	5.25 (0.85-32.43)
6	SLC17A5	73615534	T	rs628038	T/T	T/C	20	21	5.76229	11.131	0.01163	0.17 (0.04-0.66)
6	SLC17A3	25841975	C	rs4712976	C/T	C/C	16	37	5.29883	9.9135	0.01172	0.45 (0.14-1.52)
14	SLC25A29	100303601	G	rs2281892	G/G	G/A	43	9	9.67478	4.46	0.01179	11.11 (1.27-96.87)
13	SLC15A1	98748029	C	rs883962	T/T	C/T	36	17	6.84697	12.597	0.01187	0.43 (0.13-1.44)
3	SLC9A9	143266258	G	rs1046706	G/T	G/G	32	10	9.02771	5.051	0.01197	4.53 (0.83-24.76)
15	SLC51B	65052881	C	rs74981303	C/C	C/T	47	7	7.7399	15.077	0.01204	0.13 (0.02-1.21)
14	SLC25A21	37075306	C	rs74604833	C/T	C/C	13	38	5.65758	8.9175	0.01204	0.21 (0.05-0.92)
2	SLC23A3	219164178	A	rs6753739	G/G	A/A	20	11	5.94312	14.078	0.01229	0.20 (0.04-1.01)
14	SLC25A21	36808958	C	rs75412800	C/C	C/T	47	6	9.03361	4.411	0.01233	5.68 (0.61-52.43)
15	SLCO3A1	91931055	T	rs6496873	T/G	T/T	25	6	6.82375	14.776	0.01233	0.05 (0.00-1.03)
5	SLC9A3	502506	C	rs72702730	C/C	C/T	30	20	5.9886	10.834	0.01271	0.44 (0.14-1.41)
1	SLC2A7	9006604	T	rs57628014	T/C	T/T	16	30	10.3548	6.1642	0.01289	2.87 (0.82-10.10)
5	SLC6A7	150222404	A	rs112896997	G/G	A/A	9	26	4.46001	10.952	0.01295	0.22 (0.04-0.12)

14	SLC7A7	22788980	T	rs2065134	T/T	T/G	47	7	7.7399	15.808	0.01308	0.13 (0.02-1.21)
12	SLC16A7	59742017	T	rs10783999	T/T	C/C	2	41	17.5586	8.3806	0.01329	4.76 (0.22-105.42)
3	SLC9A9	143614188	G	rs16853938	G/G	G/T	20	26	5.71257	9.3542	0.01331	0.26 (0.08-0.93)
18	SLC14A1	45690617	T	rs35191019	T/T	C/C	31	5	7.79388	15.077	0.0138	0.06 (0.00-1.30)
15	SLC12A1	48229719	T	rs3784614	C/C	T/C	39	14	6.20964	11.013	0.01383	0.27 (0.07-1.05)
5	SLC36A1	151491042	G	rs165370	G/G	G/A	32	18	5.94312	9.9247	0.01384	0.3 (0.09-1.01)
3	SLC25A26	66180117	G	rs36149310	A/A	G/A	33	19	10.7738	5.6454	0.01401	2.05 (0.65-6.54)
9	SLC25A51	37889861	C	rs2025439	C/C	C/T	47	7	9.03361	4.8301	0.0142	7.42 (0.83-66.63)
6	SLC22A23	3413063	T	rs6926754	T/T	C/C	18	4	7.65759	3.8915	0.01449	5.86 (0.27-125.58)
3	SLC9A9	143355772	C	rs16853475	C/C	C/T	49	5	9.03361	4.4892	0.01488	13.44 (0.70-256.42)
18	SLC14A1	45690617	T	rs35191019	T/C	C/C	17	5	7.57527	15.077	0.01496	0.08 (0.00-1.70)
15	SLCO3A1	92080405	T	rs12101738	C/C	T/C	4	13	14.9378	8.3806	0.01513	10.38 (0.47-231.64)
1	SLC2A1	42935828	C	rs841852	C/C	T/T	31	2	7.7399	17.907	0.01515	0.16 (0.00-3.73)
3	SLC41A3	126025230	T	rs879385	T/T	T/C	13	28	5.65758	9.3575	0.01519	0.22 (0.05-1.00)
14	SLC25A29	100305146	T	rs12432201	T/T	T/C	37	17	6.59816	11.131	0.01523	0.28 (0.08-0.97)
2	SLC9A2	102650834	T	rs59608009	T/T	T/C	46	6	8.91755	3.4673	0.01523	5.95 (0.64-55.03)
12	SLC38A4-AS1	46779280	G	rs180425	A/A	G/A	35	17	9.93593	4.4892	0.01535	4.06 (1.17-14.15)
6	SLC2A12	134032171	G	rs73774198	G/G	G/A	44	9	7.28183	15.077	0.0154	0.21 (0.04-1.17)
2	SLC8A1	40591489	C	rs2193560	C/T	T/T	22	25	10.8341	6.1187	0.01545	4.5 (1.31-15.42)
8	SLCO5A1	69672393	T	rs1138541	T/C	T/T	23	21	7.7399	13.756	0.01547	0.45 (0.14-1.56)
8	SLC30A8	117041250	A	rs1394874	G/G	A/G	32	16	7.19951	13.769	0.01548	0.2 (0.05-0.76)
5	SLC36A1	151456584	A	rs11954054	A/G	A/A	17	33	5.75702	10.774	0.01548	0.35 (0.11-1.20)
2	SLC8A1-AS1	40099002	T	rs6726508	C/C	T/T	26	5	10.9522	4.8301	0.01584	20.26 (1.01-107.37)
3	SLC7A14-AS1	170670518	A	rs16847978	A/A	G/G	26	5	5.9886	14.58	0.01584	0.09 (0.01-0.97)
3	SLC6A1	11021488	A	rs4038396	A/A	G/G	17	8	5.75702	12.295	0.01586	0.07 (0.01-0.79)
11	SLC5A12	26703231	C	rs12575103	C/C	C/T	50	3	7.98074	16.539	0.01614	0.12 (0.01-2.49)
3	SLC41A3	126024828	T	rs905455	C/C	T/C	47	6	9.03361	4.3652	0.01618	16.02 (0.85-300.74)
1	SLC44A3-AS1	94741234	G	rs7539060	G/G	A/A	37	2	8.38065	17.362	0.01619	0.21 (0.01-4.69)
5	SLC27A6	128659361	C	rs4836389	C/C	C/A	12	31	4.79017	9.0336	0.01634	0.27 (0.06-1.21)
15	SLCO3A1	92020187	T	rs7178228	T/T	T/C	48	5	7.76689	16.539	0.0164	0.07 (0.00-1.47)
4	SLC39A8	102279965	A	rs28505163	G/G	A/A	4	31	14.5549	7.7939	0.01646	10.86 (0.53-218.95)
1	SLC1A7	53140212	T	rs1288367	C/C	T/T	21	9	9.93593	4.4567	0.01652	7 (1.14-42.97)
14	SLC25A21	37090622	AG	rs149255736	AG/AG	AG/A	33	20	6.11868	11.123	0.01652	0.49 (0.15-1.52)
1	SLC44A5	75437642	G	rs1891902	G/A	G/G	31	17	10.8944	6.1187	0.01662	5.04 (1.39-18.2)
1	SLC35F3	233937764	A	rs140376234	A/A	A/G	46	8	7.65759	15.922	0.01667	0.10 (0.01-0.96)
10	SLC29A3	71336490	G	rs7914132	G/G	G/A	48	6	7.76689	14.738	0.01713	0.16 (0.01-1.55)
2	SLC35F5	113735947	C	rs10199704	T/T	C/C	12	16	9.80536	5.9833	0.01724	6.6 (1.23-35.43)
6	SLC22A2	160219447	G	rs73600483	G/G	G/A	51	3	8.80148	3.3173	0.01736	7.85 (0.39-159.86)
20	SLC52A3	765779	TCTGC	rs11467076	CC/GT	CTGCC	23	21	12.5971	6.1187	0.0177	7.08 (1.87-26.71)
2	SLC8A1	40530651	A	rs13017846	A/A	A/G	42	12	9.9247	5.882	0.01782	17.87 (2.10-151.89)
2	SLC8A1	40533825	T	rs17026156	T/T	T/C	42	12	9.9247	5.882	0.01782	17.87 (2.10-151.89)
17	SLC39A11	72776752	C	rs16977385	C/C	A/A	35	2	8.38065	3.2265	0.01802	5.28 (0.23-118.03)
11	SLC17A6	22374024	G	rs12280547	G/A	A/A	26	7	8.29768	14.58	0.01822	0.4 (0.06-2.44)
3	SLC6A6	14405037	C	rs1156568	T/T	C/T	44	10	7.65759	15.361	0.01827	0.19 (0.03-0.99)
15	SLC28A2-AS1	45263965	T	rs8023604	T/T	C/C	15	6	5.76756	16.458	0.01835	0.05 (0.00-0.604)
15	SLC28A1	84904404	A	rs2290271	A/A	A/C	26	21	10.404	5.6576	0.01922	5.12 (1.42-18.37)
5	SLC6A3	1409905	T	rs8179031	T/C	T/T	18	36	5.70646	10.344	0.01925	0.24 (0.07-0.84)
8	SLCO5A1	69692837	G	rs2380563	G/G	G/T	34	15	5.94312	9.9135	0.01967	0.27 (0.07-0.98)
8	SLCO5A1	69709973	T	rs2933045	T/C	T/T	27	2	7.79388	18.756	0.0197	0.18 (0.00-4.24)
6	SLC22A23	3428874	G	rs76762931	G/G	G/A	49	5	9.03361	4.2705	0.01984	13.44 (0.70-256.94)
11	SLC1A2	35339728	C	rs149756507	C/C	C/T	50	4	8.91755	3.6925	0.01988	10.53 (0.53-205.94)
8	SLC26A7	91273494	A	rs7812726	G/G	A/A	7	21	12.5971	6.5982	0.0199	15 (1.47-152.49)

Supplementary Table E: *ABC* Gene family variants with P values ≤ 0.01 in association with moxifloxacin drug PK AUC values at T2.

Chr	Gene Name	POS	Ref/ Alt	SNP	Group 1	Group 2	Group 1 (n)	Group 2 (n)	Group 1 Median	Group 2 Median	P Value	OR, 95% CI
9	ABCA1	104818199	C/T	rs58866705	C/T	C/C	8	40	5.52	10.26	0.0002	0.12 (0.01-1.04)
21	ABCG1	42304243	G/A	rs15661	G/A	G/G	21	6	9.63	15.81	0.0015	0.07 (0.00-1.40)
21	ABCG1	42304243	G/A	rs15661	A/A	G/G	22	6	8.07	15.81	0.0021	0.05 (0.00-0.90)
1	ABCA4	94086591	A/G	rs138785943	A/A	A/G	45	4	10.18	5.37	0.0023	10.26 (0.52-201.64)
1	ABCD3	94502789	C/T	rs12130330	C/C	C/T	31	15	8.02	13.07	0.0025	0.23 (0.059-0.89)
2	ABCB11	168969927	A/G	rs80049571	A/A	G/G	34	2	8.27	20.40	0.0032	0.13 (0.02-2.82)
1	ABCB10	229524044	T/C	rs10916508	T/T	C/C	25	6	10.33	5.44	0.0033	5.42 (0.55-53.27)
16	ABCA17P	2404539	A/G	rs4261541	G/G	A/G	45	3	9.57	17.57	0.0036	0.13 (0.01-2.57)
7	ABCB5	20678540	T/C	rs7805940	T/T	T/C	34	12	8.71	13.02	0.0039	0.21 (0.05-0.91)
7	ABCA13	48638485	G/A	rs1918583	A/A	G/A	43	5	9.09	17.94	0.0041	0.07 (0.00-1.39)
9	ABCA1	104815253	G/T	rs73519841	G/T	G/G	16	33	6.88	11.17	0.0046	0.22 (0.06-0.82)
16	ABCC11	48216918	C/T	rs28684834	C/T	T/T	24	4	10.33	5.91	0.0069	14.68 (0.71-304.33)
9	ABCA1	104894789	G/A	rs3905000	G/A	A/A	22	2	9.36	17.75	0.0072	0.14 (0.01-3.28)
7	ABCA13	48486157	G/A	rs11981491	G/G	G/A	42	7	8.71	14.15	0.0083	0.05 (0.00-0.85)
2	ABCB11	168927218	C/T	rs1521808	C/C	C/T	46	3	9.60	17.57	0.009	0.12 (0.01-2.46)
12	ABCC9	21923026	T/G	rs10770871	G/G	T/T	27	2	8.02	19.55	0.0099	0.16 (0.01-3.68)
21	ABCG1	42210739	T/TC	rs11372946	TC/TC	T/T	16	9	7.72	11.45	0.0116	0.06 (0.01-0.59)
17	ABCA8	68941291	C/T	rs960929	C/C	C/T	29	18	11.45	7.63	0.0128	3.27 (0.95-11.24)
7	ABCB4	87462869	G/A	rs2302387	A/A	G/G	9	13	6.34	11.45	0.0138	0.18 (0.03-1.23)
1	ABCB10	229524044	T/C	rs10916508	T/C	C/C	18	6	10.13	5.44	0.0148	6.25 (0.60-64.86)
7	ABCB1	87550154	C/T	rs34767983	C/T	C/C	2	47	4.86	9.92	0.0153	0.19 (0.01-4.21)
21	ABCC13	14290796	C/T	rs117289299	C/C	C/T	47	2	9.63	18.05	0.0153	0.18 (0.01-3.87)
10	ABCC2	99799337	A/G	rs17222674	A/A	A/G	46	3	10.05	5.37	0.016	7.62 (0.37-155.88)
16	ABCC11	48193978	G/A	rs141897765	G/G	G/A	46	3	9.60	17.94	0.016	0.12 (-0.01-2.46)
13	ABCC4	95053053	C/T	rs7324065	C/T	C/C	6	43	5.63	10.18	0.0167	0.17 (0.02-1.62)
12	ABCC9	21923026	T/G	rs10770871	T/G	T/T	20	2	10.00	19.55	0.0173	0.2 (0.01-4.69)
10	ABCC2	99792950	T/G	rs148713114	T/G	T/T	4	45	5.14	10.18	0.018	0.32 (0.03-3.30)
2	ABCG8	43853903	T/C	rs4077440	T/C	T/T	27	4	11.42	5.76	0.0188	4.36 (0.40-47.62)
2	ABCB11	168969927	A/G	rs80049571	A/G	G/G	13	2	11.71	20.40	0.019	0.42 (0.02-10.75)
1	ABCB10	229555533	A/C	rs12067825	A/C	A/A	15	34	7.63	10.75	0.0197	0.39 (0.11-1.40)

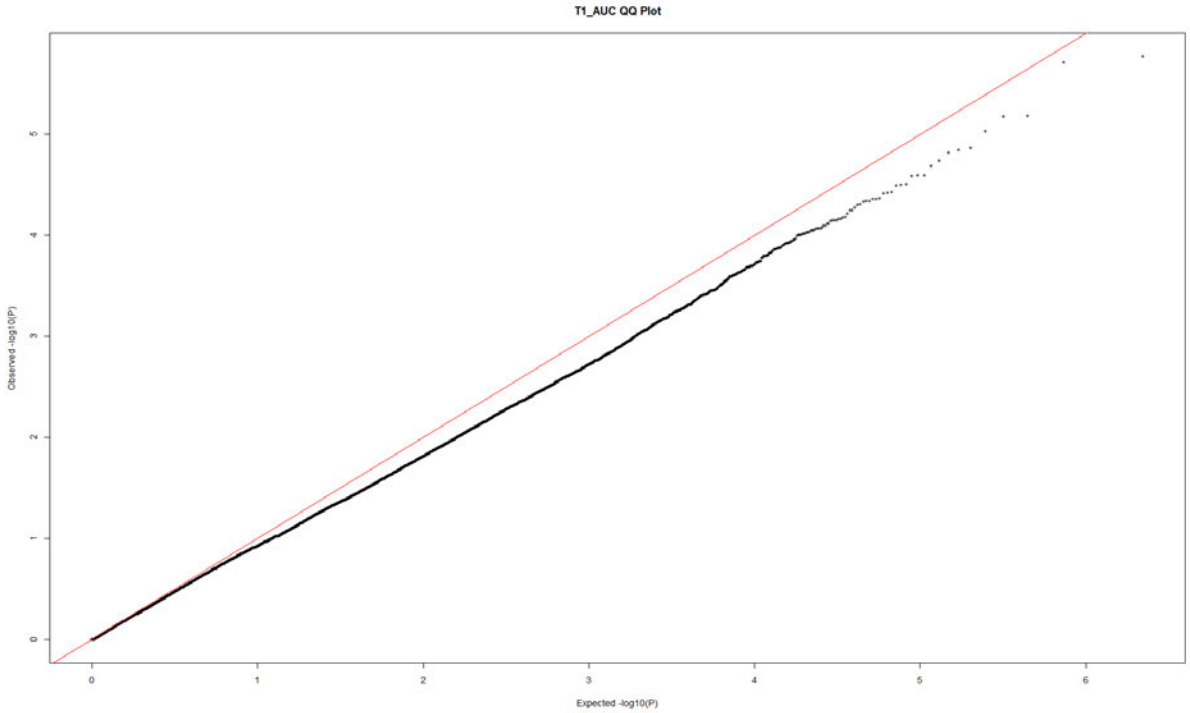
Supplementary Table F: *SLC* Gene family variants with P values ≤ 0.01 in association with moxifloxacin drug PK AUC values at T2.

Chr	Gene Name	POS	Ref/ Alt	ALT	SNP	Group 1	Group 2	Group 1 (n)	Group 2 (n)	Group 1 Median	Group 2 Median	P Value	OR, 95% CI
1	SLC44A3-AS1	94651182	T/C	C	rs75457665	T/C	T/T	15	34	5.90	11.44	7.21e-05*	0.15 (0.04-0.65)
14	SLC25A21	36900653	A/G	G	rs17177739	A/G	A/A	7	42	5.15	10.33	0.00014	0.05 (0.00-0.94)
20	SLC17A9	62956015	C/T	T	rs2427456	C/C	T/T	31	4	8.32	18.05	0.00015	0.09 (0.00-1.86)
7	SLC26A4	107666761	C/T	T	rs2701688	C/T	T/T	22	6	11.58	6.51	0.00016	26.86 (1.33-542.72)
11	SLC1A2	35341405	C/T	T	rs7942395	T/T	C/C	4	20	5.15	10.25	0.00038	0.07 (0.00-1.59)
20	SLC24A3	19516069	G/A	A	rs16980713	G/G	G/A	44	5	9.33	17.94	0.00043	0.06 (0.00-1.33)
15	SLC28A2-AS1	45262069	C/A	A	rs1060896	A/A	C/C	3	36	4.78	9.77	0.00044	0.14 (0.00-2.96)
16	SLC38A8	84023351	C/T	T	rs9783761	C/T	C/C	28	10	8.27	14.10	0.00049	0.02 (0.00-0.04)
1	SLC35D1	67021693	C/A	A	rs61780038	C/C	C/A	22	25	7.50	12.70	0.00049	0.14 (0.04-0.53)
2	SLC25A12	171855597	T/TG	TG	rs3836018	T/TG	TG/TG	25	19	7.35	12.70	0.00055	0.08 (0.02-0.35)
16	SLC7A5	87854200	G/A	A	rs7193392	G/G	G/A	20	28	6.90	10.89	0.00058	0.23 (0.07-0.81)
1	SLC35D1	67019351	A/G	G	rs2815355	A/A	A/G	20	26	7.50	12.20	0.00063	0.14 (0.03-0.55)
11	SLC67A1-AS	2901263	G/A	A	rs4930027	G/G	G/A	26	18	7.73	12.07	0.0007	0.12 (0.03-0.51)
3	SLC9A9	143614188	G/T	T	rs16853938	G/G	T/T	18	7	7.93	13.62	0.00075	0.03 (0.00-0.39)
15	SLCO3A1	92061084	G/A	A	rs74030450	G/G	G/A	39	8	8.32	13.88	0.00118	0.03 (0.00-0.69)
7	SLC37A3	140354303	T/C	C	rs73735215	T/T	T/C	46	3	9.60	18.16	0.00119	0.12 (0.01-2.46)
3	SLC9A9	143587286	C/T	T	rs56156039	T/T	C/T	5	19	17.57	7.91	0.00136	18.33 (0.88-380-72)
2	SLC8A1-AS1	40165980	A/G	G	rs75402228	A/A	A/G	45	4	10.18	5.26	0.00146	10.25 (0.52-201.64)
17	SLC47A1P2	19629642	G/A	A	rs115397975	G/G	G/A	43	5	9.09	17.33	0.00157	0.06 (0.00-1.27)
9	SLC44A1	105299600	C/A	A	rs142868271	C/A	C/C	2	47	4.57	9.92	0.0017	0.19 (0.01-4.21)
8	SLC25A37	23555426	A/G	G	rs2978470	A/G	G/G	27	12	7.91	13.00	0.00171	0.08 (0.01-0.47)
17	SLC39A11	73041047	G/A	A	rs66529051	G/G	G/A	39	10	10.33	5.43	0.00172	5.17 (0.97-27.60)
19	SLC8A2	47443142	A/G	G	rs57112810	A/A	A/G	32	16	8.27	13.47	0.00173	0.2 (0.05-0.76)
19	SLC8A2	47454274	C/T	T	rs73568302	C/C	C/T	32	16	8.27	13.47	0.00173	0.2 (0.05-0.76)
15	SLC28A1	84923740	C/T	T	rs7167591	C/C	C/T	20	24	7.63	11.56	0.00175	0.12 (0.03-0.50)
17	SLC25A19	75289806	G/A	A	rs56090937	G/A	G/G	2	46	4.57	10.05	0.00177	0.18 (0.01-4.04)
14	SLC25A21	37075306	C/T	T	rs74604833	C/T	C/C	12	34	7.48	11.30	0.00184	0.05 (0.01-0.49)
1	SLC35A3	100008541	C/T	T	rs57970573	C/T	C/C	11	38	6.44	10.75	0.00202	0.01 (0.01-0.56)
1	SLC66A1	19314182	T/C	C	rs859218	C/C	T/C	24	24	7.48	11.44	0.00235	0.42 (0.13-1.36)
2	SLC5A7	108006866	T/G	G	rs6720783	T/T	T/G	29	17	10.33	7.62	0.00246	5.31 (1.38-20.48)
12	SLC6A12	195133	A/G	G	rs12318264	A/A	A/G	34	13	7.87	12.73	0.00253	0.09 (0.02-0.52)
1	SLC16A1-AS1	112979492	C/T	T	rs17030832	C/C	C/T	39	10	8.32	13.92	0.00254	0.06 (0.01-0.60)
8	SLC7A2	17561630	G/A	A	rs10104505	G/A	G/G	11	37	5.90	10.33	0.00258	0.16 (0.03-0.89)
1	SLC44A5	75226099	C/T	T	rs679258	C/T	T/T	3	45	5.10	9.92	0.00266	0.13 (0.00-2.80)
1	SLC2A7	9006604	T/C	C	rs57628014	T/T	C/C	29	6	9.09	13.65	0.00275	0.04 (0.00-0.93)
20	SLC23A2	4911111	A/G	G	rs113907830	A/A	A/G	45	4	9.57	16.48	0.00278	0.08 (0.00-1.76)
12	SLC38A1	46184372	A/C	C	rs3742059	A/C	C/C	18	29	7.17	11.42	0.00284	0.23 (0.07-0.84)
15	SLC28A1	84899572	T/C	C	rs56690118	C/C	T/T	7	25	5.37	10.18	0.00298	0.13 (0.01-1.25)
5	SLC22A5	132387596	T/C	C	rs2073643	T/C	C/C	25	10	8.32	13.34	0.00298	0.14 (0.02-0.81)
14	SLC7A7	22784968	C/T	T	rs61974192	C/T	C/C	5	44	5.37	10.25	0.00318	0.07 (0.00-1.46)
11	SLC1A2	35341405	C/T	T	rs7942395	T/T	C/T	4	25	5.15	9.81	0.0032	0.12 (0.01-2.46)
17	SLC5A10	18953583	G/A	A	rs4924949	A/A	G/G	18	4	8.12	16.99	0.00328	0.07 (0.00-1.55)
17	SLC39A11	73083794	G/A	A	rs151082107	G/A	G/G	2	47	4.73	9.92	0.0034	0.19 (0.01-4.21)
15	SLCO3A1	92145728	A/G	G	rs758136	A/G	G/G	26	8	7.97	13.91	0.0038	0.14 (0.02-0.90)
3	SLC9C1	112263077	T/C	C	rs9809404	C/C	T/C	11	26	6.34	11.44	0.00382	0.27 (0.06-1.28)
5	SLC25A48	135594000	A/G	G	rs17168703	A/A	A/G	39	10	8.22	12.69	0.00404	0.06 (0.01-0.60)
14	SLC7A7	22777766	T/G	G	rs4981438	T/T	G/G	24	6	7.87	13.18	0.00415	0.03 (0.00-0.79)
1	SLC30A10	219896518	A/G	G	rs17006810	A/G	G/G	23	6	10.33	6.39	0.00419	7.77 (0.78-77.93)
6	SLC25A20P1	52247010	C/T	T	rs761167	C/T	C/C	17	32	7.62	11.30	0.00427	0.11 (0.03-0.48)
11	SLC6A5	20638211	G/A	A	rs2298826	G/A	G/G	19	24	7.63	11.56	0.00431	0.17 (0.05-0.67)
14	SLC35F4	57823483	G/T	T	rs10142225	G/G	G/T	18	25	6.30	10.33	0.00438	0.19 (0.05-0.75)
15	SLC28A1	84927061	G/T	T	rs2122859	G/T	G/G	8	41	5.43	10.18	0.00443	0.28 (0.05-1.60)
14	SLC25A21	37118109	C/T	T	rs66862591	C/C	C/T	14	26	6.77	11.30	0.00469	0.14 (0.04-0.65)
1	SLC44A5	75589413	C/T	T	rs396106	C/C	C/T	14	27	6.90	10.33	0.00472	0.16 (0.03-0.72)
16	SLC38A8	84023351	C/T	T	rs9783761	C/C	T/T	10	11	14.10	7.62	0.00479	24.81 (1.17-527.15)
2	SLC25A12	171916023	C/T	T	rs11695887	C/T	C/C	28	11	7.63	13.07	0.00505	0.20 (0.04-0.97)

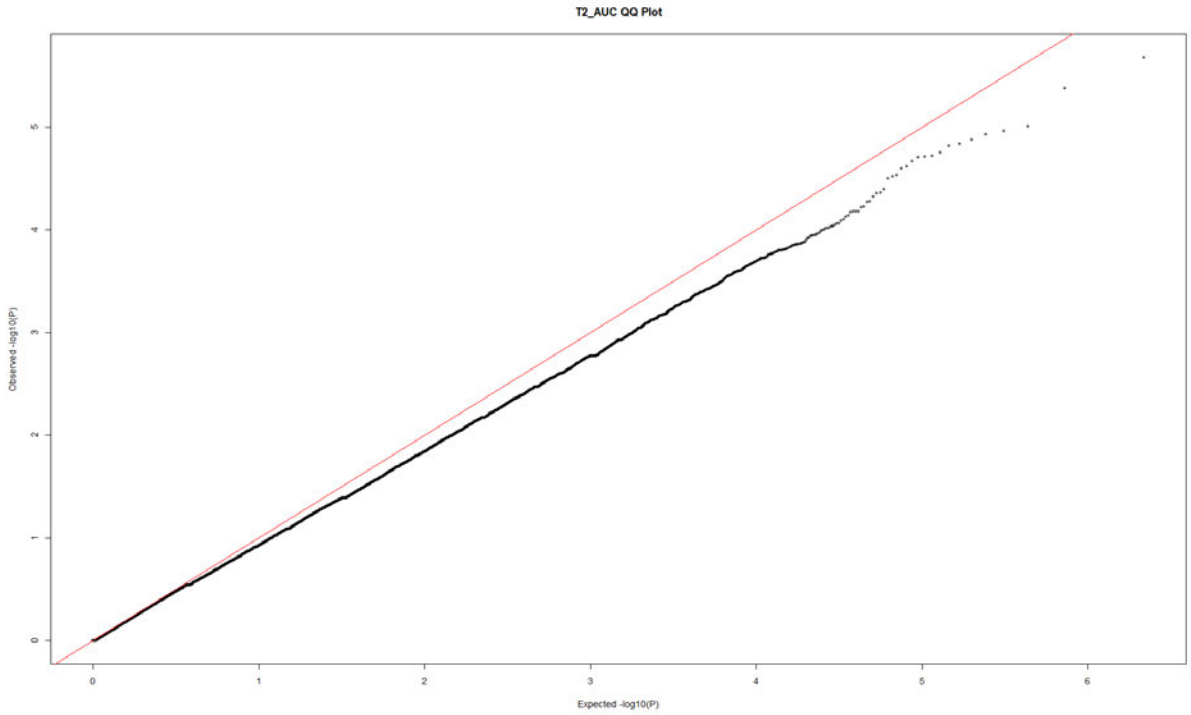
19	SLC39A3	2733252	C/T	T	rs61736898	C/C	C/T	44	5	10.25	6.15	0.00559	13.14 (0.69-252.17)
16	SLC38A8	84013301	G/A	A	rs13336065	G/A	A/A	21	9	7.82	13.07	0.00586	0.02 (0.00-0.44)
12	SLC41A2	104832472	A/G	G	rs2642120	A/G	G/G	14	3	7.98	13.62	0.00588	0.06 (0.00-1.45)
9	SLC28A3	84325099	A/G	G	rs7033971	A/A	G/G	21	8	8.22	16.60	0.00589	0.20 (0.03-1.27)
11	SLC1A2-AS1	35269781	A/G	G	rs10836361	A/G	G/G	27	4	10.18	5.55	0.00597	9.66 (0.47-1.27)
3	SLC9A9	143504202	T/G	G	rs9863900	T/T	T/G	39	9	11.17	6.99	0.00606	11.5 (1.30-101.19)
6	SLC29A1	44235246	A/G	G	rs6929249	A/A	A/G	20	24	12.20	7.87	0.00608	4.66 (1.30-16.76)
9	SLC1A1	4509590	T/C	C	rs10974584	T/T	T/C	39	10	8.22	13.47	0.00624	0.17 (0.03-0.93)
1	SLC44A3-AS1	94698742	T/C	C	rs4847341	C/C	T/T	7	26	6.15	10.87	0.00657	0.12 (0.01-1.17)
3	SLC9A9	143679268	A/G	G	rs147237793	A/A	A/G	45	4	10.18	5.77	0.00677	10.25 (0.52-201.64)
15	SLC28A2-AS1	45262069	C/A	A	rs1060896	A/A	C/A	3	10	4.78	10.75	0.00699	0.09 (0.00-2.42)
11	SLC1A2	35412379	T/C	C	rs3847624	T/C	C/C	13	3	9.63	14.53	0.00714	0.12 (0.00-2.87)
1	SLC6A17	110170913	C/T	T	rs4839205	C/C	C/T	44	5	9.33	16.65	0.00724	0.20 (0.02-2.02)
18	SLC14A2	45634334	G/A	A	rs112746495	G/G	G/A	46	3	10.05	5.37	0.00727	7.62 (0.37-155.88)
6	SLC22A2	160216207	T/C	C	rs596881	T/C	C/C	21	21	7.62	11.42	0.00729	0.30 (0.09-1.09)
8	SLC25A32	103411066	G/A	A	rs3134285	G/A	G/G	26	5	8.27	16.65	0.00739	0.04 (0.00-0.99)
3	SLC9A9	143838590	T/C	C	rs57082048	T/T	T/C	45	2	9.81	18.90	0.0074	0.19 (0.01-4.21)
8	SLCO5A1	69801997	C/T	T	rs112343884	C/T	C/C	7	42	6.34	10.25	0.00752	0.13 (0.01-1.25)
21	SLC19A1	45559605	G/A	A	rs1475596	G/G	G/A	9	23	15.63	8.32	0.00755	12.44 (1.32-117.03)
12	SLC38A4-AS1	46779280	G/A	A	rs180425	A/A	G/A	33	14	11.45	7.87	0.00775	26 (3.00-225.21)
1	SLC30A10	219930137	C/T	T	rs7523576	C/T	C/C	4	45	5.50	10.18	0.0079	0.09 (0.00-1.92)
3	SLC9A9	143713081	T/C	C	rs9289667	T/T	C/C	20	3	8.98	5.20	0.00791	5.78 (0.26-126.49)
2	SLC8A1-AS1	39955252	C/A	A	rs720796	C/C	A/A	30	2	8.71	19.24	0.00806	0.13 (0.01-3.06)
6	SLC35F1	118185042	A/G	G	rs205937	A/G	G/G	22	8	8.98	14.10	0.00848	0.09 (0.01-0.95)
15	SLC28A1	84927143	GA/G	G	rs36018131	GA/GA	G/G	5	24	5.37	11.95	0.00857	0.15 (0.01-1.56)
5	SLC9A3	503369	C/T	T	rs6420045	C/C	C/T	5	11	6.15	10.33	0.0087	0.07 (0.00-1.72)
3	SLC9A9	143587286	C/T	T	rs56156039	C/C	T/T	25	5	9.81	17.57	0.00881	0.08 (0.00-1.68)
16	SLC12A4	67967406	C/T	T	rs76481097	C/C	C/T	46	3	10.05	5.20	0.00901	7.62 (0.37-155.88)
16	SLC7A6	68270489	C/A	A	rs8056893	A/A	C/A	46	3	10.05	5.20	0.00901	7.62 (0.37-155.88)
11	SLC1A2-AS1	35269781	A/G	G	rs10836361	A/A	G/G	17	4	10.33	5.55	0.00902	12.6 (0.59-270.99)
18	SLC66A2	79946425	T/C	C	rs28549295	T/T	T/C	32	16	11.30	7.30	0.00913	5 (1.31-19.07)
2	SLC8A1-AS1	40227805	C/A	A	rs402307	C/C	C/A	32	15	7.97	12.70	0.00913	0.13 (0.03-0.56)
5	SLC12A7	1107313	C/A	A	rs4535497	C/A	A/A	23	12	7.82	12.20	0.00917	0.17 (0.04-0.85)
2	SLC8A1	40268794	A/G	G	rs13419087	A/A	G/G	25	3	11.42	6.34	0.00977	12.15 (0.57-261.61)
5	SLC12A7	1113129	G/T	T	rs4580814	G/G	T/T	23	7	8.32	13.31	0.00992	0.10 (0.01-1.04)
3	SLC25A26	66361023	A/G	G	rs80126361	A/G	A/A	10	39	6.39	10.33	0.0102	0.19 (0.04-1.03)
12	SLC16A7	59711173	T/C	C	rs73111818	T/T	T/C	47	2	9.92	4.97	0.0102	5.21 (0.24-114.41)
17	SLC38A10	81273798	G/A	A	rs72850613	G/G	G/A	47	2	9.63	19.36	0.0102	0.17 (0.01-3.87)
3	SLC9A9	143442395	C/T	T	rs62267817	C/C	C/T	44	5	9.60	17.33	0.01044	0.20 (0.02-2.02)
1	SLC44A3-AS1	94823893	A/G	G	rs11590142	A/G	A/A	4	45	5.67	10.18	0.0106	0.09 (0.00-1.92)
17	SLC16A3	82233285	T/C	C	rs12602929	C/C	T/T	13	11	12.73	6.99	0.01071	6 (1.02-35.38)
3	SLC9A9	143631999	G/A	A	rs140421046	G/G	G/A	46	3	10.05	5.20	0.01107	7.62 (0.37-155.88)
6	SLC44A4	31869289	C/T	T	rs11966200	C/C	C/T	46	3	9.60	17.33	0.01107	0.12 (0.01-2.46)
17	SLC5A10	18953558	A/G	G	rs2472693	A/G	A/A	23	10	7.62	13.00	0.01108	0.13 (0.02-0.78)
8	SLC30A8	116953011	C/CT	CT	rs145579955	C/C	C/CT	43	6	9.57	16.23	0.01125	0.15 (0.02-1.47)
12	SLC6A12	201550	T/G	G	rs76737142	T/T	T/G	43	6	9.57	14.89	0.01125	0.15 (0.02-1.47))
2	SLC19A3	227703884	T/C	C	rs55683694	C/C	T/C	10	18	6.44	12.07	0.01161	0.12 (0.02-0.78)
2	SLC9A4	102490605	C/T	T	rs1523201	T/T	C/T	41	7	9.09	17.57	0.01174	0.11 (0.01-1.07)
2	SLC25A12	171981343	G/A	A	rs1443700	G/A	A/A	30	6	7.87	13.65	0.01175	0.11 (0.01-1.12)
2	SLC25A12	171862874	AGAA/	T	rs35917877	TAGAA/T	T/T	28	10	7.92	13.02	0.01179	0.27 (0.06-1.31)
1	SLC45A1	8329094	G/A	A	rs80046868	G/G	G/A	40	9	9.33	14.53	0.01183	0.21 (0.04-1.15)
14	SLC35F4	57823483	G/T	T	rs10142225	G/G	T/T	18	6	6.30	14.35	0.01184	0.05 (0.01-0.64)
16	SLC6A2	55697252	A/G	G	rs12924088	A/A	A/G	45	4	9.57	15.40	0.0122	0.08 (0.00-1.76)
1	SLC44A3-AS1	94689057	G/A	A	rs4000657	A/A	G/G	4	27	6.09	11.42	0.01233	0.07 (0.00-1.58)
12	SLC2A14	7825127	T/C	C	rs12312135	C/C	T/T	20	3	9.72	5.50	0.01242	5.78 (0.26-126.49)
11	SLC22A10	63279526	C/T	T	rs1193906	C/T	C/C	6	43	5.32	10.18	0.01246	0.17 (0.02-1.62)
16	SLC7A6	68297987	G/A	A	rs2863979	A/A	G/A	43	6	10.18	5.55	0.01246	5.75 (0.62-53.43)
1	SLC35F3	234234907	G/A	A	rs10910390	G/G	G/A	26	21	8.12	11.42	0.01247	0.31 (0.09-1.04)

12	SLCO1B1	21178665	T/C	C	rs4149057	T/T	T/C	38	9	8.71	13.31	0.01255	0.20 (0.04-1.14)
5	SLC6A7	150203254	A/G	G	rs2240792	A/G	A/A	20	24	7.97	11.44	0.01271	0.13 (0.04-0.52)
4	SLC4A4	71564233	TGTT/(	C	rs3217637	C/C	GTT/CT(	17	10	9.09	14.89	0.01289	0.17 (0.03-1.09)
1	SLC35F3	234089771	A/G	G	rs145716512	A/G	A/A	7	41	5.15	10.18	0.0129	0.34 (0.06-1.99)
1	SLC35F3	234218424	G/A	A	rs7520297	G/A	A/A	23	9	8.22	13.07	0.01305	0.06 (0.01-0.63)
15	SLC28A1	84905841	T/C	C	rs4247408	T/T	T/C	8	18	13.80	8.27	0.01329	11 (1.10-109.68)
19	SLC25A23	6461740	T/C	C	rs168348	C/C	T/T	23	2	9.63	4.97	0.01333	4.6 (0.20-106.31)
18	SLC14A2	45569278	T/C	C	rs4890548	T/T	C/C	23	8	8.02	13.16	0.01347	0.03 (0.01-0.75)
17	SLC25A19	75279345	T/G	G	rs28516463	T/G	G/G	26	8	9.69	13.61	0.01349	0.28 (0.05-1.69)
18	SLC14A2	45633508	C/T	T	rs1484875	C/T	T/T	30	18	7.73	11.95	0.0137	0.28 (0.08-0.99)
1	SLC25A34-AS1	15749573	T/C	C	rs11580478	T/T	T/C	36	11	8.12	11.71	0.01377	0.26 (0.06-1.18)
2	SLC8A1	40530651	A/G	G	rs13017846	A/A	A/G	39	10	10.33	5.92	0.01389	5.17 (0.97-27.60)
2	SLC8A1	40533825	T/C	C	rs17026156	T/T	T/C	39	10	10.33	5.92	0.01389	5.17 (0.97-27.60)
8	SLC45A4	141273370	C/A	A	rs147931917	C/C	C/A	45	4	10.18	6.24	0.01394	10.25 (0.52-201.64)
15	SLCO3A1	92033696	A/C	C	rs12594537	C/C	A/A	32	2	9.74	4.97	0.01426	5 (0.22-112.35)
1	SLC26A9	205941755	G/T	T	rs7415921	G/T	G/G	27	20	12.73	8.02	0.01447	2.18 (0.67-7.09)
1	SLC35F3	234082401	A/C	C	rs12032360	A/A	A/C	36	13	8.27	12.73	0.01454	0.10 (0.02-0.54)
9	SLC1A1	4542097	TCTCTC	ATCTCTGAA	rs144510603	C/C	ATCTCT(	41	8	9.09	16.99	0.01455	0.26 (0.05-1.45)
9	SLC1A1	4545685	A/G	G	rs10815015	A/A	A/G	41	8	9.09	16.99	0.01455	0.26 (0.05-1.45)
12	SLC2A13	39920244	A/G	G	rs7970063	A/A	A/G	41	6	9.09	14.47	0.01465	0.06 (0.00-1.15)
11	SLC22A25	63183592	C/A	A	rs1525072	C/A	A/A	5	44	5.15	10.05	0.01467	0.22 (0.02-2.21)
8	SLCO5A1	69672393	T/C	C	rs1138541	T/C	T/T	20	21	7.82	11.42	0.01469	0.13 (0.03-0.53)
17	SLC25A10	81715528	T/C	C	rs2072700	C/C	T/C	23	17	7.91	11.45	0.01483	0.16 (0.04-0.67)
2	SLC8A1-AS1	40002657	C/T	T	rs17024801	T/T	C/C	15	13	7.62	12.70	0.01484	0.41 (0.09-1.91)
2	SLC19A3	227703884	T/C	C	rs55683694	C/C	T/T	10	21	6.44	9.81	0.01506	0.27 (0.05-1.62)
3	SLC9A9	143686482	A/G	G	rs1868175	A/G	G/G	23	14	8.32	13.73	0.01518	0.25 (0.06-1.07)
5	SLC1A3	36654391	T/G	G	rs3776579	T/T	T/G	44	4	10.05	5.85	0.01518	9.83 (0.50-193.63)
17	SLC39A11	73014638	C/T	T	rs73350914	C/C	C/T	47	2	9.92	5.12	0.01531	5.21 (0.24-114.41)
7	SLC12A9	100860471	C/G	G	rs7801190	G/G	C/G	6	14	6.66	11.44	0.01538	0.04 (0.00-0.95)
1	SLC2A5	9062899	T/C	C	rs115978069	T/C	T/T	18	29	7.92	11.17	0.01548	0.17 (0.05-0.67)
2	SLC8A1-AS1	40183129	T/C	C	rs379410	T/T	T/C	39	9	8.32	12.73	0.01554	0.07 (0.01-0.69)
2	SLC8A1	40338671	A/C	C	rs2192766	A/A	C/C	12	11	7.48	12.73	0.01561	0.07 (0.01-0.55)
12	SLC17A8	100375049	G/A	A	rs10778050	G/G	A/A	14	15	13.19	7.82	0.01568	5 (1.03-24.28)
4	SLC30A9	41996030	G/T	T	rs56119253	G/T	G/G	12	37	7.49	10.33	0.01573	0.25 (0.06-1.09)
10	SLC39A12	17961044	G/A	A	rs7074683	G/G	G/A	37	11	11.17	6.99	0.01578	6.6 (1.25-34.95)
17	SLC39A11	72899452	A/G	G	rs72847961	A/G	A/A	26	11	8.36	12.73	0.01584	0.27 (0.06-1.28)
12	SLCO1A2	21416312	CA/C	C	rs142436768	CA/CA	CA/C	45	4	9.57	16.05	0.01589	0.08 (0.00-1.76)
7	SLC37A3	140319385	AC/A	A	rs35927568	AC/AC	AC/A	46	3	9.60	16.65	0.01596	0.12 (0.01-2.46)
6	SLC35A1	87475732	T/C	C	rs2268995	T/C	C/C	19	4	9.57	17.11	0.01604	0.08 (0.00-1.73)
2	SLC30A6	32206038	C/A	A	rs212679	C/A	A/A	26	18	8.27	12.88	0.01641	0.15 (0.04-0.60)
18	SLC14A2	45564677	T/C	C	rs13382004	T/C	C/C	25	10	11.17	6.84	0.01668	9.75 (1.07-88.88)
3	SLC9A9	143713081	T/C	C	rs9289667	T/C	C/C	26	3	11.44	5.20	0.01697	9.43 (0.44-201.19)
3	SLC6A1	11035670	G/A	A	rs17033838	G/G	G/A	41	8	10.33	7.03	0.01716	3.47 (0.63-19.28)
12	SLCO1B3	20901435	G/A	A	rs3764006	G/G	A/A	29	2	9.57	18.15	0.0172	0.16 (0.01-3.71)
14	SLC24A4	92404706	T/C	C	rs59778360	T/T	T/C	29	15	10.33	7.63	0.01741	3.89 (1.00-15.21)
4	SLC4A4	71564233	TGTT/(	C	rs3217637	CTGTT/C	GTT/CT(	22	10	8.95	14.89	0.01752	0.17 (0.03-1.01)
1	SLC6A17	110159653	T/G	G	rs60970016	T/T	T/G	40	9	10.25	6.15	0.01756	4.27 (0.79-23.19)
6	SLC35A1	87475732	T/C	C	rs2268995	T/T	C/C	26	4	9.45	17.11	0.01759	0.09 (0.00-1.96)
2	SLC30A6	32222649	C/T	T	rs177083	T/T	C/C	23	4	12.70	6.39	0.01766	19.8 (0.94-416.68)
18	SLC14A2	45231762	T/C	C	rs8090126	T/T	C/C	31	3	9.57	17.33	0.01771	0.10 (0.00-2.19)
8	SLCO5A1	69801670	A/G	G	rs28414396	A/G	A/A	19	25	8.22	12.70	0.01788	0.38 (0.11-1.33)
7	SLC26A4	107666761	C/T	T	rs2701688	C/C	C/T	21	22	7.91	11.58	0.01794	0.35 (0.10- 1.22)
15	SLC27A2	50213519	C/T	T	rs58486088	T/T	C/C	3	20	5.10	9.33	0.01807	0.25 (0.01-5.68)
3	SLC9A9	143523029	T/G	G	rs12489418	T/T	T/G	44	5	9.60	16.65	0.01818	0.20 (0.02-2.02)
18	SLC14A2	45508135	C/T	T	rs72912692	C/C	C/T	44	5	9.33	13.67	0.01818	0.06 (0.00-1.33)
1	SLC6A9	43998533	A/G	G	rs3791124	G/G	A/G	44	5	10.25	6.44	0.01818	13.14 (0.69-252.17)
4	SLC4A4	71079906	G/A	A	rs76009864	G/G	G/A	43	5	9.57	14.15	0.01821	0.07 (0.00-1.39)

15	SLC30A4-AS1	45481727	A/C	C	rs6493153	A/C	A/A	24	24	7.97	10.89	0.01828	0.36 (0.11-1.16)
3	SLC9A9	143300991	T/G	G	rs6794548	G/G	T/G	35	14	8.32	12.39	0.01828	0.16 (0.04-0.69)
5	SLC25A48	135576825	T/C	C	rs73300419	T/C	T/T	15	33	7.63	10.33	0.0183	0.26 (0.07-1.02)
6	SLC35F1	118270163	C/T	T	rs79386861	C/C	C/T	29	16	8.02	12.54	0.01832	0.36 (0.10-1.29)
1	SLC44A3-AS1	94698742	T/C	C	rs4847341	C/C	T/C	7	16	6.15	9.86	0.01836	0.16 (0.02-1.72)
6	SLC35F1	118088740	A/G	G	rs9398483	G/G	A/G	43	6	9.57	15.74	0.01839	0.15 (0.02-1.47)
4	SLC4A4	71489507	G/A	A	rs10805088	A/A	G/G	19	10	10.18	7.28	0.01854	5.5 (0.91-33.18)
14	SLC7A8	23161212	A/G	G	rs73598496	A/A	A/G	32	16	11.30	7.63	0.01869	5 (1.31-19.07)
17	SLC43A2	1584194	A/G	G	rs11871321	A/A	A/G	25	18	7.63	13.02	0.01884	0.33 (0.09-1.18)
2	SLC8A1-AS1	39941438	G/A	A	rs11124710	G/G	G/A	46	3	9.60	16.65	0.01889	0.12 (0.01-2.46)
10	SLC16A12	89478945	C/A	A	rs7909696	C/C	C/A	46	3	9.60	17.57	0.01889	0.12 (0.01- 2.46)
10	SLC16A12	89529252	T/C	C	rs74146962	T/T	T/C	46	3	9.60	17.57	0.01889	0.12 (0.01-2.46
17	SLC39A11	73045450	A/G	G	rs903101	G/G	A/A	17	8	7.82	12.24	0.01895	0.10 (0.01-0.72)
14	SLC24A4	92423645	A/G	G	rs4414423	A/G	A/A	29	14	8.02	12.12	0.01903	0.21 (0.05-0.84)
1	SLC35F3	234031490	T/C	C	rs9435544	T/C	T/T	13	2	7.91	18.90	0.01905	0.09 (0.00-2.41)
3	SLC9A9	143517920	C/T	T	rs113677298	C/C	C/T	42	6	10.33	6.71	0.01905	17.21 (0.91-325.39)
1	SLC25A3P1	53437160	G/A	A	rs1288595	G/G	A/A	21	4	8.22	16.60	0.01913	0.16 (0.01-1.91)
19	SLC1A6	14967912	C/T	T	rs62113278	T/T	C/T	7	24	6.44	11.44	0.01936	0.06 (0.01-0.68)
6	SLC29A1	44235582	G/A	A	rs62401155	G/A	A/A	20	6	9.60	13.65	0.01939	0.05 (0.00-1.06)
16	SLC6A2	55664751	C/T	T	rs28618083	C/T	T/T	20	6	12.22	8.27	0.01939	15 (1.40-161.05)
15	SLC28A1	84899572	T/C	C	rs56690118	C/C	T/C	7	17	5.37	10.33	0.01949	0.14 (0.01-1.51)
1	SLC35F3	233991356	G/A	A	rs79990527	G/A	G/G	15	33	7.63	10.33	0.0195	0.26 (0.07-1.02)



Supplementary Figure 1: QQ plot of GWAS results illustrating the distribution of observed versus expected p-values for SNP associations with moxifloxacin levels at timepoint 1 (T1), highlighting the limited potential genetic influences on drug concentration due to the limited sample size, n=58.



Supplementary Figure 2: QQ plot representing GWAS results illustrating the distribution of observed versus expected p-values for SNP associations with moxifloxacin levels at timepoint 2 (T2), highlighting the limited potential genetic influences on drug concentration due to the limited sample size, n=58.

Chapter 6

6.1 Synthesis and conclusion

Africa has often been regarded as a disease burdened continent (30). Just in 2024 there were approximately 200 outbreaks across Africa. Whilst Africa is home to approximately 17% of the global population, it accounted for just over 25% of global TB cases in 2022 (2). In addition to TB, there are several health concerns that have been curbing the population such as HIV, Malaria and viral outbreaks from time to time. In general healthcare infrastructure across the continent has lots of room for development in order to manage public health and its challenges effectively. South Africa of all countries in Africa is considered to have substantial healthcare system in place yet the burden of TB still poses and a challenge affecting patients daily with up to 280,000 new cases annually 56,000 mortalities reported in 2023 (32). Adherence, sub-optimal therapies and a lack in revision of treatment regimens bring an additional layer of complications for TB patients (31). Genomics is only a single perspective in addition to factors mentioned yet plays a crucial role in determining patient outcomes.

To date several studies across the globe have intensely investigated TB pharmacodynamics (PD) and pharmacokinetics (PK) in Asian and European patients with overrepresentation in genetic databases (11, 29). The lack of African specific genetic data is slowly being addressed by several researchers both on and off continent committed to increasing representation and help understand health disparities in greater detail.

In light of the above the research aims of this study was committed and determined to further investigate the potential effects of genetic variants occurring in drug transporter genes mainly with additional focus on genome-wide markers that lead to variance in treatment response in African populations. We first investigated the role of genetic polymorphisms in drug transporter genes—particularly those belonging to the ATP-binding cassette (*ABC*) and solute carrier (*SLC*) gene families—in modulating individual variability in response to anti-tuberculosis (TB) therapy. These transporters have been found to play a crucial role in the absorption, distribution, metabolism, and excretion (ADME) of anti-TB drugs by regulating drug movement across cellular membranes. We found that comparative analyses showed that allele frequencies and functional variants within ABC and SLC transporter genes differ

considerably among African, Asian, and European populations. Such inter-population differences may contribute to variations in drug responses and treatment outcomes across ethnic groups. For instance, certain polymorphisms that are rare or absent in European populations may be relatively common in African cohorts, potentially altering the pharmacokinetics and pharmacodynamics of key first-line TB drugs such as rifampicin, isoniazid, and ethambutol. A review by Ramachandran et al., 2012 concluded that there is a critical need to incorporate pharmacogenomic evaluations into tuberculosis clinical trials to better characterize the genetic factors that influence treatment outcomes and susceptibility to adverse drug effects (39).

We then investigated the impact of selected genetic variants within the *ABCB1* and *SLCO1B1* drug transporter genes on the pharmacokinetics (PK) of moxifloxacin in South African patients undergoing anti-tuberculosis (TB) therapy. Three single nucleotide polymorphisms (SNPs) rs10261685, rs1045642, and rs11045819 were examined due to their hypothesized associations with TB recurrence, HIV-1 acquisition, and alterations in plasma drug concentration. Our results reported that the AA/AC genotypes of rs10261685 and AG/GG genotypes of rs1045642 were modestly associated with variations in moxifloxacin levels at specific time points. Notably, individuals with the CC genotype of rs11045819 exhibited significantly lower drug concentrations ($P = 0.0232$, $R^2 = 0.1117$) compared to the CA genotype, suggesting a potential rapid metabolizer phenotype for moxifloxacin. In light of this finding, our data indicates that patients harboring the CC genotype at rs11045819 may be classified as rapid metabolizers of moxifloxacin. This genotype may serve as an exploratory biomarker with potential clinical relevance; however, its utility requires validation in larger cohorts with greater statistical power, particularly in the context of dose optimization to achieve improved therapeutic outcomes. In contrast, a study conducted in the United States reported that Caucasian and Afro-American individuals with the AC genotype exhibited increased rifampin clearance compared to those with the CC genotype at rs11045819 (37). This observation contrasts with our pharmacokinetic findings for moxifloxacin, where the CC genotype was associated with enhanced metabolism, suggesting that this single nucleotide polymorphism may exert drug-specific and population-dependent effects, particularly among individuals of African ancestry. Although there were no significant associations identified between the investigated SNPs and TB recurrence or HIV-1 acquisition, the observed pharmacokinetic variability indicates that genetic polymorphisms within *ABCB1* and *SLCO1B1* may contribute to interindividual differences in moxifloxacin metabolism among African populations. However in our study

these findings underscored the importance of incorporating population-specific pharmacogenomic analyses into anti-TB treatment optimization.

Single gene SNP assays have been a great legacy method for investigating variants however over the past few years innovation by industry has brought several commercially available microarrays that target multi-ethnic genome backbones with commonly studied mutations in Asian, European and African American populations, which may be missing key variants relative to African populations. To explore our study population further we then performed our first GWAS for this study to identify potential genome-wide markers associated with HIV-1 infection, TB recurrence and infection. We successfully identified understudied and unreported genetic variants with strong associations such as rs121908780 variant, with the homozygous reference genotype, was strongly associated with HIV-1 infection ($P = 8.56\times10^{-35}$, OR >99) which has also been previously linked to cystic fibrosis. Similarly, rs886040602, also in the homozygous reference form, showed a strong association with TB infection ($P = 9.33\times10^{-87}$, OR >99) and has been reported as a pathogenic germline variant in breast/ovarian cancer. Even though these variants identified play differential roles across various populations, the findings suggest that understudied genetic biomarkers may contribute to disease susceptibility in African populations and warrant further investigation through large-scale genome-wide association studies (GWAS) to capture population-specific genetic diversity.

Lastly we performed the GWAS analysis in the context of better understanding the genetic variants in ABC and SLC drug transport gene families from South African patients undergoing anti-TB treatment with moxifloxacin at two different timepoints that were one month apart. Our observations reported several genome-wide markers in both *ABC* and *SLC* variants that showed suggestive significance ($p \leq 1.0\times10^{-5}$) however require additional investigation, with allele frequencies that contradicted from European and Asian populations. In particular, intronic solute carrier variants rs4414423 and rs75457665 were nominally significant at both time points, with homozygous reference genotypes A/A [$p = 9.50\times10^{-5}$, OR = 0.18 (0.04–0.71)] and T/T [$p = 7.21\times10^{-5}$, OR = 0.15 (0.04–0.65)], respectively, associated with higher median moxifloxacin levels. Several genome-wide variants with known clinical associations exhibited effects differing from the phenotypes assessed in this study. These differences can be expected due to the great level of genetic diversity observed in African populations, differences in variants and frequencies are likely to be associated with various phenotypes. These findings suggest that both known and novel genetic markers contribute to inter-individual variability in

drug pharmacokinetics across populations. Larger studies focusing on intronic solute carrier variants and their impact on moxifloxacin pharmacokinetics are warranted.

The genome-wide association study data from this study provides several notable strengths and advances to the field of research focused on African populations. This study is among the very few GWAS investigations conducted to date on South African Zulu populations, addressing a major gap in genomic research on African cohorts. Within this study, we identified novel genetic variants associated with HIV and TB disease, as well as variants with genotypic patterns that suggest metabolizer status in patients undergoing anti-TB therapy. These findings underscore a critical point: optimal treatment outcomes in African patients cannot be expected when therapeutic strategies and even the drugs themselves are largely designed based on genetic profiles of Caucasian populations. This surfaces discussions on the importance of population specific therapeutic drug monitoring (TDM) for optimal treatment outcomes. A recent study by Sarkar et al., 2024 underscored that TDM is a routine clinical approach that integrates pharmacokinetic and pharmacodynamic parameters to guide dose optimization. (38) Variability in drug pharmacokinetics is a major determinant of suboptimal therapeutic outcomes and contributes to the emergence of acquired drug resistance (38).Our results highlight the urgent need for more population-specific genomic studies to guide the development of therapeutics that are as genetically diverse and unique as the African patients who require them.

In summary even though several of the variants identified in this study are unreported intronic variants that have not been published or linked to clinical associations, it is strongly anticipated that these variants are unobserved in whole exome sequencing (WES) studies that lack intron targets and should not be overlooked in future research due to their potential ability to affect RNA splicing leading to potential loss of gene function in some cases. Large national genomic initiatives and population genetic studies focused on pan African cohorts can bring invaluable understanding of African genomes and the role of these genetic factors in clinical associations .Several reviews underscore the greater need for genomic initiatives to address the African diversity in the best interest of precision medicine (34). *Nature Genetics* publication referenced in Appendix B summarizes the efforts undertaken across the African continent to date to enhance the capacity of scientists to address the underrepresentation of African populations in genomic databases in the near future, primarily through the annual organization of workshops focused on genomics and bioinformatics (36). A recent publication by Equils et al., 2023 stated

that the advancing in biomedical technologies, the accessibility of extensive genetic and multi-omic datasets, along with the application of artificial intelligence have the potential to enhance public health interventions towards infections diseases from a precision medicine approach (35).Not only will these efforts and future population genetic studies add to underrepresented African profiles in global genetic databases but vastly improve content on GWAS microarrays that will incorporate comprehensive pharmacogenomic marker panels that will ensure that the products of precision medicine as just as unique as the patients receiving them.

6.2 Limitations and Recommendations

The research performed in this study provides context of treatment diversity in African patients fighting the burden of tuberculosis. One of the limitations within this study was the sample size however representation from various cohorts in Africa does provide a considerable level diversity to an extent. A single study in isolation would not suffice in order to achieve a greater level of understanding of African specific pharmacodynamics. Additional collaborative efforts are needed to combine studies that are performed in isolation which can aid in understanding the complete overview of African genetic diversity with regards to precision medicine for TB patients. In some instances, there was a shortage of samples due to a loss to follow up for some of patient samples from clinical trials included in this study. Nevertheless, the genome-wide association study conducted within this larger study did produce data to serve as statistical power. Genomics in Africa is expanding at an exponential rate, hence this study adds to the pivotal lack in research that is conducted to date on African populations. The data concludes that several large national genomic initiatives (LNGI) need to be conducted to gain deeper insights into the African genome and further our understanding of how to deliver optimal treatment and dosing therapies that are unique as the patients who will receive them. Other key factors that could potentially impact or influence TB patient metabolizer or metabolizer status such as age, gender, common health complications and co-morbidities need to be considered at greater length for future studies. With Africa being a large continent, it is likely that population genetics differ based on the different geographies of the continent. In the near future, there will be additional anti-TB therapies such as delamanid and pretomanid which will be integrated into treatment regimens and need to be investigated also. Despite this research study being limited by sample numbers, it is evident that the data suggested provides valuable insights in understanding the biology of anti-TB patients based on their drug transporter genes and overall genetic landscape.

Chapter 7

7.1 References:

1. Harishankar M, Selvaraj P, Bethunaickan R. Influence of Genetic Polymorphism Towards Pulmonary Tuberculosis Susceptibility. Front Med (Lausanne). 2018;5:213.
2. Naidoo A, Chirehwa M, Ramsuran V, McIlleron H, Naidoo K, Yende-Zuma N, et al. Effects of genetic variability on rifampicin and isoniazid pharmacokinetics in South African patients with recurrent tuberculosis. Pharmacogenomics. 2019;20(4):225-40.
3. Dowdy D.W. RMC. Basic and Descriptive Epidemiology of Tuberculosis. 2021.
4. Health SADo. South African National Tuberculosis Management Guidelines. 2014.
5. Chigutsa E, Visser ME, Swart EC, Denti P, Pushpakom S, Egan D, et al. The SLCO1B1 rs4149032 polymorphism is highly prevalent in South Africans and is associated with reduced rifampin concentrations: dosing implications. Antimicrob Agents Chemother. 2011;55(9):4122-7.
6. Naidoo A, Ramsuran V, Chirehwa M, Denti P, McIlleron H, Naidoo K, et al. Effect of genetic variation in UGT1A and ABCB1 on moxifloxacin pharmacokinetics in South African patients with tuberculosis. Pharmacogenomics. 2018;19(1):17-29.
7. McIlleron H, Wash P, Burger A, Norman J, Folb PI, Smith P. Determinants of rifampin, isoniazid, pyrazinamide, and ethambutol pharmacokinetics in a cohort of tuberculosis patients. Antimicrob Agents Chemother. 2006;50(4):1170-7.
8. Levison ME, Levison JH. Pharmacokinetics and pharmacodynamics of antibacterial agents. Infect Dis Clin North Am. 2009;23(4):791-815, vii.
9. Franke RM, Gardner ER, Sparreboom A. Pharmacogenetics of drug transporters. Curr Pharm Des. 2010;16(2):220-30.
10. Delmas O, Holmes EC, Talbi C, Larrous F, Dacheux L, Bouchier C, et al. Genomic diversity and evolution of the lyssaviruses. PLoS One. 2008;3(4):e2057.
11. Tucci S, Akey JM. The long walk to African genomics. Genome Biol. 2019;20(1):130.
12. Ellard GA, Gammon PT. Pharmacokinetics of isoniazid metabolism in man. J Pharmacokinet Biopharm. 1976;4(2):83-113.
13. Azuma J, Ohno M, Kubota R, Yokota S, Nagai T, Tsuyuguchi K, et al. NAT2 genotype guided regimen reduces isoniazid-induced liver injury and early treatment failure in the 6-month four-drug standard treatment of tuberculosis: a randomized controlled trial for pharmacogenetics-based therapy. Eur J Clin Pharmacol. 2013;69(5):1091-101.
14. Borst P, Evers R, Kool M, Wijnholds J. A family of drug transporters: the multidrug resistance-associated proteins. J Natl Cancer Inst. 2000;92(16):1295-302.
15. Dean M, Hamon Y, Chimini G. The human ATP-binding cassette (ABC) transporter superfamily. J Lipid Res. 2001;42(7):1007-17.
16. Gottesman MM, Ambudkar SV. Overview: ABC transporters and human disease. J Bioenerg Biomembr. 2001;33(6):453-8.
17. Lin JH, Yamazaki M. Role of P-glycoprotein in pharmacokinetics: clinical implications. Clin Pharmacokinet. 2003;42(1):59-98.

18. Jonker JW, Buitelaar M, Wagenaar E, Van Der Valk MA, Scheffer GL, Scheper RJ, et al. The breast cancer resistance protein protects against a major chlorophyll-derived dietary phototoxin and protoporphyria. Proc Natl Acad Sci U S A. 2002;99(24):15649-54.
19. Sparreboom A, van Asperen J, Mayer U, Schinkel AH, Smit JW, Meijer DK, et al. Limited oral bioavailability and active epithelial excretion of paclitaxel (Taxol) caused by P-glycoprotein in the intestine. Proc Natl Acad Sci U S A. 1997;94(5):2031-5.
20. Fromm MF. The influence of MDR1 polymorphisms on P-glycoprotein expression and function in humans. Adv Drug Deliv Rev. 2002;54(10):1295-310.
21. Lockhart AC, Tirona RG, Kim RB. Pharmacogenetics of ATP-binding cassette transporters in cancer and chemotherapy. Mol Cancer Ther. 2003;2(7):685-98.
22. Marzolini C, Paus E, Buclin T, Kim RB. Polymorphisms in human MDR1 (P-glycoprotein): recent advances and clinical relevance. Clin Pharmacol Ther. 2004;75(1):13-33.
23. Hediger MA, Romero MF, Peng JB, Rolfs A, Takanaga H, Bruford EA. The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteinsIntroduction. Pflugers Arch. 2004;447(5):465-8.
24. Hagenbuch B, Meier PJ. The superfamily of organic anion transporting polypeptides. Biochim Biophys Acta. 2003;1609(1):1-18.
25. Hsiang B, Zhu Y, Wang Z, Wu Y, Sasseville V, Yang WP, et al. A novel human hepatic organic anion transporting polypeptide (OATP2). Identification of a liver-specific human organic anion transporting polypeptide and identification of rat and human hydroxymethylglutaryl-CoA reductase inhibitor transporters. J Biol Chem. 1999;274(52):37161-8.
26. Dompreh A, Tang X, Zhou J, Yang H, Topletz A, Adu Ahwireng E, et al. Effect of Genetic Variation of NAT2 on Isoniazid and SLCO1B1 and CES2 on Rifampin Pharmacokinetics in Ghanaian Children with Tuberculosis. Antimicrob Agents Chemother. 2018;62(3).
27. Hennig S, Naiker S, Reddy T, Egan D, Kellerman T, Wiesner L, et al. Effect of SLCO1B1 Polymorphisms on Rifabutin Pharmacokinetics in African HIV-Infected Patients with Tuberculosis. Antimicrob Agents Chemother. 2016;60(1):617-20.
28. Weiner M, Peloquin C, Burman W, Luo CC, Engle M, Prihoda TJ, et al. Effects of tuberculosis, race, and human gene SLCO1B1 polymorphisms on rifampin concentrations. Antimicrob Agents Chemother. 2010;54(10):4192-200.
29. Fan, J., He, HY., Li, HH. *et al.* Associations between UGT1A1, SLCO1B1, SLCO1B3, BLVRA and HMOX1 polymorphisms and susceptibility to neonatal severe hyperbilirubinemia in Chinese Han population. *BMC Pediatr* 24, 82 (2024). https://doi.org/10.1186/s12887-024-04537-0
30. de-Graft Aikins A, Unwin N, Agyemang C, Allotey P, Campbell C, Arhinful D. Tackling Africa's chronic disease burden: from the local to the global. Global Health. 2010 Apr 19;6:5. doi: 10.1186/1744-8603-6-5. PMID: 20403167; PMCID: PMC2873934.
31. Sazali MF, Rahim SSSA, Mohammad AH, Kadir F, Payus AO, Avoi R, Jeffree MS, Omar A, Ibrahim MY, Atil A, Tuah NM, Dapari R, Lansing MG, Rahim AAA, Azhar ZI. Improving Tuberculosis Medication Adherence: The Potential of Integrating Digital Technology and Health Belief Model. Tuberc Respir Dis (Seoul). 2023 Apr;86(2):82-93. doi: 10.4046/trd.2022.0148. Epub 2022 Dec 23. PMID: 36597583; PMCID: PMC10073608.

32. WHO. Global Tuberculosis Report 2024 2024. https://www.who.int/teams/ global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2024 (ac- cessed November 25, 2024).
33. Antimicrobial Resistance Collaborators. The burden of bacterial antimicrobial resistance in the WHO African region in 2019: a cross-country systematic analysis. Lancet Glob Health. 2024 Feb;12(2):e201-e216. doi: 10.1016/S2214-109X(23)00539-9. Epub 2023 Dec 19. PMID: 38134946; PMCID: PMC10805005.
34. Sibomana O. Genetic Diversity Landscape in African Population: A Review of Implications for Personalized and Precision Medicine. Pharmgenomics Pers Med. 2024 Nov 11;17:487-496. doi: 10.2147/PGPM.S485452. PMID: 39555236; PMCID: PMC11566596.
35. Equils O, Bakaj A, Wilson-Mifsud B, Chatterjee A. Restoring Trust: The Need for Precision Medicine in Infectious Diseases, Public Health and Vaccines. Hum Vaccin Immunother. 2023 Aug 1;19(2):2234787. doi: 10.1080/21645515.2023.2234787. Epub 2023 Jul 19. PMID: 37465958; PMCID: PMC10361134.
36. Sharaf, A., Nesengani, L.T., Hayah, I. et al. Establishing African genomics and bioinformatics programs through annual regional workshops. Nat Genet 56, 1556–1565 (2024). https://doi.org/10.1038/s41588-024-01807-6
37. Kwara A, Cao L, Yang H, Poethke P, Kurpewski J, Tashima KT, Mahjoub BD, Court MH, Peloquin CA. Factors associated with variability in rifampin plasma pharmacokinetics and the relationship between rifampin concentrations and induction of efavirenz clearance. Pharmacotherapy. 2014 Mar;34(3):265-71. doi: 10.1002/phar.1388. Epub 2014 Jan 13. PMID: 24420746; PMCID: PMC3947488.
38. Sarkar, Malay & Sarkar, Jasmine. (2024). Therapeutic drug monitoring in tuberculosis. European Journal of Clinical Pharmacology. 80. 1659-1684. 10.1007/s00228-024-03749-8.
39. Ramachandran G, Swaminathan S. Role of pharmacogenomics in the treatment of tuberculosis: a review. Pharmgenomics Pers Med. 2012;5:89-98. doi: 10.2147/PGPM.S15454. Epub 2012 Sep 13. PMID: 23226065; PMCID: PMC3513231.

Appendix A

Preface:

The previous chapters of this thesis outlined the genetic diversity and its potential impact and association with disease susceptibility, TB and HIV co-infection as well as the influence on variable drug pharmacodynamics. Apart from drug transporter genes that were investigated in this study, it is evident that there are several other factors that can possibly influence disease susceptibility and progression. Human Leukocyte Antigen (HLA) typing currently remains a cornerstone of modern immunology, utilized as a critical tool for understanding antigen presentation, immune recognition, and pathogen interactions with the host. Global HLA research has been mainly focused on populations of European and Asian ancestry, leaving a substantial gap in high-resolution HLA typing data from African cohorts.

This gap warrants investigation due to African populations harboring the highest levels of HLA genetic diversity worldwide. The lack of representation of African patients in immunogenetic studies limits our understanding of how this diversity drives immune responses to major infectious diseases that remain a burden to the African continent. Tuberculosis (TB), HIV infection, and their frequent co-occurrence surface compelling contexts in which HLA variation may influence disease susceptibility, progression, and recurrence. Accumulating evidence suggests that specific HLA alleles may shape host immunity to TB and HIV, however the full spectrum of these associations remains poorly understood in African populations.

This chapter explores diversity in HLA typing within South African patients as an opportunity to discover novel immunological insights. By examining how HLA diversity associates with TB infection, HIV/TB co-infection, and TB recurrence, we aim to approach these persistent public health challenges from an alternate perspective. Broadening the immunogenetic landscape through comprehensive HLA profiling in African cohorts has potential not only to further our understanding from a biological perspective but also to guide future diagnostic, therapeutic, and vaccine strategies tailored to the populations that are most affected.

Exploring the Human Leukocyte Antigen (HLA) Landscape in Tuberculosis-Affected South African Populations: Genetic Associations with TB Infection, HIV Co-infection, and Recurrence

Arumugam T., Naidoo A., **Tulsi S**., and Ramsuran V.

School of Laboratory Medicine and Medical Sciences, University of KwaZulu-Natal, Durban, South Africa

Manuscript ID: To be submitted

Abstract

With only 11% of the world's population; Africa accounts for over 25% of active Tuberculosis (TB) cases. Genetic variation within immune-regulatory genes may contribute to Africa's disproportionate TB burden. The human leukocyte antigen (HLA) region plays an important role in immune responses by presenting pathogen-derived peptides to T cells. It has been implicated in the pathophysiology of several non-communicable and communicable diseases. Diversity within the HLA region may contribute to the disproportionate burden of the disease; however there is limited HLA typing information on African populations. Furthermore, the diversity of the HLA region may be associated with HIV co-infection and the recurrence of TB. Thus, this study aimed to investigate the association of HLA diversity within TB and HIV-affected South African populations. Class I and II HLA alleles of Black South Africans (n=1000; TB-positive and negative) were imputed using the Axiom™ Precision Medicine Diversity Array on the GeneTitan MC Fast Scan Instrument. We identified 123 Class I HLA alleles and 113 Class II HLA allele. Among TB-positive and negative individuals, 56 unique Class I and 45 unique Class II HLA alleles were identified. Furthermore, HLA-A*2601, HLA-A*7401, HLA-B*0702, HLA-C*0702, HLA-C*0401, HLA-C*1701, HLA-C*1801, HLA-DPB1*0402, HLA-DQA1*0101, HLA-DQA1*0102, HLA-DQA1*0301, HLA-DQA1*0501, HLA-DRB3*0202, HLA-DRB3*0301, HLA-DRB4*0103 were found to strongly associate with increased risk to TB infection; while HLA-B*0801, HLA-B*1302, HLA-B*1401, HLA-B*3910, HLA-C*0804.HLA-DQB1*0202, HLA-DQB1*0319, HLA-DRB1*0301, HLA-DRB1*0401, HLA-DRB3*0101 was found to be protective against TB infection. Within the TB-positive individuals, HLA-DPB1*0402, HLA-B*5703, and HLA-C*1203 were associated

with an increased risk of HIV co-infection, and HLA-DRB1*1501, HLA-DRB1*0901, and HLA-DPB1*1701 protected against co-infection. HLA-C*1203 and HLA-DQB1*0609 showed the strongest association with disease recurrence, after curative treatment. We have identified several unique HLA alleles associated with either protection or risk against TB infection, HIV co-infection, and the recurrence of TB infection. HLA associations may enable risk stratification of those with HIV infection and TB recurrence, enabling targeted interventions. Moreover, protective HLA alleles could inform the design of future TB vaccines, enhancing their efficacy in African populations.

Introduction

Tuberculosis (TB) has re-emerged as the leading cause of death from a single infectious agent following the COVID-19 pandemic. In 2023, an estimated 10.8 million individuals fell ill with TB, and 1.25 million deaths occurred worldwide (87). Despite being preventable and treatable in most parts of the world, TB remains widespread in Sub-Saharan Africa, where it accounts for a quarter of all active TB cases (88-90). Moreover, in regions with diverse ethnic populations, individuals of African ancestry experience a higher incidence and more severe clinical manifestations of TB (91-93). Numerous social and environmental factors, such as HIV co-infection, poverty, and inadequate healthcare infrastructure, significantly contribute to the disproportionate burden experienced by African communities due to the disease (94-96). In Sub-Saharan Africa, the impact of TB is exacerbated by the HIV epidemic. People living with HIV (PLWHIV) are 15-21 times more likely to progress to active TB infection, thereby increasing MTB transmission rates in this region (97).

Globally, South Africa has the eighth highest TB incidence and has the highest burden of TB/HIV co-infection, with 55.4% of TB cases occurring in people living with HIV (PLWHIV) (98). HIV-induced depletion of CD4+ T cells compromises immune recognition of *Mycobacterium tuberculosis* (MTB), resulting in an increased risk of TB and death (99). In addition, TB recurrence, either due to relapse (reactivation of the same MTB strain) or reinfection (new exposure to MTB), poses a major public health challenge in countries with high incidences of TB (100, 101).

While environmental and social factors may contribute to the disproportionate burden of TB within Sub-Saharan Africa, genetic factors may also be a key contributor to TB susceptibility (102-105). Several studies have found that genes responsible for immune regulation, such as

the histocompatibility leukocyte antigen (HLA) system, chemokine production, and signalling receptors, have been linked to an increased risk of developing TB(106-111).

The HLA system plays an important role in immune responses by presenting pathogen-derived peptides to T cells and has been associated with the pathophysiology of several non-communicable and communicable diseases including TB (112). Thus, the highly polymorphic HLA region can influence immune recognition, pathogen clearance, and disease susceptibility. The HLA region is grouped into 3 classes with more than 34,000 different possible HLA allelic variants (113). The extensive HLA diversity ensures that individuals within a population display a broad and variable spectrum of pathogen-derived peptides, as each allele binds and presents a distinct repertoire of epitopes (114, 115).

Several studies have reported the association of HLA class II polymorphisms, specifically HLA-DQ alleles, with TB susceptibility and protection; however, the results differed in different ethnic groups. Among Iranian TB patients, HLA-DQA1*0101 significantly increased, while HLA-DQA1*0301 and HLA-DQA1*0501 significantly decreased the risk of TB, while HLA-DQB1*0502 allele conferred a higher risk of TB in Polish individuals (116, 117). In Koreans, HLA-DQB1*0601 allele was associated with the disease progression of TB (118). In addition, HLA-DQB1*0503 was associated with TB among the Cambodian population (119).

Given that South African populations have distinct genetic diversity, there is a need to characterize population-specific HLA-TB associations. Furthermore, the high incidence of TB/HIV infection in South Africa also complicates HLA-TB interactions by weakening the CD4+ T cell response, which is crucial for MTB control (97, 98). Some HLA alleles are linked to faster HIV progression, while others enhance MTB antigen presentation and T cell-mediated immunity (120-122). However, the extent to which specific HLA alleles modify TB susceptibility in the context of HIV remains unclear. Additionally, HLA's role in TB recurrence is underexplored, despite evidence suggesting that certain HLA variants may influence reinfection risk and immune memory responses.

Given the high burden of TB, HIV co-infection, and recurrent TB in South Africa, there is an urgent need to identify host genetic factors influencing disease susceptibility. This study aims to investigate the association of HLA alleles associated with TB susceptibility and disease

outcomes (TB/HIV co-infection and recurrent TB infections) within a South African Black population.

Method

Study Population

This study investigated the association of HLA alleles associated with susceptibility to TB infection, TB/HIV co-infection, and recurrent TB infections. A total of 722 stored samples [buffy coats or peripheral mononuclear cells (PBMCs)] were HLA typed across various cohorts – TB negative (n=534) and TB positive (n=592). All individuals were adults 18 years and older, from KwaZulu-Natal, South Africa and of African descent.

Individuals with either a current or previous history of active TB infection were considered TB positive (smear-positive or Xpert-positive). These individuals were also tested for HIV (TB+/HIV+ = 479 and TB-/HIV- = 119). Out of 350 individuals that were followed up, 16 had a recurrent TB infection.

DNA extraction and HLA typing

DNA was extracted from all samples using the QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany) as per manufacturer specifications. Extracted DNA quality and concentration was assessed using the NanoDrop™ 8000 Spectrophotometer (ThermoFisher Scientific, Massachusetts, USA). An absorbance ratio of 260/280 and a purity index of > 1.8 were considered optimal.

The extracted DNA was genotyped using the Axiom Precision Medicine Diversity Array (PMDA) (ThermoFisher Scientific, Massachusetts, USA). The PMDA array enabled the genotyping of more than 850,000 variants, including more than 9,000 HLA markers. Genotyping was conducted according to the PharmacoScan™ Assay 96-Array Format. The genotyping assay was then analyzed on the Applied Biosystems GeneTitan™ Multi-Channel Instrument for automated, hands-free processing including hybridization, staining, washing and imaging.

Data analysis

Quality control (QC) analysis of data was performed using the Best Practice workflow using the Axiom Array Analysis Suite, version 4.0.3.3 (ThermoFisher Scientific, Massachusetts, USA). The workflow performs quality control analysis for samples and only genotypes those

samples which pass the defined QC thresholds, and then categorizes the probe sets to identify those whose genotypes are recommended for statistical tests in downstream study. Thereafter, the Axiom HLA Analysis 1.2 software (ThermoFisher Scientific, Massachusetts, USA) was used to infer HLA types.

Statistical analyses were performed using Microsoft excel, Graphpad Prism (version 8), and R software (version 4.2.2). The frequencies of class I (HLA-A, HLA-B and HLA-C) and class II (HLA-DPA1, HLA-DPB1, HLA-DQ1, HLA-DQB1, HLA-DRB1, HLA-DRB3, HLA-DRB4, and HLA-DRB5) were determined by direct counting. The chi-square test or Fisher's exact test was used, as appropriate, assess the association of each HLA allele with susceptibility of TB infection, TB/HIV co-infection and recurrence of TB infection. Bonferroni correction was applied by multiplying p values to calculate corrected p values. Odds ratios (OR) with 95% confidence intervals (CIs) were used to quantify the effect of each HLA allele and condition.

Results

Class I HLA alleles are significantly associated with TB infection

We evaluated the frequency of HLA alleles within individuals infected with TB (TB positive) and those with no history of TB (TB negative). Across the entire study population, we identified 47 HLA-A alleles, 44 HLA-B alleles, and 32 HLA-C alleles (Supplementary Table 1). There were 56 unique Class I HLA alleles between TB-positive and negative individuals. Thirty Class I HLA alleles are significantly associated with TB susceptibility (Table 1).

Seven HLA-A alleles were significantly associated with TB infection (Table 1). HLA-A*01:23 ($p = 0,024$), HLA-A*29:11 ($p < 0,001$), HL A-A*30:09 (0,024), HLA-A*43:01 ($p < 0,001$), and HLA-A*74:00 ($p < 0,001$) were exclusively found in TB negative individuals, suggesting a strong protective effect. Individuals with HLA-A*26:01 ($p = 0,006$) and HLA-A*74:01($p = 0,001$) had a significantly higher risk of contracting TB. After multiple testing corrections, only HLA-A*30:09, HLA-A*43:01, and HLA-A*74:00 remained significantly associated with TB protection.

Eleven HLA-B alleles were significantly associated with TB susceptibility (Table 1). HLA-B*07:02 ($p = 0,024$), HLA-B*39:01 ($p = 0,004$), HLA-B*39:09 ($p = 0,010$), and HLA-B*39:10 ($p= 0,012$) was most frequent among the TB positive individuals. HLA-B*08:01 ($p = 0,008$), HLA-B*13:02 ($p = 0,047$), HLA-B*14:01 ($p<0,001$), HLA-B*27:05 ($p<0,001$), HLA-B*42:02 ($p<0,001$), HLA-B*45:07 ($p = 0,005$), and HLA-B*51:01 ($p = 0,025$) were significantly

associated with protection against TB. However, after multiple testing corrections, only HLA-B*27:05 (p<0,001) and HLA-B*42:04 (p<0,001) remained significant.

Twelve HLA-C alleles were significantly associated with TB infection. HLA-C*18:01 showed the strongest association, with individuals carrying HLA-C*18:01 being 7 times more likely to develop TB (p<0,001). HLA-C*02:02 (p<0,001), HLA-C*04:01 (p = 0,013), HLA-C*07:02 (p = 0,001), and HLA-C*17:01 (p <0,001) were also associated with increased TB susceptibility (Table 1). HLA-C*02:10 (p<0,001), HLA-C*02:17 (p <0,001), HLA-C*07:20 (p = 0,032), HLA-C*15:05 (p = 0,032), HLA-C*17:00 (p<0,001), and HLA-C*18:00 (p<0,001) were exclusively found in TB negative individuals. After multiple testing corrections HLA-C*0202 (p<0,001), HLA-C*0210 (p<0,001), HLA-C*17:00 (p<0,001), HLA-C*17:01 (p<0,001), HLA-C*18:00 (p<0,001), and HLA-C*18:01 (p<0,001) remained significantly associated with TB susceptibility.

Table 1: Class I HLA alleles significantly associated with TB infection.

Alleles	TB Positive (%)	TB Negative (%)	OR	p-Value	Adjusted p-Value	CI
HLA-A*01:23	0 (0,00)	5 (0,47)	0,000	0,024	1,000	0,00 - 0,00
HLA-A*26:01	47 (3,97)	21 (1,97)	2,061	0,006	1,000	1,22 - 3,47
HLA-A*29:11	0 (0,00)	16 (1,50)	0,000	<0,001	**0,001**	0,00 - 0,00
HLA-A*30:09	0 (0,00)	5 (0,47)	0,000	0,024	1,000	0,00 - 0,00
HLA-A*43:01	0 (0,00)	31 (2,90)	0,000	<0,001	**<0,001**	0,00 - 0,00
HLA-A*74:00	0 (0,00)	37 (3,46)	0,000	<0,001	**<0,001**	0,00 - 0,00
HLA-A*74:01	66 (5,57)	29 (2,72)	2,115	0,001	0,177	1,36 - 3,30
HLA-B*07:02	74 (8,13)	59 (5,56)	1,505	0,024	1,000	1,06 - 2,14
HLA-B*08:01	37 (4,07)	73 (6,87)	0,574	0,008	1,000	0,38 - 0,86
HLA-B*13:02	5 (0,55)	16 (1,51)	0,361	0,047	1,000	0,13 - 0,99
HLA-B*14:01	5 (0,55)	32 (3,01)	0,178	<0,001	**0,009**	0,07 - 0,46
HLA-B*27:05	21 (2,31)	0 (0,00)	inf	<0,001	**<0,001**	0,00 - inf
HLA-B*39:01	7 (0,77)	0 (0,00)	inf	0,004	1,000	0,00 - inf
HLA-B*39:09	6 (0,66)	0 (0,00)	inf	0,010	1,000	0,00 - inf
HLA-B*39:10	6 (0,66)	22 (2,07)	0,314	0,012	1,000	0,13 - 0,78
HLA-B*42:02	0 (0,00)	25 (2,35)	0,000	<0,001	**<0,001**	0,00 - 0,00
HLA-B*45:07	0 (0,00)	9 (0,85)	0,000	0,005	1,000	0,00 - 0,00
HLA-B*51:01	1 (0,11)	9 (0,85)	0,129	0,025	1,000	0,02 - 1,02
HLA-C*02:02	63 (6,47)	15 (1,41)	4,845	<0,001	**<0,001**	2,74 - 8,57
HLA-C*02:10	0 (0,00)	78 (7,32)	0,000	<0,001	**<0,001**	0,00 - 0,00
HLA-C*02:17	0 (0,00)	7 (0,66)	0,000	0,016	1,000	0,00 - 0,00
HLA-C*04;01	125 (12,83)	100 (9,38)	1,422	0,013	1,000	1,08 - 1,88
HLA-C*07:02	106 (10,88)	70 (6,57)	1,738	0,001	0,157	1,27 - 2,38

HLA-C*07:20	0 (0,00)	6 (0,56)	0,000	0,032	1,000	0,00 - 0,00
HLA-C*08:04	28 (2,87)	52 (4,88)	0,577	0,022	1,000	0,36 - 0,92
HLA-C*15:05	0 (0,00)	6 (0,56)	0,000	0,032	1,000	0,00 - 0,00
HLA-C*17:00	0 (0,00)	82 (7,69)	0,000	<0,001	**<0,001**	0,00 - 0,00
HLA-C*17:01	131 (13,45)	72 (6,75)	2,145	<0,001	**<0,001**	1,59 - 2,90
HLA-C*18:00	0 (0,00)	43 (4,03)	0,000	<0,001	**<0,001**	0,00 - 0,00
HLA-C*18:01	49 (5,03)	8 (0,75)	7,006	<0,001	**<0,001**	3,30 - 14,87

Class I HLA alleles are significantly associated with TB infection

We identified 11 HLA-DPA1, 28 HLA-DPB1, 15 HLA-DQA1, 17 HLA-DQB1, 34 HLA-DRB1, 4 HLA-DRB3, 2 HLA-DRB4, and 2 HLA-DRB5 within our study population (Supplementary Table 2). There were 45 unique class II alleles between TB positive and negative groups. Twenty-seven class II HLA alleles were significantly associated with TB infection (Table 2).

HLA-DQA1*03:01 conferred the strongest risk, with individuals carrying this allele being 23 times more likely to develop TB (p <0,001). HLA-DPB1*04:02 (p <0,001), HLA-DQA1*01:01 (p <0,001), HLA-DQA1*01:02 (p =0,034), HLA-DQA1*05:01 (p <0,001) also conferred risk to developing TB. HLA-DPB140:01 (p = 0.010) was exclusively found in TB-positive individuals, suggesting a possible association with TB susceptibility. Only HLA-DPB1*04:02 (p<0,001), HLA-DQA1*01:01 (p<0,001), and HLA-DQA1*03:01 (p<0,001) remained significant after multiple testing correction.

Several HLA Class II alleles were significantly associated with protection against TB, as they were more frequent in TB-negative individuals: HLA-DQB1*02:02 (p <0,001), HLA-DQB1*03:19 (p <0,001), HLA-DRB1*03:01 (p = 0,003), HLA-DRB1*04:01 (p <0,001), HLA-DRB3*01:01 (p <0,001), HLA-DRB3*02:02 (p <0,001), HLA-DRB3*03:01 (p = 0,013), HLA-DRB4*0101 (p <0,001). Class II HLA alleles that were found exclusively within TB negative individuals included: HLA-DPA1*01:05 (p <0,001), HLA-DPA1*02:07 (p <0,001), HLA-DPB1*10:50 (p <0,001), HLA-DPB1*10:60 (p <0,001), HLA-DPB1*13:10 (p = 0,017), HLA-DPB1*34:01 (p = 0,002), HLA-DPB1*55:01 (p<0,001), HLA-DQA1*01:04 (p <0,001), HLA-DQA1*01:05 (p <0,001), HLA-DQA1*03:03 (p <0,001), HLA-DQA1*05:02 (p <0,001), HLA-DQA1*05:05, (p <0,001), HLA-DRB1*14:54 (p <0,001). After multiple testing comparisons, HLA-DPB1*13:10, HLA-DPB1*34:01, HLA-DQA1*05:01, HLA-DRB1*03:01, HLA-DRB1*14:54, HLA-DRB3*03:01 lost significance.

Table 2: Class II HLA alleles significantly associated with TB infection.

Alleles	TB Positive (%)	TB Negative (%)	OR	p-Value	Adjusted p-Value	CI
HLA-DPA1*01:05	0 (0,00)	9 (1,72)	0,000	<0,001	**0,008**	0,00 - 0,00
HLA-DPA1*02:07	0 (0,00)	8 (1,53)	0,000	<0,001	**0,025**	0,00 - 0,00
HLA-DPB1*04:02	172 (18,70)	9 (1,72)	13,158	<0,001	**<0,001**	6,67 - 25,96
HLA-DPB1*10:50	0 (0,00)	106 (20,23)	0,000	<0,001	**<0,001**	0,00 - 0,00
HLA-DPB1*10:60	0 (0,00)	15 (2,86)	0,000	<0,001	**<0,001**	0,00 - 0,00
HLA-DPB1*13:10	0 (0,00)	4 (0,76)	0,000	0,017	1,000	0,00 - 0,00
HLA-DPB1*34:01	0 (0,00)	6 (1,15)	0,000	0,002	0,540	0,00 - 0,00
HLA-DPB1*40:01	11 (1,20)	0 (0,00)	inf	0,010	1,000	0,00 - inf
HLA-DPB1*55:01	0 (0,00)	10 (1,91)	0,000	<0,001	**0,009**	0,00 - 0,00
HLA-DQA1*01:01	115 (11,69)	18 (3,47)	3,676	<0,001	**<0,001**	2,21 - 6,11
HLA-DQA1*01:02	290 (29,47)	126 (24,32)	1,300	0,034	1,000	1,02 - 1,66
HLA-DQA1*01:04	0 (0,00)	8 (1,54)	0,000	<0,001	**0,047**	0,00 - 0,00
HLA-DQA1*01:05	0 (0,00)	38 (7,34)	0,000	<0,001	**<0,001**	0,00 - 0,00
HLA-DQA1*03:01	117 (11,89)	3 (0,58)	23,166	<0,001	**<0,001**	7,33 - 73,25
HLA-DQA1*03:03	0 (0,00)	55 (10,62)	0,000	<0,001	**<0,001**	0,00 - 0,00
HLA-DQA1*05:01	148 (15,04)	47 (9,07)	1,774	0,001	0,221	1,25 - 2,51
HLA-DQA1*05:02	0 (0,00)	8 (1,54)	0,000	<0,001	**0,047**	0,00 - 0,00
HLA-DQA1*05:05	0 (0,00)	52 (10,04)	0,000	<0,001	**<0,001**	0,00 - 0,00
HLA-DQB1*02:02	29 (3,86)	50 (9,84)	0,367	<0,001	**0,006**	0,23 - 0,59
HLA-DQB1*03:19	28 (3,72)	45 (8,86)	0,398	<0,001	**0,046**	0,24 - 0,65
HLA-DRB1*03:01	40 (4,46)	44 (8,46)	0,506	0,003	0,794	0,32 - 0,79
HLA-DRB1*04:01	9 (1,00)	22 (4,23)	0,230	<0,001	**0,026**	0,10 - 0,50
HLA-DRB1*14:54	0 (0,00)	8 (1,54)	0,000	<0,001	0,077	0,00 - 0,00
HLA-DRB3*01:01	255 (23,48)	124 (34,73)	0,577	<0,001	**0,010**	0,44 - 0,75
HLA-DRB3*02:02	302 (27,81)	153 (42,86)	0,514	<0,001	**<0,001**	0,40 - 0,66
HLA-DRB3*03:01	136 (12,52)	64 (17,93)	0,655	0,013	1,000	0,47 - 0,91
HLA-DRB4*0101	95 (8,50)	83 (63,36)	0,054	<0,001	**<0,001**	0,04 - 0,08

HLA alleles are significantly associated with TB/HIV co-infection among South African Black population

We analyzed HLA alleles associated with TB/HIV co-infection among TB-positive individuals (Supplementary Table 3 and 4). A total of four HLA class I alleles and four HLA class II alleles were shown to significantly associate with either risk or protection against TB/HIV co-infection (Table 3).

HLA-A*31:01 (p=0,038), HLA-DPB1*26:01 (p=0,044), and HLA-DQB1*04:01 (p=0,037) were exclusively found within individuals with only TB, suggesting that they may have a

protective role against co-infection. HLA-A*02:02 (p = 0,040), HLA-DPB1*17:01 (p = 0,046), and HLA-DRB1*15:01 (p = 0,048) were also associated with protection against TB/HIV co-infection. HLA-B*07:02 (p<0,001) and HLA-C*07:02 (p=0,039) were associated with increased risk of TB/HIV co-infection; however, only HLA-B*07:02 (p=0,023) remained significant after multiple comparison testing.

Table 3: Class I and II HLA alleles significantly associated with TB/HIV co-infection.

Alleles	HIV Positive (%)	HIV Negative (%)	OR	p-Value	Adjusted p-Value	CI
HLA-A*02:02	12 (1,26)	8 (3,45)	0,357	0,040	1	0,14 -0,88
HLA-A*31:01	0 (0,00)	2 (0,86)	0,000	0,038	1	0,00 – 0,00
HLA-B*07:02	65 (8,88)	0 (0,00)	inf	<0,001	**0,023**	0,00 - inf
HLA-C*07:02	93 (11,92)	13 (6,70)	1,885	0,039	1	1,03 – 3,44
HLA-DPB1*17:01	16 (2,20)	10 (5,15)	0,415	0,046	1	0,19 - 0,93
HLA-DPB1*26:01	0 (0,00)	2 (1,03)	0,000	0,044	1	0,00 - 0,00
HLA-DQB1*04:01	0 (0,00)	2 (1,37)	0,000	0,037	1	0,00 - 0,00
HLA-DRB1*15:01	4 (0,55)	4 (2,33)	0,233	0,048	1	0,06 - 0,94

*HLA-C*12:03 increases the risk of recurrent TB infections*

Next, we evaluated the association of HLA alleles and the recurrence of TB infection (Supplementary Tables 5 and 6). HLA-C*12:03 was significantly associated with an increased risk of recurrent TB infections (OR = 7.667, p = 0.025, 95% CI: 1.52–38.68). However, this association did not remain statistically significant after multiple testing corrections, suggesting that further validation in larger cohorts is needed.

Discussion

The HLA system is a crucial component of the host immune response. Specific HLA alleles have been significantly associated with susceptibility and resistance to a range of diseases, including TB (116-119, 123). Although numerous case-control studies have investigated HLA polymorphisms associated with TB across various populations, their findings have often been heterogeneous or inconclusive, with risk or protective effects varying by population (116-119). In this study, we examined the association of both class I and class II HLA alleles with

susceptibility to TB infection, TB/HIV co-infection, and TB recurrence in a Black South African population.

We evaluated HLA associations in TB negative (n=534) and TB positive (n=592) Black South Africans. We report 30 HLA class I and 27 HLA class II alleles significantly associated with TB infection; however, after multiple corrections, testing 12 and 19 class I and II alleles were shown to associate with some conferring protection and others increasing susceptibility. Notably, HLA-C*18:01 (p<0,001, OR: 7,006, CI: 3,30 – 14,87), HLA-DQA1*03:01 (p<0,001, OR: 23,116, CI: 7,33 - 73,25), and HLA-DPB1*04:02 (p<0,001, OR: 13,158, CI: 6,67 - 25,96) showed the strongest associations with TB risk. HLA-C*0202, C*1701, HLA-DQA1*0101, HLA-DQA1*0102, and HLA-DQA1*0501 were also shown to be significantly associated with TB infection, while HLA-A*74:00, HLA-B*27:05, and HLA-DRB3*01:01 were strongly protective. The majority of studies thus far, have only identified class II alleles to be associated with TB infection (124-127).

HIV is a major driver of TB infection and accounts for 13% of TB infections globally; however, 75% of all TB/HIV co-infections occur in Africa (128, 129). In South Africa, the epicenter of the HIV epidemic, more than half of all TB cases occur in PLWHIV (98). There is a plethora of evidence regarding the association of HLA alleles with either TB or HIV susceptibility; however, only a few studies identified alleles associated with dual infection (122, 130-133). HLA-A*10, and homozygosity of HLA-DRB1 alleles were shown to associate with the development of TB in PLWHIV from India and Zimbabwe, respectively (130, 131). Among PLWHIV who had a latent TB infection and then developed active TB, Seedat *et. al.,* (2021) found that HLA-A*01:01 (p = 0.049), HLA-A*34:02 (p = 0.049), and HLA-DRB1*11:01 protected against TB infection in PLWHIV (122). HLA-B*41:01 (p = 0.05), HLA-B*42:02 (p = 0.026), HLA-DRB1*04:01 (p = 0.045), and HLA-DRB1*15:01 were shown to increase susceptibility to TB in PLWHIV who were previously latently infected (122). HLA-A*36:01, HLA-B*40:06, DRB1*15:02, HLA-DRB1*04:03:01, HLADRB1*09:01:02, and HLA-DRB1*14:01:03, HLA-DRB1*10, HLA-DQB1*06:01 were also associated with increased risk of TB/HIV infection (130-133).

Our findings were novel. Among the TB-positive cases, we evaluated the association of HLA alleles with HIV co-infection (Supplementary Tables 3 and 4). Eight HLA alleles were significantly associated with dual infection (Table 3). HLA-A*31:01 (p=0,038), HLA-DPB1*26:01 (p=0,044), and HLA-DQB1*04:01 (p=0,037) were exclusively found within individuals with only TB, suggesting that they may have a protective role against co-infection.

HLA-A*02:02 (p = 0,040), HLA-DPB1*17:01 (p = 0,046), and HLA-DRB1*15:01 (p = 0,048) were also associated with protection against TB/HIV co-infection. HLA-B*07:02 (p<0,001) and HLA-C*07:02 (p=0,039) were associated with increased risk of TB/HIV co-infection; however, only HLA-B*07:02 (p=0,023) remained significant after multiple comparison testing.

Individuals with a previous history of TB are at a greater risk of future TB infection than individuals who are TB naïve. The recurrence of TB infection occurs primarily via two main mechanisms: (i) by relapse of the original TB infection or through the re-infection of a different strain of MTB. In high TB prevalence regions, cases of recurrence occur within the first year after treatment completion (134, 135). Within 52 districts in South Africa, recurrence rates ranged from 7,6% and 40% of all confirmed TB cases (136-138). Only two studies have previously evaluated the association of HLA alleles with recurrent TB infection. HLA-DRB1*08:03 and DQB1*06:01 are associated with higher risk of recurrence in Koreans; while in Indonesians HLA-B*18:02, HLA-B*40:01, and HLA-DRB1*11:01 were associated with recurrence (118, 139). We evaluated the association of HLA alleles and the recurrence of TB infection within our South African population (Supplementary Tables 5 and 6). HLA-C*12:03 was significantly associated with an increased risk of recurrent TB infections (OR = 7.667, p = 0.025, 95% CI: 1.52–38.68). However, this association did not remain statistically significant after multiple testing corrections, suggesting that further validation in larger cohorts is needed.

This study has several limitations that should be acknowledged. First, as the analysis was restricted to a South African Black population, the findings may not be generalizable to other populations with different genetic backgrounds. Second, although the sample size was relatively large, low-frequency alleles may have been underpowered to detect true associations, and some significant findings did not withstand correction for multiple testing. Third, the case–control design does not allow us to establish causality, and we did not include functional assays to directly test the impact of the identified alleles on antigen presentation or T cell responses. Future research should aim to replicate these associations in larger and ethnically diverse cohorts, particularly across different African populations, and integrate functional studies such as peptide-binding assays or T cell activation experiments to validate the biological relevance of these alleles. Moreover, incorporating genomic, transcriptomic, and epigenetic data will help elucidate the broader immunogenetic networks influencing TB susceptibility and progression.

Despite the limitations, our study provides novel insights into the role of HLA polymorphisms in susceptibility to TB, TB/HIV co-infection, and TB recurrence in a South African Black population. By identifying both risk and protective alleles, we highlight the importance of host

genetic factors in shaping immune responses to MTB. Future studies should validate these associations in larger and ethnically diverse cohorts and incorporate functional assays to unravel the biological mechanisms underlying HLA-disease interactions. Ultimately, these findings may contribute to precision medicine approaches, inform vaccine design, and guide host-directed interventions aimed at reducing the burden of TB and TB/HIV co-infection in Africa and beyond.

Conclusion

This study provides novel evidence that both HLA class I and class II alleles contribute to risk of TB, TB/HIV co-infection, and the recurrence of TB infection in a South African Black population. We identified several alleles with strong associations, including HLA-C18:01, HLA-DQA103:01, and HLA-DPB104:02 as risk markers, and HLA-A74:00, HLA-B27:05, and HLA-DRB301:01 as protective markers against contracting TB. importantly, our findings expand current knowledge by demonstrating distinct HLA signatures in the context of dual TB/HIV infection and highlight potential genetic contributors to TB recurrence. While further replication and functional validation are required, these results underscore the critical role of host immunogenetics in TB pathogenesis and emphasize the importance of studying genetically diverse African populations. Ultimately, such insights may inform precision medicine approaches, guide vaccine design, and contribute to improved strategies for TB and TB/HIV control in high-burden settings.

1. Organization WH. WHO Global tuberculosis report. 2018.
2. Health SADo. South African National Tuberculosis Management Guidelines. 2014.
3. Chigutsa E, Visser ME, Swart EC, Denti P, Pushpakom S, Egan D, et al. The SLCO1B1 rs4149032 polymorphism is highly prevalent in South Africans and is associated with reduced rifampin concentrations: dosing implications. Antimicrob Agents Chemother. 2011;55(9):4122-7.
4. Naidoo A, Chirehwa M, Ramsuran V, McIlleron H, Naidoo K, Yende-Zuma N, et al. Effects of genetic variability on rifampicin and isoniazid pharmacokinetics in South African patients with recurrent tuberculosis. Pharmacogenomics. 2019;20(4):225-40.
5. Naidoo A, Ramsuran V, Chirehwa M, Denti P, McIlleron H, Naidoo K, et al. Effect of genetic variation in UGT1A and ABCB1 on moxifloxacin pharmacokinetics in South African patients with tuberculosis. Pharmacogenomics. 2018;19(1):17-29.
6. McIlleron H, Wash P, Burger A, Norman J, Folb PI, Smith P. Determinants of rifampin, isoniazid, pyrazinamide, and ethambutol pharmacokinetics in a cohort of tuberculosis patients. Antimicrob Agents Chemother. 2006;50(4):1170-7.
7. Gandhi NR, Nunn P, Dheda K, Schaaf HS, Zignol M, van Soolingen D, et al. Multidrug-resistant and extensively drug-resistant tuberculosis: a threat to global control of tuberculosis. Lancet. 2010;375(9728):1830-43.
8. (WHO) WHO. Global Tuberculosis Report 2023. Geneva: WHO. 2023.
9. Naidoo K, Perumal R, Cox H, Mathema B, Loveday M, Ismail N, et al. The epidemiology, transmission, diagnosis, and management of drug-resistant tuberculosis—lessons from the South African experience. The Lancet Infectious Diseases. 2024;24(9):e559-e75.
10. Franke RM, Gardner ER, Sparreboom A. Pharmacogenetics of drug transporters. Curr Pharm Des. 2010;16(2):220-30.
11. Kerb R. Implications of genetic polymorphisms in drug transporters for pharmacotherapy. Cancer Lett. 2006;234(1):4-33.
12. Glavinas H, Krajcsi P, Cserepes J, Sarkadi B. The role of ABC transporters in drug resistance, metabolism and toxicity. Curr Drug Deliv. 2004;1(1):27-42.
13. Lin JH, Yamazaki M. Role of P-glycoprotein in pharmacokinetics: clinical implications. Clin Pharmacokinet. 2003;42(1):59-98.
14. Borst P, Evers R, Kool M, Wijnholds J. A family of drug transporters: the multidrug resistance-associated proteins. J Natl Cancer Inst. 2000;92(16):1295-302.
15. Jonker JW, Buitelaar M, Wagenaar E, Van Der Valk MA, Scheffer GL, Scheper RJ, et al. The breast cancer resistance protein protects against a major chlorophyll-derived dietary phototoxin and protoporphyria. Proc Natl Acad Sci U S A. 2002;99(24):15649-54.
16. Sparreboom A, van Asperen J, Mayer U, Schinkel AH, Smit JW, Meijer DK, et al. Limited oral bioavailability and active epithelial excretion of paclitaxel (Taxol) caused by P-glycoprotein in the intestine. Proc Natl Acad Sci U S A. 1997;94(5):2031-5.
17. Fromm MF. The influence of MDR1 polymorphisms on P-glycoprotein expression and function in humans. Adv Drug Deliv Rev. 2002;54(10):1295-310.
18. Lockhart AC, Tirona RG, Kim RB. Pharmacogenetics of ATP-binding cassette transporters in cancer and chemotherapy. Mol Cancer Ther. 2003;2(7):685-98.
19. Marzolini C, Paus E, Buclin T, Kim RB. Polymorphisms in human MDR1 (P-glycoprotein): recent advances and clinical relevance. Clin Pharmacol Ther. 2004;75(1):13-33.

20. Bruckmueller H, Cascorbi I. ABCB1, ABCG2, ABCC1, ABCC2, and ABCC3 drug transporter polymorphisms and their impact on drug bioavailability: what is our current understanding? Expert Opin Drug Metab Toxicol. 2021;17(4):369-96.
21. Khan A, Abbas M, Verma S, Verma S, Rizvi AA, Haider F, et al. Genetic Variants and Drug Efficacy in Tuberculosis: A Step toward Personalized Therapy. Glob Med Genet. 2022;9(2):90-6.
22. Ahmed S, Zhou Z, Zhou J, Chen SQ. Pharmacogenomics of Drug Metabolizing Enzymes and Transporters: Relevance to Precision Medicine. Genomics Proteomics Bioinformatics. 2016;14(5):298-313.
23. Fung KL, Gottesman MM. A synonymous polymorphism in a common MDR1 (ABCB1) haplotype shapes protein function. Biochim Biophys Acta. 2009;1794(5):860-71.
24. Gomez F, Hirbo J, Tishkoff SA. Genetic variation and adaptation in Africa: implications for human evolution and disease. Cold Spring Harb Perspect Biol. 2014;6(7):a008524.
25. Saiz-Rodriguez M, Valdez-Acosta S, Borobia AM, Burgueno M, Galvez-Mugica MA, Acero J, et al. Influence of Genetic Polymorphisms on the Response to Tramadol, Ibuprofen, and the Combination in Patients With Moderate to Severe Pain After Dental Surgery. Clin Ther. 2021;43(5):e86-e102.
26. Mizuarai S, Aozasa N, Kotani H. Single nucleotide polymorphisms result in impaired membrane localization and reduced atpase activity in multidrug transporter ABCG2. Int J Cancer. 2004;109(2):238-46.
27. Sileshi T MG, Makonnen E, Aklillu E. Effect of Genetic Variations in Drug-Metabolizing Enzymes and Drug Transporters on the Pharmacokinetics of Rifamycins: A Systematic Review. Pharmgenomics Pers Med 2022;15:561-571. 2022.
28. Huerta-Garcia AP, Medellin-Garibay SE, Salazar-Gonzalez RA, Ortiz-Alvarez A, Magana-Aquino M, Rodriguez-Pinal CJ, et al. Anthropometric and Genetic Factors Associated With the Exposure of Rifampicin and Isoniazid in Mexican Patients With Tuberculosis. Ther Drug Monit. 2019;41(5):648-56.
29. Ramesh K, Hemanth Kumar AK, Kannan T, Vijayalakshmi R, Sudha V, Manohar Nesakumar S, et al. SLCO1B1 gene polymorphisms do not influence plasma rifampicin concentrations in a South Indian population. Int J Tuberc Lung Dis. 2016;20(9):1231-5.
30. Ammour A, Aberkane M, Boudjema A, Boughrara W, Benchekor SM. The Frequencies distribution of CYP3A5 rs776746 and ABCB1 rs1045642 polymorphisms in the west Algerian population and relationships with pharmacogenetics. Afr Health Sci. 2024;24(1):307-12.
31. Swart M, Ren Y, Smith P, Dandara C. ABCB1 4036A>G and 1236C>T Polymorphisms Affect Plasma Efavirenz Levels in South African HIV/AIDS Patients. Front Genet. 2012;3:236.
32. Hediger MA, Romero MF, Peng JB, Rolfs A, Takanaga H, Bruford EA. The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteinsIntroduction. Pflugers Arch. 2004;447(5):465-8.
33. Hagenbuch B, Meier PJ. The superfamily of organic anion transporting polypeptides. Biochim Biophys Acta. 2003;1609(1):1-18.
34. Hsiang B, Zhu Y, Wang Z, Wu Y, Sasseville V, Yang WP, et al. A novel human hepatic organic anion transporting polypeptide (OATP2). Identification of a liver-specific human organic anion transporting polypeptide and identification of rat and human

hydroxymethylglutaryl-CoA reductase inhibitor transporters. J Biol Chem. 1999;274(52):37161-8.
35. Mohzari YA, Musawa MA, Alrashed A, Zaeri B, Damfu N, Cahusac P, et al. The clinical impact of rifampicin-based anti-TB therapy and tenofovir alafenamide-containing ARV regimen drug Interaction in people living with HIV: Case series report. J Infect Public Health. 2023;16(12):2026-30.
36. Acocella G. Clinical pharmacokinetics of rifampicin. Clin Pharmacokinet. 1978;3(2):108-27.
37. Sanni S, Wachinou A, Merle C, Baba-Moussa L, Affolabi D. SLCO1B1 , NAT2 Polymorphisms and Pharmacokinetic Variability of Rifampicin and Isoniazid in Tuberculosis Patients from Sub-Saharan Africa. Journal of Tuberculosis Research. 2022;10:111-23.
38. Stieger B, Hagenbuch B. Organic anion-transporting polypeptides. Curr Top Membr. 2014;73:205-32.
39. Kalliokoski A, Niemi M. Impact of OATP transporters on pharmacokinetics. Br J Pharmacol. 2009;158(3):693-705.
40. Litjens CHC, van den Heuvel J, Russel FGM, Aarnoutse RE, Te Brake LHM, Koenderink JB. Rifampicin Transport by OATP1B1 Variants. Antimicrob Agents Chemother. 2020;64(10).
41. Shu Y, Brown C, Castro RA, Shi RJ, Lin ET, Owen RP, et al. Effect of genetic variation in the organic cation transporter 1, OCT1, on metformin pharmacokinetics. Clin Pharmacol Ther. 2008;83(2):273-80.
42. Lin L, Yee SW, Kim RB, Giacomini KM. SLC transporters as therapeutic targets: emerging opportunities. Nat Rev Drug Discov. 2015;14(8):543-60.
43. Konig J, Seithel A, Gradhand U, Fromm MF. Pharmacogenomics of human OATP transporters. Naunyn Schmiedebergs Arch Pharmacol. 2006;372(6):432-43.
44. Dompreh A, Tang X, Zhou J, Yang H, Topletz A, Adu Ahwireng E, et al. Effect of Genetic Variation of NAT2 on Isoniazid and SLCO1B1 and CES2 on Rifampin Pharmacokinetics in Ghanaian Children with Tuberculosis. Antimicrob Agents Chemother. 2018;62(3).
45. Weiner M, Peloquin C, Burman W, Luo CC, Engle M, Prihoda TJ, et al. Effects of tuberculosis, race, and human gene SLCO1B1 polymorphisms on rifampin concentrations. Antimicrob Agents Chemother. 2010;54(10):4192-200.
46. Hennig S, Naiker S, Reddy T, Egan D, Kellerman T, Wiesner L, et al. Effect of SLCO1B1 Polymorphisms on Rifabutin Pharmacokinetics in African HIV-Infected Patients with Tuberculosis. Antimicrob Agents Chemother. 2016;60(1):617-20.
47. Medwid S, Deckert R, Gryn SE, Kim RB. SLCO1B1 variants in a patient of African ancestry presenting with rosuvastatin-induced rhabdomyolysis: A case report. Br J Clin Pharmacol. 2025;91(1):232-5.
48. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68-74.
49. Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV, et al. A "silent" polymorphism in the MDR1 gene changes substrate specificity. Science. 2007;315(5811):525-8.
50. Woodahl EL, Hingorani SR, Wang J, Guthrie KA, McDonald GB, Batchelder A, et al. Pharmacogenomic associations in ABCB1 and CYP3A5 with acute kidney injury and

chronic kidney disease after myeloablative hematopoietic cell transplantation. Pharmacogenomics J. 2008;8(4):248-55.
51. Mpeta B, Kampira E, Castel S, Mpye KL, Soko ND, Wiesner L, et al. Differences in genetic variants in lopinavir disposition among HIV-infected Bantu Africans. Pharmacogenomics. 2016;17(7):679-90.
52. Yu N, Chen FC, Ota S, Jorde LB, Pamilo P, Patthy L, et al. Larger genetic differences within africans than between Africans and Eurasians. Genetics. 2002;161(1):269-74.
53. Wonkam A, Adeyemo A. Leveraging our common African origins to understand human evolution and health. Cell Genom. 2023;3(3):100278.
54. Zhang S, Xiao K, He X, Hao M, Ma Y, Tian Y, et al. A novel efficient algorithm for common variants genotyping from low-coverage sequencing data. Computational and Structural Biotechnology Reports. 2025;2:100046.
55. Alagarasu K, Selvaraj P, Swaminathan S, Raghavan S, Narendran G, Narayanan PR. Mannose binding lectin gene variants and susceptibility to tuberculosis in HIV-1 infected patients of South India. Tuberculosis (Edinb). 2007;87(6):535-43.
56. Selvaraj P, Alagarasu K, Swaminathan S, Harishankar M, Narendran G. CD209 gene polymorphisms in South Indian HIV and HIV-TB patients. Infect Genet Evol. 2009;9(2):256-62.
57. Indrati AR, Sumarpo A, Haryanto J, Rosmiati NMD, Munaya S, Turbawaty DK, et al. Identification of cytokine signatures in HIV-infected individuals with and without Mycobacterium tuberculosis co-infection. Biomed Rep. 2024;21(3):131.
58. Tamene W, Marconi VC, Abebe M, Wassie L, Belay Y, Kebede A, et al. Differential expression of chemokine receptors on monocytes in TB and HIV S. Heliyon. 2023;9(6):e17202.
59. McCombie WR, McPherson JD. Future Promises and Concerns of Ubiquitous Next-Generation Sequencing. Cold Spring Harb Perspect Med. 2019;9(9).
60. Petrovski S, Fellay J, Shianna KV, Carpenetti N, Kumwenda J, Kamanga G, et al. Common human genetic variants and HIV-1 susceptibility: a genome-wide survey in a homogeneous African population. Aids. 2011;25(4):513-8.
61. Shevchenko AK, Zhernakova DV, Malov SV, Komissarov A, Kolchanova SM, Tamazian G, et al. Genome-wide association study reveals genetic variants associated with HIV-1C infection in a Botswana study population. Proceedings of the National Academy of Sciences. 2021;118(47):e2107830118.
62. An P, Winkler CA. Host genes associated with HIV/AIDS: advances in gene discovery. Trends in Genetics. 2010;26(3):119-31.
63. Williamson C, Loubser SA, Brice B, Joubert G, Smit T, Thomas R, et al. Allelic frequencies of host genetic variants influencing susceptibility to HIV-1 infection and disease in South African populations. Aids. 2000;14(4):449-51.
64. Galvani AP, Novembre J. The evolutionary history of the CCR5-Δ32 HIV-resistance mutation. Microbes and Infection. 2005;7(2):302-9.
65. Sibomana O. Genetic Diversity Landscape in African Population: A Review of Implications for Personalized and Precision Medicine. Pharmgenomics Pers Med. 2024;17:487-96.
66. Sosnay PR, Siklosi KR, Van Goor F, Kaniecki K, Yu H, Sharma N, et al. Defining the disease liability of variants in the cystic fibrosis transmembrane conductance regulator gene. Nat Genet. 2013;45(10):1160-7.

67. Castellani C, Cuppens H, Macek M, Jr., Cassiman JJ, Kerem E, Durie P, et al. Consensus on the use and interpretation of cystic fibrosis mutation analysis in clinical practice. J Cyst Fibros. 2008;7(3):179-96.
68. Marco A, Montales MT, Mittadodla P, Mukasa L, Bhaskar N, Bates J, et al. Tuberculosis reinfection in a pregnant cystic fibrosis patient. Respiratory Medicine Case Reports. 2015;16:57-9.
69. Mehrgou A, Akouchekian M. The importance of BRCA1 and BRCA2 genes mutations in breast cancer development. Med J Islam Repub Iran. 2016;30:369.
70. Moosazadeh M, Bahrampour A, Nasehi M, Khanjani N. The incidence of recurrence of tuberculosis and its related factors in smear-positive pulmonary tuberculosis patients in Iran: A retrospective cohort study. Lung India. 2015;32(6):557-60.
71. Vree M, Huong NT, Duong BD, Sy DN, Van LN, Hung NV, et al. Survival and relapse rate of tuberculosis patients who successfully completed treatment in Vietnam. Int J Tuberc Lung Dis. 2007;11(4):392-7.
72. Datiko DG, Lindtjørn B. Tuberculosis recurrence in smear-positive patients cured under DOTS in southern Ethiopia: retrospective cohort study. BMC Public Health. 2009;9:348.
73. Chang SY, Chen ML, Lee MR, Liang YC, Lu TP, Wang JY, et al. SP110 Polymorphisms Are Genetic Markers for Vulnerability to Latent and Active Tuberculosis Infection in Taiwan. Dis Markers. 2018;2018:4687380.
74. Mabunda N, Alvarado-Arnez LE, Vubil A, Mariamo A, Pacheco AG, Jani IV, et al. Gene polymorphisms in patients with pulmonary tuberculosis from Mozambique. Mol Biol Rep. 2015;42(1):71-6.
75. Mroziewicz M, Tyndale RF. Pharmacogenetics: a tool for identifying genetic factors in drug dependence and response to treatment. Addict Sci Clin Pract. 2010;5(2):17-29.
76. Parvez MM, Kaisar N, Shin HJ, Lee YJ, Shin JG. Comprehensive Substrate Characterization of 22 Antituberculosis Drugs for Multiple Solute Carrier (SLC) Uptake Transporters In Vitro. Antimicrob Agents Chemother. 2018;62(9).
77. Alwi ZB. The Use of SNPs in Pharmacogenomics Studies. Malays J Med Sci. 2005;12(2):4-12.
78. Tiberi S, Munoz-Torrico M, Duarte R, Dalcolmo M, D'Ambrosio L, Migliori GB. New drugs and perspectives for new anti-tuberculosis regimens. Pulmonology. 2018;24(2):86-98.
79. da Rocha JEB, Othman H, Botha G, Cottino L, Twesigomwe D, Ahmed S, et al. The Extent and Impact of Variation in ADME Genes in Sub-Saharan African Populations. Front Pharmacol. 2021;12:634016.
80. Berry KM, Rodriguez CA, Berhanu RH, Ismail N, Mvusi L, Long L, et al. Treatment outcomes among children, adolescents, and adults on treatment for tuberculosis in two metropolitan municipalities in Gauteng Province, South Africa. BMC Public Health. 2019;19(1):973.
81. Omotoso OE, Teibo JO, Atiba FA, Oladimeji T, Adebesin AO, Babalghith AO. Bridging the genomic data gap in Africa: implications for global disease burdens. Global Health. 2022;18(1):103.
82. Di Martino MT, Arbitrio M, Leone E, Guzzi PH, Rotundo MS, Ciliberto D, et al. Single nucleotide polymorphisms of ABCC5 and ABCG1 transporter genes correlate to

irinotecan-associated gastrointestinal toxicity in colorectal cancer patients: a DMET microarray profiling study. Cancer Biol Ther. 2011;12(9):780-7.
83. Fan S, Spence JP, Feng Y, Hansen MEB, Terhorst J, Beltrame MH, et al. Whole-genome sequencing reveals a complex African population demographic history and signatures of local adaptation. Cell. 2023;186(5):923-39.e14.
84. Barbalic M, Schwartz GL, Chapman AB, Turner ST, Boerwinkle E. Kininogen gene (KNG) variation has a consistent effect on aldosterone response to antihypertensive drug therapy: the GERA study. Physiol Genomics. 2009;39(1):56-60.
85. Scionti F, Agapito G, Caracciolo D, Riillo C, Grillone K, Cannataro M, et al. Risk Alleles for Multiple Myeloma Susceptibility in ADME Genes. Cells. 2022;11(2).
86. Hertzog A, Selvanathan A, Farnsworth E, Tchan M, Adams L, Lewis K, et al. Intronic variants in inborn errors of metabolism: Beyond the exome. Front Genet. 2022;13:1031495.
87. Organization WH. Global tuberculosis report 2024. 2024.
88. Atake E-H. Health system productivity in Sub-Saharan Africa: tuberculosis control in high burden countries. Cost Effectiveness and Resource Allocation. 2023;21(1):90.
89. Alele FO, C. FR, I. ET, and Leggat P. Occupational tuberculosis in healthcare workers in Sub-Saharan Africa: A systematic review. Archives of Environmental & Occupational Health. 2019;74(3):95-108.
90. Mbuya AW, Mboya IB, Semvua HH, Mamuya SH, Msuya SE. Prevalence and factors associated with tuberculosis among the mining communities in Mererani, Tanzania. PLoS One. 2023;18(3):e0280396.
91. Hayward S, Harding RM, McShane H, Tanner R. Factors influencing the higher incidence of tuberculosis among migrants and ethnic minorities in the UK. F1000Research. 2018;7:461.
92. Nahid P, Horne DJ, Jarlsberg LG, Reiner AP, Osmond D, Hopewell PC, et al. Racial differences in tuberculosis infection in United States communities: the coronary artery risk development in young adults study. Clinical infectious diseases. 2011;53(3):291-4.
93. Kim S, Cohen T, Horsburgh Jr CR, Miller JW, Hill AN, Marks SM, et al. Trends, mechanisms, and racial/ethnic differences of tuberculosis incidence in the US-born population aged 50 years or older in the United States. Clinical Infectious Diseases. 2022;74(9):1594-603.
94. Olivier C, Luies L. WHO goals and beyond: managing HIV/TB co-infection in South Africa. SN Comprehensive Clinical Medicine. 2023;5(1):251.
95. Zondo NN. Environmental Factors Associated with Tuberculosis Incidence and Mortality in Tshwane District, Gauteng Province, South Africa: University of Johannesburg (South Africa); 2022.
96. Cioboata R, Balteanu MA, Osman A, Vlasceanu SG, Zlatian OM, Mitroi DM, et al. Coinfections in Tuberculosis in Low-and Middle-Income Countries: Epidemiology, Clinical Implications, Diagnostic Challenges, and Management Strategies—A Narrative Review. Journal of Clinical Medicine. 2025;14(7):2154.
97. Ganesan K, Mwesigwa R, Dear N, Esber AL, Reed D, Kibuuka H, et al. Epidemiology of Tuberculosis Among People Living With HIV in the African Cohort Study From 2013 to 2021. J Acquir Immune Defic Syndr. 2023;92(5):359-69.

98. Kubjane M, Osman M, Boulle A, Johnson LF. The impact of HIV and tuberculosis interventions on South African adult tuberculosis trends, 1990-2019: a mathematical modeling analysis. Int J Infect Dis. 2022;122:811-9.
99. Walker NF, Meintjes G, Wilkinson RJ. HIV-1 and the immune response to TB. Future Virol. 2013;8(1):57-80.
100. Dookie N, Naidoo K. Insights into Recurrent Tuberculosis: Relapse Versus Reinfection and Related Risk Factors. In: Kayembe J-MN, editor. Tuberculosis. Rijeka: IntechOpen; 2018.
101. Shanmugam S, Bachmann NL, Martinez E, Menon R, Narendran G, Narayanan S, et al. Whole genome sequencing based differentiation between re-infection and relapse in Indian patients with tuberculosis recurrence, with and without HIV co-infection. International Journal of Infectious Diseases. 2021;113:S43-S7.
102. Varshney D, Singh S, Sinha E, Mohanty KK, Kumar S, Barik SK, et al. Systematic review and meta-analysis of human Toll-like receptors genetic polymorphisms for susceptibility to tuberculosis infection. Cytokine. 2022;152:155791.
103. Oki NO, Motsinger-Reif AA, Antas PRZ, Levy S, Holland SM, Sterling TR. Novel human genetic variants associated with extrapulmonary tuberculosis: a pilot genome wide association study. BMC Research Notes. 2011;4(1):28.
104. Wang F, Huang G, Shen L, Peng Y, Sha W, Chen ZW, et al. Genetics and Functional Mechanisms of STAT3 Polymorphisms in Human Tuberculosis. Frontiers in Cellular and Infection Microbiology. 2021;Volume 11 - 2021.
105. Phelan J, Gomez-Gonzalez PJ, Andreu N, Omae Y, Toyo-Oka L, Yanai H, et al. Genome-wide host-pathogen analyses reveal genetic interaction points in tuberculosis disease. Nature communications. 2023;14(1):549.
106. Naranbhai V. The Role of Host Genetics (and Genomics) in Tuberculosis. Microbiology Spectrum. 2016;4(5):10.1128/microbiolspec.tbtb2-0011-2016.
107. Schurz H, Naranbhai V, Yates TA, Gilchrist JJ, Parks T, Dodd PJ, et al. Multi-ancestry meta-analysis of host genetic susceptibility to tuberculosis identifies shared genetic architecture. eLife. 2024;13:e84394.
108. Abel L, El-Baghdadi J, Bousfiha AA, Casanova JL, Schurr E. Human genetics of tuberculosis: a long and winding road. Philos Trans R Soc Lond B Biol Sci. 2014;369(1645):20130428.
109. Phelan J, Gomez-Gonzalez PJ, Andreu N, Omae Y, Toyo-Oka L, Yanai H, et al. Genome-wide host-pathogen analyses reveal genetic interaction points in tuberculosis disease. Nature Communications. 2023;14(1):549.
110. Sveinbjornsson G, Gudbjartsson DF, Halldorsson BV, Kristinsson KG, Gottfredsson M, Barrett JC, et al. HLA class II sequence variants influence tuberculosis risk in populations of European ancestry. Nature genetics. 2016;48(3):318-22.
111. Souza de Lima D, Morishi Ogusku M, Porto dos Santos M, de Melo Silva CM, Alves de Almeida V, Assumpção Antunes I, et al. Alleles of HLA-DRB1* 04 associated with pulmonary tuberculosis in Amazon Brazilian population. PLoS One. 2016;11(2):e0147543.
112. Monos DS, Winchester RJ. The major histocompatibility complex. Clinical immunology: Elsevier; 2019. p. 79-92. e1.
113. Robinson J, Barker DJ, Georgiou X, Cooper MA, Flicek P, Marsh SG. Ipd-imgt/hla database. Nucleic acids research. 2020;48(D1):D948-D55.

114. Dendrou CA, Petersen J, Rossjohn J, Fugger L. HLA variation and disease. Nature Reviews Immunology. 2018;18(5):325-39.
115. Sanchez-Mazas A. A review of HLA allele and SNP associations with highly prevalent infectious diseases in human populations. Swiss medical weekly. 2020;150(1516):w20214-w.
116. Amirzargar AA, Yalda A, Hajabolbaghi M, Khosravi F, Jabbari H, Rezaei N, et al. The association of HLA-DRB, DQA1, DQB1 alleles and haplotype frequency in Iranian patients with pulmonary tuberculosis. Int J Tuberc Lung Dis. 2004;8(8):1017-21.
117. Dubaniewicz A, Moszkowska G, Szczerkowska Z, Hoppe A. Analysis of DQB1 allele frequencies in pulmonary tuberculosis: preliminary report. Thorax. 2003;58(10):890-1.
118. Kim HS, Park MH, Song EY, Park H, Kwon SY, Han SK, et al. Association of HLA-DR and HLA-DQ Genes With Susceptibility to Pulmonary Tuberculosis in Koreans: Preliminary Evidence of Associations With Drug Resistance, Disease Severity, and Disease Recurrence. Human Immunology. 2005;66(10):1074-81.
119. Goldfeld AE, Delgado JC, Thim S, Bozon MV, Uglialoro AM, Turbay D, et al. Association of an HLA-DQ allele with clinical tuberculosis. Jama. 1998;279(3):226-8.
120. Altfeld M, Kalife ET, Qi Y, Streeck H, Lichterfeld M, Johnston MN, et al. HLA alleles associated with delayed progression to AIDS contribute strongly to the initial CD8+ T cell response against HIV-1. PLoS medicine. 2006;3(10):e403.
121. Ramirez de Arellano E, Diez-Fuertes F, Aguilar F, de la Torre Tarazona HE, Sanchez-Lara S, Lao Y, et al. Novel association of five HLA alleles with HIV-1 progression in Spanish long-term non progressor patients. PLoS One. 2019;14(8):e0220459.
122. Seedat F, James I, Loubser S, Waja Z, Mallal SA, Hoffmann C, et al. Human leukocyte antigen associations with protection against tuberculosis infection and disease in human immunodeficiency virus-1 infected individuals, despite household tuberculosis exposure and immune suppression. Tuberculosis. 2021;126:102023.
123. Liu B, Shao Y, Fu R. Current research status of HLA in immune-related diseases. Immunity, Inflammation and Disease. 2021;9(2):340-50.
124. Chihab LY, Kuan R, Phillips EJ, Mallal SA, Rozot V, Davis MM, et al. Expression of specific HLA class II alleles is associated with an increased risk for active tuberculosis and a distinct gene expression profile. HLA. 2023;101(2):124-37.
125. Croock DA, Swart Y, Schurz H, Petersen DC, Möller M, Uren C. Confirmation of HLA-II associations with TB susceptibility in admixed African samples. eLife. 2025;13:RP99200.
126. Wamala D, Buteme HK, Kirimunda S, Kallenius G, Joloba M. Association between human leukocyte antigen class II and pulmonary tuberculosis due to mycobacterium tuberculosis in Uganda. BMC Infectious Diseases. 2016;16(1):23.
127. Duarte R, Carvalho C, Pereira C, Bettencourt A, Carvalho A, Villar M, et al. HLA class II alleles as markers of tuberculosis susceptibility and resistance. Revista Portuguesa de Pneumologia (English Edition). 2011;17(1):15-9.
128. Bruchfeld J, Correia-Neves M, Källenius G. Tuberculosis and HIV Coinfection. Cold Spring Harbor Perspectives in Medicine. 2015;5(7).
129. Bruyn ED, Wilkinson RJ. The Immune Interaction between HIV-1 Infection and Mycobacterium tuberculosis. Microbiology Spectrum. 2016;4(6):10.1128/microbiolspec.tbtb2-0012-2016.

130. Shankarkumar U, Pawar A, Prabu G, Ghosh K. Role of HLA class I (HLA-A, B) and HLA class II (HLA-DRB, DQB) in HIV-1 patients with and without pulmonary tuberculosis. JAIDS Journal of Acquired Immune Deficiency Syndromes. 2009;51(5):640-1.
131. Louie LG, Hartogensis WE, Jackman RP, Schultz KA, Zijenah LS, Yiu CH-Y, et al. Mycobacterium tuberculosis/HIV-1 coinfection and disease: role of human leukocyte antigen variation. The Journal of infectious diseases. 2004;189(6):1084-90.
132. Raghavan S, Selvaraj P, Swaminathan S, Alagarasu K, Narendran G, Narayanan PR. Haplotype analysis of HLA-A, -B antigens and -DRB1 alleles in south Indian HIV-1-infected patients with and without pulmonary tuberculosis. International Journal of Immunogenetics. 2009;36(3):129-33.
133. Raghavan S, Selvaraj P, Swaminathan S, Narendran G. association of HLA-A* 1101 with resistance and B* 4006 with susceptibility to HIV and HIV-TB: an in silico analysis of promiscuous T cell epitopes. AIDS research and human retroviruses. 2009;25(10):1023-8.
134. Guerra-Assunção JA, Houben RMGJ, Crampin AC, Mzembe T, Mallard K, Coll F, et al. Recurrence due to Relapse or Reinfection With Mycobacterium tuberculosis: A Whole-Genome Sequencing Approach in a Large, Population-Based Cohort With a High HIV Infection Prevalence and Active Follow-up. The Journal of Infectious Diseases. 2014;211(7):1154-63.
135. Mirsaeidi M, Sadikot RT. Patients at High Risk of Tuberculosis Recurrence. The International Journal of Mycobacteriology. 2018;7(1):1-6.
136. Marx FM, Cohen T, Lombard C, Hesseling AC, Dlamini SS, Beyers N, et al. Notification of relapse and other previously treated tuberculosis in the 52 health districts of South Africa. The International Journal of Tuberculosis and Lung Disease. 2019;23(8):891-9.
137. Wademan DT, Tila M, Melleh G, Helen A, Kwame S, Linda M, et al. 'TB is a disease which hides in the body': Qualitative data on conceptualisations of tuberculosis recurrence among patients in Zambia and South Africa. Global Public Health. 2022;17(8):1713-27.
138. Hermans SM, Zinyakatira N, Caldwell J, Cobelens FGJ, Boulle A, Wood R. High Rates of Recurrent Tuberculosis Disease: A Population-level Cohort Study. Clinical Infectious Diseases. 2020;72(11):1919-26.
139. Yuliwulandari R, Sachrowardi Q, Nakajima H, Kashiwase K, Hirayasu K, Mabuchi A, et al. Association of HLA-A, -B, and -DRB1 with pulmonary tuberculosis in western Javanese Indonesia. Human Immunology. 2010;71(7):697-701.

Appendix B

Preface:

The preceding chapters of this thesis have highlighted the remarkable genetic diversity unique to African genomes and reiterated the underrepresentation of African individuals in global genomic databases. This gap in genetic databases severely limits our understanding of human genetic variation and hinders the development of equitable scientific insights and precision medicine strategies. The growing installation-base of next-generation sequencing (NGS) platforms across the African continent have indeed played a crucial role in genomic surveillance during the COVID-19 pandemic which has demonstrated the continent's ability to generate high-quality genomic data at scale locally. With this momentum, it would be beneficial to further strengthen genomic expertise across Africa, that will enable scientists to utilize these technologies and contribute to filling long-standing gaps in genetic data.

To address the challenge of underrepresentation in both non-human bio-diversity and human genomic datasets, well-coordinated capacity-building efforts are essential. The Africa Biogenome Project (AfricaBP) has emerged as a key initiative in this regard. With the implementation of annual training workshops held across the continent, AfricaBP is equipping both emerging and established researchers with the practical and bioinformatic skills needed to advance genomics research within African institutions. These efforts will be vital to support the generation of high-quality genomic data and enable African scientists to drive future research studies that reflect the continent's unique biodiversity and population structure. This chapter builds upon these themes, highlighting current progress made in a broader vision of sustainable, locally rooted genomic capacity that has the potential to transform research landscapes and help bridge the gaps in the global representation of genetic data.

Establishing African genomics and bioinformatics programs through annual regional workshops

Sharaf A, Nesengani LT, Hayah I, Kuja JO, Mdyogolo S, Omotoriogun TC, Odogwu BA, Beedessee G, Smith RM, Barakat A, Moila AM, El Hamouchi A, Benkahla A, Boukteb A, Elmouhtadi A, Mafwila AL, Abushady AM, Elsherif AK, Ahmed B, Wairuri C, Ndiribe CC, Ebuzome C, Kinnear CJ, Ndlovu DF, Iraqi D, El Fahime E, Assefa E, Ouardi F, Belharfi FZ, Tmimi FZ, Markey FB, Radouani F, Zeukeng F, Mvumbi GL, Ganesan H, Hanachi M, Nigussie H, Charoute H, Benamri I, Mkedder I, Haddadi I, Meftah-Kadmiri I, Mubiru JF, Domelevo Entfellner JK, Rokani JB, Ogwang J, Daiga JB, Omumbo J, Ideozu JE, Errafii K, Labuschagne K, Komi KK, Tonfack LB, Hadjeras L, Ramantswana M, Chaisi M, Botes MW, Kilian M, Kvas M, Melloul M, Chaouch M, Khyatti M, Abdo M, Phasha-Muchemenye M, Hijri M, Mediouni MR, Hassan MA, Piro M, Mwale M, Maaloum M, Mavhunga M, Olivier NA, Aminou O, Arbani O, Souiai O, Djocgoue PF, Mentag R, Zipfel RD, Tata RB, Megnekou R, Muzemil S, Paez S, Salifu SP, Kagame SP, Selka S, Edwards S, Gaouar SBS, Reda SRA, Fellahi S, Khayi S, Ayed S, Madisha T, **Sahil T**, Udensi OU, Ras V, Ezebuiro V, Duru VC, David X, Geberemichael Y, Tchiechoua YH, Mungloo-Dilmohamud Z, Chen Z, Happi C, Kariuki T, Ziyomo C, Djikeng A, Badaoui B, Mapholi N, Muigai A, Osuji JO, Ebenezer TE. Establishing African genomics and bioinformatics programs through annual regional workshops. Nat Genet. 2024 Aug;56(8):1556-1565. doi: 10.1038/s41588-024-01807-6. Epub 2024 Jul 8. **PMID: 38977855.**

nature genetics

Perspective

https://doi.org/10.1038/s41588-024-01807-6

Establishing African genomics and bioinformatics programs through annual regional workshops

Received: 27 January 2024

Accepted: 22 May 2024

Published online: 8 July 2024

Check for updates

A list of authors and their affiliations appears at the end of the paper

The African BioGenome Project (AfricaBP) Open Institute for Genomics and Bioinformatics aims to overcome barriers to capacity building through its distributed African regional workshops and prioritizes the exchange of grassroots knowledge and innovation in biodiversity genomics and bioinformatics. In 2023, we implemented 28 workshops on biodiversity genomics and bioinformatics, covering 11 African countries across the 5 African geographical regions. These regional workshops trained 408 African scientists in hands-on molecular biology, genomics and bioinformatics techniques as well as the ethical, legal and social issues associated with acquiring genetic resources. Here, we discuss the implementation of transformative strategies, such as expanding the regional workshop model of AfricaBP to involve multiple countries, institutions and partners, including the proposed creation of an African digital database with sequence information relating to both biodiversity and agriculture. This will ultimately help create a critical mass of skilled genomics and bioinformatics scientists across Africa.

The African BioGenome Project (AfricaBP) (https://africanbiogenome.org) is a coordinated Pan-African effort established in 2021 to sequence the genomes of 105,000 endemic biological species (plants, animals, fungi, protists and other eukaryotes) to improve food systems, conservation, and data sharing and benefits[1]. The Open Institute for Genomics and Bioinformatics is the knowledge exchange program for AfricaBP, which aims to overcome barriers to capacity-building efforts such as curriculum development using local languages and biological species, support African research and academic institutions that have fewer resources for their research needs, and enable scientific entrepreneurial opportunities to support Africa's bioeconomy through its distributed African regional workshops on biodiversity genomics and bioinformatics[2]. Here, biodiversity genomics is defined as the application of genomics in the assessment, conservation and improvement of biological diversity[3–5].

One of the priorities of the AfricaBP Open Institute is the exchange of knowledge at the grassroots level, as well as lowering visa and travel barriers that limit participation in scientific engagements[2,6,7]. For example, although bioinformatics in Africa was already established in 1996 through the South African National Bioinformatics Institute and more extensive capacity-building efforts have subsequently been made by H3ABioNet (the bioinformatics network for the Human Heredity and Health in Africa consortium) and the Alliance for Accelerated Crop Improvement in Africa[8–10], there is still much work to be done in scaling bioinformatics capacity building at the grassroots level in the biodiversity genomics and bioinformatics area. In 2023, the AfricaBP Open Institute implemented 28 awareness and hands-on practical workshops on biodiversity genomics and bioinformatics in 11 African countries across the 5 African geographical regions (Fig. 1). The fundamental objectives were to bring relevant technologies closer to African researchers by integrating awareness building with hands-on practical workshops and to exemplify the principles of public–private partnerships (PPPs) in delivering genomics and bioinformatics workshops in low-resource settings. These regional workshops had 3,788 registered participants with 408 African scientists trained hands-on.

Here, we discuss the implementation of transformative strategies aimed at genomic innovations and knowledge exchange at the grassroots level by deploying the AfricaBP Open Institute regional workshops through PPPs. We also propose measures to strengthen the impact of the AfricaBP Open Institute, including the creation of fellowships dedicated to biodiversity genomics and bioinformatics,

e-mail: bouabidbadaoui@gmail.com; maphon@unisa.ac.za; awmuigai@yahoo.co.uk; julian.osuji@uniport.edu.ng; thankgod1980@yahoo.co.uk

https://doi.org/10.1038/s41588-024-01807-6

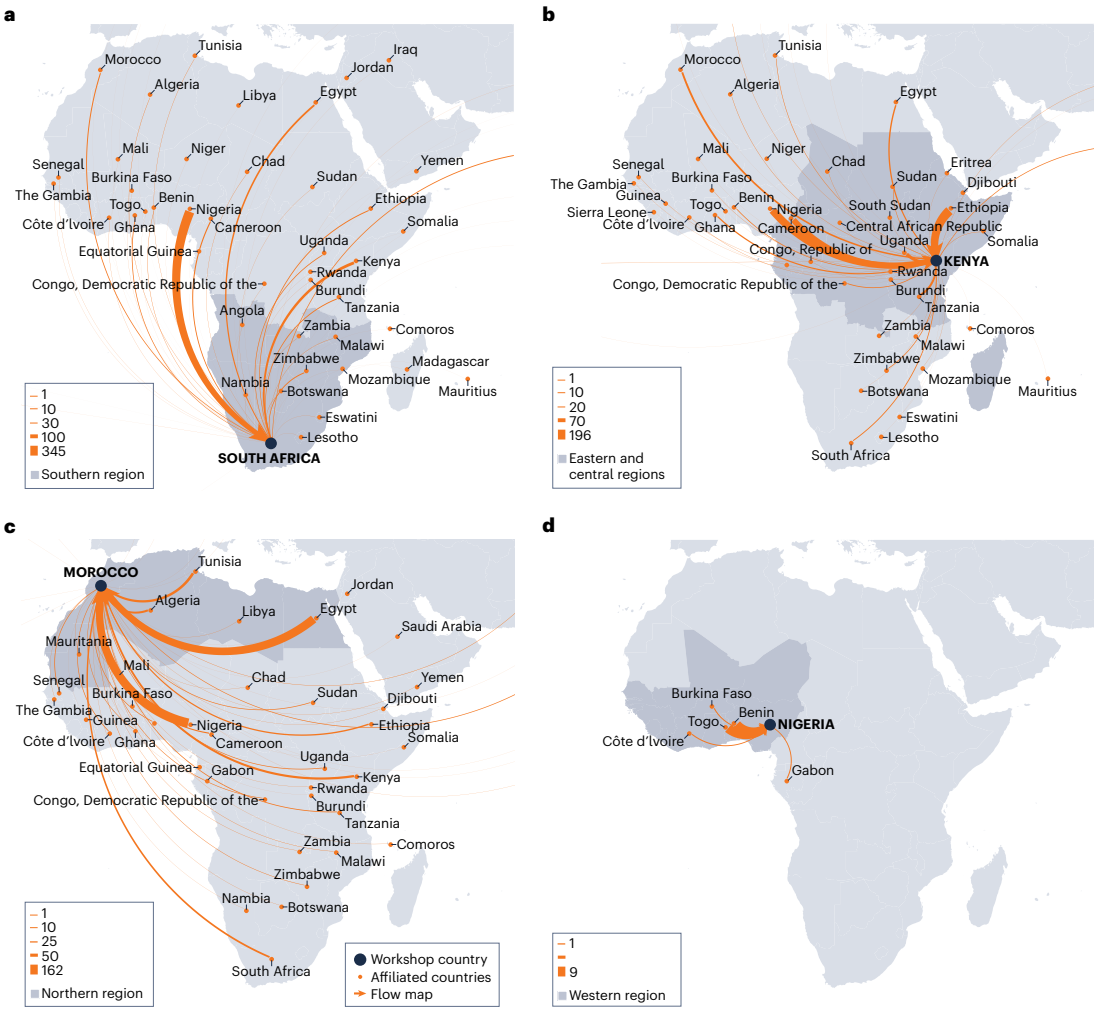


Fig. 1 | Locations of the AfricaBP Open Institute 2023 regional workshops. **a**–**d**, Illustration of the countries of origin of registered participants for the southern Africa (**a**), eastern and central Africa (**b**), northern Africa (**c**) and western Africa (**d**) regional workshops. The map was generated with AfricaBP registration information using QGIS software (QGIS, 2024), courtesy of ArcWorld Supplement. Sections of the map colored in dark gray represent African geographical regions with overlapping common travel areas. The width of the flow lines reflects the proportion of participants affiliated with a particular country.

the establishment of group leader positions for early-career and established researchers, and the creation of an African digital database containing sequence information relating to biodiversity and agriculture. In addition, we also make recommendations on how this workshop model could be adopted by other scientific communities across Africa, such as the African microbiome community[11], hopefully enabling African researchers to be reached more effectively.

Workshop structure and key outcomes

The 2023 series (inaugural edition) of the AfricaBP Open Institute regional workshops were held across the five regions of Africa (Fig. 1) at different times of the year through a 'hub-and-spoke' model (Fig. 2): (1) western Africa regional workshop, coordinated and hosted by the University of Port Harcourt, Nigeria, from 15 to 16 February 2023; (2) southern Africa regional workshop, coordinated and hosted by the University of South Africa, South Africa, from 21 to 25 August 2023; (3) northern Africa regional workshop, coordinated and hosted by the Mohammed V University in Rabat, Morocco, from 9 to 13 October 2023; and (4) eastern and central Africa regional workshop, coordinated and hosted by the International Livestock Research Institute, Kenya, from 4 to 8 December 2023 (Figs. 1 and 2).

Each regional workshop lasted 5 days, except for the West Africa workshop which lasted 2 days. Together, they attracted 3,788 African scientists, students and researchers and were sponsored through a

https://doi.org/10.1038/s41588-024-01807-6

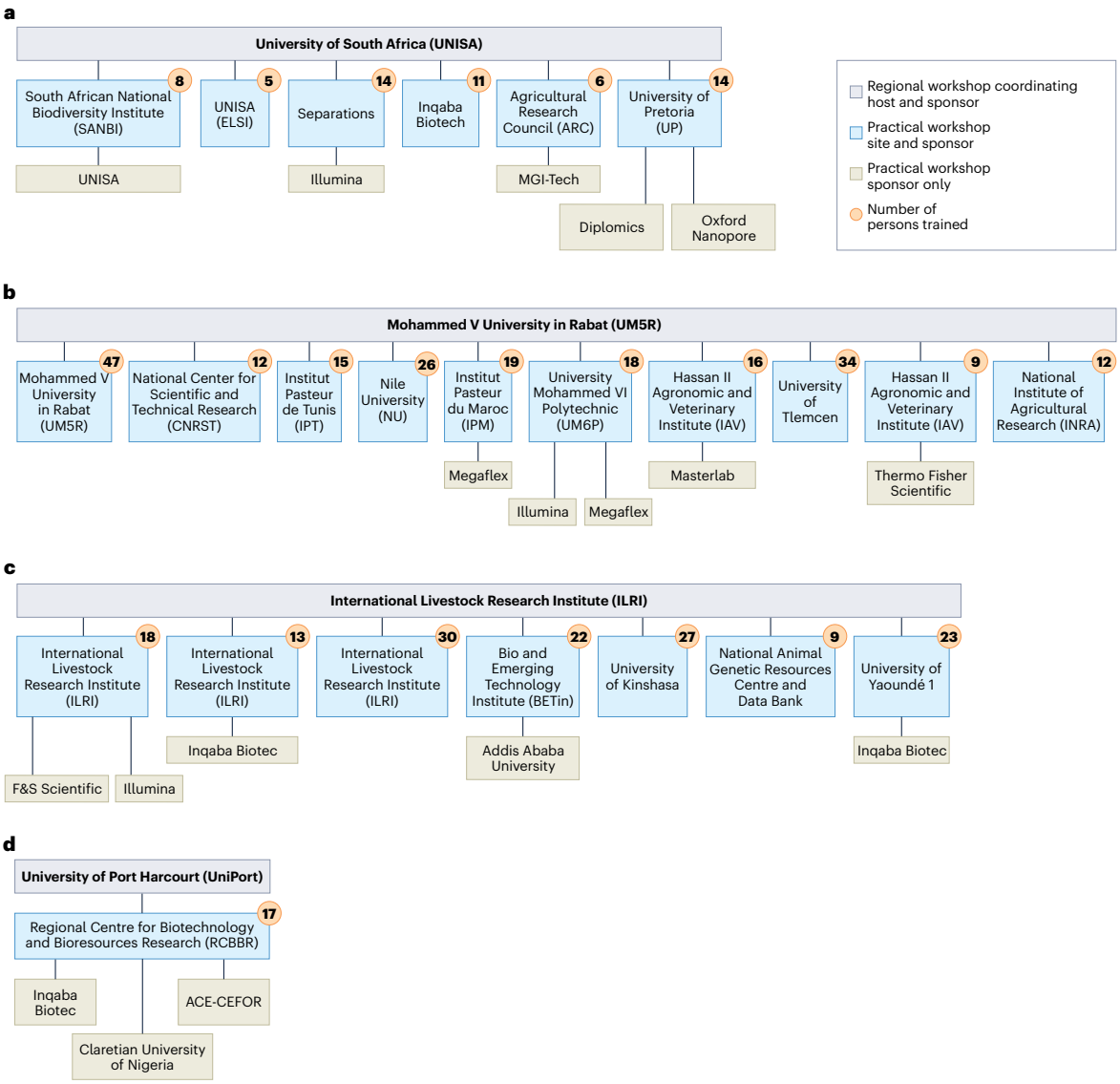


Fig. 2 | Local resources used in the 2023 regional workshop series. Illustrated here are 'hub-and-spoke' models for the delivery of the different AfricaBP Open Institute 2023 workshops. **a**, Southern Africa regional workshop. **b**, Northern Africa regional workshop. **c**, Eastern and central Africa regional workshop. **d**, Western Africa regional workshop. The diagrams also highlight interactions between genomic technology and service providers, genomic technology distribution partners, regional representatives or country representatives, and academic and research institutions providing mutual knowledge exchange. ELSI, ethical, legal and social issues; ACE-CEFOR, African Centre of Excellence Centre for Oilfield Chemicals Research.

PPP between AfricaBP, African institutions and industry sponsors. The first 2 days included presentations on diverse aspects—such as biodiversity genomics and bioinformatics, sample collection, as well as ethics and policy issues—by Africa- and non-Africa-based speakers and were selected to focus on understanding the current landscape of biodiversity across each region. The last 3 days involved hands-on practical sessions on genomics and bioinformatics processes; these were hosted in parallel across different locations in each of the regions. For each of the workshops, participants for the hands-on practical sessions were selected through a competitive short-listing process, whereas attendance at the lectures was open to anyone interested. Below, we highlight five key outcomes of the workshops.

Launch of comprehensive fellowships to drive important change

One outcome of the workshop series was the launch of a comprehensive fellowship program and proposed group leader and professorial chair positions in biodiversity genomics and bioinformatics across Africa by the AfricaBP Open Institute (Supplementary Fig. 2). Here, the regional

workshops paved the way for initial investments in short-term and residential training 'Fellowships in Biodiversity Genomics', which was launched through a partnership between the AfricaBP Open Institute and the African Genome Center, University Mohammed VI Polytechnic in Morocco. For the short-term fellowship, in May 2024, 10 selected African scientists will spend 2 weeks at the African Genome Center laboratory in Morocco to generate a publishable sequenced genome of *Vachellia gummifera*, a species native to Morocco. This plant has a crucial role in soil conservation and can thrive in harsh desert environments[12].

Similarly, the AfricaBP Open Institute partnered with the Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, USA, to launch the 'African Biodiversity Fellowship for Emerging Genomics Leaders 2024', which aims to empower five African researchers to develop curricula, research programs and leadership in biodiversity genomics and bioinformatics[2]. The eligibility for this program focuses on African scientists with permanent positions in African institutions and whose institutions organized workshops, to ensure the availability of trained personnel to continue delivering the annual practical workshops as well as address the brain-drain challenges experienced by the H3ABioNet network[1].

Bringing local genomics and bioinformatics to the mainstream

The regional workshops delivered by the AfricaBP Open Institute used entirely local infrastructure and personnel across the 28 locations and 11 African countries and were organized in African regions with initiatives that encourage the free movement of people without visa requirements for African researchers[13]. For example, the East African regional workshop in Kenya hosted researchers from the eastern African states of Burundi, Ethiopia, Rwanda, Tanzania, Uganda and Kenya; these researchers did not require visas to travel to Kenya, but this would not have been the case if they had to travel to other African regions (Fig. 1).

Furthermore, the workshops demonstrated that African countries with comparably limited genomics and bioinformatics capacities can benefit from other African countries that have more advanced genomics and bioinformatics capabilities. Specifically, during the central/eastern and northern Africa workshops, the AfricaBP Open Institute workshops reached African institutions and scientists from the Democratic Republic of Congo (DR Congo), Cameroon and Algeria, which do not constitute primary destinations for most African genomics and bioinformatics capacity-building efforts[1,2,14] (Fig. 2 and Supplementary Fig. 2).

Integrating PPPs is beneficial to workshop delivery

The AfricaBP Open Institute successfully integrated PPPs into the delivery of the regional workshops (Fig. 2) involving 9 industry sponsors and 24 African institutions; this benefited both the Open Institute and public African institutions (for example, the University of Yaounde 1 in Cameroon, which did not need to apply for grant funding to host the workshops), the industry sponsors and the African researchers trained who were able to participate at no cost to them. The PPP arrangements involved establishing 5 levels of sponsorship (Supplementary Table 1) depending on the activities sponsored during the first 2 days, such as catering, videography, information technology and/or media, or during the practical workshop, such as organizing and hosting these workshops, provision of laboratory consumables, reagents, computing resources and/or laboratory spaces.

Here, the PPP was relevant to the success of the workshop as it allowed the AfricaBP Open Institute to quickly mobilize resources (Supplementary Table 1) to host regional workshops without applying for grant funding to funding agencies, which not only takes considerable time but also has a low success rate[15,16]. Among the benefits for the industry sponsors are establishing third-party partnerships and increasing awareness of their technologies, products and services, not only among the registered participants (Supplementary Figs. 3 and 4) across 80 countries (Supplementary Fig. 1) but also among the additional followers across AfricaBP's social media platforms.

Engaging participants by lowering language barriers

One barrier to delivering effective training in Africa is the language used, which some participants may not fully understand[2]. Therefore, the presentations during the first 2 days of the eastern and central Africa regional workshop were primarily delivered in English in Nairobi, Kenya, and translated into French for the participants in Kinshasa, DR Congo (a French-speaking country), whereas the last 3 days of the hands-on practical workshops were delivered in English in Kenya, Ethiopia, Uganda and Cameroon and in French in Kinshasa, DR Congo (Box 1). In the future, delivering such training in Kikongo, one of the Bantu languages in DR Congo, is expected to increase training success further, as has been previously proposed[2].

Establishing an African digital sequence information data bank for biodiversity and agriculture

The current realities of digital sequence information and the Global Biodiversity Framework[17–19] call for a more federated database configuration that maximizes data ownership at national levels across Africa in line with national biodiversity regulatory frameworks[2,20,21]. The AfricaBP Open Institute initiated proposals for the creation of the African digital sequence information data bank for biodiversity and agriculture (Supplementary Fig. 2) in the form of federated repositories for storage, computing, processing and visualization of African biodiversity and agricultural genomic data. Such federated repositories could involve databases from national storage centers, being supplied from those, or could mirror the original datasets in the countries of sample origin. In this way, the countries where these samples originated would retain ownership of the data while promoting open science. An example of such a setup is the Korean BioData Station, a repository for biological data generated from research projects funded by the Korean government with multiple databases for individual data types[22].

Summary of the course contents of the regional workshops

Here, we provide a brief summary of the hands-on practical sessions (Fig. 2), which were classified into five categories: (1) genome sequencing and bioinformatics analysis; (2) bioinformatics analysis only; (3) basic molecular biology; (4) ethical, legal and social (ELS) issues; and (5) sample collection and biobanking, as outlined below.

Genome sequencing and bioinformatics analysis

Illumina technology genome sequencing. Illumina and its Africa-based partners (Separations, Megaflex and F&S Scientific) conducted hands-on training in South Africa, Morocco and Kenya, respectively. Practical training on day 1 began with an introduction to the basic principles of Illumina next-generation sequencing (NGS) technology, covering various experimental approaches and a practical session focusing on library preparation using Illumina DNA Prep workflows[23] for downstream whole-genome sequencing on Illumina NGS platforms. Experiments to optimize, quantify and quality control libraries were also performed hands-on by the trainees. The second day consisted of highlighting the best practices to pool libraries, as well as troubleshooting common issues, which was followed by practical guidance for sequencing on the Illumina platform. The final day was purely dedicated to data analysis and bioinformatics tools that are vital to genome assembly, annotation and analyzing genetic variation using the Illumina BaseSpace Sequencing Hub[24].

Inqaba Biotec genome sequencing. Inqaba Biotec gave practical sessions held at Inqaba Biotec in Pretoria, South Africa, and the

BOX 1

Lessons learned

- Persistent inclusion of local funding sources is key to facilitating the uptake and assimilation of genomics and bioinformatics.

One of Africa's problems is its overdependence on international grants to fund science[48,49]. For example, the African Union established a target of 1% of the gross domestic product to be committed to research and development; however, most African countries commit only approximately 0.38% (United Nations Conference on Trade and Development, 2021)[50]. In 2023, the AfricaBP Open Institute brokered opportunities for all participating African institutions to cosponsor and host the regional workshops, either through in-kind contributions, institutional part-funding or institutional self-funding, to minimize the financial burden on the scientists or institutions (Fig. 2 and Supplementary Table 1).

- Strategic planning is key to workshop success.

It is crucial to acknowledge that the inaugural West African regional workshop was considered a pilot run. It had fewer participants than the subsequent workshops and encountered challenges such as low numbers of industry sponsors and hands-on practical sessions, limited publicity for the workshop, its scheduling during the electoral period in Nigeria and a short planning time. This served as a valuable learning experience for subsequent regional workshops in 2023.

- Lack of travel and accommodation funding does not drastically limit participation.

The AfricaBP Open Institute did not fund travel and accommodation associated with workshop attendance. Instead, we prioritized sponsoring additional practical sessions for genome sequencing and holding these in African countries, where there are no such courses currently. Hopefully, this will encourage commercial sponsorship of satellite sites that might reflect new markets within Africa.

- Impact can be maximized through increased virtual engagement.

Given the limited places for in-person practical sessions, leveraging virtual participation can considerably extend the reach and impact of regional workshops. Hosting workshops in various African regions and offering online bioinformatics training sessions can help ensure that all participants, regardless of location or financial constraints, have access to valuable learning opportunities (Supplementary Fig. 5). Additionally, exploring virtual laboratories or demonstrations, especially using emerging technologies such as imaging and artificial intelligence, might enhance engagement and knowledge dissemination, ultimately maximizing the impact of the workshops.

- Customization of the curriculum is important.

A crucial lesson learned is the importance of tailoring the workshop curriculum, including language and available resources, to the specific needs and context of African biodiversity researchers. For instance, whereas some participants were trained in PacBio HiFi genome sequencing at the Inqaba Biotec facility in Pretoria, South Africa, participants at the University of Yaounde 1 in Yaounde, Cameroon, were trained in basic molecular biology techniques as there is no PacBio HiFi technology in Cameroon. This tailored approach ensured that the workshops were relevant (Supplementary Fig. 2) and directly addressed the challenges faced by participants, such as limited access to resources and expertise. By focusing on practical skills and tools relevant to the research interests of participants, the impact of the workshops can be maximized by empowering participants.

- Inclusivity and diversity should be fostered.

Creating an inclusive environment by actively including researchers from diverse backgrounds, including women, early-career researchers and those from underrepresented regions (Supplementary Fig. 2), enriches the learning experience for all participants. This inclusivity not only enhances the immediate impact of the workshop but also contributes to the long-term sustainability of the format and the broader impact within the research community.

- Continuous evaluation and adaptation are necessary.

It is essential to continuously evaluate workshops and adapt them based on feedback and evolving needs. Such an approach ensures that the workshops remain responsive to the changing needs and priorities of African biodiversity researchers, thus helping maintain their effectiveness and relevance over time.

International Livestock Research Institute in Nairobi, Kenya. Day 1 started with a theoretical introduction of the PacBio single-molecule real-time (SMRT) technology and an overview of the protocol. Participants were given either a bacterial genomic DNA sample (in South Africa) or fish tissue (in Kenya) for processing. Sample quantification, DNA extraction and quality control were performed using a PacBio Nanobind tissue kit and a Qubit high-sensitivity DNA Assay kit. Day 2 started with SMRTbell library preparation using the PacBio SMRTbell prep kit 3.0 (refs. 25,26), following which participants were introduced to PacBio's SMRTLink software for sample setup and run design. Finally, on day 3, participants were introduced to the PacBio SMRT Link analysis software for performing data quality control and genome assemblies.

MGI-Tech genome sequencing. The MGI-Tech workshop focused on their single-tube long fragment read (stLFR) technology, which entailed library preparation and sequencing of two plant species using the MGI DNBSEQ-G400 medium- to high-throughput sequencer[27] located at the Agricultural Research Council in Pretoria, South Africa. Day 1 began with an introductory lecture to MGI's DNBSEQ and stLFR technologies, before participants commenced wet laboratory work with sample quality control, followed by library preparation, which was completed on day 2. The final day started with the sequencing of constructed libraries on the DNBSEQ-G400, followed by a presentation on primary bioinformatics analyses, such as nucleotide signal detection, base calling, FASTQ writing and interpretation of preliminary and final sequencing reports.

Oxford Nanopore Technology genome sequencing. DIPLOMICS, Oxford Nanopore Technologies (ONT) and the University of Pretoria Genomics Laboratory hosted the workshop in Pretoria, South Africa.

Day 1 focused on introducing the principles of the ONT sequencing platform and library preparation. Participants completed library preparations; these were loaded onto MinION flow cells and sequenced using both MinION Mk1B and Mk1C instruments[28]. Day 2 focused on project design, quality control, planning of ONT sequencing run experiments, including the DNA isolation methods, and selection of the most relevant kits. The final day focused on in-depth analyses of all the sequencing run reports, data quality and yield. The analysis software was introduced, followed by extended practical metagenomic and SARS-CoV-2 data analysis sessions as well as exposure to the various Oxford Nanopore community resources and applications, including metagenomics solutions[29–31].

Bioinformatics analysis only

The Institut Pasteur de Tunis, Tunisia, organized a 'Data Carpentry Genomics' workshop focused on managing and analyzing genomics data using cloud computing during the North Africa regional workshop. The workshop leveraged a preconfigured Amazon machine image for Amazon Web Services and equipped participants with essential skills through four key modules: cloud computing for genomics navigation and data transfer; project organization and data management, such as accessing data on the National Center for Biotechnology Information Sequence Read Archive; command line fundamentals for file system navigation and automation; and data wrangling, quality control, read alignment and visualization of genetic variation. The workshop used data from the tempo and mode of genome evolution experiment[32].

Basic molecular biology

This workshop was organized and sponsored by Inqaba Biotec Central Africa in collaboration with the Biotechnology Center, University of Yaounde 1, Cameroon, during the East and Central Africa workshops. Primer sequences were obtained from the Biotechnology Center, synthesized and used to detect the presence of malaria parasites and drug-resistant biomarkers in blood samples. Day 1 focused on a presentation on sample collection, nucleic acid extraction and sequencing technologies and a hands-on session on DNA extraction from blood samples using the Quick-DNA Miniprep Plus Kit from Zymo Research. Day 2 focused on DNA quantification, quality assessment, PCR amplification and subsequent quality controls using the OneTaq Quick-Load 2× Master Mix from New England Biolabs. On day 3, agarose gel electrophoresis was performed on the PCR products and lectures were held on sample conditioning for sequencing, Sanger DNA sequencing and an introduction to bioinformatics analysis.

ELS issues

Navigating South Africa's access and benefit sharing and ELS issues. The main objective of this workshop held at the University of South Africa was to provide participants with a general overview of the ethical and legal requirements for research projects that involve the collection of genetic samples from African fauna and flora. This is to ensure the protection of Africa's biodiversity and its stakeholders by following a typical research project from conception through dissemination. Day 1 focused on the initial stages of a project, including conception and initiation, definition and planning, by using the San–*Hoodia* case[33] as a case study. The knowledge of the San people about the appetite-suppressant properties of *Hoodia*, a succulent plant used as a substitute for food and water during hunting expeditions, led to agreements with the Council for Scientific and Industrial Research, a South African research institution, to share any benefits arising from the use of this knowledge as well as the challenges in developing and implementing these agreements[33]. Participants explored how the lack of proper research planning, failed attempts to obtain informed consent and lack of any benefit sharing negatively influenced the project; they also learned how to effectively engage with stakeholders (including indigenous people), informed consent, data collection tools and permits. Day 2 explored regulatory instruments, such as material transfer agreements (MTAs), data transfer agreements (DTAs), access and benefit sharing (ABS) of the Nagoya Protocol, the Cartagena Protocol on Biosafety, the Convention on Biological Diversity and the South African Biodiversity Act[34–36], and how these complement and support one another. The final day addressed benefit sharing, research dissemination, and commercialization of genetic materials and how this can be resolved within MTAs and DTAs.

Navigating Uganda's ABS policy during sample acquisition. The National Animal Genetic Resources Centre and Data Bank organized a practical workshop in Entebbe, Uganda, tailored to building capacity in ELS issues associated with acquiring and/or moving genetic resources. Day 1 focused on an overview of national, international and global regulatory frameworks related to the conservation of genetic resources, including an in-depth account of the Convention on Biological Diversity, ABS of the Nagoya Protocol, Global Plan of Action on Animal Genetic Resources and the Interlaken Declaration[35–37], and how they were to be localized by countries. Day 2 focused on the completion of ethical compliance documents, in particular templates relating to research ethics, such as justification of using animals in research following the 3Rs (replacement, reduction and refinement), as well as drafting an MTA and a data-sharing agreement. The final day focused on training on obtaining blood samples from cattle for different purposes, including biobanking, diagnosis, DNA extraction and serum or plasma extraction, according to the sample requirements of respective projects.

Sample collection and biobanking

This hands-on workshop at the South African National Biodiversity Institute (SANBI) provided an opportunity for participants to learn about sample collection and processing. Day 1 focused on the value of biodiversity biobanks and collections, as well as additional topics, including ethics, permits and other considerations, such as ABS of the Nagoya Protocol and Convention on International Trade in Endangered Species of Wild Fauna and Flora import and export permits[38,39]. In addition, an overview of the Biodiversity Biobanks South Africa project, the SANBI Plant Biobank and the Millennium Seed Bank was also provided. Day 2 focused on taxon sampling, endo- and ectoparasite sampling, and e-vouchering. Sampling for genetic analysis was discussed, such as suitability for high-quality extractions required for whole-genome sequencing, as well as forensic sampling and the use of chain-of-custody methods. Day 3 focused on plant sampling for DNA and herbarium collections, introduction to biodiversity data, minimum data standards and the Barcode of Life Data system.

Participants' feedback

As noted above, the workshops attracted more than 3,700 scientists, students and researchers. Notably, approximately 90% of attendees were not affiliated with AfricaBP (Supplementary Fig. 4), suggesting that this workshop was successful in attracting participants who were not already involved in AfricaBP. More than 40% of all attendees across the four regional workshops were women, achieving a commendable gender balance in a field in which female representation is traditionally low (Box 1). Interestingly, participants noted ethical compliance for sample collection, data and biomaterial specimen sharing, and a lack of clear policy as important factors slowing down their genomics research work. Machine learning analyses of the postworkshop survey (Supplementary Fig. 5) pointed to an overwhelmingly positive assessment of the workshops, with 97% of the 1,030 survey respondents expressing a desire for the return of the regional workshops in 2024.

Future directions

Based on the outcome of the inaugural workshop and the feedback received, we developed four strategic directions for future regional workshops, as outlined below.

Increasing locations for practical sessions to build critical mass

To widen the reach of the regional workshops, we aim to add further locations for the practical sessions in various African countries. For example, in 2024, we plan to organize at least 33 workshops and reach researchers in island states, such as Cabo Verde, Sao Tome and Principe, and Seychelles, which did not participate in 2023 (Supplementary Fig. 1), as well as maximize the use of infrastructures covering infectious diseases, such as the African Center of Excellence for Genomics of Infectious Diseases[40]. Similarly, we encourage the institutions hosting the regional workshops to schedule these during the graduate (PhD and MSc) research showcase week in 2024 and beyond to ensure that their schedules align strongly with the academic calendars and research activities of the participants.

Creating further fellowship opportunities and progression to group leaders and professorial chairs

The 2023 regional workshops paved the way for the award of 14 fellowships to African scientists in 2024. We aim to ensure that the 2024 regional workshops will create an additional 35 fellowship opportunities in 2025 and enable sustained progression from fellowship recipients to early-career group leader and professorial chair positions (Supplementary Fig. 2) to firmly root the disciplines of biodiversity genomics and bioinformatics across Africa. AfricaBP will increase fellowships and create group leader and professorial chair positions by strengthening its role as a negotiator of opportunities that overlap between genomics funding agencies, industry partners, and African institutions and scientists.

Realizing the African digital sequence information data bank for biodiversity and agriculture

The AfricaBP Open Institute will work with African governments and policymakers to secure government agreements through existing (or new) national or regional regulations that govern the scientific parastatals of the African Union, African economic area or national institutions for the African digital sequence information data bank for biodiversity and agriculture. Compared to the member databases of the International Nucleotide Sequence Database Collaboration (INSDC), this has better chances of easing regulatory hurdles with regard to the sharing of sequence information data, as well as their access and ownership, across Africa[38]. For example, the framework that established the member databases of the INSDC is anchored in national or regional institutions and ratified through legislative instruments in Japan, Europe or the USA[41–44]; however, no African nation is a signatory of the agreements that established these member databases.

Engaging with genomics funders to target African scientists

The AfricaBP Open Institute will encourage genomics funding organizations to target African scientists by promoting on-the-ground genome sequencing and efforts to raise awareness and build capacities (Box 1). For example, the Illumina Agricultural Greater Good Initiative (GGI), which started in 2010 and offered reagents and consumables to researchers totaling US$350,000 for up to 1,000 genomes in 2023 (refs. 45,46), could help African scientists more by facilitating sequencing on the continent and establishment of professorial chair positions to advance the GGI. Similarly, future iterations of the 2024 Wellcome Trust scoping and workshop exercise, which aims to further the integration of international genomics research and practices focusing on ELS contexts of genomics[47], could benefit from distributed workshops across the five African regions to maximize engagement with Wellcome's initiatives. For instance, 81% of the 1,374 participants who attended the AfricaBP Open Institute regional workshops in East and Central Africa in 2023 did not attend the regional workshops in South Africa (Supplementary Fig. 3).

Conclusions

The current model of the regional workshops organized by the AfricaBP Open Institute addresses issues facing African scientists, such as visa requirements, costs, travel, time and language barrier, by delivering training using local resources and creating opportunities for both the researchers and industry partners. A multipartner and multicountry framework exemplified by this initiative is a cost-effective and viable strategy for boosting the current capacity-building strategies across Africa. In the future, the regional workshops will contain practical workshops to address specific bioinformatics problems, such as modular and Internet-agnostic analytical techniques, through hackathons.

The regional workshops held in 2023 allowed the AfricaBP Open Institute to identify several grassroots scientists who are well positioned to have a local impact in genomics and bioinformatics. Building on these successes will necessitate holding the regional workshops annually to widen the pool of trained researchers throughout Africa. However, this will require core funding and sustained investments. The AfricaBP Open Institute welcomes investments from both industry partners and science funding agencies at different levels, from sponsoring one or more practical sessions, long-term funding of fellowships, group leader and professorial chair positions, to seed funding for the establishment of an African digital sequence information data bank for biodiversity and agriculture, or by providing secretarial and administrative funding for the AfricaBP Open Institute to coordinate these efforts continuously. We will continuously implement new ways, and optimize capacity-building and development approaches, that fuel in-continent genome sequencing, analysis and impacts.

References

1. Ebenezer, T. E. et al. Africa: sequence 100,000 species to safeguard biodiversity. *Nature* **603**, 388–392 (2022).
2. Sharaf, A. et al. Bridging the gap in African biodiversity genomics and bioinformatics. *Nat. Biotechnol.* **41**, 1348–1354 (2023).
3. Lewin, H. A. et al. The Earth BioGenome Project 2020: starting the clock. *Proc. Natl Acad. Sci. USA* **119**, e2115635118 (2022).
4. Heuertz, M. et al. The application gap: genomics for biodiversity and ecosystem service management. *Biol. Conserv.* **278**, 109883 (2023).
5. Theissinger, K. et al. How genomics can help biodiversity conservation. *Trends Genet.* **39**, 545–559 (2023).
6. Moore, B. et al. Ten simple rules for organizing a bioinformatics training course in low- and middle-income countries. *PLoS Comput. Biol.* **17**, e1009218 (2021).
7. Bandara, S. et al. Imagining a future in global health without visa and passport inequities. *PLOS Glob. Public Health* **3**, e0002310 (2023).
8. Mulder, N. J. et al. The development of computational biology in South Africa: successes achieved and lessons learnt. *PLoS Comput. Biol.* **12**, e1004395 (2016).
9. Ras, V. et al. Using a multiple-delivery-mode training approach to develop local capacity and infrastructure for advanced bioinformatics in Africa. *PLoS Comput. Biol.* **17**, e1008640 (2021).
10. BecA-ILRI Hub. *BecA-ILRI Hub 2017 Annual Report: Unlocking Agricultural Prosperity in Africa* cgspace.cgiar.org/server/api/core/bitstreams/ad8685bb-26f6-4018-b99f-4f78b1c9b0e7/content (International Livestock Research Institute, 2018).
11. Makhalanyane, T. P. et al. African microbiomes matter. *Nat. Rev. Microbiol.* **21**, 479–481 (2023).
12. Bennis, M. et al. Plant growth promoting activities of *Pseudomonas* sp. and *Enterobacter* sp. isolated from the rhizosphere of *Vachellia gummifera* in Morocco. *FEMS Microbiol. Ecol.* **99**, fiad114 (2023).

13. Hirsch, A. A strategic consideration of the African Union Free Movement of Persons Protocol and other initiatives towards the freer movement of people in Africa. www.jstor.org/stable/resrep29589 (2021).
14. Inzaule, S. C., Tessema, S. K., Kebede, Y., Ogwell Ouma, A. E. & Nkengasong, J. N. Genomic-informed pathogen surveillance in Africa: opportunities and challenges. *Lancet Infect. Dis.* **21**, e281–e289 (2021).
15. Herbert, D. L., Barnett, A. G., Clarke, P. & Graves, N. On the time spent preparing grant proposals: an observational study of Australian researchers. *BMJ Open* **3**, e002800 (2013).
16. Parrilla, J. M. ChatGPT use shows that the grant-application system is broken. *Nature* **623**, 443 (2023).
17. Li, Q., Ge, Y. & Sayer, J. A. Challenges to implementing the Kunming–Montreal global biodiversity framework. *Land* **12**, 2166 (2023).
18. Obura, D. The Kunming–Montreal Global Biodiversity Framework: business as usual or a turning point? *One Earth* **6**, 77–80 (2023).
19. Smith, D., Ryan, M. J. & Buddie, A. G. *The Role of Digital Sequence Information in the Conservation and Sustainable Use of Genetic Resources for Food and Agriculture: Opportunities and Challenges* Background Study Paper No. 73 (Commission on Genetic Resources for Food and Agriculture, Food and Agriculture Organization, 2023).
20. Serwadda, D., Ndebele, P., Grabowski, M. K., Bajunirwe, F. & Wanyenze, R. K. Open data sharing and the Global South—who benefits? Limited capacity, deep mistrust pose challenges to sharing. *Science* **359**, 642–643 (2018).
21. Hutson, M. The race to save the world's DNA. *The New Yorker* www.newyorker.com/science/elements/the-race-to-save-the-worlds-dna (2023).
22. Lee, B. et al. Introduction of the Korea BioData Station (K-BDS) for sharing biological data. *Genomics Inform.* https://doi.org/10.5808/gi.22073 (2023).
23. Mitchell, P. K. et al. Multi-laboratory evaluation of the Illumina iSeq platform for whole genome sequencing of *Salmonella*, *Escherichia coli* and *Listeria*. *Microb. Genom.* https://doi.org/10.1099/mgen.0.000717 (2022).
24. Cycil, L. M. et al. Metagenomic insights into the diversity of halophilic microorganisms indigenous to the Karak salt mine, Pakistan. *Front. Microbiol.* **11**, 1567 (2020).
25. Hon, T. et al. Highly accurate long-read HiFi sequencing data for five complex genomes. *Sci. Data* **7**, 399 (2020).
26. Kim, C., Pongpanich, M. & Porntaveetus, T. Unraveling metagenomics through long-read sequencing: a comprehensive review. *J. Transl. Med.* **22**, 111 (2024).
27. Shirasawa, K., Kobayashi, N., Nakatsuka, A., Ohta, H. & Isobe, S. Whole-genome sequencing and analysis of two azaleas, *Rhododendron ripense* and *Rhododendron kiyosumense*. *DNA Res.* https://doi.org/10.1093/dnares/dsab010 (2021).
28. Player, R. et al. Optimization of Oxford Nanopore Technology sequencing workflow for detection of amplicons in real time using ONT-DART tool. *Genes (Basel)* https://doi.org/10.3390/genes13101785 (2022).
29. Chapman, R. et al. Nanopore-based metagenomic sequencing in respiratory tract infection: a developing diagnostic platform. *Lung* **201**, 171–179 (2023).
30. Latorre-Pérez, A., Villalba-Bermell, P., Pascual, J. & Vilanova, C. Assembly methods for nanopore-based metagenomic sequencing: a comparative study. *Sci. Rep.* **10**, 13588 (2020).
31. Liu, L., Yang, Y., Deng, Y. & Zhang, T. Nanopore long-read-only metagenomics enables complete and high-quality genome reconstruction from mock and complex metagenomes. *Microbiome* **10**, 209 (2022).
32. Tenaillon, O. et al. Tempo and mode of genome evolution in a 50,000-generation experiment. *Nature* **536**, 165–170 (2016).
33. Wynberg, R. & Chennells, R. in *Indigenous Peoples, Consent and Benefit Sharing: Lessons from the San–Hoodia Case* (eds Wynberg, R. et al.) 89–124 (Springer, 2009).
34. Cherry, M. South Africa—serious about biodiversity science. *PLoS Biol.* **3**, e145 (2005).
35. Secretariat of the Convention on Biological Diversity. *Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity* www.cbd.int/abs/doc/protocol/nagoya-protocol-en.pdf (United Nations, 2011).
36. Secretariat of the Convention on Biological Diversity. *Convention on Biological Diversity* www.cbd.int/doc/legal/cbd-en.pdf (United Nations, 2011).
37. Commission on Genetic Resources for Food and Agriculture. *Global Plan of Action for Animal Genetic Resources and the Interlaken Declaration* www.fao.org/3/a1404e/a1404e.pdf (Food and Agriculture Organization of the United Nations, 2007).
38. Ahmad, M. 3. Trade in endangered species. *Yearb. Int. Environ. Law* **31**, 121–124 (2020).
39. Challender, D. W. S. et al. Identifying species likely threatened by international trade on the IUCN Red List can inform CITES trade measures. *Nat. Ecol. Evol.* **7**, 1211–1220 (2023).
40. Folarin, O. A., Happi, A. N. & Happi, C. T. Empowering African genomics for infectious disease control. *Genome Biol.* **15**, 515 (2014).
41. Tateno, Y. & Gojobori, T. DNA Data Bank of Japan in the age of information biology. *Nucleic Acids Res.* **25**, 14–17 (1997).
42. Stoesser, G. et al. The EMBL nucleotide sequence database. *Nucleic Acids Res.* **26**, 8–15 (1998).
43. Smith, K. *A Brief History of NCBI's Formation and Growth* www.ncbi.nlm.nih.gov/books/NBK148949/ (National Center for Biotechnology Information, 2013).
44. Arita, M., Karsch-Mizrachi, I. & Cochrane, G. The international nucleotide sequence database collaboration. *Nucleic Acids Res.* **49**, D121–D124 (2021).
45. Kuster, R. D., Yencho, G. C. & Olukolu, B. A. ngsComposer: an automated pipeline for empirically based NGS data quality filtering. *Brief. Bioinform.* **22**, bbab092 (2021).
46. Bredeson, J. V. et al. Chromosome evolution and the genetic basis of agronomically important traits in greater yam. *Nat. Commun.* **13**, 2001 (2022).
47. Wellcome Trust. Beyond the sequence: ethical, legal and social contexts in genomics. wellcome.org/what-we-do/our-work/beyond-sequence-ethical-legal-and-social-contexts-genomics (2023).
48. Caelers, D. & Okoth, D. Research funding in Africa: navigating sustainability and shifting perspectives. *Nat. Afr.* https://doi.org/10.1038/D44148-023-00360-4 (2023).
49. Cerdeira, J., Mesquita, J. & Vieira, E. S. International research collaboration: is Africa different? A cross-country panel data analysis. *Scientometrics* **128**, 2145–2174 (2023).
50. United Nations Conference on Trade and Development. *Technology and Innovation Report 2021* https://unctad.org/system/files/official-document/tir2020_en.pdf (United Nations, 2021).

Acknowledgements

We would like to thank all listed partners and sponsors (Fig. 2) who made the 2023 series of the AfricaBP Open Institute regional workshops possible, as well as G. E. Robinson, Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA, for reviewing the initial draft of the manuscript.

The African Genome Center (AGC)–AfricaBP Open Institute Fellowship is funded by grant number AS-77 from the Office Chérifien des Phosphates Group and the University Mohammed VI Polytechnic in Morocco awarded to the AGC.

Author contributions

Conceived and initiated the regional workshops: T.E.E. and J.O.O. Designed the regional workshop model: J.O.O. and T.E.E. Discussed the designed model and implemented and executed the regional workshops: T.E.E., J.O.O., N.M., B.B., A.M., L.T.N., I. Hayah, J.O.K., S. Mdyogolo, T.C.O., B.A.O., R.M.S, A. Barakat, A.M.M., A.E.H., A. Benkahla, A. Boukteb, A.E., A.L.M., A.M.A., A.K.E., B.A., C.W., C.C.N., C.E., D.I., E.E.F., E.A., F.O., F.Z.B., F.Z.T., F.R., F.Z., G.L.M., H.G., M. Hanachi, H.N., H.C., I.B., I.M., I. Haddadi, I.M.K., J.F.M., J.-B.K.D.E., J.B.R., J. Ogwang, J.B.D., K.E., K.L., K.K.K., L.B.T., L.H., M.R., M. Chaisi, M.W.B., M. Kilian, M. Kvas, M. Melloul, M. Chaouch, M. Khyatti, M.A., M.P.-M., M. Hijri, M.R.M., M.A.H., M.P., M. Mwale, M. Maaloum, M. Mavhunga, N.A.O., O. Aminou, O. Arbani, O.S., P.F.D., R. Mentag, R.D.Z., R.B.T., R. Megnekou, S.P.K., S.S., S.E., S.B.S.G., S.R.A.R., S.F., S.K., S.A., T.M., T.S., O.U.U., V.E., X.D., Y.G., Y.H.T., Z.C., C.Z. and A.D. Drafted the manuscript: T.E.E., A.S., K.L., J. Ogwang, M.W.B., A. Benkahla, M. Chaouch, X.D., A.M.M., N.A.O., M.P.-M., R.B.T. and V.R. Analyzed data: T.E.E., G.B., A.S. and R.M.S. Reviewed the manuscript: T.E.E., J.O.O., A.M., G.B., C.J.K., D.-F.N., F.B.M., F.R., F.Z., J. Ogwang, J.E.I., K.E., K.K.K., L.B.T., M. Chaouch, M.W.B., M. Hijri, M.P., M. Mwale, R.B.T., S. Muzemil, S.P., S.P.S., S.B.S.G., S.K., O.U.U., V.E., V.C.D., X.D., Z.M.-D., C.H. and T.K. Approved the manuscript: all authors.

Competing interests

A.M.M. and H.G. are employees of Inqaba Biotechnical Industries. Z.C., M.A., S.E., L.H., T.S. and X.D. are employees of Illumina, Inc. M. Kilian, M. Kvas and C.W. are employees of Separations (Pty) Ltd. M. Maaloum and F.Z.T. are employees of Megaflex. M.P.-M. is an employee of MGI-Tech. R.B.T. is an employee of Inqaba Biotec Central Africa. T.E.E. is an independent contractor for the Wellcome Trust. All other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41588-024-01807-6.

Correspondence should be addressed to Bouabid Badaoui, Ntanganedzeni Mapholi, Anne Muigai, Julian O. Osuji or ThankGod Echezona Ebenezer.

Peer review information *Nature Genetics* thanks Sofonias Tessema and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Abdoallah Sharaf[1,2,65], Lucky Tendani Nesengani[3,65], Ichrak Hayah[4,65], Josiah Ochieng Kuja[5,65], Sinebongo Mdyogolo[3], Taiwo Crossby Omotoriogun[6,7], Blessing Adanta Odogwu[8,9], Girish Beedessee[10], Rae Marvin Smith[3], Abdelhamid Barakat[11], Acclaim M. Moila[12], Adil El Hamouchi[11], Alia Benkahla[13], Amal Boukteb[14], Amine Elmouhtadi[15], Antoine Lusala Mafwila[16], Asmaa Mohammed Abushady[2,17], Assem Kadry Elsherif[17], Bulbul Ahmed[18], Charles Wairuri[19], Charlotte C. Ndiribe[20], Chukwuike Ebuzome[21], Craig J. Kinnear[22], Deborah-Fay Ndlovu[23], Driss Iraqi[15], Elmostafa El Fahime[24], Ermias Assefa[25], Faissal Ouardi[26], Fatima Zohra Belharfi[27], Fatim Zohra Tmimi[28], Fatu Badiane Markey[23,29], Fouzia Radouani[11], Francis Zeukeng[30], Georges Lelo Mvumbi[16], Hamilton Ganesan[12], Mariem Hanachi[13], Helen Nigussie[31], Hicham Charoute[11], Ichrak Benamri[11], Ikram Mkedder[27], Imane Haddadi[27], Issam Meftah-Kadmiri[32], Jackson Franco Mubiru[33], Jean-Baka Kodjo Domelevo Entfellner[34], Joan Bayowa Rokani[33], Joel Ogwang[33], Jude Bigoga Daiga[30], Judy Omumbo[23], Justin Eze Ideozu[35], Khaoula Errafii[18], Kim Labuschagne[36], Komi Koukoura Komi[37], Libert Brice Tonfack[30], Lydia Hadjeras[38], Madeleine Ramantswana[39], Mamohale Chaisi[36], Marietjie W. Botes[40], Mariëtte Kilian[19], Marija Kvas[19], Marouane Melloul[24], Melek Chaouch[13], Meriem Khyatti[11], Michael Abdo[38], Mmatshepho Phasha-Muchemenye[41], Mohamed Hijri[18], Mohammed Rida Mediouni[27], Mohammed Ahmed Hassan[42], Mohammed Piro[43], Monica Mwale[36], Mossaab Maaloum[28], Mudzuli Mavhunga[36], Nicholas Abraham Olivier[44,45], Oumaima Aminou[43], Oumayma Arbani[46], Oussema Souiai[13], Pierre François Djocgoue[30], Rachid Mentag[15], Renate Dorothea Zipfel[45,47], Rolland Bantar Tata[48], Rosette Megnekou[30], Sadik Muzemil[49], Sadye Paez[50], Samson Pandam Salifu[51], Samuel Paul Kagame[52], Sarra Selka[27], Sean Edwards[38], Semir Bechir Suheil Gaouar[27], Shaimaa Roshdy Abdullah Reda[17], Siham Fellahi[43], Slimane Khayi[15], Soumia Ayed[27], Thabang Madisha[39], Tulsi Sahil[38], Ogbuagu Ugorji Udensi[53], Verena Ras[54], Victor Ezebuiro[8,9], Vincent C. Duru[55], Xavier David[38], Yonas Geberemichael[25], Yves H. Tchiechoua[56], Zahra Mungloo-Dilmohamud[57], Zhiliang Chen[38], Christian Happi[58], Thomas Kariuki[23], Cathrine Ziyomo[34], Appolinaire Djikeng[3,34,59], Bouabid Badaoui[4,60] ✉, Ntanganedzeni Mapholi[3] ✉, Anne Muigai[61,62] ✉, Julian O. Osuji[8,9,63] ✉ & ThankGod Echezona Ebenezer[64] ✉

[1]SequAna Core Facility, Department of Biology, University of Konstanz, Konstanz, Germany. [2]Genetics Department, Faculty of Agriculture, Ain Shams University, Cairo, Egypt. [3]College of Agriculture and Environmental Sciences, University of South Africa, Florida, South Africa. [4]Laboratory of Biodiversity, Ecology, and Genome, Department of Biology, Faculty of Sciences, Mohammed V University in Rabat, Rabat, Morocco. [5]Washington State University, Global Health, Nairobi, Kenya. [6]Department of Biological Sciences, Elizade University, Ilara-Mokin, Nigeria. [7]A. P. Leventis Ornithological Research Institute, University of Jos, Jos, Nigeria. [8]Regional Centre for Biotechnology and Bioresources Research, University of Port Harcourt, Port Harcourt, Nigeria. [9]South–South Zonal Centre of Excellence, National Biotechnology Development Agency, Port Harcourt, Nigeria. [10]Department of Applied Sciences, Faculty of Health and Life Sciences, Northumbria University, Newcastle-upon-Tyne, UK. [11]Research Department, Institut Pasteur du Maroc, Casablanca, Morocco. [12]Inqaba Biotec, Pretoria, South Africa. [13]Laboratory of Bioinformatics, Biomathematics and Biostatistics-LR16IPT09, Institut Pasteur de Tunis, Université de Tunis El Manar, Tunis, Tunisia. [14]Field Crops Laboratory, National Institute of Agricultural Research of Tunisia (INRAT), University of Carthage, Tunis, Tunisia. [15]Biotechnology Research Unit, Regional Center of Agricultural Research of Rabat, National Institute of Agricultural Research, Rabat, Morocco. [16]Laboratory of Molecular Biology, Department of Basic Sciences, University of Kinshasa, Kinshasa, Democratic Republic of Congo. [17]Biotechnology School, Nile University, Giza, Egypt. [18]African Genome Center, University Mohammed VI Polytechnic (UM6P), Ben Guerir, Morocco. [19]Separations (Pty) Ltd, Johannesburg, South Africa. [20]University of Lagos, Lagos, Nigeria. [21]Finima Nature Park, Port Harcourt, Nigeria. [22]South African Medical Research Council Genomics Platform, Cape Town, South Africa. [23]Science for Africa Foundation, Nairobi, Kenya. [24]National Center for Scientific and Technical Research, Rabat, Morocco. [25]Bio and Emerging Technology Institute, Addis Ababa, Ethiopia. [26]Faculty of Sciences, Mohammed V University, Rabat, Morocco. [27]Applied Genetics in Agriculture, Ecology and Public Health Laboratory, University of Abou Bekr Belkaid Tlemcen, Tlemcen, Algeria. [28]Megaflex, Casablanca, Morocco. [29]Rutgers University-Newark, Newark, NJ, USA. [30]Biotechnology Centre, University of Yaoundé 1, Yaoundé, Cameroon. [31]Department of Microbial Cellular and Molecular Biology, Addis Ababa University, Addis Ababa, Ethiopia. [32]Plant and Microbial Biotechnology Center, Moroccan Foundation for Advanced Science, Innovation and Research, University Mohammed VI Polytechnic, Ben Guerir, Morocco. [33]Department of Breeding and Reproduction, National Animal Genetic Resources Centre and Data Bank, Entebbe, Uganda. [34]International Livestock Research Institute, Nairobi, Kenya. [35]AbbVie Inc., North Chicago, IL, USA. [36]Foundational Biodiversity Science, South African National Biodiversity Institute, Pretoria, South Africa. [37]Laboratoire des Sciences Biomédicales, Alimentaires et de Santé Environnementale (LaSBASE), Département des Analyses Biomédicales (AMB), Ecole Supérieure des Techniques Biologiques et Alimentaires (ESTBA), Université de Lomé, Lomé, Togo. [38]Illumina, Inc., Evry, France. [39]Agricultural Research Council, Biotechnology Platform, Pretoria, South Africa. [40]Division of Medicine, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa. [41]MGI-Tech, Pretoria, South Africa. [42]Lakes and Fish Resources Protection and Development Agency (LFRPDA), Cairo, Egypt. [43]Veterinary Genetic Analysis Laboratory, Hassan II Agronomy and Veterinary Institute (IAV), Rabat, Morocco. [44]Department of Plant and Soil Sciences, University of Pretoria, Pretoria, South Africa. [45]Forestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria, South Africa. [46]Department of Veterinary Pathology and Public Health, Hassan II Agronomy and Veterinary Institute (IAV), Rabat, Morocco. [47]Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa. [48]Inqaba Biotec Central Africa, Yaoundé, Cameroon. [49]University of Warwick, Coventry, UK. [50]Department of Neurogenetics of Language, Rockefeller University, New York, NY, USA. [51]Faculty of Bioscience, College of Science, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana. [52]Department of Biology, Ghent University, Ghent, Belgium. [53]Department of Genetics and Biotechnology, University of Calabar, Calabar, Nigeria. [54]University of Cape Town, Cape Town, South Africa. [55]Department of Parasitology and Entomology, Nnamdi Azikiwe University, Awka, Nigeria. [56]Department of Biology, Chemistry and Pharmacy, Free University Berlin, Berlin, Germany. [57]University of Mauritius, Reduit, Mauritius. [58]African Centre of Excellence for Genomics of Infectious Diseases, Redeemer's University, Ede, Nigeria. [59]Centre for Tropical Livestock Genetics and Health (CTLGH), Roslin Institute, University of Edinburgh, Edinburgh, UK. [60]African Sustainable Agriculture Research Institute (ASARI), Mohammed VI Polytechnic University (UM6P), Laâyoune, Morocco. [61]National Defence University-Kenya, Nakuru, Kenya. [62]Jomo Kenyatta University of Agriculture and Technology, Juja, Kenya. [63]Department of Plant Science and Biotechnology, University of Port Harcourt, Port Harcourt, Nigeria. [64]Early Cancer Institute, Department of Oncology, School of Clinical Medicine, University of Cambridge, Cambridge, UK. [65]These authors contributed equally: Abdoallah Sharaf, Lucky Tendani Nesengani, Ichrak Hayah, Josiah Ochieng Kuja. ✉e-mail: bouabidbadaoui@gmail.com; maphon@unisa.ac.za; awmuigai@yahoo.co.uk; julian.osuji@uniport.edu.ng; thankgod1980@yahoo.co.uk

Addendum A: Ethical approval

UNIVERSITY OF KWAZULU-NATAL
INYUVESI YAKWAZULU-NATALI

22 May 2025

Mr Sahil Tulsi (200309440)
School of Laboratory Medicine & Medical Science
Medical School

Dear Mr Tulsi,

Protocol reference number: BREC/00004110/2022
Project title: Identifying the effect of polymorphisms within specific drug transporter genes within an African Mycobacterium tuberculosis cohort
Degree: PhD

RECERTIFICATION APPLICATION APPROVAL NOTICE

Approved:	**28 June 2025**
Expiration of Ethical Approval:	**27 June 2026**

I wish to advise you that your application for Recertification for the above protocol has been **noted and approved** by a sub-committee of the Biomedical Research Ethics Committee (BREC) for another approval period. The start and end dates of this period are indicated above.

If any modifications or adverse events occur in the project before your next scheduled review, you must submit them to BREC for review. Except in emergency situations, no change to the protocol may be implemented until you have received written BREC approval for the change.

The committee will be notified of the above approval at its next meeting to be held on 10 June 2025.

Yours sincerely

[redacted]

Ms A Marimuthu
(for) Prof S Singh
Chair: Biomedical Research Ethics Committee

Biomedical Research Ethics Committee
Chair: Professor S Singh
UKZN Research Ethics Office Westville Campus, Govan Mbeki Building
Postal Address: Private Bag X54001, Durban 4000
Email: BREC@ukzn.ac.za
Website: http://research.ukzn.ac.za/Research-Ethics/Biomedical-Research-Ethics.aspx

Addendum B: Guidelines for Thesis submission

GUIDELINES FOR PRESENTATION OF MASTERS AND PHD DISSERTATIONS/THESES BY RESEARCH

1. Purpose

The purpose of this document is to provide guidance to students and supervisors on how to prepare a dissertation/thesis for Masters by Research and PhD degrees using the manuscript or publication format..

2. Introduction

These guidelines must be read together with the College of Health Sciences (CHS) Handbook as well as the Jacobs documents on examination policies and procedures for PhD degrees. The rules on thesis format are based on modification of point 1 of the definition of terms section in the Jacobs document. In this section a thesis is defined as *"the supervised research component of all PhD degrees, whether by supervised research only, or coursework and research, or by papers that are either published or in manuscript form (the supervised research component of the PhD degree by paper(s) comprises the introduction, literature review, account of the methodology, selection of manuscripts, and conclusion)."* A dissertation is defined as "*the supervised research component of all Masters degrees, whether by supervised research only, or coursework and research, or by papers that are either published or in manuscript form (the supervised research component of the Masters degree by paper(s) comprises the introduction, literature review, account of the methodology, selection of manuscripts, and conclusion*)."

2.1 PhD thesis

In the CHS Handbook the rules for a PhD thesis are not in one place; they are stated in DR8 a i & ii, DR9 c and CHS 16. DR8 a i & ii and direct that a thesis be presented in the standard format together with one published paper or an unpublished manuscript that has been submitted to an accredited journal, arising from the doctoral research. CHS16 (thesis by publications states that the thesis may comprise of at least three published papers or in press in accredited journals; such papers must have the student as the prime author. The same CHS16 provides for a thesis by manuscripts that may have at least 3 papers with the student as the prime author that have not yet been published but are in the form of manuscripts; at least two of such papers must constitute original research. In both cases (thesis by publications and manuscripts), there must be introductory and concluding integrative material sections.

The standard type thesis is being phased out in many African countries in favour of the other options that originate from the Scandinavian countries. While this format ensures that all details of the work done for the doctoral degree are captured and thoroughly interrogated, they often remain as grey literature which is mainly useful to other students, usually within the same university, although with digitization of theses, such work may become more accessible beyond the source university. Apart from the risk of losing good work because of it not being on the public domain, as students rarely publish such work after graduating, this approach denies the college additional productivity units (PUs) emanating from publications.

The thesis by publication encourages students to publish key aspects of their doctoral research as they will not graduate if the papers are not published or in press. This approach ensures that the work of the student enters the public domain before the thesis is examined, providing the examiner with some assurance of prior peer review. The thesis must constitute a full study of the magnitude expected of a PhD with the papers providing a sound thread or storyline. Furthermore, the college maximizes the students' work as PUs are awarded for the papers as well as for graduating. However, this approach may negatively affect throughput and frustrate students as

they cannot graduate unless all the papers are published or in press, in addition to the synthesis chapter demonstrating the story line of the thesis.

The option of a thesis by manuscripts ensures that students make efforts to start publishing. The risk of not passing because of failure to publish all papers (as in the thesis by publication) does not exist under this option. However, the PUs emanating from publications from the doctoral work are not guaranteed as the submitted papers may eventually be rejected. Thus there is a possibility of the doctoral work remaining on the university library shelves as is the case for the standard thesis format. The standard thesis does have the advantage that more details of the doctoral work are usually included.

In view of the above, the best option for the college is that of a thesis by publication. However, in the interim, the attractive option is that of thesis by manuscripts, as it provides the possibility of publication without putting the student at risk of delayed graduation when some of the manuscripts are not published/accepted, which also disadvantages the college in terms of PU earnings. The standard thesis option should ultimately be phased out for the stated reasons and students are not encouraged to present their theses in that format. Consequently this document does not describe the standard thesis.

2.2 MSc dissertation

The rules on presentation of MSc dissertations are presented in CR13 (course work), CHS 14 (course work) and MR9 (research) in the CHS Handbook. CR13 c and MR9 c direct that a dissertation "may comprise one or more papers of which the student is the prime author, published or in press in peer-reviewed journals approved by the relevant college academic affairs board or in manuscripts written in a paper format, accompanied by introductory and concluding integrative material." Such a dissertation should include a detailed description of the student's own distinct contribution to the papers. Both CHS14 and CR13 specify that reviews and other types of papers in addition to original research paper/s may be included, provided they are on the same topic.

3 ***Length of thesis and dissertation by word count***

Table 1 provides a guide of the length of a thesis or dissertation by word count excluding preliminary pages and annexes.

Table 1: Thesis length by word count

Sections				
	Minimum	Maximum	Minimum	Maximum
Introduction	2700	2700	2000	2000
Chapters	10000	25000	6000	11000
synthesis	2000	2000	1700	1700
bridging	300	300	300	300
Total	15000	30000	10000	15000

4. *Intention to submit*

A written intention to submit a thesis or dissertation should be submitted to the appropriate postgraduate office with endorsement of the supervisor at least three months before the actual date of submission which should be before November if the student intends to graduate in the following year. The actual submission will under normal circumstances require approval of the supervisor.

5. Format for theses/dissertation

There is little variation in the actual format of the PhD thesis and Masters dissertation for the various types described above. The box below summarise the outline of a thesis/dissertation for the thesis by manuscripts and thesis by publications.

Box 1: Outline of thesis

Preliminary pages

i. Title page
ii. Preface and Declaration
iii. Dedication
iv. Acknowledgements
v. Table of contents
vi. List of figures, tables and acronyms (separately presented)
vii. Abstract

Main Text

1. Chapter 1: Introduction
   Introduction including literature review
   Research questions and/or objectives
   Brief overview of general methodology including study design
2. Chapter 2
   First manuscript/publication
3. Chapter 3
   Second manuscript/publication
4. Chapter n
   Final manuscript/publication
5. Chapter n+1: Synthesis
   Synthesis
   Conclusions
   Recommendations
6. References Appendices

NB. Between the manuscripts or publications there must be a 1 page (maximum) bridging text to demonstrate the link between them

6. Details for thesis/dissertation subheadings

This section summarizes what is expected under each subheading shown in Boxes 1 and indicates where there might be variations between a Masters Dissertation and PhD Thesis.

6.1 Title Page

The officially approved title that is concise (Fewest words that adequately describe the contents of the thesis/dissertation – usually 15 or fewer words) is presented at the top. This should be followed by the candidate's name in a new line. At the bottom the thesis statement should be presented. The thesis statement may be stated as "*Submitted in fulfillment of the requirements for the degree of ____ in the School of _____, University of KwaZulu-Natal"* for a PhD/Masters by Research thesis. In the case of a Masters Dissertation it should be stated as "Submitted as the dissertation component in partial fulfilment (% stated) for the degree of _____ in the School of _____, University of KwaZulu-Natal". For both Masters and PhD the date of submission must be stated.

6.2 Preface (Optional)

The preface merely states the reason (motivating factors) why the study was conducted without getting into details of what was investigated.

6.3 Declaration

This must be structured as follows:

I, Dr/Mr _____, declare as follows:

1. That the work described in this thesis has not been submitted to UKZN or other tertiary institution for purposes of obtaining an academic qualification, whether by myself or any other party.

   Where a colleague has indeed prepared a thesis based on related work essentially derived from the same project, this must be stated here, accompanied by the name, the degree for which submitted, the University, the year submitted (or in preparation) and a concise description of the work covered by that thesis such that the examiner can be assured that a single body of work is not being used to justify more than one degree.

2. That my contribution to the project was as follows:
   This is followed by a concise description of the candidate's personal involvement in and contribution to the project, in sufficient detail that the examiner is in no doubt as to the extent of their contribution.

3. That the contributions of others to the project were as follows:
   This is followed by a list of all others who contributed intellectually to the project, each accompanied by a concise description of their contribution. This does not include people who ordinarily would be "acknowledged" as opposed to considered for authorship.

4. Signed _____ Date_____

6.4 Dedication

This is an optional section. Should it be included it must be very brief merely indicating to whom the work is dedicated. Avoid anything too flowery

6.5 Acknowledgements

This section acknowledges all individuals, groups of people or institutions that the candidate feels indebted to for the support they rendered. The funding source for the work should also be acknowledged.

6.6 Table of contents
Table of contents must be inserted after the preliminary sections and must capture all major sections of the thesis at the various levels (primary, secondary, tertiary subheadings). It should be electronically generated and should be able to take the reader to specific headings in the thesis.

6.7 Lists of figures, tables and acronyms
These lists must be presented separately. All titles of figures presented in the thesis/dissertation must be listed indicating on what page they appear. Similarly for tables the titles must be presented indicating on what page they appear. In the case of acronyms, the acronym is stated and all the words describing the acronym are presented. Only key acronyms should be stated. In some cases they may not be listed as long as full text is presented whenever the acronym is used for the first time.

6.8 Abstract
The abstract should summarize the thesis mainly stating the purpose of the study, highlights of chapters and the new knowledge contributed by the thesis. The abstract must be approved by the supervisor of the thesis and should not be more than 350 words in length.

6.9 Introduction
The introductory chapter for both types of thesis is similar. The section should include literature review and have the following information. Headings are used as appropriate and need not correspond exactly to the following.

i. Background and the context of the study
ii. Description of the core research problem and its significance
iii. A comprehensive, critical, coherent overview of the relevant literature leading to clearly defined knowledge gaps
iv. A coherent problem statement highlighting the nature and magnitude of the problem, the discrepancy, knowledge gaps therein and possible factors influencing the problem.
v. Clear and SMART research questions, objectives and hypothesis and/or theoretical framework
vi. A conceptual framework (optional)
vii. Description of the study area and general methodology *(in a standard thesis this should be a stand-alone section)*
viii. Layout of the thesis (thesis structure) indicating what chapters are presented in the thesis and how they address the objectives.

6.10 Literature review
This section is subsumed in the introduction within the stipulated word count for a thesis or dissertation.

6.11 Methodology
A standalone section is not needed as the methods are adequately described in each manuscript/publication.

6.12 Data chapters/manuscripts/publications
The full published paper or manuscript submitted for publication should be presented as published or submitted to the journal. The actual published paper should be scanned and inserted

in the chapter. There should be a separator page between chapters that has text linking the previous chapter to the next and providing details of the next manuscript/publication indicating publication status.

6.13 General discussion/Synthesis chapter

This is a general discussion that demonstrates the logical thread that runs across the various manuscripts/publications (synthesis). There should be no doubt that the manuscripts/publications complement each other and address the original objectives stated in the general introduction of the thesis. The general discussion/synthesis chapter should end with a conclusion and recommendations where necessary.

6.14 References

Only references cited in the introduction and synthesis chapters should be listed as all other references should be within the manuscripts presented under data chapters.

6.15 Annexes

All information (questionnaires, diagrams, ethics certificates, etc) considered important but not essential for inclusion in the actual thesis is put in this section as reference material. In addition papers that emanated from the work but not directly contributing to the thesis may be included.

7. Thesis formatting

For standardisation of thesis the following formatting specifications should be followed.

7.1 Font

Times New Roman 11pt should be used throughout the thesis. However, major headings may be made bigger (12pt) but using the same font type

7.2 Paper size and margins

A4 (297 x 210 mm) should be used and in the final thesis both sides of the paper should be used. However, the loose bound copy submitted for examination should be printed on only one side. The recommended margins are 30mm for all the left, right, top and bottom margins.

7.3 Line spacing

The copy submitted for examination should have 1.5 line spacing but the final copy should have single line spacing. Paragraphs should be separated by a blank line. Published or submitted manuscripts should remain in their original format in all aspects as they are inserted in their published format in appropriate places.

7.4 Headings

A consistent numbering system and captions should be maintained with first level being in CAPS and centred, second level being **normal bold** font and third level being ***italics bold***. If there is need for 4th level it should be *normal italics*.

7.7 Pagination

Page numbers should be centred at the bottom of the page. All preliminary pages should be numbered in lower case Roman numerals and subsequent pages should be numbered as indicated in the Box The title page should not be numbered.

The body of the thesis (chapter 1 onwards) should be numbered consecutively with Arabic numerals. The numbers should continue consecutively from the introduction through the through the publications or submitted manuscripts and subsequent sections. The published papers will therefore bear two numbers: a set specific to the manuscript (it is recommended to place these in the upper right hand corner) or published paper, as well as the consecutive numbers belonging to the thesis as a whole. Care must be taken to distinguish these in terms of position and font.

7.8 Referencing

Supervisors have the freedom to decide the type of citation of references but there must be consistency. This is mainly applicable to the standard type of thesis. In the case of thesis by manuscripts or publications, individual papers will maintain the reference system of the journal but the supervisor can decide on the type of referencing for the introductory and synthesis chapters.

8. Final thesis submission

The thesis should be submitted for examination in a loose bound form accompanied by a PDF copy. After the examination process the final version PDF copy of the thesis must be submitted to PG office for onward submission to the library. It is not a requirement to submit a copy fully bound in leather cloth or similar material.

Addendum C: Turnitin Digital Receipt

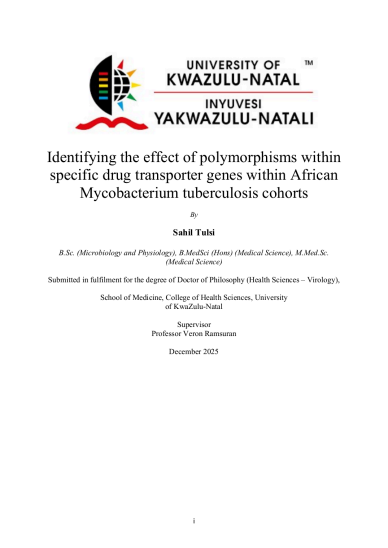
turnitin

Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Sahil Tulsi
Assignment title:	Publications
Submission title:	Sahil Tulsi_PhD Thesis_27Nov_2025
File name:	Sahil_Thesis_PhD_27Nov_2025.docx
File size:	17.51M
Page count:	166
Word count:	42,421
Character count:	248,946
Submission date:	27-Nov-2025 08:53PM (UTC+0200)
Submission ID:	2645973258



UNIVERSITY OF KWAZULU-NATAL
INYUVESI YAKWAZULU-NATALI

Identifying the effect of polymorphisms within specific drug transporter genes within African Mycobacterium tuberculosis cohorts

By

Sahil Tulsi

B.Sc. (Microbiology and Physiology), B.MedSci (Hons) (Medical Science), M.Med.Sc. (Medical Science)

Submitted in fulfilment for the degree of Doctor of Philosophy (Health Sciences – Virology),

School of Medicine, College of Health Sciences, University of KwaZulu-Natal

Supervisor
Professor Veron Ramsuran

December 2025

i

***Actual Similarity = 12%**, due to 10% originating from a previous draft of this thesis submitted to TurnItIn/ UKZN by myself, Sahil Tulsi which remains on the database hence the similarity index of 22% below.