INTEGRATED GENOMIC, TRANSCRIPTOMIC AND PHENOTYPIC ANALYSIS OF *MYCOBACTERIUM TUBERCULOSIS* TO IDENTIFY POTENTIAL THERAPEUTIC TARGETS AND BIOMARKERS

TESFAYE GEBREYOHANNIS HAILEMARIAM

(B. PHARM, MSC)

MAY 2025

ADDIS ABABA UNIVERSITY

ADDIS ABABA, ETHIOPI

Integrated Genomic, Transcriptomic and Phenotypic analysis of *Mycobacterium Tuberculosis* to identify potential Therapeutic Targets and Biomarkers

By: -

Tesfaye Gebreyohannis Hailemariam

(B. Pharm, MSc)

A Dissertation submitted to the Department of Pharmacology and Clinical Pharmacy, School of Pharmacy, College of Health Sciences, Addis Ababa University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy (PhD) in Pharmacology.

Principal Supervisor

Professor Ephrem Engidawork, PhD.

Department of Pharmacology and Clinical Pharmacy

Addis Ababa University

Co Supervisors

Dr. Kidist Bobosha, PhD.

Armauer Hansen Research Institute

Professor Simon J Waddell

Department of Global Health and Infection, Brighton and Sussex Medical School, University of Sussex, UK

Dr. Workineh Shibeshi, PhD.

Department of Pharmacology and Clinical Pharmacy, Addis Ababa University

Dr. Yonas Kassahun Hirutu, PhD.

Armauer Hansen Research Institute

MAY 2025

Addis Ababa University

Addis Ababa, Ethiopia

Declaration

I, the undersigned, hereby affirm that this Ph.D. thesis is the result of my original research and has not been submitted for a degree at any other university.

Name: Tesfaye Gebreyohannis Hailemariam

Signature: _____

This thesis has been submitted for examination with my approval as a Principal Supervisor.

Name: Professor Ephrem Engidawork

Signature: _____

Place and date of submission: Department of Pharmacology and Clinical Pharmacy, School of Pharmacy, College of Health Sciences, Addis Ababa University, Addis Ababa, Ethiopia, January,2025

Approved by the examining Board:

Name	**Date**	**Signature**
Professor Ephrem Engidawork (**Principal supervisor**)	----------	----------
Dr. Workineh Shibeshi (**co supervisor**)	----------	----------
Dr. Kidist Bobosha (**co supervisor**)	----------	----------
Professor Simon J Waddell (**co supervisor**)	----------	----------
Dr. Yonas Kassahun Hirutu (**co supervisor**)	----------	----------
Dr. Endalkachew Alemu (**External examiner**)	----------	----------
Prof. Tadesse Eguale (**Internal examiner**)	----------	----------
Dr. Teshome Nedi (**Chairperson**)	----------	----------

List of Scientific Papers

This dissertation is based on the following published and unpublished works:

1. **Tesfaye Gebreyohannis Hailemariam**, Melaku Tilahun, Abay Atinafu Ayele, Tesfaye Gelanew, Tewodros Tariku Gebresilase, Mekides Alemu, Abaysew Ayele, Shewki Moga Siraj, Workineh Shibeshi, Kidist Bobosha, Liya Wassie, Yonas K. Hirutu, Ephrem Engidawork. **Comparative growth kinetics and drug susceptibility of *Mycobacterium tuberculosis* lineages prevalent in Ethiopia: Implications for tuberculosis treatment and management.** ***Frontiers in Microbiology***. 2025. https://doi.org/10.3389/fmicb.2024.1512580
2. **Tesfaye Gebreyohannis Hailemariam**, Abaysew Ayele, Tesfaye Gelanew, Michael Brennan, Zemedkun Abebe Debella, Workineh Shibeshi, Kidist Bobosha, Yonas K. Hirutu, Simon J Waddell, Ephrem Engidawork. **Transcriptomic and Genomic analysis of Ethiopian successful *Mycobacterium tuberculosis* sub-lineage 4.2.2.2 revealed unique differentially expressed genes.** Submitted to Scientific reports journal
3. **Tesfaye Gebreyohannis Hailemariam**, Abaysew Ayele, Tesfaye Gelanew, Abay Atnafu, Melaku Tilahun, Dawit Hailu Alemayehu, Zemedkun Abebe Debella, Yared Merid, Workineh Shibeshi, Yonas Hirutu, Abraham Assefa, Simon J Waddell, Kidist Bobosha, Ephrem Engidawork. **Integrating Genomic, Transcriptomic, and Phenotypic information to Explore Drug Resistance in *Mycobacterium tuberculosis* sub-lineage 4.2.2.2.** ***Journal of Applied Microbiology***, Oxford Academy**.** **doi.org/10.1093/jambio/lxaf063**

Abstract

Integrated Genomic, Transcriptomic and Phenotypic analysis of *Mycobacterium Tuberculosis* to identify potential Therapeutic Targets and Biomarkers

Tesfaye Gebreyohannis Hailemariam (B. Pharm, Msc)

Addis Ababa University, May, 2025

Introduction: *Mycobacterium tuberculosis* (*Mtb*), the pathogen responsible for tuberculosis (TB), remains a major global health challenge, particularly due to increasing drug resistance. Sub-lineage 4.2.2.2 is the most prevalent and frequently isolated strain among the various *Mtb* lineages and sub-lineages in treatment refractory Ethiopian TB patients, making treatment a daunting task. Beyond, the well-characterized mutation, the mechanisms involved in driving drug resistance appear to be more complex. The study findings provided valuable opportunities for identifying potential targets for developing new drugs or vaccines or biomarkers, supporting efforts to combat TB effectively.

Objectives: To investigate the growth, drug susceptibility, gene expression, and molecular mechanisms of diverse *Mtb* lineages, with a focus on Ethiopian sub-lineage 4.2.2.2, to enhance understanding of TB pathogenesis and drug resistance, inform treatment strategies, and identify potential therapeutic targets.

Methods: Drug susceptibility was assessed through whole-genome analysis and phenotypic testing using the BACTEC MGIT™ 960 system. RNA sequencing analysis was performed by isolating RNA from thirty-six *Mtb* strains during the mid-logarithmic growth phases. Quality control and taxonomic identification of reads were performed, followed by alignment and quantification using Hisat2 and featureCounts. Differential gene expression (DGE) was analyzed

with DESeq2 in R, involving normalization, dispersion estimation, and filtering. DEGs were identified using a log2(fold change) > 0.85 as a threshold.

Results: A discrepancy was observed between the phenotypic resistance profiles and the predictions based on whole-genome data, with the latter indicating a wider range of resistance. For instance, the missense mutations in rpoB (p.Ser450Leu) and katG (p.Ser315Thr) were detected, but there was no corresponding change in phenotypic drug sensitivity to rifampicin and isoniazid, respectively. RNA sequencing revealed reduced expression of six genes (Rv0096, Rv2780, Rv3136, Rv3136A, Rv3137, and Rv3230c) among drug-resistant *Mtb* isolates, although direct link to known resistance mechanisms is lacking. In *Mtb* sub-lineage 4.2.2.2 clinical isolates, seven DEGs with unique SNPs were identified, six upregulated (Rv0331, Rv0720, Rv1993c, Rv2030c, Rv2034, and Rv3129) and one downregulated (Rv0997a), previously associated with virulence, lipid metabolism, stress response, and metal transport; however, functional validation in this context is necessary. Sub-lineage 4.2.2.2 exhibited enhanced DosR regulon gene expression compared to other lineages, which may contribute to its adaptability in Ethiopia; however, this association requires further validation. Sub-lineages 4.1.2.1 and 4.2.2.2 displayed higher maximum growth concentrations (Cmax), indicating superior growth efficiency and adaptability, possibly enhancing pathogenicity and resistance. While all strains were phenotypically susceptible, minimum inhibitory concentration (MIC) values varied, with sub-lineages 4.1.2.1 and 4.2.2.2 matching WHO's critical thresholds, except for rifampicin. Lineage 3 showed increased drug sensitivity, requiring lower concentrations of rifampicin, isoniazid, and streptomycin.

Conclusions: The identified differentially expressed genes and their associated networks could be useful in unraveling the complexities of *Mtb* drug resistance and in understanding the impact that drug resistance conferring mutations have on the physiology of drug-resistant *Mtb*. Moreover,

potential targets for future therapeutic development have been presented. These results emphasize the need to account for lineage-specific variations in *Mtb* isolates to optimize treatment regimens and enhance tuberculosis control strategies, especially in regions with genetically diverse *Mtb* populations such as Ethiopia.

Key words: *Mycobacterium tuberculosis; lineage; drug-resistance; variation; gene expression; Ethiopia; optimized treatment*

Acknowledgments

First and foremost, I am deeply grateful to the **Almighty God** for granting me the strength, health, and perseverance to reach this significant milestone.

I would like to express my heartfelt appreciation to **Professor Ephrem Engidawork** for his exceptional mentorship and guidance throughout my academic journey. His support and encouragement have been fundamental to my progress and personal growth.

My sincere thanks go to **Dr. Workineh Shibeshi** for his invaluable supervision, consistent guidance, and availability over the years. His mentorship provided clarity and direction during critical stages of my work.

I am especially thankful to **Dr. Kidist Bobosha**, whose mentorship and support during my laboratory work at AHRI/TBGENE have been deeply appreciated. Her steady guidance, patience, and problem-solving approach helped me navigate both the scientific and personal challenges of my research. I am sincerely grateful for her dedication and generosity in sharing her time and expertise.

I extend my sincere thanks to **Dr. Yonas Kassahun**, who recognized my potential from the outset and played a pivotal role in my academic development at AHRI. His belief in my abilities has been truly motivating.

I am also grateful to Addis Ababa University and the Armauer Hansen Research Institute, particularly the TBGENE team, for the opportunity and support to pursue my Ph.D.

My appreciation extends to **Professor Simon J. Waddell** and his team at the Department of Global Health and Infection, Brighton and Sussex Medical School, University of Sussex, UK, for their valuable support and collaboration. Your contributions have been instrumental to the success of this work.

Finally, I would like to express my deepest gratitude to my wife, **Almaz Abebe Debela**, for her unwavering support, patience, and encouragement throughout this journey. Your strength and understanding have been a constant source of inspiration. I also thank my children—**Nathan, Yemariam**, and **Yoadan**—for their love and patience, and for understanding the time and energy this endeavor required. Your support has meant more to me than words can express.

Table of Contents

List of Figures

List of Tables

Abbreviations/Acronyms

AHRI	Armauer Hansen Research Institute
CFU	Colony-Forming Units
DEGs	Differentially Expressed Genes
DGE	Differential gene expression
DosR	Dormancy Regulon Genes
DST	Drug Susceptibility Testing
FDR	False Discovery Rate
GO	Gene Ontology
MDR	Multi Drug Resistance
MGIT	Mycobacterium Growth Indicator Tube
MIC	Minimum Inhibitory Concentration
Mtb	*Mycobacterium tuberculosis*
OD_{600}	Optical Density at 600 nm
SNPs	Single nucleotide polymorphisms
TB	Tuberculosis
TBGEN	Tuberculosis Genomic Study
WGS	Whole Genome Sequencing
XDR	Extensively Drug Resistance

Chapter 1: Introduction

1.1 Background

Tuberculosis (TB) caused by *Mycobacterium tuberculosis* (*Mtb*) is believed to have caused the deaths of 1 billion people over the past two centuries and remains one of the deadliest pathogens globally. The rise of drug resistance in *Mtb* poses a significant challenge to health systems worldwide, with some strains exhibiting extensive resistance that renders them untreatable using current antibiotic therapies (Torres Ortiz et al., 2021). As an ancient pathogen, *Mtb* has evolved sophisticated strategies to evade immune detection and establish persistent infections in its hosts. Today, it is estimated that one-quarter of the global population harbors latent *Mtb* infection, with 3–10% of those infected at risk of developing active TB during their lifetime (Coppola & Ottenhoff, 2018).

Mtb, discovered by Robert Koch over 125 years ago, is a Gram-positive, acid-fast, aerobic bacterium with a G+C-rich genome and a complex cell envelope. It is characterized by dormancy, slow growth (24-hour generation time), and intracellular pathogenesis. These traits contribute to the chronic nature of TB, requiring lengthy treatments and posing challenges for researchers. *Mtb's* dormancy, likely due to metabolic shutdown from the immune response, allows it to persist in tissues, evading eradication. Reactivation can occur decades later when immunity declines due to aging or immunosuppression, leading to disease outbreaks. While the molecular mechanisms behind dormancy and reactivation remain unclear, they are believed to involve genetically programmed intracellular signaling pathways (Cole et al., 1998).

The *Mycobacterium tuberculosis* complex (MTBC) is categorized into ten lineages, which are classified as either "ancient" (lineages 1, 5, 6, 7, 8, 9 and 10) or "modern" (lineages 2, 3, and 4) (Cerezo-Cortés et al., 2022; Guyeux et al., 2024). *Mtb* genotypes vary across populations and are highly structured geographically. Different *Mtb* lineages exhibit distinct characteristics and virulence phenotypes, influencing host immune response regulation, transmissibility, disease severity, drug resistance profiles, and vaccine efficacy (Le Hang et al., 2021).

The "modern" clade, distinguished by the *Mtb* specific deletion 1 (TbD1) and its adaptation to humans, represents a monophyletic group (a single phylogenetic lineage) comprising Lineage 2

(East-Asian), Lineage 3 (East-African Indian), and Lineage 4 (Euro-American) (Mireia Coscolla & Gagneux, 2014; Moopanar & Mvubu, 2020). These lineages have experienced relatively recent diversification compared to other MTBC strains. Among them, Lineage 2 and Lineage 4 stand out for their broader global distribution and heightened virulence compared to other lineages with more geographically confined ranges. (Mireia Coscolla & Gagneux, 2014). The ancient clade, which is paraphyletic and retains the TbD1 region, includes lineages adapted to both humans and animals: Lineage 1 (Indo-Oceanic), Lineage 5 and Lineage 6 (*M. africanum* West Africa 1 and 2), Lineage 7 (Ethiopia), Lineage 8 (Rwanda), Lineage 9 (associated with East African lineages) (Mireia Coscolla & Gagneux, 2014; Mireia Coscolla et al., 2021; Moopanar & Mvubu, 2020; Ngabonziza et al., 2020) and Lineage 10 (Central Africa) (Guyeux et al., 2024). Lineages 1, 2, and 3 lack the region of difference 239 (RD239), RD105, and RD750 regions, respectively, while Lineage 4 is characterized by a deletion in the pks15/1 gene. Lineages 5 and 6 lack the RD9 region, with Lineage 6 additionally missing RD7, RD8, and RD10. Moreover, Lineages 5 and 6 possess unique markers, RD711 and RD702 (Moopanar & Mvubu, 2020). Newly sequenced genomes of Lineage 10 also lacked RD7, RD8, and RD10, while exhibiting four IS6110 copies at a unique position not found in other lineages (Guyeux et al., 2024).

Africa is the only continent where all MTBC lineages coexist, with Ethiopia serving as a significant hub for most of them (Ngabonziza et al., 2020). According to the systematic review of 4,371 clinical isolates over 20 years found that, 99.5% were *Mtb* and 0.5% were *M. bovis*. The distribution of lineages was L4 (62.3%), L3 (21.7%), L1 (7.9%), and L7 (3.4%). The most prevalent sub-lineages were L4.2.2.ETH/SIT149 (L4.2.2.2) at 14.4%, L4.10/SIT53 at 9.7%, L3. ETH1/SIT25 at 7.2%, and L4.6/SIT37 at 5.5% (Mekonnen et al., 2019). Lineage 4, particularly sub-lineage 4.2.2.2, is rapidly expanding in Ethiopia and is associated with a higher prevalence of multidrug-resistant (MDR) TB infections (Mekonnen et al., 2023), hinting at distinctive virulence traits.

In Ethiopia, the geographically confined *Mtb* sub-lineage 4.2.2.2 (4.2.2.ETH) is characterized by its successful and high transmissibility and frequent link to MDR (Mekonnen et al., 2019). In northwest Ethiopia (from 2020 to 2022), the dominant drug-resistant sub-lineage was L4.2.2.ETH, which accounted for 34.5% of cases, representing 50% of drug-resistant isolates (Mekonnen et al., 2023). These findings highlight L4.2.2.2 as a sub-lineage marked by high transmissibility and drug

resistance. The molecular mechanisms and factors driving the enhanced survival, resistance development, and widespread transmission of sub-lineage 4.2.2.2 remain poorly understood. Previous research suggests that the high transmissibility may be associated with bacterial hyper-virulence (Aiewsakun et al., 2021).

The expression levels of DosR in *Mtb* are linked to bacterial virulence (Badillo-López et al., 2010). The DosR regulon, consisting of 48 co-regulated genes, plays a vital role in *Mtb* persistence and is controlled by a two-component system involving the sensor kinases DosS (Rv3132c) and DosT (Rv2027c), along with the response regulator DosR (Rv3133c) (Chen et al., 2013). The continuous overexpression of the DosR regulon in W/Beijing strains under standard *in vitro* conditions, even in the absence of stimuli like nitric oxide, may account for the remarkable survival capabilities of these strains (Fallow, Domenech, & Reed, 2010). Highly transmissible *Mtb* strains are more likely to develop MDR or XDR, leading to numerous cases (Emane et al., 2021).

Laboratory diagnosis is essential for effective TB treatment. Sputum smear microscopy and sputum culture are the current gold standards for TB diagnosis, as recommended by the World Health Organization (WHO). While sputum smear microscopy is cost-effective and easy to perform, it has limitations, including low specificity (14-47%) and low sensitivity (requiring at least 10^3 bacterial cfu/mL) that make early TB diagnosis challenging. In contrast, sputum culture can accurately differentiate between dead and live *Mtb*, but it is labor-intensive, requires skilled personnel, and takes 2-8 weeks due to the slow doubling time of the bacteria, hindering its use for early diagnosis (He et al., 2017).

The period from 1950 to 1970 marked a pivotal moment in the fight against tuberculosis, as most of the current anti_TB drugs were discovered, and new treatment regimens made TB a curable disease. However, the initial optimism within the TB-control community began to fade as drug-resistant *Mtb* strains emerged (Caminero et al., 2010). Strains resistant to an expanding range of second-line drugs used to treat MDR-TB are increasingly posing a global public health threat (Shah et al., 2007).

TB that responds to standard drugs is curable, with most patients successfully treated in six months using an affordable regimen of rifampicin and isoniazid. However, multidrug-resistant TB (MDR-TB), which is resistant to both rifampicin and isoniazid, and extensively drug-resistant TB (XDR-

TB), which is resistant to those two drugs as well as fluoroquinolones and at least one second-line injectable (such as capreomycin, kanamycin, or amikacin), are far more challenging to treat. The current treatment for MDR-TB and XDR-TB involves a longer duration of therapy with additional drugs, resulting in higher toxicity, poorer outcomes, and significantly increased costs (Caminero et al., 2010; Nunn et al., 2014).

The recommended treatment regimen for MDR and XDR TB involves a combination of at least four drugs to which the *Mtb* strain is likely susceptible. These drugs are selected through a stepwise process across five groups based on efficacy, safety, and cost. The first group consists of oral first-line drugs, including high-dose isoniazid, pyrazinamide, and ethambutol, which serve as adjuncts in treating MDR and XDR TB. The second group includes fluoroquinolones, with high-dose moxifloxacin as the preferred choice. The third group is injectable drugs, starting with capreomycin, followed by kanamycin and then amikacin. The fourth group includes second-line drugs, to be used in the following order: thioamides, cycloserine, and aminosalicylic acid. The fifth group comprises drugs with limited effectiveness or clinical data, and should be used in the following sequence: clofazimine, amoxicillin with clavulanate, linezolid, carbapenems, thioacetazone, and clarithromycin (Caminero et al., 2010).

1.2 Statement of the Problem

Drug resistance in *Mtb* poses a major threat to current and future public health. As a highly adaptable and stealthy human pathogen, *Mtb* quickly develops resistance to commonly used anti-TB drugs. The rise of drug-resistant strains jeopardizes the efficacy of TB control programs globally (Miotto et al., 2022). Treatment success rates on a global scale range from 85% for drug-sensitive TB to 56% for MDR-TB, further dropping to 39% for XDR (Wan et al., 2023).

While well-characterized canonical mutations are known to drive high-level drug resistance in *Mtb*, emerging evidence suggests that the mechanisms of resistance are more complex than the acquisition of these mutations alone. Recent studies reveal that *Mtb* can develop non-canonical resistance-associated mutations, which confer survival advantages in the presence of certain drugs and may act as intermediates in the evolution toward high-level resistance (Martini et al., 2022).

The development of new drugs is urgently needed to reduce TB treatment duration and combat the increasing prevalence of drug-resistant strains. However, identifying effective targets for drug development remains a significant challenge due to our limited knowledge of how *Mtb* responds to disruptions in target functions, even with currently available drugs. The ability to predict targets for binding of with bactericidal agents is hindered by an incomplete understanding of the bacillus's physiology and its remarkable ability to adapt to changing environments (Boshoff et al., 2004).

As the host's response to *Mtb* infection has evolved, *Mtb* has simultaneously adapted to those immune defenses. The geographic confinement of specific strains and human populations may have created isolated regions of stable host-pathogen interactions, where the genomic characteristics of both the host and pathogen have reached a state of balance (Reed et al., 2007). Various studies have shown that geographically localized strains exhibit high transmission rates and an increased incidence of MDR cases (Emane et al., 2021; Thipkrua et al., 2025; Vyazovaya et al., 2020). Lineage and sublineage specific differences significantly influence the evolution of drug resistance in *Mtb*. Notably, the modern lineage 2 and one of its recently diversified sub-lineage 2.2 (Beijing lineage) tend to acquire de novo resistance mutations at elevated rates. These strains do so without incurring detectable fitness costs, which facilitates the persistence and transmission of resistant variants with in population (Silcocks et al., 2022)

According to the MTBC-6 model, sublineage 4.2.2.2 was likely introduced into Ethiopia around the 16th century, with a 95% highest posterior density interval estimating its arrival between 167 and 313 years ago. This sub-lineage is characterized by notable epidemiologic clustering (indicating recent transmission), frequent associations with drug resistance, evidence of increasing prevalence or expansion over time (Comas et al., 2015). The notable prevalence and geographic restriction of the L4.2.2.2 sub-lineage, along with its potential role in local tuberculosis transmission dynamics, suggest that this strain may be particularly adapted for efficient spread within the Ethiopian population. Furthermore, the high proportion of L4.2.2.2 clusters among MDR-TB patients underscores its clinical significance. The rising prevalence of this pathogenic strain could pose a major challenge to tuberculosis control efforts.

Given the limited genetic diversity among *Mtb* strains, phenotypic variations are believed to arise from differences in the regulation of biochemical networks by key transcriptional regulators

(Gomez-Gonzalez et al., 2019). This study explored these regulatory mechanisms by integrating transcriptomic profiles through RNA sequencing with whole-genome and phenotypic analyses, focusing on *Mtb* sub-lineage 4.2.2.2, which is notably more prevalent and drug-resistant than other lineages and sub-lineages. Gaining insights into its gene expression patterns, molecular characteristics, prevalence, transmissibility, and resistance mechanisms is essential. These findings provide valuable opportunities for understanding resistance mechanisms and identifying potential targets for developing new drugs or vaccines, supporting efforts to combat TB effectively.

This study also aimed to enhance TB treatment strategies by generating data on the growth kinetics and minimum inhibitory concentrations (MIC) of the first line anti TB drugs against the most prevalent *Mtb* lineages circulating in Ethiopia. By measuring bacterial load using colony-forming units (CFU), turbidity, and optical density at 600 nm (OD600), the research provides valuable insights into growth rates and resistance patterns, offering evidence for potential personalized TB treatment options.

1.3 Research Questions

- Is there a significant difference in growth kinetics and drug susceptibility among *Mtb* lineages in Ethiopia?
- Does gene expression significantly differ between the dominant drug-resistant *Mtb* sub-lineage 4.2.2.2 and other *Mtb* lineages?
- Is there a notable difference in gene expression between drug-resistant and drug-susceptible *Mtb* sub-lineage 4.2.2.2?

1.4 Objectives

1.4.1 General Objective

To better understand the pathogenesis and drug resistance mechanisms of *Mtb* lineages in Ethiopia, focusing on sub-lineage 4.2.2.2

1.4.2 Specific objectives

- To assess the growth kinetics and drug susceptibility of diverse *Mtb* lineages in Ethiopia to inform treatment strategies and improve TB management.
- To analyze gene expression and molecular traits of *Mtb* sub-lineage 4.2.2.2 using RNA sequencing and genomic analysis.
- To explore the differential gene expression and underlying mechanisms of drug resistance in Ethiopian *Mtb* sub-lineage 4.2.2.2 isolates

1.5 Significance of the Study

This study uncovered biologically relevant transcriptomic alterations, offering deeper insights into the pathogenesis of *Mycobacterium tuberculosis* (*Mtb*), particularly in drug-resistant strains. The identification of specific druggable targets and mechanisms provides promising avenues for addressing drug-resistant TB infections and advancing therapeutic strategies. Additionally, the observed differences in growth kinetics and drug susceptibility among various lineages highlight the potential for optimizing and personalizing TB treatment, ultimately enhancing disease management outcomes. The study also identified drug resistance genes that could serve as functional markers for drug resistance, equipping policymakers with tools for rapid diagnosis, treatment outcome prediction, and the refinement of therapeutic strategies. These findings lay the foundation for future research and the development of innovative agents to control disease progression, significantly contributing to global efforts in combating TB.

Chapter 2: Literature Review

2.1 Host-Pathogen Interactions and Immune Evasion Mechanisms of *Mycobacterium tuberculosis*

Mtb primarily resides within host macrophages during infection and spreads via respiratory droplets released when individuals with TB cough. Upon entering the airways of a susceptible host, *Mtb* is engulfed by alveolar macrophages, initiating an immune response that forms granulomas to contain the bacteria. The ability of *Mtb* to survive and proliferate within macrophages significantly influences the bacterial load in droplets and enhances its transmission potential (Aiewsakun et al., 2021).

Upon entering the alveolar passages of exposed individuals via aerosol droplets, *Mtb* is believed to first interact with resident macrophages. However, it is also possible that alveolar epithelial type II pneumocytes, which outnumber macrophages in the alveoli, may be the initial cells to engulf the bacteria. Dendritic cells also play a crucial role in the early stages of infection. As superior antigen-presenting cells compared to macrophages, dendritic cells are essential for activating T cells with specific *Mtb* antigens. Furthermore, their migratory nature, unlike the stationary differentiated macrophages, suggests that dendritic cells may significantly contribute to the dissemination of TB within the body (Smith, 2003).

During *Mtb* infection in the lungs, the progressive development of an immune response leads to the accumulation of immune cells, forming a structured aggregation known as a granuloma. This structure consists of macrophages at the core, surrounded by giant cells, T-lymphocytes, neutrophils, and fibroblasts. A defining characteristic of *Mtb* is its ability to remain hidden within host cells or within the caseous necrotic centers of granulomas, allowing it to persist throughout the prolonged latency phase of tuberculosis (Queval, Brosch, & Simeone, 2017).

Mtb employs multiple strategies to evade host immune responses and survive within macrophages. It prevents phagosome maturation, thereby avoiding the formation of bactericidal phagolysosomes. Unlike typical microbial ingestion, *Mtb* phagocytosis does not trigger the usual rise in cytosolic Ca^{2+} levels in macrophages, which inhibits phagosome-lysosome fusion and enhances bacterial survival. This failure is linked to *Mtb's* inhibition of sphingosine kinase, a critical enzyme for phagosome maturation, representing a novel molecular mechanism of

pathogenesis (Malik et al., 2003). Additionally, *Mtb's* preference for early endosomes reduces host immune surveillance. This is achieved by downregulating MHC-II protein expression and antigen presentation, limiting CD4+ T cell activation. These effects may be mediated by the secreted 9-kDa lipoprotein, which interacts with toll-like receptor (TLR)-2 during early bacterial entry (Smith, 2003).

2.2 Genetic and Phenotypic Diversity in *M. tuberculosis* Complex and Its Implications for Virulence and Drug Resistance

One of the most remarkable biochemical characteristics of this slow-growing bacterium is the high lipid content in its cell wall, comprising up to 60% of its dry weight, along with glycolipids and polysaccharides. Key cell wall components, including mycolic acids, mycocerosic acid, phenolthiocerol, lipoarabinomannan, and arabinogalactan, are produced through unique biosynthetic pathways. The elevated lipid content contributes to several distinctive properties, such as reduced permeability to stains, acid-fastness, exceptional resistance to acids and alkalis, prolonged survival, induction of inflammatory host responses, resistance to various bactericidal antibiotics, and overall pathogenesis. Additionally, *Mtb* demonstrates a pronounced preference for lipids as a nutrient source (Cole et al., 1998; Rachman & Kaufmann, 2007).

Unlike many bacteria, *Mtb* does not acquire drug resistance through plasmids or horizontal gene transfer. Instead, it relies on intrinsic mechanisms such as its lipid-rich cell wall, drug-modifying enzymes, efflux pumps, and target-modifying mutations. A notable feature of *Mtb's* physiology, critical to drug susceptibility, is its ability to alter gene expression and transcriptional regulation. Prolonged antibiotic exposure has driven *Mtb* to evolve complex resistance mechanisms, including post-translational modifications (PTMs). These PTMs enable rapid adaptation to drug pressure by modulating transcription, translation, metabolism, and enzyme activity. Approximately one-third of *Mtb* proteins undergo PTMs, influencing its growth, metabolism, gene regulation, and virulence (Miotto et al., 2022; Parsa & Hasnain, 2015; Sun, Ge, & Li, 2022).

Drug resistance in *Mtb* was initially explained by classical model based on Darwinian selection, where spontaneous, pre-existing mutation in the bacterial population are selected under drug pressure. This model suggested that resistant mutation inevitably emerge and dominate once susceptible bacteria are eliminated. However, this model fails to fully explain several observed

phenomena, such as relapse in drug susceptible TB and the frequent emergence of MDR (Jones et al., 2022). Gaining a deeper understanding of these resistance mechanisms is essential for designing drugs that can bypass existing resistance or target alternative biochemical pathways. Such insights will help prevent the emergence of resistance and identify genetic factors linked to resistance against new treatments (Johnson et al., 2006).

Although the MTBC is genetically uniform and exhibits less than a 0.05% sequence difference between lineages, this small percentage equates to approximately 2000 genetic variations within a 4Mb genome. Given that MTBC species typically carry around 4000 genes, it implies that nearly half of these genes could potentially be influenced by these genetic variations. While some of these alterations may be inconsequential, affecting noncoding regions or resulting in silent mutations, it's intriguing to consider the possibility that certain missense or nonsense mutations could potentially lead to lineage-specific traits or even impact the range of hosts for species adapted to both humans and animals (Yruela et al., 2016).

Apart from the genetic variations among clinical strains of MTBC, numerous experimental investigations have altered the prevailing viewpoint by revealing the tangible consequences of this genetic diversity. These studies have highlighted variations among clinical strains in terms of their transcriptomic profiles, protein and metabolite levels, methylation patterns, drug susceptibility, and cell wall composition, illustrating the phenotypic impact of genetic diversity(Borrell et al., 2019).

Various MTBC strains exhibit distinct cellular and clinical characteristics. Lineage 5 and lineage 6 strains demonstrate variations in metabolism, slower growth rates, and reduced virulence. Conversely, some lineage 2 and lineage 4 strains are more virulent in terms of disease severity and human to human transmission. Modern *Mtb* lineages exhibit rapid replication within human macrophages and often trigger a reduced and delayed pro inflammatory response upon infection. These phenotypic variations ultimately stem from genetic differences, with different genomic regions being associated with various cellular and clinical traits (Chae & Shin, 2018: Gagneux, 2017).

The enhanced virulence observed in certain Beijing strains has been associated with the production of a phenolic glycolipid (PGL) with immune-modulating properties. However, in many isolates, a

7bp deletion in the pks 15/1 gene prevents the synthesis of PGL (Homolka et al., 2010). PGLs are components of the cell wall in specific pathogenic mycobacteria, such as the W-Beijing family of MTB strains and *Mycobacterium leprae*, and they have the ability to modulate the immune system. PGLs function by reducing the antibacterial activity of nitric oxide synthase (iNOS) in macrophages infected with mycobacteria. This downregulation of iNOS is dependent on complement receptor 3 and is induced similarly by both bacterial and purified PGLs. Besides the decrease in iNOS, this process also leads to impaired production of TRIF-dependent IFN-β and CXCL10 in TLR4-stimulated macrophages. Hence, PGLs manipulate the immune response of host macrophages, affecting both their ability to combat bacteria and their inflammatory responses (Oldenburg et al., 2018).

Lipid metabolism plays a crucial role in the metabolic processes of MTB, enabling it to utilize host-derived lipids as an alternative source of nutrition during infection or as a survival and pathogenesis strategy. Numerous studies have highlighted variations in lipid metabolism among clinical strains of MTB belonging to different lineages within the MTBC. For example, The East-Asian and Euro-American lineages exhibit unique cell wall-associated lipids compared to the less transmissible Ethiopian lineage, potentially providing these lineages competitive advantage for phenotypic differences observed during infection. Within these lineages, mutations occurring in genes associated with propionyl-CoA metabolism, such as PrpR, have the potential to confer drug resistance and impact growth. Likewise, lineage-specific mutations in genes like cyp125 have repercussions on virulence and cytotoxicity (Moopanar & Mvubu, 2020).

Mycobacteria's unique qualities largely stem from mycolic acids, which play a crucial role in their growth, resilience, and pathogenicity (K. Raman, Rajagopalan, & Chandra, 2005). A recent investigation identified distinctive genetic mutations in the L7 lineage, specifically a stop-gain mutation in the mmaA3 gene responsible for encoding methoxy-mycolic acid synthase. These mutations have an impact on the production of cell wall lipids known as methoxy-mycolic acids. The study demonstrates that the absence of methoxy-mycolates in L7 has consequences, including disruptions in cell structure, changes in colony appearance on particular growth media, and an influence on biofilm formation. This deficiency might also lead to alterations in how L7 interacts with its host, potentially offering insights into the limited geographic distribution of L7 and

contributing to the understanding of the global dissemination of MTBC lineages (Hailu et al., 2023).

The outcomes of tuberculosis infection and disease vary widely and are primarily influenced by host factors, the genomic diversity of the MTBC, and environmental conditions. Different MTBC strains show variations in virulence and immunogenicity, with increasing evidence indicating that strain diversity may contribute to human TB. However, further research is needed to understand the extent to which MTBC genomic diversity affects disease progression in clinical settings and drug resistance (Mireilla Coscolla & Gagneux, 2010).

Chapter 3: Materials and Methods

3.1 Study M. tuberculosis strains

The *Mtb* isolates used for the present research work were obtained from active pulmonary TB patients enrolled in the tuberculosis genetic study (TBGEN study). A total of 36 clinically characterized *Mtb* isolates, identified through whole-genome sequencing for their genetic attributes, including drug resistance profiles, were obtained from the Armauer Hansen Research Institute (AHRI) TB laboratory. Of these, 11 samples were selected based on high-quality genome and RNA sequencing data, representing different *Mtb* sub-lineage prevalent in Ethiopia. To minimize lineage-related variability, we ensured a balanced distribution of both the number and drug resistance phenotypes of sub-lineage 4.2.2.2 between resistant and susceptible groups.

3.2 Study Setting

The 11 selected Mycobacterium tuberculosis isolates were obtained from various regions in Ethiopia, including Tulubolo, Hawassa, Adama, Gambella, Addis Ababa, Wolaita Sodo, and Wondo Genet. DNA extraction and sequencing were performed at the AHRI, where the isolates were also cultured. RNA extraction and quality assessment were conducted at Brighton and Sussex Medical School, while RNA sequencing was carried out at the University of Leeds. All bioinformatics and statistical analyses were completed at AHRI.

3.3 Standardization of Bacterial Load and In Vitro Analysis of *M. tuberculosis* Growth and Drug Response

To ensure all Mycobacterium tuberculosis strains, including clinical isolates and the H37Rv reference strain (ATCC 27294), began at a comparable growth phase, frozen stocks were thawed and cultured in MGIT tubes for 14 days. Following incubation, each isolate was adjusted to a standardized optical density (OD600) of 0.2. Isolates with higher initial concentrations were diluted with appropriate volumes of growth medium to achieve the target OD600.

3.3.1 *M. tuberculosis* Growth Kinetics

A 100 µL suspension of six isolates, including the H37Rv reference strain, each standardized to an OD600 of 0.2, was further cultured in duplicate. Each was grown in 10 mL of Middlebrook 7H9 broth (Difco, Detroit, MI, USA) supplemented with 10% OADC (oleic acid, albumin,

dextrose, and catalase; Becton Dickinson and Company, Sparks, MD, USA), 0.2% glycerol, and 0.05% Tween 80. The inoculated cultures were incubated at 37°C for 21 days in a Panasonic incubator (MCO-801C-PE). Visual inspection for contamination was performed within the first 24 h after inoculation (Sarkar et al., 2012).

To determine the colony-forming unit concentration (CFU/mL), a 100 µL bacterial suspension (from stock vials with an OD_{600} of 0.2) was serially diluted and inoculated in duplicate onto Middlebrook 7H11 agar plates (Difco, Detroit, MI, USA) supplemented with 0.5% glycerol and 10% OADC growth enrichment (Becton Dickinson, Cockysville, MD, USA). Throughout the 21-day of incubation, turbidity, CFU/mL, and OD_{600} measurements were taken every three days (Sarkar et al., 2012). The CFU/mL results were systematically correlated with McFarland standard data and OD_{600} readings. OD_{600} of the bacterial suspensions was measured using a BioSpectrophotometer (Eppendorf, AG, 22331), while the turbidity of the suspensions was evaluated through nephelometry (DEN-1). The *Mtb* colonies were carefully counted and adjusted according to the dilution factor.

3.3.2 Phenotypic drug susceptibility testing

The phenotypic drug susceptibility profiles of the isolates were assessed using the BACTEC MGIT™ 960 system with the SIRE kit. Frozen isolates were thawed and sub-cultured in MGIT tubes. The day the MGIT system signaled a positive result was designated as day 0. The tubes were then incubated for an additional day, and drug susceptibility testing (DST) was initiated on the following day. For signals detected on days 1 or 2, no dilution was necessary. However, for signals appearing between days 3 and 5, a 1:5 dilution with sterile saline was performed before DST inoculation. For growth detected after day 5, the samples were vortexed, diluted 1:100 with sterile saline, and 0.5 mL of the diluted sample was inoculated into an MGIT tube. All procedures were conducted in accordance with established standard operating protocols (S. Siddiqi & S. Rüsch-Gerdes, 2006; Tilahun et al., 2023).

Drug susceptibility testing (DST) for first- and second-line drugs was performed following the WHO technical manual for DST. The drug concentrations used were: streptomycin (STM) 1.0 µg/mL, isoniazid (INH) 0.1 µg/mL, rifampicin (RIF) 1.0 µg/mL, ethambutol (EMB) 5.0 µg/mL, bedaquiline 1.0 µg/mL, delamanid 0.06 µg/mL, moxifloxacin 0.25 µg/mL, levofloxacin 1.0

µg/mL, linezolid 1.0 µg/mL, ofloxacin 2.0 µg/mL, and clofazimine 1.0 µg/mL. The DST process was finalized when the growth control reached a growth unit (GU) value of 400. The result was recorded as "susceptible" if the GU value in the drug-containing tube was below 100 at this point. Conversely, if the GU value exceeded 100, the result was classified as "resistant." Results were invalidated if the growth control GU value reached 400 in fewer than 4 days or failed to reach 400 within 21 days, generating error codes X400 and X200, respectively. Quality assurance was maintained by testing each batch of MGIT medium and SIRE kits with the pan-susceptible reference strain *M. tuberculosis* H37Rv (S. H. Siddiqi & S. Rüsch-Gerdes, 2006).

3.3.3 Determination of Minimum inhibitory concentration

The minimum inhibitory concentrations (MICs) were determined as the lowest drug concentrations, following twofold serial dilution, that inhibited over 99.0% of the bacterial population of the tested *Mtb* clinical isolates. The assay was conducted on suitable broth method with an incubation period of 14 to 21 days at 37°C (Schönfeld et al., 2012). MIC of the first-line anti-TB drugs INH, RIF, EMB, and STR was determined using the microdilution broth method (Schön et al., 2021). Initially, *Mtb* clinical isolates were cultured and adjusted to a turbidity equivalent to the 1.0 McFarland standard (approximately 1.97×10^6 CFU/mL) by suspending bacterial colonies in sterile saline and visually comparing the turbidity to the standard. The bacterial suspension was then diluted in Middlebrook 7H9 broth to achieve the final inoculum density required for the assay (Peñuelas-Urquides et al., 2013).

Serial dilutions of each drug were prepared in sterile 96-well microtiter plates, starting with concentrations based on critical thresholds: 0.1 µg/mL for INH, 1 µg/mL for RIF, 5 µg/mL for EMB, and 2 µg/mL for STR. MIC assays were performed in triplicate. Each well received 10 µL of bacterial suspension, and the plates were incubated at 37°C for 7 to 14 days, depending on the growth rate of the strains [37]. Each plate included a growth control (without drug) and a sterile control (without inoculum). The growth controls consisted of a 1:100 dilution of the 10^{-2} inoculum of a 1 McF suspension (representing 1% of the inoculum in the drug-containing wells, GC1%) and an undiluted 10^{-2} suspension of the 1 McF inoculum (representing 100% of the inoculum in the drug-containing wells, GC100%) (Wikler, 2006). MIC assays were performed in triplicate and average value was recorded.

The MIC was identified as the lowest drug concentration that prevented visible bacterial growth, as shown by clear wells without turbidity. The findings were evaluated using the susceptibility thresholds outlined in the WHO's technical guidelines for the microdilution broth method.

3.3.4 Total RNA extraction

To achieve logarithmic-phase growth, thirty-six different *Mtb* sub-lineage isolates with an initial OD_{600} of 0.2 were selected. These strains were cultured at a 1:4 ratio (total volume of 32 mL) in Middlebrook 7H9 broth supplemented with 10% OADC enrichment, 0.5% (v/v) glycerol, and 0.05% Tween 80. The cultures were incubated at 37°C for 11 to 18 days until reaching the mid-logarithmic growth phase, with OD600 values between 0.3 and 0.5 (Sarkar et al., 2012).

A 30 mL mid-log phase culture was transferred to 50 mL screw-cap tubes. The tubes were centrifuged at 2,200 g at 4°C for 10 min to collect mycobacterial cells, discarding the supernatant. The resulting cell pellet was mixed with 1.2 mL TRIzol™ Reagent (Thermo Fisher Scientific, USA). The mixture was then transferred to a Lysing Matrix B tube (MP Biomedicals, Santa Ana, CA, USA) containing 0.1 mm silica beads. Bead beating was performed using a Fast Prep-24 5G instrument (MP Biomedicals) at a speed setting of 6.5 for 45 seconds, followed by chloroform clean-up as previously described (Waddell et al.; Wildner, Gould, & Waddell, 2018).

Samples were purified and DNase-treated using RNeasy® Mini Columns (Qiagen) following the manufacturer's protocol. The purified RNA samples were quantified using the Nanodrop Spectrophotometer and RNA quality was assessed using the Agilent 4150 TapeStation system. Quality of the RNA samples was ensured by including samples with an RNA integrity number (RIN) greater than 7 and an absorbance A260/A280 ratio between 1.8 and 2.1.

Ribosomal RNA was removed using the Illumina Ribo-Zero Plus Microbiome rRNA Depletion Kit (Illumina, USA) and libraries prepared for sequencing using TruSeq Stranded Total RNA kit designed for Illumina sequencing systems. Sequencing was performed using the Illumina NextSeq2000 platform, utilizing a P3 100-cycle protocol.

3.3.5 DNA extraction

For whole genome sequencing (WGS) of *Mtb*, high-quality DNA ranging from 0.5 – 5 µg with a concentration above 20 ng/µL was extracted using the RFLP method. In brief, mycobacterial

colonies grown on Lowenstein Jensen growth media were harvested into distilled water, followed by heat treatment and incubation with lysozyme, proteinase K, and NaCl/CTAB solutions. After DNA extraction with chloroform/isoamyl alcohol, the nucleic acids were precipitated with isopropanol and subsequently washed with ethanol. The DNA was air dried, dissolved in molecular-grade water, and integrity was assessed by agarose gel electrophoresis and Nanodrop spectrophotometer. DNA concentration was quantified using a Qubit fluorometer to ensure that the DNA met the necessary quality standards for sequencing (Dashti et al., 2009; Riaz et al., 2016).

3.4 Data Analysis

3.4.1 Growth Kinetics and Correlation analysis

Data analysis was conducted using GraphPad Prism version 10.00 for Windows (GraphPad Software, www.graphpad.com). Two-way ANOVA was employed to assess the growth curves and growth rates of *Mtb* clinical isolates, as well as the correlation between turbidity, CFU/mL, and OD600. Tukey's multiple comparisons test was subsequently used to evaluate differences in growth rates between strains. Statistical significance was determined using an adjusted p-value threshold of < 0.05.

3.4.2 Bioinformatic analyses

3.4.2.1 Whole genome analysis

The quality of the raw sequence reads was first evaluated using FASTQC (v0.12.1) (Yang et al., 2013). Trimmomatic (v0.39) was employed to remove low-quality reads, short fragments, and adapter sequences (Daniyarov et al., 2020). The filtered reads were then aligned to the H37Rv reference genome (NC 000962.3) using BWA (v0.7.18) (Li, 2013). SNPs were identified using the bcftools (v1.8)(Phelan et al., 2019). Drug resistance profiles and lineages were predicted with TBProfiler (v6.2.0) (Verboven et al., 2022). SNP based Phylogenetic trees were constructed using IQ-REE (v2.0.3) with maximum likelihood model by replicating 1000 bootstraps value, and annotated and visualized using interactive Tree of Life (iTOL) (v6.9) (Yu, 2020).

3.4.2.2 RNAseq analysis

FASTQC (v0.12.1) (Qi, Schlapbach, & Rehrauer, 2017) was used to assess the quality of the raw reads. Reads with low-quality sequences (Phred score < 20), short fragments (< 50 base-pairs), or adapter contamination were filtered and trimmed using Fastp (v0.23.4) (Mohideen, Johansen, & Babiak, 2020). Taxonomic identification and relative abundance of reads were determined using the Kraken software (Lu et al., 2017). Clean reads were then aligned to the complete genome of *Mtb* H37Rv (NC_000962.3) from the NCBI database using Hisat2 (v2.1.0) and quantified with featureCounts (v2.0.6) (Patro, Duggal, & Kingsford, 2015). Differential gene expression was analyzed using the DESeq2 package in R (v4.3.2), which included processes such as quantification, normalization, dispersion estimation, and transcript filtering (S. Liu et al., 2021). Differentially expressed genes (DEGs) were identified with a threshold of absolute log2(fold change) > 0.85. P-values were adjusted for multiple testing using the Benjamini-Hochberg procedure to control the false discovery rate (FDR) at $p < 0.05$ (Anders & Huber, 2012). Finally, Gene Ontology (GO) analysis, functional enrichment, and pathway analysis were conducted.

3.5 Ethical clearance

The study was approved by the Institutional Review Board (IRB) of the College of Health Sciences, Addis Ababa University (Approval No. 072/21/Sop). This study was part of the TBGENE project, which also received approval from the AHRI Institutional Ethics Review Committee (P031/18).

Chapter 4: Results

4.1. Characteristics of the *Mycobacterium tuberculosis* clinical isolates

After conducting whole genome and RNA quality analyses, we selected 11 *Mtb* isolates from a pool of 36 isolates, representing diverse lineages and sub-lineages. Our study included one isolate from lineage 3, six from sub-lineage 4.2.2.2, one from sub-lineage 7, one from sub-lineage 4.6.3, one from sub-lineage 4.1.2.1, and the reference strain H37Rv. Among these, eight isolates (including three from sub-lineage 4.2.2.2) were drug-sensitive, while the remaining three isolates from sub-lineage 4.2.2.2 exhibited drug resistance. Figure 1 presents their whole-genome phylogeny alongside their phenotypic drug sensitivity and resistance profiles.

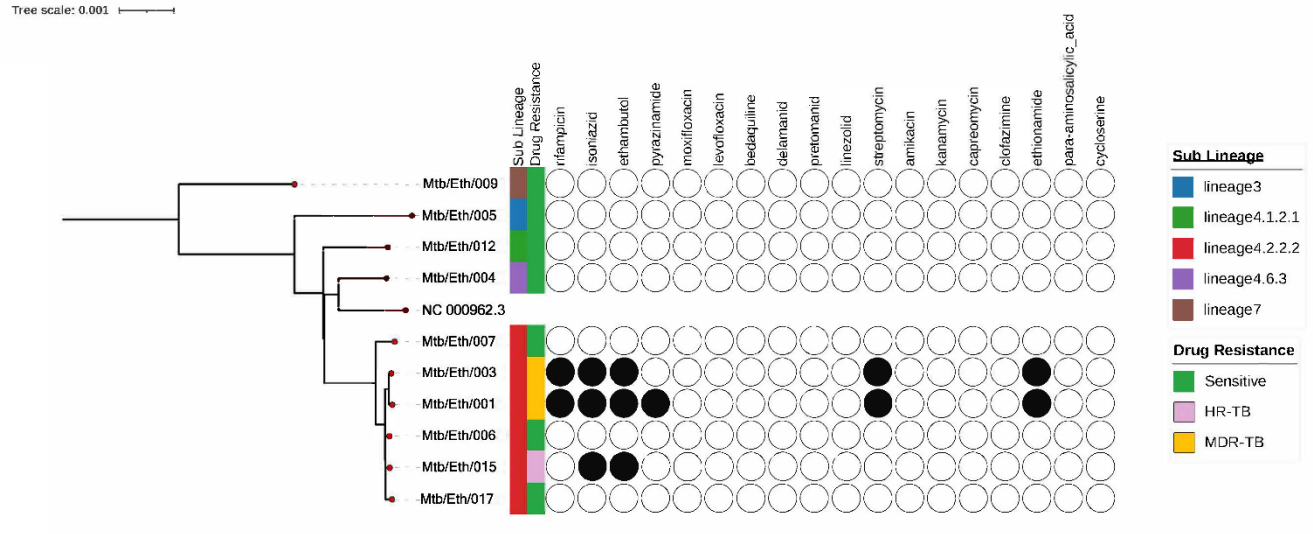

Figure 1: Phylogenic tree with drug susceptibility of *Mycobacterium tuberculosis* clinical isolates. NC 000962.3 is a *Mtb* H37Rv reference genome downloaded from NCBI. Shaded circles indicate a prediction of drug-resistance from the WGS analysis

4.2. Correlation between *Mtb* concentration measurement methods

A significant correlation was found between CFU/mL and OD600 measurements for all clinical isolates, with one lineage 3, one lineage 7 and three lineage 4 isolates (sub-lineages 4.2.2.2, 4.6.3, and 4.1.2.1), and H37Rv *Mtb* isolates, with an R^2 value of 0.9644 (Figure 2A). Additionally, strong correlations were observed between CFU/mL and the McFarland standard, as well as between OD600 and the McFarland standard, across all samples (R^2 values of 0.871 and 0.7452, respectively) (Figure 2B and 2C).

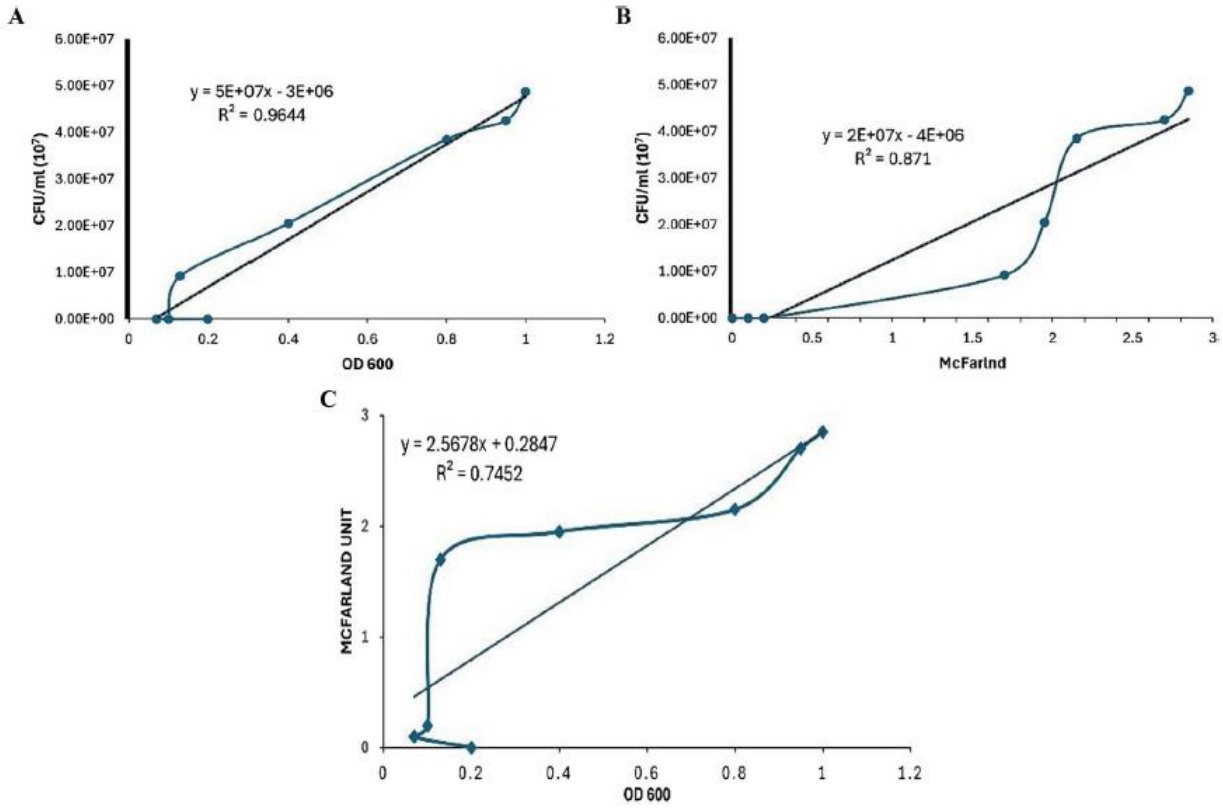

Figure 2: Comparison of the correlation between various *Mycobacterium tuberculosis* concentration measurement methods. **A**) Correlation between colony-forming units (CFU) per mL and optical density at 600 nm (OD600) **B**) Correlation between colony-forming units (CFU) per mL and McFarland standard. **C**) Correlation between McFarland standard and optical density at 600 nm (OD600)

4.3. *Mtb* growth kinetics

Mtb clinical isolates from different lineages (L3, L7, and three L4 isolates (sub-lineages 4.2.2.2, 4.6.3, and 4.1.2.1)) prevalent in Ethiopia and H37Rv were cultured in liquid 7H9 Middlebrook broth to evaluate growth rates across these lineages. Growth was observed in all *Mtb* clinical isolates starting from day 6 (Figure 3). Sub-lineage 4.1.2.1 and lineage 7 began growing on day 9, with sub-lineage 4.1.2.1 showing rapid growth kinetics (exponential phase) over the next six days (Figure 3). By day 15, all *Mtb* clinical isolates began to slow their growth and entered the stationary phase (Figure 3). When comparing Cmax (the maximum point on the growth curve) across

different sub-lineages, sub-lineage 4.1.2.1 exhibited a highest Cmax, indicating a greater bacterial concentration compared to sub-lineage 4.2.2.2, lineage 7, lineage 3, sub-lineage 4.6.3, and H37Rv.

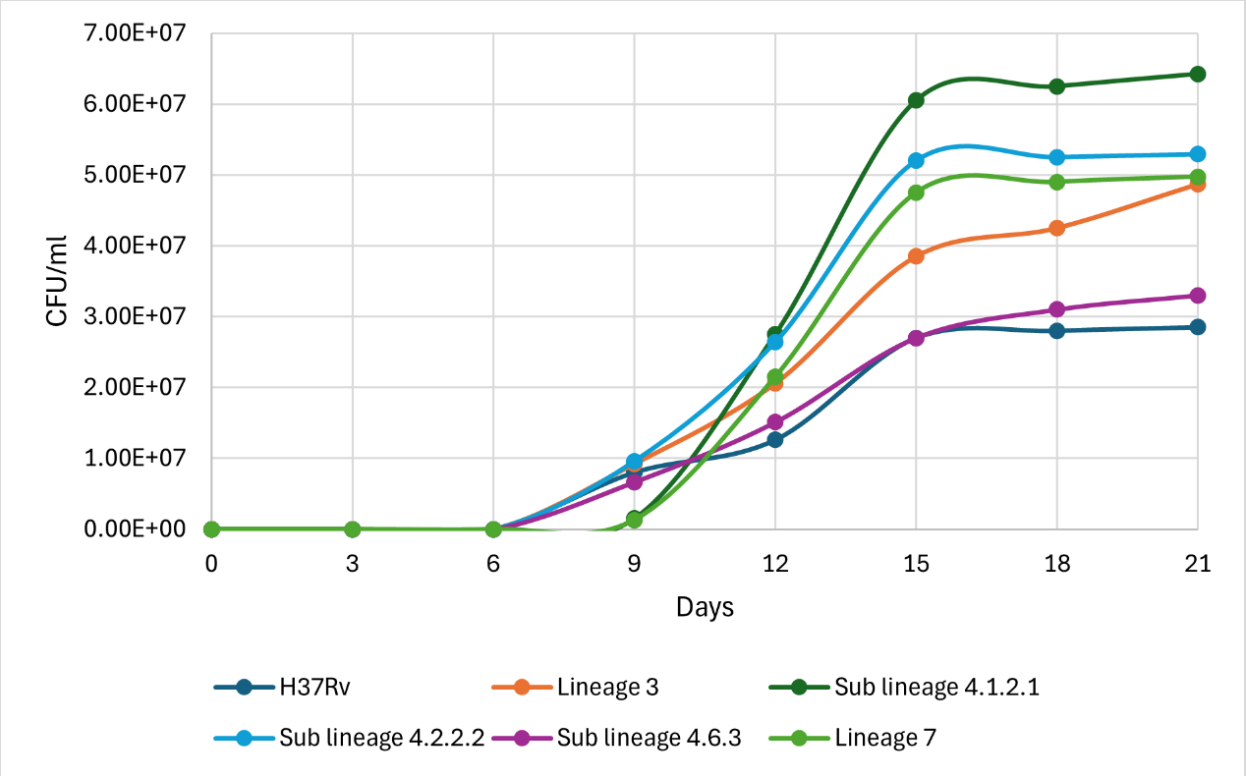

Figure 3: Growth kinetics of different clinical strains of *Mycobacterium tuberculosis* prevalent in Ethiopia, along with H37Rv (ATCC 27294)

4.3.1. Multiple comparison on growth rate difference

Sub-lineage 4.2.2.2 showed a significant difference in growth rate or CFU compared to all other strains in the study from day 9 or day 12 until the conclusion of the study (Table 1). Similarly, Sub-lineage 4.1.2.1 displayed a significant difference in growth rate or CFU compared to H37Rv, Sub-lineage 4.2.2.2, and Sub-lineage 4.6.3 starting from day 9 onward (Table 1). Other sub-lineages exhibited inconsistent but significant differences in growth rate or CFU when compared to each other on various study days. No significant differences in growth rate or CFU were observed between H37Rv and Sub-lineage 4.6.3 throughout the study period (Table 1).

Table 1. Turkey's multiple comparisons test for growth differences among *Mycobacterium tuberculosis* sub-lineages, measured by CFU over the study period.

Tukey's multiple comparisons test	Adjusted p values					
	Day 6	Day 9	Day 12	Day 15	Day 18	Day 21
H37rv vs. Sub-lineage 4.1.2.1	ns	ns	0.0001	0.0001	0.0001	0.0001
H37rv vs. Sub-lineage 4.2.2.2	ns	0.0105	0.0001	0.0001	0.0001	0.0001
H37rv vs. Sub-lineage 4.6.3	ns	ns	ns	ns	ns	ns
H37rv vs. Lineage 3	ns	ns	0.0009	0.0001	0.0001	0.0001
H37rv vs. Lineage 7	ns	0.0075	0.0002	0.0001	0.0001	0.0001
Sub-lineage 4.1.2.1 vs. Sub-lineage 4.2.2.2	ns	0.0007	ns	0.0004	0.0001	0.0001
Sub-lineage 4.1.2.1 vs. Sub-lineage 4.6.3	ns	ns	0.0001	0.0001	0.0001	0.0001
Sub-lineage 4.1.2.1vs. Lineage 3	ns	ns	ns	0.0001	0.0001	ns
Sub-lineage 4.1.2.1 vs. Lineage 7	ns	0.0005	ns	ns	ns	ns
Sub-lineage 4.2.2.2 vs. Sub-lineage 4.6.3	ns	ns	0.0001	0.0001	0.0001	0.0001
Sub-lineage 4.2.2.2 vs. Lineage 3	ns	0.0015	0.005	0.0001	0.0001	0.0001
Sub-lineage 4.2.2.2 vs Lineage 7	ns	ns	0.0219	0.0001	0.0001	0.0001
Sub-lineage 4.6.3 vs. Lineage 3	ns	ns	0.0477	0.0001	0.0001	0.0001
Sub-lineage 4.6.3 vs. Lineage 7	ns	ns	0.012	0.0001	0.0001	0.0001
Lineage 3 vs. Lineage 7	ns	0.001	ns	0.0001	0.0103	ns

Key: ns – not significant difference of bacterial concentration

4.4. Minimum Inhibitory Concentration

The study assessed the MIC for first-line anti-TB drugs against the *Mtb* clinical isolates (L3, L7, and three L4 isolates: sub-lineages 4.2.2.2, 4.6.3, and 4.1.2.1 prevalent in Ethiopia and H37Rv) using the microdilution broth method. While all isolates were susceptible to these drugs, the MIC values varied among the isolates (Table 2). Sub-lineages 4.1.2.1 and 4.2.2.2 had MIC values that corresponded to the critical concentration for most first-line anti-TB drugs, except for RIF, where the MIC was half of the critical concentration. In contrast, lineages 3 and 7 demonstrated

heightened sensitivity to RIF and STM, requiring only half of the standard critical concentrations used in the microdilution broth method. Additionally, lineage 3 showed increased sensitivity to INH, needing only half of the typical critical concentration established for the same method.

Table 2. Minimum Inhibitory Concentration for the first-line anti-TB drugs against all *Mycobacterium tuberculosis* clinical isolates, along with critical concentration for the microdilution broth method.

First line ant-TB drugs	Minimum Inhibitory Concentration (µg/ml)						
	Sub-lineage 4.1.2.1	Sub-lineage 4.2.2.2	Sub-lineage 4.6.3	Lineage 3	Lineage 7	H37Rv	Critical concentration
INH	0.1	0.1	0.1	0.05	0.1	0.1	0.1
RIF	0.5	0.5	0.5	0.25	0.25	0.5	1
EMB	4	4	4	4	4	4	5
STM	2	2	1	1	1	1	2

4.5. Transcriptomic profiles of *Mycobacterium tuberculosis*

Following quality control and trimming procedures, around 12 to 16 million single-ended reads p er sample were aligned to the reference genome sequence (H37Rv), yielding an average alignmen t rate of 85.48%. All samples were quantified for the same set of 3971 and 4008 genes with nonz ero counts for the comparison of sub-lineage 4.2.2.2 with other lineages, and between drug-resist ant and drug-sensitive sub-lineage 4.2.2.2, respectively.

4.6. Differentially Expressed Genes

Differential gene expression (DGE) analysis was conducted on eight *Mtb* isolates representing di verse lineages and sub-lineages—specifically, four from sub-lineage 4.2.2.2 and one each from li neages 3, 7, 4.6.3, and 4.1.2.1—to compare the transcriptional profiles of sub-lineage 4.2.2.2 with other groups. In addition, to investigate the expression differences between drug-resistant and dru g-sensitive strains, a separate DGE analysis was performed on six sub-lineage 4.2.2.2 isolates, co

mprising three drug-resistant (at least isoniazid-resistant) and three drug-sensitive isolates, as con firmed by both phenotypic drug susceptibility testing and whole genome sequencing.

The analysis identified seven DEGs in *Mtb* sub-lineage 4.2.2.2 compared to other lineages/sub-li neages (L7, L3, L4.1.2.1, and L4.6.3). Of these, six genes (Rv0331, Rv0720, Rv1993c, Rv2030c, Rv2034c, and Rv3129) were upregulated, while one gene (Rv0997a) was downregulated. These DEGs were visualized using clustered heat maps and volcano plots (Figures 4, Table 4). Notably, in sub-lineage 4.2.2.2, dormancy survival regulon (DosR) genes showed a slight but notable incre ase in expression compared to other lineages (Supplementary Table 1).

On the other hand, the analysis identified six genes with differential expression in the drug-resista nt group compared to the drug-susceptible group. All six genes Rv0096, Rv2780, Rv3136, Rv313 6A, Rv3137, and Rv3230c were downregulated in the resistant group. These DEGs were visualiz ed using clustered heat maps and volcano plots, as shown in Figures 4. The DEGs were classified by functional categories using http://www.geneontology.org/, as outlined in Table 3a and Table 3 b.

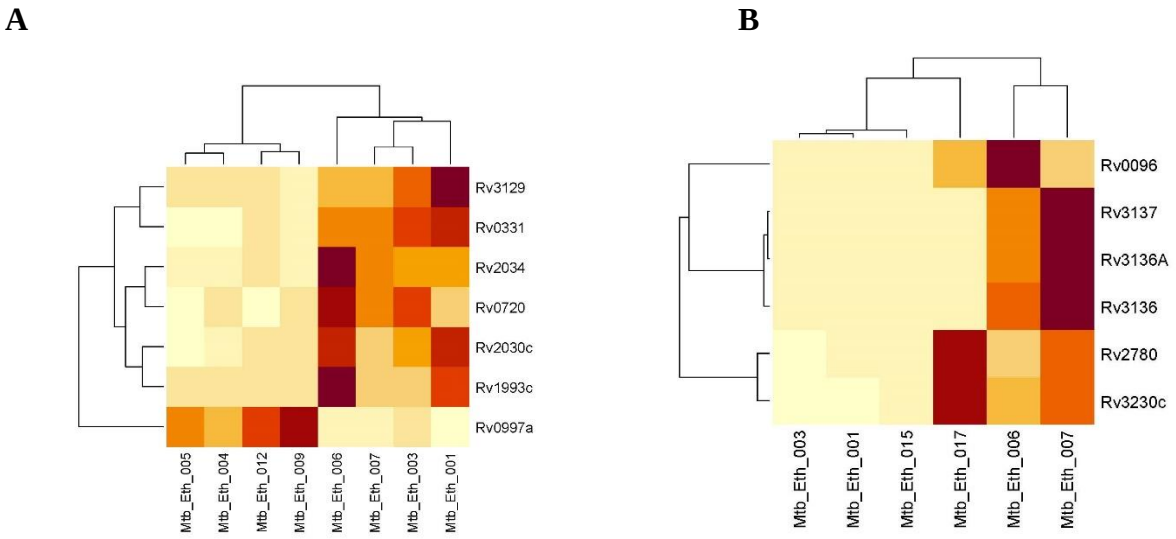

Figure 4: Clustered Heat map of differentially expressed genes: between sub-lineage 4.2.2.2 and other *Mtb* strains (**A**) and between drug resistance and drug-sensitive *Mtb* sub-lineage 4.2.2.2 (**B**). This clustered heat map illustrates gene expression levels across various samples, with each row representing a specific gene (e.g., Rv0096, Rv3230c) and each column corresponding to a sample or condition (e.g., Mtb_Eth_007, Mtb_Eth_001). The color gradient reflects expression levels,

ranging from light yellow for lower expression to dark red for higher expression. Hierarchical clustering is depicted by dendrograms along the top and left, grouping genes and samples based on similarities in their expression profiles. This clustering highlights patterns of gene expression, suggesting potential functional or regulatory connections between the genes and the isolates.

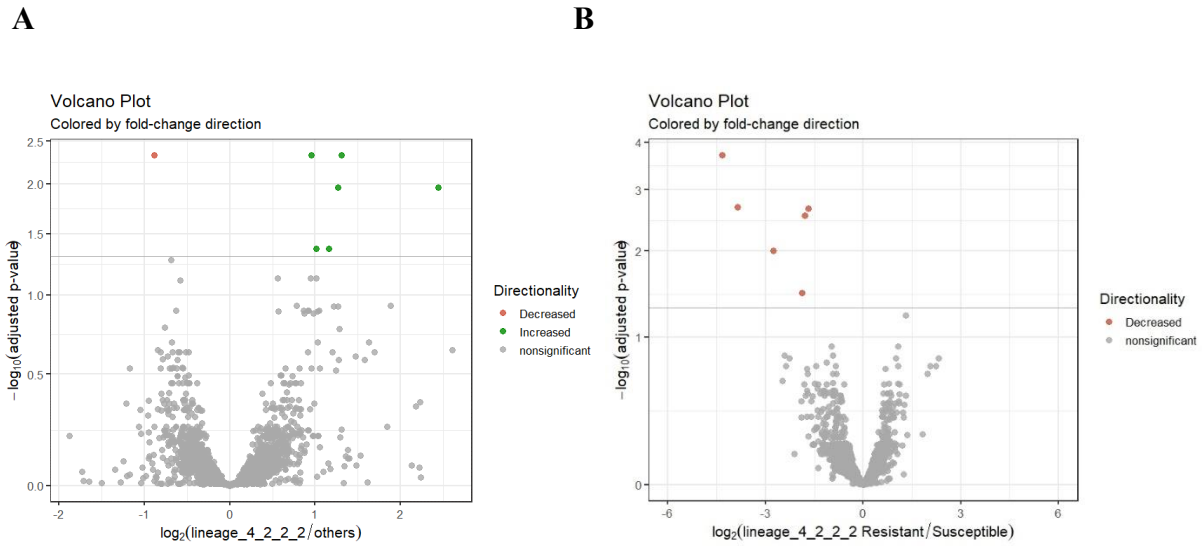

Figure 5: Volcano plot illustrating differentially expressed genes: between sub-lineage 4.2.2.2 and other *Mtb* strains (**A**) and between drug resistance and drug-sensitive *Mtb* sub-lineage 4.2.2.2 (**B**). The x-axis represents the log2 fold change, indicating the magnitude of change in gene expression, while the y-axis displays the −log10 p-value, reflecting the statistical significance of these changes. Positive x-axis values correspond to upregulated genes, whereas negative values indicate downregulated genes. The plot uses color to highlight gene significance: red points mark significantly downregulated genes with large fold changes and low p-values, green points identify significantly upregulated genes with similar criteria, and black points denote genes with no significant differential expression. Genes positioned farther from the origin and higher on the y-axis are both more statistically significant and exhibit greater expression changes, helping to pinpoint key genes relevant to the conditions being studied.

Table 3a. Functional categories of differentially expressed genes between sublineage 4.2.2.2 and other lineages/sub-lineages *Mtb* with their log2 fold change

Gene	Base Mean	Log 2Fold Change	p-value	p adj	Functional categories of genes

Rv0331	295.3324	0.955056	1.38e-06	0.0047	Possible dehydrogenase reductase
Rv0720	1527.1656	1.013034	6.87e-05	0.0427	50s ribosomal protein (RplR)
Rv0997a	436.3130	-0.885102	3.49e-06	0.0047	Hypothetical protein
Rv1993c	548.2774	2.437279	1.32e-05	0.0109	metallochaperone
Rv2030c	3736.0910	1.311698	3.59e-06	0.0047	Hypothetical protein
Rv2034	61.4288	1.163800	7.56e-05	0.0427	Transcriptional regulator and antigenic
Rv3129	1735.7153	1.265100	1.38e-05	0.0109	putative nitroreductase

Table 3b. Functional categories of differentially expressed genes in drug resistance *Mycobacterium tuberculosis* strains compared to drug sensitivity with their log2 fold change

Upregulated /downregulated gene	**Base Mean**	**log2Fold Change**	***p* values**	**Adjusted *p* values**	**Functional categories of genes**
Rv0096	180.4938	-1.879261	5.21e-05	0.0343	putative member of the PPE
Rv2780	1612.6887	-1.797150	2.94e-06	0.0029	L-alanine dehydrogenase (*ald*)
Rv3136	1925.3672	-3.848731	1.05e-06	0.0020	PPE51 protein
Rv3136A	785.9035	-4.322569	5.06e-08	0.0002	Hypothetical protein
Rv3137	543.0928	-2.754843	1.25e-05	0.0098	Histidinol phosphate phosphatase (HolPase)
Rv3230c	961.0902	-1.691440	1.65e-06	0.0022	oxidoreductase

4.7. Variant call analysis of differentially expressed genes

The variant call analysis identified 50 single nucleotide polymorphisms (SNPs) in the DEGs from the comparative study of sub lineage 4.2.2.2 versus other isolates (Supplementary Table 2). Among these, eight SNPs were distinct and consistently observed in sublineage 4.2.2.2 at specific chromosomal positions, along with the associated genes (Supplementary Table 2). Key SNPs were identified at positions 396634 (A>G), 401678 (C>A), and 401693 (C>G) in Rv0331; 811637 (C>A) in Rv0720; 2237059 (G>T) in Rv1993c; 2277272 (A>G) and 2280734 (G>A) in Rv2030c and Rv2034; and 3497369 (G>A) in Rv3129. Additionally, SNPs at positions 2282787 (C>T) and 2285251 (C>A) were consistently found across all isolates, regardless of lineage or phenotype, underscoring their potential significance. However, the analysis of DEGs identified no consistent SNPs at specific chromosomal positions across either drug-resistant or drug-susceptible strains.

4.8. Discordance of phenotypic and whole genome-based drug resistance determination in *Mycobacterium tuberculosis* clinical isolates

WGS was conducted on six *Mtb* sublineage 4.2.2.2 clinical isolates, including three drug-resistant and three drug-sensitive isolates. From WGS analysis, two isolates (Mtb/Eth/001 and Mtb/Eth/003) were predicted as MDR (INH and RIF resistance), while one isolate (Mtb/Eth/015) was predicted to be resistant to INH and EMB (Figure 1). All three resistant strains shared the same INH resistance conferring katG p.Ser315Thr mutation (Table 4). Additionally, the two isolates (Mtb/Eth/001 and Mtb/Eth/003) resistant to ethionamide carried an identical ethA p.Met1 mutation (loss of start codon). WGS analysis revealed distinct mutations associated with resistance to RIF in the two MDR-TB isolates (Table 4). Interestingly, phenotypic drug sensitivity testing indicated discrepancies, identifying only one isolate as RIF resistant and therefore MDR (isolate Mtb/Eth/001) as the WGS predicted RIF and STM resistance was not observed in phenotypic DST for isolate Mtb/Eth/003 (Table 5).

Table 4. Whole genome-based drug resistance test with corresponding mutation for the selected *Mycobacterium tuberculosis* sub-lineages 4.2.2.2

Samples	Drug resistance type	Num_dr_variants	Num_other_variants	Rifampicin	Isoniazid	Ethambutol	Pyrazinamide	Ethionamide	Streptomycin
Mtb/Eth/006	Sensitive	0	24	-	-	-	-	-	
Mtb/Eth/003	MDR-TB	6	25	rpoB p.His445Cys (1.00)	katG p.Ser315Thr (1.00)	embA c.-16C>T (1.00) embB p.Met306Ile (1.00)		ethA p.Met1? (1.00)	
Mtb/Eth/001	MDR-TB	8	58	rpoB p.Ser450Leu (1.00)	katG p.Ser315Thr (1.00)	embB p.Asp1024Asn (1.00) embB p.Gly406Ala (0.98)	pncA p.Val130Gly (1.00)	ethA p.Met1? (1.00)	gid p.Gly69Asp (1.00) rpsL p.Lys88Thr (1.00)
Mtb/Eth/015	HR-TB	2	32	-	katG p.Ser315Thr (1.00)	embB p.Met306Ile (1.00)	-	-	
Mtb/Eth/017	Sensitive	0	24	-	-	-	-	-	
Mtb/Eth/007	Sensitive	0	24	-	-	-	-	-	

Table 5. Phenotypic drug sensitivity of *Mycobacterium tuberculosis* 4.2.2.2 sub-lineage isolates. Asterisk (*) marks discrepancies between phenotypic DST results and genotypic WGS predictions.

Isolate identifier	Isoniazid	Rifampicin	Ethambutol	Streptomycin	Second line Anti Tb Drugs
Mtb_Eth_006	Sensitive	Sensitive	Sensitive	Sensitive	Sensitive
Mtb_Eth_007	Sensitive	Sensitive	Sensitive	Sensitive	Sensitive
Mtb_Eth_017	Sensitive	Sensitive	Sensitive	Sensitive	Sensitive
Mtb_Eth_001	Resistant	Resistant	Sensitive*	Resistant	sensitive
Mtb_Eth_003	Resistant	Sensitive*	Sensitive*	Sensitive	Sensitive
Mtb_Eth_015	Resistant	Sensitive	Sensitive*	Sensitive	Sensitive

Key: The second-line anti-TB drugs analyzed in this phenotypic study include bedaquiline, delamanid, moxifloxacin, levofloxacin, linezolid, ofloxacin, and clofazimine

Chapter 5: Discussion

Mtb demonstrates significant variation in growth kinetics, virulence, and drug susceptibility across different lineages and sub-lineages. After infecting the host airway, *Mtb* survives and multiplies within macrophages, influencing its transmission and virulence (Aiewsakun et al., 2021). The study showed that sub-lineages 4.2.2.2 and 4.1.2.1 exhibit notable growth advantages, higher bacillary loads, and faster replication rates compared to others, potentially contributing to their pathogenicity and widespread transmission. Sub-lineage 4.1.2.1 is globally distributed and represents approximately 10% of *Mtb* cases in Ethiopia, while sub-lineage 4.2.2.2, prevalent in Ethiopia (15%), is associated with heightened virulence and drug resistance.

Drug susceptibility tests using the microdilution broth method revealed lineage-specific differences in response to first-line anti-TB drugs. While all strains were susceptible, lineages 3 and 7 showed increased sensitivity to RIF and STM, requiring lower doses for inhibition. Lineage 3 also exhibited heightened sensitivity to INH. Notably, lineage 3 exhibited reduced growth kinetics, bacillary load, and MIC values compared to sub-lineages 4.2.2.2, 4.1.2.1, and L7. This increased sensitivity in lineage 3 may be linked to its lower bacillary load or the specific strain variants studied.

Lineage 3 accounts for 21% of *Mtb* cases in Ethiopia, with a 24.1% INH resistance rate. In contrast, lineage 4 is more prevalent (70%) and shows a significantly higher INH resistance rate of 74.1% (Getahun et al., 2024). Among lineage 4 sub-lineages, L4.2.2.2 is the most common drug-resistant group, responsible for 50% of resistant cases and 34.5% of MDR+ cases (Mekonnen et al., 2023). The substantial drug resistance in sub-lineage 4.2.2.2 poses significant challenges for standard dosing strategies, emphasizing the importance of developing tailored, lineage-specific treatment approaches.

A missense mutation in rpoB p.Ser450Leu aligns with phenotypic drug resistance in MDR isolate Mtb/Eth/001, whereas the rpoB p.His445Cys missense mutation did not confer phenotypic resistance to RIF in isolate Mtb/Eth/003. The rpoB p.Ser450Leu variant has been documented in the WHO Catalogue of Mutations in the MTBC and their association with drug resistance, demonstrating a sensitivity of 64.4% and a specificity of 99.3% for detecting RIF-resistant phenotypes. However, rpoB p.His445Cys variant has a sensitivity of 0.4% and a specificity of

100% for detecting RIF-resistant phenotypes (WHO, 2023). The substitution of the polar amino acid serine with the non-polar leucine at position 450 may disrupt RIF binding to rpoB, leading to drug resistance (Go & MIYAZAWA, 1980). No additional variants or SNP differences were detected in the rpoB gene between the two clinical isolates, both of which were classified as MDR-TB based on WGS analysis. The katG p.Ser315Thr missense mutation was associated with phenotypic drug resistance in all drug resistant isolates and is also reported in WHO Catalogue with a sensitivity of 77.8% and a specificity of 99.1% for detecting INH-resistant bacilli (WHO, 2023).

Our phenotypic drug sensitivity results indicate susceptibility to EMB (Mtb/Eth/001 and Mtb/Eth/003), whereas WGS-based drug sensitivity analysis identified resistance conferring mutations, with different mutations occurring at various positions for different samples. In line with our finding, routine phenotypic analysis failed to detect EMB resistance in 91.4% of resistant isolates, highlighting the challenges of EMB phenotypic testing. The inability of culture-based methods to reliably identify true EMB resistance negatively impacts TB control programs. However, a significant proportion of phenotypically EMB-resistant isolates (approximately 30%) still lack identifiable mutations in *embB*, underscoring the need for a comprehensive understanding of EMB resistance mechanisms in clinical isolates (Johnson et al., 2006). Differences between molecular and phenotypic EMB resistance results are likely due to limitations in conventional susceptibility testing methods (Plinke et al., 2009). Our study found higher drug resistance predicted through WGS compared to phenotypic drug sensitivity testing. This discrepancy may stem from the difficulties in generating accurate and reproducible drug sensitivity data from pathogenic clinical isolates in the laboratory. Given these discrepancies, an improved and an integrative approach combining phenotypic and genotypic methods is essential for accurately characterizing drug resistance and optimizing TB treatment strategies.

This study identified six downregulated genes (Rv0096, Rv2780, Rv3136, Rv3136A, Rv3137, and Rv3230c) in drug-resistant *Mtb* sub-lineage 4.2.2.2 isolates, none of which are direct targets of current anti-TB drugs. While INH and RIF resistance mechanisms are well-established and linked to mutations in *katG*, InhA, and the RIF resistance-determining region (RRDR) (Johnson et al., 2006), no common SNPs were found in differentially expressed genes within drug-resistant strains. This suggests that *Mtb* may employ non-SNP-related mechanisms to regulate gene expression and

drug resistance (Wan et al., 2023). Although the exact regulatory processes remain unclear, the findings underscore the complexity of gene expression changes contributing to drug resistance in *Mtb*.

This study also showed seven differentially expressed genes (DEGs) in *Mtb* sub-lineage 4.2.2.2 in comparison with other lineages and sub lineages, six of which (Rv0331, Rv0720, Rv1993c, Rv2030c, Rv2034, and Rv3129) were upregulated, while Rv0997a was downregulated. Additionally, unique SNPs specific to sub-lineage 4.2.2.2 and shared SNPs across all lineages were observed, though their role in gene expression differences remains unclear. These genetic variations may impact gene regulation by altering promoter regions, enhancers, or transcription factor binding sites. The identified DEGs, excluding Rv0997a, could contribute to sub-lineage 4.2.2.2's characteristics, including increased transmission, hyper-virulence, adaptability, persistence, and potential MDR.

This study highlights the reduced expression of two key genes, Rv0096 (PPE1) and Rv2780, in drug-resistant *Mtb* sub-lineage 4.2.2.2. Rv0096, part of the PPE family, is encoded within the rv0096–rv0101 operon (Sao Emani & Reiling, 2024), which synthesizes an isonitrile lipopeptide associated with biofilm formation and copper acquisition (Jinich et al., 2022). Downregulation of Rv0096 has been linked to drug resistance in RIF-resistant *Mtb* strains (de Knegt et al., 2013) and is associated with D-cycloserine resistance (Sao Emani & Reiling, 2024).

Rv2780 encodes L-alanine dehydrogenase, an enzyme critical for oxidative deamination of L-alanine and subsequent peptidoglycan synthesis, with additional roles in nitrogen and carbon metabolism (Sao Emani & Reiling, 2024). *Mtb* leverages Rv2780 to deplete L-alanine in macrophages, suppressing antimicrobial peptide (AMP) production and enhancing bacterial survival (Peng et al., 2024). Inhibition of Rv2780 has shown promise, as studies demonstrate that deletion or targeted inhibition of this gene improves resistance to D-cycloserine and boosts bacterial clearance when combined with L-alanine or RIF. The small-molecule and Rv2780 inhibitor GWP-042 significantly reduces the survival of *Mtb*, including MDR strains, making it a promising adjunctive therapy (Desjardins et al., 2016).

The Rv3136 gene, which encodes the PPE51 protein, is notably downregulated in drug-resistant *Mtb* strains. This gene is critical for disaccharide uptake, and mutations in its coding sequence confer resistance to bactericidal thio-disaccharides. Restoring a functional ppe50-ppe51 operon,

which includes Rv3136, reinstated drug sensitivity in resistant strains, underscoring its role in drug susceptibility (Korycka-Machała et al., 2020). Our study also observed a downregulation of Rv3135 (*ppe50*), with a log fold change of 2, though the difference was not statistically significant (supplementary table 2). The ppe50-ppe51 operon is critical for INH and RIF tolerance phenotypes in RNase J (Rv2752c) mutant *Mtb* strains. Deleting this operon induced drug tolerance and overexpression restored drug sensitivity, highlighting its role in drug susceptibility (Martini et al., 2022). Additionally, Bellerose identified a crucial role of *ppe50-ppe51* mutants in enhancing survival and contributing to the development of resistance (Bellerose, 2020).

Similarly, the Rv3136A gene, encoding a protein of unknown function (Ramakrishnan et al., 2015), was found to have reduced expression in INH-resistant samples compared to susceptible ones. The Rv3137 gene, part of the inositol monophosphatase (IMPase) family, was also downregulated in resistant strains. This gene encodes histidinol phosphate phosphatase (HolPase), vital for de novo histidine biosynthesis in *Mtb* (Jha et al., 2018). Histidine biosynthesis allows *Mtb* to evade host immune defenses, which deplete intracellular histidine via interferon gamma (IFN-γ)-mediated pathways (Dwivedy et al., 2021). Reduced expression of Rv3137 in INH-resistant strains and successful *Mtb* lineages, such as the MtZ strain, highlights its role in drug resistance and pathogenesis (Comín et al., 2023). Since humans lack a histidine biosynthesis pathway (Jha et al., 2018), targeting bacterial enzymes in this pathway might offers a promising approach for anti-tuberculosis drug development.

The Rv3230c gene encodes an oxidoreductase that, in partnership with the stearoyl-CoA desaturase DesA3, converts saturated stearic acid into unsaturated oleic acid using molecular oxygen and NADPH. Oleic acid is essential for maintaining membrane fluidity, physiological functions, and mycolic acid synthesis, which are crucial for *Mtb's* cell membrane and cell wall integrity (Rehberg et al., 2019). A genomic study identified a 1,285 bp deletion within the desA3 and Rv3230c genes in extensively drug-resistant tuberculosis strains of the Proto-Beijing genotype. This deletion is linked to resistance against the second-line anti-TB drug isoxyl and may influence virulence and evolutionary adaptations (Srilohasin et al., 2020). Additionally, a 1-bp deletion in Rv3230c was associated with clustered strains harboring the katG-S315T mutation, a marker of isoniazid resistance (Hang et al., 2019).

A gene network analysis indicated that alternative sigma factors might regulate the downregulated genes. Specifically, sigma M (SigM) suppresses the expression of Rv3136 and Rv3137 while also influencing the regulation of Rv0096 (S. Raman et al., 2006). As a member of the extracytoplasmic function subfamily of alternative sigma factors, *Mtb* SigM is expressed at low levels in vitro and does not appear to play a role in stress response regulation. Instead, SigM positively regulates genes involved in the synthesis of surface or secreted molecules. Its role in repressing virulence-associated surface lipids while upregulating Esx family secreted proteins and nonribosomal peptide synthetase genes suggests that SigM may be involved in long-term adaptation to specific host environments during infection (S. Raman et al., 2006). The regulation of three out of six genes by SigM implies a potential role in the adaptation to drug resistance in lineage 4.2.2.2. Additionally, the alternative sigma factor SigD, encoded by Rv3414c, negatively regulates Rv3230c (S. Raman et al., 2004). Further research should focus on elucidating the mechanisms underlying anti-TB drug resistance, particularly pathways differentially regulated in drug-resistant isolates that may help *Mtb* adapt to metabolic changes caused by drug-resistance-conferring mutations.

The Rv2034 gene, overexpressed in *Mtb* sub-lineage 4.2.2.2 when compared to other lineages/sublineages, encodes the arsenic transcriptional repressor (ArsR) protein. This immunogenic protein regulates lipid metabolism, hypoxic stress responses, and dormancy. It positively influences the expression of key genes like phoP, groEL2, and DosR, which are critical for *Mtb's* virulence, stress adaptation, and persistence (Gao, Yang, & He, 2011, 2012). PhoP regulates lipid biosynthesis for the cell envelope and contributes to virulence, stress responses, and hypoxia adaptation via crosstalk with DosR (Walters et al., 2006). Elevated PhoP levels in virulent strains highlight its pathogenic significance (WU et al., 2014). Rv2034 also controls GroEL2, whose cleaved form impairs dendritic cell (DC) functions, weakening the host immune response and facilitating *Mtb* survival (Georgieva et al., 2018). Rv2034 regulates additional genes, such as Rv2622, a methyltransferase contributing to drug resistance (Cappelli et al., 2006), and Rv1997, encoding CtpF, which manages calcium efflux to influence survival mechanisms like autophagy (Garg et al., 2020). Furthermore, Rv2034 directly regulates Rv3133c, encoding DosR, essential for dormancy and stress responses (Gao, Yang, & He, 2012).

Variant analysis revealed unique single nucleotide polymorphisms (SNPs) upstream of Rv2034 in sub-lineage 4.2.2.2, potentially impacting genes involved in hypoxia, virulence, and stress adaptation. These findings suggest Rv2034's pivotal regulatory role in sub-lineage-specific phenotypes, enhancing adaptation to hostile environments (Hu & Coates, 2011; Kumar et al., 2007; Kumari et al., 2020; Rubinstein et al., 2023; Snášel et al., 2021). Rv2034's regulatory and immunogenic properties underscore its potential as a target for vaccine and drug development. However, further research is necessary to elucidate its precise mechanisms and connections to the phenotypic traits of sub-lineage 4.2.2.2. This could advance our understanding of *Mtb's* pathogenesis and resistance strategies, enabling the development of novel therapeutic approaches.

This study highlighted the significant upregulation of dormancy-associated genes, particularly Rv2030c and Rv3129, in *Mtb* sub-lineage 4.2.2.2, alongside increased expression of several DosR genes. These findings suggest molecular adaptations that enhance this sub-lineage's survival and persistence under stress. Rv2030c encodes a protein induced by hypoxia, nitric oxide (NO), and carbon monoxide (CO), containing a phosphoribosyltransferase (PRT) domain essential for nucleotide metabolism and salvage pathway (Mushtaq et al., 2015). It inhibits autophagy via histone modifications, including hypermethylation and deacetylation, disrupting host immune responses (Sengupta et al., 2021). Rv2030c also belongs to the erythromycin esterase superfamily, contributing to erythromycin tolerance (Mushtaq et al., 2015). Unique single nucleotide polymorphisms (SNPs) upstream of Rv2030c were identified in sub-lineage 4.2.2.2, potentially influencing genes like DosT, pfkB, and hspX, which play roles in virulence and stress adaptation (Hu & Coates, 2011; Kumar et al., 2007; Kumari et al., 2020; Rubinstein et al., 2023; Snášel et al., 2021).

Rv3129, encoding a nitroreductase, is highly expressed and crucial for nitrate metabolism and pyridoxal 5′-phosphate (PLP) salvage, essential for *Mtb* survival (Selvaraj et al., 2012). PLP, a cofactor in numerous biochemical reactions, may act as an antioxidant, aiding *Mtb* resilience (Ankisettypalli et al., 2016). PLP levels decrease in response to antibiotics like isoniazid, prompting *Mtb* to upregulate pyridoxamine 5′-phosphate oxidases to counteract this stress (Ebadi, Gessert, & Al-Sayegh, 1982; Prosser & de Carvalho, 2013). A distinct SNP in Rv3129 and its neighboring genes, such as tgs1 (Rv3130c) and the DevR-DevS operon, further supports the unique phenotypic traits of sub-lineage 4.2.2.2. The DevR-DevS operon, regulated by Rv3134c, is

pivotal for *Mtb's* response to hypoxia and nitric oxide, activating the dormancy regulon and enabling survival under hostile conditions (Bagchi et al., 2005). Additionally, genes like Rv3130c, a triacylglyceride synthase, are linked to lipid storage and stress adaptation, aligning with studies showing triacylglyceride accumulation in Beijing strains (Kwon et al., 2017).

DosR expression is closely linked to bacterial virulence. In the *Mtb* Beijing strain, dosR is upregulated during exponential growth but downregulated under hypoxic conditions. Conversely, H37Rv and *M. canettii* exhibit lower dosR expression during exponential growth but show increased expression under hypoxia compared to the *Mtb* Beijing strain (Badillo-López et al., 2010). The observed upregulation of DosR-regulated genes during exponential growth in our study may play a role in the enhanced survival, virulence, and unique phenotypic traits of sub-lineage 4.2.2.2.

The gene Rv0331 encodes a dehydrogenase/reductase enzyme involved in sulfur metabolism during biofilm formation. Biofilms, clusters of microorganisms adhering to surfaces, create an extracellular matrix that includes polysaccharides, proteins, and DNA. These biofilms help *Mtb* survive antimicrobial treatments by harboring drug-tolerant populations, even under high antibiotic concentrations. Additionally, *Mtb* in biofilms displays resistance to host immune system modulation, enhancing its persistence (Ma et al., 2021). Unique missense SNPs were identified in sub-lineage 4.2.2.2 at chromosomal positions 401678 (C>A) and 401693 (C>G), potentially affecting genes such as PE6, Rv0331, and others involved in *Mtb* virulence and adaptation. However, no significant expression differences were noted in the study, indicating a need for further investigation into these SNPs' impact.

The Rv1993 gene encodes a metallochaperone associated with metal transport, helping *Mtb* expel cadmium (Cd) and lead (Pb) (Mehra & Kaushal, 2009; Neyrolles, Mintz, & Catty, 2013). A distinct missense SNP at chromosome position 2237059 (G>T) was identified in sub-lineage 4.2.2.2, affecting several genes involved in metal ion regulation and stress response. In particular, Rv1994c encodes CmtR, a repressor sensitive to cadmium and lead, which is activated under oxidative stress and starvation. De-repression of CmtR activates transporters like ctpG and Rv1993, which may serve as potential therapeutic targets for regulating metal ions (Mehra & Kaushal, 2009; Neyrolles, Mintz, & Catty, 2013). Other genes, such as Rv1997 (CtpF), are involved in calcium efflux,

influencing host responses (Garg et al., 2020), while Rv1999c may aid in nitrogen acquisition, crucial for *Mtb* metabolism and immune evasion (Gouzy, Poquet, & Neyrolles, 2014).

Rv0720 encodes an antigenic 50S ribosomal protein L18 (RPLR) involved in protein synthesis and folding (Kerns et al., 2014), with significantly higher expression in sub-lineage 4.2.2.2. RPLR has been linked to rifampicin resistance and may serve as a potential drug target for extensively drug-resistant *Mtb* (Zahra, Jamil, & Uddin, 2021). A unique missense SNP at chromosome position 811637 (C>A) affects several ribosomal genes, indicating a possible role in the sub-lineage's drug resistance. Additionally, the resilience regulator (ResR)/mycobacterial cell division regulator (McdR) is involved in the rplN operon control, which impacts ribosomal protein production and antibiotic resistance (Cortes & Cox, 2015). Mutations in ResR are associated with faster recovery from drug exposure, contributing to antibiotic treatment failure and resistance development (Q. Liu et al., 2022).

5.1. Limitations of the study

This study provides important insights into the genetic expression influencing the phenotypic traits and drug resistance mechanisms of *Mtb* sub-lineage 4.2.2.2, but it has certain limitations. The sample size was relatively small due to high experimental costs and difficulties in obtaining representative samples from various *Mtb* lineages. The findings of this study have not yet been validated through functional assays, highlighting the need for future studies to confirm their biological relevance.

5.2. Conclusions

These findings highlight the importance of considering lineage-specific differences in *Mtb* strains for optimizing treatment regimens and improving TB control strategies, particularly in regions with diverse *Mtb* populations like Ethiopia.

The differentially expressed genes in sub-lineage 4.2.2.2 and drug-resistant isolates, along with their associated networks, facilitate this sub-lineage’s ability to adapt and evade immune defenses. These findings provide valuable insights into the factors influencing its physiological adaptation, transmissibility, and resistance mechanisms.

These findings underscore the value of combining phenotypic, genomic, and transcriptomic data to deepen our understanding of drug resistance mechanisms and uncover promising drug targets. They provide crucial insights into the molecular biology of *Mtb*, paving the way for the development of more effective therapeutics and vaccines. Future studies should prioritize exploring the specific roles of these unique DEGs in TB pathogenesis and drug resistance to advance targeted intervention strategies.

References

Aiewsakun, P., Prombutara, P., Siregar, T. A. P., Laopanupong, T., Kanjanasirirat, P., Khumpanied, T., . . . Srilohasin, P. (2021). Transcriptional response to the host cell environment of a multidrug-resistant Mycobacterium tuberculosis clonal outbreak Beijing strain reveals its pathogenic features. *Scientific reports, 11*(1), 3199.

Anders, S., & Huber, W. (2012). Differential expression of RNA-Seq data at the gene level–the DESeq package. *Heidelberg, Germany: European Molecular Biology Laboratory (EMBL), 10*, f1000research.

Ankisettypalli, K., Cheng, J. J. Y., Baker, E. N., & Bashiri, G. (2016). PdxH proteins of mycobacteria are typical members of the classical pyridoxine/pyridoxamine 5′-phosphate oxidase family. *FEBS letters, 590*(4), 453-460.

Badillo-López, C., González-Mejía, A., Helguera-Repetto, A. C., Salas-Rangel, L. P., Rivera-Gutiérrez, S., Cerna-Cortés, J. F., & González-y-Merchand, J. A. (2010). Differential expression of dnaA and dosR genes among members of the Mycobacterium tuberculosis complex under oxic and hypoxic conditions. *Int Microbiol, 13*(1), 9-13.

Bagchi, G., Chauhan, S., Sharma, D., & Tyagi, J. S. (2005). Transcription and autoregulation of the Rv3134c-devR-devS operon of Mycobacterium tuberculosis. *Microbiology, 151*(12), 4045-4053.

Bellerose, M. M. (2020). *Genetic Identification of Novel Mycobacterium tuberculosis Susceptibility and Survival Mechanisms During Antibiotic Treatment.*

Borrell, S., Trauner, A., Brites, D., Rigouts, L., Loiseau, C., Coscolla, M., . . . Kato-Maeda, M. (2019). Reference set of Mycobacterium tuberculosis clinical strains: A tool for research and product development. *PLoS one, 14*(3), e0214088.

Boshoff, H. I., Myers, T. G., Copp, B. R., McNeil, M. R., Wilson, M. A., & Barry, C. E. (2004). The transcriptional responses of Mycobacterium tuberculosis to inhibitors of metabolism: novel insights into drug mechanisms of action. *Journal of Biological Chemistry, 279*(38), 40174-40184.

Caminero, J. A., Sotgiu, G., Zumla, A., & Migliori, G. B. (2010). Best drug treatment for multidrug-resistant and extensively drug-resistant tuberculosis. *The Lancet infectious diseases, 10*(9), 621-629.

Cappelli, G., Volpe, E., Grassi, M., Liseo, B., Colizzi, V., & Mariani, F. (2006). Profiling of Mycobacterium tuberculosis gene expression during human macrophage infection: upregulation of the alternative sigma factor G, a group of transcriptional regulators, and proteins with unknown function. *Research in microbiology, 157*(5), 445-455.

Cerezo-Cortés, M. I., Rodríguez-Castillo, J. G., Mata-Espinosa, D. A., Bini, E. I., Barrios-Payan, J., Zatarain-Barrón, Z. L., . . . Del Portillo, P. (2022). Close related drug-resistance Beijing isolates of Mycobacterium tuberculosis reveal a different transcriptomic signature in a murine disease progression model. *International Journal of Molecular Sciences, 23*(9), 5157.

Chen, T., He, L., Deng, W., & Xie, J. (2013). The Mycobacterium DosR regulon structure and diversity revealed by comparative genomic analysis. *Journal of cellular biochemistry, 114*(1), 1-6.

Cole, S., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., . . . Barry Iii, C. (1998). Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature, 393*(6685), 537.

Comas, I., Hailu, E., Kiros, T., Bekele, S., Mekonnen, W., Gumi, B., . . . Robertson, B. D. (2015). Population genomics of Mycobacterium tuberculosis in Ethiopia contradicts the virgin soil hypothesis for human tuberculosis in Sub-Saharan Africa. *Current Biology, 25*(24), 3260-3266.

Comín, J., Campos, E., Gonzalo-Asensio, J., & Samper, S. (2023). Transcriptomic profile of the most successful Mycobacterium tuberculosis strain in Aragon, the MtZ strain, during exponential and stationary growth phases. *Microbiology spectrum, 11*(6), e04685-04622.

Coppola, M., & Ottenhoff, T. H. (2018). *Genome wide approaches discover novel Mycobacterium tuberculosis antigens as correlates of infection, disease, immunity and targets for vaccination.* Paper presented at the Seminars in immunology.

Cortes, T., & Cox, R. A. (2015). Transcription and translation of the rpsJ, rplN and rRNA operons of the tubercle bacillus. *Microbiology, 161*(Pt_4), 719-728.

Coscolla, M., & Gagneux, S. (2010). Does M. tuberculosis genomic diversity explain disease diversity? *Drug Discovery Today: Disease Mechanisms, 7*(1), e43-e59.

Coscolla, M., & Gagneux, S. (2014). *Consequences of genomic diversity in Mycobacterium tuberculosis.* Paper presented at the Seminars in immunology.

Coscolla, M., Gagneux, S., Menardo, F., Loiseau, C., Ruiz-Rodriguez, P., Borrell, S., . . . Sánchez-Busó, L. (2021). Phylogenomics of Mycobacterium africanum reveals a new lineage and a complex evolutionary history. *Microbial Genomics, 7*(2), 000477.

Daniyarov, A., Molkenov, A., Rakhimova, S., Akhmetova, A., Nurkina, Z., Yerezhepov, D., . . . Akilzhanova, A. (2020). Whole genome sequence data of Mycobacterium tuberculosis XDR strain, isolated from patient in Kazakhstan. *Data in Brief, 33*, 106416.

Dashti, A. A., Jadaon, M. M., Abdulsamad, A. M., & Dashti, H. M. (2009). Heat treatment of bacteria: a simple method of DNA extraction for molecular techniques. *Kuwait Med J, 41*(2), 117-122.

de Knegt, G. J., Bruning, O., ten Kate, M. T., de Jong, M., van Belkum, A., Endtz, H. P., . . . de Steenwinkel, J. E. (2013). Rifampicin-induced transcriptome response in rifampicin-resistant Mycobacterium tuberculosis. *Tuberculosis, 93*(1), 96-101.

Desjardins, C. A., Cohen, K. A., Munsamy, V., Abeel, T., Maharaj, K., Walker, B. J., . . . Salazar, A. (2016). Genomic and functional analyses of Mycobacterium tuberculosis strains implicate ald in D-cycloserine resistance. *Nature genetics, 48*(5), 544-551.

Dwivedy, A., Ashraf, A., Jha, B., Kumar, D., Agarwal, N., & Biswal, B. K. (2021). De novo histidine biosynthesis protects Mycobacterium tuberculosis from host IFN-γ mediated histidine starvation. *Communications Biology, 4*(1), 410.

Ebadi, M., Gessert, C., & Al-Sayegh, A. (1982). Drug-pyridoxal phosphate interactions. *Drug Metabolism and Drug Interactions, 4*(4), 289-332.

Emane, A. K. A., Guo, X., Takiff, H. E., & Liu, S. (2021). Highly transmitted M. tuberculosis strains are more likely to evolve MDR/XDR and cause outbreaks, but what makes them highly transmitted? *Tuberculosis, 129*, 102092.

Fallow, A., Domenech, P., & Reed, M. B. (2010). Strains of the East Asian (W/Beijing) lineage of Mycobacterium tuberculosis are DosS/DosT-DosR two-component regulatory system natural mutants. *Journal of bacteriology, 192*(8), 2228-2238.

Gao, C.-H., Yang, M., & He, Z.-G. (2011). An ArsR-like transcriptional factor recognizes a conserved sequence motif and positively regulates the expression of phoP in mycobacteria. *Biochemical and biophysical research communications, 411*(4), 726-731.

Gao, C.-h., Yang, M., & He, Z.-G. (2012). Characterization of a novel ArsR-like regulator encoded by Rv2034 in Mycobacterium tuberculosis. *Plos one, 7*(4), e36255.

Garg, R., Borbora, S. M., Bansia, H., Rao, S., Singh, P., Verma, R., . . . Nagaraja, V. (2020). Mycobacterium tuberculosis calcium pump CtpF modulates the autophagosome in an mTOR-dependent manner. *Frontiers in cellular and infection microbiology, 10*, 461.

Georgieva, M., Sia, J. K., Bizzell, E., Madan-Lala, R., & Rengarajan, J. (2018). Mycobacterium tuberculosis GroEL2 modulates dendritic cell responses. *Infection and immunity, 86*(2), 10.1128/iai. 00387-00317.

Getahun, M., Beyene, D., Mollalign, H., Diriba, G., Tesfaye, E., Yenew, B., . . . Ameni, G. (2024). Population structure and spatial distribution of Mycobacterium tuberculosis in Ethiopia. *Scientific reports, 14*(1), 10455.

Go, M., & MIYAZAWA, S. (1980). Relationship between mutability, polarity and exteriority of amino acid residues in protein evolution. *International journal of peptide and protein research, 15*(3), 211-224.

Gomez-Gonzalez, P. J., Andreu, N., Phelan, J. E., de Sessions, P. F., Glynn, J. R., Crampin, A. C., . . . Clark, T. G. (2019). An integrated whole genome analysis of Mycobacterium tuberculosis reveals insights into relationship between its genome, transcriptome and methylome. *Scientific reports, 9*(1), 5204.

Gouzy, A., Poquet, Y., & Neyrolles, O. (2014). Nitrogen metabolism in Mycobacterium tuberculosis physiology and virulence. *Nature Reviews Microbiology, 12*(11), 729-737.

Guyeux, C., Senelle, G., Le Meur, A., Supply, P., Gaudin, C., Phelan, J. E., . . . Sola, C. (2024). Newly Identified Mycobacterium africanum Lineage 10, Central Africa. *Emerging Infectious Diseases, 30*(3), 560.

Hailu, E., Cantillon, D., Madrazo, C., Rose, G., Wheeler, P. R., Golby, P., . . . Gordon, S. V. (2023). Lack of methoxy-mycolates characterizes the geographically restricted lineage 7 of Mycobacterium tuberculosis complex. *Microbial genomics, 9*(5).

Hang, N. T. L., Hijikata, M., Maeda, S., Thuong, P. H., Ohashi, J., Van Huan, H., . . . Seto, S. (2019). Whole genome sequencing, analyses of drug resistance-conferring mutations, and correlation with transmission of Mycobacterium tuberculosis carrying katG-S315T in Hanoi, Vietnam. *Scientific reports, 9*(1), 15354.

He, J., Ou, Q., Liu, C., Shi, L., Zhao, C., Xu, Y., . . . Gu, D. (2017). Differential expression of long non-coding RNAs in patients with tuberculosis infection. *Tuberculosis, 107*, 73-79.

Homolka, S., Niemann, S., Russell, D. G., & Rohde, K. H. (2010). Functional genetic diversity among Mycobacterium tuberculosis complex clinical isolates: delineation of conserved core and lineage-specific transcriptomes during intracellular survival. *PLoS pathogens, 6*(7), e1000988.

Hu, Y., & Coates, A. R. (2011). Mycobacterium tuberculosis acg gene is required for growth and virulence in vivo. *Plos one, 6*(6), e20958.

Jha, B., Kumar, D., Sharma, A., Dwivedy, A., Singh, R., & Biswal, B. K. (2018). Identification and structural characterization of a histidinol phosphate phosphatase from Mycobacterium tuberculosis. *Journal of Biological Chemistry, 293*(26), 10102-10118.

Jinich, A., Nazia, S. Z., Tellez, A. V., Wu, A. M., Smith, C. M., & Rhee, K. (2022). Genome-wide co-essentiality analysis in Mycobacterium tuberculosis reveals an itaconate defense enzyme module. *bioRxiv*, 2022.2009. 2027.509804.

Johnson, R., Streicher, E. M., Louw, G. E., Warren, R. M., Van Helden, P. D., & Victor, T. C. (2006). Drug resistance in Mycobacterium tuberculosis. *Current issues in molecular biology, 8*(2), 97-112.

Jones, R. M., Adams, K. N., Eldesouky, H. E., & Sherman, D. R. (2022). The evolving biology of Mycobacterium tuberculosis drug resistance. *Frontiers in Cellular and Infection Microbiology, 12*, 1027394.

Kerns, P. W., Ackhart, D. F., Basaraba, R. J., Leid, J. G., & Shirtliff, M. E. (2014). Mycobacterium tuberculosis pellicles express unique proteins recognized by the host humoral response. *Pathogens and disease, 70*(3), 347-358.

Korycka-Machała, M., Pawełczyk, J., Borówka, P., Dziadek, B., Brzostek, A., Kawka, M., . . . Strapagiel, D. (2020). PPE51 is involved in the uptake of disaccharides by Mycobacterium tuberculosis. *Cells, 9*(3), 603.

Kumar, A., Toledo, J. C., Patel, R. P., Lancaster Jr, J. R., & Steyn, A. J. (2007). Mycobacterium tuberculosis DosS is a redox sensor and DosT is a hypoxia sensor. *Proceedings of the national academy of Sciences, 104*(28), 11568-11573.

Kumari, B., Saini, V., Kaur, J., & Kaur, J. (2020). Rv2037c, a stress induced conserved hypothetical protein of Mycobacterium tuberculosis, is a phospholipase: role in cell wall modulation and intracellular survival. *International journal of biological macromolecules, 153*, 817-835.

Kwon, K. W., Kim, W. S., Kim, H., Han, S. J., Hahn, M.-Y., Lee, J. S., . . . Shin, S. J. (2017). Novel vaccine potential of Rv3131, a DosR regulon-encoded putative nitroreductase, against hyper-virulent Mycobacterium tuberculosis strain K. *Scientific reports, 7*(1), 44151.

Le Hang, N. T., Hijikata, M., Maeda, S., Miyabayashi, A., Wakabayashi, K., Seto, S., . . . Thuong, P. H. (2021). Phenotypic and genotypic features of the Mycobacterium tuberculosis lineage 1 subgroup in central Vietnam. *Scientific Reports, 11*(1), 13609.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.

Liu, Q., Zhu, J., Dulberger, C. L., Stanley, S., Wilson, S., Chung, E. S., . . . Hicks, N. D. (2022). Tuberculosis treatment failure associated with evolution of antibiotic resilience. *Science, 378*(6624), 1111-1118.

Liu, S., Wang, Z., Zhu, R., Wang, F., Cheng, Y., & Liu, Y. (2021). Three differential expression analysis methods for RNA sequencing: limma, EdgeR, DESeq2. *JoVE (Journal of Visualized Experiments)*(175), e62528.

Lu, J., Breitwieser, F. P., Thielen, P., & Salzberg, S. L. (2017). Bracken: estimating species abundance in metagenomics data. *PeerJ Computer Science, 3*, e104.

Ma, F., Zhou, H., Yang, Z., Wang, C., An, Y., Ni, L., . . . Yu, L. (2021). Gene expression profile analysis and target gene discovery of Mycobacterium tuberculosis biofilm. *Applied Microbiology and Biotechnology, 105*(12), 5123-5134.

Malik, Z. A., Thompson, C. R., Hashimi, S., Porter, B., Iyer, S. S., & Kusner, D. J. (2003). Cutting edge: Mycobacterium tuberculosis blocks Ca2+ signaling and phagosome maturation in human macrophages via specific inhibition of sphingosine kinase. *The Journal of Immunology, 170*(6), 2811-2815.

Martini, M. C., Hicks, N. D., Xiao, J., Alonso, M. N., Barbier, T., Sixsmith, J., . . . Shell, S. S. (2022). Loss of RNase J leads to multi-drug tolerance and accumulation of highly structured mRNA fragments in Mycobacterium tuberculosis. *PLoS Pathogens, 18*(7), e1010705.

Mehra, S., & Kaushal, D. (2009). Functional genomics reveals extended roles of the Mycobacterium tuberculosis stress response factor σH. *Journal of bacteriology, 191*(12), 3965-3980.

Mekonnen, D., Derbie, A., Chanie, A., Shumet, A., Biadglegne, F., Kassahun, Y., . . . Munshea, A. (2019). Molecular epidemiology of M. tuberculosis in Ethiopia: A systematic review and meta-analysis. *Tuberculosis, 118*, 101858.

Mekonnen, D., Munshea, A., Nibret, E., Adnew, B., Getachew, H., Kebede, A., . . . Benito, A. (2023). Mycobacterium tuberculosis Sub-Lineage 4.2. 2/SIT149 as dominant drug-resistant clade in northwest Ethiopia 2020–2022: In-silico whole-genome sequence analysis. *Infection and Drug Resistance*, 6859-6870.

Miotto, P., Sorrentino, R., De Giorgi, S., Provvedi, R., Cirillo, D. M., & Manganelli, R. (2022). Transcriptional regulation and drug resistance in Mycobacterium tuberculosis. *Frontiers in Cellular and Infection Microbiology, 12*, 990312.

Mohideen, A. M., Johansen, S. D., & Babiak, I. (2020). High-throughput identification of adapters in single-read sequencing data. *Biomolecules, 10*(6), 878.

Moopanar, K., & Mvubu, N. (2020). Lineage-specific differences in lipid metabolism and its impact on clinical strains of Mycobacterium tuberculosis. *Microbial pathogenesis, 146*, 104250.

Mushtaq, K., Sheikh, J. A., Amir, M., Khan, N., Singh, B., & Agrewala, J. N. (2015). Rv2031c of Mycobacterium tuberculosis: a master regulator of Rv2028–Rv2031 (HspX) operon. *Frontiers in Microbiology, 6*, 351.

Neyrolles, O., Mintz, E., & Catty, P. (2013). Zinc and copper toxicity in host defense against pathogens: Mycobacterium tuberculosis as a model example of an emerging paradigm. *Frontiers in cellular and infection microbiology, 3*, 89.

Ngabonziza, J. C. S., Loiseau, C., Marceau, M., Jouet, A., Menardo, F., Tzfadia, O., . . . Fissette, K. (2020). A sister lineage of the Mycobacterium tuberculosis complex discovered in the African Great Lakes region. *Nature Communications, 11*(1), 2917.

Nunn, A. J., Rusen, I., Van Deun, A., Torrea, G., Phillips, P. P., Chiang, C.-Y., . . . Meredith, S. K. (2014). Evaluation of a standardized treatment regimen of anti-tuberculosis drugs for patients with multi-drug-resistant tuberculosis (STREAM): study protocol for a randomized controlled trial. *Trials, 15*(1), 353.

Oldenburg, R., Mayau, V., Prandi, J., Arbues, A., Astarie-Dequeker, C., Guilhot, C., . . . Demangel, C. (2018). Mycobacterial phenolic glycolipids selectively disable TRIF-dependent TLR4 signaling in macrophages. *Frontiers in Immunology, 9*, 2.

Parsa, K., & Hasnain, S. E. (2015). Proteomics of multidrug resistant Mycobacterium tuberculosis clinical isolates: a peep show on mechanism of drug resistance & perhaps more. *Indian Journal of Medical Research, 141*(1), 8-9.

Patro, R., Duggal, G., & Kingsford, C. (2015). Salmon: accurate, versatile and ultrafast quantification from RNA-seq data using lightweight-alignment. *BioRxiv, 10*, 021592.

Peng, C., Cheng, Y., Ma, M., Chen, Q., Duan, Y., Liu, S., . . . Bu, W. (2024). Mycobacterium tuberculosis suppresses host antimicrobial peptides by dehydrogenating L-alanine. *Nature Communications, 15*(1), 4216.

Peñuelas-Urquides, K., Villarreal-Treviño, L., Silva-Ramírez, B., Rivadeneyra-Espinoza, L., Said-Fernández, S., & León, M. B. d. (2013). Measuring of Mycobacterium tuberculosis growth: a correlation of the optical measurements with colony forming units. *Brazilian journal of microbiology, 44*, 287-290.

Phelan, J. E., Lim, D. R., Mitarai, S., de Sessions, P. F., Tujan, M. A. A., Reyes, L. T., . . . Jie, S. (2019). Mycobacterium tuberculosis whole genome sequencing provides insights into the Manila strain and drug-resistance mutations in the Philippines. *Scientific reports, 9*(1), 9305.

Plinke, C., Cox, H. S., Kalon, S., Doshetov, D., Rüsch-Gerdes, S., & Niemann, S. (2009). Tuberculosis ethambutol resistance: concordance between phenotypic and genotypic test results. *Tuberculosis, 89*(6), 448-452.

Prosser, G. A., & de Carvalho, L. P. S. (2013). Kinetic mechanism and inhibition of M ycobacterium tuberculosis d-alanine: d-alanine ligase by the antibiotic d-cycloserine. *The FEBS journal, 280*(4), 1150-1166.

Qi, W., Schlapbach, R., & Rehrauer, H. (2017). RNA-seq data analysis: from raw data quality control to differential expression analysis. In *Plant Germline Development: Methods and Protocols* (pp. 295-307): Springer.

Queval, C. J., Brosch, R., & Simeone, R. (2017). The macrophage: a disputed fortress in the battle against Mycobacterium tuberculosis. *Frontiers in microbiology, 8*, 2284.

Rachman, H., & Kaufmann, S. H. (2007). Exploring functional genomics for the development of novel intervention strategies against tuberculosis. *International Journal of Medical Microbiology, 297*(7-8), 559-567.

Ramakrishnan, G., Ochoa-Montano, B., Raghavender, U. S., Mudgal, R., Joshi, A. G., Chandra, N. R., . . . Srinivasan, N. (2015). Enriching the annotation of Mycobacterium tuberculosis H37Rv proteome using remote homology detection approaches: insights into structure and function. *Tuberculosis, 95*(1), 14-25.

Raman, K., Rajagopalan, P., & Chandra, N. (2005). Flux balance analysis of mycolic acid pathway: targets for anti-tubercular drugs. *PLoS computational biology, 1*(5), e46.

Raman, S., Hazra, R., Dascher, C. C., & Husson, R. N. (2004). Transcription regulation by the Mycobacterium tuberculosis alternative sigma factor SigD and its role in virulence. *Journal of bacteriology, 186*(19), 6605-6616.

Raman, S., Puyang, X., Cheng, T.-Y., Young, D. C., Moody, D. B., & Husson, R. N. (2006). Mycobacterium tuberculosis SigM positively regulates Esx secreted protein and nonribosomal peptide synthetase genes and down regulates virulence-associated surface lipid synthesis. *Journal of bacteriology, 188*(24), 8460-8468.

Reed, M. B., Gagneux, S., DeRiemer, K., Small, P. M., & Barry III, C. E. (2007). The W-Beijing lineage of Mycobacterium tuberculosis overproduces triglycerides and has the DosR dormancy regulon constitutively upregulated. *Journal of bacteriology, 189*(7), 2583-2589.

Rehberg, N., Omeje, E., Ebada, S. S., van Geelen, L., Liu, Z., Sureechatchayan, P., . . . Kalscheuer, R. (2019). 3-O-methyl-alkylgallates inhibit fatty acid desaturation in Mycobacterium tuberculosis. *Antimicrobial Agents and Chemotherapy, 63*(9), 10.1128/aac. 00136-00119.

Riaz, M., Mahmood, Z., Javed, M. T., Javed, I., Shahid, M., Abbas, M., & Ehtisham-ul-Haque, S. (2016). Drug resistant strains of Mycobacterium tuberculosis identified through PCR-RFLP from patients of Central Punjab, Pakistan. *International journal of immunopathology and pharmacology, 29*(3), 443-449.

Rubinstein, M., Makhon, A., Losev, Y., Valenci, G. Z., Gatt, Y. E., Margalit, H., . . . Zeevi, D. A. (2023). Prolonged survival of a patient with active MDR-TB HIV co-morbidity: insights from a Mycobacterium tuberculosis strain with a unique genomic deletion. *Frontiers in Medicine, 10*, 1292665.

Sao Emani, C., & Reiling, N. (2024). The efflux pumps Rv1877 and Rv0191 play differential roles in the protection of Mycobacterium tuberculosis against chemical stress. *Frontiers in microbiology, 15*, 1359188.

Sarkar, R., Lenders, L., Wilkinson, K. A., Wilkinson, R. J., & Nicol, M. P. (2012). Modern lineages of Mycobacterium tuberculosis exhibit lineage-specific patterns of growth and cytokine induction in human monocyte-derived macrophages.

Schön, T., Werngren, J., Machado, D., Borroni, E., Wijkander, M., Lina, G., . . . Giske, C. (2021). Multicentre testing of the EUCAST broth microdilution reference method for MIC determination on Mycobacterium tuberculosis. *Clinical Microbiology and Infection, 27*(2), 288. e281-288. e284.

Schönfeld, N., Bergmann, T., Vesenbeckh, S., Mauch, H., Bettermann, G., Bauer, T. T., & Rüssmann, H. (2012). Minimal inhibitory concentrations of first-line drugs of multidrug-resistant tuberculosis isolates. *Lung India, 29*(4), 309-312.

Selvaraj, S., Sambandam, V., Sardar, D., & Anishetty, S. (2012). In silico analysis of DosR regulon proteins of Mycobacterium tuberculosis. *Gene, 506*(1), 233-241.

Sengupta, S., Nayak, B., Meuli, M., Sander, P., Mishra, S., & Sonawane, A. (2021). Mycobacterium tuberculosis phosphoribosyltransferase promotes bacterial survival in macrophages by inducing histone hypermethylation in autophagy-related genes. *Frontiers in cellular and infection microbiology, 11*, 676456.

Shah, N. S., Wright, A., Bai, G.-H., Barrera, L., Boulahbal, F., Martín-Casabona, N., . . . Lepe, R. (2007). Worldwide emergence of extensively drug-resistant tuberculosis. *Emerging infectious diseases, 13*(3), 380.

Siddiqi, S., & Rüsch-Gerdes, S. (2006). MGIT™ procedure manual Geneva. *Switzerland: Foundation for Innovative New Diagnostics*.

Siddiqi, S. H., & Rüsch-Gerdes, S. (2006). MGIT procedure manual. *Geneva, Switzerland: Foundation for innovative new diagnostics*, 41-51.

Silcocks, M., Chang, X., Thuong Thuong, N. T., Qin, Y., Minh Ha, D. T., Khac Thai, P. V., . . . Lan, N. H. (2022). Rapid Acquisition and Transmission of Drug Resistance Amongst Beijing Lineage Mycobacterium tuberculosis in Vietnam. *medRxiv*, 2022.2011. 2003.22281872.

Smith, I. (2003). Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence. *Clinical microbiology reviews, 16*(3), 463-496.

Snášel, J., Machová, I., Šolínová, V., Kašička, V., Krečmerová, M., & Pichová, I. (2021). Phosphofructokinases A and B from mycobacterium tuberculosis display different catalytic

properties and allosteric regulation. *International journal of molecular sciences, 22*(3), 1483.

Srilohasin, P., Prammananan, T., Faksri, K., Phelan, J. E., Suriyaphol, P., Kamolwat, P., . . . Leechawengwongs, M. (2020). Genomic evidence supporting the clonal expansion of extensively drug-resistant tuberculosis bacteria belonging to a rare proto-Beijing genotype. *Emerging Microbes & Infections, 9*(1), 2632-2641.

Sun, M., Ge, S., & Li, Z. (2022). The role of phosphorylation and acylation in the regulation of drug resistance in mycobacterium tuberculosis. *Biomedicines, 10*(10), 2592.

Thipkrua, N., Disrathakit, A., Chongsuvivatwong, V., Mahasirimongkol, S., Ruangchai, W., Palittapongarnpim, P., . . . Suthum, K. (2025). A large geno-spatial cluster of multi-drug resistant tuberculosis outbreak in a western district of Thailand. *Infection, Genetics and Evolution, 128*, 105715.

Tilahun, M., Wegayehu, T., Wondale, B., Gebresilase, T. T., Gebreyohannes, T., Tekola, A., . . . Nassir, M. F. (2023). Phenotypic and genotypic drug susceptibility patterns of Mycobacterium tuberculosis isolates from pulmonary tuberculosis patients in Central and Southern Ethiopia. *Plos one, 18*(9), e0285063.

Torres Ortiz, A., Coronel, J., Vidal, J. R., Bonilla, C., Moore, D. A., Gilman, R. H., . . . Grandjean, L. (2021). Genomic signatures of pre-resistance in Mycobacterium tuberculosis. *Nature communications, 12*(1), 7312.

Verboven, L., Phelan, J., Heupink, T. H., & Van Rie, A. (2022). TBProfiler for automated calling of the association with drug resistance of variants in Mycobacterium tuberculosis. *Plos one, 17*(12), e0279644.

Vyazovaya, A., Proshina, E., Gerasimova, A., Avadenii, I., Solovieva, N., Zhuravlev, V., . . . Mokrousov, I. (2020). Increased transmissibility of Russian successful strain Beijing B0/W148 of Mycobacterium tuberculosis: Indirect clues from history and demographics. *Tuberculosis, 122*, 101937.

Waddell, S., Garton, N. J., Sherratt, A. L., Lee, S.-M., Smith, R. J., Senner, C., . . . Besra, G. S. Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum.

Walters, S. B., Dubnau, E., Kolesnikova, I., Laval, F., Daffe, M., & Smith, I. (2006). The Mycobacterium tuberculosis PhoPR two-component system regulates genes essential for virulence and complex lipid biosynthesis. *Molecular microbiology, 60*(2), 312-330.

Wan, L., Hu, P., Zhang, L., Wang, Z.-X., Fleming, J., Ni, B., . . . Tan, Y. (2023). Omics analysis of Mycobacterium tuberculosis isolates uncovers Rv3094c, an ethionamide metabolism-associated gene. *Communications Biology, 6*(1), 156.

WHO. (2023). *Catalogue of mutations in Mycobacterium tuberculosis complex and their association with drug resistance*: World Health Organization.

Wikler, M. A. (2006). Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically: approved standard. *Clsi (Nccls), 26*.

Wildner, L. M., Gould, K. A., & Waddell, S. J. (2018). Transcriptional profiling Mycobacterium tuberculosis from patient sputa. *Antibiotic Resistance Protocols*, 117-128.

WU, B., WU, F., Zhang, L., WU, J., Zhang, H., Zhang, C., . . . HE, L. (2014). Correlation study between Mycobacterium tuberculosis PhoPR two-component system and pathogenicity of different virulent Mycobacterium tuberculosis. *Chinese Journal of Immunology*, 1320-1324, 1329.

Yang, X., Liu, D., Liu, F., Wu, J., Zou, J., Xiao, X., . . . Zhu, B. (2013). HTQC: a fast quality control toolkit for Illumina sequencing data. *BMC bioinformatics, 14*, 1-4.

Yruela, I., Contreras-Moreira, B., Magalhaes, C., Osorio, N. S., & Gonzalo-Asensio, J. (2016). Mycobacterium tuberculosis complex exhibits lineage-specific variations affecting protein ductility and epitope recognition. *Genome biology and evolution, 8*(12), 3751-3764.

Yu, G. (2020). Using ggtree to visualize data on tree-like structures. *Current protocols in bioinformatics, 69*(1), e96.

Zahra, N. U. A., Jamil, F., & Uddin, R. (2021). Protein integrated network analysis to reveal potential drug targets against extended drug-resistant Mycobacterium tuberculosis XDR1219. *Molecular Biotechnology, 63*(12), 1252-1267.

Annex: Published/Accepted/Submitted works and supplementary materials